



Universidade do Estado do Rio de Janeiro

Centro de Tecnologia e Ciências

Escola Superior de Desenho Industrial

Raphael Argento de Souza

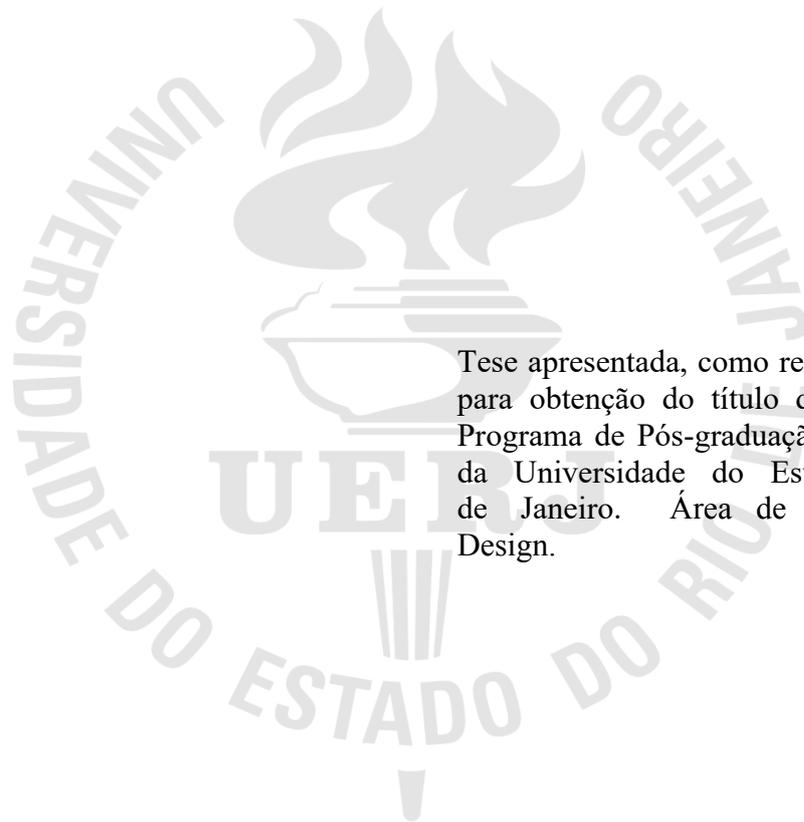
**Do realismo ao cartum: sistematização de um
modelo de conversão para captura de movimentos**

Rio de Janeiro

2019

Raphael Argento de Souza

**Do realismo ao cartum: sistematização de um modelo
de conversão para captura de movimentos**



Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-graduação em Design, da Universidade do Estado do Rio de Janeiro. Área de concentração: Design.

Orientador: Prof. DSc. André Soares Monat

Coorientador: Prof. Dr. Washington Dias Lessa

Rio de Janeiro

2019

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/BIBLIOTECA CTC/G

S729

Souza, Raphael Argento de.

Do realismo ao cartum: sistematização de um modelo de conversão para captura de movimentos / Raphael Argento de Souza. - 2019.

246 f.: il.

Orientador: Prof. DSc. André Soares Monat.

Coorientador: Prof. Dr. Washington Dias Lessa

Tese (Doutorado). Universidade do Estado do Rio de Janeiro, Escola Superior de Desenho Industrial.

1. Animação (Cinematografia) - Teses. 2. Animação 3D –Teses. 3. Design de animação - Teses. I. Monat, André Soares. II. Universidade do Estado do Rio de Janeiro. Escola Superior de Desenho Industrial. III. Título.

CDU 778.534.6

Bibliotecária: Marianna Lopes Bezerra CRB7/6386

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta tese, desde que citada a fonte.

Assinatura

Data

Raphael Argento de Souza

**Do realismo ao cartum: sistematização de um modelo
de conversão para captura de movimentos**

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-graduação em Design, da Universidade do Estado do Rio de Janeiro. Área de concentração: Design.

Aprovada em 22 de março de 2019.

Banca Examinadora:

Prof. Dr. Andre Soares Monat (Orientador)
Escola Superior de Desenho Industrial - UERJ

Prof. Dr. Washington Dias Lessa (Coorientador)
Escola Superior de Desenho Industrial - UERJ

Prof. Dr. André Ribeiro de Oliveira
Escola Superior de Desenho Industrial - UERJ

Prof. Dr. Fernando Reiszal Pereira
Escola Superior de Desenho Industrial - UERJ

Prof. Dr. Celso Pereira Guimarães
Universidade Federal do Rio de Janeiro

Prof. Dr. Gabriel Aprigliano Fernandes
Instituto Federal do Rio de Janeiro

Rio de Janeiro

2019

DEDICATÓRIA

Aos meus pais, Edmilson Mendes de Souza e Maria do Rosário Argento de Souza, por terem me incentivado e apoiado, durante toda minha vida, nos meus estudos.

AGRADECIMENTOS

Aos meus pais, Maria do Rosário Argento de Souza, e Edmilson Mendes de Souza, por me apoiarem em todos os momentos desta e de outras jornadas. Sem o apoio deles, este caminho seria bem mais atribulado. Eles estiveram comigo em meus momentos pessoais difíceis e sempre me estimularam a prosseguir nos estudos. Se grandes descobertas se apoiaram “sobre os ombros de gigantes”, posso dizer que os melhores trabalhos acadêmicos são realizados na companhia dos melhores pais. Agradeço por todas as vitórias que eles tiveram e aquelas de que precisaram abdicar para me permitir crescer, por me mostrarem que as limitações que encontramos em nossas vidas são apenas desafios que nos preparam para as conquistas e, finalmente, por terem dado a mim e aos meus irmãos um lar seguro onde pudemos crescer, acima das águas que nos assombravam todos os anos.

Aos meus irmãos, Rômulo Argento de Souza e Rodrigo Argento de Souza, por terem me incentivado a vida inteira a estudar e a escrever, mesmo quando as ideias pareciam nebulosas e o cansaço estava para me abater e por terem sido os melhores amigos que tive, tenho e terei na vida.

Ao meu orientador, Andre Monat, que soube me guiar pelos caminhos da computação, me oferecendo os melhores conhecimentos nesta área. Por ter sido um grande amigo, durante essa jornada, não apenas confiando nesta pesquisa, mas ofertando suporte intelectual e boas conversas sempre que necessário. Poucas pessoas nesse mundo são tão generosas ao compartilhar conhecimento quanto ele. Por ser um ser humano que faz com que a palavra “humano” tenha, de fato, um significado.

Ao meu coorientador, Washington Dias Lessa, por toda a paciência que teve, por ter me guiado pelas questões pertinentes à animação e ao design, por ser um intelectual inspirador e um designer que questiona o meio em que atuamos, e o faz através da sua atuação acadêmica, campo tão distante de muitos que atuam em nossa profissão.

Aos professores do PPD da ESDI – todos profissionais exemplares e sempre dedicados à formação de pesquisadores na área do design – Lígia Medeiros, Fernando Reiszal, Sydney Freitas, André Ribeiro.

Aos amigos da ESDI, Anna Tereza, Denise Fillipo e Maurício Teitel, por todo apoio durante este processo.

Aos amigos designers, Gabriel Filipe Santiago Cruz, Daniele Spada Tavares e Fábio Forti, por terem me apoiado e incentivado a acelerar a redação deste trabalho.

À Cássia Mousinho, por ter analisado o texto da minha qualificação e proposto alterações que deixaram o texto mais fluido.

À Luana Adria Vieira Monteiro, por todo o carinho, apoio e por ter sido minha querida “cobaia”, durante os experimentos aqui realizados.

Ao sempre meu professor Celso Pereira Guimarães, por ter me aberto as portas para o universo da computação gráfica, quando me aceitou como bolsista, durante a minha graduação.

Aos colegas Luiz Franca (*in memoriam*), Bruno Sérgio, Júlia Gianella, Caco, Helena de Barros, Leo Caldi, Almir Mirabeau, Francisco Valle, Yghor Kerscher, Alessandro Valério, Helga Spitz, pelo companheirismo.

Aos colegas de trabalho: Fábio Soares da Silva, Estevão Christian da Silva Leite, Marcela Botelho Tavares, Jaqueline Gomes de Jesus, Priscila Scovino, Samantha Deodato e Fábio Pires, pelo apoio.

Aos colegas da ESPM, Eliana Formiga, Maurício Vidal, Luciano Tardin, Leo Marques, por terem inserido a disciplina de captura de movimentos no currículo do design daquela instituição, me permitindo “praticar”, enquanto lecionava.

À ESDI/UERJ, que apesar da crise do governo estadual conseguiu resistir, mantendo suas atividades de pesquisa, através do seus exemplares corpos docente e técnico.

À FAPERJ, por ter disponibilizado os recursos necessários para a aquisição dos equipamentos utilizados nesta tese.

Aos que se foram durante essa caminhada, mas que estarão sempre presentes em pensamento: meu tio-avô Messias Teixeira, minha tia Anice e meu avô Joaquim Teixeira Mendes.

O conhecido é finito, o desconhecido, infinito. Intelectualmente nós estamos em uma ilha no meio de um oceano ilimitado de inexplicabilidade. Nossa obrigação, em todas as gerações, é recuperar um pouco mais de terra.

Carl Sagan

RESUMO

ARGENTO, Raphael de Souza. *Do realismo ao cartum: sistematização de um modelo de conversão para captura de movimentos*. 2019. 246 f. Tese (Doutorado em Design) – Escola Superior de Desenho Industrial, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

A animação de personagens 3D é um dos campos tecnológicos de atuação do designer, que utiliza os métodos desenvolvidos pela animação tradicional para entregar suas produções. Desde a década de 1970, entretanto, um outro campo vem se desenvolvendo paralelamente a este processo: a captura de movimentos. Inicialmente rejeitada pelos profissionais deste setor, hoje ela é parte integrante – e dependendo do tipo de produção, como os *games* – e cada vez mais utilizada para a realização de animações. Neste trabalho, foi realizada uma pesquisa envolvendo captura de movimentos e animação de personagens 3D para comprovar nossa hipótese e proposta de criação de um sistema de cartunização de capturas de movimentos de semi-automática. Nosso método utilizou o modelo proposto pela Universidade de Illinois e, através dele, percorremos as seguintes etapas: levantamento bibliográfico sobre animação e seus estilos, com vistas à aplicação no sistema proposto e questionamento dos “princípios de animação” (propondo uma classificação sua); levantamento bibliográfico exploratório no campo da visão computacional, relatando o histórico desta área e o estado da arte dos sistemas de reconhecimento de movimentos; captura de movimentos, por meio de equipamento eletromagnético; experimento empírico, utilizando o reconhecimento de movimentos, através de um algoritmo desenvolvido na linguagem de programação C#, com aplicação de lógica Fuzzy para reconhecimento e classificação de movimentos através de uso de linguagem natural; aplicação de movimentos *cartoons* com estilos de animação levantados anteriormente sobre a captura de movimentos em *software* de animação 3D. Como resultados, obtivemos: o reconhecimento bem sucedido de dois movimentos similares capturados (corrida e caminhada), sua conversão em movimentos *cartoons*; uma proposta de classificação dos princípios de animação; uma proposta de classificação de animações.

Palavras-chave: Design. Animação 3D. Captura de movimentos. Lógica fuzzy. Reconhecimento de movimentos.

ABSTRACT

ARGENTO, Raphael de Souza. *From realism to cartoon: model systematics for cartooning motion capture*. 2019. 246 f. Tese (Doutorado em Design) – Escola Superior de Desenho Industrial, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

3D character animation is one of the technological fields where designers can work. It uses methods developed by traditional animation to deliver its productions. Since the 1970's, however, another field has been developed alongside this process: motion capture. Initially rejected by animation professionals working in the industry today it is an integral part of it – specially on certain types of production such as games – and it is increasingly being used for producing animation. In this work a research involving motion capture and animation of 3D characters was carried out to prove our hypothesis and proposal of creating a semi-automatic motion capture cartoon system. We used as research method the model proposed by the University of Illinois and by it we went through the following steps: bibliographic survey about animation and its styles, aiming at the application in the proposed system and questioning the “principles of animation” (proposing a self classification); exploratory bibliographic survey in the field of computer vision, reporting the history of this area and the state of the art of motion recognition systems; motion capture using an electromagnetic equipment based system; an empirical experiment using motion recognition with a developed C# programming language algorithm, with Fuzzy logic applied to motion recognition and classification through natural language; cartoon movements application through previously know animation styles on motion capture in 3D animation software. As results we obtained: successful recognition of two similar captured movements (running and walking), their conversion into cartoon movements; a new proposal for animation principles classification; a proposal for animation classification.

Keywords: Design. 3D animation. Motion capture. Fuzzy logic. Movement recognition.

LISTA DE FIGURAS

Figura 1 –	Fase 1 do método de pesquisa proposto para investigações no campo do Design pelos pesquisadores da Universidade de Illinois.....	30
Figura 2 –	Fase 2 do método de pesquisa proposto para investigações no campo do Design pelos pesquisadores da Universidade de Illinois.....	31
Figura 3 –	Termos pesquisados na revisão da literatura.....	34
Figura 4 –	Artigos encontrados na pesquisa sistemática realizada.....	35
Figura 5 –	Pôster do Vaudeville Hurly-Burly Extravaganza, 1899.....	40
Figura 6 –	Ilustração mostrando o funcionamento da Lanterna Mágica de Athanasius Kircher.....	42
Figura 7 –	Lanterna Mágica e seus acessórios.....	43
Figura 8 –	Taumatoscópio, brinquedo que explicaria o fenômeno da persistência da visão.....	45
Figura 9 –	Ilustração de Joseph Plateau (1833), mostrando o posicionamento de imagens no disco do Fenacistoscópio.....	46
Figura 10 –	Fenacistoscópio.....	46
Figura 11 –	Gravura representando a câmara obscura e seu uso.....	48
Figura 12 –	Máquina de Silhueta.....	49
Figura 13 –	A primeira fotografia, criada por Joseph Nicephore-Niépce em uma placa de estanho, em 1826.....	50
Figura 14 –	Dispositivo inventado por Muybridge para acionar as câmeras fotográficas e capturar imagens em movimento.....	51
Figura 15 –	Eadward Muybridge. O cavalo em movimento.....	51
Figura 16 –	Fotografia “História de um salto”. De Thomas Eakins. 1844-45.....	52
Figura 17 –	Frame da Animação Betty Boop, 1933.....	53
Figura 18 –	Caricatura de John Bull, de 1709.....	54
Figura 19 –	Quadro de "A Cigarra e as Formigas", animação em stop-motion de 1911 criada por Wladislav Starevicz.....	55
Figura 20 –	Balão de diálogo usado no Gato Félix.....	56
Figura 21 –	Frame da Animação "Matches an Appeal", de Melbourne Cooper. 1899...	59
Figura 22 –	"The Grasshoper and the Ant". Ladislav Starewich, 1911.....	59

Figura 23 –	Personagens e formatos relacionados às formas geométricas: círculo, quadrado e triângulo.....	61
Figura 24 –	Frame da animação Pingu and The Birthday, com personagem comemorando seu aniversário.....	62
Figura 25 –	Aproximação da câmera no curta dos Três Porquinhos é uma apropriação da linguagem cinematográfica.....	64
Figura 26 –	Frame da animação "Asterix, Le Gaulois", 1967.....	67
Figura 27 –	Frames da animação pintada em aquarela "Onde está mamãe?", de Te Wei (1960).....	68
Figura 28 –	Frame da animação "Zhu Baizhe come a melancia", de Wan Guchan (1958).....	68
Figura 29 –	Frame da animação indiana "The Banyan Deer", apresentando o Rei (1957).....	69
Figura 30 –	Frame da Animação "A Viagem de Chihiro", de Hayao Miyazaki (2001).....	69
Figura 31 –	Frame da animação "Sinfonia Amazônica", com o curupira como personagem (1951).....	70
Figura 32 –	Frame da animação Guida, de Rosana Urbes, 2016.....	70
Figura 33 –	Frame da animação Meow, de Marcos Magalhães, 1981.....	70
Figura 34 –	Frame da animação Crônicas del Caribe, de Francisco López, 1982.....	71
Figura 35 –	Frame da animação Passajeros de la Noche (1991).....	71
Figura 36 –	Frame de Feline Follies (1919), episódio do Gato Félix.....	77
Figura 37 –	Evolução do Mickey Mouse através dos tempos.....	77
Figura 38 –	Concept de personagens do filme "Como treinar seu Dragão".....	78
Figura 39 –	Etapas de Produção de Animação.....	79
Figura 40 –	CupHEad - Game de 2016 que usa o Robber Hose animation como estilo.....	81
Figura 41 –	Exemplo de Ripple Action.....	81
Figura 42 –	Frame de animação ironizando a Alemanha nazista.....	82
Figura 43 –	Frames da animação Gerald McBoingBoing (1951).....	82
Figura 44 –	Frames da animação "A lebre e a tartaruga", Walt Disney Studios (1934).....	84

Figura 45 –	Frames da animação "Die Abenteuer des Prinzen Achmed" - Lotte Reiniger – 1926.....	84
Figura 46 –	Frames da animação em Stop-Motion "Kela mälu", de Ülo Pikkov (2011).....	85
Figura 47 –	Usos da animação.....	86
Figura 48 –	Quadro da animação "Uma cilada para Roger Rabbit".....	91
Figura 49 –	Experimento realizado por Muybridge para capturar o movimento sequencial de um cavalo.....	98
Figura 50 –	Dispositivo projetado por Muybridge para capturar um cavalo em movimento.....	98
Figura 51 –	Rifle cronofotográfico de Etienne-Jules Marey.....	99
Figura 52 –	(a) Frame de vídeo com representação de ser humano em movimento apresentado como experimento do Dr.Johansson (1971).(b) Plotagem dos movimentos de determinadas partes do corpo a partir do vídeo gravado.....	104
Figura 53 –	(A) Luva de LED desenvolvida pelo MIT; (B) Luva de fibra ótica desenvolvida por Sayre, DeFanti e Sandin na Universidade de Illinois.....	105
Figura 54 –	Modo de funcionamento do Graphical Marionette.....	106
Figura 55 –	Oclusão dos marcadores em sistemas de captura de movimentos com apenas uma câmera.....	107
Figura 56 –	Hierarquização dos segmentos do corpo humano proposta por Maxwell.....	108
Figura 57 –	Representação diagramática do corpo humano para efeitos de reconhecimento de poses.....	109
Figura 58 –	Frames do programa televisivo onde Waldo C. Graphic era exibido com seus controles eletromecânicos.....	110
Figura 59 –	Dozo, personagem de computação gráfica.....	111
Figura 60 –	Bombeiro Mario aparece no Rockfeller Center.....	112
Figura 61 –	Luva de fibra ótica.....	118
Figura 62 –	Esquema de captura óptica de movimentos.....	119
Figura 63 –	Sistema de captura de movimentos passivo, onde a luz infravermelha é emitida por LEDs ao redor da câmera.....	119
Figura 64 –	Sistemas eletromecânicos de captura de movimentos.....	120

Figura 65 –	Tela de Calibração do Neuron Mocap, sistema eletromagnético de captura de movimentos.....	121
Figura 66 –	Cirurgiões verificam os exames de imagem de um paciente sem precisar tocar nas imagens. Esta tecnologia facilitou o trabalho nas salas de cirurgia.....	125
Figura 67 –	Relação entre poses, gestos e ações.....	127
Figura 68 –	Sequência de imagens de um indivíduo sentando.....	132
Figura 69 –	Representação do Motion Energy Image. Bobick e Davis (2001).....	132
Figura 70 –	Motion History Image.....	133
Figura 71 –	Nuvem de pontos do Kinect, sistema baseado em vídeo	134
Figura 72 –	(a) Nintendo Wii Remote com Nunchuck. (b) Microsoft Kinect.....	135
Figura 73 –	Captura de movimentos com sensores ativos e luz estruturada (Raskar et al., 2007).....	136
Figura 74 –	<i>Framework</i> do reconhecimento de ações e comportamentos.....	137
Figura 75 –	Modelos usados para o reconhecimento de gestos humanos.....	138
Figura 76 –	Inferência de poses a partir de um plano traçado entre elas.....	139
Figura 77 –	Reconhecimento de ações através de "Formas de Espaço-Tempo".....	139
Figura 78 –	Segmentação do corpo humano, e reconhecimento das partes para nomear as diversas juntas.....	140
Figura 79 –	Classificação das pesquisas na área de reconhecimento de atividades humanas.....	140
Figura 80 –	Reconhecimento de pose pelo Kinect através da comparação entre as partes do corpo.....	144
Figura 81 –	(a) Reconhecimento de Ações feitos pelo sistema de [19]. O reconhecimento é realizado através da análise de frames sequenciais. (b) Proposta do autor: Reconhecimento de poses através de sequências de frames com lacunas temporais.....	146
Figura 82 –	O Uncanny Valley, gráfico que mostra o vale entre a familiaridade e a semelhança com um ser humano.....	147
Figura 83 –	(a) Asterix o Gaulês (Quadrinho); (b) Pernalonga, personagem de animação; (c) Will E. Coyote, personagem de animação 3D. Todos apresentam exageros característicos de desenhos que são aplicados em animação.....	149

Figura 84 –	Frames das animações "O Expresso Polar" e "Marte Precisa de Mães". Exemplos de Uncanny Valley com cartuns.....	149
Figura 85 –	Créditos finais do Filme Ratatouille, apresentando a mensagem "100% Animação genuína! Sem motion capture ou outros atalhos de performance foram usados na produção deste filme".....	150
Figura 86 –	Storyboards do Filme Ratatouille.....	151
Figura 87 –	Animatic do Filme "Como treinar seu dragão".....	152
Figura 88 –	Animador Carlos Baena interpretando um personagem para usar como referência no Filme Ratatouille.....	152
Figura 89 –	Atores realizado " <i>performance capture</i> " para o filme “Guerra pelo Planeta dos Macacos”.....	153
Figura 90 –	Captura de Movimentos do Ator Andy Zerkis realizada para o filme "As aventuras de Tintin".....	154
Figura 91 –	Etapas do processo de cartunização.....	156
Figura 92 –	Fluxo de produção de uma captura de movimentos.....	158
Figura 93 –	Frame de Tin Toy, curta animado da Pixar.....	159
Figura 94 –	Modelagem da malha do personagem.....	160
Figura 95 –	Texturização do Personagem. Personagem Finalizado (A). Textura da roupa (B). Textura da face (C).....	161
Figura 96 –	Criação do esqueleto e controles do personagem.....	162
Figura 97 –	Processo de skin do personagem com aplicação de zonas de influência do esqueleto (geometrias brancas dentro do personagem) sobre o modelo 3D (área laranja-avermelhada).....	162
Figura 98 –	Modelo 3D do personagem utilizado para os testes.....	163
Figura 99 –	Biped, esqueleto pronto para animação e captura de movimentos no 3ds Max.....	165
Figura 100 –	Esqueletos criados para receber a captura de movimentos e os movimentos cartunizados. À esquerda: Biped. Ao centro, o Esqueleto para receber animações cartunizadas. À direita o esqueleto que move o modelo 3D do personagem.....	165
Figura 101 –	Neuron Mocap e sua configuração para o experimento.....	166
Figura 102 –	Fotografia do indivíduo em uma das poses que calibram o sistema.....	167
Figura 103 –	Imagem da tela do Neuron Mocap com personagem sendo capturado.....	168

Figura 104 –	Indivíduo sendo capturado nas poses de caminhada (sequência de cima) e corrida (sequência de baixo).....	169
Figura 105 –	Exemplo de arquivo BVH, mostrando a hierarquia do corpo capturado.....	171
Figura 106 –	Exemplo de arquivo BVH mostrando as posições e rotações dos membros capturados.....	171
Figura 107 –	Fluxo de aplicação de movimentos no personagem.....	172
Figura 108 –	Caminhada normal e double bounce.....	173
Figura 109 –	Diagrama etapas do reconhecimento de ações usando lógica Fuzzy.....	177
Figura 110 –	Fluxo de trabalho adotado para uso com a lógica Fuzzy em nosso algoritmo.....	178
Figura 111 –	Gráfico indicando a posição do COG de acordo com o tipo de movimento realizado.....	181
Figura 112 –	Gráfico indicando a abertura das pernas nos movimentos analisados.....	182
Figura 113 –	Tela do Visual Studio, ambiente de desenvolvimento desenvolvido e disponibilizado pela Microsoft.....	184
Figura 114 –	Interface gráfica criada para o reconhecimento e análise de movimentos...	185
Figura 115 –	Declaração de variáveis para uso no banco de dados Fuzzy.....	186
Figura 116 –	Variáveis linguísticas sendo associadas aos valores das variáveis obtidas pelo sistema.....	186
Figura 117 –	Variáveis linguísticas sendo associadas aos valores das variáveis obtidas pelo sistema (cont.).....	187
Figura 118 –	Gráfico relacionando a pertinência dos ângulos da perna esquerda à abertura. Valores menores que -30° são relativos às aberturas grandes. A área compreendida entre -30° e -20° são consideradas zonas de transição de aberturas pequenas para aberturas maiores.....	189
Figura 119 –	Funcionamento do software de reconhecimento de movimentos. Os destaques em vermelho mostram o feedback do sistema para o usuário.....	190
Figura 120 –	Fluxo de etapas realizadas no 3ds Max para cartunização de personagens.....	191
Figura 121 –	Interface de usuário do “ArgenToon”, script para cartunização de personagens.....	192
Figura 122 –	Interface do ArgenToon, após a abertura do arquivo do reconhecimento de movimentos.....	193

Figura 123 –	Captura de Movimentos aplicada ao Biped e influenciando o esqueleto que atua sobre o modelo.....	194
Figura 124 –	Sequência de frames com a captura de movimentos influenciando o esqueleto do modelo 3D.....	194
Figura 125 –	Movimento cartum aplicado ao esqueleto cartum e influenciando o esqueleto que movimenta o modelo 3D.....	195
Figura 126 –	Controle de ajuste de cartunização no ArgenToon.....	196
Figura 127 –	Problema encontrado durante a transição do movimento realista para o movimento cartum.....	197
Figura 128 –	Lógica Booleana – Gráfico de exemplo com temperaturas.....	213
Figura 129 –	Pertinência de Temperaturas, segundo a lógica Fuzzy. Há um nível de transição entre os estados “Quente” e “Frio”, que permite que uma temperatura seja considerada “um pouco quente” e “um pouco fria”.....	215
Figura 130 –	Fluxo de trabalho em Lógica Fuzzy.....	218
Figura 131 –	Relações entre as variáveis linguísticas e as dimensões de uma cozinha.....	220
Figura 132 –	Relações entre as variáveis linguísticas e os recursos de um indivíduo.....	221
Figura 133 –	Gráfico com números Fuzzy triangulares.....	222
Figura 134 –	Gráfico com números Fuzzy trapezoidais.....	223
Figura 135 –	O funcionamento do sistema de inferência com duas funções de entrada associadas às variáveis "temperatura" e "Pressão".....	224
Figura 136 –	Gráfico apresentando o processo de defuzzificação.....	225

LISTA DE TABELAS

Tabela 1 –	Ranking dos 20 Títulos de maior bilheteria.....	73
Tabela 2 –	Divisão dos períodos históricos da Animação.....	74
Tabela 3 –	Estilos e características dos movimentos de animação.....	80
Tabela 4 –	Técnicas utilizadas em animações.....	88
Tabela 5 –	Princípios de animação e suas características.....	92
Tabela 6 –	Proposta de classificação de Princípios de animação.....	93
Tabela 7 –	Artigos relacionados à pesquisa do termo "captura de movimento", na base de periódicos da CAPES.....	100
Tabela 8 –	Análise dos arquivos gravados. Na primeira coluna estão os movimentos realizados. Na segunda, os ângulos máximos encontrados entre as pernas durante os movimentos e, na terceira, a posição Y do COG (movimento de cima para baixo de vice-versa).....	180
Tabela 9 –	Artigos relacionando reconhecimento de movimentos e atividades humanas.....	227

SUMÁRIO

	INTRODUÇÃO.....	18
1	METODOLOGIA.....	30
2	ANIMAÇÃO, ESTILOS E LINGUAGEM.....	39
2.1	Animação e cinema: o nascimento de uma indústria do entretenimento audiovisual.....	39
2.2	Desenvolvimento da animação: período e estilos.....	65
2.3	Animações: usos, técnicas e princípios.....	83
2.3.1	<u>Usos da animação.....</u>	85
2.3.2	<u>Técnicas de animação.....</u>	88
2.3.3	<u>Princípios de animação.....</u>	89
3	REVISÃO BIBLIOGRÁFICA DE CAPTURA DE MOVIMENTOS.....	95
3.1	Antecedentes da visão computacional.....	97
3.2	Histórico.....	100
3.3	Tecnologias usadas em capturas de movimentos.....	116
3.3.1	<u>Sistemas de captura de movimentos acústica e por fibra ótica.....</u>	117
3.3.2	<u>Captura de movimentos óptica.....</u>	118
3.3.3	<u>Captura de movimentos eletromecânica.....</u>	120
3.3.4	<u>Captura de movimentos eletromagnética.....</u>	121
3.3.5	<u>Captura de movimentos por vídeo.....</u>	122
3.4	Poses, gestos e ações – o reconhecimento de atividades humanas.....	125
3.5	A evolução no reconhecimento de movimentos.....	127
3.6	O desafio para a indústria do audiovisual e do entretenimento: o uncanny valley.....	146
3.7	O processo de animação e o processo de reconhecimento de poses: caminhos possíveis.....	150
3.8	Lacunas existentes e possibilidades de pesquisas.....	154
4	EXPERIMENTAÇÃO: DO REALISMO AO CARTUM.....	156
4.1	O processo de captura de movimentos e transferência para o personagem.....	157
4.1.1	<u>Criação do personagem 3D.....</u>	158
4.1.2	<u>Preparação dos esqueletos para captura de movimentos.....</u>	163

4.2	Captura de movimentos.....	166
4.3	A criação dos conjuntos de arquivos: capturados e cartuns.....	171
4.4	Lógica Fuzzy: motivações para aplicação na pesquisa.....	173
4.5	Reconhecendo movimentos através de uma perspectiva imprecisa (Fuzzy).....	175
4.5.1	<u>Desenvolvimento do <i>software</i> de identificação de movimentos.....</u>	177
4.5.2	<u>Aplicação no <i>software</i> de animação.....</u>	191
4.6	Reflexões.....	196
	CONCLUSÃO.....	199
	REFERÊNCIAS.....	204
	ANEXO I – Lógica Fuzzy.....	212
	ANEXO II – Lista de artigos sobre visão computacional.....	227

INTRODUÇÃO

A animação é um campo de atuação muito específico para o designer. Nas últimas décadas, o mercado de animação mundial¹ e nacional² se expandiu de maneira acelerada, gerando empregos, possibilitando o surgimento de novos estúdios e a expansão das carreiras ligadas à área da indústria criativa e de áreas tecnológicas que dão suporte a esta indústria.

Este crescimento se deu principalmente através do surgimento de novos estúdios de animação que, usando técnicas e tecnologias surgidas nas décadas de 70 e 80, conseguiram atrair a atenção do público consumidor. Dentre estas tecnologias destacam-se a técnica de animação 3D e a estereoscopia, e seus desdobramentos, através da captura de movimentos e de efeitos especiais aprimorados através das pesquisas na área da computação gráfica.

Some-se a isto a preocupação desses estúdios em criar histórias que atraíssem e agradassem tanto o público infantil quanto o público adulto, através de roteiros mais flexíveis. Assim, a animação passou a ser um entretenimento voltado para públicos de diversas idades e segmentos. Histórias criadas por estúdios como a Dreamworks™ (*Shrek*) e Disney/Pixar™ (“Os incríveis”) são exemplos de tramas criadas para atender a públicos de diversas faixas etárias.

O mercado da indústria do entretenimento (em especial o de animação) vem mostrando crescimento e força nos últimos anos e a tendência para os próximos anos é de crescimento em diversos países, entre eles o Brasil.^{1,2}

Assim como este mercado vem crescendo, toda a indústria de tecnologia associada a ele acompanha seu crescimento, tendo como motor as necessidades do setor, em termos de contratação de profissionais, geração de ferramentas e soluções para atender as demandas dele surgidas. Nos últimos vinte anos, algumas novas carreiras surgiram e empresas cresceram tendo por foco o atendimento às necessidades dessa indústria. Dentre as profissões

¹ A expansão dos mercados de animação aparece nos relatórios da AJA (The Association of Japanese Animation) <http://aja.gr.jp/english/japan-anime-data>, acesso em: 8 Nov 2018; <https://www.grandviewresearch.com/industry-analysis/3d-animation-market>, acesso em: 8 Nov 2018; <http://www.digitaljournal.com/pr/3685981>, acesso em 8 Nov. 2018; <https://markets.businessinsider.com/news/stocks/global-animation-vfx-games-industry-report-2018-2020-1022993410>, acesso em: 8 Nov. 2018.

² O mercado brasileiro de animação vem se expandindo, ano a ano, incentivado por investimentos públicos e privados, como pode ser visto em matérias publicadas na imprensa local: <http://bravi.tv/brasil-mostra-a-forca-de-sua-industria-de-animacao-no-festival-de-annecy/>, acesso em: 8 Nov. 2018; http://observatoriodaimprensa.com.br/diretorio-academico/_ed746_a_evolucao_da_animacao/, acesso em: 8 Nov. 2018; <http://revistadecinema.com.br/2018/02/brasil-lanca-maior-numero-de-animacoes-em-22-anos/>, acesso em: 8 Nov. 2018; <http://www.festivaldegramado.net/crescimento-do-cinema-de-animacao-no-brasil-se-reflete-em-gramado/>, acesso em: 8 Nov. 2018; <http://www.ebc.com.br/cultura/2016/02/o-brasil-exportava-talentos-hoje-exporta-animacao-leia-entrevista-com-cesar-coelho>, acesso em: 8 Nov. 2018.

surgidas neste período, destacam-se a de profissionais criativos e de tecnologia, como modeladores, escultores digitais, iluminadores digitais, *technical directors*, programadores de *interfaces*, entre outros.

Empresas de *software* como a Autodesk™ e a Adobe™ também se beneficiaram da ampliação desse mercado, tendo ampliado suas áreas de atuação. Outras aproveitaram a oportunidade para se expandir, como a própria Autodesk™ e algumas surgiram, como a Pixologic™ e a Maxon™ (criadoras do ZBrush™ e do Cinema 4D™), e hoje ocupam um espaço de mercado restrito na década de 1980.

O meio acadêmico também se beneficiou desse setor, pois o desenvolvimento de algoritmos, *softwares* e *plugins* ocorreu tanto dentro das empresas já consolidadas quanto daquelas surgidas de produtos advindos de pesquisas nas universidades. O melhor exemplo desse tipo de interação é o algoritmo desenvolvido por Edwin Catmull em sua tese, em 1974, para gerar superfícies tridimensionais sombreadas (CATMULL, 1974) – um problema para a época.

No Brasil, o setor de animação começou a ser percebido como um potencial mercado de trabalho e de produção de cultura, através do surgimento de festivais. Sua expansão se deu a partir de incentivos governamentais e da iniciativa de alguns empreendedores, de forma individual ou coletiva.

Assim, festivais, como o Anima Mundi, o Baixada Animada e o AnimArte!, atraíram o público brasileiro e acenderam holofotes sobre esta área, não apenas apresentando filmes, mas também o potencial dessa indústria, em termos de geração de renda e de empregos nos países onde é fomentada.

Ao considerarmos o fomento ao setor, é possível notar um crescimento nos últimos anos, alimentado principalmente por órgãos governamentais, como a Ancine e a RioFilmes. Estas agências estatais lançaram, no início dos anos 2000, editais objetivando o incentivo às produções nesse setor. Através de seu apoio, alguns estúdios brasileiros cresceram e outros surgiram. Entre eles, destacam-se a “2D Lab” e o “Copa Studio”, responsáveis pela criação de séries brasileiras e que continuam tendo importância no cenário nacional.

Essas ações não se deram por acaso. Desde 1994, países com economias desenvolvidas procuraram identificar setores da economia que auxiliassem o crescimento da economia, através da distribuição pulverizada de renda. Foram identificados, então, diversos setores que promovem esse tipo de crescimento, através da comercialização de capital intelectual. O conjunto desses setores combinados foi denominado de “economia criativa”.

O último relatório publicado pelo Programa de Desenvolvimento das Nações Unidas (UNDP) (UNDP, 2013³) em conjunto com a UNESCO revela que as áreas constituintes da economia criativa variam de contexto a contexto. Entretanto, podemos dizer que algumas áreas são encontradas na maior parte das classificações: artes visuais, design, arquitetura, *games*, música, literatura, publicidade, *software*, artes e antiguidades, filme, moda, televisão e rádio e festivais. É nessas áreas que a animação se insere melhor.

Em termos mundiais, o setor de animação é acompanhado do setor de tecnologia. Desde seu surgimento, a animação utiliza-se de dispositivos e ferramentas para auxiliar os artistas na criação da ilusão de movimento e vida.

Estúdios possuem setores⁴ responsáveis pela criação de ferramentas e de aplicações que auxiliam e aumentam a qualidade do produto exibido, diferenciando suas produções das criadas por seus concorrentes. Um dos exemplos mais conhecidos no setor de animação é o *plugin* RenderMan™⁵, criado pela Pixar™ para realizar a finalização das cenas deste estúdio e cujas versões são disponibilizadas comercialmente. Desta forma, tecnologias proprietárias, desenvolvidas com exclusividade para os estúdios tornam-se uma também fontes de renda, através da comercialização destas.

No Brasil, apenas alguns setores do audiovisual, como as emissoras de televisão – entre elas, a Rede Globo⁶, o SBT, a TV Record, e a Rede Bandeirantes – e os órgãos de pesquisa – como o Instituto Nacional de Tecnologia⁷ (INT) – investem na criação de soluções tecnológicas próprias, pois enxergam a existência de setores de P&D (pesquisa e desenvolvimento) como fatores de diferenciação em relação a seus concorrentes. A maior parte dos estúdios de animação existentes no Brasil infelizmente não possui esses setores, valendo-se de soluções prontas encontradas no mercado e da adaptabilidade dos profissionais por eles contratados às ferramentas disponíveis.

Esta situação se torna mais crítica em setores de alta tecnologia, como a animação 3D e a captura de movimentos, objetos de pesquisa desta Tese. A captura de movimentos, ou

³ UNDP (Human Development Report) 2013: The Rise of the South: Human Progress in a Diverse World. United Nations Development Programme, 2013.

⁴ Como exemplos, podemos citar o setor de tecnologia da Blue Sky (<https://medium.com/blue-sky-tech-blog>, acesso em: 8 Nov. 2018) e o setor de tecnologia da Pixar (<https://www.pixar.com/technology>, acesso em: 8 Nov. 2018).

⁵ <https://renderman.pixar.com/>, acesso em 8 de novembro de 2018.

⁶ A Rede Globo, por exemplo, possui setores específicos de pesquisa e inovação para sua produção, como pode ser visto em http://redeglobo.globo.com/TVGlobo/Comunicacao/Institucional/SiteFolder/tvg/g_tecnologia_inovacao/0,,0,00.html (acesso em: 8 Nov. 2018).

⁷ O INT fundou, em outubro de 2017, o Centro de Tecnologia do Movimento, instituição responsável por realizar estudos e pesquisas no desenvolvimento de atletas de alto desempenho para competições, usando tecnologias de captura de movimentos (<http://www.int.gov.br/component/content/article?id=8110>, <http://www.faperj.br/?id=3483.2.0> e <http://www.faperj.br/?id=3647.2.1>, acessos em: 8 Nov. 2018).

MoCap (*Motion capture*) é uma tecnologia que realiza o reconhecimento de movimentos de seres e objetos. Quando é usada com humanos, os sistemas existentes reconhecem as partes do corpo e os convertem em dados (valores numéricos) que serão aplicados a um personagem 3D. A princípio, a diferença entre um sistema de MoCap e a gravação de um vídeo (que também captura movimentos) é o fato de a captura permitir a análise e edição dos dados de movimentos armazenados, alterando-o. Os vídeos, por outro lado, permitem apenas cortes, fusões, alteração na velocidade de exibição e na aplicação de efeitos sobre eles, não possibilitando a alteração do movimento gravado do indivíduo.

Os estúdios brasileiros tradicionalmente não utilizam tecnologias de capturas de movimentos para produzir animações, devido aos custos de sua aquisição e à parca disponibilidade de profissionais qualificados no setor.⁸ Em geral, os investimentos nessa tecnologia estão restritos às universidades e aos centros de pesquisa que recebem recursos oriundos de editais governamentais e emissoras de televisão que, em função da relação custo-benefício, aplicam essas tecnologias às suas produções.

Boa parte da indústria da animação nacional e internacional realiza suas produções, através de “fórmulas” já consagradas pela prática. Os personagens de séries, curtas e longas-metragens são animados segundo os “princípios de animação” (Thomas e Johnston, 1981), uma série de regras desenvolvidas nos estúdios de animação durante o século XX que tornam o movimento de personagens cartuns mais agradáveis. Considerando que a tecnologia de captura de movimentos entrega animações com movimentos realistas e que, para cartunizar estes movimentos, é necessário esforço e tempo de produção, os estúdios optam por utilizar diretamente o trabalho de animadores profissionais, relegando a captura de movimentos a poucas produções de animação, bem como às áreas de *games* e filmes.

Entretanto, nos últimos anos, há um esforço cada vez maior por parte da indústria do audiovisual para a utilização da tecnologia de captura de movimentos em suas produções. Isto se dá graças aos avanços nesse campo que começam a reduzir consideravelmente o tempo de produção e a qualidade dos movimentos capturados. O personagem Golum, do filme “O Senhor dos anéis”, é um exemplo do uso desta tecnologia. Embora exista certo temor de profissionais de animação, quanto à redução do seu mercado de trabalho, devido à utilização dessa tecnologia, alguns pesquisadores relatam que não acreditam que isso acontecerá, especialmente devido à natureza criativa do meio e às ainda existentes dificuldades técnicas no setor (KITAGAWA; WINDSOR, 2008, p. 47). Ademais, há a crença geral no meio do

⁸ O custo dos equipamentos nesta área do conhecimento é um problema recorrente. Em 2002, Zhongxiang, Yueting, Feng et al, já apontavam este gargalo nas pesquisas realizadas (ZHONGXIANG et al., 2002).

audiovisual que a captura de movimentos e a animação de personagens são áreas distintas da produção, mas que se entrelaçam em alguns pontos (KITAGAWA; WINDSOR, 2008, p. 1).

Pesquisas realizadas de 2000 até 2018 apresentam algumas tentativas de aplicação de princípios de animação às capturas de movimentos em personagens tridimensionais, como as vistas em Wang, Drucker, Agrawala, e Cohen (2006) e Kwon e Lee (2011). Os algoritmos criados para estes fins são apropriados para uma gama limitada de movimentos como caminhar e correr, porém, não possuem a possibilidade de expansão, adicionando outros tipos de movimentos, como aqueles realizados cotidianamente tais quais “chutar”, “desferir um soco” ou “se espantar”, movimentos comuns em animações de personagens.

Assim, existiria uma forma de capturar e reconhecer esses movimentos, aplicando os “princípios de animação” diretamente a eles? A existência de um sistema com esta natureza poderia incentivar a adoção de tecnologias de captura de movimentos, por parte da indústria da animação? E de que maneira um sistema assim proposto poderia tornar o processo de utilização dessas tecnologias mais eficiente? Ao mesmo tempo, um sistema que realizasse esse trabalho poderia expandir a atuação do designer na criação e na produção de animações?

A presente pesquisa foi pensada para responder a estas perguntas, e os objetivos gerais e específicos abaixo definidos foram elaborados de maneira a atendê-las.

Objetivos da Tese

Considerando que a geração de tecnologias inovadoras é um diferencial para as empresas que operam com a animação e que, historicamente, as empresas que se destacaram neste meio foram aquelas que investiram no setor de pesquisa e desenvolvimento aplicados às suas produções, este trabalho se propõe desenvolver uma nova tecnologia para a aplicação de movimentos cartuns à captura de movimentos.

Em sua primeira etapa, é realizada uma discussão sobre a animação e seus estilos, através da história. A segunda apresenta o desenvolvimento da captura de movimentos para que, ao final, a ferramenta proposta de cartunização de movimentos capturados consiga unificar os dois conceitos. Acreditamos que o uso desta ferramenta poderá aumentar a produtividade dos estúdios e permitir ao profissional de design trabalhar diretamente com estilos de cartuns de animação de personagens.

Como será abordado no capítulo 3, a tecnologia de captura de movimentos desenvolveu-se nos últimos quarenta anos, desde as primeiras pesquisas sobre a percepção

do movimento, na qual foram utilizadas lâmpadas fixadas ao corpo de indivíduos para acompanhar seus movimentos (JOHANSSON, 1973), passando pela criação de sistemas profissionais para utilização nos setores acadêmicos e militares (CUTTING, 1978; BALLARD; BROWN, 1982; MAXWELL, 1983; LEE; CHEN, 1985; STURMAN; ZELTZER, 1994; BOBICK; DAVIS, 2001). Sua utilização se popularizou, através da utilização de equipamentos como o Kinect™ e o Wii Motion™, ambos aplicados na indústria de *games* (TANAKA et al., 2012).⁹ Aliada a esta realidade, a redução de custos de produção de produtos eletrônicos, bem como as iniciativas de criação de sistemas de captura de movimentos alternativos, como o Neuron Mocap™,¹⁰ permitiram o barateamento dos equipamentos utilizados neste setor e, conseqüentemente, o surgimento de novos desenvolvedores de aplicações para esta tecnologia.

As instituições de pesquisa e o mercado receberam bem estas iniciativas. As primeiras, porque permitiram que mais pesquisadores pudessem se envolver com a temática da captura de movimentos. Os últimos, porque os estúdios que outrora não tinham acesso a esta tecnologia passaram a usá-la em suas produções.

Como qualquer tecnologia, diversas correções e aprimoramento são desenvolvidos, ao longo do tempo, para aprimorar e tornar mais eficiente o uso desses sistemas. A produção de *hardware* pode ser realizada tanto pela academia, através das pesquisas (CARSE et al., 2013; BUYS et al., 2014) quanto pelos atores de mercado.¹¹ Já a produção de *softwares* e aplicativos para os *hardwares* existentes pode ser considerada mais acessível aos pesquisadores, devido à facilidade de acesso às ferramentas de desenvolvimento para os *hardwares* existentes – por vezes disponibilizadas pelos próprios fabricantes, os *Kits* de Desenvolvimento de *Software* (SDK)¹² – e a utilização de formatos abertos de arquivos, como o BVH, que permitem sua leitura, análise e gravação. Assim, um dos ramos mais explorados por pesquisadores, na academia e na indústria, é o da criação de ferramentas para esses sistemas.

Considerando esse contexto, esta Tese tem como objetivo geral o desenvolvimento de uma ferramenta de análise e reconhecimento de movimentos adquiridos através de captura de movimentos para sua conversão automática em movimentos cartunizados. Através de testes,

⁹ Sobre a importância do Kinect, ver a entrevista de Alex Kipman, diretor de incubação do Microsoft Xbox, em <https://www.popularmechanics.com/culture/gaming/a6234/how-microsoft-kinect-was-made/>.

¹⁰ <https://neuronmocap.com/>, acesso em 18 Out. 2018.

¹¹ Dentre estes atores destacamos, em 2018, as empresas Vicon (<https://www.vicon.com/>), Optitrack (<https://optitrack.com/>), Perception Neuron (<https://neuronmocap.com/>).

¹² Alguns exemplos de SDK disponibilizados pelos fabricantes de equipamentos de captura de movimentos podem ser encontrados em <https://optitrack.com/software/>

pretende-se provar que tal ferramenta pode ser aplicada, de maneira eficiente, em pesquisas e na indústria do audiovisual.

Acredita-se que o pioneirismo deste trabalho no campo do design está na conjunção que propõe entre a pesquisa do design de animação e o desenvolvimento de tecnologias para a captura de movimentos, em um esforço que pode contribuir para o fomento às pesquisas neste setor ainda incipientes no Brasil, mas bastante incentivadas nas universidades e empresas estrangeiras.

Objetivos específicos

Os objetivos específicos desta pesquisa são:

1. Identificar as principais lacunas existentes na área de captura de movimentos e animação, através de revisão da literatura;
2. Identificar os principais componentes da pose, durante a captura de movimentos, utilizando uma técnica de inteligência artificial (lógica Fuzzy) para tal;
3. Reconhecer o movimento realizado durante a captura para identificar ações e compará-las com uma base de conhecimento de movimentos;
4. Testar a possibilidade de aplicar de movimentos cartunizados sobre movimentos capturados, fazendo a conversão automática do movimento capturado em um movimento cartum;
5. Testar e validar o sistema proposto, do ponto de vista tecnológico (verificar a eficiência do sistema).

Justificativa e relevância

Esta Tese se propõe a analisar a evolução dos estilos de animação e apresentar uma solução tecnológica que permitirá a utilização mais eficiente dos equipamentos de MoCap (*Motion capture*) na área de animação, com vista a possíveis reduções de custos neste setor, ao aumento da eficiência no tempo de produção e ao aprimoramento do uso de sistemas de Mocap em produções de animação. A proposta aqui é criar um protótipo de *software* que fará a conversão de capturas de movimentos realistas em movimentos cartuns.

Acreditamos que este trabalho se justifica por uma série de motivos entre os quais destacamos a união entre a tecnologia de captura de movimento e os estilos de animação cartum, sendo o primeiro tema notadamente explorado em áreas como pesquisa em esportes, medicina, fisioterapia e segurança, e o segundo presente em pesquisas na área de cinema, design e artes.

No campo do design já existem pesquisas realizados na área de captura de movimentos, como os de Streit (2013), que avaliam parâmetros biomecânicos, utilizando o Kinect e Kinerd (2012) e estudam o reconhecimento e movimentos humanos (movimentos de

artes marciais), por meio da captura de movimentos. Entretanto, elas não são relacionadas à área específica do design de animação.

Além disso, acredita-se que este trabalho pode ser utilizado como uma referência para áreas afins, como os *games*, a realidade virtual e a realidade aumentada, que se utilizam do design de animação em suas pesquisas e produções. Outras aplicações poderão ser pensadas, como as da animação na indústria da realidade virtual, que vem se expandindo, em especial, com a criação de dispositivos como o Oculus Rift.¹³

A criação de tecnologias que relacionem o design de animação e a captura de movimentos ainda é uma área pouco explorada em pesquisas e no mercado brasileiro. Poucos são os profissionais que trabalham neste setor que conecta design e tecnologia, principalmente, devido à formação dos designers, que pouco aborda conceitos de programação. Embora esta tese não avance na questão educacional, acreditamos que ela pode contribuir para uma reflexão sobre a importância da inserção de disciplinas básicas de programação na área e esperamos que seja um incentivador para que designers brasileiros se insiram neste mercado.

Outro ponto a ser destacado neste trabalho é a criação de um método de utilização de sistemas de animações cartuns e de sistemas de capturas de movimentos simples e úteis para designers de animação que poderão replicá-los em suas produções. Esta área, em plena expansão, pode ser bastante explorada por esses profissionais, na elaboração de animações com cartunização automática de movimentos.

Por fim, o processo de animação, apesar de fascinante, é ainda praticamente artesanal, demandando equipes de designers de animação experientes que precisam ter muitas habilidades que se adquirem através da prática e do treinamento contínuo, tornando a produção demorada e custosa. A automação de parte de processo poderá torná-lo mais eficaz e acessível a estúdios e indivíduos interessados em criar histórias, usando a tecnologia de captura de movimentos. Acreditamos que este trabalho pode ser uma iniciativa para fomentar tal processo.

¹³ O Oculus Rift é um dispositivo eletrônico similar aos óculos que possibilita a imersão em ambientes de realidade virtual. Mais informações sobre ele podem ser encontradas em sua página https://www.oculus.com/?locale=pt_BR, acesso em: 18 Out. 2018.

Método

O método de pesquisa escolhido para o desenvolvimento desta Tese será descrito e detalhado *a posteriori*. De maneira geral, partiu-se de um questionamento acerca da utilização de capturas de movimentos em animações cartuns para iniciarmos uma pesquisa bibliográfica sistemática, procurando lacunas e possibilidades de avanços nesta área do conhecimento e suas interseções e validando, assim, a relevância da pesquisa.

A partir da revisão sistemática da literatura apresentada no capítulo 2, foram identificados escassos trabalhos acadêmicos e de mercado que relacionavam a captura de movimentos com a animação, o que demonstra o potencial de exploração dessa área do conhecimento e a necessidade de aprofundamentos no setor do audiovisual.

A revisão mostrou que o campo do conhecimento que mais possui artigos publicados em captura de movimentos atende pelo nome de *computer vision* (“visão computacional”), uma subárea da ciência da computação que pesquisa maneiras de fazer com que computadores interpretem imagens. Dentro desta área, um enorme esforço está sendo realizado por pesquisadores em centros de pesquisa e universidades para identificar movimentos humanos. As aplicações dão-se nos mais diversos horizontes do conhecimento onde a captura de movimentos é explorada, como a medicina, a fisioterapia, a interação homem-computador, a própria ciência da computação, a segurança de ambientes, a reabilitação de pessoas que sofreram ferimentos físicos ou possuem doenças degenerativas, as aplicações militares, entre diversas outras.

Identificar uma ausência de soluções nessa área foi gratificante e desafiador, pois destacou a relevância desta pesquisa, ao indicar que há muitos passos a serem dados em direção ao objetivo de gerar uma das possíveis soluções para o problema descoberto. Entre estes problemas, identificou-se a necessidade de geração de uma linguagem que consiga traduzir os movimentos adquiridos através de capturas para uma forma textual de identificação, objetivando um uso posterior. No nosso caso específico, seria para cartunizar esses movimentos. A tradução entre a linguagem do sistema de captura de movimentos e a linguagem textual foi realizada através da lógica Fuzzy, um sistema apropriado para essa tarefa, como será abordado nos capítulos 4 e no anexo I da Tese.

O método geral para o desenvolvimento desta baseou-se no modelo proposto pelo Instituto de Tecnologia de Illinois (Poggenpohl e Sato, 2003). Este método prevê algumas etapas: a identificação de uma pergunta de partida; a validação desta pergunta; a pesquisa bibliográfica aprofundada e, por fim, a análise, a síntese e a validação do conhecimento

adquirido ou do produto proposto. Esta forma de realizar pesquisa e conhecimento vem se mostrando bastante promissora no desenvolvimento de teses no campo do design desta instituição e, portanto foi aplicada aqui.

A partir do protótipo desenvolvido para o reconhecimento dos movimentos capturados, foram realizados testes que mostraram a eficiência do sistema e sua aplicação em design de animação em uma aplicação prática dentro de um *software* de animação 3D.

Delimitações deste trabalho

Considerando o escopo e os objetivos desta Tese, apresentamos abaixo o que não consideramos ser necessário abordar aqui, mas que podem ser temas de pesquisas futuras:

- Um sistema amplo de reconhecimento de gestos. Por se tratar de um protótipo, apenas alguns gestos foram escolhidos para o reconhecimento. Em um primeiro momento serão os movimentos/ações de caminhar e correr. Estes movimentos foram escolhidos pelas suas similaridades e para testar a precisão do sistema.
- Pesquisas futuras poderão ampliar a eficiência do sistema e pensamos em utilizarmos outros gestos, como “chutar”, “socar”, “chutar bola”, “raquetear uma bola de tênis”, ou seja, movimentos de espectro bem amplo;
- Não pretendemos ser um guia de estilos de animação. Embora muitos sejam identificados neste trabalho, acreditamos que a área está em constante mudança e novos estilos poderão surgir. Trabalhos históricos na área de animação estão descobrindo, a cada ano, estilos em países que antes eram ignorados pela indústria internacional, como a Colômbia, o México e a China (BENDAZZI, 2016);

Estrutura desta Tese

De acordo com as razões estabelecidas acima, esta Tese está estruturada da seguinte maneira:

Capítulo 1: Apresentação do trabalho que será realizado, justificando a razão dele ser uma tese de doutorado. Explicação de seus objetivos gerais e específicos;

Capítulo 2: Será mostrada a metodologia usada para a realização da pesquisa bibliográfica e das demais etapas desta Tese. Encontra-se também a revisão bibliográfica da literatura de forma sistemática. Nesta seção, será apresentado um histórico do desenvolvimento da captura de movimentos em sistemas computacionais. Além disso, serão pontuadas as principais pesquisas realizadas para o reconhecimento de gestos, bem como as lacunas encontradas no conhecimento que serão aqui solucionadas.

Capítulo 3: Será apresentada uma revisão histórica da animação e sua relação com o desenho em movimento. A diferença desta em comparação com o cinema – que apresenta fotogramas em movimento – é utilizada para diferenciar o cartum em movimento da imagem fotografada em movimento. Além disso, são identificados estilos de animação que serão utilizados no protótipo do sistema de cartunização proposto;

Capítulo 4: Será apresentada a revisão bibliográfica da captura de movimentos, indicando a evolução dessa área de conhecimento e sua relação com a visão computacional. Nesta etapa, serão levantadas as questões relevantes a serem pesquisadas no trabalho.

Capítulo 5: Será apresentado o protótipo do sistema proposto em lógica Fuzzy, bem como os resultados alcançados por ele e sua validação.

Capítulo 6: Conclusão.

Anexo I: Lógica Fuzzy. Será apresentado seu conceito e como ela é utilizada no capítulo 5, há a necessidade de explicação dos que regem sua lógica.

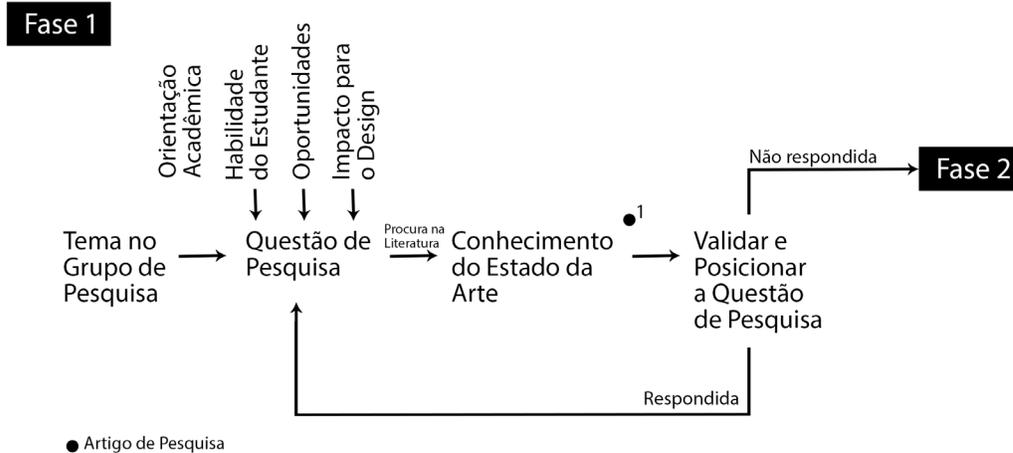
Anexo II: Artigos selecionados na revisão bibliográfica de captura de movimentos.

1 METODOLOGIA UTILIZADA

Por se tratar de uma investigação empírica no campo do design, utilizamos aqui o método proposto pela Universidade de Illinois. Elaborado por Sato e Poggenpohl (2003), ele sugere três modelos para a elaboração de pesquisas nesse campo: o primeiro focado em estudos de caso, o segundo em pesquisas teóricas e o terceiro em pesquisas metodológicas.

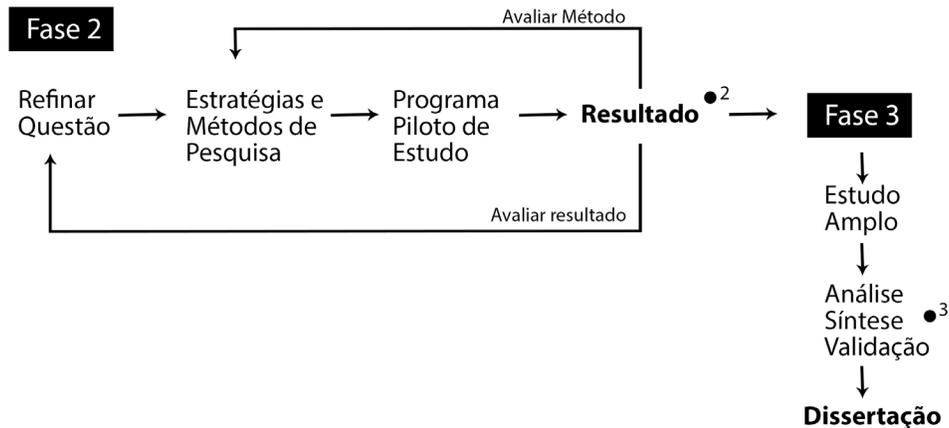
O foco detalhado nos processos de levantamento e de validação de questões para pesquisa bem como seu tratamento posterior o tornam um método apropriado para esta pesquisa. A Figura 1 mostra a primeira fase e a Figura 2, a segunda e a terceira fases que o constituem.

Figura 1 – Fase 1 do método de pesquisa proposto para investigações no campo do Design pelos pesquisadores da Universidade de Illinois.



Fonte: (POGGENPOHL; SATO, 2003. Tradução nossa).

Figura 2 - Fase 2 do método de pesquisa proposto para investigações no campo do design pelos pesquisadores da Universidade de Illinois.



Fonte: (POGGENPOHL; SATO, 2003. Tradução nossa).

Segundo este modelo, na primeira etapa do trabalho, geralmente realizada por um pesquisador atuante em um grupo de pesquisa, levanta-se uma questão relevante. Esta deve apresentar um “impacto sobre o campo do Design” (POGGENPOHL; SATO, 2003). Para alcançá-la, é necessário realizar um levantamento bibliográfico extensivo, procurando por áreas ainda não exploradas ou que apresentam lacunas relevantes para o conhecimento. A esta perspectiva, acrescentamos que, no caso de pesquisas no campo do design, também devem ser levadas em consideração as questões referentes às práticas profissionais das diversas subdivisões do design, visto que as descobertas oriundas dessas pesquisas podem revelar novos métodos de trabalho, produtos e orientações conceituais aplicáveis diretamente ao labor do designer.

Na pesquisa que aqui está sendo apresentada, a motivação nasceu da área prática, de um questionamento acerca da possibilidade de criação de um método para cartunizar personagens a partir de dados adquiridos através de captura de movimentos. Este questionamento partiu da experiência empírica do autor no mercado de trabalho (como designer e animador) e no meio acadêmico, em que leciona disciplinas relacionadas à animação de personagens.

A partir desta questão, verificou-se a necessidade de procurar sua validade na literatura existente e as lacunas existentes no conhecimento já levantadas por pesquisas anteriores, de forma a obter uma pesquisa inovadora dentro do campo do design.

A identificação dessa necessidade nos levou à revisão sistemática da literatura descrita por Dresch et al. (2015, p. 141), como base fundamental do método conhecido por *Design Science Research*. Ao propor a delimitação e a especificação dos termos da pesquisa a ser realizada, a revisão sistemática permite um levantamento rápido de questões em áreas práticas, como o design. O resultado deste levantamento será encontrado no quarto capítulo da pesquisa.

Assim, foi realizada uma pesquisa bibliográfica sistemática cujo objetivo era procurar na literatura a validade da questão apresentada e, caso esta não se mostrasse relevante, averiguar se os artefatos criados para solucioná-la abordavam, no todo ou em parte, a questão, possibilitando que outras lacunas na área do conhecimento pudessem ser exploradas. Neste caso, a procura na literatura por artefatos que apresentam soluções para o problema proposto é um tipo de pesquisa sistemática, que procura por indícios específicos acerca de soluções que advém do mundo prático.

Ao encontrarmos algumas questões – além da principal – passíveis de investigação por meio de um projeto de pesquisa, nossa pesquisa foi ampliada e, a partir daí, delimitou-se como objetivos, a identificação das lacunas e a criação de um artefato para solucioná-las, no todo ou em parte. Segundo o modelo de Illinois, estaríamos, assim, determinando os métodos e as estratégias para que a pesquisa fosse conduzida com eficiência.

Apesar de representadas na Figura 1, explicitamos abaixo as etapas do modelo de Illinois:

1. Tema pesquisado pelo grupo de pesquisa;
2. Questão de pesquisa;
3. Procura na literatura;
4. Conhecimento do estado da arte;
5. Validar e posicionar a questão de pesquisa;
6. Refinamento da questão;
7. Métodos e estratégias de pesquisa;
8. Programa piloto de estudo;
9. Estudo;
10. Estudo amplo;

11. Análise, síntese e validação;
12. Validação.

Assim, a partir da questão inicial, realizamos uma busca na literatura através de uma revisão sistemática, objetivando delimitar o estado da arte na área da captura e reconhecimento de movimentos, compreender as lacunas existentes e levantar os diversos estilos de animação e as técnicas existentes para executá-los.

A revisão sistemática da literatura é uma das principais ferramentas utilizadas para o levantamento do estado da arte, pois, através dela, potenciais artefatos e classes de problemas são evidenciados para que se possa propor um projeto de pesquisa para solucioná-los (DRESCH; LACERDA; ANTUNES JR, 2015).

Partindo daquela premissa, os seguintes termos foram pesquisados na base da Capes:⁷ *animation principles*, *3D motion capture*, *cartoon animation*, *motion capture*. Estes termos foram escolhidos, uma vez que, por meio de sua procura, eles retornavam outros trabalhos de pesquisa relevantes. Os termos foram pesquisados, usando a língua inglesa porque a maior parte da bibliografia acadêmica nas áreas pesquisadas se concentrava nessa língua. Além disso, um fator determinante para a busca foi o fato de sabermos que os estadunidenses possuem a principal cultura consumidora de animação e captura de movimentos, sendo os líderes do mercado mundial em tais áreas.

Cada busca por estes termos retornou um determinado número de artigos que foram posteriormente filtrados segundo critérios estabelecidos para o escopo do trabalho.

O termo *animation principles* foi procurado na base de dados para verificar se o assunto já estava consolidado no meio acadêmico, assim como já está nos estúdios que produzem na área. Além disso, é importante determinar se há alguma interpretação nova acerca deles, já que parte deste trabalho tratará de uma classificação sua, a ser apresentada no próximo capítulo.

⁷ A base de Periódicos da CAPES foi utilizada como ferramenta de pesquisa porque apresenta, além dos periódicos nacionais, diversos periódicos e publicações presentes em outras bases, como por exemplo, a *Web of science*. Desta forma, ela tornou-se uma ferramenta importante para esta pesquisa. A lista completa dos periódicos cobertos por ela pode ser encontrada no seguinte link: http://www.periodicos.capes.gov.br/index.php?option=com_pcollection&Itemid=105. Acesso em: 18 Out. 2018.

3D motion capture foi utilizado como um termo de busca, pois nossa pesquisa procura compreender as tecnologias existentes para a realização de capturas de movimentos e as principais aplicações propostas na área, relacionando-a a personagens tridimensionais.

O termo *cartoon animation* foi pesquisado pela mesma razão que *animation principles*, ou seja, para determinar quais questões já haviam sido levantadas anteriormente e verificar as lacunas que estariam presentes no conhecimento.

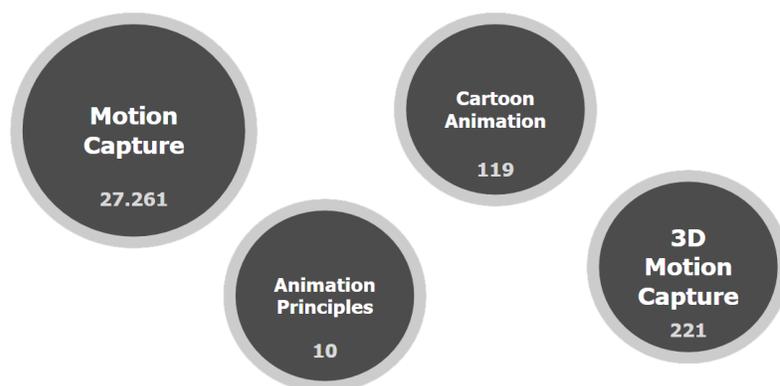
Por fim, o termo *motion capture* foi pesquisado justamente por ser a área de conhecimento diretamente relacionada a esta pesquisa, tendo sido, na pesquisa da base da Capes, o termo que mais retornou resultados de busca. Isto se explica pela existência de pesquisas nesta área há mais de 30 anos, bem como de diversos campos do conhecimento que se utilizam desta tecnologia para validar estudos e gerar novos artefatos.

Encontramos como áreas de destaque na pesquisa em captura de movimentos a fisioterapia, as aplicações militares, o audiovisual, os esportes, a medicina, os sistemas de segurança, entre outras, que contribuem, de maneira decisiva, para o reconhecimento de poses, cada qual com questões específicas a serem respondidas.

Verificamos que alguns produtos oriundos dessas pesquisas podem ser aplicados diretamente no reconhecimento de ações para animação, enquanto outras, como as avaliações de DOF (*Degrees of Freedom* – os ângulos máximos de flexão dos membros do corpo humano), realizadas pela área da fisioterapia e de reabilitação destinam-se apenas às suas áreas específicas.

As Figuras 3 e 4 mostram os termos pesquisados e a quantidade de artigos retornados para cada um deles:

Figura 3 - Termos pesquisados na revisão da literatura.



Fonte: Pesquisa realizada pelo autor na página de periódicos da CAPES.

Da primeira busca, percebe-se que a quantidade de artigos encontrados nas áreas de *motion capture*, *cartoon animation* e *3D motion capture* é deveras extensa. Estas três áreas responderam pelo total de 27.601 artigos dos quais 27.261 eram ligados à área de *motion capture*.

Esta área foi, portanto, a que apresentou o maior número de artigos devido ao grande número de aplicações em diversas áreas do conhecimento, como visto anteriormente. Muitos desses artigos não se aplicavam ao escopo da pesquisa e, portanto, foram filtrados para que apenas os mais relevantes permanecessem.

Esta foi, portanto, a primeira área a ser refinada no processo de revisão sistemática de literatura. O objetivo inicial foi filtrar os artigos por áreas diretamente relacionadas à nossa pesquisa. Desta forma, foram selecionados subtermos onde a *detecção* e o *reconhecimento de poses* poderiam ser a temática principal. Após esta separação, passamos a uma segunda etapa na qual os artigos encontrados foram filtrados por *técnicas de motion capture*, *tecnologias*, *biomecânica*, *medicina esportiva* e *biomedicina*.⁸ Por meio deste processo, reduzimos o quantitativo para 5.179 artigos relevantes, ou seja, aumentamos a relevância dos artigos encontrados e reduzimos em 82% a quantidade de triagem a ser realizada.

Figura 4 – Artigos encontrados na pesquisa sistemática realizada na base da CAPES.

Tema Pesquisado	Quant. de artigos	1º Refinamento	Quant. de artigos	2º Refinamento	Quant. de artigos	Filtros	Quant. de artigos
<i>Motion capture</i>	27.261	Técnicas	5179	<i>Animation</i>	2498	<i>Animation</i>	11
		Tecnologias		<i>Motion perception</i>		<i>Computer simulation</i>	
		Biomecânica		<i>Motion capture</i>		Revisados por pares	
		Medicina Esportiva		<i>Kinematics</i>		(-)Biometria (-)Saúde	
		Biomedicina					

⁸ Em itálico estão as áreas disponibilizadas como opções pelo portal de periódicos da Capes para a seleção desta pesquisa.

<i>Animation principles</i>	10	-	-	-	-	-	-
<i>Cartoon animation</i>	119	-	-	-	-	-	-
<i>3D motion capture</i>	221	<i>3D motion capture</i>	70	-	-	-	-
		<i>Kinematics</i>		-	-	-	-
		<i>Three dimensional mocap</i>		-	-	-	-
Livros	5	-	-	-	-	-	-
Anais de Congresso	1	-	-	-	-	-	-

Fonte: Pesquisa realizada pelo autor.

Um segundo refinamento foi realizado com os artigos apresentados e com as opções de filtragem disponíveis na base da Capes. Desta vez, foram separados artigos relacionados à *animação*, à *percepção do movimento*, à *motion capture* e à *cinemática*⁹, resultando em um total de 2.498 do qual uma grande parte estava diretamente relacionada à pesquisa.

Em seguida, um novo filtro foi aplicado, sendo selecionados apenas artigos *revisados por pares* (ou seja, aqueles cuja relevância já foi atestada por pesquisadores no meio acadêmico) e excluídos os assuntos *biometria* e *saúde*, devido às suas incompatibilidades com a pesquisa. Finalmente, selecionamos apenas aqueles que possuíam dois termos relacionados com ela: *animation* e *computer simulation*, resultando em onze artigos totais.

Os demais termos, *animation principles* e *cartoon animation*, não passaram por filtros, uma vez que apresentaram uma quantidade reduzida de artigos (129 no total) e este número por si demonstra a importância de mais uma pesquisa na área de animação.

Em seguida, o termo *3D motion capture* foi procurado na mesma base. Após uma triagem dos artigos, restaram apenas 81¹⁰ artigos ligados às áreas correlatas e de interesse da pesquisa.

⁹ Idem.

¹⁰ Os artigos podem ser encontrados no Anexo II desta pesquisa.

A relevância desta pesquisa bibliográfica sistemática se confirma ao verificarmos que uma parte das buscas encontradas possui uma aplicação na indústria do cinema e da animação, porém não são propostas através de pesquisadores da área do Design, mas por pesquisadores da área da Ciência da computação, em um campo específico conhecido por “Visão computacional” (*Computer vision*).

Desta forma, pode-se vislumbrar um grande campo de aplicação das pesquisas na área do design cujos desdobramentos podem ser a criação de produtos voltados para o entretenimento e o reconhecimento de ações humanas para diversos fins, como a segurança institucional, a sinalização de ambientes, o *wayfinding*, o reconhecimento de gestos em sistemas cuja interface do usuário poderá ser mais responsiva ou intuitiva para o usuário, para a ergonomia ou em sistemas de entretenimento, com a interatividade do usuário sendo testada e gerando produtos personalizados.

A partir desta etapa de filtragem e seleção de artigos, foram levantadas as lacunas no conhecimento e as questões que serão respondidas através da pesquisa. Paralelamente, percebeu-se que a criação de artefatos que relacionassem capturas de movimentos com animação de personagens cartuns ainda é um campo a ser explorado.

Após a definição da questão de pesquisa – o desenvolvimento de uma ferramenta de análise e o reconhecimento de movimentos adquiridos através de captura de movimentos (um modelo de algoritmo) para a conversão automática de movimentos realistas em movimentos cartunizados e de outras questões associadas – foi definida a segunda etapa da pesquisa que, segundo o modelo de Illinois, se baseia no refinamento da questão a ser procurada e da definição de métodos e estratégias de pesquisa.

Sendo nossa pesquisa de natureza empírica e com um trabalho experimental (segundo o modelo de Illinois), além da tese, propomos o desenvolvimento de um *software* relacionando design e animação. Para criar este produto utilizamos os seguintes recursos:

1. O equipamento de capturas de movimentos *Neuron MoCap*, da Noitom;
2. Microsoft Visual Studio para gerar o software de reconhecimento de ações, utilizando a linguagem de programação C#;
3. Biblioteca *Accord.Fuzzy* para C#;

4. Software 3DS Max, para aplicação dos movimentos cartuns à captura de movimentos.

Utilizamos movimentos simples e similares – a “caminhada” e a “corrida” – como base de testes para a eficiência de nosso sistema. Em um primeiro momento, realizamos a prototipação do sistema de identificação dos movimentos através do Kinect, que utiliza linguagem C# em seu SDK (*Software Development Kit*) disponibilizado pela Microsoft. Posteriormente, este protótipo foi utilizado como base para a programação do sistema de reconhecimento de capturas de movimentos advindas do *Neuron MoCap*.

Uma base de conhecimento, baseada em lógica *Fuzzy* foi implementada para que os movimentos realizados pelos usuários pudessem ser reconhecidos. Esta base de conhecimentos limitou-se aos movimentos listados no parágrafo anterior e, posteriormente, pretende-se que, com a continuidade das pesquisas, ela seja ampliada com outros movimentos. A lógica *Fuzzy* foi escolhida por possuir uma lógica de trabalho similar ao modo como seres humanos se comunicam (Anexo I – Lógica *Fuzzy*).

Após a identificação do tipo de movimento realizado, o sistema realizou a exportação dos dados, bem como a aplicação de movimentos cartuns sobre um personagem 3D no 3ds Max.

Para definirmos os estilos de movimentos cartuns utilizados, foi necessário realizar também uma extensa pesquisa bibliográfica de base. Esta será apresentada no capítulo 2 e relaciona estilos a momentos históricos.

Por fim, apresentamos os resultados de nosso experimento empírico, mostrando a eficiência de nosso sistema. Segundo o Modelo de Illinois, esta etapa corresponde à terceira e última fase: a análise, a síntese e a validação de nosso modelo, gerando um trabalho acadêmico que busca contribuir não apenas para a área do design, mas para o universo acadêmico como um todo.

2 ANIMAÇÃO, ESTILOS E LINGUAGEM

Neste capítulo será apresentado um panorama geral da animação e de sua inserção na indústria do audiovisual. Isto inclui sua relação com as artes a que se associa, como o cinema e os quadrinhos, que apresentam características formais e narrativas comuns, embora possuam pontos divergentes, em especial nas suas linguagens.

Além disso, serão apresentados exemplos de obras animadas, objetivando verificar a variedade de estilos e temas, de maneira a tornar aparente que o estilo norte-americano não é o único empregado, embora seja o mais utilizado devido à influência econômica e política dos Estados Unidos. Isto de certa forma uniformiza a produção de animação no mundo. Porém, mesmo dentro da massa que segue o mesmo estilo, existem variações.

A seguir, serão apresentadas as técnicas de animação. Aquelas mais utilizadas atualmente são o desenho animado e a animação 3D, porém outras se destacaram na história da produção audiovisual animada e ainda são utilizadas.

Por fim, será indicado, dentro desse panorama, qual o recorte contextual utilizado aqui, em termos técnicos e estilísticos, para a produção do sistema de reconhecimento de movimentos para animação. O conhecimento desses estilos e técnicas permitirá também que futuros trabalhos na área possam reconhecê-los e utilizá-los em sistemas similares.

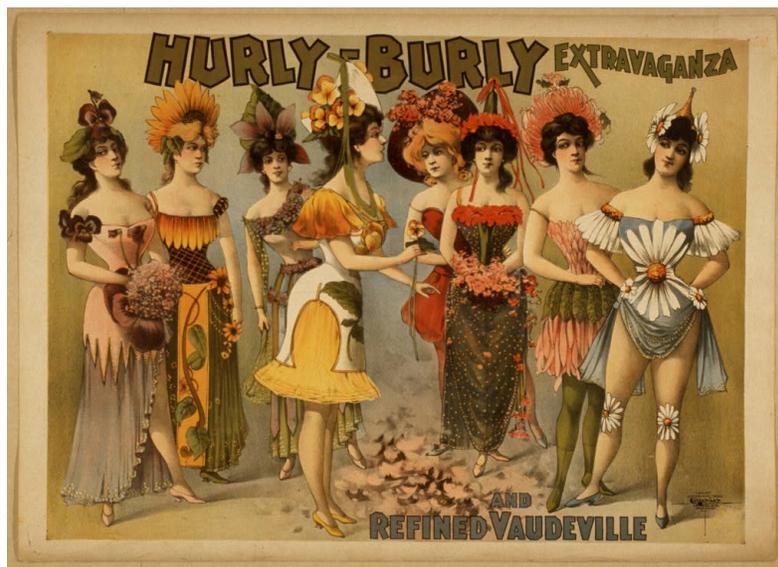
O objetivo deste capítulo é, portanto, apresentar resumidamente a história e as características da animação para, a partir desse ponto de partida, determinar quais são fundamentais para construir o sistema de cartunização que será utilizado posteriormente no trabalho.

2.1 Animação e cinema: o nascimento de uma indústria do entretenimento audiovisual

O cinema e a animação fazem parte de uma indústria maior, a do entretenimento, cujo nascimento remonta ao século XIX, voltada para a diversão popular, os cafés na Europa e os teatros de *vaudeville* nos Estados Unidos (Figura 5). A mistura de diversas formas de entretenimento com ambas faz com que iniciem suas abordagens narrativas com o objetivo de divertir o público e competir por sua atenção.

Ao contrário dos teatros, cinemas, circos, dentre outros espaços de entretenimento atualmente existentes, no século XIX, os locais onde a população se reunia para se divertir não eram especializados em apenas um tipo de *performance*. Segundo Costa (MASCARELLO, 2006) em um único espaço era possível encontrar “*performances* de acrobacia, declamações de poesia, encenações dramáticas, exposições de animais amestrados e sessões de lanterna mágica”.

Figura 5 - Poster do *vaudeville* Hurly-Burly Extravaganza, 1899.



Fonte: *Theatrical poster collection*, Library of The Congress (online), LC-USZ62-74102, <https://www.loc.gov/item/2014635695/>

Nesses locais, onde as apresentações competiam com a algazarra do público, aquelas que logo adiante se metamorfoseariam no cinema e na animação se constituíam apenas como um pequeno casulo, armazenando o embrião de ambas as artes. As projeções que gerariam o cinema eram, pois, parte de um sistema mais amplo que envolvia, além das atrações anteriormente descritas, o desenvolvimento de tecnologias voltadas para a diversão do público.

As duas tecnologias que suscitaram o surgimento do cinema e da animação foram o projetor e o cinematógrafo. O primeiro, derivado das lanternas mágicas e de seus vidros de projeção, era amplamente usado por mágicos e ilusionistas para encantar o público. O segundo era um aprimoramento das lanternas mágicas e de sua união com a fotografia.

O cinematógrafo, inventado pelos irmãos Lumière, fez parte da cadeia de apresentações e espetáculos heterogêneos nos Cafés, e esta participação lhe foi importante, pois, através dela, seus realizadores puderam encantar o futuro público consumidor, além de obterem suas primeiras histórias. Segundo Gaudreault et al. (2012, p. 93),

[...] In the beginning, the kinematograph combined several histories: that of the Théâtrophone, as we have seen, along with the magic lantern, photography, the comic strip, vaudeville, pantomime, the phonograph, popular song, the operetta, popular theater, etc. Hybridity was its *raison d'être* at a time when the mixing of cultural series occurred in response to the anthropological mutations taking place in the face of society's technological turn.³

Uma pergunta que se coloca, entretanto, é: de onde, surgiram, de fato, tanto a animação quanto o cinema? Embora a indústria cinematográfica veja a animação como mais um de seus estilos e a indústria da animação se autointitule como um meio de comunicação que porta sua própria linguagem e seja anterior ao cinema (e, portanto, seu precursor na exibição de imagens em movimento), ambos possuíram origens comuns e percorreram seus caminhos de mãos dadas, apropriando-se de diversas fontes e as adaptando aos seus próprios objetivos.⁴

Suas diferenças semânticas foram dadas apenas por suas formas de representação: uma baseada no desenho em movimento; e a outra na fotografia (a imagem natural) em movimento. O início dessa história é comum a ambas e remonta à utilização dos primeiros sistemas de projeção conhecidos como lanternas mágicas.

A história dos dispositivos de projeção de imagens começa, pois, com a invenção da lanterna mágica. Elaborada no século XVII, sua autoria é creditada a diversos inventores de diferentes países: Athanasius Kircher, Christiaan Huygens, Robert Hooke, Johannes Zahn, Samuel Rhanaeus, Petrus von Musschenbroek e Edme Gilles-Guyot (MACHADO, 2005, p. 12) são alguns dos nomes que aparecem nas publicações que a relatam. O local de sua

³ [...] No começo, o cinematógrafo combinou diversas histórias: aquelas do teatrôfono, como vimos, junto com a lanterna mágica, fotografia, a tira em quadrinhos, vaudeville, pantomima, o fonógrafo, canção popular, a opereta, teatro popular etc. O hibridismo foi sua *raison d'être* em um tempo quando a mistura de séries culturais aconteceu em resposta às mutações antropológicas que estavam tomando parte em face da virada tecnológica na sociedade. (GAUDRAULT et al., 2012, p.93, tradução nossa)

⁴ O Oscar (<https://oscar.go.com/>), maior festival de Cinema do planeta considera a Animação uma subcategoria, isolando-a das demais. Já o artigo de Crafton (2011), afirma que na verdade o Cinema é uma derivação ou “um caso especial da entidade abrangente, a animação.” Diversos autores irão discorrer sobre esta genealogia, porém no imaginário popular, a animação permanece sendo uma ramificação do Cinema.

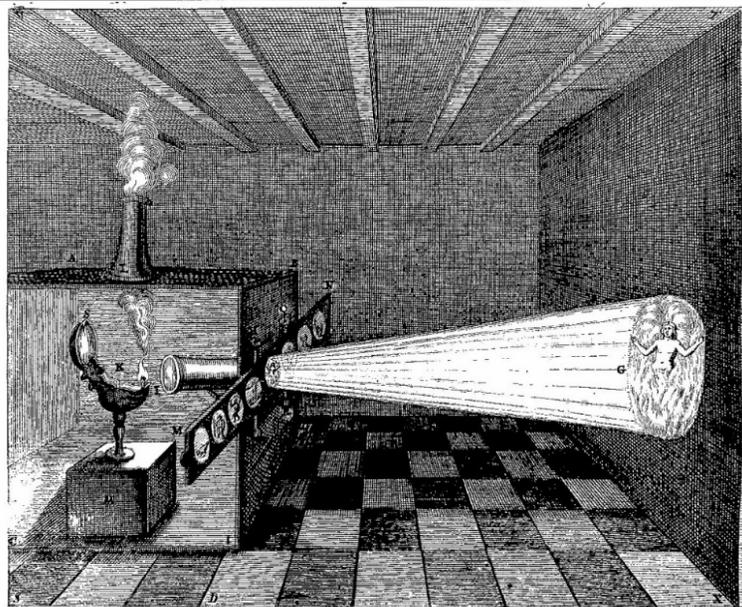
origem, entretanto, interessa pouco no contexto deste trabalho, mas sim, as funções que esse dispositivo desempenhou.

Constituída por uma câmara escura na qual, internamente, uma lamparina emitia uma luz que, ao atravessar uma placa de vidro ilustrada, projetava em uma parede a imagem ampliada, a lanterna mágica (Figura 6) realizava a exata função do projetor moderno.

Inicialmente, foi utilizada como um instrumento para a apresentação das histórias de Cristo por Athanasius Kircher (jesuíta, matemático e inventor) e sua primeira descrição data de 1646 em uma publicação do próprio Kircher, intitulada *Ars magna lucis et umbrae*, onde ele apresentava suas descobertas em relação às propriedades da luz (KIRCHER, 1646).

É importante lembrar que a fotografia ainda não havia sido inventada e, pode-se afirmar que as ilustrações nas placas de vidro não eram representações fidedignas do natural, mas sim a interpretação do artista sobre um tema que poderia ser um lugar, uma pessoa, um fenômeno, um conto ou uma lenda. Assim, elas eram representadas segundo o estilo ilustrativo da época, as convenções gráficas, as intenções narrativas e as habilidades do artista que as elaborava.

Figura 6 – Ilustração mostrando o funcionamento da lanterna mágica de Athanasius Kircher.

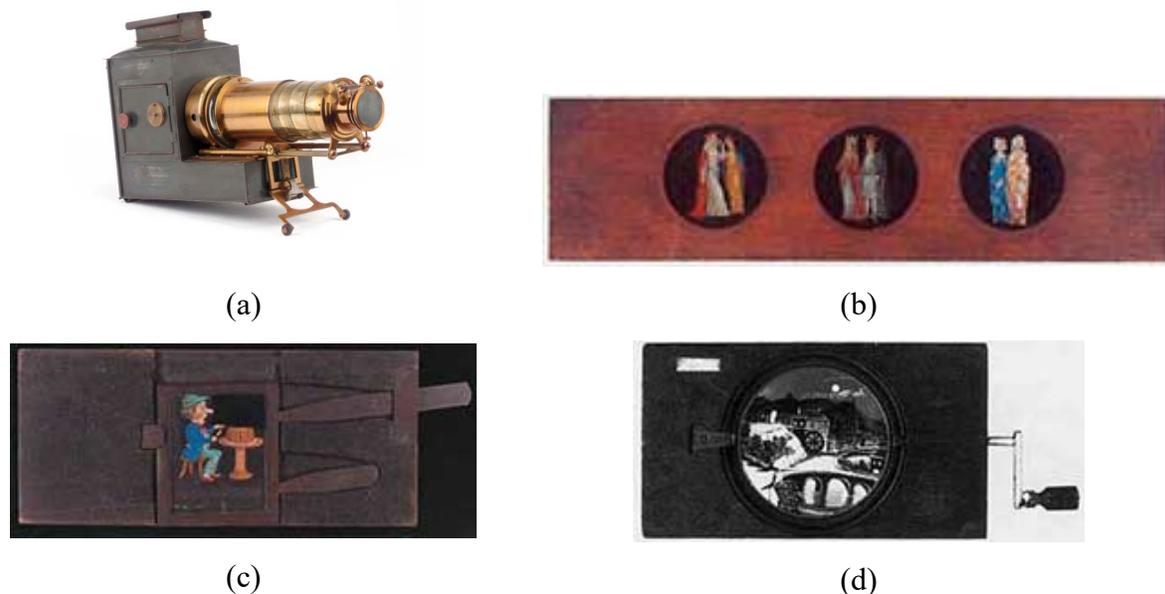


Fonte: Imagem de domínio público.

As lanternas mágicas tornaram-se muito populares na Europa e na Ásia, onde inúmeras apresentações foram feitas com placas de projeção (Figura 7 – superior). No Japão, por exemplo, foi criada uma variação desses dispositivos, feitos em madeira, denominados “Furo”, a partir de onde eram projetadas imagens em uma tela feita de papel de arroz (WILLIS, 2012). Devido à sua transparência natural, esse suporte permitia que as imagens fossem exibidas perdendo pouco de sua nitidez.

A passagem de uma imagem à outra era realizada com o simples percorrer de uma placa que deslizava de um lado para o outro da lanterna. Com o passar do tempo, elas foram ganhando mecânicas mais complexas, como engrenagens e alavancas, que permitiram que novas camadas de ilustrações sobrepostas fossem deslizadas e gerassem a ilusão de movimento para o espectador (Figura 7 – [a] e [b]).

Figura 7 – A lanterna mágica e seus acessórios.



Legenda: (a) Lanterna mágica; (b) Slides múltiplos da lanterna mágica; (c) e (d) Dispositivos mecânicos utilizados para criar ilusão de movimento.

Fonte: www.magiclantern.org.uk. Acesso em: 31 Jan. 2018.

A ilusão de movimento apresentada por estes dispositivos dependia exclusivamente do operador do equipamento. Caso ele movimentasse ou rotacionasse as alavancas e

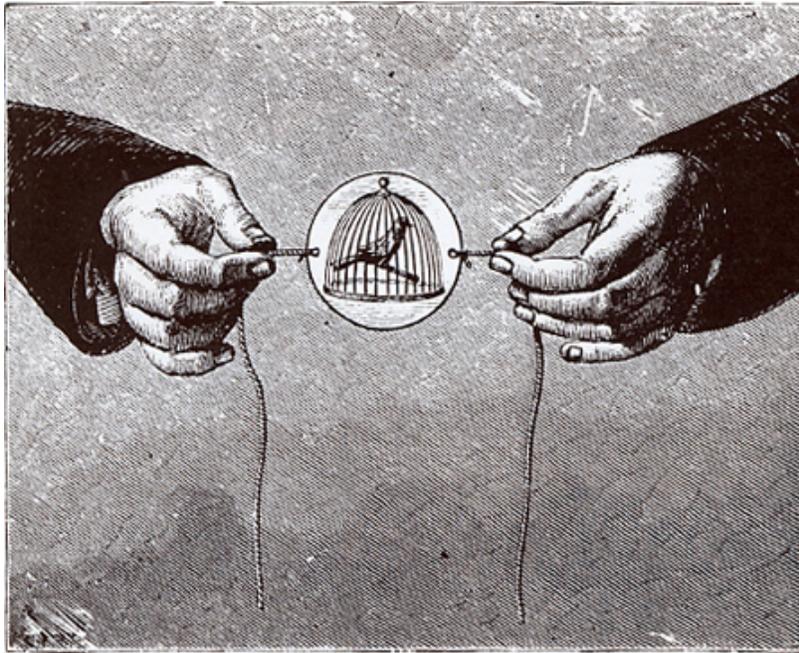
manivelas rapidamente, o personagem ou objeto aparentaria se mover rapidamente. O *timing* não era, assim, controlado por um fator de velocidade objetivo, como um motor ou um conjunto de engrenagens, mas subjetivo, que dependia do operador e alterava a percepção da ação.

Elas foram usadas igualmente por mágicos e prestidigitadores em espetáculos de entretenimento. Um deles, intitulado *Fantasmagorie*, foi elaborado por Etienne Gaspard Robert, no final do século XVIII. Este espetáculo fez bastante sucesso e era assim denominado porque simulava a aparição de espíritos no palco onde era exibido, em geral nos cafés e *vaudevilles* até o século XIX.

A evolução dos sistemas de projeção acompanhou os estudos sobre a percepção de movimentos. A procura por explicações científicas deste assunto foi tema de diversas discussões, visto que alguns fenômenos naturais ainda eram objetos de dúvidas naquela época. Segundo Barbosa Júnior (2001, p. 33), “[...] as pessoas se questionavam a respeito do fenômeno de perceber os raios das rodas das carruagens girando ao contrário ou mesmo ficando paradas, quando o veículo se movia rapidamente”.

Os primeiros a proporem teses sobre a percepção humana de movimentos foram Peter Mark Roget, em 1824 e, Joseph Plateau, em 1829. Roget publicou um estudo, intitulado *The persistence of the vision with regard to moving objects*, no qual afirmou que o fenômeno de percepção de movimentos em imagens exibidas sequencialmente ocorria devido a uma característica física do olho humano que, segundo ele, retinha as imagens na retina por uma fração de segundo. A comprovação de sua tese era realizada através de brinquedos como o taumatoscópio (Figura 8). Plateau propôs o mesmo conceito e seu trabalho permitiu a criação de instrumentos como o fenacístoscópio, o zootrópio e, posteriormente, o cinematógrafo (MACHADO, 2005, p. 19).

Figura 8 - Taumatroscópio, brinquedo que explicaria o fenômeno da persistência da visão.



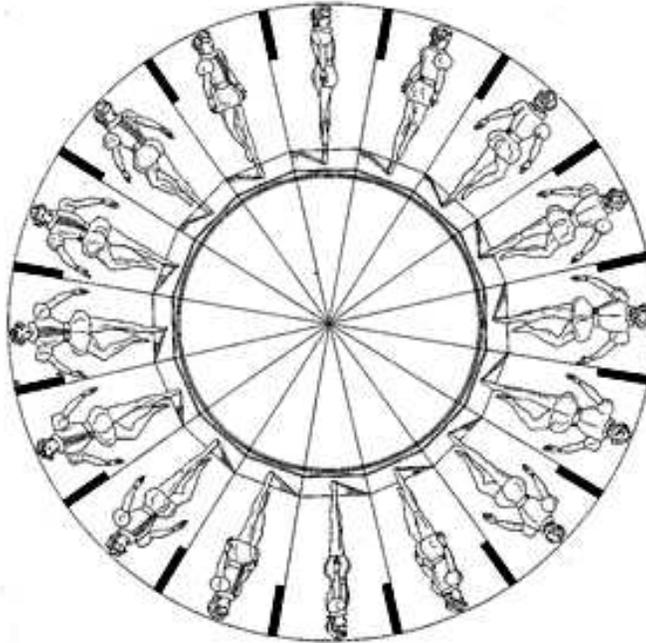
Fonte: Imagem de domínio público.

Sua teoria era a de que o olho humano seria capaz de reter imagens na retina por uma fração de segundo antes delas se desvanecerem, assim como acontecia com as primeiras fotografias, antes de serem fixadas nas placas de metal. A esse fenômeno ele também denominou “persistência da visão”.

Após propor esta teoria, Plateau criou o fenacístoscópio, que consistia de um disco composto por uma sequência de desenhos posicionados de pé nas extremidades do mesmo (Figura 9). Este disco era posicionado no centro de um eixo e, defronte a ele havia outro disco com frestas. Ao rotacionar o disco com os desenhos, obtinha-se a ilusão do movimento (Figura 10).

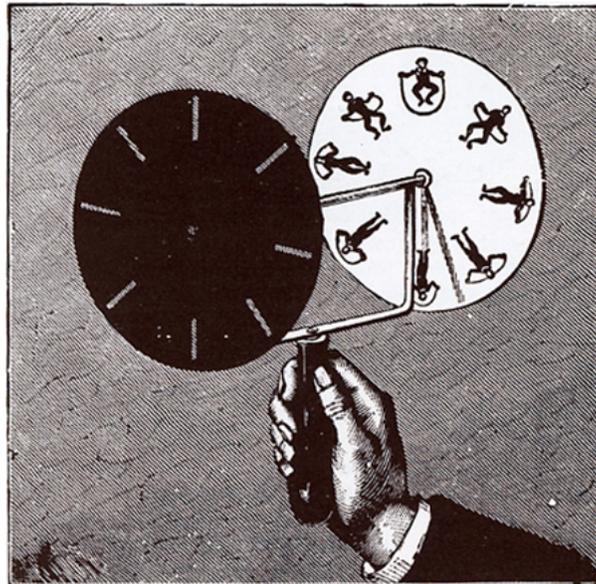
É importante salientar que essas primeiras animações rudimentares eram realizadas com ilustrações, já que a fotografia como nós a conhecemos ainda não existia naquele momento. Logo, as primeiras imagens em movimento não foram criadas a partir de imagens da natureza, mas do imaginário de artistas criativos, retratando personagens em ações cotidianas como dançar, caminhar ou saltar obstáculos.

Figura 9 - Ilustração de Joseph Plateau (1833), mostrando o posicionamento de imagens no disco do fenacístoscópio (1883).



Fonte: Imagem de domínio público retirada da Wikipédia
(<https://commons.wikimedia.org/wiki/File:Phenakistiscope.jpg>).

Figura 10 – Fenacístoscópio



Fonte: Imagem de domínio público, retirada de <http://uchihahyral.blogspot.com/2013/02/the-phenakistiscope.html>.

Ambas as teorias foram refutadas posteriormente, pois pesquisas mais recentes, como a de Anderson e Anderson (1993), indicam que o fenômeno da percepção humana de movimentos é causado por um fator psicológico. Essas ideias, porém, continuam sendo propagadas, pois se constituíram como “mitos” do cinema e da animação. Williams (2002, p. 13), por exemplo, advoga que a persistência da visão foi determinante para a existência da animação.

Embora a persistência da visão fosse um conceito que veio a ser contestado, as tecnologias a ele relacionadas se mostraram bastante proficuas em termos práticos. A partir dele, foram desenvolvidos os instrumentos óticos descritos anteriormente, os quais, por sua vez, transformaram-se em aparelhos de entretenimento, em brinquedos e, por fim, nos projetores. Atualmente, o festival Anima Mundi utiliza esses brinquedos como forma educação e popularização da animação.

Porém tais instrumentos possuíam como continência principal a limitação da quantidade de imagens que podia ser utilizada e a conseqüente necessidade de realização da animação em ciclos, para que os movimentos tivessem uma seqüência lógica para o espectador.

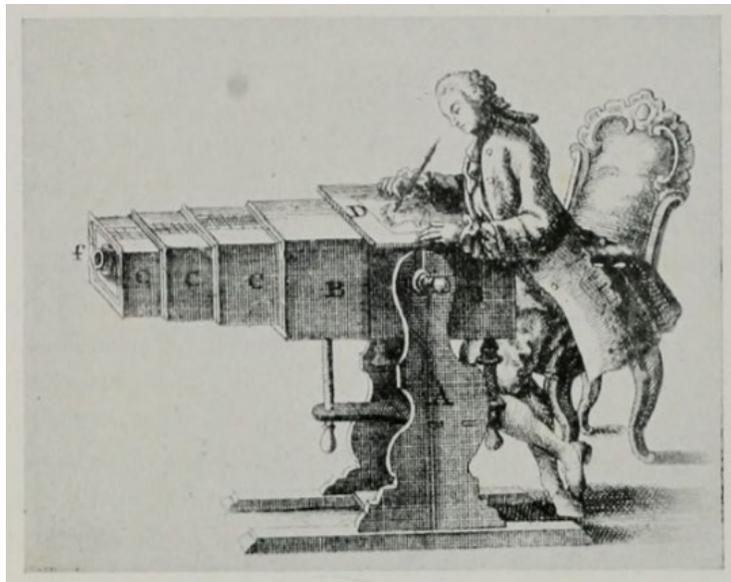
O primeiro projetor capaz de exibir imagens em seqüências e não apenas ciclos foi inventado por Émile Reynaud, no século XIX. Seu praxinoscópio foi responsável por dois marcos históricos. O primeiro foi sua própria invenção como um sistema capaz de superar a limitação da quantidade de imagens exibidas em seqüência. O segundo marco foi o da primeira exibição pública de animação, em 1892, com o teatro óptico e filmes desenhados, chamados de *pantomimes lumineuses* (BARBOSA JÚNIOR, 2001). A data da primeira exibição pública é comemorada todo dia 28 de outubro, com exibições de curtas em diversos locais do mundo, em um evento conhecido como o Dia Internacional da Animação.

Além do desenvolvimento do projetor, que teve um papel de base na história da animação e do cinema, a invenção e a consolidação da fotografia como um meio de registro de imagens foi necessário para o desenvolvimento deste último. O cinema herdou da fotografia suas técnicas de captação e de fixação de imagens, assim como sua linguagem. Posteriormente, ele partiria em direção a outro caminho.

O surgimento da fotografia se dá na primeira metade do século XIX, em um momento em que a ciência e a tecnologia apresentavam seus métodos à sociedade e ganhavam cada vez mais espaço devido às descobertas nestas áreas, que viriam a mudar a forma como os indivíduos se relacionariam com as máquinas e os meios de reprodução de imagens.

O século XIX foi um período na história da humanidade que deu ensejo ao surgimento de grandes invenções. Porém, a necessidade de registrar imagens fidedignas da natureza não surgiu nesse século, pois é o resultado de uma série de estudos e experimentos realizados desde o renascimento, que incluem não só técnicas artísticas, como a invenção da perspectiva por Brunelleschi, o *sfumato* de Leonardo da Vinci e o *chiaroscuro* de Rembrand (GOMBRICH, 1972, p. 228, 170, 333), mas também a invenção de dispositivos como as câmaras obscuras (Figura 11) (NEWHALL, 1949), as máquinas de silhuetas (ROSENBLUM, 1997) e as descobertas dos compostos químicos alterados pela ação da luz (NEWHALL, 1949).

Figura 11 - Gravura representando a câmara obscura e seu uso.



Fonte: Newhall (1949, p. 11).

Figura 12 - Máquina de silhueta.



Fonte: Roseblum (1997, p. 40).

Além disso, o fortalecimento de uma classe média europeia e norte-americana consumidora de entretenimento e ávida por retratos – até então, exclusivos da classe nobre – cria as condições econômicas ideais para o desenvolvimento da fotografia. Como aponta NEWHALL (1949, tradução nossa), “[...] A classe média queria retratos baratos; dispositivos mecânicos para eliminar a necessidade de longos treinamentos artísticos foram postos em suas mãos, de maneira que todo homem pudesse se tornar um pouco artista”.

O primeiro registro de uma imagem natural através da luz data de 1826, quando o francês Joseph Nicéphore Niépce, criou um processo rudimentar, chamado de heliografia, que permitia a impressão de imagens através da sensibilização, pela luz, de uma placa metálica e o uso de betume da Judeia (NEWHALL, 1949; GERNSHEIM; GERNSHEIM, 1965). Essa primeira “fotografia” levou oito horas para ser realizada e não era nítida, como se pode observar na Figura 13:

Figura 13 - A primeira fotografia, criada por Joseph Nicéphore Niépce em uma placa de estanho, em 1826.



Fonte: GERNSHEIM e GERNSHEIM (1952).

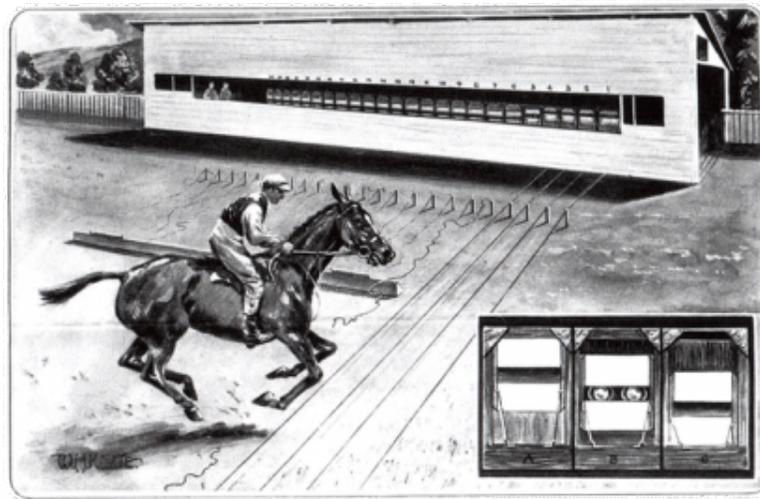
Segundo Newhall (1949), Pollack (1977), Rosenblum (1997) e Hannavy (2008), o desenvolvimento posterior da fotografia se deu por meio de uma série de inventos, a saber: o daguerreótipo, do francês Louis Daguerre, que inventou um processo de reprodução em placas metálicas mais rápido e nítido; o talbótipo, do inglês Henry Fox Talbot, que utilizou o papel como suporte, além de inventar um processo que utilizava imagens negativas para a reprodução; o cianótipo, do inglês John Herschel; o colódio úmido, inventado pelo também inglês Frederick Scott Archer, que reduziu o tempo de exposição; e, finalmente, a industrialização do filme flexível, pelo estadunidense George Eastman, fundador da Kodak.

Todas estas descobertas aconteceram do momento em que Niépce inventava sua heliografia, em 1826, até 1890. Entretanto, para que essas tecnologias pudessem ser colocadas a favor da imagem em movimento, foi necessário que outras barreiras fossem transpostas, como a velocidade de exposição do negativo à luz.

Os pioneiros no registro e na análise de movimentos foram Eadward Muybridge e Thomas Eakins nos Estados Unidos, Étienne Jules-Marey na França e Ottomar Anschütz na Alemanha. Muybridge e Eakins chegaram a se corresponder e ambos foram atraídos pelo

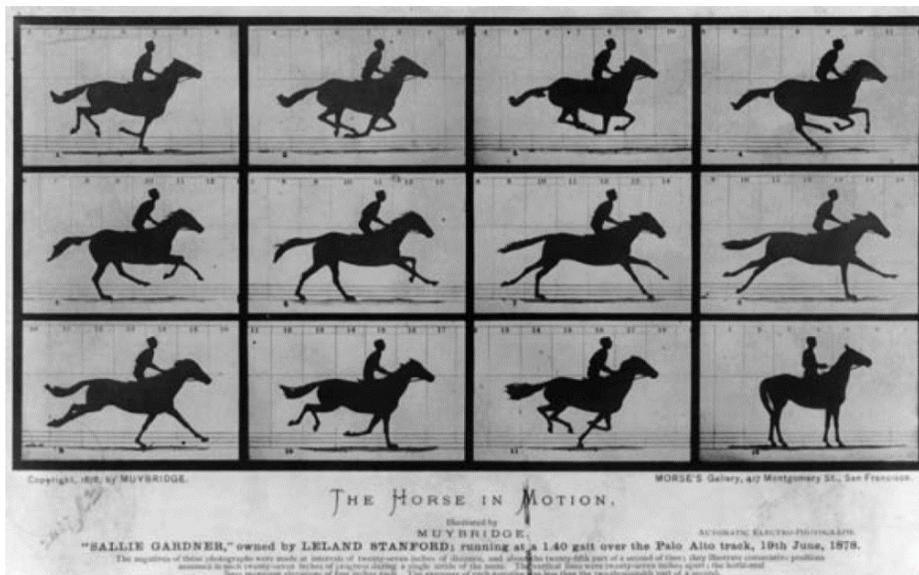
desafio de capturar cavalos em movimento (Figura 14, Figura 15 e Figura 16). Contudo, Muybridge foi o primeiro que conseguiu realizá-lo, através da criação de um conjunto de câmeras e disparadores (ROSENBLUM, 1997, p. 249).

Figura 14 - Dispositivo inventado por Muybridge para acionar as câmeras fotográficas e capturar imagens em movimento.



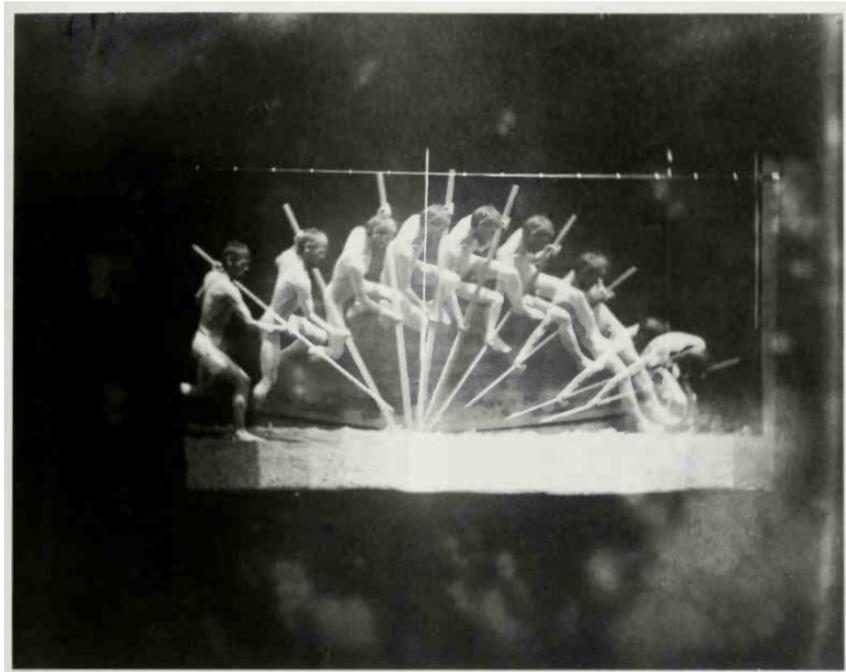
Fonte: <https://www.prestocentre.org/eadward-muybridge-horse-motion>. Acesso em: 18 Fev. 2018.

Figura 15 - Eadward Muybridge. O cavalo em movimento.



Fonte: Präkel (2010, p. 170)

Figura 16 – “História de um salto”. Fotografia de Thomas Eakins. 1844-45.



Fonte: Rosenblum (1997)

Ao registrar imagens em movimento, esses pioneiros, além de conseguirem estudar a progressão dos corpos no tempo, também possibilitaram que essas tecnologias fossem aplicadas em conjunto com o desenvolvimento do projetor – o último passo crucial para a existência da sétima arte.

O cinema e a animação possuem, portanto, a mesma origem e utilizam as mesmas tecnologias de projeção, embora não compartilhem uma mesma “essência”. Enquanto o primeiro está diretamente relacionado à imagem fotográfica, a linguagem da animação está relacionada ao desenho, ao cartum e à *charge*, assim como ao sequenciamento presente nos quadrinhos, e, em outro quadro de referência, à caracterização formal-visual e à movimentação própria do teatro de bonecos.

Essas influências implicarão em algumas características específicas das formas dos cartuns. Do desenho caricatural, do cartum e da *charge*, a animação herdará as deformações de corpo e a comicidade, que farão dela um instrumento para a narrativa cômica e a paródia. Estas características já serão notadas nos desenhos animados, desde o início do século XX. Os exemplos são variados: a *Betty Boop* (Figura 17), dos irmãos Fleischer, em que a personagem principal apresenta a forma de seu corpo estilizada enquanto sua cabeça

é desproporcional ao seu corpo; a personagem Olivia Palito, da animação *Popeye*; o personagem Mickey Mouse; o personagem Gajah, da animação *Tromba trem*, entre outros.

O exagero das formas, presente nas charges e nos cartuns, é projetado e aplicado ao design de personagens. Historicamente, essa característica pode ser encontrada em outras manifestações. A forma do personagem principal na caricatura de John Bull, de 1709 (Figura 18) na qual este é retratado como um mostrengo glutão – é similar àquela do personagem Bafo da Onça, nas animações de Mickey Mouse, ainda que, no caso específico da animação, ele seja um personagem animalizado.

Figura 17 - Frame da animação Betty Boop, 1933.



Fonte: *The old man of the mountain*. <https://www.youtube.com/watch?v=SoJkxNa6v14>. Acesso em: 17 Jan. 2018.

Figura 18 - Caricatura de John Bull, de 1709.



Fonte: Maurice e Cooper (1904, p. 8).

Dos quadrinhos e da literatura, a animação herdará os modos de narração, as situações inusitadas, a forma de contar histórias de universos fantásticos e, por diversas vezes, impossíveis de serem encontrados no mundo real e, especificamente, dos quadrinhos, a narrativa por meio de imagens. Esse tipo de narrativa exige do desenhista de quadrinhos ou do animador a habilidade para transmitir uma mensagem através dos elementos simplificados passíveis de serem reconhecidos pelo público, com base em seu repertório visual progressivo. A narrativa visual depende igualmente da cultura adquirida por um habitante de determinada região. Ao relatar a relação entre o público e os quadrinhos, Eisner (1999), afirma que:

As imagens sem palavras, embora aparentemente representem uma forma mais primitiva de narrativa gráfica, na verdade exigem certo refinamento por parte do leitor (ou espectador). A experiência comum e um histórico de observação são necessários para interpretar os sentimentos mais profundos do autor.

Algumas das primeiras animações foram baseadas em fábulas e pode-se inferir que isso se deu devido ao fato de que elas já estavam inseridas no universo cultural dos indivíduos, facilitando a conexão com o público. Essas histórias, diversas vezes, utilizaram animais como personagens para transmitirem uma lição de moral, recomendações sobre como o leitor deveria agir ou anedotas. O uso de personagens animalizados em animações

não é, portanto, apenas uma coincidência. Ele decorre das influências dessas narrativas. As fábulas de Esopo, o folclore brasileiro idem e as histórias indianas utilizam animais como personagens de suas histórias.

A fábula “A cigarra e a formiga”, de La Fontaine, com todas suas nuances, é um exemplo da criação desses universos com personagens animalizados cuja origem se deu na literatura. Sua adaptação para a animação utilizou os mesmos personagens. Esse uso dos personagens animalizados, no início do século XX, abriu caminho para a união de dois universos ficcionais: as fábulas e as animações.

No *stop-motion* “A cigarra e as formigas”, produzido pelo polonês Wladislav Starevicz em 1911, os personagens são animais reais cujos esqueletos foram reconstruídos e animados pelo artista (Figura 19). Neste caso, a história contada se vale destes personagens para transmitir uma lição de moral bem conhecida pelo público espectador: a necessidade de trabalhar para obter o sustento.

Figura 19 - Quadro de "A cigarra e as formigas", animação em *stop-motion* de 1911, criada por Wladislav Starevicz.



Fonte: <https://www.youtube.com/watch?v=mn4lnX3DDX8>. Acesso em: 31 Jan. 2018.

A relação entre as fábulas, os quadrinhos, e a animação é reforçada pelos elementos usados nas primeiras animações. Os primeiros curtas-metragens absorveram dos quadrinhos

suas características formais. Os balões de diálogo, as onomatopeias, os enquadramentos e a disposição das poses dos personagens, aliadas à inexistência do som na animação fazem com que o intercâmbio entre as linguagens seja frequente. O Gato Félix, personagem popular de Pat Sullivan, por exemplo, complementava suas ações, no início do século, através desses recursos gráficos, como pode ser visto no *frame* extraído de *Feline follies* (Figura 20).

Figura 20 - Balão de diálogo usado no “Gato Félix”.



Fonte: *Feline follies* (1919), <https://www.youtube.com/watch?v=7HskWL82GeQ>. Acesso em: 19 Fev. 2018.

Além disso, o teatro de bonecos (ou fantoches) também contribuirá para a formação de uma linguagem da animação. Esta forma de arte em movimento possui duas características que determinam sua existência: a estilização e a relação com os objetos cotidianos (GIESEN, 2019; POSNER; ORENSTEIN; BELL, 2014). A estilização é originária de sua história e suas relações com as manifestações religiosas. Segundo Posner et al. (2014), os primeiros bonecos foram utilizados para representar deuses na Alexandria e seus criadores os estruturavam de maneira que permitissem sua movimentação, para que estas entidades aparentassem ter vida.

O teatro de bonecos conecta-se também com outra tradição que será relacionada por Giesen (2019) à filosofia e por Posner et. al (2014) à cultura: o teatro de sombras. A utilização de sombras como atores de uma história remonta à cultura popular chinesa, que se valia dessa arte para transmitir suas histórias. Sua eficiência nessa empreitada foi tão

bem-sucedida, que ela se espalhou pelos diversos territórios chineses, quando da expansão das dinastias, para que o processo de comunicação entre as diferentes culturas pudesse ser concretizada. Ao mesmo tempo, houve momentos em que esta expressão cultural foi proibida, pois poderia disseminar um sentimento de revolta entre os camponeses.

A relação entre o teatro de sombras e a filosofia é exemplificada por Posner et. al (2014, p. 8) por meio da relação entre as diversas correntes filosóficas e destas com essa forma de expressão cultural. Os autores atravessam, de maneira breve, a filosofia grega (a alegoria platônica da caverna) e passam pelos questionamentos kantianos sobre a natureza da experiência humana, o existencialismo heideggeriano, entre outros, objetivando elevar o teatro de sombras a uma categoria cultural mais elevada.

A prática do teatro de sombras possibilitou a conexão entre a cultura oriental e a ocidental, quando passou a utilizar-se de silhuetas animadas para transmitir ideias e pensamentos. Apesar dessa conexão ter acontecido durante a idade moderna na Europa, seu expoente se deu através do trabalho da alemã Lotte Reiniger que, através da animação de silhuetas recortadas em papel, conseguiu, na década de 1926, realizar narrativas extensas utilizando esta técnica.

Reiniger não foi a única a explorar a técnica de animação através de fantoches. A combinação de formas antropomórficas e mecanismos para movimentá-los, que leva a um questionamento sobre o que caracteriza a vida e o que caracteriza o humano, havia sido explorado, no início do século XX, por animadores ingleses e franceses, ao relatarem suas fábulas.

A informalidade com que os personagens são criados por meio dessa técnica faz com que possuam “anatomias” bastante singulares, não encontradas em quaisquer outros indivíduos e, ao mesmo tempo, reduzidas aos elementos formais mais básicos, como as esferas e os círculos, tornando-se, por esta razão, universais, ou seja, capazes de casar identificação rápida e eficaz com o público.

The puppet achieves its elemental qualities of impersonality, unreality, and universality through the stylizations imposed upon it by its own limitations. It is a mistake to imagine that the more lifelike or natural a puppet can be, the more effective it is. Indeed, the opposite is often the case. A puppet that merely imitates nature inevitably fails to equal nature; the puppet only justifies itself

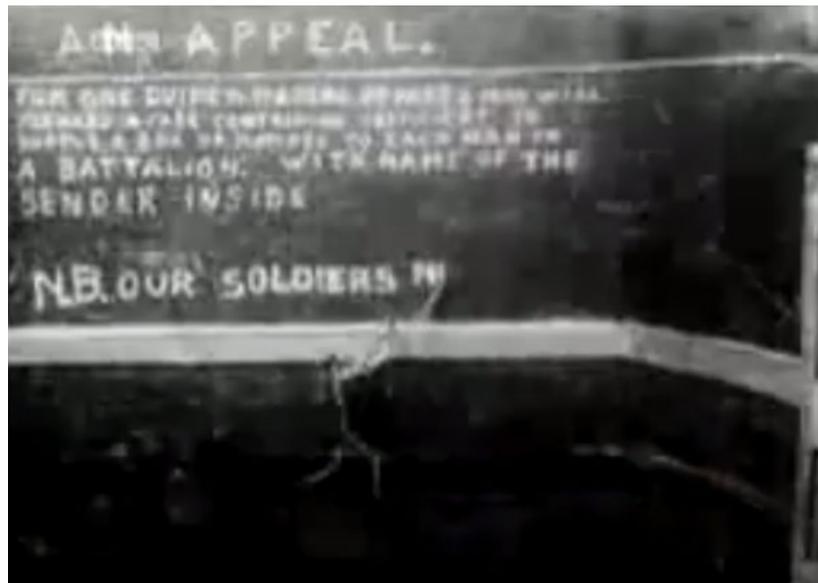
when it adds something to nature—by selection, by elimination, or by caricature. Some of the most effective puppets are the crudest (PUPPETRY, 2017).⁵

Estas artes – os quadrinhos, as *charges*, o teatro de fantoches – isoladamente não se constituem enquanto animação, mas, ao terem seus elementos transformados em imagens em movimentos por pioneiros como James Stuart Blackton, contribuíram para o surgimento de uma nova forma de contar histórias através de personagens.

Desde o início do século XX, diversos tipos de personagens foram criados para contar histórias: criados a partir de objetos do cotidiano como palitos de fósforo (*Matches – an appeal*, de Arthur Melbourne Cooper, 1899, Figura 21), animais dissecados (*The grasshopper and the ant*, de Ladislav Starewicz, 1911, Figura 22), desenhos dos mais diversos tipos, como animais e seres humanos, ou seja, tudo aquilo que faz parte do conhecimento humano ou das invenções dele resultantes (como os seres extraterrestres), transformam-se em personagens para animação. Para que estes personagens possam ser objetos de identificação com o público, é necessário atentar para que algumas características estejam presentes nos personagens, sendo sua humanização a principal.

⁵ “A marionete atinge suas qualidades elementares de impessoalidade, irrealidade e universalidade através das estilizações impostas a ele por suas próprias limitações. É um erro imaginar que, quanto mais realista ou natural uma marionete puder ser, mais eficaz ele será. De fato, o oposto é frequentemente o caso. Uma marionete que meramente imita a natureza inevitavelmente falha em igualar-se à natureza; a marionete só se justifica quando acrescenta algo à natureza – por seleção, por eliminação ou por caricatura. Algumas das marionetes mais eficazes são as mais cruéis” (PUPPETRY, 2017, tradução nossa).

Figura 21 - *Frame da animação Matches – an appeal*, de Melbourne Cooper, 1899.



Fonte: <https://www.youtube.com/watch?v=HD9NwzFNLY&index=49&t=0s&list=LLjU5teWn934I7ZDV8g7oo4Q>. Acesso em: 17 Jan. 2018.

Figura 22 - *The grasshopper and the ant*, Ladislav Starewich, 1911.



Fonte: <https://www.youtube.com/watch?v=mn4lnX3DDX8&index=46&t=385s&list=LLjU5teWn934I7ZDV8g7oo4Q>. Acesso em: 17 Jan. 2018.

Ser humano é um critério para se contar histórias. Porém, a forma como estas são transmitidas não precisa ser feita através de um indivíduo. No caso da animação, é comum que se utilizem personagens estilizados que possuem um comportamento humano.

Este modo de agir estará presente nos personagens que, de fato, são representações humanas, como homens, mulheres e crianças, mas precisará ser mais evidenciado naqueles que não são, como animais e objetos do cotidiano. Cada qual possui suas vantagens e desvantagens na narração de histórias.

Personagens humanos não apresentam problemas de identificação pelo público. Animais, porém, precisarão apresentar características humanas explícitas ou, como esperado em animação, exageradas. Isto exigirá do animador o reconhecimento dos elementos mínimos necessários para que uma pose, uma feição, um gesto, uma ação, realizados por um indivíduo sejam reconhecidos pelo espectador.

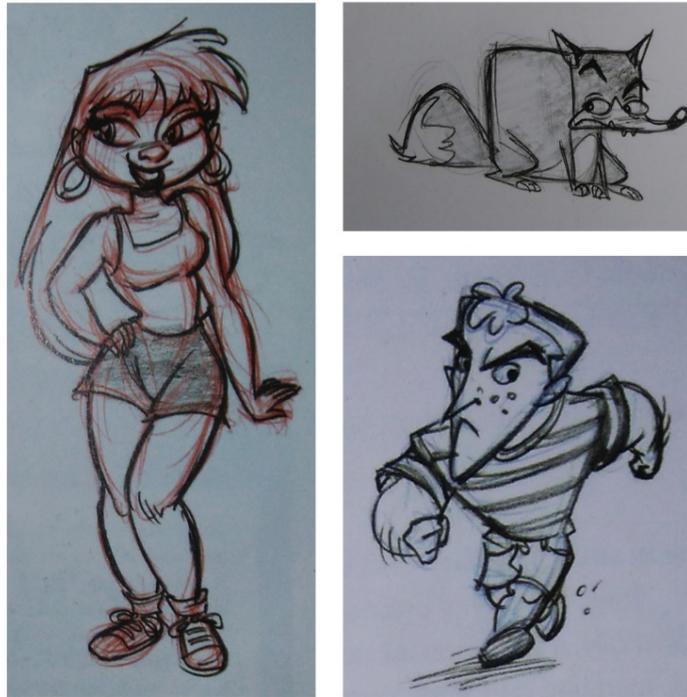
Entretanto, esta característica dos animais personagens que, em um primeiro momento, aparenta ser uma desvantagem, pode converter-se no oposto. Pelo fato de não se referirem a um sujeito específico, eles podem ser associados a qualquer indivíduo. McCloud (2005, p. 34-46) argumenta que personagens cujas características são mais icônicas, ou seja, menos realistas, possuem a capacidade de ampliar a identificação com o seu público, fazendo com que este se coloque em seu lugar. Segundo ele, essa “[...] é a razão principal do nosso fascínio por desenhos animados, embora outros fatores como simplicidade e características infantis de muitos personagens de desenhos animados também desempenhem um papel”.

Personagens são projetados de acordo com seu papel na narrativa. Designers de personagens, em geral, seguem algumas fórmulas testadas anteriormente em longas e curtas. Algumas delas são bem conhecidas pelos profissionais da área, muitas se relacionando com formas geométricas básicas, como círculos, quadrados e triângulos. Bancroft (2006a, p. 33-35) apresenta essa relação. Segundo ele,

Circles evoke appealing, good characters, and are typically used to connote cute, cuddly, friendly types. [...] Squares usually depict characters who are dependable or solid, or play the heavy. [...] Triangles easily lend themselves to more sinister, suspicious types and usually represent the bad guy or villain in character designs.⁶

⁶ Círculos evocam atração, bons personagens, e são tipicamente usados para conotar tipos bonitinhos, fofinhos e amigáveis. [...] Quadrados geralmente retratam personagens que são confiáveis ou sólidos, ou jogam

Figura 23 - Personagens e formatos relacionados às formas geométricas: círculo, quadrado e triângulo.



Fonte: Bancroft (2006a).

Podemos afirmar que a forma de um personagem influencia seus movimentos e a maneira pela qual ele age no contexto do filme, mas não determina suas ações. Um humano personagem, mas sem mãos e dedos pode, no contexto da história, “segurar” um objeto e mover-se tal qual um ser humano real (Figura 24), ou seja, a história supõe liberdade quanto ao “formato” do personagem, assim como a verossimilhança da movimentação desses personagens pode ser percebida mesmo a partir de seus elementos inverossímeis.

pesado. [...] Os triângulos se prestam facilmente a tipos mais sinistros e suspeitos e geralmente representam os designs de personagens do vilão ou bandido (BANCROFT, 2006, tradução nossa).

Figura 24 - *Frame* da animação *Pingu and the birthday*, com um personagem comemorando seu aniversário.



Fonte: *Pingu and the birthday*. Official Pingu Youtube Channel. Acesso em: 21 Fev. 2018.

No contexto da indústria do entretenimento, a animação caminha ao lado do cinema já que ambas são formas narrativas que focam principalmente na natureza humana em suas obras e, embora suas produções tenham etapas diferentes, seus caminhos históricos se complementam.

Elas ocupavam simultaneamente os mesmos espaços: inicialmente os *vaudevilles* e cafés, e, posteriormente, os *nickelodeons*, espaços baratos (o ingresso custava um níquel, daí seu nome) para onde a classe média se dirigia para assistir exibições das novas artes em movimento (BECKERMAN, 2003). Foi através da demanda por novas histórias atuadas por personagens animadas, bem recebidas e, posteriormente, requisitadas pelo público, que a animação começou a se consolidar enquanto produto audiovisual.

Entretanto, a animação e o cinema não possuem o mesmo reconhecimento pela indústria do audiovisual. O cinema geralmente classifica a animação como um de seus gêneros, colocando-a, lado a lado, com o terror, a ação, o suspense, a comédia, entre outros. A indústria cinematográfica não consegue compreender que a narrativa animada é igualmente realizada em qualquer um dos gêneros cinematográficos e, ainda assim, é considerada animação.

Os fatores que levaram o cinema a ter um *status* privilegiado sobre a animação na indústria do entretenimento são muitos, mas, historicamente, pode-se afirmar que, enquanto o primeiro se consolidou enquanto uma linguagem e encontrou intelectuais dispostos a

refletir sobre a narrativa naquele meio e a desenvolvê-lo, a animação não teve o mesmo tratamento, sendo vista como uma variação “infantilizada do cinema”. Intelectuais que trabalhavam na criação de filmes, como Eisenstein e D. W. Griffith (MASCARELLO, 2006, p. 46), atuavam também na contextualização da linguagem cinematográfica. Assim, na primeira década do século XX:

A indústria cinematográfica começou a ganhar respeitabilidade, dirigindo uma parcela cada vez maior do público para teatros luxuosos e mais caros. Poucos anos depois, em 1917, a maioria dos estúdios norte-americanos já se localizava em Hollywood e a duração dos filmes tinha aumentado de um rolo para 60 ou 90 minutos” (MASCARELLO, 2006, p. 49).

Nas décadas seguintes, diversos estilos cinematográficos surgiram: o expressionismo alemão, o impressionismo francês, a montagem soviética, o surrealismo, o neorealismo italiano, dentre outros, que diversificariam a linguagem cinematográfica e levariam a discussões sobre temáticas regionais e estilos individuais de direção.

O cinema se consolidou rapidamente enquanto uma indústria tanto na Europa quanto nos Estados Unidos. No início do século XX, além do aumento na duração dos filmes, as primeiras produtoras já existiam e competiam entre si (MASCARELLO, 2006, p. 38-40).

Enquanto isso, a animação ainda procurava uma linguagem em que se apoiar. O processo de criação de animação era deveras artesanal, o que acarretava em produções longas e custosas, em contraste com as produções cinematográficas, que levavam apenas alguns meses (e, algumas vezes, apenas algumas semanas) para serem lançadas. Mesmo após a sistematização do processo de animação, a produção de curtas animados ainda era demorada e custosa e, posteriormente, essa condição se reproduziu na produção de longas-metragens de animação que, até hoje, levam em média quatro anos para serem produzidos.

Assim como seu companheiro de telas, a animação também contou com pioneiros que pensaram em maneiras de tornar esse entretenimento eficiente, rentável para seus produtores e atraente para o público. Entretanto, os procedimentos adotados pela indústria apenas começaram a serem delineados em 1914, com Raoul Barré, John Randolph Bray e Earl Hurd, três estadunidenses criadores de cartuns, interessados em produzir curtas em série assim como se fazia no cinema. Eles patentearam diversos dispositivos – o sistema de furos nas folhas de papel para manter o registro dos desenhos, a repetição nos fundos dos desenhos e a utilização de folhas de celulose transparente – que facilitaram a produção

animada, além de implementarem um sistema de divisão do trabalho em atividades especializadas (BECKERMAN, 2003, p. 12-17).

Os curtas de animação eram comumente exibidos antes do filme principal, o que reforçava a impressão de subordinação desta técnica ao cinema. Devido ao desenvolvimento tardio, a distribuição dos curtas produzidos pelos estúdios surgidos no início do século XX ficou dependente dos distribuidores de filmes, que começavam a aumentar a quantidade de rolos por produção, criando narrativas cada vez mais complexas e experimentando novos métodos de contar histórias. Dentre as técnicas que fizeram parte da nova forma de contar histórias, destacam-se: a mudança de enquadramentos, o corte mais curto das cenas, a aproximação da câmera dos rostos dos atores e a montagem dos filmes.

Enquanto buscavam uma linguagem própria dos *cartoons*, os estúdios se apropriavam das convenções cinematográficas, aplicando-as diretamente sobre os curtas animados e, posteriormente, aos longas. A diferença de linguagem entre as produções da década de 1910 e as da década de 1930 são visíveis no grafismo dos personagens, na maneira como se movimentam e, perceptivelmente, na linguagem que simula a câmera presente nas cenas animadas (Figura 25).

Figura 25 - Aproximação da câmera no curta dos “Três porquinhos” é uma apropriação da linguagem cinematográfica.



Fonte: “Os três porquinhos” (HAND, 2009).

A linguagem da animação pode ser compreendida a partir de diversos aspectos: os movimentos dos personagens, sua forma, o tipo de narrativa proposta e suas relações culturais, e o momento histórico (que dependerá da própria história do país e seu papel na história mundial), dentre outros. Algumas características foram impostas pelo poderio econômico e cultural de cada nação. Outras foram ressaltadas pela necessidade de cada país de disseminar sua cultura em contraposição ao domínio cada vez mais forte das produtoras de Hollywood. O contraste entre as produções de cada país se dará através dos estilos, das formas e das histórias contadas, mas não do movimento.

Na próxima seção serão apresentados alguns exemplos de produções nacionais e como cada uma delas se diferenciou de seus “concorrentes”, em especial dos Estados Unidos. Diversos países passaram por questionamentos internos que os levaram a produzir animações com características formais locais. Entretanto, o estilo de movimento que impôs sua soberania sobre os demais foi, naturalmente, o estadunidense. Mas, mesmo este foi alvo de críticas e releituras dentro de seu próprio território, como veremos a seguir.

2.2 Desenvolvimento da animação: período e estilos

A animação surgiu na França e de lá se espalhou para os demais países europeus, atravessando o Atlântico, onde encontrou na América um terreno fértil para sua proliferação. Logo depois, desenvolveu-se na Ásia, na Oceania e, por fim, chegou à África (BENDAZZI, 2016a). As características dos primeiros curtas eram bastante similares naqueles países, que ousaram experimentar a arte e a técnica de dar vida a cartuns. Não havia um estilo definido, já que os primeiros desenhos animados procuravam, ainda sem referências, apenas dar vazão ao fascínio das primeiras experimentações na nova forma de narrar histórias.

Mesmo sem um estilo definido, essas primeiras realizações já mostravam que a animação era a arte da fantasia, dos universos fantásticos, da possibilidade de colocar na mesma cena seres que viveram em épocas diferentes, como um ser humano e um dinossauro (como o fez Winsor McCay), de fazer com que um desenho tivesse vida. Foi a partir desses primeiros experimentos que a consciência de uma linguagem da animação começou a se formar.

O desenvolvimento dessa linguagem pode ser compreendido segundo características regionais, técnicas, econômicas, por profissionais, por estúdios e de maneira temporal (esta última, por vezes confusa, devido à justaposição entre a produção de alguns profissionais que desenvolveram estilos específicos e determinado período histórico). Independentemente do ponto de vista adotado, as publicações que descrevem o início da história da animação possuem um ponto em comum: uma história que se inicia diversa e é rapidamente adotada pela cultura estadunidense, sendo que logo os estilos e as técnicas desenvolvidas nos Estados Unidos são adotados (ou impostos) pela indústria ali constituída e difundido como uma “história oficial”.

Não é coincidência, portanto, que tantas descobertas no século XIX sejam creditadas aos franceses. A animação, o cinema e a fotografia possuem suas raízes históricas neste país que, devido à sua cultura cosmopolita, ensejou esse tipo de avanço tecnológico. E tampouco é espantoso que a indústria da animação se desenvolva nos Estados Unidos, país que passa progressivamente a ocupar uma posição hegemônica no século XX e no qual se desenvolve uma poderosa indústria cinematográfica.

A semente da animação que, inicialmente brota na França, logo se difunde para os demais países europeus, em especial a Grã-Bretanha e se diversifica em diversas técnicas (BARBOSA JÚNIOR, 2001, p. 36-41). Estas não nascem ou são exclusivas de países específicos. Como resultados de experimentações com desenhos, marionetes, bonecos e objetos do cotidiano, tornam-se comuns a diversas culturas. O que as diferenciara são as influências narrativas das histórias locais e os recursos visuais de cada cultura.

Assim, é possível encontrar a técnica do *stop-motion* em países com culturas diferentes, como a França, a Inglaterra, a Rússia, e a Alemanha. O mesmo se dá com o desenho animado, que floresce nos Estados Unidos, embora sua semente seja europeia, com os trabalhos de Emile Reynauld. Os estilos gráficos utilizados por cada nação, porém, diferem entre si e, igualmente, com as narrativas.

É notável, por exemplo, a influência dos quadrinhos e da *bande dessinée* na França e nos Estados Unidos (BENDAZZI, 2016a, p. 38), representados em animações como o Gato Félix, o marinheiro Popeye e o gaulês Asterix (Figura 26). As primeiras séries animadas foram baseadas em quadrinhos, como descreve Barrier (2003, p. 11):

“[...] The first attempt at such a series was, like McCay’s Little Nemo cartoon, **based on a comic strip**. The French cartoonist Emile Cohl, who in the preceding four years had made dozens of short films—some containing animated drawings—for Gaumont and Pathe, came to the Eclair studios in Fort Lee, New Jersey, in the fall of 1912 to make a series **based on** George McManus’s **comic strip**, “The Newlyweds.” It began appearing in theaters in March 1913, by which time McCay had made only his first two films [...]”⁷

Figura 26 - *Frame da animação Asterix, le gaulois, 1967.*



Fonte: Goossens (1967).

No outro extremo do planeta, na China, a pintura, as lendas locais e a tradição cultural darão a tonalidade às animações, através do uso da pintura para dar vida a personagens (Figuras 27 e 28). De maneira análoga, a Índia trará sua tradição religiosa e as suas histórias locais para o desenho animado (Figura 29). No Japão, acontecerá a transposição direta do *anime* (um estilo de quadrinhos japonês) para o cartum animado (Figura 30). Nos demais países asiáticos vizinhos a estas grandes nações influenciadoras (econômicos e culturais) a procura por histórias nas culturas locais será tema das produções animadas.

⁷ “A primeira tentativa de tal série foi, como o cartum ‘Little Nemo’ de McCay, baseado em uma tirinha de quadrinhos. O cartunista francês Emile Cohl, que nos quatro anos anteriores fez dúzias de curtas-metragens – alguns contendo desenhos animados – para a Gaumont e a Pathe, veio para os estúdios Eclair, em Fort Lee, Nova Jérsei, no outono de 1912, para fazer uma série baseada na tirinha em quadrinhos de George McManus, ‘Os Recém-Casados’. Ela começou a aparecer nos cinemas em março de 1913, e naquela época McCay tinha feito apenas seus dois primeiros filmes” (BARRIER, 2003, tradução nossa”).

Figura 27 - *Frames* da animação pintada em aquarela "Onde está mamãe?", de Te Wei (1960).



Fonte: <https://www.youtube.com/watch?v=N3kZCein4BQ>. Acesso em: 31 Jan. 2018.

Figura 28 - *Frame* da animação "Zhu Baizhe come a melancia", de Wan Guchan (1958).



Fonte: <https://www.youtube.com/watch?v=tb2PxmGhxEE>. Acesso em: 31 Jan. 2018.

Figura 29 - *Frame* da animação indiana *The banyan deer*, apresentando o Rei (1957).



Fonte: (SWAMY, 1997)

Figura 30 - *Frame* da Animação "A viagem de Chihiro", de Hayao Miyazaki (2001).



Fonte: <http://site.studioghibli.com.br/wp-content/uploads/2009/04/Chihiro02.jpg>. Acesso em: 31 Jan. 2018.

Na América Latina, houve um esforço para resgatar as identidades culturais locais. No Brasil, as animações abordam o folclore nacional (Figura 31), a vida cotidiana (Figura 32), a crítica ao imperialismo estadunidense (Figura 33) e seu esforço para se firmar como um país produtor de conteúdo e não apenas consumidor de material estrangeiro.

Figura 31 - *Frame* da animação "Sinfonia amazônica", com o curupira como personagem (1951).



Fonte: <https://www.youtube.com/watch?v=176ucJPE2ao>. Acesso em: 31 Jan.2018.

Figura 32 - *Frame* da animação *Guida*, de Rosana Urbes, 2016.



Fonte: <https://www.youtube.com/watch?v=c5xB5b3dQK8>. Acesso em: 31 Jan. 2018.

Figura 33 - *Frame* da animação *Meow*, de Marcos Magalhães, 1981.



Fonte: <https://www.youtube.com/watch?v=q6p1XNokmVk>. Acesso em: 22 Fev. 2018.

Bendazzi (2016b) mostra que os demais países latino-americanos, como o México, a Argentina e a Colômbia, sofreram diretamente a influência estadunidense, mas souberam, através do esforço de seus produtores, promover suas próprias histórias e culturas. Destacaram-se, no México, *Crônicas del Caribe* (1982 – Figura 34), do diretor porto-riquenho Francisco López, que relata a invasão espanhola na América Caribenha e o média-metragem colombiano *El pasajero de la noche* (1991 – Figura 35), que critica a falta de liberdade do homem moderno.

Figura 34 - *Frame* da animação *Crônicas del Caribe*, de Francisco López, 1982.



Fonte: https://www.youtube.com/watch?v=FpdqR_GiInM&t=1371s. Acesso em: 22 Fev. 2018.

Figura 35 - *Frame* da animação *Passajeros de la noche* (1991).



Fonte: <https://www.youtube.com/watch?v=8J5HIEbj2-k>. Acesso em: 22 Fev. 2018.

No continente africano, o desenvolvimento de produções animadas foi muito mais tardio que no restante do planeta. As primeiras obras somente foram iniciadas após a década de 1950. Isso aconteceu devido ao contexto histórico de exploração do povo africano, por centenas de anos. Após a independência daqueles países, as ex-metrópoles assumiram o controle dos meios de comunicação, em especial das emissoras de televisão, passando a transmitir suas próprias produções, sem relação com as raízes históricas e culturais de cada um dos países da África, dando continuidade à colonização e ao domínio daquelas populações através da comunicação e da cultura.

Assim como na América Latina, há esforços locais no sentido de resgatar a história e a identidade cultural daqueles países através da animação. Seguindo esta direção, destacam-se algumas obras como os filmes *Mumtaz* – que denunciou a corrupção no Egito – e *Racheltjie de Beer* (1982), o primeiro longa realizado na África do Sul (BENDAZZI, 2016b).

Em resumo, é necessário observar que o contexto cultural local influencia as produções de animação, mas não necessariamente a forma como os personagens se movimentam.

A divisão histórica da animação é naturalmente bastante tendenciosa e influenciada por autores nascidos nos Estados Unidos, visto que este é o país que atualmente mais produz animações e as distribui no mundo. Segundo o Anuário Estatístico do Cinema Brasileiro de 2015 (2016, p. 12), dentre os vinte títulos de maior bilheteria nos *cinemas brasileiros* naquele ano, cinco eram animações, sendo todos originários dos Estados Unidos, o que demonstra a forte influência da cultura daquele país no Brasil (Tabela 1). Isto não é uma exclusividade do nosso país. Na Europa, o volume de animações estadunidenses chega a 71,6% das bilheterias (PUMARES et al., 2015, p. 11). Isto também se reflete na quantidade de publicações sobre animação.

Tabela 1 - *Ranking* dos vinte títulos de maior bilheteria.

Nº	TÍTULO	GÊNERO	PAÍS	DISTRIBUIDORA	DATA DE LANÇAMENTO	SALAS NO LANÇAMENTO	PÚBLICO 2015	RENDA (R\$) 2015
1	Vingadores: A Era de Ultron	Ficção	Estados Unidos	Disney	23/04/2015	1.356	10.129.071	146.184.931,00
2	Velozes e Furiosos 7	Ficção	Estados Unidos	Universal	02/04/2015	1.046	9.857.946	142.465.883,37
3	Minions	Animação	Estados Unidos	Universal	25/06/2015	1.084	8.912.094	119.998.320,79
4	Cinquenta Tons de Cinza	Ficção	Estados Unidos	Universal	12/02/2015	1.087	6.685.086	87.741.026,57
5	Jurassic World: O mundo dos dinossauros	Ficção	Estados Unidos	Universal	11/06/2015	1.001	6.356.559	90.707.161,19
6	Star Wars: Episódio VII - O despertar da Força	Ficção	Estados Unidos	Disney	17/12/2015	1.505	5.558.321	90.448.267,00
7	Jogos Vorazes: A esperança - O final	Ficção	Estados Unidos	Paris	19/11/2015	1.718	4.392.977	63.524.902,67
8	Cinderela	Ficção	Estados Unidos	Disney	26/03/2015	931	4.199.697	50.079.019,00
9	Divertida Mente	Animação	Estados Unidos	Disney	18/06/2015	879	3.780.325	45.644.493,00
10	Loucas pra Casar	Ficção	Brasil	Downtown/Paris	08/01/2015	604	3.726.547	45.688.069,53
11	Bob Esponja: Um Herói Fora d'água	Animação	Estados Unidos	Paramount	05/02/2015	818	3.719.487	48.261.491,00
12	Vai que Cola - O Filme	Ficção	Brasil	H2O Films	01/10/2015	636	3.307.837	41.803.908,21
13	Hotel Transilvânia 2	Animação	Estados Unidos	Sony	24/09/2015	577	3.272.742	42.073.475,85
14	Os Pinguins de Madagascar	Animação	Estados Unidos	Fox	15/01/2015	911	3.133.452	40.478.965,00
15	Homem-Formiga	Ficção	Estados Unidos	Disney	16/07/2015	860	2.927.606	41.539.259,00
16	A Série Divergente - Insurgente	Ficção	Estados Unidos	Paris	19/03/2015	1.101	2.819.817	39.037.020,04
17	Missão impossível - Nação secreta	Ficção	Estados Unidos	Paramount	13/08/2015	947	2.711.422	37.270.687,00
18	Uma Noite no Museu 3: O Segredo da Tumba	Ficção	Estados Unidos, Reino Unido	Fox	01/01/2015	610	2.707.296	30.904.464,00
19	O Último Caçador de Bruxas	Ficção	Estados Unidos	Paris	29/10/2015	1.050	2.658.580	34.438.709,33
20	Meu Passado Me Condena 2	Ficção	Brasil	Downtown/Paris	02/07/2015	618	2.639.935	32.941.689,75

Fonte: (OBSERVATÓRIO BRASILEIRO DO CINEMA E DO AUDIOVISUAL, 2016).

Assim, as diversas publicações sobre animação apontam divisões históricas e estilísticas, levando em consideração a realidade estadunidense e, por vezes, desconsiderando os demais países produtores, como o Reino Unido, a França, a Rússia, os países latino-americanos, os países africanos, dentre outros. Autores como Barrier (2003), Thomas e Johnston (1981), Richard Williams (2002) e Chong (2008) descrevem uma sequência histórica que se inicia com a exibição da primeira obra de animação do francês Emile Reynaud e que, de maneira quase que imediata, foi absorvida e desenvolvida na América do Norte, onde nomes como James Stuart Blackton e Winsor McCay são apontados como sendo os pioneiros da produção de desenhos animados, seguidos por Walt Disney, pelos irmãos Fleischer, por Pat Sullivan e por outros animadores.

Igualmente é enfatizada a consolidação dos grandes estúdios e distribuidoras (como a Warner) estadunidenses, com destaque para a fundação da UPA (United Productions of America), o trabalho do diretor Chuck Jones, o surgimento da computação gráfica, a criação da Pixar e o desenvolvimento da animação para Web.

Embora esses autores não estejam incorretos, pode-se dizer que a informação por eles publicada possui diversas lacunas, pois exclui boa parte da produção de animação realizada em outros países. Essas lacunas são preenchidas pelo pesquisador italiano Giannalberto Bendazzi, que publica *Animation: a world history* (BENDAZZI, 2016a), uma série de livros sobre animação que abordam a divisão desta técnica em algumas épocas, apresentando as descobertas de estilos e de técnicas, além de procurar indicar as principais produções nacionais em cada continente.

Ainda que fortemente influenciada pela história da animação estadunidense, sua obra amplia o panorama geral, até então existente, sobre a participação de outras nações nesta linha do tempo. A divisão histórica proposta por Bendazzi (2016a) é definida em seis períodos, cujos marcos principais estão resumidos na Tabela 2:

Tabela 2 - Divisão dos períodos históricos da Animação.

Períodos	Intervalo	Fatos relevantes
Primeiro	1892 – 1908	Cobre as primeiras experimentações até a exibição de <i>Fantasmagorie</i> , de Émile Cohl (1908).
Segundo	1908 – 1928	Período dos primeiros curtas silenciosos até a exibição do primeiro curta sonorizado, <i>Steamboat Willie</i> (1928).
Terceiro	1928 – 1951	O “Período de ouro” da animação, quando Walt Disney domina o mercado de produção neste gênero.
Quarto	1951 – 1960	Engloba o período que vai da exibição de Gerald McBoing Boing, em 1951, até a constituição do Festival de Annecy, na França.
Quinto	1960 – 1991	Chamado de “Os três mercados”, compreende um período de estabilidade durante a Guerra Fria. Cada um dos mercados corresponde ao que convencionou chamar a cultura de massas (TV), a cultura média (os longas feitos para o Cinema) e a alta cultura (os curtas experimentais dos festivais).
Sexto	1991 – 2015	O período posterior à queda dos regimes comunistas e que viu a expansão do mercado global de animação e a participação cada vez maior da Ásia nele.

Fonte: Baseado em Bendazzi (2016a, p. 2).

Em termos culturais e econômicos, os Estados Unidos conseguiram, de fato, impor suas técnicas e seu modelo de produção de cinema e de animação, devido aos fatores históricos relevantes que aconteceram na primeira metade do século XX e que definiram seu papel dominante país na geopolítica mundial. Dentre estes, destacam-se: as duas grandes Guerras Mundiais ocorridas na Europa (então, o centro do poder político e econômico mundial), que enfraqueceram os países deste continente e deslocaram o eixo econômico para a América; a sua rápida industrialização; a constituição de uma classe média consumidora de produtos industrializados (HOBSBAWN, 1995, p. 101); e o estímulo ao desenvolvimento de uma indústria do audiovisual que foi inicialmente protegida por mecanismos que dificultavam a entrada de produções estrangeiras naquele país (MASCARELLO, 2006).

Este contexto e o fato de possuir uma sociedade cujo modelo de crescimento é baseado no consumo foram, portanto, determinantes para que os Estados Unidos produzissem avanços consideráveis na indústria do audiovisual, em especial na animação. Após a Primeira Guerra Mundial, seu predomínio econômico já era evidente:

[...] se era possível responsabilizar, ao menos parcialmente, as perturbações da Europa na Guerra e no pós-guerra ou, pelo menos nos países beligerantes da Europa, pelos problemas econômicos ali ocorridos, os EUA tinham estado muito distantes do conflito, embora por um curto e decisivo período tivessem se envolvido nele. Assim, longe de perturbar sua economia, a Primeira Guerra Mundial, como a Segunda, beneficiou-os espetacularmente.[...] (HOBSBAWN, 1995)

Além dos fatores anteriormente descritos, os estúdios existentes naquele país passaram a focar na qualidade das produções ofertadas ao público. Isto ocorreu devido à maior exigência do público, à crise que abateu os estúdios, no início dos anos 20, e à necessidade de suprir a demanda das distribuidoras de filmes, como a Paramount e a Pathé (BARRIER, 2003). Uma das soluções encontradas para isto foi a criação de séries animadas que permitiram que os estúdios pudessem sobreviver fornecendo cartuns animados, de maneira frequente, para estas distribuidoras.

Através da experiência adquirida através da criação dessas séries, os estúdios passaram a se preocupar com diversas características das animações, como os enquadramentos, o traço, a forma dos personagens e, principalmente, as características de seus movimentos, que foram gradualmente sendo adaptadas às necessidades da produção

em larga escala e à busca por uma maior empatia do público, sendo que um dos caminhos para alcançar estes objetivos foi sua naturalização gradual. Um dos estúdios que trilhou este caminho e definiu algumas “regras” para a naturalização dos movimentos foi o de Walt Disney. Thomas e Johnston (1981, p. 37) relatam as orientações que recebiam de Disney no sentido de criar personagens cartuns cujos movimentos tivessem como referências atuações reais:

[...] “*Caricature and exaggeration*” were two favorite words to stimulate the animator’s approach to his scene. These words could be misinterpreted as a request for wild, uncontrolled action, but that course always ended up with, “Look, you’re not getting the idea of what we’re after here!” The action had to be based on **realism**, had to fit the story situation, put over the point of the scene, and be in character with other things being done in different scenes[...].⁸

Assim, de personagens que se movimentavam de maneira desordenada – por vezes ininteligíveis, como nas primeiras animações do “Gato Felix”, de Pat Sullivan (Figura 36) – os cartuns passaram por transformações formais e semânticas no período que vai dos anos 1900 até 1930, que fizeram com que uma linguagem própria (e em constante mutação) pudesse surgir e evoluir até chegar às convenções adotadas atualmente no meio da animação de personagens.

Segundo Thomas e Johnston (1981), no início não existia controle por parte dos animadores sobre os movimentos realizados pelos personagens:

[...] There was no movement in the figures in early animation besides a single progression across the paper. No one knew how to get any change of shape or flow of action from one drawing to another. There was no relationship of forms, just the same little cartoon figure in a new position on the next piece of paper [...].⁹

⁸ “[...] ‘Caricatura e exagero foram as duas palavras favoritas para estimular a abordagem do animador para sua cena. Estas palavras poderiam ser mal interpretadas, como um pedido para a ação selvagem, descontrolada, mas este caminho sempre terminou com, ‘Veja, você não está pegando a ideias do que nós estamos fazendo aqui!’ A ação tinha que ser baseada no realismo, tinha que se adequar à situação da história, colocar em questão o ponto da cena, e se envolver com outras coisas sendo feitas em cenas diferentes.” (THOMAS & JOHNSTON, 1981, tradução nossa).

⁹ “[...] Não havia movimento nas figuras das primeiras animações, além de uma singela progressão através do papel. Ninguém sabia como alcançar uma mudança de forma ou fluxo de ação de um desenho a outro. Não havia relações de formas, somente o mesmo pequeno desenho cartum em uma nova posição no próximo pedaço de papel [...]

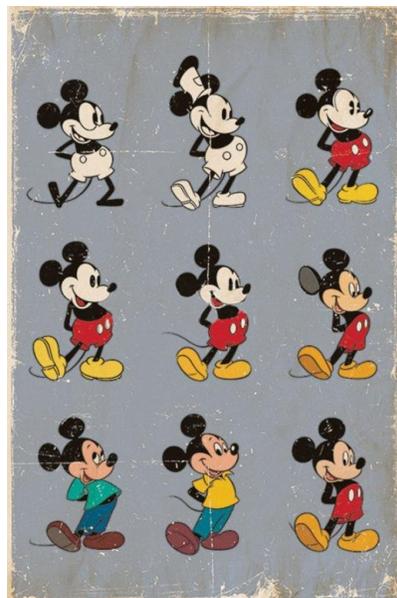
Figura 36 - *Frame de Feline follies* (1919), episódio do “Gato Félix”,



Fonte: <https://www.youtube.com/watch?v=fsRullIZs9M>. Acesso em: 28 Fev.2018.

A simplicidade no traço, com poucos detalhes e um número reduzido de membros, como os dedos das mãos, a procura pelo máximo contraste tonal para se destacar do fundo branco sobre o qual eram animados, a ausência de roupas e de suas dobras, eram algumas das estratégias usadas pelos animadores para conseguir realizar seu trabalho em equipes reduzidas e em um curto tempo, mas que se tornaram ícones de uma época da animação (Figura 36 e 37).

Figura 37 - Evolução do Mickey Mouse através dos tempos.



Fonte: Pôster. <https://www.europosters.pt/posters/mickey-mouse-evolution-v21418>. Acesso em: 29 Jan. 2018.

Conforme o processo de animação foi sendo dividido em etapas – onde cada artista era responsável por uma parte do processo – e os estúdios contratavam mais profissionais para nela trabalharem. As transformações formais foram se desenrolando, ao longo do tempo, e aqueles cuja atribuição era criar os personagens adotaram diversos expedientes para fazer com que suas criaturas se diferenciasssem e destacassem das produções de seus concorrentes. Diversas dessas tentativas transformaram-se em fórmulas para a criação de personagens, como as descritas por Hart (2005) e Bancroft (2006b), ou aquelas encontradas em *artbooks* de filmes de animação (Figura 38).

Figura 38 - *Concept* de personagens do filme "Como treinar seu dragão".



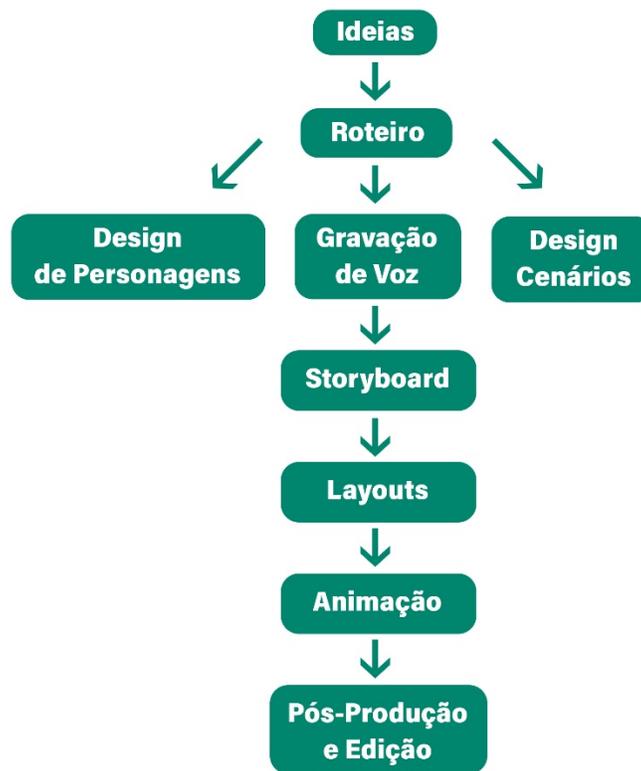
Fonte: (MILLER-ZARNEKE; COWELL, 2010, p. 64).

Tais “fórmulas” foram aplicadas tanto nas formas quanto nos movimentos dos personagens e alteradas conforme novos modelos eram experimentados e descobertos, em um processo de tentativa e erro que culminou com sua consolidação e aceitação pelos profissionais e pelo público consumidor.

Esses equacionamentos técnicos de linguagem absorveram igualmente as transformações conceituais ocorridas no cinema e nas artes de onde a animação tirava suas referências. Conceitos como a montagem e o enquadramento, as influências teatrais (em especial a pantomima), as formas dos desenhos dos quadrinhos e das ilustrações infantis, todos se apresentaram através de um prisma que acabou por uni-los, filtrá-los e mesclá-los através dos *frames* animados ao longo do tempo.

A partir destes, um sistema de produção foi adotado na indústria da animação, tal qual uma linha de montagem. O quadro da Figura 39 resume este sistema.

Figura 39 - Etapas da produção de animação.



Fonte: Baseado em Wright (2005).

Ao mesmo tempo em que este sistema uniformizou a forma de produção do setor de animação, permitiu que os esforços dos diretores passassem a se concentrar na “atuação” do personagem em cena, que passou de repetições de movimentos a *performances* complexas, similares às do teatro ou do cinema, porém representadas de maneira exagerada, típica dos cartuns.

Essa evolução é descrita nas obras de Barrier (2003), Thomas e Johnston (1981), Bendazzi (2016a, 2016b, 2016c) e Barbosa Junior (2001) e pode ser resumida na Tabela 3.

Tabela 3 - Estilos e características dos movimentos de animação

Estilo	Características
Experimental	Curtas realizados por apenas um autor. Movimentos limitados e repetitivos. Ausência de aceleração nas ações. Falta de “gravidade”. Não há pausa entre as ações. Personagens parecem não ter peso. Presença de onomatopeias e símbolos gráficos para representar sons e emoções dos personagens (influência direta dos quadrinhos).
<i>Loops (Ciclos)</i>	<i>Loops</i> ou ações repetitivas eram bastante comuns no início dos desenhos animados porque facilitavam o trabalho do animador e aumentavam a produção do estúdio. No início, os movimentos eram mais maleáveis e as ações repetitivas. Mesmo estúdios mais robustos, como o de Walter Lantz (da série do “Pica-pau”) usavam este recurso. Por serem repetitivos, muitos personagens aparentam ter movimentos robóticos. Foram muito usados no início do século XX e, aos poucos, sendo aplicados pontualmente em movimentos como corridas de personagens, que se repetiam nas animações. Além disso, são extremamente usados em <i>games</i> .
Rotoscopia	Desenhos feitos diretamente sobre a filmagem de um ator. Movimentos naturais, mas, em geral, destoantes da linguagem do cartum. Melhor aproveitamento do volume e de posições complicadas para desenhar o personagem, como o escorço de membros do corpo.
<i>Rubber hose animation</i>	Tronco e membros de personagens realizam movimentos repetitivos e ondulados, semelhantes a uma mangueira de borracha esguichando água. Este tipo de movimento permitia fluidez nos movimentos, mas fazia com que os personagens aparentassem não ter músculos ou ossos.
<i>Ripple-action</i>	Uma variação do <i>Rubber hose</i> , na qual os personagens exageram movimentos ondulatórios que fazem durante o caminhar ou a corrida. Nota-se que esse foi um dos precursores do que veio a se denominar, mais tarde, de movimentos em arcos (<i>arcs</i>).
“Naturalista”	Baseado nos movimentos naturais, porém aplicando características cartuns aos movimentos, como deformações no volume e na velocidade dos movimentos.
Princípios de animação	Usados nos estúdios Disney, baseiam-se em algumas características dos movimentos “naturalistas”, como a flexibilidade das formas, sendo acrescidos de convenções cinematográficas de enquadramento e edição; manutenção do volume durante as deformações; utilização dos movimentos naturais como referências e não cópias, tal qual a rotoscopia; exagero de formas e de <i>timing</i> ;
Warner	Exagero de formas; poses extremas como as dos comediantes dos filmes em preto e branco; uso de convenções dos quadrinhos, como onomatopeias e rastros de movimentos; diferenciação através de linguagem (incluindo o uso de violência); inversão de alguns princípios de animação da Disney, como a antecipação e o <i>timing</i> .
UPA (United Productions of America)	Questionamento visual dos princípios de animação em função do design (forma) dos personagens. Movimentos limitados – muitas vezes, em função do orçamento – que se tornaram um estilo.

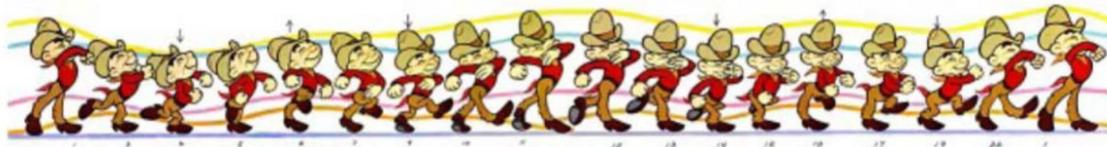
Fonte: Baseado em Barrier (2003), Thomas e Johnston (1981), Bendazzi (2016a, 2016b, 2016c) e Barbosa Junior (2001).

Figura 40 - CupHEad - *Game* de 2016 que usa o *Robber hose animation* como estilo.



Fonte: *Frame* do game CupHead.

Figura 41 - Exemplo de *Ripple action*.



Fonte: Blair (1994, p. 141)

Após a Segunda Guerra, a supremacia dos cartuns produzidos nos Estados Unidos se evidenciou devido à força geopolítica e econômica deste país. A maior parte dos países europeus estava reconstruindo suas economias e muitos estavam endividados, especialmente com a nova potência global. Tratava-se de uma dominação não apenas econômica, mas cultural também, e os países derrotados (na verdade, todos os que faziam parte do mundo ocidental) logo experimentaram isso. Hollywood e seus cartuns invadiram os cinemas de todo o mundo.

Durante a guerra, a animação estadunidense não havia atuado apenas como divertimento para as massas esquecerem aquele momento. Os estúdios se favoreceram naquele período. Alguns criaram animações instrucionais para o exército, outros, curtas para fazer propaganda das forças armadas, enquanto alguns ironizavam seus inimigos (Figura 42). Algumas divisões do exército ianque eram responsáveis pela comunicação e a

partir dali se organizaram os artistas que fundaram um estúdio bastante influente no cenário da animação: a United Productions of America.

Este, assim como a Warner, iria propor novas fórmulas para a animação, questionando tanto o estilo gráfico dos personagens e cenários quanto a forma como eles se movimentavam. E os princípios de animação da Disney seriam finalmente questionados e colocados à prova, em especial pela indústria da animação. O Oscar de 1951 refletiu essa visão alternativa da animação ao premiar o curta *Gerald McBoingBoing* da UPA, com seu estilo minimalista e movimentos limitados (Figura 39).

Figura 42 - *Frame* de animação ironizando a Alemanha nazista.



Fonte: *Donald Duck – Der Fuehrer's face* (1943).

Figura 43 - *Frames* da animação *Gerald McBoingBoing* (1951).



Fonte: <https://www.youtube.com/watch?v=uNsyQDmEopw>. Acesso em: 22 Fev 2018.

Nos anos seguintes a própria Disney procurou incorporar outros estilos gráficos à sua animação, porém os trilhos por onde a animação iria trafegar já estavam bifurcados e, tanto internamente – nos Estados Unidos – quanto externamente – na Europa e nos demais países – essa forma de arte seria desenvolvida para além da indústria, ganhando novos ares nos festivais, em especial em Annecy, na França, onde a linguagem experimental encontrou um público ávido por recebê-la (BENDAZZI, 2016b).

Esses festivais estimularam igualmente os animadores independentes, que encontraram neles locais para difundir seus trabalhos e realimentar a inovação na indústria da animação. Historicamente, os festivais são os locais onde os criadores podem apresentar formas não tradicionais de contar histórias, que é uma das principais bases da animação.

Assim, pode-se concluir que, tal qual a linguagem do cinema se consolidou com os pioneiros do início do século XX, como D. W. Griffith e Sergei Eisenstein, a dos cartuns foi amplamente influenciada por Walt Disney e seus princípios de animação, que permitiu a sua larga difusão no meio audiovisual. Entretanto, enquanto a Disney consolidava sua maneira “Disney” de animar, outras revoluções estilísticas baseadas nas culturas locais e no próprio questionamento de Hollywood a essa maneira de dar vida a personagens possibilitaram que a animação se tornasse um meio plural de transmissão de ideias e, principalmente, de criação de histórias, como será apresentado a seguir.

2.3 Animações: usos, técnicas e princípios.

Uma das funções primordiais da animação é contar histórias. Não há outra razão pela qual os espectadores mudam seu cotidiano para se dirigirem aos cinemas e passarem em média duas horas assistindo personagens movendo-se em uma tela. Entretanto, o meio e a técnica da animação não se restringem apenas à animação de personagens, embora seja através dela que a animação se fez mais popular. Desde seu nascimento, são as histórias conduzidas por personagens que atraem o público para as salas de cinema, para os festivais e para frente das televisões e dos computadores e dispositivos móveis, à procura de entretenimento.

A maior parte das histórias contadas no meio audiovisual é feita através de estruturas de construção narrativa, herdadas da tradição literária. Não foi à toa que os

primeiros curtas animados usaram histórias infantis conhecidas como base para suas narrativas: “A cigarra e a formiga” (*The grasshoper and the ant* – Figura 22), “A lebre e a tartaruga” (*The tortoise and the hare* – 1934 – Figura 44), “As aventuras do Príncipe Achmed (*Die Abenteuer des Prinzen Achmed* – 1926, de Lotte Reiniger – Figura 45), dentre outros que aproveitaram a popularidade e a concisão destas histórias para adaptá-las para o novo meio.

Figura 44 - *Frames* da animação "A lebre e a tartaruga", Walt Disney Studios (1934).



Fonte: <https://www.youtube.com/watch?v=GcAgo9U3vQY>. Acesso em: 23 Fev. 2018.

Figura 45 - *Frames* da animação *Die Abenteuer des Prinzen Achmed* (1926).



Fonte: <https://www.youtube.com/watch?v=VBnefAxbfvI>. Acesso em: 24 Fev.2018.

Além das histórias baseadas em fábulas, a pesquisa acerca das estruturas narrativas gerou fórmulas que passaram a ser utilizadas extensivamente por Hollywood, incluindo a animação de personagens. Há diversas maneiras de se contar uma história. Para a animação, a mais utilizada comercialmente é a estrutura do *monomito*, conhecida como “Jornada do herói”. Descrita por Joseph Campbell, esta apresenta uma estrutura de passos comumente encontrada nas histórias populares e, segundo ela, a história precisa ser feita para um protagonista (o herói) que precisará atravessar algumas adversidades e será guiado para alcançar um objetivo (LUOMALA; CAMPBELL, 1950).

Animações como a do “Rei Leão”, “A bela e a fera”, “Cinderela”, “A princesa e o sapo”, “Rio”, dentre outras, seguem esta estrutura. Os curtas-metragens exibidos em festivais como o da Siggraph ou Annecy procuram propor outras formas narrativas, como *Kela mälu* (“Memória corporal”, 2011 – Figura 46), animação de Ülo Pikkov, que apresenta, de forma metafórica, as deportações ocorridas na Estônia pela União Soviética.

Figura 46 - *Frames* da animação em *stop-motion* "Kela mälu", de Ülo Pikkov (2011).



Fonte: <https://www.youtube.com/watch?v=OLKPmoMsD9g>.

Assim como nas fábulas e na literatura, uma história animada geralmente envolve um personagem principal, que é um indivíduo ou um ser com características humanizadas. O cinema e a animação se caracterizam por serem artes baseadas em narrativas. Como apresentado anteriormente, as bases que compõem a animação são os quadrinhos, o teatro e o próprio cinema.

As histórias interessam aos indivíduos porque permitem a compreensão de seu passado para que possam prever o futuro. Fábulas e histórias infantis costumam usar metáforas para ensinar lições de moral e ética às crianças. A animação vai no mesmo sentido, ampliando o alcance das mensagens, devido ao uso de elementos visuais para transmiti-las. Desta maneira, ela atua na capilarização das histórias. Não é incomum nesse meio a adaptação de narrativas originárias na literatura que, após sua exibição nos cinemas, tornaram-se mais populares e procuradas pelo público.

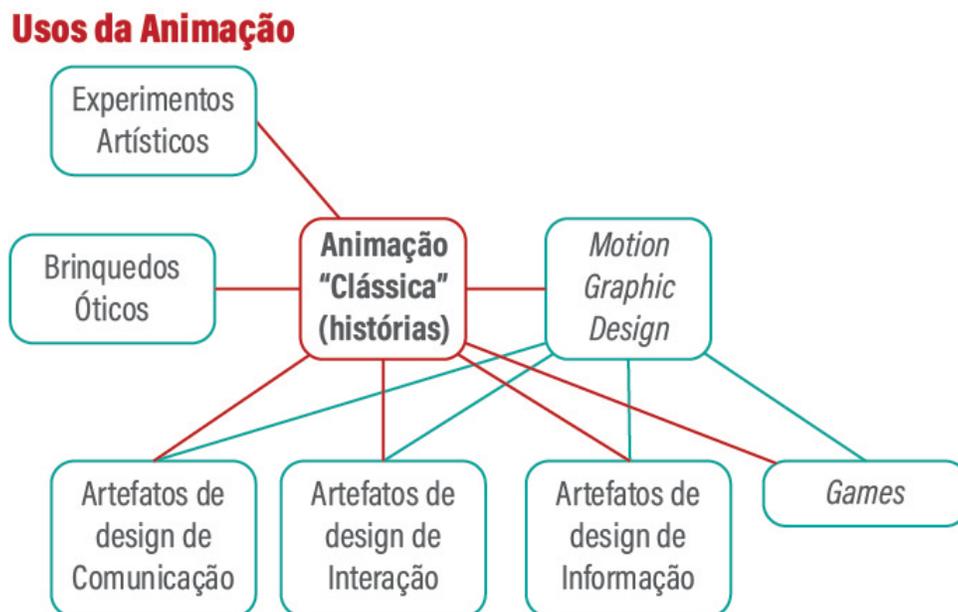
2.3.1 Usos da animação

A animação é, portanto, uma forma de contar histórias e, embora esta não seja sua única função, é através dela que os personagens atuam. Considerando que a animação pode ter diversos usos e que ela se relaciona com o design, principalmente através do *Motion*

graphic design, proponho a classificação da Figura 47 para identificar diversos usos da animação em diferentes tipos de suporte e através de diversos meios.

O *Motion graphic design* é um campo de atuação do designer, no qual ele utiliza os elementos do design para transmitir mensagens através da animação.

Figura 47 - Usos da animação.



Fonte: Diagrama do autor.

- *Animação clássica*: animações de personagens, em geral, humanizados que se utilizam da narrativa para contar histórias para o público;
- *Motion graphic design*: compreende o design gráfico animado, com o uso de elementos abstratos e personagens para a representação de conceitos, como nas aberturas de programas televisivos, *games* e publicidade;
- *Brinquedos óticos*: constituíram o primeiro conjunto de dispositivos mecânicos projetados para gerar imagens em movimento. Foram comercializados no século XVIII como brinquedos e atualmente continuam sendo explorados com esta função e como dispositivos para o ensino de animação e design (CRUZ, 2017);
- *Experimentos artísticos*: consistem no uso da animação como um meio para expressão visual artística por meio do vídeo. O início do século XX, possibilitaram

o surgimento de uma gama desses experimentos através de movimentos como o cubismo e o construtivismo;

- *Artefatos de design de comunicação*: são representações animadas em design utilizadas para trazer ampliar a comunicação em áreas do design como a identidade visual e o *branding*;

- *Artefatos de design de interação*: são as representações gráficas de sites e aplicativos que utilizam animação como uma forma de classificar as informações e atrair a atenção do usuário;

- *Artefatos de design de informação*: é a representação da visualização da informação, de maneira animada. Encontramos esses artefatos representados por gráficos animados nos mais diversos meios, como a televisão, a internet e a visualização científica;

- *Games*: constituem uma área muito específica da aplicação da animação clássica e do *Motion graphic design*, pois se utilizam de ambos para entreter e contar histórias interativas.

Como exposto na Figura 47, a classificação apresentada considera que cada uma das áreas pode interagir com as demais, gerando produtos híbridos. Caberá ao designer decidir qual das abordagens usará de acordo com o conceito desenvolvido e o objetivo da mensagem a ser transmitida. A vantagem dessa classificação reside no fato de que permite diferenciar as animações com objetivos narrativos de outras expressões em movimento, como o *Motion graphic design* e os experimentos artísticos de vanguarda que se utilizam de animação para transmitir uma ideia.

O *Motion graphic design* se destaca como um caso em particular do uso da animação porque essa é utilizada, de maneira utilitária, para a produção de projetos gráficos em movimento. Objetiva, pois, transmitir mensagens de maneira gráfica, tendendo a priorizar textos e formas abstratas em detrimento da animação de personagens em suas produções.

As experiências vanguardistas com movimento possuem uma natureza experimental. São obras de arte que buscam explorar e questionar a mudança das formas no

tempo, objetivando sensibilizar o público através do som e da imagem, sem necessariamente preocupar-se com a estrutura lógica narrativa.

A animação clássica, por outro lado, são histórias com personagens que possuem variações de temas e técnicas. Os temas abordarão as culturas nacionais, as histórias locais e até as histórias internacionais, podendo não ser influenciadas por culturas estrangeiras. Neste sentido, a influência estadunidense é tão impositiva que é comum que animadores de países tropicais utilizem como referências as animações produzidas naquele país. Uma das características mais marcantes desse processo de colonização é a presença de neve nestas produções, elemento não encontrado em países tropicais, como o Brasil.

2.3.2 Técnicas de animação

Já as técnicas de animação são transferidas de país a país, geralmente através da disseminação da tecnologia e da comunicação através da exibição dos filmes entre os países. Bendazzi (2016a, 2016b, 2016c), Chong (2008) e Barbosa Junior (2001) descrevem diversas técnicas utilizadas nos curtas e longas, que são característicos deles, sem que exista a possibilidade de determinar onde cada qual nasceu. Dentre estas se destacam (com suas variações):

Tabela 4 - Técnicas utilizadas em animações.

Técnica	Variações	Características
<i>Desenho animado</i>	-	Desenho animado diretamente sobre acetato.
	Pintura sobre vidro	Pintura realizada diretamente sobre vidro. Cada <i>frame</i> é levemente alterado em relação ao anterior.
	Gravura	É realizada a gravação do desenho em placas de madeira, que são regravadas a cada <i>frame</i> da animação.
<i>Stop-motion</i>	-	
Tridimensional	<i>Stop-action</i>	Atores reais são fotografados em momentos específicos de seus movimentos e estes são exibidos sequencialmente, gerando sequências fragmentadas com “efeitos especiais”.

	Marionetes	As marionetes são animadas através de controles que movem suas partes.
Bidimensional	Animação de recortes	
	Animação em areia	
<i>Pin-animation</i>		Animação realizada através das sombras de pinos que são levantados e abaixados a cada <i>frame</i> .
Rotoscopia	-	O movimento é desenhado diretamente sobre a filmagem realizada.
Animação digital	-	
	Animação 3D	Animação similar à do <i>stop-motion</i> , em que marionetes são movimentadas através de esqueletos e controles. Isto, no entanto, é realizado através de <i>softwares</i> de computadores.

Fonte: Elaborada pelo autor.

Todas estas técnicas reforçam o caráter de criação de universos fantasiosos e de entretenimento da animação, reafirmando a linguagem do cartum, ao mesmo tempo em que mostram que as técnicas são menos importantes que as histórias narradas e as atuações dos personagens.

2.3.3 Princípios de animação

A animação de personagens remonta, com suas poses, ao cartum dos quadrinhos ao utilizar-se de exageros expressivos que são aplicados a diversos elementos presentes no deslocamento do personagem, como o tempo em que ele leva para realizar uma ação, sua postura, seus trejeitos faciais, que se assemelharão àquelas das artes já consolidadas: a pantomima, o cinema, o teatro, os quadrinhos, seguindo suas convenções e ampliando seu valor com o uso do tempo.

Como apresentado na Tabela 3, essas convenções são frutos das descobertas dos animadores no século XX e dos avanços realizados na forma e nos movimentos dos

cartuns, com o objetivo de fazer com que as histórias animadas atraíssem e satisfizessem o gosto do público.

Pode-se afirmar, com base nas descrições das primeiras animações feitas por Bendazzi (2016a) e Thomas e Johnston (1981), que, nos primeiros curtas – antes das convenções dos Estúdios Disney serem adotadas pelo mercado de cartuns – os animadores possuíam mais “liberdade expressiva” para criarem suas produções. Devido à característica artesanal do processo, era mais comum a mão “invisível” do animador ser mais aparente nos traços e nos trejeitos dos personagens que naquelas animações que foram realizadas de 1930 em diante, quando o processo de produção da animação já começava a se consolidar.

Isto não significa que as experiências individuais bem-sucedidas não se tornaram estilos. Antes de 1930, porém, estes eram mais subjetivos. O *rubber hose*, estilo marcante nos curtas animados nas décadas de 1910 e 1920, foi amplamente disseminado, sendo reproduzido por animadores em curtas de estúdios que eram concorrentes: os Estúdios Disney e os Estúdios Fleischer. Contudo, ele não era o único. Havia uma variedade formal entre os curtas destes estúdios e a animação do Gato Félix, por exemplo. Algo que não se verifica nas animações 3D atuais, que pasteurizaram o tipo de movimento esperado dos personagens, com poucos lampejos de diferenciação entre os grandes estúdios.

Por fim, com o passar do tempo, o objetivo dos animadores passou a ser o exagero de formas e movimentos. As formas ficaram sob a tutela do referencial quadrinístico, enquanto os movimentos foram atrelados às referências naturais (às *performances* de atores e animadores). Ressalte-se aqui que o uso de referências não consiste em uma cópia de movimentos, mas uma apreensão destes sob a ótica do exagero.

Sob tais condições, percebe-se a adaptação daquilo que é “natural” no contexto do ator para o que é “natural” no universo do cartum: ao mesmo tempo em que se utilizavam as referências naturalistas para dar vida às personagens, estes apresentavam diferenças radicais de movimentos em relação àquelas, ou seja, criam-se exageros de expressão que se complementam e, por vezes, se opõem ao naturalismo, distorcendo formas e movimentos, como pode ser visto na Figura 48, *frame* da animação “Uma cilada para Roger Rabbit”, em que o personagem principal é desenhado em uma pose impossível de ser realizada por um ator real, mas perfeitamente plausível e esperada para um cartum.

Além disso, este *frame* mostra que os animadores criaram convenções que se tornaram comuns no meio, como os olhos saltando do personagem e a boca extremamente aberta, vista nesta animação.

Figura 48 - Quadro da animação "Uma cilada para Roger Rabbit".



Fonte: "Uma Cilada para Roger Rabbit" (ZEMECKIS, 1988).

A utilização de referências reais em conjunção com características apropriadas de outros meios como os quadrinhos, a pantomima, o teatro e o cinema gerou o que se convencionou chamar de doze princípios de animação. Estes, descritos por Thomas e Johnston (1981) tornaram-se recomendações para animadores sobre como criar animações de personagens que fossem mais naturais. Tornaram-se um conjunto de fórmulas a serem utilizadas pelos profissionais da área para alcançarem o estilo Disney (atualmente o mais empregado nas produções, devido à influência deste estúdio).

Como em toda arte, os princípios da animação foram questionados pelos profissionais que trabalhavam no setor e objetivavam concorrer com a Disney. Como visto anteriormente, os animadores que trabalharam na Warner e na UPA foram os principais agentes reativos a este estilo, propondo animações, como a do Coiote e a do Papa-léguas, que faziam com que estes princípios fossem distorcidos e questionados através de movimentos que em si eram antíteses do que pressupunham os princípios. Estes princípios são apresentados na Tabela 5.

Tabela 5 - Princípios de animação e suas características.

Princípio de animação	Características
<i>Squash and stretch</i>	Deformação volumétrica do personagem ao movimentar-se ou sofrer.
<i>Anticipation</i>	Realização de movimento na direção oposta ao movimento principal para acumular energia e atrair a atenção do público.
<i>Staging</i>	Apresentação clara de ideias através do posicionamento do personagem em cena.
<i>Straight ahead action and pose to pose</i>	São métodos de animação: <i>Straight ahead action</i> (“animação direta”) – método em que o animador coloca o personagem em movimento de maneira direta e sequencial; <i>Pose to pose</i> (“pose a pose”) – método em que o animador desenha poses e vai completando com os desenhos entre elas.
<i>Follow through and overlapping action</i>	Referem-se à representação de movimentos de partes maleáveis dos personagens ou apêndices a estes, quando são puxadas ou empurradas.
<i>Slow in and slow out</i>	Este princípio rege os movimentos do personagem no que tange à aceleração e à desaceleração de movimentos, sejam eles realizados pelo personagem como um todo, como por suas partes.
<i>Arcs</i>	Recomendação de que os movimentos dos personagens sejam realizados em formas de arcos para que sejam mais fluidos e graciosos.
<i>Secondary action</i>	A “ação secundária” diz respeito a movimentos realizados simultaneamente pelo personagem. Por exemplo, ajeitar os óculos enquanto levanta de uma mesa.
<i>Timing</i>	Refere-se ao tempo entre cada pose do personagem.
<i>Exaggeration</i>	Exagero expressivo do personagem.
<i>Solid drawing</i>	É o domínio do desenho do personagem em qualquer posição e situação.
<i>Appeal</i>	É a preocupação com a forma do personagem (seu design) em relação à sua atratividade ou repulsão para o público.

Fonte: Baseado em Thomas e Johnston (1981).

Da forma como são apresentados, estes princípios carecem, em conjunto, de uma lógica que os unifique. Alguns tratam diretamente da maneira como o desenho é feito, outros se referem à cinematografia, enquanto outros ainda se referem à metodologia de

animação, ou seja, são um conjunto não uniforme de elementos agrupados sem uma unidade.

Utilizá-los como um todo coeso para propor um sistema de classificação e aplicação de movimentos cartuns seria incoerente. Uma proposta mais apropriada seria dividi-los e classificá-los de maneira a extrair deste conjunto apenas aqueles que se relacionam diretamente aos movimentos dos personagens.

Assim, propomos a divisão destes princípios segundo algumas configurações comuns: os métodos, os tipos de movimentos, a característica cinematográfica e as formas dos personagens, apresentados na Tabela 6.

Tabela 6 – Proposta de classificação de Princípios de animação

Característica relacionada	Princípios
Métodos de animação	<i>Straight ahead e pose to pose.</i>
Movimentos	<i>Timing; antecipação; slow in e slow out; follow through and overlapping action; arcs; secondary action.</i>
Cinematografia	<i>Staging.</i>
Forma do personagem	<i>Appeal; solid drawing; squash and stretch.</i>

Fonte: Elaborado pelo autor.

Há de se notar que o princípio do exagero não se encontra relacionado a qualquer dos itens de classificação propostos. Embora Thomas e Johnston (1981) deixem subentendido que o exagero descrito em sua obra refira-se ao exagero nas emoções dos personagens, é comum na prática profissional do animador o uso deste termo para se referir aos demais princípios de animação. Assim, pode-se exagerar o *squash and stretch* ou o *timing* do personagem, bem como suas formas.

Por ser um termo genérico, o próprio exagero poderia ser uma categoria de uso para classificação dos princípios. Entretanto, esta discussão poderá ser desenvolvida em pesquisas posteriores.

A própria classificação dos princípios permite subclassificações, como por exemplo, o *staging*, que se refere ao posicionamento do personagem em cena, mas que, pode ser

trabalhado através de regras cinematográficas, como a regra dos terços, o uso da perspectiva para guiar o olhar do público ou a utilização de enquadramentos específicos, quando o personagem participa de um diálogo.

Algumas dessas categorias abrem um leque tão amplo de possibilidades que geraram especializações dentro do universo da animação, como o *appeal*, explorado pelos *concept artists*, quando da criação dos personagens para animação.

Entretanto, após realizar esta classificação, é importante ressaltar que nem todos os doze princípios de animação adotados pela indústria serão base para este trabalho. Para a elaboração do sistema de cartunização de personagens a ser apresentado no capítulo 4 da tese, apenas aqueles relacionados aos movimentos serão abordados.

Acreditamos, entretanto, que as classificações aqui propostas podem se constituir como uma base mais apropriada para a análise dos princípios de animação, à luz de uma visão mais racional dos mesmos.

A seguir, apresentaremos a revisão bibliográfica da captura de movimentos para, então, relacionar as descobertas relacionadas neste campo do conhecimento à criação do sistema de conversão de movimentos realistas em movimentos cartuns.

3 REVISÃO BIBLIOGRÁFICA DA CAPTURA DE MOVIMENTOS

A captura de movimentos aplicada à animação em estilo cartum não é um tema de investigação recente. Pesquisadores o têm abordado através de diversas perspectivas, desde o desenvolvimento de algoritmos capazes de aplicar alguns princípios de animação em partes de personagens *cartoons* a filtros programados para reduzir a quantidade de ruído presentes nas capturas para uso posterior, como pode ser verificado nas pesquisas de Aguiar (2003) e Kwon (2011).

A captura de movimentos ou *motion capture* (aqui tratada pela abreviação Mocap, como é comumente conhecida) faz parte de um subcampo do conhecimento chamado de *computer vision* (ou “visão computacional”), pertencente a uma área maior, a ciência da computação. Nas últimas décadas, diversas tecnologias vêm sendo desenvolvidas neste campo, objetivando a solução de problemas cotidianos. Dentre esses problemas, encontra-se o reconhecimento do ser humano e suas ações, tema deste capítulo.

Permitir que os computadores (e *smartphones*, atualmente) consigam reconhecer as ações humanas é um desafio para programadores e cientistas, visto que, ao contrário do ser humano, essas máquinas não possuem um cérebro capaz de diferenciar as sutilezas das imagens e das ações que são apresentadas pela natureza aos seus sensores (na maior parte das vezes, câmeras que fazem o papel dos olhos para elas).

Não é sem razão que diversos pesquisadores se debruçam sobre o problema de reconhecer imagens e movimentos. A evolução dessa área já permite que carros inteligentes sejam projetados (AMIR et al., 2017), pessoas possam ser rapidamente assistidas em caso de um acidente (AMINIAN; NAJAFI, 2004), a segurança de aeroportos seja aprimorada (LIN; SUN, 2008), estudos ergonômicos sejam feitos com maior precisão (DUTTA, 2012), a animação de personagens possa se valer das tecnologias de capturas de movimentos (WITTENBERG, 2013), o desenvolvimento de aplicações militares seja mais precisa (WILLIAMS et al., 2008), a medicina possa se beneficiar do uso de sistemas interativos dentro de centros cirúrgicos (O’HARA et al., 2014a, 2014b), o monitoramento de pacientes idosos seja feito com a urgência necessária às suas necessidades (AMINIAN; NAJAFI, 2004) e diagnósticos mais eficientes possam ser alcançados, através das aplicações que são desenvolvidas no campo (KARAMANIDIS; ARAMPATZIS, 2009).

No final da década de 1990 e na primeira década de 2000, a visão computacional deu um salto, em termos de pesquisas e produtos gerados, tendo seu uso ampliado para as mais diversas áreas do conhecimento. Isto se deu graças a alguns fatores: a redução de custos para a aquisição de equipamentos que realizam capturas (o *hardware*, em especial); a disseminação do conhecimento em programação, através da internet, que se tornou mais acessível à população mundial, de maneira geral; e à popularização de dispositivos eletrônicos da área de *games* cujos sensores reconhecem, total ou parcialmente, os movimentos humanos, e que passaram a ser usados por pesquisadores.

Os telefones celulares também se destacaram entre estes dispositivos, pois, se antes eram usados apenas para comunicação, com o surgimento dos *smartphones*, passaram a realizar uma série de funções que apenas os computadores desempenhavam (como por exemplo, agendar tarefas, editar textos e planilhas, fazer apresentações, etc.) e a serem equipados com diversos sensores para identificar, reconhecer e responder a interação humana, como acelerômetros e câmeras.

É importante destacar que esta pesquisa foca justamente na análise do movimento humano, através da sua identificação e compreensão, ignorando os demais movimentos realizados por outros seres ou fenômenos da natureza. Nela, também não serão abordadas as tecnologias presentes em dispositivos móveis, embora já se saiba que muitos *smartphones* estão agregando as funções que antes eram exclusivas de sistemas de capturas de movimentos, como as câmeras de profundidade, cuja definição será abordada adiante.

Pretende-se, neste capítulo, apresentar o histórico da captura de movimentos, sua relação com a visão computacional, as funcionalidades dos sistemas existentes e os avanços das aplicações desenvolvidas nas últimas décadas, focando naquelas que, direta ou indiretamente, influenciaram as pesquisas na área da indústria do entretenimento e a fizeram avançar, especialmente no campo da animação computadorizada.

Ver-se-á que os sistemas de captura de movimentos evoluíram e foram apropriados, de tal forma, por outras áreas do conhecimento nas últimas décadas que, se antes estavam restritos aos meios militares e acadêmicos, hoje são comumente aplicados na indústria do cinema, da animação, dos *games* e do design, tendo sido sua “naturalização” ampliada pela divulgação dos *sets* de filmagens, como formas de promover os filmes e *games*, os conhecidos *making of*.

Por fim, espera-se que a revisão bibliográfica ora apresentada seja esclarecedora no que tange a importância desta área para o campo do design de animação, apontando as lacunas existentes no conhecimento e as principais questões ainda não respondidas neste campo, fornecendo, assim, mais justificativas para a realização da pesquisa, bem como, abrindo possibilidades para a realização futura de outras.

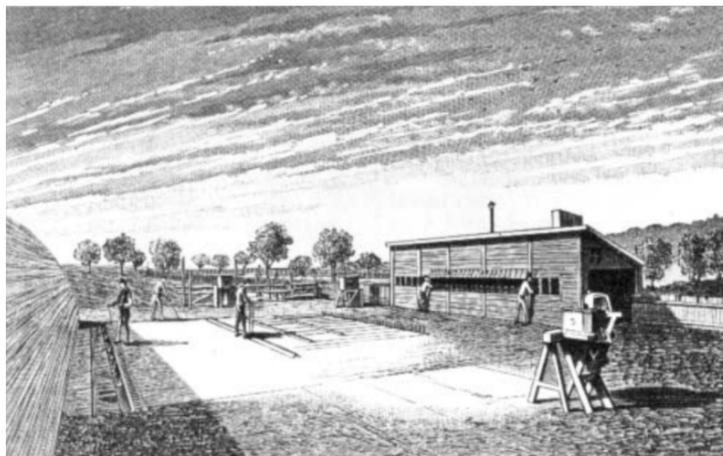
3.1 Antecedentes da visão computacional

De maneira bem idílica, pode-se dizer que o estudo dos movimentos humanos e de suas representações começou quando os primeiros seres humanos, munidos apenas de pedaços de carvão, desenharam suas caçadas nas paredes de cavernas. Embora essas representações sejam rudimentares, encontram-se ali os embriões do que seriam os primeiros passos na compreensão sobre os movimentos humanos.

No entanto, as pesquisas aprofundadas acerca da compreensão dos movimentos humanos e os equipamentos que as possibilitarão foram efetivadas no final do século XIX e deram origem ao que futuramente ficaria conhecido como captura de movimentos. Este período, fértil em descobertas, permitiu o desenvolvimento da câmera fotográfica e do filme flexível, tecnologias sem as quais dificilmente se registrariam os movimentos naturais. Como visto no capítulo anterior, foi graças a esses registros que se tornou possível compreender como os movimentos são realizados. Essas primeiras pesquisas fundamentaram as bases para o conhecimento acerca da captura de movimentos e foram fundamentais para seu amadurecimento posterior.

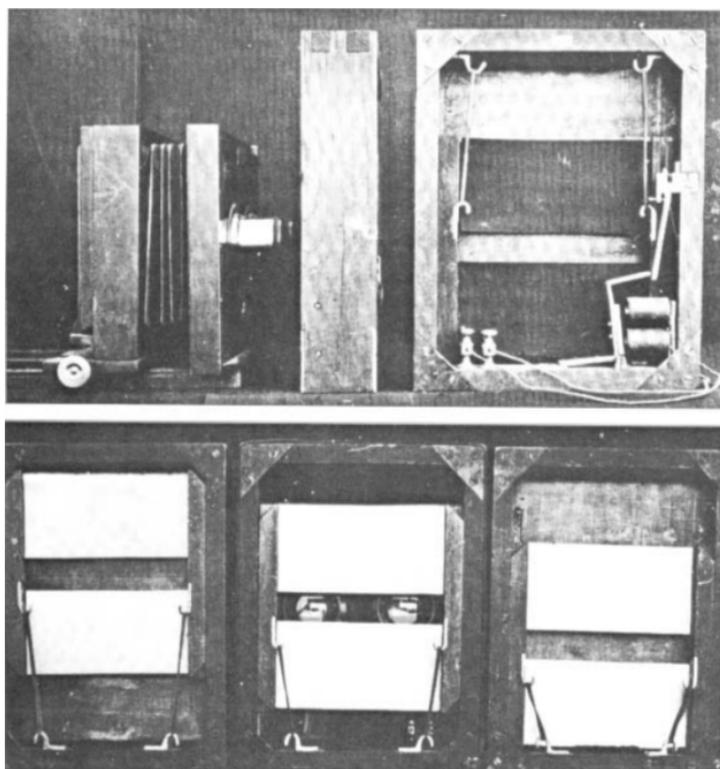
Dentre os pioneiros nesta área do conhecimento, estão o inglês Eadward Muybridge (1830 - 1904) e o francês Étienne Jules-Marey (1830 - 1904) (KITAGAWA; WINDSOR, 2008, p. 2–4). O primeiro criou, por volta de 1878, um sistema composto por doze câmeras que, colocadas lado a lado, eram disparadas, conforme o sujeito a ser capturado movimentava-se em frente a elas (Figura 49). O segundo criou diversos dispositivos mecânicos para compreender a fisiologia de seres humanos e animais não humanos. Um deles, o rifle cronofotográfico (1882), consistiu de um rifle que fazia rotacionar uma série de papéis fotográficos através de um tambor (Figura 50). Este permitiu a captura de imagens de indivíduos e animais em movimento.

Figura 49 - Experimento realizado por Muybridge para capturar o movimento sequencial de um cavalo.



Fonte: (TOSI, 2005, p. 47).

Figura 50 - Dispositivo projetado por Muybridge para capturar um cavalo em movimento.



Fonte: Tosi (2005, p. 47).

Figura 51 - Rifle cronofotográfico de Étienne-Jules Marey.



Fonte: By David Monniaux (*Own work*) [GFDL (<http://www.gnu.org/copyleft/fdl.html>), CC-BY-SA-3.0 (<http://creativecommons.org/licenses/by-sa/3.0/>)]

Tais dispositivos carregam em sua essência a ação investigativa dos pesquisadores, aliada à disposição destes em desenvolver outros que auxiliassem a ciência a responder as questões de seu tempo, através da utilização criativa dos meios disponíveis, para criar aparatos tecnológicos que lhes proporcionassem uma nova forma de observar os fenômenos naturais. Esta visão pragmática da realidade é a mesma que guia hoje os pesquisadores da visão computacional.

Apesar destes esforços serem importantes para a construção de uma visão genealógica que apresenta seu aparecimento e de outras tecnologias como as bases responsáveis pela captura de movimentos, o objetivo desta seção da pesquisa é apresentar aquelas que fundamentaram o uso de dispositivos eletrônicos, em especial os computadores, para realizar essas análises no que se convencionou chamar de visão computacional. Dentro desta área, interessam particularmente as utilizadas para capturar e reconhecer os movimentos. A aplicação aqui desse conhecimento consistirá em tentar relacionar o uso destas tecnologias e de suas aplicações com a animação cartum.

A seguir, a serão abordados três temas importantes para o entendimento da visão computacional: o histórico do reconhecimento de ações e atividades humanas; as tecnologias usadas para a captura de movimentos e suas aplicações; e as descobertas recentes da área e as questões que ainda estão abertas à pesquisa e ao desenvolvimento.

3.2 Histórico

Desfiar a história da captura de movimentos e do reconhecimento de gestos possui alguns objetivos específicos nesta Tese. O primeiro é, como a maior parte dos trabalhos acadêmicos, seguir uma metodologia que apresenta as fontes do conhecimento do qual se extraem as questões abordadas e as possibilidades de avanço na área do conhecimento. Para tais fins, a revisão bibliográfica aqui realizada busca apontar as lacunas existentes no conhecimento, determinando o escopo da pesquisa, ou seja, quais daquelas perguntas serão investigadas assim como seus limites, deixando algumas dessas perguntas sem resposta, para que possam ser futuramente pesquisadas. Esta via de atuação também permite que o conhecimento acerca da área seja aprofundado, de maneira que possa servir como base para estudos futuros.

O segundo objetivo é o de ampliação desse conhecimento tanto em língua portuguesa quanto na área de design. Ao procurar por artigos e publicações sobre a captura de movimentos em nossa língua, descobriu-se que poucos eram os que abordavam diretamente o tema, fazendo-se necessária a presença de pesquisas na área, em especial no campo do design, tão pouco atendido nesse meio. Esta pesquisa procura também firmar o posicionamento do design na área. Para ter uma noção da abrangência desse conhecimento, fez-se uma procura pelo termo exato “captura de movimento” no portal de periódicos da CAPES, em dezembro de 2018, que retornou apenas nove artigos. Nenhum deles na área de design (Tabela 7) e alguns publicados em inglês e espanhol.

Tabela 7 - Artigos relacionados à pesquisa do termo "captura de movimento" na base de periódicos da CAPES.

Nome do artigo	Área do conhecimento	Tipo de publicação	Data da publicação
Projeto de pesquisa: <i>Software: glossário de informática com aplicação de libras e de tecnologia de captura de movimento 3D</i>	Educação	Artigo	Agosto de 2016

<i>Estudo comparativo prospectivo para a avaliação da reabilitação de usuários de próteses com amputações transtibiais</i>	Saúde	Artigo	Outubro de 2016
<i>Efeito do treino neuromuscular na rotação do joelho durante a aterrissagem em mulheres</i>	Medicina do esporte	Artigo	Abril de 2016
<i>Upper extremity joint stresses during walker-assisted ambulation in post-surgical patients</i> (“Estresse articular no membro superior durante marcha assistida por andador em pacientes pós-cirúrgicos”).	Fisioterapia	Artigo	Agosto de 2011
<i>Using leap motion to create a physically interactive musical interface</i> (“Usando leap motion para criar uma interface musical interativa fisicamente”).	Engenharia e música	Tese	Agosto de 2017
<i>Paradigmas do jogar: interação, corpo e imersão nos videogames</i>	Jogos digitais	Artigo	2010
<i>Upper extremity joint stresses during walker-assisted ambulation in post-surgical patients.</i> (“Estresses da articulação da extremidade superior durante a deambulação assistida por andador em pacientes pós-cirúrgicos”).	Fisioterapia	Artigo	2011
<i>Influência das variáveis cinemáticas sobre o desempenho do chute giro dorsal de atletas de karatê de alto</i>	Esportes	Artigo	2015

rendimento.

Biomecânica do saque no tênis de campo: "estado da arte" e tendência dos estudos. Esportes Artigo 2014

Fonte: Portal de Periódicos da CAPES, consulta realizada em dezembro de 2018.

Já a busca pelo termo exato “captura de movimento” resultou em apenas 67 documentos na Biblioteca Digital Brasileira de Teses e Dissertações, plataforma mantida pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), do Ministério da Ciência, Tecnologia, Inovações e Comunicações.¹⁰

Por fim, objetivou-se também compartilhar o conhecimento nessa área de saber. Ao iniciar as pesquisas em captura de movimentos, acreditava-se que as publicações acerca desse assunto seriam formadas por artigos e teses publicados nos últimos dez ou vinte anos, ou seja, surgidas no final da década de 1990. Ledo engano, pois logo foram encontrados artigos indicando que as raízes que fundaram esse conhecimento foram plantadas na década de 1970, utilizando sistemas eletrônicos rudimentares e que, por si só, já se constitui em uma experiência instigante e esclarecedora. Compartilhar esse conhecimento é o mínimo que se pode fazer para contribuir com essa área no Brasil, como será visto mais à frente.

A visão computacional é, pois, uma área do conhecimento que trata da forma como dispositivos eletrônicos são programados para “automatizar e integrar um amplo campo de processos e representações usados na percepção da visão” (BALLARD; BROWN, 1982, p. 2). A percepção da visão é uma área bastante abrangente, pois pode incluir também disciplinas que não se relacionam diretamente com o reconhecimento digital de imagens, como a psicologia. Ballard e Brown (1982) complementam esta definição, ao afirmarem que a visão computacional “inclui como partes muitas técnicas que são úteis por si mesmas, como o processamento de imagens [...], e a classificação estatística de padrões. [...] Mais importante para nós, ela inclui técnicas para a modelagem geométrica e o processamento

¹⁰ Procura realizada em 15 Dez. 2018.

cognitivo”. Assim, como processamento cognitivo, pode-se incluir o reconhecimento de gestos humanos.

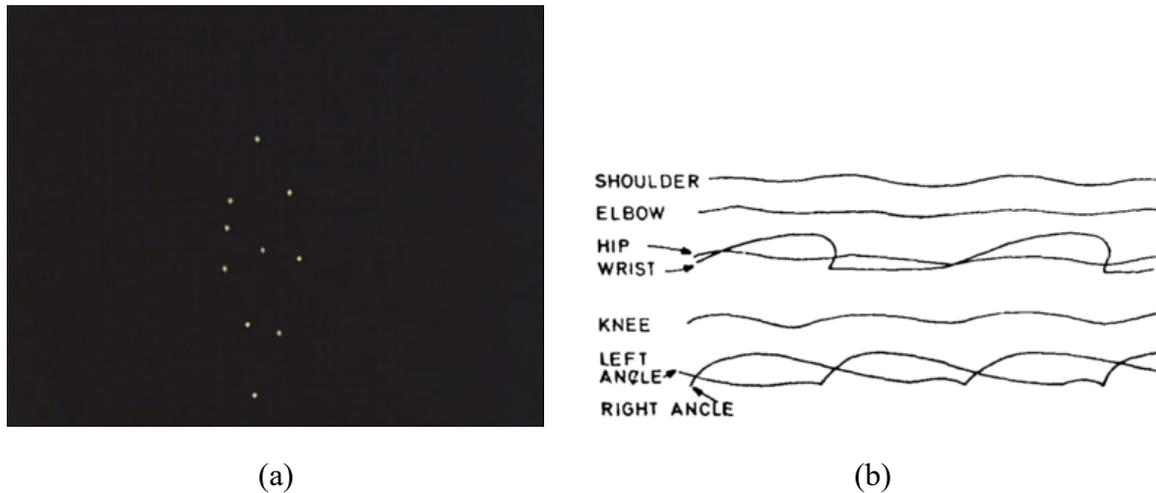
Embora seja possível o acesso às interfaces interativas disponibilizadas por meio dos produtos resultantes do conhecimento dessa área através dos dispositivos que utilizamos cotidianamente (como computadores e *smartphones*), as primeiras pesquisas datam de um período anterior ao da popularização da informática, que se deu a partir da década de 1980, com a invenção do computador pessoal (*personal computer* ou PC).

O desenvolvimento da visão computacional teve suas origens na década de 1970, quando alguns pesquisadores, trabalhando em laboratórios universitários, começaram a questionar quais seriam os elementos mínimos necessários para o reconhecimento de um movimento. Dentre esses pioneiros destacaram-se o sueco Dr. Gunnar Johansson (JOHANSSON, 1973) e o estadunidense Dr. James E. Cutting (CUTTING, 1978), que realizaram experimentos na área da percepção de padrões visuais de movimentos biológicos. O objetivo de suas pesquisas era compreender como indivíduos reconheciam movimentos realizados por meios desses elementos mínimos.

Na década de 1970, acreditava-se que a percepção dos movimentos possuía uma relação direta com a silhueta do indivíduo observado. Os estudos de Johansson indicaram que, por meio da utilização de poucos elementos visuais (pontos luminosos colocados em partes específicas do corpo), era possível identificar os movimentos realizados por seres humanos. Tais quais os experimentos de Muybridge e Jules-Marey, suas pesquisas ampliaram o conhecimento acerca do movimento humano, assim como ajudaram a questionar o consenso geral da época.

O experimento (Figura 52) foi gravado em vídeo, sendo considerado como uma referência na área da psicologia e da visão computacional (Blake e Shiffrar, 2007), já que comprovou a possibilidade de reconhecermos um indivíduo e sua movimentação, utilizando apenas um conjunto limitado de elementos – algumas luzes posicionadas nas juntas do corpo humano. Apesar de usar um equipamento rudimentar, o objetivo de Johansson era analisar estes movimentos, posteriormente, utilizando sistemas computacionais.

Figura 52 - (a) *Frame* de vídeo com a representação de um ser humano em movimento apresentado como experimento do Dr. Johansson (1971); (b) plotagem dos movimentos de determinadas partes do corpo a partir do vídeo gravado.



Fonte: (a) <https://www.youtube.com/watch?v=1F51CP9SYLU>. Acesso em: 22 Dez. 2017; (b) Johansson (1973, p. 203)

Pode-se dizer que a extensão de seu trabalho se deu com a pesquisa realizada por James Cutting (CUTTING, 1978), que, em 1978, publicou uma pesquisa na área da fisioterapia sobre a análise da marcha (*gait analysis*) na qual tratava dos diagnósticos de pacientes através da análise da forma como eles caminhavam.

Seu experimento foi um dos primeiros a utilizar o computador para realizar uma síntese da caminhada através do reconhecimento de pontos luminosos. Ele veio a fazê-lo por meio da criação de um algoritmo em Fortran em um computador Tektronix 604, recriando, assim, os movimentos de algumas partes do corpo e ampliando a compreensão da percepção do movimento.

Além do uso do computador, outro aspecto pioneiro na análise dos movimentos dos corpos foi a utilização da engenharia reversa, no momento em que o pesquisador resolveu reconstituir os movimentos gravados. Para realizar este feito, Cutting apropriou-se de estudos anteriores sobre o campo da visão computacional e simulou os movimentos de partes do corpo humano. No relato de seu experimento, ele apresentou suas impressões:

[...] movimentos dos ombros e dos quadris são elipsoidais (ver Carlsöö, 1972) e os movimentos dos braços e das pernas são pendulares. O corpo inteiro se move

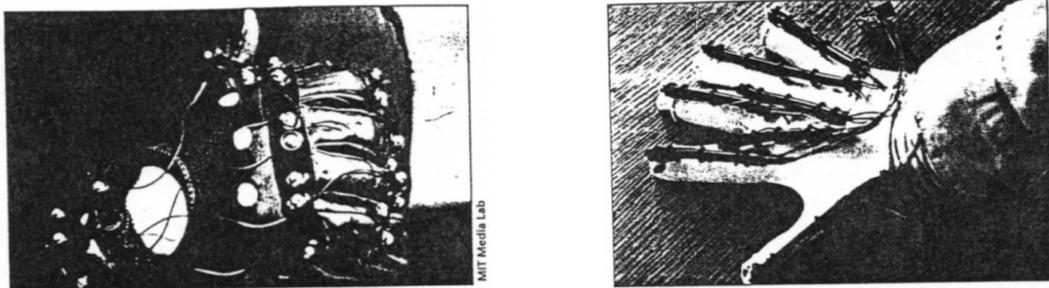
lateralmente através da tela, de acordo com o tamanho do passo, e sobe e desce em um movimento sutil.(CUTTING, 1978).

Este estudo permitiu a abertura das portas no campo da análise e da simulação de movimentos por meios de computadores, que foram iniciados e aprimorados na década seguinte. Os institutos de pesquisa nos Estados Unidos, na época os mais interessados em desenvolver *softwares* e *hardwares* para computadores, souberam como aproveitá-lo.

No início da década de 1980, procurando obter uma maior precisão das capturas, o pesquisador Tom Calvert, da Universidade Simon Fraser, adicionou potenciômetros e goniômetros a um exoesqueleto (um esqueleto mecânico) preso ao corpo de um voluntário para capturar seus movimentos e aplicá-los a um personagem digital (Sturman e Zeltzer, 1994, p. 37) e (PAIVA, 2014, p. 15).

Paralelamente, o interesse na captura de movimentos levou pesquisadores de diversas outras universidades a desenvolverem dispositivos capazes de realizar sua leitura. Um desses centros foi o Laboratório de Visualização Eletrônica na Universidade de Illinois. Através da pesquisa de Sayre, DeFanti e Sandin (Sturman e Zeltzer, 1994, p. 32), eles conseguiram criar, em 1976, uma solução barata para a captura de movimentos das mãos, uma luva de fibra ótica (Figura 53).

Figura 53 - (a) luva de LED desenvolvida pelo MIT; e (b) luva de fibra ótica desenvolvida por Sayre, DeFanti e Sandin na Universidade de Illinois.



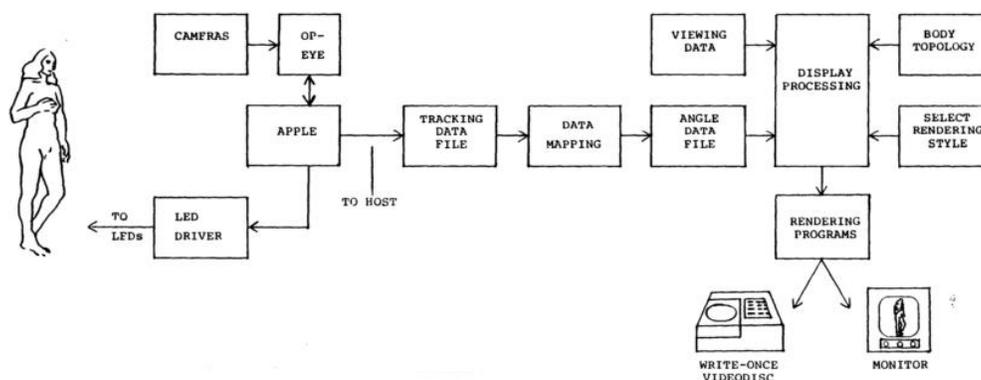
Fonte: Sturman e Zeltzer (1994, p. 32).

Ainda na década de 1980, o MIT, através do seu Media Lab, também desenvolveu um dispositivo capaz de capturar o movimento das mãos que, devido ao projeto anteriormente utilizado por Illinois, não acabou não sendo pioneiro. Entretanto, a luva do MIT Media Lab continha uma novidade eletrônica que seria aplicada futuramente pelos fabricantes de equipamentos de MoCap: o uso de LEDs (diodos emissores de luz) nas

juntas das mãos, permitindo o acompanhamento (*tracking*) dos movimentos das mãos e dos dedos através de câmeras que visualizam os pontos por eles representados. Esse tipo de sistema tornou-se popular entre os sistemas de capturas de movimentos devido à sua eficiência, sendo usado, até hoje, por aqueles mais precisos.

Assim, já preconizando o pioneirismo nessa área de pesquisa, o primeiro equipamento comercial de captura de movimentos foi, a seguir, desenvolvido no MIT por Delle Maxwell, no ano de 1983 (Ginsberg e Maxwell, 1983), em sua dissertação de mestrado. Nela, Maxwell utilizou duas câmeras para reconhecer os movimentos do corpo e um conjunto de LEDs conectados aos indivíduos. As câmeras filmavam os LEDs e as posições destes eram enviadas a um sistema de *software* que as seguia e convertia em ângulos de rotação. Esses ângulos foram, então, comparados às representações de um corpo humano (a topologia do corpo) para que o sistema apresentasse a captura em um monitor. O processo geral é apresentado na Figura 54:

Figura 54 – Modo de funcionamento do *graphical marionette*.

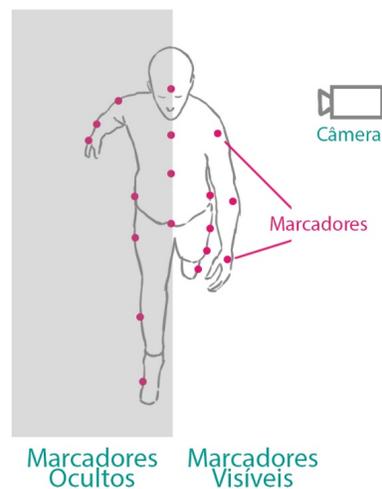


Fonte: (MAXWELL, 1983).

É importante salientar que esse sistema apresentou o *primeiro problema* advindo da utilização dos sistemas baseados em vídeo: o da oclusão dos membros capturados (Figura 55). Esse tipo de problema ocorre quando as câmeras não conseguem visualizar algum dos LEDs. Ele pode acontecer devido ao posicionamento das câmeras no ambiente ou devido ao posicionamento do ator, escondendo de maneira não intencional um dos LEDs. Imagine, por exemplo, um sistema hipotético que possui apenas uma câmera. Quando um indivíduo atravessa o campo de visão dessa câmera, caminhando de um lado a outro, o braço e a

perna que não estão de frente para ela ficam ocultos pelo corpo, durante parte do movimento. Tal problema é comumente encontrado em sistemas que utilizam poucas câmeras para identificar os pontos de LED e, identificado por Maxwell, permeia este tipo de sistemas de capturas de movimentos, até os dias atuais. Em sua pesquisa, Maxwell procura solucioná-lo por meio da criação de um sistema de predição de posição das juntas do corpo, através de um *software*, que, porém, não tem se mostrado muito efetivo.

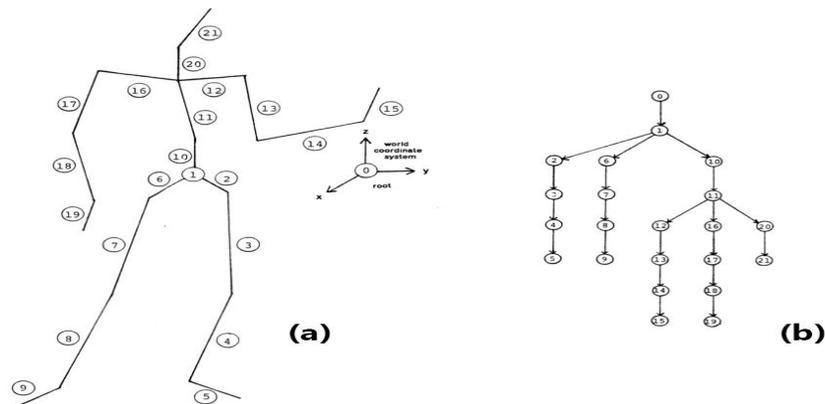
Figura 55 - Oclusão dos marcadores em sistemas de captura de movimentos com apenas uma câmera.



Fonte: Ilustração do autor.

Apesar de não resolverem de todo o problema da captura de movimentos, algumas outras inovações são apresentadas no trabalho de Maxwell, como a representação do esqueleto humano através uma série de nós e a hierarquização entre eles que foi a base para a representação que conhecemos atualmente (Figura 56).

Figura 56 - Hierarquização dos segmentos do corpo humano proposta por Maxwell.

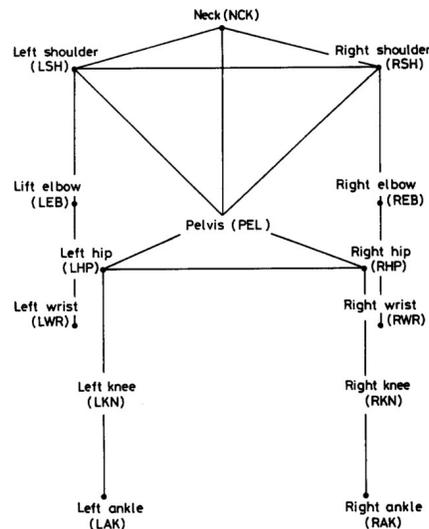


Fonte: Maxwell (1983, p. 45 e 46).

Outra inovação presente naquele trabalho foi a implementação de uma etapa de inicialização do sistema, através de um *software* cujo objetivo era reconhecer a posição inicial das juntas do corpo humano para realizar o posterior acompanhamento (*tracking*) das mesmas. Essa etapa é, até hoje, replicada em sistemas e aplicativos que fazem o reconhecimento do corpo para capturar movimentos.

Pouco tempo depois, mais um aprimoramento foi registrado nas etapas que compõem a etapa de reconhecimento corporal: a determinação das juntas necessárias para que exista um “corpo” (um esqueleto mínimo necessário para captura de movimentos). Em 1985, uma tentativa de desenvolvimento de um método computacional para o reconhecimento de segmentos e juntas em 3D de um indivíduo, a partir de uma única vista (ou vídeo), foi realizada por Lee e Chen (1985). Em seu artigo, eles descrevem um método para determinação de poses do corpo humano, desde que algumas de suas partes estejam visíveis e o tamanho da cabeça e dos membros do corpo seja conhecido. É importante observar que, enquanto um protótipo funcional, o sistema desenvolvido por eles parte da necessidade de determinar elementos mínimos para a representação do esqueleto humano visando o reconhecimento de poses. O esquema que propuseram (Figura 57), além de ser baseado nas pesquisas anteriores, tornou-se um padrão para a indústria, sendo utilizado, até hoje, em sistemas de reconhecimento de junções humanas como o Kinect.

Figura 57 - Representação diagramática do corpo humano para efeitos de reconhecimento de poses.



Fonte: (Lee e Chen, 1985, p. 150)

O método utilizado por Lee e Chen para determinar um conjunto de juntas em 3D do corpo humano consiste no reconhecimento das coordenadas de seis pontos característicos da cabeça e de suas projeções. A partir deste reconhecimento, perfazem, por meio de uma matriz de transformações, o estabelecimento entre as quatorze juntas do corpo.

Ambas as pesquisas, tanto a de Maxwell (1983) quanto de Lee e Chen (1985), contribuíram, portanto, para o estabelecimento da base tecnológica e conceitual que será utilizada nos sistemas atuais de capturas de movimentos. Delas, surgiram as bases para a industrialização dos sistemas ópticos de reconhecimento que trarão à indústria do audiovisual algumas soluções comerciais, como o OptiTrack™, o Vicon™, nas décadas de 1980 e 1990, e os sistemas baseados em vídeo de apenas uma câmera, como o popular dispositivo Kinect™, da Microsoft, no final da primeira década dos anos 2000.

A disseminação desse conhecimento levou a produtora Jim Henson Productions a criar a primeira marionete virtual televisionada em 1988, chamada de Waldo C. Graphic (Figura 58). Esta marionete era controlada por um ator cujos movimentos eram capturados através de dispositivos eletromecânicos. Os movimentos eram, então, processados e

aplicados à movimentação de um personagem tridimensional em tempo real (STURMAN, 2016).

Figura 58 - *Frames* do programa televisivo onde Waldo C. Graphic era exibido com seus controles eletromecânicos.



Fonte: Canal do Youtube de The Jim Henson Company (<https://www.youtube.com/watch?v=dP6TUB7KQc4&t=24s>). Acesso em: 17 Dez. 2016.

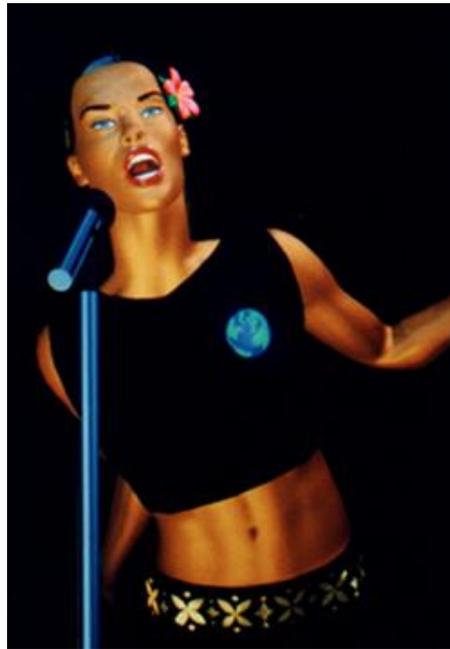
Segundo Suppa (2006, p. 145), Thalmann (2004, p. 46) e Sturman (1998, p. 39), Waldo representou um marco para a captura de movimentos e a popularização dos produtos advindos desta tecnologia, pois permitiu a visualização de um personagem tridimensional em tempo real, interagindo com atores reais. Suas limitações eram principalmente as articulações do exoesqueleto ao qual estava associado (Figura 58). Este tipo de abordagem, utilizando sensores eletromecânicos e um exoesqueleto, incentivou o desenvolvimento dos sistemas mecânicos de capturas de movimentos.¹¹

Em 1989, surgiu Dozo (Figura 59), a primeira cantora virtual animada por captura de movimentos. Criada pela recém-criada companhia Kleiser-Walczak, teve grande recepção por parte do público. Em termos tecnológicos, seus movimentos foram capturados através da utilização de um sistema óptico (o sistema *motion analysis*) e, apesar de seus

¹¹ Em 2013, a série britânica *Black mirror*, lançou um episódio chamado “Waldo moment”, em homenagem ao Waldo da década de 1980. Nessa história uma marionete 3D animada, controlada por um comediante, passa a atuar na política do país. O dispositivo eletromecânico representado na série para controlar o personagem era uma referência ao usado em Waldo C. Graphic.

movimentos serem considerados fluidos para a época, sua captura apresentou o mesmo problema dos sistemas ópticos anteriores: a oclusão dos marcadores, que limitou o uso de apenas trinta segundos de movimentos que precisaram ser repetidos durante o clipe musical. (MENACHE, 2011, p. 12).

Figura 59 - Dozo, personagem de computação gráfica.



Fonte: <https://www.youtube.com/watch?v=8ovn8qRezPA>. Acesso em: 22 Dez. 2017.

Em 1991, a produtora francesa Videosystem criou o personagem Mat, o fantasma, para o programa infantil Canaille Pelluche. Mat era um personagem digital tridimensional, gravado com um fundo verde para depois ser inserido sobre outro vídeo gravado por atores, simulando uma interação. Este personagem foi manipulado através de diversos dispositivos como os *joystickers*, o *tracking* magnético, entre outros (STURMAN, 1994).

No ano seguinte a empresa SimGraphics desenvolveu um sistema de *tracking* facial chamado *face waldo*. Tal como o sistema que controlava Waldo, este utilizava dispositivos eletromecânicos que possibilitavam a transferência dos movimentos de um ator diretamente para o personagem, em tempo real (posteriormente, a empresa focou neste nicho de mercado). Seu primeiro personagem de sucesso foi o bombeiro Mario da Nintendo™ (Magnenat-Thalmann e Thalmann, 2006, p. 46).

Figura 60 - Bombeiro Mario aparece no Rockfeller Center.



Fonte: <http://www.simg.com/Video-1.html>

Para além da utilização e do desenvolvimento dessas tecnologias pela indústria da computação gráfica na qual os movimentos eram transmitidos diretamente aos personagens, as pesquisas com sistemas de capturas de movimentos e de visão computacional foram se aprimorando, principalmente na área de desenvolvimento de *softwares* para o reconhecimento de poses. Em 1992, Yamato et al. (1992), apresentaram um método de reconhecimento de ações humanas que utilizava modelos ocultos de Markov (*Hidden Markov Models* - HMM) em vídeos.

Os HMMs são modelos estatísticos usados (dentre outras aplicações) em sistemas de inteligência artificial para a análise de séries temporais (como, por exemplo, a análise de vídeos) e que procuram antecipar o estado de um determinado parâmetro (oculto), a partir dos parâmetros observáveis. Esses modelos procuram prever qual o próximo estado de um determinado sistema baseado apenas no estado anterior. Tal decisão foi tomada para determinar as possibilidades de que um novo estado aconteça através da determinação de pesos para cada novo estado que indicarão a maior ou a menor probabilidade de que determinado evento aconteça (WIKIPEDIA).

Em seu experimento, Yamato et al. (1992) destacaram que esses modelos podem ser os mais apropriados para o reconhecimento de gestos humanos, já que podiam ser aplicados

com eficiência a dados temporais-sequenciais. Além disso, eles abriam outras possibilidades de aplicações, por exemplo, em sistemas de aprendizagem de reconhecimento.

O sistema por eles desenvolvido, embora funcionasse bem com uma base de vídeos específica criada para seus experimentos, apresentou uma *performance* muito aquém da esperada, quando o sistema foi testado com vídeos de indivíduos diferentes daqueles apresentados em suas bases.

Esta questão foi posteriormente debatida por outras pesquisas e se tornou um desafio a ser ultrapassado, quando sistemas de segurança baseados em vídeo passaram a ser desenvolvidos. Como estes precisavam reconhecer ações de indivíduos, independentemente da base de vídeos (ou da posição da câmera) que os alimentam, diversas pesquisas passaram a direcionar suas atenções para a solução do problema.

Em 1993, a Acclaim, uma fabricante de *games*, “surpreendeu o público com uma animação realista e complexa com dois personagens feita completamente com a captura de movimentos” (STURMAN, 1994). O sistema por eles criado permitia o *tracking* de até 100 pontos, simultaneamente, o que aumentou a precisão da reprodução de movimentos.

Além disso, a Acclaim foi responsável pela criação de dois formatos de arquivos importantes para a indústria de *games* e para o desenvolvimento da tecnologia em captura de movimentos, o ASF (*Acclaim Skeleton File*) e o AMC (*Acclaim Motion Capture data*). Estes arquivos separavam os dados da configuração do esqueleto dos dados dos movimentos capturado, que consistiam na rotação das juntas do esqueleto (KITAGAWA e WINDSOR, 2008, p. 181).

Até o momento, as capturas de movimentos deveriam ser feitas com atores que possuíssem estatura similar aos personagens capturados, pois os sistemas de capturas faziam a correspondência entre as posições das juntas. Devido à separação de esqueleto e movimento criada pela Acclaim, os movimentos capturados puderam ser aplicados em personagens com tamanhos diferentes. Esta forma de tratamento da captura permitiu que o mesmo movimento pudesse ser aproveitado em diversas versões de um *game* e ou no mesmo *game* em personagens diferentes, tornando a aplicação mais eficiente para a indústria.

As aplicações da captura de movimentos não se restringiram ao setor do audiovisual e aos sistemas de segurança. O setor militar, tradicionalmente um dos que mais avançaram neste campo, também se valeu da tecnologia de captura de movimentos. Em 1995, a missão espacial EuroMIR 95, cujo objetivo era levar astronautas europeus à estação espacial Russa, utilizou um sistema passivo de captura de movimentos aplicado aos tripulantes para monitorar seus movimentos no espaço (ALIVERTI et al., 2011, p. 26).

Este sistema seria posteriormente aprimorado para a utilização na Estação Espacial Internacional (*International Space Station – ISS*) com o objetivo de observar e analisar os movimentos do corpo humano, quando exposto às condições de microgravidade. Tais estudos mostraram a confiabilidade dos sistemas óptico-eletrônicos no acompanhamento de movimentos de pessoas.

Na metade da década de 1990, outras abordagens começaram a ser utilizadas com o uso de múltiplas câmeras para reconhecer o corpo humano e acompanhar seus movimentos. Neste sentido, Gavrilá e Davis (1996) descreveram um sistema computacional que aborda o problema da recomposição do movimento humano. Em sua pesquisa, utilizaram várias câmeras para recompor a pose do corpo em três dimensões (3D), através de um modelo. Utilizando a inteligência artificial, este modelo era, então, comparado a uma base de dados composta por 2500 imagens de diversas ações humanas (por eles chamada de *humans in action*), gravadas por meio de quatro vistas ortogonais. Esta comparação permitiu o reconhecimento de diversas poses nos vídeos usados como entrada do sistema.

No final da década de 1990, Gavrilá publicou o artigo “The visual analysis of human movement: a survey” (GAVRILA, 1999), apresentando os estudos mais recentes no campo da visão computacional. Este ressaltava a importância do reconhecimento de poses humanas (problema ainda não resolvido pelos sistemas existentes na época), suas aplicações em diversos campos do conhecimento, como a realidade virtual, os sistemas inteligentes de segurança, as interfaces de usuário avançadas (aquelas controladas por gestos), as análises de movimentos e a codificação baseada em modelo.

Gavrilá desmembrou cada uma dessas áreas em subcampos específicos, sendo que a realidade virtual acabou ficando composta por: mundos virtuais interativos, *games*, estúdios virtuais, animação de personagens e teleconferência (filme, propaganda e uso caseiro) (GAVRILA, 1999, p. 83). Apesar da atuação do designer poder ser realizada em qualquer

dos campos principais do conhecimento apontados por ele (através do design de interfaces), percebe-se nessa subdivisão que o domínio no qual ele pode atuar com maior representatividade é justamente aquele denominado *realidade virtual*, no qual ele será responsável pelo desenvolvimento de projetos que envolverão modelos, análises de usuários, testes e projetos visuais. É nesse importante campo que ele se insere e pode contribuir para os estudos de visão computacional. Não à toa uma das subáreas tem a aplicação desta tese.

Para Gavrila (1999), os sistemas de captura de movimentos podem ser divididos naqueles cuja abordagem são:

- 2D sem formas de modelos explícitos nos quais são utilizados modelos estatísticos simples, sem a tentativa de reconhecimento das partes do corpo para o reconhecimento de ações humanas;
- 2D com formas de modelos explícitos nos quais o corpo humano é subtraído do fundo e suas partes segmentadas para que o sistema reconheça o corpo e, através da movimentação de suas partes, as ações a ele relacionadas e;
- 3D, a partir de vídeos, nos quais existe a tentativa de reconstruir as juntas do corpo humano e descobrir seus ângulos em relação a um ponto central. Neste caso, não são consideradas as técnicas que usam marcadores ou sensores ativos.

Para cada um dos sistemas apresentados, havia soluções em termos de algoritmos, cada qual com desafios ainda não solucionados (muitos, até hoje, ainda carecem de solução), que o autor dividiu de acordo com sua abordagem de sistemas:

- para os sistemas 2D foram: a “segmentação da imagem, o uso de modelos, o *tracking* versus inicialização, múltiplas pessoas, oclusão e custo computacional” e “mostrar que as abordagens podem ser escalonadas para permitir a recuperação da pose, através de um amplo conjunto de movimentos advindos de diferentes pontos de vista” (GAVRILA, 1999); e

- para os sistemas 3D, os sistemas ainda deveriam tratar da aquisição de modelos, de detalhes da modelagem e da obtenção do chão verdadeiro.

Este último foi um dos fatores limitadores para o uso do Kinect da Microsoft pela pesquisa aqui apresentada e levou-a a abandonar essa solução para o reconhecimento de movimentos.

Apesar destas limitações, no início do século XXI, os sistemas de captura de movimentos já estavam mais desenvolvidos, embora algumas questões ainda não estivessem resolvidas. Os sistemas comerciais já faziam parte do dia a dia de centros de pesquisa e estúdios de audiovisual. Entre eles, destacam-se os sistemas baseados em vídeo, em áudio, em sensores magnéticos, em sistemas de emissão de luzes e os eletromecânicos. A seguir, serão apresentados os sistemas existentes na época e suas vantagens e desvantagens. A partir de sua apresentação, será possível notar que alguns deles se complementam. Outros não podem ser aplicados a determinadas situações e um foi descontinuado, devido à sua pouca eficiência.

Ao final da breve explicação sobre seus funcionamentos, retomaremos o histórico da captura de movimentos.

3.3 Tecnologias usadas em capturas de movimentos

As tecnologias utilizadas atualmente para capturas de movimentos foram desenvolvidas nos últimos quarenta anos. A descontinuidade de algumas delas se deve ao surgimento de tecnologias com maior precisão e ao barateamento de outras. Algumas se tornaram populares, devido à sua utilização no mercado de *games*, do cinema e da realidade virtual. Atualmente, as mais utilizadas são a captura de movimentos óptica, a eletromagnética, a mecânica (KITAGAWA e WINDSOR, 2008, p. 8) e a que é feita através de vídeo.

A captura de movimentos óptica também é realizada através de câmeras, porém foi separada da que é por vídeo, porque, nesta pesquisa, verificou-se que a captura de movimentos óptica exige que alguns marcadores sejam colocados no corpo do indivíduo, enquanto que esta última é realizada apenas pela dedução do sistema, quanto à posição do indivíduo, sem a interação direta do sistema com ele.

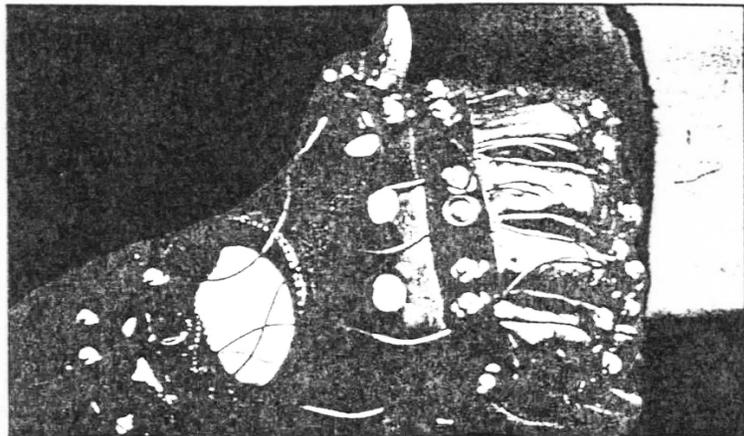
Alguns sistemas tinham por objetivo capturar apenas as mãos para o aprimoramento da interação do homem com o computador. Segundo Sturman e Zeltzer (1994), a interação com o computador é realizada principalmente pelas mãos. Entretanto os dispositivos usados criados para isto são grosseiros e deveriam ser substituídos por outros que permitissem a leitura direta das mãos. Tendo este objetivo em mente, alguns dispositivos *hardwares* foram desenvolvidos no sentido de aprimorar essa interação. Atualmente, a maior parte dos sistemas de captura de movimentos existentes consegue capturar tanto o corpo quanto as mãos.

Apesar de suas características diferentes, algumas soluções encontradas em um sistema podem ser aplicadas em outros. Abaixo, descrevemos cada um deles, bem como suas vantagens e desvantagens em relação ao seu uso em personagens *cartoons*. Por fim, justificaremos a utilização de uma delas na pesquisa.

3.3.1 Sistemas de captura de movimentos acústica e por fibra ótica

As capturas através de fibras óticas foram utilizadas nos *hardwares* para capturas de movimentos dos dedos – os sistemas de luvas – no início da década de 80 (*glove system*, [Sturman e Zeltzer, 1994]) . Seu princípio de funcionamento era baseado na emissão e na recepção de raios luminosos através das fibras óticas, que eram posicionadas sobre os dedos. De acordo com seus ângulos de abertura e de fechamento, a fibra ótica dobrava mais ou menos, fazendo com que a luz que a atravessava chegasse com uma maior ou menor intensidade aos sensores. Através do cálculo dessa intensidade, era possível determinar o tipo de movimento realizado (Figura 61).

Figura 61 - Luva de fibra ótica.



Fonte: Sturman e Zeltzer (1994, imagem contrastada para melhor visualização).

Já os sistemas acústicos foram baseados na técnica de triangulação de sinal. Estes são compostos por fontes acústicas emissoras e por sensores receptores de sinais (microfones). As fontes emissoras são posicionadas no corpo humano, enquanto as fontes receptoras são colocadas no ambiente no qual se dará a captura de movimentos, em locais pré-determinados. O posicionamento de cada um dos sensores é determinado por meio da intensidade do sinal capturado pelas fontes receptoras.

Uma das limitações desse tipo de sistema é a acústica do ambiente em que se encontra, pois é comum ocorrer reflexões e absorções de ondas acústicas nas paredes, que interferem como ruídos nas capturas.

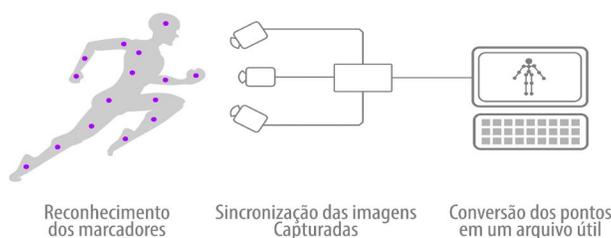
3.3.2 Captura de movimentos óptica

A captura de movimentos óptica foi uma das primeiras a serem desenvolvidas. Ela se baseava na colocação de pontos emissores ou refletores de luz ao longo do corpo humano e em um sistema de câmeras que visualiza esses pontos para remontá-los em um *software*. Sturman e Zeltzer (1994) relatam que o uso desse tipo de sistema remonta à década de 80.

Os pontos de luz são chamados de *marcadores*, que podem ser ativos (quando emitem luz) ou passivos (quando refletem luz). Os ativos são aqueles que usam LEDs presos nas roupas usadas pelos indivíduos sendo capturados. Necessitam de uma fonte de

energia para alimentá-los e cada um possui um número identificador. Este facilita o processamento dos movimentos pelo sistema. Funcionam através da sincronização das imagens capturadas e da reconstrução do corpo (em 3D), através de um *software* (Figura 62).

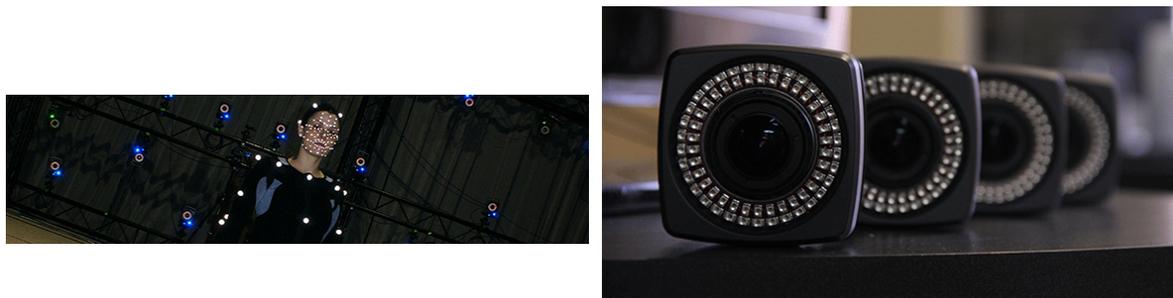
Figura 62 - Esquema de captura óptica de movimentos.



Fonte: Ilustração do autor.

Os marcadores passivos são posicionados nas juntas do indivíduo e iluminados por uma fonte de luz infravermelha que não pode ser vista pelo olho humano, mas é capturada por câmeras sensíveis ao espectro da luz. As fontes emissoras de luz infravermelha ficam posicionadas ao redor das lentes das câmeras para garantir que o marcador reflita o máximo de luz possível, aumentando a precisão de reconhecimento do sistema (Imagem 2).

Figura 63 - Sistema de captura de movimentos passivos no qual a luz infravermelha é emitida por LEDs ao redor da câmera.



Fonte: <https://www.vicon.com/what-is-motion-capture>. Acesso em: 07 Jan.2017.

As vantagens desse sistema são a alta precisão com que se capturam os marcadores e a grande flexibilidade de movimentos por parte dos atores. As desvantagens são: a oclusão dos marcadores, que pode acontecer de acordo com os movimentos realizados; o custo dos equipamentos (*hardware* e *software*); a necessidade de um grande processamento computacional paralelo ou posterior à captura, tanto para a inicialização do sistema quanto para a análise da grande quantidade de dados adquiridos; a exigência de um local grande o

suficiente para a colocação das câmeras; e a ocorrência de oclusão e a dificuldade de reconhecimento de marcadores pelo sistema, quando há mais de um indivíduo sendo capturado.

Um exemplo de aplicação dessa tecnologia fora da indústria do cinema é na área da saúde, na qual os pesquisadores procuram mensurar as atividades mecânicas dos membros do corpo humano e a fadiga causada pelo esforço (Wang, Frame, Ozimek, Leib, e Dugan, 2013).

3.3.3 Captura de movimentos eletromecânica

Os sistemas de captura de movimentos eletromecânicos são aqueles que contam com um exoesqueleto acoplado ao corpo do indivíduo (Figura 64). Os dados são a eles transmitidos, de acordo com a movimentação deste último. Já o exoesqueleto é formado por hastes em cujas junções são colocados sensores que, de acordo com a rotação corporal realizada, variam a eletricidade que por eles passa, o que possibilita, através dessa variação, calcular sua rotação nas juntas do corpo.

Estes sistemas são extremamente precisos, já que se encontram diretamente presos ao corpo humano e, devido a essa característica, conseguem capturar mesmo os movimentos mais sutis. A Figura 64 mostra sistemas eletromecânicos acoplados ao corpo humano.

Figura 64 - Sistemas eletromecânicos de captura de movimentos.



Fonte: <https://www.dextarobotics.com/> e <https://www.youtube.com/watch?v=LgDbCxSg7hA>. Acesso em: 15 Mar. 2016.

É possível notar as desvantagens desse tipo de sistema, a saber: a presença e o peso relativo de um exoesqueleto que pode limitar os movimentos dos atores e a impossibilidade de realizar determinados movimentos como, por exemplo, uma posição deitada ou agachada que poderia danificar o sistema (KITAGAWA; WINDSOR, 2008, p. 12).

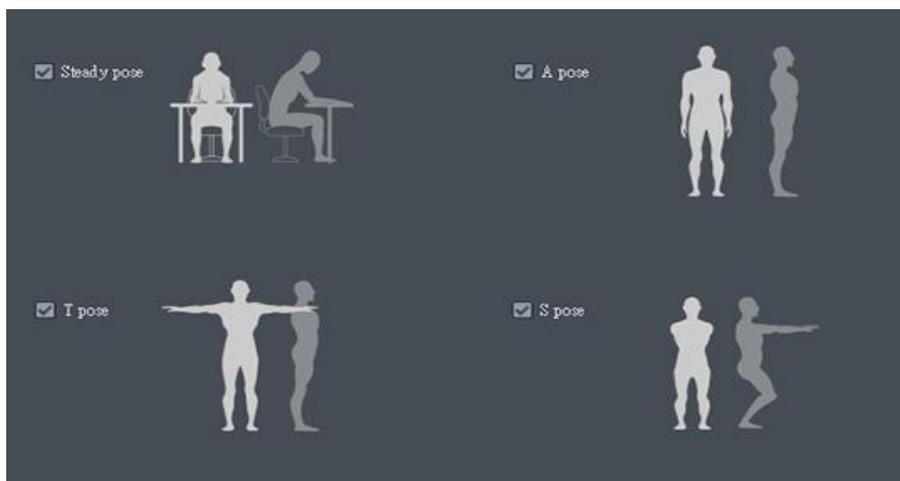
3.3.4 Captura de movimentos eletromagnética

De forma similar aos sistemas ópticos, os sistemas de captura eletromagnéticos também exigem a colocação de sensores no corpo dos indivíduos. Entretanto, a captura de movimentos é feita através de magnetômetros e acelerômetros que capturam as rotações e posições de cada junta do corpo.

Para trabalhar de forma satisfatória e precisa, é necessário fazer uma calibragem precisa do sistema onde este reconhece as posições relativas de cada parte do corpo e as distâncias entre cada uma delas. Após a calibragem, o sistema está pronto para o uso.

A Figura 65 mostra a tela de calibração de um destes sistemas, no qual o usuário precisa se colocar em cada uma das poses ilustradas para que o sistema tenha precisão na captura de movimentos.

Figura 65 - Tela de calibração do Neuron Mocap, sistema eletromagnético de captura de movimentos.



Fonte: Tela de calibragem do sistema Neuron Mocap.

As vantagens dos sistemas magnéticos são: a velocidade de calibração do sistema, bastante rápida comparada aos demais (nesta pesquisa, levamos, em média, trinta minutos

para a calibragem do sistema, já calculando os erros ocorridos); o custo; a não limitação de espaço de utilização, quando usamos um sistema com conexão sem fio; e a não ocorrência de problemas de oclusão, já que os sensores não precisam ser visualizados por câmeras.

As desvantagens desses sistemas são as interferências magnéticas do ambiente que podem alterar a precisão dos sensores e fazer com que os movimentos sejam capturados com ruídos. Frequentemente, é necessário realizar calibrações entre as sessões de captura, pois as interferências são constantes e atrapalham o sistema. Há alguns anos atrás, eles sofriam interferências das próprias estruturas dos prédios onde eram realizados os testes. Os materiais ferrosos utilizados nas construções geravam interferências difíceis de serem eliminadas. Embora os sistemas atuais sofram menos com esses ruídos, eles ainda são bem frequentes.¹²

Apesar das limitações, consideramos adequadas a relação custo-benefício e a precisão desse tipo de sistema para aplicações em pesquisas em animação e captura de movimentos. Assim, este foi o sistema escolhido para utilização na pesquisa.

3.3.5 Captura de movimentos por vídeo

Como apresentado anteriormente, a captura de movimentos por vídeo foi a primeira técnica a ser explorada no âmbito da pesquisa acadêmica. Dentre suas aplicações mais recorrentes, estão os sistemas de segurança e de monitoramento, que precisam reconhecer os movimentos de seres humanos através de imagens de câmeras de vídeo, os sistemas voltados para a área da saúde e, nas últimas duas décadas, os sistemas voltados para a indústria do audiovisual, em especial, o cinema e a área de *games*.

Como apontado anteriormente, os sistemas de captura óptica de movimentos também são baseados em vídeos. A diferença entre esses sistemas e os abordados nesta seção está na maneira como os movimentos são capturados. Em um deles, há a necessidade de uma implantação de marcadores no indivíduo (sistemas ópticos) e no outro, não (sistemas baseados em vídeo). Estes últimos exigem o uso de sistemas mais inteligentes e computadores com maior capacidade de processamento.

¹² Em julho de 2018 fomos até o Centro de Tecnologia do Movimento (CEMOV), do Instituto Nacional de Tecnologia. Ao conversarmos com os pesquisadores daquele centro, fomos informados que eles pararam de usar os sistemas de capturas eletromagnéticos, devido às interferências do ambiente sobre os sensores. Atualmente, eles utilizam um sistema óptico de alta precisão.

Apesar de tais limitações, muitas pesquisas são realizadas com esse tipo de equipamento, visto que eles possuem aplicabilidades mais abrangentes e necessitam de uma menor intervenção sobre o usuário do sistema. Outra vantagem do uso deste tipo de sistema refere-se às necessidades técnicas de sua instalação. Enquanto os sistemas ópticos são facilmente reconhecidos no ambiente, os baseados em vídeo podem ficar escondidos, detectando as reações de usuários, suas ações e movimentos, enquanto sugerem modos de atuar para outros sistemas ou determinam que tipo de atitude deve ser tomada por outros usuários.

Uma das premissas para o funcionamento destes equipamentos é a separação entre o indivíduo a ser capturado e o ambiente que o cerca. No momento em que essa etapa é concluída, o sistema prossegue com a segmentação da imagem resultante para que cada parte possa ser comparada a um modelo previamente armazenado e, após a efetiva comparação, para que sejam determinadas quais são as partes do corpo na imagem. Essa etapa é normalmente chamada de *labeling* (criação de “rótulos” de identificação ou “rotulação”) na qual cada parte do corpo recebe um nome. A partir daí, o sistema realiza o *tracking* (acompanhamento) delas para traduzir seus deslocamentos em valores que podem ser de posições, de rotações ou ambas. A maior parte dos sistemas utiliza a cabeça, o tronco e os membros como referências para gravar os dados.

As etapas iniciais podem ser concluídas com esse reconhecimento, mas para que os dados tenham utilidade para aplicações específicas é necessário realizar o processamento e o reconhecimento de movimentos humanos – sua parte complexa e uma das temáticas desta pesquisa.

Embora a pesquisa na área de visão computacional remonte à década de 70, pode-se afirmar que a popularização de suas tecnologias se deu com seu uso em *smartphones* e por meio dos *games*, em especial com o surgimento do console Xbox, da Microsoft, que trazia consigo o Kinect, em 2009.

O Kinect, de fato, representou uma quebra no paradigma da captura de movimentos em vídeo, por ser um equipamento barato e acessível, quando comparado aos demais sistemas presentes no mercado. Após seu lançamento, a Adafruit, uma empresa de Nova York que vendia equipamentos *open hardware* ofereceu uma “recompensa” para quem conseguisse criar *drivers open source* para o Kinect (BORENSTEIN, 2012), objetivando

tornar esse *hardware* acessível a qualquer um que necessitasse desenvolver aplicações para utilizá-lo. Dessa forma, um ano após o desafio, surgiu o OpenNI¹³ (*Open Natural Interaction*) – um conjunto de *drivers* abertos – fruto do trabalho conjunto de pesquisadores e *hackers*.

A existência do OpenNI representou um desafio para a Microsoft, já que, através dele, se tornou possível utilizar grande parte das funcionalidades de seu *hardware* em qualquer computador. Percebendo que o desenvolvimento de aplicações para o Kinect poderia se tornar um grande nicho de mercado e que deveria investir nele antes de perder o pioneirismo no setor, a Microsoft resolveu disponibilizar, de forma gratuita, as bibliotecas de programação do Kinect, permitindo a qualquer desenvolvedor utilizar seu *hardware* e seu *software* em qualquer máquina na qual seu sistema operacional (o WindowsTM) estivesse instalado.

Esta visão empresarial possibilitou o desenvolvimento de várias pesquisas e aplicações que utilizaram o Kinect, de forma comercial e experimental. Algumas delas foram bem-sucedidas em diversos meios como na medicina, uma área em que foram criados sistemas que permitiam aos médicos observar exames na sala de cirurgia sem a necessidade de esterilizar as mãos¹⁴ (Figura 66).

¹³ O OpenNI pode ser encontrado em <https://structure.io/openni>.

¹⁴ https://msdnshared.blob.core.windows.net/media/MSDNBlogsFS/prod.evol.blogs.msdn.com/CommunityServer.Blogs.Components.WeblogFiles/00/00/01/32/81/7485.Kinect-medicine_486px.jpg

Figura 66 - Cirurgiões verificam os exames de imagem de um paciente sem precisar tocar nas imagens. Esta tecnologia facilitou o trabalho nas salas de cirurgia.



Fonte: https://msdnshared.blob.core.windows.net/media/MSDNBlogsFS/prod.evol.blogs.msdn.com/CommunityServer.Blogs.Components.WeblogFiles/00/00/01/32/81/7485.Kinect-medicine_486px.jpg. Acesso em: 17 Dez. 2016.

Apesar da popularidade do Kinect, ele é o produto de outras pesquisas que foram realizadas no início dos anos 2000 às quais abordaremos, a seguir.

3.4 Poses, gestos e ações: o reconhecimento de atividades humanas

“Pose” e “ação” são termos comumente usados por animadores ao se referirem ao planejamento em animação. Esses profissionais trabalham criando a linguagem expressiva do cartum e do movimento mediante a qual transmitirão uma ideia através de poses. Qualquer ação realizada por um personagem animado passa pela criação de uma pose.

Na visão computacional, existe a necessidade de definir os elementos básicos mínimos para o reconhecimento de atividades humanas. Para as máquinas e os *softwares* utilizados por elas, é importante partir de um conjunto de regras pré-definidas que possibilitariam aos programadores tratarem os dados que recebem e devolver uma resposta ao usuário. No caso da visão computacional, os elementos básicos encontrados na literatura são as poses, os gestos e as ações, que apresentaremos a seguir.

Poucos são os autores que definem o que é uma pose, mas muitos utilizam definições de uso prático na área da visão computacional de onde é possível extrair um sentido para o termo. Para Holte, Tran, Trivedi, e Moeslund (2012, p. 540): “A estimativa e *tracking* de poses baseadas em vídeo de corpos humanos articulados é um problema de estimar parâmetros cinemáticos do modelo do corpo (como posições e ângulos das juntas), advindas de uma imagem estática ou de uma sequência de vídeo”. Já Shotton et al. (2013, p. 116) definem a pose como sendo simplesmente “as posições 3D das juntas do corpo”. Poppe (2007, p. 4) afirma que:

A estimativa de uma pose do corpo humano, ou estimativa de pose, brevemente, é o processo no qual a configuração das partes do corpo é estimada da entrada de um sensor. Quando as poses são estimadas ao longo do tempo, o termo análise de movimento humano é utilizado (POPPE, 2007, p. 4).

Logo, podemos definir a “pose” como o elemento básico do reconhecimento de movimentos, o conjunto de relações espaciais entre as juntas do corpo que resultam em uma forma conhecida.

Em animação, as poses são os elementos básicos que determinam a forma como o personagem se movimentará. Elas são o primeiro elemento planejado no processo de execução de uma cena que permitirá ao espectador ou ao usuário de um sistema (como em um *game*) compreender como um personagem está realizando um movimento.

A importância de definir o que é uma pose reside no fato de que nem sempre o personagem estará se movimentando exageradamente em uma cena. Por exemplo, quando sentado em um sofá assistindo televisão, ele estará em uma pose, mas não necessariamente se movimentando. Para os fins desta pesquisa, as poses que interessam são aquelas que, quando agrupadas em sequência, constituem um movimento, ou seja, um personagem estático não será aqui analisado.

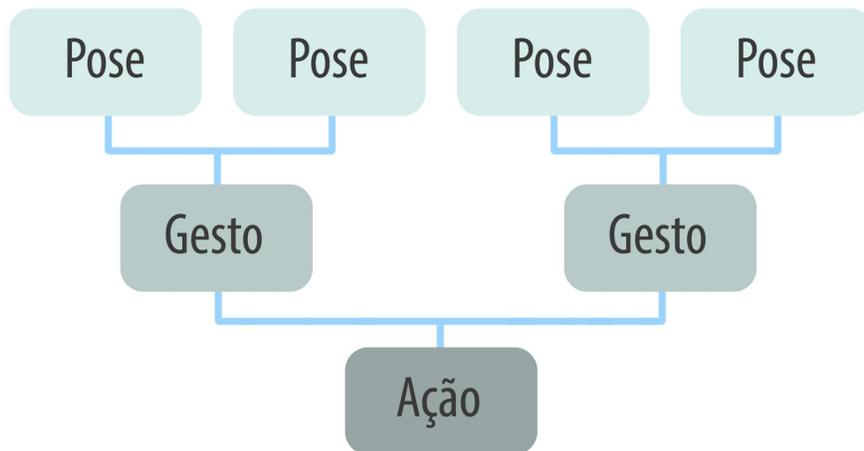
Os gestos são, segundo Aggarwal (2011), “movimentos elementares da parte do corpo de uma pessoa e [...] os componentes atômicos que descrevem o movimento significativo de uma pessoa”. Esticar as pernas e os braços é um exemplo de gesto. Abrir as mãos é um gesto. Contrair os dedos representa um gesto.

O mesmo autor define que o conjunto coordenado de gestos de um indivíduo compõe suas ações. Dessa forma, as ações de esticar e comprimir braços e pernas, quando

estes são articulados de maneira coordenada, podem ser traduzida nas ações “caminhar”, “correr” ou “desferir um golpe”.

Estes três elementos são considerados, assim, as estruturas básicas para o reconhecimento de ações de um indivíduo (Figura 67).

Figura 67 - Relação entre poses, gestos e ações.



Fonte: Ilustração do autor (2016).

Estas definições serão retomadas mais adiante, quando forem discutidas as características do reconhecimento de poses e ações realizadas, de maneira automática, por sistemas de *motion capture*. Desde já, pode-se dizer que toda pose irá gerar uma ação que pode ser estática (ou semi-estática, já que nenhum ser vivo fica, em algum momento, completamente parado) ou em movimento. Determinadas poses podem gerar ações exageradas, enquanto outras podem gerar ações sutis. Durante o planejamento de animação, as poses determinam o tipo de ação que será realizado. Durante a análise de uma captura de movimentos, elas indicarão o movimento realizado. Uma refere-se ao futuro, a outra ao passado. Compreender essas relações é a base para o projeto de sistemas de reconhecimento de movimentos.

3.5 A evolução no reconhecimento de movimentos

O reconhecimento de movimentos e seu desdobramento, o reconhecimento de ações humanas, continuam a ser um campo bastante amplo de atuação na área de visão

computacional. Na virada do século XX para o XXI, o acesso à internet tornou-se mais amplo. Grupos de pesquisa divulgavam seus resultados *online* e diversas revistas passaram a ser publicadas no meio digital. Ao mesmo tempo, houve um barateamento cada vez maior dos computadores e de seus periféricos que favoreceu igualmente o avanço das pesquisas na área da visão computacional.

No início dos anos 2000, Moeslund e Granum (2001), dois renomados pesquisadores na área de visão computacional, relataram o crescente interesse no campo do reconhecimento de atividades humanas através de uma pesquisa bibliográfica exploratória, apresentando as principais pesquisas desenvolvidas desde os anos 1980 até o ano 2000 (MOESLUND; GRANUM, 2001).

Sua abordagem apresentou o desenvolvimento das pesquisas em sistemas de reconhecimento de ações, focando naqueles baseados em vídeo. Eles indicam três áreas principais em que o uso desses sistemas é explorado: o da *vigilância* em que são monitoradas ações de indivíduos e cujo resultado é a segurança do ambiente; o do *controle*, em que o MoCap é usado como uma interface para controlar *games*, animações ou sistemas locais, remotamente; e o da *análise* detalhada de dados para estudos clínicos.

Estas áreas se interpenetram, mas, em geral, são estudadas como campos isolados do conhecimento nos quais são comumente apresentados os sistemas de captura de movimentos (*hardwares* e *softwares*) e as bases de movimentos usados como entradas dos sistemas. As bases de movimentos foram criadas para que os pesquisadores da área de reconhecimento de ações pudessem ter parâmetros de comparação entre suas pesquisas. Usando as mesmas bases, diversos grupos podem apresentar o nível de confiabilidade e precisão de seus sistemas. Entre essas bases, pode-se encontrar todo tipo de movimentos. Desde aqueles retirados de filmes (Kuehne, Jhuang, Garrote, Poggio, e Serre, 2011) até aqueles que procuram fazer o reconhecimento de objetos em cenários (Russell, Torralba, Murphy e Freeman, 2008).

Assim, para a área da fisioterapia, os vídeos contendo o deslocamento das juntas do braço são relevantes. Nos esportes, aqueles que apresentam jogadores realizando ações são mais relevantes (por exemplo, para a análise de uma jogada de vôlei, a forma como o jogador se projeta para saltar na rede será utilizada para análise). Para os sistemas de

segurança, é importante a existência de vídeos contendo gestos sutis, como “uma mala sendo abandonada no aeroporto” ou um pacote sendo deixado em uma lixeira.

Dessa forma, pode-se dizer que cada pesquisa relacionada a uma área do conhecimento precisará de um conjunto de movimentos característicos para seu uso, ou seja, será possível, através do conhecimento da base de movimentos utilizada para a pesquisa, saber a área do conhecimento à qual ela se refere e, assim, propor novas bases para o estudo de ações.

Os sistemas analisados por Moeslund e Granum (2001) procuram resolver problemas de reconhecimento de ações. Segundo eles, os sistemas tratam os dados de entrada (os vídeos) através de um conjunto de etapas semelhantes que seguem os seguintes passos: a *inicialização* (calibragem da câmera, reconhecimento do ambiente e inicialização do modelo), o *tracking* (reconhecimento das partes do corpo e seu acompanhamento ou, ainda, o acompanhamento de objetos), a *estimativa da pose* (cada parte do corpo está hierarquicamente conectada e essas conexões são reconhecidas pelo sistema e convertidas em poses, através das relações entre as partes que estão sendo acompanhadas [*tracking*]) e o *reconhecimento de ações* (interpretação dos dados de entrada pelo sistema). Esta última etapa é realizada através de uma série de algoritmos que usam técnicas de probabilidade para estimar as poses apresentadas e, com isso, identificar as ações. Nos sistemas pesquisados, uma ação se constitui de um conjunto de poses reconhecidas em sequência.

Compreendemos que, até a primeira década do século XXI, o desafio desses sistemas baseados em vídeo era reconhecer um indivíduo e realizar a separação do ambiente no qual se encontrava. Considerando que muitas das bases utilizadas para as pesquisas nessas áreas eram vídeos obtidos de câmeras em ambientes reais, estes eram, em geral, uma imagem em movimento muito ruidosa para o sistema, o que consistia em um desafio para os pesquisadores.

Outras pesquisas, identificando as dificuldades em reconhecer indivíduos em movimento, quando as câmeras que os filmavam também estavam em movimento, partiram do pressuposto de que reconhecer movimentos com câmeras paradas trazia resultados mais consistentes, tanto em aplicações mais gerais quanto para estimar poses humanas (LEE; COHEN, 2006) em aplicações mais específicas, como o reconhecimento de atletas em movimento (WEINHANDL et al., 2010).

Um dos equipamentos que utiliza esta abordagem é o Kinect, da Microsoft, que realiza as etapas enumeradas por Moeslund e Granum (2001) para o reconhecimento de poses e somente as realiza quando o equipamento está parado, ou seja, o fundo da imagem capturada é estático. Assim, os algoritmos por ele utilizados reconhecem o fundo dos vídeos como uma imagem que não se alterará através do tempo, facilitando a separação do indivíduo do ambiente em que está localizado.

O processamento de imagens com a comparação entre dados de entrada em uma base pré-gravada já tinha sido experimentada antes. Sistemas de Karaokê utilizavam algoritmos para o reconhecimento de poses e ações, através da comparação com *templates* (modelos) pré-gravados (Li, Zhang, e Liu, 2008) em suas memórias. Quando o usuário se movimentava em frente a eles, sua imagem era capturada e comparada com cada um dos *templates* para determinar as poses e, assim, reconhecer ações.

Outros algoritmos baseavam-se em histogramas de orientações de uma imagem (MOESLUND; GRANUM, 2001). Os histogramas são uma forma gráfica de representar dados que mostra onde eles ocorrem com maior frequência. Neste caso, os sistemas reconheciam as poses, determinavam sua probabilidade de pertencerem a uma categoria de poses (corrida, caminhada, parada, etc.), extraíam esses histogramas e, de acordo com a distribuição deles, apontavam o movimento realizado. Este procedimento ainda é utilizado em sistemas de reconhecimento de ações, pois permite que as imagens possam ser novamente processadas e reconhecidas em subcategorias (como, por exemplo, dentro de uma base de dados de uma caminhada, pode haver uma mais rápida e outra mais lenta).

Dentre as lacunas apontadas pela revisão de Moeslund e Granum (2001), estão a falta de um “alfabeto” constituído por “entidades-movimento” que facilitaria o reconhecimento de ações, pois “transformariam o problema de estimativa de uma pose em um problema de reconhecimento, por exemplo, o reconhecimento de uma sequência de símbolos”.

Outra questão que ainda é atualmente um desafio para o reconhecimento de ações é a necessidade de identificação do estado (posição) inicial do indivíduo. Esse tipo de situação não ocorre quando do reconhecimento de poses pelo cérebro humano, ou seja, no ambiente natural, e não deveria ocorrer no artificial, visto que, a depender da aplicação, o

sistema utilizado não deveria considerar que o indivíduo está posicionado em frente a ele para, a partir dessa pose inicial, iniciar o *tracking* de suas juntas.

Essa etapa de ajuste não interessa a esta pesquisa, mas sim a etapa subsequente, o de reconhecimento de poses e ações. A partir destas, será explorado o conceito de criação de um alfabeto, ou de um conjunto cognoscível de poses e movimentos, que será útil para a etapa de interpretação de movimentos capturados por parte do sistema aqui proposto.

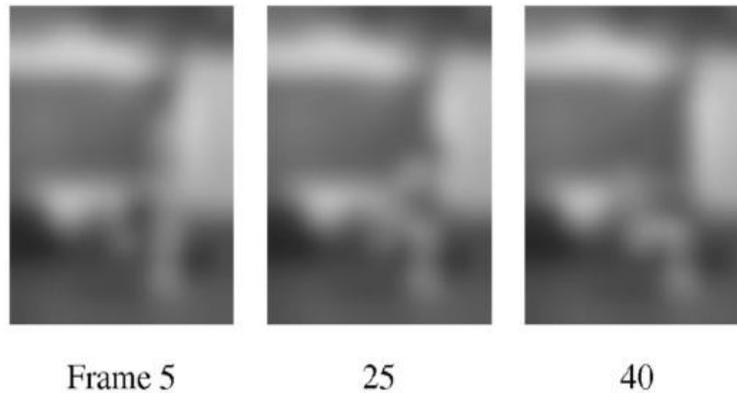
Apesar dessas questões de reconhecimento de poses (e, conseqüentemente, de ações) terem sido apontadas no início dos anos 2000, as primeiras soluções eficientes para elas começaram a aparecer apenas em 2016. Uma delas constituiu-se por meio da criação de um banco de dados de imagens descritivas (utilizando o conceito de *labeling*), criado pela Microsoft Research (HUANG et al., 2016). Entretanto, o reconhecimento de ações humanas nesse banco de dados não foi realizado por meio de aplicações de inteligência artificial, mas sim por indivíduos seguindo orientações dos pesquisadores, o que aumentou significativamente a precisão da identificação de poses por parte do sistema por eles proposto.

O reconhecimento de poses é uma temática que permeia diversos estudos e, como indicamos anteriormente, não foi assunto de apenas um grupo de pesquisa. Dentre as pesquisas que proporcionaram um avanço nessa área está a de Bobick e Davis (2001) que propuseram um sistema de reconhecimento de movimentos humanos, utilizando *templates* temporais. Seu sistema recebia como entrada um vídeo e, em vez de analisar cada um de seus *frames*, eles criaram uma metodologia que identificava a sequência temporal de eventos (Bobick e Davis, 2001) e reconhecia o movimento realizado.

Sua solução era baseada na percepção humana, ou seja, na lógica de reconhecimento de uma imagem por um ser humano. No momento em que publicaram sua pesquisa, já era conhecido no meio da visão computacional que imagens desfocadas representavam um grande desafio para os sistemas de reconhecimento de poses. Embora um algoritmo não fosse capaz de reconhecer um movimento desfocado, um indivíduo era completamente capaz de reconhecê-lo já que, naturalmente, a percepção humana não reconhece o movimento somente por meio da percepção da disposição das juntas de um indivíduo, mas pela memória do deslocamento temporal de um corpo movimentando-se no

espaço. Este é o mesmo princípio que nos permite perceber o movimento de um indivíduo correndo através de uma janela embaçada.

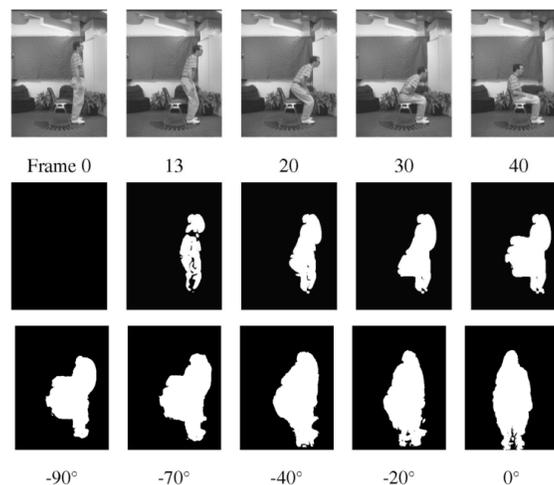
Figura 68 - Sequência de imagens de um indivíduo sentando.



Fonte: Bobick e Davis (2001).

A solução encontrada por eles, para realizar o reconhecimento destes movimentos, foi desmembrar um mesmo vídeo em duas sequências de imagens: uma com o movimento real e outra extremamente contrastada. Denominaram essa criação abstrata de “*templates* temporais” (*temporal templates*) que eram formados por dois conjuntos de imagens, denominadas *Motion Energy Image* (MEI) e *Motion History Image* (MHI). A primeira sendo a sequência de imagens reais e a segunda uma sequência binária de imagens (binária no sentido de que cada *frame* é representado apenas como sendo preto ou branco), como visto na Figura 68 e na Figura 69.

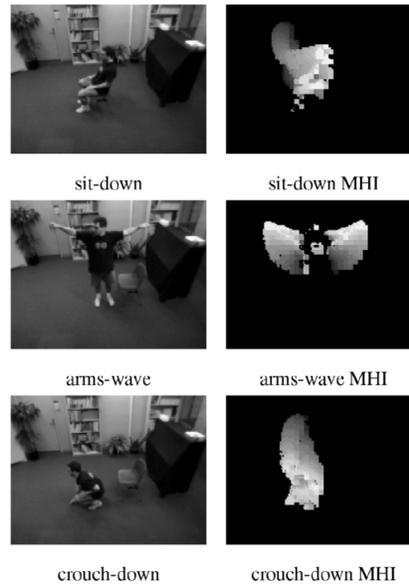
Figura 69 - Representação do *Motion Energy Image* (MEI), Bobick e Davis (2001).



Fonte: Bobick e Davis (2001).

De acordo com o movimento, as formas das MEI e das MHI terão diferentes representações visuais (Figura 70). A comparação dessas formas com bancos de dados pré-configurados resulta no reconhecimento dos movimentos realizados pelo indivíduo.

Figura 70 - *Motion History Image*.



Fonte: Bobick e Davis (2001).

A pesquisa Bobick e Davis alterou o conceito de reconhecimento por meio de algoritmos. Se antes se acreditava no reconhecimento de poses individuais, a partir de sua proposta, o paradigma do sistema foi alterado. Do reconhecimento de poses individuais – cujo conjunto ordenado é considerado um movimento – passou-se para o reconhecimento de *templates* (e modelos) de movimentos cujos padrões são “percebidos” e comparados, de maneira mais “intuitiva” e eficiente pelo computador.

Porém, os sistemas baseados em vídeo somente viriam a apresentar uma melhora em sua confiabilidade, a partir das pesquisas que acrescentaram as câmeras de profundidade, trabalhando em conjunto com as câmeras de vídeo. As câmeras de profundidade têm esta nomenclatura porque procuram identificar a profundidade de objetos e indivíduos no ambiente filmado. Baseiam-se no reconhecimento de padrões pré-determinados e cujo funcionamento será explicado, a seguir. Elas são constituídas por uma fonte luminosa infravermelha (invisível ao olho humano) e uma câmera que reconhece esse espectro luminoso. A fonte luminosa emite uma grade cujo padrão é previamente

conhecido pelo sistema e a câmera captura a cena, com todas as deformações causadas pela projeção do padrão sobre os objetos existentes. De acordo com a distorção da grade, o algoritmo existente nesses sistemas reconhece cada ponto que a constitui como uma distância calculada a partir da fonte emissora (Figura 71).

Figura 71 - Nuvem de pontos do Kinect, sistema baseado em vídeo.



Fonte: Imagem gerada pelo autor.

Alguns dispositivos usados em *games* surgiram dessas pesquisas, como o Wii Remote™ com seu Nunchuck (Nintendo™, Figura 72 [a]) e o Kinect (Microsoft™, Figura 72 [b]), conhecidos por permitirem a interação entre os usuários e seus consoles, através do reconhecimento de seus gestos. Estes equipamentos, inicialmente projetados para o entretenimento, passaram a ser utilizados pela comunidade acadêmica como recursos baratos para pesquisas com captura de movimentos.

Figura 72 - (a) Nintendo Wii Remote com Nunchuck; e (b) Microsoft Kinect.



Fontes: http://www.theregister.co.uk/2006/12/11/nintendo_sued_over_wii_remote/ e <http://www.xbox.com>.
Acesso em: 17 Dez. 2016.

Neles, o desenvolvimento de *hardwares* e *softwares* foi bastante acentuado, tanto por meio do aprimoramento dos primeiros utilizado para a captura de movimentos quanto das aplicações utilizadas para o reconhecimento de gestos.

Outras pesquisas permitiram a precisão das capturas. Um deles foi apresentado na Siggraph,¹⁵ em 2007, e utilizava marcadores ativos e câmeras de alta velocidade para capturar mais detalhes do movimento capturado (Figura 73). Este sistema baseava-se na utilização de *leds* cujas luzes (infravermelhas) são invisíveis para as câmeras RGB e, através da captura da posição destes, por um sistema de câmeras e de luz estruturada (RASKAR et al., 2007).

A vantagem da utilização desses marcadores ativos era o fato deles continuarem sendo acompanhados (*tracking*), mesmo após passarem por uma oclusão de elementos presentes no ambiente ou pelas próprias partes do corpo. Esta característica permitiu a redução da quantidade de câmeras utilizadas nos testes para apenas uma. Dentre as aplicações sugeridas por seus autores para este sistema, estava a área de efeitos especiais no

¹⁵ A Siggraph é a maior feira de computação gráfica mundial. É organizada anualmente pela ACM (Association for Computing Machinery), que possui um Grupo Especial de Interesse (SIG – Special Interest Group) em computação gráfica e técnicas interativas. Ela acontece desde 1967 e funciona como um local de troca de informações sobre as descobertas mais recentes e inovadoras no setor. Nesta feira, é organizado também o festival da Siggraph, que apresenta filmes com o estado da arte em tecnologia de computação gráfica. Mais informações podem ser encontradas no site oficial da organização: <https://www.siggraph.org/>.

cinema. Entretanto, nada nos impede de dizer que todo o setor do audiovisual (incluindo a área da animação) também poderia se beneficiar dessa pesquisa, devido à precisão alcançada na captura através destes marcadores.

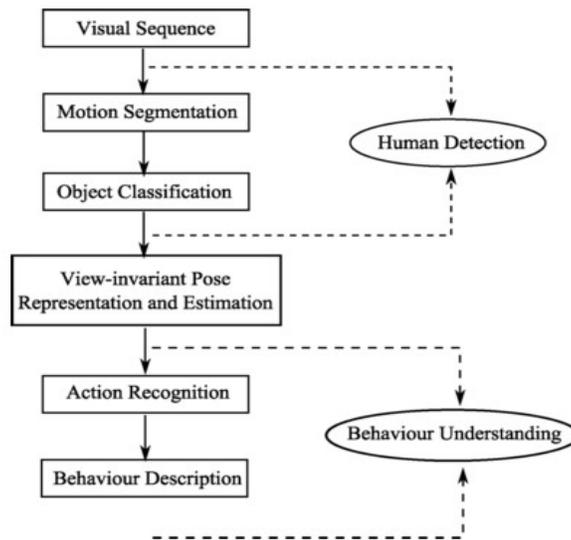
Figura 73 – Captura de movimentos com sensores ativos e luz estruturada (RASKAR et al., 2007).



Fonte: Raskar et al. (2007)

Em 2010, uma revisão das principais descobertas realizadas, até então, na área da visão computacional foi realizada por Ji e Liu (2010). Nessa revisão, eles se limitaram aos autores que focaram suas pesquisas em capturas realizadas através de vistas invariantes, que são aquelas cujo reconhecimento de poses e ações é feito, a partir de quaisquer dispositivos para os quais a posição do indivíduo observado não é relevante para o sistema (ou seja, a câmera que o visualiza pode estar muito acima ou abaixo do seu horizonte de visão). A partir de seu levantamento, eles apresentam uma estrutura (*framework*) do processo de reconhecimento de ações através do esquema a seguir (Figura 74):

Figura 74 - *Framework* do reconhecimento de ações e comportamentos.

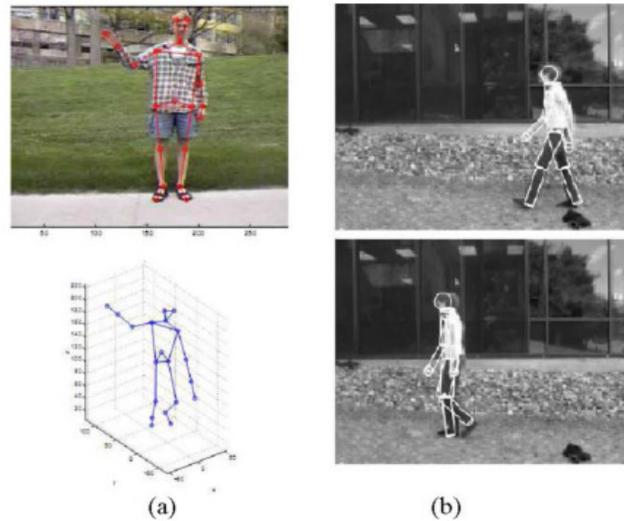


Fonte: Ji e Liu (2010).

Segundo este esquema, o processo de reconhecimento de uma ação passa pela detecção de um ser humano em uma cena, com a segmentação do movimento e a classificação de suas poses. Em seguida, há o reconhecimento do comportamento humano através do reconhecimento de ações e da descrição do comportamento, ou seja, há a segmentação do processo em duas etapas. A etapa de descrição (*labeling*) é atualmente um dos caminhos adotados pelos sistemas de inteligência artificial para reconhecer objetos, pessoas e ações, de maneira semântica. Apesar desse estudo não ser pioneiro neste sentido, é importante apresentá-lo como um dos caminhos que já se delineavam no período.

Já Sidenbladh et al. (2000) mostram que os principais avanços realizados no reconhecimento de ações humanas se deram através do reconhecimento das partes do corpo humano e de suas relações. Abaixo (Figura 75), vemos a imagem de dois tipos de modelos humanos utilizados para o reconhecimento de poses (SIDENBLADH; BLACK; FLEET, 2000).

Figura 75 - Modelos usados para o reconhecimento de gestos humanos.

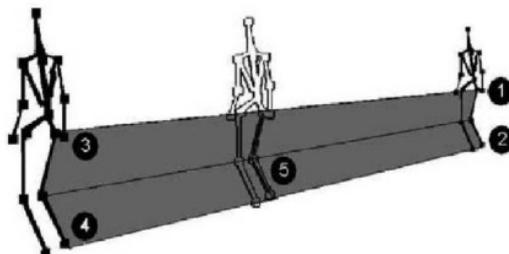


Fonte: Sidenbladh, Black e Flett (2000).

Nestes modelos, observam-se duas formas distintas de abstração de estruturas humanas utilizadas para o reconhecimento de partes do corpo humano. No primeiro deles (Figura 75a), o corpo humano é representado como uma estrutura semelhante àquela apresentada por Lee e Chen (1985), o que facilita a identificação das posições de cada junta do corpo e a plotagem em um sistema de representação tridimensional. No segundo (Figura 75b), cada parte do corpo é representada como um conjunto de formas geométricas que permite ao sistema inferir com maior precisão quando um membro fica ocluído por outro (como, por exemplo, durante a caminhada).

Em seu trabalho, eles apresentam também alguns dos processos mais eficientes descobertos para o reconhecimento de ações. Um deles é o reconhecimento de poses-chave (as poses principais) e a determinação do que seria uma pose intermediária (JI; LIU, 2010). Na imagem abaixo (Figura 76), o reconhecimento de um movimento não é feito a partir de todos os membros do corpo, mas apenas dos da parte inferior do corpo. Os números 1 e 2, bem como o 3 e o 4, mostram as poses principais. A pose intermediária (número 5) é inferida pelo sistema.

Figura 76 - Inferência de poses a partir de um plano traçado entre elas.



Fonte: Ji & Liu (2010).

Outra forma de reconhecimento de ações remete aos MHI e MEI propostos por Bobick e Davis (2001), ao reconhecer padrões tridimensionais e representá-los nas dimensões X, Y e T (tempo), sem preocupar-se com a terceira dimensão (Z) da imagem. Esses padrões, ao se constituírem em *templates*, podem ser utilizados para comparações com movimentos capturados (BLANK, 2005) (Figura 77).

Figura 77 - Reconhecimento de ações através de "Formas de espaço-tempo".



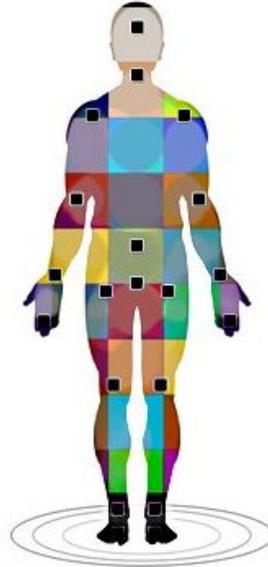
Fonte: Blank (2005).

Em 2009, a Microsoft apresentava ao público o Kinect para o Xbox 360, um dispositivo formado por duas câmeras, uma RGB e uma RGB-D (D de *depth*), dois microfones laterais e um emissor de infravermelho. Seu funcionamento se dá através da emissão de uma grade de pontos infravermelhos e do reconhecimento desta grade pela câmera RGB-D. De acordo com a dimensão destes pontos, o Kinect consegue, através de um algoritmo, reconhecer os pontos que estão mais próximos ou mais afastados do sensor.

Por suas características, é possível perceber que a Microsoft aplicou ao seu Kinect a maior parte dos resultados de pesquisas em tecnologias para capturas de movimentos através de vídeos que estavam em desenvolvimento até aquele momento. Para o reconhecimento do corpo humano, de suas partes e a conversão destas em um esqueleto, é utilizado um algoritmo de divisão de partes do corpo, similar aos apresentados nas

pesquisas anteriores, que segmentam e o corpo humano e as nomeiam (*labeling*), de forma a perceber as posições das juntas para o *trackeamento* (Figura 78).

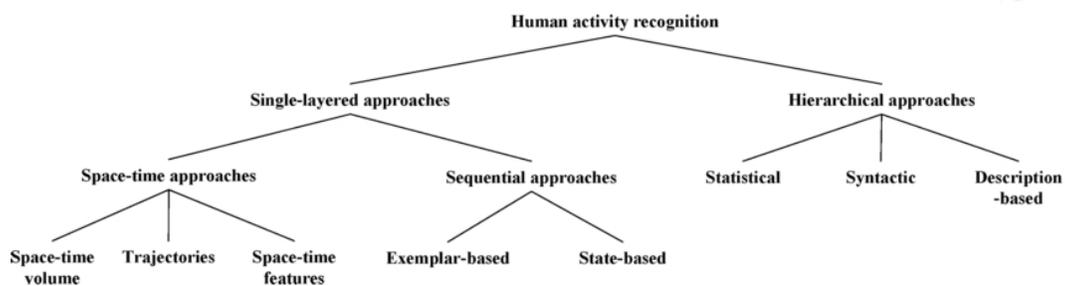
Figura 78 - Segmentação do corpo humano, e reconhecimento das partes para nomear as diversas juntas.



Fonte: Jana (2012).

Em 2011, Aggarwal (2011) realizou uma nova revisão da literatura para a ACM (*Association for Computer Machinery*) acerca das pesquisas realizadas, até então, sobre o reconhecimento de atividades humanas (outro termo utilizado para o reconhecimento de gestos) na área de visão computacional. Em sua revisão, ele apresentou uma proposta de classificação das abordagens realizadas pelos diversos artigos pesquisados e as lacunas no conhecimento que ainda estavam latentes. Esta classificação pode ser vista na Figura 79.

Figura 79 - Classificação das pesquisas na área de reconhecimento de atividades humanas.



Fonte: Aggarwal (2011).

Essa classificação dividiu as técnicas em dois troncos principais. O primeiro, o *Single-layered approaches*, que reconhece atividades humanas através da análise de vídeos, bastante utilizada em sistemas de segurança. O segundo, o *Hierarchical approaches*, que aborda o reconhecimento de atividades humanas através da análise de atividades simples (chamadas de *subeventos*) que, combinadas, podem gerar interpretações de alto nível pelo sistema, usando diversas técnicas para o reconhecimento de atividades como a análise estatística, a construção de modelos e algoritmos baseados em sintaxe gramatical. Uma atividade como carregar um objeto, por exemplo, poderia ser reconhecida como a combinação de algumas interações como segurar um objeto e caminhar.

Embora apresentasse uma nomenclatura nova e categorizasse o conhecimento nesta área, Aggarwal estava, na verdade, apresentando uma revisão das diversas pesquisas com foco em análises em vídeo descritas anteriormente e classificando-as, segundo um critério por ele estabelecido.

A importância de seu trabalho residiu na descrição dos diversos métodos e algoritmos usados para o reconhecimento de poses em vídeo, ou seja, a partir de um ponto vista bidimensional. Do nosso ponto de vista, acreditamos que tais abordagens contribuem para a reflexão sobre os métodos mais eficazes para o reconhecimento de movimentos, utilizando sistemas de captura de movimentos eletromagnéticos que são usados nesta pesquisa.

Entre esses métodos, encontram-se o uso dos “Modelos ocultos de Markov”,¹⁶ abordagens baseadas em descrições – que são o reconhecimento do movimento realizado pelo indivíduo, sua fragmentação em diversos gestos e a interpretação (descrição) das ações realizadas a partir destes; abordagens de espaço-tempo – em que são analisados os movimentos dos indivíduos realizados em X, Y e T (tempo), e definidos algoritmos que comparam os vídeos apresentados com exemplos ou modelos armazenados no sistema.

¹⁶ “Modelos ocultos de Markov” são uma representação probabilística de um sistema. Eles funcionam da seguinte maneira: considerando um sistema A, que está em um determinado estado, este pode ser mudado para os estados B, C e D (A, B, C e D podem ser ruas, passos de um animal, o estado físico da matéria, etc). Os “Modelos ocultos de Markov” são a probabilidade que A tem de se transformar em B, C ou D, de acordo com um “peso” determinado para cada possibilidade. Esta lógica é replicada, então, para cada estado seguinte no qual o sistema se encontra. Devido à sua aleatoriedade, é comumente conhecido como “o passo do bêbado”. Em termos deste estudo, é o que determinará, por exemplo, se uma pose pode levar um indivíduo a dar um passo ou a sentar. No caso da pose, são analisadas todas as juntas segundo os “Modelos ocultos de Markov”. Assim, um sistema pode prever a pose seguinte. Isto é bastante usado em sistemas de segurança, mas pode ser também usado para o reconhecimento de poses para animação.

O uso destes métodos foi relevante para o reconhecimento de atividades humanas em vídeos e Aggarwal (2011) fez uma revisão extensiva dos métodos existentes, chegando à conclusão de que, apesar de todos os avanços feitos nesta área do conhecimento, ainda há muita dificuldade, por parte dos sistemas existentes, em reconhecer as segmentações temporais de atividades e gestos.

Em 2015, uma nova revisão publicada pelos pesquisadores Afsar et al. (2015) analisou as pesquisas de sistemas que detectam o comportamento humano, entre os anos 2000 e 2014. Ao contrário de Aggarwal, que classificou os artigos encontrados segundo as abordagens do *Single-layer* e do *Hierarchical*, de Afsar et al. foi mais simples, concentrando-se em três tópicos: as técnicas de detecção, os bancos de dados e as aplicações. Este tipo de classificação é mais abrangente, pois cada um dos três tópicos pode englobar múltiplas áreas de pesquisa, consistindo em um apanhado geral dos sistemas de capturas de movimentos. Ao compararmos as duas revisões, podemos dizer que a de Aggarwal pode ser inserida dentro das de Afsar et al. na categoria de “aplicações” desses sistemas.

Nos últimos cinco anos, outros algoritmos foram desenvolvidos no campo da visão computacional para o reconhecimento de gestos humanos e poses. Um estudo crítico conduzido por Santofimia et al. (2014a) mostrou que os algoritmos apresentados até ali, como os de Laptev, Marszaek, Schmid, & Rozenfeld (2008), e os de (KUEHNE et al., 2011), ambos baseados em aprendizagem de máquina (onde o algoritmo é programado com uma base de vídeos de teste), eram mais precisos somente quando utilizados para análise de bases de vídeos específicas, ou seja, após o treinamento do sistema, ao serem testados em situações cotidianas, sua precisão era extremamente afetada e, portanto, não suficientemente eficientes para serem aplicados em outras situações como, por exemplo, em sistemas de segurança. Os algoritmos eram fortemente influenciados pela base de treinamento que somente conseguia reconhecer vídeos similares a eles.

Como vimos, as bases de vídeos padronizadas são utilizadas extensivamente em pesquisas de visão computacional porque permitem que seus estudiosos consigam comparar seus sistemas e suas análises entre si, a partir do mesmo referencial, aferindo a eficácia de cada sistema. Entretanto, essas bases são constituídas, em grande parte, por vídeos gravados em situações controladas ou semicontroladas, como o interior de estúdios ou quadras de

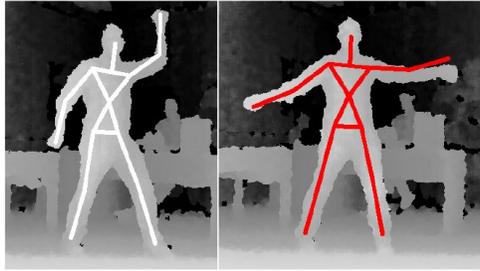
esportes, onde os fundos das imagens quase não se alteram com o tempo. Não há, em muitas destas bases, uma preocupação com situações atípicas (o fundo se movimentando, ou a câmera mudando sua posição, por exemplo), o que faz com que a eficácia desses sistemas seja comprometida, quando aplicadas para interpretar vídeos de outras bases ou vídeos reais capturados ao vivo.

Santofimia et al. (2014a) propuseram um método para a análise de vídeos através de uma abordagem da psicologia da forma. Afirmaram que o reconhecimento de movimentos humanos deveria ser abordado da mesma maneira que fazem os indivíduos, ou seja, baseado em experiências prévias adquiridas de vivências e situações, do conhecimento acerca do ambiente que os rodeia, das intenções do ator/sujeito e das expectativas que indivíduos possuem acerca do movimento, quando este é realizado. Para chegar a esse nível de interpretação, seria necessário usar técnicas de aprendizagem de máquina nas quais o sistema que reconhecerá o movimento é treinado com um conjunto de dados e, em seguida, apresentado a novos dados para classificá-los.

Foi proposta uma abordagem em duas etapas: na primeira, o algoritmo de aprendizado de máquina reconhece o corpo e seus deslocamentos e, a partir destes, apresenta como saída cinco tipos mais prováveis de movimentos que podem ser reconhecidos naquele vídeo, associando cada um a uma probabilidade de ocorrência e permitindo, dessa maneira, uma seleção de alternativas mais prováveis. Na segunda, um sistema de raciocínio e senso-comum qualifica os movimentos propostos no primeiro passo, apresentando uma descrição do movimento realizado e aumentando, assim, a precisão do sistema.

Esse tipo de abordagem tornou-se apropriado para a criação de protótipos de sistemas de reconhecimentos de ações humanas, principalmente aqueles simplificados como o proposto por Borenstein (2012, p. 279–299), onde é criado um algoritmo para ser usado com o Kinect como equipamento de reconhecimentos de poses. Em seu sistema, propõe-se que as posições dos membros de um indivíduo sejam comparadas entre si e que, automaticamente, questões como: “A posição do braço direito é menor que a posição do braço esquerdo?” sejam respondidas pelo sistema (Figura 80).

Figura 80 - Reconhecimento de pose pelo Kinect através da comparação entre as partes do corpo.



Fonte: Borenstein (2012).

Nesta situação, existe um sistema cujas perguntas lhe são apresentadas tais quais as fazemos com linguagem natural como, por exemplo: “O pulso direito está acima da cabeça? O quadril direito está acima do esquerdo?”. O uso de comparações entre as posições dos membros traz uma vantagem: torna mais fácil a descrição daquilo que o sistema deverá “ver” e facilita a programação para indivíduos com pouca prática na área. Ao contrário de sistemas inteligentes, não há necessidade de “treinamento” por meio de bases de dados previamente estabelecidas.

A programação é uma área do conhecimento humano que cada vez mais se insere em diversas áreas do conhecimento e o design é uma das que vem se beneficiando com a abertura deste campo de atuação. O surgimento de linguagens de programação mais intuitivas, como o Processing, e *hardwares* abertos, como o Arduino, permitiram que o design de interação, antes restrito a profissionais altamente qualificados e formados em áreas exatas, pudesse ser explorado, de maneira mais ampla, pelos designers. Mesmo com essas facilidades, não é esperado que o designer tenha conhecimentos matemáticos e de criação de algoritmos suficientemente aprofundados para gerar um sistema capaz de prever o comportamento de indivíduos e seus movimentos.

Nos últimos anos, algumas publicações, como o livro de Borenstein (2012), voltadas para a popularização da visão computacional para artistas e designers, têm mudado essa realidade, porém, a lógica de trabalho, a pesquisa, a educação e a produção do designer não seguem a mesma lógica utilizada para o ensino de programação, que é a realização de tarefas repetitivas e não criativas. Apesar dessa realidade estar aos poucos sendo alterada,¹⁷

¹⁷ O currículo da graduação em Desenho Industrial da Escola Superior de Desenho Industrial da UERJ inclui atualmente o ensino de programação e interação em sua grade. Entretanto, é uma exceção em relação às demais instituições de ensino de design no Brasil.

atualmente, a maior parte dos designers não possui habilidades para resolver problemas de design utilizando programação em conjunto com matemática complexa.

Porém, como apontado anteriormente, uma boa parte das pesquisas desenvolvidas na área da visão computacional, como a de Santofimia (2014a), começou a investigar a forma pela qual indivíduos interpretam imagens estáticas e em movimento, e como esse conhecimento pode ser convertido em funções, para que o computador interprete um movimento.

Assim, poderia a maneira como o designer-animador interpreta uma cena ser compatível com a criação de um método para reconhecer poses e ações humanas? E esta seria útil em sistemas de *motion capture* aplicados à animação? Para o campo do design, quais seriam as abordagens mais apropriadas para um processo de reconhecimento de ações? É o que pretendemos responder, através de uma pesquisa prática, a ser realizada, por meio de um experimento com *motion capture*, no próximo capítulo.

A partir deste breve histórico, percebe-se que as pesquisas na área de captura de movimentos humanos e o reconhecimento de atividades humanas não são áreas exploradas recentemente. Ao publicar uma revisão classificando as principais aplicações e métodos na primeira década do século XXI, Aggarwal et al. (2011) fizeram menção a estudos datados da década de 70. Em sua revisão, os pesquisadores destacaram a utilização de câmeras RGB, que são câmeras de uso comum, encontradas em sistemas de segurança e também aquelas disponíveis ao público em geral.

Percebe-se que uma grande parte do desenvolvimento das tecnologias naquele período destinava-se a diferenciar a figura humana do fundo em que ela se encontrava. Por diversas razões, isso era importante, já que as aplicações resultantes de muitos desses trabalhos eram os sistemas de segurança, cujo reconhecimento das atividades depende principalmente do isolamento do indivíduo e do reconhecimento do tipo de ação que está sendo realizada.

De fato, o reconhecimento de atividades humanas feitas por sistemas de visão computacional têm um percentual muito baixo de efetividade no reconhecimento dos movimentos humanos. A de Santofimia et al. (2014b) indicava que, apesar de toda as pesquisas realizadas até então, os sistemas existentes não eram capazes de reconhecer com eficiência os vídeos gravados no mundo real.

O que Santofimia et al. (2014b) e Linn (2016) propuseram foi o uso de janelas sequenciais de *frames* para o reconhecimento de ações. A partir do reconhecimento da pose, elas usariam os *frames* adjacentes para determinar o movimento. Além de mais preciso, este método torna a procura mais eficiente, já que não é necessário varrer todos os *frames* de uma sequência (Figura 81).

Figura 81 - (a) Reconhecimento de ações feitos pelo sistema de [19]. O reconhecimento é realizado através da análise de *frames* sequenciais; (b) proposta do autor: reconhecimento de poses através de sequências de *frames* com lacunas temporais.



Fonte: Ilustração do autor, baseada nas ideias de Santofimia (2014b) e Linn (LINN, 2016).

Acreditamos que este é um caminho que pode ser seguido em pesquisas futuras, mas procuraremos métodos mais simples para reconhecer ações.

Entendemos que a necessidade de pesquisas na área é enorme e o design pode contribuir com uma abordagem que mescla o tecnológico com o visual, incrementando a eficiência destes sistemas. Este trabalho destina-se, em parte, a contribuir nesta direção.

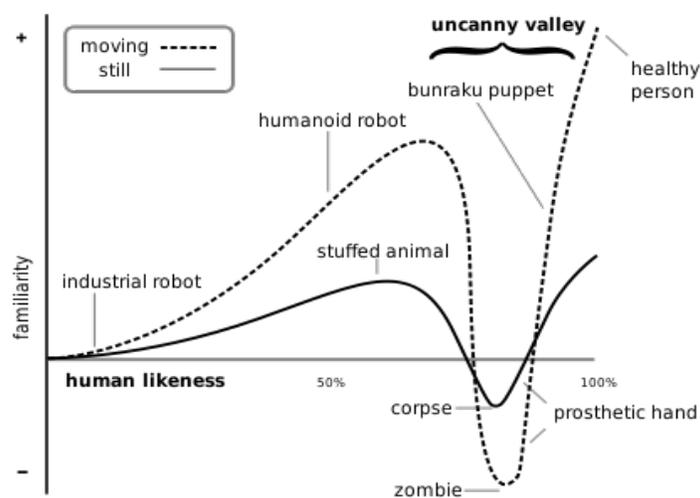
3.6 O desafio para a indústria do audiovisual e do entretenimento: o *uncanny valley*

Na área do entretenimento e, em particular, nas indústrias do cinema, da animação e de *games*, o uso de sistemas de comerciais de capturas de movimentos passou a ser usado de maneira mais constante e com resultados mais satisfatórios, objetivando, ora proporcionar experiências mais realistas para seus usuários, ora reproduzir movimentos humanos realistas em avatares 3D. Como exemplos destas aplicações são os *games FIFA soccer™*, *Battlefield*, *Assassin's creed*, e alguns filmes como *Beowulf* (2007), o “Expresso polar” (2004) e “O curioso caso de Benjamin Button” (2008), cujos produtores utilizaram intensivamente a captura de movimentos para levar às telas a sensação de realismo.

Embora tivessem avançado bastante, os sistemas de MoCap ainda existentes dependiam bastante da pós-produção, ou seja, do tratamento dado pelo animador aos

movimentos capturados. E, quando este tratamento não era bem realizado, os movimentos não ficavam suficientemente realistas. Para toda situação em que havia uma tentativa de criar uma imagem ou um movimento que simulasse o realismo fotográfico e em que o mesmo não era alcançado, causando uma sensação de falsidade para os expectadores, havia um termo que foi bastante utilizado na indústria do entretenimento, o *uncanny valley* (“vale da estranheza”).¹⁸ Para explicar este termo, alguns autores usaram um gráfico para descrever o que ele seria (Figura 82).

Figura 82 - O *uncanny valley*, gráfico que mostra o vale entre a familiaridade e a semelhança com um ser humano.



Fonte: MacDorman (2005, apud Cheetham, Wu, Pauli, e Jancke, 2015, p. 2).

Neste gráfico, o eixo horizontal representa a semelhança do objeto ou ser com o homem e o eixo vertical é a familiaridade ou a percepção de quem observa a imagem, sua empatia com a mesma. A linha cheia do gráfico representa o personagem parado (*still*), uma imagem estática que pode ser uma ilustração. A linha tracejada corresponde às mesmas representações, mas em movimento (*moving*).

Próximas à origem do gráfico, estão representações que utilizam características humanas, mas sem tentar se aproximar do realismo fotográfico, como robôs industriais e robôs humanoides. Neste caso, a semelhança com o ser humano é mínima. Entretanto, à medida que são adicionadas características humanas às representações, como braços mais

¹⁸ O termo *uncanny valley* chegou a ser usado pelo cineasta David Cameron, quando ele se referiu ao desafio de criar o realismo fantástico do filme *Avatar* (2009). Sua entrevista pode ser vista no canal Discovery News, no YouTube, neste link: <https://www.youtube.com/watch?v=Vt-XCDjyDNs>.

cilíndricos, pelos e olhos, em uma busca pela aproximação com a fisicalidade humana e a representação mais próxima à fotográfica, os problemas visuais começam a aparecer.

Como humanos, estamos acostumados a ver outros como nós em fotografias e vídeos, realizando as mais diversas ações. Quando algum modelo, imagem, foto ou avatar possui características próximas a elas, porém incompletas, tendemos a perceber que é algo falso ou inverossímil. Conforme mostra o gráfico, a percepção humana tende a perceber nessa representação algo como um cadáver (*corpse*), caso esteja parado, ou como um zumbi (*zombie*), caso esteja em movimento.

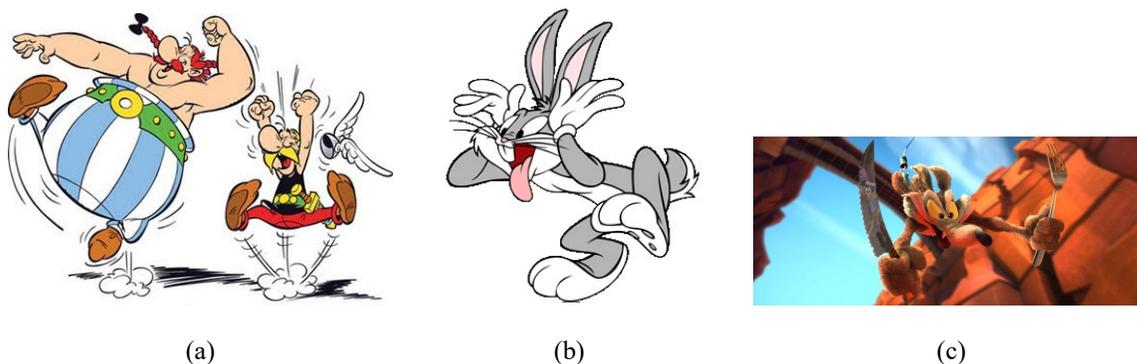
Na primeira década dos anos 2000, um dos desafios da computação gráfica era gerar personagens que se assemelhassem e se movimentassem tais quais indivíduos fotografados ou filmados, ou seja, que ultrapassassem o *uncanny valley* e alcançassem o realismo fotográfico – no gráfico representado pela “pessoa saudável” (*healthy person*).

As produções hollywoodianas que utilizavam a computação gráfica para seus efeitos especiais tinham essa preocupação, embora alguns diretores ousassem usar modelos e movimentos tridimensionais não tão realistas, em troca de serem os pioneiros no uso destas crescentes tecnologias. Somente os maiores estúdios possuíam recursos humanos e financeiros para criar personagens que não caíssem no *uncanny valley*, como a Disney, a Dreamworks e a BlueSky, embora alguns pequenos estúdios, ao realizarem curtas, também arriscassem produzir efeitos de realismo fotográfico. Assim como o cinema, alguns estúdios de animação ousaram aplicar movimentos considerados realistas à animação de personagens, especialmente os cartuns.

Quando um personagem cartunado recebe um movimento produzido, através da captura, por um ser humano, ele aparenta movimentar-se, de maneira estranha e antinatural. Isto pode ser explicado pela expectativa do público ao observar um personagem cartunado. Durante sua vida, um indivíduo está em contato com cartuns através de diversos meios: animações, quadrinhos, embalagens de alimentos e personagens temáticos que algumas marcas usam para promover seus produtos, entre outros. Eles se movimentam de maneira bastante peculiar, usando o que se convencionou chamar, na indústria e na academia, de princípios de animação, que são exageros aplicados aos movimentos e poses de personagens que os tornam mais cômicos (Figura 83).

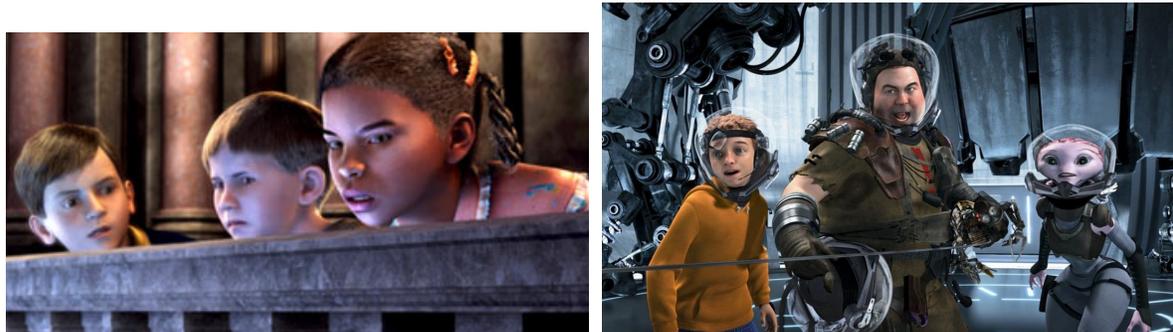
Quando um personagem não apresenta estes princípios, o público tende a estranhar seu movimento, da mesma forma como acontece com a computação gráfica, quando esta tenta se aproximar do realismo fotográfico e não consegue, ou seja, acontece um *uncanny valley* cartum, como pode ser assistido nas animações “O expresso polar” e “Marte precisa de mães” (Figura 84).

Figura 83 - (a) Asterix, o gaulês (Quadrinho); (b) Pernalonga, personagem de animação; (c) Will E. Coyote, personagem de animação 3D. Todos apresentam exageros característicos de desenhos que são aplicados em animação.



Fontes: (a) Asterix, o gaulês (2004); (b) Adesivo do Pernalonga à venda no Ebay; e (c) [completar] <https://www.awn.com/news/road-runner-wile-e-coyote-return-3-d-theatrical-shorts>. Acesso em: 22 Dez. 2017.

Figura 84 - Frames das animações "O expresso polar" e "Marte precisa de mães". Exemplos de *uncanny valley* com cartuns.



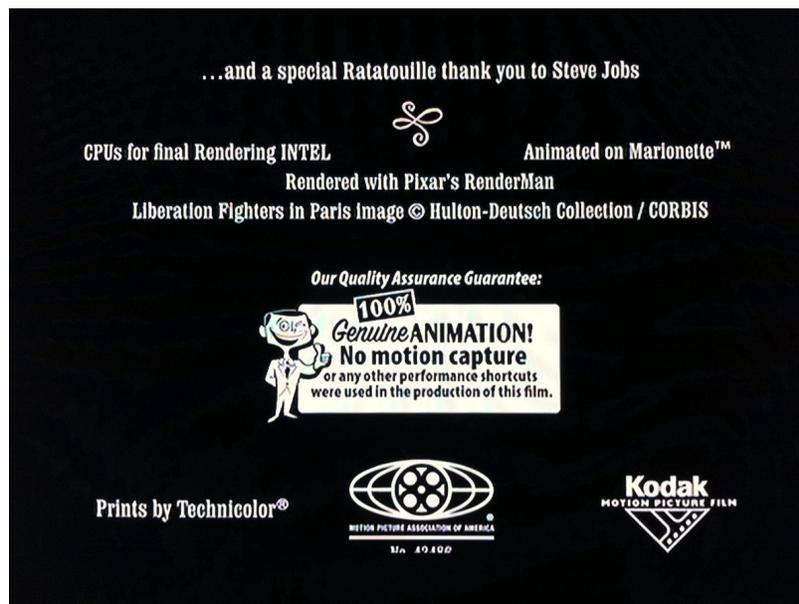
Fontes: DVDs “O expresso polar”, Warner Home Video (2005) e “Marte precisa de mães”, Walt Disney (Sonopress, 2011).

Decerto, todas estas dificuldades não passaram despercebidas e o setor do audiovisual procurou cada vez mais o auxílio da tecnologia em comunhão com a habilidade

dos artistas digitais (animadores, designers, pintores digitais) para superá-las, fosse através de ferramentas digitais, fosse através do trabalho minucioso desses profissionais.

Apesar de todos os avanços descritos anteriormente, alguns estúdios de animação ainda se orgulham de não utilizar a captura de movimentos em seus filmes, como por exemplo, a Pixar (Figura 85).

Figura 85 - Créditos finais do filme *Ratatouille*, apresentando a mensagem "100% animação genuína! Nenhuma *motion capture* ou quaisquer atalhos de *performance* foram usados na produção deste filme".



Fonte: DVD *Ratatouille*, 2007.

Apesar desta visão tradicionalista da animação, que vê o uso da captura de movimentos como um “atalho para a *performance*”, cada vez mais estúdios, influenciados pelas experiências exitosas na área de *games*, têm procurado esta tecnologia para trazer ao público uma experiência mais atraente. Um dos objetivos desta pesquisa é justamente apontar um caminho para o uso da captura de movimentos com a animação tradicional.

3.7 O processo de animação e o processo de reconhecimento de poses: caminhos possíveis

A criação de filmes de animações 3D pelo designer/animador é um processo complexo, realizado em diversas etapas, durante as quais participam diversos profissionais

especializados em cada uma delas. Em países onde a indústria da animação é bem desenvolvida (e o audiovisual, em geral), é comum encontrarmos profissionais responsáveis apenas pelo roteiro da história, outros pelo desenho dos personagens, alguns somente com a preocupação pela iluminação dos cenários e outros ainda pela modelagem dos personagens. A animação é apenas uma das etapas do processo de criação do filme animado. A seguir, descrevemos uma sequência de atividades geralmente presentes na produção de uma animação 3D, para que possamos compreender onde o reconhecimento de poses pode ser explorado em nossa pesquisa.

Em geral, após a criação do roteiro e dos personagens da história, o animador recebe um *storyboard* (Figura 86) de uma cena, com as instruções sobre os enquadramentos, as poses dos personagens que serão animados, os diálogos e os movimentos de câmera. Em alguns estúdios, lhe será disponibilizado o *layout* – uma cena em 3D com o cenário, a posição da câmera (já animada, quando necessário), e o personagem posicionado segundo o enquadramento. Um *animatic* (um *storyboard* animado usado como referência) é também disponibilizado para que o animador consiga ter uma ideia de como o diretor espera que a cena seja animada (Figura 87). Por fim, os diálogos transcritos e os áudios-base são entregues.

Figura 86 - *Storyboards* do filme *Ratatouille*.



Fonte: Paik (2007, p. 73).

Figura 87 - *Animatic* do filme "Como treinar seu dragão".



Fonte: Miller-Zarneke (2010, p. 140).

A partir deste conjunto de informações, o animador procurará referências para animar seu personagem. Essas referências incluem vídeos profissionais ou amadores, outras animações e suas próprias gravações interpretando o personagem (Figura 88).

Figura 88 - Animador Carlos Baena interpretando um personagem para usar como referência no filme *Ratatouille*.



Fonte: https://www.youtube.com/watch?v=4x_1aTERMuo. Acesso em: 22 Dez. 2018.

Neste processo, é importante salientar que, enquanto a cena não está animada, o profissional conta com apenas o *storyboard* e o *animatic* como suas principais referências visuais. O trabalho do animador torna-se descobrir quais movimentos são necessários para melhor transmitir as emoções do personagem. E ele faz isso tendo apenas como base as imagens geradas por estas duas referências.

Quando há o trabalho com *motion capture* em conjunto com a animação, o animador recebe as cenas animadas e precisa trabalhar para aprimorar esses movimentos. Caso seja uma cena cujo objetivo é substituir um personagem real, ele deverá refinar os movimentos vindos da *performance capture*¹⁹ e corrigir as imperfeições resultantes do sistema de captura de movimentos (Figura 89 e Figura 90).

Figura 89 - Atores realizando *performance capture* para o filme “O planeta dos macacos – a guerra”.



Fonte: <https://www.youtube.com/watch?v=a8sdVLnj99U>. Acessado em 22 de dezembro de 2018.

Caso seja uma cena capturada para um filme cartum animado, ele deverá trabalhar para alterar o movimento, de acordo com os princípios de animação. Neste caso, seu trabalho é bem mais intenso, pois precisa seguir o *storyboard* e o *animatic*, além de modificar praticamente todos os *frames* da animação recebida. O trabalho do animador começa quando a captura de movimentos termina. Neste sentido, qualquer ferramenta que aprimore a produção de animação com esta tecnologia será bem recebida pelo setor.

¹⁹ No Cinema, é comum os diretores usarem o termo *performance capture* para se referirem à captura de movimentos utilizada para substituir atores em uma cena. Eles consideram que são capturados os movimentos e a *performance* do ator que está interpretando o personagem.

Figura 90 - Captura de movimentos do ator Andy Serkis realizada para o filme "As aventuras de Tintin".



Fonte: Revista NeoPixel (<https://www.youtube.com/watch?v=5OGWPtaUOok>). Acesso em: 22 Dez. 2018.

Adiante apontaremos as lacunas encontradas e as possibilidades de pesquisa para esta tese.

3.8 Lacunas existentes e possibilidades de pesquisas

A partir desta revisão, foram encontradas as seguintes lacunas existentes no conhecimento acerca da captura de movimentos e do reconhecimento de ações humanas:

1. ausência de um alfabeto visual constituído por entidades-movimento (MOESLUND; GRANUM, 2001);
2. ausência de métodos simples – aqui referimo-nos à simplicidade de elaboração de algoritmos, em especial para áreas como o design, cuja formação não engloba programação na maior parte das instituições de ensino – e eficazes de reconhecimento de gestos (como apresentado anteriormente, as soluções existentes possuem taxas de erro elevadas, dependendo das bases de dados utilizadas); e
3. inexistência de sistemas que fazem a cartunização de quaisquer movimentos, não se limitando a alguns específicos como as corridas e saltos (KWON; LEE, 2011) ou à mecanização de alguns princípios de animação (WANG et al., 2006) e que, por vezes, não permitem a expansão dos tipos de movimentos reconhecidos ou a aplicação de estilos específicos de animação ao personagem cartum.

As seguintes questões serão abordadas adiante para tentar solucionar as lacunas encontradas:

1. reconhecimento de ações através da criação de uma base de conhecimento que poderá ser continuamente alimentada pelos usuários do sistema;
2. identificação de parâmetros para o reconhecimento de poses;
3. criação de uma base de dados que poderá se constituir em uma forma de alfabeto visual, definido por poses com possibilidades de expansão de forma rápida, eficiente e prática;
4. criação de uma base de dados de movimentos cartuns, construída a partir da relação entre a proporção entre o indivíduo capturado e o personagem cartunizado;
5. realização da cartunização de um personagem a partir do reconhecimento de poses, gestos e ações, através da relação estabelecida entre movimentos reais e movimentos cartuns gravados anteriormente; e
6. consideração desta proposta se tornar um método para cartunização de captura de movimentos para o setor da animação.

Estes são os desafios que pretendemos abordar a partir do próximo capítulo.

4 EXPERIMENTAÇÃO: DO REALISMO AO CARTUM

Para responder às questões levantadas no capítulo anterior e contribuir para o avanço das pesquisas na área de design de animação, propomos um método para cartunizar personagens que será comprovado através de experimento prático, relacionando captura de movimentos à movimentação cartum.

Nosso método é composto por quatro etapas: a primeira consiste na captura de movimentos, utilizando um sistema apropriado; a segunda é o reconhecimento dos movimentos capturados; a terceira é a criação de uma série de movimentos cartunizados; e a quarta, a realização da cartunização em um *software* de animação 3D, através de um *script*. Destas etapas, a única não serializada é a terceira, visto que não é dependente de nenhuma das demais.

Figura 91 - Etapas do processo de cartunização.



Fonte: Esquema criado pelo autor.

Para nosso experimento foram escolhidas duas ações: caminhar e correr, em razão de sua similaridade. O algoritmo projetado reconhece as ações realizadas e as classifica, utilizando a lógica Fuzzy. Em seguida, ele entrega, em sua saída, um arquivo com os dados analisados, para ser utilizado pelo *software* de animação 3D (optamos pelo 3ds Max). Através deste mesmo *software*, são criadas diversas animações cartunizadas que, salvas em uma pasta, serão acessadas para realizar a substituição do movimento realista pelo movimento cartunizado. Finalmente, um *script* importa a captura de movimentos e utiliza os dados advindos do algoritmo, para cartunizar os movimentos realistas. A seguir, apresentaremos detalhadamente cada etapa deste processo.

4.1 O processo de captura de movimentos e transferência para o personagem

O processo de animação por captura de movimentos é realizado através de uma série de procedimentos técnicos que buscam transformar a imagem, gravada ou reconhecida por câmeras, em dados numéricos que podem ser armazenados, analisados e manipulados, segundo a conveniência do usuário.

De maneira geral, a maior parte dos sistemas de captura de movimentos realiza as seguintes etapas, durante o processo de conversão de movimentos em dados: reconhecimento do corpo humano; subtração do corpo reconhecido do fundo onde se encontra; *tracking* (acompanhamento) das partes do corpo; reconhecimento e segmentação destas partes (como a cabeça, o tronco e os membros); e conversão destes em dados. Esses dados são posições e rotações das juntas do indivíduo que realizou os movimentos. Eles são sintetizados e transferidos para um personagem de maneira que o usuário possa visualizar o movimento detalhadamente ou simplesmente plotar as informações em um gráfico para análise. Esses dados são brutos, ou seja, não possuem nenhum tipo de correção que corrija falhas durante o processo de captura. Dependendo da aplicação, eles podem passar por um processo de refinamento no qual os ruídos são filtrados.

Para esta pesquisa, a etapa de reconhecimento e subtração em relação ao fundo – normalmente encontrada em sistemas que usam câmeras de vídeo para a captura de movimentos – é “transparente”, isto é, invisível ou irrelevante, visto que, em nossos testes, utilizamos um equipamento de captura de movimentos que não utiliza câmeras, mas sensores eletromagnéticos colocados nas juntas do corpo humano para capturar os dados.

Como protótipo para testes nesta pesquisa foram usados personagens 3D, pois eles são comumente preparados com um esqueleto projetado para receber os dados de capturas de movimentos e, a partir daí, realizar o deslocamento (animação) do personagem. A Figura 92 apresenta a sequência de transferência de um movimento humano para um personagem:

Figura 92 - Fluxo de produção de uma captura de movimentos.



Fonte: Elaborado pelo autor.

Distinguem-se duas etapas neste processo: o primeiro é a criação do personagem no qual os movimentos serão aplicados. O segundo é a captura de movimentos realizada por um equipamento e um *software* especializados em realizar esta tarefa. Tais processos, embora pareçam, não são independentes. É necessário conhecer previamente o sistema de captura de movimentos que será utilizado, durante a produção de animação, para projetar um esqueleto adequado a ele. Cada equipamento trabalha com formatos de arquivos diferentes e cabe ao profissional que trabalhará com o sistema entender a constituição deste arquivo, bem como as vantagens e as limitações para a transferência da captura para o personagem. Alguns formatos facilitam a conexão entre o sistema de captura e o personagem.

4.1.1 Criação do personagem 3D

A preparação de personagens 3D para qualquer finalidade se constitui como um processo bem definido no mercado de produção de filmes e séries de animação. A estruturação desse processo se inicia na década de 1980, quando os primeiros curtas-metragens começaram a ser criados para a animação de testes de tecnologias aplicadas à computação gráfica. Muitos foram exibidos tanto em programas de televisão quanto em festivais, como a SIGGRAPH (Figura 93).

Figura 93 - *Frame de Tin Toy*, curta animado da Pixar.

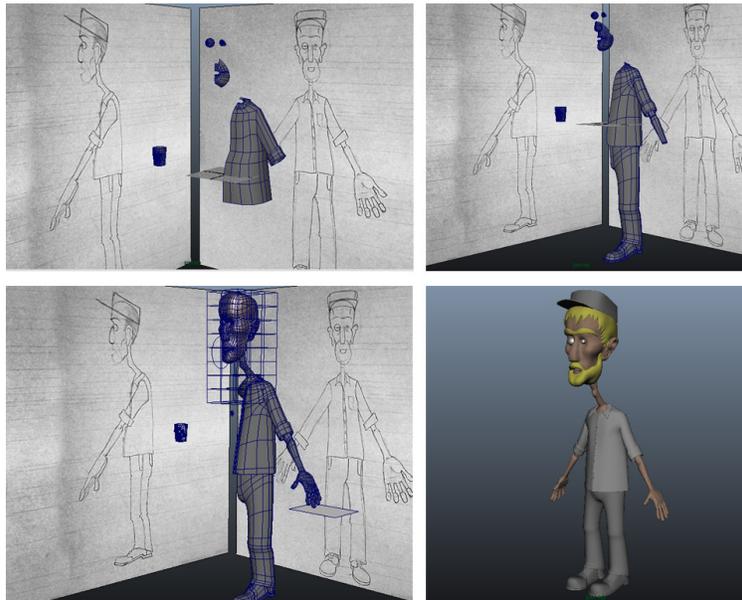


Fonte: (PIXAR, 1988).

O fluxo de criação de um personagem, nos dias atuais, é descrito, a seguir, e sua compreensão é útil para que se entenda o ponto exato no qual a captura de movimentos é aplicada. Esse fluxo varia de tempos em tempos, quando novas tecnologias são adicionadas às etapas de produção de personagens. Um exemplo disso é a escultura digital, técnica recentemente inserida na modelagem de personagens. Podemos dizer que as etapas envolvidas na criação de um personagem são:

1 – *Modelagem*: processo de criação da forma digital tridimensional do personagem. Neste processo, parte-se de uma geometria básica (em geral, de esferas ou cubos) e, a partir da movimentação, da deformação, da subtração e da adição de polígonos, chega-se à forma de um personagem. A prática na indústria aponta para a utilização de um *concept* (desenho conceitual criado por um designer de personagens) como base para a modelagem. O modelador – profissional responsável por esta etapa – procura seguir o desenho do *concept* de maneira a concretizar digitalmente a forma criada pelo artista. Ele é responsável por fazer com que o modelo tridimensional corresponda, o máximo possível, ao *concept*, limitado apenas pelas características dos *softwares* 3D (Figura 94);

Figura 94 - Modelagem da malha do personagem.



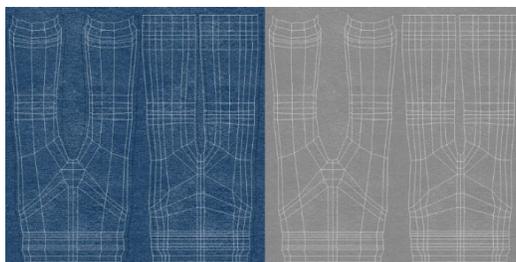
Fonte: Personagem do autor.

2 – *Texturização*: Após a etapa de modelagem, o personagem ganha texturas, imagens que são aplicadas ao modelo tridimensional para simulação de elementos superficiais, como os tecidos, a textura da pele, a cor dos olhos, etc. Durante esta etapa, o artista digital possui duas opções, de acordo com as ferramentas que está utilizando. Ele pode planificar a geometria do personagem, gerando uma imagem digital para pintá-la em um *software* de edição/pintura de imagens ou pode pintar diretamente a geometria do personagem em *softwares* específicos, como o ZBrush. Assim como o modelador, o profissional de texturização também segue um *concept* para guiar seu trabalho (Figura 95); e

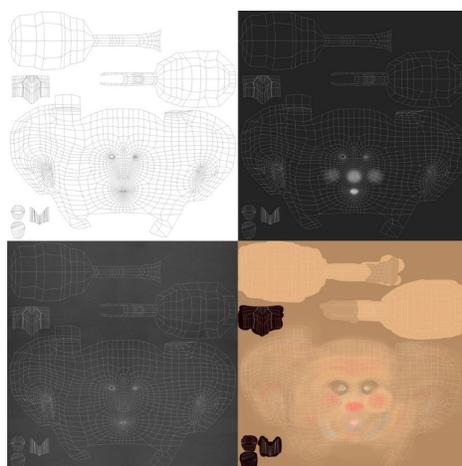
Figura 95 – Texturização do personagem: a) personagem com roupas e pele texturizados; b) planificação da geometria da roupa para texturização em *software* de edição de imagens; e c) planificação da geometria facial do personagem para texturização.



(a)



(b)

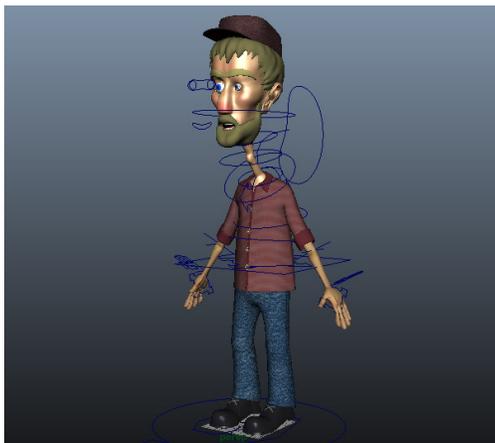


(c)

Fonte: Personagem do autor.

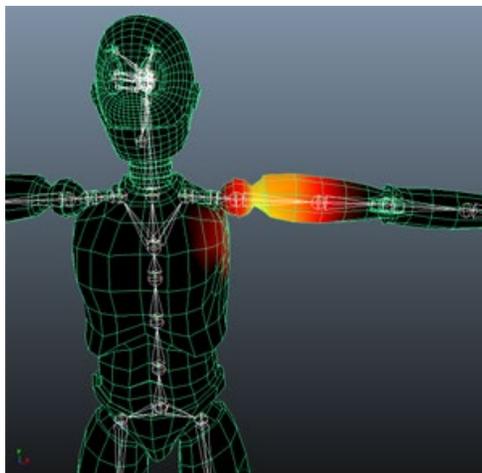
3 – *Criação de esqueleto e controles para animação*: Após a etapa de texturização, é necessário criar uma estrutura para movimentar a geometria tridimensional do personagem. Este processo é conhecido como *rigging* e se realiza em duas etapas: a primeira consiste na criação de um esqueleto que moverá a geometria; e a segunda é a criação de controles que movimentarão o esqueleto. A criação destas estruturas permite que o animador movimente o personagem tal qual uma marionete, porém digital (Figura 96 e Figura 97).

Figura 96 – Criação do esqueleto e de controles do personagem.



Fonte: Personagem do autor, com controles expostos.

Figura 97 – Processo de *skin* do personagem com aplicação de zonas de influência do esqueleto (geometrias brancas dentro do personagem) sobre o modelo 3D (área laranja-avermelhada).

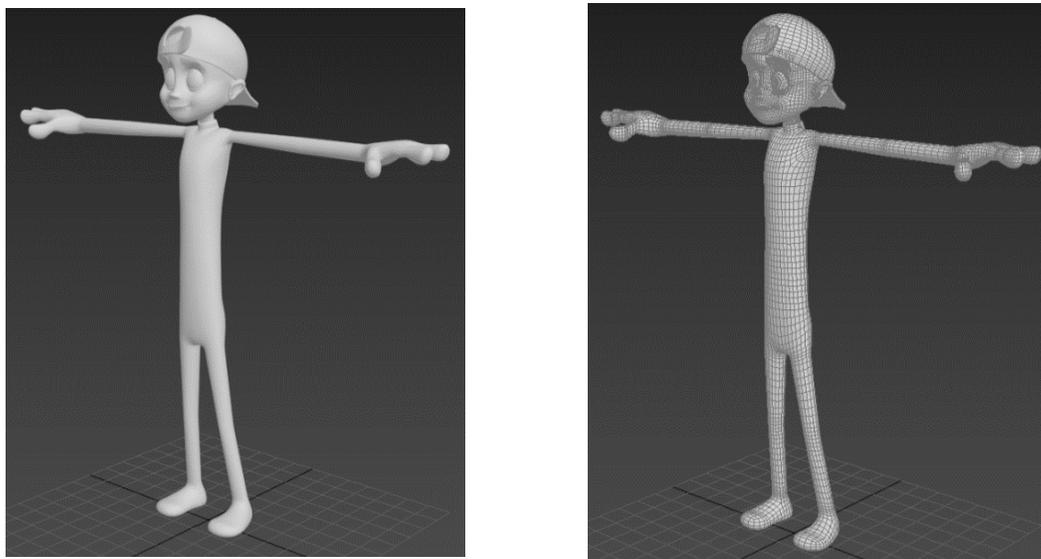


Fonte: Autodesk Maya Help (<http://help.autodesk.com/view/MAYAUL/2017/ENU/?guid=GUID-EEFC3BE1-B386-4664-BD19-8688C58D1618>). Acesso em: 5 Fev. 2018.

Ao final destas etapas, o animador recebe o personagem pronto para movimentação. Nesta pesquisa, foi utilizado um personagem projetado para um curta-metragem de animação

experimental do autor deste trabalho. Portanto, o mesmo já se encontrava previamente modelado, texturizado e com controles (Figura 98) que permitiriam o avanço no que realmente concerne a um dos objetos deste trabalho: a aplicação dos movimentos capturados sobre o personagem.

Figura 98 - Modelo 3D do personagem utilizado para os testes.



Fonte: Arquivo do autor.

4.1.2 Preparação dos esqueletos para captura de movimentos

A preparação do personagem para a captura de movimentos exige a criação de, pelo menos, dois esqueletos. Um deles recebe os movimentos das capturas e o outro é similar ao esqueleto descrito na seção anterior, sendo controlável pelo animador para que sejam feitas correções e aplicações de movimentos específicos, como os princípios de animação – quando necessário – sobre o personagem.

Apesar do mercado possuir soluções que recebem a captura de movimentos e a apliquem diretamente sobre o esqueleto para edição,¹⁸ optamos nesta pesquisa por criar três esqueletos para validar nosso protótipo: um para receber as capturas de movimentos, outro para receber os movimentos cartuns e um terceiro, controlado pelos dois anteriores, para movimentar a geometria do personagem. Esta configuração apresenta algumas vantagens

¹⁸ Uma das soluções é o *software* Motion Builder, da empresa Autodesk.

sobre as soluções limitadas já existentes no mercado, pois permite que o animador possa decidir o quanto usará dos movimentos realistas capturados ou o quão cartum ele deseja.

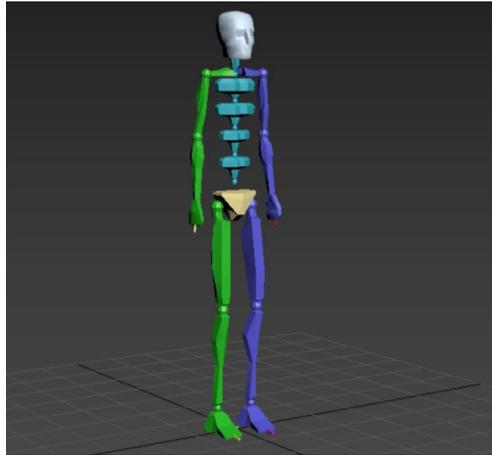
Em nossa pesquisa, concluímos que, para utilizarmos movimentos vindos de fontes de animação/movimentos diferentes, é necessário criarmos um esqueleto para cada uma destas fontes, como descrevemos, a seguir.

Nossos testes indicaram não ser possível utilizar a captura de movimentos diretamente sobre esqueletos construídos com *bones* do 3ds Max, devido às limitações deste *software*: as rotações das capturas de movimentos não eram compatíveis com as rotações do sistema de *bones*. No processo de transferência dos movimentos das capturas para o personagem, utilizamos um esqueleto existente dentro do 3ds Max, o Biped (Figura 99), que agilizou esta etapa. Esta solução, entretanto, é específica do 3ds Max. Outros *softwares* de animação, como o Autodesk Maya, utilizam sistemas internos para a aplicação da captura de movimentos diretamente sobre o esqueleto construído pelo animador, por meio de um procedimento conhecido na indústria como *retargeting*.

Nosso propósito inicial era usar apenas o Biped, porém nos deparamos com limitações que nos forçaram a procurar alternativas:

- 1 – Em nossos testes, não conseguimos elaborar uma maneira de exportar movimentos cartuns, a partir do esqueleto, para a criação de um banco de dados desse tipo de movimento. A criação deste banco era necessária para a aplicação sobre o personagem, durante a etapa de substituição dos movimentos realistas por movimentos cartuns; e
- 2 – O esqueleto Biped não admite ser controlado por outro Biped ou esqueleto personalizado, o que invalidaria a proposta de utilização de três esqueletos, separando as animações cartuns das realistas.

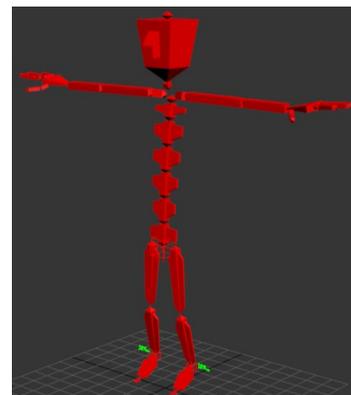
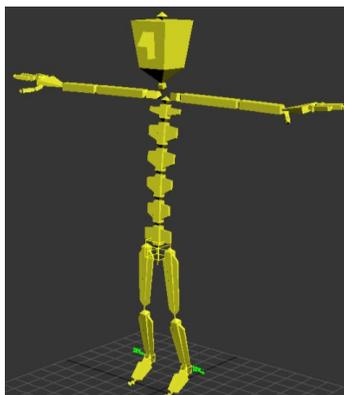
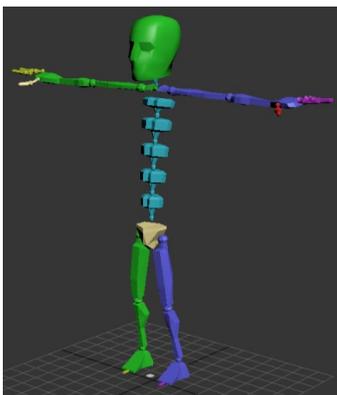
Figura 99 - Biped, esqueleto pronto para animação e captura de movimentos no 3ds Max.



Fonte: Captura de tela do 3ds Max.

Assim, elaboramos o esquema descrito anteriormente com três esqueletos trabalhando em conjunto (Figura 100) e, para que não ocorressem problemas durante sua utilização (como diferenças entre tamanhos e posições da estrutura esquelética), construímos todos, tendo por base o esqueleto Biped, e alinhando as rotações dos *bones* às rotações de suas juntas.

Figura 100 - Esqueletos criados para receber a captura de movimentos e os movimentos cartunizados. À esquerda: Biped; ao centro, o esqueleto para receber animações cartunizadas; e à direita, o esqueleto que move o modelo 3D do personagem.



Fonte: Elaborado pelo autor.

Como as animações cartuns eram realizadas pelo “esqueleto cartum”, foi feita uma cópia do esqueleto para criar as animações cartunizadas, de maneira que pudéssemos usá-lo para gerar o banco de cartuns para uma posterior aplicação na cena final.

4.2 Captura de movimentos

Os movimentos utilizados nesta pesquisa foram capturados por meio do equipamento Neuron Mocap adquirido para nossos testes. Este equipamento é um sistema de capturas completo, composto por um *hardware* e um *software* que, trabalhando em conjunto, reconstituem digitalmente os movimentos humanos em meio eletrônico. Ele é constituído por sensores – acelerômetros e magnetômetros – que reconhecem, interpretam e exportam os movimentos dos membros (juntas) do corpo em formato de ângulos e posições. O sistema realiza o trabalho de armazenar os sinais captados pelos sensores e de traduzi-los em posições e rotações dos membros (Figura 101).

Figura 101 - Neuron Mocap e sua configuração para o experimento.



Fonte: Fotografia do autor.

Cada sessão de captura de movimentos se inicia com o posicionamento dos sensores em pontos específicos do corpo do indivíduo, por meio de um traje. Em seguida, é realizada a calibração do sistema, com o voluntário a ser capturado colocando-se em poses pré-determinadas pela orientação do *software* (Figura 102). Estas etapas não são exclusivas do sistema Neuron Mocap e podem ser observadas nas capturas com câmeras abordadas nos trabalhos de aquisição e de trackeamento do corpo humano, usando voxels, de Mikic et al. (2003), na revisão de literatura sobre processos de captura de movimentos por vídeo, realizada por Borges et al. (2013) e de criação de *pipeline* de animação para a captura de movimentos, de Mattsson e Sundh (2014).

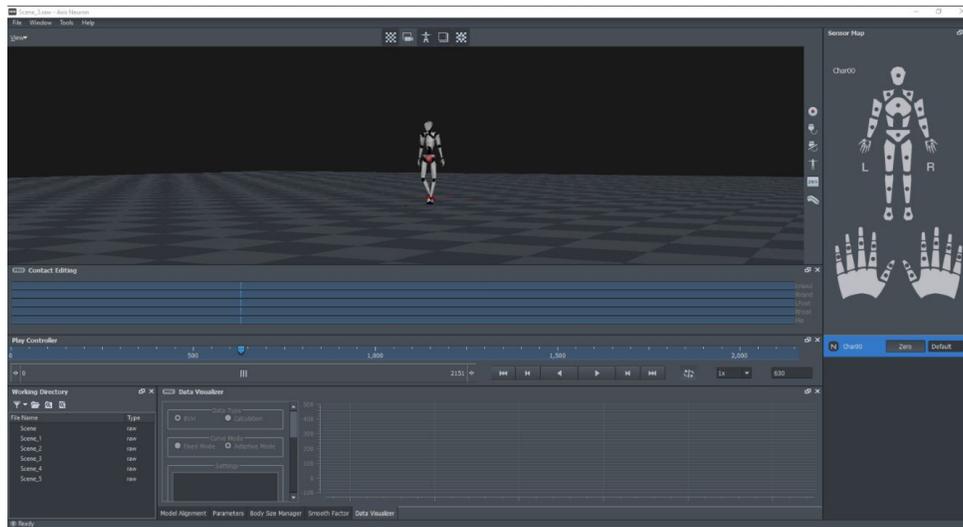
Figura 102 - Fotografia do indivíduo em uma das poses que calibram o sistema.



Fonte: Fotografia do autor.

Após a calibração, o sistema está pronto para o uso. Quando o voluntário se movimenta, os sensores são trackeados e transmitem as alterações percebidas em suas posições e rotações para o *software* que interpreta e transmite os dados. Esta etapa pode ser feita via rede sem fio ou através de um cabo conectado a um computador. Na estação de trabalho, as aquisições são reconhecidas por outro *software* específico e aplicadas em tempo real sobre o personagem que representa os movimentos capturados (Figura 103).

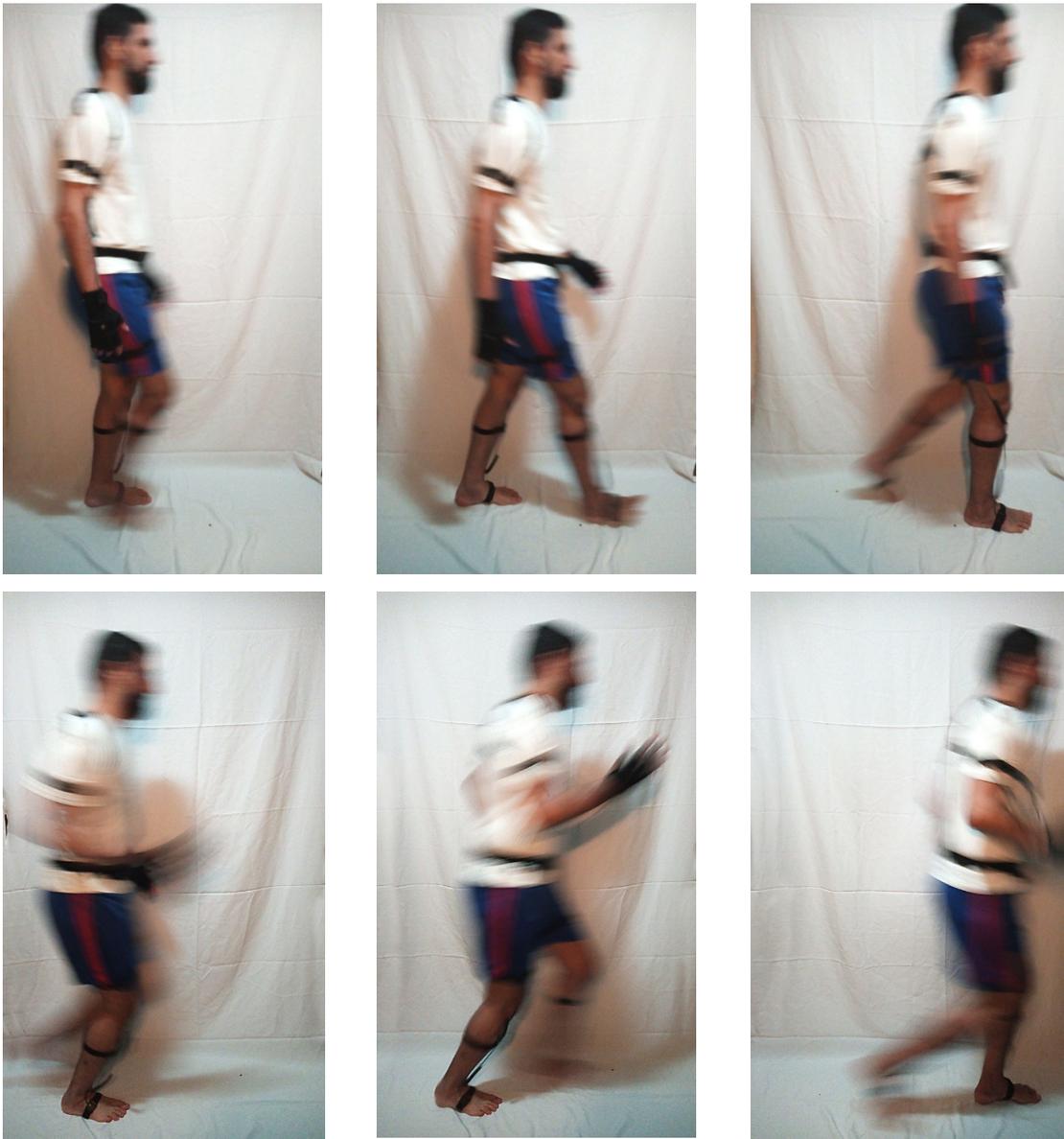
Figura 103 - Imagem da tela do Neuron Mocap com personagem sendo capturado.



Fonte: Captura de tela do *software* Neuron Mocap.

Em nossa pesquisa, escolhemos os movimentos da caminhada e da corrida humanas (Figura 104). Esta escolha se deu em função de dois fatores específicos: objetivávamos criar um sistema de reconhecimentos que fosse preciso o suficiente para reconhecer e diferenciar os movimentos. Por sua similaridade, estes movimentos são ideais para esta análise. O outro fator decisivo para a escolha dos movimentos foi sua presença constante em animações de personagens.

Figura 104 - Indivíduo sendo capturado nas poses de caminhada (sequência de cima) e corrida (sequência de baixo).



Fonte: *Frames* do experimento gravado.

Ao final da sessão de captura, tivemos a opção de gravar os dados em diversos formatos de arquivos de capturas de movimentos. Neste experimento, o BVH foi escolhido para ser utilizado pelas suas características, a seguir, descritas:

1 – *Popularidade*: é um dos formatos mais antigos e conhecidos de captura de movimentos. Embora a empresa que o criou (BioVision) esteja extinta, este

formato continua sendo utilizado no mercado de captura de movimentos e no meio acadêmico;

2 – *Formato aberto*: é um formato que facilita a extração, a classificação, o tratamento e a alteração de dados. O algoritmo que programamos consegue ler seus dados como um simples arquivo de texto, facilitando a importação e análise; e

3 – *Organização*: o arquivo é separado em dois blocos de informações. O primeiro armazena as posições das juntas do corpo humano, de acordo com as dimensões do indivíduo. O segundo armazena as transformações – alterações nas rotações das juntas e na posição do centro do corpo – ocorridas durante a captura de movimentos. Devido a esta organização, a procura e filtragem dos dados é facilitada.

Destacamos algumas características presentes nos arquivos BVH, que facilitam seu uso:

1 – Todos os arquivos contêm a posição e a rotação do quadril (*Hips* ou COG), visto que toda a hierarquia dos esqueletos está conectada a ele, como apresentamos em nossa revisão bibliográfica (Figura 105);

2 – Todos os arquivos contêm o afastamento entre cada membro do indivíduo, o que permite reconstruir o personagem (OFFSET, Figura 105);

3 – A quantidade de *frames* capturados é armazenada, bem como a duração de cada *frame* (*frame time*), em milissegundos (Figura 106); e

4 – Em todos os *frames* são gravados as posições e rotações do quadril. Os demais membros do corpo somente possuem suas rotações gravadas. Esta sequência de valores pode ser vista na Figura 106.

Figura 105 - Exemplo de arquivo BVH, mostrando a hierarquia do corpo capturado.

```
HIERARCHY
ROOT Hips
{
  OFFSET 0.000 93.594 0.000
  CHANNELS 6 Xposition Yposition Zposition Yrotation Xrotation Zrotation
  JOINT RightHip
  {
    OFFSET -9.500 -1.594 0.000
    CHANNELS 3 Yrotation Xrotation Zrotation
    JOINT RightKnee
```

Fonte: Arquivo BVH.

Figura 106 - Exemplo de arquivo BVH mostrando as posições e rotações dos membros capturados.

```
MOTION

Frames: 668
Frame Time: 0.008
-115.75 93.25 6.31 21.91 -2.87 0.74 2.30 -8.17 -4.87 -0.00 26.37 -0.00 -15.30 -16.50 2.73 18.17 -4.02 -4.07 0.00
22.26 0.00 -13.65 -14.06 7.85 -3.18 4.59 -0.97 -0.80 1.14 -0.22 -1.34 1.88 -0.37 0.33 -0.26 -0.02 0.33 -0.27 -0.02 6.07
-4.77 -0.66 -1.89 -1.74 1.00 18.18 11.39 -17.58 -75.18 17.81 -13.71 -17.92 -18.15 -11.16 -21.86 -0.30 0.00 10 -0.00 10
-0.00 0 -0.00 -2.55 -0.20 0.00 0 -0.24 0.00 0 -0.20 0.00 0 -0.00 0 -0.20 0.00 0 -0.00 0 -0.00 3 -0.20 0.00 0
-0.24 0.00 0 -0.20 0.00 0 -0.00 8.83 -0.20 0.00 0 -0.24 0.00 0 -0.20 0.00 9.22 2.04 -0.33 4.94 -1.32 -3.17 -97.67 -20.41 -15.94
6.07 10.99 -5.81 -21.86 0 -30.00 10.00 0 0.00 10.00 0 0.00 0.00 0 0.00 -2.55 0 -20.00 0.00 0 -24.00 0.00 0 -20.00
```

Fonte: Arquivo BVH.

4.3 A criação dos conjuntos de arquivos: capturados e cartuns

Além das etapas anteriormente descritas, para a realização do experimento foram criados dois conjuntos de arquivos (um conjunto de captura de movimentos e outro de animações cartunizadas), um *software* para reconhecer os movimentos gravados no banco de capturas e um *script* para interpretar a saída do *software* de reconhecimento de movimentos e aplicar o movimento cartum equivalente no personagem.

O conjunto de capturas constituiu-se de movimentos de caminhada e corrida, para realizarmos a análise, a identificação e a diferenciação entre estes movimentos. O reconhecimento dos movimentos foi realizado através do uso da lógica Fuzzy, por meio de um conjunto de regras que interpretavam as relações entre as diversas juntas capturadas (a seção 4.5.1 descreve este processo).

O conjunto de movimentos cartuns foi criado no *software* 3ds Max, através do esqueleto criado para esse propósito, como visto na Figura 100. O diagrama da Figura 107 mostra o fluxo de funcionamento do sistema que propomos.

Figura 107 - Fluxo de aplicação de movimentos no personagem.



Fonte: Esquema do autor.

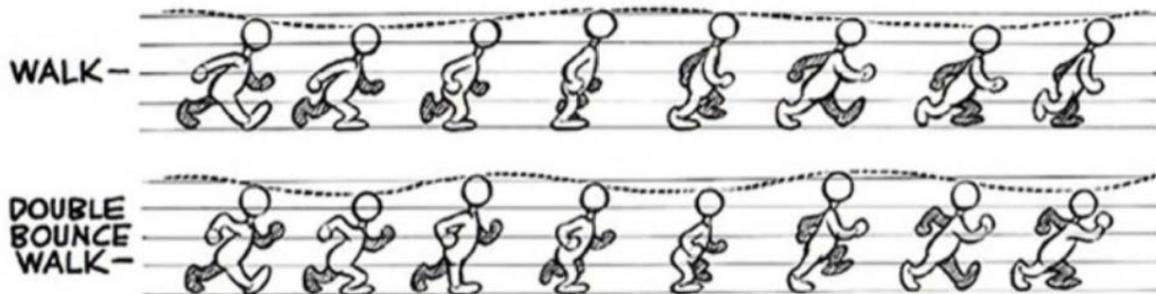
As capturas de movimentos foram aplicadas sobre o esqueleto Biped e exportadas como arquivos BIP. Este procedimento foi necessário pois, ao aplicarmos o arquivo BVH diretamente sobre o Biped, este reconhecia as dimensões do indivíduo capturado e alterava as dimensões do Biped, deformando nosso modelo 3D.

Ao exportarmos os movimentos como arquivos BIP e depois importarmos o mesmo sobre o esqueleto Biped, ele reconhecia apenas as rotações das juntas, não alterando as dimensões do esqueleto e, portanto, não redimensionando o personagem.

Já as animações cartuns foram criadas, utilizando o esqueleto cartum e exportadas para uma pasta como arquivos de animações do 3ds Max, o XAF (*3ds Max Animation File*). O objetivo dessa etapa foi possibilitar a criação de diversos tipos de movimentos que pudessem ser importados via *script* em formato nativo do 3ds Max, além de permitir que qualquer animador pudesse criar seu próprio conjunto de animações.

Os estilos de caminhada e de corrida foram escolhidos entre aqueles apresentados no capítulo 2 e optamos por uma caminhada estilo *double-bounce* e um cartum tradicional (Figura 108) através de referências que foram retiradas de Blair (1980). Estas foram selecionadas porque são bem exageradas e podem ser facilmente discernidas dos movimentos capturados, ao serem aplicadas sobre o personagem. Em seguida, começamos a programação do sistema de reconhecimento de movimentos.

Figura 108 - Caminhada normal e *double bounce*.



Fonte: Blair (1980, p. 69)

4.4 Lógica Fuzzy: motivações para aplicação na pesquisa

Classificar conjuntos é uma das aplicações mais frequentes dos campos da inteligência artificial, do aprendizado de máquina e, por consequência, da visão computacional, que utiliza o conhecimento adquirido nestas áreas para aplicá-lo no reconhecimento de gestos e ações humanas.

Como seres humanos, aprender a classificar a natureza nos fez compreendê-la para melhor dominá-la. No campo do design – que trata dos artefatos criados pelo homem – classificar é uma forma de compreender as maneiras através das quais aproveitamos as pesquisas geradas por outras áreas do conhecimento para aplicá-las nas nossas próprias, transformando a maneira de nos relacionarmos com os artefatos e propormos novas soluções para os problemas existentes.

Apesar da importância de realizarmos classificações através da observação, da análise e da síntese de fenômenos, nos deparamos frequentemente com a incerteza, visto que, ao contrário das ciências exatas, o universo que nos rodeia não possui valores precisos que se

encaixam nas classificações propostas pelo homem. Não à toa, a maior parte das regras existentes, seja nas áreas das ciências exatas, seja nas da saúde e nas humanas (essas principalmente), trazem consigo um conjunto de exceções que são similares entre si, dificultando o processo de classificação de atividades.

Ao observarmos um conjunto de veículos, podemos tentar adivinhar seus pesos e, apesar de não termos ideia de como fazê-lo sem uma balança, sabemos intuitivamente que um carro popular pesa bem menos que uma SUV e que esta pesa bem menos que um caminhão. Seguindo uma lógica, pode-se dizer que um carro popular tem um peso “leve”, enquanto uma caminhonete tem um peso “intermediário” e que um caminhão é “pesado”. Sob essa perspectiva, como classificar outro modelo de veículo, como uma Kombi, que não é necessariamente “leve” nem “intermediária”? Ou, refletindo sobre características físicas, dentre diversos espectros de cores, como classificar aquelas que se encontram em um matiz entre o verde e o azul? Seriam verdes ou azuis?

Para estas questões, de cunho difuso, foi elaborada uma lógica de classificação por pertinência, a lógica Fuzzy (ZADEH, 1965). Segundo esta lógica, um elemento não precisa pertencer exclusivamente a um conjunto. Ele pode ser pertinente a um e a outros. No exemplo anterior, nossa Kombi pode ter 25% de pertinência ao conjunto dos veículos leves e 75% de pertinência ao conjunto dos veículos intermediários. Um verde-água pode ter 65% de pertinência ao conjunto das cores verdes e 35% de pertinência ao conjunto das cores azuis. Ambos terão uma classificação dentro dos questionamentos levantados: a Kombi é um veículo com peso intermediário (apesar de não ser completamente “intermediário”, nem tampouco “leve”) e o verde-água é considerado verde, mas possui alguma pertinência não excludente com o azul.

A lógica de conjuntos Fuzzy foi utilizada nesta pesquisa porque uma de suas características principais é a aplicação de termos da linguagem natural,¹⁹ sob a forma de regras linguísticas de comparação, para classificar elementos de conjuntos cujas propriedades não são definidas em termos de afirmações verdadeiras ou falsas. Ao contrário, encontram-

¹⁹ Linguagem natural é a linguagem aquela usada no cotidiano. A maior parte das linguagens de programação tenta se aproximar da linguagem natural para facilitar a compreensão dos programadores. Porém, sabemos que algumas relações lógicas realizadas com computadores exigem conhecimentos muito específicos da ciência da computação e da matemática, o que faz com que, com o aumento da complexidade dos programas, os códigos criados pelos programadores tornem-se menos “naturais”. A lógica Fuzzy é uma lógica de conjuntos que procura utilizar termos da linguagem natural para classificar os elementos de conjuntos.

se em situações de incerteza ou onde as fronteiras de conjuntos não são bem definidas, como vimos nos exemplos anteriores.

O Anexo I desta pesquisa descreve as principais características da lógica Fuzzy, que a tornam apropriada para classificar os conjuntos que serão analisados pelo sistema proposto neste trabalho. Destacamos que a escolha desta lógica se deu porque os dados a serem classificados são posições e rotações dos membros do corpo humano presentes nas capturas e, de acordo com os valores de cada um destes parâmetros, o sistema pode confundir o movimento realizado. Fuzzy auxilia no processo de determinação do tipo de movimento realizado.

A “confusão” ao classificarmos se dá porque cada grupo de valores de posições e rotações, armazenados *frame a frame*, poderá “pertencer a” ou indicar um ou mais movimentos, porém, com o uso da lógica Fuzzy, poderemos determinar qual o maior ou menor grau de pertinência dos valores a um conjunto de movimentos, conseguindo classificá-los.

Outra característica da lógica Fuzzy que determinou sua utilização neste projeto é a sua escalabilidade, ou seja, a possibilidade de ampliar a quantidade de movimentos analisados, de acordo com a necessidade do usuário ou do projeto. A lógica de funcionamento dos conjuntos Fuzzy permite a adição de novas regras linguísticas através da utilização dos parâmetros existentes (as rotações e as posições das juntas do corpo humano), simplificando o processo de reconhecimento de poses e movimentos.

4.5 Reconhecendo movimentos através de uma perspectiva imprecisa (Fuzzy)

Como visto no capítulo 3, o reconhecimento de poses de um indivíduo em movimento já foi estudado por Bregler (1997), Gleicher (1997), Gavrilin (1999), Bobick e Davis (2001) e Santofimia et al (2014), dentre outros, que utilizaram diversos métodos probabilísticos para alcançar resultados precisos nas áreas de segurança, saúde (SHULTZ; BIRMINGHAM; JENKYN, 2011), fisioterapia (SCHLOEMER et al., 2013), esportes (LÄDERMANN et al., 2014), para a pesquisa de diversos assuntos, tais como o estudo dos movimentos de atletas amadores, a análise das juntas de jogadores de tênis (com vistas a evitar lesões nas

articulações), a análise de caminhar para determinar o melhor movimento para uma sessão de fisioterapia, além de outras aplicações.

Na maior parte dos métodos propostos por estes pesquisadores, nota-se o uso de sistemas que exigem um alto esforço em processamento de dados, assim como o uso de modelos probabilísticos complexos, além das limitações já apresentadas no capítulo 3. Dentre estes métodos, podemos citar os modelos ocultos de Markov e os filtros de Kalman recursivo (BREGLER, 1997); métodos probabilísticos (SIDENBLADH; BLACK; FLEET, 2000); os “estados definidos de máquina” (*Definite state machine*) para o reconhecimento de movimentos (ZHONGXIANG et al., 2002); o reconhecimento de ações através de silhuetas, usando “Modelos de Mistura Gaussianos” (*Gaussian Mixture Models – GMMs*); a análise de componente principal (NAIEL; ABDELWAHAB; EL-SABAN, 2011); os descritores de bancos de dados (HUANG et al., 2016), dentre outros.

Todos os sistemas propostos e as análises realizadas por pesquisadores nesta área apresentam vantagens e limitações oriundas tanto dos *hardwares* utilizados para o reconhecimento de poses e atividades quanto dos algoritmos propostos para realizar tais tarefas. De fato, apesar de ser uma área que já possui várias direções traçadas há algumas décadas, podemos dizer que, por desenvolver e depender de tecnologias, ainda há muito a ser pesquisado e diversos problemas a serem solucionados, bem como a perspectiva do surgimento de novas questões nos próximos anos. Nossa questão, entretanto, não passa por muitos dos problemas por eles pesquisados, visto que usamos um sistema de capturas de movimentos que não depende de câmeras para tal. Logo, problemas de oclusão não serão relevantes.

No protótipo projetado para este trabalho, optou-se por utilizar a lógica Fuzzy da seguinte maneira: foram escolhidos dois movimentos bem similares – a corrida e a caminhada – de um mesmo indivíduo. Após serem capturados, foram inseridos como os dados de entrada de um *software* responsável por reconhecê-los e classificá-los.

Quando cartunizados, estes movimentos se diferiram mutuamente. Entretanto, quando capturados, algumas de suas características eram similares, como, por exemplo, o movimento do quadril e a abertura das pernas do indivíduo, visto que, dependendo da intensidade da caminhada ou da corrida, seus movimentos podiam ser bastante similares (uma caminhada rápida pode se assemelhar a uma corrida lenta). Escolhemos, pois, estes

movimentos, justamente para avaliar a precisão do sistema em diferenciá-los e reconhecê-los. E para isto utilizamos somente estes dois fatores (o movimento do quadril e a abertura das pernas) como fontes de análise.

4.5.1 Desenvolvimento do *software* de identificação de movimentos

As etapas que o algoritmo de reconhecimento dos movimentos percorrerá estão descritas no diagrama a seguir (Figura 109) e podem ser encontradas nos arquivos dos códigos-fonte que acompanham esta tese, no DVD em anexo.

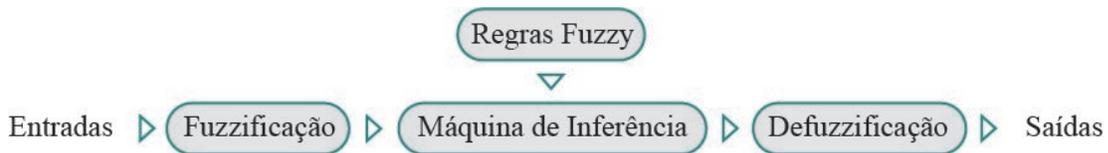
Figura 109 - Diagrama etapas do reconhecimento de ações usando lógica Fuzzy.



Fonte: Diagrama elaborado pelo autor.

A etapa do algoritmo proposto responsável pelo reconhecimento dos movimentos seguiu um fluxo de funcionamento, conforme o apresentado no Anexo I e reproduzido abaixo. Para uma melhor compreensão do funcionamento dos conjuntos Fuzzy e suas aplicações, verificar o Anexo I.

Figura 110 - Fluxo de trabalho adotado para uso com a lógica Fuzzy em nosso algoritmo.



Fonte: Ilustração do autor.

Considerando o diagrama, nosso algoritmo foi projetado para seguir as seguintes etapas:

- 1 – Leitura dos dados de entrada do sistema. Em nosso caso, são os parâmetros de posição do COG (*Center of Gravity* – “Centro de Gravidade do Corpo”), rotações do COG e das juntas do corpo humano, extraídas do arquivo BVH. Cada parâmetro possui pelo menos três valores a serem armazenados, aqueles referentes aos eixos X, Y e Z;
- 2 – Armazenamento dos dados de entrada do sistema em variáveis. O armazenamento em variáveis permite o uso destas para a análise subsequente do movimento realizado;
- 3 – *Fuzzificação*. Nesta etapa criamos as variáveis linguísticas, associamos as mesmas aos valores linguísticos e definimos as funções de pertinência.
- 4 – Criação das regras Fuzzy. É nesta seção que são criadas as condições para o reconhecimento dos movimentos realizados, de acordo com as entradas e a fuzzificação;
- 5 – Atuação da máquina de inferência. Não projetamos a máquina de inferência. Utilizamos uma biblioteca pronta em C# para realizar a associação das variáveis linguísticas e das funções de pertinência com as entradas do sistema. Resolvemos utilizar uma solução já existente, visto que nosso propósito não é replicar ou aprimorar uma máquina de inferência;

6 – *Defuzzificação*. Assim como na etapa anterior, utilizamos uma biblioteca pronta para realizar a defuzzificação dos dados recebidos da máquina de inferência. Também aqui resolvemos usar uma biblioteca pronta para realizar esta tarefa, visto que a programação do bloco de código constituiria um esforço desnecessário, tendo em vista os objetivos deste trabalho; e

6 – *Saídas*. Nesta etapa, os dados já passaram pela análise e pela classificação nas etapas anteriores, e o sistema entrega o tipo de movimento identificado nos arquivos (dados) de entrada. No nosso experimento, além da identificação do movimento, entregamos um arquivo que contém os *frames* onde acontecem as poses extremas do movimento identificado, para a utilização na etapa posterior de cartunização do movimento pelo *script* criado no *software* de animação 3D.

Criado com regras linguísticas de fácil assimilação, o algoritmo projetado para separar os conjuntos de dados em movimentos distintos necessitou de parâmetros bem definidos para atuar com eficiência. Assim, antes de iniciarmos a programação do bloco de *softwares* responsável pelo reconhecimento dos movimentos advindos do arquivo BVH e, portanto, da captura de movimentos, separamos os dados referentes aos membros que desejávamos analisar (o movimento do COG e a abertura das pernas) e criamos a Tabela 8. Ela apresenta os dados dos movimentos do eixo Y do COG – parâmetro principal utilizado na análise dos movimentos escolhidos – e a abertura das pernas do indivíduo.

É necessário salientar que, para cada tipo de movimento, pode ser necessário focar em um ou mais parâmetros específicos do corpo (as rotações de determinadas juntas, como as dos braços, da cabeça ou a posição do COG). Considerando que a base de parâmetros é relacionada à quantidade de juntas do corpo humano, concluímos que esta não aumentará com o passar do tempo e as relações entre eles estarão mapeadas (como, por exemplo, as aberturas possíveis entre o braço e o antebraço, as rotações da coluna em relação ao quadril, as rotações das pernas em relação aos ombros, etc.), é possível prever que, após a inserção constante de dados advindos das capturas de movimentos, em algum momento, grande parte dos movimentos humanos mais comuns será mapeado e reconhecido pelo algoritmo. Esta propriedade do sistema reforça sua característica de escalabilidade, mostrando sua vantagem.

Para este experimento, inicialmente foram gravadas mais de vinte amostras de cada movimento. Entretanto, apenas seis movimentos de caminhada e dez de corrida foram aproveitados. Isto se deu devido às limitações do equipamento utilizado para realizar as capturas.²⁰

Tabela 8-Análise dos arquivos gravados. Na primeira coluna, estão os movimentos realizados. Na segunda, os ângulos máximos encontrados entre as pernas, durante os movimentos, e, na terceira, a posição Y do COG (movimento de cima para baixo e vice-versa).

Gravação	Abertura das pernas	Posição do COG
Caminhada001_Char00_bipe	14.24	93.46
Caminhada002_Char00_Bip_bipe	26.82	93.73
Caminhada003_Char00_bip_bipe	23.36	93.62
Caminhada004_Char00_bip_bipe	30.80	93.57
Caminhada005_Char00_bip_bipe	25.28	93.52
Caminhada006_Char00_bipe	33.80	93.65
Corrida_1_Char00_bip_bipe	54.32	98.32
Corrida_2_Char00_bipe	42.50	102.04
Corrida_3_Char00_bipe	47.25	103.55
Corrida_4_Char00_bipe	42.94	104.92
Corrida_5_Char00_bipe	35.13	101.56
Corrida_6_Char00_bip_bipe	35.86	98.96
Corrida_7_Char00_bip_bipe	43.77	101.15
Corrida_8_Biped_Char00_bipe	29.06	100.23
Corrida_9_Char00_bip_bipe	49.17	98.08
Corrida_Char00_bipe	56.62	98.62

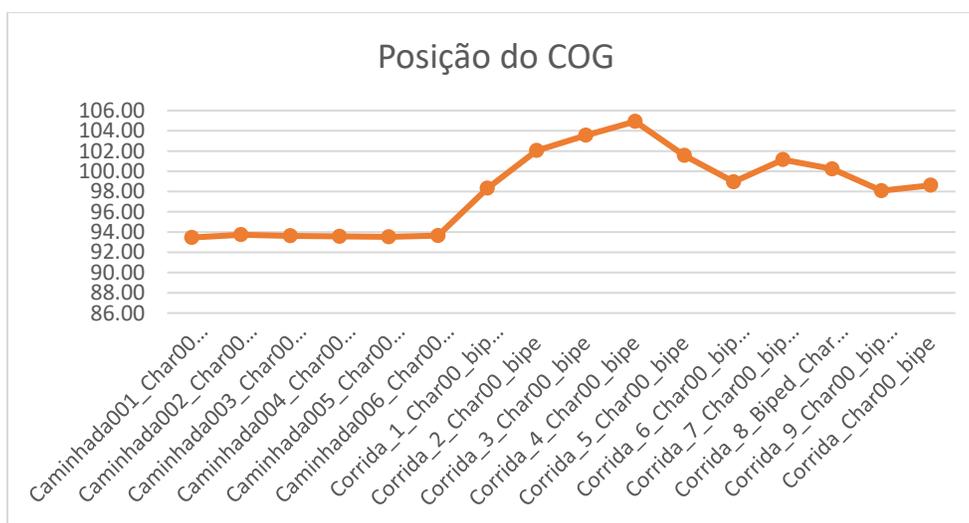
Fonte: Dados coletados pelo autor (2018).

²⁰ Ao utilizar magnetômetros, o sistema fica vulnerável às interferências causadas por metais e aparelhos eletrônicos presentes no ambiente. A frágil conexão entre os cabos que conectam os magnetômetros também fizeram com que diversos membros deixassem de ser capturados, durante a execução dos movimentos. Tentamos nos afastar das fontes metálicas, porém, por diversas vezes, o sistema sofria interferências. Portanto, das gravações realizadas, apenas algumas puderam ser usadas com sucesso nesta tese. Acreditamos que, com equipamentos mais robustos, essa limitação possa ser superada.

Como é possível observar, os valores das aberturas das pernas nas caminhadas são menores que os valores equivalentes nas corridas, o mesmo acontecendo com a posição do COG. Apesar desses valores estarem dentro de algumas faixas de valores, notamos, na realização do experimento, que alterações sutis, ocorridas durante a caminhada ou a corrida, faziam com que, quando nosso algoritmo fosse construído usando a lógica booleana – na qual o julgamento de valores de entrada é feito através de comparações que retornam valores “verdadeiro” e “falso” –, ele, por vezes, confundia o movimento realizado.²¹

Para facilitar a compreensão dos movimentos com os quais estávamos lidando, plotamos os valores da posição Y do COG e da abertura das pernas em formato de gráfico (figuras Figura 111 e Figura 112), e fizemos, a partir deles, nossa análise visual. Nota-se que, enquanto quase não há variações na posição do COG na seção do gráfico equivalente ao movimento de caminhada (os seis primeiros valores), na corrida, as variações são bem acentuadas.

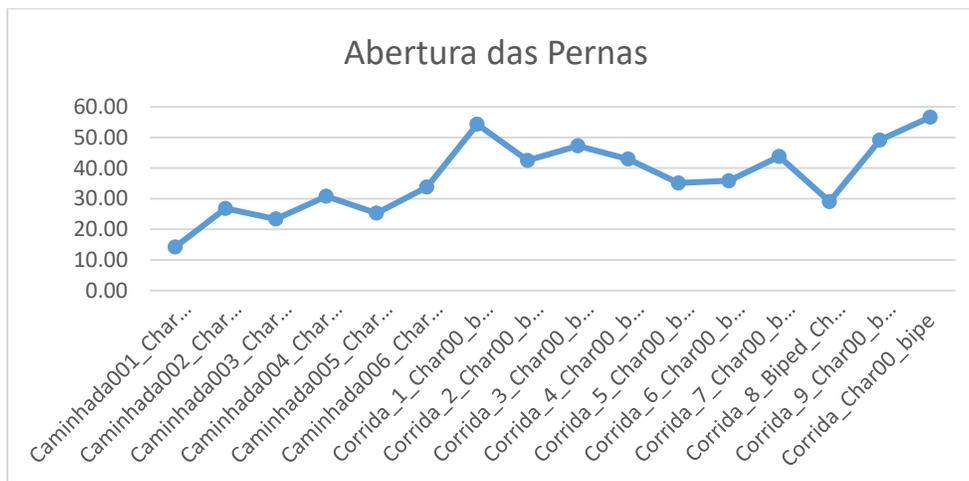
Figura 111 – Gráfico indicando a posição do COG de acordo com o tipo de movimento realizado.



Fonte: Dados do autor.

²¹ Usando a lógica tradicional, determinávamos um movimento apenas através de afirmações condicionais, como “SE o valor da posição do COG for *MAIOR QUE* 45,00, ENTÃO o movimento *É UMA CAMINHADA*”. Porém, como afirmamos anteriormente, esses valores poderiam ser facilmente confundidos pelo nosso algoritmo, caso o indivíduo fizesse um movimento um pouco mais exagerado durante uma caminhada ou resolvesse correr de maneira mais suave.

Figura 112 - Gráfico indicando a abertura das pernas nos movimentos analisados.



Fonte: Dados do autor.

No primeiro gráfico, estão representadas as posições Y do COG, por sua vez, representadas por valores que variam entre 86 cm e 106 cm no eixo vertical. No eixo horizontal, encontram-se as descrições dos arquivos de onde esses dados foram obtidos.

É possível observar que as maiores diferenças existem nos valores da posição Y do COG. Os seis primeiros movimentos representam a caminhada do indivíduo. Verifica-se uma linha praticamente horizontal cujos valores oscilam próximos aos 94 cm, com poucas variações. Logo a seguir, apresentam-se as dez corridas, onde as variações das posições do COG são bem mais acentuadas.

Esta plotagem dos valores foi fundamental para definir os pontos de início de cada movimento no sistema de lógica Fuzzy, visto que, ainda que visualmente eles sejam bem diferentes, em termos absolutos, a variação não é tão acentuada (a diferença entre o valor máximo de uma caminhada e o valor mínimo realizada por uma corrida é de apenas 4 cm, o que equivale a menos de 5% do valor total aferido pelo sistema de captura).

Nesta situação, em que os valores dos parâmetros estão difusos, a lógica Fuzzy é apropriada para reconhecer os movimentos realizados. Por meio dela, estes valores podem ser relativizados e, de acordo com as relações definidas por meio da análise dos dados, aumentamos a precisão do algoritmo de reconhecimento.

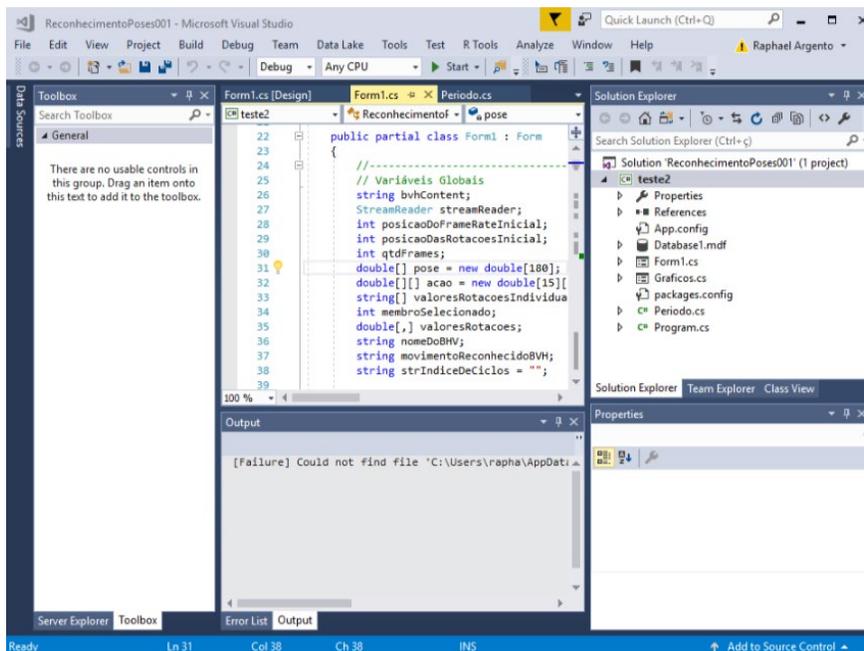
Além disso, ao observarmos os valores da abertura das pernas da captura Caminhada006_Char00_bipe com a Corrida_8_Biped_Char00_bipe, perceberemos que, se usássemos apenas este parâmetro como referência, o segundo movimento poderia facilmente ser confundido com uma caminhada, visto que seu valor equivale aos daqueles aferidos nos movimentos de caminhada. Neste caso, a lógica de conjuntos Fuzzy, utilizando mais de uma regra para comparação dos parâmetros, nos auxiliou a construir o protótipo que classifica a captura de movimentos “Corrida_8_Biped_Char00_bipe” como uma corrida.

Partindo desta análise, passamos à configuração e à programação do sistema. Nosso algoritmo foi programado em C#, uma linguagem de programação criada pela Microsoft e que apresenta uma interoperabilidade entre diversos sistemas diferentes, através do *software* disponibilizado gratuitamente pela empresa, o Visual Studio (Figura 113).

Outra razão para a utilização desta linguagem de programação foi a utilização do Kinect como um equipamento de captura de movimentos, quando começamos nossos experimentos. O Kinect é programado usando bibliotecas disponibilizadas pela Microsoft na linguagem de programação C#. E boa parte do nosso algoritmo foi desenvolvido para esta tecnologia, usando o Visual Studio. Porém, durante a pesquisa, percebemos as limitações de captura com ele²² e resolvemos migrar para o Neuron Mocap, aproveitando o código já desenvolvido em C#. Além disso, todos os *softwares* utilizados nesta pesquisa rodam em ambiente Windows™, que possui uma integração direta e eficiente com o Visual Studio.

²² Apesar de capturar movimentos bem, o Kinect apresenta uma limitação quanto ao seu campo (*range* de atuação), que varia de 0,4m a 4m. Não conseguimos capturar uma corrida ou uma caminhada usando este sistema, em um espaço tão curto. Além disso, o Kinect exige que o indivíduo a ser capturado esteja em uma pose pré-definida para realizar a calibração do sistema, ou seja, a cada novo movimento é necessário realizar uma nova calibração, a partir dessa pose estática, o que nos impediu de capturar uma corrida, por exemplo.

Figura 113 - Tela do Visual Studio, ambiente de desenvolvimento desenvolvido e disponibilizado gratuitamente pela Microsoft™.



Fonte: Tela do *software*.

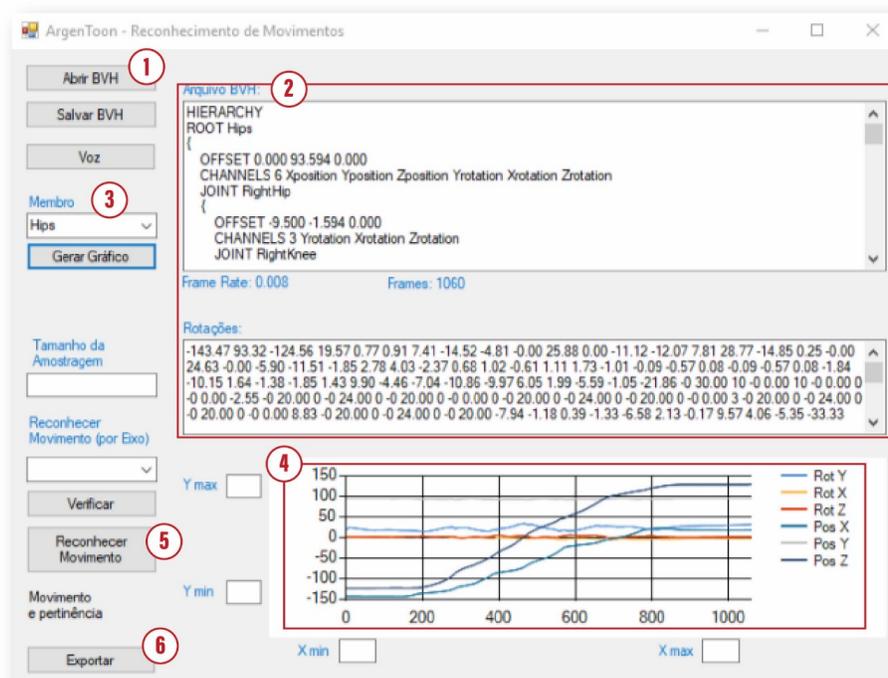
Em nosso experimento, instalamos uma biblioteca disponibilizada para o Visual Studio, a *Accord.Net*²³, que possui um pacote Fuzzy amplamente utilizado e validado. Considerando que necessitávamos da implementação da lógica, não da programação do algoritmo responsável por executá-la, resolvemos utilizar esta biblioteca.

Para que o usuário pudesse interagir com o sistema, criamos uma interface cuja função era permitir que o reconhecimento de movimentos fosse realizado e os movimentos das diversas juntas do corpo pudessem ser analisados de forma gráfica (Figura 114).

Através deste *software*, foi possível: (1) abrir o arquivo BVH, observar sua estrutura e os valores das rotações dos membros do corpo humano; (2) separar cada membro capturado; (3) apresentá-los de forma gráfica, explicitando as posições e rotações nos eixos X,Y e Z; e, finalmente (4), reconhecer o movimento e (5) exportar o arquivo com os dados para uso posterior no 3ds Max.

²³ A Biblioteca Accord.Net pode ser baixada e instalada no Visual Studio por meio deste endereço: <http://accord-framework.net/>.

Figura 114 - Interface gráfica criada para o reconhecimento e análise de movimentos.



Fonte: Interface gráfica programada pelo autor.

Após realizarmos a identificação de cada junta no arquivo BVH e as selecionarmos, armazenamos seus valores de posição e de rotação em variáveis para, a partir destas, definirmos outras variáveis, as linguísticas, usadas pela lógica Fuzzy. As duas variáveis são associadas, constituindo um banco de dados (Figura 115). As variáveis utilizadas nesta pesquisa podem ser encontradas no código-fonte que está no CD em anexo a este trabalho na seção “// O banco de dados”.

Em nosso caso, adicionamos ao banco de dados Fuzzy, as variáveis que criamos e denominamos “posCOG” (posição do COG), “freqCOG” (frequência com que os movimentos se repetiam), “anguloPernaDireita” (o ângulo entre o COG e a perna direita), “anguloPernaEsquerda” (o ângulo entre o COG e a perna esquerda), “tipoDeMovimento” (os possíveis movimentos realizados – corrida ou caminhada) e “aberturaDasPernas” (a diferença entre cada uma das aberturas, resultando em um valor que varia de acordo com o movimento realizado).

Figura 115 - Declaração de variáveis para uso no banco de dados Fuzzy.

```
// O banco de dados
Database posicoesDB = new Database();
posicoesDB.AddVariable(posCOG);
posicoesDB.AddVariable(freqCOG);
posicoesDB.AddVariable(anguloPernaDireita);
posicoesDB.AddVariable(anguloPernaEsquerda);
posicoesDB.AddVariable(tipoDeMovimento);
posicoesDB.AddVariable(aberturaDasPernas);
```

Fonte: Algoritmo criado pelo autor.

A associação destas variáveis com as regras linguísticas Fuzzy foi feita através de funções trapezoidais (sobre funções trapezoidais, ver Anexo I), que definiram a pertinência de cada valor de variável a uma característica a ser identificada, como “baixo”, “alto”, “AnguloPequeno”, “AnguloGrande”, etc. Na Figura 116, são apresentadas as variáveis associadas aos conjuntos Fuzzy.

Figura 116 - Variáveis linguísticas sendo associadas aos valores das variáveis obtidas pelo sistema.

```
// -----
// Conjuntos FUZZY
// --> Transformar isto em uma classe
// Conjuntos Fuzzy do COG (linguística)
FuzzySet fsCogBaixo = new FuzzySet("Baixo", new TrapezoidalFunction(94, 98, TrapezoidalFunction.EdgeType.Right));
FuzzySet fsCogAlto = new FuzzySet("Alto", new TrapezoidalFunction(94, 98, TrapezoidalFunction.EdgeType.Left));

// Valores possíveis do COG (input)
LinguisticVariable posCOG = new LinguisticVariable("PosicaoDoCOG", 0, 140);
posCOG.AddLabel(fsCogBaixo);
posCOG.AddLabel(fsCogAlto);

// --> Transformar isso em uma classe
// Conjuntos Fuzzy da Frequência do COG
FuzzySet fsCogFreqBaixa = new FuzzySet("FrequenciaBaixa", new TrapezoidalFunction(60, 80, TrapezoidalFunction.EdgeType.Right));
FuzzySet fsCogFreqAlta = new FuzzySet("FrequenciaAlta", new TrapezoidalFunction(60, 80, TrapezoidalFunction.EdgeType.Left));

// Valores possíveis da Frequência do COG (input)
LinguisticVariable freqCOG = new LinguisticVariable("FrequenciaDoCOG", 20, 200);
freqCOG.AddLabel(fsCogFreqBaixa);
freqCOG.AddLabel(fsCogFreqAlta);

// --> Transformar isto em uma classe
// Conjuntos Fuzzy da Perna Direita (linguística)
FuzzySet fsPernaDirAnguloPequeno = new FuzzySet("AnguloPequeno", new TrapezoidalFunction(-30, -20, TrapezoidalFunction.EdgeType.Right));
FuzzySet fsPernaDirAnguloGrande = new FuzzySet("AnguloGrande", new TrapezoidalFunction(20, 30, TrapezoidalFunction.EdgeType.Left));

// Valores possíveis da Perna Direita (input)
LinguisticVariable anguloPernaDireita = new LinguisticVariable("AnguloDaPernaDireita", -90, 90);
anguloPernaDireita.AddLabel(fsPernaDirAnguloPequeno);
anguloPernaDireita.AddLabel(fsPernaDirAnguloGrande);
```

Fonte: Algoritmo criado pelo autor.

Figura 117 - Variáveis linguísticas sendo associadas aos valores das variáveis obtidas pelo sistema (cont.).

```
// --> Transformar isto em uma classe
// Conjuntos Fuzzy da Perna Esquerda (linguística)
FuzzySet fsPernaEsqAnguloPequeno = new FuzzySet("AnguloPequeno", new TrapezoidalFunction(-30, -20, TrapezoidalFunction.EdgeType.Right));
FuzzySet fsPernaEsqAnguloGrande = new FuzzySet("AnguloGrande", new TrapezoidalFunction(20, 30, TrapezoidalFunction.EdgeType.Left));

// Valores possíveis da Perna Esquerda (input)
LinguisticVariable anguloPernaEsquerda = new LinguisticVariable("AnguloDaPernaEsquerda", -90, 90);
anguloPernaEsquerda.AddLabel(fsPernaEsqAnguloPequeno);
anguloPernaEsquerda.AddLabel(fsPernaEsqAnguloGrande);

// --> Transformar isto em uma classe
// Conjunto Fuzzy da Abertura entre as pernas (linguística)
FuzzySet fsAberturaDasPernasPequena = new FuzzySet("AberturaDasPernasPequena", new TrapezoidalFunction(33, 35, TrapezoidalFunction.EdgeType.Right));
FuzzySet fsAberturaDasPernasGrande = new FuzzySet("AberturaDasPernasGrande", new TrapezoidalFunction(33, 35, TrapezoidalFunction.EdgeType.Left));
// FuzzySet fsAberturaDasPernasMedia = new FuzzySet("AberturaDasPernasMedia", new TrapezoidalFunction(33, 40, 100, 112));

// Valores possíveis para a abertura entre as pernas (input)
LinguisticVariable aberturaDasPernas = new LinguisticVariable("AberturaDasPernas", 0, 80);
aberturaDasPernas.AddLabel(fsAberturaDasPernasGrande);
aberturaDasPernas.AddLabel(fsAberturaDasPernasPequena);
// aberturaDasPernas.AddLabel(fsAberturaDasPernasMedia);
```

Fonte: Algoritmo criado pelo autor.

Fizemos com que cada conjunto Fuzzy criado correspondesse a uma variável como, por exemplo, o conjunto dos ângulos pequenos da perna esquerda (fsPernaEsqAnguloPequeno, Figura 117), que corresponde a todos os movimentos realizados por este membro do corpo que apresentam valores menores que -20° . Para esta variável, utilizamos uma função trapezoidal de pertinência entre -30° e -20° (em nosso caso, as aberturas das pernas são valores que variam entre -20° e aproximadamente -30° , dependendo do tipo de movimento realizado). Isto faz com que valores mais próximos a -20° sejam menos pertinentes que aqueles situados próximo à -30° ou menores que este. O gráfico da Figura 118 apresenta essa relação, utilizando uma função trapezoidal. Quanto mais o ângulo se aproxima dos -20° , menor consideramos a abertura desta perna.

Isto significa que, quanto maior for o ângulo entre as pernas, maior será o movimento e, conseqüentemente, teremos aí a diferenciação entre corridas e caminhadas. A partir da determinação destas relações, criamos as regras Fuzzy para serem utilizadas no reconhecimento de movimentos, reproduzidas, a seguir, e que se encontram no código-fonte criado nesta pesquisa:

// Regra 1

// Se a frequência do COG for *Alta* e o Ângulo da Perna Direita for *Alto*, então é uma *Corrida*

```
IS.NewRule("Regra 1", "IF FrequenciaDoCOG IS FrequenciaAlta AND AnguloDaPernaDireita IS AnguloGrande THEN TipoDeMovimento IS Corrida");
```

// Regra 2

// Se a posição do COG for *Baixa* e o Ângulo da Perna Direita for *Baixo*, então é uma *Caminhada*

```
IS.NewRule("Regra 2", "IF FrequenciaDoCOG IS FrequenciaBaixa AND AnguloDaPernaDireita IS AnguloPequeno THEN TipoDeMovimento IS Caminhada");
```

// Regra 3

// Se a posição do COG for *Alta* e o Ângulo da Perna Esquerda for *Alto*, então é uma *Corrida*

```
IS.NewRule("Regra 3", "IF FrequenciaDoCOG IS FrequenciaAlta AND AnguloDaPernaEsquerda IS AnguloGrande THEN TipoDeMovimento IS Corrida");
```

// Regra 4

// Se a posição do COG for *Baixa* e o Ângulo da Perna Esquerda for *Baixo*, então é uma *Caminhada*

```
IS.NewRule("Regra 4", "IF FrequenciaDoCOG IS FrequenciaBaixa AND AnguloDaPernaEsquerda IS AnguloPequeno THEN TipoDeMovimento IS Caminhada");
```

// Regra 5

// Se as posições do COG forem *Baixas* e o ângulo entre as pernas forem *Baixos*, então é uma *Caminhada*

```
IS.NewRule("Regra 1", "IF PosicaoDoCOG IS Baixo AND AberturaDasPernas IS AberturaDasPernasPequena THEN TipoDeMovimento IS Caminhada");
```

// Regra 6

// Se as posições do COG forem *Altas* e o ângulo entre as pernas forem *Baixos*, então é uma *Corrida*

```
IS.NewRule("Regra 2", "IF PosicaoDoCOG IS Alto AND AberturaDasPernas IS AberturaDasPernasGrande THEN TipoDeMovimento IS Corrida");
```

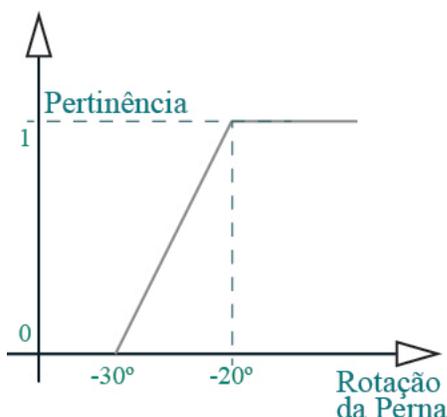
// Regra 7

// Se as posições do COG forem *Baixas* e o ângulo entre as pernas forem *Médios*, então é uma *Caminhada*

```
IS.NewRule("Regra 3", "IF PosicaoDoCOG IS Baixo AND AberturaDasPernas IS AberturaDasPernasMedia THEN TipoDeMovimento IS Caminhada");
```

```
// Regra 8
// Se as posições do COG forem Altas e o ângulo entre as pernas forem Médios, então é uma Corrida
// IS.NewRule("Regra 4", "IF PosicaoDoCOG IS Alto AND AberturaDasPernas IS AberturaDasPernasMedia
THEN TipoDeMovimento IS Corrida");
```

Figura 118 - Gráfico relacionando a pertinência dos ângulos da perna esquerda à abertura. Valores menores que -30° são relativos às aberturas grandes. A área compreendida entre -30° e -20° é considerada uma zona de transição de aberturas pequenas para aberturas maiores.



Fonte: Gráfico do autor.

A mesma lógica se aplicou às demais variáveis, de acordo com os dados representados na Figura 111 (gráfico da posição do COG). Ao final deste processo, o sistema estava programado e pronto para prosseguir à próxima fase: a defuzzificação.

O processo de defuzzificação, consiste na interpretação dos valores de entrada do sistema, em conjunto com as variáveis do conjunto Fuzzy, já condicionadas ao processo de fuzzificação dos conjuntos. Nesta etapa, também utilizamos a biblioteca Fuzzy da *Accord.NET*. Após alguns ajustes nos valores das variáveis de entrada, o sistema foi capaz de reconhecer os movimentos inseridos no sistema. O código deste processo pode ser visto na seção “Saídas”, do código-fonte, encontrado no DVD que acompanha esta pesquisa.

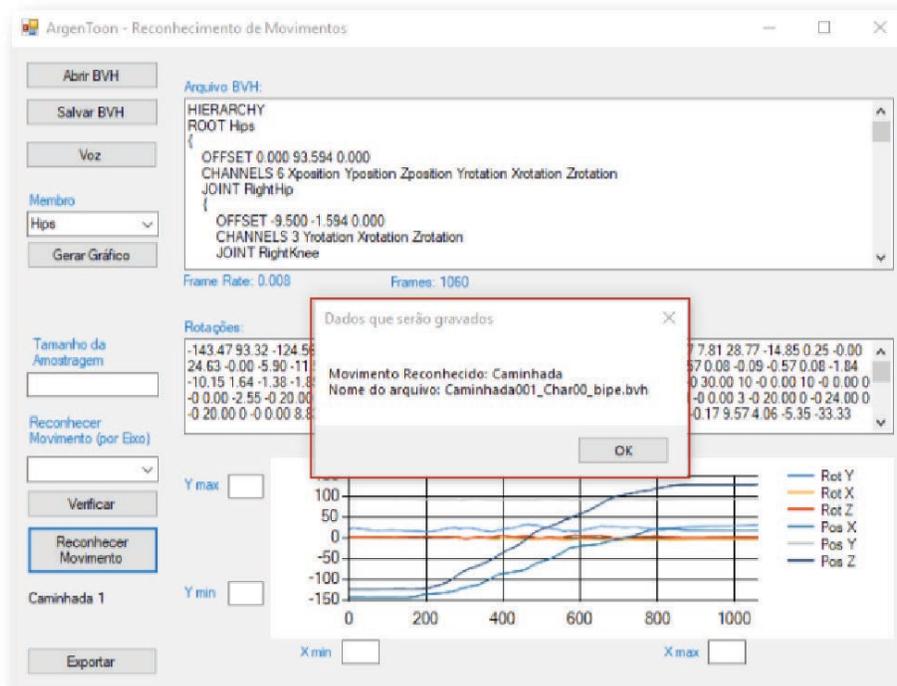
Considerando que esta é a última etapa, ao abrirmos um arquivo BVH com a captura de movimentos, o sistema facilmente identificava o movimento, conforme visto na Figura 119, através de uma caixa de diálogo do *software*.

A partir dessa saída, gravamos um arquivo texto contendo as seguintes informações: o nome do arquivo, o tipo de movimento realizado e onde foram encontradas as poses

extremas do movimento do indivíduo. Esses dados são importantes, pois permitem que a etapa de implementação do algoritmo de aplicação do movimento no 3ds Max (o *script* programado para aplicação dos movimentos), seja realizada de maneira simples e eficiente.

Uma das vantagens no uso da Lógica Fuzzy para o reconhecimento de movimentos humanos em animação é a possibilidade de expansão do sistema por parte do usuário, conforme a necessidade de reconhecimento de novos movimentos se apresentar. Para isto, decidem-se quais variáveis serão utilizadas (quais rotações de juntas do corpo ou as posições do COG) e criam-se regras Fuzzy para determinar quais relações corresponderão a uma pose específica.

Figura 119 - Funcionamento do *software* de reconhecimento de movimentos. Os destaques em vermelho mostram o *feedback* do sistema para o usuário.

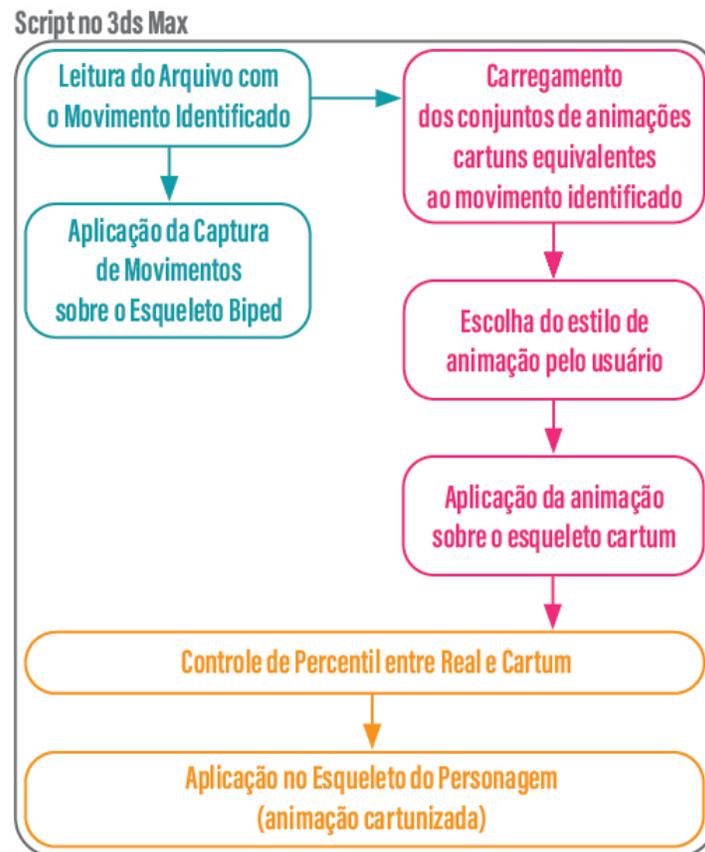


Fonte: *Software* criado pelo autor.

4.5.2 Aplicação no software de animação

Após o principal estágio de nosso sistema de reconhecimento de movimentos ser finalizado e operar de maneira eficiente, criamos um *script* no 3ds Max para aplicarmos o conjunto de animações cartuns à captura de movimentos. Conforme apresentado na Figura 106, o planejamento para aplicarmos a animação cartum ao personagem envolve algumas etapas, entre elas, a importação dos movimentos para cada um dos esqueletos presentes no *software* de animação: o esqueleto cartum e o esqueleto Biped. A Figura 120 mostra o fluxo deste processo.

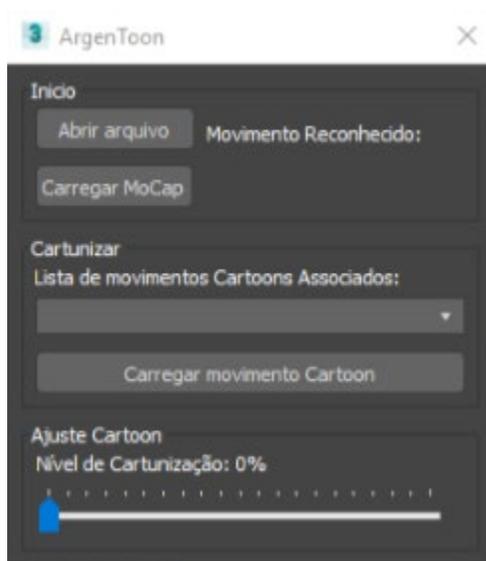
Figura 120 - Fluxo de etapas realizadas no 3ds Max para cartunização de personagens.



Fonte: Esquema elaborado pelo autor.

Dentro do 3ds Max, foi criada uma interface para que o usuário pudesse escolher qual arquivo de captura de movimentos seria importado, escolher qual a lista de movimentos cartuns associados a ele e carregar o movimento cartum selecionado (Figura 121).

Figura 121 - Interface de usuário do ArgenToon, *script* para cartunização de personagens.



Fonte: Interface criada pelo autor.

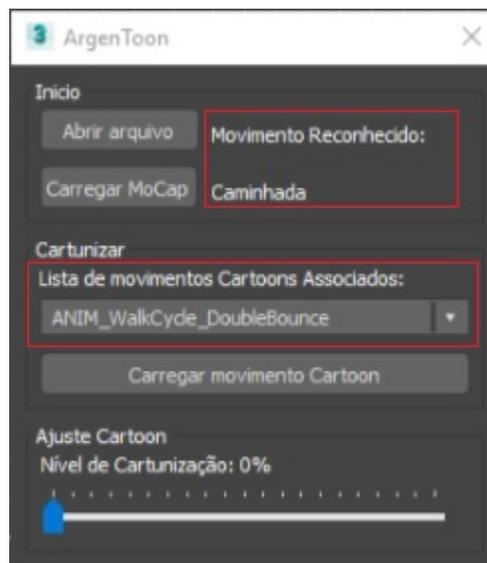
A primeira etapa foi realizada através da leitura do arquivo texto gerado pelo *software* de reconhecimento de movimentos. Em nossa interface, essa leitura é feita através do botão “Abrir arquivo” que, quando pressionado, permite ao usuário escolher o arquivo texto gerado pelo *software* de reconhecimento de movimentos. Ao abrir o arquivo, são armazenadas as informações gerais do sistema (como a localização do diretório onde estão os demais arquivos) e o armazenamento dos dados gravados no arquivo, como o movimento reconhecido e os *frames* nos quais cada uma das poses foi encontrada. O código-fonte do *script* pode ser encontrado no DVD anexado a esta pesquisa. Essa etapa específica encontra-se no bloco de código denominado “1 Abre o arquivo da captura de movimentos...”.

Após a leitura do movimento reconhecido, o *script* realiza uma busca, no diretório onde se encontram os arquivos, por aquele correspondente à captura de movimentos e o importa para o 3ds Max no qual apresenta ao usuário os dados reconhecidos. Caso tenha se enganado ao escolher seu arquivo, o usuário pode abrir outro arquivo com o reconhecimento.

Ao selecionar Carregar MoCap, nosso script importa as capturas gravadas em formato BIP para o esqueleto Biped e, ao mesmo tempo, carrega as animações cartuns associadas ao

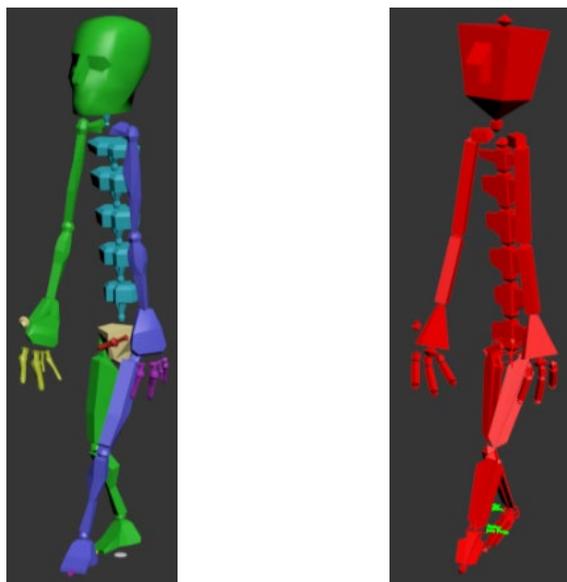
movimento, presentes no conjunto de animações cartunizadas (Figura 122). Em nosso exemplo, é possível observar o movimento “Caminhada” e um dos movimentos listados pelo *script* de busca, o “ANIM_WalkCycle_DoubleBounce”. O bloco de código associado a estas ações está no código-fonte, no bloco iniciado por “3 - Importar a captura de movimentos para o Biped padrão do 3DS Max”.

Figura 122 - Interface do ArgenToon, após a abertura do arquivo do reconhecimento de movimentos.



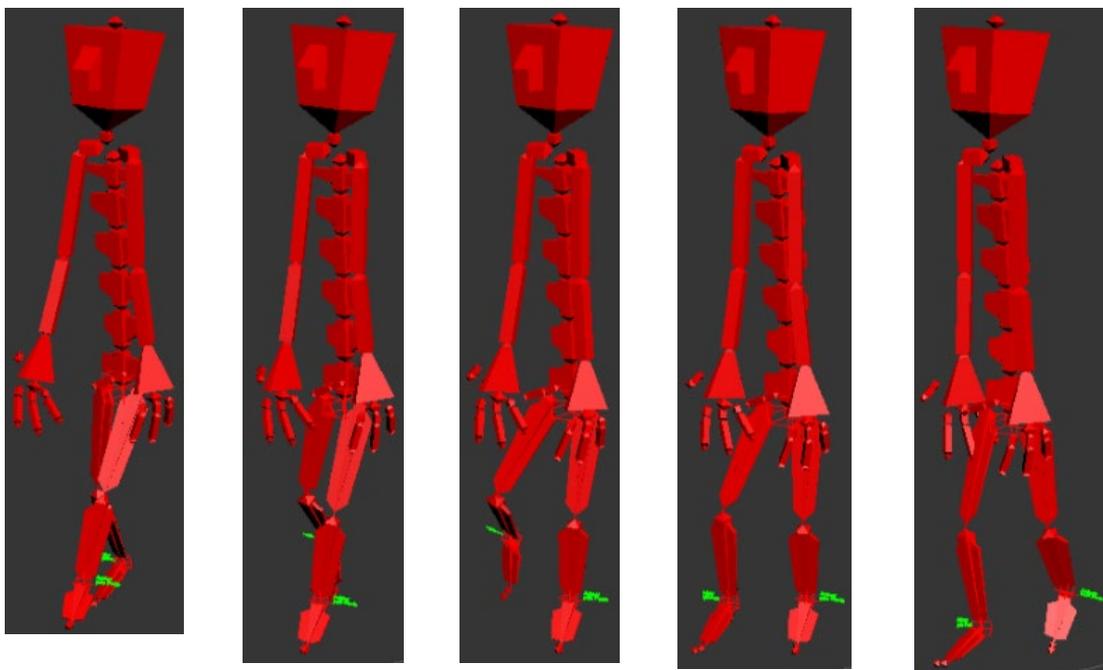
Fonte: Interface criada pelo autor.

Figura 123 - Captura de movimentos aplicada ao Biped e influenciando o esqueleto que atua sobre o modelo.



Fonte: Captura de tela do *software* 3ds Max.

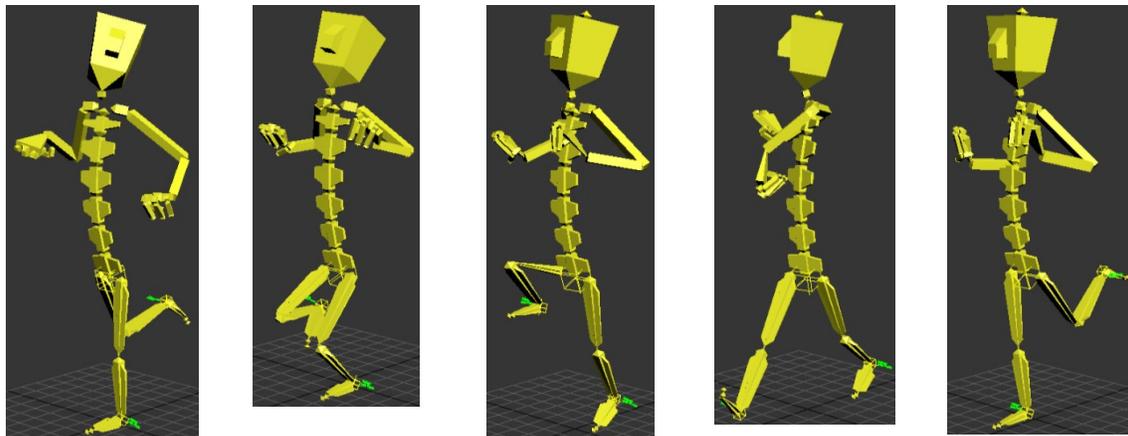
Figura 124 - Sequência de *frames* com a captura de movimentos influenciando o esqueleto do modelo 3D.



Fonte: Captura de tela do *software* 3ds Max com o *script* de inserção de movimentos funcionando.

Após selecionar o movimento cartum e pressionar o botão “Carregar movimento Cartum”, nosso *script* ArgenToon percorre a *timeline* do 3ds Max e, nos *frames* indicados pelo arquivo de reconhecimento de movimentos, insere as animações cartunizadas, completando o processo de cartunização. No código-fonte, este bloco de algoritmo pode ser encontrado sob o nome “4 - Criação das camadas de animação onde as animações do banco de animações serão inseridas. Inserção em cada um dos índices”.

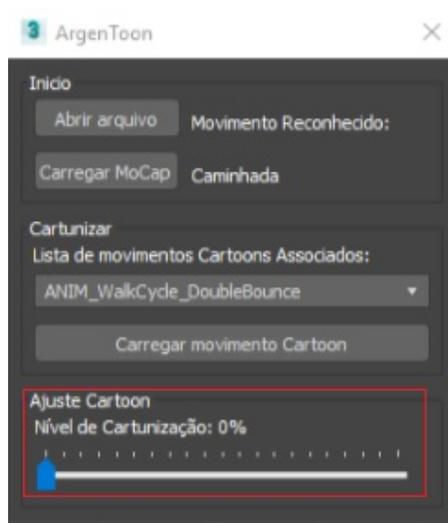
Figura 125 - Movimento cartum aplicado ao esqueleto cartum e influenciando o esqueleto que movimenta o modelo 3D.



Fonte: Captura de tela do *software* 3ds Max com o *script* de inserção de movimentos cartuns funcionando.

Para permitir ao usuário decidir o nível de cartunização do personagem, foi criado um controle deslizante “Ajuste Cartum” (Figura 125), que ajusta o quanto cada esqueleto (o do Biped e o cartum) controlará o esqueleto do personagem através de *constraints* (controles de restrição) dentro do 3ds Max.

Figura 126 - Controle de ajuste de cartunização no ArgenToon.



Fonte: Interface criada pelo autor.

Através deste experimento, comprova-se a validade do método proposto para cartunizar capturas de movimentos, além de provarmos que é possível automatizar essa cartunização, possibilitando ao animador definir estilos.

4.6 Reflexões

Com este experimento, criamos um método eficaz de cartunização de personagens. Por meio da lógica Fuzzy foi possível reconhecer movimentos para utilização em *softwares* de animação 3D.

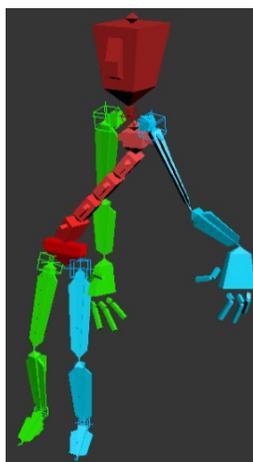
Apesar de nosso sistema apresentar uma solução eficiente para o reconhecimento de movimentos e a conversão para cartuns, alguns pontos devem ser relatados aqui para que possam ser solucionados futuramente.

O reconhecimento de movimentos foi realizado com uma quantidade mínima de membros do corpo. Quanto mais membros forem utilizados, tende-se a ter uma maior precisão no reconhecimento de movimentos, porém, mais regras Fuzzy precisam ser adicionadas para que o sistema os reconheça. Com o tempo, a quantidade de regras Fuzzy tende a crescer, o que pode indicar que será necessário subdividir o sistema de reconhecimento em etapas, fazendo com que, em cada uma delas, o algoritmo encaminhe o

processo para uma subseção Fuzzy. Para realizar isto, serão necessários mais testes e, conseqüentemente, mais pesquisas na área.

O processo de transição entre a animação realista e a animação cartum também não é completamente preciso. Como esta utiliza uma série de controles (denominados *constraints*) do 3ds Max que acionam os dois esqueletos, alguns ruídos foram notados durante a rotação dos *bones* do esqueleto do personagem (Figura 127).

Figura 127 - Problema encontrado durante a transição do movimento realista para o movimento cartum.



Fonte: Captura de tela do *software* 3ds Max.

Acreditamos que o próximo passo para o avanço desta pesquisa é o reconhecimento de ações com o uso de janelas temporais, como indica a pesquisa de Santofimia (2014), algo que pode ser desenvolvido futuramente. Além disso, a criação das regras Fuzzy para o reconhecimento dos movimentos (e, posteriormente, o reconhecimento de ações), podem se constituir como um caminho para a criação do “alfabeto-base” necessário para o aprimoramento da área.

Os conceitos advindos da área da visão computacional aplicados ao design de animação constituem-se como caminhos a serem explorados. O uso de algoritmos de aprendizado de máquina, através de treinamento do sistema, pode aprimorar o reconhecimento de movimentos e sua subsequente aplicação para a cartunização de personagens.

Desejamos que esta e outras ferramentas sejam projetadas para aprimorar a linguagem da animação em seus mais diversos aspectos, permitindo que o design seja beneficiado pelos avanços neste setor.

Por fim, acreditamos que as animações que aqui foram criadas localmente podem transformar-se em um banco global de animações, disponibilizados *online*, onde animadores e estúdios poderão fazer o *upload* e o *download* de animações em diversos estilos para aplicação em diversas produções, permitindo que estes profissionais foquem em outros aspectos da produção como, por exemplo, na criação de histórias e elementos visuais mais atraentes.

CONCLUSÃO

Ao propormos uma pesquisa em design e tecnologia, procuramos levantar algumas lacunas existentes nesta área do saber, de maneira que pudéssemos contribuir para o aprimoramento do setor e permitir o avanço do conhecimento com uma singela, mas (acreditamos que) significativa contribuição.

Ao optarmos por explorar os meandros de uma tecnologia que é apresentada para o público espectador como algo tão próximo (os filmes de heróis utilizam intensamente no cinema a captura de movimentos) e para os pesquisadores e profissionais da área como algo ainda distante – devido aos seus custos e aos seus processos de aquisição no Brasil – esperávamos reduzir o hiato entre esses dois públicos e, ao mesmo tempo, propor uma forma inovadora de se pensar animação associada aos processos de produção tecnológicos utilizados na indústria do audiovisual.

Escolhemos um método que consideramos apropriado para as pesquisas no campo do design. Alguns dos desafios que enfrentaríamos nesta jornada já eram conhecidos e outros foram se apresentando, conforme a pesquisa foi avançando. O primeiro deles foi encontrar uma lacuna na área do conhecimento para explorarmos. O uso da tecnologia de captura de movimentos no setor de animação e design não se constitui como uma novidade. É algo já explorado, há alguns anos, como apresentamos em nossa revisão bibliográfica. Precisamos pesquisar a literatura para, de fato, encontrarmos um campo a ser expandido.

E o fizemos em duas frentes: a primeira, contrapondo a animação, enquanto uma sequência de desenhos em movimento, ao cinema, enquanto fotografia em movimento. A animação, explorando a imagem do desenho e se afastando do realismo, o cinema, usando a base do realismo para desenvolver suas narrativas e cada qual desenvolvendo sua própria linguagem e se alimentando com a outra para aprimorar o processo narrativo.

Neste processo, ao explorarmos a história da animação e as formas pelas quais sua linguagem se formou, nos deparamos com os “princípios de animação” que se constituem como a base de conhecimento necessária para que o animador possa dar vida a um personagem. Esse conhecimento tão fundamental para a animação apresentou-se confuso em seus conceitos, em especial no exagero que, segundo nossa interpretação, pode ser

aplicado a qualquer um daqueles principais. Ao analisarmos sua divisão, refletimos e propusemos uma nova maneira de interpretá-los, para clarear nossa visão sobre eles.

Sob uma ótica mais abrangente, separamos os que são baseados em métodos dos que são baseados em movimentos, dos que se referem à forma do personagem e dos que são uma consequência direta da influência do teatro e da linguagem cinematográfica sobre a animação.

A partir desta classificação, acreditamos que levantamos a discussão sobre a lógica (ou a falta desta) por trás dos princípios, além de possibilitarmos que outros autores se debruceem sobre o “exagero”, enquanto um princípio de animação, e sua evidente presença nos demais.

A segunda frente de revisão bibliográfica se deu nos campos da captura de movimentos e da visão computacional, áreas que naturalmente se entrelaçam. Por meio desta revisão, descobrimos as lacunas existentes naquela área do conhecimento e suas descobertas recentes, como a ausência de uma linguagem para a interpretação dos movimentos realizados por seres humanos (lacuna) e a rotulação de trechos de vídeos (avanço no campo do conhecimento).

Resolvemos explorar a interpretação de movimentos humanos para cartunizá-los, unindo, dessa maneira, as duas frentes de revisão bibliográfica, através do uso de uma tecnologia inovadora. Para fazê-lo, precisamos desenvolver um protótipo de testes.

A aquisição do equipamento para criarmos nosso protótipo foi outro desafio, pois precisávamos de um que suprisse nossas necessidades (precisão dos movimentos capturados e facilidade de uso) e limitações (orçamentárias). Inicialmente, tentamos utilizar o Kinect™, da Microsoft como um sistema para a captura de movimentos. Apesar de se apresentar muito eficiente para esse fim, se mostrou ineficiente em nossos testes, quando o usuário estava em frente a ele. Precisávamos capturar movimentos que exigissem um deslocamento maior do que o limite alcançável pelo Kinect. Após diversas tentativas malsucedidas, optamos por usar outro sistema, o Neuron Mocap, que se mostrou bastante eficaz para esse propósito.

Nosso avanço na área do conhecimento que escolhemos para pesquisar se deu ao inserirmos um método para reconhecer e cartunizar os movimentos adquiridos por esse equipamento, utilizando uma lógica de programação – a Fuzzy – que é parametrizada, por

meio de comandos que se assemelham à linguagem natural (ou seja, a comunicação humana).

Desenvolver uma tecnologia aplicável ao processo de produção de animação e que permita o reconhecimento e a conversão de movimentos adquiridos por sistemas de Mocap é algo que o setor da animação ainda carece tanto em termos acadêmicos quanto mercadológicos. O que se observa atualmente é a utilização da força de trabalho humana para corrigir os movimentos adquiridos dos sistemas de Mocap, fazendo com que o trabalho dos profissionais do setor se assemelhe a algo que poderia ser realizado de forma mecânica.

Nossa abordagem identificou este problema, bem como outras lacunas existentes na área e, através de uma pesquisa aplicada, propôs uma das muitas soluções possíveis para ele. Assim como diversos pesquisadores na área do design e da visão computacional propuseram métodos, modelos e propostas para soluções em suas áreas, acreditamos que esta pesquisa apresenta uma contribuição para o avanço do conhecimento, ao propor um método que pode ser reproduzido e ampliado por outros pesquisadores da área, obtendo resultados similares e avanços.

O campo do design é bastante amplo, se considerarmos sua história e seus desdobramentos fomentados pelo avanço das tecnologias, em especial, a dos computadores e dispositivos móveis. Atualmente, as divisões do design e suas nomenclaturas são tantas que nos é difícil enumerá-las. A animação e a captura de movimentos são áreas de exploração do designer, especialmente no Brasil, onde diversas instituições de ensino formam profissionais com um perfil de produção voltado para o audiovisual – direção que vem se intensificando com a popularização dos dispositivos móveis e a expansão desse setor no Brasil.

Em nosso país, começamos a verificar a formação de equipes de P&D em empresas e estúdios, objetivando criar soluções que atendessem às suas necessidades específicas e diferenciassem seus produtos daqueles criados por seus concorrentes, assim como fazem os estúdios e as empresas internacionais.¹

¹ Como exemplo, destacamos no site da TV Globo – maior emissora do Brasil – uma área dedicada à inovação e à infraestrutura, responsável pela criação e pelo desenvolvimento de novas tecnologias para o audiovisual (<http://redeglobo.globo.com>, subseção “tecnologia”).

Por outro lado, na academia, o design é uma das áreas do conhecimento que mais se beneficiam, quando se utilizam de pesquisa aplicada, pois esta contribui para propor uma geração de conhecimento e de produtos cuja relevância é cada vez mais reconhecida pela sociedade, devido ao retorno mais imediato às demandas sociais, através do processo de geração de conhecimento acadêmico, por meio do rigor metodológico da pesquisa.

Assim, acreditamos que esta pesquisa contribui para o avanço do conhecimento tanto na área do design quanto no meio acadêmico, por seu caráter inovador, em alguns aspectos. O primeiro deles é a utilização do modelo de pesquisa para o campo do design proposta pela Universidade de Illinois e a efetiva geração de um resultado positivo para uma pesquisa empírica que envolveu o desenvolvimento de uma nova tecnologia.

Finalmente, o uso da lógica Fuzzy para o reconhecimento de movimentos advindos de capturas é o ponto mais importante do trabalho, visto que sua utilização com a finalidade aqui proposta não foi encontrada nas publicações pesquisadas, o que, mais uma vez, demonstra sua relevância.

Por outro lado, ela deixa, para uma resolução posterior, algumas questões que não puderam ser solucionadas, devido à limitação do escopo do trabalho e ao tempo disponível para que lhe fossem propostas novas respostas. Dentre as ainda latentes e que não puderam ser nela contempladas, encontra-se a criação de um alfabeto completo de entidades-movimento, que pretendemos construir em pesquisas futuras. Apesar desta questão não ter sido inteiramente resolvida, é possível, a partir do protótipo proposto, ampliar seu escopo, para que este possa reconhecer uma quantidade maior de movimentos humanos.

Além disso, acreditamos que a lógica Fuzzy pode ser aplicada em diversas etapas para a classificação de movimentos capturados. Essas etapas poderiam ser definidas por meio de outros algoritmos, como as árvores de decisões ou outros de inteligência artificial, que serviriam como uma pré-classificação de movimentos.

A discussão sobre os princípios de animação também pode ser aprofundada, formando um arcabouço teórico e crítico sobre seus usos e um possível desmembramento desses princípios em unidades lógicas que façam mais sentido para seus usuários. Tal desmembramento pode ser utilizado como base para uma nova compreensão destes e de suas aplicações acadêmicas e no mercado do audiovisual.

Por fim, esperamos que esta pesquisa contribua para o avanço do conhecimento no campo do design e incentive outros pesquisadores a seguirem o mesmo caminho, propondo novas questões para uma reflexão, através de pesquisas aplicadas, tão necessária para o estreitamento de nossas relações com a sociedade e com os profissionais de nossa área.

REFERÊNCIAS

Formatado: Inglês (Estados Unidos)

ABDELWAHAB, M. M.; EL-SABAN, M.; NAIEL, M. A. Multi-view human action recognition system employing 2DPCA. 2011 **IEEE Workshop on Applications of Computer Vision**, WACV 2011, p. 270-275, 2011.

AFSAR, P.; CORTEZ, P.; SANTOS, H. Automatic visual detection of human behavior: A review from 2000 to 2014. **Expert Systems with Applications**, v. 42, n. 20, p. 6935–6956, 2015. Internet. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417415003516>>.

AGGARWAL, J. K. Human activity analysis: a review. **ACM Comput. Surv. Article**, v. 43, n. 16, p. 43, 2011.

AGUIAR, E. De. **Character animation from a motion capture database**. 2003. 87p. Dissertação (Master Thesis in Computer Science) – Computer Graphics Group. Max-Planck-Institut für Informatik, Saarbrücken, 2003.

AGUIAR, H., OLIVEIRA, J., CALDEIRA, A. M., MACHADO, M. A. S., SOUZA, R. C., & Tanscheit, R. (2007). **Inteligência computacional aplicada à administração, economia e engenharia em Matlab**. São Paulo: Thomson Learning, 2007.

ALIVERTI, A. et al. Functional evaluation and rehabilitation engineering. **IEEE Pulse**, v. 2, n. 3, p. 24-34, 2011.

ANDERSON, J.; ANDERSON, B. The myth of the persistence of vision revisited. **Journal of Film and Video**, v. 45, n. 1, p. 3-12, 1993.

ANTUNES JR, J. A. V; DRESCH, A.; LACERDA, D. P.. **Design science research: a method for science and technology advancement**. Heidelberg: Springer International Publishing Switzerland, 2015.

ASTERIX: o gaulês. Direção de Ray Goossens. França e Bélgica: Studios Belvision, 1967. DVD (68 min), son., color., legendado.

Formatado: Inglês (Estados Unidos)

BALLARD, D. H.; BROWN, C. M. **Computer vision**. 1. ed. New Jersey: Prentice-Hall, Inc., 1982.

BANCROFT, T. **Creating characters with personality: for Film, TV, animation, video games, and graphic novels**. New York: Watson-Guption Publications, 2006.

BARBOSA JÚNIOR, A. L. **Arte da animação: técnica e estética através da história**. 3ª ed. São Paulo: Editora Senac, 2001.

BARRIER, M. **Hollywood cartoons: american animation in its golden age**. Oxford: Oxford University Press, 2003.

- BECKERMAN, H. **Animation: the whole story**. 15. ed. New York: Allworth Press, 2003.
- BELL, J.; ORENSTEIN, C.; POSNER, D. N. (ed.). **The Routledge companion to puppetry and material performance**. New York: Routledge, 2014.
- BENDAZZI, G. **Animation: a world history**, Volume I: Foundations - golden age. New York: Focal Press, 2016a.
- BENDAZZI, G. **Animation: a world history**, Volume II: The birth of a style – the three markets. New York: Focal Press, 2016b.
- BENDAZZI, G. **Animation: a world history**, Volume III: Contemporary times. New York: Focal Press, 2016c.
- BIRMINGHAM, T. B.; JENKYN, T. R.; SHULTZ, R. Differences in neutral foot positions when measured barefoot compared to in shoes with varying stiffnesses. **Medical Engineering and Physics**, v. 33, n. 10, p. 1309-1313, 2011. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.medengphy.2011.05.009>>.
- BLAIR, P. **Advanced animation: learn how to draw animated cartoons**. Laguna Hills: Walter Foster Publishing, 1980.
- BLAIR, P. **Cartoon animation**. Laguna Hills: Walter Foster Publishing, Inc., 1994.
- BLAKE, R.; SHIFFRAR, M. Perception of human motion. **Annual Review of Psychology**, v. 58, p. 47–73, 2007. Internet. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16903802>><<http://www.annualreviews.org/doi/abs/10.1146/annurev.psych.57.102904.190152>><<http://www.ncbi.nlm.nih.gov/pubmed/16903802>>.
- BLACK, M. J.; FLEET, D. J.; SIDENBLADH, H. Stochastic tracking of 3D human figures using 2D image motion. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000. v. 1842.
- BLANK, M. G. L. S. E. I. M. B. R. Actions as space-time shapes. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 29, n. 12, p. 2247-2253, 2005.
- BOBICK, A. F.; DAVIS, J. W. The recognition of human movement using temporal templates. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 23, n. 3, p. 257-267, 2001.
- BORENSTEIN, G. **Making things see: 3D vision with Kinect, Processing, Arduino and Makerbot**. Sebastopol: Maker Media, 2012.
- BREGLER, C. Learning and recognizing human dynamics in video sequences. **Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, n. June, p. 568-574, 1997.

CAMPBELL, J. ;LUOMALA, K. **The hero with a thousand faces**. 2. ed. Novato, CA:New World Library, 2008.

Formatado: Inglês (Estados Unidos)

CHEETHAM, M. et al. Arousal, valence, and the uncanny valley: psychophysiological and self-report findings. **Frontiers in Psychology**, v. 6, n. August, 2015. Internet. Disponível em: <<http://journal.frontiersin.org/Article/10.3389/fpsyg.2015.00981/abstract>>.

CHONG, A. **Digital animation**. Lausanne: AVA Publishing, 2008.

COHEN, I. LEE, M. W. A model-based approach for estimating human 3D poses in static images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 28, n. 6, p. 905-916, 2006.

COOPER, F. T.; MAURICE, A. B. **The history of the nineteenth century in caricature**. New York: Dodd, Mead and Company, 1904.

COWELL, C.; MILLER-ZARNEKE, T. **The art of how to train your dragon**. New York: Newmarket Press, 2010.

CRAFTON, D. The veiled genealogies of animation and cinema. **Animation**, v. 6, n. 2, p. 93-110, 2011.

COX, E. (1992). Fuzzy fundamentals. **IEEE Spectrum**, 29 (10), p. 58-61. <https://doi.org/10.1109/6.158640>

Formatado: Português (Brasil)

Código de campo alterado

Formatado: Português (Brasil)

CRUZ, G. F. S. **Brinquedos óticos animados e o ensino de design**. 2017. 115f. Tese (Doutorado em Design). Departamento de Artes e Design, Pontifícia Universidade Católica, Rio de Janeiro, 2017.

Formatado: Português (Brasil)

CUTTING, J. E. A program to generate synthetic walkers as dynamic point-light displays. **Behavior Research Methods & Instrumentation**, v. 10, n. 1, p. 91-94, 1978. Internet. Disponível em: <<http://www.springerlink.com/index/10.3758/BF03205105>>.

DAVIS, L. S.; GAVRILA, D. M. 3-D model-based tracking of humans in action: a multi-view approach. In: **Proceedings of the 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, 1, California. Anais... California: IEEE Computer Society Press, 1996. Internet. Disponível em: <<http://www.umiacs.umd.edu/users/fgavrila,lsdg/>>.

Formatado: Inglês (Estados Unidos)

DULAC, N.; GAUDREAU, A.; HIDALGO, S. **A companion to early cinema**. West Sussex: Wiley-Blackwell, 2012.

DUTTA, T. Evaluation of the Kinect(tm) sensor for 3-D kinematic measurement in the workplace. **Applied Ergonomics**, v. 43, n. 4, p. 645-649, 2012. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.apergo.2011.09.011>>.

EISNER, W. **Quadrinhos e arte sequencial**. 3ª ed. São Paulo: Martins Fontes, 1999.

GAVRILA, D. M. The visual analysis of human movement: a survey. **Computer Vision and Image Understanding**, v. 73, n. 1, p. 82-98, 1999.

Formatado: Português (Brasil)

GERNSHEIM, A.; GERNSHEIM, H.A **concise history of photography**. New York: Grosset & Dunlap, 1965.

GERNSHEIM, A.; GERNSHEIM, H. Re-discovery of the world's first photograph. **Journal of Photography of the George Eastman House**, v. 1, n. 6, 1952.

GIESEN, R. **Puppetry, puppet animation and the digital age**. Boca Raton, FL: CRC Press / Taylor & Francis Group, 2019.

GLEICHER, M. Motion editing with spacetime constraints. **Proceedings of the 1997 symposium on Interactive 3D graphics SI3D 97**, v. pp, 139-?, 1997. Internet. Disponível em: <<http://portal.acm.org/citation.cfm?doid=253284.253321>>.

Formatado: Português (Brasil)

GLIMPSES OF INDIAN ANIMATION. Direção de R. Swamy. India: Indian Films Division, 1997. DVD, son., color.

GOMBRICH, E. H. **História da arte**. 14ª ed. São Paulo: Círculo do Livro, 1972.

Formatado: Inglês (Estados Unidos)

GRANUM, E. A; MOESLUND, T. B. survey of computer vision-based human motion capture. **Computer Vision and Image Understanding**, v. 81, n. 3, p. 231-268, 2001. Internet. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S107731420090897X>>.

Formatado: Português (Brasil)

HART, C. **Cartoon cool: how to draw new retro-style characters**. New York: Watson-Guptill Publications, 2005.

HANNAVY, J. **Encyclopedia of nineteenth-century photography**. New York: Taylor & Francis Group, 2008.

HOBSBAWN, E. **Era dos extremos: o breve século XX: 1914-1991**. 2ª ed. São Paulo: Companhia das Letras, 1995.

HOLTE, M. B. et al. Human pose estimation and activity recognition from multi-view videos: comparative explorations of recent developments. **Ieee Journal of Selected Topics in Signal Processing**, v. 6, n. 5, p. 538-552, 2012.

Formatado: Português (Brasil)

HUANG, T.-H. et al. Visual storytelling. **Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies**, p. 7, 2016. Internet. Disponível em: <<http://research.microsoft.com/apps/pubs/default.aspx?id=264715>>.

Formatado: Inglês (Estados Unidos)

Formatado: Português (Brasil)

JANA, A. **Kinect for Windows SDK programming guide**. 1. ed. Birmingham: PACKT Publishing, 2012.

JI, X. J. X.; LIU, H. L. H. Advances in view-invariant human motion analysis: a review. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, v. 40, n. 1, p. 13-24, 2010.

JOHANSSON, G. Visual perception of biological motion and a model for its analysis. **Perception & Psychophysics**, v. 14, n. 2, p. 201-211, 1973. Internet. Disponível em: <papers3://publication/uuid/9F0D433F-719B-496B-AAA3-28542970C1E2>.

JOHNSTON, O.; THOMAS, F. **The illusion of life: Disney animation**. New York: Walt Disney Productions, 1981.

JONES, A.; OLIFF, J. **Thinking animation: bridging the gap between 2D and CG**. Boston: Thomson Course Technology, 2007.

KINERD, N. Motion capture study of human movement recognition. 2012. **Internet**. Disponível em: <http://tigerprints.clemson.edu/all_theses>.

KIRCHER, A. *Ars magna lucis et umbrae*. 3. ed. [s.l.: s.n.]

KITAGAWA, M.; WINDSOR, B. **MoCap for artists: workflow and techniques for motion capture**. Burlington: Focal Press, 2008.

KUEHNE, H. et al. HMDB: A large video database for human motion recognition. In: **2011 IEEE International Conference on Computer Vision**, Barcelona. Anais... Barcelona: IEEE, 2011.

Formatado: Inglês (Estados Unidos)

KWON, J. Y.; LEE, I. K. An animation bilateral filter for slow-in and slow-out effects. **Graphical Models**, v. 73, n. 5, p. 141-150, 2011. Disponível em: <http://dx.doi.org/10.1016/j.gmod.2011.02.002>.

Formatado: Alemão (Alemanha)

LÄDERMANN, A. et al. Kinematics of the shoulder joint in tennis players. **Journal of Science and Medicine in Sport**, v. 19, n. 1, p. 56-63, 2014. Internet. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1440244014002175>.

Formatado: Português (Brasil)

LAPTEV, I. et al. Learning realistic human actions from movies. In: **26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR**, Anais.2008.

Formatado: Inglês (Estados Unidos)

LEE, H.-J.; CHEN, Z. Determination of 3D human body postures from a single view. **Computer Vision, Graphics, and Image Processing**, v. 30, p. 148-168, 1985.

LI, W.; LIU, Z.; ZHANG, Z. Expandable data-driven graphical modeling of human actions based on salient postures. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 18, n. 11, p. 1499-1510, 2008.

LIN, W.; SUN, M. Activity recognition using a combination of category components and local models for video surveillance. **IEEE transactions on circuits and and Systems for video technology**, v. 18, n. 8, p. 1128-1139, 2008. Internet. Disponível em: <http://ieeexplo

Formatado: Português (Brasil)

re.ieee.org/xpls/abs_all.jsp?arnumber=4543872>.

LINN, A. The AI Blog. Teaching computers to describe images as people would. 2016. **Internet**. Disponível em: <<https://blogs.microsoft.com/ai/teaching-computers-to-describe-images-as-people-would/>>. Acesso em 20 out. 2017.

Formatado: Português (Brasil)

MACHADO, A. **Pré-cinemas & pós-cinemas**. 3. ed. Campinas: Papyrus, 2005.

Formatado: Português (Brasil)

MAGNENAT-THALMANN, N.; THALMANN, D. **Handbook of virtual humans**. West Sussex: John Wiley & Sons, 2006.

MARTINEZ-DEL-RINCON, J.; NEBEL, J. C.; SANTOFIMIA, M. J. Episodic reasoning for vision-based human action recognition. **Scientific World Journal**, 2014.

Formatado: Português (Brasil)

MASCARELLO, F. **História do cinema mundial**. Campinas: Papyrus, 2006.

Formatado: Português (Brasil)

MAXWELL, D. R. **Graphical marionette**. Proc. ACM SIGGRAPH/SIGART Workshop, p. 172-179, 1983.

Formatado: Inglês (Estados Unidos)

MCCLOUD, S. **Desvendando os quadrinhos**. São Paulo: M. Books do Brasil Ltda., 2005.

Formatado: Português (Brasil)

MENACHE, A. **Understanding motion capture for computer animation**. 2nd. ed. Burlington, MA: Morgan Kaufmann, 2011.

NEWHALL, B. **The history of photography: from 1839 to the present day**. New York: The Museum of Modern Art, 1949.

OBSERVATÓRIO BRASILEIRO DO CINEMA E DO AUDIOVISUAL. **Anuário estatístico do Cinema Brasileiro**, 2015 [s.l.] ANCINE, 2016.

Formatado: Português (Brasil)

OS TRÊS PORQUINHOS: DVD Disney Animation Collection. Direção de David Hand. Walt Disney Studios, 2009. DVD (61 min), son., color., legendado.

OHYA, J.; ISHII, K.; YAMATO, J. Recognizing human action in time-sequential images using hidden Markov model. **Computer Vision and Pattern Recognition**, p. 379-385, 1992. **Internet**. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=223161%5Cnhttp://ieeexplore.ieee.org/ielx2/418/5817/00223161.pdf?tp=&arnumber=223161&isnumber=5817%5Cnhttp://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=223161&contentType=Conference+Publication>.

PAIK, K. **The art of Ratatouille**. San Francisco: Chronicle Books LLC, 2007.

PAIVA, J. P. T. M. The use of motion capture in non-realistic **Animation**. n. September, 2014.

POGGENPOHL, S.; SATO, K. Models of dissertation research in design. **3rd Doctoral Education in Design Conference**, 2003.

POLLACK, P. **The picture history of photography**: from the earliest beginnings to the present day. New York: Harry N. Abrams, Inc. Publishers, 1977.

POPPE, R. Vision-based human motion analysis: an overview. **Computer Vision and Image Understanding**, v. 108, n. 1-2, p. 4-18, 2007.

Formatado: Inglês (Estados Unidos)

PRÄKEL, D. **The visual dictionary of photography**. Lausanne: AVA Publishing SA, 2010.

PUMARES, M. J. et al. **Mapping the animation industry in Europe**. Strasbourg: European Audiovisual Observatory (Council of Europe), 2015.

Formatado: Inglês (Estados Unidos)

PUPPETRY. **Encyclopedia Britannica online**, 21 de fevereiro de 2018. Internet. Disponível em <https://www.britannica.com/art/puppetry>. Acesso em 21 de fevereiro de 2018.

Formatado: Português (Brasil)

RASKAR, R. et al. Prakash: lighting aware motion capture using photosensing markers and multiplexed illuminators. **ACM Transactions on Graphics**, v. 26, n. 3, p. 36, 2007. Internet. Disponível em: <http://dl.acm.org/citation.cfm?id=1276377.1276422%5Cnhttp://portal.acm.org/citation.cfm?doid=1276377.1276422>.

Formatado: Inglês (Estados Unidos)

Formatado: Inglês (Estados Unidos)

ROSENBLUM, N. **A world history of photography**. 3. ed. New York, London, Paris: Abbeville Press, 1997.

Formatado: Inglês (Estados Unidos)

RUSSELL, B. C. et al. LabelMe: a database and web-based tool for image annotation. **International Journal of Computer Vision**, v. 77, n. 1-3, p. 157-173, 2008.

SCHLOEMER, S. et al. Time-to-contact demonstrates modulation of postural control during a dynamic lower extremity task. **Gait & posture**, v. 38, n. 4, p. 658-662, 2013. Internet. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23522669>.

Formatado: Português (Brasil)

SHOTTON, J. et al. Real-time human pose recognition in parts from single depth images. **Studies in Computational Intelligence**, v. 411, p. 119-135, 2013.

Formatado: Português (Brasil)

STREIT, P. **Comparação de parâmetros biomecânicos entre sistemas de captura de movimentos**: avaliação do Microsoft Kinect. 2013. Universidade do Estado do Rio de Janeiro, 2013.

STURMAN, D. J. A brief history of motion capture for computer character animation. **Internet**. Disponível em: https://www.siggraph.org/education/materials/HyperGraph/animation/character_animation/motion_capture/history1.htm. Acesso em: 25 maio. 2016.

Formatado: Português (Brasil)

Formatado: Inglês (Estados Unidos)

STURMAN, D. J. Computer puppetry. **IEEE Computer Graphics and Applications**, v. 18, n. February, p. 38-45, 1998.

Formatado: Inglês (Estados Unidos)

STURMAN, D. J.; ZELTZER, D. A survey of glove-based input. **IEEE Computer Graphics and Applications**, v. 14, n. 1, p. 30-39, 1994.

SUPPA, R. Thinking animation: bridging the gap between 2D and CG. [s.l.: s.n.]???

TOSI, V. Cinema before cinema: the origins of scientific cinematography, 2005. . Internet. Disponível em: <<http://books.google.dk/books?id=ka0fAQAIAAJ>>.

Formatado: Português (Brasil)

WANG, H. et al. The effects of load carriage and muscle fatigue on lower-extremity joint mechanics. **Research Quarterly for Exercise and Sport**, v. 84, n. November 2013, p. 305-12, 2013. Internet. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24261009>>.

Formatado: Inglês (Estados Unidos)

WANG, J. et al. The cartoon animation filter. **ACM Transactions on Graphics**, v. 25, n. 3, p. 1169, 2006.

Formatado: Inglês (Estados Unidos)

WEINHANDL, J. T. et al. Validation of a single camera three-dimensional motion tracking system. **Journal of biomechanics**, v. 43, n. 7, p. 1437-40, 2010. Internet. Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-77951979086&partnerID=tZOtx3y1>>.

Formatado: Alemão (Alemanha)

Formatado: Inglês (Estados Unidos)

WIKIPEDIA. Cadeias de Markov. **Internet**. Disponível em: <https://pt.wikipedia.org/wiki/CadeiasDe_Markov>. Acesso em: 16 dez. 2017.

WILLIAMS, R. **The animator's survival kit**. London: Faber & Faber, 2002.

WILLIS, K. D. D. A pre-history of handheld projector-based interaction. **Personal and Ubiquitous Computing**, v. 16, n. 1, p. 5-15, 2012.

WRIGHT, J. A. **Animation writing and development from script to pitch**. Burlington, MA: Focal Press, 2005.

Formatado: Português (Brasil)

UMA CILADA PARA ROGER RABBIT. Direção de Robert Zemeckis. Walt Disney (Sonopress), 1988, DVD (104 min), son., color., legendado.

Formatado: Português (Brasil)

ZADEH, L. A. (1965). Fuzzy sets. **Information and Control**, 8 (3), p. 338-353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X).

Código de campo alterado

Formatado: Inglês (Estados Unidos)

Formatado: Inglês (Estados Unidos)

ZHONGXIANG, L. et al. Multiple animated characters motion fusion. **Journal of Visualization and Computer Animation**, v. 13, n. 5, p. 275-285, 2002.

ANEXO I – Lógica Fuzzy

A lógica Fuzzy é uma das muitas técnicas utilizadas na área de inteligência computacional. Ela foi elaborada para aprimorar o modo como os sistemas computacionais tratam conjuntos de dados.

Tradicionalmente, os cientistas da computação analisavam os dados que a eles eram confiados, através da comparação e da lógica condicional. Os dados eram comumente tratados, através de funções e da lógica booleana. Essas relações previam que um sistema receberia os dados, os trataria (ou interpretaria) e devolveria uma saída.

A lógica por trás do tratamento de dados baseia-se em relações que podem ser verdadeiras ou falsas. O resultado da comparação entre dois ou mais valores que são entradas de um sistema somente pode ser assumido como sendo verdadeiro ou falso.

Imaginemos um conjunto A e um conjunto B, ambos compostos por diversos valores numéricos. É comum encontrarmos em algoritmos uma condição do tipo: “Se o valor de um elemento de A for maior que o valor de um elemento de B, então execute uma função”, ou “Se o valor de um elemento de A for verdadeiro, então escreva algo para o usuário”.

Essa lógica tradicional permitiu o avanço dos sistemas computacionais por décadas (e, ainda hoje, é utilizada em diversos *softwares*), considerando que os valores encontrados nos conjuntos A e B pertencem a esses conjuntos e que as relações entre eles são dadas através da teoria de conjuntos estudada em Matemática. Assim, os conceitos de “união”, “interseção”, “exclusão”, dentre outros, são utilizados para tratar dados e fornecer uma resposta ao usuário.

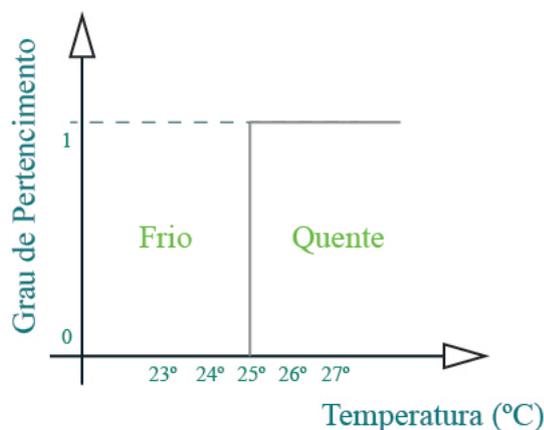
Apesar de resolver uma série de situações, a lógica tradicional de conjuntos exclui condições mais próximas à realidade cotidiana do usuário. Considere, por exemplo, um conjunto contendo uma série de valores referentes à altura de uma população adulta e que estes valores variem entre 1,60m e 2,00m. Levando em conta que a maior parte dos indivíduos tenha uma altura média de 1,70m, pode-se dizer que um indivíduo com 1,80m é considerado “alto”, enquanto que um indivíduo com 1,60m é considerado “baixo”.

Entretanto, onde seria inserido um indivíduo com 1,68m? Ele é considerado “um pouco baixo”? Considerando as classificações rígidas da teoria de conjuntos,

torna-se difícil classificar este indivíduo pois, comparando com um de 1,70m, ele é considerado alto, mas ao ser colocado próximo a outro de 1,80m, ele é baixo. Além disso, para classificar os elementos do cotidiano, as pessoas usam, em geral, conceitos linguísticos que se aproximam mais da experiência sensorial que de um conjunto restrito e limitado de valores.

Outra situação cotidiana também pode levar a confusões em um sistema que utiliza a lógica tradicional. Assim como a altura, podemos definir que um dia está quente se apresenta temperaturas em torno de 25°C ou maiores e frio se elas estiverem abaixo deste valor. Porém, sabemos que em uma temperatura de 24°C faz com que um dia esteja muito mais próximo do quente que uma temperatura de 17°C. Segundo a lógica booleana, só poderíamos classificar estes conjuntos usando valores absolutos, ou seja, estes intervalos seriam classificados apenas como frios ou quentes, não existindo possibilidade de um dia “estar um pouco quente”, ou “um pouco frio”, ou “muito quente”, ou “muito frio”.

Figura 128 - Lógica booleana – Gráfico de exemplo com temperaturas.



Fonte: Ilustração do autor.

De fato, em muitas situações cotidianas, é difícil, senão impossível, realizar classificações em termos de “verdadeiro” e “falso”. Geralmente, as fronteiras entre os fenômenos são um pouco mais difusas. É nestas situações que a lógica Fuzzy torna-se relevante. Segundo seu criador, Lofti Zadeh:

Mais comum que o contrário, as classes de objetos encontrados no mundo físico real não têm critérios precisos de pertinência. Por exemplo, as classes de animais claramente incluem cachorros, cavalos, pássaros, etc.

como seus membros e claramente excluem objetos tais como pedras, fluidos, plantas, etc. Entretanto, objetos tais como estrelas do mar, bactéria, etc. têm uma condição ambígua em relação à classe de animais... De fato, o fato é que tais “classes” definidas imprecisamente representam um papel importante no pensamento humano, particularmente nos domínios do reconhecimento de padrões, comunicação de informação e abstração (Zadeh, 1965).

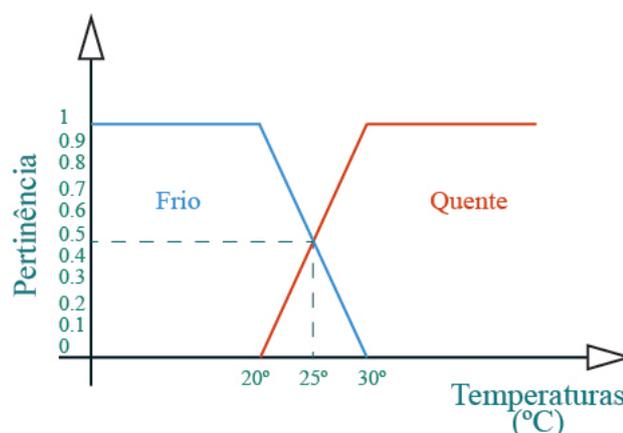
Concretamente, a lógica Fuzzy foi bastante adotada pelas áreas de controle em sistemas e tomadas de decisões, devido à sua simplicidade, facilidade de adoção e sofisticação. Um de seus estudiosos, Cox (1992), afirma que “requisitos complexos podem ser implementados, por meio de controles surpreendentemente simples, facilmente mantidos e pouco custosos” (COX, 1992, tradução nossa).

Segundo Aguiar e Oliveira Junior (2007): “A lógica Fuzzy tem por finalidade o estudo dos princípios formais do raciocínio aproximado” e estabelece que cada elemento de um conjunto fuzzy possui uma relação de pertinência com o conjunto do qual faz parte.

Assim, considerando o primeiro dos dois exemplos referidos anteriores, pode-se considerar que aquele com 1,68m poderia ser pouco pertinente a um conjunto de pessoas altas e muito pertinente a um de pessoas baixas. A lógica Fuzzy determina que essas relações de pertinência possuem valores (que podem ser fracionários) que variam entre 0 e 1. Dessa forma, um indivíduo com 1,68m poderia ter 0,15 de pertinência em um conjunto de pessoas altas e 0,75 de pertinência a um conjunto de pessoas baixas, ou seja, pertence a ambos os conjuntos, estando, porém menos “agregado” a um que ao outro.

Considerando o segundo exemplo, temperaturas podem variar suas gradações entre o quente e o frio, de acordo com seu grau de pertinência (Figura 129), ou seja, uma temperatura de 23°C pode ser considerada um pouco fria, mas não exatamente fria, pois há um pouco de pertinência dela ao conjunto de temperaturas quentes.

Figura 129 - Pertinência de temperaturas, segundo a lógica Fuzzy. Há um nível de transição entre os estados “Quente” e “Frio” que permite que uma temperatura seja considerada “um pouco quente” e “um pouco fria”.



Fonte: Ilustração do autor.

Entretanto, o diferencial da lógica Fuzzy está no uso das funções de pertinência, que são relacionadas a variáveis linguísticas. No nosso exemplo da Figura 129, os dois trapézios representados pelas linhas azul e vermelha correspondem às variáveis “Frio” e “Quente”, respectivamente.

Apesar de nossos exemplos tratarem de problemas que envolvem grandezas da mesma natureza entre si (altura e temperatura), devido ao uso das variáveis linguísticas, a lógica Fuzzy permite a criação de relações entre grandezas que podem não estar relacionadas entre si, como por exemplo, a quantidade de capital para investir em um negócio e o número de profissionais (valor monetário e mão de obra) disponíveis para nele trabalhar podem ser relacionados entre si para determinarmos a viabilidade de um negócio ou não .

Representação²⁴

Conjuntos numéricos são geralmente representados conforme a notação abaixo: $A = \{ 1, 2, 3, 4, 5 \}$. Já um conjunto Fuzzy apresenta a seguinte notação: $A = \{ (1 / 0,5), (2 / 0,75), (3 / 0,4), (4 / 0,9), (5 / 0,1) \}$, onde o primeiro algarismo entre parên-

²⁴ Os conceitos desenvolvidos nesta seção foram baseados no conteúdo presente no livro *Inteligência computacional* (Aguiar et al., 2007). Apesar de não serem diretamente copiados, toda lógica presente aqui foi apreendida desta publicação.

teses representa o elemento do conjunto e o segundo seu grau de pertinência a este conjunto. A notação para estes é o par $(x, \mu_a(x))$.

O objetivo de usar conjuntos e dados é realizar análises e comparações entre eles. As relações que se estabelecem entre os diversos conjuntos, como a união, a interseção, o reconhecimento de valores mínimos e máximos, são utilizadas para classificar as entradas dos sistemas e apresentar uma resposta ao usuário (a saída do sistema).

Portanto, compreender estas relações é imprescindível para que se criem normas de classificação de conjuntos. No caso específico desta pesquisa, para o reconhecimento das poses dos movimentos capturados. A seguir, serão descritas as relações utilizadas nesta pesquisa.

Norma-T²⁵

As normas são as regras lógicas que determinam as operações entre os conjuntos Fuzzy. Abordaremos aqui somente aquelas que são necessárias para a criação das relações entre os conjuntos Fuzzy utilizadas na pesquisa. Apesar destas não estarem explícitas durante o experimento aqui realizado, acreditamos ser importante sua apresentação nesta seção.

Mínimo

A Norma-T mínimo entre dois grupos é composta pelos elementos de cada grupo, acompanhadas da pertinência mínima entre eles. Consideremos dois conjuntos A e Z, representados abaixo:

$$A(X) = \{ (0 / 0), (1 / 0,2), (2 / 0,7), (4 / 1), (5 / 0,1) \}$$

$$Z(X) = \{ (0 / 0,5), (1 / 0,1), (2 / 0,7), (4 / 0,3), (5 / 0,8) \}$$

A Norma-T mínimo entre estes grupos será representada por:

²⁵ As referências às Normas-T foram retiradas do vídeo do canal Hackeando Tec (<https://www.youtube.com/watch?v=VbWH1QO9VLU>. Acesso em: 2 Jan. 2019).

$$T\text{-min}(A,Z) = \{ (0 / 0), (1 / 0,1), (2 / 0,7), (4 / 0,3), (5 / 0,1) \}$$

Conorma-T

Máximo

A Conorma-T máximo entre dois conjuntos é composta pelo elemento de cada grupo, acompanhada da pertinência máxima entre eles. Apresentamos os conjuntos A e Z novamente:

$$A(X) = \{ (0 / 0), (1 / 0,2), (2 / 0,7), (4 / 1), (5 / 0,1) \}$$

$$Z(X) = \{ (0 / 0,5), (1 / 0,1), (2 / 0,7), (4 / 0,3), (5 / 0,8) \}$$

A Norma-T mínimo entre eles será representada como:

$$T\text{-max}(A,Z) = \{ (0 / 0,5), (1 / 0,2), (2 / 0,7), (4 / 1), (5 / 0,8) \}$$

União de conjuntos Fuzzy

A União de conjuntos Fuzzy é realizada através da escolha das maiores pertinências dos elementos pertencentes a cada conjunto. Exemplificamos abaixo:

$$A(X) = \{ (0 / 0), (1 / 0,2), (2 / 0,7), (4 / 1), (5 / 0,1) \}$$

$$Z(X) = \{ (0 / 0,5), (1 / 0,1), (2 / 0,7), (4 / 0,3), (5 / 0,8) \}$$

$$A \cup Z = \{ (0 / 0,5), (1 / 0,2), (2 / 0,7), (4 / 1), (5 / 0,8) \}$$

Interseção de conjuntos Fuzzy

Já a interseção de conjuntos Fuzzy é dada pela seleção dos elementos que possuem a menor função de pertinência. Considerando os conjuntos A e Z, exemplificamos abaixo esta relação:

$$A(X) = \{ (0 / 0), (1 / 0,2), (2 / 0,7), (4 / 1), (5 / 0,1) \}$$

$$Z(X) = \{ (0 / 0,5), (1 / 0,1), (2 / 0,7), (4 / 0,3), (5 / 0,8) \}$$

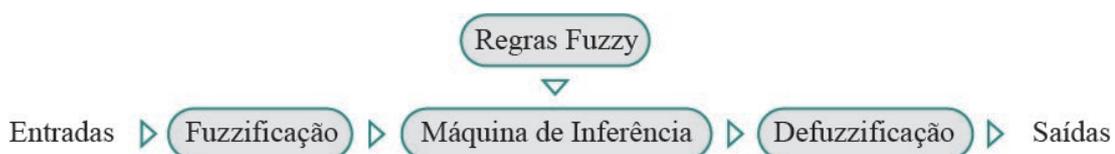
$$A \cap Z = \{ (0 / 0), (1 / 0,1), (2 / 0,7), (4 / 0,3), (5 / 0,1) \}$$

Essas operações de conjuntos Fuzzy mostram diferenças significativas com a lógica tradicional de conjuntos, permitindo que novas relações entre elementos sejam estabelecidas e permitem que os sistemas computacionais se aproximem da percepção humana.

Sistema Fuzzy

A implementação de um sistema Fuzzy pode ser dividido nas etapas representadas pela Figura 130 , descritas adiante.

Figura 130 - Fluxo de trabalho em Lógica Fuzzy.



Fonte: Ilustração do autor

Entradas

Os dados são as entradas de qualquer algoritmo. Em programação de computadores, alguns são comuns: valores inteiros, números reais, valores booleanos, palavras, caracteres, entre outros. Os dados podem ser inseridos nos sistemas através da interação humana, de sensores (como termostatos e sensores de proximidade) ou podem ser advindos de outros sistemas, como arquivos de capturas de movimentos, (que contêm dados de rotação e posição das juntas do corpo).

Regras Fuzzy

Esta é a etapa do processo em que as relações entre os dados de entrada são estabelecidas. O estabelecimento de regras entre grandezas diferentes é feito por meio de operações lógicas e linguísticas.

As operações lógicas são regras condicionais, começadas pela palavra “SE” (IF) e concluídas com a palavra “ENTÃO” (THEN), intercaladas por uma série de restrições.

Consideremos um indivíduo que precisa adquirir um fogão para sua casa. Nesta situação, ele precisa considerar o espaço que possui em sua cozinha e os recursos que dispõe para realizar este gasto.

De acordo com o espaço disponível, ele poderá adquirir um fogão de quatro, cinco ou seis bocas. Por outro lado, o número de bocas irá variar de acordo com os recursos que possui. Considerando estes aspectos, podemos definir algumas regras Fuzzy para aquisição do fogão:

1. **SE** o espaço disponível é pequeno **E** os recursos que possui são escassos, **ENTÃO** seu fogão será de quatro bocas;
2. **SE** o espaço disponível é médio **E** os recursos que possui são escassos, **ENTÃO** seu fogão será de quatro bocas;
3. **SE** o espaço disponível é grande **E** os recursos que possui são escassos, **ENTÃO** seu fogão será de quatro bocas;
4. **SE** o espaço disponível é pequeno **E** os recursos que possui são abundantes, **ENTÃO** seu fogão será de cinco bocas;
5. **SE** o espaço disponível é médio **E** os recursos que possui são abundantes, **ENTÃO** seu fogão será de cinco bocas;
6. **SE** o espaço disponível é grande **E** os recursos que possui são abundantes, **ENTÃO** seu fogão será de seis bocas.

Pelas condições apresentadas, pode-se notar que não há referências a valores, mas sim a relações entre as diversas variáveis apresentadas como entrada. A etapa seguinte a este processo será criar as relações entre o conjunto possível de valores de

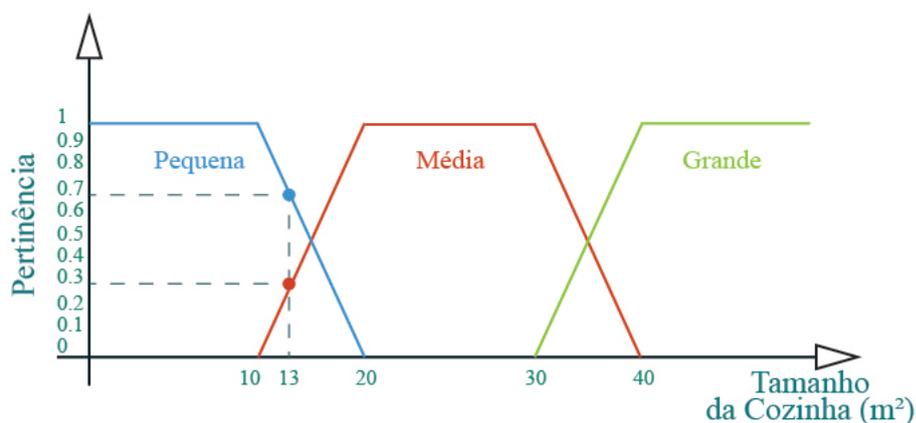
entrada e sua pertinência com cada um dos conjuntos em que se encontram. Este processo é denominado fuzzificação.

Fuzzificação

A fuzzificação é o processo através do qual o computador interpreta as regras linguísticas criadas pelo usuário, as relaciona a funções e as aplica aos valores de entrada, gerando assim, as relações de pertinência do sistema.

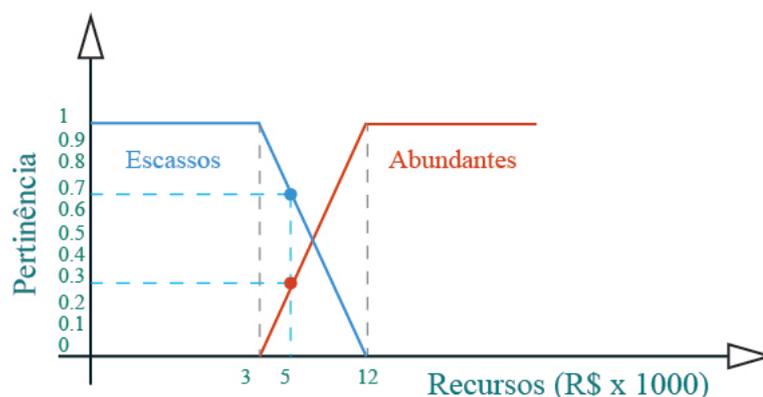
Retomando nosso exemplo da aquisição do fogão, a primeira etapa do processo de fuzzificação é a associação das variáveis de entrada e saída a valores e funções, visto que o computador interpreta valores para estabelecer relações entre os conjuntos de entrada. As figuras Figura 131 e Figura 132 apresentam essas associações.

Figura 131 - Relações entre as variáveis linguísticas e as dimensões de uma cozinha.



Fonte: Gráfico criado pelo autor.

Figura 132 - Relações entre as variáveis linguísticas e os recursos de um indivíduo.



Fonte: Gráfico criado pelo autor.

Nota-se que para diferentes tamanhos de cozinha serão associadas diferentes funções que, em algum momento, se interpolam. Em nosso exemplo, o valor de 13 m² possui 70% (0,7) de pertinência com o conjunto de cozinhas pequenas e 30% (0,3) de pertinência ao conjunto das cozinhas médias. O mesmo acontece com o conjunto de recursos pertencentes a um indivíduo (Figura 132). Ao inserirmos estes valores em um sistema de fuzzificação, ele retornará essas pertinências, definindo um valor difuso, ou seja, pertencente a dois conjuntos, mas em proporções diferentes.

Em nosso exemplo, vemos o uso de funções trapezoidais para definir as relações entre os valores presentes nos conjuntos e as pertinências destes a cada um de seus conjuntos. Algumas funções são utilizadas frequentemente para representar os conjuntos de números Fuzzy e suas pertinências, compondo o processo de fuzzificação. Entre elas, as funções triangulares, trapezoidais, singleton, sigmóides e gaussianas (Aguiar et al., 2007). Para os fins desta pesquisa descreveremos apenas os números Fuzzy triangulares e os trapezoidais, visto que são aqueles que utilizamos em nosso experimento. Entretanto, é importante saber que cada variável linguística pode ter uma função a ela associada.

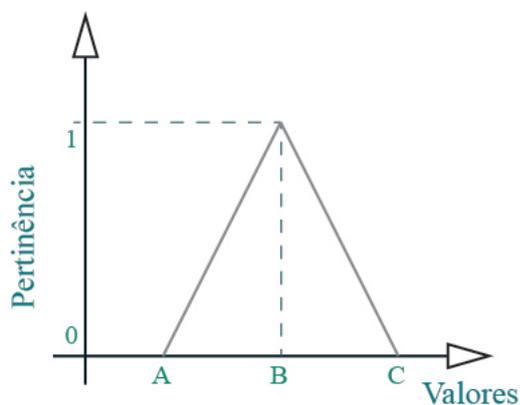
Números Fuzzy

Como vimos anteriormente, cada número Fuzzy é a associação de um valor e sua pertinência a um conjunto. Essa relação de pertinência é dada por funções. A seguir veremos os números Fuzzy triangulares e trapezoidais.

Números Fuzzy triangulares

Os números Fuzzy triangulares são aqueles definidos por funções como as apresentadas na Figura 133, em que os valores das pertinências crescem até o máximo (1.0) e depois decrescem até o mínimo (0.0).

Figura 133 - Gráfico com números Fuzzy triangulares.

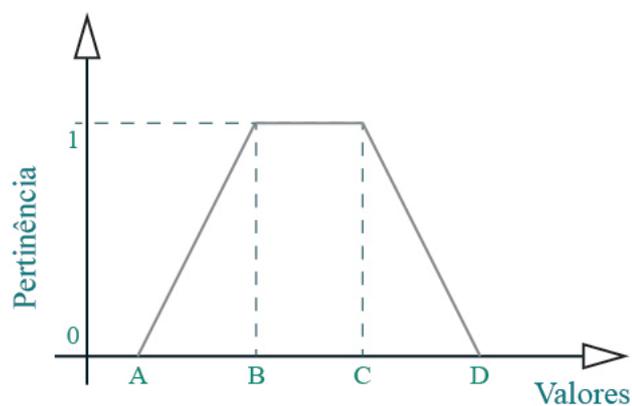


Fonte: Ilustração do autor.

Números Fuzzy trapezoidais

Números Fuzzy trapezoidais são aqueles que possuem duas regiões de transição de pertinência entre uma região de pertinência invariável. A Figura 134 apresenta um número Fuzzy trapezoidal. Entre A e B, encontra-se a zona de transição crescente de pertinência, enquanto entre B e C encontra-se a zona invariável. Por fim, entre C e D encontra-se a zona decrescente de pertinência. Essas zonas representam o quanto os valores de um conjunto pertencem ou não a ele.

Figura 134 - Gráfico com números Fuzzy trapezoidais.



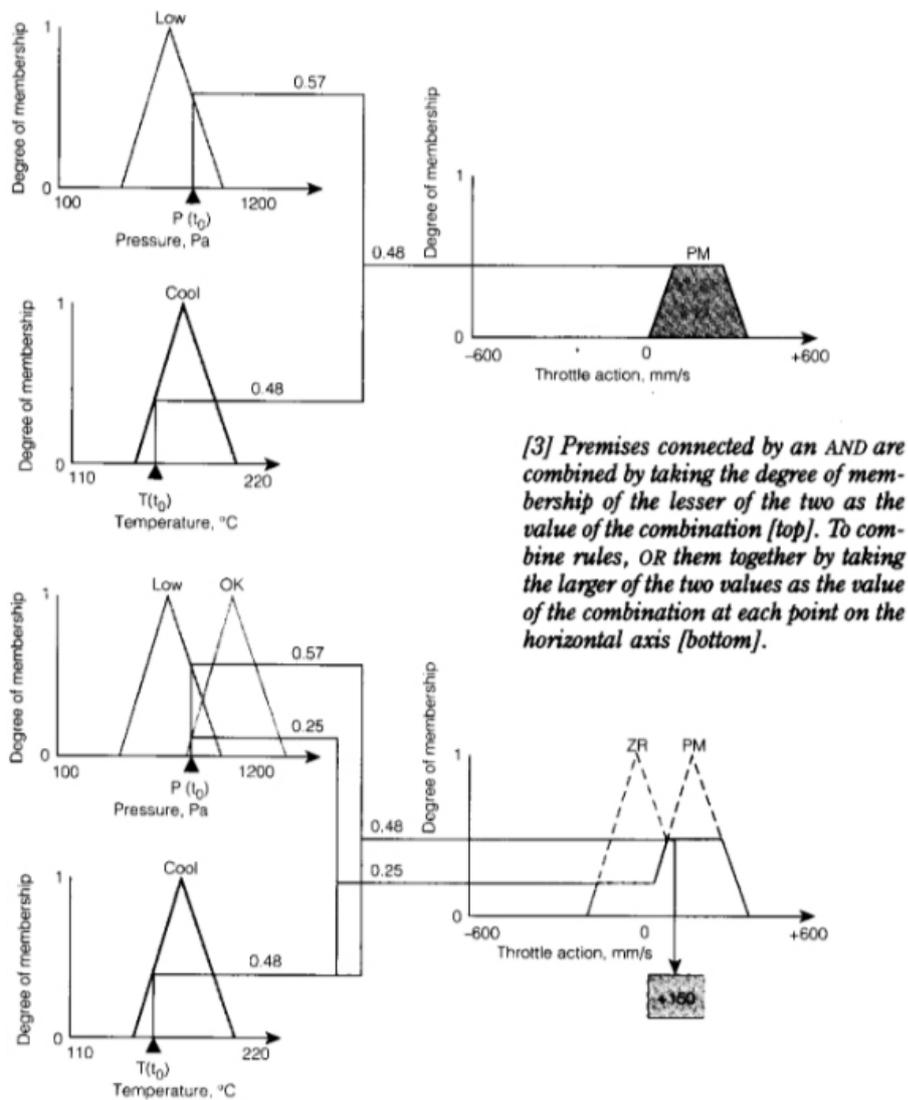
Fonte: Ilustração do autor.

Sistemas de inferência Fuzzy

Um sistema de Inferência Fuzzy é um algoritmo que realiza a tradução entre as funções de entrada, gerando funções de saída que representam as relações entre cada uma das variáveis de entrada (Figura 135).

Dois modelos são comumente utilizados e têm seus nomes oriundos dos pesquisadores que os desenvolveram: o modelo de Mamdani e o modelo de Takagi-Sugeno (Aguilar et al., 2007).

Figura 135 - O funcionamento do sistema de inferência com duas funções de entrada associadas às variáveis "Temperatura" e "Pressão".



Fonte: Cox (1992, p. 59).

Modelo de Mamdani

Segundo Aguiar et al. (2007, p. 47), “uma regra típica deste modelo é: se x é A e y é B (onde A e B são conjuntos Fuzzy), então z é C (onde C é conjunto Fuzzy)”.

Modelo de Takagi-Sugeno

Segundo Aguiar et al. (2007, p. 47), “uma regra típica desse modelo é: ‘se x é A e y é B , então $z = f(x, y)$ ’, onde A e B são conjuntos Fuzzy e f é uma função real de x e y ”.

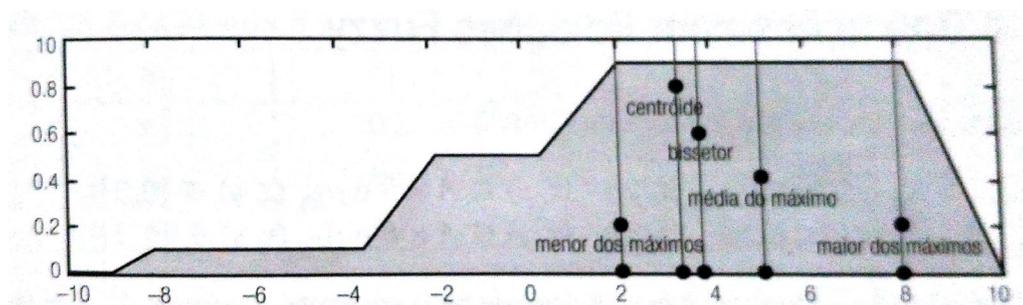
Defuzzificação

A defuzzificação é o processo através do qual os dados de entrada são transformados em saídas do sistema, de forma discursiva. Segundo Aguiar et al. (2007):

“[...] Conceitua-se defuzzificação como o processo de transformar um conjunto fuzzy em um elemento do universo de discurso (em geral, um número real), ou seja, o inverso da fuzzificação. Entre as técnicas utilizadas para tal processo, a mais usual é a do centroide” (Aguiar et al., 2007).

Vários modelos matemáticos podem ser empregados para determinar as saídas resultantes das interações entre os números Fuzzy. A Figura 136 mostra os resultados (saídas) advindos de vários destes modelos, de acordo com as funções de entrada. Como ressaltado anteriormente por Aguiar et al., o centroide é o mais utilizado em sistemas Fuzzy.

Figura 136 - Gráfico apresentando o processo de defuzzificação.



Fonte: Aguiar et al.(2007, p. 41).

Os resultados são valores reais que, associados às variáveis linguísticas, geram as saídas esperadas de um sistema Fuzzy. Assim, ao entrarmos com valores reais no sistema, ele nos retornará uma interpretação destes, realizando a tomada de decisão de forma automática.

Consideramos estes conceitos como a base para a compreensão dos sistemas que utilizam lógica Fuzzy. O aprofundamento desse conhecimento pode ser encontrado na bibliografia referenciada nesta pesquisa.

ANEXO II – PESQUISAS RELEVANTES

A seguir, apresentamos uma tabela com os artigos relevantes encontrados nesta pesquisa e suas contribuições para o campo do reconhecimento de atividades humanas. Apesar de muitos não constarem na redação de nossa pesquisa, consideramos relevante apresentá-los, pois representam os avanços encontrados nesta área do conhecimento. No total, perfazem 81 artigos que relacionam a detecção de poses com a captura de movimentos. Nela, destacamos os que eram representativos de técnicas específicas de reconhecimento do corpo e os que suscitaram questões que apontaram as lacunas do conhecimento, algumas das quais serão destacadas ao final deste capítulo para a aplicação prática. Excluimos aqueles que apenas se limitavam a fazer o reconhecimento e *tracking* do corpo, sem interpretar poses ou atividades humanas.

Tabela 9 - Artigos relacionando reconhecimento de movimentos e atividades humanas.

Artigo	Autor	Título	Contribuição para o reconhecimento de poses
1	(JOHANS SON, 1973)	Visual perception of biological motion and a model for its analysis.	O primeiro artigo a propor uma quantidade mínima de elementos necessários para o reconhecimento de ações humanas.
2	(CUTTIN G, 1978)	A program to generate synthetic walkers as dynamic point-light displays	O autor usa as propriedades exploradas por Johansson (1973) para criar um <i>software</i> em FORTRAN para a síntese de caminhadas naturais. Encontra os primeiros problemas relacionados à visão computacional com a oclusão de membros do corpo humano.
3	(MAXWE LL, 1983)	<i>Graphical marionette</i> (tese)	O autor apresenta uma hierarquia para os membros do corpo humano, com o quadril como centro de partida de todos os movimentos. Cria um sistema para gravar movimentos humanos e replicá-los em um “boneco de palito”.
4	(HOGG, 1984)	<i>Interpreting images of a known moving object</i> (tese)	Uma das primeiras teses na área da visão computacional a propor um sistema de reconhecimento de movimentos, replicação em uma estrutura volumétrica (cilindros representando partes do corpo). Mostra que a representação hierárquica do corpo pode auxiliar no desenvolvimento de uma solução para a oclusão de membros do corpo não vistos por uma câmera.
5	(LEE; CHEN, 1985)	Determination of 3D human body postures from a single view.	O autor apresenta um método para o reconhecimento de posturas do corpo humano, a partir de uma única vista (câmera). É importante notar que ele não se atém ao reconhecimento de poses, mas apresenta uma representação do esqueleto humano usada até hoje em sistemas de reconhecimento de junções humanas, como o Kinect. Fornece um método para determinar um conjunto de juntas em 3D do corpo humano. Usa as coordenadas de seis pontos característicos da cabeça e suas projeções. A partir disto, descobre uma matriz de transformações e, com esta, estabelece relações entre as quatorze juntas do

6	(YAMATO; OHYA; ISHII, 1992)	Recognizing human action in time-sequential images using hidden Markov model.	Os autores utilizam os modelos ocultos de Markov. Estes são "um número de estados em que a cada um é atribuída uma probabilidade de transição de um estado para outro estado. Com o tempo, as transições de estado acontecem estocasticamente. Assim como os modelos de Markov, os estados em qualquer momento dependem somente do estado que os precedem. Criam <i>templates</i> , a partir de ações específicas realizadas por indivíduos e utiliza partes dessas ações como treinamento para o sistema. O restante é para testá-lo.
7	(STURMAN; ZELTZER, 1994)	A survey of glove-based input	Os autores apresentam os principais desenvolvimentos em tecnologias para a captura de movimentos das mãos (<i>glove-based input</i>), com luvas eletrônicas. Mostram que os estudos buscavam novas formas de interação com o computador que fossem mais naturais que o <i>mouse</i> e o teclado.
8	(STURMAN, 1994)	A brief history of motion capture for computer character animation	Nesta página da internet, o autor apresenta o resultado de um minicurso ministrado na Siggraph de 1994 em que aponta os principais usos da captura de movimentos em aplicações comerciais até o ano de 1994. Destacam-se as referências ao personagem Waldo e à personagem Dozo.
9	(GAVRILAS; DAVIS, 1996)	3-D model-based tracking of humans in action: a multi-view approach	Os autores propõem um sistema para recompor o movimento humano em 3D, a partir de imagens capturadas por diversas câmeras. Para recompor o movimento, utilizam um método de procura por poses e síntese das mesmas. Usam quatro componentes para fazer a estimativa de uma pose: predição, síntese, análise da imagem e estimativa do estado (da pose). A base de vídeos usada para o experimento é denominada pelos autores de <i>Humans-In-Action</i> (HIA).
10	(DIBERNARDO; GONCALVES; PERONA, 1996)	Monocular tracking of the human arm in 3D: real-time implementation and experiments	Os autores desenvolvem um sistema monocular capaz de acompanhar (<i>tracking</i>) um braço humano no espaço 3D em tempo real.
11	(GLEICHER, 1997)	Motion editing with spacetime constraints	O autor aponta que, muitas vezes, a captura de movimentos precisa de um tratamento posterior por parte do animador e propõe um método para alterar uma pose do movimento capturado, de maneira que as poses adjacentes a ela sejam modificadas tanto em termos espaciais quanto no tempo (<i>timing</i>) da animação. Seu método apresenta uma eficiência limitada, mas aceita pelo autor.
12	(BREGLER, 1997)	Learning and recognizing human dynamics in video sequences	O autor deste artigo foi um dos primeiros pesquisadores a utilizar a probabilidade, a partir de diversos modelos conhecidos (modelos ocultos de Markov, Kalman recursivo, EM e outros) para prever e identificar movimentos em uma sequência de vídeos. Usa apenas vídeos de caminhadas para realizar sua predição. Consegue reconhecimentos positivos que variam de 86% a 93%.
13	(HARITAOGLU;	W4: Who? When? Where? What? A real	Os autores propõem um método, para o reconhecimento de ações humanas baseada em vídeo, que analisa a

	HARWOOD; DAVIS, 1998)	time system for detecting and tracking people	cabeça, as mãos, os pés e o torso de um indivíduo em imagens de baixa resolução e sem cores. Seu sistema consegue reconhecer e acompanhar o movimento de pessoas nesses vídeos.
14	(GLEICHER, 1998)	Retargetting motion to new characters	Este estudo, da empresa Autodesk, identifica o problema de utilizar uma captura de movimentos em diversos personagens diferentes (interações com outros personagens e objetos que não casam), e propõe um método para solucionar este problema que, apesar de ser um primeiro passo, não é totalmente eficiente.
15	(ISARD; BLAKE, 1998)	Condensation – conditional density propagation for visual tracking	Os autores criticam o uso dos filtros de Kalman para o trackeamento de curvas e propõem um novo algoritmo, o <i>condensation</i> , que utiliza amostragens fatoriais para aumentar a precisão do trackeamento de movimentos rápidos. A vantagem de seu algoritmo é rodar quase em tempo real.
16	(GAVRIL A, 1999)	The visual analysis of human movement: a survey	O pesquisador faz uma revisão das aplicações promissoras, focando naquelas que abordam o reconhecimento de indivíduos e suas atividades. Limita seu estudo às aplicações que reconhecem o corpo humano e o movimento das mãos. Propõe uma classificação dos domínios nos quais a visão computacional pode se atuar, inserindo a animação sob a realidade virtual. Classifica as pesquisas naquelas que realizam abordagens 2D (com ou sem a criação de modelos de formas) e as abordagens 3D (representando o corpo através de sólidos geométricos). Apresenta fluxo de trabalho para o reconhecimento de imagens em movimento.
17	(RITTSCHER; BLAKE, 1999)	Classification of human body motion	Os autores propõem um algoritmo de classificação de movimentos humanos usando um filtro de condensação, descrito por Blake e Isard (1998), para o aumento da eficiência do reconhecimento de movimentos humanos. Seu algoritmo possui uma taxa de erro de classificação de 10%, usando dois movimentos: um “salto puro” (sem movimentos laterais de braços ou pernas) e uma “meia estrela” (salto com braços e pernas abertos em meia flexão).
18	(SIDENBLADH; BLACK; FLEET, 2000)	Stochastic tracking of 3D human figures using 2D image motion	Os autores propõem um método probabilístico para o reconhecimento de partes do corpo humano, a partir da recomposição dos seus membros, representando-os como cilindros. Esta recomposição é feita a partir de uma única vista (uma única câmera). A partir da geração deste modelo humano em 3D, o algoritmo proposto por eles consegue reconhecer e acompanhar movimentos como o caminhar humano, estimando as posições dos membros ocultos.
19	(BOBICK; DAVIS, 2001)	The recognition of human movement using temporal templates.	O autor propõe um método para o reconhecimento de ações humanas através do movimento e não da construção de um modelo para, então, reconhecer o movimento que está sendo realizado.
20	(MOESLUND; GRANUM, 2001)	A survey of computer vision-based human motion capture.	Os autores apresentam diversos artigos que tratam do reconhecimento de ações humanas. Trata-se de uma pesquisa (neste caso, uma revisão bibliográfica) sobre o avanço da tecnologia de captura de movimentos, entre a década de 1980 e a metade dos anos 2000. Focam numa visão geral de taxonomia de sistemas de capturas de

			<p>movimentos e apresentam as quatro etapas de quebra do processo de captura de movimentos: a inicialização; o <i>tracking</i>; a estimativa de pose; e o reconhecimento. Estes conceitos serão utilizados para criação de sistemas mais complexos de captura de movimentos.</p>
21	(PERALES, 2001)	Human motion analysis and synthesis using computer vision and graphics techniques. State of art and applications	<p>Trata-se de uma revisão que apresenta os equipamentos utilizados para realizar capturas de movimentos, bem como a diferença entre <i>motion capture</i> e <i>performance animation</i>. Para o autor, o primeiro termo é usado para descrever as tecnologias usadas no processo, enquanto o segundo refere-se à <i>performance</i> usada pelos designers de animação para trazer personagens à vida. Aborda as diversas técnicas aplicadas, até então, para reconhecer ações humanas, incluindo aquelas já apresentadas por Bobick e Davis (2001) e Moeslund e Granum (2001).</p>
22	(MEREDITH; MADDOK; ROAD, 2001)	Motion capture file formats explained	<p>Artigo técnico que apresenta os formatos de arquivos usados em captura de movimentos e suas principais características. Foca nos seguintes formatos: ASC, ASF e AMC, ASK e SDL, BVA e BVH, BRD, C3D, CSM, DAT, GTR, HTR e TRC, MOT e SKL.</p>
23	(GAGVANI; SILVER, 2001)	Animating volumetric models	<p>Os autores apresentam um método para animar modelos volumétricos. Estes modelos permitem que as partes internas do modelo de um personagem sejam animados. Capturas de movimentos podem ser utilizadas pelos animadores para animação, através do esqueleto <i>biped</i> presente no 3ds Max.</p>
24	(PULLEN; BREGLER, 2002)	Motion capture assisted animation: texturing and synthesis	<p>Os autores propõem um método oposto ao que propomos em nossa pesquisa: aplica captura de movimentos a animações criadas por profissionais experientes. O processo é semi-automatizado, permitindo que os animadores possam escolher quais juntas do corpo serão animadas. Dividem as animações realizadas pelo animador e aquelas realizadas pela captura de movimentos em bandas de frequência para posteriormente casá-las através de um processo automatizado por algoritmo. Permitem que mais detalhes sejam adicionados ao personagem animado, possibilitando que este tenha mais “textura”.²⁶ Apresentam problemas com os esqueletos usados em seus experimentos.</p>
25	(JONATHAN HOWELL; BUXTON, 2002)	Active vision techniques for visually mediated interaction	<p>Os autores criam um sistema que reconhece gestos através do uso de uma base de dados de gestos constituída por quatro exemplos, coletados de três pessoas, tendo 48 sequências de imagens em cada.</p>
26	(ZHONGXIANG et al., 2002)	Multiple animated characters motion fusion	<p>Os autores propõem um método para fusão de personagens na mesma cena, interagindo entre si. Segundo eles, não existia (na época em que o artigo foi redigido) uma forma de fazer com que mais de um personagem interagisse em cena. Utilizam um algoritmo de percepção do ambiente virtual e de outros personagens, bem como um algoritmo de <i>DSM (Definite State Machine)</i> para definir um tipo de movimento para cada personagem na cena. As limitações de seu estudo</p>

²⁶ Em animação, “textura” refere-se aos detalhes nos movimentos presentes no personagem animado.

			são a falta de uma base bem estruturada de movimentos para comparar com os movimentos realizados por personagens e as restrições do ambiente virtual usado por eles.
27	(PLÄNKE RS; FUA, 2003)	Articulated soft objects for multiview shape and motion capture	Os autores propõem um método para reconstrução de formas e movimentos do corpo humano de maneira a aproximá-lo da realidade, incluindo aí os choques entre suas diversas partes. Em vez de simples cilindros, os membros do corpo humano são apresentados como <i>metaballs</i> (corpos esféricos que interagem entre si), apresentando maior fidelidade ao movimento capturado. Os movimentos do corpo e as formas são capturadas e processadas através de três câmeras.
28	(WANG; HU; TAN, 2003)	Recent developments in human motion analysis	Trata-se de uma revisão da literatura que aponta o que foi descoberto até o momento, bem como analisa outras revisões anteriormente realizadas. Destaca três áreas principais de exploração da <i>HMA (Human Motion Analysis)</i> : segurança visual, interface de usuário avançada e diagnóstico e identificação, baseadas em movimentos. Além disso, apresenta outras áreas de aplicação, como os games, animação de personagens, realidade virtual e conferências virtuais.
29	(SWARTZ et al., 2003)	Face-mask removal: movement and time associated with cutting of the loop straps.	Nesta pesquisa na área de esportes, os autores utilizam um sistema de captura de movimentos para determinar a maneira mais eficiente de remover capacetes de jogadores de futebol americano, sem comprometer suas colunas cervicais.
30	(OWINGS et al., 2003)	Influence of ball velocity, attention, and age on response time for a simulated catch	Os autores utilizam um equipamento de captura de movimentos para determinar o tempo de resposta de um grupo de jovens (8 a 16 anos), de acordo com a velocidade de lançamento de uma bola de baseball. Inclui, entre os parâmetros analisados, a atenção e a idade como fatores determinantes para o reflexo dos indivíduos.
31	(NAWOC ZENSKI et al., 2003)	Three-dimensional shoulder kinematics during a pressure relief technique and wheelchair transfer	Os autores deste estudo na área de biomecânica utilizam equipamentos de MoCap para avaliar a dinâmica de funcionamento dos ombros durante uma transferência de cadeira de rodas e alívio de pressão. Graças ao estudo chegam à conclusão que os profissionais que fazem este trabalho rotineiramente correm risco de lesão ou progressão de dores nos ombros.
32	(KEIONE N; KAURAN EN; VANHAR ANTA, 2003)	The relationship between anthropometric factors and body-balancing movements in postural balance	Os autores utilizam equipamento de MoCap para avaliar as diferenças na estabilidade do corpo, através das relações antropométricas entre as diversas partes (tronco, membros, etc.), quando os indivíduos estão com os olhos abertos e com os olhos fechados. Com seu estudo, esperam avaliar os motivos por meio dos quais os indivíduos caem e reduzir as lesões em seus corpos.
33	(AGUIAR, 2003)	<i>Character animation from a motion capture database</i> (dissertação)	O autor discute métodos para aplicar movimentos de uma base dados de captura de movimentos para aprimorar o movimento realista em personagens, a partir de alguns frames dados pelo animador.
34	(CERVERI ; PEDOTTI; FERRIGN)	Robust recovery of human motion from video using Kalman filters and virtual	Os autores propõem um método para reconstruir as posições de um indivíduo capturado por meio de filtros de Kalman, um método estatístico utilizado para calcular grandezas ao longo do tempo, independente do ruído.

	O, 2003)	humans	Eles utilizam um modelo biomecânico para associar aos membros do corpo reconhecidos. Seu método retorna resultados confiáveis e precisos.
35	(Crawford, 2003)	The digital turn: animation in the age of information technologies	Neste capítulo de livro, a pesquisadora aborda a mudança de paradigma das tecnologias analógicas de produção de animações para a produção digital, indicando as novas relações que se estabelecem entre o tradicional e o digital.
36	(GEROCH, 2004)	Motion capture for the rest of us	Este artigo, com um título bastante instigante, aborda a visão de um pesquisador/professor, apresentando um panorama do setor de captura de movimentos e ressaltando sua importância para o ensino de diversas áreas, em especial, de animação nas escolas.
37	(AMINIAN; NAJAFI, 2004)	Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications	Os autores apresentam uma novidade no desenvolvimento de sistemas de captura de movimentos: o uso de acelerômetros e giroscópios para reconstrução do corpo humano e seus movimentos. Como vantagens de seu sistema, apontam a possibilidade de uso fora de um ambiente controlado, como um ambiente externo a um laboratório de pesquisas. Indicam caminhos para o uso deste tipo de sistema, como o monitoramento de idosos, a <i>performance</i> de atletas e, resultados ortopédicos.
38	(BLANK, 2005)	Actions as space-time shapes.	Os autores consideram o reconhecimento de ações em vídeos como sendo derivados dos movimentos do torso e dos membros a ele associados. Propõem um método para converter as silhuetas dos indivíduos reconhecidos em formas tridimensionais que se movimentam.
39	(LU; LITTLE, 2006)	Simultaneous tracking and action recognition using the PCA-HOG descriptor	Os autores propõem um algoritmo para acompanhar (<i>tracking</i>) e reconhecer as ações de atletas, usando <i>templates</i> para a comparação de movimentos. Apesar dessa abordagem já ter sido utilizada anteriormente, os autores advogam que a sua se diferencia das demais por unir as duas etapas – a de acompanhar e a de reconhecer. Testam seu algoritmo com atletas de futebol e <i>hockey</i> .
40	(WEINLAND; RONFARD; BOYER, 2006)	Free viewpoint action recognition using motion history volumes	Os autores propõem um sistema de interpretação de gestos e ações humanas baseado no que eles denominam <i>Motion History Volumes (MHV)</i> . Seu sistema consegue interpretar diversos movimentos realizados por indivíduos independentemente da posição da câmera no ambiente (<i>free viewpoint</i>).
41	(LEE; COHEN, 2006)	A model-based approach for estimating human 3D poses in static images	Apesar de não utilizar imagens em movimento como base para sua proposta de algoritmo, este artigo foi incluído nesta lista por sua relevância. Seus autores criam um algoritmo capaz de estimar poses humanas em 3D usando cadeias de Markov Monte Carlo. Eles utilizam o conceito de mapas de proposta, um método eficiente de gerar poses candidatas em 3D, durante o processo de procura através da técnica de Monte Carlo.
42	(BLAKE; SHIFFRARR, 2007)	Perception of human motion	Nesta pesquisa, na área de psicologia, os autores mostram a evolução dos sistemas de visão computacional na área de percepção do movimento humano. Apontam estas mudanças de acordo com o avanço das tecnologias no tempo. Relatam algumas experiências importantes, como o reconhecimento de ações humanas a partir de dois

			frames apenas e ressaltam que a percepção humana é totalmente dependente do ponto de vista do observador.
43	(RASKAR et al., 2007)	Prakash: lighting aware motion capture using photosensing markers and multiplexed illuminators	Os pesquisadores criam um sistema composto por <i>leds</i> posicionados em pontos específicos do corpo humano para realizar a captura de movimentos através da sua observação. Através da análise de parâmetros visíveis, eles conseguem reconstruir o movimento realizados. Seu sistema apresenta algumas características que os diferencia daqueles existentes no mercado: explora um grupo de transmissores para computar a localização de um ponto no espaço 3D; computa a orientação dos rótulos de cada um dos sensores; computa a reflectância intrínseca de cada ponto; imperceptível dependendo das condições do ambiente; e cria novos algoritmos para apoiar aplicações em CG na área de quantidades geométricas e fotométricas.
44	(YAMAZOE et al., 2007)	A body-mounted camera system for head-pose estimation and user-view image synthesis	Os autores criam um <i>hardware</i> para capturar as experiências dos usuários através de duas câmeras: uma que captura os movimentos da cabeça de um usuário e outra que captura a imagem frontal do indivíduo. Eles advogam que o sistema proposto sobrepõe as desvantagens normalmente encontradas nas câmeras montadas em cabeças, como a facilidade de vesti-las e a reprodução de uma imagem similar àquela experienciada pelo usuário.
45	(FATHI; MORI, 2008)	Action recognition by learning Mid-level Motion Features.	Os autores recortam uma sequência de cinco frames e estabilizam a figura no centro do quadro. Características de nível médio são extraídas de fluxos ópticos de baixo nível, sendo que cada característica foca em um pequeno cuboide dentro do volume centrado na figura. Distinguem entre pares de classes. Criam classes de ações. Conseguem 99% de precisão em bases de imagens selecionadas.
46	(LIN; SUN, 2008)	Activity recognition using a combination of category components and local models for video surveillance	Os autores utilizam o reconhecimento de atividades humanas na área de segurança em vídeo. Para realizar esta tarefa, propõem um algoritmo que trabalha em duas etapas: na primeira, “representa uma atividade pela combinação de componentes categorizados”; na segunda, reconhece uma atividade em apenas um <i>frame</i> da sequência de vídeo para posteriormente classificar os demais.
47	(LAPTEV et al., 2008)	Learning realistic human actions from movies	Os autores reconhecem uma limitação no reconhecimento de ações humanas naturais devido à ausência de vídeos realistas e comentados. Para solucionar isto, eles criam uma base de vídeos, comentada para auxiliarem no reconhecimento de ações. Em um primeiro momento, dividem seu problema em duas etapas: na primeira selecionam uma série de filmes e seus roteiros e associam as sequências em que cada cena acontece com ações encontradas nos <i>scripts</i> , gerando anotações para cada uma delas. A partir desta base de dados de vídeos, propõem um algoritmo que reconhece ações do “mundo real”.
48	(LI; ZHANG; LIU, 2008)	Expandable data-driven graphical modeling of human actions based on salient postures	Os autores apresentam um método de reconhecimento de ações humanas através de silhuetas. Criam seus métodos baseados em gráficos de ações cujos movimentos são reconhecidos através da aplicação de métodos de

-
- | | | | |
|----|-------------------------------------|---|--|
| | | | <p>probabilidade das poses seguintes. Como a cada ação corresponde um número de poses, estas podem ser utilizadas em várias ações. Aventam serem os primeiros a criarem um método em que é possível adicionar novas ações de treinamento ao sistema, criando escalabilidade. O fazem usando um método de <i>Gaussian Mixture Models (GMMs)</i>.</p> |
| 49 | (MURPHY - CHUTORIAN; TRIVEDI, 2009) | Head pose estimation in computer vision: a survey. | Os pesquisadores levantam um histórico da estimativa de poses de cabeça na área da visão computacional, indicando a importância destes estudos para a área da visão computacional e apontando as vantagens e as desvantagens de cada abordagem realizada pelas pesquisas desenvolvidas até aquele momento. |
| 50 | (JI; LIU, 2010) | Advances in view-invariant human motion analysis: a review. | Trata-se de uma revisão apresentando o que havia sido descoberto em termos de reconhecimento de ações humanas e enfatizando a detecção humana, a representação e estimativa de poses com vista invariante, e a compreensão de comportamento. Lacunas na área do conhecimento: 1) Estimativa de poses 3D de sequências monoculares; 2) A inferência de poses 3D usando métodos baseados em exemplos é difícil, devido principalmente à perspectiva; 3) HMMs requerem muito processamento. CRFs devem substituí-los; e 4) A compreensão de comportamento é complexa. Um mesmo comportamento pode ter diferentes significados, dependendo da cena e da tarefa sendo realizada em determinado contexto. Pesquisas em padrões de comportamentos são direções futuras. |
| 51 | (POPPE, 2010) | A survey on vision-based human action recognition | Nesta revisão, o autor apresenta os avanços realizados no reconhecimento de ações humanas. Define que a reconstrução de poses e movimentos humanos, a partir de sequências de imagens, pode ser considerado um problema de regressão, enquanto o reconhecimento de movimentos é um problema de classificação. O <i>labeling</i> (anotação) é um procedimento necessário para o reconhecimento de ações humanas. Apresenta as bases de vídeos públicas usadas até então para avaliação dos sistemas de reconhecimento de ações. Classifica os diversos meios utilizados para reconhecer ações e aponta que, apesar do processo de <i>labeling</i> poder ser realizado e conferido de maneira manual, ele se tornará proibitivo, conforme as bases de vídeos forem aumentando. |
| 52 | (SIGAL; BALAN; BLACK, 2010) | HumanEva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion | Os autores apresentam uma grande base de dados, constituída por vários movimentos (em torno de 40.000 frames) capturados. Esta base apresenta por vantagem, ter <i>ground-truth</i> , uma forma de captura utilizada por sistemas profissionais de Mocap (no caso deles, um Vicon), em que a representação do chão vem definida no arquivo. Além disso, eles descrevem um algoritmo-base para o <i>tracking</i> (acompanhamento) de indivíduos usando volumes 3D por meio de um método Bayesiano. Sua limitação é quanto ao uso de equipamentos limitados a laboratórios de Mocap. |
| 53 | (SCHINDLER et al., 2010) | Automatic detection and tracking of pedestrians from a moving stereo rig | Os autores criam um sistema de <i>tracking</i> (acompanhamento) de pedestres, por meio de um modelo probabilístico que se baseia em imagens armazenadas previamente e reconstrução 3D de movimentos. Destacam a necessidade da existência de pesquisas na |
-

			<p>área de sistemas de segurança automotivos e navegação autônoma. Criam uma montagem (<i>rig</i>) de câmeras que consegue detectar pedestres no espaço tridimensional. Conseguem um resultado bom (73% de precisão) de reconhecimento de pedestres.</p>
54	(LI; ZHANG; LIU, 2010)	Action recognition based on a bag of 3D points	<p>Os autores criam um algoritmo capaz de reconhecer com precisão (acima de 90%) os movimentos realizados por um indivíduo, utilizando como <i>hardware</i> uma câmera de profundidade (<i>depth maps</i>). Seu método utiliza o <i>bag of point</i>, um conjunto de informações armazenadas na forma de gráfico de ações, para identificar os movimentos realizados.</p>
55	(NAIEL; ABDELW AHAB; EL-SABAN, 2011)	Multi-view human action recognition system employing 2DPCA.	<p>Os autores baseiam-se na tecnologia de criação de <i>Motion History Images</i> e <i>Motion Energy Images</i>, criada por Bobick (2001) e aplicam uma análise de componente principal bidimensional. Alegam que esse método "reduz a complexidade computacional por um fator de pelo menos 66, alcançando a maior precisão de reconhecimento por câmera.</p>
56	(KWON; LEE, 2011)	An animation bilateral filter for slow-in and slow-out effects	<p>Os autores reconhecem as curvas de animação de um personagem animado e, a partir delas, alteram suas acelerações e desacelerações, fazendo com que o princípio de animação conhecido como <i>slow-in</i> e <i>slow-out</i> seja aplicado via algoritmo ao personagem.</p>
57	(AGGAR WAL, 2011)	Human activity analysis: a review.	<p>Apresenta uma revisão bibliográfica contendo as principais descobertas na área de Reconhecimento de ações humanas. Divide o reconhecimento de poses entre abordagens de camadas simples e abordagens hierárquicas. Os primeiros reconhecem as atividades humanas através da análise de vídeos. Os segundos reconhecem atividades humanas através da análise de atividades simples (subeventos) que, combinadas, podem gerar a compreensão de interações de alto nível, como o lutar (exemplo do autor) que pode ser reconhecido como a combinação de muitas interações de chutes e socos. A ideia das <i>hierarchical approaches</i> é fazer com que o sistema seja capaz de reconhecer e analisar micro-movimentos chamados por Aggarwal de <i>atomic actions</i>. O sistema, então, será capaz de reconhecer ações através da sequencialização de múltiplas poses.</p>
58	(ALIVERT I et al., 2011)	Functional evaluation and rehabilitation engineering	<p>Os autores apresentam um breve histórico da criação de captura de movimentos, suas áreas de aplicação e a importância da engenharia de reabilitação, aquela em que os movimentos dos indivíduos com algum tipo de limitação são analisados para que uma solução seja projetada.</p>
59	(KUEHNE et al., 2011)	HMDB: a large video database for human motion recognition	<p>Os autores criam um banco de dados extenso, dividido em 51 categorias de ações nas quais encontram-se em torno de 7000 vídeos manualmente identificados. Esses vídeos são disponibilizados para que pesquisadores possam usá-los para comparar a eficiência de seus sistemas, utilizando uma mesma base de dados.</p>
60	(BORENS TEIN, 2012)	Making things see.	<p>O autor apresenta diversos projetos usando o Kinect, o Processing e uma biblioteca aberta para o reconhecimento de poses e partes do corpo humano. Sua abordagem simplificada auxilia aqueles que estão entrando no universo Maker.</p>

61	(SPEHR; WINKELBACH; WAHL, 2012)	Hierarchical pose estimation for human gait analysis	Os autores pesquisadores criam uma técnica de estimativa de poses para análise de caminhadas humanas, usando métodos probabilísticos através do compartilhamento de modelos gráficos hierárquicos. Estes são construídos através de campos de Markov aleatórios associados a <i>nodes</i> , que são convertidos para ?
62	(DUTTA, 2012)	Evaluation of the Kinect™ sensor for 3-D kinematic measurement in the workplace	O autor compara a precisão do equipamento Kinect™ com o sistema de captura de movimentos Vicon, para determinar sua viabilidade de uso como sistema de captura de movimentos portátil em avaliações ergonômicas no ambiente de trabalho. Conclui que, com alguns aprimoramentos, ele pode ser usado com tal finalidade.
63	(BARNACHON, 2012)	Human actions recognition from streamed motion capture	Os autores propõem um algoritmo para reconhecimento de ações humanas utilizando histogramas de poses de ação e equipamento de Mocap. O método proposto por eles possui precisão graças ao que descrevem como “algoritmo de programação dinâmica e computação de histogramas incremental”.
64	(GUTIÉRREZ-DÁVILA et al., 2013)	Effect of uncertainty on the reaction response in fencing.	Os autores utilizam duas plataformas de força e um sistema de captura de movimentos com onze marcadores para investigar como a incerteza da posição do alvo, durante o ataque na esgrima, pode afetar os parâmetros temporais de reação de resposta na coordenação do ataque. Apesar dos estudos demonstrarem pouca alteração dos reflexos dos atletas durante o movimento realizado, eles concluem que o tempo de resposta é influenciado pela incerteza da posição e que, em atletas com pouca experiência, esses estudos podem apresentar alterações mais consideráveis.
65	(IOSIFIDIS; TEFAS; PITAS, 2013)	Multi-view human action recognition: A survey	Este <i>paper</i> é uma revisão bibliográfica que aborda as diferentes técnicas usadas para o reconhecimento de ações humanas por meio de múltiplas câmeras (<i>multi-view</i>). Eles classificam os métodos entre aqueles que usam números fixos ou arbitrários de câmeras, revisam os processos já vistos anteriormente e apresentam as bases de dados de movimentos utilizadas para pesquisas na área.
66	(ZHOU et al., 2013)	Kernel-based sparse representation for gesture recognition	Os autores identificam o problema de processamento e reconhecimento de gestos humanos através das abordagens anteriormente exploradas por outras pesquisas – como o <i>Dynamic Time Warping</i> e o <i>Longest Common Subsequence</i> – e propõem o uso de “Classificações Baseadas em Representações Esparsas” (<i>Sparse Representation-based Classification - SRC</i>) para realizar o reconhecimento. Propõem uma técnica de extração de características chamada de <i>Covariance Matrix Singular Value Decomposition for Kernelization (CovSVDK)</i> , que é robusta para o reconhecimento de gestos humanos e eficiente em termos computacionais. Usam bases de vídeos da Georgia-Tech (universidade) e da linguagem australiana de sinais para determinar a eficiência do método proposto, comparado com outros, o que se confirma com dezenove de vinte bancos de dados, sendo reconhecidos de forma satisfatória.
67	(WANG et al., 2013)	The effects of load carriage and muscle	Neste exemplo de aplicação militar de um sistema de captura de movimentos, os pesquisadores utilizam um

		fatigue on lower-extremity joint mechanics	sistema Vicon para avaliar os efeitos do transporte de cargas sobre a fadiga muscular de soldados, durante o caminhar. Este estudo é relevante para a área, pois esse tipo de lesão acomete muitos militares. Em sua conclusão, mostram que, ao caminhar com cargas excessivas, os militares prejudicam a mecânica de seus membros inferiores.
68	(SANTOFIMIA; MARTINEZ-DEL-RINCON; NEBEL, 2014)	Episodic reasoning for vision-based human action recognition	As autoras apresentam uma nova visão acerca da visão computacional. Elas acreditam que a simples análise de poses do corpo não é suficiente para a interpretação do movimento. Para elas, "essa tarefa deveria ser apoiada pelo conhecimento profundo da natureza de atuação humana e sua estreita conexão com as razões e as motivações que a explicam. A combinação deste conhecimento e o conhecimento sobre o funcionamento do mundo é essencial para o reconhecimento e a compreensão das ações humanas sem comprometer os erros comuns de falta de sentido". Para tanto, elas usam um sistema de reconhecimento de atividades humanas que utiliza duas estratégias: a <i>computer vision</i> e o senso comum: "O sistema proposto executa um reconhecimento primário de ações, o qual é baseado somente nas capacidades de análise de imagens. O primeiro estágio calcula as cinco ações mais prováveis, de acordo com as posturas dos corpos dos atores. Estas ações são utilizadas como entradas ao sistema de racionalização de senso comum. Em um segundo estágio, o modelo de racionalização de senso comum executa algumas tarefas de racionalização sobre as ações sugeridas ao sistema de <i>computer vision</i> . Estas operações são suportadas sobre um modelo formal de conhecimento, também proposto e formalizado aqui. Essencialmente, três abstrações conceituais são propostas neste modelo de maneira a replicar o processo mental seguido por humanos em um sistema computacional. As noções de ação, crença e expectativa articulam os mecanismos de racionalização implementados, de acordo com a teoria davidsoniana das ações."
69	(BARNACHON et al., 2014)	Ongoing human action recognition with motion capture	Os autores utilizam o método de <i>ongoing human action recognition</i> para classificar movimentos advindos de capturas de movimentos. Utilizam métodos baseados em histogramas para o reconhecimento dos dados. Uma das vantagens do método apresentado é o reconhecimento do movimento no momento em que ele começa a ser realizado e não apenas realizando um pós-processamento. Testam seu método em diversas bases de dados de vídeos, apresentando resultados satisfatórios em muitas delas.
70	(PAQUETE, 2014)	Using Mocap as reference	Neste artigo em seu <i>blog online</i> , Pete Paquete – animador – mostra como o uso de captura de movimentos pode ser utilizada como base para animações na área de <i>games</i> .
71	(PAIVA, 2014)	<i>The use of motion capture in non-realistic animation</i> (dissertação)	Nesta dissertação, o autor cria uma animação a partir de captura de movimentos. Entretanto, ele foca na <i>performance</i> do ator para obter movimentos cartunizados.
72	(ZIAEEFARD; BERGEVI)	Semantic human activity recognition: A literature review	Neste artigo, os autores fazem uma revisão da literatura apresentando a importância do reconhecimento semântico de atividades humanas. Este tipo de

	N, 2015)		abordagem auxilia os algoritmos a diferenciarem movimentos semelhantes entre si. Eles classificam os métodos utilizados por pesquisadores da área, dividindo-os em “poses”, <i>poselets</i> , <i>scene</i> , <i>HOI</i> , <i>attribute</i> e <i>linguistic descriptions</i> . Mostram como os principais autores tratam da descrição semântica de ações, subdividindo-as em poses e as “interpretando” para gerar um texto que define a ação sendo realizada.
73	(CHEN; KOSKELA, 2015)	Skeleton-based action recognition with extreme learning machines	Neste artigo, os pesquisadores identificam os principais problemas que ocorrem no reconhecimento de ações feitas a partir de câmeras RGBD (câmeras de profundidade) e propõem três métodos de modelos para o pós-processamento das imagens obtidas. Como resultado, concluem que o modelo de probabilidade de sequências é o mais apropriado para o reconhecimento de ações.
74	(LO PRESTI; LA CASCIA, 2016)	3D skeleton-based human action classification: a survey	Revisão da literatura acerca da classificação de ações humanas baseadas em esqueletos 3D. Seu foco é, segundo os autores, o “pré-processamento de dados, as referências públicas disponíveis e as medidas de precisão comumente usadas”. Focam, de maneira incisiva, nas pesquisas sobre mapas de profundidade e métodos para estimar o esqueleto.
75	(KOLLER; FORSTER; NEY, 2015)	Continuous sign language recognition: towards large vocabulary statistical recognition systems handling multiple signers.	Neste artigo, os autores desenvolvem um método para o reconhecimento eficiente da linguagem de sinais. Eles utilizam um algoritmo que é capaz de acompanhar (<i>trackear</i>) as mãos e o rosto do usuário, através de modelos ocultos de Markov. Embora não seja detalhado no reconhecimento de poses, usam métodos específicos para <i>trackear</i> as mãos e fazer o reconhecimento de características da face.
76	(JI et al., 2016)	Robust multi-body feature tracker: a segmentation-free approach	Neste artigo, os autores propõem um método inovador de <i>tracker</i> de livre-segmentação de multicorpo. Sua abordagem perpassa um passo visto em outras abordagens de <i>trackeamento</i> de poses, o passo de atribuição de movimento. Para isto, eles usam o subespaço restrito derivados diretamente das restrições epipolares de múltiplos movimentos (investigar isto): "Como resultado, nosso algoritmo não necessita de um conhecimento anterior do número de movimentos. Além disso, isto nos permite formular o <i>tracking</i> como um problema de otimização cujos subproblemas têm todos soluções de forma fechada."
77	(HUANG et al., 2016)	Visual storytelling	Neste artigo, os autores criaram um imenso banco de dados de imagens classificados por humanos que descrevem o que é apresentado em cada imagem. Também fazem a descrição da história por trás de uma sequência de imagens e validam um sistema como base para a classificação de imagens. A base pode ser um início para a interpretação de imagens por computadores. Esta pesquisa, se interessa, sobretudo, pelo o método usado por eles.
78	https://www.mpg.de/10722890/3d-animation-incredibly-easy?print=	3D animation – incredibly easy new software makes it possible to generate animated three-dimensional figures of animals from short	Neste artigo de divulgação da Max-Planck-Gesellschaft, é apresentado o resultado de uma pesquisa na área de ciência da computação: um <i>software</i> que converte vídeos curtos de animais em figuras tridimensionais animadas.

	yes	vídeos	
79	(LINN, 2016)	Teaching computers to describe images as people would	Neste artigo, o autor apresenta um trabalho no qual pesquisadores da Microsoft estão criando um algoritmo para “ensinar” computadores a descreverem imagens, através de um sistema que “conta as histórias” dos vídeos. Para isto, eles propõem um modelo que utiliza a visão computacional e o processamento de linguagem natural.
80	(HERATH; HARANDI ; PORIKLI, 2017)	Going deeper into action recognition: a survey	Neste trabalho de revisão da bibliografia os autores refazem o caminho traçado por outros pesquisadores que revisaram o reconhecimento de ações, como Moeslund e Granum (2006) e Wang et al. (2016), utilizando fontes como o dicionário Oxford para reforçar algumas definições, como a de “ações”. Ele identifica cada taxonomia sobre “movimentos” e “análise de ações” encontrada nos <i>papers</i> analisados. Apresenta, ainda, os principais algoritmos utilizados para reconhecimento de vídeos e suas limitações. Por fim, mostra que a área de reconhecimento de ações humanas é bastante promissora nos próximos anos e apresenta os possíveis caminhos para o desenvolvimento de algoritmos baseados nas pesquisas encontradas de 2014 a 2016.
81	(MEHTA et al., 2017)	VNect: real-time 3D human pose estimation with a single RGB camera ACM reference format: VNect: real-time 3D human pose estimation with a single RGB camera	Os pesquisadores do Instituto Max Planck criaram um sistema de estimativa de poses, usando utilizando redes neurais convolucionais. Seu método permite que apenas uma câmera RGB seja utilizada para reconstruir a pose de um indivíduo em 3D. As etapas utilizadas por seu sistema são: reconhecimento e recorte do indivíduo; regressão através de redes neurais convolucionais; aplicação de um filtro temporal para estimar os pontos do corpo em 2D; estimativa da pose 3D; ajuste do esqueleto reconhecido ao esqueleto 3D. Para que seu sistema seja eficiente, eles o treinaram através de conjuntos de mais de 10.000 poses humanas 3D com notações. Não reconhecem movimentos e reconstroem poses.

Fonte: Pesquisa realizada no portal de Periódicos da CAPES, entre 6 de março de 2014 e 20 de dezembro de 2018.

Fontes bibliográficas dos artigos encontrados:

ABDELWAHAB, M. M.; EL-SABAN, M. NAIEL, M. A.; Multi-view human action recognition system employing 2DPCA. 2011 **IEEE Workshop on Applications of Computer Vision**, 2011, p. 270–275, 2011.

AGGARWAL, J. K. Human activity analysis: a review. **ACM Comput. Surv. Article**, v. 43, n. 16, p. 43, 2011.

AGRAWAL, A.; BATRA, D.; DEVLIN, J; GALLEY, M.; GIRSHICK, R.; HE, X.; HUANG, T.-H.; FERRARO, F.; KOHLI, P.; MISRA, I; MITCHELL, M.; MOSTAFAZADEH, N.; PARIKH, D.; VANDERWENDE, L ZITNICK; C. L. Visual storytelling. **Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies**, p. 7, 2016. Internet. Disponível em: <<http://research.microsoft.com/apps/pubs/default.aspx?id=264715>>.

AGUIAR, E. De. Character animation from a motion capture database. 2003. **Dissertação (Master Thesis in Computer Science)**. Computer Science Department. University of Saarland.

ALIVERTI, A.; FRIGO, C.; ANDREONI, G.; BARONI, G.; BONARINI, a.; CERVERI, P.; CRIVELLINI, M.; DELLACÁ, R.; FERRIGNO, G.; GALLI, M.; PEDROCCHI, a.; RODANO, R.; SANTAMBROGIO, G. C.; TOGNOLA, G.; PEDOTTI, A. Functional evaluation and rehabilitation engineering. **IEEE Pulse**, v. 2, n. 3, p. 24–34, 2011.

AMINIAN, K.; NAJAFI, B. Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications. **Computer Animation and Virtual Worlds**, v. 15, n. 2, p. 79–94, 2004.

ANTONIO, R.; GUTIÉRREZ-DÁVILA, M.; JAVIER ROJAS, F.; NAVARRO, E. Effect of uncertainty on the reaction response in fencing. **Research Quarterly for Exercise and Sport**, v. 84, n. 1, p. 16–23, 2013. Internet. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23611004>>.

ARMSTRONG, C. W.; KLEINER, D. M.; NORKUS, S. A.; SWARTZ, E. E. Face-mask removal: movement and time associated with cutting of the loop straps. **Journal of Athletic Training** (National Athletic Trainers' Association), v. 38, n. 2, p. 120–125, 2003. Internet. Disponível em: <<http://search.ebscohost.com.proxy-ub.rug.nl/login.aspx?direct=true&db=cin20&AN=2004016853&site=ehost-live&scope=site>>.

BALAN, A. O.; BLACK, M. J.; SIGAL, L. HumanEva: synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion. **International Journal of Computer Vision**, 2010.

BARNACHON, M. Human actions recognition from streamed motion capture. **International Conference on Pattern Recognition**, n. Icp, p. 3807–3810, 2012.

BARNACHON, M.; BOUAKAZ, S.; BOUFAMA, B.; GUILLOU, E. Ongoing human action recognition with motion capture. **Pattern Recognition**, v. 47, n. 1, p. 238–247, 2014. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.patcog.2013.06.020>>.

BEKAERT, P.; BARNWELL, J.; BATRA, D.; BRANZOI, V.; BRUNS, E.; DEDECKER, B.; DIETZ, P.; HASHIMOTO, Y.; INAMI, M.; MOORE, D.; NAYAR, S.; NIL, H.; NOLAND, M.; RASKAR, R.; SUMMET, J.; WESTHUES, J.; ZHAO, Y.; Prakash: lighting aware motion capture using photosensing markers and multiplexed illuminators. **ACM Transactions on Graphics**, v. 26, n. 3, p. 36, 2007. Internet. Disponível em: <<http://dl.acm.org/citation.cfm?id=1276377.1276422%5Cnhttp://portal.acm.org/citation.cfm?doid=1276377.1276422>>.

BERGEVIN, R.; ZIAEEFARD, M.; Semantic human activity recognition: a literature review. **Pattern Recognition**, v. 48, n. 8, p. 2329–2345, 2015. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.patcog.2015.03.006>>.

BLAKE, R.; SHIFFRAR, M. Perception of human motion. **Annual Review of Psychology**, v. 58, p. 47–73, 2007. Internet. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16903802%5Cnhttp://www.annualreviews.org/doi/abs/10.1146/annurev.psych.57.102904.190152>>.

%5Cn<http://www.ncbi.nlm.nih.gov/pubmed/16903802>>.

BLACK, M. J.; FLEET, D. J.; SIDENBLADH, H. Stochastic tracking of 3D human figures using 2D image motion. **Springer Berlin Heidelberg**, 2000. v. 1842

BLAKE, A.; RITTSCHER, J. Classification of human body motion. **Proceedings of the Seventh IEEE International Conference on Computer Vision**, p. 634–639 vol.1, 1999. Internet. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=791284>>.

BLANK, M. G. L. S. E. I. M. B. R. Actions as space-time shapes. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 29, n. 12, p. 2247–2253, 2005.

BOBICK, A. F.; DAVIS, J. W. The recognition of human movement using temporal templates. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 23, n. 3, p. 257–267, 2001.

BORENSTEIN, G. **Making things see: 3D Vision with Kinect, Processing, Arduino and MakerBot**. Sebastopol: Makermedia, 2012.

BORSTAD, J. D.; CLOBES, S. M.; GORE, S. L.; LUDEWIG, P. M.; NAWOCZENSKI, D. A.; NEU, J. L.; OLSEN, J. E. Three-dimensional shoulder kinematics during a pressure relief technique and wheelchair transfer. **Arch Phys Med Rehabil**, v. 84, n. 9, p. 1293–1300, 2003. Internet. Disponível em: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=13680564><http://www.ncbi.nlm.nih.gov/pubmed/13680564>>.

BOYER, E.; RONFARD, R.; WEINLAND, D. Free viewpoint action recognition using motion history volumes. **Computer Vision and Image Understanding**, v. 104, n. 2–3 SPEC. ISS., p. 249–257, 2006.

BREGLER, C. Learning and recognizing human dynamics in video sequences. **Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, n. June, p. 568–574, 1997.

BREGLER, C.; PULLEN, K. Motion capture assisted animation: texturing and synthesis. **ACM Transactions on Graphics**, v. 21, p. 501–508, 2002.

BUXTON, H. JONATHAN HOWELL, A.; Active vision techniques for visually mediated interaction. **Image and Vision Computing**, v. 20, n. 12, p. 861–871, 2002.

CARRILLO, R. E.; BARNER, K. E.; KIAMILEV, F.; LIU, K.; ZHOU, Y. **Kernel-based sparse** representation for gesture recognition. **Pattern Recognition**, v. 46, n. 12, p. 3208–3222, 2013. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.patcog.2013.06.007>>.

CASAS, D.; MEHTA, D.; RHODIN, H.; SEIDEL, H.; SHAFIEI, M.; SOTNYCHENKO, O.; SRIDHAR, S.; THEOBALT, C.; XU, W. VNect: Real-time 3D human pose estimation with a single RGB camera ACM reference format: VNect: Real-time 3D human pose estimation with a single RGB camera. **Tog**, v. 36, n. 14, p. 1–13, 2017. Internet. Disponível em: <<https://doi.org/http://dx.doi.org/10.1145/3072959.3073596>><http://dx.doi.org/10.1145/3072959.3073596>>.

CERVERI, P.; FERRIGNO, G.; PEDOTTI, A.; Robust recovery of human motion from video using Kalman filters and virtual humans. **Human Movement Science**, v. 22, n. 3, p. 377–404, 2003.

CHEN, X.; KOSKELA, M. Skeleton-based action recognition with extreme learning machines. **Neurocomputing**, v. 149, n. Part A, p. 387–396, 2015. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.neucom.2013.10.046>>.

CUTTING, J. E. A program to generate synthetic walkers as dynamic point-light displays. **Behavior Research Methods & Instrumentation**, v. 10, n. 1, p. 91–94, 1978. Internet. Disponível em: <<http://www.springerlink.com/index/10.3758/BF03205105>>.

DAVIS, L. S.; HARITAOGLU, I.; HARWOOD, D.; W4: Who? When? Where? What? a real time system for detecting and tracking people. **European Conference on Computer Vision**, p. 877–892, 1998. Internet. Disponível em: <http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=698720>.

DUGAN, E. L.; FRAME, J.; LEIB, D.; OZIMEK, E.; WANG, H.; The effects of load carriage and muscle fatigue on lower-extremity joint mechanics. **Research Quarterly for Exercise and Sport**, v. 84, n. November 2013, p. 305–12, 2013. Internet. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24261009>>.

DI BERNARDO, E.; GONCALVES, L.; PERONA, P. Monocular tracking of the human arm in 3D: real-time implementation and experiments. In: **Proceedings - International Conference on Pattern Recognition**, Anais...1996.

DUTTA, T. Evaluation of the Kinect(tm) sensor for 3-D kinematic measurement in the workplace. **Applied Ergonomics**, v. 43, n. 4, p. 645–649, 2012. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.apergo.2011.09.011>>.

FATHI, A.; MORI, G. Action recognition. **Learning Mid-level Motion Features**. Sfu.Ca, 2008. Internet. Disponível em: <http://www.sfu.ca/~alirezaf/publication/cvpr2008/Fathi_Mori_CVPR_2008.pdf>.

FENG, L.; YUETING, Z.; YUNHE, P.; ZHONGXIANG, L. Multiple animated characters motion fusion. **Journal of Visualization and Computer Animation**, v. 13, n. 5, p. 275–285, 2002.

FUA, P.; PLÄNKERS, R. Articulated soft objects for multiview shape and motion capture. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 25, n. 9, p. 1182–1187, 2003.

GAGVANI, N.; SILVER, D. Animating volumetric models. **Graphical Models**, v. 63, n. 6, p. 443–458, 2001. Internet. Disponível em: <<http://portal.acm.org/citation.cfm?id=766307#>>.

GARROTE, E.; JHUANG, H.; KUEHNE, H.; POGGIO, T.; SERRE, T. HMDB: A large video database for human motion recognition. In: 2011 **IEEE International Conference on Computer Vision**, Barcelona. Anais... Barcelona: IEEE, 2011.

GAVRILA, D. M. The visual analysis of human movement: a survey. **Computer Vision and Image Understanding**, v. 73, n. 1, p. 82–98, 1999.

GAVRILA, D. M.; DAVIS, L. S. 3-D model-based tracking of humans in action: a multi-view approach. In: **Proceedings - 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, 1, California. Anais... California: IEEE Computer Society Press, 1996. Internet. Disponível em: <<http://www.umiacs.umd.edu/users/fgavrila,lsdg/>>.

GEROCH, M. S. Motion capture for the rest of us. **Journal of Comp. Sciences in Colleges**, v. 19, n. 3, p. 157–164, 2004.

GLEICHER, M. Motion editing with spacetime constraints. **Proceedings of the 1997 Symposium on Interactive 3D graphics SI3D 97**, v. pp, p. 139-, 1997. Internet. Disponível em: <<http://portal.acm.org/citation.cfm?doid=253284.253321>>.

GLEICHER, M. Retargetting motion to new characters. **Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '98**, p. 33–42, 1998. Internet. Disponível em: <<http://portal.acm.org/citation.cfm?doid=280814.280820>>.

GRABINER, M. D; LAMPE, E. M.; LANCIANESE, S. L.; OWINGS, T. M.. Influence of ball velocity, attention, and age on response time for a simulated catch. **Medicine and Science in Sports and Exercise**, v. 35, n. 8, p. 1397–1405, 2003.

GRANUM, E.; MOESLUND, T. B. A survey of computer vision-based human motion capture. **Computer Vision and Image Understanding**, v. 81, n. 3, p. 231–268, 2001. Internet. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S107731420090897X>>.

HARANDI, M.; HERATH, S.; PORIKLI, F. Going deeper into action recognition: a survey. **Image and Vision Computing**, v. 60, p. 4–21, 2017.

HOGG, D. C. Interpreting images of a known moving object. Tese de Doutorado. University of Sussex, 1984.

HOSAKA, K.; UTSUMI, A.; YACHIDA, M.; YAMAZOE, H. A body-mounted camera system for head-pose estimation and user-view image synthesis. **Image and Vision Computing**, v. 25, n. 12, p. 1848–1855, 2007. Internet. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0262885606002885>>.

HU, W.; TAN, T.; WANG, L. Recent developments in human motion analysis. **Pattern Recognition**, v. 36, n. 3, p. 585–601, 2003.

ISHII, K.; OHYA, J.; YAMATO, J. Recognizing human action in time-sequential images using hidden Markov model. **Computer Vision and Pattern Recognition**, p. 379–385, 1992. Internet. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=223161%5Cnhttp://ieeexplore.ieee.org/ielx2/418/5817/00223161.pdf?tp=&arnumber=223161&isnumber=5817%5Cnhttp://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=223161&contentType=Conference+Publication>.

IOSIFIDIS, A.; PITAS, I. TEFAS, A.; Multi-view human action recognition: a survey. **Proceedings - 2013 9th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IHH-MSP 2013**, p. 522–525, 2013.

ISARD, M.; BLAKE, A. Condensation: conditional density propagation for visual tracking. **International Journal of Computer Vision**, v. 29, n. 1, p. 5–28, 1998. Internet. Disponível em: <<http://link.springer.com/article/10.1023/A:1008078328650%5Cnhttp://link.springer.com/10.1023/A:1008078328650>>.

JI, P.; LI, H.; SALZMANN, M.; ZHONG, Y. Robust multi-body feature tracker: a segmentation-free **Approach**, 2016. Internet. Disponível em: <<http://arxiv.org/abs/1603.00110>>.

JI, X. J. X.; LIU, H. L. H. Advances in view-invariant human motion analysis: a review. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, v. 40, n. 1, p. 13–24, 2010.

JOHANSSON, G. Visual perception of biological motion and a model for its analysis. **Perception & Psychophysics**, v. 14, n. 2, p. 201–211, 1973. Internet. Disponível em: <<papers3://publication/uuid/9F0D433F-719B-496B-AAA3-28542970C1E2>>.

KAURANEN, K.; KEIONEN, P.; VANHARANTA, H. The relationship between anthropometric factors and body-balancing movements in postural balance. **Archives of Physical Medicine and Rehabilitation**, v. 84, p. 17–22, 2003.

KOLLER, O.; FORSTER, J.; NEY, H. Continuous sign language recognition: towards large vocabulary statistical recognition systems handling multiple signers. **Computer Vision and Image Understanding**, v. 141, p. 108–125, 2015. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.cviu.2015.09.013>>.

KWON, J. Y.; LEE, I. K. An animation bilateral filter for slow-in and slow-out effects. **Graphical Models**, v. 73, n. 5, p. 141–150, 2011. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.gmod.2011.02.002>>.

LA CASCIA, M.; LO PRESTI, L. 3D skeleton-based human action classification: a survey. **Pattern Recognition**, v. 53, p. 130–147, 2016. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.patcog.2015.11.019>>.

LAPTEV, I.; MARSZALEK, M.; ROZENFELD, B.; SCHMID, C. Learning realistic human actions from movies. In: **26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR, Anais...2008**.

LEE, H.-J.; CHEN, Z. Determination of 3D human body postures from a single view. **Computer Vision, Graphics, and Image Processing**, v. 30, p. 148–168, 1985.

LEE, M. W.; COHEN, I. A model-based approach for estimating human 3D poses in static images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 28, n. 6, p. 905–916, 2006.

LEIBE, B.; SCHINDLER, K.; ESS, A.; VAN GOOL, L. Automatic detection and tracking of pedestrians from a moving stereo rig. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 65, n. 6, p. 523–537, 2010. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.sprsjprs.2010.06.006>>.

LI, W.; ZHANG, Z.; LIU, Z. Expandable data-driven graphical modeling of human actions based on salient postures. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 18, n. 11, p. 1499–1510, 2008.

LI, W.; ZHANG, Z.; LIU, Z. Action recognition based on a bag of 3D points.pdf. **Computer Vision and Pattern Recognition Workshops**, 2010 IEEE Computer Society Conference on, p. 9–14, 2010.

LIN, W.; SUN, M. Activity recognition using a combination of category components and local models for video surveillance. **IEEE transactions on circuits and and Systems for video technology**, v. 18, n. 8, p. 1128–1139, 2008. Internet. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4543872>.

LINN, A. Teaching computers to describe images as people would. Internet. Disponível em <https://blogs.microsoft.com/ai/teaching-computers-to-describe-images-as-people-would/>.

LITTLE, J. J.; LU, W. L. Simultaneous tracking and action recognition using the PCA-HOG descriptor. **Third Canadian Conference on Computer and Robot Vision**, 2006, v. 2006, 2006.

MADDOCK, S.; MEREDITH, M.; ROAD, P. Motion capture file formats explained. **Motion Capture Stuff**, p. 1–36, 2001.

MAXWELL, D. R. Graphical marionette. Proc. **ACM SIGGRAPH/SIGART Workshop**, p. 172–179, 1983.

MURPHY-CHUTORIAN, E.; TRIVEDI, M. M. Head pose estimation in computer vision: a survey. **IEEE transactions on pattern analysis and machine intelligence**, v. 31, n. 4, p. 607–26, 2009. Internet. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19229078>>.

PAIVA, J. P. T. M. The use of motion capture in non-realistic animation. Dissertação de Mestrado. Escola das Artes da Universidade Católica Portuguesa. Mestrado em Som e Imagem. September, 2014.

PERALES, F. Human motion analysis and synthesis using computer vision and graphics techniques: state of art and applications. **Proc. World Multiconf. on systemics, cybernetics and ...**, 2001. Internet. Disponível em: <http://pdf.aminer.org/000/351/715/analysis_and_synthesis_of_human_motion_from_external_measurements.pdf>.

POPPE, R. A survey on vision-based human action recognition. **Image and Vision Computing**, v. 28, n. 6, p. 976–990, 2010. Internet. Disponível em: <<http://dx.doi.org/10.1016/j.imavis.2009.11.014>>.

SANTOFIMIA, M. J.; MARTINEZ-DEL-RINCON, J.; NEBEL, J. C. Episodic reasoning for vision-based human action recognition. **The Scientific World Journal**, Vol 2014, 2014.

SPEHR, J.; WAHL, F. M.; WINKELBACH, S. Hierarchical pose estimation for human gait analysis. **Computer Methods and Programs in Biomedicine**, v. 106, n. 2, p. 104–113, 2012. Internet. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0169260711003270>>.

STURMAN, D. J. A brief history of motion capture for computer character animation. **Internet**. Disponível em: <https://www.siggraph.org/education/materials/HyperGraph/animation/character_animation/motion_capture/history1.htm>. Acesso em: 25 maio. 2016.

STURMAN, D. J.; ZELTZER, D. A Survey of glove-based input. **IEEE Computer Graphics and Applications**, v. 14, n. 1, p. 30–39, 1994.