

Gustavo Eggert Boehs

**ANIMAÇÃO DE PERSONAGENS ANIMAIS NÃO HUMANOS
POR CAPTURA DE MOVIMENTOS**

Tese submetida ao Programa de Pós-
Graduação em Design da Universidade
Federal de Santa Catarina para a obtenção
do Grau de Doutor em Design.
Orientador: Prof. Dr. Milton Luiz Horn
Vieira

Florianópolis
2018

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Boehs, Gustavo Eggert
ANIMAÇÃO DE PERSONAGENS ANIMAIS NÃO HUMANOS POR
CAPTURA DE MOVIMENTOS / Gustavo Eggert Boehs ;
orientador, Milton Luiz Horn Vieira, 2018.
154 p.

Tese (doutorado) - Universidade Federal de Santa
Catarina, Centro de Comunicação e Expressão,
Programa de Pós-Graduação em Design, Florianópolis,
2018.

Inclui referências.

1. Design. 2. Design. 3. Mídias Digitais. 4.
Captura de Movimentos. 5. Aprendizado de Máquin. I.
Vieira, Milton Luiz Horn. II. Universidade Federal
de Santa Catarina. Programa de Pós-Graduação em
Design. III. Título.

ANIMAÇÃO DE PERSONAGENS ANIMAIS NÃO-HUMANOS POR CAPTURA DE MOVIMENTOS

Esta Dissertação/Tese foi julgada adequada para obtenção do Título de
Doutorado e aprovada em sua forma final pelo Programa de Pós-
Graduação em Design da Universidade Federal de Santa Catarina

Florianópolis, 09 de março de 2018.

Prof. Luiz Fernando Figueiredo Gonçalves, Dr. Eng.
Coordenador do Curso

Banca Examinadora:

Prof. Milton Luiz Horn Vieira, Dr. Eng.
Orientador
Universidade Federal de Santa Catarina

Prof. Ricardo Triska, Dr.
Universidade Federal de Santa Catarina

Prof. Milton José Cinelli, Dr.
Universidade do Estado de Santa Catarina

Prof. Mauro Roisenberg, Dr.
Universidade Federal de Santa Catarina

Rodrigo Marques Almeida da Silva, Dr.

Este trabalho é dedicado à minha
esposa e à minha família.

AGRADECIMENTOS

Agradeço ao professor Milton Luiz Horn Vieira, por me receber como orientando, por disponibilizar de seu tempo, conhecimentos e infraestrutura de laboratório para a execução desta pesquisa. Aos participantes da banca, doutor Rodrigo Marques, professor Mauro Roisenberg, professor Ricardo Triska e professor Milton Cinelli, por compartilharem de seu tempo e conhecimento em prol do enriquecimento deste trabalho e de minha formação. Aos docentes do programa de Pós-Graduação em Design da Universidade Federal de Santa Catarina pelos conhecimentos compartilhados. A equipe da secretaria do Pós-Design, que sempre esteve disponível para me auxiliar de maneira célere e cordial. Ao programa de Pós-Graduação Ciências da Computação por permitir minha participação como aluno em algumas de suas disciplinas. Ao departamento de Expressão Gráfica que, compreensivamente, cedeu a este servidor o tempo necessário para a condução desta pesquisa. Aos contribuintes, que através de órgãos governamentais como o MEC, a CAPES e o CNPq financiaram esta pesquisa, ainda que de modo involuntário. Ao laboratório de captura de movimentos da Universidade Carnegie Mellon e seus pesquisadores que disponibilizam gratuitamente uma extensa base de dados construída ao longo de vários anos. Ao graduando Victor Honório pelo comprometimento e colaboração continuada em determinadas etapas desta pesquisa. A todos os participantes voluntários que se submeteram a participar da etapa de coleta de dados.

“O limite do aprendizado humano não é conhecer uma única coisa, mas combinar a fração do conhecimento de uma ou algumas coisas com o conhecimento geral de muitas coisas”

(John Stewart Mill, em discurso inaugural à Universidade de St. Andrews em 1 de fevereiro de 1867)

RESUMO

Personagens são a principal interface entre ser humano e conteúdo em determinadas mídias digitais. Um dos atrativos destes agentes é seu caráter lúdico e a possibilidade de representação de formas fantásticas. A captura de movimentos é uma ferramenta utilizada para animar personagens digitais que pode resultar em ganho de velocidade e redução de custos de produção, no entanto, seu uso tem sido restrito a certos nichos. A animação de personagens animais, não humanos, é um destes nichos onde a produção de animação por captura de movimentos tem se mostrado tecnicamente complexa e está em estágio embrionário. Além da dificuldade existente para tratar os dados obtidos por captura de movimentos, devido ao fato de serem altamente correlacionados à estrutura anatômica capturada, há a dificuldade adicional de que a animação de personagens animais é costumeiramente feita com diferenciados graus de antropomorfismo. A presente pesquisa teve como objetivo propor e desenvolver um *framework* computacional para a utilização dos dados obtidos pela captura de movimentos de atores, não treinados, na animação de personagens animais, não humanoides, capaz de controlar personagens com graus de liberdade inferiores, iguais ou superiores aos do ser humano de quem os movimentos são capturados. Para tanto, o movimento animal em animação foi estudado e tipificado de modo a delimitar os resultados pretendidos para o sistema. Técnicas de aprendizado de máquina foram utilizadas para decompor o movimento humano em dois componentes, um expressivo e outro não-expressivo, e então transpor estes dois componentes ao personagem animal. O método proposto foi implementado na forma de um protótipo funcional capaz de processar dados provenientes de equipamentos de captura de movimentos em tempo real. As principais contribuições do método apresentado são a capacidade de transpor o movimento humano a personagens animais não humanos a partir da captura de movimento de atores leigos e capacidade de fazer tal transposição a qualquer personagem animal, independentemente do número de membros de sua anatomia. Por meio da aplicação de um teste piloto foi possível verificar que ações transpostas a personagens animais pelo *framework* desenvolvido foram corretamente identificadas por seres humanos. Conclui-se que a utilização de tal sistema permite expansão da adoção da tecnologia de captura de movimentos a um contexto mais amplo de produção de conteúdo digital animado. Isto, por sua vez, permite trazer a redução de custos e a flexibilidade proporcionados pela tecnologia de captura de movimentos, a estes novos contextos de aplicação.

Palavras-chave: Design. Mídias Digitais. Captura de Movimentos.
Aprendizado de Máquina. Personagens Digitais.

ABSTRACT

Characters are the main interface between humans and content in certain digital media. One of the main reasons for the use of such agents is their ludic aspect and the possibility of representing fantastic figures. Motion capture is a tool used in the animation of digital characters which may result in significant reduction of a production's duration and costs, although the use of such tool has been restricted to certain niches. The animation of non-human animal character is one of such niches where the production through motion capture has still been not practical. One of the difficulties is the fact that motion capture data is tightly correlated to the anatomy of the captured performer. Another difficulty is the fact that, in animation, animal characters are represented with varying degrees of anthropomorphism. The objective of this research was to propose and develop a computational framework for the utilization of motion data captured of untrained human actors in the control of non-humanoid animal characters of any number of limbs. To define and delimit the behavior of this framework we studied and typified the animal movement as represented in animated media. Then, machine learning was used to decompose the human movement into an expressive and a non-expressive component, which are in turn mapped onto the animal character. This method was implemented as a fully functioning prototype capable of running in real-time. The principal contributions of the proposed method are its capacity to transpose human movement onto non-human animal characters in a scenario where the human actor is untrained and where the anatomy of the animal character may be composed of any number of limbs. In a preliminary test, human subjects accurately recognized animations generated by the proposed system. In conclusion the proposed system allows the adoption of motion capture technology in a wider context of the production of animated digital content. This in turn enables this wider field to enjoy the flexibility and time saving benefits of motion capture technology.

Keywords: *Design. Digital media. Motion capture. Machine Learning. Digital characters.*

LISTA DE FIGURAS

Figura 1 – Captura de movimentos inercial com cavalo	32
Figura 2 – Captura de movimentos com humano em personagem animal	33
Figura 3 – O personagem Burro apresenta características animais e humanas.....	33
Figura 4 – Abordagens: cinética (à esquerda) onde as rotações das juntas do ator são transferidas diretamente ao personagem com proporções modificadas; e dinâmica (à direita) onde a distribuição do peso sobre o corpo do personagem com as proporções modificadas é levada em conta para deformar a pose original.....	37
Figura 5 – Geometrias humanoides controlando modelos não humanoides a partir da correlação entre poses chave.....	39
Figura 6 – Interface gráfica dos <i>softwares</i> Matlab e <i>Neural Network Toolbox</i> realizando o treinamento de uma das redes neurais utilizadas na implementação do protótipo.....	48
Figura 7 – Interface gráfica da plataforma de desenvolvimento <i>Fabric Engine</i> utilizada para o pré e pós processamento dos dados de captura de movimentos.....	48
Figura 8 – Interface gráfica da plataforma Maya, utilizada para visualizar os personagens finais animados.....	49
Figura 9 – Escala de antropomorfismo de movimento em personagens animais	55
Figura 10 – Esquema gráfico do funcionamento do <i>framework</i> proposto e seus diferentes módulos.....	60
Figura 11 – Topologia de uma rede com duas camadas escondidas.....	62
Figura 12 – Etapa de treinamento do Módulo de Dissociação.....	65
Figura 13 – Etapa de inferência em tempo real do Módulo de Dissociação	65
Figura 14 – Representação visual dos parâmetros dinâmicos.....	67
Figura 15 – Topologia da rede neural profunda para regressão da pose do ator a partir de parâmetros dinâmicos do movimento. Dimensionalidade das camadas entre parênteses.....	69
Figura 16 – Dois exemplos de dissociação do movimento. Os esqueletos à esquerda representam o movimento original, os esqueletos ao centro representam o CED e os esqueletos à direita representam o CPE. A imagem superior mostra um sujeito limpando uma janela, a imagem inferior mostra um sujeito lançando uma bola de basquete.....	71
Figura 17 – Treinamento da classificação de tipo	74
Figura 18 – Treinamento da regressão do índice de progressão.....	75

Figura 19 – Inferência em tempo real de tipo e progressão	75
Figura 20 – Topologia da rede neural profunda para a classificação de tipos locomotivos. Dimensionalidade das camadas entre parênteses....	77
Figura 21 – Representação esquemática da transformação temporal baseada em quadros chave.	78
Figura 22 – À esquerda o esqueleto do sujeito capturado e os vetores de velocidade do quadril. À direita o resultado da camada da rede neural de de-estilização do movimento.	79
Figura 23 – Topologia da rede profunda para a regressão da fase do movimento. Dimensionalidade das camadas entre parênteses.	79
Figura 24 – Comparação entre valores de fase filtrados e não filtrados na classificação de um ciclo de caminhada para um valor $u = 0,9$	81
Figura 25 – Diferentes poses classificadas como ‘parado em pé’.....	83
Figura 26 – A partir do topo esquerdo no sentido horário: caminhando, correndo, passos à direita, caminhando de costas.	83
Figura 27 – Regressão de fase para uma caminhada neutra. ‘Controle’ representa os valores classificados por um especialista, ‘NN’ os valores calculados com o método Nearest Neighbor e ‘Este Protótipo’ o valor calculado pelo método desenvolvido nesta pesquisa.	85
Figura 28 – Regressão de progressão de uma caminhada com tropeços (estilizada). ‘Controle’ representa os valores classificados por um especialista, ‘NN’ os valores calculados com o método Nearest Neighbor e ‘Este Protótipo’ o valor calculado pelo método desenvolvido nesta pesquisa.....	85
Figura 29 – Ator move em diferentes direções dentro de um ambiente.	86
Figura 30 – Ator anda, corre e pula.....	87
Figura 31 – Etapa de treinamento do mapeamento de poses expressivas	90
Figura 32 – Etapa de inferência do mapeamento de poses expressivas	90
Figura 33 – Poses que exemplificam diferentes amplitudes das articulações do braço.....	93
Figura 34 – Poses que exemplificam diferentes movimentos coordenados entre dois membros do corpo.....	93
Figura 35 – Poses que exemplificam contatos entre superfícies do mesmo corpo.....	93
Figura 36 – Coletânea de poses transpostas ao personagem do tubarão. De cima para baixo, na primeira linha, exemplos de poses em que os braços se movem na mesma direção, na segunda linha, exemplos em que os braços se movem em direções diferentes.	94
Figura 37 – Casos onde o ator realiza contato entre duas superfícies do seu corpo mal representados no mapeamento de poses.	95

Figura 38 – Mapeamento de movimentos expressivos em dois graus diferenciados de antropomorfismo.....	96
Figura 39 – Mapeamento de poses expressiva comparado, com e sem o uso de um <i>autoencoder</i> , em caso conhecido a priori.....	98
Figura 40 – Mapeamento de poses expressiva comparado, com e sem o uso de um <i>autoencoder</i> , em caso desconhecido a priori.	99
Figura 41 – À esquerda o CED, ao centro o CPE e a direita o resultado da soma dos dois componentes.	101
Figura 42 – Interpolação entre poses do CED e do CPE aplicada de modo global, com peso $\alpha = 0,5$ em todas as juntas do corpo.	101
Figura 43 – Interpolação em função do índice de expressividade local	102
Figura 44 – Diagrama exemplificando as regiões anatômicas dos modelos utilizados quanto ao tipo de fusão escolhido.	103
Figura 45 – Composição com vários movimentos obtidos por captura transpostos a modelos animais pelo protótipo desenvolvido.....	108
Figura 46 – Interface do formulário utilizado na coleta de dados.	114

LISTA DE TABELAS

Tabela 1 – Comparação entre métodos de transposição de captura de movimentos.....	42
Tabela 2 –Comparação com trabalhos anteriores cuja a aplicação é o controle de personagens não bípedes por meio da classificação de movimentos.....	73
Tabela 3 – Comparação do erro entre o método proposto e o método de Seol <i>et al.</i> (2013) na classificação da locomoção interpretada de modo neutro. Onde ‘Este’ representa o presente trabalho.....	84
Tabela 4 – Comparação do erro entre o método proposto e o método de Seol <i>et al.</i> (2013) na classificação da locomoção interpretada de modo estilizado. Onde ‘Este’ representa o presente trabalho.....	84
Tabela 5 – Matriz de classificação de tipos para todos os quadros de todos os cliques testados. Onde ‘Este’ representa o presente trabalho....	85
Tabela 6 – Resultados da coleta e análise de dados do teste piloto. Movimentos em destaque foram positivamente reconhecidos pela população entrevistada tanto no grupo de teste como no de controle, tendo, nestes casos a hipótese nula sido rejeitada com um grau de confiança de 95%.....	115

LISTA DE ABREVIATURAS E SIGLAS

ACM – *Association for Computing Machinery*
AM – Aprendizado de Máquina
CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CPE – Componente Potencialmente Expressivo
CED – Componente de Equilíbrio e Deslocamento
CPU – *Central Processing Unit*
GPLVM – *Gaussian Process Latent Volume Method*
NN – *Nearest Neighbor linear interpolation*
PCA – *Principal Component Analysis*
PIBIC – Programa Institucional de Bolsas de Iniciação Científica
Pós-Design UFSC – Programa de Pós-Graduação em Design da Universidade Federal de Santa Catarina
UFSC – Universidade Federal de Santa Catarina
SVM – *Support Vector Machine*
TRAA – Tipificação da Representação Animal em Animação

SUMÁRIO

Introdução.....	25
1.1 Problema de Pesquisa.....	27
1.2 Objetivos.....	28
1.3 Justificativa.....	29
1.4 Aderência ao programa Pós-Design.....	29
1.5 Estrutura do Trabalho.....	30
2 Revisão Bibliográfica.....	31
2.1 A Representação dos Animais em Animação.....	31
2.2 Métodos Computacionais para a Transposição da Captura de Movimentos.....	35
3 Procedimentos Metodológicos.....	43
3.1 Tipificação da Representação Animal em Animação.....	43
3.2 Desenvolvimento Conceitual e Construção do Protótipo.....	46
4 Tipificação da Representação Animal em Animação.....	51
4.1 Apresentação e Análise dos Casos.....	51
4.2 Comparação Entre Casos.....	52
4.3 Aplicabilidade dos Resultados.....	57
5 Desenvolvimento Conceitual e Construção do Protótipo.....	59
5.1 Módulo de Dissociação.....	63
5.1.1 Treinamento.....	65
5.1.2 Inferência em Tempo Real.....	70
5.1.3 Aplicabilidade dos Resultados.....	70
5.2 Módulo de Classificação.....	72
5.2.1 Treinamento.....	76
5.2.2 Inferência em Tempo Real.....	80
5.2.3 Aplicabilidade dos Resultados.....	82
5.3 Módulo de Mapeamento.....	87

5.3.1	Treinamento	90
5.3.2	Inferência em Tempo Real	92
5.3.3	Aplicabilidade dos Resultados.....	92
5.4	Módulo de Fusão.....	100
5.4.1	Aplicabilidade dos Resultados.....	102
6	Resultados e Discussão.....	105
6.1	Principais Contribuições	109
6.2	Validação dos Resultados	112
6.3	Limites e Trabalhos Futuros.....	118
6.4	Considerações Finais.....	121
	Referências.....	123
	Apêndice A – Análise De Obras De Animação	131
	Procurando Dory	131
	Zootopia	133
	Kung Fu Panda	134
	Madagascar 3	136
	Pets – A Vida Secreta dos Bichos.....	137
	Shrek 2	139
	Valente	141
	Up – Altas Aventura	142
	Frozen – Uma Aventura Congelante	143
	Apêndice B – Poses Chave Ser Humano	145
	Apêndice C – Poses Chave Tubarão.....	147
	Apêndice D – Poses Chave Cavalo Tipo 1.....	149
	Apêndice E – Poses Chave Cavalo Tipo 2.....	151
	Apêndice F – Termo de Consentimento Livre e Esclarecido	153

INTRODUÇÃO

Personagens animados, criados em ambientes virtuais tridimensionais, são parte de um número significativo de obras utilizadas para o entretenimento, a educação e a comunicação. Podem ser apresentados sob a forma de filmes, vídeos, jogos, livros eletrônicos e outras obras criadas em meio digital, nas quais a principal peça de mediação entre ser humano e conteúdo são as ações de um ou mais personagens.

A palavra animação origina-se do radical latino *animo* acrescido do sufixo *are* que, em conjunto, significam dar vida, ou ainda dar aparência de vida. No caso destes ambientes virtuais, criados com recursos digitais, é possível separar as técnicas mais populares para a animação dos personagens em dois grupos: (1) o personagem pode ser manipulado por um profissional como uma marionete virtual, o que é costumeiramente denominado como animação quadro-a-quadro; (2) o personagem tem seus movimentos derivados da interpretação de um ator, o que é geralmente denominado animação por captura de movimentos (Izani *et al.*, 2003, p.177-178).

No primeiro grupo o especialista, doravante denominado animador, é responsável por criar as poses do movimento do personagem em diferentes instantes das ações que este executa. O número de poses costuma variar entre 24 e 60 por segundo, dependendo do suporte escolhido. Destas poses, o animador costuma conceber algumas, denominadas chave, as quais o computador usa como dados de entrada para gerar o movimento das poses restantes, por interpolação. A quantidade de poses criada pelo animador depende da qualidade e estilo da animação.

No segundo grupo, o da captura de movimentos, um equipamento é utilizado para determinar a posição e orientação do esqueleto de um ator em diversos instantes de tempo, novamente variando de acordo com o suporte pretendido. Estas informações são mapeadas para o esqueleto virtual do personagem que se move de maneira a reproduzir os movimentos do ator. A escolha entre estas duas técnicas tem impactos significativos no método de trabalho e no resultado da obra.

Quanto ao método de trabalho quadro-a-quadro, o uso de animadores treinados exige geralmente muitas horas de trabalho para a produção de alguns segundos de conteúdo. Winder e Dowlatabadi (2011) estimam que, considerando um conteúdo de animação com qualidade suficiente para ser consumido no meio televisivo, um animador é capaz de produzir aproximadamente 5 segundos de animação por dia de

trabalho. Sendo assim é costumeiro, nesta prática, que um planejamento detalhado prescindia a produção, evitando desperdícios no trabalho executado pelo animador. Quanto à qualidade final da animação e suas características formais, estas dependem da capacidade do animador e do tempo disponível para execução da obra, uma vez que o modelo do personagem pode ser manipulado arbitrariamente, quadro-a-quadro.

A técnica de captura de movimentos, por sua vez, possui processos menos onerosos no que se refere ao tempo despendido pelos profissionais envolvidos. De fato, o trabalho é executado em tempo real, cabendo ao ator executar a interpretação do personagem virtual conforme esta tenha sido anteriormente planejada. Por isso a animação realizada por captura de movimentos é dotada de um processo mais flexível que permite a improvisação e a coautoria, estabelecida entre o diretor do projeto, o ator, e o editor do conteúdo capturado. Quanto ao resultado final, a captura de movimentos mostra-se mais restrita que a animação quadro-a-quadro, pois depende das ações que o ator seja capaz de executar, bem como de suas características anatômicas (Kitagawa e Windsor, 2012, p. 19-21).

Assim, ao longo dos anos em que estas duas técnicas vêm sendo utilizadas, nichos foram estabelecidos, dadas as especificidades dos procedimentos metodológicos e dos resultados que proporcionam. O uso da animação quadro-a-quadro é comumente associado aos seguintes cenários: a representação de personagens humanoides caricaturados (os *cartoons*) e a representação de criaturas animais ou fantásticas. Já a captura de movimentos é associada à representação precisa de movimentos humanos e a redução de custos (Freedman, 2012); além disso é a única técnica capaz de ser aplicada em tempo real, sendo a única alternativa para produção de conteúdo ao vivo.

Os nichos de aplicação atualmente existentes na captura de movimentos, dentro do contexto da criação de animações para mídias digitais, não são necessariamente intransponíveis. De fato, parte destes obstáculos acabam sendo transpostos pelo uso de um refinamento dos dados executados por animadores, utilizando a técnica quadro-a-quadro como um pós-processamento das capturas (Kitagawa e Windsor, 2012, p. 39). Mas isto impacta de modo negativo um dos benefícios do uso da captura, a desconexão do processo. Outra possibilidade é o uso de técnicas computacionais para o tratamento dos dados obtidos por captura, o que evita este trabalho manual e especializado. Um exemplo de tais técnicas é o trabalho seminal de Gleicher (1998) que propôs um sistema capaz de adaptar o movimento de atores a personagens humanoides com corpos de proporções diferenciadas.

Diante deste panorama e das restrições anteriormente expostas esta tese de doutorado propõe o desenvolvimento de um *framework* computacional para o tratamento de dados obtidos por captura de movimentos, com vistas à animação de personagens virtuais tridimensionais animais e não humanóides.

A motivação para esta pesquisa se fundamenta em dois eixos. O primeiro é a proeminência deste tipo de personagem dentro do contexto de obras de animação; o segundo eixo é que as técnicas existentes para animação deste tipo de personagem, por captura de movimentos, ainda não dão conta de solucionar o problema em sua integralidade, como será discutido mais adiante neste trabalho.

1.1 PROBLEMA DE PESQUISA

O processo de animação é parte necessária à construção de personagens digitais tridimensionais. Na criação deste tipo de conteúdo duas abordagens são possíveis, a animação quadro-a-quadro realizada por especialistas em animação, ou a captura de movimentos interpretados por atores. Dada a impossibilidade de se representar certos movimentos e estéticas, utilizando a captura de movimentos, o uso desta tecnologia ainda está restrito a alguns nichos.

Um nicho para o qual a captura de movimentos apresenta especial restrição é a animação de personagens animais não humanos, como acontece, em filmes de animação. Esta restrição se dá pela alta correlação entre os dados de captura de movimentos e a estrutura anatômica dos seres capturados. É possível capturar o movimento de um animal e usá-lo para animar um personagem animal, ou capturar o movimento de um ser humano e usá-lo para animar um personagem humanoíde. No entanto, em um conjunto significativo de obras de animação os personagens apresentam um comportamento híbrido de características animais e humanas que será detalhado mais adiante na seção 2.1.

Desde 2009, alguns autores procuram solucionar parte dos problemas inerentes a transferência do movimento humano a personagens animais não-humanóides (como detalhado na seção 2.2), no entanto, outros problemas ainda carecem de solução. Nenhum dos métodos encontrados permite ao ator agir naturalmente, exigindo deste algum tipo de treinamento quanto à forma como deve desempenhar suas ações, ou ainda, exigindo que este execute certos gestos, pré-determinados, para controlar o personagem. Este problema apresenta duas implicações práticas, a primeira é a necessidade de atores treinados para esta tarefa, a segunda é que os dados utilizados para controlar um personagem não

podem ser utilizados para controlar outro personagem de anatomia diferente; uma vez que os gestos são executados com maneirismos dedicados exclusivamente ao controle de uma determinada forma anatômica. Além disso algumas das soluções de transposição de movimentos encontradas na literatura são limitadas a personagens animais com número de membros igual ou menor que os dos seres humanos, enquanto outros casos são indiferentes quanto ao número de membros do personagem. Em casos onde técnicas são combinadas com vistas a superação destas limitações a combinação foi feita de modo heurístico e arbitrário.

Assim, estabeleceu-se o seguinte problema de pesquisa: no contexto da animação de personagens digitais, como utilizar os dados obtidos a partir da captura de movimentos de atores não treinados, para controlar personagens animais não humanos, sem limites quanto às ações interpretadas bem como quanto ao número de membros ou a forma anatômica destes?

1.2 OBJETIVOS

O objetivo geral deste trabalho é propor e desenvolver o protótipo de um *framework* computacional para a utilização dos dados obtidos pela captura de movimentos de atores, não treinados, na animação de personagens animais, não humanoides, capaz de controlar personagens com graus de liberdade inferiores, iguais ou superiores aos do ser humano de quem os movimentos são capturados.

Para tanto são propostos os seguintes objetivos específicos:

1. Descrever e analisar sistematicamente a representação de animais em obras de animação, no sentido de estabelecer uma tipificação deste fenômeno.
2. Estudar como os métodos existentes para a transposição de movimentos obtidos por captura à personagens não humanos se limitam ao contexto de animais representados com diferenciados graus de antropomorfismo e propor técnicas e procedimentos que possam ser adotados para superar tais limitações.
3. Discutir as soluções propostas em função de resultados apresentados por trabalhos anteriores e do contexto de aplicação efetivamente pretendido.

1.3 JUSTIFICATIVA

Partiu-se do princípio que os resultados apresentados em pesquisas anteriores apontam para possíveis soluções à questão de pesquisa anteriormente formulada. Quanto à possibilidade de transpor movimentos humanos a personagens animais não humanos, Baran *et al.* (2009) Ishigaki *et al.* (2009), Yamane *et al.* (2010), Seol *et al.* (2013), Rhodin *et al.* (2014) Rhodin *et al.* (2015) apresentam métodos para o mapeamento de poses e a classificação e transposição de movimentos entre estruturas anatomicamente diferenciadas. Cada um destes métodos possui limitações, detalhadas na seção 2.2, em relação ao contexto de aplicação proposto pela presente pesquisa. Assim está se buscando uma solução capaz de abranger as demandas aqui postas, a saber: o caráter leigo dos atores, a possibilidade de representação de um conjunto diverso de ações e a animação de personagens com número de membros menor, igual ou maior que os de um ser humano.

Acredita-se que a realização dos objetivos aqui estabelecidos sinalizará, com bastante clareza, a possibilidade do uso da captura de movimentos em um conjunto amplo de aplicações das quais hoje a tecnologia está restrita. Isto não só permitirá uma redução de tempo e custos na produção de mídias digitais, como também, espera-se, ensinará novas formas de expressão.

1.4 ADERÊNCIA AO PROGRAMA PÓS-DESIGN

A presente tese foi desenvolvida dentro do programa de Pós-Graduação em Design da Universidade Federal de Santa Catarina (Pós-Design UFSC). Tal programa é constituído de duas linhas de pesquisa: Gestão e Mídia. A linha de pesquisa Mídia é assim definida:

Reúne pesquisas com base nas mídias e suas inter-relações, envolvendo: interatividade, interação, usabilidade, informação e comunicação, dentro das ações de branding, comunicação, educação e entretenimento.(Pós-Design, 2017)

Além disso, o referido programa possui uma ênfase em Tecnologia, que é assim definida: “Reúne pesquisas de Design com ênfase no uso de tecnologias de prototipagem, simulação e experimentação nos métodos, processos e serviços” (Pós-Design, 2017).

A presente pesquisa trata do estudo de métodos para a produção de conteúdo midiático, mais especificamente dentro do contexto do entretenimento e da comunicação. São abordados os aspectos da interatividade e da interação, no processo de criação de conteúdo de animação digital. O pesquisador define claramente o contexto de estudo e, posteriormente, realiza a prototipagem de métodos e processos que permitem alcançar os objetivos propostos.

Desta forma, a presente pesquisa adere explicitamente a linha de pesquisa Mídia, com ênfase em Tecnologia, no programa de Pós-Design UFSC.

1.5 ESTRUTURA DO TRABALHO

O conteúdo da presente tese está organizado da seguinte maneira: (1) neste primeiro capítulo foram apresentados o contexto de aplicação pretendido, o problema, os objetivos e as justificativas para a realização da presente pesquisa, além do seu enquadramento dentro do programa Pós-Design UFSC; (2) no segundo capítulo são apresentadas revisões bibliográficas referentes ao estudo da Representação dos Animais em Animação e aos Métodos Computacionais para a Transposição da Captura de Movimentos; (3) no capítulo 3 trata-se dos procedimentos metodológicos adotados por esta pesquisa no intuito de atingir os objetivos propostos; (4) o capítulo 4 relata os resultados obtidos com o desenvolvimento de uma pesquisa qualitativa que culminou na criação de uma ferramenta teórica para a tipificação do fenômeno da representação dos animais em obras de animação; (5) o capítulo 5 relata o desenvolvimento conceitual e a implementação de um *framework* para a transposição de movimentos em conformidade com os objetivos desta tese; (6) o sexto capítulo apresenta os resultados obtidos com o uso do *framework* desenvolvido, apresenta também os resultados de uma validação preliminar de tal sistema, sumariza as principais contribuições desta tese e aponta direções para futuras pesquisas neste campo.

2 REVISÃO BIBLIOGRÁFICA

No presente capítulo apresenta-se a contextualização teórica da presente pesquisa, dividida em duas partes: a primeira trata do estudo da representação dos animais em obras de animação e a segunda trata dos métodos computacionais para a transposição de captura de movimentos.

A representação de animais em obras de animação não é uma representação fiel de como estas criaturas se comportam na natureza. Portanto, para delimitar o problema da criação de um *framework* que transponha o movimento humano a estes personagens é necessário descrever adequadamente tal modo de representação.

Já a revisão dos métodos computacionais para a transposição de captura de movimentos visa relatar os trabalhos anteriores cuja os objetivos guardam algum grau de semelhança com a presente tese. Ou seja, buscou-se expor e descrever trabalhos que abordam o tratamento de dados obtidos por captura de movimentos com vistas a animação de personagens cuja anatomia apresenta diferenças em relação ao ator capturado; relatando as potencialidades e limitações de cada trabalho em relação ao contexto desta pesquisa.

2.1 A REPRESENTAÇÃO DOS ANIMAIS EM ANIMAÇÃO

Em um levantamento realizado na base de dados Box Office Mojo (Box Office Mojo, 2016) foi possível verificar que da lista das 50 obras de longa-metragem de animação com maior bilheteria, 36 possuíam em seus *trailers* animais como personagens e 14 não possuíam. Em 15, destas 36 obras, os animais figuram como protagonistas e os seres humanos, quando presentes, atuam como coadjuvantes; nas outras 21 obras, humanos, ou outras criaturas, são protagonistas enquanto os animais são coadjuvantes.

Estes dados mostram como, dentro de um conjunto de obras de animação comercialmente relevantes, a representação de figuras animais é altamente recorrente. Entretanto, estas representações não são feitas de forma homogênea, ou seja, uma mesma espécie de animal pode ser representada de modos diferentes de acordo com a obra.

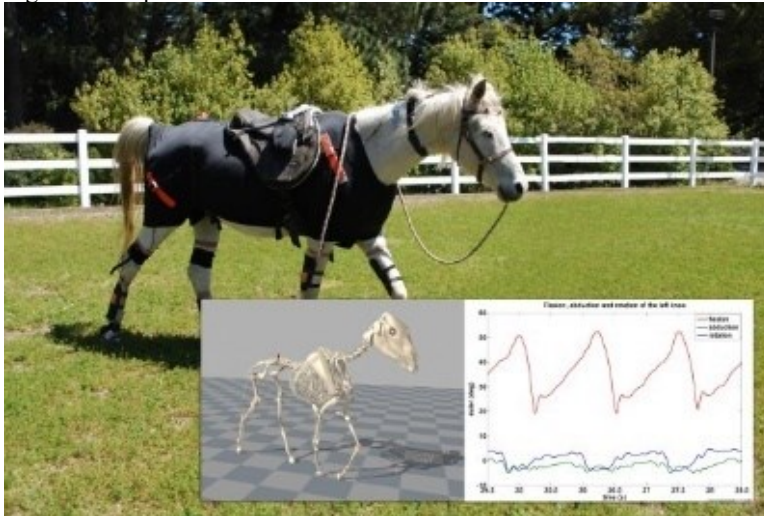
No que concerne o emprego da captura de movimentos, como ferramenta de animação, uma variável importante no modo de representação destes personagens é a maior ou menor presença de características humanas em seus movimentos.

Os dados da captura de movimentos tem alta correlação com a anatomia da criatura capturada (Kitagawa e Windsor, 2012, p. 19-21). Se

em uma animação um personagem animal é representado à semelhança da criatura como esta existe no mundo real, em sua anatomia e no modo que se move, a captura de movimentos pode servir como ferramenta para a animação.

Não há restrições técnicas de maior monta nos equipamentos profissionais de capturas de movimentos quanto a anatomia do ser capturado, na Figura 1, por exemplo, se vê um cavalo em uma seção de captura de movimentos. Se um personagem é animado com as mesmas características de um ser humano (ainda que sua aparência seja de animal) um ator humano pode ser utilizado para as seções de captura de movimentos. Na Figura 2 se vê um exemplo desta situação, onde os movimentos de um humano são utilizados para animar um macaco.

Figura 1 – Captura de movimentos inercial com cavalo



Fonte: Hufgefluester (2011)

Figura 2 – Captura de movimentos com humano em personagem animal



Fonte: Hart (2012)

Figura 3 – O personagem Burro apresenta características animais e humanas



Fonte: Adamson *et al.* (2004)

No entanto, em obras de animação, é comum encontrar casos diferentes destas duas situações ideais à captura de movimentos. Personagens como o burro na animação *Shrek* (Adamson e Jenson, 2001), Figura 3, são animados com um misto de características animais e humanas de modo a manter traços do movimento da espécie animal do personagem mas permitir que este expresse intencionalidade e emoções humanas. Este fenômeno, de atribuir características humanas a objetos e criaturas não humanas, recebe o nome de antropomorfismo.

O antropomorfismo é encontrado em diversas épocas e culturas, encarnado em mitos, artefatos e no folclore em geral. A esfinge de Gizé, por exemplo, representa a figura de um leão com a cabeça de um homem. O mito grego do Minotauro, descreve um homem com cabeça de touro. Nos contos folclóricos germânicos, registrados pelos irmãos Grimm, *Chapeuzinho Vermelho* interage com um lobo que, como homem, pensa e fala. Também na época contemporânea, na civilização ocidental, podemos encontrar exemplos de representações antropomórficas. Na literatura, um destes exemplos é a obra “*A Revolução dos Bichos*” (Orwell, 1945), onde animais, que pensam e falam, servem à representação de um evento histórico, a revolução russa, sem que seja necessário mencioná-lo diretamente. A história em quadrinhos “*Maus*” (Spiegelman, 1991), de modo similar, representa o holocausto judeu como uma história ocorrida entre gatos e ratos que pensam, falam e andam sobre duas patas. Neste contexto cultural, a animação é um formato onde a representação antropomórfica tem singular recorrência. Uma das primeiras obras deste gênero, da qual se tem registro, “*Gertie the Dinosaur*” (McCay, 1914), representa um dinossauro que se comporta de modo humano ao dançar sobre duas patas e ao chorar. Com o advento do período que alguns historiadores chamam de “a era de ouro da animação” muitos outros personagens antropomórficos surgiram, como: *Mickey Mouse* (Disney e Iwerks, 1928), *o Pato Donald* (Jackson, 1934), *o Pica-Pau* (Lantz e Lovy, 1940), *Dumbo* (Sharpsteen, 1941), dentre outros. Este uso de características humanas em figuras animais continua na animação até os dias de hoje.

Alguns estudos buscam descrever e analisar o fenômeno das representações antropomórficas em animação. Jardim (2013) pondera que, dado que o público infantil é uma das principais audiências das obras de animação, uma hipótese para a recorrência do antropomorfismo neste meio é a propensão das crianças a atitude animista. Ou seja, a atitude de interpretar os objetos do mundo como agentes pensantes capazes de agir intencionalmente. Ainda em Jardim (2013) outra hipótese levantada para a recorrência do fenômeno antropomórfico é o diferencial do cinema de

animação, em oposição ao cinema tradicional, de ser um meio capaz de representar o fantástico e o lúdico. Em “Understaing Animation”, Wells (1998), discute o uso do antropomorfismo no contexto da animação como forma de trazer às obras símbolos e conceitos culturalmente pré-estabelecidos, facilitando o desenvolvimento da narrativa. Por exemplo, um animal selvagem pode representar perigo, enquanto que um animal doméstico pode representar fidelidade. O autor também discute os símbolos humanos sobrepostos à figura do animal, como os trejeitos e adereços utilizados para representar os gêneros masculino e feminino em personagens como Mickey e Minnie Mouse.

Em “The Animated Bestiary, Animals, Cartoons, and Culture” (2008) Wells faz uma revisão de diferentes representações de animais em animação, ao longo da história deste meio, buscando estabelecer as condições culturais que fomentam tais representações, e as funções narrativas desempenhadas por estes personagens. Outros autores ainda, como Baker (1993) e Leventi-Perez (2011), estão interessados em estudar animações antropomórficas como um documento histórico capaz de prover indícios para a análise da percepção da figura do animal em uma determinada cultura.

Um tema sub-representado na literatura é a descrição do fenômeno do antropomorfismo na animação com vistas às características qualitativas da representação dos movimentos dos personagens. Tal descrição é importante no contexto do desenvolvimento deste trabalho pois, em uma análise preliminar, é possível verificar que as características do movimento em figuras animais, não humanoides, variam significativamente de uma obra para outra. Assim é possível intuir que estas variações afetem de modo diferenciado a produção de animação e suas técnicas.

2.2 MÉTODOS COMPUTACIONAIS PARA A TRANSPOSIÇÃO DA CAPTURA DE MOVIMENTOS

Nesta seção, apresenta-se a síntese dos resultados de uma revisão realizada no sentido de encontrar trabalhos que propusessem soluções aos desafios impostos pelo uso da captura de movimentos na animação de personagens animais humanoides e não humanoides. As buscas foram realizadas no período compreendido entre os dias primeiro de agosto de 2016 e o dia primeiro de setembro de 2016 nas bases ACM Library, Periódicos CAPES e Google Acadêmico. A relevância de cada artigo para com o tema da presente pesquisa foi auferida pela análise dos respectivos resumos.

Além deste primeiro conjunto de trabalhos encontrados a partir das buscas em bases de dados, um segundo conjunto de trabalhos foi selecionado a partir das citações encontradas no primeiro conjunto de artigos. No total foram analisados, para a ocasião desta revisão, 19 trabalhos.

De acordo com Kitagawa e Windsor (2012, p. 73), ainda que a estrutura anatômica do personagem capturado e do personagem animado seja similar, na maioria dos casos as proporções entre o sujeito e o personagem são diferentes. No caso, os autores se referem a captura de movimentos para a animação de personagens humanoides, ainda que sejam animais, como no exemplo da Figura 2. Ainda segundo Kitagawa e Windsor (2012, p. 73) estas diferenças de proporção podem gerar problemas como a sensação de que o movimento do personagem é duro, ou desajeitado, além de causar a impressão de que os pés do personagem deslizam sobre o cenário.

Um dos primeiros trabalhos a descrever um método para solucionar os problemas causados por diferenças nas proporções do esqueleto capturado e do esqueleto animado foi o de Gleicher (1998). A solução proposta pelo autor permite a adaptação da animação ao tamanho dos ossos e a posição das passadas do ator, evitando que o personagem deslize durante a animação. De modo complementar Shin *et al.* (2001) verificam que quando o personagem e um ator tem diferenças de proporção, além do problema dos contatos dos pés sobre o chão, ocorre problema similar quando as mãos do ator entram em contato com algum objeto, fazendo com que o personagem atravesse ou não alcance o objeto pretendido. Os referidos autores propõem então, um método computacional para análise das poses do ator, permitindo determinar de modo heurístico quando este manipula objetos, ativando no sistema de animação funções que adaptam o movimento do personagem. Al-Asqhar *et al.* (2013) propõe uma maneira alternativa de descrever a cinética do movimento que soluciona de modo generalizado o contato do personagem com o piso, com outros objetos da cena e até a interação com outros personagens em movimento.

Todas estas soluções são cinemáticas, ou seja, levam em conta apenas a trajetória dos movimentos e não suas propriedades dinâmicas. Trabalhos propostos por Wrotek *et al.* (2006), Lee *et al.* (2010) e Geijtenbeek *et al.* (2012), modelam o problema da transferência de movimento entre ator e personagem como um problema dinâmico e, assim, possibilitam resultados mais adequados mesmo a adaptações anatômicas extremas (Figura 4). Além de possibilitar a consideração de outras propriedades anatômicas do personagem, como o peso, que altera

o equilíbrio do corpo e, conseqüentemente, a maneira como o personagem se movimenta.

Figura 4 – Abordagens: cinética (à esquerda) onde as rotações das juntas do ator são transferidas diretamente ao personagem com proporções modificadas; e dinâmica (à direita) onde a distribuição do peso sobre o corpo do personagem com as proporções modificadas é levada em conta para deformar a pose original.



Fonte: Geijtenbeek *et al.* (2012)

Além da transposição de movimentos, uma técnica desenvolvida no processamento de dados de captura de movimentos é a transferência de características estéticas da representação de um movimento a outros. A premissa das técnicas que visam tal aplicação é alterar a estética de um movimento, utilizando uma segunda captura como exemplo e sintetizando movimentos inéditos.

Um trabalho pioneiro neste campo foi publicado por Amaya *et al.* (1996) e propõe que dada uma animação “neutra” de uma determinada ação e uma animação interpretada com características estéticas ligadas a um estado emotivo (ex.: agitado, nervoso, inseguro, etc...), é possível estabelecer correlações temporais entre as duas animações e encontrar a diferença entre os dois estados. Estas diferenças, de velocidade e amplitude, podem então ser aplicadas a outras animações neutras. O sistema, inicialmente apresentado por Amaya *et al.* (1996) apresenta limitações, como estar restrito à aplicação em pequenos ciclos de animação, com início e fim determinados.

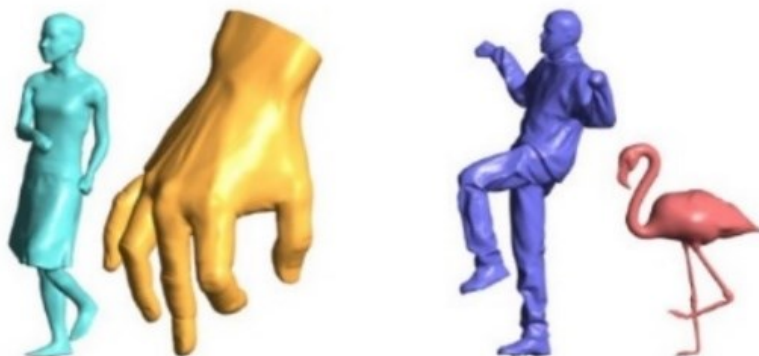
Hsu *et al.* (2005) expandem o trabalho de Amaya *et al.* (1996) de modo a permitir seu uso em longas sequências de movimento e de modo a compensar problemas advindos das modificações aplicadas ao movimento, como pés que deslizam por sobre o cenário. O trabalho de Min *et al.* (2010), utilizando um grande banco de dados de captura de movimentos, propõe um método para separar as diferenças de velocidade e amplitude em dois componentes, um componente descrevendo características estéticas médias dos movimentos de uma população, e um segundo descrevendo as características estéticas únicas ao movimento de cada indivíduo desta população. O trabalho publicado por Xia *et al.* (2015) propõe uma solução onde uma determinada característica estética do movimento não é aprendido globalmente, mas localmente em relação a um conjunto de ações de treinamento, ou seja, as características que definem uma determinada estética podem variar em função da ação desempenhada. Na ocasião do uso, quando o movimento do ator é capturado, o sistema procura, dentro do conjunto de treinamento, os exemplos mais parecidos para determinar a solução adequada. Yumer e Mitra (2016) expandem o trabalho de Xia *et al.* (2015) propondo uma solução para casos onde a ação desempenhada não encontra correspondência direta no conjunto de treinamento.

No entanto, como descrito item 2.1, na representação de animais em animação são recorrentes os casos onde a anatomia dos personagens e a maneira como atuam não são do tipo humanoide. Uma das primeiras abordagens ao problema da transposição de movimentos entre anatomias significativamente diferenciadas foi proposta por Baran *et al.* (2009). Neste trabalho os autores apresentam um método capaz de estabelecer correlações entre duas geometrias¹, dado um determinado conjunto de poses em cada uma delas. Assim o sistema proposto é capaz de sintetizar a animação em uma geometria a partir da análise de outra geometria (Figura 5). As limitações deste trabalho incluem: (1) operar sobre a geometria do modelo animado, pois isso impede que a animação seja capturada por um equipamento de captura de movimentos, exigindo tecnologias mais onerosas, como sistemas de escaneamento 3d; (2) dependendo do movimento pretendido o ator deve interpretar a ação

¹ Em animações realizadas por computação gráfica é comum que um personagem seja composto por dois grandes componentes a geometria e o sistema de articulações. A geometria é o conjunto de superfícies que descrevem a aparência do personagem e o sistema de articulações é uma espécie de esqueleto virtual responsável pela deformação da geometria (O'Neill, 2015, p.64-68). No caso do trabalho proposto por Baran *et al.* (2009) o modelo não possui um sistema de articulações e todos os procedimentos são aplicados estritamente sobre seu componente geométrico.

mimetizando o personagem animal, impedindo a transposição de um movimento capturado a múltiplos modelos animais diferenciados; (3) o modelo animado deve conter, no máximo, o mesmo número de graus de liberdade do ator; ou seja, não é possível, estabelecer uma correlação adequada entre um ator humano e um personagem com mais membros que um humano.

Figura 5 – Geometrias humanoides controlando modelos não humanoides a partir da correlação entre poses chave



Fonte: Baran *et al.* (2009)

Yamane *et al.* (2010) descrevem um sistema similar ao de Baran *et al.* (2009), no entanto, propõe a correlação entre poses armazenadas no sistema de articulações dos personagens e não em seus modelos geométricos, permitindo assim o uso de dados obtidos por captura de movimentos². Além disso, este modelo resolve alguns outros problemas como, por exemplo, o contato dos pés com o piso da cena. Assim como o método proposto por Baran *et al.* (2009) esta abordagem também exige que o ator mimetize algumas ações à moda do animal interpretado e é incapaz de representar modelos com mais membros que os atores capturados. Além disso nenhum dos dois sistemas é capaz de estabelecer relações temporais entre o movimento de atores e os personagens, uma vez que guardam apenas informações referentes às correlações espaciais dos modelos, ou seja, suas poses. Chen *et al.* (2012) propõe um terceiro, e simplificado, modelo utilizando apenas uma pose no modelo de entrada

² Armazenar as informações de pose no sistema de articulações do personagem é o procedimento padrão na maioria dos programas de animação quadro-a-quadro e de captura de movimentos.

e uma pose no modelo de saída para manter uma correlação entre ator e personagem. Apesar de ter um tempo reduzido de configuração os resultados do sistema são limitados quando comparados aos modelos de Baran *et al.* (2009) e de Yamane *et al.* (2010).

Visando um sistema capaz de utilizar animações obtidas por captura de movimentos para controlar modelos com número de membros superiores aos de um ator humano, Seol *et al.* (2013) propuseram um sistema capaz de estabelecer correlações entre pares de movimentos animados, ao invés apenas de poses chave. O trabalho foi inspirado, em parte, no modelo proposto por Ishigaki *et al.* (2009) para controlar personagens humanos completos a partir da captura parcial de movimentos. O sistema de Seol *et al.* (2013) tem duas partes: uma primeira parte estabelece arbitrariamente a relação entre membros no modelo de entrada e saída mapeando poses de movimentos simples; e, uma segunda parte, que classifica movimentos desempenhados pelo ator acionando movimentos equivalentes e pré-animados no modelo animal. A escolha pelo uso de um ou outro método de transposição dos movimentos se dá de modo arbitrário: quando o sistema não é capaz de classificar o movimento do ator, recai no uso do mapeamento de movimentos (Seol *et al.*, 2013, p.217).

Rhodin *et al.* (2015) propõe um modelo capaz de classificar e distinguir a amplitude, a fase e a frequência dos movimentos em gestos realizados com a mão e sugerem a utilização destes dados para controlar ações cíclicas em personagens animais.

Apesar de estes dois trabalhos (Seol *et al.*, 2013; Rhodin *et al.*, 2015) apresentarem soluções pertinentes ao mapeamento de movimentos entre personagens com graus de liberdade maiores que os dos seres humanos, cabe ressaltar aqui as limitações encontradas. O trabalho de Rhodin *et al.* (2015) pressupõem a existência de um modelo animado do animal para cada ação que se deseje representar, e se limita a representação de ações de locomoção. Apesar de Seol *et al.* (2013) tentarem mitigar este problema acoplando um sistema para o mapeamento de poses à técnica proposta, a escolha pela método de transposição adequado se dá de modo arbitrário, em função das limitações do sistema de classificação de movimentos. Tais limitações são pouco significativas no contexto de aplicação em que os trabalhos citados se apresentam, a saber, o controle de personagens em jogos eletrônicos. No entanto, o uso destas soluções em contextos mais amplos, onde o ator está livre para executar um conjunto maior de ações, implicaria na necessidade de criação de uma grande base de animações para o treinamento dos referidos sistemas.

A Tabela 1 sumariza as características, potencialidades e limitações dos trabalhos capazes de realizar a transposição de movimentos entre anatomias diferenciadas encontrados nesta revisão. Em síntese constituem-se como desafios em aberto a trabalhos que busquem viabilizar o uso da técnica de captura de movimentos na representação de animais não humanoides e antropomórficos em animação: (1) a transposição movimentos humanos a modelos animais com número diferenciado de membros em seus corpos considerando um conjunto de ações tão irrestrito quanto possível; e (2) a possibilidade de que os sujeitos capturados possam atuar naturalmente, permitindo o uso do sistema de transposição de movimentos por indivíduos não treinados. Tais lacunas anteriormente elencadas são alguns dos desafios explorados e superados nesta tese.

Tabela 1 – Comparação entre métodos de transposição de captura de movimentos

Autor e ano	Método	Anatomia Humana	Anatomia Não-humana	Graus de liberdade ilimitados	Contato c/ solo	Contato c/ objetos	Sim. de dinâmica
Gleicher, 1998	Otimização com restrições						
Shin et al., 2001	Cinemática inversa						
Wrotek et al., 2006	Simulação dinâmica						
Ishigaki et al., 2009	Classificação SVM						
Baran et al., 2009	Regressão linear						
Yamane et al., 2010	Regressão por GPLVM						
Lee et al., 2010	Simulação dinâmica						
Chen et al., 2012	Transformação linear						
Geijtenbeek et al., 2012	Simulação dinâmica e otimização com restrições						
Al-Asqhar et al., 2013	Representação cinemática alternativa						
Seol et al., 2013	Classificação SVM						
Rhodin et al., 2015	Filtros de Garbor						

Fonte: O autor.

3 PROCEDIMENTOS METODOLÓGICOS

No presente capítulo são apresentados os procedimentos metodológicos que guiaram os dois momentos distintos desta pesquisa, a saber: (1) a tipificação da representação animal em animação; e (2) o desenvolvimento conceitual e construção do protótipo. O primeiro momento está diretamente ligado ao seguinte objetivo específico da pesquisa: (1) Descrever e analisar sistematicamente a representação de animais em obras de animação, no sentido de estabelecer uma tipificação deste fenômeno. O segundo momento está diretamente ligado ao seguinte objetivo específico desta pesquisa: (2) Estudar como os métodos existentes para a transposição de movimentos obtidos por captura à personagens não humanos se limitam ao contexto proposto e propor técnicas e procedimentos que possam ser adotados para superar tais limitações.

3.1 TIPIFICAÇÃO DA REPRESENTAÇÃO ANIMAL EM ANIMAÇÃO

A Tipificação da Representação Animal em Animação (TRAA) consiste de um estudo qualitativo de caráter exploratório realizado com diferentes obras de animação, cada qual compreendida como um caso. A escolha por esta abordagem se deu pelo entendimento que as variáveis que compõe as qualidades antropomórficas do movimento, como representadas nos desenhos animados, não são claramente definidas na literatura e, portanto, precisam ser exploradas. Segundo Creswell (2012, p. 17-18) tais condições caracterizam predisposições típicas de um estudo de abordagem qualitativa. Ainda segundo o autor (Creswell, 2012, p. 61-64) o método de estudo de múltiplos casos permite a descrição detalhada de cada uma das obras em si, além de permitir a comparação das características encontradas entre essas diferentes obras. Por meio desta comparação foi possível estabelecer uma escala dos diferentes graus do movimento antropomórfico em peças de animação. Tal escala colabora com a delimitação do problema da proposição e construção do protótipo.

A seleção dos casos foi determinística, pois, buscou-se representar instâncias qualitativamente diferenciadas do movimento antropomórfico em peças de animação, dentro de um determinado contexto cultural e técnico. A restrição do conjunto de animações analisadas ocorreu em função do campo, para com o qual, esta pesquisa objetiva contribuir, isto é, a produção profissional de animações digitais 3d. Por produção profissional pretende-se a prática da animação como uma atividade

econômica em um sistema de divisão do trabalho. Assim, a bilheteria foi tomada como indicador quantitativo do sucesso da obra³.

Quanto ao caráter digital, optou-se por limitar a análise a obras realizadas dentro do período histórico mais recente, em que o ferramental para a produção de animação digital 3d tornou-se ubíquo. Utilizou-se como limite arbitrário de tempo o ano 2000, momento em que, segundo Yoon e Malecki (2009, p. 257), o baixo custo de hardware e software tornaram a produção digital de animação 3d lugar comum.

Assim resultou que o conjunto das animações analisadas foram aquelas de maior bilheteria, em âmbito mundial, no período compreendido entre os anos de 2000 e 2016. A base de dados consultada para determinar a bilheteria das obras, bem como as respectivas datas de publicação, foi a plataforma digital Box Office Mojo (Box Office Mojo, 2016). Dentro deste conjunto, foram analisados os *trailers* das obras de modo a permitir uma pré-seleção daquelas que apresentavam as características desejadas por esta pesquisa, ou seja, a representação de animais não humanos. A presença destes animais no *trailer* foi tomada como indício suficientemente forte de que estes personagens têm atuação significativa e não acidental ou cênica na obra. Foi feita uma opção por excluir obras onde o antropomorfismo era encontrado unicamente em objetos inanimados ou em criaturas fantásticas, pois estes casos estão fora do escopo da presente pesquisa. No total 9 obras de animação foram analisadas; o nome das obras e os critérios detalhados de seleção estão apresentados na seção 4.1.

A análise de cada um dos casos foi feita em três etapas, como propostas por Penn (2003, p. 340-341):

- A primeira etapa consistiu da própria seleção de obras.
- A segunda etapa consistiu da descrição dos elementos denotativos das obras, neste caso as características anatômicas dos personagens e a maneira como se movem, em diferentes situações. Com características mais ou menos antropomórficas.
- A terceira etapa consistiu na descrição dos elementos conotativos, ou seja, como os movimentos

³ Com isso os autores não visam minimizar a importância da animação realizada como arte, desapegada de considerações econômicas. Entende-se esta como forma de expressão culturalmente relevante, mas obra do indivíduo que Von Mises (1949, p. 138-140) classifica como gênio criativo; ou seja, um indivíduo para quem a realização de sua obra é objetivo e não meio e, portanto, para quem a lógica da divisão do trabalho não se aplica.

antropomórficos encontrados podem contribuir com a significação da obra.

Nota-se que, como afirma a autora (Penn, 2003), a última etapa da análise, a etapa conotativa, é um exercício de interpretação empenhado pelos pesquisadores e que, portanto, depende da capacidade destes em elucidar os diferentes signos observados. As efetivas análises de cada uma das obras encontram-se descritas no Apêndice A.

Além da análise de cada caso foi feita uma análise comparativa entre os casos, visando o estabelecimento de uma escala ordinal da presença do antropomorfismo. A priori entende-se que uma escala do antropomorfismo na representação de movimentos de personagens animais tem localizados em seus extremos dois tipos ideais⁴ bem definidos: em um dos extremos o movimento é completamente naturalista, ou seja, os movimentos dos animais são representados como se apresentam na natureza, sem qualquer característica humana; no extremo oposto o movimento é absolutamente antropomórfico, ou seja, todos os movimentos dos personagens podem ser entendidos como humanos, ainda que a anatomia ou aparência do personagem não o sejam. As obras foram então analisadas comparativamente, de modo a estabelecer, entre elas, uma ordem de ocorrência do fenômeno do antropomorfismo na representação dos movimentos de diferentes personagens. Esta escala comparativa e ordinal se justifica, em oposição a uma escala cardinal, pela impossibilidade de se determinar que o personagem A é duas vezes mais antropomórfico que o personagem B. Tudo que se pode saber é se A é mais ou menos antropomórfico que B. A comparação entre casos e definição da escala ordinal de antropomorfismo encontra-se descrita na seção 4.2.

Realizada tal análise foi possível então enquadrar os métodos computacionais existentes para a transposição de movimentos obtidos por captura (descritos na seção 2.2) à tipificação recém-criada, evidenciando assim as lacunas existente no corpo teórico deste campo e delimitando o

⁴ Utiliza-se aqui o conceito de tipo ideal como proposto por Weber. Segundo o autor o tipo ideal é uma categoria estritamente teórica composta das características mais acentuadas de fenômenos que ocorrem no mundo real (Weber, 1949, p. 89). É um conceito similar ao da ideia pura encontrado na obra de Platão, exceto que o filósofo grego não entende a ideia pura como uma ferramenta estritamente teórica, mas como algo que existe em si mesmo. Um exemplo de tipo ideal é o da figura do triângulo equilátero. Na definição matemática a figura tem os três ângulos iguais, mas nas representações feitas no mundo real, aquilo que chamamos de equilátero, nunca tem os ângulos precisamente iguais.

escopo da construção do protótipo. Estes resultados são apresentados na seção 4.3.

3.2 DESENVOLVIMENTO CONCEITUAL E CONSTRUÇÃO DO PROTÓTIPO

O momento do Desenvolvimento Conceitual e da Construção do Protótipo consistiu de um estudo de caráter experimental que buscou testar soluções para as seguintes limitações encontradas nos métodos de transposição de movimentos existentes na literatura:

- Inexistência de um método para dissociar movimentos eminentemente antropomórficos de outros movimentos
- Incapacidade de classificar movimentos humanos em um contexto de alta estilização
- Incapacidade de mapear poses de atores humanos a personagens não-humanos para um conjunto irrestrito de ações em tempo real

Cada um destes problemas foi analisado individualmente. Buscou-se na literatura existente problemas com características similares, em outros contextos de aplicação, adequando-os quando possível ao contexto aqui proposto. As soluções encontradas baseiam-se em sua totalidade no uso de técnicas de Aprendizado de Máquina (AM).

O AM permite que um sistema computacional detecte padrões relevantes em um conjunto de dados, “aprendendo” a calcular uma resposta aproximada ao problema pretendido sem que seja necessário a programação detalhada e explícita da tarefa por um humano (Shalev-Shwartz e Ben-David, 2014, p. 7). Esta abordagem é relevante em casos onde apesar de humanos serem capazes de realizar uma dada tarefa tem dificuldade em refletir e tornar explícitos todos os procedimentos necessários a sua execução (Shalev-Shwartz e Ben-David, 2014, p. 21-22). Tal é o caso dos problemas em questão, ou seja: distinguir movimentos que devem ser transpostos ao personagem de modo antropomórfico de outros movimentos, classificar movimentos humanos e mapear poses entre modelos anatômicos diferenciados.

Para a criação do protótipo focou-se em uma abordagem bio-inspirada de AM conhecida como redes neurais, descrita em maior detalhe no capítulo 5. Ainda quanto as redes neurais utilizadas na construção do protótipo, todas elas foram implementadas com auxílio do software MATLAB (Mathworks, 2016a) e de um complemento para tal

software intitulado *Neural Network Toolbox* (Mathworks, 2016b), Figura 6.

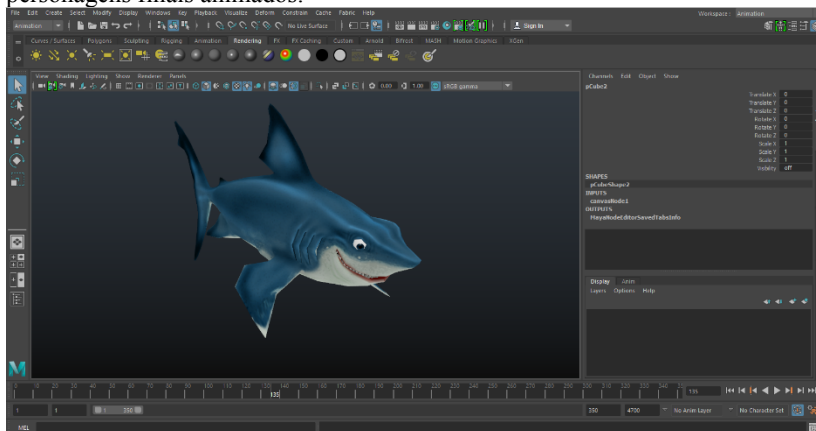
Os dados de captura de movimentos utilizados para treinar tais redes neurais são provenientes da base *CMU Graphics Lab Motion Capture Database* (Cmu, 2017). Tal base de dados consiste em um conjunto de 2605 sessões de capturas de movimentos de 142 indivíduos de tamanhos e gêneros diferentes, organizadas em 23 categorias. A escolha por tal conjunto de dados se deu pelos seguintes fatores: tamanho da amostra e diversidade de temas representados, categorização das atividades representadas e a ampla utilização da base no meio científico, o que permite a comparação com outros trabalhos.

Ainda a respeito dos dados da captura de movimentos faz-se saber que o *framework* proposto e implementado nesta pesquisa se limita a tratar a captura de movimentos do corpo humano, excluindo deliberadamente os movimentos da face. Tal limitação fundamenta-se em dois motivos: (1) que a representação facial de personagens em animação é altamente humanizada; e (2) que a natureza dos dados de captura de movimentos faciais e corpóreos são significativamente diferentes⁵.

O processamento dos dados de captura de movimentos tanto nas etapas que antecederam quanto nas que sucederam o treinamento das redes neurais foi feito em uma plataforma de desenvolvimento voltada para aplicações de computação gráfica intitulada *Fabric Engine* (Fabric Software Inc., 2017), Figura 6. E, finalmente, a visualização dos modelos finais animados foi feita dentro da plataforma de animação *Autodesk Maya* (Autodesk, 2017), Figura 7.

⁵ Enquanto os movimentos do corpo são comumente representados pela orientação das juntas do personagem as expressões faciais são comumente obtidas através de uma combinação de diferentes modelos do rosto, denominados *face shapes*, combinados por interpolação linear (O'Neill, 2015, p.167-168).

Figura 8 – Interface gráfica da plataforma Maya, utilizada para visualizar os personagens finais animados.



Fonte: O autor.

O funcionamento do *framework* proposto pode ser entendido de modo compartimentado, em módulos. Um primeiro Módulo de Dissociação, responsável por separar os movimentos humanos ligados ao equilíbrio e ao deslocamento dos demais movimentos corpóreos, ditos potencialmente expressivos⁶. Um segundo módulo (Módulo de Classificação) é responsável pela transposição dos movimentos de equilíbrio e deslocamento, e um terceiro módulo (Módulo de Mapeamento) é responsável pela transposição dos movimentos potencialmente expressivos. Como a transposição dos movimentos se dá em duas etapas diferentes foi necessária também a criação de um último módulo (Módulo de Fusão) capaz de realizar a fusão dos componentes do movimento. Este último módulo, diferentemente dos demais, não deriva de uma limitação existente na literatura atual, mas sim de uma característica da abordagem proposta, a saber: de separar o movimento em dois componentes distintos.

Ainda sobre estes módulos é possível dizer que o Módulo de Dissociação consiste de um sistema inédito. Os componentes dissociados representam diferentes pontos no espectro de tipificação do movimento animal (proposto no capítulo 4), permitindo a síntese de animações em diferentes categorias de tal espectro.

O Módulo de Classificação consiste de um sistema de classificação de movimentos que é utilizado para o controle de uma máquina de estados

⁶ Uma definição precisa destes componentes pode ser encontrada no capítulo 5.

de um personagem animal animado. Tal abordagem é semelhante ao que foi proposto por trabalhos anteriores (Seol *et al.*, 2013; Rhodin *et al.*, 2015), no entanto, o sistema aqui apresentado tem a inédita capacidade de classificar o movimento de atores não treinados executando ações altamente estilizadas. Tais casos são importantes à animação de personagens e eram limitações de trabalhos anteriores.

O Módulo de Mapeamento consiste de um sistema para o mapeamento de poses entre um esqueleto digital humano e um esqueleto digital animal não-humano. Nesse sentido o sistema apresentado é similar a aquele proposto por Yamane *et al.* (2010). No entanto, no presente trabalho tratou-se da limitação, anteriormente existente, da necessidade de uso de um especialista em animação na criação de poses para o personagem animal a cada seção de captura. Tratada esta limitação o sistema aqui proposto se diferencia do trabalho original de Yamane *et al.* (2010) por ser capaz de atuar em tempo real com um ator não treinado.

Por fim o Módulo de Fusão combina o resultado dos dois módulos de mapeamento em um modelo único.

O relato das atividades realizadas no Desenvolvimento Conceitual e Construção do protótipo encontram-se no capítulo 5.

4 TIPIFICAÇÃO DA REPRESENTAÇÃO ANIMAL EM ANIMAÇÃO

Este capítulo, consiste do estudo qualitativo realizado com diferentes obras de animação, cada qual compreendida como um caso. Os procedimentos adotados na análise aqui relatada estão descritos em detalhes na seção 3.1. Os resultados das análises intra-obras estão detalhados na seção 4.1, já os resultados da comparação inter-obras estão descritos na seção 4.2. Por fim a seção 4.3 enquadra os métodos computacionais para o tratamento de captura de movimentos à tipificação da representação animal proposta, colaborando com a delimitação da construção do protótipo detalhada no capítulo 5.

4.1 APRESENTAÇÃO E ANÁLISE DOS CASOS

Aplicando o conjunto de procedimentos estabelecidos na seção 3.1, foram selecionadas as trinta obras de animação com maior bilheteria global, como disponíveis em relatório da base Box Office Mojo (Box Office Mojo, 2016). Destas, removeram-se as que utilizavam a técnica de animação 2d ou obras que haviam sido produzidas antes do ano 2000, além daquelas que não apresentavam, em seus trailers, animais como parte da história. Portanto dentre as trinta obras de maior bilheteria restaram treze que cumpriam todos os quesitos de seleção para a análise.

No entanto, algumas destas obras eram sequencias de outros filmes, por exemplo, o filme Shrek (Adamson e Jenson, 2001) e todas suas três sequencias aparecem nesta lista. Como era interesse deste estudo analisar obras com qualidades diversas, e as sequencias filmicas apresentam demasiada semelhança entre si, optou-se por analisar sempre a obra com maior bilheteria dentre todas as sequencias.

Restaram, portanto, nove obras, que foram todas analisadas. Destas, cinco são obras onde os animais são protagonistas e os humanos aparecem na história de modo secundário ou nem aparecem, são elas: Procurando Dory (Stanton, 2016), Zootopia (Howard e Moor, 2016), Pets – A Vida Secreta dos Bichos (Renaud e Cheney, 2016), Madagascar 3 (Darnell *et al.*, 2012) e Kung Fu Panda (Osborne e Stevenson, 2008). As outras 4 obras, onde os seres humanos (ou criaturas fantásticas) são protagonistas e os animais aparecem como coadjuvantes na história, são: Shrek 2 (Adamson *et al.*, 2004), Frozen (Buck e Lee, 2012), Up – Altas Aventuras (Docter, 2009) e Valente (Andrews e Chapman, 2012).

As análises de cada uma destas nove obras estão dispostas no Apêndice A e são compostas dos seguintes itens:

- Resumo da narrativa: Uma breve descrição da história narrada durante o filme.
- Anatomia dos personagens: Descrição das espécies e anatomia dos personagens animais presentes na obra.
- Movimentos Antropomórficos: as características mais ou menos antropomórficas encontradas nos movimentos dos personagens animais.
- Antropomorfismo na narrativa: a interpretação realizada acerca da função exercida pela qualidade dos movimentos dos personagens na narrativa.

Após a análise de todos os casos, estabelece-se, na seção 4.2, uma comparação entre os movimentos antropomórficos encontrados nas diferentes obras.

4.2 COMPARAÇÃO ENTRE CASOS

Como estabelecido nas considerações realizadas a respeito do método, item 3.1, pretende-se comparar como os animais são representados, em diferentes obras de animação, quanto as qualidades antropomórficas de seus movimentos. Entende-se, para tanto, que as características destes movimentos se localizam entre dois tipos ideias o naturalista e o antropomórfico. Nota-se que em cada uma das obras diferentes personagens apresentam movimentos com diferentes aspectos e, por vezes, até um mesmo personagem pode ser representado de diferentes modos em função da narrativa. Agruparam-se os personagens em categorias de acordo com a semelhança com que elementos antropomórficos foram encontrados em seus movimentos. Nos parágrafos seguintes, detalham-se as características de cada categoria e os personagens que as exemplificam.

Na categoria 1 temos uma série de animais secundários à narrativa, como cavalos, cachorros e outros, nas obras Valente e Frozen, os quais são representados de modo estritamente naturalista, ou seja, seus movimentos são aqueles característicos de suas espécies. Nota-se que a função destes animais, nas referidas obras, é apenas criar uma ambientação assim como fazem outros elementos cênicos. Quando um personagem, nestas histórias, vai do ponto A ao ponto B a cavalo, o animal está lá para representar o costume de andar a cavalo que poderia ter sido substituído, por exemplo, por uma longa caminhada. Além destes podemos citar casos de animais com movimentos representados de modo completamente naturalista, como os ursos de Valente, quando atçados

pelo cheiro de comida, e os animais de Zootopia quando em modo selvagem; estas representações, no entanto, são motivadas por eventos narrativos.

A categoria 2, tem um grau de antropomorfismo imediatamente maior que a primeira categoria e é composta por animais como o alce de Frozen e os cachorros de Up. Ambos os casos, coincidentemente, de animais que servem de companhia a personagens solitários que tem dificuldade em manter relações com outros seres humanos. Nestes dois casos são utilizados artifícios para comunicar intenções e sentimentos dos personagens: no caso do alce, um ventríloquo e no caso dos cachorros, uma coleira que lê pensamentos. Os movimentos antropomórficos dos personagens restringem-se a expressões faciais e movimentos da cabeça que corroboram e complementam o sentido indicado pelos artifícios de fala. A mãe urso em Valente apresenta diferentes graus de antropomorfismo, no menor grau, antes de atingir uma postura completamente naturalista, a personagem se encaixa neste grupo pois utiliza apenas a face e os movimentos da cabeça como forma de expressão. Os cavalos de Madagascar, a arraia em Procurando Dory, o canário em Pets e o corvo em Valente se comportam de modo similar aos demais representantes desta categoria, exceto que utilizam a boca para falar.

A categoria 3 pode ser delimitada por aqueles personagens que além de utilizarem as expressões faciais e a cabeça para comunicação eventualmente fazem uso de uma ou mais extremidades do corpo para manipular objetos, interagir com outros personagens e gesticular. Este é o caso do Burro e dos sapos em Shrek, dos cachorros e gatos em Pets, dos peixes, polvo e outras criaturas marítimas em Procurando Dory, da cobra em Kung Fu Panda e da zebra em Madagascar. Note-se que existe, entre estes casos algumas diferenças. Em Shrek, os sapos e o Burro precisam de ao menos três patas para manterem-se equilibrados, quando estes personagens gesticulam antropomorficamente com duas mãos há uma justificativa como, o fato de estarem sentados. Enquanto em Procurando Dory, apesar dos personagens constantemente utilizarem as nadadeiras para se manterem flutuando, se há necessidade de que os personagens gesticulem com as duas mãos nenhum outro elemento justifica o equilíbrio dos mesmos. Os cachorros e gatos domesticados de Pets e a cobra de Kung Fu Panda comportam-se como o Burro e os sapos de Shrek, ou seja, o uso de um ou mais membros está condicionado a alguma limitação quanto ao equilíbrio do personagem. A zebra em Madagascar e os animais abandonados em Pets comportam-se de modo mais parecido aos peixes de Procurando Dory, no sentido que não há consideração pela

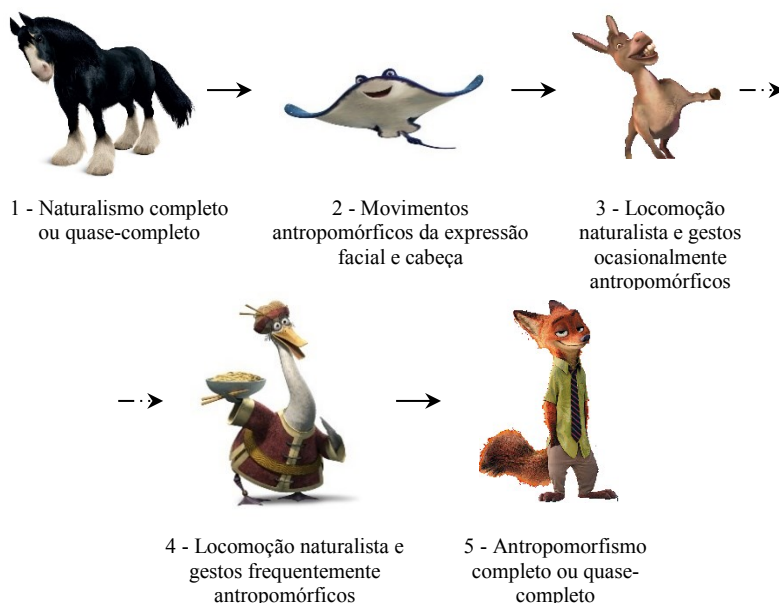
plausibilidade física do equilíbrio caso o uso de um ou mais membros, de modo antropomórfico, se faça necessário à narrativa.

A categoria 4 de antropomorfismo é aquele em que os personagens fazem uso de seus membros para manipular objetos e expressar intenções com grande frequência, em oposição a categoria 3 onde isto ocorre eventualmente. É o caso das aves de Kung Fu Panda, da águia em Pets e dos pinguins em Madagascar. Nestes casos, por serem naturalmente bípedes, os personagens locomovem-se de maneira naturalista, mas têm sempre os membros superiores livres para gesticular e manipular objetos. Nota-se que a águia em Pets tem um comportamento menos antropomórfico que os demais membros do grupo, pois utiliza as asas apenas como forma de gesticular sem utiliza-las para manipular objetos; o carpo da asa é utilizado para representar gestos comuns ao ombro humano. Os demais animais desta categoria utilizam as asas como braços com mãos capazes de manipular com destreza objetos, como seres humanos o fariam. Um caso diferenciado dentro deste grupo é o do louva-a-deus em Madagascar que não caminha de modo naturalista, pois o faz utilizando quatro patas quando, na natureza, animais de sua espécie o fazem com seis; de todo modo tampouco caminha como humano. Tal adaptação faz com que o personagem tenha sempre duas patas que pode utilizar para manipular objetos, ou gesticular, sem perder o equilíbrio, como os demais casos desta categoria.

A categoria 5, é aquela que tem o maior grau de antropomorfismo e é representada pelos personagens vertebrados quadrúpedes de Kung Fu Panda, os personagens civilizados de Zootopia, o coelho em Pets e uma miscelânea de personagens de Madagascar 3. Estes personagens, em sua totalidade quadrúpedes vertebrados com caudas, utilizam as patas traseiras como pernas humanas e as patas dianteiras como braços. Locomovem-se e equilibram-se como bípedes, falam e tem expressões faciais humanizadas, utilizam as patas dianteiras para manipular objetos, interagir com outros personagens e gesticular. Nos casos de Kung Fu Panda e Zootopia os personagens desconhecem o conceito de ser humano, o espectador é livre para se colocar no lugar dos personagens e é plausível supor que o alto grau de antropomorfismo contribua para isso. Em Pets e Madagascar os personagens não só tem ciência da existência de humanos como tem problemas de relacionamento com eles, neste caso é possível supor que o antropomorfismo contribui para que o espectador compreenda as motivações dos animais. Em todos os casos destes grupos, eventualmente, os personagens agem de modo menos antropomórfico e mais animal; as vezes por motivação própria ou outras vezes de modo reativo.

Distinguiram-se, portanto, os personagens das obras analisadas em cinco diferentes categorias, com base nas qualidades expressivas de seus movimentos. Para tornar os conceitos que definem as categorias mais claros e sucintos, cada uma delas foi nomeada de acordo com suas propriedades. Os nomes das categorias são aqui apresentados em uma escala ordinal sendo a primeira categoria aquela com maior grau de naturalismo e menor grau de antropomorfismo e a última categoria aquela com menor grau de naturalismo e maior grau de antropomorfismo: (1) naturalismo completo ou quase-completo, (2) movimentos antropomórficos da expressão facial e cabeça, (3) locomoção naturalista e gestos ocasionalmente antropomórficos, (4) locomoção naturalista e gestos frequentemente antropomórficos, (5) antropomorfismo completo ou quase-completo. A Figura 9 representa graficamente as 5 categorias com exemplos de personagens presentes nas obras analisadas.

Figura 9 – Escala de antropomorfismo de movimento em personagens animais



Fontes: Esquema geral criado pelo próprio o autor (2016), imagens dos personagens da esquerda para a direita de cima para baixo: Andrews e Chapman (2012), Stanton (2016), Adamson *et al.* (2004), Osborne e Stevenson (2008), Howard e Moor (2016).

Com relação a motivação para o uso de um grau maior ou menor de antropomorfismo a presente análise permite-nos apontar algumas tendências:

1. Em animações como Valente, Up e Frozen, onde os seres humanos são personagens principais e os animais coadjuvantes, em geral, os animais estão localizados nos grupos um e dois, exceto a mãe urso por razões justificadas na própria narrativa.
2. Em animações como Zootopia e Kung Fu Panda, onde o conceito de homem não existe, os animais estão localizados em geral na categoria cinco, exceto os selvagens de Zootopia por razões justificadas na própria narrativa.
3. Animações onde animais são os personagens principais, mas convivem com seres humanos, no caso, Pets e Madagascar, apresentaram personagens com graus três, quatro e cinco de antropomorfismo.

Apesar de algumas obras como Up apresentarem artifícios interessantes para comunicar as intenções dos personagens animais, sem conferir a estes movimentos com características demasiado humanas, nota-se que quanto maior a profundidade dos personagens e protagonismo destes na narrativa, maior a necessidade de comunicar as intenções que os motivam e maior a propensão ao uso do antropomorfismo. Ou seja, quanto mais secundário e inexpressivo o papel do animal na trama menor a necessidade, dos expectadores, em entenderem suas motivações e, portanto, menor a necessidade do uso do antropomorfismo ou de outros recursos que possam exercer este papel.

Uma outra consideração a respeito do uso do antropomorfismo é possível a partir da análise da obra Pets. Nesta obra um grupo de animais domesticados tem um grau de antropomorfismo três, enquanto outro grupo, dos animais abandonados, tem integrantes com graus de antropomorfismo variando entre três e cinco. No entanto, todos têm profundidade e papel relevante na trama. É possível sugerir algumas hipóteses para este comportamento: uma delas é a de que como o segundo grupo precisa viver por conta própria e não pode contar com a ajuda humana, o antropomorfismo pode simbolizar uma maior capacidade ou necessidade destes personagens em cuidarem de si próprios; outra hipótese ainda seria o caráter lúdico da motivação destes personagens, no caso, de tramar uma conspiração contra a raça humana.

Em algumas obras como *Shrek 2* e *Madagascar 3* as motivações ou tendências para o uso do antropomorfismo são mais nebulosas e, talvez, menos ligadas a narrativa. No caso de *Shrek 2*, sendo uma obra satírica, a representação de muitos dos personagens da animação visa estabelecer elos com obras famosas realizadas anteriormente. Assim enquanto o Burro tem um grau de antropomorfismo três, o gato de botas é representado em grau cinco, pois, no passado é assim que tal personagem foi representado e, portanto, é assim que é reconhecido pelo público. Já em *Madagascar 3* cada personagem parece ser animado de maneira diversa sem que fosse possível levantar hipóteses coerentes a este respeito.

4.3 APLICABILIDADE DOS RESULTADOS

Como exposto na seção 2.2 (Métodos Computacionais para o Mapeamento de Capturas de Movimentos) existem atualmente duas abordagens para a transposição de movimentos humanos a personagens não-humanos: a primeira utiliza o mapeamento de poses e a segundo utiliza a classificação de movimentos. O mapeamento de poses apresenta, dentre outras limitações a incapacidade de traduzir movimentos em que o personagem tem número maior de extremidades do corpo que o ser humano e a necessidade que o ator mimetize os movimentos animais; já a classificação de movimentos é limitada na incapacidade de classificar ações para as quais não foi explicitamente configurada.

A Tipificação da Representação Animal em Animação apresentada neste capítulo permitiu verificar que o conjunto de métodos de classificação de movimentos é suficiente para a representação de personagens à moda da Categoria 1. Nesta categoria os animais estão limitados a representar um pequeno conjunto de movimentos atrelados ao equilíbrio de seus corpos e ao deslocamento destes pelo espaço, o que é condizente com a limitação dos métodos citados. No entanto, esta técnica é inadequada para a representação de movimentos como gestos antropomórficos, pois estes se mostram altamente diversos o que implicaria na necessidade de um banco de animações tão amplo quanto as ações que se pretenda representar.

As categorias 2, 3 e 4 apresentam movimentos antropomórficos, em maior ou menor grau, ao modo natural de locomoção e equilíbrio dos personagens. O conjunto de trabalhos que trata do mapeamento de poses e mais especificamente o trabalho proposto por (Yamane *et al.*, 2010) parece o mais adequado a este tipo de movimento. Apesar da citada limitação do método, de estar limitado a mapear modelos com número de

extremidades menor ou igual aos de seres humanos, ficou claro no estudo relatado neste capítulo que os movimentos antropomórficos costumam ser representados por um conjunto limitado de membros do personagem. Mesmo em casos onde o personagem possui diversas extremidades móveis apenas a cabeça e as duas extremidades mais próximas a ela são utilizadas, frequentemente, para executar gestos expressivos. Além disso nos casos em que os movimentos são antropomórficos não há, evidentemente, a necessidade de o ator mimetizar os movimentos do animal.

Assim, a luz do estudo exposto neste capítulo, conclui-se que os movimentos dos personagens das categorias 2, 3 e 4 podem ser em parte representados pela classificação de movimentos e em parte pelo mapeamento de poses. O protótipo proposto e detalhado no capítulo 5 apresenta um método para a dissociação de movimentos humanos em dois componentes, um ligado aos movimentos de locomoção e equilíbrio e o outro aos demais movimentos, de modo a possibilitar o mapeamento de cada tipo de ação por meio da técnica mais adequada. Esta dissociação dos movimentos e sua posterior combinação permite a superação de limites encontrados atualmente na representação de personagens das referidas categorias. Para finalizar cabe citar que os personagens da categoria 5 podem ser adequadamente controlados por captura de movimentos utilizando métodos já existentes e detalhados nas seções 2.1 e 2.2.

5 DESENVOLVIMENTO CONCEITUAL E CONSTRUÇÃO DO PROTÓTIPO

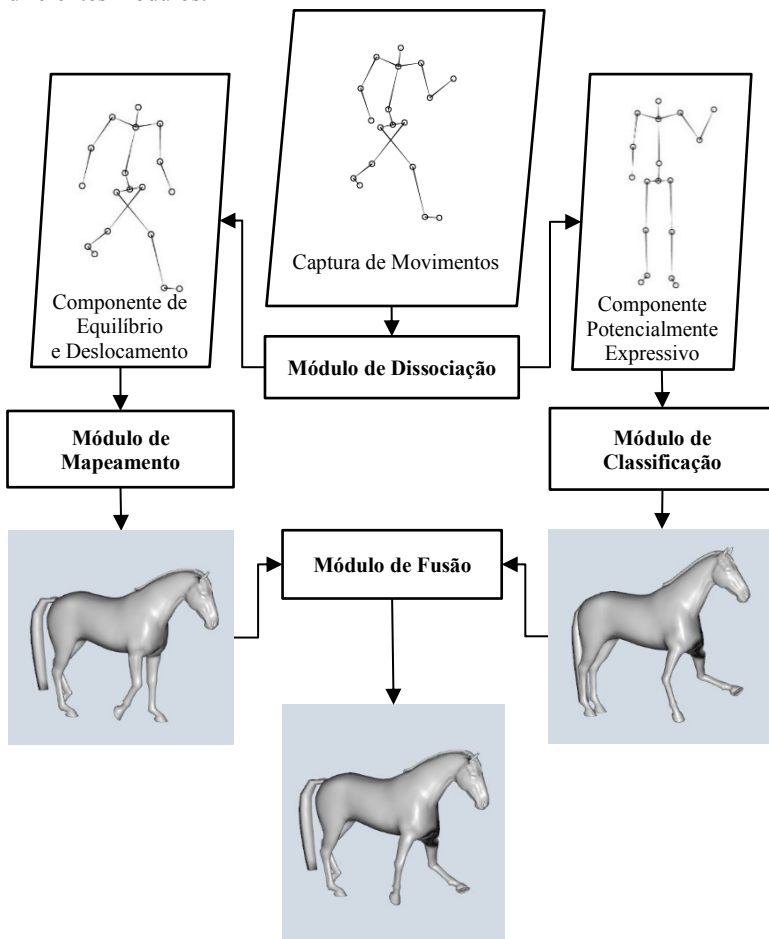
No sentido de atingir o objetivo desta pesquisa propõe-se aqui um *framework* para o mapeamento dos dados de captura de movimentos a estruturas anatômicas animais não-humanas, o qual pode ser implementado na forma de um protótipo composto por um conjunto de técnicas e procedimentos computacionais. O presente capítulo descreve este *framework* e os detalhes de sua implementação. Ainda quanto a este *framework*, o mesmo foi desenvolvido no sentido de permitir a representação dos movimentos corpóreos de personagens nas categorias 2, 3 e 4 da Tipificação da Representação Animal em Animação (TRAA) descrita no capítulo 4. Isto porque, os casos 1 e 5 já são solucionados de maneira satisfatória por métodos existentes (como relatado nas seções 2.1 e 2.2).

Em comparação com métodos anteriormente apresentados na literatura o presente sistema se diferencia por combinar dois métodos de transposição, um por classificação dos movimentos e outro por mapeamento de poses, sendo esta combinação feita em função do tipo de ação desempenhada pelo ator. Tal combinação de métodos de transposição permite ao usuário determinar, para cada personagem, o modo como os movimentos devem ser sintetizados. Quanto às suas limitações o método trata exclusivamente dos movimentos corpóreos do ator capturado, não dando conta de suas expressões faciais. Isto porque além dos dados de movimento do corpo e do rosto serem de naturezas diferentes⁵ a representação da face em personagens animais encontrados nas obras de animação analisadas (no capítulo 4) costumam ser altamente humanizadas.

O *framework* proposto é composto de quatro módulos. O primeiro é o Módulo de Dissociação, responsável por separar os movimentos ligados ao equilíbrio e ao deslocamento dos demais movimentos corpóreos. Os movimentos de equilíbrio e deslocamento são aqueles que compõe a quase totalidade das ações desempenhadas pelos personagens da categoria 1, na TRAA. Já os demais movimentos, por não estarem ligados diretamente ao equilíbrio e ao deslocamento do personagem, podem ter características intencionais e expressivas. São estes gestos que, atuados de modo antropomorfizado, aparecem em grau crescente nas categorias 2, 3, 4 e 5 da TRAA. Enquanto alguns movimentos deste componente, como a movimentação da cabeça não são necessariamente expressivos ou intencionais, outros como a representação de gestos simbólicos com as mãos o são. Assim, denominam-se estes componentes

do movimento de Componente de Equilíbrio e Deslocamento (CED) e Componente Potencialmente Expressivos (CPE). As seções 5.1 e 5.1.1 descrevem em maior detalhe os critérios adotados para distinguir estes dois componentes do movimento humano.

Figura 10 – Esquema gráfico do funcionamento do *framework* proposto e seus diferentes módulos.



Fonte: O autor.

Estes componentes são então tratados por outros dois módulos: o **Módulo de Classificação** que utiliza uma técnica inédita de classificação de movimentos para classificar e transpor os movimentos do CED; e o

Módulo de Mapeamento que utiliza uma técnica inédita de mapeamento de poses para transpor os movimentos do CPE. Por fim, um último módulo, o Módulo de Fusão, combina os resultados do Módulo de Classificação e do Módulo de Mapeamento em uma única figura, em função do grau de antropomorfismo desejado. A Figura 10 traz um esquema gráfico dos módulos do *framework* e da maneira como se relacionam.

Em síntese os Módulos de Dissociação, Classificação e Mapeamento buscam encontrar correlações entre diferentes conjuntos de dados, no caso do Módulo de Dissociação entre o movimento do ator e sua expressividade, no do Módulo de Classificação entre o movimento do ator e uma classificação nominal de tal movimento e no Módulo de Mapeamento entre uma pose do ator e uma pose do personagem. O movimento do ator e suas poses são dados multidimensionais, ou seja, uma série de parâmetros são necessários para descrevê-los completamente. Estes parâmetros podem ter diferentes graus de correlação com aquilo que se busca determinar, por exemplo: o movimento de uma mão tem pouca influência na classificação de um movimento de locomoção, mas muita influência na classificação de um movimento gestual. Assim é possível dizer que entre as diferentes dimensões dos dados e aquilo que buscamos determinar há uma relação não linear. Uma rede neural é um modelo computacional bio-inspirado capaz de “aprender” correlações não-lineares (Haykin, 1999, p. 23) existentes entre estes conjuntos de dados.

Haykin define formalmente o conceito de rede neural do seguinte modo:

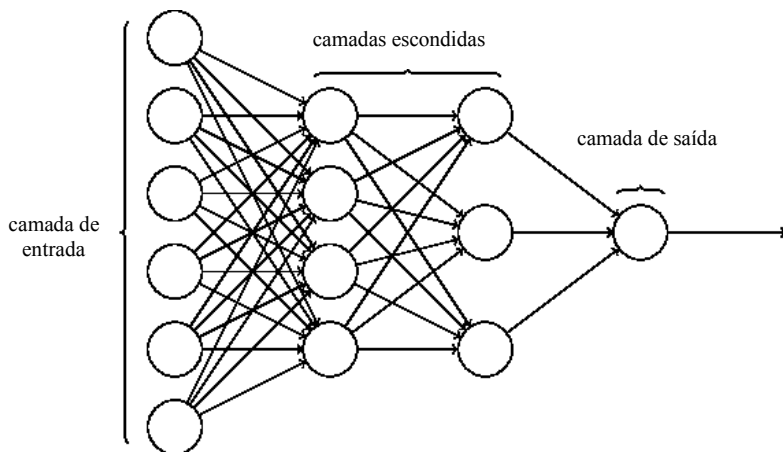
"A rede neural é um processador massivamente paralelizado construído de pequenas unidades [neurônios artificiais], que tem propensão natural a armazenar conhecimento empírico, tornando-o disponível para uso." Haykin (1999, p. 24)

Segundo Kriesel (2007, p.10) a aplicabilidade das redes neurais é derivada do fato destas aprenderem a associar dados, permitindo a aplicação deste aprendizado a problemas similares para os quais a rede não foi explicitamente treinada.

A Figura 11 representa visualmente a topologia de uma rede neural composta de 4 camadas, sendo uma de entrada, duas escondidas e uma de saída. A camada de entrada é aquela onde se inserem os dados conhecidos e a camada de saída é aquela que exhibe os resultados gerados pela rede.

A camada escondida, ou um conjunto de camadas escondidas, em conjunto com a camada de saída compõe a região da rede que efetivamente realiza o processamento dos dados. Os círculos na Figura 11 representam os neurônios artificiais e as linhas que ligam um neurônio a outro são as conexões sinápticas da rede, o processo de aprendizado consiste na modificação dos pesos destas conexões de modo que a rede seja capaz de produzir os resultados desejados (Haykin, 1999, p. 24).

Figura 11 – Topologia de uma rede com duas camadas escondidas.



Fonte: (Nielsen, 2015).

Além dos pesos sinápticos aprendidos pela rede, a maneira como os neurônios são ligados uns aos outros, ou seja, a topologia da rede determina seu funcionamento. De modo simplificado podemos dizer que a rede que possui uma única camada escondida é uma rede rasa, enquanto uma rede com múltiplas camadas escondidas é uma rede profunda (Kriesel, 2007, p.96); no entanto, outras modificações topológicas são possíveis no contexto das redes neurais. Ao longo da descrição da implementação de cada módulo detalham-se as características das redes neurais utilizadas em cada caso.

Nas seções 5.1, 5.2, 5.3 e 5.4 apresenta-se a efetiva implementação de cada um dos módulos do *framework* proposto sob a forma de um protótipo funcional. Para cada módulo são apresentados os trabalhos anteriores relevantes para o contexto, os procedimentos adotados e os resultados obtidos.

5.1 MÓDULO DE DISSOCIAÇÃO

A dissociação de movimentos em subcomponentes é assunto já explorado no contexto da literatura de edição de dados obtidos por captura de movimentos. A premissa é que dissociando o movimento em subcomponentes e recombinação tais componentes torna-se possível a criação de novas animações a partir de dados existentes, sem a necessidade da gravação de novas seções de captura.

Uma abordagem proeminente neste campo é a separação do movimento em dois componentes que Etemad *et al.* (2015) denomina como o Tema Primário (TP) e o Tema Secundário (TS). Segundo o autor o:

“o tema primário [TP] refere-se a ações específicas como andar e correr, enquanto que o tema secundário [TS] se refere a afetações, estilos e características particulares de um indivíduo que interpreta tais ações”
(Etemad *et al.*, 2015, p. 60)

A ideia de separar tais componentes de dados obtidos por capturas de movimentos foi inicialmente introduzida por Amaya (1996) e posteriormente avançada por outros autores (Brand e Hertzmann, 2000; Hsu *et al.*, 2005; Min *et al.*, 2010; Yumer e Mitra, 2016), como discutido na seção 2.2.

Na presente pesquisa propõe-se um método para a dissociação de dados de captura de movimentos diferente dos anteriormente propostos, onde os componentes resultantes são intitulados: (1) Componente de Equilíbrio e Deslocamento (CED) e (2) Componente Potencialmente Expressivo (CPE).

O CED é definido como sendo aquela parte da pose do ator estritamente relacionada a maneira como este se desloca no espaço e mantém o equilíbrio do corpo. Assume-se que a partir de um certo conjunto de propriedades dinâmicas é possível estimar o CED adequado, dada a análise de um largo conjunto de exemplos. Os critérios para a escolha de tais exemplos estão detalhados na seção 5.1.1.

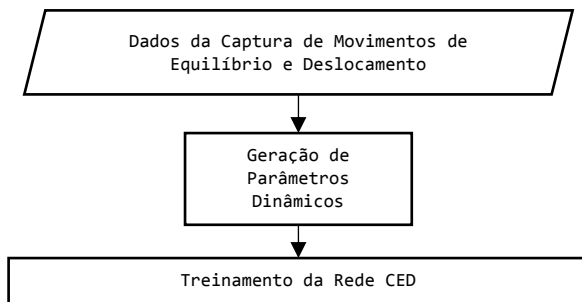
O CPE, por sua vez, é composto por qualquer parte da pose humana que não seja uma função direta de seu deslocamento ou equilíbrio. Os movimentos do CPE incluem, portanto, uma série de movimentos como gestos executados com as mãos, movimentos da cabeça, arqueamento da coluna, e movimentos com os pés. Alguns destes

movimentos podem ser considerados expressivos e, em geral, há uma sobreposição significativa entre os movimentos do CPE e aqueles movimentos antropomórficos encontrados em grau crescente nas categorias 2, 3, 4 e 5 da TRAA.

Em trabalhos anteriores o TP refere-se a uma ação qualquer e o TS a maneira como tal ação é interpretada, o TS contém também características individuais do ator. No caso deste trabalho o CED refere-se aos movimentos que mantêm o corpo em equilíbrio, ou que proporcionam seu deslocamento; enquanto o CPE contempla todos os demais movimentos, dentre os quais muitos são expressivos. Nota-se que é possível que o CED e o CPE contém em si TP e TS; por exemplo: é possível que o CED represente uma caminhada (TP) executada de forma rápida (TS), ou ainda, que o CPE represente a ação desferir socos (TP) executada de modo lento (TS).

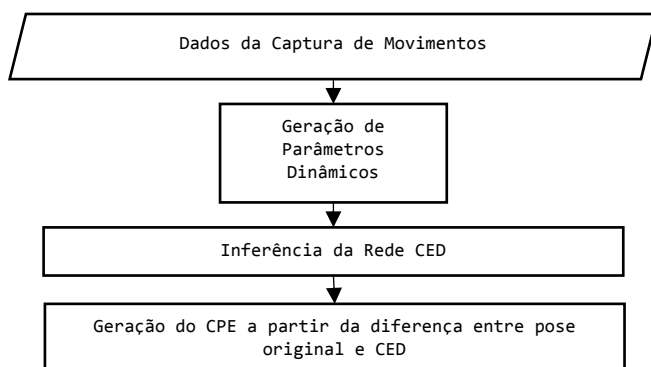
O funcionamento deste Módulo de Dissociação pode ser entendido em duas etapas, uma de treinamento e uma de inferência em tempo real. Durante a etapa de treinamento (Figura 12) uma série de movimentos foram selecionados em função de sua contribuição exclusiva ao componente do CED (ver critérios de seleção na seção 5.1.1), um conjunto de parâmetros dinâmicos foi criado para representar estes movimentos e uma rede neural foi treinada para associar estes parâmetros dinâmicos às poses do ator. Na etapa de inferência em tempo real (Figura 13) os parâmetros dinâmicos são gerados a partir das poses obtidas na captura de movimentos, estes dados alimentam a rede neural que calcula a pose do CED mais representativa para aquele conjunto de dados, finalmente subtrai-se do movimento original os elementos responsáveis pelo equilíbrio e deslocamento o que resulta na síntese dos demais elementos, potencialmente expressivos. Estas duas etapas são discutidas em maior detalhe nas seções 5.1.1 e 5.1.2, por fim apresentamos a aplicabilidade dos resultados obtidos com a implementação deste Módulo de Dissociação na seção 5.1.3.

Figura 12 – Etapa de treinamento do Módulo de Dissociação.



Fonte: O autor.

Figura 13 – Etapa de inferência em tempo real do Módulo de Dissociação



Fonte: O autor.

5.1.1 Treinamento

O Módulo de Dissociação foi implementado por meio do treinamento e inferência de uma rede neural. Como discutido nos procedimentos metodológicos (item 4.2) os dados utilizados são provenientes da base de dados de captura de movimentos mantida pela Carnegie Mellon University. Esta é uma biblioteca extensa, desenvolvida ao longo de vários anos e dotada atualmente de 2605 movimentos diferentes (Cmu, 2017).

Como, na implementação deste módulo, pretendia-se estabelecer a correlação entre um conjunto de parâmetro dinâmicos e as poses do ator exclusivamente no CED foi necessário selecionar, dentre os diversos

movimentos disponíveis na base da CMU, aqueles que atendessem a uma das seguintes definições: (1) movimentos ocasionados pela aplicação de uma força externa ao corpo e a reação do corpo para equilibrar-se a tal perturbação; e (2) movimentos ocasionados pelo deslocamento do corpo no espaço. Foram descartados, portanto, quaisquer gestos que não se adequassem uma destas definições.

Além disso foram descartados movimentos que se adequavam apenas parcialmente às definições. Por exemplo, no caso de uma caminhada (deslocamento) ser realizado de modo concomitante a algum gesto atuado com a cabeça ou com as mãos o movimento seria descartado.

O sistema de captura de movimentos gera dois dados que descrevem a pose do ator: a posição e a orientação de suas juntas. A representação escolhida para a caracterização da pose do ator no treinamento da rede neural foi a orientação das juntas. O uso da orientação, ao contrário da posição, é benéfico por não variar em função do tamanho e das proporções anatômicas do ator.

Como outras bases de captura de movimentos a base CMU consiste de dados exclusivamente cinemáticos, ou seja, descrevem a trajetória do movimento não suas propriedades dinâmicas. No entanto, como definido na seção 5.1 o CED varia em função de parâmetros dinâmicos do movimento. Assim foi necessário calcular, a partir das trajetórias, alguns parâmetros dinâmicos do movimento, como: o vetor de velocidade do quadril do ator no último instante (v_t) e o vetor de velocidade do quadril no último meio segundo ($v_{t'}$). Além disso, outros parâmetros não propriamente dinâmicos se mostraram relevantes na descrição da dinâmica do corpo e desambiguação do CED: o tempo em que os pés estão ou não em contato com o chão, a altura relativa entre o quadril e o chão e a orientação das coxas do ator.

O vetor de velocidade do quadril representa características importantes do movimento como o estar parado ou em deslocamento, além da direção e velocidade deste movimento. A combinação deste parâmetro com um segundo vetor de velocidade capturado em um intervalo de tempo mais prolongado permite a detecção de alterações na direção do movimento, como ocorre em saltos ou viradas.

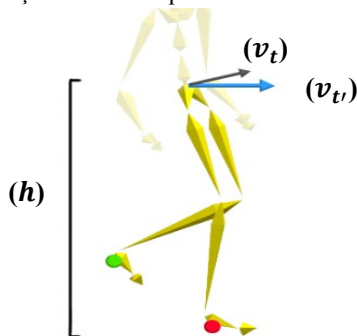
O uso da detecção de contato dos pés contra o solo, facilita a desambiguação de poses corpóreas de outra forma indistinguíveis, como proposto por Holden *et al.* (2016). O detector de contato aqui utilizado baseia-se no trabalho de Lee *et al.* (2002), ou seja, o contato é considerado ativo quando os pés do ator encontram-se abaixo de um certo limite de velocidade e altura. Estendendo o trabalho de Holden *et al.* (2016), verificou-se na presente pesquisa que o uso do registro de tempo em que

os pés estão ou não em contato com o piso colabora positivamente com a distinção de poses similares. Este parâmetro adicional carrega informações implícitas a respeito dos estados anteriores do ator.

A altura entre o piso e o quadril do ator é dita relativa pois é normalizada em função da altura do quadril em um estado neutro. Este parâmetro se mostra importante pois o corpo do ator tende a equilibrar-se de maneira distinta em diferentes situações como, estar em pé, agachado ou na ponta dos pés.

Por fim, são utilizados como parâmetros de desambiguação a orientação das coxas do ator, visto que mesmo dentro de uma certa condição geral de movimento e equilíbrio o ator pode encontrar-se em diferentes estágios do movimento.

Figura 14 – Representação visual dos parâmetros dinâmicos.



Fonte: O autor.

Utilizou-se então uma rede neural para determinar a correlação não linear entre os parâmetros dinâmicos anteriormente descritos à pose do ator, na forma da orientação de juntas de seu corpo. Tanto os parâmetros dinâmicos como a pose do ator podem ser matematicamente entendidos como vetores de muitas dimensões, sendo 23 dimensões no caso dos parâmetros dinâmicos e 124 dimensões no caso das poses do ator.

Uma rede neural é capaz de encontrar as correlações não-lineares existentes entre estes dois vetores, com algum grau de erro. Em testes realizados nesta pesquisa foi possível verificar que erros numericamente pequenos podem resultar em impactos visuais significativos. Uma razão para este fenômeno é a natureza hierárquica da representação esquelética em computação gráfica. Por causa desta forma de representação os erros

se acumulam dos pontos mais altos da hierarquia (como a espinha) até as extremidades (mãos, pés e cabeça) onde mostram-se mais evidentes.

Para superar tal limitação uma representação intermediária da pose do ator foi construída utilizando uma segunda rede neural do tipo *denoise autoencoder*, como proposta por Vincent *et al.* (2008).

Um *denoise autoencoder* consiste de uma rede neural que correlaciona dois conjuntos de dados iguais, exceto pelo fato de que um dos conjuntos é parcial e propositadamente corrompido. Este tipo de treinamento permite que a rede infira o estado mais provável de um conjunto de dados com erros, desde que haja redundância suficiente nos dados.

O *denoise autoencoder* foi treinado com um conjunto de 235 cliques randomicamente selecionados da base CMU. A orientação de 31 juntas do ator, descrevendo sua pose, foram coletadas a 30 quadros por segundo, resultando em 33.531 amostras, sendo cada quadro uma amostra. A orientação de cada junta é descrita por 4 dimensões, o vetor da camada de entrada desta rede possui, portanto, um total de 124 dimensões. Para o treinamento desta rede foi utilizado o algoritmo *Scaled Gradient Descent* como disponibilizado no software MATLAB⁷.

O melhor desempenho alcançado foi um erro quadrático médio⁸ de 0,01301, possuindo a camada escondida do *autoencoder* 300 neurônios. Notou-se que o uso de um número de neurônios na camada escondida superior ao número de neurônios da camada de entrada (23 neurônios) proporcionou os melhores resultados. Comparativamente os melhores erros quadráticos médios obtidos para *autoencoders* com 150, 50 e 15 neurônios foram de, respectivamente: 0,0135; 0,0156 e 0,0196. Tal observação é coerente com os resultados obtidos por Vincent *et al.* (2008, p. 6) e segundo os autores justifica-se no fato de o processo de *denoising* ser capaz de gerar filtros relevantes para descrição dos dados.

Uma segunda rede foi então treinada para encontrar a correlação entre os parâmetros descritos na seção 5.1.1 e a pose do ator, como descrita pelo *autoencoder*. Os referidos parâmetros dinâmicos e, portanto,

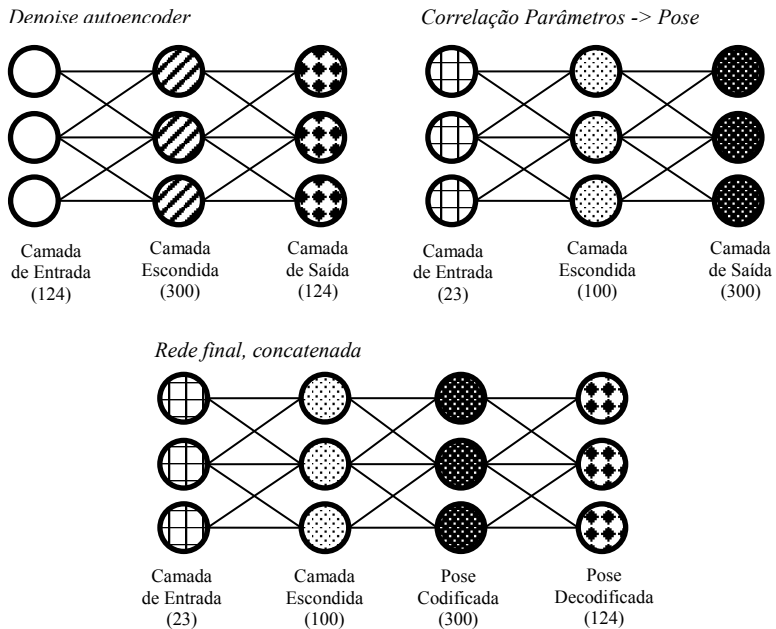
⁷ De acordo com a documentação do software MATLAB (Matlab, 2017), o algoritmo foi implementado em acordo com o método descrito por (Moller, 1993).

⁸ O erro quadrático médio é dado pela soma do quadrado das diferenças entre o resultado esperado e o resultado calculado pelo modelo preditivo, neste caso a rede neural. Esta é uma medida comumente usada para descrever o desempenho de uma rede que faça regressão de dados. Apesar do treinamento de uma rede presumir a minimização de tal erro nem sempre a rede de menor erro será mais adequada a uma dada aplicação; o super-treinamento da rede pode implicar na perda de outra característica desejável, a sua capacidade de generalização.

a camada de entrada desta rede, possuem 29 dimensões. No total 96 cliques de mecânica corporal foram selecionados da base de dados CMU. Os parâmetros dinâmicos de cada clique foram gerados a 60 quadros por segundo resultando em 18.009 amostras, sendo cada quadro uma amostra. Destas amostras 30% foram reservadas para testes e os 70% restantes usados exclusivamente para o treinamento. Assim como no *autoencoder* o algoritmo utilizado foi o *Scaled Gradient Descent*. O melhor desempenho obtido para esta rede apresentou um erro quadrático médio de 0,037.

Concatenou-se a camada de saída da rede tipo *denoise autoencoder* ao final da rede que estabelece a correlação entre os parâmetros dinâmicos e a pose codificada, de modo que a saída final da rede concatenada resulta na orientação das juntas do ator. A topologia da rede resultante desta concatenação está ilustrada na Figura 15.

Figura 15 – Topologia da rede neural profunda para regressão da pose do ator a partir de parâmetros dinâmicos do movimento. Dimensionalidade das camadas entre parênteses.



Fonte: O autor.

5.1.2 Inferência em Tempo Real

A síntese do CED e do CPE são operações que podem ser executadas em tempo real, durante uma seção de captura de movimentos, ou como uma etapa de pós processamento a registros armazenados em um banco de dados. O CED é formado pelo conjunto de dados resultante da aplicação da rede neural descrita na seção 5.1.1 aos parâmetros dinâmicos calculados do modo descrito no mesmo item. Como corolário da definição do CPE apresentada na seção 5.1 pode-se resumir o processo de sua síntese como sendo o cálculo da diferença entre a pose original do ator e o seu CED, gerado pela já citada rede neural.

A representação matemática do cálculo do CPE pode ser feita do seguinte modo:

$$E = O.M^{-1} \quad (1)$$

Onde E são as orientações do corpo (pose) em CPE, O é o movimento original e M são as orientações do corpo (pose) em CED. Em softwares de animação é comum a representação de orientações na forma de matrizes, por isso, o cálculo é feito em álgebra linear onde a diferença é dada pela multiplicação de um elemento pelo inverso do outro.

Para a ocasião deste trabalho os procedimentos necessários à geração dos parâmetros dinâmicos, a inferência do CED pela rede neural e a obtenção de CPE pela diferença entre a pose original e o CED foram implementados utilizando a plataforma de desenvolvimento Fabric Engine (Fabric Software Inc., 2017). Nos testes realizados todos os procedimentos anteriormente citados são calculados em um tempo médio de 1,09 milissegundos⁹ por pose capturada.

5.1.3 Aplicabilidade dos Resultados

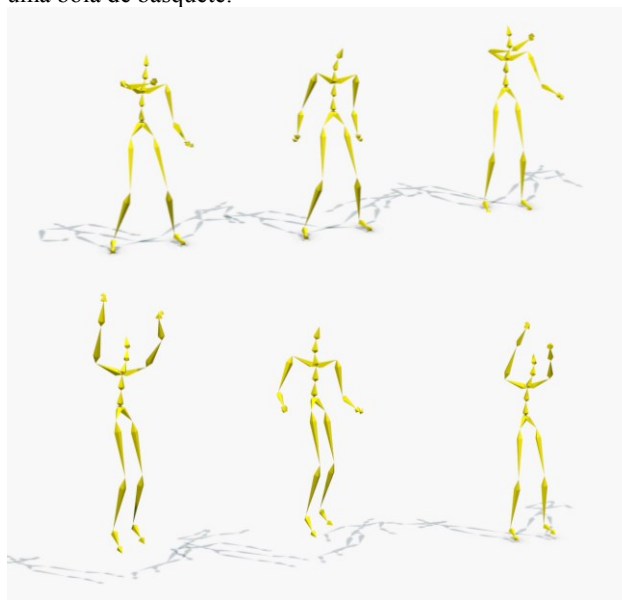
Após a aplicação dos procedimentos descritos resulta dissociado o movimento humano registrado por captura de movimentos. Essa dissociação é caracterizada pelo fato de que um componente carrega movimentos considerados uma função do equilíbrio e do deslocamento, e um segundo componente que carrega os demais elementos do movimento, potencialmente intencionais e expressivos.

⁹ Sendo o processamento executado na CPU de um notebook com processador Intel Core i7-3537U.

O cálculo do CED, é feito por meio de uma rede neural treinada por um amplo conjunto de movimentos capturados de seres humanos com diferentes gêneros, idades e características anatômicas. Assim estes resultados não são específicos de um único sujeito, mas representam o comportamento médio de uma população. No caso, das pessoas que tiveram seus movimentos capturados pelo laboratório de captura de movimentos da Universidade de Carnegie Mellon (Cmu, 2017).

Sendo assim, sabe-se que há uma diferença entre o CED calculado e o CED verdadeiramente existente em cada indivíduo. No entanto, o comportamento visual do CED e do CPE atendem adequadamente às demandas desta tese, ou seja, de dissociar movimentos para posteriormente mapeá-los a estruturas anatômicas não bípedes utilizando técnicas de classificação de movimentos e mapeamento de poses. A Figura 16 apresenta dois exemplos de ações quotidianas registradas por captura de movimentos e dissociadas em CED e CPE pelos procedimentos propostos neste trabalho.

Figura 16 – Dois exemplos de dissociação do movimento. Os esqueletos à esquerda representam o movimento original, os esqueletos ao centro representam o CED e os esqueletos à direita representam o CPE. A imagem superior mostra um sujeito limpando uma janela, a imagem inferior mostra um sujeito lançando uma bola de basquete.



Fonte: O autor.

Além da aplicação primariamente pretendida nesta tese outras aplicações desta técnica são possíveis no campo da animação e da computação gráfica, como a modulação da potencial expressividade em tempo real, a transição assíncrona de movimentos de deslocamento e expressivos e a combinação de componentes de capturas distintas em novas animações. Estas aplicações são apresentadas e discutidas fora do escopo desta tese, em artigo correlato submetido para revisão e publicação (Boehs e Vieira, 2018a).

5.2 MÓDULO DE CLASSIFICAÇÃO

Na última década, equipamentos de baixo custo para a captura de movimentos fomentaram o desenvolvimento de procedimentos para classificação de movimentos humanos com vistas ao controle de personagens humanoides (Ishigaki *et al.*, 2009) e não-humanoides (Seol *et al.*, 2013; Rhodin *et al.*, 2015).

O tipo de dado gerado por tais equipamentos de baixo custo é similar aos dados gerados por equipamentos profissionais para a captura de movimentos. No entanto, o volume útil de captura de tais equipamentos costuma ser muito menor, o que limita a atuação do ator que costuma agir de modo não natural. Ao invés de caminhar, por exemplo, um sujeito deve levantar as pernas simulando uma caminhada, sem sair do lugar para não deixar o volume de captura.

Além disso, os trabalhos existentes na literatura não levam em conta o fato de que uma mesma ação pode ser interpretada por um ator de diferentes modos. De fato, na aplicação pretendida por esta tese, da animação de personagens, é comum que a interpretação do ator seja carregada de trejeitos, maneirismos e particularidades que modificam os dados, caracterizando um contexto complexo para a classificação dos movimentos. Movimentos carregados destas particularidades serão doravante denominados estilizados.

Para a classificação de ações de equilíbrio e deslocamento de seres humanos propõe-se, portanto, um conjunto de procedimentos para classificação de movimentos em um contexto tão natural quanto possível, em que um sujeito possa atuar uma mesma ação nos mais diferentes modos, sejam eles neutros ou estilizados.

Dois redes neurais foram utilizadas para tal tarefa. A primeira com o objetivo de classificar o tipo da ação interpretada e a segunda com o objetivo de determinar um índice da progressão da ação desempenhada em função do movimento do ator. Entende-se por índice de progressão

um valor numérico, variando entre 0 e 1 que indica o estado temporal da ação; sendo 0 o início, 0,5 o meio e 1 o fim.

A Tabela 2 resume a maneira como o sistema proposto se diferencia dos trabalhos de Seol *et al.* (2013) e Rhodin *et al.* (2015), ambos descritos anteriormente no item 2.2.

Tabela 2 – Comparação com trabalhos anteriores cuja a aplicação é o controle de personagens não bípedes por meio da classificação de movimentos.

	Seol	Rhodin	Boehs
<i>Dados</i>			
<i>Atuação não é restrita a pequenos volumes</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Não requer que o usuário se comporte de modo não natural</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Permite ao usuário agir de modo estilizado</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Compatível com equipamento de baixo custo</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<i>Compatível com equipamento profissional</i>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Capacidades</i>			
<i>Não requer configuração</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Detecta ação do usuário</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Detecta progressão da ação</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Detecta amplitude do movimento</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<i>Permite controlar criaturas não bípedes</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Independente do tamanho e proporções do usuário</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Independente do estilo de interpretação do usuário</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Previne deslizamento dos pés do personagem</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<i>Aplicações possíveis</i>			
<i>Jogos eletrônicos</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Pré-Visualização de Animação em Tempo Real</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Animação com qualidade 'final'</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Suporte total , Suporte parcial , suporte inexistente

Fonte: O autor.

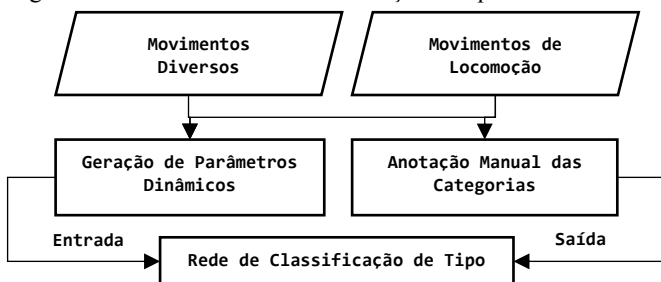
Por se tratar de um sistema que faz uso de redes neurais, esta abordagem pode ser dividida em duas partes: a primeira é o Treinamento e a segunda é a Inferência em Tempo Real. O treinamento pode ser dividido em duas subpartes: o treinamento da classificação de tipos (Figura 17) e o treinamento da regressão do índice de progressão (Figura 18). Entenda-se regressão como a tarefa de encontrar uma relação entre dois componentes de um conjunto de dados (Shalev-Shwartz e Ben-David, 2014, p. 47); neste caso os parâmetros dinâmicos do movimento e o índice de progressão.

O treinamento da classificação é feito com um conjunto de capturas de movimentos de sujeitos interpretando diversas ações como: ficar parado, andar, correr e pular, nos mais diferentes modos. O treinamento da regressão do índice de progressão é realizado com dois conjuntos de capturas de movimentos, um de movimentos estilizados e outro de movimentos neutros. O conjunto de movimentos neutros é constituído por um movimento selecionado arbitrariamente de cada categoria. A regressão do índice de progressão se dá em duas etapas: primeiro determinam-se as correlações entre os dados ditos estilizados e os ditos neutros e então a correlação entre os dados ditos neutros e o valor da progressão.

A inferência em tempo real (Figura 19) ocorre quando parâmetros dinâmicos gerados a partir dos dados de captura de movimentos são processados pelas redes neurais de tipo e progressão. Os valores resultantes dessas operações são filtrados para maximizar a consistência temporal e então são utilizados para controlar o personagem não bípede que se deseje. O controle efetivo do personagem se dá utilizando o tipo de locomoção para carregar uma animação pré-existente de tipo igual ou similar no personagem, e o valor do índice de progressão para determinar o instante da ação a ser reproduzido.

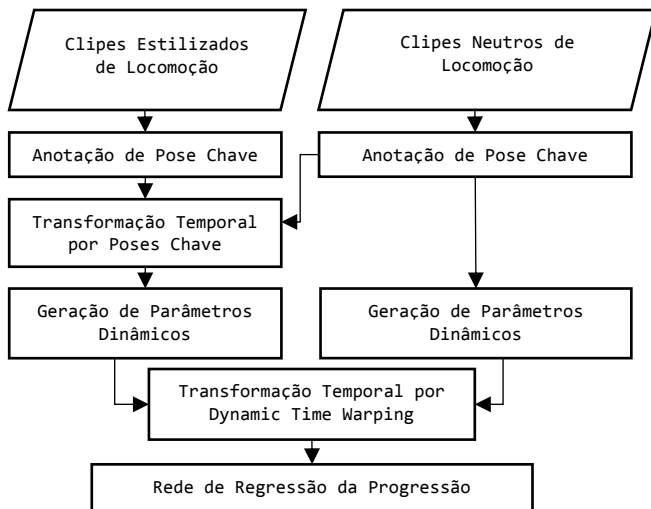
Os parâmetros utilizados para alimentar as redes neurais deste sistema de classificação de movimento são semelhantes aos descritos para o sistema de dissociação do movimento e que estão detalhados na seção 5.1.1. A única diferença entre os parâmetros deste e daquele componente é que neste são utilizadas as orientações de todos os membros inferiores do ator, enquanto naquele são utilizados apenas os das coxas. Isto se justifica, pois, a classificação precisa do índice de progressão do movimento se beneficia destes dados adicionais.

Figura 17 – Treinamento da classificação de tipo



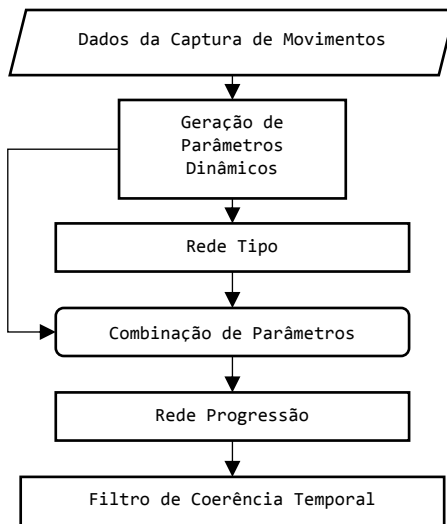
Fonte: O autor.

Figura 18 – Treinamento da regressão do índice de progressão



Fonte: O autor.

Figura 19 – Inferência em tempo real de tipo e progressão



Fonte: O autor.

5.2.1 Treinamento

5.2.1.1 Tipo de Locomoção

Para classificar diferentes tipos de locomoção uma rede neural foi treinada para encontrar a correlação entre os parâmetros dinâmicos anteriormente descritos (itens 5.1.1 e 5.2) e as anotações do tipo de movimento. A anotação consiste de um vetor com N dimensões, onde N é o total de tipos locomotivos a serem classificados. A rede foi treinada para classificar 7 tipos de locomoção: parado, caminhando, caminhando de costas, passo à esquerda, passo à direita, correndo, e pulando para frente.

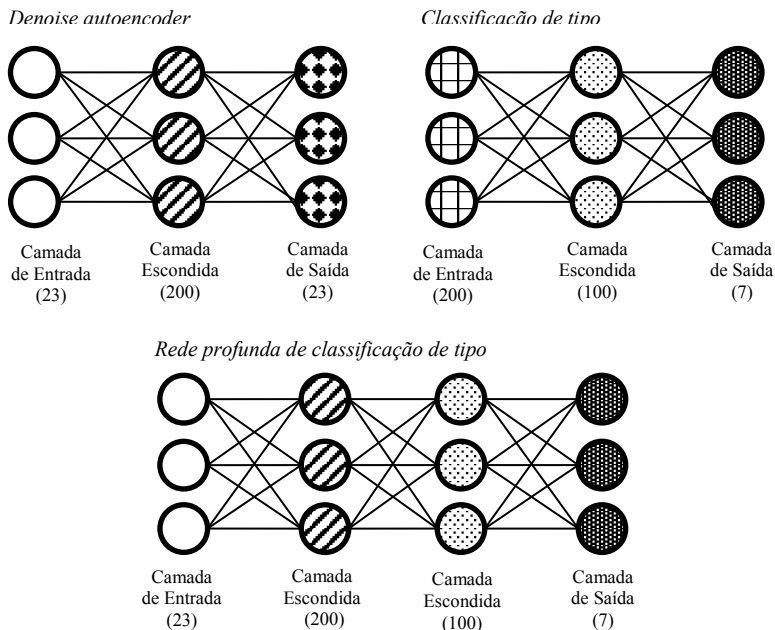
Neste caso, primeiramente, um *denoise autoencoder* foi treinado com os parâmetros dinâmicos de diversos atores, interpretando variadas ações, inclusive ações para as quais o classificador não foi treinado. Estes parâmetros totalizam 23 entradas na rede. Foram utilizados 227 clipes de captura randomicamente selecionados da base CMU (Cmu, 2017), totalizando 33.531 amostras, sendo uma amostra por quadro a 30 quadros por segundo. De modo similar ao *denoise autoencoder* treinado para o Módulo de Dissociação (detalhado anteriormente no item 5.1.1) os melhores desempenhos foram obtidos com um maior número de dimensões na camada escondida, sendo que o melhor desempenho um erro quadrático médio de 0,0124, para uma camada escondida com 200 neurônios.

Uma segunda rede foi treinada então para estabelecer uma correlação entre os parâmetros dinâmicos codificados pelo *denoise autoencoder* e o vetor de anotação dos tipos de movimento. Para esta segunda rede foram utilizados 100 clipes anotados quanto ao tipo de movimento, totalizando 19.621 amostras, sendo uma amostra por quadro a 60 quadros por segundo, destas 30% foram reservadas exclusivamente para testes.

Finalmente, as camadas de entrada e escondida do *denoise autoencoder* foram concatenadas às camadas escondida e de saída da segunda rede. O treinamento foi executado utilizando o algoritmo *Scale Gradient Descent* como disponibilizado no software MATLAB. O melhor desempenho alcançado para a rede foi um erro de entropia-cruzada¹⁰ de 0.00039.

¹⁰ O erro de entropia cruzada é comumente utilizado como medida de desempenho em redes neurais cuja o objetivo seja a classificação de um conjunto de dados. A função penaliza severamente classificações extremamente inaccuradas e penaliza menos severamente classificações pouco inaccuradas (Mathworks, 2013).

Figura 20 – Topologia da rede neural profunda para a classificação de tipos locomotivos. Dimensionalidade das camadas entre parênteses.



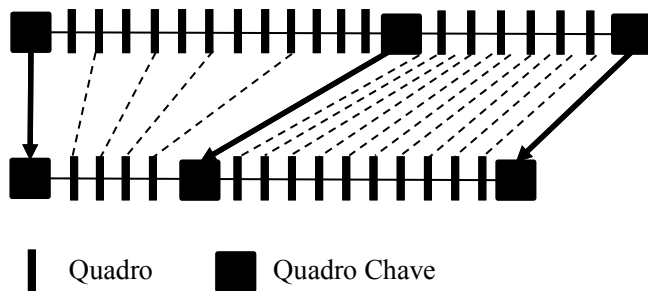
Fonte: O autor.

5.2.1.2 Índice de Progressão

Para treinar a rede e capacitá-la a inferir o valor do índice de progressão, 76 cliques foram anotados quanto aos tipos locomotivos neles representados. O critério para a seleção destes cliques foi a representação do maior número possível de estilos diferenciados da representação de uma ação. Então, cada um destes cliques foi anotado quanto a seu índice de progressão. Progressão é um número real p ($0 \geq p \leq 1$), $p = 0$ determina o início de um movimento, $p = 0,5$ determina a metade do movimento e $p = 1$ determina o fim de um movimento. Algumas ações locomotivas são perfeitamente cíclicas, como a caminhada, neste caso início e fim devem ser determinados arbitrariamente. Outras, no entanto, como ‘pular’, tem um início e fim bem definido. O estado ‘parado’ representa a ausência de locomoção e, portanto, neste caso o valor de p será sempre igual a zero.

Para anotar cada quadro com seu valor próprio de p um processo semi-automatizado foi construído. Para cada tipo de movimento foi criado um conjunto de poses chave, claramente identificáveis. Na caminhada, por exemplo, estas poses eram três: (1) o momento em que o pé esquerdo toca o chão, (2) o momento em que o pé direito toca o chão e (3) o momento em que o pé esquerdo está prestes a tocar o chão novamente. Para cada tipo de ação um dos cliques, desprovido de estilização, foi selecionado como sendo um padrão. Os cliques de movimentos estilizados foram então transformados de modo que o tempo entre suas poses-chave coincidissem com aqueles do clipe padrão para o mesmo tipo de movimento, como mostra a Figura 21. Uma segunda transformação temporal foi aplicada aos cliques, com vistas a estabelecer a correlação entre os quadros não chave. Para tanto foi utilizada a técnica de *Dinamic Time Warping* (DTW) que compara dois sinais de diferentes minimizando a distância entre correspondências encontradas entre tais sinais (Sakoe e Chiba, 1990). Para o propósito deste trabalho foi utilizado o algoritmo de DTW disponível no software MATLAB.

Figura 21 – Representação esquemática da transformação temporal baseada em quadros chave.



Fonte: O autor.

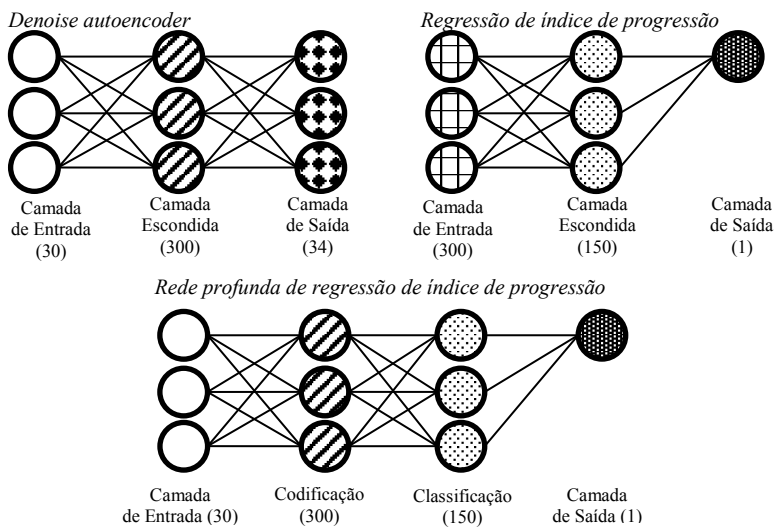
Uma rede neural foi então criada para estabelecer a correlação entre os parâmetros dinâmicos dos cliques estilizados e temporalmente transformados aos parâmetros dinâmicos dos cliques padrão. No total 76 cliques estilizados e 7 cliques padrão foram utilizados, totalizando 4.583 amostras; sendo que 30% das amostras foram separadas para testes. A Figura 22 apresenta o resultado do processamento de um movimento analisado por esta rede. Nota-se que a pose original da caminhada do indivíduo é conformada a um formato desprovido de estilização.

Figura 22 – À esquerda o esqueleto do sujeito capturado e os vetores de velocidade do quadril. À direita o resultado da camada da rede neural de destilização do movimento.



Fonte: O autor.

Figura 23 – Topologia da rede profunda para a regressão da fase do movimento. Dimensionalidade das camadas entre parênteses.



Fonte: O autor.

Uma segunda rede foi treinada para estabelecer a correlação entre os cliques padrão e as anotações de progressão. Finalmente as camadas de entrada e escondida da primeira rede foram concatenadas as camadas escondida e de saída da segunda rede, formando uma rede capaz de fazer a regressão de fase a partir de dados de movimentos estilizados. A Figura 23 ilustra a topologia final da rede. A rede foi treinada com o algoritmo *Scale Gradient Descent* e o melhor desempenho obtido foi um erro quadrático médio de 0,0017.

5.2.2 Inferência em Tempo Real

Após o treinamento das redes de tipo e progressão a classificação do movimento pode ser feita em tempo real. Os dados de captura de movimentos são processados de modo a gerar os parâmetros dinâmicos detalhados nas seções 5.1.1 e 5.2. O vetor resultante é utilizado para alimentar a rede de classificação de tipos. O vetor resultante da rede de classificação de tipo é concatenado ao primeiro vetor e alimentado à rede de regressão do índice de progressão.

Como o tipo e a progressão são classificados quadro a quadro e não é possível saber os estados futuros do ator, os resultados das redes apresentam ruídos e descontinuidades. Para compensar este comportamento indesejado são utilizadas técnicas para filtrar as estimativas de tipo e progressão.

O vetor tipo é filtrado do seguinte modo: primeiro atribui-se à dimensão vencedora o valor 1 e as demais dimensões o valor zero; na sequência interpolam-se o vetor do quadro atual com o do quadro anterior do seguinte modo:

$$\vec{m}_t = (\vec{m}_t * (1 - w)) + (\vec{m}_{t-1} * w) \quad (2)$$

Onde \vec{m} é o vetor da classificação de movimento e w é o peso de interpolação. Apesar do fato de que a filtragem implica em atraso na classificação, nos testes realizados não foram notados problemas visualmente significativos usando um valor baixo para w . No caso dos exemplos apresentados na seção 5.2.3 o valor usado para w foi de 0,016.

Para realizar a estabilização do valor da progressão utilizou-se um filtro similar ao proposto por Rhodin et al. (Rhodin *et al.*, 2015). O objetivo de tal filtro é estabilizar a taxa de variação da progressão e não o valor da progressão diretamente. Para tanto, o funcionamento do filtro pode ser dividido em duas etapas, na primeira estima-se qual deveria ser o valor do índice de progressão para o instante atual, caso a taxa de

variação do instante anterior fosse mantida constante. Esse valor é calculado do seguinte modo:

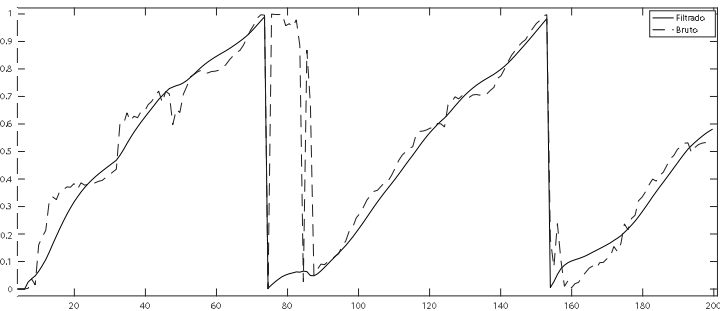
$$p'_t = 2p_{t-1} - p_{t-2} \quad (3)$$

Neste caso p' refere-se a progressão que se quer estimar, que é calculada com base na variação entre as progressões estimadas pela rede neural nos instantes $t - 1$ e $t - 2$. Na sequência, interpolam-se o valor de progressão estimado pela equação 3 e o valor inferido pela rede neural, do seguinte modo:

$$p''_t = (p_t * (1 - u)) + (p'_t * u) \quad (4)$$

Onde u representa o peso da interpolação e p'' o valor final da progressão, após o processo de filtragem. O comportamento do filtro em função da variável u é o seguinte: um valor alto desta variável minimiza a descontinuidade temporal enquanto um valor baixo maximiza a sincronicidade. A Figura 24 compara valores filtrados e não filtrados utilizando um valor de $u = 0,9$.

Figura 24 – Comparação entre valores de fase filtrados e não filtrados na classificação de um ciclo de caminhada para um valor $u = 0,9$.



Fonte: O autor.

Todo o processo de classificação de tipos e progressão leva aproximadamente 2,2 milissegundos por quadro¹¹. A extração dos parâmetros dinâmicos é executada em aproximadamente 0,2

¹¹ Sendo o processamento executado na CPU de um notebook com processador Intel Core i7-3537U.

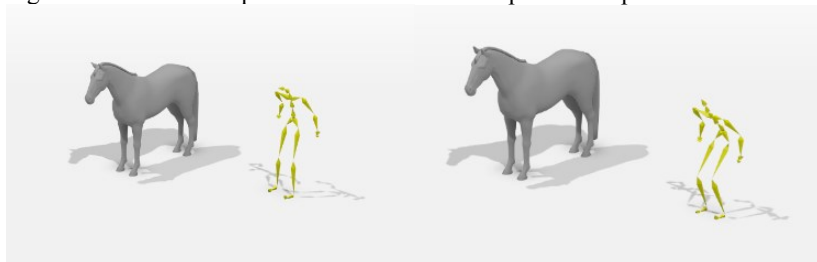
milissegundos; a simulação das redes neurais é executada em aproximadamente 1,5 milissegundos e, os filtros para estabilização temporal dos resultados são calculados em aproximadamente 0,5 milissegundos. Tal desempenho permite a execução de todos os processos em tempo real.

5.2.3 Aplicabilidade dos Resultados

As redes de tipo e progressão foram treinadas para classificar sete diferentes tipos de locomoção: parado em pé, andando para frente, andando de costas, passos à esquerda, passos à direita, correndo e pulando para frente. Para que seja possível o mapeamento dos movimentos e controle de um personagem virtual é necessário que tal personagem disponha de uma biblioteca de movimentos que contemple todas as ações classificáveis. Além disso, é necessário convencionar uma correspondência entre a anotação de progressão utilizada no treinamento da rede neural e na animação do personagem. Assim, para os casos mostrados a seguir foi utilizado o modelo de um cavalo. Todas as animações deste modelo foram criadas de modo coerente com as anotações de progressão do treinamento da rede. Isto é, no caso da caminhada a animação do cavalo inicia-se com a pata frontal esquerda tocando o chão e finda no instante imediatamente anterior a um novo passo com tal pata; seguindo, portanto, os mesmos critérios de anotação da caminha descritos na seção 5.2.1.2.

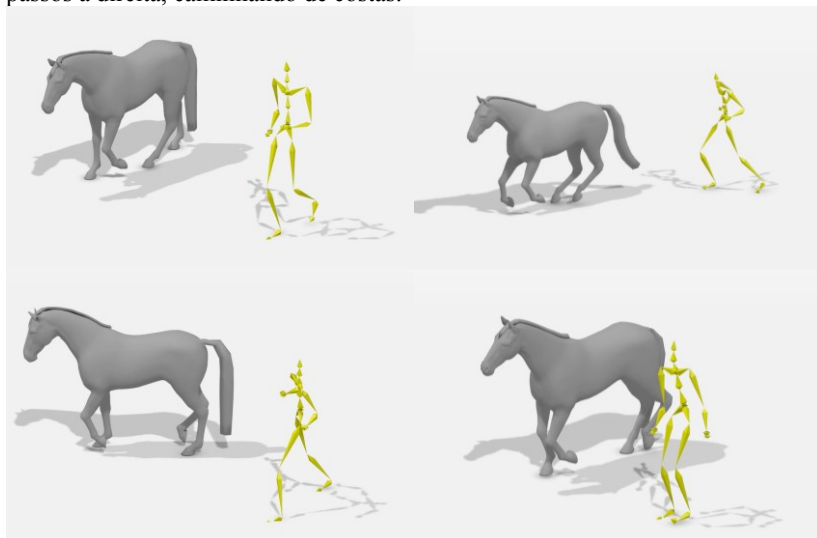
Como resultado dos diversos movimentos estilizados e não estilizados utilizados na etapa de treinamento, as redes apresentam grande robustez a variações de interpretação nos diferentes modos de locomoção. No caso da não locomoção, ou seja, da ação ‘parado em pé’ nota-se que a classificação se mantém estável mesmo quando o ator se mexe sem sair do lugar, enquanto alterna o pé de apoio ou faz outros pequenos movimentos que alteram o equilíbrio do corpo, como mostra a Figura 25. No caso dos demais tipos de locomoção nota-se a robustez do método proposto a variações na direção do movimento e a diversidade de estilos de interpretação, como mostra a Figura 26.

Figura 25 – Diferentes poses classificadas como ‘parado em pé’.



Fonte: O autor.

Figura 26 – A partir do topo esquerdo no sentido horário: caminhando, correndo, passos à direita, caminhando de costas.



Fonte: O autor.

Para comparar os resultados do método aqui apresentado com a classificação de tipos por SVM e a inferência de fase por *Nearest Neighbors* (NN)¹² proposta por Seol *et al.* (2013) dois conjuntos de movimentos, um de (1) neutros e outro de (2) estilizados, foram selecionados para testes. A Tabela 3 e a Tabela 4 mostram os erros de classificação apresentados pelos dois métodos no reconhecimento de diferentes movimentos, neutros e estilizados, respectivamente. A Tabela

¹² Algoritmo que procura a menor distância entre o dado de entrada e um conjunto de dados pré-determinado.

5 mostra uma matriz de classificação para todos os quadros de todos os cliques testados. A Figura 27 e a Figura 28 mostram respectivamente a inferência de fase realizadas pelos dois métodos em um caso de caminhada neutra e outro de caminhada com tropeços.

Apesar de o método de Seol *et al.* (2013) apresentar bons resultados quando as ações são interpretadas de modo neutro, o método aqui descrito se mostrou mais adequado a interpretação de caminhadas neutras com mudanças de direção. Além disso, no caso de ações estilizadas, o método aqui proposto se mostra significativamente superior ao de Seol *et al.* (2013). Tanto em casos de movimentos neutros como em casos de movimentos estilizados os valores de progressão inferidos pelo método proposto apresentam maior coerência temporal.

Tabela 3 – Comparação do erro entre o método proposto e o método de Seol *et al.* (2013) na classificação da locomoção interpretada de modo neutro. Onde ‘Este’ representa o presente trabalho.

Erro por Clip – Cliques Neutros		
<i>Clipe</i>	<i>SVM</i>	<i>Este</i>
Caminhada em linha reta 1	0,0%	0,0%
Caminhada em linha reta 2	0,0%	0,0%
Caminhada em 180º	4,7%	0,0%
Caminhada em forma de 8	25,1%	0,0%
Corrida em linha reta 1	0,0%	0,0%
Corrida em linha reta 2	0,0%	0,0%
Corrida em 180º	0,0%	0,0%
Corrida em forma de 8	0,0%	0,0%

Fonte: O autor.

Tabela 4 – Comparação do erro entre o método proposto e o método de Seol *et al.* (2013) na classificação da locomoção interpretada de modo estilizado. Onde ‘Este’ representa o presente trabalho.

Erro por Clip – Cliques Estilizados		
<i>Clipe</i>	<i>SVM</i>	<i>Este</i>
Caminhada feliz	74,8%	0,0%
Caminhada deprimido	0,0%	0,0%
Caminhada com tropeços	35,4%	0,0%
Caminhada sorrateira	100,0%	0,0%
Corrida de basquete	11,1%	0,0%
Corrida com passadas largas	72,7%	0,0%
Corrida com passos cruzados	61,5%	3,7%
Corrida agachado	3,5%	0,0%

Fonte: O autor.

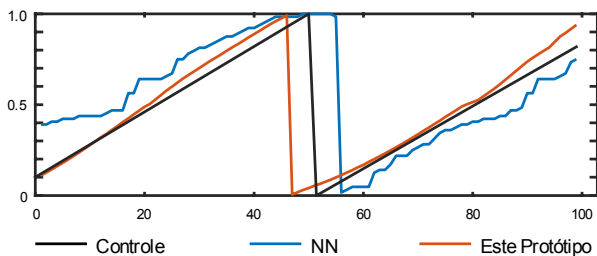
Tabela 5 – Matriz de classificação de tipos para todos os quadros de todos os cliques testados. Onde ‘Este’ representa o presente trabalho.

Classificação de Tipo a Cada Quadro – SVM				
Tipo Real	Tipo Classificado			
		Parado em Pé	Caminhando	Correndo
	Caminhando	9 - 1%	715 - 73%	258 - 26%
Correndo	5 - 1%	62 - 9%	598 - 90%	

Classificação de Tipo a Cada Quadro – Este				
Tipo Real	Tipo Classificado			
		Parado em Pé	Caminhando	Correndo
	Caminhando	0	990 - 100%	0
Correndo	3 - 0,4%	0	744 - 99,6%	

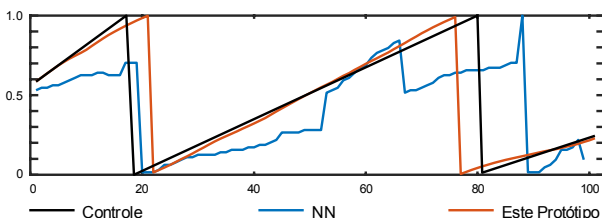
Fonte: O autor.

Figura 27 – Regressão de fase para uma caminhada neutra. ‘Controle’ representa os valores classificados por um especialista, ‘NN’ os valores calculados com o método Nearest Neighbor e ‘Este Protótipo’ o valor calculado pelo método desenvolvido nesta pesquisa.



Fonte: O autor.

Figura 28 – Regressão de progressão de uma caminhada com tropeços (estilizada). ‘Controle’ representa os valores classificados por um especialista, ‘NN’ os valores calculados com o método Nearest Neighbor e ‘Este Protótipo’ o valor calculado pelo método desenvolvido nesta pesquisa.

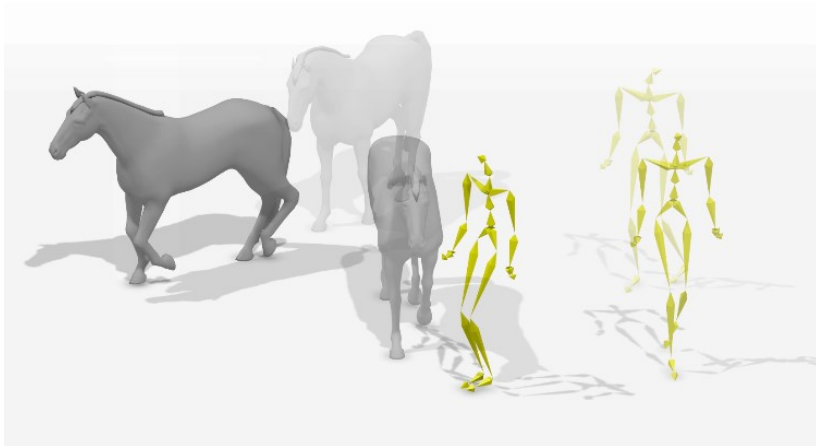


Fonte: O autor.

Além da inédita capacidade de classificação de movimentos interpretados de modo estilizado, o método proposto apresenta propriedades benéficas à aplicação pretendida no fenômeno da concatenação de atividades de locomoção. Combinar um movimento locomotor com seus antecessores e sucessores pode ser uma tarefa complicada visto que existe a necessidade de que as poses de transição entre os movimentos sejam visualmente semelhantes. Para solucionar tal restrição diferentes métodos sofisticados de transição foram anteriormente propostos, sendo um dos exemplos mais significativos neste campo uma técnica conhecida como *motion graphs* (Kovar *et al.*, 2002; Chiu, 2007; Kovar *et al.*, 2008; Min e Chai, 2012).

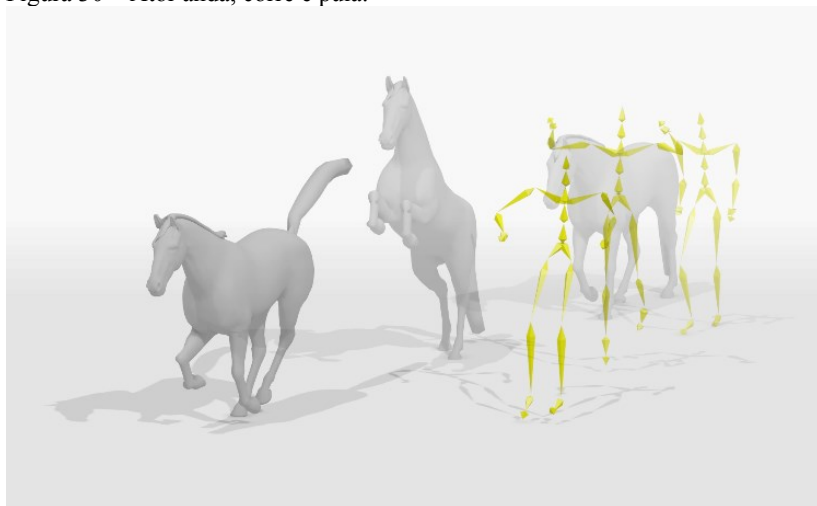
No entanto, no método aqui apresentado a regressão do índice de progressão é feita com base em parâmetros dinâmicos do movimento (aqueles descritos nas seções 5.1.1 e 5.2). Assim, nos testes realizados, notou-se que as progressões inferidas na transição entre uma ação locomotiva e outra costumam ser visualmente similares, como mostram a Figura 29 e a Figura 30.

Figura 29 – Ator move em diferentes direções dentro de um ambiente.



Fonte: O autor.

Figura 30 – Ator anda, corre e pula.



Fonte: O autor.

5.3 MÓDULO DE MAPEAMENTO

O mapeamento de poses entre modelos de anatomias significativamente diferenciadas foi proposto inicialmente por Baran *et al.* (2009). No citado trabalho os autores descrevem um método capaz de estabelecer correlações entre duas geometrias, dado um determinado conjunto de poses chave em cada uma delas. Yamane *et al.* (2010) descrevem um sistema similar ao de Baran *et al.* (2009), no entanto, propõe a correlação entre poses armazenadas no sistema de articulações dos personagens e não em seus modelos geométricos, permitindo assim o uso de dados obtidos por captura de movimentos.

Em oposição a transposição de movimentos por classificação de ações utilizada para os movimentos de equilíbrio e deslocamento e detalhada na seção 5.2, a transposição de movimentos pelo mapeamento de poses é mais propensa a representação de gestos e outros movimentos potencialmente expressivos por não estar restrita a um conjunto limitado de ações possíveis. Neste tipo de abordagem, métodos computacionais são utilizados para determinar a melhor pose no modelo a ser controlado, dado um conjunto de poses exemplo (ou chave) nos modelos de entrada e de saída.

Em comum estes métodos apresentam algumas limitações. Uma delas é que o modelo animado deve conter, no máximo, o mesmo número

de graus de liberdade do ator, ou seja, não é possível controlar todos os membros de um personagem que possua maior número de extremidades que um ser humano. No caso das poses antropomórficas como existentes nas obras de animação analisadas no capítulo 4 esta limitação não é significativa. Isto se deve ao fato de que foi observado que quando os personagens das categorias 2, 3 e 4 da TRAA atuam de forma antropomórfica o fazem comumente por meio da cabeça e dos dois membros superiores mais próximos à cabeça; podendo estes serem patas, nadadeiras, ou braços dependendo da anatomia do personagem.

Outra limitação comum aos métodos de mapeamento de poses, é que o conjunto de poses de exemplo deve ser suficientemente grande, de modo a representar tantas poses quanto se desejem no personagem a ser controlado. No entanto, tais poses são criadas por um animador. É necessário, portanto, acomodar duas demandas que se contrapõe: (1) cobrir um espectro de mapeamento de poses tão grande quanto possível, para que o usuário do sistema possa ser um leigo; e (2) reduzir o número de poses chave necessárias a um número pequeno de poses, para não tornar a utilização da técnica onerosa.

Nos casos citados de Baran *et al.* (2009) e Yamane *et al.* (2010) os autores estão interessados na transposição de movimentos de uma seção de captura de movimentos já existente e conhecida. No entanto, no caso deste protótipo pretende-se permitir ao ator leigo representar movimentos expressivos a priori desconhecidos. A solução encontrada para este problema foi inspirada em um método de mapeamento de movimentos proposto por Rhodin *et al.* (2014).

Rhodin *et al.* (2014) propõe um sistema para o controle de personagens animais animados em jogos eletrônicos a partir da captura de movimentos humanos. Diferentemente de trabalhos anteriores, os referidos autores optaram por criar uma base de dados com diversas animações do personagem animal a ser controlado, ao invés de um pequeno conjunto de poses chave. Estas animações foram utilizadas então para criar um espaço latente da representação de poses no animal, utilizando a técnica estatística de Análise de Componentes Principais (ACP).

Um espaço latente é um espaço composto de um conjunto de variáveis não diretamente observáveis (variáveis latentes), mas que podem ser inferidos a partir de um modelo. A disposição de peças em um jogo de xadrez pode ser usada como um exemplo análogo; a posição das peças pode ser descrita tanto pela configuração observável destas quanto pelo conjunto de movimentos que levaram a atual organização do tabuleiro.

Segundo os autores, esta representação alternativa do movimento contém, implicitamente, informações a respeito das limitações das juntas dos animais e das orientações em que estas se movem Rhodin *et al.* (2014, p. 5). No contexto do referido trabalho isto simplifica o problema da correlação, exigindo um menor número de poses exemplo no modelo humano.

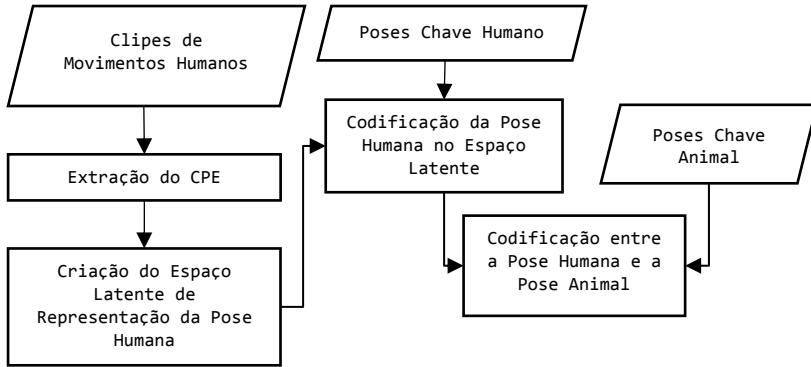
No contexto da presente pesquisa a limitação quanto a disponibilidade de dados é diferente daquela encontrada em Rhodin *et al.* (2014). Aqui presume-se a disponibilidade de uma larga base de dados de movimentos humanos, combinada a meta de reduzir, tanto quanto possível, a criação de poses no modelo animal. Assim um espaço latente das poses humanas foi criado, mais especificamente um espaço latente exclusivo de movimentos do CPE que são aqueles que se objetiva transpor ao modelo animal neste módulo. Estabeleceu-se então a correlação entre poses chave do modelo animal e poses chave do modelo humano, sendo as poses deste último modelo codificadas no referido espaço latente.

O funcionamento do método de mapeamento de movimentos aqui proposto pode ser compreendido em dois momentos distintos, um primeiro do Treinamento e um segundo dá Inferência em Tempo Real.

No treinamento, representado pela Figura 31, os movimentos humanos de um subconjunto da base de dados CMU foram dissociados em CPE e CED pelo método proposto na seção 5.1; o subcomponente expressivo dos movimentos foi utilizado para determinar um espaço latente de representação dos movimentos humanos. Na sequência, foi estabelecida a correlação entre um conjunto de poses chave, do ser humano e do personagem a ser animado, utilizando o espaço latente como uma representação intermediária das poses humanas.

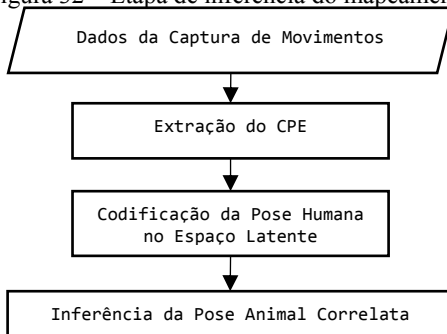
Na inferência em tempo real, representada pela Figura 32, o movimento obtido por captura de movimentos é dissociado pelo método proposto na seção 5.1, o CPE das poses humanas é codificado no espaço latente criado durante a fase de treinamento e, utilizando as correlações estabelecidas na fase de treinamento gera-se a pose no modelo a ser controlado. Os itens 5.3.1 e 5.3.2 descrevem em maior detalhe as etapas de treinamento e inferência. Os resultados da aplicação dos procedimentos propostos são exibidos na seção 5.3.3.

Figura 31 – Etapa de treinamento do mapeamento de poses expressivas



Fonte: O autor.

Figura 32 – Etapa de inferência do mapeamento de poses expressivas



Fonte: O autor.

5.3.1 Treinamento

A etapa de treinamento do módulo de correlação de poses iniciou-se com a criação de um espaço latente para a representação de poses humanas expressivas. A criação desta representação intermediária do movimento se deu pelo fato de que ela permite explorar um grande conjunto de dados não categorizados para inferir implicitamente as variações admissíveis na pose do modelo (Rhodin *et al.*, 2014, p. 5). Diversos modelos, como a já referida Análise de Componentes Principais (ACP), são capazes de gerar um espaço latente a partir de um determinado conjunto de dados. Redes neurais do tipo *autoencoder* também possuem esta propriedade. A camada escondida de uma rede neural tipo

autoencoder pode ser compreendida como o conjunto de variáveis latentes que codificam os dados que se busca representar.

Assim o presente trabalho utilizou uma rede neural do tipo *denoise autoencoder* para criar um espaço latente a partir de um conjunto 223 cliques randomicamente selecionados da base CMU. A orientação das juntas do ator, descrevendo sua pose, foram coletadas a 30 quadros por segundo resultando em 83.167 amostras, sendo cada quadro uma amostra. Cada um dos cliques foi então espelhado no plano sagital, de modo que os movimentos possuíssem igual distribuição nos membros, direito e esquerdo, dos corpos dos sujeitos capturados. Assim o número total de amostras, após o espelhamento, foi de 166.334. Para o treinamento desta rede foi utilizado o algoritmo *Scaled Gradient Descent* como disponibilizado no software MATLAB. O melhor desempenho alcançado foi um erro quadrático médio de 0,00452, para uma camada escondida de 300 neurônios.

Para encontrar a correlação entre as poses humanas e as poses no personagem animal, 34 poses foram selecionadas da base de dados CMU. O critério para seleção de tais poses foi encontrar instâncias qualitativamente diferenciadas do movimento humano, em que as partes do corpo se movimentam em diferentes direções e amplitudes. A lista completa de poses está representada no Apêndice B. Para cada pose humana uma pose animal correlata foi criada, os Apêndices C, D e E mostram as poses criadas para dois personagens animais diferentes.

Uma segunda rede neural foi então treinada para encontrar a correlação entre poses chave do ator, como descritas pelo *autoencoder*, e poses chave do personagem animal, descritas pela orientação das juntas de seu esqueleto virtual. Como o número de poses é reduzido optou-se por utilizar o algoritmo de treinamento de Regressão Bayesiana. Este algoritmo é eficaz em evitar o problema de *overfitting*¹³ (Mathworks, 2017), comumente associado ao treinamento de redes neurais com pequenos conjuntos de dados. Uma rede precisa ser treinada para cada personagem que se pretenda controlar, para a ocasião desta tese foram feitos testes com poses de um modelo quadrúpede (cavalo) e de um modelo cordato (tubarão). Para o modelo do quadrúpede dois conjuntos diferentes de poses foram utilizados, como será melhor discutido na seção 5.3.3. O menor erro quadrático médio obtido no modelo quadrupede foi de 0,0013 para o primeiro conjunto de poses e 0,00106 para o segundo

¹³ O *overfitting* é a condição de treinamento em que uma rede neural apresenta pouca capacidade de generalização.

conjunto de poses. O menor erro quadrático médio obtido no modelo cordato foi de 0,000308.

5.3.2 Inferência em Tempo Real

A inferência das poses animais pode ser feita em tempo real, a partir de dados obtidos por captura de movimentos ou a partir da leitura de dados de captura previamente armazenados em arquivos. Primeiramente a pose humana é dissociada em dois componentes, CED e CPE como descrito na seção 5.1. O CPE é então codificado no espaço latente descrito na seção 5.3.1. O vetor resultante desta operação é utilizado então para alimentar a rede neural que determina a correlação entre poses humanas e as poses do animal, resultando assim mapeada a pose pretendida.

Todo este processo é executado em 3,23 milissegundos¹⁴; sendo 1,09 milissegundos para a dissociação dos componentes do movimento, 1,07 milissegundos para a codificação da pose humana pelo *autoencoder* e, finalmente, 1,07 milissegundos para a inferência da pose animal.

5.3.3 Aplicabilidade dos Resultados

Como descrito na seção 5.3.1, 34 poses chave foram utilizadas para estabelecer a correlação entre o personagem humano e o personagem animal. Foi possível perceber no processo de geração e seleção das poses que a quantidade e natureza destas influenciam qualitativamente a capacidade do método em representar movimentos.

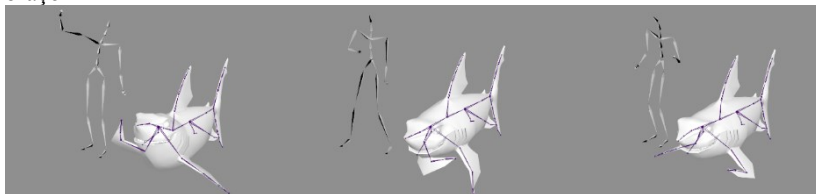
No entanto, é difícil mensurar objetivamente a influência de cada pose no resultado. Isto se deve em parte a natureza de “caixa preta” das redes neurais, que não fornecem informações explícitas a respeito dos componentes que influenciam um resultado específico (Rabunal e Dorado, 2006, p. 17). Além disso o erro quadrático médio, do treinamento de cada rede só é válido para as poses de treinamento, não dando conta do desempenho da rede na maioria dos casos em que componentes de cada uma destas poses chave são interpolados.

As 34 poses chave utilizadas para estabelecer correlação entre poses humanas e poses animais representam instâncias significativamente diferentes dos movimentos humanos. Além da pose de repouso, foram testadas a influência de 3 tipos diferentes de poses: as poses que exemplificam a amplitude de um membro do corpo atuando

¹⁴ Em um notebook com processador Core i7-3537U de dois núcleos físicos.

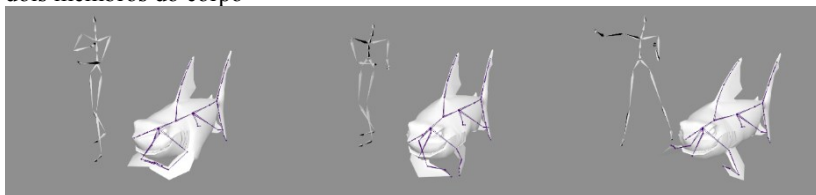
individualmente (Figura 33), as poses que exemplificam movimentos coordenados de dois ou mais membros do corpo (Figura 34) e as poses que exemplificam o contato entre superfícies do mesmo corpo (Figura 35).

Figura 33 – Poses que exemplificam diferentes amplitudes das articulações do braço



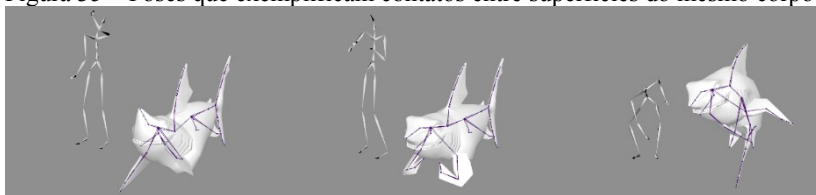
Fonte: O autor

Figura 34 – Poses que exemplificam diferentes movimentos coordenados entre dois membros do corpo



Fonte: O autor

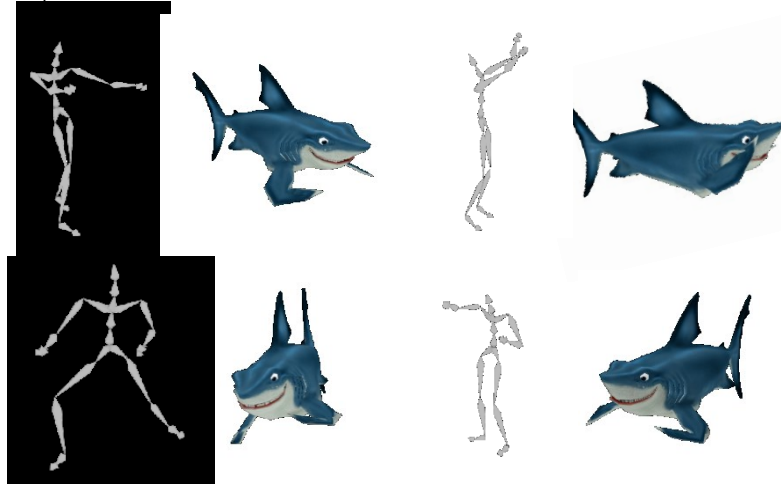
Figura 35 – Poses que exemplificam contatos entre superfícies do mesmo corpo



Fonte: O autor

É possível observar, nos exemplos da Figura 36, que a direção e amplitude dos movimentos do ator são representadas de modo adequado no modelo do personagem animal. Sendo assim tanto em casos em que os braços se movem de modo coordenado, em uma mesma direção, quanto em casos que cada braço se move em uma direção.

Figura 36 – Coletânea de poses transpostas ao personagem do tubarão. De cima para baixo, na primeira linha, exemplos de poses em que os braços se movem na mesma direção, na segunda linha, exemplos em que os braços se movem em direções diferentes.



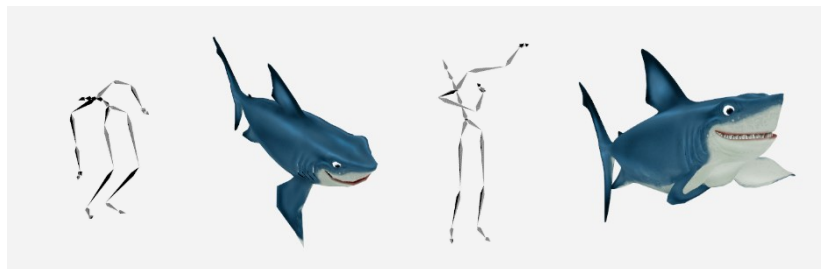
Fonte: O autor.

No entanto, não foi observado um comportamento igualmente adequado nos casos em que o personagem realiza contato entre duas superfícies do seu corpo. Os erros que acontecem neste caso são de dois tipos: um contato não acontece quando deveria acontecer, ou um contato acontece quando não deveria acontecer. A Figura 37 exemplifica estes dois casos no contexto de uma captura de movimentos onde o ator mimetiza as ações de uma pessoa tomando banho e esfregando seu corpo. No caso onde o ator toca o pé com a mão, era esperado, com base nas poses chave criadas, que o tubarão tocasse sua calda com a nadadeira, no entanto, há uma grande distância entre as duas superfícies. Já no caso em que o ator esfrega o peito esperava-se que o modelo do tubarão reproduzisse o movimento com a nadadeira, no entanto, esta atravessa o peito do personagem.

Supõe-se que isto se deva ao menos em parte ao fato de que a rede neural treinada não possui nenhuma informação, ainda que simplificada, a respeito da superfície geométrica propriamente dita. O contato entre superfícies, portanto, não é estimulado ou penalizado durante a fase de treinamento. Em teoria seria possível acrescentar termos adicionais ao treinamento da rede neural, utilizando os dados da camada de saída para

calcular a posição de cada vértice da superfície do personagem e então verificar se há ou não contato entre diferentes pontos da superfície. No entanto, propostas de novos algoritmos de treinamento para redes neurais estão fora do escopo desta tese.

Figura 37 – Casos onde o ator realiza contato entre duas superfícies do seu corpo mal representados no mapeamento de poses.



Fonte: O autor

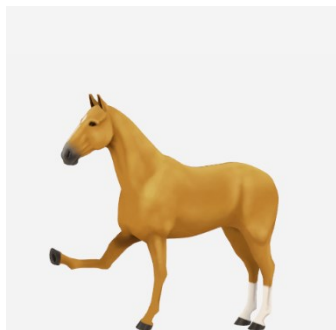
As poses chave escolhidas também são determinantes no tipo de representação de movimentos que se pretende para o modelo animal não humano. Como foi possível verificar durante a tipificação da representação animal proposta no capítulo 4, por vezes, um mesmo tipo anatômico pode ser representado de modos diferentes em obras diferentes. Foi o caso do burro e da zebra nas obras *Shrek* (Adamson *et al.*, 2004) e *Madagascar* (Darnell *et al.*, 2012). No primeiro caso o personagem atua majoritariamente como um quadrúpede, utilizando uma das patas frontais quando necessário para se expressar, e eventualmente utilizando as duas patas frontais. Quando as duas patas frontais são utilizadas o personagem precisa assumir a pose de bípede, evitando a perda de equilíbrio. Já no caso da zebra de *Madagascar* (Darnell *et al.*, 2012) o personagem se move como quadrúpede, mas sempre que precisa gesticular, ainda que com apenas um membro, o faz como um bípede.

O tipo de relação existente entre o movimento do ator e o resultado no personagem é controlado pelas poses chave, que podem ser utilizadas para definir estes modos diferenciados de representação. A Figura 38 compara o resultado do mapeamento de poses utilizando conjuntos de poses chave diferentes em um mesmo personagem quadrúpede, sendo à moda da representação encontrada na obra *Shrek* (Adamson *et al.*, 2004), poses do Apêndice D, e o outro à moda da representação encontrada na obra *Madagascar* (Darnell *et al.*, 2012), poses do Apêndice E.

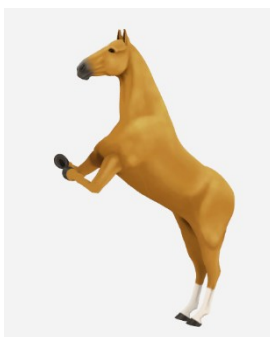
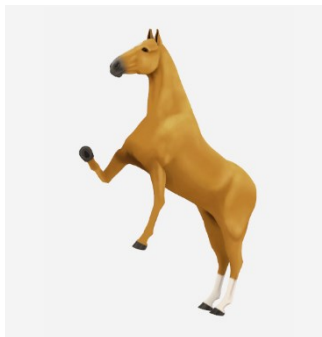
Figura 38 – Mapeamento de movimentos expressivos em dois graus diferenciados de antropomorfismo.



Mapeamento de poses expressivas com poses chave à moda Shrek



Mapeamento de poses expressivas com poses chave à moda Madagascar



Tanto o método aqui proposto, quanto aquele proposto por Yamane *et al.* (2010) presumem que uma regressão não linear é realizada a partir de um conjunto de poses chave em dois modelos anatomicamente diferentes. O principal diferencial do método aqui proposto é que as poses humanas são codificadas por uma representação intermediária, gerada por um *autoencoder* a partir de uma larga base de movimentos humanos. Isto amplia a capacidade de generalização do método aqui proposto. Para comparar estas duas abordagens, do mapeamento de poses com e sem a codificação das mesmas por um *autoencoder*, testes foram realizados em dois contextos: (1) utilizando poses chave geradas a partir de uma cena já conhecida, como ocorre em Yamane *et al.* (2010) (Apêndice F); e (2) utilizando as poses chave apresentadas no Apêndice B em uma cena diferente das quais as poses chave foram coletadas, em função da necessidade de que o protótipo aqui desenvolvido funcione em tempo real.

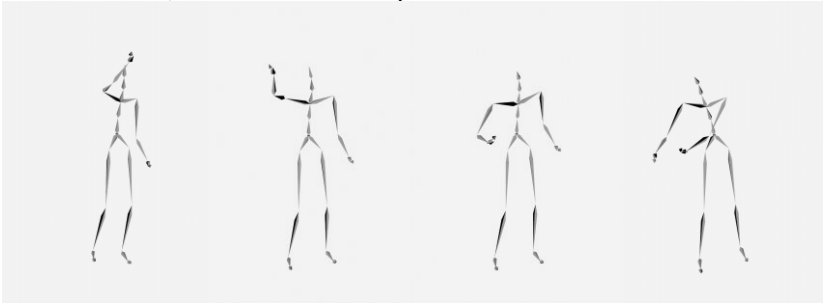
Quando a cena em que se pretende transpor o movimento obtido por captura é conhecida ambas as abordagens, com e sem a codificação da pose humana em um espaço latente, apresentam resultados similares. A Figura 39 apresenta 4 instantes de uma cena onde um ator mimetiza os movimentos de uma pessoa limpando uma janela e, posteriormente, torcendo um pano; e os mapeamentos produzidos com e sem o uso do *autoencoder*. Para a transposição desta animação de 12 segundos 9 poses chave do movimento foram escolhidos e reproduzidos no personagem animal, o que significa que uma pose chave foi criada a cada 1,66 segundos. Esta proporção é comparável ao número de poses chaves utilizadas por Yamane *et al.* (2010). Na ocasião daquela pesquisa, em uma determinada atividade os autores criaram uma pose chave a cada 1,58 segundos.

No entanto, quando as poses chave utilizadas são compostas de um conjunto mais amplo de poses que visam representar um largo espectro de movimentos e, além disso, quando se objetiva transpor movimentos a priori desconhecidos a abordagem que faz uso do *autoencoder* mostrou melhores resultados. A Figura 40 apresenta 4 poses de uma cena onde um ator executa os movimentos de um jogador de baseball. O conjunto de poses chave utilizado para o treinamento das redes foi aquele apresentado no Apêndice B. Vale ressaltar que nenhuma das poses do Apêndice B pertence ao movimento representado na Figura 40.

Portanto, o método de mapeamento de poses aqui proposto se mostra adequado à transposição de movimentos expressivos entre duas anatomias diferentes e a partir de um conjunto de poses de exemplo, em casos onde os movimentos atuados são a priori desconhecidos. Foi

possível transpor adequadamente os movimentos de ações que não constavam das poses exemplo iniciais, exceto nos casos em que há contato entre diferentes pontos da superfície do personagem, o que é uma limitação do método proposto.

Figura 39 – Mapeamento de poses expressiva comparado, com e sem o uso de um *autoencoder*, em caso conhecido a priori.



Mapeamento de poses sem *autoencoder*

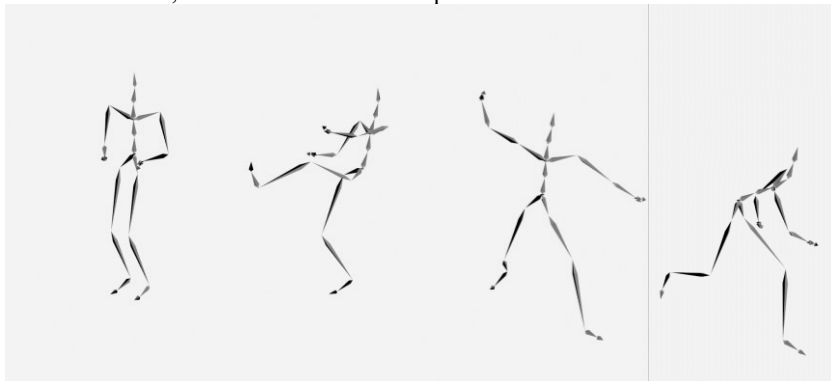


Mapeamento de poses com *autoencoder*

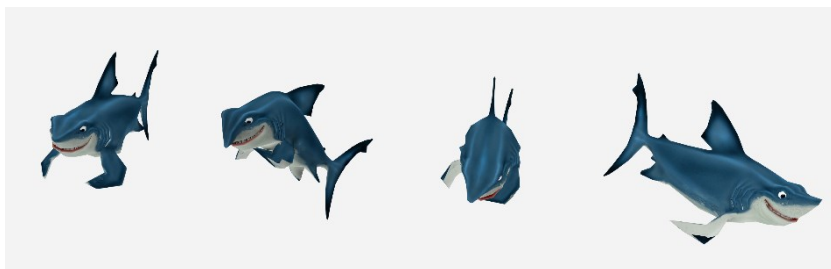


Fonte: O Autor.

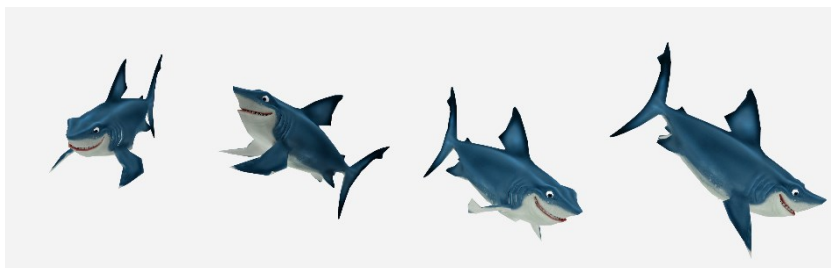
Figura 40 – Mapeamento de poses expressiva comparado, com e sem o uso de um *autoencoder*, em caso desconhecido a priori.



Mapeamento de poses sem *autoencoder*



Mapeamento de poses com *autoencoder*



Fonte: O autor.

Tendo sido descritos separadamente os métodos para a transposição dos movimentos do CED e do CPE cabe agora descrever como estes dois conjuntos de dados são sintetizados na animação final do personagem.

5.4 MÓDULO DE FUSÃO

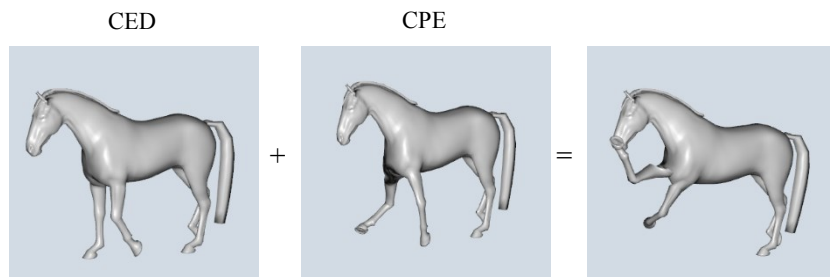
Como corolário do desenho do sistema proposto, que dissocia o movimento humano em dois componentes distintos e faz a transposição destes componentes por meio de técnicas distintas, faz-se necessário, ao final deste processo, fundir os dois componentes nos movimentos de um único personagem animado.

A dissociação do movimento e sua posterior fusão é objetivo de diversos trabalhos, como Amaya *et al.* (1996); Brand e Hertzmann (2000); Hsu *et al.* (2005); Min *et al.* (2010); Yumer e Mitra (2016), anteriormente expostos nas seções 2.2 e 5.1. Além disso a fusão de diferentes “camadas” de movimentos é um recurso comum a alguns softwares de animação quadro-a-quadro como o Autodesk Maya (Autodesk, 2107).

De modo simplificado a fusão dos componentes de animação pode ser de dois tipos: (1) por interpolação ou (2) por adição. Na interpolação a animação final é o resultado da combinação dos componentes em função de um determinado peso α , sendo a influência de um dos componentes o próprio peso α e a do outro seu inverso $(1 - \alpha)$. Na adição, um dos componentes é tomado como base e o outro é adicionado ao primeiro total ou parcialmente em função de um determinado peso β .

A fusão por adição apresentou problemas em alguns casos no contexto do desenvolvimento deste protótipo. Isto porque nestes casos a ação transposta no CED conflita, em partes do modelo, com a pose transposta no CPE. Assim resulta que quando combinada por adição a pose torna-se visualmente incompreensível, como exemplifica o diagrama da Figura 41 onde o movimento de caminhar e o movimento de gesticular com as patas são combinados no modelo do cavalo.

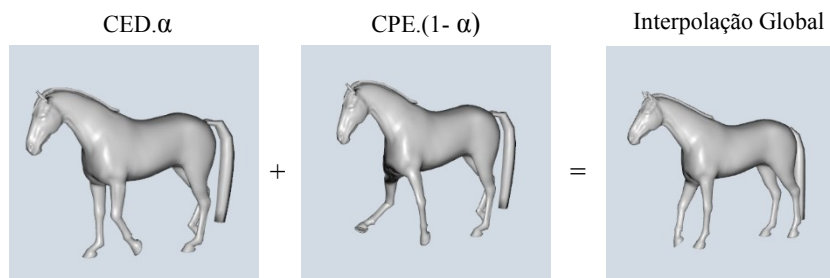
Figura 41 – À esquerda o CED, ao centro o CPE e a direita o resultado da soma dos dois componentes.



Fonte: O autor

Já na fusão por interpolação é necessário determinar o peso adequado da interpolação para que a pose final resulte coerente. Este peso não deve ser aplicado de modo global, visto que nem todas as juntas do CPE possuem informação desejável; de fato se o ator não executa nenhum movimento expressivo espera-se que a influência do CPE e, portanto, o peso da interpolação seja zero. A Figura 42 mostra o que acontece quando CED e CPE são combinados com um peso global de interpolação no valor de 0,5, o que significa que o resultado da pose final sofre 50% de influência do CED e 50% de influência do CPE. Nota-se que nas patas dianteiras, onde há a representação de um gesto o peso 0,5 é insuficiente para que o CPE do movimento seja visível na pose final. Já nas patas traseiras, onde o CPE não possui nenhum gesto relevante, o peso 0,5 acaba alterando indevidamente o CED.

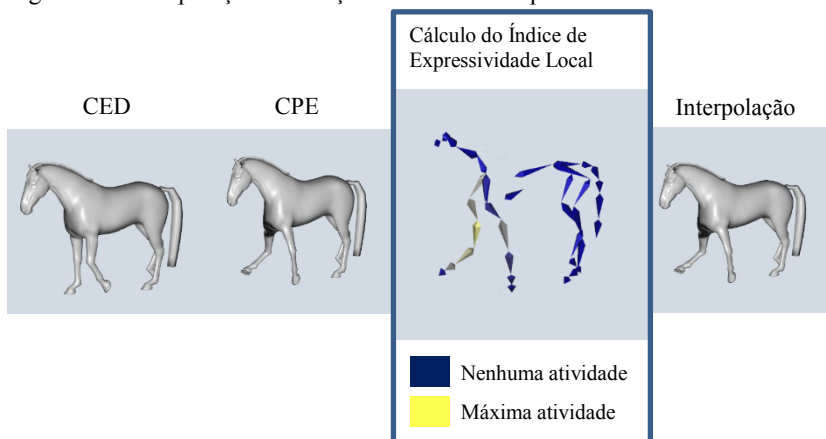
Figura 42 – Interpolação entre poses do CED e do CPE aplicada de modo global, com peso $\alpha = 0,5$ em todas as juntas do corpo.



Fonte: O autor

Assim nos casos em que combinação por adição não era possível e a fusão por interpolação se fez necessário um peso foi calculado a partir de um índice de expressividade local. Este índice é tomado como sendo a diferença angular entre a orientação de uma determinada junta na pose de repouso do personagem e a pose da mesma junta após a transposição da pose pelo Módulo de Mapeamento, descrito na seção 5.3. Quanto maior o ângulo entre estas duas poses maior o peso de interpolação em favor do CPE. O diagrama da Figura 43 ilustra os elementos necessários a esta fusão por interpolação baseada na expressividade local.

Figura 43 – Interpolação em função do índice de expressividade local



Fonte: O autor

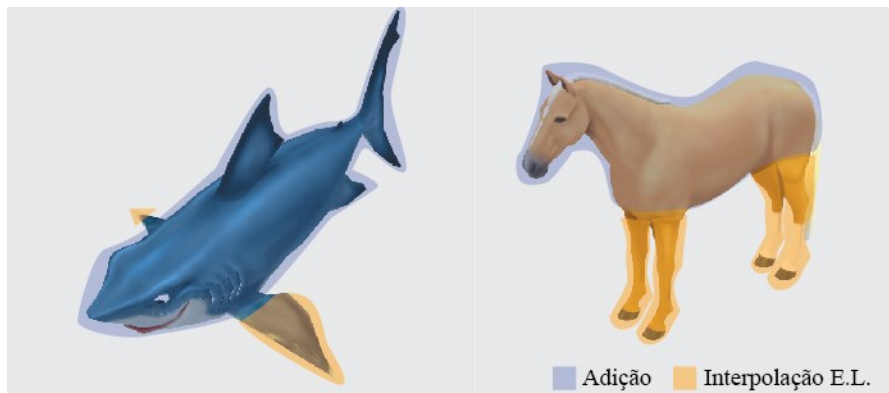
5.4.1 Aplicabilidade dos Resultados

Para a ocasião da aplicação das diferentes abordagens de fusão discutidas anteriormente, verificou-se que em parte das anatomias dos personagens animais testados, uma ou outra técnica se mostram mais adequadas. A fusão por adição tem como característica positiva o fato de que combina dados das duas poses originais; já a fusão por interpolação baseada na expressividade local permite a escolha entre um ou outro componente, em função de um índice de expressividade atribuído a cada membro do personagem.

Em relação aos personagens do tubarão e do cavalo, apresentados na seção 5.3.3, notou-se que nas nadadeiras do tubarão e nas patas do cavalo o método de fusão mais adequado aos diferentes movimentos testados foi o da interpolação baseada na expressividade local. Já no caso

da espinha, pescoço e cabeça destes personagens o uso da fusão por adição se mostrou mais satisfatório. O diagrama da Figura 44 aponta em cada anatomia as regiões separadas pelo tipo de fusão utilizado.

Figura 44 – Diagrama exemplificando as regiões anatômicas dos modelos utilizados quanto ao tipo de fusão escolhido.



Fonte: O autor.

No capítulo 6 o resultado da dissociação, transposição e fusão dos movimentos são apresentados, exemplificados e discutidos de modo integrado.

6 RESULTADOS E DISCUSSÃO

O objetivo geral deste trabalho foi propor e desenvolver o protótipo de um *framework* computacional para a utilização dos dados obtidos pela captura de movimentos de atores, não treinados, na animação de personagens animais, não humanoides, capaz de controlar personagens com graus de liberdade superiores, iguais ou inferiores aos do ser humano de quem os movimentos são capturados.

No capítulo 4 foi apresentada uma análise da representação de animais em obras de animação, no sentido de delimitar o desenvolvimento do protótipo a este contexto de aplicação. No capítulo 5 foram descritas as características do protótipo, também foram detalhados os procedimentos adotados no desenvolvimento de cada um de seus quatro módulos.

No presente capítulo são apresentados os resultados obtidos com o uso do *framework* proposto no processamento de movimentos humanos obtidos por captura de movimentos, visando transpor tais movimentos a personagens animais não humanos. Os movimentos utilizados são, em sua totalidade, provenientes da base de dados CMU (2017).

Quanto aos personagens digitais animais não-humanos, foram utilizados dois modelos: sendo um de anatomia cordata, um tubarão, e outro de anatomia quadrúpede, um cavalo. Procurou-se tomar as medidas necessárias para que os referidos modelos pudessem ser usados com o protótipo desenvolvido. Estas medidas envolvem: (1) a criação das animações necessárias a transposição de movimentos por classificação de ações e (2) a criação das poses necessárias a transposição de movimentos por mapeamento de poses.

As animações utilizadas para a transposição dos movimentos do CED do cavalo incluem os movimentos de caminhar, correr, caminhar para os lados esquerdo e direito, caminhar para trás e saltar. Já para a transposição dos do CED do tubarão os movimentos de caminhar assim como caminhar para os lados são um e o mesmo e apresentam o tubarão batendo as nadadeiras e fazendo um movimento senoidal com o corpo. O movimento de correr é similar ao movimento de caminhar, no entanto com o dobro da velocidade e maior amplitude. Além destes movimentos foi criado também um movimento para o salto. As poses utilizadas para a transposição dos movimentos expressivos do tubarão estão representadas no Apêndice C, já as poses utilizadas para o modelo do cavalo encontram-se nos Apêndices D e E.

Quanto a maneira como os movimentos dos personagens do tubarão e do cavalo se adequam a tipificação proposta no capítulo 4,

ambos estão enquadrados na categoria 3 de antropomorfismo. Tal categoria é definida pelo uso de uma locomoção naturalista combinada a gestos ocasionalmente antropomórficos. No entanto, o modelo do cavalo foi representado de dois modos diferentes quanto a maneira como articula tais gestos. No primeiro modo (A), o cavalo assume em todos os casos uma pose quadrúpede exceto quando há a necessidade de representar um gesto executado pelo humano com as duas mãos, nestes casos o cavalo fica ereto sobre as duas patas traseiras. Já no segundo modo (B) de representação, sempre que há a necessidade de representar um gesto humano o cavalo assuma uma pose bípede, independentemente de quantas mãos foram utilizadas pelo humano para executar o gesto. Estes modos de representação foram inspirados, respectivamente, pelos personagens do burro e da zebra das obras *Shrek* (Adamson *et al.*, 2004) e *Madagascar* (Darnell *et al.*, 2012).

A Figura 45 apresenta um movimento obtido da já referida base de dados CMU onde o ator humano executa de modo concatenado uma série de atividades ligadas ao CED e ao CPE: fica parado, gesticula com um braço, gesticula com os dois braços, anda para trás, corre e gesticula novamente com um dos braços. Ainda na Figura 45 apresenta-se o resultado da transposição do movimento humano às figuras do tubarão e do cavalo, sendo o cavalo representado nos modos A e B.

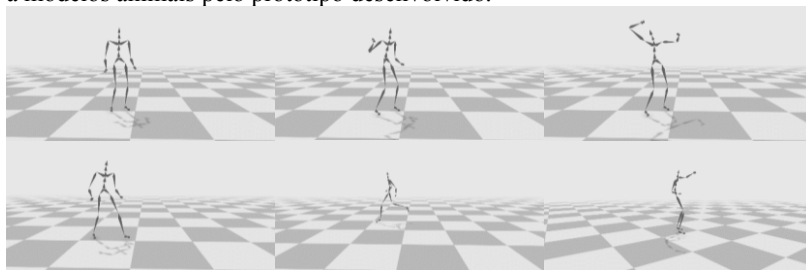
Apesar de a natureza dos dados ser mais propícia a visualização em vídeo¹⁵, é possível ver pelas imagens da Figura 45 que os gestos sintetizados nos modelos dos animais possuem coerência com os gestos realizados pelo ser humano de quem o movimento foi capturado. É possível notar, especialmente no modelo do cavalo, que o modo de locomoção do humano é coerente com o modo de locomoção do animal. Além disso verifica-se que a postura do cavalo varia em função da existência ou não de gestos expressivos no modelo animal e na maneira como foram configurados os modos A e B de representação do cavalo.

O conjunto de atividades representados na Figura 45, a maneira como tais atividades são executados pelo ator e o modo como se encontram concatenadas em uma sequência, configuram um caso não trivial, o qual não tem condições de ser tratado por nenhuma técnica de transposição de movimentos anteriormente proposta. Ainda tratando das atividades representadas na Figura 45, nota-se que o ator que executa as ações não foi treinado para se movimentar de um modo que facilitasse a utilização do protótipo, ele não imita os movimentos dos personagens

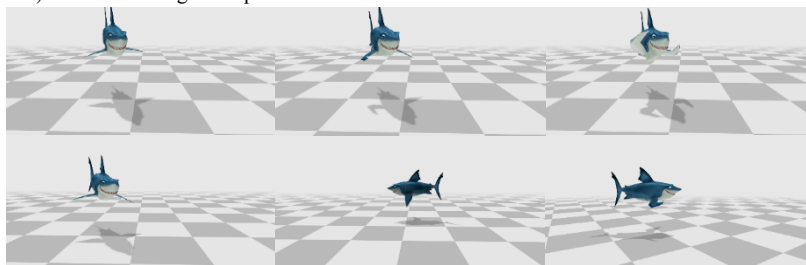
¹⁵ Um vídeo contendo os movimentos representados na Figura 45 pode ser acessado por meio do endereço: <http://www.tinyurl.com/tgeb-vid1>.

para os quais o movimento será transposto, nem executa ações com vistas a um padrão pré-configurado de movimentos. A transposição dos movimentos é feita em tempo real, ou seja, nenhum animador foi utilizado para escolher e reproduzir as poses chave do movimento. Salienta-se também que o *framework* proposto foi capaz de representar uma mesma figura, o cavalo, em diferentes modos de representação no que se refere ao grau de antropomorfismo do personagem animado. Estas características, que diferenciam o sistema desenvolvido nesta tese de trabalhos anteriores, são discutidas na seção seguinte.

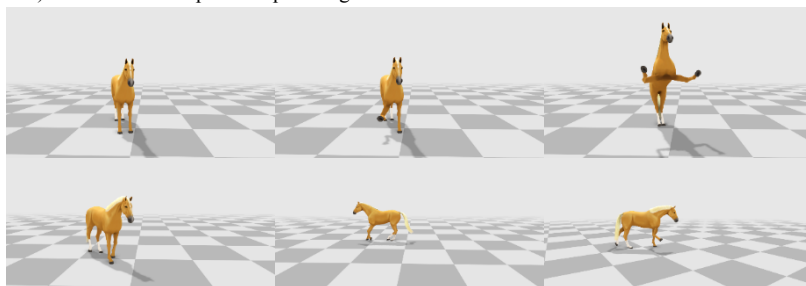
Figura 45 – Composição com vários movimentos obtidos por captura transpostos a modelos animais pelo protótipo desenvolvido.



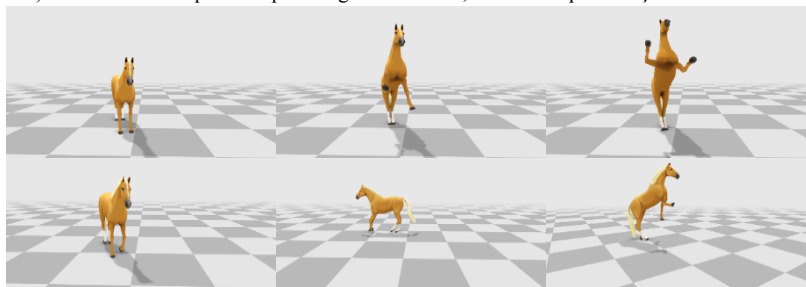
a) Movimento original capturado do ser humano



b) Movimento transposto ao personagem do tubarão



c) Movimento transposto ao personagem do cavalo, modo de representação A



d) Movimento transposto ao personagem do cavalo, modo de representação B

Fonte: O autor.

6.1 PRINCIPAIS CONTRIBUIÇÕES

Salientam-se aqui as principais contribuições do *framework* proposto em comparação com métodos de transposição de movimentos já existentes. Como preconizado no objetivo geral, o sistema desenvolvido prova que é possível fazer a transposição de movimentos de sujeitos leigos, ou seja, de sujeitos que não foram instruídos a atuar de um determinado modo para que seus movimentos possam ser transpostos a personagens de anatomia animal. Esta afirmação se fundamenta no fato de que em nenhum dos movimentos utilizados nos testes o ator imita um animal ou age de maneira limitada com vistas a restrições do sistema desenvolvido. De fato, como as capturas são em sua totalidade provenientes da base de dados CMU e antecedem a presente pesquisa seria impossível instruir os sujeitos capturados a agirem desta ou daquela maneira.

Ainda com relação ao objetivo geral desta pesquisa nota-se que a diferença nos graus de liberdade dos personagens animados não são um impeditivo à transposição dos movimentos. Foram tratados com sucesso: (1) casos em que o personagem possuía menor número de graus de liberdade que o ator, (2) casos em que o personagem e o ator possuíam um número igual de graus de liberdade e (3) casos em que o personagem possuía maior número de graus de liberdade que o ator.

Quanto aos casos onde o número de graus de liberdade do personagem é menor que o do ator cita-se como exemplo testado com sucesso a transposição dos movimentos de pescoço e cabeça humanos à cabeça do personagem do tubarão, sendo o conjunto de membros do primeiro modelo composto de duas articulações e o do segundo de apenas uma.

Quanto aos casos onde o número de graus de liberdade do personagem e do ator são iguais cita-se como exemplo a transposição dos movimentos dos braços humanos às nadadeiras do tubarão, sendo estes iguais em número de membros e articulações.

Já nos casos onde o número de graus de liberdade do personagem é maior que o do ator cita-se como exemplo a locomoção no modelo do cavalo, onde os movimentos das duas pernas do ator humano são utilizados para controlar as quatro patas do modelo do personagem; sendo

as pernas e pés do ator humano compostas de quatro articulações e as patas do cavalo compostas de seis articulações¹⁶.

O sucesso na obtenção de tais características, do caráter leigo do ator e das não-restrições anatômicas do modelo, deve-se à adoção de um conjunto de técnicas para a transposição de movimentos que constituem a principal contribuição deste trabalho.

No que se refere ao caráter leigo da atuação foram propostas várias adequações e melhoramentos a técnicas existentes de transposição de movimentos por classificação de ações (apresentadas inicialmente na seção 5.3) e de transposição de movimentos por mapeamento de poses (apresentadas inicialmente na seção 5.2).

As técnicas de transposição de movimentos por classificação de ações até então existentes tinham como principal contexto de aplicação o controle de personagens em jogos eletrônicos. Por se tratarem de jogos, os sujeitos adequavam sua atuação de modo a obter o resultado desejado na tela. Em alguns casos o sistema de classificação de movimentos contava com uma etapa de configuração onde o usuário era explicitamente instruído a respeito de como deveria ser sua atuação. Também foi constatado nos exemplos apresentados por tais trabalhos que as atuações deveriam ocorrer sempre da mesma forma, ou seja, as ações não podiam ser interpretadas de modos estilizados, mas sempre de um modo padrão; o que constitui uma limitação dentro do contexto da criação de conteúdo digital animado.

No método desenvolvido nesta pesquisa uma grande base de dados de movimentos foi categorizada no sentido de identificar diferentes instâncias estilísticas de um mesmo movimento. Após tal ordenação, estes dados foram registrados temporalmente e correlacionados por meio de redes neurais tornando possível a identificação e classificação de uma mesma atividade quando interpretada em diferentes estilos. Como resultado o método desenvolvido se mostrou adequado à classificação de um largo conjunto de movimentos estilisticamente diferenciados. Isso torna o sistema desenvolvido mais adequado a criação de conteúdo digital animado que métodos anteriormente disponíveis, dadas as características intrínsecas a este contexto de aplicação.

Em relação a transposição de movimentos por mapeamento de poses alguns trabalhos existentes permitem o mapeamento em tempo real e outros o fazem como uma etapa posterior a seção de captura de

¹⁶ Nota-se que as articulações dos modelos virtuais utilizadas para controlar tais personagens são uma simplificação dos modelos anatômicos reais. Portanto o número de articulações virtuais nem sempre corresponde ao número de articulações reais.

movimentos. Nos métodos em que o mapeamento em tempo real é possível, um pequeno conjunto de poses de exemplo é utilizado para configurar o sistema. Nos métodos onde o mapeamento é posterior a captura um animador seleciona as poses chave do movimento humano capturado e cria poses correlatas no personagem animal. A desvantagem da primeira abordagem é a incapacidade de representar uma grande variedade de movimentos, a desvantagem da segunda abordagem é a necessidade do trabalho de um especialista a cada nova seção de captura de movimentos.

No método desenvolvido e apresentado nesta tese uma grande base de dados de movimentos foi utilizada para criar um espaço latente do movimento humano. Tal espaço latente foi então utilizado para codificar a representação de 34 poses chaves do movimento humano; e 34 poses equivalentes foram criadas nos modelos dos personagens animais. A correlação entre as poses humanas e as poses dos personagens foi feita utilizando redes neurais. Notou-se que a correlação feita entre as poses humanas, codificadas no espaço latente, e a pose dos animais se mostrou qualitativamente superior a correlação feita sem o uso do espaço latente. A abordagem desenvolvida se mostrou adequada ao mapeamento de um amplo espaço de poses, em tempo real, e sem a necessidade do trabalho de um animador a cada nova seção de captura. Tais características representam uma potencial redução de custos, uma vez que é possível representar muitos movimentos em tempo real e sem a necessidade do trabalho de um especialista.

Finalmente, a possibilidade da transposição de movimentos do ator ao modelo animal não-humano independentemente dos graus de liberdade encontrados num ou noutro modelo, é possível devido o método proposto para a dissociação do movimento humano em dois componentes sendo um ligado aos movimentos de equilíbrio e deslocamento e o outro aos demais movimentos. Tal dissociação permitiu o uso da técnica de transposição de movimentos por classificação de ações aos movimentos de equilíbrio e deslocamento e o uso da técnica de transposição de movimentos por mapeamento de poses aos demais movimentos.

Seol *et al.* (2013) propuseram, pioneiramente, a combinação de dois modos de transposição de movimentos no sentido de superar as limitações individuais de cada modelo. No entanto, a escolha pelo uso de um ou outro método se deu naquele trabalho de modo arbitrário, quando o sistema não era capaz de classificar um movimento recaia no uso do mapeamento de poses (Seol *et al.*, 2013, p.217).

Já o *framework* apresentado nesta tese foi desenvolvido em acordo com a representação dos personagens animais não-humanos em animação

(tipificada no capítulo 4). Ou seja, assume-se que algumas ações são mais adequadas a representações pouco antropomórficas e outras mais adequadas a representações bastante antropomórficas; e utiliza-se o módulo de dissociação dos movimentos para separar um conjunto do outro. A abordagem aqui proposta permite maior generalização que a abordagem heurística proposta por Seol *et al.* (2013) e contribui para a redução da perda de dados importantes para a transposição da pose do ator ao modelo do personagem.

6.2 VALIDAÇÃO DOS RESULTADOS

Os exemplos de captura e transposição de movimentos apresentados na Figura 45 sugerem que o *framework* desenvolvido é capaz de criar imagens dos personagens animais não humanos que possuem alta correlação com a atuação capturada de um ator humano. Tal resultado enseja questões relativas a validação do citado *framework*, e cuja as respostas encontram-se fora das delimitações deste projeto, questões como:

- As ações interpretadas pelo ator e os movimentos transpostos ao modelo animal seriam entendidos por um observador como equivalentes?
- Existe alguma perda significativa de detalhes na transposição dos movimentos?
- As animações transpostas são percebidas como esteticamente agradáveis?
- Como as animações transpostas automaticamente pelo sistema proposto se comparam a animações transpostas por um animador profissional?

Apesar de respostas epistemologicamente adequadas a tais perguntas se encontrarem fora do escopo da presente tese, um estudo piloto foi conduzido no sentido de avaliar, de forma preliminar, dentre estas a questão mais simples e fundamental, qual seja: “sujeitos humanos são capazes de identificar as ações interpretadas após sua transposição ao modelo animal não-humano?”. Para responder tal pergunta um conjunto de 254 indivíduos, divididos em um grupo de teste e outro de controle, foi sensibilizado a vídeos produzidos com o *framework* desenvolvido neste trabalho. Os procedimentos e resultados deste estudo estão descritos a seguir.

Quanto aos estímulos utilizados, um conjunto de dez movimentos da base de dados CMU foi selecionado para a realização deste estudo.

Procurou-se tomar os cuidados necessários para que tais movimentos não contivessem ações que se encontrassem fora dos limites de atuação do protótipo, limites estes apresentados em detalhes na seção 6.3. Como algumas restrições afetam diferentemente os modelos do tubarão e do cavalo, para a ocasião deste teste piloto apenas o modelo do tubarão foi testado; sendo este o modelo que melhor se adequa aos atuais limites do sistema desenvolvido.

Buscou-se representar nestes dez movimentos selecionados ações com variados graus de antropomorfismo, de modo a testar o comportamento do *framework* em diferentes contextos. As ações selecionadas foram as seguintes:

- Anda para frente e volta
- Anda para frente, volta e pula
- Anda e corre
- Discutindo
- Praticando artes marciais
- Lançamento de basquete
- Lançamento de baseball
- Recebendo e lançando bola de futebol americano
- Estica e espreguiça

Para cada uma das dez ações foram gerados vídeos do movimento original executado pelo ator humano e do movimento transposto ao modelo do tubarão¹⁷. Assim foi possível sensibilizar os sujeitos a estes dois estímulos, sendo o vídeo do ator o estímulo de controle e o vídeo do tubarão o estímulo de teste.

A coleta dos dados foi feita por meio digital. Após aceitar o Termo de Compromisso Livre e Esclarecido (Apêndice 6) cada um dos 254 participantes foi sensibilizado, randomicamente, a um vídeo contendo uma das dez ações anteriormente listadas e interpretadas pelo tubarão ou pelo ator humano. As condições em que o participante assistiu aos vídeos e preencheu os formulários não foram controladas. Para minimizar a possibilidade de erros na visualização do conteúdo permitiu-se ao participante repetir a exibição do estímulo quantas vezes acreditasse necessário.


Em um formulário o participante era convidado a classificar o estímulo a que foi sensibilizado com a seguinte pergunta: “Qual das seguintes legendas melhor classifica o movimento representado no vídeo

¹⁷ Os vídeos de controle e teste podem ser visualizados por meio do endereço: <http://www.tinyurl.com/tgeb-vid2>.

que você acabou de assistir?”. O participante podia então selecionar uma dentre as dez classificações apresentadas anteriormente. O número limitado de respostas possíveis induz, em certo grau, a resposta do participante, potencialmente reduzindo seu erro. No sentido de mitigar tal desvio nos dados duas medidas foram tomadas: (1) o participante poderia selecionar uma opção adicional intitulada “Nenhuma das outras opções”; e (2) cada participante foi sensibilizado a um único estímulo, evitando que este escolhe-se a resposta por eliminação de alternativas. A Figura 46 mostra a interface do formulário digital utilizado. Após responder à pergunta o participante era informado que não deveria preencher o formulário novamente, uma vez que isso faria introduzir um desvio na pesquisa; ou seja, permitiria que o sujeito escolhesse a resposta por eliminação.

Figura 46 – Interface do formulário utilizado na coleta de dados.

Assista ao vídeo quantas vezes acreditar necessário, prestando atenção aos movimentos interpretados pelo personagem, e então responda à pergunta.



Qual das seguintes legendas melhor classifica o movimento representado no vídeo que você acabou de assistir? *

- Estica e espreguiça
- Praticando artes marciais
- Recebendo e lançando bola de futebol americano
- Anda e corre
- Anda para frente, volta e pula
- Discutindo
- Anda para frente e volta
- Lançamento de basquete
- Nenhuma das outras opções
- Desferindo socos
- Lançamento de baseball

Fonte: O autor.

As respostas referentes a cada movimento, em cada grupo, foram analisadas pelo teste do qui-quadrado para verificar se, em cada caso, era possível refutar a hipótese nula. Ou seja, buscou-se verificar se as respostas divergiram significativamente de uma distribuição randômica, dado um grau de confiança de 95%. Considerando que em cada pergunta a resposta certa representava 1/11 das respostas, e as respostas erradas 10/11 das respostas a seguinte equação foi aplicada:

$$X^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (5)$$

Onde O representa a distribuição observada e E a distribuição esperada em cada caso. A Tabela 6 apresenta o número de respondentes de cada questão em cada grupo, a porcentagem de respondentes que classificou o movimento na categoria esperada (acerto) e se a hipótese nula foi ou não refutada.

Tabela 6 – Resultados da coleta e análise de dados do teste piloto. Movimentos em destaque foram positivamente reconhecidos pela população entrevistada tanto no grupo de teste como no de controle, tendo, nestes casos a hipótese nula sido rejeitada com um grau de confiança de 95%.

Movimentos	Grupo de Controle (Humano)			Grupo de Teste (Tubarão)		
	Nº Part.	h1 (p=.05)	Acerto	Nº Part.	h1 (p=.05)	Acerto
<i>Lançamento de basquete</i>	16	<input type="checkbox"/>	100,00%	12	<input type="checkbox"/>	83,33%
<i>Anda para frente, volta e pula</i>	11	<input type="checkbox"/>	90,91%	15	<input type="checkbox"/>	80,00%
<i>Estica e espreguiça</i>	15	<input type="checkbox"/>	93,33%	12	<input type="checkbox"/>	75,00%
<i>Anda para frente e volta</i>	12	<input type="checkbox"/>	100,00%	18	<input type="checkbox"/>	72,22%
<i>Desferindo socos</i>	12	<input type="checkbox"/>	91,67%	9	<input type="checkbox"/>	55,56%
<i>Praticando artes marciais</i>	8	<input type="checkbox"/>	100,00%	11	<input type="checkbox"/>	45,45%
Lançamento de baseball	13	<input type="checkbox"/>	76,92%	11	x	18,18%
Anda e corre	15	<input type="checkbox"/>	80,00%	12	x	8,33%
Discutindo	11	<input type="checkbox"/>	90,91%	15	x	6,67%
Recebendo e lançando bola de futebol americano	14	<input type="checkbox"/>	42,86%	12	x	0,00%
Total de Participantes						
			127			
				127		

Fonte: O autor.

Quanto ao grupo de controle é possível analisar que o nível de acerto verificado em todos os casos se mostra estatisticamente diferenciado de respostas dadas ao acaso; ou seja, em todas as respostas pode-se refutar a hipótese nula. No entanto, é notório também que alguns casos se mostram inequívocos, tendo um índice de acerto de 100%, como ocorreu no “lançamento de basquete”, no “anda para frente e volta” e no “praticando artes marciais”; enquanto outros casos se mostram mais dúbios como o “lançamento de baseball” com 76,92% de acerto e o “recebendo e lançando bola de futebol americano” com 42,86% de acerto. É possível levantar diversas hipóteses do porquê, dentro do grupo de controle, há uma amplitude tão grande na porcentagem de acertos: (1) alguns movimentos podem não fazer parte do repertório dos participantes; (2) o nome dado pelos pesquisadores a um fenômeno pode não corresponder ao nome que o participante dá ao mesmo fenômeno; (3) o ângulo de visão no qual o vídeo foi produzido, e a ação representada, não privilegiam a adequada compreensão do conceito; (4) a falta de elementos de contexto como um cenário e os objetos com os quais o ator interage prejudicam a compreensão do movimento.

Em relação a esta última hipótese, é possível supor que havendo a imagem de uma bola nos casos “lançamento de baseball” e o “recebendo e lançando bola de futebol americano” um maior índice de acerto poderia ser observado; no entanto, era objetivo desta análise verificar se uma ação poderia ser compreendida apenas pelos movimentos executados pelo ator. Ainda quanto a esta hipótese, nota-se que o “lançamento de basquete” contou com cem por cento de acerto, mesmo sem estes elementos de contexto.

Quanto ao grupo de teste foi possível verificar um grau de acerto que permitiu refutar a hipótese nula em seis dos dez casos. Estes casos possuem qualidades diferenciadas quanto ao grau de antropomorfismo e expressividade representados em cada ação. Os casos “anda para frente e volta” e “anda para frente, volta e pula” estão estritamente ligados ao equilíbrio corporal e ao deslocamento do corpo no espaço, nestes casos os movimentos são transpostos ao personagem de modo naturalista, ou seja, como na categoria 1 do TRAA. Nos casos “lançamento de basquete” e “desferindo socos” os personagens não estão se deslocando pelo espaço, executando apenas gestos transpostos ao personagem de modo antropomórfico. Já os casos “praticando artes marciais” e “estica e espreguiça” apresentam uma combinação de movimentos gestuais e de deslocamento, sendo transpostos ao personagem concomitantemente de modo naturalista e antropomórfico. Destes três, o último caso, da combinação de movimentos antropomórficos e naturalistas é o que

melhor representa o cenário de ineditismo do *framework* apresentado por esta tese.

Quanto aos casos em que não foi possível verificar um grau de acerto que permitisse refutar a hipótese nula, estes foram: “anda e corre”, “discutindo”, “lançamento de baseball” e “recebendo e lançando bola de futebol americano”.

O caso “anda e corre” não apresenta gestos expressivos; este caso apresenta a peculiaridade de que foi usado no modelo do tubarão o mesmo movimento para a ação de andar e de correr, um movimento caracterizado pelo bater das nadadeiras e um movimento senoidal do corpo do animal. A única diferença entre o “andar” e o “correr” adotada na configuração do personagem dentro do protótipo é a amplitude e a velocidade em que estes movimentos são executados. Assim, uma hipótese para o baixo índice de acertos nesta resposta é que o movimento configurado para ação andar e correr não tenham sido significativamente diferenciados para sensibilizar os sujeitos de pesquisa. Outra hipótese é que os sujeitos não associem propriamente os conceitos andar e correr ao deslocamento do tubarão que em verdade apenas nada em diferentes velocidades.

Os casos “discutindo”, “lançamento de baseball” e “recebendo e lançando bola de futebol americano” configuram movimentos que combinam componentes expressivos a alteração do equilíbrio e ao deslocamento do personagem pelo espaço. Cabe salientar que os casos “lançamento de baseball” e “recebendo e lançando bola de futebol americano” foram aqueles que, no grupo de controle, mostraram os resultados menos assertivos; assim já havia por parte dos participantes uma dificuldade com estes movimentos, dificuldades para as quais uma série de hipóteses foram anteriormente levantadas. Além disso levanta-se a hipótese de que a representação de tais movimentos na figura de um personagem não humano torna ainda mais complexa sua identificação, já que este não é o contexto em que habitualmente se encontram tais movimentos.

Quanto ao caso “discutindo”, além das dificuldades já abordadas do entendimento e classificação das ações por parte dos respondentes, cabem outras duas colocações. A presente pesquisa e o protótipo implementado focaram exclusivamente na transposição de movimentos corpóreos e não faciais do personagem, portanto, nos materiais desenvolvidos para este teste o personagem possui uma única expressão facial, que é a expressão facial em que o personagem foi modelado e configura um sorriso. É possível que o sorriso do personagem tenha influenciado a compreensão dos respondentes, uma vez que tal expressão facial se opõe ao conceito de discussão. Além disso, nota-se que 60% dos

respondentes desta categoria (“discutindo”) classificaram o movimento como “anda para frente e volta”. De fato, no movimento original, enquanto o ator executa gestos que mimetizam uma discussão, ele também caminha para frente e para trás. Assim é possível supor que, dadas as dificuldades já tratadas da falta de contexto, combinadas a expressão facial conflitante com o movimento, os respondentes tenham focado estritamente no componente de equilíbrio e deslocamento movimento.

Portanto, diante dos resultados do teste aqui apresentados é possível constatar que o *framework* desenvolvido se mostrou adequado ao objetivo de transpor o movimento de atores, não treinados, a personagens animais, não humanoides, e que após tal transposição uma série de movimentos que se mostram reconhecíveis e identificáveis quando interpretados pelo ator humano permanecem assim após a transposição ao personagem não humano. Também foram apontados os casos em que os movimentos deixaram de ser corretamente reconhecidos pelos participantes e foram levantadas as hipóteses para tal fenômeno. Além disso os procedimentos adotados neste teste piloto servem como ponto de partida para futuros trabalhos que visem avaliar o *framework* desenvolvido.

Em razão do exposto o próximo item aborda os limites do *framework* proposto bem como as oportunidades de superação de tais limites e potencialização dos resultados aqui obtidos por trabalhos futuros.

6.3 LIMITES E TRABALHOS FUTUROS

No presente trabalho buscou-se propor e desenvolver um *framework* para, de modo automatizado, transpor o movimento capturado de seres humanos a personagens animais não-humanos, superando os limites inerentes a sistemas até então existentes e conhecidos. Atualmente é comum que personagens não-humanos sejam animados de modo manual por animadores que definem as poses de tais personagens utilizando quadros chave. Winder e Dowlatabadi (2011) estimam que, considerando um conteúdo de animação com qualidade suficiente para ser consumido no meio televisivo, um animador é capaz de produzir aproximadamente 5 segundos de animação por dia de trabalho. Este processo minucioso, permite a representação de movimentos detalhados até às menores sutilezas, no entanto, de forma muito onerosa.

O *framework* desenvolvido nesta tese apresenta uma solução adequada ao problema de pesquisa proposto e aos objetivos traçados;

possibilitando uma redução drástica no tempo e custo de produção de conteúdo de animação para mídias digitais. Por outro lado, cabe estabelecer os limites de aplicação do sistema altamente automatizado, desenvolvido nesta tese, com trabalho manual realizado por especialistas de animação. Durante o relato da implementação de cada um dos módulos do sistema desenvolvido foram apresentados, caso a caso, tais limites. Estes são retomados aqui de modo condensado e estabelecendo uma ponte com oportunidades para futuros trabalhos de pesquisa.

Quanto a transposição dos movimentos ditos de equilíbrio e deslocamento, no estado atual de desenvolvimento do *framework* apresentado, só é possível realizar a transposição de ações do ator para que o sistema tenha sido explicitamente configurado. No caso dos resultados apresentados o sistema foi configurado para transpor as seguintes ações: ficar parado, andar, andar de costas, dar passos à esquerda e à direita, correr e pular para frente. Outras ações de deslocamento ou alteração do equilíbrio do ator como sentar, levantar e deitar, podem ser configuradas no sistema em futuras versões e aperfeiçoamentos. Além disso, o referido módulo de transposição de movimentos não-expressivos é capaz de adequar a direção geral e a temporalidade da reprodução do movimento do personagem à interpretação do ator, mas não é capaz de adequar outros aspectos espaciais do movimento como, por exemplo, sua amplitude.

Em parte tal restrição deriva do fato de que o Módulo de Classificação, como proposto neste trabalho, estabelece uma correlação ente o movimento capturado e uma classificação de tipos e um índice de progressão do movimento. Tal abordagem apresentou benefícios, como robustez, coerência temporal e a possibilidade de usar uma mesma rede neural para qualquer personagem que se deseje. Entretanto, é possível estabelecer uma correlação entre a camada intermediária da rede e uma pose no modelo do personagem a ser animado. Entende-se que isto permitirá que poses sejam geradas inclusive em casos para os quais a rede não foi treinada. Um trabalho nesse sentido poderá averiguar a capacidade da rede em representar estes movimentos que se encontram em posição intermédia aos exemplos usados no treinamento. Além disso, quanto a possibilidade de representar outras características do movimento, para além do seu tipo e progresso, um futuro trabalho poderá combinar o sistema proposto no módulo de transposição dos movimentos com sistemas de decomposição do movimento em temas primários e secundários como os de Amaya *et al.* (1996); Brand e Hertzmann (2000); Hsu *et al.* (2005); Min *et al.* (2010); Yumer e Mitra (2016), discutidos na seção 2.2.

Em relação à transposição de movimentos potencialmente expressivos, como proposta neste trabalho, um limite encontrado foi a incapacidade de tratar adequadamente ações em que o ator faça contato entre duas superfícies de seu próprio corpo. Um futuro trabalho neste campo poderá avaliar diferentes soluções para tal problema. Uma possibilidade seria utilizar uma técnica de registro entre geometrias tridimensionais, similar ao proposto por Bogo *et al.* (2017), para detectar colisões no corpo do ator e então forçar tais colisões no modelo do personagem. Outra alternativa seria adequar o treinamento da rede neural para calcular a pose gerada a partir dos dados produzidos pela camada de saída e penalizar o treinamento caso tal pose gere interpenetrações da geometria.

No que se refere à maneira como os componentes dos movimentos são combinados pelo Módulo de Fusão, o sistema proposto ainda não é capaz de identificar quando a combinação entre estes dois componentes é conflitante no sentido de criar uma pose de mecânica inválida, ou seja, uma pose na qual o equilíbrio do personagem não se sustenta. Uma possível solução para tal problema seria o uso de controladores dinâmicos, como os propostos em Shapiro *et al.* (2007), para verificar ou forçar a validade de uma dada pose. No entanto, isto implica na necessidade de um especialista configurar um conjunto de controladores para cada personagem de cada anatomia diferente. No caso de seres humanos, uma outra abordagem utilizada para solucionar este problema é utilizar um conjunto de dados obtidos por captura de movimentos para “aprender” o espaço das poses válidas do movimento humano, como proposto em Ho *et al.* (2013). No entanto, tal abordagem pressupõe a criação de uma ampla base de movimentos tidos como válidos para o personagem.

Os limites do sistema desenvolvido afetam diferentemente personagens de anatomias distintas. No caso do modelo do tubarão, por este flutuar não se lhe aplica a restrição do módulo de fusão e a incapacidade deste em avaliar quando as poses dos dois componentes estão em conflito no sentido de criar uma pose de mecânica inválida. Esta restrição, portanto, no caso dos exemplos aqui apresentados só se aplica ao modelo do cavalo. Por este motivo apenas o modelo do tubarão foi utilizado no teste piloto, uma vez que as poses de mecânica inválida do cavalo poderiam gerar uma sensação de estranheza do movimento adicional nos participantes.

Com o encerramento desta tese faz-se a exposição, de forma sucinta, das considerações finais sobre o tema proposto e os resultados obtidos na pesquisa.

6.4 CONSIDERAÇÕES FINAIS

Em conclusão o presente trabalho apresentou uma resposta adequada ao problema de pesquisa desta tese, o qual foi assim definido: no contexto da animação de personagens digitais, como utilizar os dados obtidos a partir da captura de movimentos de atores não treinados, para o controle de personagens animais não humanos, sem limites quanto às ações interpretadas, ao número de membros ou a forma anatômica destes?

O contexto de aplicação foi estudado e tipificado, e um *framework* para a transposição automatizada de movimentos humanos à personagens animais não humanos foi proposto e desenvolvido. O sistema desenvolvido foi comparado com métodos anteriormente propostos e os resultados gerados pelo sistema foram testados, de modo preliminar, com um conjunto de participantes voluntários. Na comparação com trabalhos anteriores o método desenvolvido nesta tese mostra avanços importantes no sentido de superar limitações existentes. Já no teste preliminar com espectadores voluntários o sistema desenvolvido mostrou, em diversos casos, ser eficaz na tarefa proposta de transpor os movimentos humanos a personagens não humanos.

Os resultados apresentados na presente tese permitem afirmar que a adoção do *framework* proposto em um contexto de produção de conteúdo de animação para mídias digitais pode produzir uma redução significativa no tempo e nos custos de produção, além de possibilitar uma nova dinâmica de trabalho dado que o sistema opera em tempo real. Assim o trabalho apresenta contribuição altamente relevante para o campo das Mídias, linha de pesquisa do Programa de Pós-Graduação em Design no qual a presente tese se enquadra.

Quanto a ênfase tecnológica do referido programa, o conjunto geral de processos que compõe o *framework* desenvolvido, bem como seus módulos apresentam elementos de inovação tecnológica importantes. Quanto ao funcionamento geral do sistema, a inovação reside no fato de que o *framework* desenvolvido faz uso de um procedimento para dissociar o movimento humano em dois componentes, o que permite a aplicação de métodos distintos de transposição de movimentos a cada um destes componentes. Em relação a estes métodos propriamente ditos, da transposição por classificação de movimentos e da transposição por mapeamento de poses, a presente tese apresentou

inovações que permitem o uso de tais técnicas por atores não treinados; o que não era possível até então. Tais inovações são apresentadas no presente documento, em três depósitos de patentes e dois artigos científicos, tendo um sido publicado (Boehs e Vieira, 2016) e outro aceito para publicação (Boehs e Vieira, 2018b).

Além da contribuição direta para o objetivo primeiro desta tese o sistema desenvolvido para a dissociação do movimento humano apresenta outras aplicações no campo da animação e da computação gráfica, como a modulação de gestos potencialmente expressivos do movimento humano em tempo real, a transição assíncrona dos componentes e a combinação destes componentes em novas animações. Estas contribuições são discutidas em artigo correlato que aguarda revisão e publicação (Boehs e Vieira, 2018a). Já a tipificação da representação animal apresenta uma contribuição teórica relevante para o design de conteúdo digital animado, tais contribuições são discutidas em artigo correlato, aceito para publicação e que aguarda revisão final (Boehs *et al.*, 2018).

Finalmente, os limites ainda existentes para a aplicação do *framework* desenvolvido ao contexto pretendido, foram discutidos e caminhos para a superação de tais limites sugeridos. Acredita-se que estes desafios remanescentes constituem oportunidades para a continuidade da pesquisa dentro deste programa de pós-graduação expandindo o pioneirismo consolidado com este trabalho.

REFERÊNCIAS

- ADAMSON, A.; ASBURY, K.; VERNON, C. **Shrek 2**: DreamWorks Pictures 2004.
- ADAMSON, A.; JENSON, V. **Shrek**: DreamWorks Pictors 2001.
- AL-ASQHAR, R. A.; KOMURA, T.; CHOI, M. G. **Relationship descriptors for interactive motion adaptation**. Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation. Anaheim, California: ACM: 45-53 p. 2013.
- AMAYA, K.; BRUDERLIN, A.; CALVERT, T. **Emotion from motion**. Proceedings of the conference on Graphics interface '96. Toronto, Ontario, Canada: Canadian Information Processing Society: 222-229 p. 1996.
- ANDREWS, M.; CHAPMAN, B. **Brave**: Walt Disney Pictures 2012.
- AUTODESK. **Autodesk Maya**: Autodesk, 2017.
- _____. Autodesk Maya Documentation, Animation layer modes. 2107. Disponível em: <
<http://help.autodesk.com/view/MAYALT/2018/ENU/?guid=GUID-BBCA0BC3-7608-4E86-8E9F-B4099C316156>>. Acesso em: 11/11/2017.
- BAKER, S. **Peacturing the Beast**. Manchester: Manchester University Press, 1993.
- BARAN, I. et al. Semantic deformation transfer. **ACM Trans. Graph.**, v. 28, n. 3, p. 1-6, 2009. ISSN 0730-0301.
- BOEHS, G. E.; VIEIRA, M. L. H. **Non-humanoid creature performance from human acting**. ACM SIGGRAPH 2016 Posters. Anaheim, California: ACM: 1-1 p. 2016.
- _____. **Decoupling Expressiveness and Body-Mechanics in Human Motion**. IEEE: Transactions on Visualizations and Computer Graphics 2018a.

_____. **Style Invariant Locomotion Classification for Character Control**. Computer Graphics Forum 2018b.

BOEHS, G. E.; VIEIRA, M. L. H.; ANDRADE, W. M. D. Antropomorphic Animal Movement in 3d Animation. **Journal for Animation History and Theory**, 2018.

BOGO, F. et al. **Dynamic FAUST: Registering Human Bodies in Motion**. 2017. 5573-5582.

BOX OFFICE MOJO. Animation Movies at The Box Office. 4/9/2016 2016. Disponível em: < <http://www.boxofficemojo.com/> >. Acesso em: 05/09/2016.

BRAND, M.; HERTZMANN, A. **Style machines**. ACM SIGGRAPH 2000: 183–192 p. 2000.

BUCK, C.; LEE, J. **Frozen**: Walt Disney Studios 2012.

CHEN, J.; IZADI, S.; FITZGIBBON, A. **KinÊtre: animating the world with the human body**. ACM SIGGRAPH 2012 Talks. Los Angeles, California: ACM: 1-1 p. 2012.

CHIU, Y.-C. **State-annotated motion graphs for behavior control**. 2007. 136 University of California, Riverside

CMU. **Carnegie-Mellon Mocap Database** 2017.

CRESWELL, J. W. **Qualitative Inquiry and Research Design: Choosing Among Five Approaches**. SAGE Publications, 2012. ISBN 9781452276847.

DARNELL, E.; VERNON, C.; MCGRATH, T. **Madagascar 3: Europe's Most Wanted**: Paramount Pictures 2012.

DISNEY, W.; IWERKS, U. **Steamboat Willie**. Burbank 1928.

DOCTER, P. **Up**: Walt Disney Pictures 2009.

EEMAD, S. A. et al. Perceptual validity in animation of human motion. **Computer Animation & Virtual Worlds**, v. 27, n. 1, p. 58-71, 2015.

FABRIC SOFTWARE INC. **Fabric Engine**. Montreal: Fabric Software Inc., 2017.

FREEDMAN, Y. Is It Real . . . or Is It Motion Capture?: The Battle to Redefine Animation in the Age of Digital Performance. **The Velvet Light Trap**, n. 69, p. 38-49, 2012.

GEIJTENBEEK, T.; PRONOST, N.; STAPPEN, A. F. V. D. **Simple data-driven control for simulated bipeds**. Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation. Lausanne, Switzerland: Eurographics Association: 211-219 p. 2012.

GLEICHER, M. **Retargetting motion to new characters**. Proceedings of the 25th annual conference on Computer graphics and interactive techniques: ACM: 33-42 p. 1998.

HART, H. When Will a Motion-Capture Actor Win an Oscar? , *Wired*, 2012. Disponível em: < <http://www.wired.com/2012/01/andy-serkis-oscars/> >. Acesso em: 26/08/2016.

HAYKIN, S. S. **Neural Networks: A Comprehensive Foundation**. Prentice Hall, 1999. ISBN 9780139083853. Disponível em: < <https://books.google.com.br/books?id=M5abQgAACAAJ> >.

HO, E. L. et al. Topology Aware Data-Driven Inverse Kinematics. **Computer Graphics Forum**, v. 32, n. 7, 2013.

HOLDEN, D.; SAITO, J.; KOMURA, T. A deep learning framework for character motion synthesis and editing. **ACM Trans. Graph.**, v. 35, n. 4, p. 1-11, 2016. ISSN 0730-0301.

HOWARD, B.; MOOR, R. **Zootopia**: Walt Disney Studios 2016.

HSU, E.; PULLI, K.; POPOVIC, J. Style translation for human motion. **ACM Trans. Graph.**, v. 24, n. 3, p. 1082-1089, 2005. ISSN 0730-0301.

HUFGEFLUESTER. World's first 3D inertial motion capture of a horse in motion. Hufgefluester, 2011. Disponível em: <
http://hufgefluester.eu/artikel/worlds-first-3d-inertial-motion-capture-of-a-horse-in-motion_4262?SID=072979798c304cb09545c5aec5c86da1fe8f75dc>.
 Acesso em: 26/08/2016.

ISHIGAKI, S. et al. **Performance-based control interface for character animation**. ACM SIGGRAPH 2009 papers. New Orleans, Louisiana: ACM: 1-8 p. 2009.

IZANI, M. et al., Eds. **Analysis of the Keyframe Animation and Motion Capture Case Studies**. Student Conference on Research and Development. Putrajaya, Malaysia: IEEE, Student Conference on Research and Development. 2003.

JACKSON, W. **The Wise Little Hen**. Silly Symphony. DISNEY, W. Burbank: United Artists 1934.

JARDIM, T. J. **Animals as character: Anthropomorphism as personality in animation**. 2013. 70 Dissertation (Maters). Arts, University of the Wirwatersrand, Johannesburg.

KITAGAWA, M.; WINDSOR, B. **MoCap for Artists: Workflow and Techniques for Motion Capture**. Taylor & Francis, 2012. ISBN 9781136139659.

KOVAR, L.; GLEICHER, M.; PIGHIN, F. Motion graphs. **ACM Trans. Graph.**, v. 21, n. 3, p. 473-482, 2002. ISSN 0730-0301.

_____. **Motion graphs**. ACM SIGGRAPH 2008 classes. Los Angeles, California: ACM: 1-10 p. 2008.

KRIESEL, D. A Brief Introduction to Neural Networks. 2007.

LANTZ, W.; LOVY, A. **Knock Knock**: Universal Pictures 1940.

LEE, J. et al. Interactive control of avatars animated with human motion data. **ACM Trans. Graph.**, v. 21, n. 3, p. 491-500, 2002. ISSN 0730-0301.

LEE, Y.; KIM, S.; LEE, J. Data-driven biped control. **ACM Trans. Graph.**, v. 29, n. 4, p. 1-8, 2010. ISSN 0730-0301.

LEVENTI-PEREZ, O. **Disney's Portrayal of Nonhuman Animals in Animated Films Between 2000 and 2010**. 2011. Thesis (PhD). Department of Communication, Georgia State University, Georgia.

MATHWORKS. Cross-Entropy. 2013. Disponível em: < <https://www.mathworks.com/help/nnet/ref/crossentropy.html> >. Acesso em: 29/11/2017.

_____. **MATLAB: MathWors.**, 2016a 2016a.

_____. **Neural Network Toolbox: MathWorks.**, 2016a 2016b.

_____. Improve Neural Network Generalization and Avoid Overfitting. 2017. Disponível em: < <https://www.mathworks.com/help/nnet/ug/improve-neural-network-generalization-and-avoid-overfitting.html> >. Acesso em: 11/11/2017.

MATLAB. Scaled conjugate gradient backpropagation. 2017. Disponível em: < <https://www.mathworks.com/help/nnet/ref/trainscg.html> >. Acesso em: 12/12/2017.

MCCAY, W. **Gertie the Dinosaur**. Chicago 1914.

MIN, J.; CHAI, J. Motion graphs++: a compact generative model for semantic motion analysis and synthesis. **ACM Trans. Graph.**, v. 31, n. 6, p. 1-12, 2012. ISSN 0730-0301.

MIN, J.; LIU, H.; CHAI, J. **Synthesis and editing of personalized stylistic human motion**. Proceedings of the 2010 ACM SIGGRAPH symposium on Interactive 3D Graphics and Games. Washington, D.C.: ACM: 39-46 p. 2010.

MOLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. **Neural Networks**, v. 6, n. 4, p. 525-533, 1993.

NIELSEN, M. A. Using neural nets to recognize handwritten digits. In: PRESS, D. (Ed.). **Neural Networks and Deep Learning**: Determination Press, 2015.

O'NEILL, R. **Digital Character Development: Theory and Practice, Second Edition**. CRC Press, 2015. ISBN 9781482250787.

ORWELL, G. **Animal Farm**. Londres: Secker and Warburg, 1945.

OSBORNE, M.; STEVENSON, J. **Kung Fu Panda**: Paramount Pictures 2008.

PENN, G. Análise Semiótica de Imagens Paradas. In: BAUER, M. W. e GASKELL, G. (Ed.). **Pesquisa Qualitativa com Texto Imagem e Som**. Petrópolis: Editora Vozes, 2003.

PÓS-DESIGN. Doutorado em Design. 30/10/2017 2017. Acesso em: 09/11/2017.

RABUNAL, J. R.; DORADO, J. **Artificial Neural Networks in Real-life Applications**. Idea Group Pub., 2006. ISBN 9781591409038. Disponível em: <
<https://books.google.com.br/books?id=nMOmPwAACAAJ>>.

RENAUD, C.; CHENEY, Y. **The Secret Life of Pets**: Universal Pictures 2016.

RHODIN, H. et al. Generalizing wave gestures from sparse examples for real-time character control. **ACM Trans. Graph.**, v. 34, n. 6, p. 1-12, 2015. ISSN 0730-0301.

RHODIN, H. et al. Interactive motion mapping for real-time character control. **Comput. Graph. Forum**, v. 33, n. 2, p. 273-282, 2014. ISSN 0167-7055.

SAKOE, H.; CHIBA, S. Dynamic programming algorithm optimization for spoken word recognition. In: ALEX, W. e KAI-FU, L. (Ed.). **Readings in speech recognition**: Morgan Kaufmann Publishers Inc., 1990. p.159-165. ISBN 1-55860-124-4.

SEOL, Y.; O'SULLIVAN, C.; LEE, J. **Creature features: online motion puppetry for non-human characters**. Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation. Anaheim, California: ACM: 213-221 p. 2013.

SHALEV-SHWARTZ, S.; BEN-DAVID, S. **Understanding Machine Learning: From Theory to Algorithms**. Cambridge University Press, 2014. ISBN 9781107057135. Disponível em: < <https://books.google.com.br/books?id=ttJkAwAAQBAJ> >.

SHAPIRO, A. et al. **A dynamic controller toolkit**. Proceedings of the 2007 ACM SIGGRAPH symposium on Video games. San Diego, California: ACM: 15-20 p. 2007.

SHARPSTEEN, B. **Dumbo**. Burbank: RKO Radio Pictures 1941.

SHIN, H. J. et al. Computer puppetry: An importance-based approach. **ACM Trans. Graph.**, v. 20, n. 2, p. 67-94, 2001. ISSN 0730-0301.

SPIEGELMAN, A. **Maus**. Nova Iorque: Pantheon Books, 1991.

STANTON, A. **Finding Nemo**: Buena Vista Pictures 2003.

_____. **Finding Dory**: Walt Disney Pictures 2016.

VINCENT, P. et al. **Extracting and composing robust features with denoising autoencoders**. Proceedings of the 25th international conference on Machine learning. Helsinki, Finland: ACM: 1096-1103 p. 2008.

VON MISES, L. **Human action: a treatise on economics**. H. Regnery Co., 1949. ISBN 9780809297436.

WEBER, M. **The Methodology of Social Sciences**. Nova Iorque: The Free Press, 1949.

WELLS, P. **Understanding Animation**. Routledge, 1998. ISBN 9780415115971.

_____. **The Animated Bestiary: Animals, Cartoons, and Culture**. Rutgers University Press, 2008. ISBN 9780813546438.

WINDER, C.; DOWLATABADI, Z. **Producing Animation 2nd Edition**. 2. Focal Press, 2011.

WROTEK, P.; JENKINS, O. C.; MCGUIRE, M. **Dynamo: dynamic, data-driven character control with adjustable balance**. Proceedings of the 2006 ACM SIGGRAPH symposium on Videogames. Boston, Massachusetts: ACM: 61-70 p. 2006.

XIA, S. et al. Realtime style transfer for unlabeled heterogeneous human motion. **ACM Trans. Graph.**, v. 34, n. 4, p. 1-10, 2015. ISSN 0730-0301.

YAMANE, K.; ARIKI, Y.; HODGINS, J. **Animating non-humanoid characters with human motion data**. Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. Madrid, Spain: Eurographics Association: 169-178 p. 2010.

YOON, H.; MALECKI, E. J. Cartoon planet: worlds of production and global production networks in the animation industry. **Industrial and Corporate Change**, v. 19, n. 1, p. 239-271, 2009.

YUMER, M. E.; MITRA, N. J. Spectral style transfer for human motion between independent actions. **ACM Trans. Graph.**, v. 35, n. 4, p. 1-8, 2016. ISSN 0730-0301.

APÊNDICE A – Análise De Obras De Animação

Este apêndice contém o relatório individualizado das obras analisadas na pesquisa de “Tipificação da Representação Animal em Animação”, detalhada no capítulo 4.

Procurando Dory

Resumo da narrativa: Procurando Dory (Stanton, 2016) é uma sequência do filme Procurando Nemo (Stanton, 2003). A história se passa no fundo do mar e narra a aventura de três peixes procurando a família de um deles, Dory. No caminho se deparam com uma série de criaturas marítimas, fazem novos amigos e superam desafios, até finalmente encontrarem os pais de Dory.

Anatomia dos personagens: Os peixes são a espécie de animais mais proeminentes durante todo o filme, outras espécies que tem participação significativa nos eventos da trama são: arraias, polvos, tubarões, baleias e leões marinhos. Alguns outros animais ainda, aparecem em um número reduzido de cenas, com contribuição apenas marginal à trama, são eles: tartarugas, ostras, pássaros, morsas e enguias. Os corpos dos personagens são desenhados com formato similar aos dos animais que representam e com número idêntico de extremidades do corpo, como nadadeiras, caudas, patas, etc.; sendo a mudança anatômica mais significativa o fato que todos os personagens tem olhos e bocas humanizados.

Movimentos Antropomórficos: Com raras exceções, os personagens falam e utilizam expressões faciais antropomórficas para se comunicar. Os olhos, boca e sobrancelhas ajudam a comunicar emoções como tristeza, alegria, espanto, entre outras. Para acentuar a expressividade de peixes e arraias, o centro de gravidade do corpo destes animais é utilizado para expressar movimentos que, nos seres humanos, acontecem no tronco. Por exemplo, em uma das cenas um peixe, está raciocinando e seu centro de gravidade balança da esquerda para a direita, com tempo e movimento similar ao de uma pessoa que joga o peso de seu corpo sobre uma ou outra perna. Em outra cena o personagem dá um pulo de espanto. Como não existem pernas que justifiquem estes movimentos, fica caracterizado claramente que os mesmos estão ali unicamente para comunicar certos conceitos aos espectadores. Este é praticamente o único movimento expressivo apresentado pela arraia, cujo corpo não possui extremidades definidas.

Quanto a locomoção os peixes, tubarões e baleias, balançam o corpo todo em um movimento senoidal, além de utilizarem as nadadeiras, de modo naturalista. Quando estão parados o corpo e as nadadeiras movimentam-se lentamente para manter o equilíbrio do personagem. As baleias, enjauladas, também utilizam as nadadeiras para se escorarem em grades, evitando a necessidade de nadar e expressando preguiça de modo humanizado.

Em diversos momentos os peixes têm de executar ações humanas como pegar e manipular objetos, tocar ou interagir com outros personagens, fazer mímicas e gestos. Em muitos destes casos apenas uma nadadeira é utilizada e a outra mantém o movimento involuntário que equilibra o corpo, no entanto, quando necessário as duas nadadeiras são utilizadas, sem prejuízo ao equilíbrio. Quando o animal tem várias nadadeiras apenas as duas maiores e mais próximas a cabeça são utilizadas. Os rabos foram utilizados para executar movimentos antropomórficos apenas três vezes, durante todo o filme, em casos onde os personagens precisavam levantar ou lançar objetos do chão.

O polvo utiliza seus tentáculos para executar ações humanizadas como apontar, acariciar superfícies, manipular e segurar objetos, expressar medo cobrindo os olhos e dirigir. Além disso, quando está parado, em muitas cenas, os tentáculos do polvo ficam dispostos de modo a fazê-los parecerem braços humanos. Em algumas cenas eles estão cruzados, em outra ele descansa o “cotovelo” sobre uma superfície e assim por diante. Em uma dada cena, o polvo coça o queixo com um dos tentáculos, como faz uma pessoa pensativa, apesar de o personagem do polvo não possuir um queixo protuberante. Quanto a locomoção, na maior parte do tempo ele se comporta como um polvo, no entanto, em alguns casos ele se pendura como um ginasta em partes dos cenários para impulsionar-se à frente.

Apenas dois animais que participam da narrativa não falam, um deles é um leão marinho e outro um pássaro. Nestes dois casos os animais se comportam com pouco ou nenhum antropomorfismo e sua participação colabora apenas acidentalmente com o desenrolar da trama, uma vez que não expressam intencionalidade.

Antropomorfismo na narrativa: O antropomorfismo parece ser utilizado nesta obra de modo livre, sempre que necessário para expressar as intenções e sentimentos dos personagens na narrativa. Quando é possível conciliar o movimento natural dos animais com a função narrativa isto é feito, quando não é possível, os movimentos antropomórficos são utilizados.

Zootopia

Resumo da narrativa: Em Zootopia (Howard e Moor, 2016) animais, uma vez selvagens, vivem em espaços urbanos e rurais, compartilhados por diversas espécies de modo civilizado. Neste contexto, a primeira policial coelho da cidade descobre uma conspiração que se armava para segregare animais que uma vez foram predadores dos demais habitantes da cidade, tornando-os selvagens novamente; desarticulada tal conspiração a cidade volta à normalidade.

Anatomia dos personagens: Quanto aos animais representados e suas anatomias é curioso que, apesar de tratar de um mundo povoado inteiramente por bichos bastante diversos, as espécies representadas são, em um sentido, restritas. Todos os animais encontrados, que atuam de modo relevante na narrativa, são animais vertebrados, de quatro patas e com uma cauda de maior ou menor protuberância. Aves, peixes, cobras, insetos e outros animais não foram encontrados nesta obra. Os animais encontrados são retratados de dois modos bastante distintos: (1) como selvagens; e (2) como animais civilizados, onde a civilização é expressa por um alto grau de antropomorfismo. Muitos signos são utilizados para configurar o conceito de civilização, como, a presença ou ausência de roupas. Porém, para a ocasião desta análise detalhamos apenas aqueles signos diretamente ligados à movimentação dos personagens.

Movimentos Antropomórficos: Quando os personagens estão selvagens, a representação de movimento é naturalista, não se nota, pela atuação dos personagens, uma intencionalidade humana; apenas os movimentos instintivos dos animais. Nestes casos os animais são apresentados como perigosos e, portanto, executavam ações como correr, espreitar e dar o bote. Em todos os casos eles executam estas ações sobre as quatro patas, com maneirismos característicos de cada espécie.

Já no modo civilizado todos os personagens são representados como bípedes com postura ereta. Patas traseiras são utilizadas como pernas humanas, para locomoção, e patas dianteiras são utilizadas para gesticular, interagir com outros personagens, manipular objetos, dentre outros. Todos os personagens têm dedos e polegares opositores nas patas dianteiras. Como modo de diferenciar um personagem dos demais, a velocidade e amplitude dos movimentos levam em conta características de cada animal, por exemplo: coelhos se movem de modo rápido e com passadas curtas, elefantes dão passadas longas e demoradas, porcos da índia tem passos curtíssimos e rebolam ao andar e assim por diante. Além desta definição geral é possível encontrar alguns outros elementos distintivos de algumas espécies. No caso específico dos elefantes a

tromba é utilizada algumas vezes de modo antropomórfico, como um terceiro braço, para manipular objetos e gesticular. Outras vezes a tromba é utilizada como uma espécie de aspirador. No caso dos coelhos e raposas, que possuem orelhas proeminentes, estas são utilizadas de modo naturalista, ou seja, assim como na natureza as orelhas levantam e giram quando algo chama a atenção destes animais.

Antropomorfismo na narrativa: Os movimentos são utilizados, no filme, como um dos muitos signos que distinguem os animais civilizados dos não civilizados. Este modo de animar os personagens é consistentemente aplicado durante toda a obra. Uma única vez, durante todo o filme um animal civilizado foi visto em modo quadrupede. No caso, a personagem principal, quando assustada por outro personagem, pula e corre nas quatro patas por um breve espaço de tempo, até recobrar sua calma e voltar à pose bípede. Apesar de breve, este momento contribui para a narrativa a medida que exprime a ideia, difusa em outros elementos da obra, de que todos animais têm um instinto selvagem latente que é desencadeado pelo medo. As intenções e emoções dos personagens são expressas sempre de modo antropomórfico e seus instintos de sobrevivência mais primitivos aparecem, pontualmente, de modo naturalista.

Kung Fu Panda

Resumo da narrativa: Kung Fu Panda (Osborne e Stevenson, 2008) é uma história que se passa em uma fictícia China imperial povoada por bichos, onde a segurança das cidades é garantida por mestres de Kung Fu. Um urso panda desajeitado e sem experiência de luta, torna-se, por indicação, o dragão guerreiro, supostamente o maior mestre de toda a China. Em um curto espaço de tempo Po, o urso panda, é treinado por seu mestre, Shifu, enfrenta e derrota Tai Lung, vilão da história e uma grande ameaça ao vilarejo de Po.

Anatomia dos personagens: Quanto a anatomia dos personagens estes são, em sua maioria, vertebrados e quadrupedes. As exceções a esta regra são alguns gansos, uma garça, uma cobra e um louva-a-deus. Porcos, coelhos e gansos representam o povo da cidadela, já outros animais como felinos, lobos, uma cobra, uma garça, um louva-a-deus, uma tartaruga e um panda vermelho são guerreiros, ou mestres de Kung Fu. O personagem principal, Po, é um urso panda gigante que vive em meio aos cidadãos comuns, até ser escolhido como o dragão guerreiro.

Movimentos Antropomórficos: Todos animais quadrupedes anteriormente citados são representados, nesta obra, como bípedes.

Utilizam as patas traseiras como pernas, para a locomoção, e as dianteiras, que possuem dedos e polegares opositores, são usadas para pegar e manipular objetos, interagir com outros personagens e gesticular, de modo antropomórfico. Além disso, no caso dos guerreiros, patas traseiras e dianteiras são utilizadas, respectivamente, como pés e mãos em cenas de luta. Ainda quanto aos personagens guerreiros nota-se que, quando estes se encontram em uma cena de ação, especialmente quando precisam correr, o fazem de modo naturalista, cada qual no modo peculiar a sua espécie. Como este filme possui muitas cenas de ação, este fenômeno ocorre com alguma frequência.

As aves são naturalmente bípedes, portanto, o fato de os gansos e garças andarem sobre duas patas não pode ser considerado antropomorfismo. Na obra, a locomoção destes animais tem aspecto naturalista, com os maneirismos esperados de tais espécies. A característica antropomórfica mais destacada nestes animais é o uso das asas como mãos para gesticular, manipular objetos e interagir com outros personagens. Além disso a garça, frequentemente, usa os pés para lutar; já nos gansos, o uso da pata para manipular objetos foi verificado em apenas uma cena e foi utilizado de modo a demonstrar que aquele personagem possuía grande destreza na tarefa desempenhada, podendo manipular um objeto sem olhar diretamente para ele. Em cenas de ação a garça voa ao invés de caminhar, enquanto os gansos sempre caminham e também não participam de cenas de ação.

Quanto ao louva-a-deus, na natureza este bicho costuma andar sobre as seis patas, recolhendo as duas patas frontais quando está parado, além de saltar para cobrir distâncias maiores. Na obra analisada o louva-a-deus salta para cobrir grandes distâncias, mas, para pequenas distâncias ele se move sobre quatro patas utilizando as duas patas posteriores para gesticular ou manipular objetos enquanto caminha, caracterizando um comportamento antropomórfico, não propriamente nas patas traseiras, mas nas duas patas dianteiras. Como este personagem é um guerreiro também usa as patas para lutar. Ao executar movimentos de luta com as patas traseiras ele agrupa as patas duas a duas, e cada um dos pares agrupados executa um movimento antropomórfico. Assim, se ele executa um chute para o lado direito, as duas patas traseiras do lado direito chutam a um só tempo, como se fossem uma perna humana, enquanto as outras duas patas, do lado esquerdo, equilibram o corpo; as patas dianteiras fazem algum outro gesto, como braços humanos, visando o equilíbrio.

Já a cobra, que não possui patas, locomove-se como um ser rastejante que é. Para gesticular utiliza a ponta do seu rabo livremente e para manipular objetos enrola a ponta do rabo em torno deles. Quanto

maior e mais pesado o objeto maior a porção do rabo enrolada, em alguns casos a cobra utiliza quase que todo seu corpo para enrolar e manipular um objeto ou personagem.

Antropomorfismo na narrativa: Os autores desta obra utilizam animais para contar uma história que poderia também ser representada por seres humanos. No entanto, como o Kung Fu é uma arte marcial cuja alguns estilos têm movimentos inspirados em animais, e como o Kung Fu é tema que permeia toda a obra a escolha por animais antropomórficos não é injustificada. O antropomorfismo é utilizado para desenvolver toda a narrativa e só não é presente quando, por ocasião da luta, os personagens usam suas características animais.

Madagascar 3

Resumo da narrativa: Esta é a terceira obra da série de filmes Madagascar, cuja a primeira obra data de 2005, e que trata das aventuras de um grupo de animais fugidos do zoológico. Neste filme os animais sentem saudades do zoológico e tentam voltar até ele. No caminho conhecem um circo de animais abandonado pelos humanos e descobrem, junto dos novos amigos, sua vocação circense.

Anatomia dos personagens: Os animais representados nesta obra são, na sua maioria, vertebrados e quadrúpedes, dos mais proeminentes na narrativa podemos citar: leão, zebra, girafa, hipopótamo, lêmures, tigre, leão marinho e onça. Dentre o núcleo principal os únicos animais que, na natureza, não são quadrúpedes são os pinguins. Quanto aos personagens secundários todos são quadrúpedes: cavalos, cachorros, porcos da Índia, macacos e um urso. Além disso, em uma única cena aparecem algumas cobras. Excetuando as cobras, todos animais falam e tem expressões faciais que comunicam seus sentimentos.

Movimentos Antropomórficos: A maneira como as diferentes espécies são representadas quanto a locomoção, bípede ou quadrúpede, não segue, aparentemente, nenhuma regra. Leões, tigres, onças, hipopótamos e cachorros por exemplo, são bípedes em praticamente toda a duração do filme. A zebra, os macacos e a girafa são quadrúpedes em grande parte das cenas e bípedes em outras; cavalos, de outro modo, são quadrúpedes em todas as cenas.

Os animais que são representados como bípedes, na maior parte das cenas, utilizam as patas traseiras como pernas, para locomoção, e as dianteiras como braços, para gesticular, interagir com outros personagens e manipular objetos. Nestes personagens as mãos possuem dedos e polegares opositores. O leão aparece algumas vezes como quadrúpede

quando, em uma cena de ação, precisa correr em grandes velocidades. Os lêmures e pinguins são representados em todas as cenas como bípedes. No caso dos pinguins as asas atuam como mãos e, mesmo sem a representação dos polegares opositores, são capazes de manipular objetos com destreza.

A zebra, que é representada em muitas cenas como quadrupede usa o movimento da cabeça para expressar sentimentos de modo antropomórfico, quando suas quatro patas estão no chão. Algumas vezes o personagem levanta uma das patas, sustentando seu equilíbrio com as outras três, para executar algum gesto. Outras vezes, quando está em um diálogo em que gesticula com as duas mãos, equilibra-se sobre duas patas, antropomorficamente. Em outro caso ainda, o animal é representado sentado dirigindo um automóvel com duas de suas patas. A zebra mesmo quando é apresentada de maneira bípede não possui dedos e polegares opositores.

A girafa, assim como a zebra é representada em alguns casos como bípede, no entanto, em frequência significativamente menor. Os macacos são bípedes em algumas cenas, onde se fantasiam de humanos para ajudar os outros animais e utilizam os dedos e os polegares opositores para gesticular e manipular objetos. No entanto, na maior parte do tempo estes animais utilizam as patas dianteiras para auxiliar na locomoção. Para ampliar a aparência de destreza da espécie o rabo é utilizado como uma quinta pata que o permite agarrar objetos e dependurar-se de superfícies. Os cavalos, representados a todo tempo como quadrupedes, utilizam, além da fala e das expressões faciais, apenas o movimento da cabeça para gesticular e expressar seus sentimentos e intenções.

Antropomorfismo na narrativa: O antropomorfismo nesta obra é utilizado para que os personagens possam se expressar de modo compreensível aos expectadores. No entanto, em alguns casos os autores optam por utilizar graus maiores e menores de antropomorfismo sem que haja uma relação clara com a narrativa ou mesmo com a personalidade dos personagens.

Pets – A Vida Secreta dos Bichos

Resumo da narrativa: Pets (Renaud e Cheney, 2016) trata da chegada de um novo cachorro ao apartamento em que Max, um terrier, vive com sua dona. A briga entre o novo e o antigo cachorro do apartamento os leva para longe de casa e suas vidas são ameaçadas por um grupo de animais abandonados que quer se vingarem dos seres humanos;

eventualmente eles são salvos com a ajuda dos animais domésticos da vizinhança, antigos conhecidos de Max.

Anatomia dos personagens: O grupo dos animais domésticos é composto principalmente por cachorros, mas também conta com um gato, um pássaro, um porco da Índia, um canário e uma águia. O grupo dos animais abandonados é liderado por um coelho e conta também com um porco, jacarés, cabras, tartarugas, alguns cachorros e diversos gatos. Todos os animais têm expressões faciais bastante humanizadas, inclusive na presença de humanos.

Movimentos Antropomórficos:

Todos os animais domesticados, na presença de humanos, não falam nem agem antropomorficamente. Para que o espectador possa entender os sentimentos dos animais alguns artifícios são utilizados. Em alguns casos o narrador diz o que o animal está sentindo. Em outros vemos o animal falando em uma cena e, na cena seguinte, a partir de um plano subjetivo do dono do animal, vemos este último latindo ou miando. Os animais abandonados comportam-se de modo naturalista junto a seres humanos, quando querem passar despercebidos e comportam-se de modo antropomórfico quando querem confrontar os seres humanos diretamente. Em nenhuma situação, no entanto, um animal é capaz de falar com um ser humano e ser compreendido.

Todos os animais domésticos se locomovem respeitando suas características anatômicas, ou seja, gatos, cachorros e porco da Índia, andam e correm como quadrúpedes e as aves andam como bípedes e voam. Os quadrúpedes usam o movimento da cabeça para ampliar a expressividade daquilo que falam ou de suas expressões faciais, em alguns casos a cabeça também é utilizada para acompanhar o som de alguma música, em uma espécie de dança, e a boca utilizada para manipular objetos, como chaves.

O gato utiliza as patas dianteiras para manipular objetos, mas sem características antropomórficas, dando pequenos tapas com a parte interna das patas como fazem os gatos. Em alguns casos as patas dianteiras são utilizadas antropomorficamente pelos cachorros, como mãos, para apontar, gesticular, lutar e desferir tapas. Uma das cadelas de estimação, Gidget, eventualmente utiliza o corpo inteiro para lutar comportando-se, temporariamente, como bípede, utilizando as patas traseiras como pernas e as dianteiras como braços. No entanto, em nenhum caso estes animais são representados com polegares opositores.

O canário, domesticado, utiliza o bico para bicar alguns objetos, não há antropomorfismo em seus movimentos. A águia, no entanto, utiliza as asas retraídas como ombros para expressar alguns sentimentos

como empolgação e descontentamento; quando suas asas estão abertas utiliza-as como braços para gesticular-se e expressar-se. A águia não utiliza suas asas como mãos e, quando necessário, manipula objetos com o bico ou com as patas.

Entre os animais não domesticados há uma maior heterogeneidade na ocorrência do fenômeno antropomórfico. O coelho, líder da trupe, gesticula, interage com outros personagens, luta e manipula objetos como bípede; quando necessário, para correr em grande velocidade, ou enganar humanos, atua como quadrupede. Os jacarés andam como quadrupedes, mas param em pé como bípedes, o que ajuda o espectador a visualizar sua grandeza em comparação com os outros animais.

Os gatos não domesticados utilizam, em muitos casos, uma das patas para manipular objetos; em alguns casos equilibram-se temporariamente sobre as patas traseiras para usar as dianteiras como braços; e, outras vezes ainda, utilizam o rabo como um braço, para manipular objetos. O porco em uma única cena anda como bípede, fantasiado com roupa de ser humano.

Antropomorfismo na narrativa: O antropomorfismo, nesta obra, é evitado dentre os personagens protagonistas e usado esporadicamente nos antagonistas. Artifícios são utilizados para que o espectador compreenda as intenções dos personagens, sempre que necessário e especialmente quando os humanos estão presentes, preservando em um grau significativo os movimentos naturais de cada espécie. Uma hipótese para a maior liberdade de uso do antropomorfismo entre os antagonistas é que as motivações destes personagens, ou seja, empreender uma vingança contra todos os seres humanos, já é bastante lúdica, não havendo, portanto, um compromisso significativo em conectar estes personagens a realidade.

Shrek 2

Resumo da narrativa: Shrek 2 (Adamson *et al.*, 2004) é a sequência de Shrek (Adamson e Jenson, 2001) que, por sua vez, é uma sátira aos desenhos de contos de fada. Em Shrek 2 um ogro e uma ogra se casaram e vão pedir a benção dos pais da noiva. A família da noiva e sua fada madrinha tentam separar o casal, Shrek e seus amigos, o Burro e o Gato de Botas, tentam impedi-los.

Anatomia dos personagens: Nesta história os animais são personagens coadjuvantes sendo destes, o Burro e o Gato de Botas os mais proeminentes na trama. Além destes dois personagens, fazem parte da história um cavalo, um lobo, sapos, ratos e porcos. Todos os animais

da história se comunicam e expressam sentimentos pela fala e pelas expressões faciais.

Movimentos Antropomórficos: O Burro se locomove sobre as quatro patas e utiliza o movimento da cabeça em conjunto com as expressões faciais e fala para comunicar-se e expressar-se. Eventualmente o burro levanta uma das patas para realizar algum gesto antropomórfico como apontar em alguma direção, ou apenas para gesticular enquanto fala; em algumas ocasiões ele dá saltos com as patas dianteiras, para poder usar as duas mãos ao mesmo tempo ao gesticular. O Burro também gesticula com as duas mãos quando está sentado, e quando o faz senta de modo antropomórfico. Algumas vezes ele usa as patas para apoiar-se ou debruçar-se sobre algum objeto enquanto dialoga com os outros personagens.

O Gato de Botas, por outro lado, encontra-se na maior parte das cenas em poses bípedes e locomove-se desta maneira. Este personagem é um hábil esgrima e, portanto, tem as duas patas dianteiras livres para manipular sua espada como se tivesse mãos. Além da esgrima usa suas patas dianteiras para comunicar e expressar sentimentos, manipular objetos e interagir com outros personagens. É representado com dedos e polegares opositores. Em alguns casos, no entanto, o gato é representado com movimentos mais naturalistas, quando se lambe, quando corre longas distâncias como um quadrupede, ou quando arranha outros personagens.

Quanto aos demais personagens o cavalo é, na verdade, o mesmo personagem que o Burro, após receber os efeitos de um feitiço; assim comporta-se do mesmo modo que o Burro. O lobo os ratos e os porcos comportam-se como bípedes em todos os momentos, utilizando as patas traseiras como pernas e as dianteiras como braços; além disso os ratos, cegos, utilizam bengalas. Os sapos se comportam como quadrupedes, utilizando a expressão facial e os movimentos da cabeça para a expressão e a comunicação, além de complementar estes com o movimento de uma das mãos, evitando a perda de equilíbrio.

Antropomorfismo na narrativa: A maior parte dos personagens de Shrek é inspirado em criaturas dos contos de fada, assim alguns animais como o gato de botas, os três porquinhos e o lobo mau seguem representações clássicas destes personagens em outras obras, tanto na forma física quanto no movimento. O Burro, que é um personagem original da obra Shrek é representado de maneira antropomórfica apenas à medida que é necessário, para a narrativa. Fazendo saber das intenções e sentimentos do personagem, mas preservando sua postura quadrupede na maior parte do tempo.

Valente

Resumo da narrativa: Valente (Andrews e Chapman, 2012) conta a história de uma princesa de um clã escocês que está prometida para casar. Não querendo este destino para si pede a uma bruxa que impeça o casamento, o que esta faz transformando a mãe da princesa e seus irmãos em ursos. Para desfazer o feitiço a garota é obrigada a rever os mitos de seu clã, enquanto esconde sua mãe de outras pessoas para garantir sua segurança. Ao final da história, mãe e filha passam a entender melhor uma à outra.

Anatomia dos personagens: O filme tem duas representações distintas de animais, os domésticos, que povoam a vila, como cavalos, cachorros, cabras e ovelhas; e os mágicos que se limitam a um corvo que acompanha a bruxa, a urso mãe e os três ursos filhos. Os animais domésticos se comportam, a todo tempo, de modo naturalista, ou seja, seus movimentos são representados a imagem do que são na natureza e sem características humanas. Uma única vez que isso não acontece é quando o cavalo de estimação da personagem principal balança a cabeça de modo negativo. O restante das situações antropomórficas fica a cargo dos animais mágicos.

Movimentos Antropomórficos: O corvo fala e tem expressões faciais, além disso ele utiliza os movimentos cabeça para expressar intenções, como a de assustar ou interrogar outros personagens. Para pegar e manipular objetos utiliza o bico. Suas asas acompanham as demais expressões que faz com a face e a cabeça, no entanto, o movimento das asas não tem características humanizadas, ou seja, ele bate as asas ou as comprime contra o corpo como asas, não como braços e mãos. Como ave ele é bípede o que em si não configura antropomorfismo.

As pessoas transformadas em ursos, ou seja, a mãe e seus três filhos, movimentam o corpo, na maior parte do tempo, antropomorficamente. Estes personagens não falam, mas tem expressões faciais, movimentam a cabeça como humanos e utilizam as patas anteriores como pernas e as posteriores como braços. Andam como bípedes e utilizam as patas anteriores para gesticular e fazer mímicas, além de expressar sentimentos como empolgação, susto e outros.

Com o passar do tempo como urso a mãe percebe que é mais confortável para ela fazer algumas coisas como urso, como andar sobre as quatro patas e comer com a boca, sem utilizar talheres, preservando expressões faciais e movimentos da cabeça antropomórficos. O cheiro da comida, no entanto, faz com ela esqueça que é humana e, a partir daí ela

se comporta de modo inteiramente naturalista deixando de lado qualquer característica antropomórfica na representação de seus movimentos.

Antropomorfismo na narrativa: O antropomorfismo dos movimentos das personagens exerce uma função narrativa clara nesta obra. A mãe, enquanto urso, apesar do corpo grande e desajeitado, mantém os costumes e a etiqueta que tinha como rainha; eventualmente ela percebe que sua nova condição já não necessita de tais costumes e que libertar-se deles lhe proporciona uma vida mais agradável.

A total ausência de antropomorfismo dos movimentos da urso mãe, no entanto, é uma preocupação para a filha que precisa achar uma cura para o feitiço. Nenhum artifício é utilizado para comunicar as intenções da mãe quando seus movimentos são naturalistas pois pretende-se justamente indicar aos expectadores que naqueles momentos o personagem não tem mais nenhuma característica humana e, portanto, age de modo instintivo.

Up – Altas Aventura

Resumo da narrativa: Up (Docter, 2009) é a história de um recém viúvo que decide realizar um grande sonho que tinha com sua esposa, viajar à América do Sul; para levar consigo as lembranças dela leva toda sua casa içando-a com balões. Chegando ao destino encontra o explorador, Muntz, ídolo de sua infância, que passou décadas na Venezuela procurando um pássaro raro. No final os dois personagens se enfrentam pois um pretende raptar e ou outro libertar o pássaro. Muntz é um velho recluso e seus únicos ajudantes são cachorros treinados.

Anatomia dos personagens: Para a ocasião desta análise, o pássaro raro foi descartado, pois sua espécie é fictícia e como criatura fantástica ele está fora do escopo desta pesquisa. Já os cachorros treinados, que fazem companhia ao personagem Muntz, foram analisados.

Movimentos Antropomórficos: Estes cachorros locomovem-se como quadrúpedes e expressam seus sentimentos, na maior parte do tempo como cachorros, ou seja, abanando o rabo, pulando, correndo, rolando, sem executar movimentos antropomórficos. O modo que os autores encontraram para fazer os espectadores compreenderem a atuação dos cachorros foi utilizar um artifício, no caso, uma coleira que lê pensamentos e emite sons, para que os animais pudessem falar sem mexer as bocas. Enquanto “falam” expressões sutis nos olhos, sobrancelhas e lábios, além do movimento da cabeça complementam a atuação dos personagens.

Alguns cachorros executam tarefas domésticas como varrer, empurrar carrinhos e servir comida e bebida a convidados. Em todas estas tarefas os objetos são manipulados com a boca, comumente exibindo falta de destreza. Em uma única cena, durante todo o filme, 3 cachorros aparecem sentados jogando baralho, nesta ocasião um dos cachorros segura suas cartas com uma das patas, utilizando-a como se fosse uma mão, inclusive com o polegar por sobre as cartas. Em outra sequência ainda alguns cachorros pilotam aviões. Eles estão sentados no *cockpit* como cachorros, não como humanos, e o espectador não consegue ver suas patas. Assim os autores se eximiram de explicar como exatamente os personagens estão executando esta tarefa, desviando a atenção do espectador para a ação da cena.

Antropomorfismo na narrativa: O antropomorfismo, nesta obra, é usado de modo bastante sutil e artificiosos, como o da coleira e o do *cockpit*, possibilitam a ocorrência de certos eventos na narrativa, sem chamar a atenção do espectador para a não naturalidade dos personagens.

Frozen – Uma Aventura Congelante

Resumo da narrativa: Frozen (Buck e Lee, 2012) narra a história de Elza uma herdeira de um trono Norueguês, portadora de um poder mágico, ela congela tudo o que toca. Incapaz de controlar este poder ela se refugia de sua cidade, não sem antes, por acidente, lançar sobre ela um inverno eterno. Sua irmã, Ana, ajudada por um homem da montanha e seu alce, vai em busca de Elza, na intenção de trazer paz a sua irmã e a toda a cidade.

Anatomia dos personagens: Nesta obra foram encontradas três espécies de animais: cavalos, lobos e alces. Cavalos e lobos aparecem de modo bastante secundário na trama e são representados de modo absolutamente naturalista, ou seja, seus movimentos não apresentam, em nenhum momento, traços antropomórficos. Já o alce, personagem coadjuvante no núcleo principal da trama, possui alguns traços antropomórficos.

Movimentos Antropomórficos: Durante toda a obra o alce locomove-se como quadrúpede, executa ações como andar, galopar e comer de modo naturalista. Para tornar o personagem mais palatável a audiência os autores optaram por fazer o alce atuar como um cachorro nos momentos da narrativa onde havia necessidade de que este expressasse afetividade para com outros personagens; assim nestes momentos ele pula, gira em torno do próprio rabo, abana o rabo e lambe outros personagens.

Não podemos classificar estes movimentos como naturalistas, pois não são naturais a um alce, nem como antropomórficos, pois não são características humanas. Este é um artifício zoomórfico para comunicar os conceitos desejados pelos autores, com maior facilidade.

Em algumas outras situações, no entanto, o alce apresenta traços de antropomorfismo. É o caso quando seu dono, ao conversar com ele, faz duas vozes, uma voz para si próprio e outra para o alce; como se fosse um ventríloquo. Nestes casos o alce mexe a cabeça corroborando, e tornando mais expressivo, aquilo que é falado pelo ventríloquo. Em outra ocasião, induzido por outro personagem, o alce vira sua cabeça para cima de modo a olhar contemplativamente o céu, de maneira antropomórfica.

Antropomorfismo na narrativa: A humanização do movimento do alce é utilizada, nesta obra, de modo delicado, caracterizando o elo de amizade que tem para com seu dono sem destitui-lo de sua função de animal na história.

APÊNDICE B – Poses Chave Ser Humano

Poses chave para o mapeamento das poses expressivas.

Personagem em Repouso



Poses de Movimento Individual de Membros



Poses de Movimento Coordenado de Membros



Poses de Contato Entre Superfícies



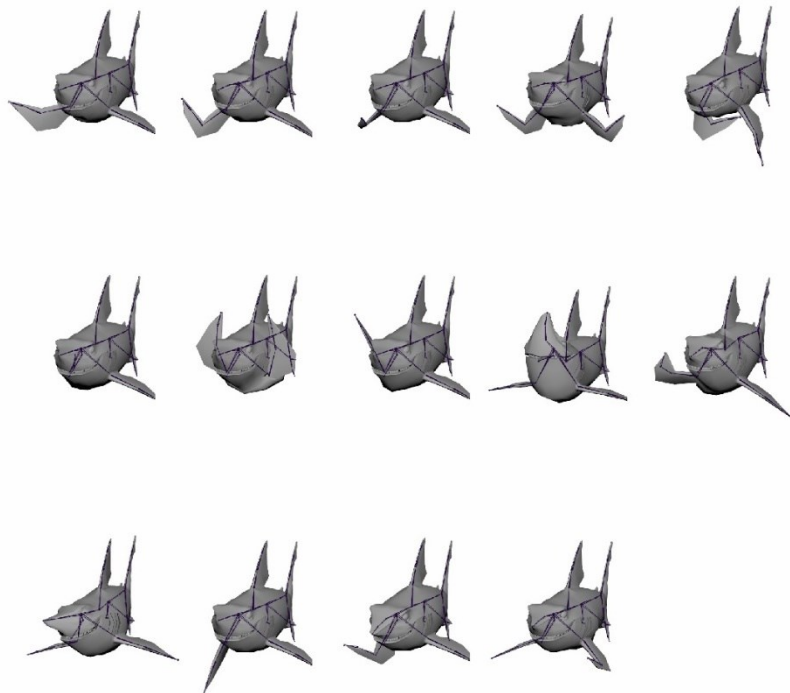
APÊNDICE C – Poses Chave Tubarão

Poses chave para o mapeamento das poses expressivas a um modelo de um tubarão.

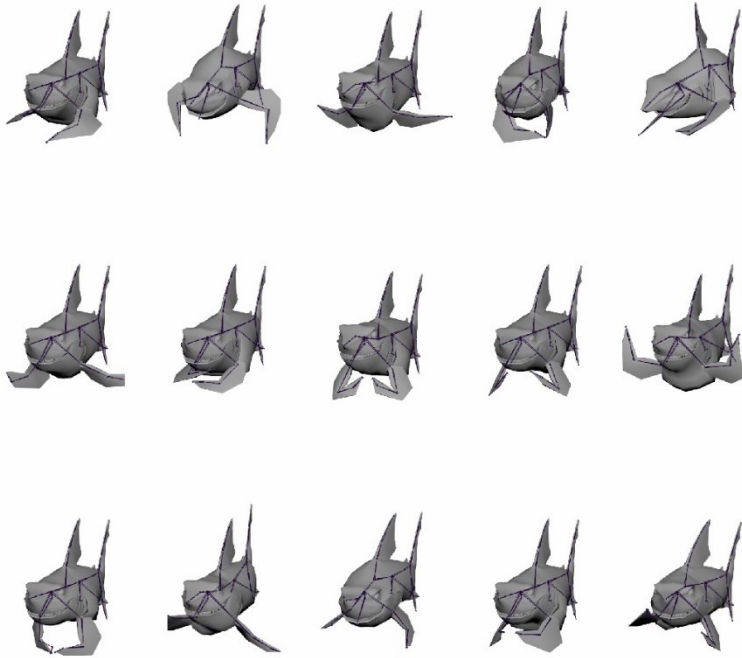
Personagem em Repouso



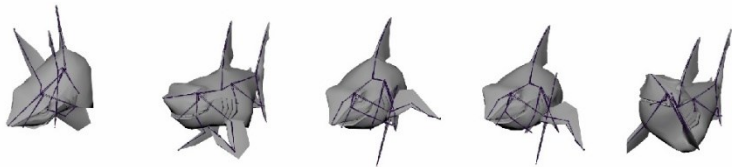
Poses de Movimento Individual de Membros



Poses de Movimento Coordenado de Membros



Poses de Contato Entre Superfícies



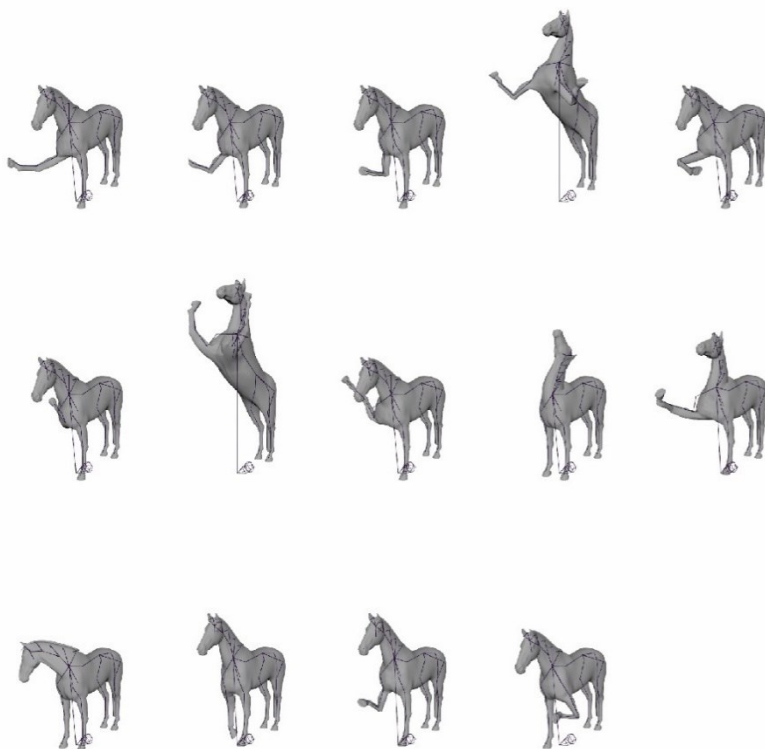
APÊNDICE D – Poses Chave Cavalos Tipo 1

Poses chave para o mapeamento das poses expressivas a um modelo quadrúpede em estilo de representação similar à do personagem burro da obra Shrek.

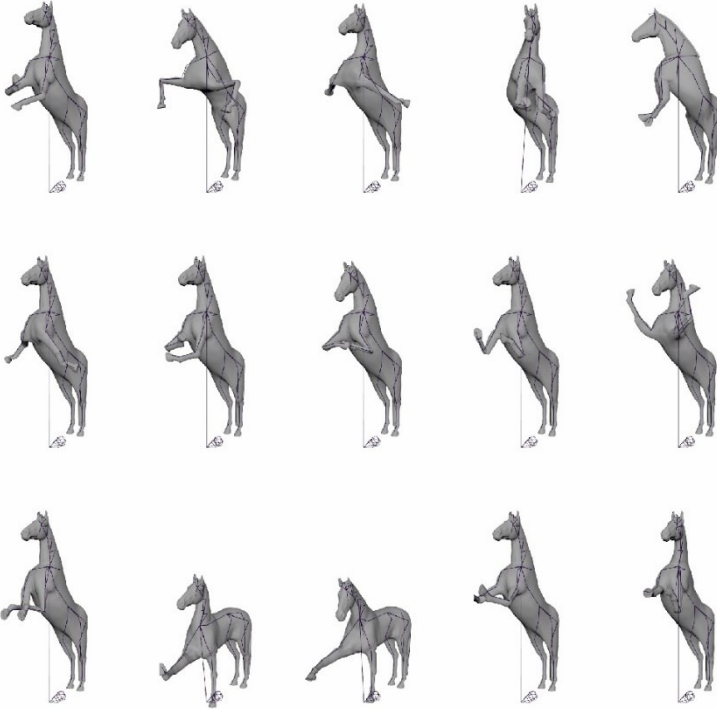
Personagem em Repouso



Poses de Movimento Individual de Membros



Poses de Movimento Coordenado de Membros



Poses de Contato Entre Superfícies



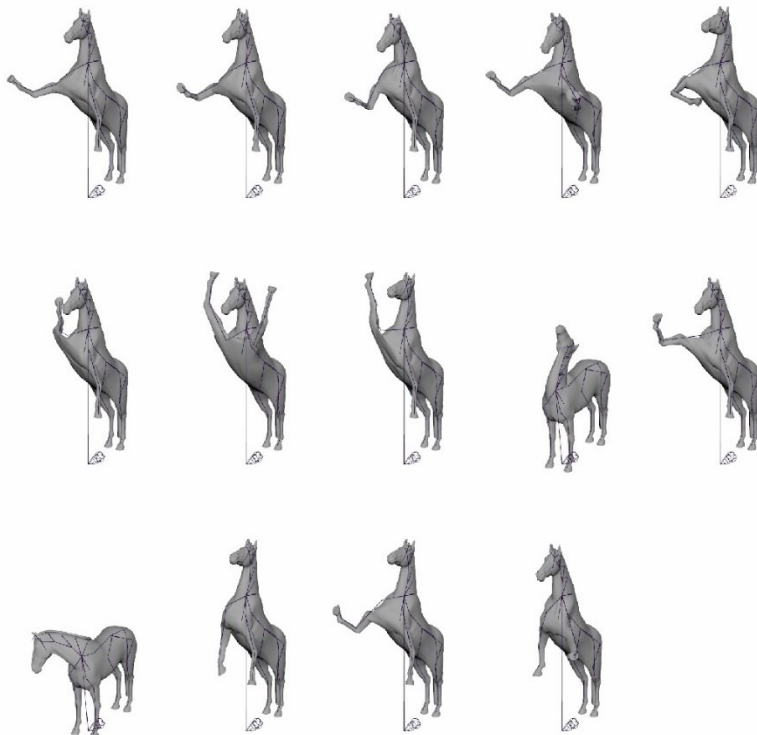
APÊNDICE E – Poses Chave Cavallo Tipo 2

Poses chave para o mapeamento das poses expressivas a um modelo quadrúpede em estilo de representação similar à do personagem zebra da obra Madagascar.

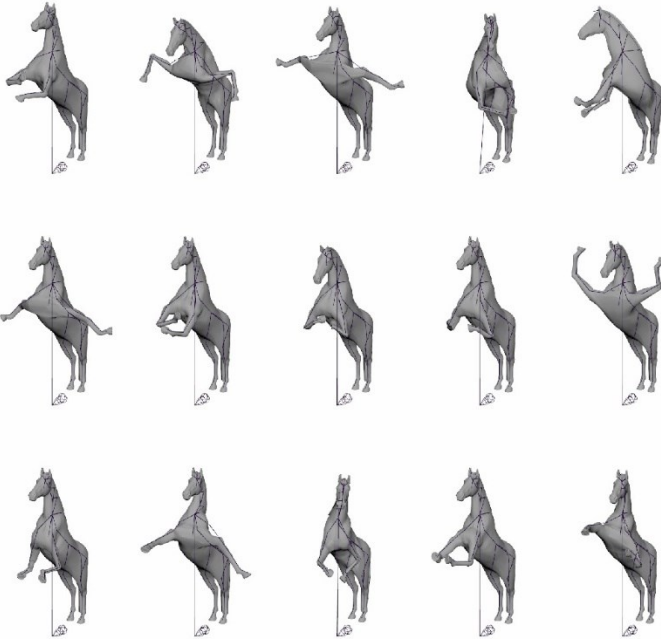
Personagem em Repouso



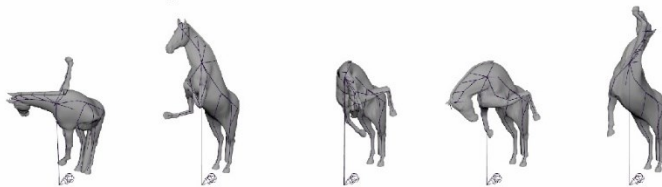
Poses de Movimento Individual de Membros



Poses de Movimento Coordenado de Membros



Poses de Contato Entre Superfícies



APÊNDICE F – Termo de Consentimento Livre e Esclarecido

Prezado (a) Senhor (a):

Você está sendo convidado(a) a participar do projeto “ANIMAÇÃO DE PERSONAGENS ANIMAIS NÃO HUMANOS POR CAPTURA DE MOVIMENTOS” de forma **totalmente voluntária**; Antes de concordar em participar desta pesquisa é muito importante que você compreenda as informações e instruções contidas neste documento; o conteúdo disponibilizado nesta página deverá esclarecer suas dúvidas antes de você decidir participar; Você tem o direito de **desistir** de participar da pesquisa a qualquer momento, sem quaisquer prejuízos e penalidades.

Objetivo e Justificativa do estudo: A presente pesquisa visa verificar a eficácia de uma nova ferramenta de desenvolvimento de animação de personagens. Caso esta ferramenta se mostre eficaz seu uso deve proporcionar redução de custos no campo da criação de conteúdo para mídias digitais.

Procedimentos: Sua participação nesta pesquisa consistirá apenas na visualização e classificação de um vídeo. Você assistira o vídeo de um personagem animado, este personagem pode ser humano ou outro animal. Posteriormente você será questionado a respeito da natureza da atividade desempenhada pelo personagem durante a animação, por meio de um formulário. Após o preenchimento do formulário encerra-se a pesquisa.

Benefícios: Esta pesquisa trará maior conhecimento para a comunidade científica sobre a eficácia de certas técnicas no processamento de dados de captura de movimentos e seu uso na animação de personagens, mas sem um benefício direto a você.

Riscos: Todos cuidados foram tomados para limitar, tanto quanto possível, a incidência de riscos aos participantes. Toda a pesquisa é realizada por meio digital, dentro deste site, evitando eventuais desconfortos ou constrangimentos que pudessem vir a ocorrer quando do deslocamento do voluntário a um laboratório de coleta de dados. De modo a evitar riscos advindos de uma eventual quebra de sigilo nos bancos de dados dos pesquisadores optou-se por não coletar nenhum dado que o identifique, como seu nome, e-mail, ou IP de conexão à internet. A pesquisa pode ser interrompida a qualquer momento de acordo com a sua necessidade. Assim, busca-se reduzir os riscos de ordem física ou psicológica para você. Caso haja qualquer risco ou dano significativo a você que estejam ou não previstos neste TCLE, o fato será comunicado imediatamente ao Sistema CPEP/CONEP. Este estudo pode ser adequado ou suspenso caso seja constatada necessidade, visando oferecer segurança ao participante.

Sigilo: Somente a equipe envolvida na pesquisa terá conhecimento dos dados. Nenhum dado pessoal seu será exigido. Eventualmente os dados referentes a classificação dos vídeos serão publicados de forma agregada sem menção a este ou aquele participante.

Custos: A pesquisa é realizada de forma voluntária e gratuita.

Ao clicar em “EU ACEITO” você, como participante afirma que está ciente de tudo que foi anteriormente exposto e concorda com sua participação no estudo.

Os pesquisadores responsáveis (Prof. Milton Luiz Horn Vieira e Prof. Gustavo Eggert Boehs), que também confirmam aceite deste termo, comprometem-se em conduzir a pesquisa de acordo com o que preconiza a Resolução 466/12 de 12/06/2012, que trata dos preceitos éticos e da proteção aos participantes da pesquisa.