



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Ackley Dias Will

**Modelagem e síntese de aspectos faciais da
Língua de Sinais Brasileira para avatares
sinalizantes**

Campinas

2021

Ackley Dias Will

Modelagem e síntese de aspectos faciais da Língua de Sinais Brasileira para avatares sinalizantes

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica, na Área de Engenharia de Computação.

Orientador: Prof. Dr. José Mario De Martino

Este trabalho corresponde à versão final da tese defendida pelo aluno Ackley Dias Will, e orientada pelo Prof. Dr. José Mario De Martino.

Campinas

2021

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

W66m Will, Ackley Dias, 1979-
Modelagem e síntese de aspectos faciais da Língua de Sinais Brasileira para avatares sinalizantes / Ackley Dias Will. – Campinas, SP : [s.n.], 2021.

Orientador: José Mario De Martino.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Animação por computador. 2. Língua Brasileira de Sinais. 3. Avatares. 4. Tecnologia assistiva. I. De Martino, José Mario, 1958-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Modeling and synthesis of facial aspects of Brazilian Sign Language for signing avatars

Palavras-chave em inglês:

Computer animation

Brazilian Sign Language

Avatars

Assistive technology

Área de concentração: Engenharia de Computação

Titulação: Doutor em Engenharia Elétrica

Banca examinadora:

José Mario De Martino [Orientador]

Léo Pini Magalhães

Antonielle Cantarelli Martins

Emely Pujólli da Silva

Messias Ramos Costa

Data de defesa: 14-12-2021

Programa de Pós-Graduação: Engenharia Elétrica

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0001-7828-100>

- Currículo Lattes do autor: <http://lattes.cnpq.br/4484563500611177>

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato: Ackley Dias Will

RA: 179529

Data de defesa: 14 de dezembro de 2021

Título da Tese: Modelagem e síntese de aspectos faciais da Língua de Sinais Brasileira para avatares sinalizantes

Prof. Dr. José Mario De Martino (Presidente)

Prof. Dr. Léo Pini Magalhães

Profa. Dra. Antonielle Cantarelli Martins

Profa. Dra. Emely Pujólli da Silva

Prof. Dr. Messias Ramos Costa

A Ata de Defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

À minha esposa Alexandra e aos meus filhos, Kevin e Richard.

Agradecimentos

Agradeço a Deus pelo dom da vida e por seu constante cuidado.

Agradeço ao professor José Mario, pela orientação, paciência e compreensão durante essa jornada.

Agradeço aos meus amigos e colegas pelo apoio nos trabalhos no laboratório.

Agradeço ao Juliano pelo desenvolvimento e aprimoramento do Avatar.

Agradeço às intérpretes de Libras Márcia Camargo e Juliana Moraes, pelas gravações e capturas e a professora Janice pelas transcrições.

Agradeço a minha família e amigos pela força e motivação. À minha esposa, Alexandra e a meus filhos, Kevin e Richard. Aos meus pais pelo incentivo constante.

Agradeço aos professores e funcionários da Faculdade de Engenharia Elétrica e de Computação da Unicamp que de alguma forma fizeram parte da minha trajetória.

“A verdadeira viagem de descobrimento não consiste em procurar novas paisagens, mas em ter novos olhos.”
(Marcel Proust)

Resumo

Na disseminação de informação na [Língua de Sinais Brasileira \(Libras\)](#), muitas vezes, são utilizados vídeos de sinalizações realizadas por surdos e intérpretes. Contudo, esse tipo de abordagem apresenta, muitas vezes, problemas na recuperação e apresentação de vídeos em tempo real. Uma alternativa utilizada e que vem ganhando a atenção nos últimos anos é o desenvolvimento de avatares tridimensionais sinalizantes em sistemas de tradução. Contudo, apesar da vantagem dessa tecnologia assistiva, ainda existem críticas quanto a apresentação de parâmetros não-manuais, em particular as expressões faciais nos avatares sinalizantes atuais. O foco deste trabalho é a modelagem e síntese de expressões faciais da [Libras](#) para avatares tridimensionais sinalizantes desenvolvidos em estilo realista. Como resultado final, foi desenvolvida uma metodologia para a criação de animações faciais para serem incorporadas à avatares sinalizantes. Os trechos das expressões que compõem a animação facial, foram criados a partir de um processo de captura de movimento facial baseado em marcadores e para isso foram desenvolvidas abordagens específicas para a criação de *layouts* otimizados de marcadores faciais e de mapeamento dos dados capturados ao rosto do avatar. Foi criado e gravado também, um corpus de sinalizações com 270 sentenças nas modalidades afirmativa, interrogativa total e interrogativa parcial. As análises das anotações deste corpus apontaram as expressões faciais relacionadas as modalidades sinalizadas e serviu como base para o desenvolvimento de regras para a criação da animação facial final. As expressões faciais obtidas através de captura de movimento são inseridas na animação final de acordo com os dados de entrada, com informações sobre a sentença a ser sinalizada, e as regras criadas. A aplicação dessa metodologia permite a criação e incorporação de animações faciais realistas para acompanhar as animações corporais e manuais em avatares sinalizantes.

Palavras-chave: Animação facial; Libras; Avatares sinalizantes; Tecnologia assistiva.

Abstract

In the dissemination of information in the Brazilian sign language (Libras), videos of signings made by deaf people and interpreters are often used. However, this type of approach often presents problems in real-time video retrieval and presentation. A commonly used alternative that has been gaining attention in recent years is the development of three-dimensional signing avatars in translation systems. However, despite the advantage of this assistive technology, there are still criticisms regarding the deficiency in the presentation of non-manual parameters, in particular the facial expressions in current signing avatars. The focus of this work is the modeling and synthesis of Libras facial expressions for three-dimensional signing avatars developed in a realistic style. As a final result, a methodology for creating facial animations to be incorporated into signing avatars was developed. The excerpts of the expressions that make up the facial animation were created from a process of facial motion capture based on markers, and for this, specific approaches were developed for the creation of optimized layouts of facial markers and mapping the captured data to the avatar's face. A corpus of signals with 270 sentences in the affirmative and interrogative modalities was also created and recorded. The analysis of the annotations of this corpus pointed out the facial expressions related to the signaled modalities and served as a basis for the development of rules for the creation of the final facial animation. The facial expressions obtained through motion capture are inserted into the final animation according to the input data, with information about the sentence to be signaled, and the rules created. The application of this methodology allows the creation and incorporation of realistic facial animations to accompany the body and manual animations in signing avatars.

Keywords: Facial animation; Libras; Signing avatars; Assistive technology.

Lista de ilustrações

Figura 2.1 – Ilustração do fenômeno <i>believability flip</i> (PARKE; WATERS, 2008a) . . .	21
Figura 2.2 – <i>Uncanny valley</i> (adaptado de Mori, MacDorman e Kageki (2012)) . . .	22
Figura 2.3 – Ossos do crânio (adaptado de Sobotta et al. (2001))	23
Figura 2.4 – Músculos da face (adaptado de Sobotta et al. (2001))	23
Figura 2.5 – Pontos característicos do padrão MPEG-4 (ISO, 2001)	26
Figura 2.6 – Monitoramento de pontos faciais articulatórios da fala (DE MARTINO, 2005)	27
Figura 2.7 – (A) Músculos faciais modelados. (B) Resultados de movimentação (adaptado de Corrêa, Magalhães e De Martino (2008))	28
Figura 2.8 – Configuração padrão Vicon de 53 marcadores para captura corporal (VICON, 2020)	30
Figura 2.9 – Sistema Vicon Cara posicionado na cabeça de uma modelo (VICON, 2015)	31
Figura 3.1 – a) <i>Layout</i> denso de captura e b) Demarcações das regiões faciais . . .	34
Figura 3.2 – Regiões do rosto (adaptado de Sobotta et al. (2001))	35
Figura 3.3 – a) Ator com os marcadores posicionados e utilizando o equipamento de captura de movimento Vicon Cara. b) Visualização das câmeras durante o processamento dos dados capturados	36
Figura 3.4 – Fluxograma do algoritmo de agrupamento e determinação de marcadores representativos	38
Figura 3.5 – Ilustração de uma sequência de agrupamentos dos marcadores da região da bochecha	39
Figura 3.6 – Fluxograma do algoritmo de ajustes de marcadores	40
Figura 3.7 – Exemplo de ajuste de agrupamento. Esquerda: Agrupamento original. Direita: Agrupamento ajustado	40
Figura 3.8 – Ilustração de espelhamento de marcadores (esquerda) e agrupamentos e <i>layouts</i> obtidos através de coeficiente de refinamentos $\text{coef}=0.9$, $\text{coef}=1.0$ e $\text{coef}=1.1$ respectivamente	41
Figura 3.9 – Visão geral do processo de <i>retargeting</i> facial (WILL; DE MARTINO; BEZERRA, 2019)	42
Figura 3.10–Exemplo de distâncias em cada eixo para os cálculos de proporcionalidades	43
Figura 3.11–Exemplo de criação dos vetores iniciais utilizados no processo de formação da base vetorial de cada <i>layout</i>	44
Figura 3.12–Visão geral do processo de aplicação dos vetores de deslocamento . . .	46

Figura 3.13–Ilustração das distâncias entre os pontos mais distantes da boca na configuração de ossos controladores e no <i>layout</i> de marcadores para o cálculo do valor proporcional	47
Figura 3.14–Resultados da aplicação do <i>retargeting</i> facial em diferentes avatares realistas	48
Figura 4.1 – Tempo de execução de uma expressão facial gramatical (REILLY, 2005)	50
Figura 4.2 – Tempo de execução de uma expressão facial de emoção (REILLY, 2005)	50
Figura 4.3 – (1) Sinalização de VAI em uma sentença interrogativa total (do tipo sim/não). (2) Sinalização de ONDE em uma sentença interrogativa parcial. (3) Sinalização de POR-QUE em uma sentença interrogativa parcial	52
Figura 4.4 – Avatares dos sistemas de tradução: Hand Talk, Rybená e VLibras (respectivamente da esquerda para a direita)	53
Figura 4.5 – Sinalização dos sinais: ALEGRIA, TRISTEZA, GORDO e MAGRO realizados por uma intérprete de Libras e por avatares sinalizantes de Libras (parcialmente retirado de (FELIPE, 2013))	55
Figura 4.6 – Sinalização de sentenças interrogativas totais (sim/não)	56
Figura 4.7 – Sinal de interrogação dos avatares dos sistema (de cima para baixo) Hand Talk, Rybená e VLibras. O símbolo em vermelho de interrogação foi inserido para representar o movimento realizado pela mão	57
Figura 4.8 – Sinalização de sentenças interrogativas parciais	58
Figura 5.1 – Visão geral da metodologia de criação de animações faciais	61
Figura 5.2 – Sessão de gravação do corpus	64
Figura 5.3 – Anotação das sentenças utilizando o ELAN	65
Figura 5.4 – Exemplo de expressões faciais executadas pela intérprete	66
Figura 5.5 – Exemplo de situação com a expressão facial do tipo B executada durante o sinalização da glosa-chave <i>QUANDO</i> e a expressão facial do tipo A durante os dois sinais anteriores	67
Figura 5.6 – Sinalizante com os marcadores e equipamento de captura de movimento	75
Figura 5.7 – Processamento dos dados no software <i>Vicon Post</i>	75
Figura 5.8 – Ilustração dos processos de ajuste de duração de animações. (A) Redução da duração da animação. (B) Aumento da duração da animação	77
Figura 5.9 – Processo de criação da animação final	78
Figura 5.10–Ilustração de representação da curva paramétrica cúbica de Hermite	80
Figura 6.1 – Arquitetura geral do sistema piloto	82
Figura 6.2 – Ilustração da anotação dos sinais e transições na animação corporal da sentença: HOMEM CONSTRUIR CASAS COMO?	84

Figura 6.3 – Deslocamento no eixo X de um marcador ao longo da animação facial de uma sentença. (A) Animações faciais inseridas na animação final. (B) Inserção das animações de transições criadas	86
Figura 6.4 – Avatar do projeto TAS utilizado para a aplicação das animações	86
Figura 6.5 – Linha 1: Expressão facial neutra. Linha 2: Expressão facial do tipo B (sobrancelhas franzidas e lábios projetados). Coluna A: Ossos controladores. Coluna B: Sistema de ossos de deformação. Coluna C: Expressões faciais resultantes	87
Figura 6.6 – Sinalização da sentença: "CASAS HOMEM EL@ VENDER VAI?". I) Sinalizações realizadas pela intérprete; II) Sinalizações realizadas pelo avatar; III) Destaque das expressões faciais executadas pelo avatar. . .	90
Figura 6.7 – Deslocamento do marcador mais ao centro da sobrancelha esquerda na sinalização da sentença: CASAS HOMEM EL@ VENDER VAI?. (A) Deslocamento durante toda a animação. (B) Deslocamento das transições criadas	91
Figura 6.8 – Sinalização da sentença: "HOMEM CONSTRUIR CASAS COMO?". I) Sinalizações realizadas pela intérprete; II) Sinalizações realizadas pelo avatar; III) Destaque das expressões faciais executadas pelo avatar. . .	92
Figura 6.9 – Deslocamento do marcador mais ao centro da sobrancelha esquerda na sinalização da sentença: HOMEM CONSTRUIR CASAS COMO?. (A) Deslocamento durante toda a animação. (B) Deslocamento das transições criadas	93
Figura C.1 – Repositório contendo os arquivos das animações de expressões faciais .	117
Figura E.1 – Tela inicial do Blender ao ser aberto o arquivo de animação corporal .	124
Figura E.2 – Quadros da animação da sentença CASAS HOMEM EL@ VENDER VAI? no Blender	124

Lista de tabelas

Tabela 3.1 – Quantidade de marcadores em cada região facial	34
Tabela 4.1 – Características dos sistemas de tradução português-Libras informadas na plataforma <i>Play Store</i> (verificado em: 14/03/2021).	53
Tabela 5.1 – Expressões faciais realizadas pela intérprete durante a sinalização das glosas-chave	65
Tabela 5.2 – Tabulamento dos dados de expressões faciais observados durante as sinalizações das sentenças do corpus	68
Tabela 5.3 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave <i>QUANDO</i>	69
Tabela 5.4 – Regras de diferentes tipos de sentenças	73
Tabela 5.5 – Sentenças selecionadas para captura das expressões faciais buscadas	74
Tabela B.1 – Dados Tabulados da modalidade interrogativa total	109
Tabela B.2 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave <i>ONDE</i>	110
Tabela B.3 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave <i>COMO</i>	111
Tabela B.4 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave <i>POR-QUE</i>	112
Tabela B.5 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave <i>O-QUE</i>	113
Tabela B.6 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave <i>QUAL</i>	114
Tabela B.7 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave <i>QUANT@</i>	115

Lista de abreviaturas e siglas

ASL *American Sign Language.*

CEP Comitê de Ética em Pesquisa.

CTS Centro de Tecnologia de Software.

EFGD Expressões Faciais Gramaticais de Distinção.

EFGI Expressões Faciais Gramaticais de Intensidade.

EFGS Expressões Faciais Gramaticais para Sentenças.

ENM Expressões Não-Manuais.

FAP *Facial Animation Parameter.*

FP Futuro do Presente.

FPs *Feature Points.*

Libras Língua de Sinais Brasileira.

Mocap *Motion Capture.*

PP Pretérito Perfeito.

PS Presente.

TAS Tecnologia Assistiva para Surdos.

Sumário

1	INTRODUÇÃO	18
1.1	Descrição e motivação	18
1.2	Contribuições e publicações	19
1.3	Organização do trabalho	20
2	ANIMAÇÃO FACIAL	21
2.1	Introdução	21
2.2	Anatomia da face	22
2.3	Técnicas de manipulação da geometria da face	24
2.3.1	Animação facial baseada em modelo tridimensional	24
2.4	Mecanismos de definição e controle da animação	28
2.5	Captura de movimento - Mocap	29
2.6	Comentários Finais	30
3	ABORDAGENS PARA CAPTURA DE MOVIMENTO FACIAL	32
3.1	Introdução	32
3.2	Posicionamento otimizado de marcadores para Mocap facial	32
3.2.1	Layout denso e regiões faciais	34
3.2.2	Captura e processamento	35
3.2.3	Matrizes de posições	35
3.2.4	Matrizes de correlações	36
3.2.5	Agrupamento dos marcadores	37
3.2.6	Ajustes dos marcadores	38
3.2.7	Resultados	39
3.3	Mapeamento dos dados de Mocap	40
3.3.1	Marcadores e ossos controladores	42
3.3.2	Alinhamento dos dados de Mocap	42
3.3.3	Vetores de deslocamento	45
3.3.4	Ajustes por região	46
3.3.5	Aplicação do <i>retargeting</i>	47
3.4	Comentários finais	47
4	EXPRESSÕES FACIAIS NA LIBRAS	49
4.1	Introdução	49
4.2	Expressões faciais na Libras	50
4.3	Expressões faciais em avatares sinalizantes de Libras	52

4.3.1	Expressões faciais como componentes gramaticais de palavras	54
4.3.2	Expressões faciais em sentenças interrogativas totais	54
4.3.3	Expressões faciais em sentenças interrogativas parciais	55
4.4	Comentários Finais	59
5	METODOLOGIA PARA CRIAÇÃO E APLICAÇÃO DE EXPRES-	
	SÕES FACIAIS NA LIBRAS	60
5.1	Introdução	60
5.2	Etapa de análise	61
5.2.1	Criação das sentenças do corpus	62
5.2.2	Gravação e anotação do corpus	63
5.2.3	Tabelamento dos dados	66
5.2.4	Análise das anotações	70
5.2.5	Criação do banco de regras	72
5.3	Etapa de aquisição	74
5.3.1	Captura de movimento	74
5.3.2	<i>Retargeting</i>	75
5.3.3	Recorte e criação do banco de expressões faciais	76
5.4	Etapa de animação	76
5.4.1	Verificação da sentença e regras	76
5.4.2	Seleção e ajustes das animações faciais	77
5.4.3	Finalização da animação	77
5.4.4	Criação das transições	79
5.5	Comentários finais	80
6	IMPLEMENTAÇÃO PILOTO	82
6.1	Introdução	82
6.1.1	Sentenças e anotações	82
6.1.2	Verificação dos dados de entrada e aplicação das regras	83
6.1.3	Seleção das animações de expressões faciais	85
6.1.4	Criação da Animação facial final	85
6.2	Execução da animação final	85
6.3	Criação e execução das animações faciais	87
6.3.1	Animação da sentença: CASAS HOMEM EL@ VENDER VAI?	88
6.3.2	Animação da sentença: HOMEM CONSTRUIR CASAS COMO?	89
6.4	Comentários finais	92
7	CONCLUSÃO	95

REFERÊNCIAS	97
APÊNDICES	104
APÊNDICE A – SENTENÇAS DO CORPUS	105
APÊNDICE B – DADOS TABULADOS	109
APÊNDICE C – EXECUÇÃO DO SISTEMA PILOTO	116
APÊNDICE D – RELATÓRIO DE EXECUÇÃO DA CRIAÇÃO DA ANIMAÇÃO	119
APÊNDICE E – EXECUÇÃO DA ANIMAÇÃO FACIAL PELO AVATAR	123
ANEXOS	125
ANEXO A – DOCUMENTOS DA PESQUISA	126

1 Introdução

1.1 Descrição e motivação

As expressões faciais desempenham um papel importante na demonstração de emoção e são comumente utilizadas para complementar a comunicação verbal, auxiliando na transmissão da mensagem. Desde muito cedo, os seres humanos são expostos à diferentes expressões faciais e, ao longo do tempo, o aperfeiçoamento dessa habilidade permite reconhecermos e interpretarmos até mesmo pequenas variações na face. Dessa forma, (PARKE, 1972) aponta que o desenvolvimento de animação facial realista se torna um desafio. Conceitos como o *believability flip* apresentado por (PARKE; WATERS, 2008a) e o *uncanny valley* pontuado por (MORI; MACDORMAN; KAGEKI, 2012) apresentam a dificuldade de alcançar este objetivo. O conceito de animação facial é caracterizado por técnicas de manipulação e transformação das características da face ao longo do tempo resultando em sensação de movimento.

Nas línguas de sinais, as expressões faciais podem ser caracterizadas, de uma maneira geral, por expressões afetivas ou gramaticais (REILLY, 2005). Segundo (QUADROS; KARNOPP, 2004), nas línguas de sinais as expressões faciais gramaticais, juntamente com os movimentos dos olhos, da cabeça e troco, fazem parte das *Expressões Não-Manuais (ENM)* e podem ser utilizadas para marcação de construções sintáticas e diferenciação de itens lexicais.

Ao longo dos últimos anos, avatares sinalizantes de língua de sinais e especialmente da *Língua de Sinais Brasileira (Libras)*, foram desenvolvidos para apresentar conteúdo escrito/falado traduzido para a língua de sinais. Contudo, ainda existem críticas quanto às limitações da representação de parâmetros não-manuais, em particular expressões faciais nos avatares sinalizantes atuais (VIEIRA et al., 2014). Tendo em vista essa deficiência e dada a importância das expressões faciais nas línguas de sinais, este trabalho apresenta uma metodologia de criação e incorporação de animação facial em avatares sinalizantes realistas para modular a sinalização juntamente com a animação corporal e manual.

A animação facial resultante da aplicação da metodologia é composta por trechos de animações de expressões faciais obtidas através de captura de movimento de sinalizações de *Libras*. A captura de movimento, ou *Motion Capture (Mocap)*, utilizou um sistema óptico baseado em marcadores. O *layout* de marcadores utilizado nas capturas foi elaborado através de uma abordagem de criação de *layouts* otimizados desenvolvida no âmbito deste trabalho. Da mesma forma, o processo utilizado de mapeamento dos dados

capturados ao rosto do avatar foi desenvolvido no contexto deste trabalho. Por fim, para demonstrar a viabilidade e potencial para a criação e aplicação de expressões faciais na [Libras](#) foi desenvolvido um protótipo baseado nos conceitos e metodologias apresentadas.

1.2 Contribuições e publicações

As principais contribuições deste trabalho são:

1. Desenvolvimento de uma abordagem para a criação de *layouts* otimizados para a captura de movimento facial;
2. Desenvolvimento de uma abordagem de mapeamento de dados de [Mocap](#) ao rosto do avatar (*retargeting*);
3. Análise de expressões faciais gramaticais da [Libras](#) através de um corpus contendo sentenças interrogativas totais e parciais;
4. Formalização de regras de aplicação de expressões faciais para um sistema de animação facial;
5. Elaboração de uma metodologia de criação de animações faciais para acompanharem animações corporais e manuais em um avatar sinalizante de língua de sinais.

Durante o desenvolvimento das pesquisas para a presente tese, foram publicados os seguintes trabalhos:

- PAIVA, F. A. S.; WILL, A. D.; DE MARTINO, J. M.; BARBOSA, P. A.; BENETTI, A. B. Incorporating non-manual expressions into a realistic signing avatar. In: IX Congreso Iberoamericano de Tecnología de Apoyo a la Discapacidad, 2017, Bogotá. Memorias IX Congreso Iberoamericano de Tecnología de Apoyo a la Discapacidad, 2017. p. 551-558.
- PAIVA, F. A. S.; BARBOSA, P. A.; DE MARTINO, J. M.; WILL, A. D.; OLIVEIRA, M. R. N. S.; SILVA, I. R.; XAVIER, A. N. Análise do Papel das Expressões Não Manuais na Intensificação em Libras. DELTA, São Paulo, v. 34, n. 4, p. 1135-1158, 2018.
- WILL, A. D.; MARTINO, J. M. D.; BEZERRA, J. An Optimized Marker Layout for 3D Facial Motion Capture. In: LIVESU, M.; PINTORE, G.; SIGNORONI, A. (Ed.). Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference. [S.l.]: The Eurographics Association, 2018. ISBN 978-3-03868-075-8. ISSN 2617-4855.

- WILL, A. D.; MARTINO, J. M. D.; BEZERRA, J. Automatic mapping of facial mocap data to a realistic signing avatar. In: Memórias X Congresso Iberoamericano de Tecnologías de Apoyo a la Discapacidad. [S.l.: s.n.], 2019. p. 431–434.
- WILL, A. D.; DE MARTINO, J. M.; Modeling and synthesizing realistic facial expressions for sign language 3D avatars. In: Eurographics 2021 Doctoral Consortium, 2021, Viena, Áustria. Eurographics 2021 Doctoral Consortium, 2021.

1.3 Organização do trabalho

O restante da tese apresenta a fundamentação teórica, motivação, metodologia proposta e implementação de um sistema piloto como forma de avaliação. Os capítulos estão organizados da seguinte maneira:

- Capítulo 2 - Animação facial: Revisão dos conceitos, técnicas e desafios associados ao desenvolvimento de animação facial por computador;
- Capítulo 3 - Abordagens para captura de movimento facial: Descrição da abordagem para a criação de *layouts* para Mocaps faciais otimizados e da abordagem de mapeamento dos dados capturados ao rosto do avatar sinalizante desenvolvidas no âmbito deste trabalho;
- Capítulo 4 - Expressões faciais na [Libras](#): Revisão sobre a importância da expressão facial na [Libras](#) e apresentação da motivação através da análise de animações de expressões faciais reproduzidas por avatares sinalizantes de [Libras](#) atuais;
- Capítulo 5 - Criação e aplicação de expressões faciais na [Libras](#): Apresentação da metodologia desenvolvida para a criação de expressões faciais na [Libras](#) desde a análise das expressões faciais e definição de regras de aplicação, aquisição da animação facial através de [Mocap](#) e criação da animação facial final;
- Capítulo 6 - Implementação piloto: Descrição e apresentação dos resultados do desenvolvimento da implementação piloto da metodologia de aplicação de expressões faciais.

2 Animação facial

2.1 Introdução

Este capítulo apresenta uma revisão dos desafios, conceitos e técnicas relacionados ao desenvolvimento de animações faciais por computador. Serão vistas estratégias de manipulação da geometria da face e os mecanismos de definição e controle da animação.

Desde o trabalho de Parke (PARKE, 1972), diversas pesquisas foram desenvolvidas na área de modelagem e animação realista de personagens tridimensionais como apresentado nos trabalhos de (WATERS, 1987), (GUENTER, 1989), (MAGNENAT-THALMANN; PRIMEAU; THALMANN, 1988), (BLANZ; VETTER, 1999), (DENG; NEUMANN, 2007), (PARKE; WATERS, 2008b), (LIN, 2011), (POTAMIAS et al., 2020).

A geração de animação facial realista e convincente é um desafio, especialmente porque os seres humanos possuem um sentido apurado para perceber até mesmo as mais sutis mudanças nas expressões faciais (PARKE, 1972). À medida que se aumenta o realismo, quanto a aparência e movimentação do personagem, pode-se chegar a um ponto em que começamos a perceber que algo não está mais certo. O personagem passa a ter uma aparência não natural, ocorrem problemas na fluidez da animação ou na sincronização labial fazendo com que paremos de acreditar no realismo do personagem (PARKE; WATERS, 2008a). Esse fenômeno é chamado de *believability flip* (Figura 2.1) e segundo o autor é mais presente quando o personagem tenta reproduzir uma pessoa conhecida como por exemplo uma celebridade.

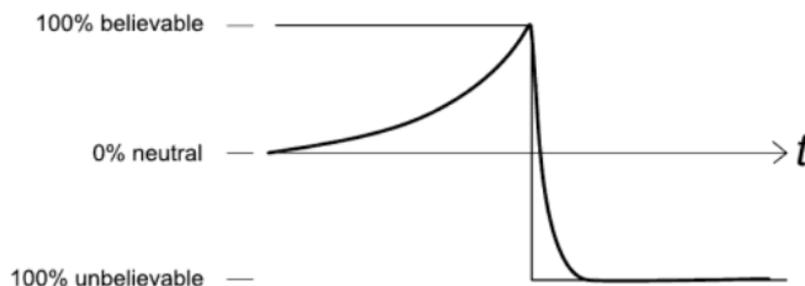


Figura 2.1 – Ilustração do fenômeno *believability flip* (PARKE; WATERS, 2008a)

Nessa mesma linha, Mori, MacDorman e Kageki (2012) criou um conceito conhecido como *uncanny valley* o qual está diretamente ligado a resposta emocional dos humanos para com os robôs ou para outras entidades não humanas. Esse conceito teoriza que quanto mais um personagem aumenta sua semelhança com um humano, em termos de aparência e movimentação, nossa empatia também aumenta até um ponto em que ocorre

uma reversão e o personagem passa a ter uma aparência repulsiva (Figura 2.2).

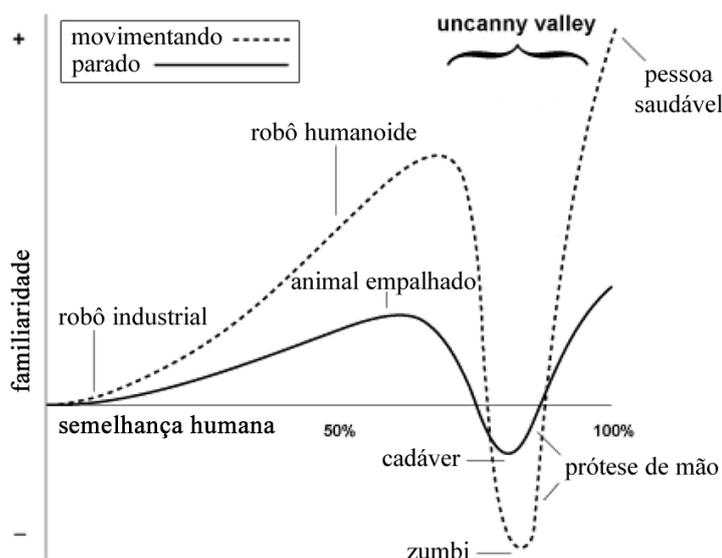


Figura 2.2 – *Uncanny valley* (adaptado de Mori, MacDorman e Kageki (2012))

As hipóteses apresentadas nos conceitos de *believability flip* e *uncanny valley* ressaltam a problemática em se desenvolver animações faciais realistas por computador. Contudo, ao longo dos anos, devido ao desenvolvimento de hardwares específicos e de avanços nas técnicas de computação gráfica muito progresso tem sido feito na área da animação facial e em especial, na animação facial realista.

2.2 Anatomia da face

Uma animação realista ideal é aquela que pode ser facilmente confundida com uma pessoa real, tanto na aparência quanto na movimentação. Para esse objetivo, é importante levar em consideração as movimentações que o modelo deve realizar já na fase do planejamento da modelagem.

Na busca pelo realismo é comum a comparação do modelo criado com um ator humano desde as fases iniciais de concepção do modelo e definição dos mecanismos de animação. Nesse contexto, o estudo da anatomia da face é de grande importância para a criação de expressões faciais realistas. Iyaniwura (2008) ressalta que as expressões faciais são realizadas através da contração e relaxamento de músculos e o entendimento da anatomia da face, o que inclui os mecanismos de movimentação da mandíbula, músculos e pele, é importante tanto para a modelagem quanto para a produção de animação facial.

Os ossos da cabeça são importantes para determinar as características físicas e individuais. Segundo Kitagawa e Windsor (2008) os ossos mais significativos que determinam essas características são a mandíbula, a maxila, o zygomatic (ossos das bochechas) e osso nasal (Figura 2.3). De Martino (2005) observa que a mandíbula serve como base para

os dentes inferiores e tem como função principal a mastigação e contribui para a produção da fala. A mandíbula é o único osso móvel da face. [Corrêa \(2007\)](#) lembra que diversas expressões faciais são compostas pela atuação desse osso, como exemplo da expressão de alegria quando, na produção do sorriso, a boca entreabre-se.

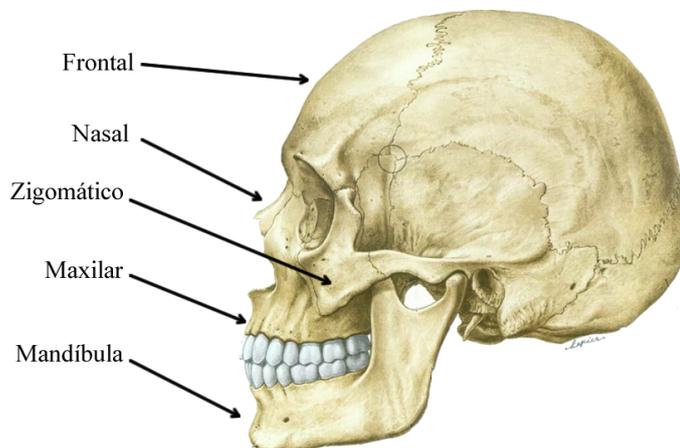


Figura 2.3 – Ossos do crânio (adaptado de [Sobotta et al. \(2001\)](#))

Os músculos faciais cobrem a parte frontal do crânio, gordura, cartilagem e outros tecidos da face (Figura 2.4). A combinação da atuação desses músculos cria as expressões faciais ([KITAGAWA; WINDSOR, 2008](#)).

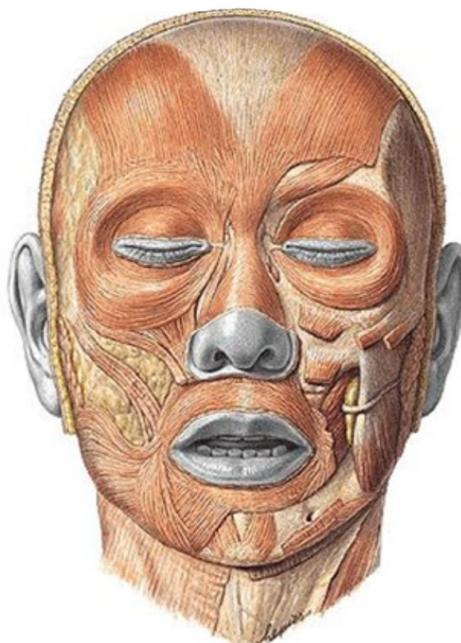


Figura 2.4 – Músculos da face (adaptado de [Sobotta et al. \(2001\)](#))

Um dos resultados que será apresentado é o desenvolvimento de uma abordagem para criação de *layout* de marcadores otimizados para a captura de movimento facial

que foi criado levando em consideração as regiões faciais determinadas por conjuntos de músculos e suas movimentações durante a reprodução de expressões faciais.

2.3 Técnicas de manipulação da geometria da face

Em um primeiro momento, as estratégias de manipulação da geometria da face para a criação da animação facial podem ser classificadas em duas linhas distintas, as baseadas em imagem onde é levado em consideração somente o espaço bidimensional e as animações baseadas em modelos tridimensionais. Dentre as técnicas baseadas em imagem destaca-se a estratégia de *morphing*, nela, imagens intermediárias são criadas através da combinação de imagens (normalmente fotográficas) que servem como pose-chave (LEE et al., 1996) (EZZAT; POGGIO, 1998) (PIGHIN et al., 1998) (BUTTFIELD, 2003). O presente trabalho tem como foco a animação facial baseada em modelo tridimensional. Esse tipo de animação é realizada através da manipulação dos polígonos que compõe a malha que descreve a face do modelo.

2.3.1 Animação facial baseada em modelo tridimensional

O trabalho de Parke (1972) é considerado o marco inicial da linha de pesquisa na área de animação facial baseada em modelo. A criação de animações baseadas em modelo tridimensional incluem as estratégias de interpolação de poses-chave (*keyframing*), parametrização geométrica, parametrização *data-driven* e simulação biomecânica (DE MARTINO, 2005) (LIN, 2011).

Interpolação de poses-chave

Na estratégia de interpolação de poses-chave, ou *keyframing*, são calculadas poses intermediárias dado um conjunto de poses-chave. Essas poses intermediárias definem, a movimentação da animação. Dessa forma, dada duas expressões faciais como poses-chave, é criada a animação da movimentação entre a pose inicial e final. A malha poligonal que define cada uma das poses-chave é definida pela posição dos vértices que compõe sua geometria. Os vértices das poses intermediárias são calculados pela interpolação das posições do vértice correspondente na pose-chave inicial e final. O trabalho pioneiro de Parke (1972) apresenta uma interpolação de poses-chave para produzir animações de expressões faciais. Em Chae e Kim (2020) é desenvolvida uma abordagem para a criação de animação de fala e de expressões de emoção usando um pequeno grupo de modelos de exemplo, onde as animações são realizadas através da aplicação de um método de interpolação de dados dispersos.

Um problema inerente a essa estratégia é o fato da necessidade de se ter

definidas a priori todas as poses que representam as poses-chave para que se possa, através de um algoritmo de interpolação, realizar as animações decorrente das poses intermediárias.

Parametrização geométrica

Diferentemente da estratégia de interpolação de poses-chave, que move diretamente os vértices da malha poligonal, a parametrização geométrica utiliza uma série de controles parametrizados de manipulação das movimentações e deformações distintas na face.

Parke (1974) apresenta o trabalho pioneiro sobre modelos paramétricos através da formulação de conceitos simples. Posteriormente, o autor apresentou uma redução de parâmetros através do método: "parametrização direta" (Parke, 1982). Nessa abordagem, uma pose facial é especificada através da manipulação de parâmetros faciais. Comparada com a estratégia de interpolação, esse método apresenta uma maneira mais intuitiva para o desenvolvimento de animação facial. Contudo, como podem existir muita diferença entre as faces, os parâmetros podem não funcionar da mesma maneira e uma nova configuração pode ser necessária (LIN, 2011).

Na abordagem de parametrização geométrica o modelo ideal seria um conjunto de parâmetros faciais que permita a reprodução de todas as possíveis expressões faciais através da seleção e manipulação de seus valores, contudo, um subconjunto de parâmetros podem permitir a reprodução de uma boa gama de expressões faciais. A busca e definição do conjunto de parâmetros é uma das dificuldades encontradas na parametrização geométrica. Os parâmetros podem ser obtidos pela simples observação da face, ou em uma abordagem mais complexa, pelo estudo das estruturas faciais. Nesse caso os parâmetros terão como base a anatomia facial. Após a observação e análise dos parâmetros necessários, é preciso desenvolver um modelo facial capaz de implementar esse conjunto de parâmetros (PARKE; WATERS, 2008a).

O padrão MPEG-4 (ISO, 2001) pode ser classificado como um modelo de parametrização geométrica. Uma face no padrão MPEG-4 é definida por 84 pontos característicos, ou *Feature Points (FPs)*, que definem a modelagem e animação da face. As animações faciais são realizadas através de parâmetros de animação facial conhecidos como *Facial Animation Parameter (FAP)* que controlam os pontos característicos. A Figura 2.5 apresenta os pontos característicos definidos pelo padrão.

Badler e Byun (2002) propõem um mecanismo de controle para expressões faciais através da aplicação de modificações paramétricas. Os autores pontuam que essa abordagem pode ser aplicada para qualquer animação facial expressa no formato geral do MPEG-4 seja proveniente de uma biblioteca de expressões faciais predefinidas, criada manualmente ou através de captura de movimento.

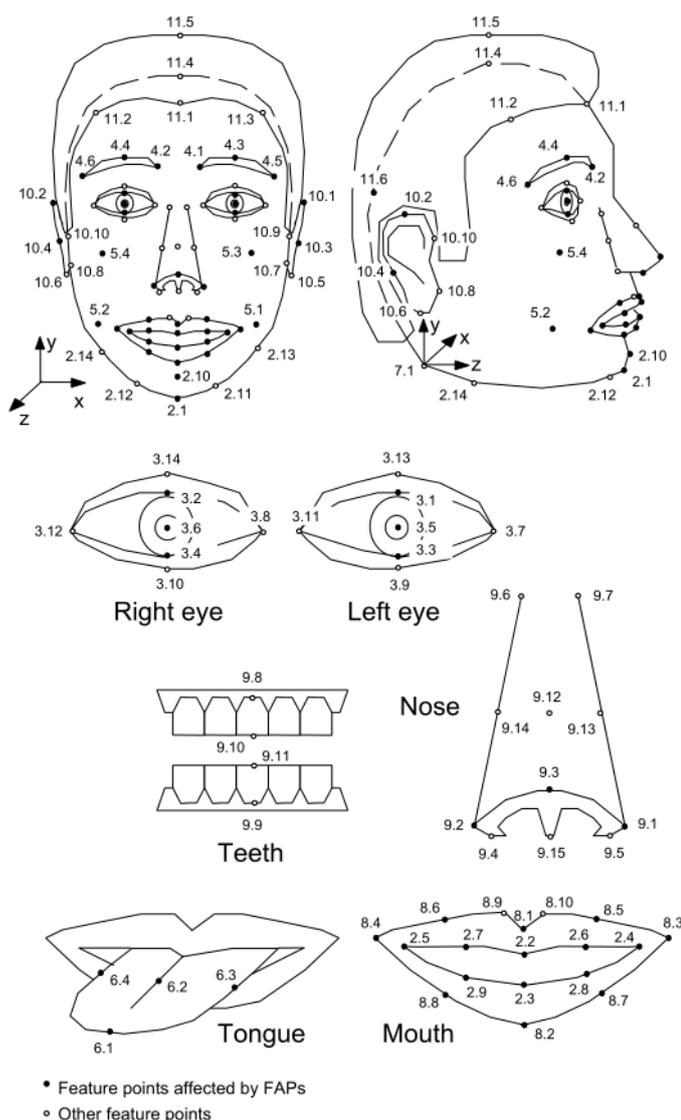


Figura 2.5 – Pontos característicos do padrão MPEG-4 (ISO, 2001)

Parametrização *data-driven*

A abordagem apresentada pela parametrização *data-driven* é uma tentativa de definição dos parâmetros ideais para movimentação do modelo através de dados de mensurações de ações humanas obtidos por equipamentos como *scanners* e sensores ou outros equipamentos que possam realizar captura de movimento (CORRÊA, 2007).

Na abordagem *data-driven*, é analisado e medido o comportamento real para a determinação dos parâmetros. De Martino (2005) explica que a parametrização *data-driven* pode permitir estabelecer poses-chave através de um conjunto de parâmetros definidos por pesos provenientes de dados de medidas e análises estatísticas de performances humanas. O autor apresenta em seu trabalho, voltado para a animação facial da fala, uma metodologia de captura de dados e definição de parâmetros através de gravações de vídeos com auxílio de um capacete desenvolvido para a calibração (Figura 2.6).



Figura 2.6 – Monitoramento de pontos faciais articulatórios da fala (DE MARTINO, 2005)

Um outro exemplo do uso de parametrização *data-driven* é o trabalho de Feng, Kim e Yu (2008) que apresenta uma interface de animação para o controle de deformação de um grande número de locais por meio de um número mínimo de pontos de controle através do método baseado em kernel de análise de correlação canônica (*CCA - Canonical Correlation Analysis*).

Recentemente, diversos avanços têm sido realizados em relação à técnica de parametrização *data-driven* conforme pode ser observado nos trabalhos de Zhang et al. (2016), Ruhland, Prasad e McDonnell (2017), Zhang, Liao e Li (2020) e Liu et al. (2020). É importante lembrar que a estratégia *data-driven* possui a limitação da impossibilidade de geração de movimentos que não tenham sido previamente medidos.

Simulação biomecânica

Na abordagem biomecânica são buscadas e aplicadas as restrições reais de execução de expressões faciais encontradas da face humana de forma a simular o comportamento real. São consideradas as características físicas com o intuito de buscar um alto grau de realismo nas animações. Nessa abordagem, os músculos faciais são modelados como parâmetros para a movimentação facial.

Dentre os primeiros trabalhos sobre a estratégia de simulação biomecânica está o trabalho de Platt e Badler (1981) que simulou o comportamento de fibras musculares para a representação facial. Outros trabalhos posteriores importantes exploraram modelos de várias camadas da pele baseado na anatomia muscular (TERZOPOULOS; WATERS, 1990) (LEE; TERZOPOULOS; WATERS, 1995).

Em Corrêa, Magalhães e De Martino (2008) é criado um modelo geométrico que emprega diferentes superfícies para representar características da fisiologia humana. São elas: a superfície externa para a pele; a superfície interna para os ossos e a superfície

intermediária para a fáscia. O sistema de movimentações neste trabalho foi restrito a região da boca, onde foram modelados músculos faciais cutâneos aplicados a um modelo tridimensional. Para isso, foram criados dois tipos de músculos diferentes: os lineares que contraem na direção do vértice e o esfíncter que contrai na direção a um centro determinado. A Figura 2.7 apresenta os músculos modelados neste sistema e resultados comparativos de movimentações produzidas.

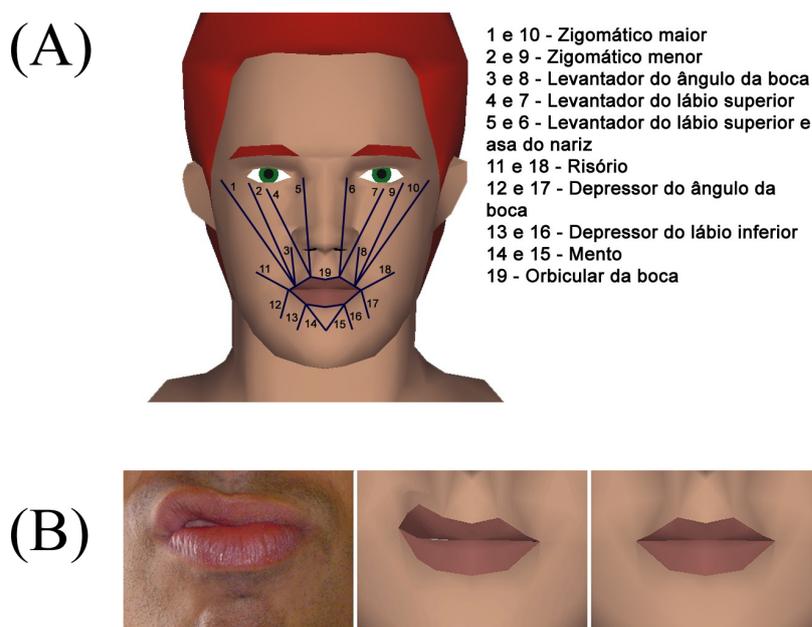


Figura 2.7 – (A) Músculos faciais modelados. (B) Resultados de movimentação (adaptado de [Corrêa, Magalhães e De Martino \(2008\)](#))

[Kim, Choi e Kim \(2020\)](#) apresentam uma abordagem de animação facial baseada em física em tempo real guiado por deformações de músculos. No sistema foram utilizadas quatro camadas: pele, subcutânea, músculos e crânio, baseado na anatomia facial humana. A camada subcutânea é assumida como um corpo elástico e a parte onde os músculos são fixados como uma posição estática.

2.4 Mecanismos de definição e controle da animação

As técnicas de manipulação da geometria da face estão diretamente ligados aos mecanismos de definição e controle da animação facial, pois disponibilizam recursos para a realização das deformações necessárias na face, e conseqüentemente, para a produção da animação.

Segundo [De Martino \(2005\)](#) os mecanismos de controle e definição da animação podem ser classificados em três diferentes abordagens que incluem os sistemas interativos, os sistemas baseados em *script* e os sistemas baseados em captura de movimento. Os

sistemas interativos se baseiam na habilidade dos animadores para criar manualmente os movimentos através de softwares de edição tridimensional. Os sistemas baseados em *script* se apoiam em um formalismo que descreve a animação, denominada *script*, que de forma geral, é composto pela descrição de comandos que acionam algoritmos previamente determinados para atuar na manipulação da geometria e conseqüentemente da animação.

Os sistemas de captura de movimento ([Mocap](#)) se baseiam no processo de gravar determinados movimentos de humanos, animais ou objetos e os transformar em termos matemáticos através do rastreamento de pontos ao longo do tempo obtendo assim uma representação tridimensional ([MENACHE, 2010](#)) ([KITAGAWA; WINDSOR, 2008](#)). As animações faciais e corporais do presente trabalho são baseadas em capturas de movimento.

2.5 Captura de movimento - Mocap

Existem diversos tipos de sistemas de [Mocap](#), os principais atualmente são os sistemas magnéticos, mecânicos e ópticos. Os sistemas ópticos são o mais comumente usados devido a fatores como flexibilidade, precisão e taxa de captura ([MENACHE, 2010](#)) ([KITAGAWA; WINDSOR, 2008](#)). Os sistemas ópticos podem ser classificados em duas principais técnicas: sem marcadores (*markerless*) e baseado em marcadores (*marker based*).

Os Mocaps corporais comumente são realizados por sistemas baseados em marcadores que pode ser passivos ou ativos. Marcadores passivos são revestidos com um material que reflete a luz emitida pelas câmeras do sistema. Marcadores ativos são compostos por LEDs portanto emitem luz. A Figura 2.8 apresenta uma configuração padrão de 53 marcadores passivos apresentada pela empresa Vicon ([VICON, 2020](#)) para captura corporal.

Em relação aos Mocaps faciais, o sistema sem marcadores permite uma captura mais detalha da superfície da face, porém é mais sensível ao ambiente e o rastreamento muitas vezes não é confiável ([GAO et al., 2007](#)). Por outro lado, no sistema baseado em marcadores, cada marcador é triangulado por mais de uma câmera e a configuração dos marcadores é esparsa, fazendo deste sistema comumente mais robusto. Neste tipo de sistema, diversos marcadores são posicionados em locais estratégicos para que os movimentos possam ser capturados e reproduzidos em um personagem tridimensional.

As capturas de movimento facial realizadas por sistemas ópticos baseado em marcadores, que são normalmente os sistemas mais utilizados, possuem a capacidade de capturar até os movimentos mais sutis ([KITAGAWA; WINDSOR, 2008](#)) ([MENACHE, 2010](#)). Para a realização de capturas faciais utilizando esse tipo de sistema, é necessário que se defina o *layout* de captura, ou seja a quantidade e o posicionamento dos marcadores no rosto do sinalizante. Nesse contexto, [Kitagawa e Windsor \(2008\)](#) argumenta que não se deve usar mais marcadores do que o necessário, como costuma acontecer em capturas corporais.



Figura 2.8 – Configuração padrão Vicon de 53 marcadores para captura corporal (VICON, 2020)

O autor alerta que marcadores extras nas capturas faciais tendem a ser descartados e, quando capturados, podem ser vistos como um único marcador pelas câmeras do sistema de captura. No entanto, o uso de poucos e mal posicionados marcadores resulta em uma baixa qualidade de captura facial. Com base nesses problemas, a busca por um *layout* otimizado torna-se necessária (LE; ZHU; DENG, 2013), (REVERDY; GIBET; LARBOULETTE, 2015). O *layout* de marcadores faciais utilizado para as capturas de movimento do presente trabalho foi desenvolvido através da abordagem descrita na Seção 3.2. A abordagem de mapeamento dos dados capturados ao rosto do avatar (processo de *retargeting*) é descrito na Seção 3.3. Ambas as abordagens fazem parte das contribuições do presente trabalho.

As animações faciais neste trabalho foram obtidas através do sistema *Vicon Cara* o qual rastreia marcadores brancos e/ou pretos fixados no rosto do sinalizante através de 4 câmeras com resolução HD em até 60 quadros por segundo (VICON, 2018). Portanto, óptico e baseado em marcadores. Nesse trabalho optou-se pela utilização do sistema *Vicon Cara* pela disponibilidade no laboratório onde a pesquisa foi desenvolvida. A Figura 2.9 apresenta uma modelo com o Sistema *Vicon Cara* posicionado em sua cabeça durante uma sessão de captura de movimento.

2.6 Comentários Finais

A geração de animações faciais realistas é um desafio devido a nossa refinada habilidade, treinada através de anos de observação, de perceber mesmo pequenas alterações



Figura 2.9 – Sistema Vicon Cara posicionado na cabeça de uma modelo (VICON, 2015)

em movimentos faciais. Conceitos como o *believability flip* e o *uncanny valley* descrevem os desafios inerentes ao desenvolvimento de animação realista.

Ao longo dos anos as técnicas de modelagem e animação facial têm se aprimorado e alinhado com a modelagem realista, atualmente muitas vezes torna-se difícil a distinção entre uma animação produzida por computador e uma cena gravada com atores reais. As técnicas de manipulação da geometria da face e os mecanismos de definição dos controles de animação aqui descritos apresenta um panorama das abordagens utilizadas para produção de animações faciais.

As animações faciais utilizadas neste trabalho para a produção de expressões faciais em avatares sinalizantes de Libras se baseiam em capturas de movimento de pessoas fluentes em Libras com o intuito de conferir realismo e agilidade ao processo de produção de animações. Aspectos relacionados as expressões faciais na Libras serão abordados no Capítulo 4.

3 Abordagens para captura de movimento facial

3.1 Introdução

As animações faciais, de modo geral, podem admitir expressões exageradas e cartunescas fugindo do formalismo realista. Contudo, as expressões faciais deste trabalho são focadas no realismo, objetivo este que requer uma atenção extra tendo em vista desafios como o apresentado pelo conceito *uncanny valley* descrito anteriormente. Na busca por realismo, muitas vezes comparações e medidas de movimentação realizada por pessoas são analisadas e utilizadas.

Com o foco em expressões faciais realistas, as animações que serão incorporadas ao avatar sinalizante são obtidas através de um sistema óptico de [Mocap](#) baseado em marcadores. A utilização de [Mocap](#), além de proporcionar movimentações realistas e naturais reduz a necessidade de trabalho manual normalmente realizado por um animador que deve ter, além de sensibilidade artística, conhecimento da fisiologia da face.

Para a utilização de um sistema de [Mocap](#) baseado em marcadores é necessário definir o *layout* de marcadores faciais a ser utilizado e determinar a estratégia de mapeamento dos dados provenientes das capturas ao rosto do avatar. Neste capítulo são apresentadas duas contribuições para o presente trabalho: O desenvolvimento de uma abordagem para o posicionamento otimizado de marcadores para Mocaps faciais (Seção [3.2](#)) e o desenvolvimento de uma abordagem para o mapeamento de dados capturados ao rosto de um personagem (Seção [3.3](#)).

3.2 Posicionamento otimizado de marcadores para Mocap facial

Diversos *layouts* empíricos de marcadores faciais foram criados e usados por pesquisadores e pela indústria ([LE; ZHU; DENG, 2013](#)). No entanto, essas soluções carecem de fundamentação sólida, dificultando a generalização e muitas delas precisam de padronização que é importante para facilitar o processo posterior de mapeamento dos dados ao rosto do personagem. No âmbito do presente trabalho foi desenvolvida uma abordagem de criação de *layouts* otimizados de marcadores para capturas faciais usando dados capturados de atores reais. Os resultados a seguir também são reportados em ([WILL; DE MARTINO; BEZERRA, 2018](#)).

Em um primeiro momento, parece interessante a utilização do maior número possível de marcadores de forma a cobrir a maior parte do rosto do ator. Contudo, um alto

número de marcadores em um *layout* denso aumenta a demanda por poder de processamento e a probabilidade de trocas de identificação de marcadores que podem ocorrer quando os marcadores estão muito próximos um do outro (KITAGAWA; WINDSOR, 2008). Por outro lado, o uso de poucos ou mal posicionados marcadores pode não permitir a captura de movimentos significativos associados às expressões faciais. Com base nesses problemas, a busca por um *layout* otimizado se torna necessária (LE; ZHU; DENG, 2013), (REVERDY; GIBET; LARBOULETTE, 2015).

O objetivo da criação de um *layout* otimizado de marcadores faciais é determinar um número mínimo de marcadores e suas respectivas posições que satisfaça um alto grau de realismo e inteligibilidade das expressões faciais capturadas. A abordagem aqui apresentada foi desenvolvida para a reprodução de expressões faciais em geral, englobando as expressões faciais da *Libras*.

O problema de otimização deste estudo também foram tratados em (LE; ZHU; DENG, 2013) e (REVERDY; GIBET; LARBOULETTE, 2015). Le, Zhu e Deng (2013) resolve um problema de otimização de pontos de controle característicos baseados nos vértices da malha de polígonos a partir de amostras de poses faciais em alta resolução. No trabalho de Reverdy, Gibet e Larboulette (2015) são usadas posições dos vértices da malha facial de todos os quadros de uma sequência de treinamento e, através de uma técnica de agrupamento, é então calculado o *layout* otimizado dos marcadores para um determinado número de grupos. Ambos os estudos calculam um *layout* otimizado a partir das posições dos vértices da malha da face quando expressões faciais são aplicadas a elas. Diferentemente, na abordagem desenvolvida no âmbito do presente trabalho, a determinação do *layout* otimizado é realizada através de dados capturados de atores reais.

Em Mocaps corporais, os marcadores são posicionados considerando as localizações das articulações e ossos do corpo humano. No entanto, como as expressões faciais são produzidas principalmente por movimentos musculares, a determinação do número de marcadores e seu posicionamento não é óbvia, nem trivial.

Na abordagem aqui apresentada, foram realizadas capturas de movimento utilizando um *layout* de marcadores denso, ou seja, com muitos marcadores posicionados no rosto de um ator e, realizados movimentos específicos para as regiões faciais: sobrancelhas e testa, olhos, nariz, bochechas, boca e queixo. A partir da análise do comportamento da movimentação dos marcadores foram observados grupos de marcadores com comportamentos semelhantes. Em cada um dos grupos determinados, apenas um marcador é escolhido para representar, no *layout* otimizado, os demais membros de seu grupo.

3.2.1 *Layout* denso e regiões faciais

Apesar de que em capturas de movimento normalmente não se deve utilizar marcadores próximos uns dos outros (KITAGAWA; WINDSOR, 2008), para o desenvolvimento do *layout* otimizado, as capturas foram realizadas em um *layout* denso de marcadores com distância entre eles de aproximadamente 1 centímetro.

Para os experimentos aqui realizados, foi assumido que a maior parte das expressões faciais são simétricas. Dessa forma, os marcadores a serem analisados foram posicionados somente em um lado e na linha central do rosto de um ator como pode ser observado na Figura 3.1-a.

Como marcadores, foram utilizados artefatos com geometria de meia esfera fixada no rosto do ator com cola hipoalergênica. Os marcadores foram pintados com tinta branca fosca para uma reflexão mais uniforme e difusa.

Para a análise, optou-se por restringir a investigação dos marcadores de acordo com a região facial de sua localização. As regiões faciais determinadas no trabalho foram baseadas, mas não restritas, às regiões apresentadas por Sobotta et al. (2001) (Figura 3.2) que se baseia na musculatura facial. Uma representação das regiões faciais e da quantidade de marcadores em cada uma pode ser vista na Figura 3.1-b. A quantidade de marcadores em cada região pode ser observada na Tabela 3.1.

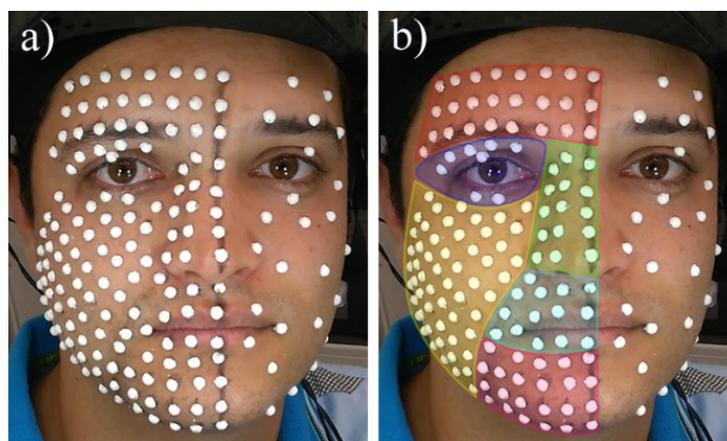


Figura 3.1 – a) *Layout* denso de captura e b) Demarcações das regiões faciais

Região facial	Quantidade de marcadores
Sobrancelha e testa	25
Olho	8
Nariz	15
Bochecha	56
Boca	14
Queixo	23

Tabela 3.1 – Quantidade de marcadores em cada região facial

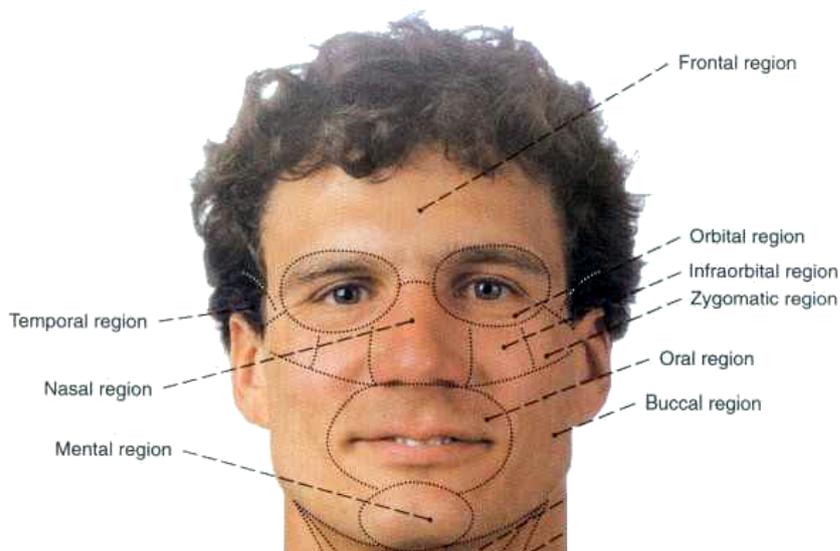


Figura 3.2 – Regiões do rosto (adaptado de Sobotta et al. (2001))

3.2.2 Captura e processamento

As capturas foram feitas com o uso do sistema *Vicon CARA*, conforme mencionado na Seção 2.5. As capturas das expressões faciais alvo foram alternadas, em intervalos de 2 segundos, com a expressão neutra (pose de descanso). Através do software *Vicon CaraPost* (VICON, 2021), os dados de captura foram verificados e processados resultando na trajetória tridimensional de cada marcador ao longo das capturas (Figura 3.3).

3.2.3 Matrizes de posições

Após o processamento das capturas, as posições de cada marcador a cada 5 quadros foram obtidas e utilizadas para a criação das matrizes ($m \times n$) de posições de cada eixo (PX , PY e PZ) para todas as regiões faciais. Abaixo é apresentada a matriz de posições do eixo x .

$$PX = \begin{bmatrix} px_1^1 & px_1^2 & \dots & px_1^n \\ px_2^1 & px_2^2 & \dots & px_2^n \\ \vdots & \vdots & \ddots & \vdots \\ px_m^1 & px_m^2 & \dots & px_m^n \end{bmatrix} \quad (3.1)$$

Em que px_j^i é a coordenada no eixo x do i -ésimo marcador no j -ésimo quadro analisado, m o número de quadros em uma determinada região e n o número de marcadores observados. As matrizes de posições PY e PZ são definidas de forma análoga.

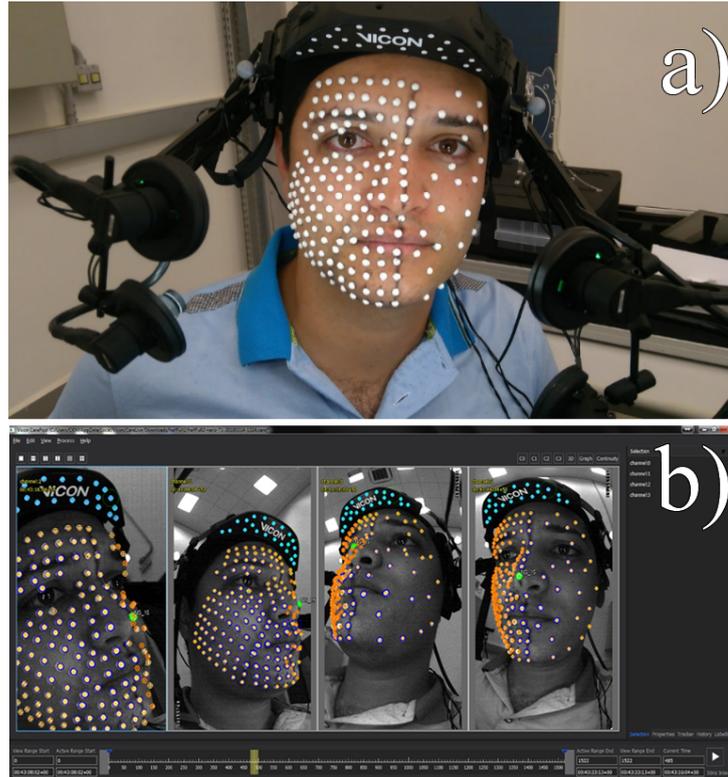


Figura 3.3 – a) Ator com os marcadores posicionados e utilizando o equipamento de captura de movimento Vicon Cara. b) Visualização das câmeras durante o processamento dos dados capturados

3.2.4 Matrizes de correlações

Como mencionado, é necessário descobrir os marcadores que possuem comportamento semelhante. Para isso, a partir dos dados de posições, foram criadas matrizes de correlação entre os marcadores de cada região facial. Para criar as matrizes de correlação foram calculadas as covariâncias entre os marcadores e acomodados os valores em um formato matricial. As covariâncias entre os marcadores foram calculadas através da equação 3.2.

$$\text{cov}(px^j, px^k) = \frac{\sum_{i=1}^n [(px_i^j - \overline{px^j}) \cdot (px_i^k - \overline{px^k})]}{(n - 1)} \quad (3.2)$$

Em que px^j e px^k são os valores de posição no eixo x de dois marcadores ao longo dos quadros observados, $\overline{px^j}$ e $\overline{px^k}$ são os vetores médios, px_i^j e px_i^k os valores das coordenadas na i -ésima posição e n o número de quadros observados. As covariâncias dos marcadores nos eixos x e y são calculadas de forma análoga.

Para cada matriz de posição (PX , PY e PZ), os valores das covariâncias entre os marcadores são acomodados formando as matrizes de covariância. Abaixo é apresentada a matriz de covariância CX resultante da matriz PX .

$$CX = \begin{bmatrix} cov(px^1, px^1) & cov(px^1, px^2) & \dots & cov(px^1, px^n) \\ cov(px^2, px^1) & cov(px^2, px^2) & \dots & cov(px^2, px^n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(px^n, px^1) & cov(px^n, px^2) & \dots & cov(px^n, px^n) \end{bmatrix} \quad (3.3)$$

Com as matrizes de covariância de cada região facial calculada, para cada eixo, são então computadas as matrizes de correlação através da Equação 3.4.

$$rx_{ij} = \frac{cx_{ij}}{\sqrt{cx_{ii} \cdot cx_{jj}}} \quad (3.4)$$

Em que rx_{ij} é o elemento da linha i e coluna j da matriz de correlação RX e cx_{ij} o elemento da linha i e coluna j da matriz de covariância CX . As matrizes de correlação RY e RZ são definidas de forma análoga.

O cálculo das matrizes de covariância e, conseqüentemente, das matrizes de correlação são realizadas a partir das matrizes de deslocamento de cada eixo (x , y e z) em cada região facial. Ao final do cálculo, as matrizes de correlação de cada eixo são consolidadas em uma matriz final através da soma das matrizes.

3.2.5 Agrupamento dos marcadores

Os marcadores são agrupados, em cada região facial, a partir da análise de seus comportamentos expressos pelas matrizes de correlação. Se um grupo de marcadores em uma determinada região facial possui comportamento semelhante, observado através de um valor mínimo de correlação, definido inicialmente pela média de todas as correlações na região facial analisada, os marcadores desse grupo são representados por um único marcador que fará parte do *layout* otimizado final.

Os agrupamentos e determinação do marcador representativo em cada região facial são realizados de acordo com o fluxograma da Figura 3.4.

No algoritmo apresentado pelo fluxograma, são definidas duas *flags* para cada marcador: agrupado e verificado. Uma vez comparada sua correlação com o marcador representativo, o marcador é definido como verificado e não será mais analisado para o agrupamento atual. Ao ser criado um novo grupo todos os marcadores têm suas marcações de verificados retiradas. Quando um marcador é adicionado a um grupo ele é definido como agrupado e não é mais verificado.

Para uma referência mais visual, a Figura 3.5 apresenta uma sequência de agrupamentos ocorridos na região da bochecha através de esferas posicionadas na face de um personagem tridimensional. As cores nos marcadores representam os diferentes grupos de marcadores. Os marcadores representativos são apresentados com um maior diâmetro.

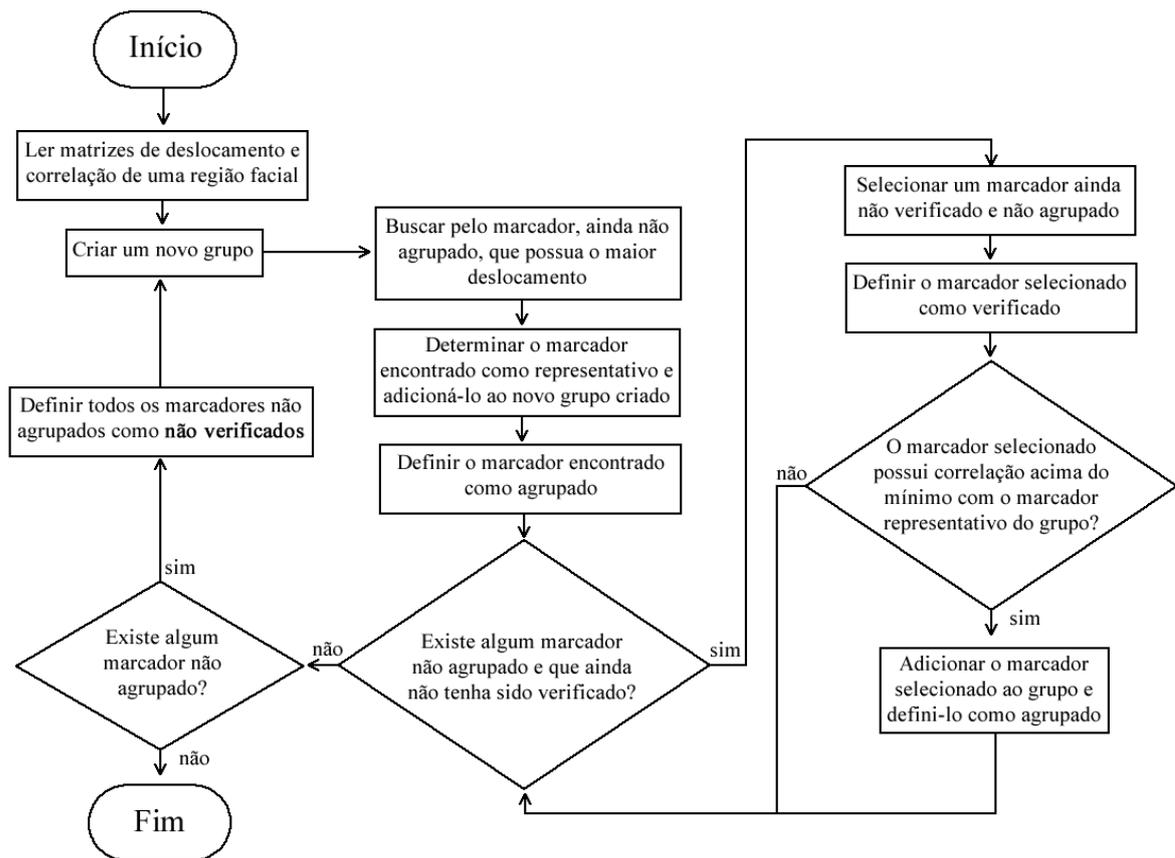


Figura 3.4 – Fluxograma do algoritmo de agrupamento e determinação de marcadores representativos

3.2.6 Ajustes dos marcadores

No processo de agrupamento, os marcadores são incluídos no grupo do primeiro marcador representativo que apresentar uma correlação dentro do limite aceitável. Dessa forma, pode acontecer de um marcador ter uma correlação maior com outro marcador representativo de um grupo criado posteriormente. Para resolver esse problema, um processo de ajustes é realizado para avaliar as correlações e, se necessário, realocar marcadores entre os grupos criados. Neste processo, são verificados todos os marcadores, com exceção dos marcadores representativos de cada grupo. O fluxograma do algoritmo criado para a realização de ajustes de marcadores pode ser observado na Figura 3.6. A Figura 3.7 apresenta um exemplo de ajuste de agrupamento.

Apesar de serem utilizados apenas os marcadores representativos na criação do *layout* otimizado, as regiões de cada grupo no rosto do avatar podem auxiliar na definição da área de influência dos marcadores do *layout* otimizado ao serem definidos os mecanismos de manipulação da malha de polígonos.

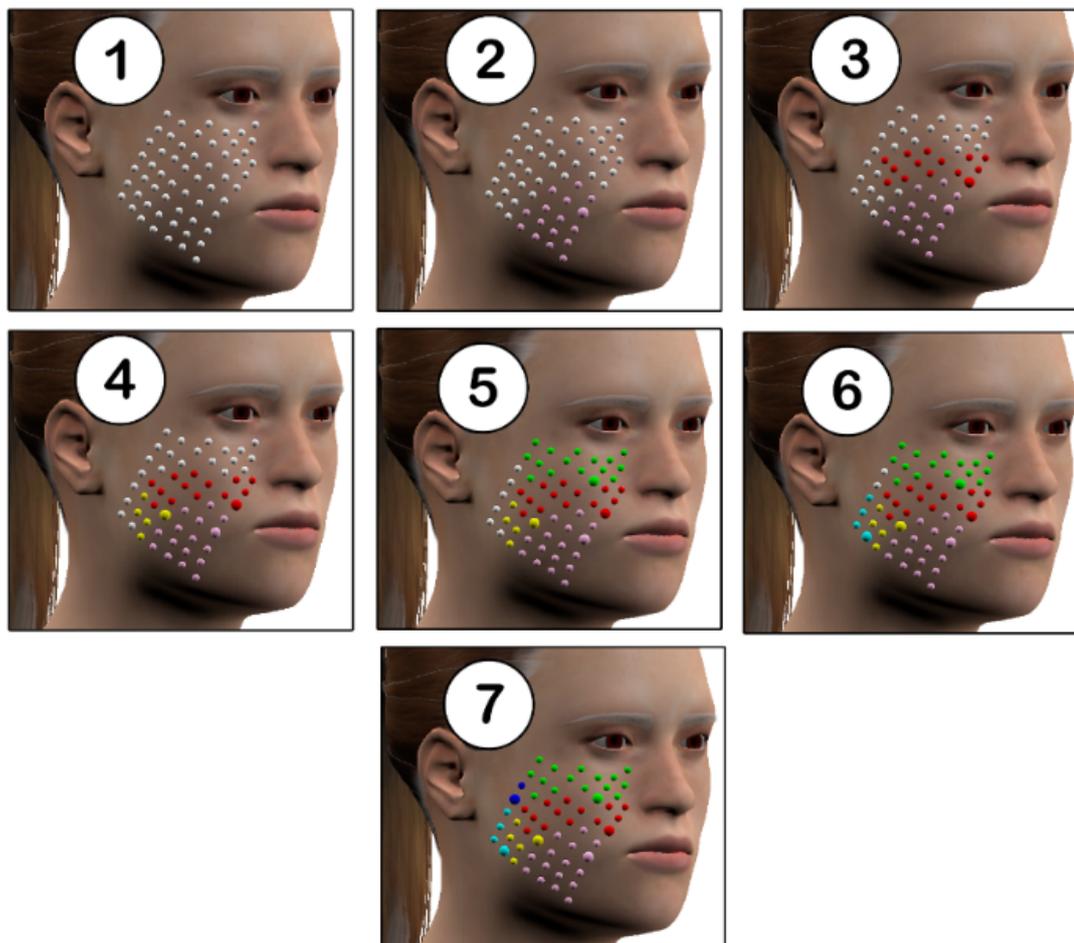


Figura 3.5 – Ilustração de uma sequência de agrupamentos dos marcadores da região da bochecha

3.2.7 Resultados

As capturas e agrupamentos foram realizados com marcadores fixados apenas do lado direito e na linha central do rosto como já mencionado. Contudo, para os resultados finais, os marcadores foram espelhados para o lado esquerdo (Figura 3.8 esquerda).

Os marcadores que devem compor o *layout* otimizado, são os marcadores representativos de cada grupo formado. Esse *layout*, pode ter sua densidade alterada através da manipulação do limite de correlação (multiplicando por um coeficiente) utilizado no processo de agrupamento. Essa opção pode ser utilizada de acordo com a necessidade de detalhamento da captura. A coluna *Layouts* na Figura 3.8 apresenta três diferentes *layouts* de marcadores obtidos através da manipulação do fator de refinamento (*coef*) que, como mencionado na Seção 3.2.5, é inicialmente obtido pela média de todas as correlações em cada região facial.

As capturas realizadas nesse trabalho terão como base o *layout* otimizado com 38 marcadores (Figura 3.8 linha inferior).

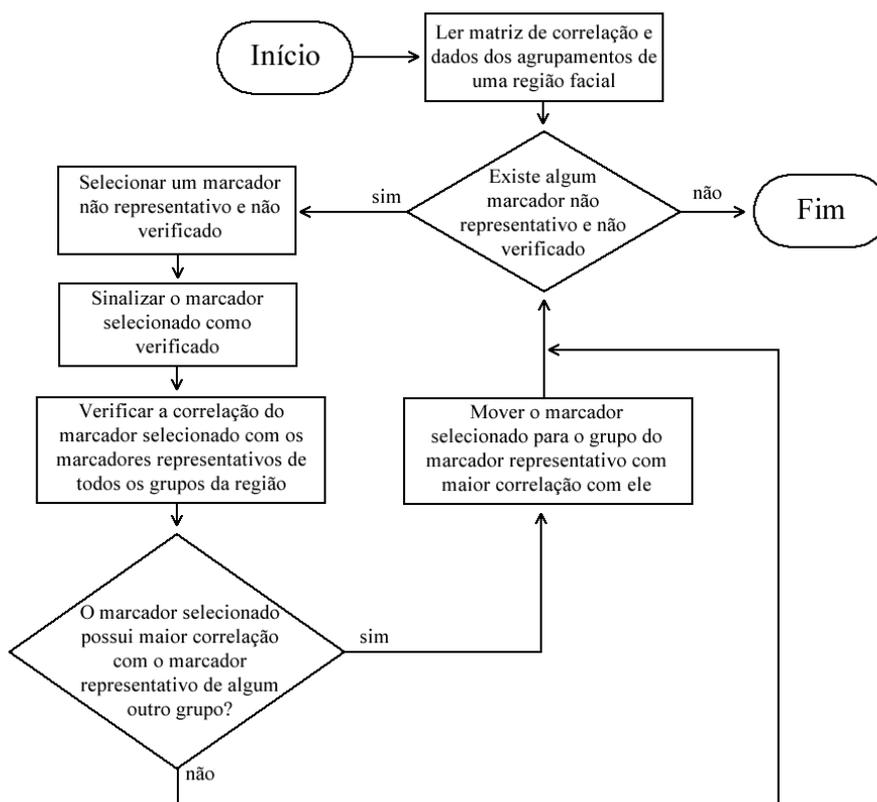


Figura 3.6 – Fluxograma do algoritmo de ajustes de marcadores

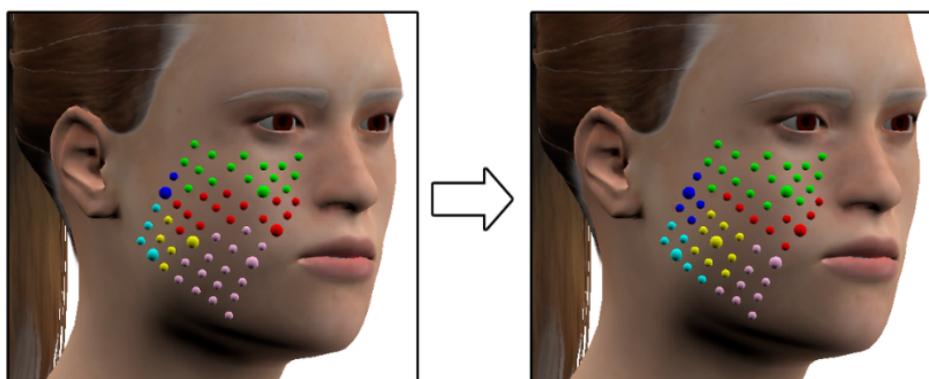


Figura 3.7 – Exemplo de ajuste de agrupamento. Esquerda: Agrupamento original. Direita: Agrupamento ajustado

3.3 Mapeamento dos dados de Mocap

Após a determinação do *layout* otimizado e realização de Mocaps é possível transferir os movimentos para o avatar de forma a reproduzir as expressões faciais. Esse processo é conhecido como *retargeting*. Em linhas gerais, o *retargeting* de animações faciais implica em transferir o movimento entre personagens diferentes ou entre dados provenientes de capturas de movimento (Mocap) e personagens. Esse processo requer a adaptação dos dados do Mocap para que o personagem reproduza os movimentos capturados pelo ator de maneira natural e convincente.

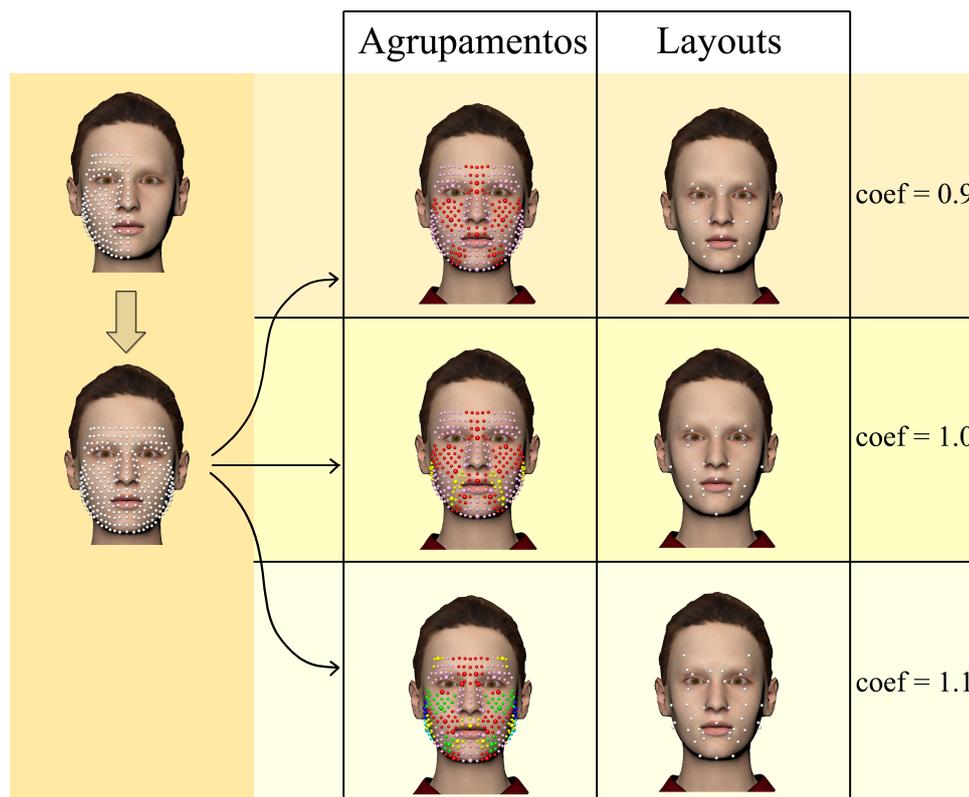


Figura 3.8 – Ilustração de espelhamento de marcadores (esquerda) e agrupamentos e *layouts* obtidos através de coeficiente de refinamentos $\text{coef}=0.9$, $\text{coef}=1.0$ e $\text{coef}=1.1$ respectivamente

Chuang (2004), Ribera et al. (2017), Pighin e Lewis (2006) complementam que o problema apresentado pelo *retargeting* facial está principalmente relacionado à diferença de propriedades físicas, como tamanho, proporção e grau de liberdade entre o ator e o personagem a ser incorporado os movimentos. Um outro fator a se considerar é o alinhamento dos eixos dos dados do *Mocap* com a face do personagem, para que os movimentos do personagem ocorram na mesma direção daqueles capturados pelo ator.

Para o presente trabalho foi desenvolvida uma abordagem de *retargeting* facial que resolve o problema de ajustes de proporções e alinhamento dos dados de *Mocap* de forma automática. Os resultados apresentados a seguir também são reportados em (WILL; DE MARTINO; BEZERRA, 2019).

A abordagem de *retargeting* aqui descrita foi desenvolvida de forma a resolver o alinhamento e mapeamento entre dados de *Mocap* e de avatares tridimensionais realistas diversos através de ajustes baseados nas características físicas e alinhamento da face do avatar. Assim, os movimentos capturados por um ator podem ser transferidos para vários avatares com características faciais distintas conforme ilustrado na Figura 3.9.

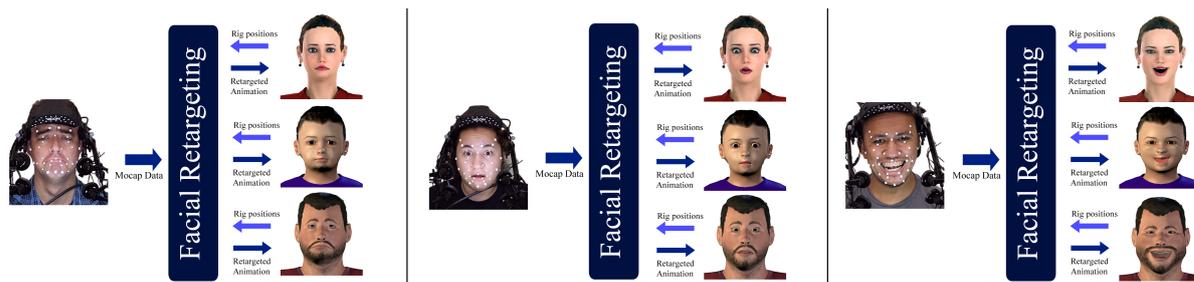


Figura 3.9 – Visão geral do processo de *retargeting* facial (WILL; DE MARTINO; BEZERRA, 2019)

3.3.1 Marcadores e ossos controladores

O *layout* de marcadores utilizado se baseou no *layout* obtido pelo processo descrito na Seção 3.2 e apresentado na Figura 3.7 com coeficiente de refinamento igual a 1.1. Esse coeficiente foi escolhido, de forma empírica, através de testes realizados.

Para reproduzir as expressões faciais no avatar a partir de dados de *Mocap*, foi utilizada uma abordagem que utiliza correspondência direta entre os dados do *Mocap* (movimentos dos marcadores) e os polígonos da malha do rosto do avatar. Como o número de marcadores é muito menor que o número de vértices de malha ($M \ll V$), essa é uma correspondência esparsa.

No rosto do avatar, foram inseridos mecanismos, denominados ossos controladores, para a manipulação da malha de polígonos. Na abordagem aqui descrita, é utilizado o mesmo *layout* para os marcadores de captura e para a configuração de ossos controladores no rosto avatar, criando assim um par correspondente entre marcadores e ossos controladores. Através desse par, cada osso controlador é associado a apenas um marcador e influencia, em graus variados, um certo número de vértices da malha do rosto do avatar.

A associação de cada osso controlador com seu marcador correspondente é realizada aplicando-se uma restrição onde os ossos controladores devem reproduzir os deslocamentos realizados pelos marcadores ao longo da animação.

3.3.2 Alinhamento dos dados de Mocap

Após uma captura de movimento, os dados processados que contêm a trajetória tridimensional dos marcadores, ou seja, a posição dos marcadores em cada quadro, estão prontos para serem analisados e o processo de *retargeting* facial é iniciado. No entanto, antes de qualquer outro processamento, deve-se notar que as posições do *layout* dos marcadores dos dados de *Mocap* diferem em escala e rotação quando comparadas à configuração de ossos controladores do avatar, impossibilitando a aplicação direta dos movimentos dos marcadores aos ossos controladores.

Ajuste de escala e proporção

O problema da diferença de tamanho entre o *layout* dos marcadores e a configuração de ossos controladores pode ser causada pelas características do sistema de captura e processamento ou pela diferença de proporção entre a face do sinalizante e a face do avatar.

Levando em consideração que o *layout* de marcadores fixados no rosto do ator e o *layout* de ossos controladores no rosto do avatar são correspondentes, para a resolução do problema de escala e proporção é calculada a distância entre pontos extremos de cada eixo (x, y, z), tanto na configuração de ossos controladores quanto no *layout* de marcadores e dessa forma, conseqüentemente, calculadas as proporções entre essas distâncias conforme apresentado na Equação 3.5.

$$prop_i = db_i/dm_i \quad (3.5)$$

Em que i é o índices do eixos (x, y, z), db_i a distância entre os pontos determinados para cada eixo no rosto do avatar, dm_i a distância entre os pontos determinados para cada eixo no rosto do ator e $prop_i$ a proporção entre o rosto do avatar e o rosto do ator em cada eixo.

Para o eixo x, foram escolhidos os pontos nas laterais da face, perto das orelhas, para o eixo y, o ponto mais interno da sobrancelha esquerda e o ponto sob o queixo à esquerda e, para o eixo z, o ponto mais próximo da orelha direita e o ponto no lado direito do nariz, como mostra a Figure 3.10.

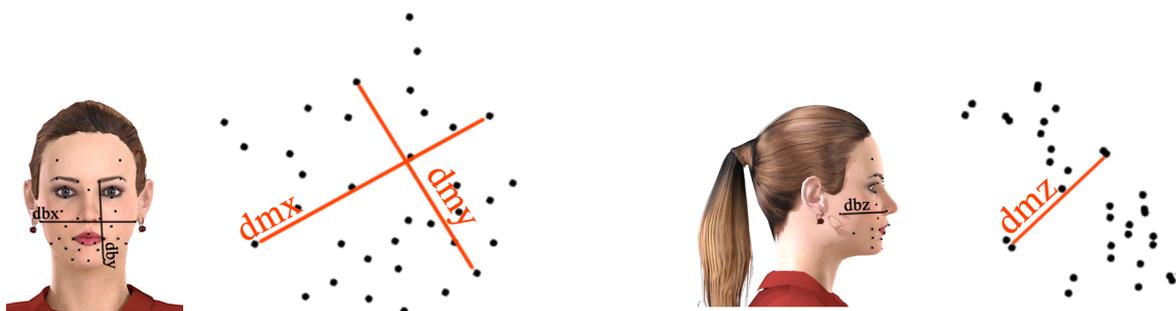


Figura 3.10 – Exemplo de distâncias em cada eixo para os cálculos de proporcionalidades

Mudança de base vetorial (Rotação)

A diferença de rotações entre a posição original do *layout* dos marcadores e a configuração de ossos controladores pode ser entendida como diferentes espaços vetoriais. A proposta para resolver esse problema é determinar uma base vetorial para o *layout* dos marcadores (A) e uma base vetorial para a configuração de ossos controladores (B) e, em

seguida, através de alteração da base vetorial, mover os marcadores para o espaço vetorial da configuração dos ossos controladores.

Inicialmente, para a construção das bases vetoriais, em cada *layout*, foram utilizados os mesmos pares de pontos utilizados no cálculo das proporções do eixos x e do eixo y aqui denominados respectivamente por (u_A, u_B) e (u_C, u_D) . Os valores das coordenadas desses pontos, obtidos na pose neutra, foram utilizados para a criação de dois vetores temporários e linearmente independentes, um na direção horizontal $v_{AB} = (u_B - u_A)$ e outro na direção vertical $v_{CD} = (u_D - u_C)$, como mostrado na Figura 3.11.

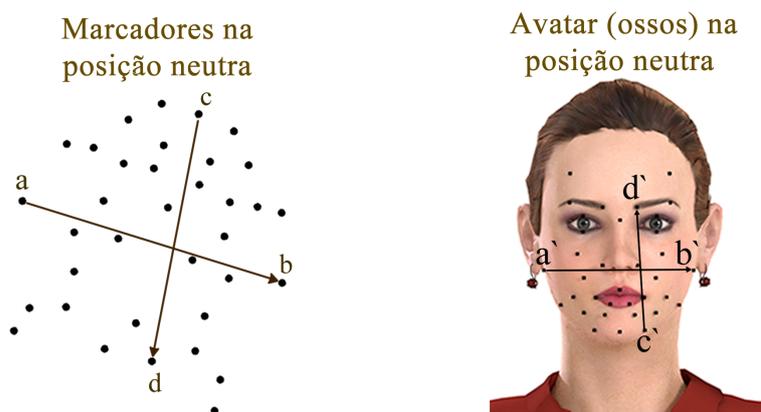


Figura 3.11 – Exemplo de criação dos vetores iniciais utilizados no processo de formação da base vetorial de cada *layout*

Tendo determinados os dois vetores iniciais em cada *layout*, foram realizadas as seguintes etapas:

- Normalização dos vetores v_{AB} e v_{CD} :

$$v = \frac{v}{|v|} \quad (3.6)$$

Em que $|v| = \sqrt{x^2 + y^2 + z^2}$ e x , y , e z são coordenadas do vetor v .

- Criação do vetor v_Z perpendicular ao vetor v_{CD} e v_{AB} :

$$v_Z = (v_{CD} \times v_{AB}) \quad (3.7)$$

- Criação do vetor v_Y perpendicular ao vetor v_Z e v_{AB} :

$$v_Y = (v_Z \times v_{AB}) \quad (3.8)$$

- Criação do vetor v_X perpendicular ao vetor v_Y e v_Z :

$$v_X = (v_Y \times v_Z) \quad (3.9)$$

Para cada *layout* (marcadores e ossos controladores), uma base vetorial $\beta = [v_X, v_Y, v_Z]$ é criada, em que v_X , v_Y e v_Z são vetores ortonormais.

Cada ponto u_A da base vetorial dos marcadores (A) é levado para a base vetorial da configuração de ossos controladores (B) aplicando a equação de transformação de base vetorial:

$$u_A = Mu_B \quad (3.10)$$

Em que M é uma matriz de mudança de base dada por:

$$M = B^{-1}A \quad (3.11)$$

Depois de aplicado o processo de transformação da base vetorial a todos os dados do [Mocap](#) (posições dos marcadores em cada quadro), o *layout* dos marcadores terá seus eixos alinhados à configuração de ossos controladores da face do personagem.

3.3.3 Vetores de deslocamento

A reprodução das expressões faciais de uma determinada captura facial por um avatar pode ser entendida e representada pela aplicação de vetores de deslocamento nos ossos controladores faciais ao longo de uma animação. Para facilitar a representação e manipulação de dados, a captura de movimento sempre começa na posição neutra. Assim, a posição dos marcadores no quadro 0 determina a origem de cada marcador.

Em cada quadro de captura de movimento, os vetores de deslocamento são obtidos calculando a diferença entre a posição atual de cada marcador e sua posição na pose neutra conforme apresentado na Equação 3.12.

$$v_j^i = (p_j^i - p_0^i) \quad (3.12)$$

Em que v_j^i é o vetor de deslocamento do marcador i no quadro j , p_j^i a posição do marcador i no quadro j e p_0^i a posição do marcador i no quadro 0.

No final do processo de *retargeting*, a animação é executada pela aplicação desses vetores de deslocamento aos ossos controladores no rosto do avatar. A Figura 3.12 apresenta uma visão geral do processo de aplicação dos vetores de deslocamento a um personagem tridimensional para produzir uma determinada expressão facial.

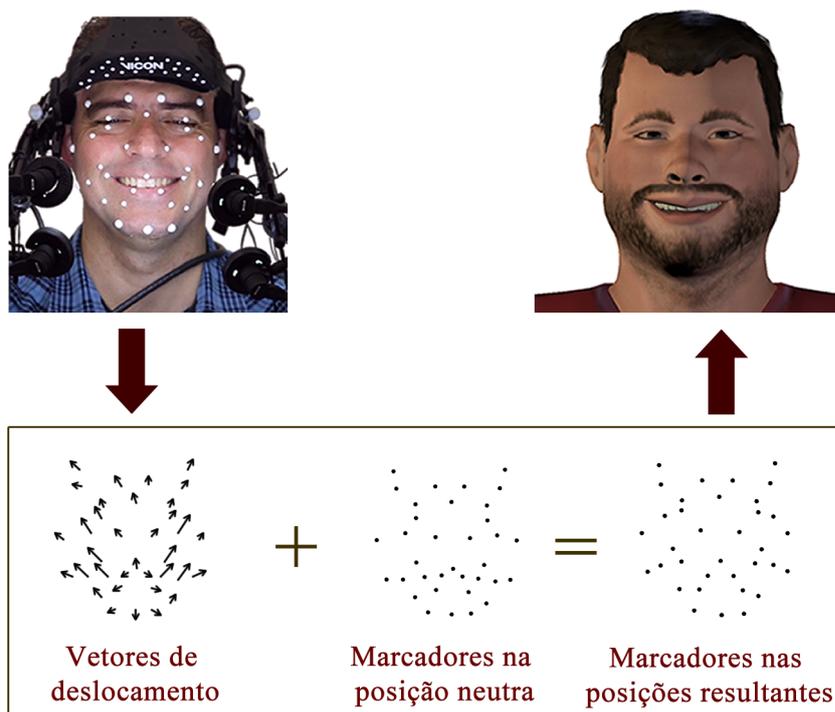


Figura 3.12 – Visão geral do processo de aplicação dos vetores de deslocamento

3.3.4 Ajustes por região

Como o foco é o *retargeting* facial para avatares realistas, em muitos casos, o ajuste do *layout* dos marcadores ao formato e proporção da face (configuração de ossos controladores) é suficiente para a adequação e reprodução de expressões faciais. No entanto, ainda podem haver diferenças entre as proporções de algumas regiões faciais comparando o rosto do ator e o rosto do personagem.

Na tentativa de resolver esse problema em cada região, são selecionados automaticamente os ossos controladores que mais se distanciam entre si e então calculada, através da Equação 3.13, a proporção entre essa distância (db) e a distância calculada entre os marcadores correspondentes a esses ossos controladores no *layout* de marcadores do *Mocap* (dm) (Figura 3.13).

$$p = \frac{db}{dm} \quad (3.13)$$

Em que p é o valor proporcional de uma região facial específica entre a configuração de ossos controladores e o *layout* do marcador.

O valor proporcional encontrado é multiplicado pelos dados de deslocamento de todos os marcadores que compõem a região facial em todos os quadros de captura.

As regiões faciais selecionadas para a aplicação dos ajustes são: boca, olhos



Figura 3.13 – Ilustração das distâncias entre os pontos mais distantes da boca na configuração de ossos controladores e no *layout* de marcadores para o cálculo do valor proporcional

(direita e esquerda), bochechas (direita e esquerda), sobrancelhas (direita e esquerda) e queixo.

3.3.5 Aplicação do *retargeting*

Para a verificação da aplicação da abordagem de *retargeting* descrita foram realizadas capturas, com três pessoas com diferentes proporções faciais, de diversas expressões faciais. Em seguida, foram aplicados os dados capturados, por meio da abordagem apresentada, à três avatares com proporções faciais também diferentes. É importante observar que não foram realizadas qualquer alteração ou ajuste manual. A Figura 3.14 apresenta os resultados dos experimentos realizados.

3.4 Comentários finais

Para a criação de expressões faciais de forma realistas e com um reduzido esforço manual, as animações de expressões faciais para este trabalho utilizaram um sistema óptico de *Mocap* baseado em marcadores. Com o intuito de padronizar e agilizar o processo de capturas e mapeamento dos dados capturados foi desenvolvida uma abordagem para a criação de um *layout* de marcadores faciais otimizado e uma abordagem para o processo de *retargeting*.

A abordagem de criação de *layout* otimizado de marcadores foi desenvolvida através da análise do comportamento de movimentação dos marcadores em um *layout* denso para a busca dos marcadores mais representativos. Na abordagem de *retargeting* foram resolvidos os problemas quanto ao alinhamento e ajustes de proporção dos dados capturados ao rosto do avatar para que as expressões faciais capturadas fossem semanticamente iguais às reproduzidas pelo avatar.

Ambas as abordagens apresentadas neste capítulo foram desenvolvidas para serem utilizadas em contexto geral, mas neste trabalho são utilizadas especificamente para capturas faciais de sinalizações em *Libras*.



Figura 3.14 – Resultados da aplicação do *retargeting* facial em diferentes avatares realistas

4 Expressões faciais na Libras

4.1 Introdução

No Brasil, o meio legal de comunicação e expressão da comunidade surda é a [Libras](#). Ela é uma língua visuoespacial com sua própria estrutura gramatical. A [Libras](#) foi oficialmente reconhecida pela lei federal número 10.436 de 24 de abril de 2002.

Segundo o censo de 2010, o Brasil possui cerca de 9.7 milhões de pessoas com algum tipo de deficiência auditiva e aproximadamente 3 milhões são analfabetos (IBGE, 2010). [Silva et al. \(2011\)](#) e [De Martino et al. \(2016\)](#) argumentam que isso se dá pela falta de professores bilíngues (português-Libras) e pela falta de material didático específico. Além disso, os surdos encontram outras dificuldades como atendimentos em estabelecimentos privados e públicos devido à barreira de comunicação, como o apontado na revisão realizada por [Souza et al. \(2017\)](#), onde são apresentados os desafios enfrentados pela comunidade surda no acesso à saúde.

Neste sentido, tecnologias assistivas podem auxiliar a comunicação entre ouvintes e surdos impulsionando o processo de letramento e de ensino e aprendizagem em geral. [Hersh e Johnson \(2006\)](#) apresentam diferentes tipos de tecnologias assistivas desenvolvidas para a comunidade surda. Atualmente, para este propósito, têm sido elaboradas tecnologias que utilizam animações tridimensionais. O trabalho de [Gibet \(1994\)](#) foi um dos primeiros a utilizar a tecnologia tridimensional na criação de animações para língua de sinais, em seu trabalho a autora apresenta um modelo que simula movimentos naturais de braço e mão guiados por comandos. Nos últimos anos, diversas pesquisas foram realizadas com o intuito de desenvolver e avaliar avatares sinalizantes que apresentam o conteúdo traduzido da língua falada para a língua de sinais ([KIPP; HELOIR; NGUYEN, 2011](#); [VIEIRA et al., 2014](#); [LOMBARDO et al., 2011](#); [BOUZID; JEMNI, 2013](#); [PATEL et al., 2020](#); [BROCK et al., 2020](#)).

No desenvolvimento de tecnologias assistivas para apresentação de conteúdo de [Libras](#) utilizando avatares sinalizantes é importante observar a estrutura das sinalizações. Os sinais em [Libras](#), assim como em outras línguas de sinais, são compostos por diversos parâmetros. Em ([STOKOE, 1960](#)) é proposta uma descrição fonológica para a língua de sinais americana, ou *American Sign Language (ASL)*, decompondo os sinais em três parâmetros: configuração, localização e movimento da mão. Posteriormente, estudos incluíram a orientação da palma ([BATTISON, 1974](#); [FRIEDMAN, 1975](#)) e expressões não manuais ([BAKER-SHENK; COKELY, 1980](#)). Entretanto, autores como ([VIEIRA et al., 2014](#)), destacam a importância da análise em separado das expressões não-manuais como

expressões faciais e corporais.

Neste capítulo é discutida a importância das expressões faciais na **Libras** e em especial sua incorporação em avatares sinalizantes. É destacada também a deficiência da apresentação de expressões faciais pelos avatares sinalizantes atuais em especial de expressões faciais características da modalidade interrogativa. Essas observações e preocupações motivaram o desenvolvimento da metodologia de modelagem e incorporação de expressões faciais apresentada no presente trabalho.

4.2 Expressões faciais na Libras

De maneira geral, na **Libras**, as expressões faciais podem ser de dois tipos distintos: afetivas (emotivas) e gramaticais. [Reilly \(2005\)](#) destaca que apesar das expressões faciais afetivas utilizarem os mesmos músculos empregados nas expressões gramaticais, existem diferenças quanto ao tempo de execução (início e duração). A autora também destaca que a execução de uma expressão facial gramatical coocorre com a sinalização manual.

As Figuras 4.1 e 4.2 apresentam uma ilustração para elucidar as diferenças de tempo de execução de expressões faciais gramaticais e de emoção.

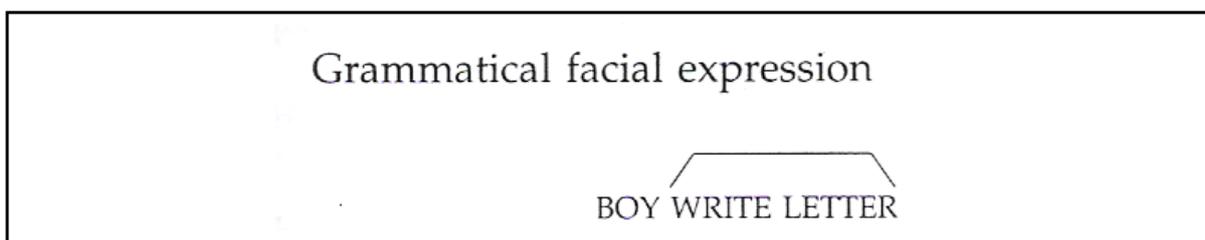


Figura 4.1 – Tempo de execução de uma expressão facial gramatical ([REILLY, 2005](#))

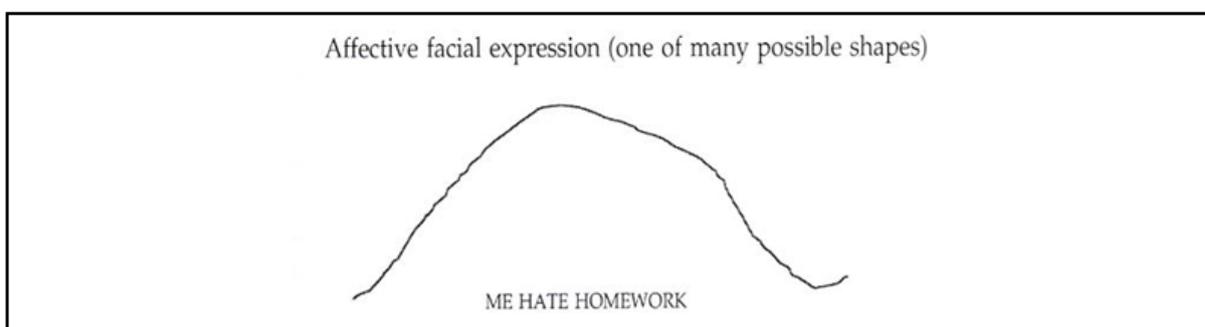


Figura 4.2 – Tempo de execução de uma expressão facial de emoção ([REILLY, 2005](#))

Pode-se observar através da ilustração da Figura 4.1 que a expressão facial gramatical coocorre com os sinais manuais relacionados, determinando o início e duração de sua execução enquanto que, conforme observado na Figura 4.2, uma expressão emotiva é executada por toda a sinalização e não apenas durante a execução de determinados

sinais, tendo seu início antes das sinalizações manuais e terminando após a conclusão da sentença.

As informações relativas ao início e duração de expressões faciais gramaticais em conjunção com os sinais manuais são especialmente importantes para a elaboração de mecanismos para a incorporação de animação facial realista em avatares sinalizantes de Libras.

No contexto gramatical, que é o escopo deste trabalho, (QUADROS; KARNOPP, 2004) argumentam que as expressões não-manuais, que inclui as expressões faciais, possuem dois papéis nas línguas de sinais: marcação de construções sintáticas e diferenciação de itens lexicais:

As expressões não-manuais que têm função sintática marcam sentenças interrogativas sim-não, interrogativas QU-, orações relativas, topicalizações, concordância e foco... As expressões não-manuais que constituem componentes lexicais marcam referência específica, referência pronominal, partícula negativa, advérbio, grau ou aspecto. (QUADROS; KARNOPP, 2004, 60)

Silva e Costa (2017), Silva (2020) e Silva et al. (2021) apresentam uma categorização das expressões faciais gramaticais da Libras em três grupos: Expressões Faciais Gramaticais para Sentenças (EFGS); Expressões Faciais Gramaticais de Intensidade (EFGI) e Expressões Faciais Gramaticais de Distinção (EFGD). Os autores descrevem que as EFGS definem o tipo da sentença e podem ser subdivididas em diferentes modalidades que envolvem as Interrogativas Parciais, usadas em sentenças com pronomes e advérbios interrogativos, as Interrogativas Totais, quando a resposta esperada é sim ou não, as Negativas, quando a sentença está na negativa dentre outras modalidades. As EFGI são responsáveis por apresentar o quantificador como, por exemplo, a palavra *bonito* tem seu sinal intensificado para *lindo* ou *muito bonito*. As EFGD complementam a sinalização para a compreensão distinguindo de outros sinais com a mesma sinalização manual, ajudando assim a distinguir o significado do sinal. Para a metodologia apresentada neste trabalho, como um recorte inicial, serão analisadas e aplicadas expressões faciais gramaticais de sentenças do tipo interrogativas. A Figura 4.3 apresenta exemplos de expressões faciais gramaticais ligadas a sentenças interrogativas.

Dada a relevância da expressão facial na Libras, é importante destacar que mesmo com a correta configuração de mãos, ponto de articulação e movimento, porém sem o emprego da expressão facial, o surdo pode não entender ou ainda interpretar de forma errada a informação (BRECAILO, 2012).

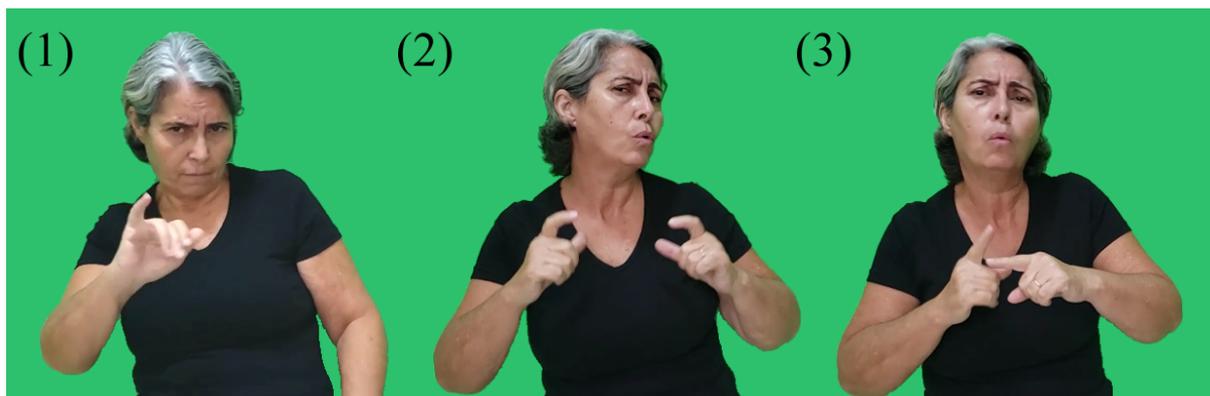


Figura 4.3 – (1) Sinalização de **VAI** em uma sentença interrogativa total (do tipo sim/não). (2) Sinalização de **ONDE** em uma sentença interrogativa parcial. (3) Sinalização de **POR-QUE** em uma sentença interrogativa parcial

4.3 Expressões faciais em avatares sinalizantes de Libras

Vídeos de sinalizações de **Libras** realizadas por humanos são muitas vezes usados para tornar o material escrito/falado acessível a pessoas surdas. No entanto, vídeos pré-gravados podem apresentar desvantagens relacionadas a problemas na recuperação e apresentação de vídeos em tempo real (KIPP; HELOIR; NGUYEN, 2011). Como alternativa ao vídeo tradicional, avatares sinalizantes tridimensionais foram desenvolvidos ao longo dos anos para apresentar conteúdo escrito/falado traduzido para a língua de sinais. Apesar das vantagens desse tipo de tecnologia assistiva, ainda existem críticas quanto às limitações de parâmetros não-manuais, em particular das expressões faciais nos avatares sinalizantes atuais (VIEIRA et al., 2014). Com o objetivo de verificar a deficiência em relação a incorporação de expressões faciais em avatares tradutores de **Libras**, no contexto do presente trabalho, foram realizadas comparações entre sinalizações feitas por uma intérprete de **Libras** e por sistemas de tradução português-Libras que utilizam avatares sinalizantes. A escolha dos sistemas tomou como base principal a popularidade do uso de dispositivos móveis e disponibilidade na loja de aplicativos *Play Store* do sistema operacional Android, dada a forte utilização deste sistema operacional no âmbito brasileiro. Segundo apresentado em (STATCOUNTER, 2021), em fevereiro de 2021, o mercado de sistemas operacionais móveis no Brasil contou com a participação de 86.03% do Sistema Android contra apenas 13.66% do sistema IOS, seu concorrente mais próximo.

Através de uma busca pelo termo: *Tradutor de Libras*, na *Play Store* foram selecionados os aplicativos *Hand Talk*, *VLibras* e *Rybená*. Outros aplicativos disponíveis nessa busca não apresentavam um avatar sinalizante próprio ou foram desenvolvidos para um nicho específico. A Tabela 4.1 apresenta as características dos sistemas escolhidos disponíveis na plataforma da *Play Store*.

VLibras é um sistema de código aberto desenvolvido através de uma parceria entre o Ministério da Economia e a Universidade da Paraíba, não fornece serviços pagos e

Aplicativo	Downloads na <i>Play Store</i>	Classificação (1-5)	Quantidade de avaliações
Hand Talk	Mais de 1 milhão	4,5	32.884
VLibras	Mais de 100 mil	4	1.187
Rybená	Mais de 50 mil	3,8	345

Tabela 4.1 – Características dos sistemas de tradução português-Libras informadas na plataforma *Play Store* (verificado em: 14/03/2021).

está disponível para uso em computadores, *smartphones* e páginas WEB (VLIBRAS..., 2021). O sistema Hand Talk é uma iniciativa privada que, além do sistema de tradução disponibilizado de forma gratuita para *smartphones*, fornece serviços pagos de tradução para ambientes web (HANDTALK, 2021). O projeto Rybená criado através de uma parceria entre o Grupo de Usuários Java do Distrito Federal e o Instituto Centro de Tecnologia de Software (CTS), disponibiliza, assim como o Hand Talk, o sistema de forma gratuita para *smartphone* e oferece serviços pagos de tradução para uso em ambientes WEB (RYBENÁ..., 2021). A Figura 4.4 apresenta os avatares sinalizantes em pose de descanso dos sistemas que foram analisados.

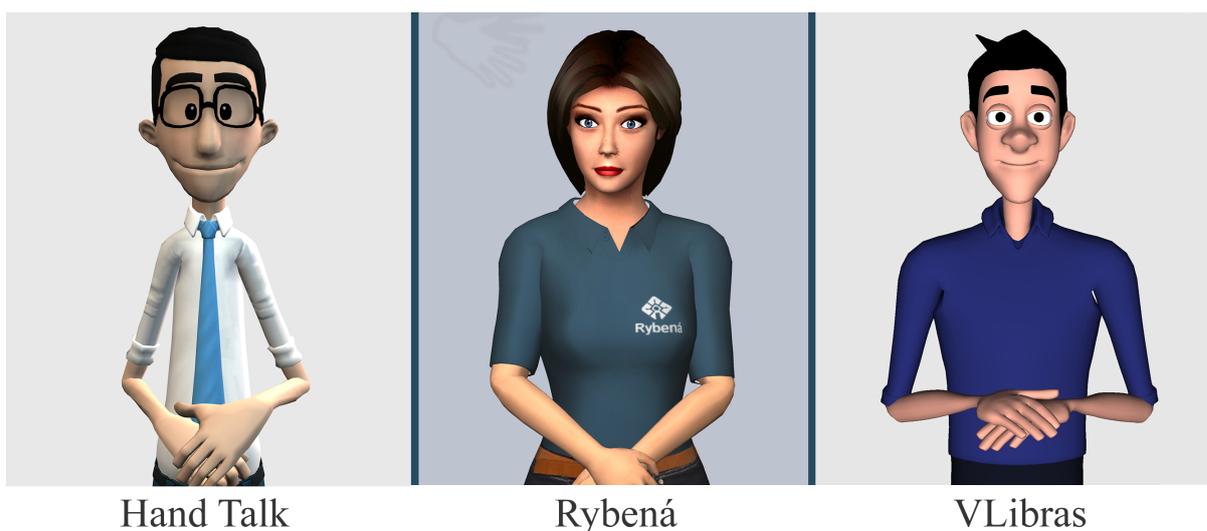


Figura 4.4 – Avatares dos sistemas de tradução: Hand Talk, Rybená e VLibras (respectivamente da esquerda para a direita)

Para as verificações, foram realizadas sinalizações de palavras que possuem a expressão facial como componente fonológico do sinal e também sinalizações de sentenças nas modalidades interrogativas parciais (com pronomes e advérbios interrogativos) e interrogativas total (do tipo sim/não). As sentenças foram também transcritas em glosas por uma pessoa fluente em Libras. Glosas são palavras de uma determinada língua oral que representam sinais manuais (MCCLEARY; VIOTTI; LEITE, 2010).

4.3.1 Expressões faciais como componentes gramaticais de palavras

O objetivo do primeiro conjunto de sinalizações é analisar a presença de expressões faciais quando elas compõem gramaticalmente o sinal. Para este conjunto foram comparadas as reproduções de algumas expressões faciais durante a sinalização dos sinais: ALEGRIA, TRISTEZA, GORDO e MAGRO. Tanto as palavras quanto as imagens da intérprete foram retiradas de (FELIPE, 2013). Conforme observado na Figura 4.5, os avatares de todos os sistemas observados reproduziram as expressões faciais de forma semanticamente semelhantes às realizadas pela intérprete.

4.3.2 Expressões faciais em sentenças interrogativas totais

Para a análise de sentenças na modalidade interrogativa total, ou seja, aquelas cujas respostas esperadas são sim ou não, foram sinalizadas duas sentenças conforme apresentadas abaixo. A intérprete recebeu e sinalizou as sentenças já transcritas em glosas.

Primeira sentença:

Português: A bicicleta quebrou?

Libras (glosas): PASSADO BICICLETA QUEBRAR?

Segunda sentença:

Português: Ana lerá o livro?

Libras (glosas): A-N-A MULHER EL@ FUTURO LER LIVRO?

Comparação

De acordo com o observado na execução realizada pela intérprete a expressão facial *sobrancelhas franzidas*, que indica a entonação da modalidade interrogativa, deveria acontecer durante a sinalização da última palavra (Figura 4.6). Contudo, os avatares não apresentaram tais expressões.

Foi observado também que durante a sinalizações dessa sentença houve algumas variações como por exemplo no sinal de *livro*. Contudo essa comparação não é o escopo deste trabalho.

Os avatares dos três sistemas ao final da sentença realizaram a sinalização da interrogação desenhando o símbolo no ar. Foi notado apenas uma variação no avatar do Rybená que após o desenho do símbolo finalizou com um movimento girando ambas as palmas das mãos para cima (Figura 4.7).

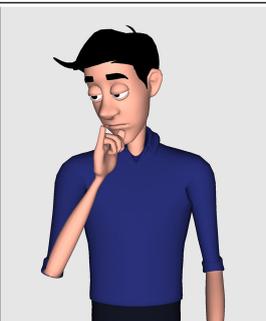
		HandTalk	Rybená	VLibras
ALEGRIA				
TRISTEZA				
GORDO				
MAGRO				

Figura 4.5 – Sinalização dos sinais: ALEGRIA, TRISTEZA, GORDO e MAGRO realizados por uma intérprete de Libras e por avatares sinalizantes de Libras (parcialmente retirado de (FELIPE, 2013))

4.3.3 Expressões faciais em sentenças interrogativas parciais

Na comparação realizada com sentenças interrogativas parciais, que são geralmente utilizadas em questões com pronomes e advérbios interrogativos, foram sinalizadas duas sentenças conforme apresentadas abaixo. A intérprete recebeu e sinalizou as sentenças já transcritas em glosas. As sentenças em português, foram escritas de maneira a deixar o pronome e o advérbio interrogativo ao final para permitir uma melhor comparação com as

- (1) Sentença em português: A bicicleta quebrou?
 Sentença em Libras (glosas): PASSADO BICICLETA QUEBRAR?
 Sinal observado: QUEBRAR



- (2) Sentença em português: Ana lerá o livro?
 Sentença em Libras (glosas): A-N-A MULHER EL@ FUTURO LER LIVRO?
 Sinal observado: LIVRO



Figura 4.6 – Sinalização de sentenças interrogativas totais (sim/não)

sentenças transcritas em glosas sinalizada pela intérprete.

Primeira sentença:

Português: O homem constrói o que?

Libras: HOMEM CONSTRUIR O-QUE?

Segunda sentença:

Português: Você carregou a mala por quê?

Libras: VOCÊ PASSADO CARREGAR-MALA, POR-QUE?

Comparação

Conforme realizado pela intérprete de **Libras** durante a sinalização da glosa



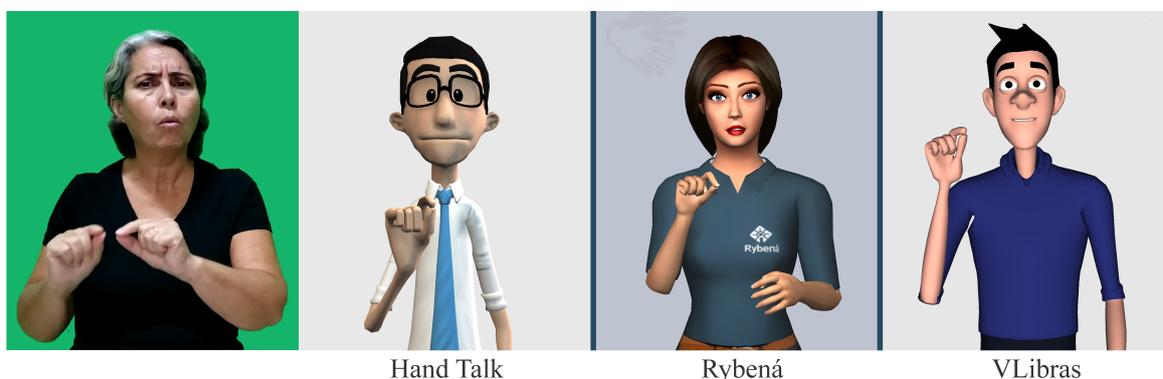
Figura 4.7 – Sinal de interrogação dos avatares dos sistema (de cima para baixo) Hand Talk, Rybená e VLibras. O símbolo em vermelho de interrogação foi inserido para representar o movimento realizado pela mão

O-QUE e da glosa *POR-QUE* a expressão facial nos avatares deveria ser: *sobrancelhas franzidas e lábios projetados* (Figura 4.8), porém os avatares não realizaram as expressões faciais esperadas.

As expressões faciais realizadas pelos avatares no momento da sinalização de *O-QUE* foram:

- No sistema Hand Talk o avatar levantou as sobrancelhas.
- No sistema Rybená o avatar levantou levemente as sobrancelhas e levantou o lábio superior.
- No sistema VLibras o avatar abriu a boca parcialmente.

- (1) Sentença em português: O homem constrói o que?
 Sentença em Libras (glosas): HOMEM CONSTRUIR O-QUE?
 Sinal observado: O-QUE



- (2) Sentença em português: Você carregou a mala por quê?
 Sentença em Libras (glosas): VOCÊ PASSADO CARREGAR-MALA, POR-QUE?
 Sinal observado: POR-QUE

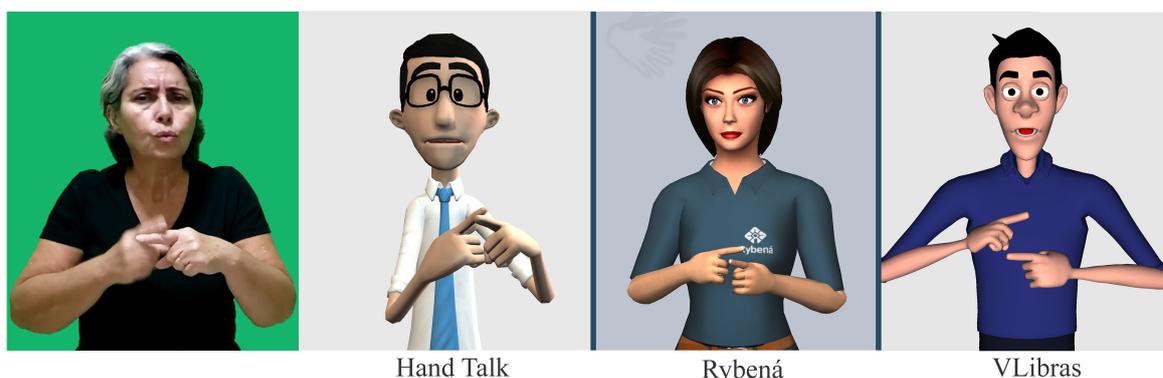


Figura 4.8 – Sinalização de sentenças interrogativas parciais

As expressões faciais realizadas pelos avatares no momento da sinalização de *POR-QUE* foram:

- No sistema Hand Talk o avatar levantou as sobrancelhas e abriu parcialmente a boca.
- No sistema Rybená o avatar permaneceu com a expressão facial neutra.
- No sistema VLibras o avatar abriu a boca.

Após a sinalização das últimas glosas das sentenças, os avatares dos sistemas Rybená e VLibras realizaram uma sinalização complementar assim como executado ao final das sentenças na modalidade interrogativa total (Figura 4.7).

Para conclusões mais assertivas é necessária uma investigação mais completa, contudo, pelas comparações e observações aqui realizadas, temos indícios das deficiências dos sistemas de traduções analisados quanto à reprodução de expressões faciais.

4.4 Comentários Finais

Como a *Libras* é uma língua visuoespacial, muitas vezes são utilizados vídeos na difusão da informação para pessoas surdas. Contudo, problemas relacionados a recuperação e apresentação de vídeos em tempo real pode atrapalhar ou impossibilitar esse processo. Como alternativa, sistemas de tradução português-Libras que empregam avatares sinalizantes são utilizados. Esse tipo de tecnologia assistiva pode trazer diversas vantagens como a possibilidade de criação da animação em tempo real. Apesar disso, ainda existem críticas quanto a problemas na apresentação de expressões faciais. Na *Libras*, as expressões faciais fazem parte da estrutura da sinalização e têm papel fundamental para determinar tipos de sentenças, intensidade e distinguir diferentes sinalizações que possuem a mesma configuração e movimentos manuais.

Neste capítulo, foi realizada uma breve análise dos principais avatares de sistemas de tradução português-Libras atuais evidenciando a incorporação de expressões faciais em sinalizações de palavras que possuem expressões faciais como componentes fonológicos, mas apresentando também deficiências na reprodução de expressões faciais características nas sinalizações de sentenças interrogativas.

Dada as atuais observações e relevância das expressões faciais em avatares sinalizantes, o objetivo principal dessa tese é desenvolver uma metodologia de incorporação de expressões faciais em avatares sinalizantes de forma a auxiliar na compreensão da mensagem sinalizada.

5 Metodologia para criação e aplicação de expressões faciais na Libras

5.1 Introdução

Neste capítulo será apresentada a metodologia de desenvolvimento e aplicação de animações de expressões faciais de **Libras** em um avatar sinalizante. Como um recorte inicial, foram analisadas e aplicadas ao avatar expressões faciais gramaticais da **Libras** relacionadas a sentenças interrogativas parciais, com pronomes e advérbios interrogativos e sentenças interrogativas totais, que têm como resposta esperada sim ou não.

As expressões faciais são componentes importantes que modulam a sinalização manual. Por isso, a metodologia apresentada busca a criação de animações faciais de sinalizações em **Libras** dada a existência da animação manual e corporal no avatar, aqui chamadas apenas de animação corporal.

A animação corporal de um sinal individual pode ser criada através da modelagem da movimentação tanto por técnicas manuais, que podem incluir processos de animação por *keyframe*, quanto por Mocaps de pessoas fluentes em **Libras**. A criação da animação corporal de uma determinada sentença pode ser feita a partir da concatenação de animações individuais de cada sinal, formando a animação final. Dessa maneira, o sistema de criação da animação corporal pode ser capaz de fornecer dados quanto ao início e duração da animação de cada sinal na animação final, importantes para guiar as animações faciais.

A metodologia desenvolvida e apresentada a seguir é composta por três etapas principais: análise, aquisição e animação como pode ser observado na ilustração apresentada pela Figura 5.1. O escopo deste trabalho não envolve, contudo, o processo de tradução de português para **Libras**.

Na etapa de análise foi criado um corpus contendo sentenças, escritas inicialmente em português, com diferentes contextos e tempos verbais. As sentenças do corpus foram transcritas para glosas com a ajuda de uma intérprete de **Libras**. Em seguida foi feita a gravação e anotação das sinalizações e por fim uma análise para entender o uso das expressões faciais nos contextos do corpus, etapa importante para auxiliar no desenvolvimento de regras para conduzir a incorporação adequada das expressões faciais no avatar.

Na etapa de aquisição foram realizadas capturas de movimento de sinalizações de sentenças selecionadas na etapa de análise que apresentam expressões faciais distintas e

capturas separadas de expressões faciais características para a composição do banco de dados de expressões faciais.

Na etapa de animação são verificadas as regras de aplicação de expressões faciais criadas na etapa de análise, realizada a seleção e recuperação das expressões faciais do banco de expressões e por fim criada a animação facial final.

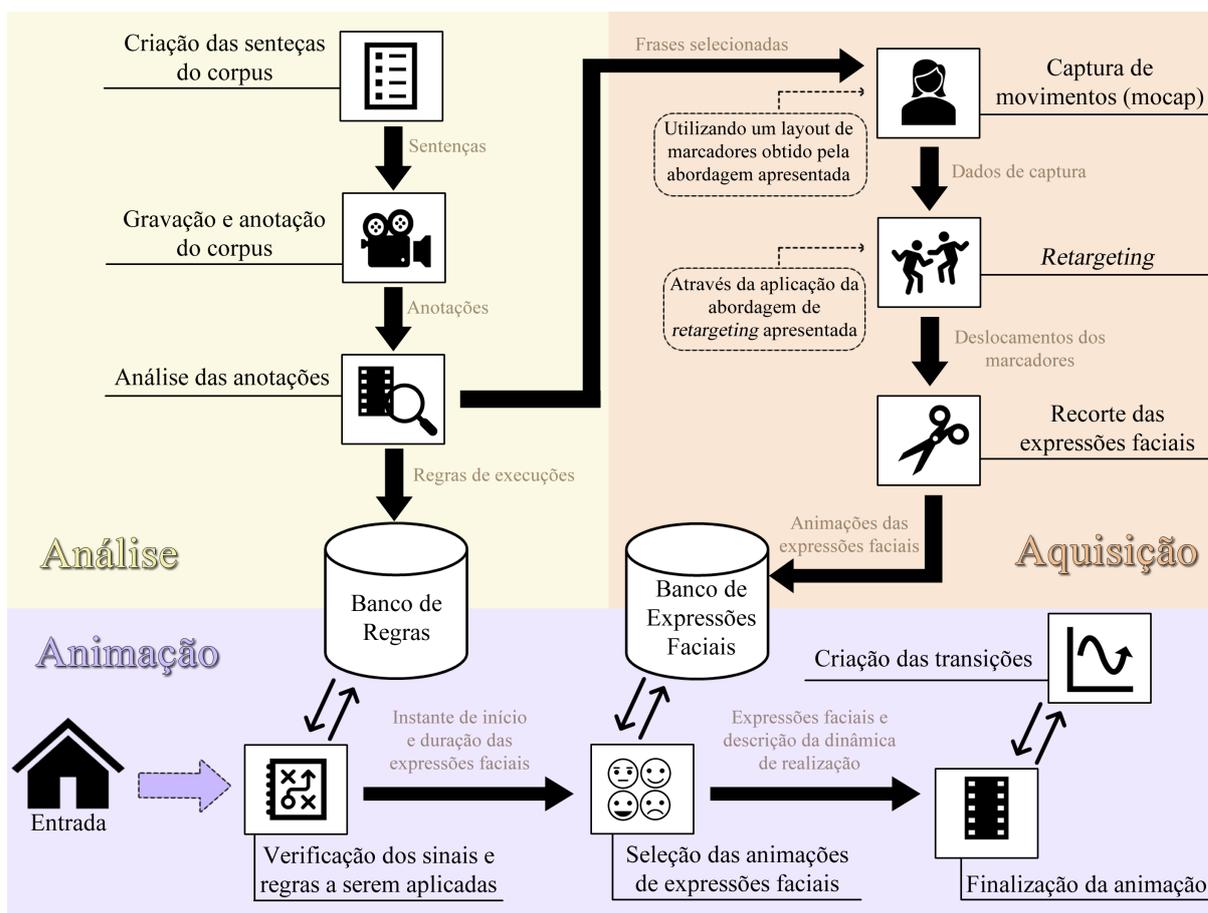


Figura 5.1 – Visão geral da metodologia de criação de animações faciais

5.2 Etapa de análise

A etapa de análise tem como objetivo investigar a execução de expressões faciais e realizar um levantamento de potenciais regras através da análise do corpus criado. Além disso, também provê um subconjunto de sentenças que apresentam uma representatividade de expressões faciais para a etapa de aquisição, onde são realizadas capturas de movimento facial. Essa etapa é composta pelas subetapas: criação das sentenças do corpus, gravação e anotação do corpus e análise das anotações.

5.2.1 Criação das sentenças do corpus

O corpus criado para os estudos deste trabalho é composto por 270 sentenças divididas em 30 grupos distintos seguindo as modalidades: afirmativa, interrogativa total e interrogativa parcial, essa última contendo 7 sentenças compostas por pronomes e advérbios interrogativos.

Cada um dos 30 grupos é formado por uma sentença base afirmativa e suas variações para atender aos diferentes tipos de sentenças. As sentenças foram escritas inicialmente em português e transcritas em glosas com a ajuda de uma intérprete de Libras. Os enunciados foram escritos nos tempos verbais: pretérito perfeito, presente e futuro do presente.

O formalismo utilizado no desenvolvimento das sentenças em glosas foi baseado nos trabalhos de Paiva (2019) e Tanya (2007). Algumas características utilizadas são listadas a seguir:

1. Os verbos são mantidos no infinitivo.
2. Não há flexão de gênero e tempo. A ausência do gênero é marcado com o caractere @. Por exemplo: a glosa EL@, pode indicar *ele* ou *ela*.
3. Na falta de um sinal correspondente é utilizado datilologia.
4. Glosas compostas que indicam um único sinal são escritas com hífen. Exemplo: CARREGAR-MALA.

Abaixo são apresentados exemplos de um grupo de sentenças do corpus:

Modalidade: Afirmativa

Português: A mulher conserta o carro

Libras: MULHER CARRO CONSERTAR

Modalidade: Interrogativa Total

Português: A mulher conserta o carro?

Libras: MULHER CARRO CONSERTAR VAI?

Modalidade: Interrogativa Parcial (Quando)

Português: Quando a mulher conserta o carro?

Libras: MULHER CARRO CONSERTAR, QUANDO?

Modalidade: Interrogativa Parcial (Onde)

Português: Onde a mulher conserta o carro?

Libras: MULHER CARRO CONSERTAR, ONDE?

Modalidade: Interrogativa Parcial (Como)

Português: Como a mulher conserta o carro?

Libras: MULHER CARRO CONSERTAR, COMO?

Modalidade: Interrogativa Parcial (Por que)

Português: Por que a mulher conserta o carro?

Libras: MULHER CARRO CONSERTAR, POR-QUE?

Modalidade: Interrogativa Parcial ((O) que)

Português: O que a mulher conserta?

Libras: MULHER CONSERTAR O-QUE?

Modalidade: Interrogativa Parcial (Qual)

Português: Qual carro a mulher conserta?

Libras: MULHER CONSERTAR, CARRO QUAL?

Modalidade: Interrogativa Parcial (Quanto)

Português: Quantos carros a mulher conserta?

Libras: MULHER CONSERTAR, CARRO QUANT@?

O corpus completo é apresentado no Apêndice A.

5.2.2 Gravação e anotação do corpus

As sentenças do corpus foram sinalizadas e gravadas por uma intérprete ouvinte e fluente em **Libras**. As gravações foram realizadas em duas sessões. A intérprete recebeu todas as sentenças em português e suas transcrições em glosas com antecedência, contudo, foi deixado livre para que ela fizesse alterações de acordo com sua maneira usual de sinalizar. As alterações e sugestões dadas pela intérprete foram aplicadas ao corpus. A intérprete

autorizou a gravação e divulgação das imagens para a pesquisa conforme documentos submetidos e aprovados pelo [Comitê de Ética em Pesquisa \(CEP\)](#) da Unicamp (Anexo A).

A Figura 5.2 mostra o momento em que a intérprete realiza uma sinalização em uma sessão de gravação. Durante as gravações foram apresentadas, em um monitor com a função de teleprompter posicionado em frente a intérprete, as sentenças em português e em glosas. A sinalizante ficou livre em tomar o tempo necessário para sinalizar cada sentença, que foi apresentada em sequência. Como equipamento de gravação dos vídeos foi utilizado um celular da marca Samsung modelo S9 posicionado em frente à intérprete.



Figura 5.2 – Sessão de gravação do corpus

As anotações dos vídeos gravados foram realizadas utilizando o software ELAN ([WITTENBURG et al., 2006](#)). Esse software permitiu que os sinais realizados em cada sentença, bem como as expressões faciais pudessem ser anotadas e através desse recurso foi possível observar a execução das expressões faciais em cada tipo de enunciado, assim como sua duração em relação as sinalizações manuais.

Para as anotações deste trabalho foram criadas três trilhas no ELAN: *Português*, *Libras* e *Expressões faciais*. Na trilha *Português* foi anotado o tempo completo de execução de cada sentença, na trilha *Libras* foram anotados os momentos de execução dos sinais em cada sentença e na trilha *Expressões Faciais* foram anotados os momentos de execuções das expressões faciais. Seguindo a convenção apresentada em ([CAPOVILLA; RAPHAEL, 2015](#)), os sinais foram anotados a partir do momento em que se realizam todos os parâmetros envolvidos no sinal até o momento em que os parâmetros da configuração de mão deixam de ser realizados. A Figura 5.3 apresenta a anotação de uma sentença utilizando o ELAN.

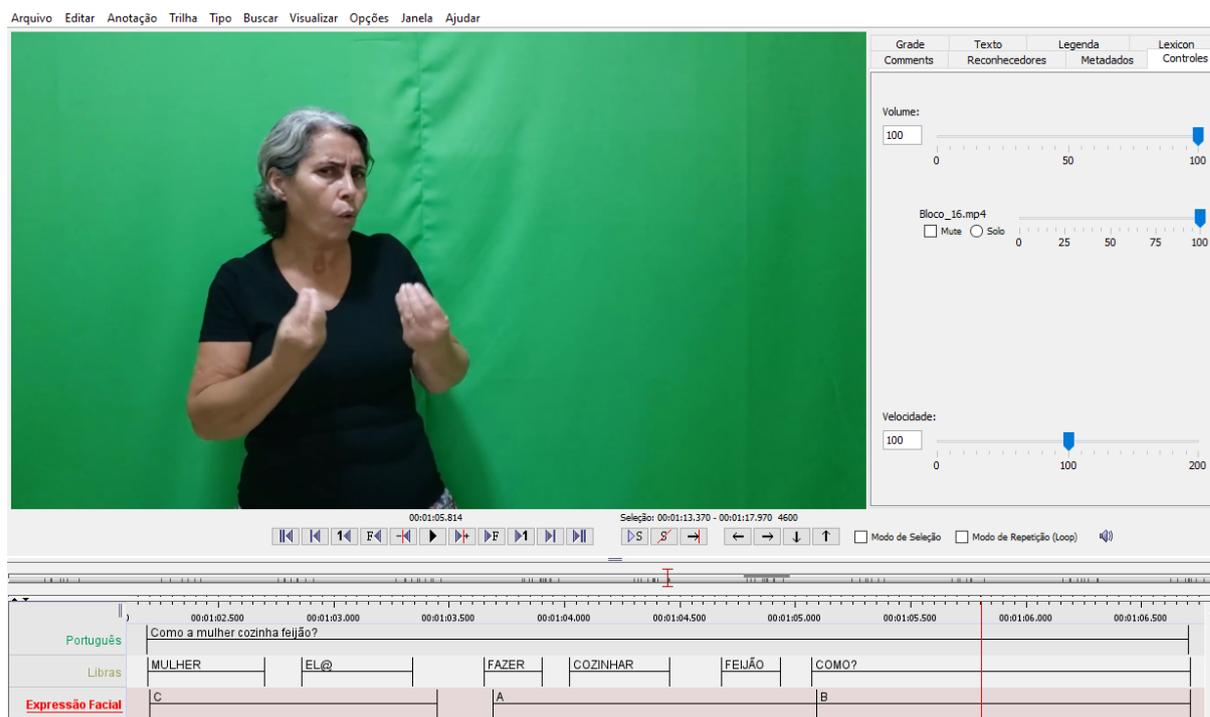


Figura 5.3 – Anotação das sentenças utilizando o ELAN

Código	Expressão Facial
A	sobrancelhas franzida
B	sobrancelhas franzidas e lábios projetados
C	Sobrancelhas levantadas
D	Sobrancelhas franzidas e centralmente levantadas

Tabela 5.1 – Expressões faciais realizadas pela intérprete durante a sinalização das glosas-chave

Expressões faciais identificadas

As expressões faciais são compostas pela combinação de modificações em cada região facial, como por exemplo alterações na boca e sobrancelhas. Neste trabalho foram anotadas as expressões faciais de forma geral resultantes da combinação das modificações de cada região. A identificação das expressões faciais foi realizada a medida que o corpus foi analisado.

As expressões faciais anotadas são as que foram realizadas simultaneamente à sinalização das glosas-chave que variavam de acordo com o tipo de sentença. Para cada expressão facial foi determinado um código para facilitar as anotações e referências (Tabela 5.1). Nas sentenças interrogativas parciais, foram tomadas como glosas-chave os advérbios e pronomes interrogativos e nas sentenças interrogativas totais, a última glosa anterior ao sinal de interrogação.

Em situações em que durante a sinalização de uma glosa foi executada a expressão facial do tipo **B** seguido por movimentos da boca relacionados a pronúncia da

glosa (*mouthing*), foi considerada a expressão facial do tipo **B** associada à glosa sinalizada. De modo geral, não foi considerada a presença de *mouthing* e de outras variações da boca com exceção das que compõem a expressão facial do tipo **B**. A Figura 5.4 apresenta exemplos das expressões faciais realizadas pela intérprete.



Figura 5.4 – Exemplo de expressões faciais executadas pela intérprete

5.2.3 Tabelamento dos dados

Para cada sentença do corpus foi anotado o tipo da expressão facial que ocorreu durante a sinalização da glosa-chave (denominada expressão facial principal), assim como a quantidade de sinais em que essa expressão facial, incluindo o sinal da glosa-chave, foi executada. Foi anotado também, como expressão secundária, o tipo da expressão facial executada imediatamente anterior a expressão principal, assim como a quantidade de sinais em que ela foi realizada. Por exemplo, durante a sinalização de uma glosa-chave em uma sentença, ocorreu a execução da expressão facial do tipo **B** e ocorreu também a expressão facial do tipo **A** durante a execução de dois sinais imediatamente anteriores.

Para a situação descrita é inserido na tabela a anotação *B1A2*. A Figura 5.5 apresenta um exemplo da situação apresentada.



Figura 5.5 – Exemplo de situação com a expressão facial do tipo **B** executada durante o sinalização da glosa-chave *QUANDO* e a expressão facial do tipo **A** durante os dois sinais anteriores

Em cada grupo de sentenças do corpus há uma sentença afirmativa, essa sentença serviu como parâmetro para verificar as expressões faciais características de cada sinal na modalidade afirmativa e averiguar as possíveis mudanças quando sinalizadas as diferentes modalidades. No processo de tabelamento dos dados foram considerados somente as expressões faciais que sofreram alterações levando em consideração a modalidade afirmativa.

Na Tabela 5.2 pode ser observado o tabelamento dos dados das expressões faciais observadas durante as sinalizações das glosas-chave em todas as sentenças do corpus. Na coluna *Grupo* estão apresentados os grupos de sentenças conforme descrito na Seção 5.2.1. Na coluna *Tempo Verbal* foram anotados os tempos verbais de cada grupo de sentenças (*Preterito Perfeito (PP)*, *Presente (PS)*, *Futuro do Presente (FP)*). Na coluna *Interrogativa Total* foram anotadas as expressões faciais relacionadas ao último sinal da sentença (imediatamente anterior ao sinal de interrogação). Nas colunas referentes à *Interrogativa Parcial*, foram anotadas as expressões faciais relacionadas aos pronomes e advérbios interrogativos.

Os dados de cada glosa-chave foram reescritos em novas tabelas para facilitar a verificação das relações entre expressões faciais e glosas-chave. As linhas nessas tabelas correspondem as sentenças em cada grupo. Nessas tabelas, para cada um dos tipos de expressões faciais (conforme apresentado na Tabela 5.1), foram inseridas duas colunas: Na primeira (**_at*) foi anotada a quantidade de sinais durante os quais a expressão facial principal foi realizada, incluindo o sinal na glosa-chave e sinais anteriores. Na segunda coluna (**_ant*) foi anotada a quantidade de sinais em que a expressão facial secundária (imediatamente anterior à expressão facial principal) foi executada simultaneamente.

Em um exemplo prático, no grupo 7 da Tabela 5.3 (Interrogativa Parcial - QUANDO) na coluna *B_at* está anotado o valor 1 e na coluna *A_ant* está anotado 2.

Grupo	Tempo Verbal	Interrogativa Total		Interrogativa Parcial							
		Último Sinal	QUANDO	ONDE	COMO	POR-QUE	O-QUE	QUAL	QUANTO		
1	PS	A4	B1A3	B1	B1A1	B1A1	B1A1	B1A1	B1C1	B1D1	
2	PS	B1A3	B1A2	B1A1	B1A3	B1A4	B1D2	B1	B1	B1A4	
3	FP	-	B1	B1D1	B1D2	B1A1	B1A1	B1C3	B1D1	B1D1	
4	PP	A5	B1D1	B1A1	B1C1	B1A1	B1	B1A1	B1A1	B1A1	
5	PS	A1	B1	B1D1	B1A1	B1A1	B1A1	B1D1	B1D1	B1A2	
6	FP	A2	B1A1	B1A1	B1A1	B1A1	B1	B1D1	B1D1	B1A1	
7	PS	A3	B1A2	B1A2	B1	A4	B1A1	B1A2	B1A2	B1A2	
8	PS	D1A4	B1A2	B1	B1A3	B1	B1A2	B1A1	B1A1	B1C3	
9	PP	A2	B1D3	B1D3	B1D1	B1A3	D2	B1D1	B1D1	B1	
10	FP	A1	B1C3	B1D2	B1A1	B1A1	-	B1D2	B1D2	B1C2	
11	PP	-	B1	B1	B1	B1	B1A1	B1	B1	B1	
12	PS	A2C1	-	B1A1	B1A1	B1A1	B1A1	B1A1	B1A1	B1A3	
13	PP	A2	B1A2	B1	B1A1	B1A1	B1D2	B1	B1	B1A2	
14	FP	A2	B1C2	B1	B1	B1D2	B1D1	B1A2	B1A2	B1A2	
15	PP	D1A3	B1A2	B1D3	B1	B1C1	B1C1	B1	B1	B1	
16	PS	C3	B1A1	B1C1	B1A3	B1A1	B1D1	B1C2	B1C2	B1C2	
17	PS	C4	B1	B1C3	B1A2	B1C3	B1A3	B1A2	B1A2	B1A2	
18	FP	A1C1	B1	B1C1	B1	B1	B1A1	B1D2	B1C1	B1C1	
19	PP	C1	B1C1	B1A1	B1	B1A1	B2	B1	B1	B1A1	
20	FP	D1	B1	B1	B1A1	A2D2	B1	B1A1	B1A1	B1A1	
21	PP	A3	B1D1	B1A1	B1A1	B1A1	B1D1	B1	B1	B1	
22	PP	A1	B1A1	B1	B1	B1D2	B1C1	B1	B1	B1	
23	FP	A4	B1A1	B1A1	B1A1	B1A1	B1A1	B1	B1	B1A1	
24	PS	A3	B1D2	B1A2	B1A2	B1D1	B1A2	B1A2	B1A2	B1A2	
25	PP	A1	B1	B1A1	B1A1	B1A1	B1	B1A1	B1A1	B1D1	
26	PS	A1	B1	B1C2	B1	B1A1	B2	B1	B1	B1	
27	FP	D5	B1A1	B1A2	B1A2	B1D1	B1A2	B1A2	B1A2	B1A2	
28	PP	A1D2	B1A1	B1A2	B1D1	B1A1	B1A1	B1A1	B1A1	B1A1	
29	FP	D1	B1A2	B1	B1	B1A2	B1A1	B1A2	B1A2	B1A2	
30	FP	A1D1	B1A2	A1D1	B1	B1A1	B1A1	B1A1	B1A1	B1	

Tabella 5.2 – Tabulamento dos dados de expressões faciais observados durante as sinalizações das sentenças do corpus

Interrogativa Parcial - QUANDO									
Grupo	T.Verb.	A_at	A_ant	B_at	B_ant	C_at	C_ant	D_at	D_ant
1	PS	0	3	1	0	0	0	0	0
2	PS	0	2	1	0	0	0	0	0
3	FP	0	0	1	0	0	0	0	0
4	PP	0	0	1	0	0	0	0	1
5	PS	0	0	1	0	0	0	0	0
6	FP	0	1	1	0	0	0	0	0
7	PS	0	2	1	0	0	0	0	0
8	PS	0	2	1	0	0	0	0	0
9	PP	0	0	1	0	0	0	0	3
10	FP	0	0	1	0	0	3	0	0
11	PP	0	0	1	0	0	0	0	0
12	PS	0	0	0	0	0	0	0	0
13	PP	0	2	1	0	0	0	0	0
14	FP	0	0	1	0	0	2	0	0
15	PP	0	2	1	0	0	0	0	0
16	PS	0	1	1	0	0	0	0	0
17	PS	0	0	1	0	0	0	0	0
18	FP	0	0	1	0	0	0	0	0
19	PP	0	0	1	0	0	1	0	0
20	FP	0	0	1	0	0	0	0	0
21	PP	0	0	1	0	0	0	0	1
22	PP	0	1	1	0	0	0	0	0
23	FP	0	1	1	0	0	0	0	0
24	PS	0	0	1	0	0	0	0	2
25	PP	0	0	1	0	0	0	0	0
26	PS	0	0	1	0	0	0	0	0
27	FP	0	1	1	0	0	0	0	0
28	PP	0	1	1	0	0	0	0	0
29	FP	0	2	1	0	0	0	0	0
30	FP	0	2	1	0	0	0	0	0
	Quant.	0	14	29	0	0	3	0	4

Tabela 5.3 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave *QUANDO*

Essa anotação indica que foi executada a expressão facial do tipo **B** somente no momento da sinalização da glosa-chave (*QUANDO*), e nos dois sinais imediatamente anteriores foi executada a expressão facial do tipo **A**.

Na última linha de cada tabela é apresentada a quantidade de sentenças em que foram executadas as expressões faciais na glosa-chave (*_at*) e em sinais imediatamente anteriores a ela (*_ant*). Nessa linha, as células em negrito indicam o maior valor das colunas *_at* e das colunas *_ant*.

As tabelas com os dados tabulados das demais glosas-chave podem ser observadas no Apêndice B.

5.2.4 Análise das anotações

A etapa de análise de anotações investigou, no contexto das sentenças do corpus, quais expressões faciais são executadas em cada modalidade e quais suas durações em relação a sinalização manual. Além disso, verificou também se existe uma relação direta com alguma outra expressão facial executada imediatamente anterior a expressão facial ligada a glosa-chave.

É importante notar que as sinalizações das sentenças do corpus foram realizadas por apenas uma intérprete, assim, este estudo não leva em consideração as possíveis variações regionais e pessoais e pode gerar alterações na ocorrência de diferentes sinalizantes.

As análises realizadas nesta etapa serviram para a construção do conjunto de regras para guiar a criação das animações faciais. Nas análises realizadas, foi anotada a relação entre uma expressão facial e a sinalização manual quando se observou a execução simultânea em mais de 50% das ocorrências.

A seguir são apresentadas as análises das expressões faciais observadas em cada glosa-chave, assim como das expressões faciais realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave. São também apresentadas as regras para a implementação de expressões faciais no sistema de animação facial para cada tipo de sentença.

Interrogativas Totais

Análise dos dados: Em 19 das 30 sentenças na modalidade interrogativa total (63,33%), a expressão facial do tipo **A** foi realizada durante a sinalização da glosa-chave (último sinal da sentença) e em 23 dessas sentenças (76,66%), nenhuma expressão facial foi realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave.

Regras para implementação: Reprodução da expressão facial do tipo **A** somente na sinalização da última glosa da sentença.

Interrogativas Parciais - QUANDO

Análise dos dados: Em 29 das 30 sentenças na modalidade interrogativa parcial do tipo *quando* (96,66%), a expressão facial **B** foi realizada durante a sinalização da glosa-chave. Em 14 dessas sentenças (46,66%), a expressão facial **A** foi realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave.

Regras para implementação: Reprodução da expressão facial do tipo **B** somente na sinalização da glosa-chave *QUANDO*.

Interrogativas Parciais - ONDE

Análise dos dados: Em 29 das 30 sentenças na modalidade interrogativa parcial do tipo *onde* (96,66%), a expressão facial **B** foi realizada durante a sinalização da glosa-chave. Em 12 dessas sentenças (40,00%), a expressão facial **A** foi realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave.

Regras para implementação: Reprodução da expressão facial do tipo **B** somente na sinalização da glosa-chave *ONDE*.

Interrogativas Parciais - COMO

Análise dos dados: Em todas as 30 sentenças na modalidade interrogativa parcial do tipo *como*, a expressão facial **B** foi realizada durante a sinalização da glosa-chave. Em 16 dessas sentenças (53,33%), a expressão facial **A** foi realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave. Desse total, em 10 sentenças (33,33%) a expressão facial foi realizada durante a execução de apenas 1 sinal anterior, em 3 sentenças (10%) durante a execução de dois sinais anteriores seguidos e em 2 sentenças (6,66%) durante a execução de três sinais anteriores seguidos.

Regras para implementação: Reprodução da expressão facial do tipo **B** somente na sinalização da glosa-chave *COMO* e a reprodução da expressão facial do tipo **A** durante a sinalização do sinal anterior a glosa-chave.

Interrogativas Parciais - POR-QUE

Análise dos dados: Em 28 das 30 sentenças na modalidade interrogativa parcial do tipo *por que* (93,33%), a expressão facial **B** foi realizada durante a sinalização da glosa-chave. Em 19 dessas sentenças (63,33%), a expressão facial **A** foi realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave. Desse total, em 16 sentenças (53,33%) a expressão facial foi realizada durante a execução de apenas 1 sinal anterior, em 1 sentença (3,33%) durante a execução de dois sinais anteriores seguidos, em 1 sentença (3,33%) durante a execução de três sinais anteriores seguidos e em 1 sentença (3,33%) durante a execução de quatro sinais anteriores seguidos.

Regras para implementação: Reprodução da expressão facial do tipo **B** somente na sinalização da glosa-chave *POR-QUE* e a reprodução da expressão facial do tipo **A** durante a sinalização do sinal anterior a glosa-chave.

Interrogativas Parciais - O-QUE

Análise dos dados: Em 28 das 30 sentenças na modalidade interrogativa parcial do tipo *o que* (93,33%), a expressão facial **B** foi realizada durante a sinalização da

glosa-chave. Em 15 dessas sentenças (50%), a expressão facial **A** foi realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave.

Regras para implementação: Reprodução da expressão facial do tipo **B** somente na sinalização da glosa-chave *O-QUE*.

Interrogativas Parciais - QUAL

Análise dos dados: Em todas as sentenças na modalidade interrogativa parcial do tipo *qual*, a expressão facial **B** foi realizada durante a sinalização da glosa-chave. Em 13 dessas sentenças (43,33%), a expressão facial **A** foi realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave.

Regras para implementação: Reprodução da expressão facial do tipo **B** somente na sinalização da glosa-chave *QUAL*.

Interrogativas Parciais - QUANT@

Análise dos dados: Em todas as sentenças na modalidade interrogativa parcial do tipo *quanto*, a expressão facial **B** foi realizada durante a sinalização da glosa-chave. Em 16 dessas sentenças (53,33%), a expressão facial **A** foi realizada em sinais imediatamente anteriores a expressão facial ligada a glosa-chave. Desse total, em 6 sentenças (20%) a expressão facial foi realizada durante a execução de apenas 1 sinal anterior, em 8 sentenças (26,66%) durante a execução de dois sinais anteriores seguidos, em 1 sentença (3,33%) durante a execução de três sinais anteriores seguidos e em 1 sentença (3,33%) durante a execução de quatro sinais anteriores seguidos.

Regras para implementação: Reprodução da expressão facial do tipo **B** somente na sinalização da glosa-chave *QUANT@* e a reprodução da expressão facial do tipo **A** durante a sinalização dos dois sinais anteriores a glosa-chave.

5.2.5 Criação do banco de regras

Com os dados das anotações analisados, foi criada uma base de dados de regras que descrevem o tipo e a forma de aplicação de expressões faciais. A catalogação inicialmente levou em consideração sentenças interrogativas totais e parciais, mas pode ser estendida a outras modalidades de sentenças e sinais.

Através da análise do corpus, como mencionado anteriormente, foi observado que podem haver situações, relacionadas a tipos de sentença, em que ocorrem alterações nas expressões faciais não somente durante a sinalização da glosa-chave da sentença, mas também durante sinalizações anteriores. Uma outra situação observada é a duração da expressão facial, que pode estender para outros sinais. Para resolver essas demandas, as

regras utilizam o seguinte formato:

$$E_1N_1E_2N_2$$

Em que E_1 indica o tipo da expressão facial a ser aplicada durante a sinalização da glosa-chave e N_1 a quantidade de sinais à esquerda da glosa-chave durante os quais a expressão facial deve ser executada. Os termos E_2N_2 são equivalentes ao descrito acima para a expressão facial secundária (realizada imediatamente antes da execução de E_1). Quando um sinal ou modalidade de sentença se relaciona com apenas um tipo de expressão facial a regra utiliza apenas os termos: E_1N_1 . Como exemplo, abaixo pode ser observado o registro relacionado a glosa POR-QUE:

POR-QUE B1A1

Pela regra descrita neste caso, será aplicada a expressão facial do tipo **B** apenas durante a sinalização da glosa-chave e a expressão facial do tipo **A** durante o sinal imediatamente anterior à glosa-chave. Ou seja, na sentença: HOMEM CONSTRUIR CASAS POR-QUE? será aplicada a expressão facial do tipo **B** durante a execução do sinal POR-QUE e a aplicação da expressão facial do tipo **A** durante a execução do sinal CASAS.

A Tabela 5.4 apresenta as regras aplicadas a cada glosa-chave das modalidades analisadas. Na modalidade interrogativa total os símbolos *? indicam que a glosa-chave é a última glosa do sentença (imediatamente anterior ao sinal de interrogação).

Modalidade	Sinal	Regra
Interrog. Total	*?	A1
Interrog. Parcial	QUANDO	B1
Interrog. Parcial	ONDE	B1
Interrog. Parcial	COMO	B1A1
Interrog. Parcial	PORQUE	B1A1
Interrog. Parcial	O-QUE	B1
Interrog. Parcial	QUAL	B1
Interrog. Parcial	QUANT@	B1A2

Tabela 5.4 – Regras de diferentes tipos de sentenças

As expressões faciais do tipo **C** e **D** não aparecem na Tabela 5.4 pois, apesar de suas execuções serem observadas durante as sinalizações das sentenças do corpus conforme apresentado na Tabela 5.2, pela análise dos dados, tiveram uma baixa frequência e consequentemente não foram relacionadas a nenhuma glosa-chave aqui investigada.

sentença	Modalidade
MULHER CARRO CONSERTAR VAI?	Inter. Total
FUTURO POLÍTICO DISTRIBUIR-DINHEIRO VAI?	Inter. Total
MENINO EL@ PASSADO QUEBRAR O-QUE?	Inter. Parcial
MENINO EL@ QUEBRAR BICICLETA PASSADO COMO?	Inter. Parcial
HOMEM CONSTRUIR O-QUE?	Inter. Parcial
VOCÊ PASSADO CARREGAR-MALA COMO?	Inter. Parcial

Tabela 5.5 – Sentenças selecionadas para captura das expressões faciais buscadas

5.3 Etapa de aquisição

Na etapa de aquisição foram realizadas capturas de movimento facial de um subconjunto de sentenças selecionadas do corpus (Tabela 5.5), além de capturas de movimento de cada expressão facial em separado. Todas as capturas de movimento foram realizadas pela mesma intérprete de Libras que participou das gravações das sinalizações do corpus.

Após o processamento e mapeamento desses dados capturados ao rosto do avatar, os trechos das animações de cada captura relacionados as expressões faciais buscadas foram recortados e salvos em uma base de expressões para serem recuperados de acordo com o mecanismo de seleção de expressões faciais da etapa de execução.

5.3.1 Captura de movimento

Como mencionado anteriormente, as capturas de movimento facial foram executadas por meio do sistema *Vicon Cara* (VICON, 2018). Para isso, foi utilizado um *layout* de marcadores baseado na abordagem descrita na Seção 3.2 e apresentada por Will, De Martino e Bezerra (2018). O *layout* empregado foi criado utilizando um coeficiente de refinamento igual a 1.1 conforme apresentado na Figura 3.8. Nesse *layout*, foi ajustado o posicionamento de alguns marcadores para melhor se adequar ao rosto da intérprete e adicionado um marcador entre os olhos para uma possível investigação, contudo, os dados deste marcador não foram utilizados na aplicação de movimentação ao avatar sinalizante.

Como marcadores foram utilizados artefatos de geometria meia-esfera pintado com tinta branca fosca para uma reflexão mais difusa. Os marcadores foram fixados no rosto do sinalizante com cola anti-alérgica. A Figura 5.6 apresenta a intérprete de Libras com os marcadores fixados no rosto e com o equipamento de captura de movimento.

Durante as sessões de capturas de movimento a transcrição por glosas das sentenças selecionadas foram apresentadas em uma televisão posicionada em frente à sinalizante. Após as capturas os dados foram verificados e processados através do software *Vicon Post* (VICON, 2021), resultando nas trajetórias tridimensionais dos marcadores



Figura 5.6 – Sinalizante com os marcadores e equipamento de captura de movimento representadas pela posição de cada marcador em cada quadro da captura (Figura 5.7).

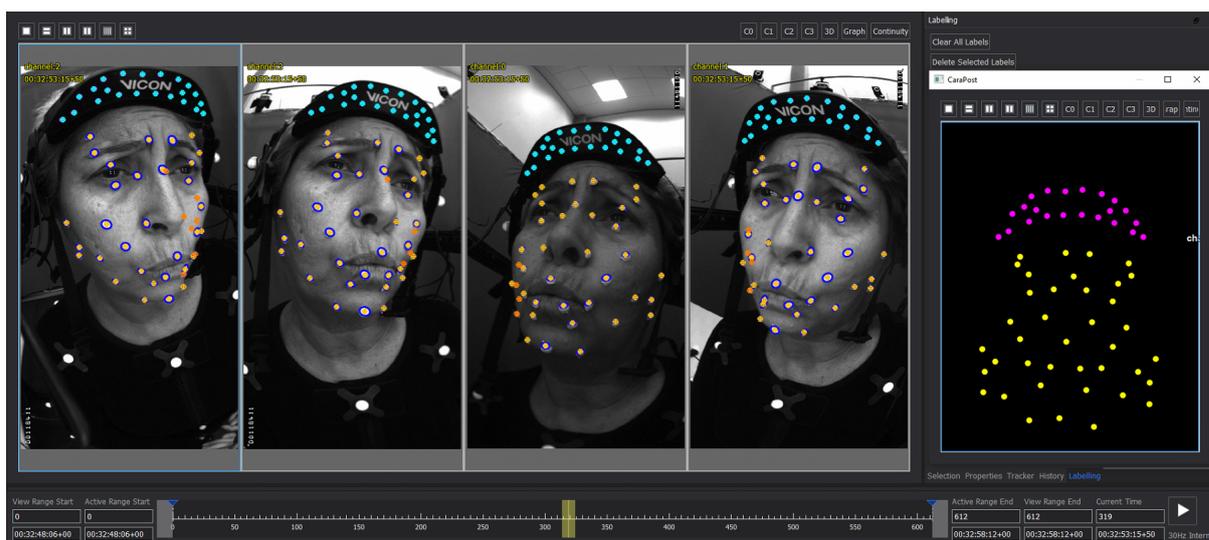


Figura 5.7 – Processamento dos dados no software *Vicon Post*

5.3.2 *Retargeting*

Na fase de *Retargeting*, que é o processo de mapeamento dos dados capturados ao rosto do avatar, os dados são ajustados para que o avatar reproduza as mesmas expressões faciais capturadas pela sinalizante. O processo de *retargeting* aplicado nessa etapa é descrito na Seção 3.3 e apresentado em Will, De Martino e Bezerra (2019).

No rosto do avatar foi criada uma configuração de ossos controladores com o mesmo *layout* utilizado pelos marcadores fixados no rosto da sinalizante. Esses ossos não deformam a malha de polígonos diretamente, mas controlam um sistema de ossos de deformação responsável pela movimentação da malha do rosto do avatar. Na Seção 6.2 são apresentados mais detalhes sobre o avatar sinalizante de Libras utilizado neste trabalho.

5.3.3 Recorte e criação do banco de expressões faciais

Em cada **Mocap**, após realizado o processo de *retargeting*, os trechos das animações contendo as expressões faciais associadas às modalidades investigadas são identificados. As coordenadas de cada marcador em cada quadro desses trechos são transformadas em vetores de deslocamento, conforme apresentado na Seção 3.3.3, e salvos no banco de expressões faciais. Para cada tipo de expressão facial buscada foram salvas animações de diferentes capturas.

5.4 Etapa de animação

Conforme já mencionado no início deste capítulo, as animações faciais são criadas a partir de informações sobre as sentenças a serem sinalizadas e sobre o instante de início e duração da sinalização de cada sinal das animações corporais aplicadas ao avatar.

Dados de Entrada

O exemplo abaixo apresenta uma sentença em glosas e informações sobre o início de cada sinal e transição dessa sentença. Idealmente essas informações deverão ser estabelecidas durante o processo de criação de animação corporal.

HOMEM ENCHER COPO QUANT@?

T375,S425,T517,S545,T720,S762,T838,S883,T993,E1070

Na segunda linha do exemplo, os valores separados por vírgula começam com um caractere indicando o início de uma transição (T), início de um sinal (S) ou o fim da sinalização (E). As transições são os momentos em que o avatar: sai de sua pose neutra e inicia os movimentos para a configuração do sinal; termina a sinalização de um sinal e parte para as configurações do próximo sinal (movimentação entre as sinalizações dos sinal); ou termina a execução de um sinal e volta à pose neutra. Os valores dos tempos são dados em quadros da animação.

5.4.1 Verificação da sentença e regras

Após a obtenção e leitura dos dados de entrada é realizada a verificação das regras. Nesse processo é executada uma busca no banco de regras para cada glosa da sentença. As regras criadas para o escopo deste trabalho foram apresentadas na Tabela 5.4, contudo, as regras podem ser estendidas para abranger outros tipos de sentenças e sinais.

Para cada glosa que tenha sido encontrada uma correspondência no banco de regras é verificado nos dados de entrada a duração de sua sinalização manual e de outros sinais anteriores de acordo com o indicado na regra.

Por exemplo, na sentença: **HOMEM ENCHER COPO QUANT@?** é verificada uma correspondência da regra B1A2 para a glosa QUANT@. Dessa forma, de acordo com o primeiro termo da regra (B1) é obtido o momento de início e duração da animação corporal do sinal QUANT@ para ser aplicada a expressão facial do tipo **B** e de acordo com o segundo termo da regra (A2), que indica a execução da expressão facial do tipo **A** durante os dois sinais a esquerda do principal (sinal imediatamente a esquerda e mais um), é verificado o início da sinalização manual do sinal ENCHER até o final da sinalização do sinal COPO.

5.4.2 Seleção e ajustes das animações faciais

Com os dados sobre o tipo e duração da expressão facial de cada sinal a ser incluído na animação final, é buscada a animação no repositório de expressões faciais, de acordo com o tipo indicado pelas regras, que apresentar maior aproximação com a duração esperada, em seguida é aplicado o processo de ajuste de duração para que a animação da expressão facial a ser inserida corresponda com a duração necessária na animação final.

Quando a animação recuperada do repositório tiver uma duração maior do que a necessária é realizada uma redução através de descarte de quadros da animação (Figura 5.8-A) e quando tiver uma duração menor são realizadas inserções de quadros interpolados (Figura 5.8-B).

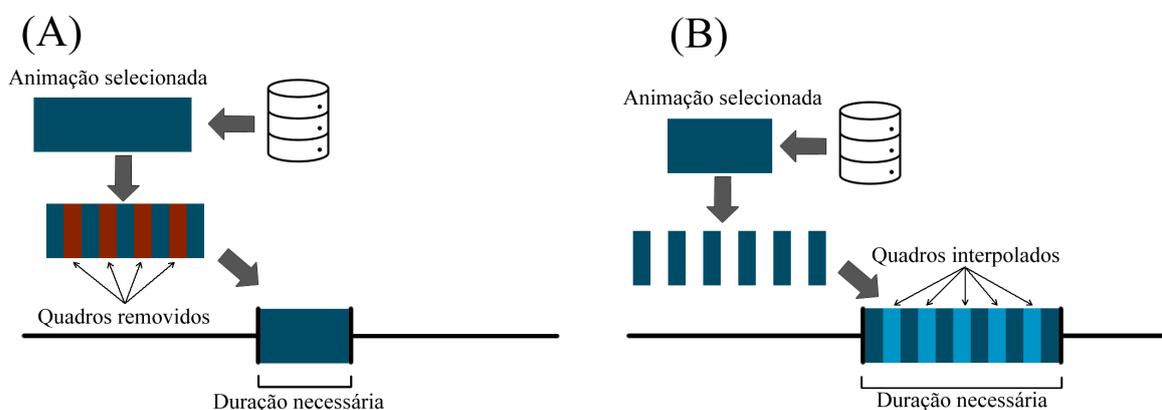


Figura 5.8 – Ilustração dos processos de ajuste de duração de animações. (A) Redução da duração da animação. (B) Aumento da duração da animação

5.4.3 Finalização da animação

Na fase de finalização da animação é criada uma animação neutra, ou seja, com todos os vetores de deslocamentos que compõe a animação nulos, e com a mesma duração da animação corporal através das informações recebidas como parâmetros iniciais.

Após a verificação das regras para cada glosa da sentença de entrada, os trechos das animações escolhidos têm suas durações ajustadas e são inseridos na animação final

nas posições calculadas de acordo com os dados de entrada e regras associadas. Em seguida são criadas e inseridas as transições entre os trechos de animações. A Figura 5.9 ilustra o processo de criação da animação facial final. Os passos do processo foram numerados em vermelho e detalhados a seguir:

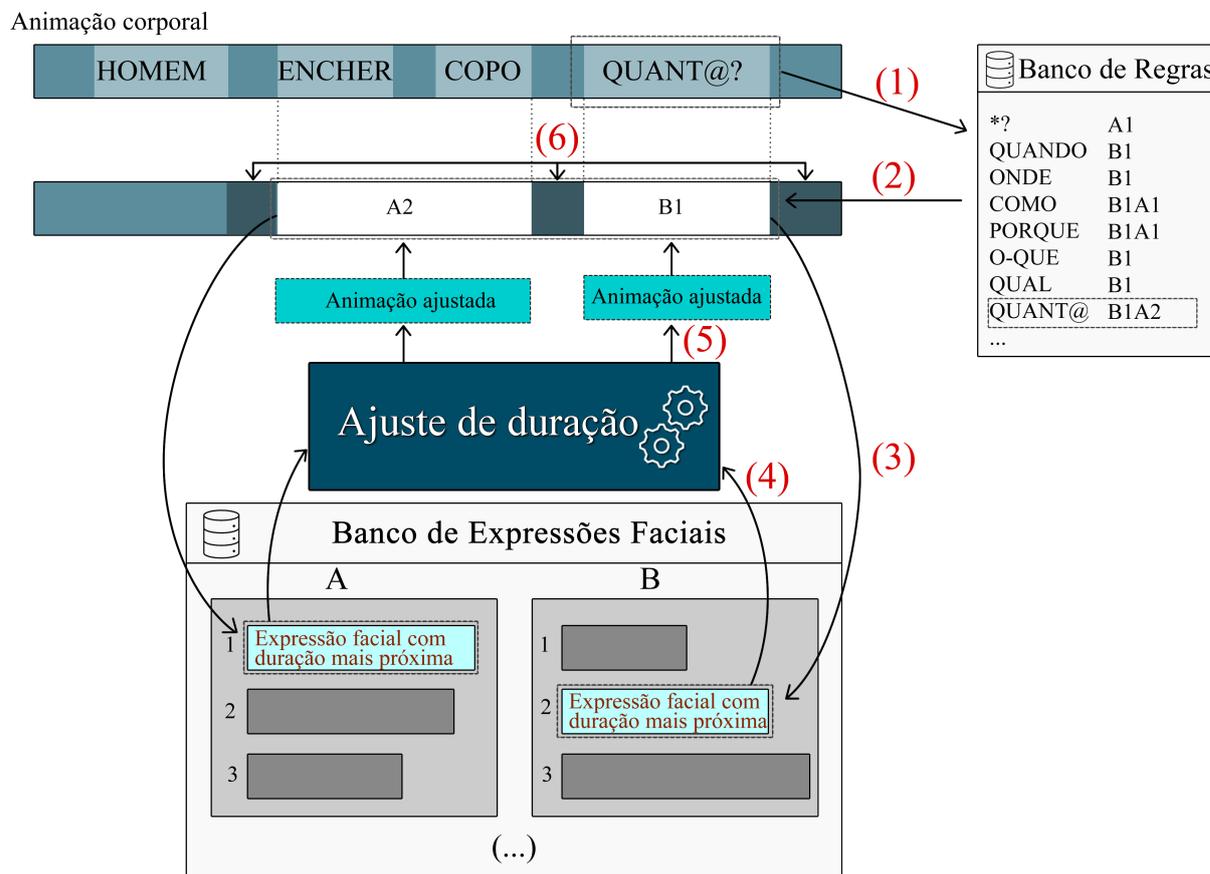


Figura 5.9 – Processo de criação da animação final

1. Para cada glosa (que é verificada da direita para a esquerda na sentença) é verificado se há correspondência no banco de regras. Em caso positivo, é recuperada e aplicada a regra associada.
2. Através das regras e dados de duração da animação corporal são calculados os intervalos, na animação final, em que os trechos das animações das expressões faciais deverão ser inseridos.
3. Para cada expressão facial indicada pelo banco de regras, é buscada no repositório de expressões faciais, de acordo com o tipo requerido, a animação com duração mais próxima do tamanho do intervalo a ser inserido.
4. Em cada animação recuperada do repositório é aplicado o processo de ajuste de duração para se adequar ao intervalo de animação calculado na etapa 2.

5. As animações faciais ajustadas são inseridas na animação final nos momentos determinados pelos dados de entrada e regras associadas.
6. São criados e inseridos os trechos de animações das transições através de representações de curvas paramétricas nos momentos de início, fim e entre as animações faciais para garantir a continuidade da animação.

5.4.4 Criação das transições

Os trechos de animações armazenados no banco de expressões faciais como já apresentado, incluem apenas a animação alvo e não as transições anteriores ou posteriores à expressão facial que são necessárias para a continuidade da animação. Essas transições são criadas através de curvas paramétricas. Para isso, foi adotado como representação da movimentação das transições a curva paramétrica cúbica de *Hermite* (FOLEY et al., 1990). A Equação 5.1 apresenta o modelo paramétrico utilizado. Na Figura 5.10 é possível observar uma ilustração de aplicação da curva paramétrica cúbica de *Hermite* com o posicionamento dos pontos de controle.

$$P(t) = \begin{bmatrix} t^3 & t^2 & t & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ \dot{P}_1 \\ \dot{P}_2 \end{bmatrix} \quad (5.1)$$

Em que:

- $P(t)$ é o ponto de interesse;
- t a variável independente da representação paramétrica, normalizada pelo intervalo entre os pontos inicial e final da curva;
- P_1 o ponto inicial da curva;
- P_2 o ponto final da curva;
- \dot{P}_1 a tangente no ponto inicial;
- \dot{P}_2 a tangente no ponto final.

Para garantir a continuidade da animação levando em consideração a duração da transição, a tangente no ponto inicial (\dot{P}_1) é criada através da Equação 5.2:

$$\dot{P}_1 = (S1_i - S1_{i-1}).L \quad (5.2)$$

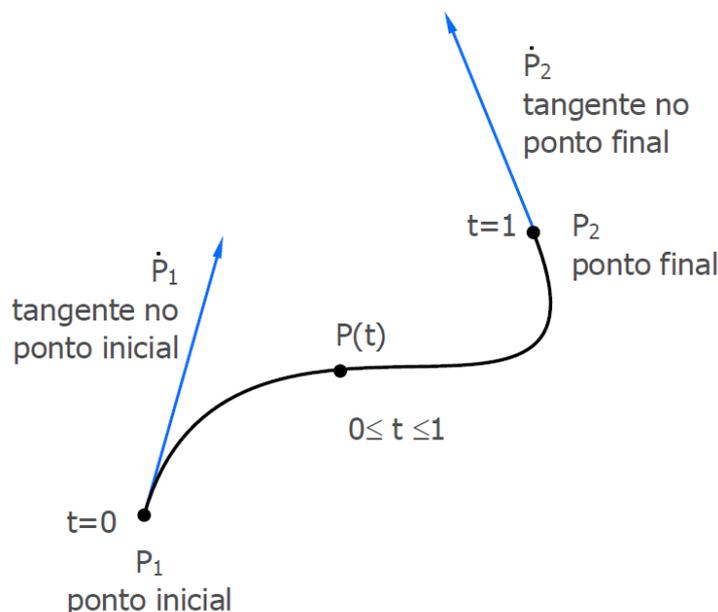


Figura 5.10 – Ilustração de representação da curva paramétrica cúbica de Hermite

Em que \dot{P}_1 é a tangente no ponto inicial da curva, $S1_i$ o último ponto do sinal anterior à transição, $S1_{i-1}$ o penúltimo ponto do sinal anterior à transição e L o comprimento da transição dado em números de quadros.

A tangente no ponto final (\dot{P}_2) é criada através da Equação 5.3:

$$\dot{P}_2 = (S2_2 - S2_1).L \quad (5.3)$$

Em que \dot{P}_2 é a tangente no ponto final da curva, $S2_2$ o segundo ponto do sinal posterior à transição, $S2_1$ o primeiro ponto do sinal posterior à transição e L o comprimento da transição dado em números de quadros.

5.5 Comentários finais

A metodologia para aplicação de expressões faciais de **Libras** descrita neste capítulo propõe o desenvolvimento de três etapas principais distintas. A primeira delas, etapa de análise, tem como objetivo a identificação de expressões faciais associadas aos tipos de sentenças apresentadas em um corpus de sinalizações de **Libras** criado e o desenvolvimento de um banco de regras (para a aplicação de expressões faciais em um avatar sinalizante) através da análise do corpus. A segunda, etapa de aquisição, tem como objetivo o desenvolvimento de um repositório de trechos de animações de expressões faciais provenientes de **Mocap** de um conjunto de sinalizações realizadas por uma intérprete de **Libras**. A terceira, Etapa de animação, é responsável pela criação da animação facial final guiada pelas regras resultantes da etapa de análise e utilizando trechos de animações

faciais providos pela etapa de aquisição.

Vale ressaltar que uma vez executadas as etapas de análise e aquisição e criado assim o banco de regras e o repositório de animações faciais, a etapa de animação pode ser realizada para a criação de diversas animações faciais. É importante notar também que apesar do escopo deste trabalho ser relacionado a expressões nas modalidades interrogativas totais e parciais com o emprego de pronomes e advérbios interrogativos, a metodologia apresentada pode ser expandida para contextos mais amplos com outras modalidades de sentenças e sinais em [Libras](#).

6 Implementação piloto

6.1 Introdução

Como ferramenta de avaliação da metodologia para aplicação de expressões faciais de **Libras** descrita no Capítulo 5 foi desenvolvido um protótipo de um sistema de animação facial. O sistema implementa a etapa de **Animação** (Seção 5.4) com o objetivo de avaliar a dinâmica da animação assim como as regras de aplicação resultantes da etapa de **Análise** (Seção 5.2) e o banco de expressões faciais provido pela etapa de **Aquisição** (Seção 5.3). A Figura 6.1 apresenta, através de um diagrama de blocos, uma visão geral do protótipo desenvolvido. Os retângulos com cantos arredondados representam a entrada e saída do sistema. Os retângulos amarelos representam as módulos principais do sistema, os retângulos azuis os módulos auxiliares, os cilindros os banco de regras e de expressões faciais e o retângulo verde os dados da animação corporal.

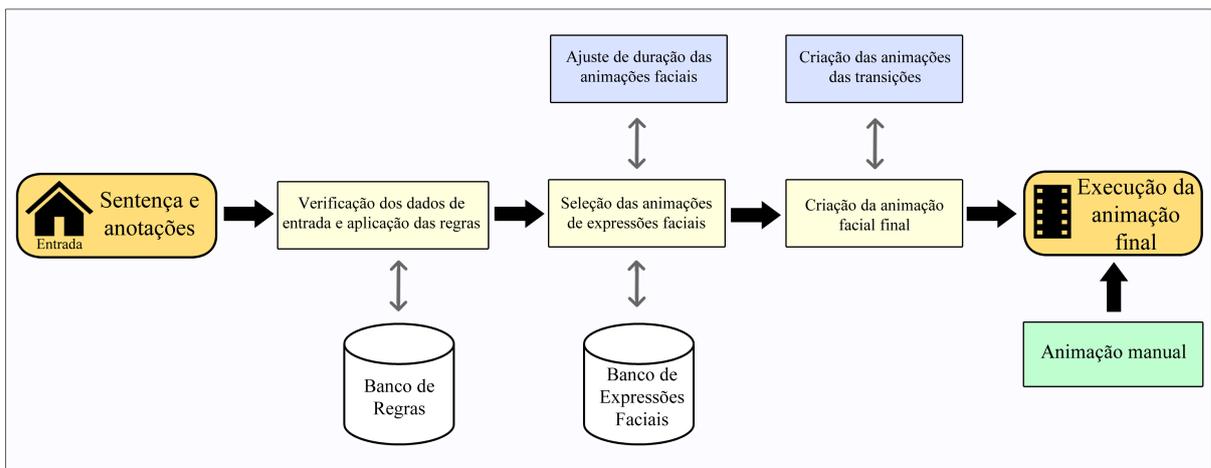


Figura 6.1 – Arquitetura geral do sistema piloto

Os códigos da implementação do sistema piloto e dos arquivos referentes ao banco de regras e de animações de expressões faciais necessários para a execução estão disponíveis na plataforma do GitHub ([GITHUB, 2021](https://github.com/ackleywill/PilotoAnimacaoFacial)) pelo endereço:

<https://github.com/ackleywill/PilotoAnimacaoFacial>.

O Apêndice C apresenta informações sobre os arquivos disponibilizados e sobre a execução do sistema.

6.1.1 Sentenças e anotações

Como entrada do sistema é esperada a sentença a ser sinalizada no formato de glosas bem como as anotações acerca da duração e início da sinalização de cada sinal da

animação corporal, conforme apresentado na Seção 5.4, para que a animação facial possa ser sincronizada e assim apresentada como componente do sinal. O sistema de criação da animação corporal, conforme mencionado na Seção 5.1, pode apresentar dados, com descrições dos instantes de início e fim de cada sinal na animação produzida, que podem ser utilizados para guiar o sistema de animação facial. O sistema piloto aqui apresentado, foi testado com dados de anotações obtidos de forma manual conforme o modelo descrito na Seção 5.4.

A geração final da animação facial pode ser realizada de forma *off-line* para uma apresentação posterior a sua criação ou de forma simultânea ao processo de criação. Na forma simultânea, a apresentação da animação deve ter uma taxa de geração de quadros e latência que permitam uma visualização confortável da animação sem a introdução de artefatos distratores como a variação da velocidade de apresentação e atrasos excessivos. O escopo deste protótipo é definido somente para a apresentação da animação de forma *off-line*.

6.1.2 Verificação dos dados de entrada e aplicação das regras

Conforme já mencionado, é necessário como dados de entrada, a sentença em glosas e informações sobre o instante de início e duração de cada sinal na animação corporal que a animação facial acompanhará. O formato utilizado desses dados é apresentado na Seção 5.4. Neste padrão são anotados os momentos em que cada sinal é constituído e os momentos de início de cada transição além do momento de término da animação.

Na Figura 6.2 é ilustrado um exemplo de de anotação dos instantes de início dos sinais (iniciados por S), início das transições (iniciado por T) e fim da animação (iniciado por E) de uma animação corporal.

Os dados anotados da sentença ilustrada na Figura 6.2 é apresentado a seguir:

HOMEM CONSTRUIR CASAS COMO?

T488,S540,T593,S630,T695,S740,T805,S835,T948,E1038

Como apresentado, este trabalho tem como escopo as expressões faciais de sentenças interrogativas, contudo a metodologia apresentada pode ser aplicada a outros tipos de expressões faciais gramaticais da *Libras*. Neste contexto, como forma de avaliação, foi observada a sinalização da sentença do exemplo acima pela intérprete e consideradas as expressões faciais executadas durante cada sinal, quando estas forem relacionadas a uma das expressões faciais levantadas (Tabela 5.1), e incluídas as regras no banco de regras. Neste exemplo, além da regra relacionada à glosa-chave da sentença interrogativa, foram consideradas as expressões faciais do tipo **C** e **A** associadas respectivamente aos sinais

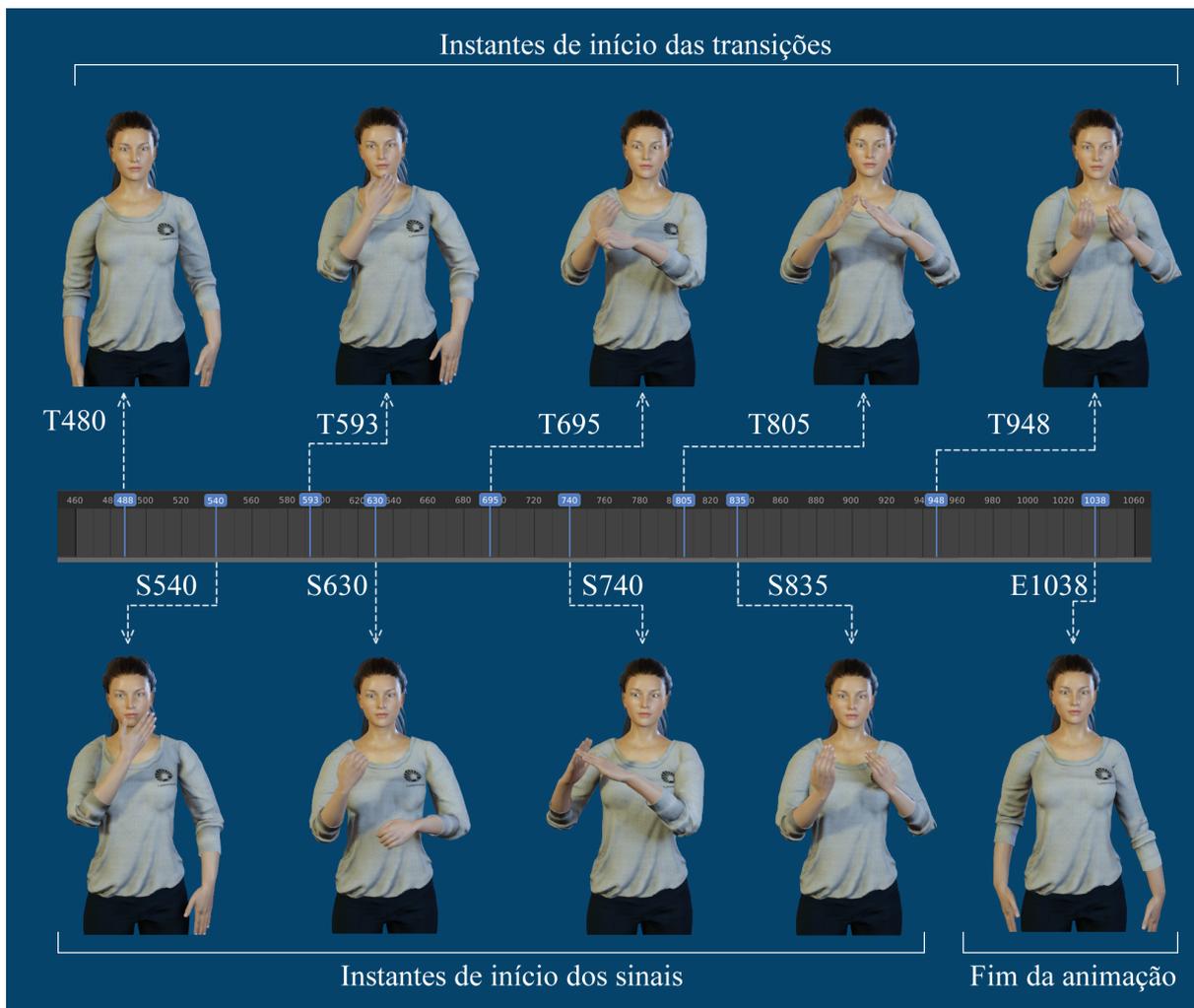


Figura 6.2 – Ilustração da anotação dos sinais e transições na animação corporal da sentença: **HOMEM CONSTRUIR CASAS COMO?**

HOMEM e CONSTRUIR. Abaixo são apresentadas as regras ligadas às glosas nos dados de entrada do exemplo:

HOMEM C1

CONSTRUIR A1

COMO? B1A1

Após a leitura dos dados de entrada, para cada sinal, são calculados: o instante de início, sua duração e a duração da transição anterior e posterior. O formato de apresentação dos dados de cada sinal pelo sistema segue a seguinte estrutura:

$$[sign, start, len, lenT1, lenT2]$$

Em que *sign* é o sinal a ser executado, *start* o instante em que a animação do sinal será posicionada na animação final, *len* a duração da animação do sinal, *lenT1* a duração da transição a esquerda do sinal e *lenT2* a duração da transição a direita do sinal.

Abaixo é apresentada a saída do sistema dessa primeira parte em uma execução com os dados de entrada de acordo com o exemplo apresentado:

```
['HOMEM', 540, 53, 52, 37], ['CONSTRUIR', 630, 65, 37, 45],  
['CASAS', 740, 65, 45, 30], ['COMO?', 835, 113, 30, 90]
```

6.1.3 Seleção das animações de expressões faciais

Para cada expressão facial a ser incluída na animação final, após a averiguação do tipo, instante de início e duração necessária, é buscado no repositório a animação que seja do mesmo tipo e que apresente a menor diferença entre sua duração e a duração necessária na animação final. As durações das animações recuperadas no banco de expressões faciais são ajustadas de acordo com o apresentado na Seção 5.4.2 para se adequarem a duração necessária.

No Apêndice D é apresentado um relatório do sistema com informações sobre o processo de seleção e ajustes das animações em uma execução com os dados de entrada do exemplo apresentado.

6.1.4 Criação da Animação facial final

Na fase de geração da animação final é criada uma animação vazia (com todos os vetores de deslocamentos nulos) com duração de acordo com a animação corporal. Seguindo os instantes de início das sinalizações manuais, as animações das expressões faciais selecionadas e ajustadas são posicionadas na animação criada. Após o posicionamento das animações das expressões faciais são criadas as transições para garantir a continuidade da animação. Para essa representação conforme já apresentado, foi utilizada a curva paramétrica cúbica de *Hermite*.

A Figura 6.3 apresenta o deslocamento no eixo X de um marcador da região da sobrelha durante a sinalização de uma sentença. O gráfico a esquerda da figura apresenta o posicionamento das animações faciais de cada expressão facial na animação final e o gráfico à direita da figura a inserção das transições criadas (faixas beges).

6.2 Execução da animação final

Para a execução das animações faciais produzidas pelo sistema piloto foi utilizado um avatar tridimensional modelado para o projeto [Tecnologia Assistiva para](#)

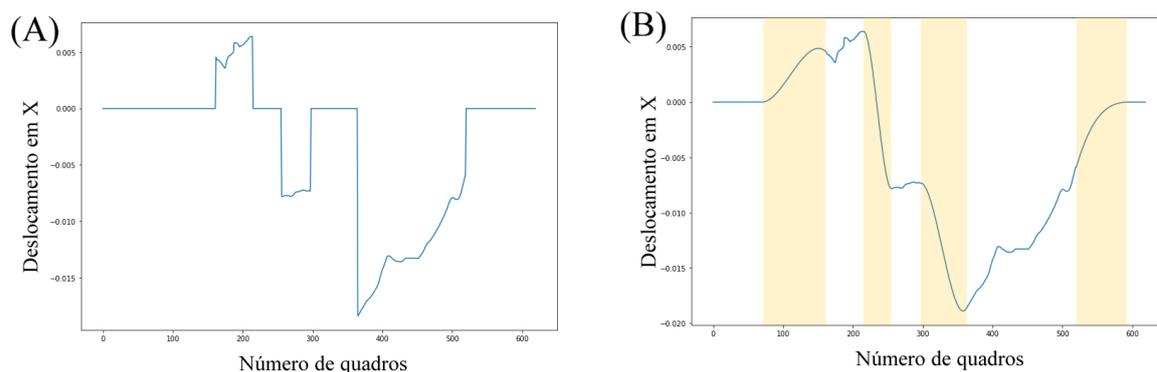


Figura 6.3 – Deslocamento no eixo X de um marcador ao longo da animação facial de uma sentença. (A) Animações faciais inseridas na animação final. (B) Inserção das animações de transições criadas

Surdos (TAS) (DE MARTINO et al., 2016) (Figura 6.4).

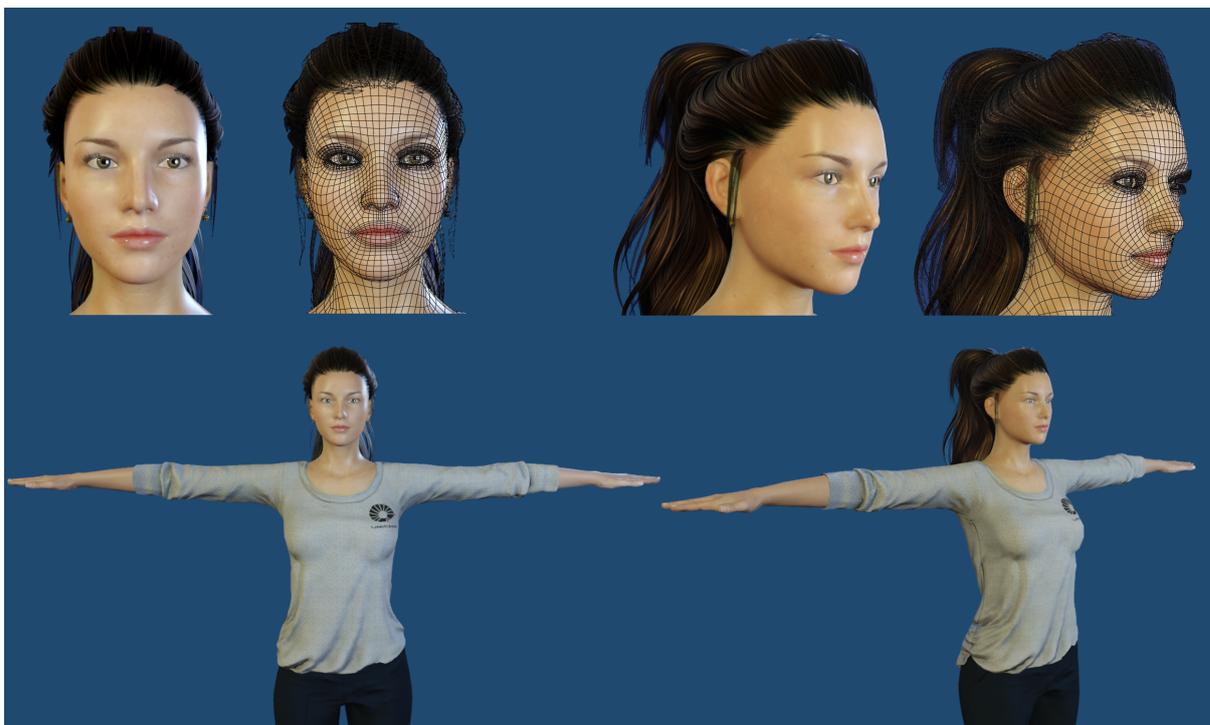


Figura 6.4 – Avatar do projeto TAS utilizado para a aplicação das animações

O avatar possui um sistema de ossos responsável por deformar a malha de vértices do rosto para produzir a movimentação facial (Figura 6.5 - coluna B). Esses ossos são acionados por um conjunto de ossos controladores (Figura 6.5 - coluna A) que seguem as movimentações dos dados de *Mocap* ajustados ao rosto do avatar através da aplicação da abordagem de *retargeting* apresentada na Seção 3.3. No *layout* dos ossos controladores foi acrescentado um osso em cada sobrancelha e dois ossos na região boca para ajustes das expressões faciais. Esses ossos controladores auxiliares não possuem um marcador correspondente no *layout* de dados do *Mocap* e seus movimentos são obtidos pelos dados

dos marcadores correspondentes aos ossos controladores vizinhos.

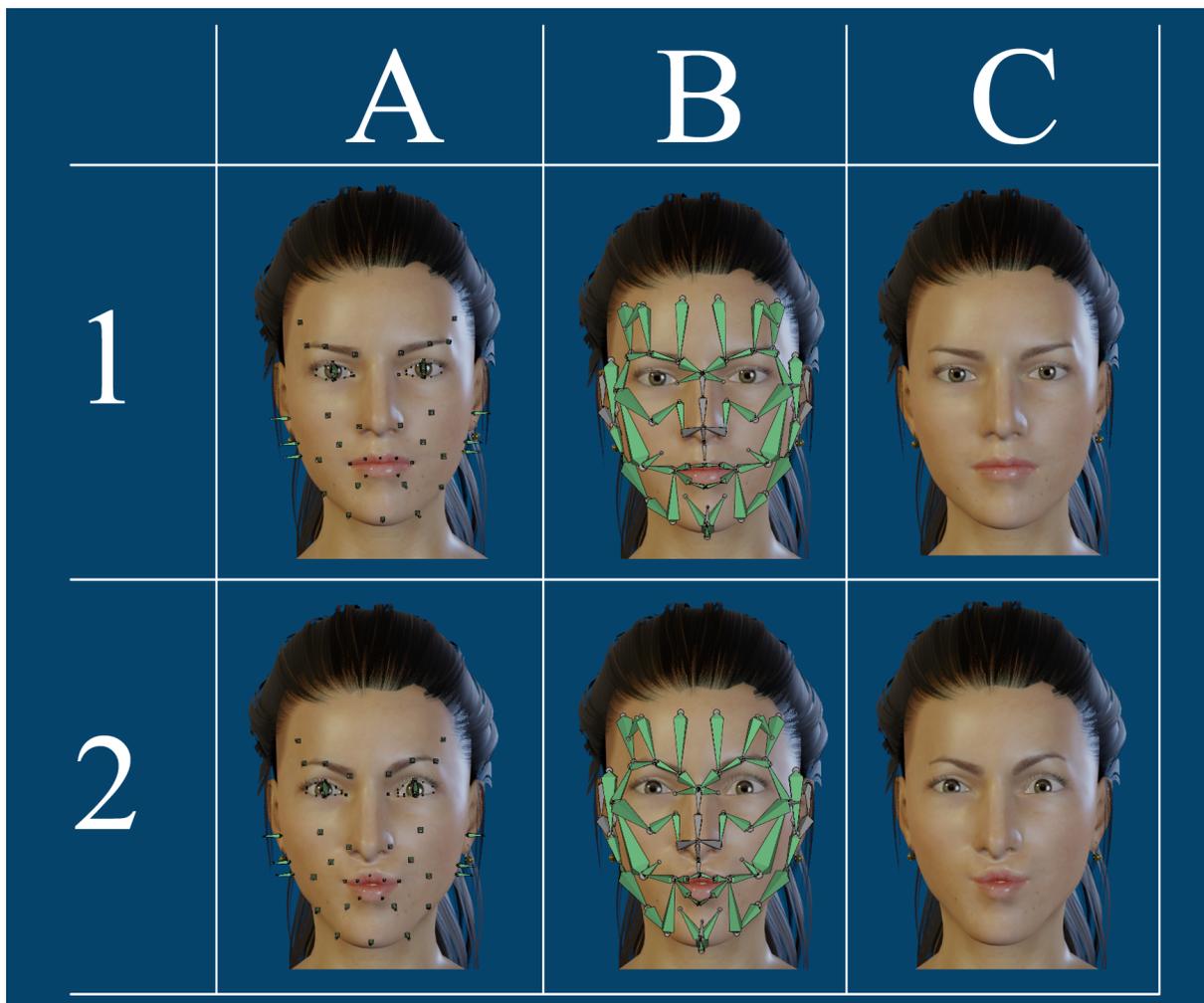


Figura 6.5 – Linha 1: Expressão facial neutra. Linha 2: Expressão facial do tipo **B** (sobrancelhas franzidas e lábios projetados). Coluna A: Ossos controladores. Coluna B: Sistema de ossos de deformação. Coluna C: Expressões faciais resultantes

6.3 Criação e execução das animações faciais

Nesta seção são apresentados os resultados de aplicações de animações de expressões faciais geradas pelo protótipo para acompanhar as animações corporais já presentes no avatar sinalizante.

Na plataforma do projeto no GitHub estão disponíveis também arquivos do avatar sinalizante (*Anim_0.rar* e *Anim_1.rar*) já com a animação corporal e com o *script* para a inclusão da animação facial. O Apêndice E apresenta informações para a realização da incorporação da animação facial no avatar sinalizante.

As expressões faciais levantadas e estudadas neste trabalho são focadas naquelas que são características de interrogativas totais e parciais. Contudo, para testar a usabilidade e extensibilidade da metodologia, nos resultados aqui apresentados, assim como realizado

na Seção 6.1.2, além das expressões faciais relacionadas as sentenças estudadas, foram consideradas as expressões faciais executadas durante cada sinal, quando relacionadas a uma das expressões faciais levantadas.

Para a avaliação do mecanismo de criação das animações faciais foram realizadas capturas de movimento corporal de algumas sentenças do corpus. Os dados dessas capturas foram processados e incorporados ao avatar sinalizante utilizado neste trabalho. Com o avatar já animado foram então realizadas anotações dos instantes de início de cada sinal e transição. Conforme apresentado anteriormente, essas anotações servem de entrada para o sistema de criação de animações faciais. As animações faciais criadas pelo protótipo foram inseridas no avatar já com as animações corporais aplicadas.

Nos experimentos apresentados a seguir, para cada uma das animações reproduzidas são exibidas imagens dos momentos da execução de cada sinal da sentença pelo avatar e a mesma sinalização realizada pela intérprete. Na parte superior das figuras são apresentados os sinais e o tipo de expressão facial sinalizada. Para efeito de comparação é apresentado em cada figura o avatar e a intérprete na pose neutra. Para cada animação é apresentado também um gráfico com a trajetória de um osso controlador da sobrancelha do rosto do avatar durante a animação e evidenciados os trechos onde foram inseridas as transições criadas pelo sistema para a verificação da continuidade da animação.

6.3.1 Animação da sentença: CASAS HOMEM EL@ VENDER VAI?

Nesta seção é exibido o resultado da criação e execução da animação facial da sentença "CASAS HOMEM EL@ VENDER VAI?". De forma a testar o sistema, neste experimento, a associação da expressão facial do tipo **A** aos sinais CASAS e EL@ e a associação da expressão facial do tipo **C** ao sinal HOMEM foram inseridos no banco de regras para a criação da animação. Essas associações foram realizadas com base na observação das expressões faciais realizadas pela intérprete durante a sinalização dos sinais.

Abaixo são apresentados os dados de entrada inseridos e processados pelo sistema assim como as regras utilizadas. O formato dos dados processados e das regras são apresentados na Seção 6.1.2.

Dados de entrada:

CASAS HOMEM EL@ VENDER VAI?

T360,S440,T535,S573,T620,S643,T691,S726,T766,S815,T914,E1023

Dados de entrada processados:

['CASAS', 440, 95, 80, 38], ['HOMEM', 573, 47, 38, 23],

['EL@' , 643, 48, 23, 35] , ['VENDER' , 726, 40, 35, 49] ,
['VAI?' , 815, 99, 49, 109]

Regras utilizadas:

CASAS A1

HOMEM C1

EL@ A1

*? A1

Animação

A Figura 6.6 apresenta uma sequência de imagens obtidas durante a animação da sinalização de cada sinal da sentença pelo avatar e pela intérprete. Na imagem relacionada a sinalização dos sinais CASAS, EL@ e VAI pode ser observada a expressão facial do tipo **A** (sobrancelhas franzidas) e durante a sinalização do sinal HOMEM a expressão facial do tipo **C** (sobrancelhas levantadas). Mesmo que mais de um sinal tenha associado o mesmo tipo de expressão facial, como no caso dos sinais CASAS, EL@ e VAI com a expressão facial do tipo **A**, não necessariamente utilizarão a mesma animação facial do repositório, pois a busca leva em consideração a duração da animação necessária de cada sinal.

O vídeo da animação (Anim_0.mp4) está disponível no repositório do projeto no GitHub e pode ser acessado pelo endereço:

https://github.com/ackleywill/PilotoAnimacaoFacial/blob/main/videos/Anim_0.mp4.

Gráficos da animação

O gráfico na parte superior da Figura 6.7 apresenta o deslocamento no eixo X do osso controlador mais ao centro da sobrancelha esquerda do rosto do avatar durante a animação. As faixas azuis marcam os trechos das execuções das expressões faciais recuperadas do repositório e as faixas bege os trechos das transições criadas pelo sistema. Os gráficos da parte inferior da figura apresentam, de forma destacada, o deslocamento do osso controlador em cada transição criada ao longo da animação.

6.3.2 Animação da sentença: HOMEM CONSTRUIR CASAS COMO?

Nesta seção é exibido o resultado da criação e execução da animação facial da sentença "HOMEM CONSTRUIR CASAS COMO?". Neste experimento, a associação das expressões faciais do tipo **C** ao sinal HOMEM e do tipo **A** ao sinal CONSTRUIR, foram inseridos no banco de regras. Essas associações foram realizadas com base nas observações das expressões faciais realizadas pela intérprete.

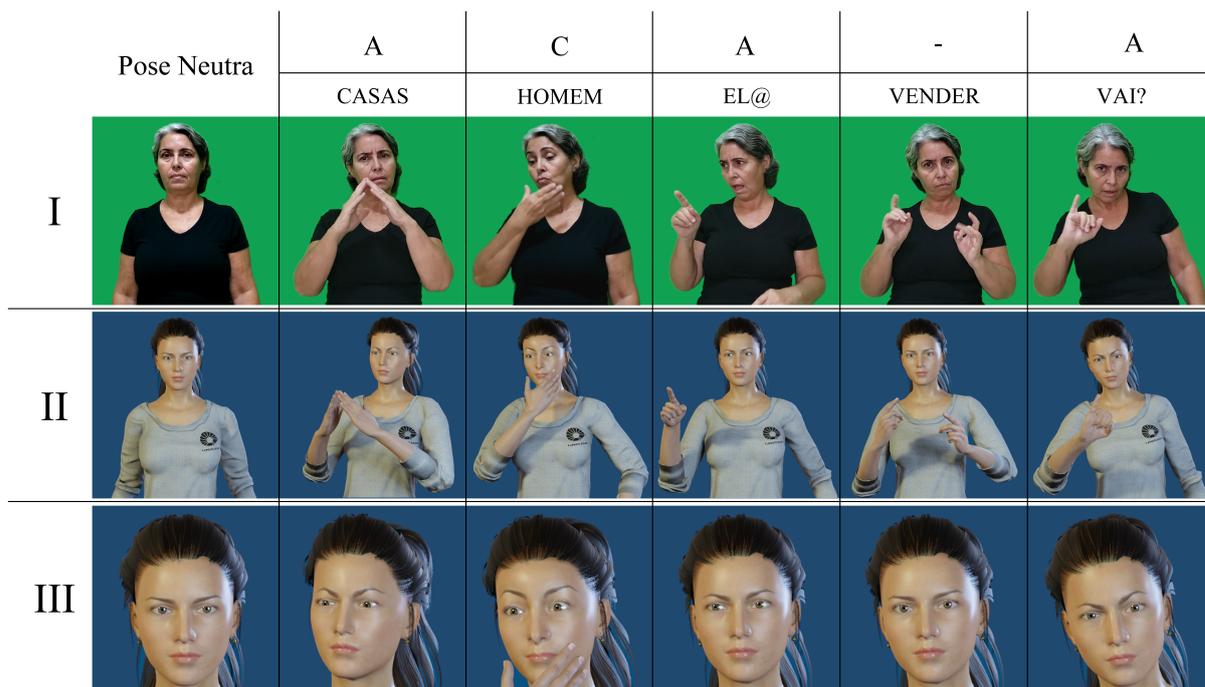


Figura 6.6 – Sinalização da sentença: "CASAS HOMEM EL@ VENDER VAI?". I) Sinalizações realizadas pela intérprete; II) Sinalizações realizadas pelo avatar; III) Destaque das expressões faciais executadas pelo avatar.

Abaixo são apresentados os dados de entrada inseridos e processados pelo sistema assim como as regras utilizadas. O formato dos dados processados e das regras são apresentados na Seção 6.1.2.

Dados de entrada:

HOMEM CONSTRUIR CASAS COMO?

T488 ,S540 ,T593 ,S630 ,T695 ,S740 ,T805 ,S835 ,T948 ,E1038

Dados de entrada processados:

['HOMEM', 540, 53, 52, 37], ['CONSTRUIR', 630, 65, 37, 45],
 ['CASAS', 740, 65, 45, 30], ['COMO?', 835, 113, 30, 90]

Regras utilizadas:

HOMEM C1

CONSTRUIR A1

COMO? B1A1

Animação

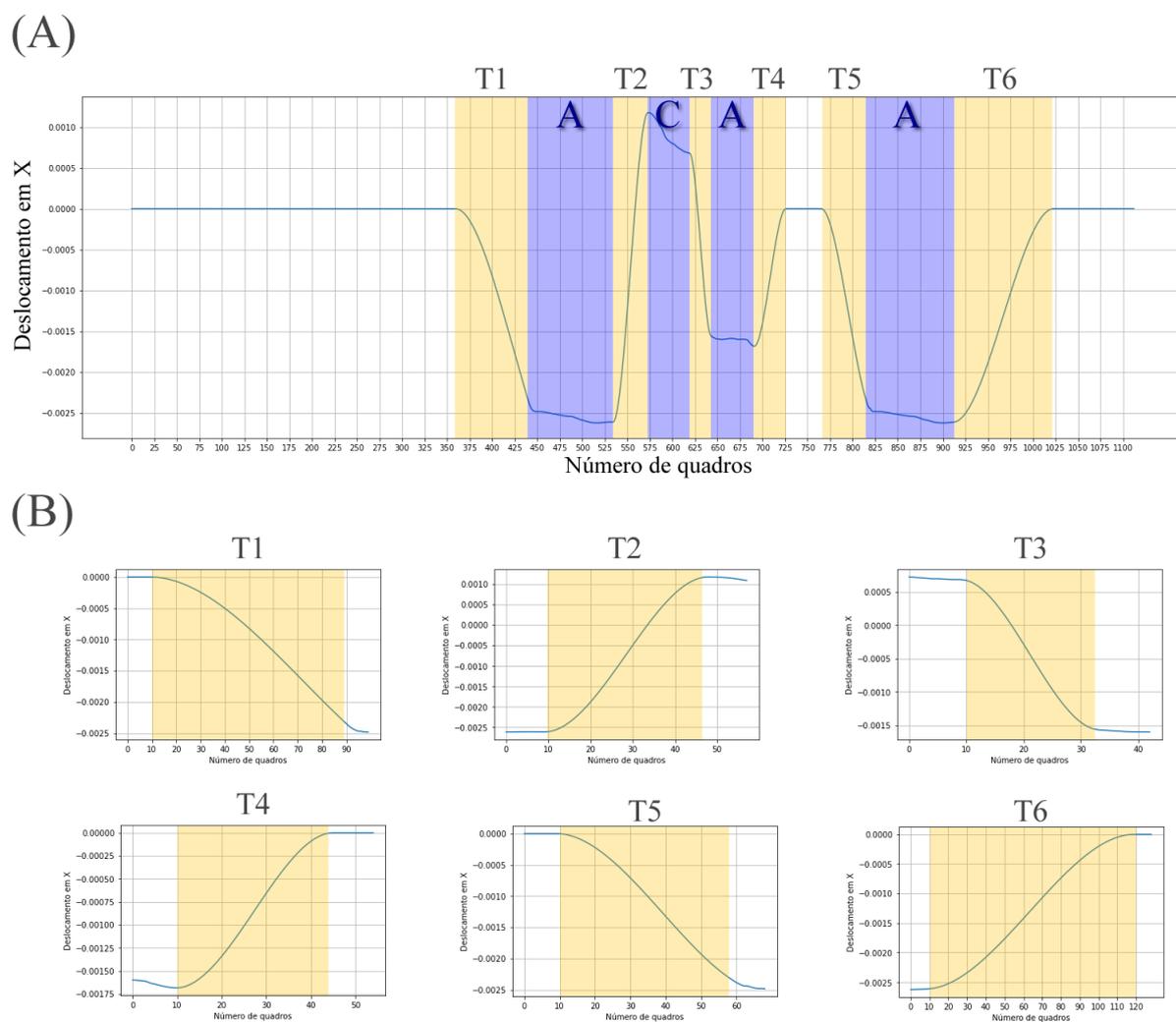


Figura 6.7 – Deslocamento do marcador mais ao centro da sobrancelha esquerda na sinalização da sentença: CASAS HOMEM EL@ VENDER VAI?. (A) Deslocamento durante toda a animação. (B) Deslocamento das transições criadas

A Figura 6.8 apresenta uma sequência de imagens obtidas durante a animação da sinalização de cada sinal da sentença pelo avatar e pela intérprete. Na imagem relacionada a sinalização do sinal HOMEM pode ser observada a expressão facial do tipo **C** (sobrancelhas levantadas), durante a sinalização dos sinais CONSTRUIR e CASAS a expressão facial do tipo **A** (sobrancelhas franzidas) e durante a sinalização do sinal COMO a expressão facial do tipo **B**. No banco de regras não existe a associação da expressão facial do tipo **A** ao sinal CASAS, contudo, a regra associada ao sinal COMO (B1A1) indica que durante a sinalização do sinal anterior a expressão facial do tipo **A** deve ser executada.

O vídeo da animação (Anim_1.mp4) está disponível no repositório do projeto no GitHub e pode ser acessado pelo endereço:

https://github.com/ackleywill/PilotoAnimacaoFacial/blob/main/videos/Anim_1.mp4.

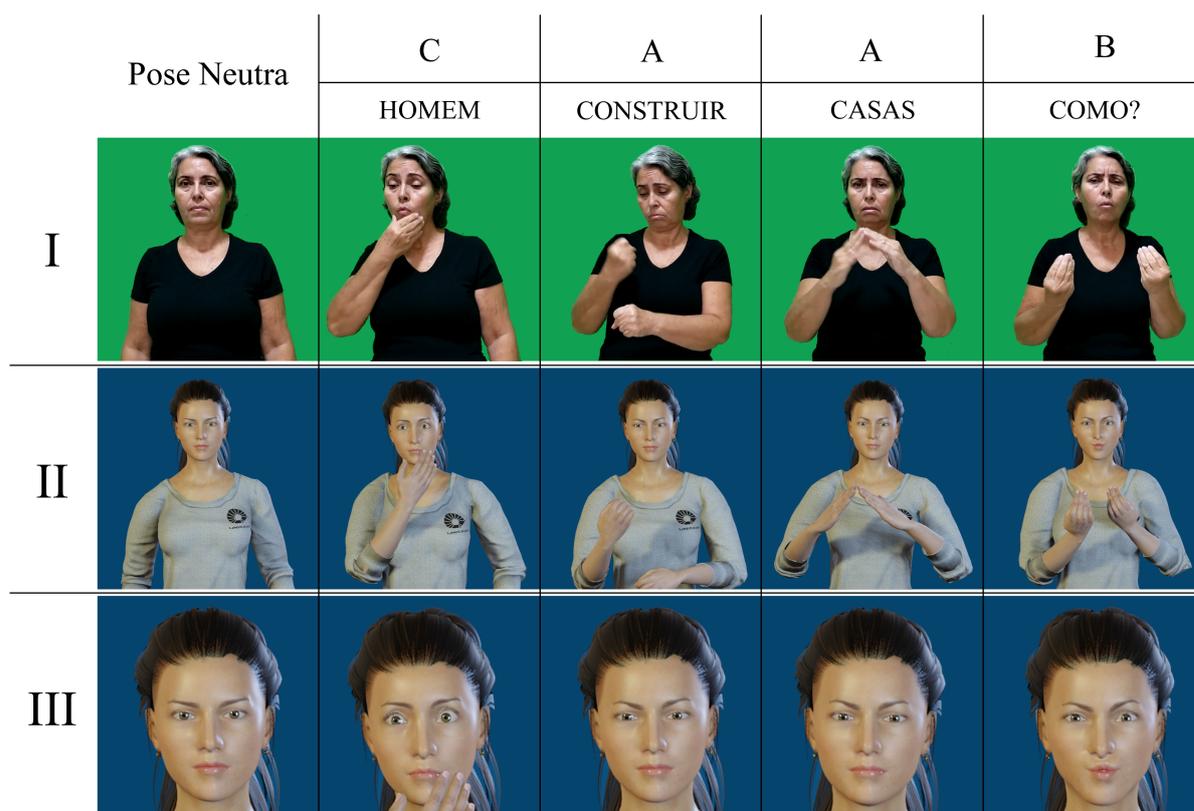


Figura 6.8 – Sinalização da sentença: "HOMEM CONSTRUIR CASAS COMO?". I) Sinalizações realizadas pela intérprete; II) Sinalizações realizadas pelo avatar; III) Destaque das expressões faciais executadas pelo avatar.

Gráficos da animação

O gráfico na parte superior da Figura 6.7 apresenta o deslocamento no eixo X do osso controlador mais ao centro da sobrancelha esquerda do rosto do avatar durante a animação. As faixas azuis marcam os trechos das execuções das expressões faciais recuperadas do repositório e as faixas bege os trechos das transições criadas pelo sistema. Os gráficos da parte inferior da figura apresentam, de forma destacada, o deslocamento do osso controlador em cada transição criada ao longo da animação.

6.4 Comentários finais

Neste capítulo foi apresentado o protótipo que implementa o processo de criação de animações faciais relatado na etapa de animação da metodologia descrita no Capítulo 5. As animações faciais criadas foram guiadas pelas regras obtidas através das análises dos dados do corpus na etapa de análise (Seção 5.2) e compostas por animações de expressões faciais obtidas na etapa de aquisição (Seção 5.3).

Na primeira etapa do sistema piloto apresentado são obtidos os dados de entrada referentes às informações de início dos sinais na animação corporal, nessa etapa

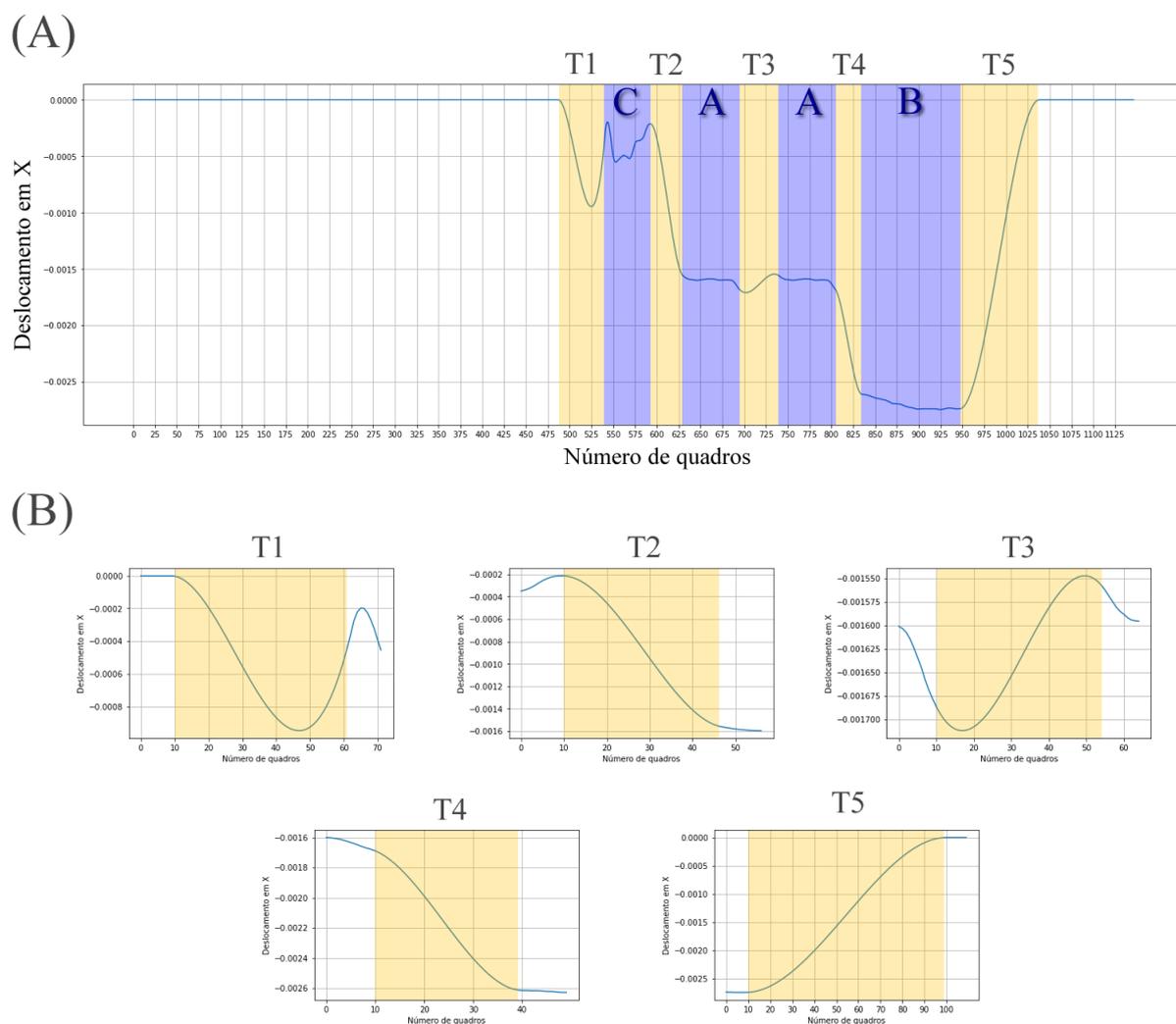


Figura 6.9 – Deslocamento do marcador mais ao centro da sobrançelha esquerda na sinalização da sentença: *HOMEM CONSTRUIR CASAS COMO?*. (A) Deslocamento durante toda a animação. (B) Deslocamento das transições criadas

é também verificado para cada sinal, se há correspondência com alguma regra presente no banco de regras. Essa etapa retorna ao sistema os instantes de início, duração e tipos de expressões faciais a serem aplicadas. Na etapa seguinte, são escolhidas no banco de expressões faciais, as animações que mais se adequam a cada expressão facial necessária de acordo com seu tipo e duração, por fim é criada a animação final com as animações individuais posicionadas e então inseridas as transições entre os sinais, criadas através de curvas paramétricas cúbicas de *Hermite*.

Foram analisadas e obtidas animações de expressões faciais associadas à diferentes tipos de sentenças interrogativas. Contudo, tanto o banco de regras quanto o banco de animações faciais podem ser ampliados para conter outros tipos de modalidades de sentenças e sinais proporcionando uma maior abrangência ao sistema proposto.

Os vídeos em que foram recortadas as imagens para os exemplos apresentados foram gerados através do protótipo apresentado e discutido acima e permitem uma avaliação

da metodologia proposta neste trabalho.

7 Conclusão

As expressões faciais em **Libras** fazem parte da sinalização e são utilizadas principalmente para marcação de construções sintáticas e diferenciação de itens lexicais (QUADROS; KARNOPP, 2004). Ainda que indiscutivelmente importantes, os sistemas de tradução português-Libras atuais possuem deficiência nas reproduções de expressões faciais em contextos importantes, como por exemplo, em modalidades interrogativas conforme observado no Capítulo 4.

O presente trabalho buscou contribuir com o desenvolvimento de uma metodologia para criação de animações de expressões faciais para complementar as sinalizações manuais em avatares realistas sinalizantes de **Libras**. As animações faciais foram obtidas através de capturas de movimento, e para isso, foram também desenvolvidas abordagens para a criação de *layouts* otimizados de marcadores faciais e para o mapeamento dos dados capturados ao rosto do avatar. Além disso, foi criado e anotado um corpus em **Libras** que auxiliou na elaboração de regras utilizadas para guiar o processo de criação das animações de expressões faciais. As contribuições e objetivos deste trabalho são revistos a seguir.

Com o intuito de criar animações faciais realistas reduzindo o trabalho de um animador, o uso de um sistema de capturas de movimento facial do tipo óptico baseado em marcadores foi utilizado. Neste contexto, foi definido um *layout* de marcadores padrão para se utilizar nas capturas. Na indústria e em áreas de pesquisa, *layout* empíricos são normalmente utilizados (LE; ZHU; DENG, 2013), porém carecem de fundamentação sólida, dificultando a generalização e padronização. Uma das contribuições deste trabalho, conforme relatado na Seção 3.2, foi o desenvolvimento de uma abordagem de criação de *layouts* otimizados de marcadores para capturas faciais. Este objetivo foi alcançado através de um processo de agrupamentos de marcadores com comportamentos similares observados em dados de capturas de movimento utilizando um *layout* denso de marcadores.

Uma outra contribuição, ainda no contexto de captura de movimento, foi a criação de uma abordagem para mapeamento dos dados capturados ao rosto do avatar (*retargeting* facial). Neste processo, conforme pode ser observado na Seção 3.3, foi buscado ajustar os dados capturados dadas as diferenças de proporção e formato entre o rosto do avatar e do ator. Além disso, essa abordagem buscou também alinhar os dados através de aplicação de mudança de espaço vetorial nos dados capturados tendo em vista o *layout* de ossos controladores no rosto do avatar.

A contribuição apresentada na abordagem de criação de *layouts* otimizados e a contribuição apresentada na abordagem de mapeamento dos dados capturados ao rosto do avatar servem para um contexto geral de animação facial, que engloba também, as

animações faciais na [Libras](#).

Para a contribuição relativa a criação e aplicação de expressões faciais na [Libras](#), que é o objetivo final desse trabalho, foi desenvolvida uma metodologia, apresentada no Capítulo 5, que possui três etapas principais: análise, aquisição e animação. Na etapa de análise foi criado e gravado um corpus de sinalizações em [Libras](#) com 270 sentenças nas modalidades afirmativa, interrogativa total e interrogativa parcial. As análises das anotações deste corpus apontou as expressões faciais relacionadas as modalidades sinalizadas e serviu como base para o desenvolvimento de regras para a criação da animação facial final. Na etapa de aquisição, foram realizadas capturas de movimento de sinalizações em [Libras](#) onde as expressões faciais definidas pela etapa de análise foram recortadas e catalogadas em uma base de dados. Por fim, na etapa de animação, a partir de dados de entrada, são verificadas as regras a serem aplicadas, selecionados e posicionados os trechos de animações necessários, criadas as transições através de curvas paramétricas e produzida então a animação facial final. A aplicação dessa metodologia proporciona a criação de animações faciais para acompanhar animações corporais em avatares sinalizantes. O desenvolvimento do protótipo foi apresentado no Capítulo 6.

Para o desenvolvimento futuro é planejado, como forma de verificação e validação, a utilização de outros operadores de correlação estatística, como o coeficiente de correlação de *Spearman* e lógica *Fuzzy*, para o cálculo das matrizes de correlação na abordagem de criação de *layouts* de marcadores otimizados faciais. Está prevista também, através de mais corpora contendo diferentes modalidades e sinais, a ampliação do banco de regras e do banco de expressões faciais, tanto em relação a quantidade de animações para cada tipo de expressão facial quanto em relação a variedade de tipos de expressões faciais. Além disso, está planejado um experimento de avaliação dos impactos das expressões faciais na inteligibilidade das sinalizações de [Libras](#) por um avatar assim como a avaliação da qualidade das expressões faciais e das regras de aplicação que envolvem os tipos e momentos de execução das expressões faciais na animação.

O presente trabalho apresenta uma linha de pesquisa com foco no desenvolvimento e aplicação de animações faciais em avatares sinalizantes de [Libras](#) e apresenta, além de contribuições nesta área em específico, contribuições na área de animação facial em geral através da criação de abordagens para auxiliar no processo de captura movimento e aplicação dos dados ao rosto de personagens tridimensionais.

Referências

- BADLER, N.; BYUN, M. Facemote: Qualitative parametric modifiers for facial animations. *Proc. ACM SIGGRAPH Symp. on Comput. Anim.*, 07 2002. Citado na página 25.
- BAKER-SHENK, C.; COKELY, D. *American Sign Language: A Teacher's Resource Text on Grammar and Culture*. [S.l.]: T.J. Publishers, 1980. (Original green books). ISBN 9780932666093. Citado na página 49.
- BATTISON, R. Phonological deletion in american sign language. *Sign Language Studies*, Project Muse, v. 1005, n. 1, p. 1–19, 1974. Citado na página 49.
- BLANZ, V.; VETTER, T. A morphable model for the synthesis of 3d faces. In: *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1999. (SIGGRAPH '99), p. 187–194. ISBN 0-201-48560-5. Citado na página 21.
- BLENDER. 2021. <<https://www.blender.org/>>. Accessed: 2021-09-24. Citado na página 123.
- BOUZID, Y.; JEMNI, M. An avatar based approach for automatically interpreting a sign language notation. In: *2013 IEEE 13th International Conference on Advanced Learning Technologies*. [S.l.: s.n.], 2013. p. 92–94. Citado na página 49.
- BRECAILO, S. de F. *Expressão Facial e Corporal na Comunicação em LIBRAS*. 2012. <https://www.academia.edu/5947147/Expressao_Facial_e_Corporal_na_comunicacao_em_LIBRAS>. Accessed: 2019-08-30. Citado na página 51.
- BROCK, H. et al. Learning three-dimensional skeleton data from sign language video. *ACM Trans. Intell. Syst. Technol.*, Association for Computing Machinery, New York, NY, USA, v. 11, n. 3, abr. 2020. ISSN 2157-6904. Citado na página 49.
- BUTTFIELD, A. A new approach to rapid image morphing for lip motion synthesis. In: OUDSHOORN, M. (Ed.). *26th Australasian Computer Science Conference*. Australia: Australian Computer Society, 2003. v. 16, p. 79–86. ISBN 14451336. A new approach to rapid image morphing for lip motion synthesis ; Conference date: 01-01-2003. Citado na página 24.
- CAPOVILLA, F.; RAPHAEL, W. *Dicionário enciclopédico ilustrado trilingüe da língua de sinais brasileira - Vol. 1 e Vol. 2*. Third. [S.l.]: Editora da Universidade de São Paulo, 2015. (Dicionário enciclopédico ilustrado trilingüe da língua de sinais brasileira). Citado na página 64.
- CHAE, W.; KIM, Y. Text-driven speech animation with emotion control. *KSII Transactions on Internet and Information Systems*, Korean Society for Internet Information (KSII), v. 14, n. 8, ago. 2020. Citado na página 24.
- CHUANG, E. S. *Analysis, Synthesis, and Retargeting of Facial Expressions*. Tese (Doutorado) — Stanford University, Stanford, CA, USA, 2004. AAI3128633. Citado na página 41.

CORRÊA, R.; MAGALHÃES, L. P.; MARTINO, J. M. D. Sistema de animação facial biomecânica. *Reverte (Indaiatuba)*, Indaiatuba, SP, Brasil, v. 1, p. 1–6, dez. 2008. Citado 3 vezes nas páginas 9, 27 e 28.

CORRÊA, R. *Animação facial por computador baseada em modelagem biomecânica*. Dissertação (Mestrado) — Universidade Estadual de Campinas, Campinas, SP, 10 2007. Citado 2 vezes nas páginas 23 e 26.

DENG, Z.; NEUMANN, U. (Ed.). *Data-Driven 3D Facial Animation*. [S.l.]: Springer London, 2007. Citado na página 21.

EZZAT, T.; POGGIO, T. Miketalk: a talking facial display based on morphing visemes. In: *Proceedings Computer Animation '98 (Cat. No.98EX169)*. [S.l.: s.n.], 1998. p. 96–102. Citado na página 24.

FELIPE, T. A. O discurso verbo-visual na língua brasileira de sinais - Libras. *Bakhtiniana: Revista de Estudos do Discurso*, scielo, v. 8, p. 67 – 89, 12 2013. ISSN 2176-4573. Citado 3 vezes nas páginas 10, 54 e 55.

FENG, W.-W.; KIM, B.-U.; YU, Y. Real-time data driven deformation using kernel canonical correlation analysis. *ACM Trans. Graph.*, Association for Computing Machinery, New York, NY, USA, v. 27, n. 3, p. 1–9, ago. 2008. ISSN 0730-0301. Citado na página 27.

FOLEY, J. D. et al. *Computer Graphics: Principles and Practice*. 2. ed. Reading, MA: Addison-Wesley, 1990. ISBN 978-0-201-12110-0. Citado na página 79.

FRIEDMAN, L. A. Phonological processes in the american sign language. *The First Annual Meeting Of The Berkeley Linguistics Society*, Linguistic Society of America, 1975. Citado na página 49.

GAO, H. et al. Marker based facial tracking application in communication disorder research. In: . [S.l.: s.n.], 2007. p. 244–247. Citado na página 29.

GIBET, S. Synthesis of sign language gestures. In: *Conference Companion on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 1994. (CHI '94), p. 311–312. ISBN 0897916514. Citado na página 49.

GITHUB. 2021. <<https://github.com/>>. Accessed: 2021-09-26. Citado na página 82.

GUENTER, B. A system for simulating human facial expression. In: MAGNENAT-THALMANN, N.; THALMANN, D. (Ed.). *State-of-the-art in Computer Animation*. Tokyo: Springer Japan, 1989. p. 191–202. ISBN 978-4-431-68293-6. Citado na página 21.

HANDTALK - tradutor de português-libras. 2021. <<https://www.handtalk.me/br>>. Accessed: 2021-03-15. Citado na página 53.

HERSH, M. A.; JOHNSON, M. A. *Assistive Technology for the Hearing-impaired, Deaf and Deafblind*. [S.l.: s.n.], 2006. ISBN 9781852338558. Citado na página 49.

IBGE - Instituto Brasileiro de Geografia e Estatística: Censo Demográfico 2010: Características gerais da população, religião e pessoas com deficiência. 2010. <<https://censo2010.ibge.gov.br/resultados.html>>. Citado na página 49.

- ISO. *Information technology - Coding of audio-visual objects - Part 2: Visual*. ISO 14496-2:2001(E). Geneva, Switzerland: Standard, 2001. Disponível em: <<https://www.iso.org/standard/36081.html>>. Citado 3 vezes nas páginas 9, 25 e 26.
- IYANIWURA, A. *Optical Motion Capture for Performance-based Facial Animation*. Dissertação (Mestrado) — Carleton University, Ottawa, Ontario Canada, 5 2008. Citado na página 22.
- KIM, J.; CHOI, M. G.; KIM, Y. J. Real-time muscle-based facial animation using shell elements and force decomposition. In: *Symposium on Interactive 3D Graphics and Games*. New York, NY, USA: Association for Computing Machinery, 2020. (I3D '20). ISBN 9781450375894. Citado na página 28.
- KIPP, M.; HELOIR, A.; NGUYEN, Q. Sign language avatars: Animation and comprehensibility. In: *Intelligent Virtual Agents*. [S.l.]: Springer Berlin Heidelberg, 2011. p. 113–126. Citado 2 vezes nas páginas 49 e 52.
- KITAGAWA, M.; WINDSOR, B. *MoCap for Artists: Workflow and Techniques for Motion Capture*. [S.l.]: Focal Press, 2008. ISBN 0240810007, 9780240810003. Citado 5 vezes nas páginas 22, 23, 29, 33 e 34.
- LE, B. H.; ZHU, M.; DENG, Z. Marker optimization for facial motion acquisition and deformation. v. 19, n. 11, p. 1859–1871, Nov. 2013. ISSN 1077-2626. Citado 4 vezes nas páginas 30, 32, 33 e 95.
- LEE, S.-Y. et al. Image morphing using deformation techniques. *The Journal of Visualization and Computer Animation*, v. 7, n. 1, p. 3–23, 1996. Citado na página 24.
- LEE, Y.; TERZOPOULOS, D.; WATERS, K. Realistic modeling for facial animation. In: *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA: Association for Computing Machinery, 1995. (SIGGRAPH '95), p. 55–62. ISBN 0897917014. Citado na página 27.
- LIN, A. J. *Three Dimensional Modeling and Animation of Facial Expressions*. Tese (Doutorado) — University of Kentucky, Lexington, Kentucky, 2011. Citado 3 vezes nas páginas 21, 24 e 25.
- LIU, S. et al. Ras: A data-driven rigidity-aware skinning model for 3d facial animation. *Computer Graphics Forum*, v. 39, n. 1, p. 581 – 594, 2020. ISSN 01677055. Citado na página 27.
- LOMBARDO, V. et al. An avatar-based interface for the italian sign language. In: *2011 International Conference on Complex, Intelligent, and Software Intensive Systems*. [S.l.: s.n.], 2011. p. 589–594. Citado na página 49.
- MAGNENAT-THALMANN, N.; PRIMEAU, E.; THALMANN, D. Abstract muscle action procedures for human face animation. *The Visual Computer*, v. 3, n. 5, p. 290–297, Mar 1988. ISSN 1432-2315. Citado na página 21.
- MARTINO, J. M. D. *Animação Facial Sincronizada com a Fala: Visemas Dependentes do Contexto Fonético para o Português do Brasil*. Tese (Doutorado) — Universidade Estadual de Campinas, Campinas, SP, 7 2005. Citado 6 vezes nas páginas 9, 22, 24, 26, 27 e 28.

- MARTINO, J. M. D. et al. Signing avatars: making education more inclusive. *Universal Access in the Information Society*, Springer Science and Business Media LLC, v. 16, n. 3, p. 793–808, nov. 2016. Citado 2 vezes nas páginas 49 e 86.
- MCCLEARY, L.; VIOTTI, E.; LEITE, T. de A. Descrição das línguas sinalizadas: a questão da transcrição dos dados. *Alfa Revista de Linguística*, v. 54, 2010. ISSN 1981-5794. Citado na página 53.
- MENACHE, A. *Understanding Motion Capture for Computer Animation, Second Edition*. 2nd. ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2010. ISBN 0123814960. Citado na página 29.
- MORI, M.; MACDORMAN, K. F.; KAGEKI, N. The uncanny valley [from the field]. *IEEE Robotics Automation Magazine*, v. 19, n. 2, p. 98–100, 2012. Citado 4 vezes nas páginas 9, 18, 21 e 22.
- PAIVA, F. A. dos S. *Tradução Automática de Português Brasileiro para Libras e Análise de Processos de Intensificação*. Tese (Doutorado) — Universidade Estadual de Campinas, Campinas, SP, 9 2019. Citado na página 62.
- Parke. Parameterized models for facial animation. *IEEE Computer Graphics and Applications*, v. 2, n. 9, p. 61–68, 1982. Citado na página 25.
- PARKE, F. I. Computer generated animation of faces. In: *Proceedings of the ACM Annual Conference - Volume 1*. New York, NY, USA: ACM, 1972. (ACM '72), p. 451–457. Citado 3 vezes nas páginas 18, 21 e 24.
- PARKE, F. I. *A parametric model for human faces*. Tese (Doutorado) — UTAH University, Salt Lake CityP, 12 1974. Citado na página 25.
- PARKE, F. I.; WATERS, K. *Computer Facial Animation*. Second. [S.l.]: AK Peters Ltd, 2008. ISBN 1568814488. Citado 4 vezes nas páginas 9, 18, 21 e 25.
- PARKE, F. I.; WATERS, K. *Computer Facial Animation*. Second. [S.l.]: AK Peters Ltd, 2008. ISBN 1568814488, 9781568814483. Citado na página 21.
- PATEL, B. D. et al. Es2isl: An advancement in speech to sign language translation using 3d avatar animator. In: *2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. [S.l.: s.n.], 2020. p. 1–5. Citado na página 49.
- PIGHIN, F. et al. Synthesizing realistic facial expressions from photographs. In: *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA: Association for Computing Machinery, 1998. (SIGGRAPH '98), p. 75–84. ISBN 0897919998. Citado na página 24.
- PIGHIN, F.; LEWIS, J. P. Facial motion retargeting. In: *ACM SIGGRAPH 2006 Courses*. New York, NY, USA: ACM, 2006. (SIGGRAPH '06). ISBN 1-59593-364-6. Citado na página 41.
- PLATT, S. M.; BADLER, N. I. Animating facial expressions. *SIGGRAPH Comput. Graph.*, Association for Computing Machinery, New York, NY, USA, v. 15, n. 3, p. 245–252, ago. 1981. ISSN 0097-8930. Citado na página 27.

- POTAMIAS, R. A. et al. *Learning to Generate Customized Dynamic 3D Facial Expressions*. 2020. Citado na página 21.
- QUADROS, R. M. de; KARNOPP, L. B. *Língua de Sinais Brasileira: Estudos Lingüísticos*. [S.l.]: Artmed Editora, 2004. ISBN 9788536311746. Citado 3 vezes nas páginas 18, 51 e 95.
- REILLY, J. How faces come to serve grammar: The development of nonmanual morphology in american sign language. In: *Advances in the Sign-Language Development of Deaf Children*. [S.l.]: Oxford University Press, 2005. p. 262–290. Citado 3 vezes nas páginas 10, 18 e 50.
- REVERDY, C.; GIBET, S.; LARBOULETTE, C. Optimal marker set for motion capture of dynamical facial expressions. In: *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*. New York, NY, USA: ACM, 2015. (MIG '15), p. 31–36. ISBN 978-1-4503-3991-9. Citado 2 vezes nas páginas 30 e 33.
- RIBERA, R. B. i. et al. Facial retargeting with automatic range of motion alignment. *ACM Trans. Graph.*, ACM, New York, NY, USA, v. 36, n. 4, p. 154:1–154:12, jul. 2017. ISSN 0730-0301. Citado na página 41.
- RUHLAND, K.; PRASAD, M.; MCDONNELL, R. Data-driven approach to synthesizing facial animation using motion capture. *IEEE Computer Graphics and Applications*, v. 37, n. 4, p. 30–41, 2017. Citado na página 27.
- RYBENÁ - tradutor de português-libras. 2021. <<https://portal.rybena.com.br/site-rybena/>>. Accessed: 2021-03-15. Citado na página 53.
- SILVA, E.; COSTA, P. Recognition of non-manual expressions in brazilian sign language. In: . [S.l.: s.n.], 2017. Citado na página 51.
- SILVA, E. P. da. *Facial expression recognition in Brazilian sign language using facial action coding system [recurso eletrônico] : Reconhecimento de expressões faciais na língua de sinais brasileira por meio do sistema de códigos de ação facial*. Tese (Doutorado) — Universidade Estadual de Campinas, Campinas, SP, 10 2020. Citado na página 51.
- SILVA, E. P. da et al. Facial action unit detection methodology with application in brazilian sign language recognition. *Pattern Analysis and Applications*, Springer Science and Business Media LLC, set. 2021. Disponível em: <<https://doi.org/10.1007/s10044-021-01024-5>>. Citado na página 51.
- SILVA, I. R. et al. [...] língua de sinais eu sei, mas o português é difícil”: reflexões sobre políticas linguísticas e de identidade no contexto da surdez no fomento de práticas de letramento diferenciadas. In: *UFBA*. Salvador: Congresso-Luso-Afro Brasileiro de Ciências Sociais, 2011. p. 1–16. Citado na página 49.
- SOBOTTA, J. et al. *Sobotta Atlas of Human Anatomy*. [S.l.]: Lippincott Williams & Wilkins, 2001. ISBN 9780781731737. Citado 4 vezes nas páginas 9, 23, 34 e 35.
- SOUZA, M. F. N. S. de et al. Principais dificuldades e obstáculos enfrentados pela comunidade surda no acesso à saúde: uma revisão integrativa de literatura. *Revista CEFAC*, FapUNIFESP (SciELO), v. 19, n. 3, p. 395–405, jun. 2017. Citado na página 49.

- STATCOUNTER global stats - Mobile Operating System Market Share Brazil. 2021. <<https://gs.statcounter.com/os-market-share/mobile/brazil>>. Accessed: 2021-03-15. Citado na página 52.
- STOKOE, W. *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf*. [S.l.]: University of Buffalo, 1960. v. 8. (Studies in Linguistics. Occasional Papers, v. 8). Citado na página 49.
- TANYA, F. A. *Libras em contexto: Curso básico, livro do estudante*. 8ª. ed. [S.l.]: Rio de Janeiro: WalPrint Gráfica e Editora, 2007. ISBN 85-99091-01-8. Citado na página 62.
- TERZOPOULOS, D.; WATERS, K. Physically-based facial modelling, analysis, and animation. *The Journal of Visualization and Computer Animation*, v. 1, n. 2, p. 73–80, 1990. Citado na página 27.
- VICON. *Vicon Cara User Guide*. 2015. <<https://docs.vicon.com/display/CD?preview=/83301483/83301486/ViconCaraUserGuide.pdf>>. Accessed: 2018-07-30. Citado 2 vezes nas páginas 9 e 31.
- VICON. *Vicon Cara*. 2018. <<https://www.vicon.com/products/camera-systems/cara-1>>. Accessed: 2018-04-16. Citado 2 vezes nas páginas 30 e 74.
- VICON. *Shogun 1.4 Documentation*. 2020. <<https://docs.vicon.com/display/Shogun14/>>. Accessed: 2020-07-26. Citado 3 vezes nas páginas 9, 29 e 30.
- VICON. *Vicon CaraPost*. 2021. <<https://www.vicon.com/software/carapost/>>. Accessed: 2021-12-14. Citado 2 vezes nas páginas 35 e 74.
- VIEIRA, M. C. et al. Análise de expressões não-manuais em avatares tradutores de língua portuguesa para libras. In: *Memórias del XIX Congresso Internacional Informática Educativa, TISE 2014*. [S.l.: s.n.], 2014. Accessed: 2019-09-22. Citado 3 vezes nas páginas 18, 49 e 52.
- VLIBRAS - Tradução de Português para Libras. 2021. <<https://www.vlibras.gov.br/>>. Accessed: 2021-03-15. Citado na página 53.
- WATERS, K. A muscle model for animation three-dimensional facial expression. *SIGGRAPH Comput. Graph.*, ACM, New York, NY, USA, v. 21, n. 4, p. 17–24, ago. 1987. ISSN 0097-8930. Citado na página 21.
- WILL, A. D.; MARTINO, J. M. D.; BEZERRA, J. An Optimized Marker Layout for 3D Facial Motion Capture. In: LIVESU, M.; PINTORE, G.; SIGNORONI, A. (Ed.). *Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference*. [S.l.]: The Eurographics Association, 2018. ISBN 978-3-03868-075-8. ISSN 2617-4855. Citado 2 vezes nas páginas 32 e 74.
- WILL, A. D.; MARTINO, J. M. D.; BEZERRA, J. Automatic mapping of facial mocap data to a realistic signing avatar. In: *Memórias X Congreso Iberoamericano de Tecnologías de Apoyo a la Discapacidad*. [S.l.: s.n.], 2019. p. 431–434. Citado 4 vezes nas páginas 9, 41, 42 e 75.

WITTENBURG, P. et al. ELAN: a professional framework for multimodality research. In: *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*. Genoa, Italy: European Language Resources Association (ELRA), 2006. Citado na página 64.

ZHANG, J.; LIAO, G.; LI, N. Combining active learning and local patch alignment for data-driven facial animation with fine-grained local detail. *Neurocomputing*, v. 398, p. 431–441, 2020. ISSN 0925-2312. Citado na página 27.

ZHANG, J. et al. Data-driven facial animation via semi-supervised local patch alignment. *Pattern Recognition*, v. 57, p. 1–20, 2016. ISSN 0031-3203. Citado na página 27.

Apêndices

APÊNDICE A – Sentenças do Corpus

	Modalidade	Português	Libras
1	Afirmativa	Ele vende casas	EL@ VENDER CASAS
	Int. Total	Ele vende casas?	CASAS EL@ VENDER VAI?
	Int. Parcial (Quando)	Quando ele vende casas?	EL@ PASSADO VENDER CASAS QUANDO?
	Int. Parcial (Onde)	Onde ele vende casas?	EL@ VENDER CASAS, ONDE?
	Int. Parcial (Como)	Como ele vende casas?	EL@ VENDER CASAS, COMO?
	Int. Parcial (Por que)	Por que ele vende casas?	EL@ VENDER CASAS, POR-QUE?
	Int. Parcial ((O) que)	O que ele vende?	EL@ VENDER O-QUE?
	Int. Parcial (Qual)	Quais casas ele vende?	EL@ VENDER, CASAS QUAL?
	Int. Parcial (Quanto)	Quantas casas ele vende?	EL@ VENDER QUANT@ CASAS
2	Afirmativa	Eu preciso estudar a matéria	EU PRECISAR ESTUDAR MATÉRIA
	Int. Total	Eu preciso estudar a matéria?	MATÉRIA EU PRECISAR ESTUDAR?
	Int. Parcial (Quando)	Quando eu preciso estudar a matéria?	EU FUTURO ESTUDAR MATÉRIA QUANDO?
	Int. Parcial (Onde)	Onde eu preciso estudar a matéria?	EU PRECISAR ESTUDAR MATÉRIA, ONDE?
	Int. Parcial (Como)	Como eu preciso estudar a matéria?	EU PRECISAR ESTUDAR MATÉRIA, COMO?
	Int. Parcial (Por que)	Porque eu preciso estudar a matéria?	EU PRECISAR ESTUDAR MATÉRIA, POR-QUE?
	Int. Parcial ((O) que)	O que eu preciso estudar?	EU PRECISAR ESTUDAR O-QUE?
	Int. Parcial (Qual)	Qual matéria eu preciso estudar?	EU PRECISAR ESTUDAR, MATERIA QUAL?
	Int. Parcial (Quanto)	Quantas matérias eu preciso estudar?	EU PRECISAR ESTUDAR, MATÉRIA QUANT@?
3	Afirmativa	Ela ganhará um prêmio	EL@ MULHER GANHAR PRÊMIO VAI.
	Int. Total	Ela ganhará um prêmio?	EL@ MULHER GANHAR PRÊMIO VAI?
	Int. Parcial (Quando)	Quando ela ganhará um prêmio?	EL@ MULHER FUTURO GANHAR PRÊMIO, QUANDO?
	Int. Parcial (Onde)	Onde ela ganhará um prêmio?	EL@ MULHER FUTURO GANHAR PRÊMIO, ONDE?
	Int. Parcial (Como)	Como ela ganhará um prêmio?	EL@ MULHER FUTURO GANHAR PRÊMIO, COMO?
	Int. Parcial (Por que)	Por que ela ganhará um prêmio?	EL@ MULHER FUTURO GANHAR PRÊMIO, POR-QUE?
	Int. Parcial ((O) que)	O que ela ganhará?	MULHER EL@ FUTURO GANHAR O-QUE?
	Int. Parcial (Qual)	Qual prêmio ela ganhará?	EL@ MULHER GANHAR, PRÊMIO FUTURO QUAL?
	Int. Parcial (Quanto)	Quantos prêmios ela ganhará?	EL@ MULHER GANHAR, PRÊMIO QUANT@?
4	Afirmativa	O menino quebrou a bicicleta	MENINO EL@ PASSADO QUEBRAR BICICLETA
	Int. Total	O menino quebrou a bicicleta?	MENINO EL@, PASSADO BICICLETA QUEBRAR?
	Int. Parcial (Quando)	Quando o menino quebrou a bicicleta?	MENINO EL@ PASSADO BICICLETA QUEBRAR, QUANDO?
	Int. Parcial (Onde)	Onde o menino quebrou a bicicleta?	MENINO EL@ PASSADO QUEBRAR BICICLETA, ONDE?
	Int. Parcial (Como)	Como o menino quebrou a bicicleta?	MENINO EL@ PASSADO BICICLETA QUEBRAR, COMO?
	Int. Parcial (Por que)	Por que o menino quebrou a bicicleta?	MENINO EL@ PASSADO QUEBRAR BICICLETA, POR-QUE?
	Int. Parcial ((O) que)	O que o menino quebrou?	MENINO EL@ PASSADO QUEBRAR O-QUE?
	Int. Parcial (Qual)	Qual bicicleta o menino quebrou?	MENINO EL@ PASSADO QUEBRAR, BICICLETA QUAL?
	Int. Parcial (Quanto)	Quantas bicicletas o menino quebrou?	MENINO EL@ PASSADO QUEBRAR, BICICLETA QUANT@?
5	Afirmativa	Jonas pesca carangueijo	HOMEM NOME J-O-N-A-S EL@ PESCAR CARANGUEIJO
	Int. Total	Jonas pesca carangueijo?	HOMEM NOME J-O-N-A-S EL@ PESCAR CARANGUEIJO VAI?
	Int. Parcial (Quando)	Quando Jonas pesca carangueijo?	HOMEM NOME J-O-N-A-S EL@ PESCAR CARANGUEIJO, QUANDO (PASSADO)?
	Int. Parcial (Onde)	Onde Jonas pesca carangueijo?	HOMEM NOME J-O-N-A-S EL@ PESCAR CARANGUEIJO, ONDE?
	Int. Parcial (Como)	Como Jonas pesca carangueijo?	HOMEM NOME J-O-N-A-S EL@ PESCAR CARANGUEIJO, COMO?
	Int. Parcial (Por que)	Por que Jonas pesca carangueijo?	HOMEM NOME J-O-N-A-S EL@ PESCAR CARANGUEIJO, POR-QUE?
	Int. Parcial ((O) que)	O que Jonas pesca?	HOMEM NOME J-O-N-A-S EL@ PESCAR O-QUE?
	Int. Parcial (Qual)	Qual carangueijo Jonas pesca?	HOMEM NOME J-O-N-A-S EL@ PESCAR, CARANGUEIJO QUAL?
	Int. Parcial (Quanto)	Quantos carangueijos Jonas pesca?	HOMEM NOME J-O-N-A-S EL@ PESCAR, CARANGUEIJO QUANT@?
6	Afirmativa	O homem cortará a árvore	HOMEM FUTURO ÁRVORE-CORTAR VAI
	Int. Total	O homem cortará a árvore?	HOMEM FUTURO ÁRVORE-CORTAR VAI?
	Int. Parcial (Quando)	Quando o homem cortará a árvore?	HOMEM FUTURO ÁRVORE-CORTAR, QUANDO?
	Int. Parcial (Onde)	Onde o homem cortará a árvore?	HOMEM FUTURO ÁRVORE-CORTAR, ONDE?
	Int. Parcial (Como)	Como o homem cortará a árvore?	HOMEM FUTURO ÁRVORE-CORTAR, COMO?
	Int. Parcial (Por que)	Por que o homem cortará a árvore?	HOMEM FUTURO ÁRVORE-CORTAR, POR-QUE?
	Int. Parcial ((O) que)	O que o homem cortará?	HOMEM FUTURO CORTAR O-QUE?
	Int. Parcial (Qual)	Qual árvore o homem cortará?	HOMEM FUTURO, ÁRVORE-CORTAR QUAL?
	Int. Parcial (Quanto)	Quantas árvores o homem cortará?	HOMEM FUTURO ÁRVORE-CORTAR QUANT@?

7	Afirmativa	O homem constrói casas	HOMEM CONSTRUIR CASAS
	Int. Total	O homem constrói casas?	HOMEM CASAS CONSTRUIR VAI?
	Int. Parcial (Quando)	Quando o homem constrói casas?	HOMEM CONSTRUIR CASAS FUTURO QUANDO?
	Int. Parcial (Onde)	Onde o homem constrói casas?	HOMEM CONSTRUIR CASAS, ONDE?
	Int. Parcial (Como)	Como o homem constrói casas?	HOMEM CONSTRUIR CASAS, COMO?
	Int. Parcial (Por que)	Por que o homem constrói casas?	HOMEM CONSTRUIR CASAS, POR-QUE?
	Int. Parcial ((O) que)	O que o homem constrói?	HOMEM CONSTRUIR O-QUE?
	Int. Parcial (Qual)	Quais casas o homem constrói?	HOMEM CONSTRUIR, CASAS QUAL?
Int. Parcial (Quanto)	Quantas casas o homem constrói?	HOMEM CONSTRUIR, CASAS QUANT@?	
8	Afirmativa	O menino gosta de comer batata	MENINO EL@ GOSTAR COMER BATATA
	Int. Total	O menino gosta de comer batata?	MENINO EL@ GOSTAR COMER BATATA?
	Int. Parcial (Quando)	Quando o menino gosta de comer batata?	MENINO EL@ GOSTAR COMER BATATA, PASSADO QUANDO?
	Int. Parcial (Onde)	Onde menino gosta de comer batata?	MENINO EL@ GOSTAR COMER BATATA, ONDE?
	Int. Parcial (Como)	Como o menino gosta de comer batata?	MENINO EL@ GOSTAR COMER BATATA, COMO?
	Int. Parcial (Por que)	Por que menino gosta de comer batata?	MENINO EL@ GOSTAR COMER BATATA, POR-QUE?
	Int. Parcial ((O) que)	O que o menino gosta de comer?	MENINO EL@ GOSTAR COMER O-QUE?
	Int. Parcial (Qual)	Qual batata o menino gosta de comer?	MENINO EL@ GOSTAR COMER, BATATA QUAL?
Int. Parcial (Quanto)	Quantas batatas o menino gosta de comer?	MENINO EL@ GOSTAR COMER, BATATA QUANT@?	
9	Afirmativa	Você carregou a mala	VOCÊ PASSADO CARREGAR-MALA
	Int. Total	Você carregou a mala?	VOCÊ PASSADO CARREGAR-MALA?
	Int. Parcial (Quando)	Quando você carregou a mala?	VOCÊ PASSADO CARREGAR-MALA, QUANDO?
	Int. Parcial (Onde)	Onde você carregou a mala?	VOCÊ PASSADO CARREGAR-MALA, ONDE?
	Int. Parcial (Como)	Como você carregou a mala?	VOCÊ PASSADO CARREGAR-MALA, COMO?
	Int. Parcial (Por que)	Por que você carregou a mala?	VOCÊ PASSADO CARREGAR-MALA, POR-QUE?
	Int. Parcial ((O) que)	O que você carregou?	VOCÊ PASSADO CARREGAR O-QUE?
	Int. Parcial (Qual)	Qual mala você carregou?	VOCÊ PASSADO CARREGAR-MALA QUAL?
Int. Parcial (Quanto)	Quantas malas você carregou?	VOCÊ PASSADO CARREGAR-MALA QUANT@?	
10	Afirmativa	O rapaz ganhará dinheiro por causa do trabalho	HOMEM JOVEM FUTURO GANHAR DINHEIRO POR-CAUSA TRABALHAR
	Int. Total	O rapaz ganhará dinheiro por causa do trabalho?	HOMEM JOVEM FUTURO GANHAR DINHEIRO POR-CAUSA TRABALHAR VAI?
	Int. Parcial (Quando)	Quando o rapaz ganhará dinheiro por causa do trabalho?	HOMEM JOVEM FUTURO GANHAR DINHEIRO POR-CAUSA TRABALHAR, QUANDO?
	Int. Parcial (Onde)	Onde o rapaz ganhará dinheiro por causa do trabalho?	HOMEM JOVEM FUTURO GANHAR DINHEIRO POR-CAUSA TRABALHAR, ONDE?
	Int. Parcial (Como)	Como o rapaz ganhará dinheiro por causa do trabalho?	HOMEM JOVEM FUTURO GANHAR DINHEIRO POR-CAUSA TRABALHAR, COMO?
	Int. Parcial (Por que)	Por que o rapaz ganhará dinheiro por causa do trabalho?	HOMEM JOVEM FUTURO GANHAR DINHEIRO POR-CAUSA TRABALHAR, POR-QUE?
	Int. Parcial ((O) que)	O que o rapaz ganhará por causa do trabalho?	HOMEM JOVEM FUTURO GANHAR DINHEIRO POR-CAUSA TRABALHAR
	Int. Parcial (Qual)	Qual dinheiro o rapaz ganhará por causa do trabalho?	HOMEM JOVEM FUTURO GANHAR POR-CAUSA TRABALHAR, DINHEIRO QUAL?
Int. Parcial (Quanto)	Quanto dinheiro o rapaz ganhará por causa do trabalho?	HOMEM JOVEM FUTURO GANHAR POR-CAUSA TRABALHAR, DINHEIRO QUANT@?	
11	Afirmativa	Você gostou de comer torta	VOCÊ PASSADO GOSTAR T-O-R-T-A COMER
	Int. Total	Você gostou de comer torta?	VOCÊ PASSADO GOSTAR COMER T-O-R-T-A?
	Int. Parcial (Quando)	Quando você gostou de comer torta?	VOCÊ PASSADO GOSTAR COMER T-O-R-T-A, QUANDO?
	Int. Parcial (Onde)	Onde você gostou de comer torta?	VOCÊ PASSADO GOSTAR COMER T-O-R-T-A, ONDE?
	Int. Parcial (Como)	Como você gostou de comer torta?	VOCÊ PASSADO GOSTAR COMER T-O-R-T-A, COMO?
	Int. Parcial (Por que)	Por que você gostou de comer torta?	VOCÊ PASSADO GOSTAR COMER T-O-R-T-A, POR-QUE?
	Int. Parcial ((O) que)	O que você gostou de comer?	VOCÊ PASSADO GOSTAR COMER O-QUE?
	Int. Parcial (Qual)	Qual torta você gostou de comer?	VOCÊ PASSADO GOSTAR COMER, T-O-R-T-A QUAL?
Int. Parcial (Quanto)	Quantas torta você gostou de comer?	VOCÊ PASSADO GOSTAR COMER, T-O-R-T-A QUANT@?	
12	Afirmativa	A mulher conserta o carro	MULHER CARRO CONSERTAR
	Int. Total	A mulher conserta o carro?	MULHER CARRO CONSERTAR VAI?
	Int. Parcial (Quando)	Quando a mulher conserta o carro?	MULHER CARRO CONSERTAR, QUANDO?
	Int. Parcial (Onde)	Onde a mulher conserta o carro?	MULHER CARRO CONSERTAR, ONDE?
	Int. Parcial (Como)	Como a mulher conserta o carro?	MULHER CARRO CONSERTAR, COMO?
	Int. Parcial (Por que)	Por que a mulher conserta o carro?	MULHER CARRO CONSERTAR, POR-QUE?
	Int. Parcial ((O) que)	O que a mulher conserta?	MULHER CONSERTAR O-QUE?
	Int. Parcial (Qual)	Qual carro a mulher conserta?	MULHER CONSERTAR, CARRO QUAL?
Int. Parcial (Quanto)	Quantos carros a mulher conserta?	MULHER CONSERTAR, CARRO QUANT@?	
13	Afirmativa	O menino conheceu o castelo	MENINO EL@ PASSADO CONHECER CASTELO
	Int. Total	O menino conheceu o castelo?	MENINO EL@ PASSADO CONHECER CASTELO?
	Int. Parcial (Quando)	Quando o menino conheceu o castelo?	MENINO EL@ PASSADO CONHECER CASTELO, QUANDO?
	Int. Parcial (Onde)	Onde o menino conheceu o castelo?	MENINO EL@ PASSADO CONHECER CASTELO, ONDE?
	Int. Parcial (Como)	Como o menino conheceu o castelo?	MENINO EL@ PASSADO CONHECER CASTELO, COMO?
	Int. Parcial (Por que)	Por que o menino conheceu o castelo?	MENINO EL@ PASSADO CONHECER CASTELO, POR-QUE?
	Int. Parcial ((O) que)	O que o menino conheceu?	MENINO EL@ PASSADO CONHECER O-QUE?
	Int. Parcial (Qual)	Qual castelo o menino conheceu?	MENINO EL@ PASSADO CONHECER, CASTELO QUAL?
Int. Parcial (Quanto)	Quantos castelos o menino conheceu?	MENINO EL@ PASSADO CONHECER, CASTELO QUANT@?	
14	Afirmativa	Ana lerá o livro	A-N-A MULHER EL@ FUTURO LIVRO-LER
	Int. Total	Ana lerá o livro?	A-N-A MULHER EL@ FUTURO LER LIVRO?
	Int. Parcial (Quando)	Quando Ana lerá o livro?	A-N-A MULHER EL@ FUTURO LER LIVRO, QUANDO?
	Int. Parcial (Onde)	Onde Ana lerá o livro?	A-N-A MULHER EL@ FUTURO LER LIVRO, ONDE?
	Int. Parcial (Como)	Como Ana lerá o livro?	A-N-A MULHER EL@ FUTURO LER LIVRO, COMO?
	Int. Parcial (Por que)	Por que Ana lerá o livro?	A-N-A MULHER EL@ FUTURO LER LIVRO, POR-QUE?
	Int. Parcial ((O) que)	O que Ana lerá?	A-N-A MULHER EL@ FUTURO LER O-QUE?
	Int. Parcial (Qual)	Qual livro Ana lerá?	A-N-A MULHER EL@ FUTURO LER LIVRO QUAL?
Int. Parcial (Quanto)	Quantos livros Ana lerá?	A-N-A MULHER EL@ FUTURO LER LIVRO QUANT@?	

15	Afirmativa	A mulher conseguiu terminar o trabalho	MULHER PASSADO CONSEGUIR TRABALHO TERMINAR
	Int. Total	A mulher conseguiu terminar o trabalho?	PASSADO MULHER CONSEGUIR TERMINAR TRABALHO?
	Int. Parcial (Quando)	Quando a mulher conseguiu terminar o trabalho?	MULHER PASSADO CONSEGUIR TRABALHO TERMINAR, QUANDO?
	Int. Parcial (Onde)	Onde a mulher conseguiu terminar o trabalho?	MULHER PASSADO CONSEGUIR TRABALHO TERMINAR, ONDE?
	Int. Parcial (Como)	Como a mulher conseguiu terminar o trabalho?	MULHER PASSADO CONSEGUIR TRABALHO TERMINAR, COMO?
	Int. Parcial (Por que)	Por que a mulher conseguiu terminar o trabalho?	MULHER PASSADO CONSEGUIR TRABALHO TERMINAR, POR-QUE?
	Int. Parcial ((O) que)	O que a mulher conseguiu terminar?	MULHER PASSADO CONSEGUIR TERMINAR O-QUE?
	Int. Parcial (Qual)	Qual trabalho a mulher conseguiu terminar?	MULHER PASSADO CONSEGUIR TERMINAR, TRABALHO QUAL?
Int. Parcial (Quanto)	Quantos trabalhos a mulher conseguiu terminar?	MULHER PASSADO CONSEGUIR TERMINAR, TRABALHO QUANT@?	
16	Afirmativa	A mulher cozinha feijão	MULHER EL@ FAZER COZINHAR FEIJÃO
	Int. Total	A mulher cozinha feijão?	MULHER EL@ FAZER COZINHAR FEIJÃO VAI?
	Int. Parcial (Quando)	Quando a mulher cozinha feijão?	MULHER EL@ FAZER COZINHAR FEIJÃO, PASSADO QUANDO?
	Int. Parcial (Onde)	Onde a mulher cozinha feijão?	MULHER EL@ FAZER COZINHAR FEIJÃO, ONDE?
	Int. Parcial (Como)	Como a mulher cozinha feijão?	MULHER EL@ FAZER COZINHAR FEIJÃO, COMO?
	Int. Parcial (Por que)	Por que a mulher cozinha feijão?	MULHER EL@ FAZER COZINHAR FEIJÃO, POR-QUE?
	Int. Parcial ((O) que)	O que a mulher cozinha?	MULHER EL@ FAZER COZINHAR O-QUE?
	Int. Parcial (Qual)	Qual feijão a mulher cozinha?	MULHER EL@ FAZER COZINHAR, FEIJÃO QUAL?
Int. Parcial (Quanto)	Quanto feijão a mulher cozinha?	MULHER EL@ FAZER COZINHAR, FEIJÃO QUANT@?	
17	Afirmativa	Ele digita a carta	HOMEM EL@ DIGITAR CARTA
	Int. Total	Ele digita a carta?	HOMEM EL@ DIGITAR A CARTA VAI?
	Int. Parcial (Quando)	Quando ele digita a carta?	HOMEM EL@ DIGITAR CARTA, QUANDO FUTURO?
	Int. Parcial (Onde)	Onde ele digita a carta?	HOMEM EL@ DIGITAR CARTA, ONDE?
	Int. Parcial (Como)	Como ele digita a carta?	HOMEM EL@ DIGITAR CARTA, COMO?
	Int. Parcial (Por que)	Por que ele digita a carta?	HOMEM EL@ DIGITAR CARTA, POR-QUE?
	Int. Parcial ((O) que)	O que ele digita?	HOMEM EL@ DIGITAR O-QUE?
	Int. Parcial (Qual)	Qual carta ele digita?	HOMEM EL@ DIGITAR, CARTA QUAL?
Int. Parcial (Quanto)	Quantas cartas ele digita?	HOMEM EL@ DIGITAR, CARTA QUANT@?	
18	Afirmativa	Ela escreverá um poema novo	MULHER EL@ FUTURO ESCREVER POEMA NOVO
	Int. Total	Ela escreverá um poema novo?	FUTURO MULHER EL@ ESCREVER POEMA NOVO VAI?
	Int. Parcial (Quando)	Quando ela escreverá um poema novo?	MULHER EL@ FUTURO ESCREVER POEMA NOVO, QUANDO?
	Int. Parcial (Onde)	Onde ela escreverá um poema novo?	MULHER EL@ FUTURO ESCREVER POEMA NOVO, ONDE?
	Int. Parcial (Como)	Como ela escreverá um poema novo?	MULHER EL@ FUTURO ESCREVER POEMA NOVO, COMO?
	Int. Parcial (Por que)	Por que ela escreverá um poema novo?	MULHER EL@ FUTURO ESCREVER POEMA NOVO, POR-QUE?
	Int. Parcial ((O) que)	O que ela escreverá?	MULHER EL@ FUTURO ESCREVER O-QUE?
	Int. Parcial (Qual)	Qual poema ela escreverá?	MULHER EL@ FUTURO ESCREVER, POEMA QUAL?
Int. Parcial (Quanto)	Quantos poemas ela escreverá?	MULHER EL@ FUTURO ESCREVER, POEMA QUANT@?	
19	Afirmativa	Ele roubou um carro amarelo	HOMEM EL@ PASSADO ROUBAR CARRO AMARELO
	Int. Total	Ele roubou um carro amarelo?	HOMEM EL@ PASSADO ROUBAR CARRO AMARELO?
	Int. Parcial (Quando)	Quando ele roubou um carro amarelo?	HOMEM EL@ PASSADO ROUBAR CARRO AMARELO, QUANDO?
	Int. Parcial (Onde)	Onde ele roubou um carro amarelo?	HOMEM EL@ PASSADO ROUBAR CARRO AMARELO, ONDE?
	Int. Parcial (Como)	Como ele roubou um carro amarelo?	HOMEM EL@ PASSADO ROUBAR CARRO AMARELO, COMO?
	Int. Parcial (Por que)	Por que ele roubou um carro amarelo?	HOMEM EL@ PASSADO ROUBAR CARRO AMARELO, POR-QUE?
	Int. Parcial ((O) que)	O que ele roubou?	HOMEM EL@ PASSADO ROUBAR O-QUE?
	Int. Parcial (Qual)	Qual carro amarelo ele roubou?	HOMEM EL@ PASSADO ROUBAR, CARRO AMARELO QUAL?
Int. Parcial (Quanto)	Quantos carros amarelos ele roubou?	HOMEM EL@ PASSADO ROUBAR, CARRO AMARELO QUANT@?	
20	Afirmativa	O homem venderá o carro	HOMEM FUTURO VENDER CARRO
	Int. Total	O homem venderá o carro?	FUTURO HOMEM VENDER CARRO VAI?
	Int. Parcial (Quando)	Quando o homem venderá o carro?	HOMEM FUTURO VENDER CARRO, QUANDO?
	Int. Parcial (Onde)	Onde o homem venderá o carro?	HOMEM FUTURO VENDER CARRO, ONDE?
	Int. Parcial (Como)	Como o homem venderá o carro?	HOMEM FUTURO VENDER CARRO, COMO?
	Int. Parcial (Por que)	Por que o homem venderá o carro?	HOMEM FUTURO VENDER CARRO, POR-QUE?
	Int. Parcial ((O) que)	O que o homem venderá?	HOMEM FUTURO VENDER O-QUE?
	Int. Parcial (Qual)	Qual carro o homem venderá?	HOMEM FUTURO VENDER, CARRO QUAL?
Int. Parcial (Quanto)	Quantos carros o homem venderá?	HOMEM FUTURO VENDER, CARRO QUANT@?	
21	Afirmativa	A mulher perdeu a bolsa no metrô	MULHER PASSADO PERDER BOLSA METRÔ
	Int. Total	A mulher perdeu a bolsa no metrô?	PASSADO MULHER PERDER BOLSA METRÔ?
	Int. Parcial (Quando)	Quando a mulher perdeu a bolsa no metrô?	MULHER PASSADO PERDER BOLSA METRÔ, QUANDO?
	Int. Parcial (Onde)	Onde a mulher perdeu a bolsa no metrô?	MULHER PASSADO PERDER BOLSA METRÔ, ONDE?
	Int. Parcial (Como)	Como a mulher perdeu a bolsa no metrô?	MULHER PASSADO PERDER BOLSA METRÔ, COMO?
	Int. Parcial (Por que)	Por que a mulher perdeu a bolsa no metrô?	MULHER PASSADO PERDER BOLSA METRÔ, POR-QUE?
	Int. Parcial ((O) que)	O que a mulher perdeu no metrô?	MULHER PASSADO PERDER METRÔ O-QUE?
	Int. Parcial (Qual)	Qual bolsa a mulher perdeu no metrô?	MULHER PASSADO PERDER METRÔ, BOLSA QUAL?
Int. Parcial (Quanto)	Quantas bolsas a mulher perdeu no metrô?	MULHER PASSADO PERDER METRÔ, BOLSA QUANT@?	
22	Afirmativa	Você tinha medo do escuro	VOCÊ PASSADO MEDO ESCURO
	Int. Total	Você tinha medo do escuro?	VOCÊ PASSADO MEDO ESCURO?
	Int. Parcial (Quando)	Quando você tinha medo do escuro?	VOCÊ PASSADO MEDO ESCURO, QUANDO?
	Int. Parcial (Onde)	Onde você tinha medo do escuro?	VOCÊ PASSADO MEDO ESCURO, ONDE?
	Int. Parcial (Como)	Como você tinha medo do escuro?	VOCÊ PASSADO MEDO ESCURO, COMO?
	Int. Parcial (Por que)	Por que você tinha medo do escuro?	VOCÊ PASSADO MEDO ESCURO, POR-QUE?
	Int. Parcial ((O) que)	O que você tinha medo? (Você tinha medo de quê?)	VOCÊ PASSADO MEDO O-QUE?
	Int. Parcial (Qual)	Qual medo você tinha?	VOCÊ PASSADO, MEDO QUAL?
Int. Parcial (Quanto)	Quanto medo você tinha?	VOCÊ PASSADO, MEDO QUANT@?	

23	Afirmativa	O professor quererá o sapato	PROFESSOR FUTURO QUERER SAPATO
	Int. Total	O professor quererá o sapato?	PROFESSOR FUTURO QUERER SAPATO?
	Int. Parcial (Quando)	Quando o professor quererá o sapato?	PROFESSOR FUTURO QUERER SAPATO, QUANDO?
	Int. Parcial (Onde)	Onde o professor quererá o sapato?	PROFESSOR FUTURO QUERER SAPATO, ONDE?
	Int. Parcial (Como)	Como o professor quererá o sapato?	PROFESSOR FUTURO QUERER SAPATO, COMO?
	Int. Parcial (Por que)	Por que o professor quererá o sapato?	PROFESSOR FUTURO QUERER SAPATO, POR-QUE?
	Int. Parcial ((O) que)	O que o professor quererá?	PROFESSOR FUTURO QUERER O-QUE?
	Int. Parcial (Qual)	Qual sapato o professor quererá?	PROFESSOR FUTURO QUERER, SAPATO QUAL?
Int. Parcial (Quanto)	Quantos sapatos o professor quererá?	PROFESSOR FUTURO QUERER, SAPATO QUANT@?	
24	Afirmativa	O homem enche o copo	HOMEM ENCHER COPO
	Int. Total	O homem não enche o copo?	HOMEM ENCHER COPO VAI?
	Int. Parcial (Quando)	Quando o homem enche o copo?	HOMEM ENCHER COPO, QUANDO (PASSADO)?
	Int. Parcial (Onde)	Onde o homem enche o copo?	HOMEM ENCHER COPO, ONDE?
	Int. Parcial (Como)	Como o homem enche o copo?	HOMEM ENCHER COPO, COMO?
	Int. Parcial (Por que)	Por que o homem enche o copo?	HOMEM ENCHER COPO, POR-QUE?
	Int. Parcial ((O) que)	O que o homem enche?	HOMEM ENCHER O-QUE?
	Int. Parcial (Qual)	Qual copo o homem enche?	HOMEM ENCHER, COPO QUAL?
Int. Parcial (Quanto)	Quantos copos o homem enche?	HOMEM ENCHER, COPO QUANT@?	
25	Afirmativa	O chefe plantou frutas	CHEFE PASSADO PLANTAR FRUTAS
	Int. Total	O chefe plantou frutas?	CHEFE PASSADO PLANTAR FRUTAS?
	Int. Parcial (Quando)	Quando o chefe plantou frutas?	CHEFE PASSADO PLANTAR FRUTAS, QUANDO?
	Int. Parcial (Onde)	Onde o chefe plantou frutas?	CHEFE PASSADO PLANTAR FRUTAS, ONDE?
	Int. Parcial (Como)	Como o chefe plantou frutas?	CHEFE PASSADO PLANTAR FRUTAS, COMO?
	Int. Parcial (Por que)	Por que o chefe plantou frutas?	CHEFE PASSADO PLANTAR FRUTAS, POR-QUE?
	Int. Parcial ((O) que)	O que o chefe plantou?	CHEFE PASSADO PLANTAR O-QUE?
	Int. Parcial (Qual)	Quais frutas o chefe plantou?	CHEFE PASSADO PLANTAR, FRUTAS QUAL?
Int. Parcial (Quanto)	Quantas frutas o chefe plantou?	CHEFE PASSADO PLANTAR, FRUTAS QUANT@?	
26	Afirmativa	O médico doa remédios	MÉDIC@ DOAR REMÉDIOS
	Int. Total	O médico doa remédios?	MÉDIC@ DOAR REMÉDIOS VAI?
	Int. Parcial (Quando)	Quando o médico doa remédios?	MÉDIC@ DOAR REMÉDIOS, QUANDO FUTURO?
	Int. Parcial (Onde)	Onde o médico doa remédios?	MÉDIC@ DOAR REMÉDIOS, ONDE?
	Int. Parcial (Como)	Como o médico doa remédios?	MÉDIC@ DOAR REMÉDIOS, COMO?
	Int. Parcial (Por que)	Por que o médico doa remédios?	MÉDIC@ DOAR REMÉDIOS, POR-QUE?
	Int. Parcial ((O) que)	O que o médico doa?	MÉDIC@ DOAR O-QUE?
	Int. Parcial (Qual)	Quais remédios o médico doa?	MÉDIC@ DOAR, REMÉDIOS QUAL?
Int. Parcial (Quanto)	Quantos remédios o médico doa?	MÉDIC@ DOAR, REMÉDIOS QUANT@?	
27	Afirmativa	Ele varrerá a casa	HOMEM EL@ VARRER CASA VAI
	Int. Total	Ele varrerá a casa?	HOMEM FUTURO EL@ VARRER CASA VAI?
	Int. Parcial (Quando)	Quando ele varrerá a casa?	HOMEM EL@ FUTURO VARRER CASA, QUANDO?
	Int. Parcial (Onde)	Onde ele varrerá a casa?	HOMEM EL@ FUTURO VARRER CASA, ONDE?
	Int. Parcial (Como)	Como ele varrerá a casa?	HOMEM EL@ FUTURO VARRER CASA, COMO?
	Int. Parcial (Por que)	Por que ele varrerá a casa?	HOMEM EL@ FUTURO VARRER CASA, POR-QUE?
	Int. Parcial ((O) que)	O que ele varrerá?	HOMEM EL@ FUTURO VARRER O-QUE?
	Int. Parcial (Qual)	Qual casa ele varrerá?	HOMEM EL@ FUTURO VARRER, CASA QUAL?
Int. Parcial (Quanto)	Quantas casas ele varrerá?	HOMEM EL@ FUTURO VARRER, CASA QUANT@?	
28	Afirmativa	O mecânico dirigiu o carro	HOMEM MECÂNICO PASSADO DIRIGIR-CARRO
	Int. Total	O mecânico dirigiu o carro?	HOMEM MECÂNICO PASSADO DIRIGIR-CARRO?
	Int. Parcial (Quando)	Quando o mecânico dirigiu o carro?	HOMEM MECÂNICO PASSADO DIRIGIR-CARRO, QUANDO?
	Int. Parcial (Onde)	Onde o mecânico dirigiu o carro?	HOMEM MECÂNICO PASSADO DIRIGIR-CARRO, ONDE?
	Int. Parcial (Como)	Como o mecânico dirigiu o carro?	HOMEM MECÂNICO PASSADO DIRIGIR-CARRO, COMO?
	Int. Parcial (Por que)	Por que o mecânico dirigiu o carro?	HOMEM MECÂNICO PASSADO DIRIGIR-CARRO, POR-QUE?
	Int. Parcial ((O) que)	O que o mecânico dirigiu?	HOMEM MECÂNICO PASSADO DIRIGIR O-QUE?
	Int. Parcial (Qual)	Qual carro o mecânico dirigiu?	HOMEM MECÂNICO PASSADO DIRIGIR-CARRO QUAL?
Int. Parcial (Quanto)	Quantos carros o mecânico dirigiu?	HOMEM MECÂNICO PASSADO DIRIGIR-CARRO QUANT@?	
29	Afirmativa	O intérprete guardará a caixa	INTÉRPRETE GUARDAR CAIXA VAI
	Int. Total	O intérprete guardará a caixa?	FUTURO INTÉRPRETE GUARDAR CAIXA VAI?
	Int. Parcial (Quando)	Quando o intérprete guardará a caixa?	INTÉRPRETE FUTURO GUARDAR CAIXA, QUANDO?
	Int. Parcial (Onde)	Onde o intérprete guardará a caixa?	INTÉRPRETE FUTURO GUARDAR CAIXA, ONDE?
	Int. Parcial (Como)	Como o intérprete guardará a caixa?	INTÉRPRETE FUTURO GUARDAR CAIXA, COMO?
	Int. Parcial (Por que)	Por que o intérprete guardará a caixa?	INTÉRPRETE FUTURO GUARDAR CAIXA, POR-QUE?
	Int. Parcial ((O) que)	O que o intérprete guardará?	INTÉRPRETE FUTURO GUARDAR O-QUE?
	Int. Parcial (Qual)	Qual caixa o intérprete guardará?	INTÉRPRETE FUTURO GUARDAR, CAIXA QUAL?
Int. Parcial (Quanto)	Quantas caixas o intérprete guardará?	INTÉRPRETE FUTURO GUARDAR, CAIXA QUANT@?	
30	Afirmativa	O político distribuirá dinheiro	HOMEM POLÍTICO FUTURO DISTRIBUIR DINHEIRO
	Int. Total	O político distribuirá dinheiro?	HOMEM POLÍTICO FUTURO DISTRIBUIR DINHEIRO VAI?
	Int. Parcial (Quando)	Quando o político distribuirá dinheiro?	HOMEM POLÍTICO FUTURO DISTRIBUIR DINHEIRO, QUANDO?
	Int. Parcial (Onde)	Onde o político distribuirá dinheiro?	HOMEM POLÍTICO FUTURO DISTRIBUIR DINHEIRO, ONDE?
	Int. Parcial (Como)	Como o político distribuirá dinheiro?	HOMEM POLÍTICO FUTURO DISTRIBUIR DINHEIRO, COMO?
	Int. Parcial (Por que)	Por que o político distribuirá dinheiro?	HOMEM POLÍTICO FUTURO DISTRIBUIR DINHEIRO, POR-QUE?
	Int. Parcial ((O) que)	O que o político distribuirá?	HOMEM POLÍTICO FUTURO DISTRIBUIR O-QUE?
	Int. Parcial (Qual)	Qual dinheiro o político distribuirá?	HOMEM POLÍTICO FUTURO DISTRIBUIR, DINHEIRO QUAL?
Int. Parcial (Quanto)	Quanto dinheiro o político distribuirá?	HOMEM POLÍTICO FUTURO DISTRIBUIR, DINHEIRO QUANT@?	

APÊNDICE B – Dados tabulados

Interrogativa Total									
Grupo	T.Verb.	A_at	A_ant	B_at	B_ant	C_at	C_ant	D_at	D_ant
1	PS	4	0	0	0	0	0	0	0
2	PS	0	3	1	0	0	0	0	0
3	FP	0	0	0	0	0	0	0	0
4	PP	5	0	0	0	0	0	0	0
5	PS	1	0	0	0	0	0	0	0
6	FP	2	0	0	0	0	0	0	0
7	PS	3	0	0	0	0	0	0	0
8	PS	0	4	0	0	0	0	1	0
9	PP	2	0	0	0	0	0	0	0
10	FP	1	0	0	0	0	0	0	0
11	PP	0	0	0	0	0	0	0	0
12	PS	2	0	0	0	0	1	0	0
13	PP	2	0	0	0	0	0	0	0
14	FP	2	0	0	0	0	0	0	0
15	PP	0	3	0	0	0	0	1	0
16	PS	0	0	0	0	3	0	0	0
17	PS	0	0	0	0	4	0	0	0
18	FP	1	0	0	0	0	1	0	0
19	PP	0	0	0	0	1	0	0	0
20	FP	0	0	0	0	0	0	1	0
21	PP	3	0	0	0	0	0	0	0
22	PP	1	0	0	0	0	0	0	0
23	FP	4	0	0	0	0	0	0	0
24	PS	3	0	0	0	0	0	0	0
25	PP	1	0	0	0	0	0	0	0
26	PS	1	0	0	0	0	0	0	0
27	FP	0	0	0	0	0	0	5	0
28	PP	1	0	0	0	0	0	0	2
29	FP	0	0	0	0	0	0	1	0
30	FP	1	0	0	0	0	0	0	1
Quant.		19	3	1	0	3	2	5	2

Tabela B.1 – Dados Tabulados da modalidade interrogativa total

Interrogativas Parciais - ONDE									
Grupo	T.Verb.	A_at	A_ant	B_at	B_ant	C_at	C_ant	D_at	D_ant
1	PS	0	0	1	0	0	0	0	0
2	PS	0	1	1	0	0	0	0	0
3	FP	0	0	1	0	0	0	0	1
4	PP	0	1	1	0	0	0	0	0
5	PS	0	0	1	0	0	0	0	1
6	FP	0	1	1	0	0	0	0	0
7	PS	0	2	1	0	0	0	0	0
8	PS	0	0	1	0	0	0	0	0
9	PP	0	0	1	0	0	0	0	3
10	FP	0	0	1	0	0	0	0	2
11	PP	0	0	1	0	0	0	0	0
12	PS	0	1	1	0	0	0	0	0
13	PP	0	0	1	0	0	0	0	0
14	FP	0	0	1	0	0	0	0	0
15	PP	0	0	1	0	0	0	0	3
16	PS	0	0	1	0	0	1	0	0
17	PS	0	0	1	0	0	3	0	0
18	FP	0	0	1	0	0	1	0	0
19	PP	0	1	1	0	0	0	0	0
20	FP	0	0	1	0	0	0	0	0
21	PP	0	1	1	0	0	0	0	0
22	PP	0	0	1	0	0	0	0	0
23	FP	0	1	1	0	0	0	0	0
24	PS	0	2	1	0	0	0	0	0
25	PP	0	1	1	0	0	0	0	0
26	PS	0	0	1	0	0	2	0	0
27	FP	0	2	1	0	0	0	0	0
28	PP	0	2	1	0	0	0	0	0
29	FP	0	0	1	0	0	0	0	0
30	FP	1	0	0	0	0	0	0	1
	Quant.	1	12	29	0	0	4	0	6

Tabela B.2 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave *ONDE*

Interrogativas Parciais - Como									
Grupo	T.Verb.	A_at	A_ant	B_at	B_ant	C_at	C_ant	D_at	D_ant
1	PS	0	1	1	0	0	0	0	0
2	PS	0	3	1	0	0	0	0	0
3	FP	0	0	1	0	0	0	0	2
4	PP	0	0	1	0	0	1	0	0
5	PS	0	1	1	0	0	0	0	0
6	FP	0	1	1	0	0	0	0	0
7	PS	0	0	1	0	0	0	0	0
8	PS	0	3	1	0	0	0	0	0
9	PP	0	0	1	0	0	0	0	1
10	FP	0	1	1	0	0	0	0	0
11	PP	0	0	1	0	0	0	0	0
12	PS	0	1	1	0	0	0	0	0
13	PP	0	1	1	0	0	0	0	0
14	FP	0	0	1	0	0	0	0	0
15	PP	0	0	1	0	0	0	0	0
16	PS	0	3	1	0	0	0	0	0
17	PS	0	2	1	0	0	0	0	0
18	FP	0	0	1	0	0	0	0	0
19	PP	0	0	1	0	0	0	0	0
20	FP	0	1	1	0	0	0	0	0
21	PP	0	1	1	0	0	0	0	0
22	PP	0	0	1	0	0	0	0	0
23	FP	0	1	1	0	0	0	0	0
24	PS	0	2	1	0	0	0	0	0
25	PP	0	1	1	0	0	0	0	0
26	PS	0	0	1	0	0	0	0	0
27	FP	0	2	1	0	0	0	0	0
28	PP	0	0	1	0	0	0	0	1
29	FP	0	0	1	0	0	0	0	0
30	FP	0	0	1	0	0	0	0	0
	Quant.	0	16	30	0	0	1	0	3

Tabela B.3 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave *COMO*

Interrogativas Parciais - POR-QUE									
Grupo	T.Verb.	A_at	A_ant	B_at	B_ant	C_at	C_ant	D_at	D_ant
1	PS	0	1	1	0	0	0	0	0
2	PS	0	4	1	0	0	0	0	0
3	FP	0	1	1	0	0	0	0	0
4	PP	0	1	1	0	0	0	0	0
5	PS	0	1	1	0	0	0	0	0
6	FP	0	1	1	0	0	0	0	0
7	PS	4	0	0	0	0	0	0	0
8	PS	0	0	1	0	0	0	0	0
9	PP	0	3	1	0	0	0	0	0
10	FP	0	1	1	0	0	0	0	0
11	PP	0	0	1	0	0	0	0	0
12	PS	0	1	1	0	0	0	0	0
13	PP	0	1	1	0	0	0	0	0
14	FP	0	0	1	0	0	0	0	2
15	PP	0	0	1	0	0	1	0	0
16	PS	0	1	1	0	0	0	0	0
17	PS	0	0	1	0	0	3	0	0
18	FP	0	0	1	0	0	0	0	0
19	PP	0	1	1	0	0	0	0	0
20	FP	2	0	0	0	0	0	0	2
21	PP	0	1	1	0	0	0	0	0
22	PP	0	0	1	0	0	0	0	2
23	FP	0	1	1	0	0	0	0	0
24	PS	0	0	1	0	0	0	0	1
25	PP	0	1	1	0	0	0	0	0
26	PS	0	1	1	0	0	0	0	0
27	FP	0	0	1	0	0	0	0	1
28	PP	0	1	1	0	0	0	0	0
29	FP	0	2	1	0	0	0	0	0
30	FP	0	1	1	0	0	0	0	0
	Quant.	2	19	28	0	0	2	0	5

Tabela B.4 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave *POR-QUE*

Interrogativas Parciais - O-QUE									
Grupo	T.Verb.	A_at	A_ant	B_at	B_ant	C_at	C_ant	D_at	D_ant
1	PS	0	1	1	0	0	0	0	0
2	PS	0	0	1	0	0	0	0	2
3	FP	0	1	1	0	0	0	0	0
4	PP	0	0	1	0	0	0	0	0
5	PS	0	1	1	0	0	0	0	0
6	FP	0	0	1	0	0	0	0	0
7	PS	0	1	1	0	0	0	0	0
8	PS	0	2	1	0	0	0	0	0
9	PP	0	0	0	0	0	0	0	2
10	FP	0	0	0	0	0	0	0	0
11	PP	0	1	1	0	0	0	0	0
12	PS	0	1	1	0	0	0	0	0
13	PP	0	0	1	0	0	0	0	2
14	FP	0	0	1	0	0	0	0	1
15	PP	0	0	1	0	0	1	0	0
16	PS	0	0	1	0	0	0	0	1
17	PS	0	3	1	0	0	0	0	0
18	FP	0	1	1	0	0	0	0	0
19	PP	0	0	2	0	0	0	0	0
20	FP	0	0	1	0	0	0	0	0
21	PP	0	0	1	0	0	0	0	1
22	PP	0	0	1	0	0	1	0	0
23	FP	0	1	1	0	0	0	0	0
24	PS	0	2	1	0	0	0	0	0
25	PP	0	0	1	0	0	0	0	0
26	PS	0	0	2	0	0	0	0	0
27	FP	0	2	1	0	0	0	0	0
28	PP	0	1	1	0	0	0	0	0
29	FP	0	1	1	0	0	0	0	0
30	FP	0	1	1	0	0	0	0	0
	Quant.	0	15	28	0	0	2	0	6

Tabela B.5 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave *O-QUE*

Interrogativas Parciais - QUAL									
Grupo	T.Verb.	A_at	A_ant	B_at	B_ant	C_at	C_ant	D_at	D_ant
1	PS	0	0	1	0	0	1	0	0
2	PS	0	0	1	0	0	0	0	0
3	FP	0	0	1	0	0	3	0	0
4	PP	0	1	1	0	0	0	0	0
5	PS	0	0	1	0	0	0	0	1
6	FP	0	0	1	0	0	0	0	1
7	PS	0	2	1	0	0	0	0	0
8	PS	0	1	1	0	0	0	0	0
9	PP	0	0	1	0	0	0	0	1
10	FP	0	0	1	0	0	0	0	2
11	PP	0	0	1	0	0	0	0	0
12	PS	0	1	1	0	0	0	0	0
13	PP	0	0	1	0	0	0	0	0
14	FP	0	2	1	0	0	0	0	0
15	PP	0	0	1	0	0	0	0	0
16	PS	0	0	1	0	0	2	0	0
17	PS	0	2	1	0	0	0	0	0
18	FP	0	0	1	0	0	0	0	2
19	PP	0	0	1	0	0	0	0	0
20	FP	0	1	1	0	0	0	0	0
21	PP	0	0	1	0	0	0	0	0
22	PP	0	0	1	0	0	0	0	0
23	FP	0	0	1	0	0	0	0	0
24	PS	0	2	1	0	0	0	0	0
25	PP	0	1	1	0	0	0	0	0
26	PS	0	0	1	0	0	0	0	0
27	FP	0	2	1	0	0	0	0	0
28	PP	0	1	1	0	0	0	0	0
29	FP	0	2	1	0	0	0	0	0
30	FP	0	1	1	0	0	0	0	0
	Quant.	0	13	30	0	0	3	0	5

Tabela B.6 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave *QUAL*

Interrogativas Parciais - QUANT@									
Grupo	T.Verb.	A_at	A_ant	B_at	B_ant	C_at	C_ant	D_at	D_ant
1	PS	0	0	1	0	0	0	0	1
2	PS	0	4	1	0	0	0	0	0
3	FP	0	0	1	0	0	0	0	1
4	PP	0	1	1	0	0	0	0	0
5	PS	0	2	1	0	0	0	0	0
6	FP	0	1	1	0	0	0	0	0
7	PS	0	2	1	0	0	0	0	0
8	PS	0	0	1	0	0	3	0	0
9	PP	0	0	1	0	0	0	0	0
10	FP	0	0	1	0	0	2	0	0
11	PP	0	0	1	0	0	0	0	0
12	PS	0	3	1	0	0	0	0	0
13	PP	0	2	1	0	0	0	0	0
14	FP	0	2	1	0	0	0	0	0
15	PP	0	0	1	0	0	0	0	0
16	PS	0	0	1	0	0	2	0	0
17	PS	0	2	1	0	0	0	0	0
18	FP	0	0	1	0	0	1	0	0
19	PP	0	1	1	0	0	0	0	0
20	FP	0	1	1	0	0	0	0	0
21	PP	0	0	1	0	0	0	0	0
22	PP	0	0	1	0	0	0	0	0
23	FP	0	1	1	0	0	0	0	0
24	PS	0	2	1	0	0	0	0	0
25	PP	0	0	1	0	0	0	0	1
26	PS	0	0	1	0	0	0	0	0
27	FP	0	2	1	0	0	0	0	0
28	PP	0	1	1	0	0	0	0	0
29	FP	0	2	1	0	0	0	0	0
30	FP	0	0	1	0	0	0	0	0
	Quant.	0	16	30	0	0	4	0	3

Tabela B.7 – Dados Tabulados da modalidade interrogativa parcial com glosa-chave *QUANT@*

APÊNDICE C – Execução do Sistema Piloto

Este apêndice descreve os arquivos, configurações e parâmetros que envolvem a execução do sistema piloto. O sistema foi escrito na linguagem de programação Python. Os códigos e arquivos necessários para a execução estão disponíveis na plataforma do GitHub pelo endereço:

<https://github.com/ackleywill/PilotoAnimacaoFacial>.

O arquivo *CreateAnimation.py* apresenta os códigos do sistema que foram comentados para melhor entendimento das funcionalidades e execução.

Arquivos de parâmetros do sistema

Como parâmetro do sistema são necessários: o arquivo de sentenças e anotações, o arquivo de regras e o caminho do repositório das expressões faciais. Caso necessário, os caminhos desses arquivos podem ser alterados na seção do código demarcada como *Parâmetros iniciais*.

Arquivo: *sentencas.txt*

O arquivo *sentencas.txt* contém as sentenças em glosas e a descrição dos instantes de início de cada transição e sinal sinalizado assim como o instante de término da animação corporal. Abaixo está listado o conteúdo do arquivo com duas sentenças e suas anotações:

```
CASAS HOMEM EL@ VENDER VAI?
T360,S440,T535,S573,T620,S643,T691,S726,T766,S815,T914,E1023
HOMEM CONSTRUIR CASAS COMO?
T488,S540,T593,S630,T695,S740,T805,S835,T948,E1038
```

Arquivo: *regras.txt*

Neste arquivo são descritas as regras associadas aos sinais da [Libras](#). Em cada linha é apresentado o sinal e sua regra associada separados por um espaço. Abaixo está listado um trecho do conteúdo do arquivo *regras.txt*:

```
ONDE? B1
QUANDO? B1
```

COMO? B1A1

POR-QUE? B1A1

Repositório de expressões faciais

Os trechos de animações de expressões faciais são salvos em arquivos independentes em um diretório que representa o repositório de expressões faciais (Figura C.1). O nome de cada arquivo é composto pelo tipo da expressão facial seguido por um número inteiro sequencial. Os diferentes tipos de expressões faciais utilizados neste trabalho são representados pelas letras: A, B, C e D.

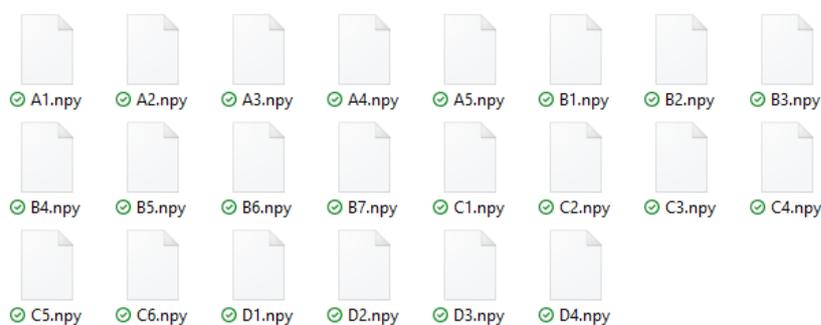


Figura C.1 – Repositório contendo os arquivos das animações de expressões faciais

Pré-requisitos

Para a execução do sistema piloto é necessário ter o software Python instalado. O arquivo de instalação pode ser obtido através do endereço: www.python.org/downloads/. É importante também adicionar o Python à configuração de *PATH* do sistema operacional. No Windows essa configuração pode ser realizada automaticamente selecionando a opção *add Python to PATH* no início da instalação do Python.

Para a execução também é necessária a instalação de alguns módulos utilizados no código. Para facilitar essa instalação foi criado o arquivo “*requerimentos.txt*” que automatiza as instalações. Para executar esse arquivo é necessário abrir o *prompt* de comando e digitar:

```
pip install -r requerimentos.txt
```

Execução

Após a instalação do Python e dos pacotes necessários o sistema piloto pode ser executado pelo *prompt* de comando. Na execução deve ser passado como parâmetro o índice da sentença no arquivo *sentenças.txt*. O exemplo abaixo apresenta a execução do

sistema utilizando o índice 0 (zero), correspondente a sentença: CASAS HOMEM EL@VENDER VAI?.

```
python CreateAnimation.py -sentenca 0
```

O sistema irá gerar o arquivo da animação facial para ser executado pelo avatar. O nome do arquivo resultante inicia com os caracteres EF seguido do índice da sentença processada e possui a extensão npy (matriz NumPy). Os arquivos das animações são salvos, por padrão, no diretório *dados\animacoesProntas* localizado na raiz do sistema.

Relatório e gráfico da execução

Durante a execução do programa é exibido um relatório textual com informações sobre a execução do sistema e ao final um gráfico com o deslocamento do eixo x de um marcador da sobrancelha ao longo dos quadros de animação. Tanto o marcador quanto o eixo podem ser alterados na seção denominada *Representação gráfica* no código. O trecho de código abaixo indica que será exibido o eixo X do marcador de índice 10.

```
axis2Print = [True, False, False]
marker2Print = 10
```

O marcador a ser exibido pode ser escolhido através da alteração da variável *marker2Print* com um valor entre 0 e 38. É possível escolher o(s) eixo(s) a ser(em) exibido(s) através da alteração dos valores da lista *axis2Print*. Essa lista contém 3 valores *booleanos* indicando a execução dos eixos x, y e z respectivamente. O valor *True* indica o eixo que deve ser exibido e o valor *False* o eixo que não deve ser exibido.

APÊNDICE D – Relatório de execução da criação da animação

Dados de entrada formatados:

```
[['HOMEM', 540, 53, 52, 37], ['CONSTRUIR', 630, 65, 37, 45],  
['CASAS', 740, 65, 45, 30], ['COMO?', 835, 113, 30, 90]]
```

Sinais com expressões faciais:

```
[['COMO', 'B1A1', 835, 113, 30, 90], ['HOMEM', 'C1', 540, 53, 52, 37],  
['CONSTRUIR', 'A1', 630, 65, 37, 45]]
```

Sinal: COMO

Expressão facial do tipo B

Transição inicial: 30

Transição final: 90

Animações (Expressões faciais) no repositório:

B1 - Duração: 185 - Diferença: 72

B2 - Duração: 86 - Diferença: 27

B3 - Duração: 149 - Diferença: 36

B4 - Duração: 86 - Diferença: 27

B5 - Duração: 23 - Diferença: 90

B6 - Duração: 45 - Diferença: 68

B7 - Duração: 28 - Diferença: 85

Animação escolhida: B2.npy

Duração esperada: 113 - Duração da animação escolhida: 86

Aplicação de aumento da duração
Tamanho Original: 86
Duração da animação ajustada: 113

Sinal: CASAS (da regra aplicada ao sinal COMO)

Expressão facial do tipo A

Transição inicial: 45

Transição final: 30

Animações (Expressões faciais) no repositório:

A1 - Duração: 26 - Diferença: 39

A2 - Duração: 82 - Diferença: 17

A3 - Duração: 169 - Diferença: 104

A4 - Duração: 67 - Diferença: 2

A5 - Duração: 23 - Diferença: 42

Animação escolhida: A4.npy

Duração esperada: 65 - Duração da animação escolhida: 67

Aplicação de redução da duração
Tamanho Original: 67
Duração da animação ajustada: 65

Sinal: HOMEM

Expressão facial do tipo C

Transição inicial: 52

Transição final: 37

Animações (Expressões faciais) no repositório:

C1 - Duração: 111 - Diferença: 58

C2 - Duração: 82 - Diferença: 29

C3 - Duração: 44 - Diferença: 9

C4 - Duração: 75 - Diferença: 22

C5 - Duração: 54 - Diferença: 1

C6 - Duração: 31 - Diferença: 22

Animação escolhida: C5.npy

Duração esperada: 53 - Duração da animação escolhida: 54

Aplicação de redução da duração

Tamanho Original: 54

Duração da animação ajustada: 53

Sinal: CONSTRUIR

Expressão facial do tipo A

Transição inicial: 37

Transição final: 45

Animações (Expressões faciais) no repositório:

A1 - Duração: 26 - Diferença: 39

A2 - Duração: 82 - Diferença: 17

A3 - Duração: 169 - Diferença: 104

A4 - Duração: 67 - Diferença: 2

A5 - Duração: 23 - Diferença: 42

Animação escolhida: A4.npy

Duração esperada: 65 - Duração da animação escolhida: 67

Aplicação de redução da duração

Tamanho Original: 67

Duração da animação ajustada: 65

APÊNDICE E – Execução da animação facial pelo avatar

Neste apêndice são apresentados os procedimentos para a execução, pelo avatar sinalizante, das animações faciais produzidas pelo sistema piloto. As animações faciais acompanham as animações corporais já inseridas no avatar sinalizante. Dessa forma, são disponibilizados arquivos criados no software Blender (BLENDER, 2021), do avatar já com as animações corporais para a execução dos procedimentos aqui apresentados. Os arquivos (Anim_0.rar e Anim_1.rar) estão disponíveis na plataforma GitHub e podem ser acessadas pelo endereço:

<https://github.com/ackleywill/PilotoAnimacaoFacial>.

Pré-requisitos

Para a execução das animações é necessário ter instalado o software Blender que pode ser obtido pelo endereço: <https://www.blender.org/>. Os experimentos aqui apresentados foram realizados na versão do Blender 2.93.4.

Execução

Ao abrir os arquivos descompactados do avatar com a animação corporal (Anim_0.blend e Anim_1.blend) no Blender é exibido, na aba *Scripting*, o avatar no painel à esquerda e o *script* responsável pela incorporação da animação facial no painel à direita. Abaixo do painel do avatar se encontra o painel de linha de tempo (Figura E.1).

No *script*, a variável *sentence* indica o nome do arquivo (sem sua extensão) que contém os dados da animação facial, criada pelo sistema piloto, a ser incorporada ao avatar. Na variável *dispPath* é armazenado o caminho do arquivo de animação. Essa variável deve ser alterada de acordo com o local em que foi salvo o arquivo.

A execução do *script* é realizada através do atalho *Alt+P* ou através de um clique no botão marcado com um retângulo vermelho (Figura E.1) na parte superior do painel de *script*. Após a execução é possível verificar as animações arrastando o cursor na barra do painel da linha de tempo ou clicando no botão de *play animation* marcado com um retângulo azul (Figura E.1). Também é possível verificar a animação no modo renderizado clicando no ícone na parte superior do painel de animação marcado com um retângulo amarelo (Figura E.1).

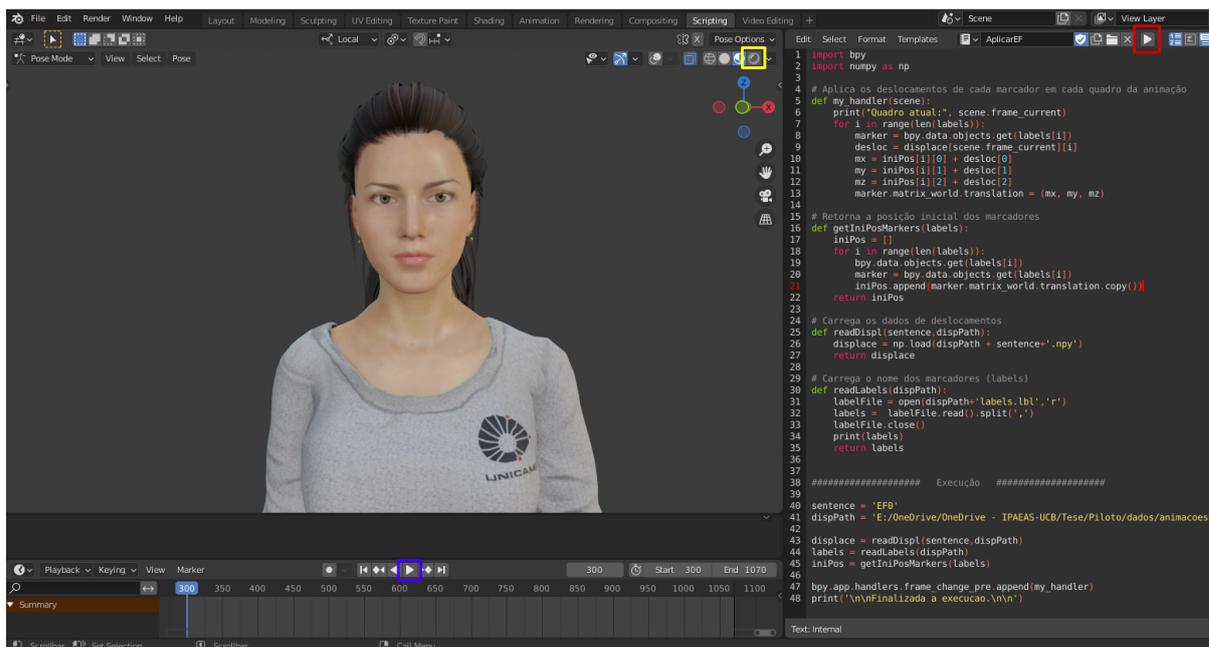


Figura E.1 – Tela inicial do Blender ao ser aberto o arquivo de animação corporal

A Figura E.2 apresenta três momentos de execução, no modo renderizado, da animação da sentença CASAS HOMEM EL@ VENDER VAI? produzida pelo sistema piloto.



Figura E.2 – Quadros da animação da sentença CASAS HOMEM EL@ VENDER VAI? no Blender

Anexos

ANEXO A – Documentos da pesquisa

O projeto de pesquisa e uma emenda foram enviados ao Comitê de Ética e Pesquisa (CEP) da Unicamp, resultando em aprovação. Abaixo são apresentados trechos do pareceres finais e o termo de autorização de uso de imagem.



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Avaliação de Animações de Expressões Faciais em Sinalizações em Libras

Pesquisador: ACKLEY DIAS WILL

Área Temática:

Versão: 2

CAAE: 29650820.1.0000.5404

Instituição Proponente: Faculdade de Engenharia Elétrica e de Computação

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.988.386

Conclusões ou Pendências e Lista de Inadequações:

Aprovado.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

CAMPINAS, 24 de Abril de 2020

Assinado por:
Renata Maria dos Santos Celeghini
(Coordenador(a))



INSTRUMENTO DE CESSÃO DE DIREITOS PARA USO E COMPARTILHAMENTO DE IMAGENS
“AUTORIZAÇÃO DE USO DE IMAGEM, VOZ E RESPECTIVA CESSÃO DE DIREITOS (LEI N. 9.610/98)”

Eu _____, portador do CPF nº:
_____, RG nº: _____, domiciliado em

nos termos da legislação de direitos autorais, especialmente os arts. 29 a 45 da Lei Federal nº 9610/98 autorizo a UNIVERSIDADE ESTADUAL DE CAMPINAS – UNICAMP, pessoa jurídica de direito público, autarquia estadual, devidamente inscrita no CNPJ sob nº 46.068.425/0001-33 e isenta de inscrição estadual, com sede na Cidade Universitária “Zeferino Vaz”, Distrito de Barão Geraldo, Campinas – SP, a utilizar a minha imagem e som de voz no projeto de pesquisa “Avaliação de Animações de Expressões Faciais em Sinalizações em Libras”, coordenado pelo aluno de doutorado Ackley Dias Will, RA: 179529, da Faculdade de Engenharia Elétrica e de Computação (FEEC) da Universidade Estadual de Campinas (Unicamp).

Em particular, autorizo e cedo, a título definitivo e gratuito e em caráter de colaboração pessoal, a utilização de minhas imagens fotográficas, som de voz e vídeo do projeto para fins de publicação, exibição, transmissão, retransmissão, distribuição, redistribuição, processamento digital e compartilhamento, para fins exclusivos de divulgação e pesquisa científica, incluindo a divulgação em mídia impressa (livros, catálogos, revista, jornal, entre outros), mídia eletrônica (programas de rádio, podcasts, vídeos e filmes para televisão aberta e/ou fechada, documentários para cinema ou televisão, entre outros), Internet, Banco de Dados Informatizado Multimídia, “home vídeo” e DVD (digital video disc). A autorização inclui a veiculação de minhas imagens e som de voz para realização de testes e avaliações pertinentes ao projeto realizadas via web (Internet).

Campinas, ____ de _____ de 2020

Ass: _____

**PARECER CONSUBSTANCIADO DO CEP****DADOS DA EMENDA**

Título da Pesquisa: Avaliação de Animações de Expressões Faciais em Sinalizações em Libras

Pesquisador: ACKLEY DIAS WILL

Área Temática:

Versão: 3

CAAE: 29650820.1.0000.5404

Instituição Proponente: Faculdade de Engenharia Elétrica e de Computação

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 4.101.894

Conclusões ou Pendências e Lista de Inadequações:

Emenda aprovada

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

CAMPINAS, 22 de Junho de 2020

Assinado por:
Renata Maria dos Santos Celeghini
(Coordenador(a))