



ESCOLA DE ARTES, CIÊNCIAS E
HUMANIDADES
EACH - USP

**Geração de expressões faciais a
partir de antropometria e
gramáticas para aplicação em
jogos psiquiátricos**

PROGRAMA INSTITUCIONAL DE BOLSAS DE
INICIAÇÃO CIENTÍFICA – PIBIC/CNPQ/USP

RELATÓRIO FINAL

Autor:
Rafael Luiz Testa

Orientadora:
Profa. Dra. Fátima L. S.
Nunes Marques

SÃO PAULO - JULHO/2014

Resumo

Processar e identificar emoções faciais é um componente essencial para estabelecer interação entre pessoas. Alguns transtornos psiquiátricos podem limitar a capacidade de um indivíduo em reconhecer emoções em expressões faciais. A fim de contribuir com a solução deste problema técnicas computacionais podem ser utilizadas compor ferramentas destinadas ao diagnóstico, avaliação e treinamento no reconhecimento de emoções faciais. Este trabalho apresenta uma abordagem que usa *splines*, linguagens formais, antropometria e *Facial Action Coding System (FACS)* para gerar caricaturas que representam as emoções neutra, satisfação, tristeza, raiva, aversão-nojo, medo e surpresa. Uma avaliação com usuários voluntários saudáveis indicou que algumas emoções são mais facilmente reconhecidas e que as caricaturas precisam ser aperfeiçoadas para outras emoções. A abordagem é promissora, uma vez que confere flexibilidade, a partir de parâmetros, para definir a intensidade da emoção que deve ser representada e, futuramente, outras características como gênero, grupo étnico e idade.

Sumário

1	Introdução	7
1.1	Objetivos	8
2	Revisão bibliográfica	10
2.1	Antropometria	10
2.2	Expressões faciais	12
2.3	Gramáticas	15
2.4	Geração de imagens	17
2.5	Trabalhos correlatos	18
3	Metodologia	21
3.1	Representação da face	21
3.1.1	Modelo antropométrico	21
3.2	<i>Splines</i>	33
3.3	Representação de emoções	33
3.3.1	O modelo de representação das expressões faciais por meio de FACS	34
3.4	Integração com a gramática	42
3.4.1	Construção do Grafo AND-OR	42
3.5	Avaliação das imagens geradas	44
4	Resultados	46
4.1	Discussão	49
4.1.1	Avaliação global das caricaturas	50
4.1.2	Variáveis demográficas e coerência das respostas	51
5	Considerações finais	53
A	Generating facial emotions using anthropometry and formal languages for individuals suffering from psychiatric disorders	
B	Avaliação de desempenho do aluno	

C Avaliação do Programa

Listas de Figuras

2.1	<i>Landmarks</i> antropométricos faciais. (FARKAS, 1994)	11
2.2	Medições antropométricas faciais. (FARKAS, 1994)	11
2.3	Exemplos das unidades de ação 1, 2 e 4.	12
2.4	<i>Exemplo de grafo AND-OR de relógios de parede.</i> A imagem ilustra a estrutura geral do grafo e grafos sintáticos de exemplos de relógios. Fonte: (LIN et al., 2009).	17
2.5	Exemplos de modificação nos pontos de controle em <i>b-splines</i> . Retirado de (FOLEY, 1996)	18
2.6	Funcionamento do sistema. Adaptado de (YANG et al., 2009)	19
2.7	Modelos masculinos variados (DECARLO; METAXAS; STONE, 1998)	20
3.1	Esquematização da metodologia.	22
3.2	<i>Landmarks</i> utilizados para a construção dos <i>sketches</i> . As marcações em vermelho representam os pontos de controle utilizados nas curvas, e as marcações em azul são pontos empregados como auxiliares para posicionamento de outros pontos de controle. Além disso, foram criados pontos com letras verdes para construir as curvas.	25
3.3	A linha vermelha interna mostra a reção entre o tamanho do envoltório e a curva contida. A linha externa delimita um envoltório que é o dobro do tamanho. A curva em preto corresponde a uma parte do rosto e em verde seus pontos de controle.	35
3.4	<i>Ilustração do grafo AND-OR parcial com a face neutra e as emoções Satisfação e Tristeza.</i>	43
3.5	Exemplo de formulário usado para coletar dados.	45
4.1	Expressões faciais utilizadas na avaliação.	47
4.2	Taxa de acerto das respostas para cada emoção avaliada.	48
4.3	Taxa de acerto das respostas para cada gênero.	48
4.4	Taxa de acerto das respostas para cada faixa etária.	49

Lista de Tabelas

2.1	Descrições da aparência de cada unidade de ação baseados em (EKMAN; FRIESEN; HAGER, 2002)	13
2.1	Descrições da aparência de cada unidade de ação baseados em (EKMAN; FRIESEN; HAGER, 2002)	14
2.1	Descrições da aparência de cada unidade de ação baseados em (EKMAN; FRIESEN; HAGER, 2002)	15
2.2	Relação entre unidades de ação e as emoções. Adaptada de (EKMAN; FRIESEN; HAGER, 2002)	15
3.1	Inferências feitas a partir da antropometria.	23
3.1	Inferências feitas a partir da antropometria.	24
3.2	Cálculo do posicionamento dos <i>landmarks</i> para os olhos.	26
3.3	Cálculo do posicionamento dos <i>landmarks</i> para sobrancelha.	27
3.4	Cálculo do posicionamento dos <i>landmarks</i> para as pálpebras.	27
3.5	Cálculo do posicionamento dos <i>landmarks</i> para o nariz.	28
3.6	Cálculo do posicionamento dos <i>landmarks</i> para os lábios.	28
3.7	Cálculo do posicionamento dos <i>landmarks</i> para as pupilas.	29
3.8	Cálculo do posicionamento dos <i>landmarks</i> para os enrugamentos dos olhos.	29
3.9	Cálculo do posicionamento dos <i>landmarks</i> para enrugamento da testa.	30
3.10	Cálculo do posicionamento dos <i>landmarks</i> para enrugamento das bochechas.	30
3.10	Cálculo do posicionamento dos <i>landmarks</i> para enrugamento das bochechas.	31
3.11	Cálculo do posicionamento dos <i>landmarks</i> para os enrugamentos do nariz.	31
3.11	Cálculo do posicionamento dos <i>landmarks</i> para os enrugamentos do nariz.	32
3.12	Cálculo do posicionamento dos <i>landmarks</i> para enrugamento das bochechas.	32
3.13	Cálculo do posicionamento dos <i>landmarks</i>	33

3.14 Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.	36
3.14 Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.	37
3.14 Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.	38
3.14 Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.	39
3.14 Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.	40
3.14 Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.	41
3.14 Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.	42
4.1 Matriz de confusão das respostas obtidas. Em negritos estão indicadas as taxas de correspondência das emoções esperadas.	46
4.2 Coerência das respostas referentes às imagens duplicadas. . . .	49

Capítulo 1

Introdução

A comunicação de uma emoção pode ser estabelecida por meio verbal (SCHE-RER, 2003) e também por meios não verbais como expressões faciais, posturas e gestos. A comunicação efetiva de uma emoção é benéfica tanto para quem está se comunicando quanto para quem está tentando compreender (HESS, 2001).

A capacidade de processar e identificar emoções faciais é um componente essencial da comunicação humana e interação social de acordo com as normas e costumes culturais, baseadas em padrões comuns. Seis emoções universais já foram estabelecidas, incluindo, satisfação, tristeza, raiva, aversão-nojo, medo e surpresa, cada uma das quais corresponde a um padrão específico dos músculos da face e tem neurocircuitos parcialmente separáveis, capazes de preparar uma resposta comportamental apropriada (EKMAN; FRIESEN, 1971; ROCCA et al., 2009; GOSSELIN; KIROUAC et al., 1995).

Segundo a revisão de Rocca et al. (2009) o processamento de expressões faciais das emoções, que varia de interpretação do estímulo para a preparação de uma resposta comportamental apropriada, é apoiado por sistemas neurais específicos. Estudos de imagens cerebrais sugerem que regiões específicas no córtex temporal inferior medeiam a percepção de faces. Doenças distintas podem se caracterizar por déficits diferentes em aspectos específicos do processamento emocional. Por exemplo, pacientes com doença de Urbach-Wiethe, que causa dano bilateral na amígdala, são incapazes de distinguir medo tanto em faces reais como imaginárias, nos transtornos psiquiátricos, o paradigma de reconhecimento de faces vem recebendo maior atenção especialmente nos transtornos do espectro autista (HARMS; MARTIN; WALLACE, 2010), transtornos do humor (ROCCA et al., 2009) e na esquizofrenia (TAYLOR; MACDONALD, 2012). Estudos mais recentes vêm inclusive avaliando a possibilidade de reconhecimento de padrões em populações de risco para transtornos mentais (MOURÃO-MIRANDA et al., 2012). Portanto, o para-

digma de reconhecimento de expressões faciais pode ser desenvolvido como uma ferramenta na avaliação de transtornos mentais. Assim uma ferramenta capaz de melhorar as habilidades de reconhecimento das expressões faciais pode ser de grande utilidade.

Estudos sugerem que indivíduos portadores de transtornos do espectro autista apresentam dificuldades na interação e comunicação social (NEWSCHAFFER; CURRAN, 2003), incluindo a habilidade de identificar a expressão facial das emoções em outras pessoas (CLARK; WINKIELMAN; MCINTOSH, 2008). O uso de ambientes de realidade virtual, de expressões 3D em ambiente multimídia interativo e de software de treinamento utilizando interface humano-computador apresentou respostas positivas no que se refere à melhora desses aspectos em indivíduos do espectro autista (LAHIRI et al., 2013; CHENG; LING, 2008; GRYNSZPAN; MARTIN; NADEL, 2008).

Com esta motivação, a presente proposta faz parte de um projeto de pesquisa mais abrangente que visa a desenvolver um jogo de reconhecimento de expressões faciais que possa ser explorado no campo de pesquisa dos transtornos mentais. Em uma versão inicial deste jogo, imagens estáticas de várias emoções devem ser geradas e mostradas ao jogador, que deve adivinhar qual é a emoção descrita naquela imagem. Posteriormente o jogador deverá identificar mudanças emocionais ocorrendo dinamicamente (progressivamente) na imagem. Também será de interesse selecionar, dependendo do perfil do jogador, características das imagens (ou seja, do personagem) do jogo, como sexo, idade e grupo étnico. Sendo assim, grande parte do desafio computacional da implementação deste jogo reside na geração destas imagens.

1.1 Objetivos

O presente projeto faz parte de um projeto maior, cujo objetivo é utilizar técnicas gramaticais e de processamento de imagens para gerar, em um primeiro momento, imagens (caricaturas) das emoções neutra, satisfação, tristeza, raiva, aversão, medo e surpresa. Com intuito de gerar imagens consistentes com uma face humana, estas serão baseadas em medidas antropométricas. O projeto é composto por duas partes. A primeira parte, com ênfase na parte gramatical, consiste na implementação da estrutura de dados de grafos AND-OR (ZHU; MUMFORD, 2007a) e a definição do grafo AND-OR (símbolos terminais e não terminais, relações, etc) para as várias emoções. A segunda parte, com ênfase em processamento de imagens, consiste na definição de pontos de controle para os vários componentes de face e implementação de um programa capaz de, a partir do grafo AND-OR definido na primeira parte, gerar as imagens utilizando técnicas de processamento de

imagens. O objetivo do projeto desta iniciação científica, ou seja, do aluno em questão, é a realização da parte dois, orientado pela Profa. Fátima de Lourdes dos Santos Nunes, especialista em processamento de imagens. A parte um está sendo realizada por outro aluno sob a supervisão da Profa. Ariane Machado Lima, especialista em gramáticas.

O presente capítulo apresenta uma introdução ao tema, justificando o trabalho e apresentando os objetivos. O capítulo 2 apresenta os conceitos básicos sobre geração de imagens, antropometria, expressões faciais e gramáticas, além de expor os trabalhos correlatos. Já o capítulo 3 mostra como as medidas antropométricas foram utilizadas na geração da face e como os conceitos da psicologia e de traçado de curvas (computação gráfica) foram empregados para representar as expressões faciais, também trata da integração com a gramática e forma de avaliação dos resultados. O capítulo 4 expõe as imagens e resultados obtidos da metodologia em questão. O capítulo 5 apresenta uma discussão sobre a abordagem adotada

Capítulo 2

Revisão bibliográfica

Neste capítulo são apresentados os conceitos referentes às áreas que compõem o tema deste projeto.

2.1 Antropometria

Antropometria é a ciência que estuda as medidas de tamanho, peso e proporções do corpo humano (FARKAS, 1994); é um método de análise obtido por meio da mensuração da população estudada. Essas medidas podem ser utilizadas na medicina na análise de irregularidades em pacientes e o planejamento de cirurgias plásticas (FARKAS, 1981), antropologia forense (KRISHAN, 2007) e no envelhecimento de imagens de faces (RAMANATHAN; CHELLAPPA; BISWAS, 2009). Esta proposta utiliza as medidas antropométricas para a geração de *sketches* que representem faces.

Para a realização das medições foi determinado um conjunto de pontos bem definidos sobre o corpo humano, chamados de *landmarks*. Por exemplo, no caso da face o ponto denominado *sn* indica o local abaixo do nariz, como ilustrado na Figura 2.1. Assim, a distância entre dois *landmarks* irá compor uma medida.

Essas distâncias podem ser dos seguintes tipos:

- menor distância entre dois *landmarks*, exemplo: *en-ex* na Figura 2.2;
- distância axial entre dois *landmarks*, exemplo: *v-tr* na Figura 2.2;
- distância geodésica entre dois *landmarks*, exemplo: *ch-t* na Figura 2.2;
- ângulo de inclinação formado de um *landmark* em relação aos eixos canônicos, exemplo: ângulo na orelha na Figura 2.2.

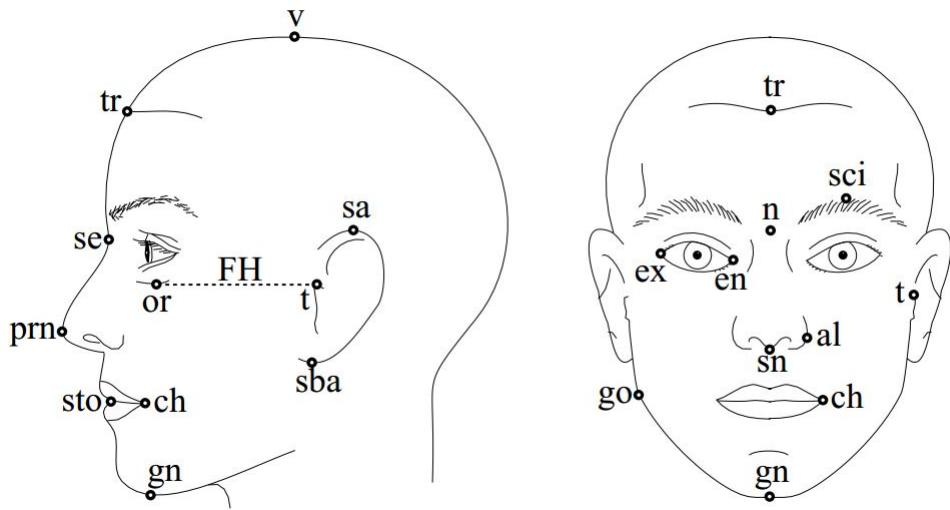


Figura 2.1: *Landmarks* antropométricos faciais. (FARKAS, 1994)

- ângulo entre retas definidas por *landmarks*, exemplo: ângulo no queixo na Figura 2.2.

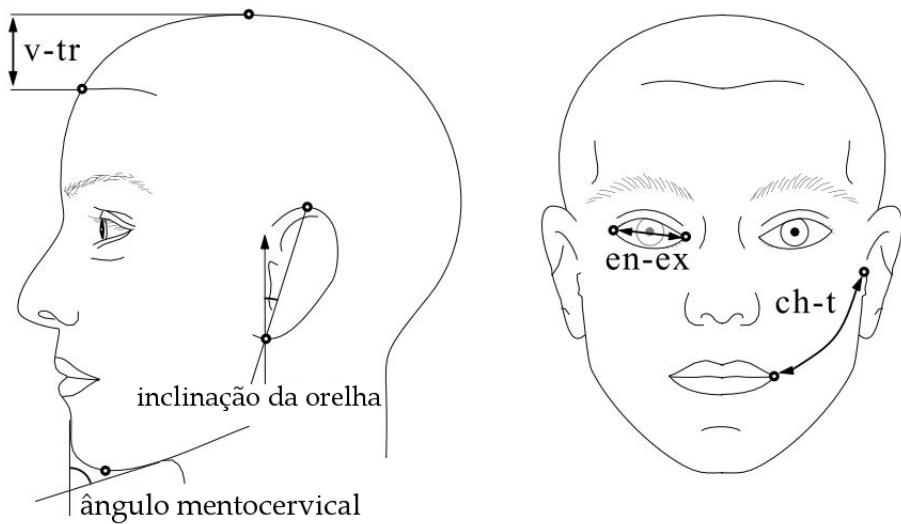


Figura 2.2: Mediçãoes antropométricas faciais. (FARKAS, 1994)

Existem fatores que podem influenciar nas variações dessas medidas como, por exemplo, as diferenças naturais entre os grupos étnicos (FARKAS; KATIC; FORREST, 2005), os gêneros (WEEDEN; TROTMAN; FARAWAY, 2001), as faixas etárias (PERISSINOTTO et al., 2002), e os tipos faciais (longo, médio, curto) (GUEDES; TEIXEIRA; CATTONI, 2010), etc.

2.2 Expressões faciais

Com base na universalidade das expressões faciais e na representação das emoções pelas mesmas, foram definidas técnicas que permitissem relacionar as medidas dos movimentos faciais com as expressões geradas. Ekman e Friesen (1976) formulam uma maneira de medir os movimentos faciais. Para isso, definiram unidades de ação (*action units*) como ações fundamentais (mínimas) de um músculo ou um grupo de músculos. Por meio das unidades de ação é possível constituir qualquer comportamento facial. Assim, pode-se compor uma expressão facial como uma combinação dessas unidades de ações.

Essas ideias foram utilizadas para a elaboração de uma metodologia denominada *Facial Action Coding System (FACS)*, capaz de classificar expressões faciais com base na identificação da presença de determinadas unidades de ação presentes no rosto (EKMAN; FRIESEN, 1978).

A fim de ilustrar as unidades de ação, a Figura 2.3 mostra as unidades de ação 1, 2 e 4. Na unidade de ação 1 (Figura 2.3(a)) os cantos internos das sobrancelhas estão erguidos; na UA 2 (Figura 2.3(b)) as partes externas das sobrancelhas estão levantadas e na UA 4 (Figure 2.3(c)) as sobrancelhas estão unidas.

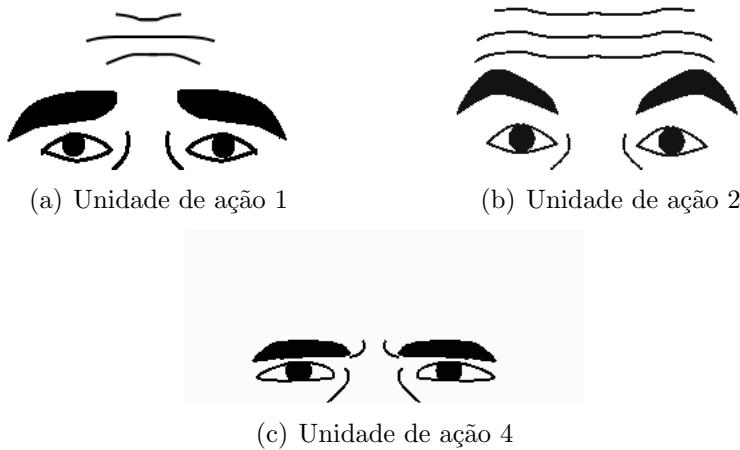


Figura 2.3: Exemplos das unidades de ação 1, 2 e 4.

A Tabela 2.1 contém uma descrição textual do movimento e da aparência gerada por cada unidade de ação relevante. Embora as unidades de ação sejam consideradas de forma independente, a combinação de algumas delas pode causar um resultado diferente daquele obtido por meio da simples junção das unidades de ação. Isto ocorre porque um músculo acaba exercendo

influência no outro.

Tabela 2.1: Descrições da aparência de cada unidade de ação baseados em (EKMAN; FRIESEN; HAGER, 2002)

Unidade de ação	Nome	Descrição
1	Inner Brow Raiser	Eleva os cantos internos da sobrancelha Pode causar enrugamento horizontal na testa
2	Outer Brow Raiser	Eleva a porção exterior das sobrancelhas Produz uma forma arqueada para as sobrancelhas Permite gerar rugas laterais entre as sobrancelhas
4	Brow Lowerer	Diminui a sobrancelha Empurra as pálpebras estreitando a altura dos olhos Aproxima as sobrancelhas Produz rugas verticais entre as sobrancelhas Pode produzir um erugamento do nariz
5	Upper Lid Raiser	Amplia a abertura dos olhos Eleva a pálpebra superior sem que ela desapareça por completo Permite expor esclera por cima da íris Apresenta um olhar fixo Amplia pálpebra inferior
6	Cheek Raiser and Lid Compressor	Puxa as bochechas em direção aos olhos Empurra a pele ao redor dos olhos em direção à cavidade ocular Aprofunda o sulco da pálpebra inferior Pode diminuir porção lateral das sobrancelhas Provoca rugas nos cantos externos da abertura dos olhos (Pés de galinha)
7	Lid Tightener	Aperta as pálpebras Diminui a abertura dos olhos Eleva a pálpebra inferior Pode causar rugas nos olhos

Tabela 2.1: Descrições da aparência de cada unidade de ação baseados em (EKMAN; FRIESEN; HAGER, 2002)

Unidade de ação	Nome	Descrição
9	Nose Wrinkler	Puxa a pele do nariz, causando rugas laterais
		Reduz a parte medial das sobrancelhas
		Limita a abertura dos olhos
		Puxa o centro do lábio superior para cima
		Pode ampliar e levantar as narinas
12	Lip Corner Puller	Puxa os cantos dos lábios para trás e para cima, de forma oblíqua
		Aprofunda o sulco nasogeniano, puxando-o lateralmente e para cima
		Pode limitar a abertura dos olhos
		Pode aumentar e alargar as narinas
		Pode achatar e esticar a pele do queixo
15	Lip Corner Depressor	Puxa os cantos dos lábios para baixo
		O lábio inferior é ligeiramente esticado horizontalmente
		Pode produzir enrugamentos nos cantos dos lábios
		Pode achatar/ produzir uma depressão no queixo
16+25	Lower Lip Depressor	Puxa o lábio inferior para baixo
		Eticam os lábios lateralmente
		Normalmente expõe mais os dentes
		Estica o queixo
		Pode causar rugas no queixo ou no lábio inferior
20	Lip Stretcher	Estica os lábio lateralmente
		Alonga a boca
		Achata os lábios
		Pode aparecer rugas nos cantos do lábios
		Puxa a porção inferior do sulco nasolabial, e puxa a pele na área inferior do sulco nasolabial lateralmente

Tabela 2.1: Descrições da aparência de cada unidade de ação baseados em (EKMAN; FRIESEN; HAGER, 2002)

Unidade de ação	Nome	Descrição
20	Lip Stretcher	Estica lateralmente a pele do queixo Pode alongar a abertura das narinas
23	Lip Tightener	Estreita os lábios
		Pode haver uma leve separação dos lábios
		Pode produzir enrugamento dos lábios
25	Lips Part	Expõe uma parte da mucosa da boca
		Pode expor os dentes
26	Jaw Drop	A mandíbula é abaixada
		Pode separar os lábios a fim de expor os dentes

A Tabela 2.2 mostra uma possível combinação das unidades de ação que compõem cada expressão facial correspondente a uma emoção.

Tabela 2.2: Relação entre unidades de ação e as emoções.
Adaptada de (EKMAN; FRIESEN; HAGER, 2002)

Emoções	Unidades de ação
Satisfação	6+12
Tristeza	1+4+15
Surpresa	1+2+5B+26
Medo	1+2+4+5+20+26
Raiva	4+5+7+23
Aversão (Nojo)	9+15+16+25

2.3 Gramáticas

A descrição de uma face pode ser feita como uma composição hierárquica multinível (olhos, boca, nariz, etc., que, por sua vez, são compostos por outros elementos mais simples – olho, por exemplo, é composto por contorno, sobrancelhas, cílios, íris, etc.). Além disso, regras definem diferentes tipos de elementos (exemplo: olho cerrado ou aberto) relacionados de diferentes maneiras (exemplo: distância entre as sobrancelhas) dependendo da emoção. Sendo assim, as diferentes faces para as várias emoções podem ser descritas

por meia de gramáticas, mais especificamente grafos AND-OR.

Um grafo AND-OR é uma sétupla que representa uma gramática de imagem $G = \langle S, V_T, V_N, R, C, P \rangle$, em que S é o símbolo inicial da gramática, V_T representa os nós terminais, V_N representa o conjunto de nós não terminais, R é o conjunto de regras ou produções, C é o conjunto de todas as configurações válidas que podem ser derivadas de G e P representa o modelo de probabilidade definido em G (LIN et al., 2009). Graficamente, essa gramática é descrita por um grafo AND-OR no qual o nó raiz representa o símbolo inicial e os nós folhas representam símbolos terminais, os nós internos representam os símbolos não terminais e as arestas verticais estão relacionadas às produções: os filhos de um nó OR indicam todas as produções possíveis para um mesmo símbolo não-terminal; e os filhos de um nó AND indicam todos os símbolos que são gerados por uma determinada produção (ZHU; MUMFORD, 2007b).

Gramáticas e grafos AND-OR já vêm sendo bastante utilizados em reconhecimento de padrões em imagens. O problema de geração de imagens utilizando gramáticas é menos explorado, e até o momento gramáticas paralelas (*L-systems*) (ROZENBERG; SALOMAA, 2001) é a forma de gramática mais utilizada. No entanto, este tipo de gramática é mais adequado para desenho de estruturas recursivas, como fractais. Acreditamos que, para o problema descrito neste projeto, grafos AND-OR são mais adequados.

Diferentemente de uma árvore, um grafo AND-OR não é necessariamente acíclico. De uma forma geral, os ciclos costumam ocorrer quando um nó possui mais de um nó pai ou quando existem arestas horizontais indicando relações entre nós de um mesmo nível.

No subgrafo referente aos ponteiros do relógio exibido na Figura 2.4, nota-se que ambos os nós AND “3 hands” e “2 hands” são pais dos nós OR “Hour hand” e “Minute hand”, pois os ponteiros das horas e dos minutos estão presentes nos dois tipos de arranjo de ponteiros. Ainda na Figura 2.4, as linhas horizontais tracejadas que ligam diversos nós do grafo indicam relações entre eles. Essas relações podem indicar distância, adjacência, oclusão, paralelismo, perpendicularidade, concentricidade, etc. entre dois nós (XU et al., 2008; ZHU; MUMFORD, 2007b).

A base para essas estruturas são os modelos gráficos conhecidos como campos aleatórios de Markov (JR; HART, 1971), que representam a estrutura de dependências entre um conjunto de variáveis aleatórias. Esse modelo gráfico é apropriado para imagens compostas por objetos que se relacionam, pois a direção da dependência é desconhecida ou irrelevante. Por esse motivo, não são suportadas dependências indiretas, portanto todas as relações entre objetos devem ser declaradas explicitamente por meio de arestas (JR; HART, 1971; ZHU; MUMFORD, 2007b).

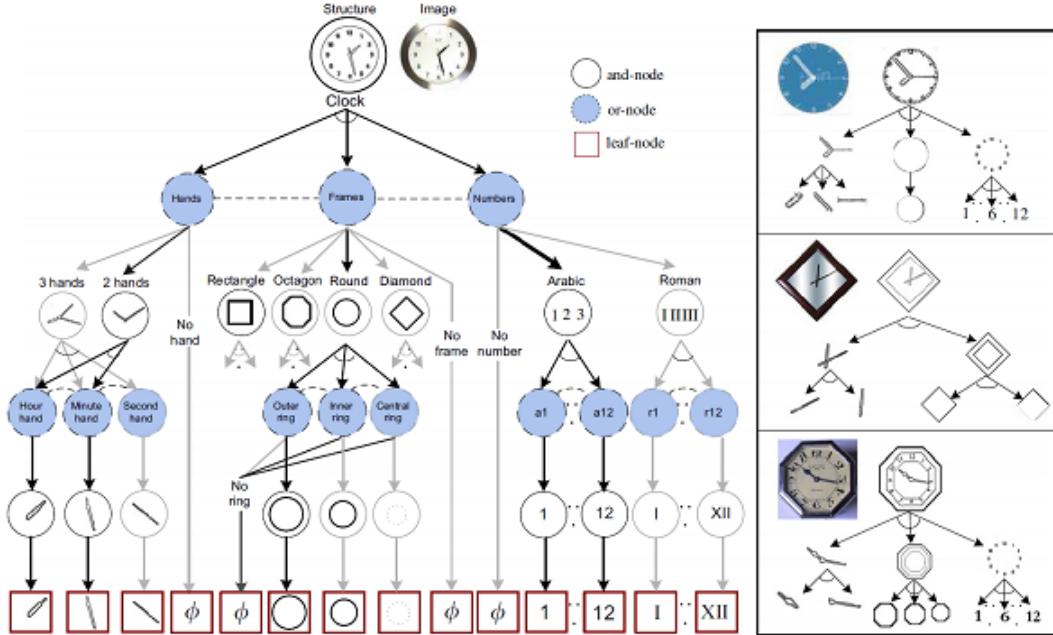


Figura 2.4: *Exemplo de grafo AND-OR de relógios de parede.* A imagem ilustra a estrutura geral do grafo e grafos sintáticos de exemplos de relógios. Fonte: (LIN et al., 2009).

As arestas horizontais acrescentam, portanto, informação de contexto entre os nós, tornando a gramática sensível ao contexto. Assim, um grafo AND-OR é uma união de uma árvore AND-OR, que provê a estrutura hierárquica, e um grafo de relações, que adiciona contexto à estrutura. As relações expandem a capacidade representativa dos grafos AND-OR em relação às árvores, permitindo que não apenas gramáticas livres de contexto, mas também gramáticas sensíveis ao contexto sejam descritas (XU et al., 2008; ZHU; MUMFORD, 2007b).

2.4 Geração de imagens

Uma imagem digital é uma representação do objeto físico que pode ser armazenada, manipulada e interpretada de acordo com as necessidades do interessado. Matematicamente, uma imagem pode ser descrita como uma função da intensidade do sinal retornado por um sensor. Em imagens digitais, considerando o domínio espacial, esta função pode ser representada por um vetor de valores discretos (GONZALEZ; WOODS; EDDINS, 2009).

O objetivo de definir matematicamente a imagem é a possibilidade de

manipular o seu conteúdo a fim de transformá-la ou retirar dela informações importantes. Ao vasto conjunto de operações que pode-se aplicar no vetor que representa uma imagem denominamos processamento de imagem.

Spline é um conceito que define uma tira flexível usada para produzir suavização da curva por meio de pontos de controle (FOLEY, 1996), ou seja, um conjunto de pontos irá indicar o formato da curva. Existem duas maneiras de implementação de *splines*: a interpolação e a aproximação. Na primeira, a curva deve passar pelos pontos de controle, já a segunda não há uma necessidade disso ocorrer.

Em um *spline*, quanto maior o grau do polinômio, mais flexibilidade a curva ganha, porém a quantidade de cálculos necessários para o traçado da curva também aumenta. A elevação na quantidade de pontos de uma dada curva irá acrescentar complexidade no polinômio que a determina. Para contornar esse problema utilizam-se seções (pequenas regiões) com polinômios de graus menores na composição dessas curvas, sendo necessário apenas fazer uma simples junção das extremidades de cada curva.

A definição, modificação e manipulação das curvas *splines* são realizadas com operações em seus pontos de controle. Assim, a reestruturação da forma da curva ocorre com o reposicionamento de algum ou todos pontos de controle (FOLEY, 1996). A Figura 2.5 mostra a influência da alteração no posicionamento dos pontos de controle.

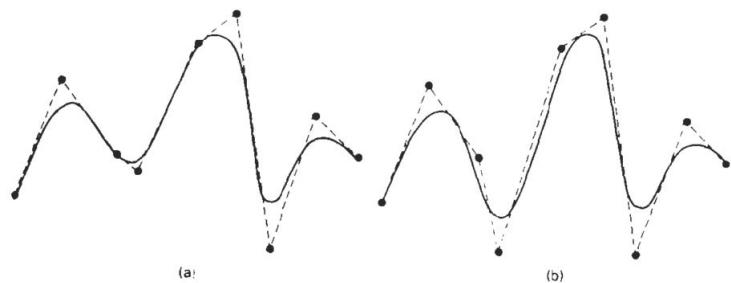


Figura 2.5: Exemplos de modificação nos pontos de controle em *b-splines*. Retirado de (FOLEY, 1996)

2.5 Trabalhos correlatos

Yang et al. (2009) apresenta um sistema interativo para manipulação de expressões faciais em *sketches*. Primeiramente construiu-se um banco de dados contendo os posicionamentos de cada ponto rastreado com base nos vídeos de

entrada. Então, a partir de vídeos contendo a movimentação foram adicionados os dados relativos a movimentos. Com esse banco de dados construído foram sorteadas as medidas que compõem a face inicial. Por fim, o usuário desloca um ponto de controle de uma das curvas, e a partir dessa manipulação o sistema propaga o movimento de acordo com o banco de dados. Esse processo é ilustrado na Figura 2.6.

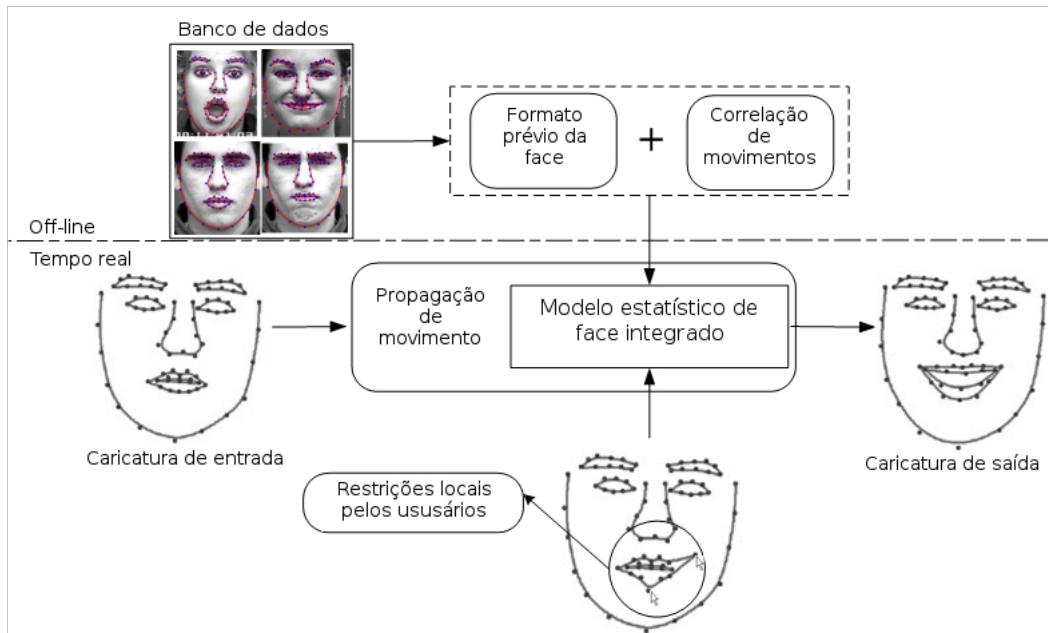


Figura 2.6: Funcionamento do sistema. Adaptado de (YANG et al., 2009)

Em (DECARLO; METAXAS; STONE, 1998) é proposto um método para variação nos modelos utilizando medidas antropométricas. Primeiramente são sorteadas medidas antropométricas que compõem o modelo facial. Depois, essas medidas são adaptadas para seguirem o modelo de restrição da superfície. Assim consegue-se gerar uma série de modelos variados, como pode ser visto na Figura 2.7.



Figura 2.7: Modelos masculinos variados (DECARLO; METAXAS; STONE, 1998)

Capítulo 3

Metodologia

As caricaturas para representar as emoções faciais foram construídas considerando que cada parte do rosto é composta por curvas *spline*. As medidas antropométricas foram empregadas para definir a dimensão de cada curva e o posicionamento correto dos seus pontos de controle. A quantidade de curvas e o posicionamento das mesmas para compor uma face inicial neutra foram definidos com base nas descrições das unidades de ação (*FACS*). Os conceitos de *FACS* também foram empregados para reposicionar os pontos de controle das curvas conforme a emoção. Esse modelo contendo um conjunto de curvas é armazenado e representado pela gramática. A partir da representação desenham-se as curvas *splines* para reproduzir as expressões faciais desejadas. Esse esquema é representado na Figura 3.1

3.1 Representação da face

A representação da face foi realizada por meio da construção de um modelo antropométrico. Esse modelo deixou de ser parte integrante do programa de síntese de expressões faciais para integrar a gramática. Esta escolha permite mais flexibilidade do modelo, como discutido na seção 3.4.

3.1.1 Modelo antropométrico

As medidas antropométricas variam de acordo com o grupo étnico e a idade dos indivíduos. Para constituição do modelo antropométrico foram considerados os dados para uma população composta de indivíduos caucasianos norte-americanos saudáveis na faixa etária de 19-25 anos, idade na qual ocorre a estabilização das medidas para adultos (FARKAS, 1994).

As medidas antropométricas referem-se às distâncias entre os *landmarks*

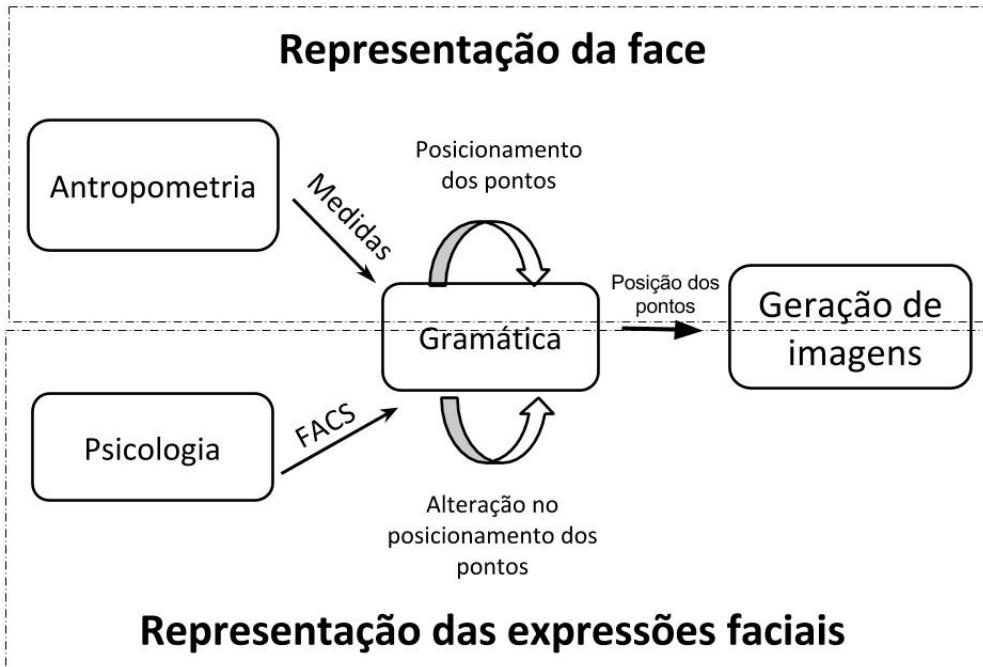


Figura 3.1: Esquematização da metodologia.

e não às posições dos mesmos. Portanto, a fim de desenhar uma face neutra inicial, foi necessário calcular o posicionamento de cada um desses pontos no espaço bidimensional da imagem com base nas relações de distância entre dois pontos, fornecidas pelas mensurações antropométricas. Finalmente, os *landmarks* foram utilizados na definição dos pontos de controle que compõem as curvas que formam cada parte do rosto.

Das medidas descritas na seção 2.1 foram utilizadas as medidas de menor distância entre dois pontos, primeiramente porque são as que contêm o maior número de dados antropométricos, e também por serem as distâncias que melhor se enquadram em um espaço bidimensional.

Vários dos pontos definidos neste trabalho não possuem *landmarks* equivalentes na antropometria. Para definir a posição desses pontos foi necessário fazer uma aproximação com base nas medidas antropométricas disponíveis, considerando também a simetria da face. Isto pode ser exemplificado com a definição do ponto *p1* (Figura 3.2). Como a sobrancelha segue a linha do olho, a coordenada *x*(posicionamento horizontal) foi definida como sendo a mesma linha do ponto *en*. Uma descrição mais detalhada para cada ponto é

encontrada na Tabela 3.1

Tabela 3.1: Inferências feitas a partir da antropometria.

Parte do Rosto	Pontos	Considerado
Sobrancelha	$p1$ e $p2$	A distância entre as extremidades das duas sobrancelhas é média da diferença de ex_ex e ft_ft
		A distância entre as extremidades internas das sobrancelhas é igual a en
		O nível mais baixo da sobrancelha tem a mesma altura que n
	os	Fica na metade da distância de $p1$ e $p2$
Olho	ps e pi	Ficam na metade da distância de en_ex
	ex e en	Ficam na mesma altura
Pálpebra inferior	$pp1$	Fica logo abaixo de en
	$pp2$	Fica a 90% de en_ex em relação a en
	$pp1$ e $pp2$	Ficam a 15% de pi_or em relação a pi
Nariz	n	Fica na metade de en_en
	mf	Fica 5% de n_sn abaixo do olho (en)
	al	Fica 90% de n_sn abaixo de n
Lábios	ls e sto	Ficam na mesma linha vertical que n
Enrugamento do olho	$po1_po3$, $po4_po6$ e $po7_po9$	As rugas possuem em média 2 mm
	$po1_po3$, $po4_po6$ e $po7_po9$	As rugas esticam até 2+1 mm
	$po4$	$po4$ fica na altura de ex
	$po1$	$po1$ fica 1mm acima de $po4$
	$po9$	$p9$ fica 1mm abaixo de $po4$
Enrugamento da testa	$pt3$	Fica no meio de ft_ft e tr_g (no meio da testa)
	$pt1$	Fica na mesma posição que ft (no final da testa)
	$pt2$	Fica à um quarto a esquerda de ft_ft e tr_g (um quarto da testa)

Tabela 3.1: Inferências feitas a partir da antropometria.

Parte do Rosto	Pontos	Considerado
Enrugamento da testa	<i>pt4</i>	Fica a um quarto a direita de <i>ft_ft</i> e <i>tr_g</i> (um quarto da testa)
	<i>pt5</i>	Fica na mesma posição que <i>ft(d)</i> (no final da testa)
	<i>pt3(c)</i>	Fica a 25% de <i>tr_g</i> em relação a <i>tr</i>
	<i>pt3(b)</i>	Fica a 75% de <i>tr_g</i> em relação a <i>tr</i>
Contorno	<i>zy</i>	Possui o mesmo posicionamento vertical que <i>or</i>
	<i>eu</i>	Possui o mesmo posicionamento vertical que <i>tr</i>
	<i>gn, tr e v</i>	Possui o mesmo posicionamento horizontal que <i>n</i>
	<i>t</i>	Possui o posicionamento vertical na metade de <i>pi_or</i>
Auxiliares	<i>n</i>	Fica na metade da distância de <i>en_en</i> e <i>mf_mf</i>
	<i>se</i>	Fica na mesma linha que <i>n</i>

Além deste ponto, os pontos *p2* da sobrancelha, *pas1* e *pas2* da pálpebra inferior, bem como os pontos do enrugamento não apresentam nenhuma descrição pela antropometria e, por isso, foram inferidos com base em outras medidas antropométricas para que fosse possível a representação de suas curvas.

Para o cálculo das posições dos pontos, primeiramente definiu-se um ponto inicial, o *landmark v*, no qual são baseados os demais pontos. O ponto *v* se posiciona bem no topo da cabeça e foi definido como sendo a extremidade superior e o centro vertical da imagem (Figura 3.2).

Para mostrar como é computada a posição de um dado ponto pode-se observar o ponto *sn*, desenhado na parte inferior do nariz. Sua coordenada *y* (posicionamento vertical) foi definida com o valor da coordenada *y* (posicionamento vertical) do ponto *n* somado à mensuração antropométrica *n-sn*. Sua coordenada *x* (posicionamento horizontal) é igual à coordenada *x* (posicionamento horizontal) de *n*. Dessa forma, são definidos todos os pontos de controle de todas as curvas que formam o rosto.

As Tabelas 3.2 a 3.13 expõem os cálculos das posições de cada *landmark* com base nas distâncias entre os pontos descritos pela antropometria. O posicionamento de um *landmark* é definido de acordo com as medidas antro-

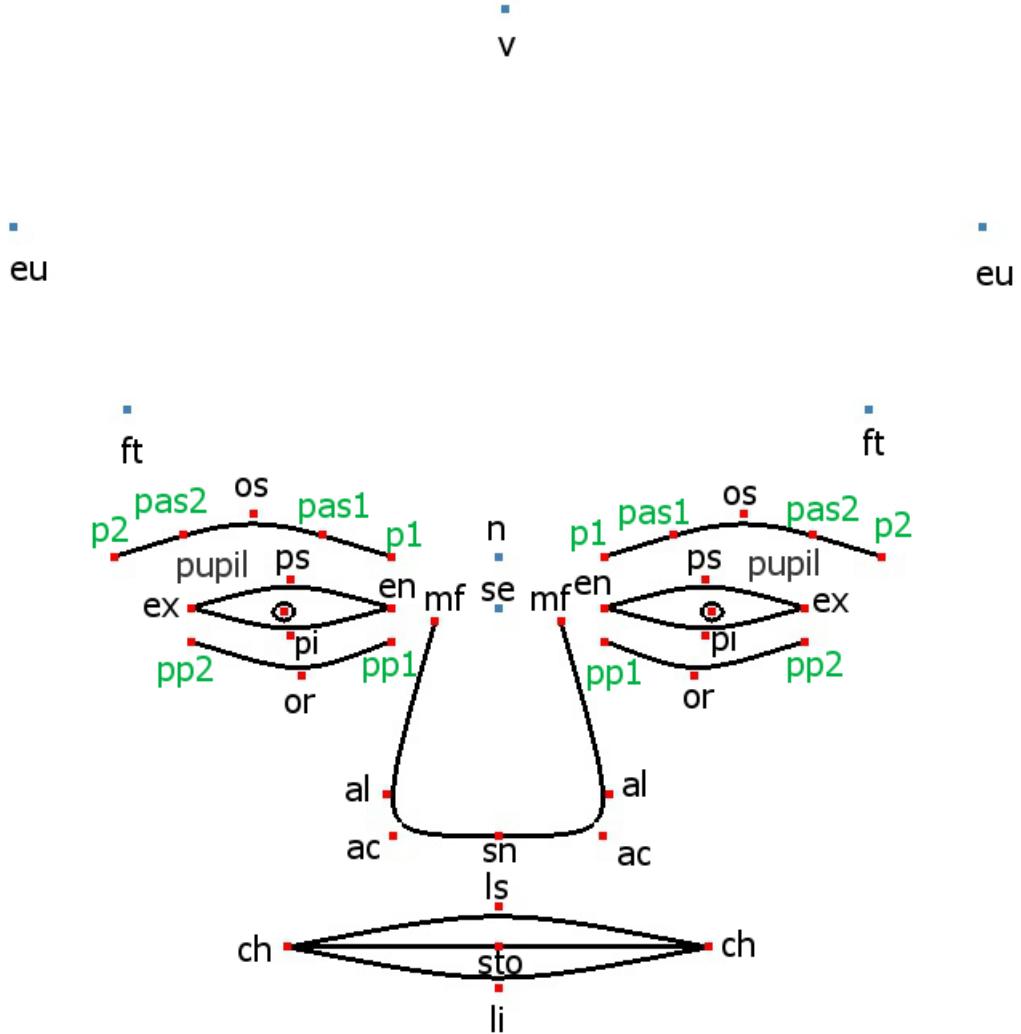


Figura 3.2: *Landmarks* utilizados para a construção dos *sketches*. As marcações em vermelho representam os pontos de controle utilizados nas curvas, e as marcações em azul são pontos empregados como auxiliares para posicionamento de outros pontos de controle. Além disso, foram criados pontos com letras verdes para construir as curvas.

pométricas e a posição de outros pontos da antropometria. A distância entre dois pontos é denotada pelo símbolo: traço inferior ($_$). Deste modo se um traço inferior está entre dois pontos, significa que está sendo considerada a distância entre esses dois pontos. Por exemplo, ps_pi indica a distância entre os *landmarks* *ps* e *pi*.

A equação 3.1 mostra como é calculada a posição de cada ponto de acordo

com os dados obtidos pela antropometria. Aonde $P1$ é o posicionamento do ponto a se calculado; $P2$ e $P3$ são as posições de outros dois *landmarks* que tomamos como base; α é uma porcentagem e $P2-P3$ representa a distância entre entre os pontos.

$$P1 = P2 + \alpha \cdot |P2 - P3| \quad (3.1)$$

O posicionamento de um ponto é relativo à posição de outro somada a distância entre eles; caso essa distância seja uma parte da distância total o elemento porcentagem estará entre zero e um, representando a porcentagem da distância total; caso contrário será um para representar a distância total.

A seguir há uma descrição da construção de cada parte do rosto representada. Para cada parte são apresentados uma tabela com os cálculos do posicionamento horizontal (para X) e vertical (para Y).

Olhos

Cada olho é composto por duas curvas de três pontos: uma superior (*en-ps-ex*) e outra inferior (*en-pi-ex*), como mostrado na Tabela 3.2.

Tabela 3.2: Cálculo do posicionamento dos *landmarks* para os olhos.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Olho esquerdo	en	$n-0,5 \cdot (\text{en_en})$	$n+v_{\text{en}}-v_n$
	ex	$n-0,5 \cdot 2.5cm \cdot (\text{en_en}) - \text{ex_en}$	$n+v_{\text{en}}-v_n$
	ps	$\text{en}-0,5 \cdot (\text{ex_en})$	$\text{en}-0,5 \cdot (\text{ps_pi})$
	pi	$\text{en}-0,5 \cdot (\text{ex_en})$	$\text{en}+0,5 \cdot (\text{ps_pi})$
Olho direito	en	$n+0,5 \cdot (\text{en_en})$	$n+v_{\text{en}}-v_n$
	ex	$n+0,5 \cdot (\text{en_en}) - \text{ex_en}$	$n+v_{\text{en}}-v_n$
	ps	$\text{en}+0,5 \cdot (\text{ex_en})$	$\text{en}-0,5 \cdot (\text{ps_pi})$
	pi	$\text{en}+0,5 \cdot (\text{ex_en})$	$\text{en}+0,5 \cdot (\text{ps_pi})$

Sobrancelhas

A sobrancelha direita é construída com os pontos *pas1*, *os* e *pas2*. E três pontos simétricos são usados para desenhar a sobrancelha esquerda (*pp1-os-pp2*), como mostrado na Tabela 3.3.

Tabela 3.3: Cálculo do posicionamento dos *landmarks* para sobrancelha.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Sobrancelha esquerda	pas1	en	n
	os	$p1-0,5 \cdot (p1+p2)$	$n+g_sn-n_sn$
	pas2	$en+0,5 \cdot (ft_ft-ex_ex)$	n
Sobrancelha direita	pas1	en	n
	os	$p1+0,5 \cdot (p1+p2)$	$n+g_sn-n_sn$
	pas2	$en-0,5 \cdot (ft_ft-ex_ex)$	n

Pálpebras

As pálpebras são formadas por três pontos: *pp1*, *or* e *pp2*. Cada posição dessas varia conforme o lado esquerdo e direito do rosto, como mostrado na Tabela 3.4.

Tabela 3.4: Cálculo do posicionamento dos *landmarks* para as pálpebras.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Pálpebra inferior esquerda	pp1	en	$pi+0,15 \cdot (pi_or)$
	or	$en-0,5 \cdot ex_en$	$pi+pi_or$
	pp2	$en-ex_en$	$pi+0,15 \cdot (pi_or)$
Pálpebra inferior direita	pp1	en	$pi+0,15 \cdot (pi_or)$
	or	$en+0,5 \cdot ex_en$	$pi+pi_or$
	pp2	$en+ex_en$	$pi+0,15 \cdot (pi_or)$

Nariz

O nariz é constituído de duas curvas simétricas interligadas, uma do lado esquerdo com os pontos *mf(e)*, *al(e)*, *ac(e)* e *sn* e outra do lado direito com os pontos *mf(d)*, *al(d)*, *ac(d)* e *sn*, como mostrado na Tabela 3.5.

Tabela 3.5: Cálculo do posicionamento dos *landmarks* para o nariz.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Nariz	mf(e)	n-0,5 · mf_mf	n+0,5 · n_sn
	mf(d)	n+0,5 · mf_mf	n+0,5 · n_sn
	al(e)	n-0,5 · al_al	n+0,85 · n_sn
	al(d)	n+0,5 · al_al	n+0,85 · n_sn
	ac(e)	n-0,5 · ac_ac	n+n_sn
	ac(d)	n+0,5 · ac_ac	n+n_sn
	sn	n	n+n_sn

Lábios

Cada lábio contém duas curvas com três pontos. Para o lábio superior as curvas são $ch(e)\text{-}ls\text{-}ch(d)$ e $ch(e)\text{-}sto\text{-}ch(d)$. Já para o direito são $ch(e)\text{-}sto\text{-}ch(d)$ e $ch(e)\text{-}li\text{-}ch(d)$. Deve-se atentar que duas curvas são repetidas para permitir a abertura dos lábios, como mostrado na Tabela 3.6.

Tabela 3.6: Cálculo do posicionamento dos *landmarks* para os lábios.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Lábios	ls	n	N+n_sto-ls_sto
	sto	n	ls+ls_sto
	li	n	ls+2 · ls_sto
	ch(e)	n-ch_sto	ls+ls_sto
	ch(d)	n+ch_sto	ls+ls_sto

Pupila

As pupilas são embasadas em circunferências conforme uma posição inicial e um raio, como mostrado na Tabela 3.7.

Tabela 3.7: Cálculo do posicionamento dos *landmarks* para as pupilas.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Pupila esquerda	pupil	se-pupil_se	or+pupil_or
	raio	$0,4 \cdot ps_pi$	
Pupila direita	pupil	se+pupil_se	or+pupil_or
	raio	$0,4 \cdot ps_pi$	

Enrugamento dos cantos externos do olhos

As rugas nos cantos externos dos olhos foram desenhadas utilizando três curvas com três pontos, como mostrado na Tabela 3.8.

Tabela 3.8: Cálculo do posicionamento dos *landmarks* para os enrugamentos dos olhos.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Enrugamento do olho esquerdo	po1	po4	po4-2
	po2	po5	po1-2
	po3	po6	po2-2
	po4	en-2	en
	po5	po4-2	po4
	po6	po5-2	po4
	po7	po4	po4+2
	po8	po5	po7+2
	po9	po6	po8+2
Enrugamento do olho direito	po1	po4	po4-2
	po2	po5	po1-2
	po3	po6	po2-2
	po4	en+2	en
	po5	po4-2	po4
	po6	po5-2	po4
	po7	po4	po4+2
	po8	po5	po7+2
	po9	po6	po8+2

Enrugamento da testa

As rugas na testa foram desenhadas utilizando três curvas com cinco pontos, como mostrado na Tabela 3.9.

Tabela 3.9: Cálculo do posicionamento dos *landmarks* para enrugamento da testa.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Enrugamento da testa	pt1(c)	$pt3(c) - (0.5 \cdot ft_ft)$	$pt3(c)$
	pt2(c)	$pt3(c) - (0.25 \cdot ft_ft)$	$pt3(c) + (0.1 \cdot tr_g)$
	pt3(c)	n	$tre + (0.5 \cdot tr_g)$
	pt4(c)	$pt3(c) + (0.25 \cdot ft_ft)$	$pt3(c) + (0.1 \cdot tr_g)$
	pt5(c)	$pt3(c) + (0.5 \cdot ft_ft)$	$pt3(c)$
	pt1(m)	$pt3(m) - (0.5 \cdot ft_ft)$	$pt3(m)$
	pt2(m)	$pt3(m) - (0.25 \cdot ft_ft)$	$pt3(m) + (0.1 \cdot tr_g)$
	pt3(m)	n	$tre + (0.5 \cdot tr_g)$
	pt4(m)	$pt3(m) + (0.25 \cdot ft_ft)$	$pt3(m) + (0.1 \cdot tr_g)$
	pt5(m)	$pt3(m) + (0.5 \cdot ft_ft)$	$pt3(m)$
	pt1(b)	$pt3(b) - (0.5 \cdot ft_ft)$	$pt3(b)$
	pt2(b)	$pt3(b) - (0.25 \cdot ft_ft)$	$pt3(b) + (0.1 \cdot tr_g)$
	pt3(b)	n	$tre + (0.5 \cdot tr_g)$
	pt4(b)	$pt3(b) + (0.25 \cdot ft_ft)$	$pt3(b) + (0.1 \cdot tr_g)$
	pt5(b)	$pt3(b) + (0.5 \cdot ft_ft)$	$pt3(b)$

Enrugamento dos cantos internos das sobrancelhas

As rugas nos cantos internos das sobrancelhas foram desenhadas utilizando três curvas com três pontos, como mostrado na Tabela 3.11.

Tabela 3.10: Cálculo do posicionamento dos *landmarks* para enrugamento das bochechas.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Enrugamento sobrancelhas	pis1(e)	$p1(e) + 0.84 \cdot (n - p1(e))$	$n - 0.06 \cdot (n_se)$
	pis2(e)	$p1(e) + 0.87 \cdot (n - p1(e))$	$n + 0.02 \cdot (n_se)$
	pis3(e)	$p1(e) + 0.62 \cdot (n - p1(e))$	$n + 0.46 \cdot (n_se)$
	pis1(c)	n	$pis1(e)$
	pis2(c)	n	$pis3(e)$

Tabela 3.10: Cálculo do posicionamento dos *landmarks* para enrugamento das bochechas.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Enrugamento sobrancelhas	pis3(c)	n	$\text{pis3}(c)-0.5 \cdot (\text{pis3}(c)-\text{pis1}(c))$
	pis1(d)	$\text{p1}(d)+0.84 \cdot (\text{n}-\text{p1}(d))$	$\text{pis1}(e)$
	pis2(d)	$\text{p1}(d)+0.87 \cdot (\text{n}-\text{p1}(d))$	$\text{pis2}(e)$
	pis3(d)	$\text{p1}(d)+0.62 \cdot (\text{n}-\text{p1}(d))$	$\text{pis3}(e)$

Enrugamento do nariz

As rugas do nariz foram desenhadas utilizando com sete curvas com três pontos, como mostrado na Tabela 3.11.

Tabela 3.11: Cálculo do posicionamento dos *landmarks* para os enrugamentos do nariz.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Enrugamento nariz sobrancelha	psn1(c)	mf(e)	$\text{n}+0.25 \cdot \text{n_se}$
	psn2(c)	n	$\text{n}+0.5 \cdot \text{n_se}$
	psn3(c)	mf(d)	$\text{n}+0.25 \cdot \text{n_se}$
	psn1(m)	mf(e)	$\text{n}+0.5 \cdot \text{n_se}$
	psn2(m)	n	$\text{n}+0.6 \cdot \text{n_se}$
	psn3(m)	mf(d)	$\text{n}+0.5 \cdot \text{n_se}$
	psn1(b)	mf(e)	$\text{n}+0.75 \cdot \text{n_se}$
	psn2(b)	n	se
	psn3(b)	mf(d)	$\text{n}+0.75 \cdot \text{n_se}$
Enrugamento nariz esquerda	pn1(e)	$\text{al}(e)+0.22 \cdot \text{al_al}$	$\text{nY}+0.27 \cdot \text{n_sn}$
	pn2(e)	$\text{al}(e)+0.24 \cdot \text{al_al}$	$\text{nY}+0.34 \cdot \text{n_sn}$
	pn3(e)	$\text{al}(e)+0.4 \cdot \text{al_al}$	$\text{nY}+0.39 \cdot \text{n_sn}$
	pn4(e)	$\text{al}(e)+0.2 \cdot \text{al_al}$	$\text{nY}+0.36 \cdot \text{n_sn}$
	pn5(e)	$\text{al}(e)+0.26 \cdot \text{al_al}$	$\text{nY}+0.39 \cdot \text{n_sn}$
	pn6(e)	$\text{al}(e)+0.28 \cdot \text{al_al}$	$\text{nY}+0.46 \cdot \text{n_sn}$
Enrugamento nariz direita	pn1(d)	$\text{al}(d)+0.22 \cdot \text{al_al}$	$\text{nY}+0.27 \cdot \text{n_sn}$
	pn2(d)	$\text{al}(d)+0.24 \cdot \text{al_al}$	$\text{nY}+0.34 \cdot \text{n_sn}$

Tabela 3.11: Cálculo do posicionamento dos *landmarks* para os enrugamentos do nariz.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Enrugamento nariz direita	pn3(d)	$al(d)+0.4 \cdot al_al$	$nY+0.39 \cdot n_sn$
	pn4(d)	$al(d)+0.2 \cdot al_al$	$nY+0.36 \cdot n_sn$
	pn5(d)	$al(d)+0.26 \cdot al_al$	$nY+0.39 \cdot n_sn$
	pn6(d)	$al(d)+0.28 \cdot al_al$	$nY+0.46 \cdot n_sn$

Enrugamento das bochechas

As rugas da bochecha foram desenhadas utilizando duas curvas com cinco pontos, como mostrado na Tabela 3.12.

Tabela 3.12: Cálculo do posicionamento dos *landmarks* para enrugamento das bochechas.

Parte do Rosto	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Enrugamento da bochecha	psn1(e)	$al(e)+0.1 \cdot (al(e)-ch(e))$	$mf(e)+0.6 \cdot (sn-mf(e))$
	pb2(e)	$al(e)+0.5 \cdot (al(e)-ch(e))$	$pb1(e)+0.15 \cdot (ch(e)-pb1(e))$
	pb3(e)	$al(e)-0.8 \cdot (al(e)-ch(e))$	$pb1(e)+0.35 \cdot (ch(e)-pb1(e))$
	pb4(e)	$al(e)-1.1 \cdot (al(e)-ch(e))$	$pb1(e)+0.55 \cdot (ch(e)-pb1(e))$
	pb5(e)	$al(e)-1.1 \cdot (al(e)-ch(e))$	$pb1(e)+0.8 \cdot (ch(e)-pb1(e))$
	pb1(e)	$al(e)+0.1 \cdot (al(e)-ch(e))$	$mf(e)+0.6 \cdot (sn-mf(e))$
	pb2(e)	$al(e)+0.5 \cdot (al(e)-ch(e))$	$pb1(e)+0.15 \cdot (ch(e)-pb1(e))$
	pb3(e)	$al(e)-0.8 \cdot (al(e)-ch(e))$	$pb1(e)+0.35 \cdot (ch(e)-pb1(e))$
	pb4(e)	$al(e)-1.1 \cdot (al(e)-ch(e))$	$pb1(e)+0.55 \cdot (ch(e)-pb1(e))$
	pb5(e)	$al(e)-1.1 \cdot (al(e)-ch(e))$	$pb1(e)+0.8 \cdot (ch(e)-pb1(e))$

Auxiliares

Os pontos auxiliares foram utilizados para embasar o cálculo dos demais pontos que formam as curvas, como mostrado na Tabela 3.13.

Tabela 3.13: Cálculo do posicionamento dos *landmarks*.

Pontos	Ponto	Cálculo da posição horizontal	Cálculo da posição vertical
Auxiliares	n	v+v_n	v
	ft(e)	n-0,5 · ft_ft	tr+0,5 · tr_g
	ft(d)	n+0,5 · ft_ft	tr+0,5 · tr_g
	v	0	0
	se	n	en

3.2 *Splines*

O desenho das curvas que representam as partes do rosto é realizado por meio da técnica de representação de curvas *b-splines*. Cada curva é caracterizada pelo vetor de pontos de controle (*landmarks* antropométricos) que a formam. No traçado das curvas foram utilizados polinômios cúbicos, pois estes oferecem um compromisso razoável entre flexibilidade e velocidade de processamento (FOLEY, 1996).

A maioria das curvas foi gerada com três pontos de controle, conforme mostrado na Figura 3.2. Vale ser frisado que foi necessária a adição de dois pontos na sobrancelha (*pas1* e *pas2*) para que o resultado da movimentação fosse o mesmo que o descrito nas unidades de ação 1, 2 e 4 (Figure 2.3). Esses dois pontos foram estabelecidos por meio do método de subdivisão em *splines* descrito em (HEARN; BAKER, 1997), no qual são adicionados mais pontos na curva sem alterá-la. As coordenadas do novo ponto são definidas como a média das coordenadas de dois pontos da curva.

3.3 Representação de emoções

Para contemplar a tarefa de reproduzir as expressões faciais nos *sketchs*, é necessário simular os movimentos provocados pelas emoções na face, e aplicar esse efeito nas curvas que compõem o rosto. A seguir é descrito como *FACS* e *splines* foram empregados para obter uma representação em forma de *sketchs* das expressões faciais. Esse modelo de representação das expressões faciais também foi incorporado às gráficas.

Os movimentos reproduzidos neste trabalho retratam a descrição mais comum de representação das expressões faciais. Apesar de poderem existir variações nessas expressões, apenas uma variação foi considerada.

Foram realizadas tentativas de construção de modelos variados para americanos afro-descendentes e indivíduos com idades inferiores a 19 anos, porém nem todas as medidas necessárias foram obtidas como no modelo para um americano caucasiano de 19 a 25 anos.

3.3.1 O modelo de representação das expressões faciais por meio de FACS

Primeiramente foram definidas quais expressões faciais seriam representadas nesse trabalho. Foram consideradas as expressões de satisfação, tristeza, raiva, aversão, medo e surpresa, por serem universais, conforme descrito no capítulo 1. Em seguida, estudou-se quais unidades de ação formam essas expressões faciais. A aparência resultante da influência que cada unidade de ação exerce no rosto foi seguida conforme as descrições encontradas em Ekman e Friesen (1976) e Ekman, Friesen e Hager (2002).

Apesar da ampla descrição de cada unidade de ação, as ciências comportamentais não estão preocupadas em medir exatamente com qual intensidade se varia um músculo em uma dada unidade de ação. Há apenas uma descrição de qual músculo se move, e a implicação dessa ação nas partes dos rostos. Por exemplo, quando se varia o músculo equivalente a unidade de ação 2, há uma descrição retratando que existe um arqueamento da sobrancelha, um esticamento da pálpebra superior e um possível enrugamento da testa. Porém, não existe nenhuma descrição do grau de arqueamento da sobrancelha, ou qual é a medida total de enrugamento da testa. O presente estudo adotou a abordagem de trabalhar apenas com as descrições e as imagens fornecidas para gerar o novo posicionamento de cada ponto antropométrico.

Para contornar o problema, foi preciso encontrar uma maneira alternativa de realizar as variações nos pontos de controle sem gerar inconsistências. Para isso definiu-se um envoltório no entorno de cada parte do rosto. Assim os pontos de controle variam seus posicionamentos com base na medida definida por esse envoltório.

O envoltório de variação é definido com o dobro do tamanho do retângulo que envolve a curva, como ilustrado na Figura 3.3. Assim, o reposicionamento de um ponto se baseia em variar um percentual da medida do envoltório correspondente, pois a geração dinâmica das expressões permite que o usuário defina a porcentagem de emoção que deve ser representada (de zero a cem por cento para cada tipo de emoção).

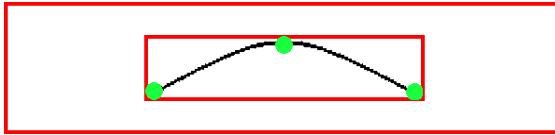


Figura 3.3: A linha vermelha interna mostra a reção entre o tamanho do envoltório e a curva contida. A linha externa delimita um envoltório que é o dobro do tamanho. A curva em preto corresponde a uma parte do rosto e em verde seus pontos de controle.

A partir da abordagem apresentada, é possível gerar desde expressões mais sutis até expressões totais a partir das variações nos pontos de controle envolvidos em cada unidade de ação da emoção facial solicitada. As transições entre duas expressões faciais diferentes são feitas por meio da passagem pela expressão neutra. Exemplos de porcentagens diferentes da mesma expressão são apresentados na Figura 4.1.

Nenhum modelo de representação da movimentação dos músculos foi utilizado, restringiu-se apenas a utilizar a aparência produzida nos movimentos dos músculos, que está descrita em cada unidade de ação. Logo, a real variação (porcentagem do envoltório) de cada ponto de controle das curvas foi realizada de forma empírica. Considerando as diferentes intensidades (EKMAN et al., 1987; EKMAN, 2007) que cada expressão facial pode produzir, cada unidade de ação oscila de zero até a porcentagem máxima.

A solução adotada para representar as transições que ocorrem entre uma expressão facial neutra e as demais expressões faciais foi variar a intensidade de cada movimento entre zero e o máximo definido para determinada ação. Portanto, durante a transição ocorre passagem por todas as intensidades de cada emoção.

Outra dificuldade encontrada na representação das emoções surgiu com o fato da antropometria não descrever nenhum dado sobre o enrugamento da face. Portanto, foi necessário fazer uma estimativa da localização desses pontos para geração dessas rugas e das variações que podem sofrer.

Para fornecer maior impressão de realidade do enrugamento, as rugas se movem conforme a movimentação da parte do rosto na unidade de ação. Por exemplo, as rugas que surgem na testa sofrem a mesma elevação vertical da sobrancelha levantando-se.

Outro efeito adotado foi o aparecimento gradual das rugas. Este efeito é importante visto que a profundidade dessas marcas de expressão aumentam conforme o movimento do músculo. Por isso, para dar um tom de aparecimento das rugas, a espessura das curvas que as formam, varia entre zero e a espessura padrão adotada para as demais linhas do desenho.

A equação 3.2 representa a variação no posicionamento dos pontos de acordo com o envoltório. No qual, PE é a posição do *landmark* em uma expressão facial, PN corresponde a posição do mesmo *landmarks* na face neutra, α é uma porcentagem, I é a intensidade a emoção e $PN2-PN3$ é o tamanho do envoltório.

$$PE = PN + \alpha \cdot I \cdot |PN2 - PN3| \quad (3.2)$$

A posição do ponto em uma dada emoção é calculada com base nas coordenadas x e y da face neutra. Na equação 3.2, o elemento *alpha* é o percentual necessário e variação para unidade de ação atual de acordo com o envoltório para a emoção que está sendo representada, I mostra o quanto intensa será a emoção e o tamanho do envoltório ($PN2-PN3$) é a maior distância para x ou para y da parte do rosto.

Na Tabela 3.14 é apresentado como este cálculo das posições se estabelece em cada unidade de ação. A relação entre as unidades de ação e as expressões faciais, bem como a descrição de cada movimento estão descrita na seção 2.2. As letras em cada elemento da equação representam os *landmarks* antropométricos: e=esquerda, d=direita, c=acima, m=meio, b=abaixo e i=intensidade.

Tabela 3.14: Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.

Unidade de ação	Parte do rosto	Equação
1	Sobrancelha esquerda	$p1eX=p1eX+0.07 \cdot i \cdot (p2eX-p1eX)$
		$p1eY=p1eY+0.09 \cdot i \cdot (p1eY-oseY)$
		$pas1eY=pas1eY+0.035 \cdot i \cdot (p1eY-oseY)$
		$oseY=oseY-0.02 \cdot i \cdot (p1eY-oseY)$
	Sobrancelha direita	$p1dX=p1eX-0.07 \cdot i \cdot (p2dX-p1dX)$
		$p1dY=p1dY+0.09 \cdot i \cdot (p1dY-osdY)$
		$pas1dY=pas1dY+0.035 \cdot i \cdot (p1dY-osdY)$
		$osdY=osdY-0.02 \cdot i \cdot (p1dY-osdY)$
	Enrugamentos Testa central	$pt2Y=pt2Y+(i-1) \cdot (pt3Y-pt2Y)$
		$pt3Y=pt3Y+(i-1) \cdot (pt3Y-pt2Y)$
		$pt4Y=pt4Y+(i-1) \cdot (pt3Y-pt2Y)$
2	Sobrancelha esquerda	$p1eX=p1eX-0.01 \cdot i \cdot (p2eX-p1eX)$
		$p2eX=p2eX-0.01 \cdot i \cdot (p2eX-p1eX)$
		$pas1eX=pas1eX-0.01 \cdot i \cdot (p2eX-p1eX)$

Tabela 3.14: Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.

Unidade de ação	Parte do rosto	Equação
2	Sobrancelha esquerda	$\text{pas2eX} = \text{pas2eX} - 0.01 \cdot i \cdot (\text{p2eX} - \text{p1eX})$
		$\text{oseX} = \text{oseX} - 0.01 \cdot i \cdot (\text{p2eX} - \text{p1eX})$
		$\text{p1eY} = \text{p1eY} + 0.01 \cdot i \cdot (\text{p1eY} - \text{oseY})$
		$\text{p2eY} = \text{p2eY} + 0.015 \cdot i \cdot (\text{p1eY} - \text{oseY})$
		$\text{pas1eY} = \text{pas1eY} + 0.015 \cdot i \cdot (\text{p1eY} - \text{oseY})$
		$\text{pas2eY} = \text{pas2eY} + 0.06 \cdot i \cdot (\text{p1eY} - \text{oseY})$
		$\text{oseY} = \text{oseY} - 0.025 \cdot i \cdot (\text{p1eY} - \text{oseY})$
	Sobrancelha direita	$\text{p1dX} = \text{p1dX} + 0.01 \cdot i \cdot (\text{p2dX} - \text{p1dX})$
		$\text{p2dX} = \text{p2dX} + 0.01 \cdot i \cdot (\text{p2dX} - \text{p1dX})$
		$\text{pas1dX} = \text{pas1dX} + 0.01 \cdot i \cdot (\text{p2dX} - \text{p1dX})$
		$\text{pas2dX} = \text{pas2dX} + 0.01 \cdot i \cdot (\text{p2dX} - \text{p1dX})$
		$\text{osdX} = \text{osdX} + 0.01 \cdot i \cdot (\text{p2dX} - \text{p1dX})$
		$\text{p1dY} = \text{p1dY} + 0.01 \cdot i \cdot (\text{p1dY} - \text{osdY})$
		$\text{p2dY} = \text{p2dY} + 0.015 \cdot i \cdot (\text{p1dY} - \text{osdY})$
	Olho: pálpebras superiores	$\text{pas1dY} = \text{pas1dY} + 0.015 \cdot i \cdot (\text{p1dY} - \text{osdY})$
		$\text{pas2dY} = \text{pas2dY} + 0.06 \cdot i \cdot (\text{p1dY} - \text{osdY})$
	Enrugamentos Testa central	$\text{osdY} = \text{osdY} - 0.025 \cdot i \cdot (\text{p1dY} - \text{osdY})$
		$\text{pseY} = \text{pseY} + 0.2 \cdot i \cdot (\text{pieY} - \text{pseY})$
		$\text{pseY} = \text{pseY} + 0.2 \cdot i \cdot (\text{pidY} - \text{psdY})$
		$\text{pt1cY} = \text{pt1cY} + (i-1) \cdot (\text{pt3cY} - \text{pt2cY})$
		$\text{pt2cY} = \text{pt2cY} + (i-1) \cdot (\text{pt3cY} - \text{pt2cY})$
		$\text{pt3cY} = \text{pt3cY} + (i-1) \cdot (\text{pt3cY} - \text{pt2cY})$
		$\text{pt4cY} = \text{pt4cY} + (i-1) \cdot (\text{pt3cY} - \text{pt2cY})$
		$\text{pt5cY} = \text{pt5cY} + (i-1) \cdot (\text{pt3cY} - \text{pt2cY})$
		$\text{pt1mY} = \text{pt1mY} + (i-1) \cdot (\text{pt3mY} - \text{pt2mY})$
		$\text{pt2mY} = \text{pt2mY} + (i-1) \cdot (\text{pt3mY} - \text{pt2mY})$
		$\text{pt3mY} = \text{pt3mY} + (i-1) \cdot (\text{pt3mY} - \text{pt2mY})$

Tabela 3.14: Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.

Unidade de ação	Parte do rosto	Equação
2	Enrugamentos Testa central	$pt4mY=pt4mY+(i-1) \cdot (pt3mY-pt2mY)$ $pt5mY=pt5mY+(i-1) \cdot (pt3mY-pt2mY)$ $pt1bY=pt1bY+(i-1) \cdot (pt3bY-pt2bY)$ $pt2bY=pt2bY+(i-1) \cdot (pt3bY-pt2bY)$ $pt3bY=pt3bY+(i-1) \cdot (pt3bY-pt2bY)$ $pt4bY=pt4bY+(i-1) \cdot (pt3bY-pt2bY)$ $pt5bY=pt5bY+(i-1) \cdot (pt3bY-pt2bY)$
4	Sobrancelha esquerda	$p1eX=p1eX-0.05 \cdot i \cdot (p2eX-p1eX)$ $p2eX=p2eX-0.01 \cdot i \cdot (p2eX-p1eX)$ $pas1eX=pas1eX-0.04 \cdot i \cdot (p2eX-p1eX)$ $pas2eX=pas2eX-0.01 \cdot i \cdot (p2eX-p1eX)$ $oseX=oseX-0.02 \cdot i \cdot (p2eX-p1eX)$ $p1eY=p1eY-0.03 \cdot i \cdot (p1eY-oseY)$ $p2eY=p2eY-0.01 \cdot i \cdot (p1eY-oseY)$ $pas1eY=pas1eY-0.025 \cdot i \cdot (p1eY-oseY)$ $pas2eY=pas2eY-0.01 \cdot i \cdot (p1eY-oseY)$ $oseY=oseY-0.04 \cdot i \cdot (p1eY-oseY)$
	Sobrancelha direita	$p1dX=p1dX+0.05 \cdot i \cdot (p2dX-p1dX)$ $p2dX=p2dX+0.01 \cdot i \cdot (p2dX-p1dX)$ $pas1dX=pas1dX+0.04 \cdot i \cdot (p2dX-p1dX)$ $pas2dX=pas2dX+0.01 \cdot i \cdot (p2dX-p1dX)$ $osdX=osdX+0.02 \cdot i \cdot (p2dX-p1dX)$ $p1dY=p1dY-0.03 \cdot i \cdot (p1dY-osdY)$ $p2dY=p2dY-0.01 \cdot i \cdot (p1dY-osdY)$ $pas1dY=pas1dY-0.025 \cdot i \cdot (p1dY-osdY)$ $pas2dY=pas2dY-0.01 \cdot i \cdot (p1dY-osdY)$ $osdY=osdY-0.04 \cdot i \cdot (p1dY-osdY)$
Olho: pálpebras superiores		$pseY=pseY-0.2 \cdot i \cdot (pieY-pseY)$ $psdY=psdY-0.2 \cdot i \cdot (pidY-psdY)$
1+2	Sobrancelha	$p1eY=p1eY-0.02 \cdot i \cdot (p1eY-oseY)$

Tabela 3.14: Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.

Unidade de ação	Parte do rosto	Equação
1+2	Sobrancelha esquerda	$p1eY = p1eY - 0.02 \cdot i \cdot (p1eY - oseY)$
		$p2eY = p2eY + 0.01 \cdot i \cdot (p1eY - oseY)$
		$pas1eY = pas1eY + 0.015 \cdot i \cdot (p1eY - oseY)$
		$pas2eY = pas2eY + 0.02 \cdot i \cdot (p1eY - oseY)$
		$oseY = oseY + 0.04 \cdot i \cdot (p1eY - oseY)$
	Sobrancelha direita	$p1dY = p1dY - 0.02 \cdot i \cdot (p1dY - osdY)$
		$p2dY = p2dY + 0.01 \cdot i \cdot (p1dY - osdY)$
		$pas1dY = pas1dY + 0.015 \cdot i \cdot (p1dY - osdY)$
		$pas2dY = pas2dY + 0.02 \cdot i \cdot (p1dY - osdY)$
		$osdY = osdY + 0.04 \cdot i \cdot (p1dY - osdY)$
1+4	Sobrancelha esquerda	$p1eX = p1eX + 0.02 \cdot i \cdot (p2eX - p1eX)$
		$pas1eX = pas1eX + 0.1 \cdot i \cdot (p2eX - p1eX)$
		$pas1eY = pas1eY + 1.5 \cdot i \cdot (p1eY - oseY)$
		$pas2eY = pas2eY - 0.5 \cdot i \cdot (p1eY - oseY)$
		$oseY = oseY + i \cdot (p1eY - oseY)$
	Sobrancelha direita	$p1dX = p1dX - 0.02 \cdot i \cdot (p2dX - p1dX)$
		$pas1dX = pas1dX - 0.1 \cdot i \cdot (p2dX - p1dX)$
		$pas1dY = pas1dY + 1.5 \cdot i \cdot (p1dY - osdY)$
		$pas2dY = pas2dY - 0.5 \cdot i \cdot (p1dY - osdY)$
		$osdY = osdY + i \cdot (p1dY - osdY)$
1+2+4	Sobrancelha esquerda	$p1eY = p1eY + 0.01 \cdot i \cdot 3.5cm(p1eY - oseY)$
		$p2eY = p2eY + 0.01 \cdot i \cdot (p1eY - oseY)$
		$pas1eY = pas1eY - 0.025 \cdot i \cdot (p1eY - oseY)$
		$pas2eY = pas2eY + 0.01 \cdot i \cdot (p1eY - oseY)$
		$oseY = oseY + 0.01 \cdot i \cdot (p1eY - oseY)$
	Sobrancelha direita	$p1dY = p1dY + 0.01 \cdot i \cdot (p1dY - osdY)$
		$p2dY = p2dY + 0.01 \cdot i \cdot (p1dY - osdY)$
		$pas1dY = pas1dY - 0.025 \cdot i \cdot (p1dY - osdY)$
		$pas2dY = pas2dY + 0.01 \cdot i \cdot (p1dY - osdY)$
		$osdY = osdY + 0.01 \cdot i \cdot (p1dY - osdY)$

Tabela 3.14: Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.

Unidade de ação	Parte do rosto	Equação
1+2+4	Olho: pálpebras superiores	$pseY=pseY-0.2 \cdot i \cdot (pieY-pseY)$
		$psdY=psdY-0.2 \cdot i \cdot (pidY-psdY)$
5	Olho: pálpebras superiores	$pseY=pseY+0.6 \cdot i \cdot (pieY-pseY)$
		$psdY=psdY+0.6 \cdot i \cdot (pidY-psdY)$
6	Olhos e pálpebras esquerda	$pseY=pseY-0.3 \cdot i \cdot (pieY-pseY)$
		$pieY=pieY+0.3 \cdot i \cdot (pieY-pseY)$
		$oreY=oreY+0.3 \cdot i \cdot (oreY-pp1eY)$
	Olhos e pálpebras direita	$psdY=psdY-0.3 \cdot i \cdot (pidY-psdY)$
		$pidY=pidY+0.3 \cdot i \cdot (pidY-psdY)$
		$ordY=ordY+0.3 \cdot i \cdot (ordY-pp1dY)$
	Enrugamentos olho esquerdo	$p1eY=p1eY+(i-1) \cdot (p8eY-p6eY)$
		$p1eX=p1eX+(i-1) \cdot (p8eX-p6eX)$
		$p2eX=p2eX+(i-1) \cdot (p8eX-p6eX)$
		$p3eX=p3eX+(i-1) \cdot (p8eX-p6eX)$
		$p4eX=p4eX+(i-1) \cdot (p8eX-p6eX)$
		$p5eX=p5eX+(i-1) \cdot (p8eX-p6eX)$
		$p6eX=p6eX+(i-1) \cdot (p8eX-p6eX)$
		$p7eX=p7eX+(i-1) \cdot (p8eX-p6eX)$
		$p8eX=p8eX+(i-1) \cdot (p8eX-p6eX)$
		$p9eX=p9eX+(i-1) \cdot (p8eX-p6eX)$
		$p10eX=p10eX+(i-1) \cdot (p8eX-p6eX)$
		$p11eX=p11eX+(i-1) \cdot (p8eX-p6eX)$
	Enrugamentos olho direito	$p1dY=p1dY+(i-1) \cdot (p8dY-p6dY)$
		$p1dX=p1dX+(i-1) \cdot (p8dX-p6dX)$
		$p2dX=p2dX+(i-1) \cdot (p8dX-p6dX)$
		$p3dX=p3dX+(i-1) \cdot (p8dX-p6dX)$
		$p4dX=p4dX+(i-1) \cdot (p8dX-p6dX)$
		$p5dX=p5dX+(i-1) \cdot (p8dX-p6dX)$
		$p6dX=p6dX+(i-1) \cdot (p8dX-p6dX)$
		$p7dX=p7dX+(i-1) \cdot (p8dX-p6dX)$
		$p8dX=p8dX+(i-1) \cdot (p8dX-p6dX)$
		$p9dX=p9dX+(i-1) \cdot (p8dX-p6dX)$
		$p10dX=p10dX+(i-1) \cdot (p8dX-p6dX)$
		$p11dX=p11dX+(i-1) \cdot (p8dX-p6dX)$
7	Olhos e pálpebras	$pseY=pseY-0.5 \cdot i \cdot (pieY-pseY)$

Tabela 3.14: Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.

Unidade de ação	Parte do rosto	Equação
7	Olhos e pálpebras esquerda	$\text{pieY} = \text{pieY} + 0.5 \cdot i \cdot (\text{pieY}-\text{pseY})$
		$\text{oreY} = \text{oreY} + 0.5 \cdot i \cdot (\text{oreY}-\text{pp1eY})$
	Olhos e pálpebras direita	$\text{psdY} = \text{psdY} - 0.5 \cdot i \cdot (\text{pidY}-\text{psdY})$
		$\text{pidY} = \text{pidY} + 0.5 \cdot i \cdot (\text{pidY}-\text{psdY})$
		$\text{ordY} = \text{ordY} + 0.5 \cdot i \cdot (\text{ordY}-\text{pp1dY})$
9	Nariz esquerda	$\text{aceX} = \text{aceX} + 0.03 \cdot i \cdot (\text{mfeX}-\text{snX})$
		$\text{aceY} = \text{aceY} + 0.06 \cdot i \cdot (\text{snY}-\text{mfeY})$
		$\text{aleY} = \text{aleY} + 0.06 \cdot i \cdot (\text{snY}-\text{mfeY})$
	Nariz direita	$\text{aceX} = \text{aceX} - 0.03 \cdot i \cdot (\text{mfeX}-\text{snX})$
		$\text{aceY} = \text{aceY} + 0.06 \cdot i \cdot (\text{snY}-\text{mfeY})$
		$\text{aleY} = \text{aleY} + 0.06 \cdot i \cdot (\text{snY}-\text{mfeY})$
	Lábio superior	$\text{lsY} = \text{lsY} + 0.06 \cdot i \cdot (\text{liY}-\text{lsY})$
12	Lábios	$\text{stoY} = \text{stoY} + 0.06 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{cheX} = \text{cheX} + 0.1 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{chdX} = \text{chdX} - 0.1 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{cheY} = \text{cheY} + 1.0 \cdot i \cdot (\text{liY}-\text{lsY})$
15	Lábios	$\text{chdY} = \text{chdY} + 1.0 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{cheY} = \text{cheY} - 1.0 \cdot i \cdot (\text{liY}-\text{lsY})$
16+25	Lábio inferior	$\text{chdY} = \text{chdY} - 1.0 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{stoY} = \text{stoY} - 1.2 \cdot i \cdot (\text{liY}-\text{lsY})$
20	Lábios	$\text{liY} = \text{liY} - 1.2 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{cheX} = \text{cheX} + 0.15 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{chdX} = \text{chdX} - 0.15 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{lsY} = \text{lsY} - 0.1 \cdot i \cdot (\text{liY}-\text{lsY})$
23	Lábios	$\text{liY} = \text{liY} + 0.1 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{cheX} = \text{cheX} + 0.004 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{chdX} = \text{chdX} - 0.004 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{lsY} = \text{lsY} - 0.27 \cdot i \cdot (\text{liY}-\text{lsY})$
25	Lábios	$\text{liY} = \text{liY} + 0.27 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{cheX} = \text{cheX} + 0.01 \cdot i \cdot (\text{cheX}-\text{chX})$
	Lábio superior	$\text{chdX} = \text{chdX} - 0.01 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{lsY} = \text{lsY} + 0.3 \cdot i \cdot (\text{liY}-\text{lsY})$
	Lábio inferior	$\text{stoY} = \text{stoY} + 0.3 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{stoY} = \text{stoY} - 0.3 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{liY} = \text{liY} - 0.3 \cdot i \cdot (\text{liY}-\text{lsY})$

Tabela 3.14: Cálculo das posições de cada ponto nas unidades de ação seguindo a descrição do envoltório.

Unidade de ação	Parte do rosto	Equação
26	Lábios	$\text{cheX} = \text{cheX} + 0.05 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{chdX} = \text{chdX} - 0.05 \cdot i \cdot (\text{cheX}-\text{chX})$
	Lábio superior	$\text{lsY} = \text{lsY} + 1.0 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{stoY} = \text{stoY} + 1.0 \cdot i \cdot (\text{liY}-\text{lsY})$
	Lábio inferior	$\text{stoY} = \text{stoY} - 1.0 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{liY} = \text{liY} - 1.0 \cdot i \cdot (\text{liY}-\text{lsY})$
27	Lábios	$\text{cheX} = \text{cheX} + 0.1 \cdot i \cdot (\text{cheX}-\text{chX})$
		$\text{chdX} = \text{chdX} - 0.1 \cdot i \cdot (\text{cheX}-\text{chX})$
	Lábio superior	$\text{lsY} = \text{lsY} + 2.0 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{stoY} = \text{stoY} + 2.0 \cdot i \cdot (\text{liY}-\text{lsY})$
	Lábio inferior	$\text{stoY} = \text{stoY} - 2.0 \cdot i \cdot (\text{liY}-\text{lsY})$
		$\text{liY} = \text{liY} - 2.0 \cdot i \cdot (\text{liY}-\text{lsY})$

3.4 Integração com a gramática

A gramática entrou nesse projeto com o objetivo de flexibilizar a maneira e a estrutura para armazenar e gerenciar a geração das caricaturas, ou seja, tem a função de oferecer um grau de liberdade maior na estruturação e geração dos *sketches*. Essa gramática é implementada por meio de um grafo AND-OR.

Primeiramente os modelos para representação da face e das emoções foram transferidos para o Grafo. Então o sistema de desenho de *splines* foi incorporado ao sistema de geração de faces com a gramática. Essa integração foi realizada por meio de uma interface no código da aplicação que chama os métodos responsáveis por desenhar as curvas *splines* e, assim, realizar a síntese das caricaturas de cada expressão facial desejada.

3.4.1 Construção do Grafo AND-OR

O grafo AND-OR é esquematizado na Figura 3.4. A definição do grafo apresenta um nó OR na raiz (nível 1), que seleciona um dentre sete diferentes nós AND, cada um simbolizando uma face com uma das seis expressões faciais básicas, além da face neutra (nível 2).

No terceiro nível encontram-se os nós OR das sete partes principais que compõem qualquer um dos tipos de faces: sobrancelhas (esquerda e direita), olhos (esquerdo e direito), nariz, boca e restante do rosto. Estes, por serem

do tipo OR, também são seletores e possuem uma gama de opções cada um no nível 4 (ex.: boca reta, boca arqueada para cima, boca arqueada para baixo, boca aberta, etc).

Para o desenho *b-spline*, basta um conjunto de pontos de controle como informação adicional ao grafo para que uma curva seja desenhada. Então, foram incluídas duas novas classes de componentes no grafo: curvas e pontos. Cada nó AND no quarto nível se decompõe em curvas, presentes no quinto nível.

As curvas neste quinto nível são nós OR, e também possuem tipos à escolha no sexto nível. Por fim, cada curva em um nó AND do sexto nível é composta pelos pontos de controle usados como parâmetro para sua geração. Os pontos são os nós folhas do grafo e contêm suas respectivas coordenadas *x* e *y* no plano bidimensional.

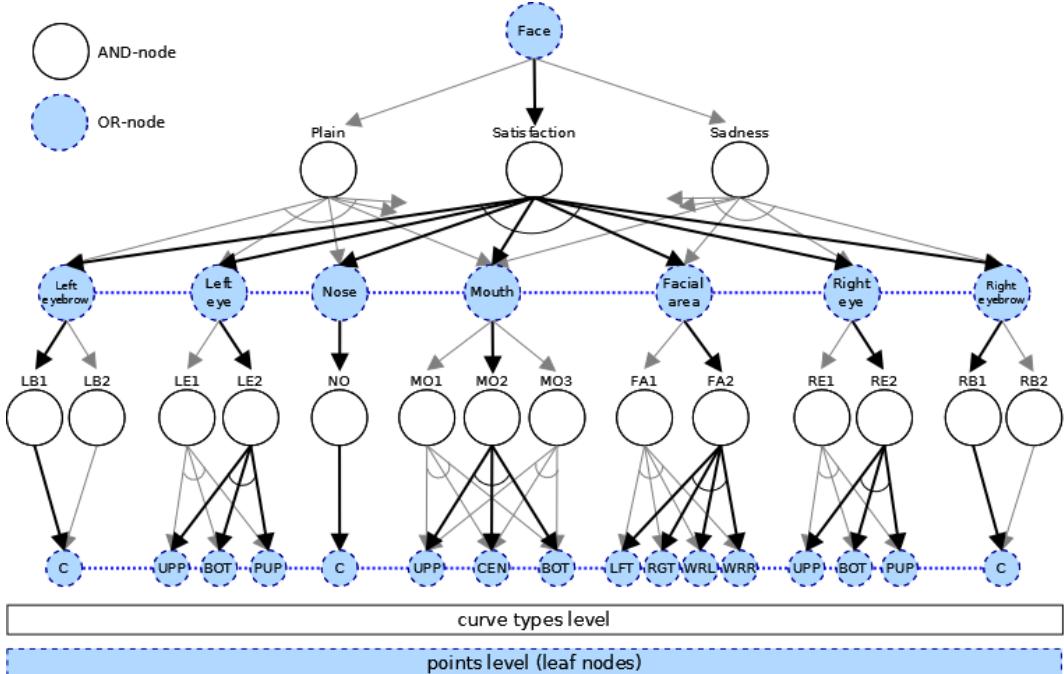


Figura 3.4: Ilustração do grafo AND-OR parcial com a face neutra e as emoções Satisfação e Tristeza.

Uma dada configuração desse grafo representa uma face inicial com determinada emoção, mas tendo nas folhas um intervalo possível de posicionamento de todos os pontos de controle. Assim, para representar uma dada imagem, os valores exatos das coordenadas *x* e *y* dos pontos de controle precisam ser definidos. Isso é feito primeiro inicializando esses valores aleatoriamente em um intervalo de um desvio padrão para mais ou para

menos em torno da média (valores definidos nas folhas com base nos dados antropométricos). Os nós folhas da face neutra tem seus posicionamentos calculados conforme o modelo antropométrico descrito na Seção 3.1.1.

Esses valores são então ajustados conforme a intensidade desejada daquela emoção. Agora, os nós folhas terão suas posições de x e y computadas de acordo com o modelo do envoltório, estes cálculos são apresentados na Tabela 3.14

Para gerar as configurações de imagens, um algoritmo deve percorrer o grafo recursivamente a partir da raiz, ramificando a busca em cada nó AND e selecionando uma opção em cada nó OR. Durante esse percurso, as relações de distância são consultadas e as posições dos pontos são sorteadas e ajustadas segundo as restrições previamente definidas nas relações. As setas em preto na Figura 3.4 indicam as arestas associadas às produções que geram uma face de satisfação.

A maior parte das arestas horizontais indicam relações espaciais de distância entre as partes, curvas e pontos. Essas distâncias, conforme mencionado, são atribuídas segundo os dados antropométricos.

3.5 Avaliação das imagens geradas

O realismo das caricaturas geradas foi avaliado por voluntários. Foram construídos formulários utilizando-se a ferramenta Google formulários *Google-Forms*¹, disponibilizados na Internet. Os formulários continham imagens de todas as emoções, sendo disponibilizada uma única imagem em cada página a fim de evitar que o avaliador se confundisse. Ao lado de cada imagem foram disponibilizadas as opções das sete emoções (incluindo a face neutra), sendo solicitado ao avaliador que selecionasse aquela que ele considerava representada na imagem.² A primeira página do formulário de apresentação é mostrada na Figura 3.5.

Cada formulário possuía quatro imagens com diferentes níveis de intensidade (25%, 50%, 75% e 100%) de cada uma das seis emoções e mais uma imagem representando a face neutra. Como mecanismo de conferência, foram repetidas as seis imagens com intensidade 75%, uma de cada emoção não neutra, em posições aleatórias dos formulários, a fim de permitir estimar o nível de coerência dos avaliadores. Cada formulário totalizou, assim, 31 imagens.

A fim de tentar minimizar a influência de fadiga do avaliador, o que poderia gerar efeitos no reconhecimento das imagens posicionadas no final

¹<http://www.google.com/google-d-s/createforms.html>

²Um formulário de demonstração está disponível em: <http://goo.gl/AqX7Ut>

Demonstração: Expressões faciais das emoções

Nosso projeto visa desenvolver um programa capaz de desenhar, de forma automática, imagens de faces humanas com expressões faciais. Inicialmente, as imagens serão apenas esboços (caricaturas), sem cores, texturas ou aparência 3D. Esse programa poderá ser utilizado para fins médicos, como um jogo ou ferramenta de diagnóstico para pacientes que, devido ao transtorno psiquiátrico que apresentam, possuem dificuldades em reconhecer corretamente as expressões faciais das emoções.

O questionário a seguir tem como objetivo avaliar a qualidade das imagens geradas até o momento. Apresentamos abaixo um conjunto de 31 imagens de faces. Cada uma delas está associada a uma das seis expressões faciais básicas – satisfação (ou alegria), tristeza, surpresa, raiva, aversão (ou nojo), medo e expressão neutra – e apresenta um certo grau de intensidade.

Solicitamos que você assinale para cada uma das imagens qual das sete expressões faciais você acredita que ela representa melhor.

*Obrigatório

Idade: *

Sexo: *

- Feminino
 Masculino



1 - Qual emoção melhor descreve a imagem acima? *

Selecione apenas uma opção

- Satisfação
 Tristeza
 Surpresa
 Raiva
 Aversão (Nojo)
 Medo
 Neutra

[Continuar >](#)

3% concluído

Powered by
 Google Forms

Este conteúdo não foi criado nem aprovado pela Google.
[Demandar abuso](#) • [Termos de Serviço](#) • [Termos Adicionais](#)

Figura 3.5: Exemplo de formulário usado para coletar dados.

do formulário, as imagens foram aleatoriamente distribuídas gerando dez diferentes versões dos formulários.

A participação dos avaliadores foi voluntária e anônima. Os formulários foram distribuídos por mensagens eletrônicas pelos participantes do projeto. Os dados de idade e sexo também foram coletados com o intuito de permitir verificar eventuais diferenças de percepção em função dessas variáveis.

Capítulo 4

Resultados

As caricaturas geradas para a avaliação são apresentadas na Figura 4.1, considerando as intensidades descritas na seção anterior.

A avaliação das imagens contou com 237 avaliadores voluntários saudáveis, cujas idades variaram de 13 a 77 anos, sendo 54% homens do total de avaliadores.

Foram considerados três grupos nos resultados apresentados: menos de 25, 26 a 55 e maior que 55 anos. O percentual de participantes em cada grupo foi, respectivamente, 32% (77 voluntários), 61% (145 voluntários) e 6% (15 voluntários).

A Figura 4.2 e a Tabela 4.1 apresentam os resultados de forma global, ou seja, sem distinção de gênero ou idade. Na Figura 4.2, cada curva representa a porcentagem de acertos para cada uma das intensidades consideradas de cada emoção (25%, 50%, 75%, 100%). A Tabela 4.1 apresenta o cruzamento de respostas entre os diversos tipos de emoções representadas nas imagens. Vale ressaltar que uma média foi feita para a intensidade 75% de cada emoção por causa das imagens duplicadas.

Tabela 4.1: Matriz de confusão das respostas obtidas. Em negritos estão indicadas as taxas de correspondência das emoções esperadas.

	Raiva	Aversão	Medo	Satisfação	Neutra	Tristeza	Surpresa
Raiva	51%	7%	6%	1%	27%	5%	3%
Aversão	38%	49%	2%	2%	5%	2%	2%
Medo	9%	7%	36%	2%	12%	13%	21%
Satisfação	0%	0%	0%	86%	10%	2%	2%
Neutra	1%	0%	2%	2%	91%	2%	2%
Tristeza	6%	5%	6%	0%	3%	79%	1%
Surpresa	0%	1%	8%	5%	13%	2%	71%

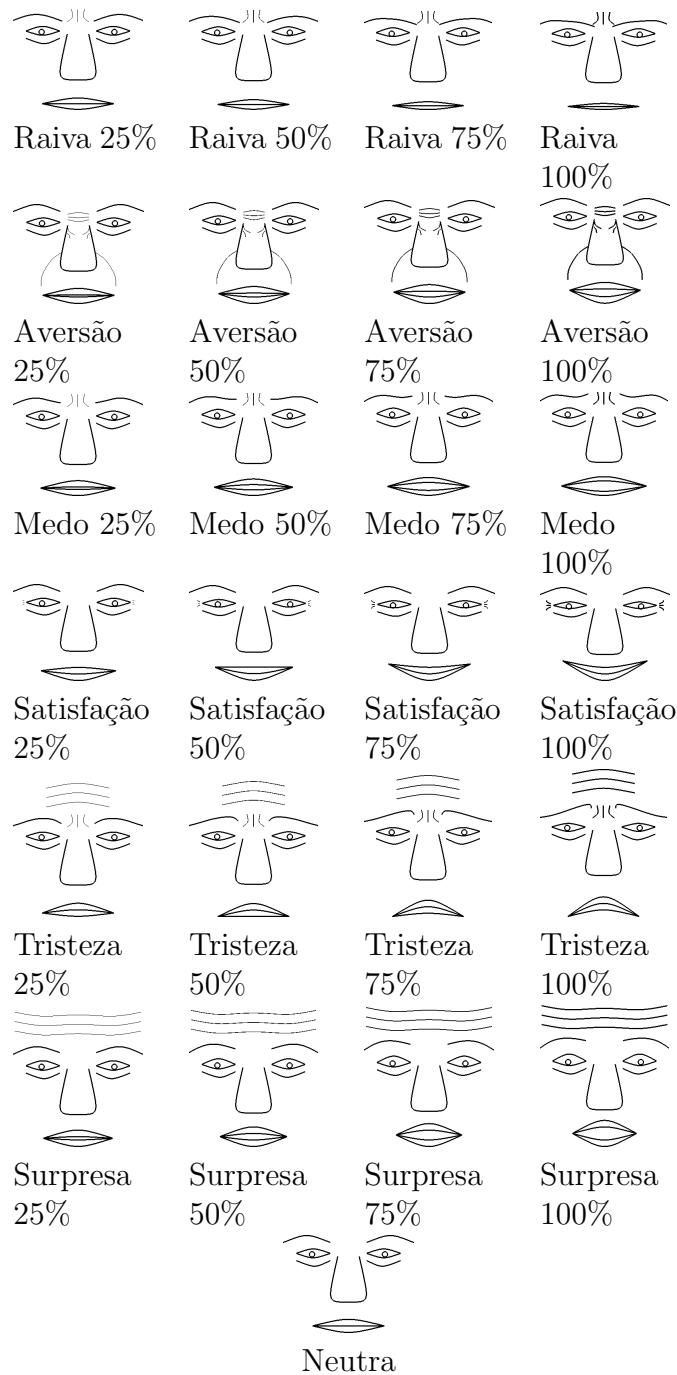


Figura 4.1: Expressões faciais utilizadas na avaliação.

A Figura 4.3 apresenta as taxas de acerto considerando-se os gêneros feminino e masculino, mas desconsiderando as intensidades.

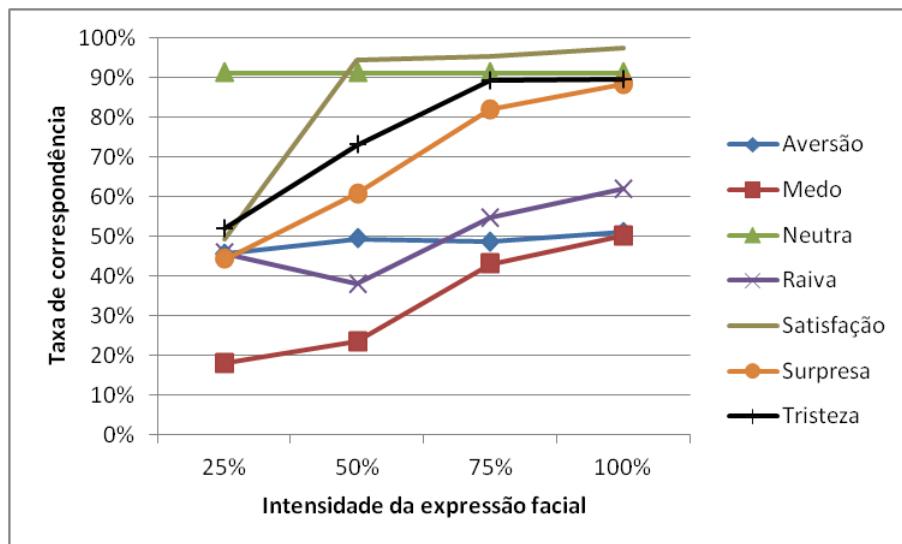


Figura 4.2: Taxa de acerto das respostas para cada emoção avaliada.

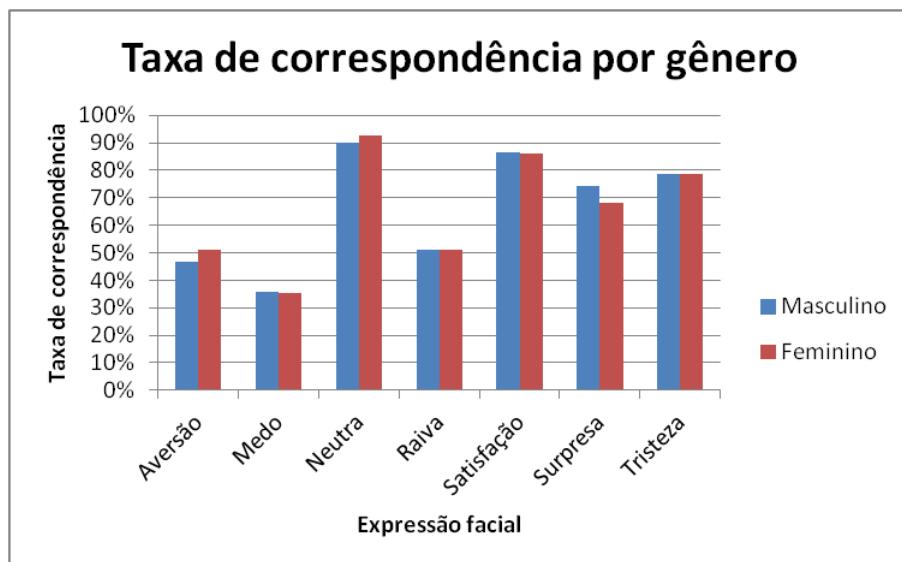


Figura 4.3: Taxa de acerto das respostas para cada gênero.

Finalmente, a Figura 4.4 apresenta as taxas de acerto dentro de cada faixa etária, considerando-se três grupos de idades.

Uma última análise foi conduzida sobre as respostas. Conforme descrito na seção anterior, imagens foram incluídas repetidamente nos formulários como uma forma de controle. A Tabela 4.2 apresenta os níveis de coerência entre as respostas de imagens repetidas, considerando-se uma análise global

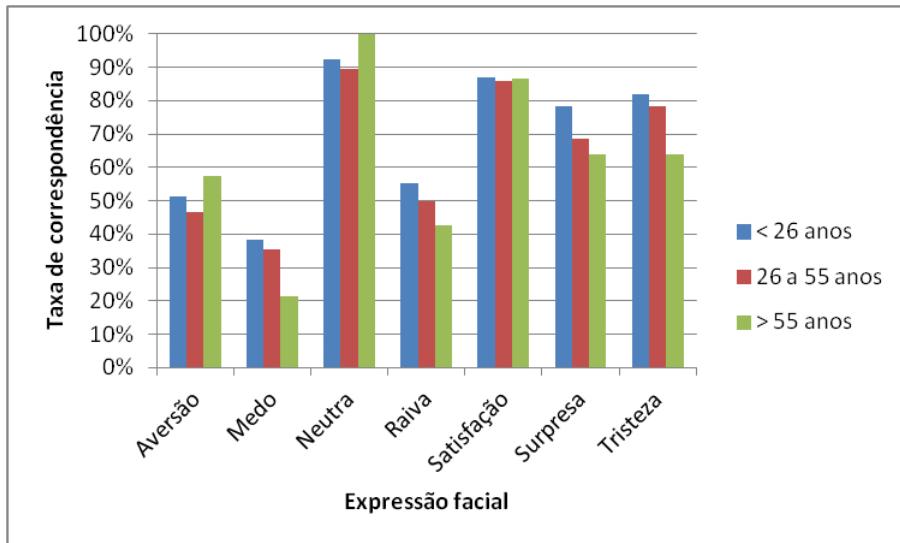


Figura 4.4: Taxa de acerto das respostas para cada faixa etária.

e divisões por gênero e por faixa etária.

Tabela 4.2: Coerência das respostas referentes às imagens duplicadas.

	<25%	26 a 50%	51 a 75%	>75%
Global	1%	28%	33%	38%
Masculino	2%	28%	29%	41%
Feminino	1%	28%	37%	34%
<25 anos	0%	24%	36%	40%
26-55 anos	2%	27%	32%	39%
>55 anos	0%	56%	25%	19%

4.1 Discussão

Apesar de utilizarmos diferentes intensidades de cada emoção nas entrevistas, consideramos como resposta correta de uma imagem a emoção relacionada àquela imagem, não havendo sido perguntado nada sobre intensidade. Vale ressaltar que utilizamos voluntários saudáveis com o objetivo de avaliar a qualidade da representação das emoções nas imagens geradas. Logo, as taxas de erros provavelmente não significam inabilidade dos avaliadores em reconhecerem emoções faciais, mas indícios de que as caricaturas precisam ser aperfeiçoadas.

4.1.1 Avaliação global das caricaturas

A partir da Figura 4.2 é possível perceber que, tristeza da face neutra e da emoção aversão, o reconhecimento correto aumentou à medida que aumentou a intensidade das emoções representadas. Outra exceção foi a emoção raiva, que apresentou uma queda da taxa de acerto na intensidade 50%, na qual essa emoção foi mais confundida com a emoção neutra. As duas emoções diferem no arqueamento da boca para baixo e na inclusão de rugas no meio da testa na emoção raiva, que são sutis na intensidades de 25 e 50%.

As imagens da emoção medo foram as que apresentaram os menores níveis de reconhecimento em todas as intensidades, seguidas pelas imagens das emoções aversão e raiva. A análise das imagens de tais emoções permite perceber que essas são as caricaturas com menos alterações na parte inferior da face (região da boca). Isso sugere que os voluntários reconhecem mais facilmente as caricaturas com alterações mais acentuadas nesta região.

Verifica-se, ainda, que todas as emoções apresentaram taxas baixas de reconhecimento com a intensidade 25%, fornecendo indícios de que as alterações nas caricaturas estão muito sutis e dificultam o reconhecimento. Isso acontece até mesmo para as emoções que apresentaram alta taxa de reconhecimento nas demais intensidades. Destaca-se que o nível 25% de intensidade produz poucas alterações na região inferior da face, o que pode reforçar a importância da alteração dos traços nesta região.

A Tabela 4.1 fornece dados mais detalhados da confusão no reconhecimento das emoções. Analisando-se a emoção medo na terceira linha desta Tabela, verifica-se a menor taxa de reconhecimento (36%). As imagens referentes a esta emoção foram reconhecidas pelos voluntários como pertencentes a quase todas as demais emoções. A emoção com a qual a emoção medo foi mais confundida foi surpresa. Ao analisar as imagens dessas duas emoções verifica-se que a parte inferior das faces são parecidas, aumentando-se a abertura da boca de acordo com a intensidade. Assim, há indícios que esta alteração de traços foi relacionada à emoção surpresa e que as caricaturas referentes à emoção Medo devem ser reavaliadas a fim de incluir ou alterar traços que possam melhor representar tal emoção.

Uma análise similar pode ser conduzida para a emoção aversão que foi a segunda com menor taxa de acerto (segunda linha da Tabela 4.1). Esta emoção foi majoritariamente confundida com a emoção raiva. As imagens indicam que as principais diferenças entre as caricaturas dessas duas emoções são as linhas da testa e o acréscimo das linhas do nariz na emoção aversão. Essas linhas no nariz representam a alteração na unidade de ação 9, mas não foram suficientes para expressar com clareza a emoção citada, fato que deve ser melhorado nas caricaturas. De forma semelhante, a emoção raiva foi

confundida pelos usuários com a face neutra, indicando que os traços desta emoção deveriam ser mais contundentes.

Com exceção das três emoções analisadas, a diagonal principal da matriz de confusão apresentada na Tabela 4.1 indica que as demais emoções apresentaram taxas satisfatórias de reconhecimento. A emoção satisfação foi a que apresentou a maior taxa de acerto, seguida das emoções tristeza e surpresa. Esses dados sugerem que as caricaturas foram devidamente reconhecidas pelos avaliadores, principalmente considerando-se as intensidades de 75 e 100%.

Esses resultados concordam com os resultados descritos em (DYCK et al., 2008), no qual as emoções satisfação, tristeza e neutra apresentaram alta taxa de reconhecimento enquanto a emoção aversão foi fracamente reconhecida. Dyck et al. (2008), no entanto, alcançaram taxas de acerto satisfatórias também para as emoções medo e raiva.

4.1.2 Variáveis demográficas e coerência das respostas

A Figura 4.3 mostra que as taxas de acerto das emoções não variaram significativamente com relação ao sexo dos avaliadores.

De forma semelhante, a análise dos resultados considerando-se os três grupos de faixa etária não fornece indícios de diferenças significativas nos resultados considerando-se esta variável. No entanto, algumas observações, ainda que incipientes, podem ser discutidas. O grupo de voluntários mais velhos (mais que 55 anos) foram os que mais acertaram a face neutra e a emoção aversão. Também foram os que menos acertaram as demais emoções, tristeza da emoção satisfação, a qual apresenta uma taxa de reconhecimento mais consistente entre as faixas etárias. O grupo mais jovem (até 25 anos) acertou mais as emoções tristeza e surpresa. Entretanto, o gráfico mostra que as porcentagens são parecidas, o que torna difícil uma análise mais definitiva da influência desta variável.

Como mencionado na seção 3.5 a avaliação com voluntários incluiu um mecanismo para verificar a coerência das respostas, repetindo a imagem com 75% de intensidade de cada emoção. A Tabela 4.2 apresenta os resultados obtidos, considerando-se uma análise global e a divisão por gênero e faixa etária. Cada coluna traz a porcentagem de indivíduos que atingiram determinada faixa de coerência. Um resultado foi considerado coerente quando o avaliador indicou a mesma emoção para duas caricaturas idênticas no formulário, mesmo que esta emoção não estivesse correta.

A primeira linha da tabela indica que mais de 70% dos voluntários atingiram mais de 50% de coerência. As linhas 2 e 3 indicam que não houve diferença na divisão por gêneros: ambos os gêneros apresentaram em torno

de 71% dos voluntários com coerência na faixa superior (50% ou maior).

No entanto, verifica-se diminuição da coerência à medida que aumenta a idade dos voluntários. A faixa etária mais jovem (até 25 anos) apresentou em torno de 76% de voluntários com coerência na faixa superior. Já os voluntários com idade entre 26 e 55 anos e idade maior que 55 anos apresentaram, respectivamente, média de 1% e 44% de coerência na faixa superior. Embora a menor quantidade de voluntários com idade superior a 55 anos possa ter influenciado nessas porcentagens, é importante destacar que os avaliadores mais velhos parecem se confundir mais com as caricaturas do que aqueles mais jovens.

Capítulo 5

Considerações finais

A abordagem utilizada para produzir caricaturas com diferentes intensidades de emoções obteve resultados animadores considerando-se a avaliação de voluntários saudáveis. Com isso a proposta de união da antropometria, psicologia, geração de imagens por meio de *splines* e gramática se mostrou uma forma eficaz de sintetizar caricaturas das emoções.

As caricaturas da face neutra e das emoções satisfação, tristeza e surpresa foram claramente definidas pelos traços usando curvas *splines*, mas as emoções medo, aversão e raiva tiveram caricaturas com reconhecimento limitado. Verificou-se uma certa semelhança dos traços, uma vez que as AUs acionadas são semelhantes. Embora este fato possa parecer uma limitação da abordagem, a versão atual do projeto usou uma quantidade de curvas que pode ser aumentada. Em trabalhos futuros, portanto, o refinamento dos traços e a inclusão de novos traços podem diminuir a confusão entre as representações dessas emoções e melhorar os resultados.

Vale ressaltar a importância do esforço em gerar imagens virtuais ao invés da utilização de imagens reais. Imagens virtuais para reconhecimento de emoções podem ser igualmente eficazes se comparadas com imagens reais em aplicações de pesquisa (DYCK et al., 2008; BEKELE et al., 2014). Além disso, as imagens virtuais facilitam a implementação de aplicativos que trabalhem a manipulação e a animação (mudanças dinâmicas entre as intensidades e emoções) das faces, permitindo uma maior precisão no seu detalhamento (BEKELE et al., 2014). Por exemplo, não há meios de definir, de forma objetiva, como uma pessoa deveria fazer uma expressão de 75% de aversão. Além disso, uma pessoa real pode não conseguir isolar os músculos específicos de uma determinada emoção, ocasionando mistura de expressões e consequentemente tornando o reconhecimento mais difícil. Como o objetivo final deste trabalho é a criação de um jogo para treinamento de pessoas com dificuldades de reconhecimento de emoção, o cenário virtual, menos complexo, seria um

preparativo para as situações reais.

Além do mais, a integração com a gramática permitiu um flexibilização do modelo, e, desta maneira, a evolução do trabalho no futuro se dará de uma forma mais dinâmica, conforme se agregar mais contribuições.

Um artigo com os resultados desse projeto está sendo elaborado e será submetido a conferência ACM SAC. Uma versão prévia é apresentada no apêndice A

Referências Bibliográficas

- BEKELE, E. et al. Assessing the utility of a virtual environment for enhancing facial affect recognition in adolescents with autism. *Journal of Autism and Developmental Disorders*, Jan 2014.
- CHENG, Y.; LING, S. 3d animated facial expression and autism in taiwan. In: IEEE. *Advanced Learning Technologies, 2008. ICALT'08. Eighth IEEE International Conference on*. [S.l.], 2008. p. 17–19.
- CLARK, T. F.; WINKIELMAN, P.; MCINTOSH, D. N. Autism and the extraction of emotion from briefly presented facial expressions: stumbling at the first step of empathy. *Emotion*, American Psychological Association, v. 8, n. 6, p. 803, 2008.
- DECARLO, D.; METAXAS, D.; STONE, M. An anthropometric face model using variational techniques. In: *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA: ACM, 1998. (SIGGRAPH '98), p. 67–74. ISBN 0-89791-999-8. Disponível em: <<http://doi.acm.org/10.1145/280814.280823>>.
- DYCK, M. et al. Recognition profile of emotions in natural and virtual faces. *PLOS One*, v. 3, n. 11, p. e3628, 2008.
- EKMAN, P. *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. [S.l.]: Macmillan, 2007.
- EKMAN, P.; FRIESEN, W. V. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, American Psychological Association, v. 17, n. 2, p. 124, 1971.
- EKMAN, P.; FRIESEN, W. V. Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, Springer, v. 1, n. 1, p. 56–75, 1976.
- EKMAN, P.; FRIESEN, W. V. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto: Consulting Psychologists Press, 1978.

EKMAN, P.; FRIESEN, W. V.; HAGER, J. C. Facs investigator's guide. *A human face*, 2002.

EKMAN, P. et al. Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, American Psychological Association, v. 53, n. 4, p. 712, 1987.

FARKAS, L. *Anthropometry of the head and face*. Raven Press, 1994. ISBN 9780781701594. Disponível em: <<http://books.google.com.br/books?id=MKVpAAAAMAAJ>>.

FARKAS, L. G. *Anthropometry of the head and face in medicine*. [S.l.]: Elsevier New York, 1981.

FARKAS, L. G.; KATIC, M. J.; FORREST, C. R. International anthropometric study of facial morphology in various ethnic groups/races. *Journal of Craniofacial Surgery*, LWW, v. 16, n. 4, p. 615–646, 2005.

FOLEY, J. D. *Computer graphics: Principles and practice*, in C. [S.l.]: Addison-Wesley Professional, 1996.

GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. L. *Digital image processing using MATLAB*. [S.l.]: Gatesmark Publishing Knoxville, 2009.

GOSSELIN, P.; KIROUAC, G. et al. Decoding facial emotional prototypes. *Canadian journal of experimental psychology= Revue canadienne de psychologie expérimentale*, v. 49, n. 3, p. 313, 1995.

GRYNSZPAN, O.; MARTIN, J.-C.; NADEL, J. Multimedia interfaces for users with high functioning autism: An empirical investigation. *International Journal of Human-Computer Studies*, Elsevier, v. 66, n. 8, p. 628–639, 2008.

GUEDES, S. P. de C.; TEIXEIRA, B. V.; CATTONI, D. M. Medidas orofaciais em adolescentes do estado do rio de janeiro segundo a tipologia facial. *Revista CEFAC*, SciELO Brasil, v. 12, n. 1, p. 68–74, 2010.

HARMS, M. B.; MARTIN, A.; WALLACE, G. L. Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychology review*, Springer, v. 20, n. 3, p. 290–322, 2010.

HEARN, D.; BAKER, M. P. *Computer graphics, C version*. [S.l.]: Prentice Hall Upper Saddle River, 1997.

- HESS, U. The communication of emotion. *Emotions, qualia and consciousness*, p. 397–409, 2001.
- JR, L. S.; HART, M. H. Random gravitational encounters and the evolution of spherical systems. ii. models. *The Astrophysical Journal*, v. 166, p. 483, 1971.
- KRISHAN, K. Anthropometry in forensic medicine and forensic science-'forensic anthropometry'. *The Internet Journal of Forensic Science*, v. 2, n. 1, 2007.
- LAHIRI, U. et al. Design of a virtual reality based adaptive response technology for children with autism. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, IEEE, v. 21, n. 1, p. 55–64, 2013.
- LIN, L. et al. A stochastic graph grammar for compositional object representation and recognition. *Pattern Recogn.*, v. 42, n. 7, p. 1297–1307, jul. 2009. ISSN 0031-3203.
- MOURÃO-MIRANDA, J. et al. Pattern recognition and functional neuroimaging help to discriminate healthy adolescents at risk for mood disorders from low risk adolescents. *PloS one*, Public Library of Science, v. 7, n. 2, p. e29482, 2012.
- NEWSCHAFFER, C. J.; CURRAN, L. K. Autism: an emerging public health problem. *Public health reports*, Association of Schools of Public Health, v. 118, n. 5, p. 393, 2003.
- PERISSINOTTO, E. et al. Anthropometric measurements in the elderly: age and gender differences. *British Journal of Nutrition*, Cambridge Univ Press, v. 87, n. 2, p. 177–186, 2002.
- RAMANATHAN, N.; CHELLAPPA, R.; BISWAS, S. Computational methods for modeling facial aging: A survey. *Journal of Visual Languages & Computing*, Elsevier, v. 20, n. 3, p. 131–144, 2009.
- ROCCA, C. C. d. A. et al. Facial emotion recognition in bipolar disorder: a critical review. *Revista Brasileira de Psiquiatria*, SciELO Brasil, v. 31, n. 2, p. 171–180, 2009.
- ROZENBERG, G.; SALOMAA, A. *Lindenmayer systems: Impacts on theoretical computer science, computer graphics, and developmental biology*. [S.l.]: Springer-Verlag New York, Inc., 2001.

- SCHERER, K. R. Vocal communication of emotion: A review of research paradigms. *Speech communication*, Elsevier, v. 40, n. 1, p. 227–256, 2003.
- TAYLOR, S. F.; MACDONALD, A. W. Brain mapping biomarkers of socio-emotional processing in schizophrenia. *Schizophrenia bulletin*, MPRC, v. 38, n. 1, p. 73–80, 2012.
- WEEDEN, J. C.; TROTMAN, C.-A.; FARAWAY, J. J. Three dimensional analysis of facial movement in normal adults: influence of sex and facial shape. *The Angle Orthodontist*, v. 71, n. 2, p. 132–140, 2001.
- XU, Z. et al. A hierarchical compositional model for face representation and sketching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 30, n. 6, p. 955–969, June 2008.
- YANG, Y. et al. Interactive facial sketch expression generation using local constraints. In: *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on*. [S.l.: s.n.], 2009. v. 2, p. 864–868.
- ZHU, S.-C.; MUMFORD, D. *A stochastic grammar of images*. [S.l.]: Now Publishers Inc, 2007.
- ZHU, S.-C.; MUMFORD, D. *A stochastic grammar of images*. [S.l.]: Now Publishers Inc, 2007.

Apêndice A

**Generating facial emotions
using anthropometry and
formal languages for individuals
suffering from psychiatric
disorders**

Generating facial emotions using anthropometry and formal languages for individuals suffering from psychiatric disorders

Rafael L. Testa
School of Arts, Sciences and Humanities
University of São Paulo
São Paulo - SP, Brazil
rafael.testa@usp.br

Rodrigo S. Dias
Department of Psychiatry
University of São Paulo
Medical School
São Paulo - SP, Brazil
rdgdias@gmail.com

Antônio H. N. Muniz
School of Arts, Sciences and Humanities
University of São Paulo
São Paulo - SP, Brazil
antonio.muniz@usp.br

Cristiana C. A. Rocca
Department of Psychiatry
University of São Paulo
Medical School
São Paulo - SP, Brazil
crisrocca@gmail.com

Fátima L. S. Nunes
School of Arts, Sciences and Humanities
University of São Paulo
São Paulo - SP, Brazil
fatima.nunes@usp.br

Omitted due to the blind review. Carpio L. U. S.
School of Arts, Sciences and Humanities
University of São Paulo
São Paulo - SP, Brazil
liseth.segundo@usp.br

Ariane Machado-Lima
School of Arts, Sciences and Humanities
University of São Paulo
São Paulo - SP, Brazil
ariane.machado@usp.br

ABSTRACT

The ability to process and identify facial emotions is an essential factor for an individual's social interaction. There are certain psychiatric disorders that can limit an individual's ability to recognize emotions in facial expressions. This problem could be confronted by making use of computational techniques in order to develop tools for the diagnosis, evaluation and training in identifying facial emotions. This paper presents an approach that uses image processing techniques, formal languages, anthropometry and Facial Action Coding System (FACS) to generate caricatures that represent facial movements related to neutral, satisfaction, sadness, anger, disgust, fear and surprise emotions. The rules that define the emotions were determined using an AND-OR graph to enable generating these images in a flexible manner. An evaluation conducted with healthy volunteer users showed that some emotions are more easily recognized, while for other emotions the caricatures need to be further improved. This is a promising approach, since the parameters used provide flexibility to define the emotional intensity that must be represented, and in the future,to include other characteristics

such as gender, ethnicity and age.

Categories and Subject Descriptors

F.4.2 [Mathematical Logic and Formal Languages]: Grammars and Other Rewriting Systems; I.2.1 [Artificial Intelligence]: Applications and Expert Systems—*Games*; I.3.3 [Computer Graphics]: Picture/Image Generation; I.4.7 [Image Processing and Computer Vision]: Feature Measurement—*Feature Representation*; I.4.10 [Image Processing and Computer Vision]: Image Representation

General Terms

Algorithms, Experimentation, Measurement.

Keywords

AND-OR Graphs, facial recognition, facial emotions, facial expression, anthropometry, computational grammars, splines, formal languages.

1. INTRODUCTION

The capacity to process and identify facial emotions is an essential factor of human communication and social interaction according to cultural norms and customs, based on common standards. Six universal emotions have been mapped, namely; satisfaction, sadness, anger, disgust, fear and surprise, which correspond to a specific pattern of facial muscles, with partially separable neural circuits that can prepare an appropriate behavioral response [5, 17, 10].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM Multimedia 2014 Orlando, Florida, USA
Copyright 20XX ACM X-XXXXX-XX-X/XX/XX ...\$15.00.

With regards to psychiatric disorders, the face recognition paradigm has received more attention, especially in Autism Spectrum Disorders (ASD) [12], mood disorders [17] and schizophrenia [19]. Computational techniques that include image processing, computer graphics, pattern recognition and virtual reality have been closely associated with developing tools for the diagnosis, evaluation and training in the recognition of emotional facial expressions [15, 2, 11]. What has been verified, however, is that most of these tools are limited in terms of flexibility to generate an adequate variety of objects. Such flexibility is necessary when the goal is to provide autonomy to healthcare professionals so they can adapt the tools for diagnosis or for training in a specific situation or for a particular patient.

This paper describes the initial results of a project intended for developing expressions (initially static images and later animations) to represent different intensities of facial emotions assigned for the training and rehabilitation of individuals with mental disorders. Several graphics processing techniques were used to create facial caricatures in order to represent different states of emotional intensities. The grammatical techniques (AND-OR graphs) used detailed the initial representation of the images that represent expressions of the six emotions cited before. The purpose for creating this new approach is, in the future, to create games that can easily manipulate not only the images that represent emotions and their various intensities, but also to parameterize avatar characteristics such as gender, ethnicity and age-groups.

The caricatures were constructed based on the concepts of anthropometry and Facial Action Coding System (FACS), presented with the AND-OR graphs definition in section 2. Related works are presented in section 3. The methodology used is presented in section 4. Section 5 describes the results obtained with the approach presented and Sections 6 and 7, respectively, describe the discussion of the results achieved and the conclusions about the proposed approach.

2. CONCEPTS

The methodology used in this work includes concepts from different areas, which are presented in this section.

Anthropometry is the science that studies the measurements of the size, weight and proportions of the human body, from the measurements of a population [8]. These measurements were determined by a set of well-defined points on the human body, called *landmarks* points. The distance between two landmarks characterizes an anthropometric measurement. Figure 1 shows a visual example of the location of the landmarks.

Based on the universality of facial emotions, this paper determined techniques to relate the measures across parts of the face with the facial movements involved in the expressions generated by emotions. In [6] the authors present the facial expressions of emotions considering the fundamental actions of a muscle or group of muscles, called Action Units (AUs). Thus, a facial expression is represented as a combination of these Action Units. These ideas were used to develop a methodology known as Facial Action Coding System (FACS), which can classify facial expressions based on the identification of certain Action Units found in the face [7].

In order to illustrate the Actions Units, Figure 2 shows the AUs 1, 2 and 4. In action unit 1 (Figure 2(a)) the inner

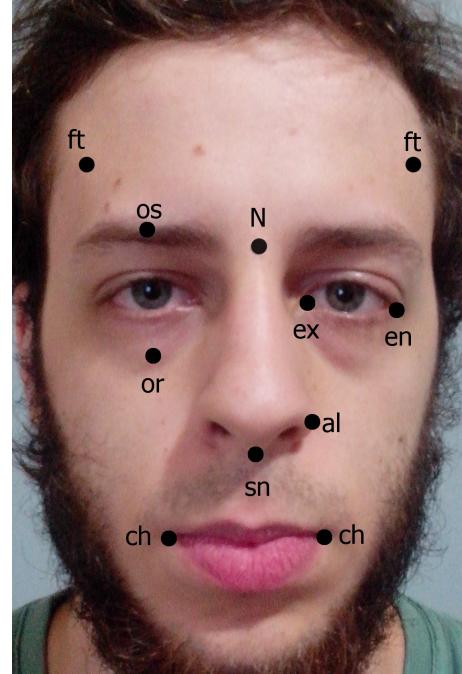


Figure 1: Example of anthropometric landmarks. (Adapted from [8])

portion of the brows is raised; in AU 2 (Figure 2(b)) the outer portion of the eyebrows is raised and in AU 4 (Figure 2(c)) the eyebrows are lowered and drawn closer.

Spline is a mathematical curve defined by two or more control points [9]. The control points create smooth curves and then shape its form. Thus, the definition, modification and manipulation of spline curves are carried out through changes in its control points [9].

An AND-OR graph is a sixfold representing the ‘grammar’ of the image $G = \langle S, V_T, V_N, R, C, P \rangle$, where S is the initial symbol of the grammar, V_T represents the terminal nodes and V_N represents the set of non-terminal nodes, R is the set of production rules productions, C is the set of all valid configurations that may be derived from the G , and P represents the probability model defined in G [16]. Graphically, this grammar is described by an AND-OR graph in which the root node represents the initial symbol and the leaf nodes represent the terminal symbols, the internal nodes represent non-terminal symbols and the vertical edges are related to the productions: the children of an OR node indicate all possible productions for the same non-terminal symbol, and the children of an AND node indicate all symbols that are generated by a particular production [22].

Different from a tree, an AND-OR graph is not necessarily acyclic. In general, the cycles usually occur when a node has more than one parent node or when there are horizontal edges that indicate relationships between single-level nodes. These relationships may indicate distance, adjacency, occlusion, parallelism, perpendicularity, concentricity, and so forth, between two nodes [20, 22]. The horizontal edges add context information between the nodes, creating context-sensitive grammars.

3. CORRELATED WORKS

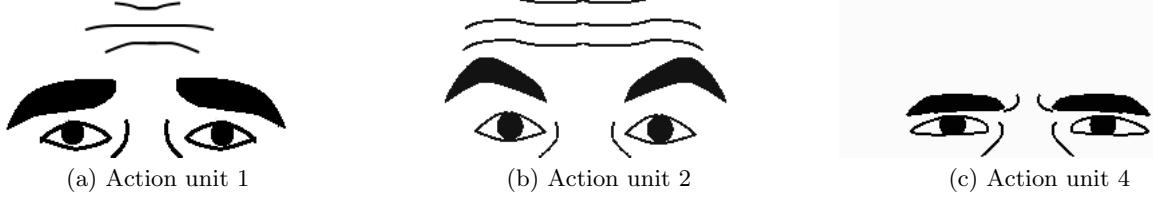


Figure 2: Example of Action Units 1, 2 and 4.

The studies aimed at generating virtual images expressing facial emotions make use of game engines or virtual reality ready packages [1, 4, 14, 2, 15]. However, with regards to flexibility the majority of these resources are limited in generating an adequate variety of emotional intensities and characteristics of avatars such as gender, ethnicity and age, including the limited effortlessness in changing such features to implement these parameters in the game.

No studies were found in the literature that use grammars to specifically generate images of facial emotions. However, one study that uses anthropometry and three that use AND-OR graphs to generate face images with different objectives are shown here to highlight the similar studies consulted in the literature.

A system for creating three-dimensional anthropometric models of the human face is described in [3]. In this work, a set of anthropometric measures are generated to compose different basic models and then these various basic models are employed to generate several combinations of possible face representations.

The article [18] uses AND-OR graphs to represent the aging of faces. The proposed AND-OR graph has three levels: the first level describes the appearance of the face and hair, the second level describes the refined components of the face (eyes, nose, etc), and the third level defines the wrinkles and creases on the skin.

In [20] an AND-OR graph is also used as a face representation model, but the goal is to reconstruct the input images and generate their caricatures. The graph has three decomposition levels: the first level represents the whole face, the intermediate level decomposes the face into its main parts (eyebrows, eyes, nose, mouth and the rest of the face) and the lowest level divides the parts into small fragments. The model has dictionaries of image primitives for every level, from which the reconstructions and generations of caricatures are constructed.

Another example of the application of AND-OR graphs in facial images proposes a tool that can apply cartoon-like styles in face images [21]. This kind of style is achieved by the intensification of colors, intensification of the contours and attenuation of textures. In this work the faces are represented by a two-level AND-OR graph: the first level has a representation of the whole face that occupies the root of the graph, the second level shows the face divided into eyebrows, eyes, nose, mouth and the rest of the face. A syntax analysis algorithm analyzes an input image in order to identify the face components to identify the edges and protrusions that will be highlighted in the next step.

In short, no examples were found related to the AND-OR graph applications to generate faces without the use of an input image. The most frequent structural applications are

recognition, reconstruction and transformation of images.

Furthermore, as observed in [20] and [21], a face AND-OR graph generally has an AND node representing the entire face as root, which immediately decomposes in its main parts. In [20] and in [16], there are graphs that have, in the last decomposition level, nodes that point directly to image primitives, in pixels, already available in libraries. Further on it will be seen that these characteristics are not suitable for the AND-OR graph of this work, and that other approaches were used to replace them.

4. METHODOLOGY

The methodology used in this work is divided into two parts: generation of facial emotion caricatures and definition of the AND-OR graph to represent these caricatures. The caricatures were generated using graphics processing concepts presented in section 2 and detailed in section 4.1. The measures and parts of faces defined in the graphics processing contains the information needed to define the AND-OR graph described in section 4.2.

4.1 Generating the Caricatures

The caricatures to represent the facial expressions were constructed considering that each part of the face is composed of spline curves. Anthropometric measurements were used to define the size of each curve and the correct position of its control points. To create a neutral face, the number of curves and their positioning were determined based on the descriptions of the Action Units (*FACS*). The *FACS* concepts were also used to reposition the control points of the spline curves in accordance with the emotion, in order to reproduce the desired facial expressions.

4.1.1 Using Splines

The design of the curves representing parts of the face is performed by using *B-splines* curves. *B-spline* is a *spline* function of a Bezier curve generalization. Each curve is characterized by the vector of control points that shape it. Cubic polynomials were used to trace the curves because they offer a reasonable compromise between flexibility and processing speed [9].

Most of the curves were generated with three control points, as shown in Figure 3. It should be noted that two points were added to the eyebrow (*pas1 e pas2*) so that the movement result would be the same as that described in the AUs 1, 2 and 4 (Figure 2). These two points were determined by the spline subdivision method described in [13], which adds more points to the curve without changing it. The coordinates of the new point are defined as the average coordinates of two points of the curve.

4.1.2 Anthropometric measurements

Anthropometric measurements vary with the group's ethnicity and with the individuals' age. This work used the data of a population composed of healthy Caucasian North American individuals (19-25 years of age), an age in which adult measurements stabilize [8].

Anthropometric measurements refer to the distances between *landmarks* and not their positions. Therefore, to draw a neutral face required calculating the position of each of these points in the two-dimensional space of the image based on the distance ratio between two points, provided by anthropometric measurements.

, the landmarks were used to define the control points that comprise the curves that make up each part of the face.

Several of the points defined in this work do not have equivalent *landmarks* in anthropometry. To set the position of these points required an approximation based on available anthropometric measurements, while taking into consideration the symmetry of the face. This can be exemplified with the definition of point *p1* (Figure 3). The eyebrow line is positioned horizontally on the same coordinate of point *en*. Besides this point, points *p2* of the eyebrow, *pp1* and *pp2* of the lower eyelid, as well as the wrinkle points, do not have any anthropometric descriptions, they were therefore inferred based on other anthropometric measures in order to represent their curves.

To calculate the positions of the points, a starting point was defined, the *landmarks* *v*, which is based on the other points. Point *v* is positioned at the top of the head and was defined as the upper end and vertical center of the image (Figure 3).

To show how the position of a given point is computed, note the *sn* point, which is drawn at the bottom part of the nose. Its vertical position was defined as the vertical position of point *n* added to the anthropometric measurement *n-sn* and its horizontal position is equal to the horizontal position of *n*. Thus, all control points of all the curves that form the face are defined.

4.1.3 Representation of emotions

To reproduce facial expressions in the *sketches*, the movements produced by facial emotions must be simulated, and this effect is then applied to the curves that compose the face.

The facial emotions to be defined, as well as the neutral expression, were first represented in this study, namely: Satisfaction, sadness, anger, disgust, fear and surprise, as they are universal expressions. Then, the action units that shape these facial expressions were studied, as well as the facial appearance resulting from the influence of each unit [7].

No descriptions were found for the varying intensity of the muscles involved in the action units. For example, in action unit 2, FACS refers to the arching of eyebrows, stretching of the upper eyelid and a possible wrinkling of the forehead (Figure 2(b)). However, the degree of arching, the wrinkle measurement and how much the eyelid is stretched are not defined. This problem is bypassed by determining an imaginary rectangle range for each curve. This rectangle range is defined in order to represent a position variation of twice the size of an internal rectangle that encloses the curve tightly, as illustrated in Figure 4. Thus, repositioning a point is based on a percentage of the corresponding measurement range, because the dynamic generation of expressions allows

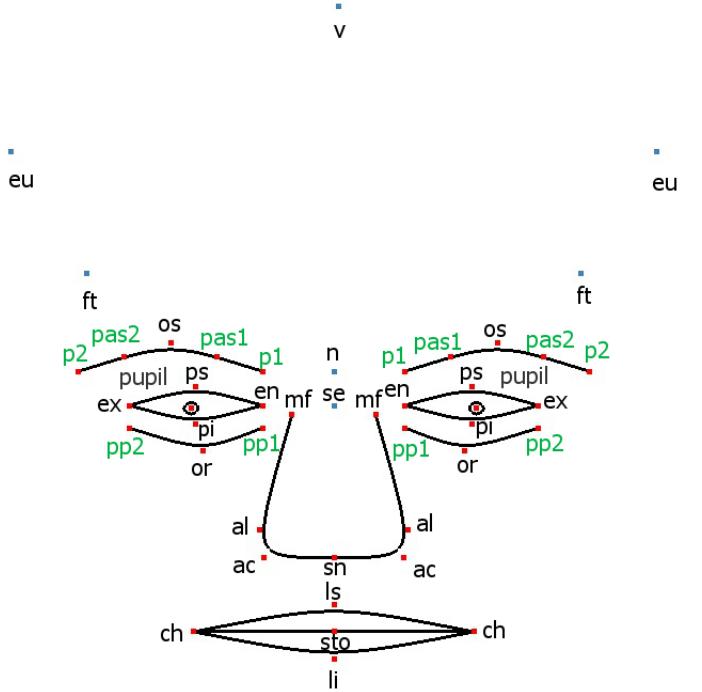


Figure 3: *Landmarks* used for the construction of sketches. The red markings represent the control points used in the curves, and the blue markings are auxiliary points used for positioning the other control points. In addition, green letters describe new points created to build the curves.

the user to set the emotion percentage that should be represented (from zero to 100% for each type of emotion).

The approach presented enables to generate from more subtle expressions to full expressions from the varying control points involved in each action unit of the facial emotion requested. The transitions between two different facial expressions are achieved passing, through the neutral expression. Examples of different percentages for the same expression are shown in Figure 7.

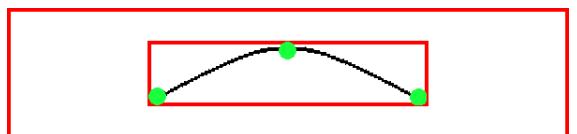


Figure 4: The internal red line shows size of an wrap in relation to the enclosed curve. The outer line delineates a wrap that is twice the size. The black curve corresponds to some part of the face and the green points are its control points.

Since anthropometry does not describe the FACS facial wrinkling, it was necessary to estimate the location of points to generate wrinkles and the variations they undergo. To provide greater realism to some expressions that include wrinkling, the wrinkles move according to that part of the facial movement in the corresponding AU. For example, the wrinkles on the forehead have the same vertical elevation as

the eyebrow to represent the *surprise* expression.

Another effect used was the gradual appearance of wrinkles according to the intensity of the emotion. This effect is important because the depth of the expression wrinkles increases according to the facial muscle movement. To simulate this effect, the thickness of the curves that form the wrinkles varies between zero and the standard thickness used for the other lines in the drawing.

4.2 Construction of the AND-OR graph

The AND-OR graph proposed herein, and described below, is only partially sketched in Figure 5 for the sake of clarity. The Figure shows only the face options of *neutral*, *satisfaction* and *sadness*, and at the intermediate levels there are nodes and edges that are not shown. The last levels (curves and points) are only symbolized in the figure for the sake of clarity.

The full definition of the graph presents an OR node at the root (level 1), which selects one of seven different AND nodes, each symbolizing a face with one of the six basic facial emotions, in addition to the neutral face (level 2). The third level has the OR nodes of the seven main parts that make up any of the face types: eyebrows (left and right), eyes (left and right), nose, mouth and the rest of the face. As these are the OR type, they are also selectors and each have a range of options at level 4 (such as straight mouth, upward curved mouth, downward curved mouth, open mouth, etc.).

In this design, the goal is to generate images by drawing curves without the use of primitive libraries. The technique for drawing the *b-spline* requires only a set of control points as additional information to draw a curve. To do this, two new classes of components were included in the graph: curves and points. Each AND node at the fourth level decomposes into curves, at the fifth level.

The remaining face area will have the curves that define the face contour, as well as several internal curves that serve as facial expression marks and other details. At this fifth level the curves are OR nodes, and also have types to choose from the sixth level. Finally, each curve in an AND node at the sixth level is composed of the control points used as a parameter to generate them. The points are the leaf nodes of the graph and contain their respective *x* and *y* coordinates in the two-dimensional plane.

A given configuration of this graph represents an initial face with a particular emotional expression, but the leaves have a possible positioning range of all control points. Thus, to represent a given image, the exact values of *x* and *y* coordinates of the control points must be defined. This is done by randomly initializing these values in a range of one standard deviation to more or less averages (values defined in the leaves based on anthropometric data). These values are then adjusted according to the desired intensity of that emotion. Therefore, many different images can be generated from a similar configuration.

To generate the image configurations, an algorithm must recursively traverse the graph from the root, extending the search in each AND node and selecting an option from each OR node. During this trajectory, the distance relationships are consulted and the positions of the points are selected and adjusted according to the predefined constraints in the relationships. The black arrows in Figure 5 indicate the edges associated with the productions that create a facial expression of satisfaction.

Most horizontal edges indicate spatial distance relations between the parts, curves and points. These distances, as mentioned earlier, are designated according to anthropometric data.

4.3 Evaluation of the images generated

Volunteers evaluated the realism of the caricatures generated. Forms were constructed using the *GoogleForms* tool, available on the Internet (<http://www.google.com/google-d-s/createforms.html>). The forms contained images of all emotional expressions, and each page showed only one image so as to not confuse the evaluator. Next to each image the choices corresponding to the seven emotions were available, and the evaluator was asked to choose the image that represented the emotion¹. The first page of a demo form is presented in Figure 6.

Demonstration: Facial expressions of emotions demonstration

Our project aims to develop a program able to draw images of human faces with facial expressions automatically. Initially, the images will be simple sketches (cartoon), no colors, textures or 3D appearance. This program can be used for medical purposes like a patient or researcher for patients with psychiatric disorders, who have difficulties in correctly recognizing the facial expressions of emotions.

The following questionnaire aims to evaluate the quality of images generated so far. Below we present a set of 31 images of faces. Each one is associated with one of the six basic social expressions - happiness, sadness, surprise, anger, disgust, fear and neutral expressions - and has a certain degree of intensity.

We request you to check each image which of the seven facial expressions you believe it is better.

Age:

Sex: Female Male

1 - What is the best emotion that describes the picture above?*

- Happiness
- Sadness
- Surprise
- Anger
- Disgust
- Fear
- Neutral

Figure 6: Example of a form used to collect data.

Each form had four images with different intensity levels (25%, 50%, 75% and 100%) of each of the six emotions and an image representing the neutral face. As a comparison mechanism, the six images with 75% intensity were repeated at random positions in the forms, one from each non-neutral emotion, in order to estimate the evaluators' consistency level. Thus, each form comprised 31 images.

In order to minimize the evaluator's fatigue influence, which could influence and affect the recognition of images at the end of the form, the images were randomly distributed, generating ten different versions of the forms.

The participation of the evaluators was voluntary and anonymous. The forms were distributed through electronic messages by the project participants. Age and gender data

¹A demo form is available at: <http://goo.gl/UhXwq7>.

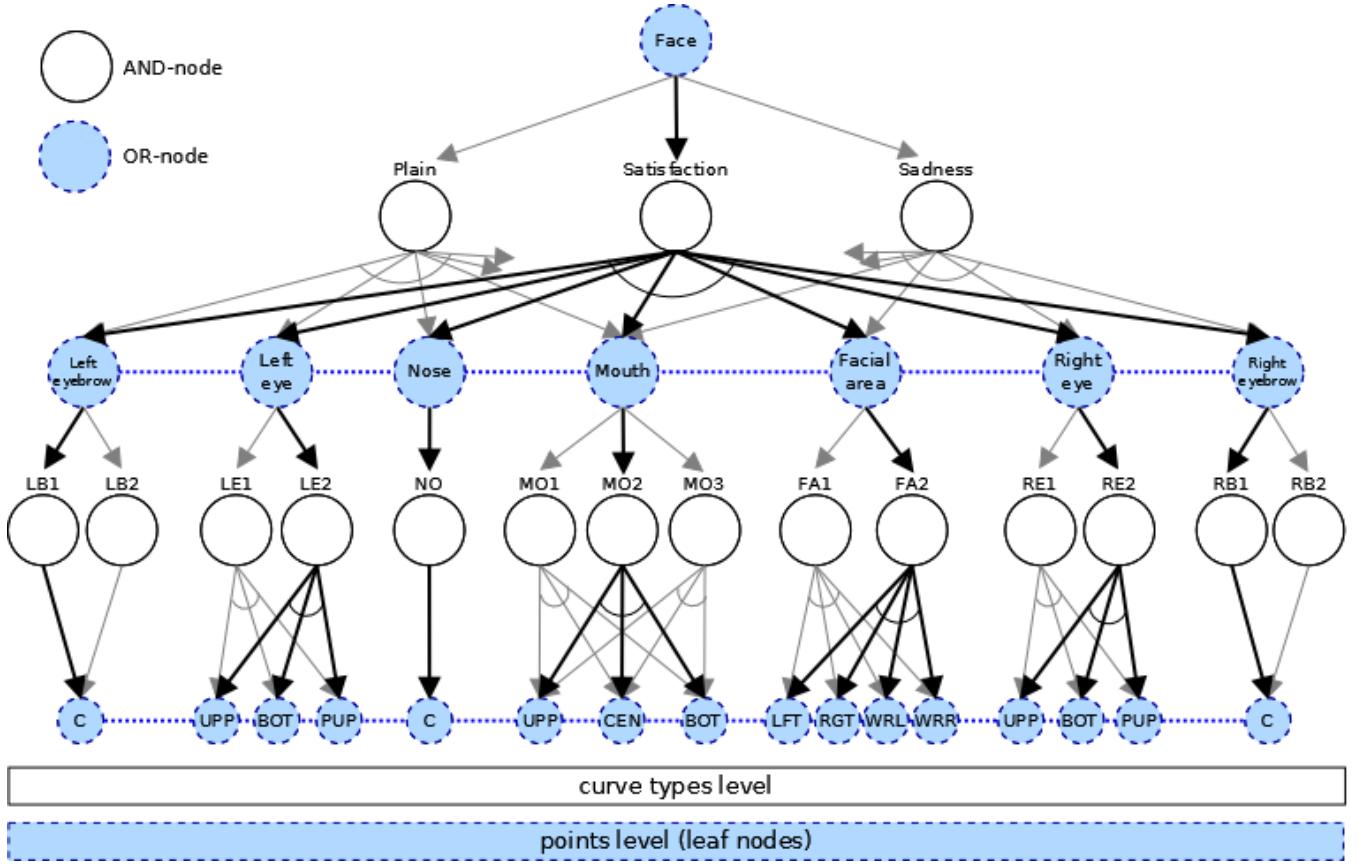


Figure 5: Illustration of partial AND-OR graph with the neutral face and the happiness and sadness emotions.

were also collected to allow verifying perception differences according to these variables.

5. RESULTS

Figure 7 illustrates the caricatures generated for the evaluation process, taking into consideration the intensities described in the previous section.

The evaluation of the images included 237 healthy volunteer evaluators, with ages ranging from 13 to 77 years, 54% of all evaluators were men.

Three age groups were considered in the results presentation: less than 25, 26 to 55 and greater than 55 years. The percentage of participants in each group was, respectively, 32% (77 volunteers), 61% (145 volunteers) and 6% (15 volunteers).

Figure 8 and Table 1 show the overall results, in other words, irrespective of gender or age. Figure 8 shows each curve that represents the percentage of correct answers for each of the intensities used for each emotion (25%, 50%, 75%, 100%). Table 1 shows the cross-referencing of responses between the different types of emotions represented in the images. It should be noted that for the 75% intensity of each emotion an average was taken because of duplicate images.

Figure 9 shows the rates of correct responses with regards to the male and female respondents, but disregarding the intensities.

Finally, Figure 10 shows the rate of correct responses

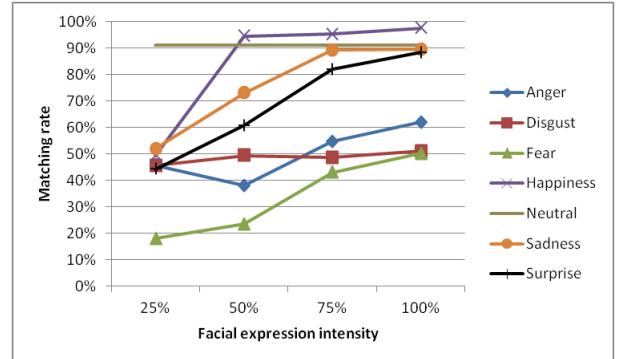


Figure 8: Rate of correct responses for each emotion evaluated.

within each age group, considering three age groups.

A final analysis was carried out for the responses. As described in the previous section, images were repeatedly included in the forms as a control form. Table 2 shows the g levels of responses of repeated images, considering an overall analysis and divisions by gender and age group.

6. DISCUSSION

Though different intensities were used for each emotion in the interviews, the correct response considered for an image

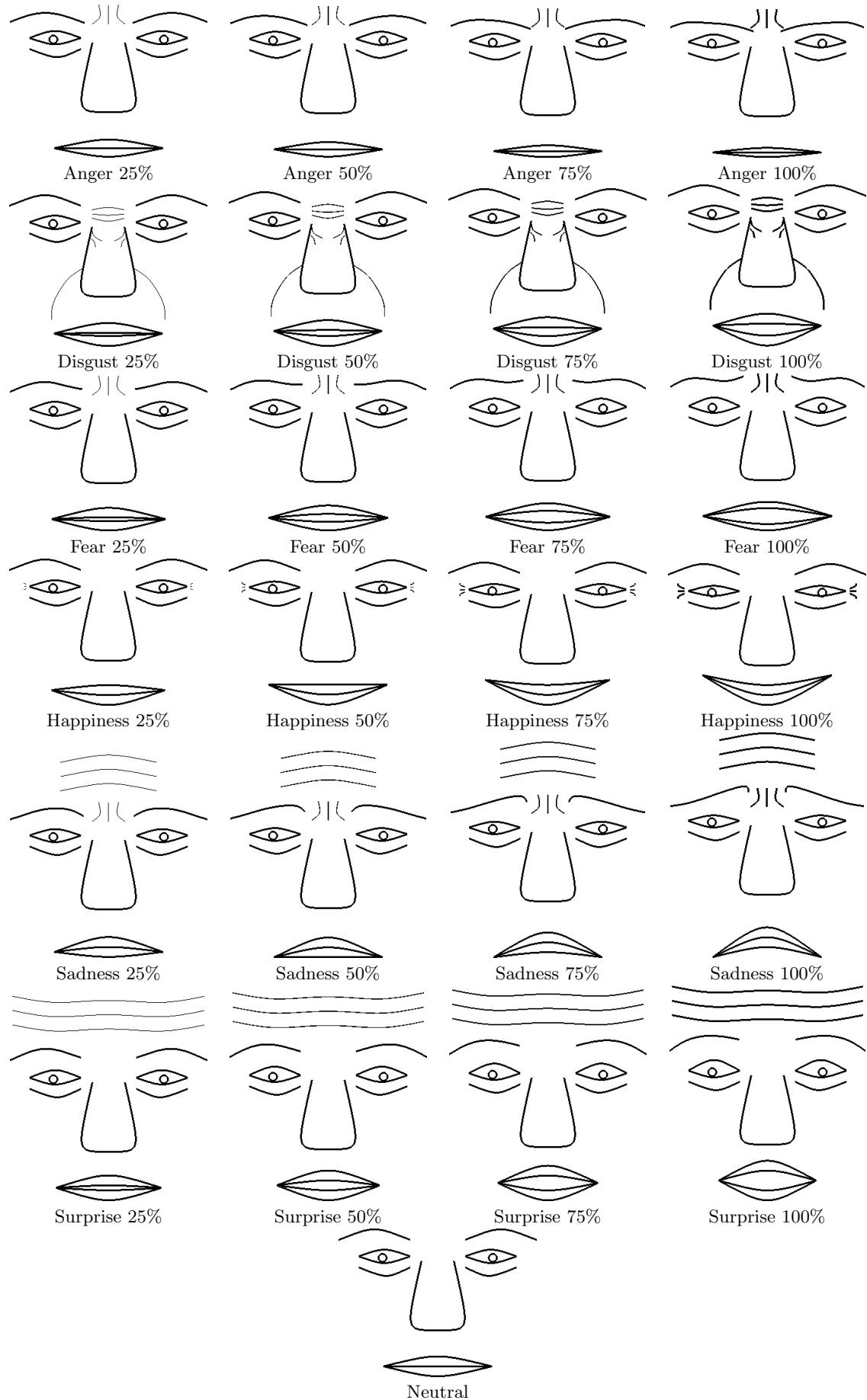


Figure 7: Facial expressions used in the evaluation.

Table 1: Confusion matrix of the responses obtained. Boldface indicates recognition rates of intended emotion.

	Anger	Disgust	Fear	Happiness	Neutral	Sadness	Surprise
Anger	51%	7%	6%	1%	27%	5%	3%
Disgust	38%	49%	2%	2%	5%	2%	2%
Fear	9%	7%	36%	2%	12%	13%	21%
Happiness	0%	0%	0%	86%	10%	2%	2%
Neutral	1%	0%	2%	2%	91%	2%	2%
Sadness	6%	5%	6%	0%	3%	79%	1%
Surprise	0%	1%	8%	5%	13%	2%	71%

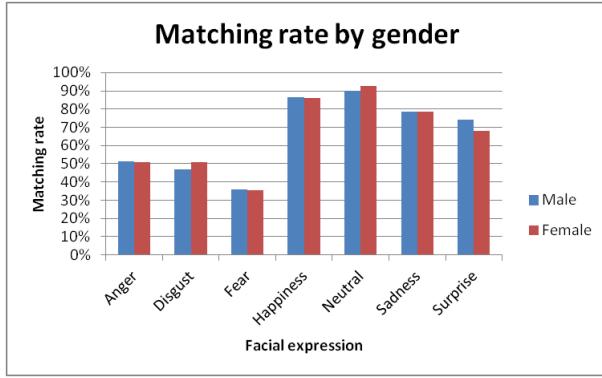


Figure 9: Rate of correct responses for each gender.

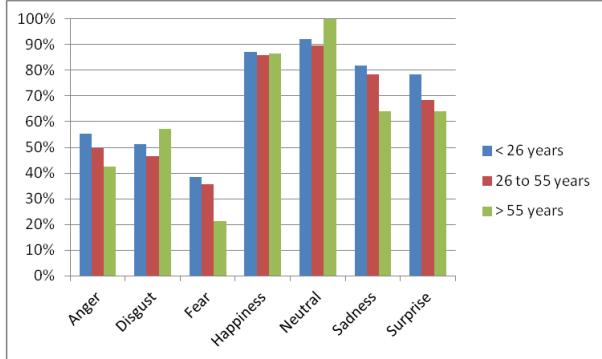


Figure 10: Rate of correct responses for each age group.

Table 2: Consistency of responses regarding the duplicate images.

	<25%	26 a 50%	51 a 75%	>75%
Global	1%	28%	33%	38%
Male	2%	28%	29%	41%
Female	1%	28%	37%	34%
<26 years	0%	24%	36%	40%
26-55 years	2%	27%	32%	39%
>55 years	0%	56%	25%	19%

was the emotion related to that image, though no intensity related questions were asked. It should be noted that since this is an ongoing work, we used healthy volunteers in order to assess the quality of the representation of emotions in the images generated. Therefore, the error rates are not indicative of the evaluators' inability to recognize facial emotions,

but rather methodology deficiencies to be remedied in the representation of emotions in the caricatures.

6.1 Overall evaluation of the caricatures

Figure 8 shows that, except for the neutral face and the *disgust* emotion, the correct recognition rate increased as the intensity of the emotions represented also increased. Another exception regards the *anger* emotion, which at the 50% intensity showed a decrease in the correct response rate, when this emotion was often confused with the *neutral* emotion. The two emotions differ in the downward arching of the mouth and the inclusion of wrinkles (frown) between the eyes in the *anger* emotion, which are subtle at intensities of 25 and 50%.

The images related to the *fear* emotion were those that had the lowest recognition levels in all intensities, followed by images related to *disgust* and *anger* emotions. The analysis of the images representing these emotions shows that these are the caricatures with the least changes in the lower face (mouth region). This suggests that the volunteers can easily recognize the caricatures with more pronounced changes in this region.

It also shows that with a 25% intensity, all emotions showed low recognition rates, evidencing that the changes in the caricatures are very subtle and hard to recognize. This is true even for the emotions with high recognition rates in the other intensities. Interestingly, the intensity level of 25% produces little change in the lower face region, which reinforces the importance of the changing traits in this region.

Table 1 provides more detailed data regarding the confusion in recognizing the emotions. Analyzing the *fear* emotion on the third line of this table demonstrates that there is a lower recognition rate (36%). The images that are related to this emotion were recognized by the volunteers as belonging to nearly all the other emotions. The emotion in which the *surprise* images was most confused regarded the *surprise* emotion. Analyzing the images of these two emotions it was noted that the lower facial sides are similar, increasing the mouth opening according to the intensity. Thus, it evidences that this facial change was related to the *surprise* emotion and that the caricatures related to the *fear* emotion should be reassessed in order to add or change features that can better represent this emotion.

A similar analysis can be done for the *disgust* emotion, which was the second one with the lowest rate of correct responses (second line of Table 1). This emotion was often confused with the *anger* emotion. The images indicate that the main differences between the caricatures of these two emotions are the forehead lines and the lines added to the nose in the *disgust* emotion. These lines on the nose represent the change in the AU 9, but not enough to clearly express the aforementioned emotion, something that should

be improved in the caricatures. Similarly, the users confused the *anger* emotion with the neutral face, indicating that the features of this emotion should be more intense.

With the exception of the three emotions analyzed, the main diagonal of the confusion matrix shown in Table 1 indicates that all other emotions presented satisfactory recognition rates. The *happiness* emotion presented the highest correct recognition rate, followed by the expressions of *sadness* and *surprise* emotions. These data suggest that the caricatures were properly recognized by the evaluators, especially with intensities of 75 and 100%.

These results agree with the results described in [4], which report that the expressions of *happiness*, *sadness* and *neutral* emotions had high recognition rates, while the *disgust* emotion was vaguely recognized. Dyck and colleagues, however, achieved satisfactory recognition rates for the *fear* and *anger* emotions [4].

6.2 Demographic variables and consistency of responses

Figure 9 shows that the correct recognition rates of the emotions did not vary significantly in terms of the evaluator's gender.

Similarly, the results analysis of the three age groups (Figure 10) provides no significant result differences for this variable. However, some observations, although preliminary, can be discussed. The group of older volunteers (over 55) was the group with the highest rate of correct responses for the neutral face and for the expression representing the disgust emotion. They were also the ones that least recognized the other emotions, except for the happiness emotion, which had a more consistent recognition rate among the age groups. The youngest group (under 25) was the one that most recognized the expressions of *sadness* and *surprise* emotions. However, the graph shows that the percentages are similar, making it difficult to attain a more conclusive analysis about the influence of this variable.

As mentioned in section 4.3, the evaluation of the volunteers included a mechanism to verify the consistency of answers by repeating the image with 75% of intensity for each emotion. Table 2 shows the results in terms of an overall analysis, gender and age group classification. Each column shows the percentage of subjects that attained a particular consistency range. A given result was regarded as consistent when the evaluator indicated the same emotion for two identical caricatures in the form, although this emotion was not correct.

The first line of the table indicates that over 70% of the volunteers attained over 50% of consistency. Lines 2 and 3 indicate that there was no significant difference by gender: both genders showed that about 70% of the volunteers demonstrated a higher consistency range (50% or higher).

However, consistency reduced as the volunteers' age increased. In the youngest age group (under 25) 76% of the volunteers showed a higher consistency range. However, the volunteers with ages between 26 and 55 as well as over 55 years showed, respectively, lower consistency averages of 71% and 44%. Although the lower number of volunteers over 55 may have influenced these percentages, it should be highlighted that the older evaluators were more confused with the caricatures than the younger ones.

7. CONCLUSION

The approach used to produce caricatures with different emotional intensities achieved encouraging results with regards to the evaluation conducted by healthy volunteers.

The caricatures of the neutral face and the expressions of *happiness*, *sadness* and *surprise* emotions were clearly defined by the traces using *spline* curves, however the *fear*, *disgust* and *anger* emotions included caricatures that had a limited recognition range. There was a certain similarity among features, since the activated AUs are similar. Although this fact may appear as a limitation to the approach, the current version of the project used a number of curves that can be increased. Therefore, we believe that refining the features and including new features can reduce the confusion in representing these emotions as well as improving the results.

The importance of undertaking the task to generate virtual images instead of using real images should be emphasized. Virtual images for identifying emotions can be equally effective when compared with real images in research applications [4, 1]. Moreover, virtual images make it easier to implement applications for the manipulation and animation (dynamic changes between intensities and emotions) of faces, hence enabling greater precision [1]. For example, it is not possible to objectively define how a 75% expression of *disgust* should be made. In addition, a real person may not be able to separate the specific muscles of a particular emotion, which results in mixed expressions, consequently making them more difficult to recognize. The specific goal of this work regards creating a game for training individuals with emotion recognition impairment, and a less complex virtual scenario could represent a preparation process for real-life situations.

8. ACKNOWLEDGMENTS

This research was supported by the State of São Paulo Research Foundation (FAPESP grant #2011/50761-2, CNPq, CAPES, NAP eScience - PRP - USP), Brazilian National Council of Scientific and Technological Development (Processes 559931/2010-7 and 401745/2013-9 and PIBIC Program), the National Institute of Science and Technology Medicine Assisted by Scientific Computing and Provost's Office for Graduate Studies of the University of São Paulo. We would like to thank the BIOINFO-Vision Laboratory (University of São Paulo) for computing facilities.

9. REFERENCES

- [1] E. Bekele, J. Crittenden, Z. Zheng, A. Swanson, A. Weitlauf, Z. Warren, and N. Sarkar. Assessing the utility of a virtual environment for enhancing facial affect recognition in adolescents with autism. *Journal of Autism and Developmental Disorders*, Jan 2014.
- [2] Y. Cheng and S. Ling. 3d animated facial expression and autism in taiwan. In *Advanced Learning Technologies, 2008. ICALT'08. Eighth IEEE International Conference on*, pages 17–19. IEEE, 2008.
- [3] D. DeCarlo, D. Metaxas, and M. Stone. An anthropometric face model using variational techniques. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '98*, pages 67–74, New York, NY, USA, 1998. ACM.

- [4] M. Dyck, M. Winbeck, S. Leiberg, Y. Chen, R. C. Gur, and K. Mathiak. Recognition profile of emotions in natural and virtual faces. *PLOS One*, 3(11):e3628, 2008.
- [5] P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2):124, 1971.
- [6] P. Ekman and W. V. Friesen. Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, 1(1):56–75, 1976.
- [7] P. Ekman and W. V. Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, 1978.
- [8] L. Farkas. *Anthropometry of the head and face*. Raven Press, 1994.
- [9] J. D. Foley. *Computer graphics: Principles and practice*, in *C*, volume 12110. Addison-Wesley Professional, 1996.
- [10] P. Gosselin, G. Kirouac, et al. Decoding facial emotional prototypes. *Canadian journal of experimental psychology*, 49(3):313, 1995.
- [11] O. Grynszpan, J.-C. Martin, and J. Nadel. Multimedia interfaces for users with high functioning autism: An empirical investigation. *International Journal of Human-Computer Studies*, 66(8):628–639, 2008.
- [12] M. B. Harms, A. Martin, and G. L. Wallace. Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychology review*, 20(3):290–322, 2010.
- [13] D. Hearn and M. P. Baker. *Computer graphics, C version*, volume 2. Prentice Hall Upper Saddle River, 1997.
- [14] J. Ku, H. J. Jang, K. U. Kim, J. H. Kim, S. H. Park, J. H. Lee, J. J. Kim, I. Y. Kim, and S. I. Kim. Experimental results of affective valence and arousal to avatar’s facial expressions. *Cyber Psychology & Behavior*, 8(5):493–503, 2005.
- [15] U. Lahiri, E. Bekele, E. Dohrmann, Z. Warren, and N. Sarkar. Design of a virtual reality based adaptive response technology for children with autism. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 21(1):55–64, 2013.
- [16] L. Lin, T. Wu, J. Porway, and Z. Xu. A stochastic graph grammar for compositional object representation and recognition. *Pattern Recogn.*, 42(7):1297–1307, July 2009.
- [17] C. C. d. A. Rocca, E. v. d. Heuvel, S. C. Caetano, and B. Lafer. Facial emotion recognition in bipolar disorder: a critical review. *Revista Brasileira de Psiquiatria*, 31(2):171–180, 2009.
- [18] J. Suo, S.-C. Zhu, S. Shan, and X. Chen. A compositional and dynamic model for face aging. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(3):385–401, March 2010.
- [19] S. F. Taylor and A. W. MacDonald. Brain mapping biomarkers of socio-emotional processing in schizophrenia. *Schizophrenia bulletin*, 38(1):73–80, 2012.
- [20] Z. Xu, H. Chen, S.-C. Zhu, and J. Luo. A hierarchical compositional model for face representation and sketching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(6):955–969, June 2008.
- [21] M. Yang, S. Lin, P. Luo, L. Lin, and H. Chao. Semantics-driven portrait cartoon stylization. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 1805–1808, Sept 2010.
- [22] S.-C. Zhu and D. Mumford. *A stochastic grammar of images*. Now Publishers Inc, 2007.

Apêndice B

Avaliação de desempenho do aluno

EACH –USP - Programa PIBIC

Orientador: Fátima de Lourdes dos Santos Nunes Marques

Aluno: Rafael Luiz Testa

Avaliação do Orientador sobre o aluno

Este é o segundo projeto de Iniciação Científica que o aluno conduz sob minha orientação. O anterior foi com bolsa PIBIT. O aluno tem uma grande capacidade técnica e iniciativa para resolver os problemas. Especificamente neste projeto ele mesmo descobriu importantes fontes de literatura sem auxílio da orientadora. Foi assíduo nas reuniões com a Orientadora e com o grupo de pesquisa. Foi possível perceber que a participação no PIBIC e o contato com outros alunos do grupo de pesquisa, principalmente com aqueles cujos trabalhos estão diretamente relacionados com o tema do seu projeto, permitiu seu crescimento intelectual e científico. Este projeto faz parte de um escopo maior, envolvendo mais uma docente e outro aluno de Iniciação Científica. O Rafael Testa tem contribuído diretamente com o desenvolvimento de tal projeto, mostrando uma generosidade ímpar no compartilhamento de informações. Além disso, tem mostrado esforço para aprender e ao longo do projeto superou muitas limitações em relação à escrita científica. Além da capacidade técnica destaca-se o interesse por estudos avançados, tendo já cursado uma disciplina do Programa de Pós-Graduação em Sistemas de Informação. Destaca-se, ainda, sua facilidade para trabalhar em equipe e acatar sugestões da orientadora e do grupo de pesquisa, características que permitiram que seu crescimento fosse além do esperado. Sua dedicação contribuiu para a obtenção de um ótimo resultado do projeto tendo gerado, inclusive, um artigo do grupo que será submetido a uma importante conferência internacional na área de Computação (*ACM Symposium on Applied Computing*, que atualmente é classificada no estrato A1 na área de Ciência da Computação pela Capes), tendo este aluno como primeiro autor. Desta forma, considero que o aluno apresentou desempenho além do esperado para um projeto PIBIC.

Apêndice C

Avaliação do Programa

Este documento tem por objetivo apresentar as vantagens e desvantagens vivenciadas pelo autor por meio do Programa de Iniciação Científica/PIBIC.

Após dois anos de aprendizado com a iniciação científica posso dizer que já conquistei conhecimento suficiente para me qualificar como pesquisador iniciante, porque pude compreender sobre quais são as bases que constituem uma pesquisa científica, como se organizar pessoalmente e a vantagem de um aprendizado compartilhado com meu colegas de laboratório e minha orientadora.

Durante esse projeto conquistei algo mais: o aprendizado interdisciplinar. Fui exposto a áreas que aparentemente não possuíam nada em comum, porém durante o desenvolvimento aprendi que existe um grande ganho de conhecimento quando se consegue integrar as informações contidas em campos de pesquisas bem distintos. Apesar das dificuldades de fazer com que diferentes áreas do conhecimento conversassem, acredito que esse tenha sido mais um grande ganho pessoal.

No futuro irei colher os frutos dessa capacitação na área científica, mas acredito que independentemente de eu seguir a carreira acadêmica ou não, ainda assim serei um profissional melhor graças ao que esse projeto me proporcionou.

São Paulo, 31 de julho de 2014.

Rafael Luiz Testa

Rafael Luiz Testa
Bolsista