

SÍDNEI AUGUSTO DROVETTO JUNIOR

RECONHECIMENTO FACIAL 3D UTILIZANDO O *SIMULATED ANNEALING* COM AS MEDIDAS *SURFACE INTERPENETRATION MEASURE* E *M-ESTIMATOR SAMPLE CONSENSUS*

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre. Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná.

Orientadora: Profa. Dra. Olga R. P. Bellon

Co-orientador: Prof. Dr. Luciano Silva

CURITIBA

2007

SÍDNEI AUGUSTO DROVETTO JUNIOR

RECONHECIMENTO FACIAL 3D UTILIZANDO O *SIMULATED ANNEALING* COM AS MEDIDAS *SURFACE INTERPENETRATION MEASURE* E *M-ESTIMATOR SAMPLE CONSENSUS*

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre. Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná.

Orientadora: Profa. Dra. Olga R. P. Bellon

Co-orientador: Prof. Dr. Luciano Silva

CURITIBA

2007

“Don’t be afraid to give up the good to go for the great”

John D. Rockefeller (1839 – 1937)

“Whether you believe you can do a thing or not, you’re right”

Henry Ford (1863 – 1947)

*“You see things; and you say ‘Why?’ But I dream things that never
were; and I say ‘Why not?’ ”*

George Bernard Shaw (1856 – 1950)

“When you reach the end of your rope, tie a knot in it and hang on”

Thomas Jefferson (1743 – 1826)

“Thought is action in rehearsal”

Sigmund Freud (1856 – 1939)

“Nam et ipsa scientia potestas est”

Francis Bacon (1561 – 1626)

“E pur si muove!”

Galileo Galilei (1564 – 1642)

AGRADECIMENTOS

A Deus, que me deu forças e saúde para que pudesse realizar este trabalho.

Aos meus orientadores, Profa. Dra. Olga R. P. Bellon e Prof. Dr. Luciano Silva, por estarem sempre dispostos a sanar minhas constantes dúvidas, pelas oportunidades de aprendizado e pela atenciosa orientação.

Aos meus pais, Sidnei Augusto Drovetto e Elizabeth Silva Drovetto, e irmãos, Cristhian Augusto Drovetto e Graziella Gabeline Drovetto, que sempre me apoiaram.

Aos amigos e colegas do Imago pela insubstituível companhia ao longo dos anos.

Ao CNPq pelo apoio financeiro.

A todas as demais pessoas que direta ou indiretamente contribuíram para que este trabalho pudesse ser concluído.

SUMÁRIO

Lista de Figuras	v
Lista de Tabelas	viii
Lista de Abreviaturas	ix
RESUMO	xi
ABSTRACT	xii
1 INTRODUÇÃO	1
1.1 MOTIVAÇÃO	1
1.2 REGISTRO	2
1.3 BIOMETRIA	3
1.4 RECONHECIMENTO FACIAL 3D	4
2 CONCEITOS	5
2.1 BIOMETRIA	5
2.1.1 Características Biométricas	7
2.1.2 Avaliação de Sistemas Biométricos de Verificação	9
2.1.2.1 Taxa de Falsa Aceitação	10
2.1.2.2 Taxa de Falsa Rejeição	10
2.1.2.3 Taxa Equivalente de Erro	12
2.1.2.4 Taxa de Verificação	15
2.1.3 Avaliação de Sistemas Biométricos de Identificação	15
2.2 RECONHECIMENTO FACIAL 3D	17
2.2.1 Aquisição	18
2.2.2 Pré-processamento	20
2.2.3 Representação das Faces	24
2.2.4 Desafios	24
2.3 <i>SIMULATED ANNEALING</i>	26
2.3.1 Esquema de Resfriamento	29
2.3.1.1 Temperatura Inicial	30
2.3.1.2 Temperatura Final	31
2.3.1.3 Decremento da Temperatura	32
2.3.1.4 Número de Iterações em Cada Temperatura	32
2.3.2 Decisões Específicas do Problema	33

2.3.2.1	Função de Custo	33
2.3.2.2	Estrutura de Vizinhaça	34
2.3.2.3	Espaço de Solução	35
2.3.2.4	Sumário	35
2.4	MEDIDAS DE ERRO	35
2.4.1	<i>M-estimator Sample Consensus</i>	36
2.4.2	<i>Surface Interpenetration Measure</i>	37
3	TRABALHOS RELACIONADOS	40
4	ALGORITMOS DESENVOLVIDOS	46
4.1	ESTÁGIOS	47
4.1.1	Pré-alinhamentos	47
4.1.1.1	Centro de Massa	48
4.1.1.2	Ponta do Nariz	48
4.1.1.3	<i>Bounding Box</i>	49
4.1.2	<i>Simulated Annealing</i> com o MSAC	49
4.1.3	<i>Simulated Annealing</i> com a SIM	51
4.2	VERSÃO PRELIMINAR	55
4.2.1	Resultados Experimentais	56
4.2.1.1	Base de Dados	56
4.2.1.2	Resultados	58
4.2.1.3	Exemplos de Resultados	60
4.3	VERSÃO ATUAL	73
4.3.1	Resultados Experimentais	77
4.3.1.1	Base de Dados	77
4.3.1.2	Resultados	78
4.4	VANTAGENS VS. DESVANTAGENS E LIMITAÇÕES	100
5	CONCLUSÕES	102
	REFERÊNCIAS BIBLIOGRÁFICAS	104
A	COMPARAÇÃO DAS ABORDAGENS: ICP, GA E SA	108

LISTA DE FIGURAS

1.1	Dimensões utilizadas para solucionar o problema de registro de imagens tridimensionais.	3
2.1	Exemplos de algumas características biométricas separadas em classes.	8
2.2	Distribuição do escore de impostores e sua FAR.	11
2.3	Distribuição do escore de clientes e sua FRR.	12
2.4	Distribuição dos escores de clientes e impostores e sua EER.	13
2.5	Curva ROC, relacionando a FAR e a FRR.	14
2.6	Curva CMC.	17
2.7	Esquema de funcionamento de aquisição baseada em visão estéreo.	18
2.8	Esquema de funcionamento de aquisição baseada em luz estruturada.	19
2.9	Esquema de funcionamento de aquisição baseada em sensor laser (triangulação).	20
2.10	Efeito da variação de iluminação na aquisição de uma imagem 3D da face	21
2.11	Exemplo de imagem 3D com “buracos” e que foram removidos através de pré-processamento	22
2.12	Exemplo de ruído do tipo <i>spike</i>	22
2.13	Exemplo de imagem 3D onde o ruído do tipo <i>spike</i> foi removido por um pré-processamento	23
2.14	Exemplo de como o mapa de textura pode ocultar imperfeições do modelo 3D	23
2.15	Mínimo global e mínimos locais.	26
2.16	Fluxograma do <i>Simulated Annealing</i>	30
2.17	Ilustração do funcionamento do MSAC.	37
2.18	Ponto interpenetrado p de A em relação a B	38
2.19	Diferentes alinhamentos de duas imagens A e B , todas com baixo MSE.	39
3.1	Segmentações do nariz utilizadas em [14, 16, 17].	41
3.2	Pontos selecionados para alinhamento grosseiro, segundo o método descrito em [26, 27]	44
3.3	Distâncias entre pontos correspondentes.	44
4.1	Erro absoluto da SIM esparsa em relação à SIM normal.	52
4.2	Erro relativo da SIM esparsa em relação à SIM normal.	53
4.3	Diagrama do SAM na sua versão preliminar.	56
4.4	Exemplos de imagens da base de dados BEE.	57
4.5	Exemplos de imagens com ruído	58
4.6	Regiões utilizadas das imagens	58
4.7	Exemplo de registro e pré-alinhamento para indivíduos iguais (com expressão).	62

4.8	Exemplo de registro e pré-alinhamento para indivíduos iguais (sem expressão).	63
4.9	Exemplo de registro e pré-alinhamento para indivíduos diferentes.	64
4.10	Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos iguais.	65
4.11	Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos iguais.	66
4.12	Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos iguais.	67
4.13	Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos iguais.	68
4.14	Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos diferentes.	69
4.15	Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos diferentes.	70
4.16	Exemplo de registro, pré-alinhamento para indivíduos iguais e com faces muito desalinhadas.	71
4.17	Faces desalinhadas da Figura 4.16 e gráficos de convergência do MSAC e da SIM	72
4.18	Tempo de execução do SAM para diferentes regiões da face.	73
4.19	Diagrama do SAM na versão atual.	75
4.20	Pesos utilizados para combinar os valores obtidos em cada segmentação da face.	77
4.21	Distribuição das combinações para a região do nariz.	79
4.22	Distribuição das combinações para a região elíptica ao redor do nariz.	80
4.23	Distribuição das combinações para a região circular ao redor do nariz.	80
4.24	Distribuição das combinações para a região dos olhos.	81
4.25	Distribuição das combinações para a região da face.	81
4.26	Distribuição das combinações para a combinação de todas as regiões.	82
4.27	Curvas de FAR e de FRR de acordo com o limiar, para a versão atual do algoritmo	83
4.28	Cruzamento das curvas de FAR e de FRR, para a versão atual do algoritmo . .	84
4.29	Curvas ROC e DET, para a versão atual do algoritmo	85
4.30	Curva CMC, para a versão atual do algoritmo.	86
4.31	Exemplo de alinhamento das segmentações utilizadas pelo algoritmo para indivíduos iguais	88
4.32	Exemplo de alinhamento das segmentações utilizadas pelo algoritmo para indivíduos iguais	89
4.33	Exemplo de alinhamento das segmentações de um mesmo indivíduo que foi falsamente rejeitado	90
4.34	Exemplo de alinhamento das segmentações de um mesmo indivíduo que foi falsamente rejeitado	91

4.35 Exemplo de alinhamento das segmentações de um mesmo indivíduo que foi falsamente rejeitado	92
4.36 Exemplo de alinhamento das segmentações de um mesmo indivíduo que foi falsamente rejeitado	93
4.37 Exemplo de alinhamento das segmentações de um mesmo indivíduo	94
4.38 Exemplo de alinhamento das segmentações de um mesmo indivíduo	95
4.39 Exemplo de alinhamento das segmentações de um mesmo indivíduo	96
4.40 Exemplo de alinhamento das segmentações de um mesmo indivíduo	97
4.41 Exemplo de alinhamento das segmentações de indivíduos diferentes	98
4.42 Alinhamento das segmentações de indivíduos diferentes, que produziram o maior valor de SIM	99
A.1 Exemplo de alinhamento 1	111
A.2 Exemplo de alinhamento 2	111
A.3 Exemplo de alinhamento 3	112
A.4 Exemplo de alinhamento 4	112
A.5 Exemplo de alinhamento 5	113
A.6 Exemplo de alinhamento 6	113

LISTA DE TABELAS

2.1	Exemplos de características biométricas, suas vantagens e desvantagens	9
2.2	Relação entre o <i>Simulated Annealing</i> e o processo real de <i>annealing</i>	27
3.1	Efeito da expressão em faces 3D no desempenho das abordagens baseadas em ICP e PCA.	40
3.2	Taxas de reconhecimento <i>rank-one</i> obtidas pelo algoritmo ARMS.	41
3.3	Taxas de reconhecimento <i>rank-one</i> obtidas pelo algoritmo descrito em [16].	42
3.4	Taxas de reconhecimento <i>rank-one</i> obtidas pelo algoritmo descrito em [26].	44
3.5	Taxas de reconhecimento <i>rank-one</i> obtidas pelo algoritmo descrito em [27].	45
4.1	Taxa de verificação, usando o registro com ICP e FAR de 0%.	59
4.2	Taxa de verificação para combinação das medidas, usando o registro com ICP e FAR de 0%.	59
4.3	Taxa de verificação usando o registro com SA e FAR de 0%.	60
4.4	Taxa de verificação para combinação das medidas, usando o registro com <i>Simulated Annealing</i> (SA) e FAR de 0%.	60
4.5	Taxa de verificação e limiar para as segmentações com FAR de 0% da base <i>Face Recognition Grand Challenge</i> (FRGC) v2.0.	78
4.6	Taxa de reconhecimento <i>rank-one</i> e número de identificações incorretas da base FRGC v2.0.	85
4.7	Comparação do SAM com os algoritmos [16, 17, 26, 27]	100
A.1	Comparação do tempo de alinhamento para os métodos ICP, GA e SA.	108
A.2	Comparação do valor médio de SIM para clientes dos métodos ICP, GA e SA.	109
A.3	Comparação do valor médio de SIM para impostores dos métodos ICP, GA e SA.	109
A.4	Comparação do valor médio de MSAC para clientes dos métodos ICP, GA e SA.	109
A.5	Comparação do valor médio de MSAC para impostores dos métodos ICP, GA e SA.	109
A.6	Comparação do valor médio de MSE para clientes dos métodos ICP, GA e SA.	110
A.7	Comparação do valor médio de MSE para impostores dos métodos ICP, GA e SA.	110
A.8	Comparação de alinhamentos para os métodos ICP, GA e SA.	110

LISTA DE ABREVIATURAS

ARMS	<i>Adaptive Rigid Multi-region Selection</i>
BEE	<i>Biometric Experimentation Environment</i>
CMC	<i>Cumulative Match Characteristic</i>
DET	<i>Detection Error Trade-off</i>
EER	<i>Equal Error Rate</i>
FA	<i>False Acceptance</i>
FAR	<i>False Acceptance Rate</i>
FR	<i>False Rejection</i>
FRGC	<i>Face Recognition Grand Challenge</i>
FRR	<i>False Rejection Rate</i>
FRVT	<i>Face Recognition Vendor Test</i>
GA	<i>Genetic Algorithm</i>
GSL	<i>GNU Scientific Library</i>
ICP	<i>Iterative Closest Point</i>
LDA	<i>Linear Discriminant Analysis</i>
MSAC	<i>M-estimator Sample Consensus</i>
MSE	<i>Mean Squared Error</i>
PCA	<i>Principal Components Analysis</i>
RANSAC	<i>Random Sample Consensus</i>
RMSE	<i>Root Mean Squared Error</i>
ROC	<i>Receiver Operating Characteristic</i>
ROI	<i>Region Of Interest</i>
SA	<i>Simulated Annealing</i>

- SAM** *Simulated Annealing Matcher*
- SIM** *Surface Interpenetration Measure*

RESUMO

Este trabalho apresenta uma nova abordagem para o problema de reconhecimento facial 3D. Muitos trabalhos presentes na literatura empregam o *Iterative Closest Point* (ICP) como técnica de alinhamento das faces e, portanto, baseia seus resultados na similaridade provida pelas impróprias medidas *Mean Squared Error* e *Root Mean Squared Error*. Este trabalho propõe a utilização do algoritmo estocástico *Simulated Annealing* baseado em duas medidas robustas, *M-estimator Sample Consensus* e *Surface Interpenetration Measure* (SIM), para obter um alinhamento preciso das faces. A conclusão acerca da similaridade das faces é feita utilizando a SIM, que provê uma maior discriminação entre imagens da face de um mesmo indivíduo e de indivíduos diferentes.

Neste trabalho utilizou-se diferentes técnicas de alinhamento, como por exemplo o ICP e algoritmo genético com a SIM, para definir um guia de referência com a intenção de comprovar a eficácia da abordagem proposta. Os experimentos foram realizados em uma conhecida base de dados de faces 3D disponível no *Biometric Experimentation Environment*.

Além da face como um todo, também foram utilizadas outras regiões da face (como a do nariz) para atingir melhores resultados e garantir a insensibilidade do algoritmo na presença de expressões faciais suaves.

Com a abordagem proposta, foram obtidas altas taxas de verificação, que estão além daquelas sugeridas no *Face Recognition Grand Challenge 2006* e tida como um dos principais objetivos da competição. Os resultados comprovam a eficácia da abordagem para o problema do reconhecimento facial 3D.

ABSTRACT

This work presents a novel approach for the 3D face recognition problem. Many works found in the literature employ the Iterative Closest Point (ICP) as the registration technique for the faces and, therefore, base their results in the similarity provided by the improper measures Mean Squared Error and Root Mean Squared Error. This work proposes the utilization of the Simulated Annealing stochastic algorithm based on two robust measures, M-estimator Sample Consensus and Surface Interpenetration Measure (SIM), to achieve a precise alignment of the faces. The conclusion concerning the similarity of the faces is done using the SIM, which provides a larger discrimination between face images from the same subject and face images from different ones.

In this work, different registration techniques were used, like ICP and genetic algorithm using the SIM, to define a baseline in order to verify the efficiency of the proposed approach. The experiments were performed on a well-known 3D face database available on the Biometric Experimentation Environment.

Besides the whole face, other regions of the face (like the nose) were also used to achieve a better result and to ensure the insensitivity of the algorithm in the presence of minor facial expressions.

With the proposed approach, high verification rate scores were achieved that are beyond the ones suggested by the Face Recognition Grand Challenge 2006 and are considered one of the main goals of the competition. The results prove the efficiency of the proposed approach for the 3D face recognition problem.

CAPÍTULO 1

INTRODUÇÃO

Atualmente, a face se destaca como uma das características do ser humano mais apropriadas para sua identificação, permitindo que pessoas sejam identificadas sem que sequer haja qualquer tipo de contato físico [1]. Neste campo de pesquisa, a informação tridimensional (3D) pode propiciar um aperfeiçoamento em relação à simples utilização de uma imagem bidimensional (2D). A área de pesquisa do reconhecimento facial 3D é relativamente nova e recentemente tem recebido cada vez mais atenção por parte dos pesquisadores.

Este trabalho apresenta uma nova abordagem para o problema de reconhecimento facial 3D. A abordagem baseia-se na utilização do *Simulated Annealing* (SA) com duas medidas de erro robustas o *M-estimator Sample Consensus* (MSAC) e a *Surface Interpenetration Measure* (SIM).

Este trabalho está organizado como se segue: Neste capítulo são introduzidos alguns conceitos referentes ao reconhecimento facial 3D; no Capítulo 2, os conceitos mais relevantes são aprofundados; os trabalhos relacionados são apresentados no Capítulo 3; no Capítulo 4 os algoritmos desenvolvidos e os resultados obtidos são evidenciados; seguido da conclusões no Capítulo 5.

1.1 MOTIVAÇÃO

A área de reconhecimento facial 3D é relativamente nova e, portanto, ainda possui grandes lacunas a serem exploradas. E, com o barateamento dos sistemas de aquisição de imagens 3D, são grandes as chances de que o reconhecimento facial 3D seja amplamente utilizado num futuro próximo.

O reconhecimento facial 3D representa um aprimoramento em relação ao 2D, uma vez que a simples utilização de imagens 3D elimina os maiores problemas do reconhecimento facial 2D, que são a variação de iluminação¹ e a variação de pose [2, 41], o que torna o sistema de reconhecimento mais robusto. No entanto, as imagens 3D também introduzem problemas inerentes a elas, como é o caso da freqüente ocorrência de ruídos, que devem ser corretamente

¹Dependendo do método de aquisição, a variação de iluminação pode influenciar negativamente o reconhecimento facial 3D.

tratados.

1.2 REGISTRO

O registro é uma tarefa fundamental em processamento de imagens cujo objetivo, de modo simples, consiste em, dadas duas, ou mais imagens de uma mesma cena, que apresentem diferenças devido ao processo de aquisição (e.g., diferença temporal, de ângulo de vista, diferença entre sensores, etc.), obter uma relação de mapeamento (conjunto de transformações geométricas) entre as imagens de tal forma que as diferenças entre elas sejam eliminadas, minimizadas, ou ainda ressaltadas, dependendo de cada aplicação [13]. Ou seja, fazer o registro de duas imagens equivale a alinhá-las.

A necessidade de se efetuar o registro de imagens aparece em muitos problemas práticos nas mais diversas áreas. O registro é comumente necessário para: (1) integração das informações obtidas de diferentes sensores; (2) encontrar diferenças entre imagens obtidas em tempos diferentes ou sob diferentes condições; (3) inferência de informação 3D; (4) e, reconhecimento de objeto baseado em modelo [13].

O registro de imagens tem sido aplicado em diferentes áreas de pesquisa, incluindo imagens médicas, visão robótica e arqueologia. A maior parte das aplicações estão focadas no desenvolvimento de técnicas para construir modelos tridimensionais precisos de objetos, preservando o máximo de informação possível [35, 36].

Atualmente, o *Iterative Closest Point* (ICP), apresentado em [7] por McKay e Besl, incluindo suas versões aprimoradas [21, 32], é o método de registro mais utilizado. No entanto, para que produzam resultados satisfatórios requerem um bom pré-alinhamento das imagens que serão registradas e, portanto, carecem de supervisão. Motivo pelo qual não deve ser utilizado como método de registro em um sistema totalmente automático.

O registro de imagens tridimensionais é um problema no qual é necessário lidar com seis dimensões, sendo, destas, três de translação e três de rotação, como visto na Figura 1.1 (a) e (b), respectivamente. A resolução do problema de registro de imagens se apresenta de forma complexa e, normalmente, demanda muito esforço computacional.

E é através do registro de imagens tridimensionais que grande parte dos algoritmos de reconhecimento facial 3D conseguem calcular a similaridade entre as faces.

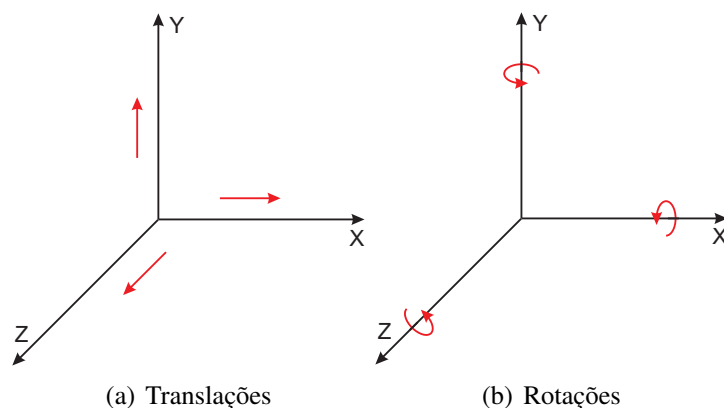


Figura 1.1: Dimensões utilizadas para solucionar o problema de registro de imagens tridimensionais. (a) três translações; (b) três rotações.

1.3 BIOMETRIA

Biometria é a ciência que estuda a utilização de características físicas e comportamentais para identificar pessoas. Estas características são chamadas de características biométricas ou, simplesmente, biometrias.

A assinatura e a caligrafia são as características biométricas utilizadas há mais tempo e, ainda hoje, são usadas na verificação de autenticidade de documentos. Posteriormente, a imagem da face e a impressão digital passaram a ser empregadas. Mais recentemente, voz, modo de andar, imagens da retina e da íris, impressão palmar e informação facial 3D passaram a ser utilizadas como características biométricas. Cada uma possuindo vantagens e desvantagens que restringem sua aplicação [2].

Com a crescente demanda por segurança, a utilização de características biométricas é vista como um meio de desenvolver novas soluções de segurança ou complementar as já existentes. Um dos grandes benefícios de sistemas que utilizam características biométricas, chamados de sistemas biométricos, é que a “chave de acesso” está sempre presente com o indivíduo e, muitas vezes, pode ser obtida à distância, o que facilita a utilização do sistema, já que não requer procedimentos especiais para sua obtenção.

Uma outra utilização para sistemas biométricos é a localização de pessoas desaparecidas e fugitivos. Neste último caso, faz-se necessária a utilização de uma característica biométrica que pode ser obtida sem a ciência e a cooperação da pessoa analisada (e.g., a face).

O reconhecimento facial é um método biométrico de identificação considerado natural, não-intrusivo e amplamente aceito e, portanto, possui o potencial para se tornar o líder na tecnologia biométrica [12]. No entanto, avaliações, como as do *Face Recognition Vendor Test (FRVT)*, mostraram que os algoritmos que representam o estado da arte no reconhecimento facial 2D,

não eram próprias para aplicações biométricas de alta demanda [12]. Assim, o reconhecimento facial 3D é uma tendência relativamente nova que pode contribuir para que o reconhecimento facial produza resultados mais precisos e robustos [12].

Maiores informações sobre a biometria e as características biométricas podem ser encontradas na Seção 2.1.

1.4 RECONHECIMENTO FACIAL 3D

O reconhecimento facial 2D foi objeto de extensa pesquisa por muitos anos. Recentemente, o reconhecimento facial 3D tem ganhado cada vez mais atenção e, o avanço tecnológico na área² de obtenção dessas imagens é um dos principais fatores para esse aumento de interesse [5].

Recentemente, algumas abordagens para o reconhecimento facial 3D utilizaram o registro de imagens para efetuar o alinhamento das faces e realizar a identificação de indivíduos [14, 17, 26, 27]. Estas abordagens são geralmente baseadas na média quadrática de erro, *Mean Squared Error* (MSE), ou na raiz média quadrática de erro, *Root Mean Squared Error* (RMSE), algumas vezes combinadas com outras medidas, para avaliar a qualidade do registro das faces. Em todas estas abordagens, o ICP é aplicado para obter o registro das imagens.

O ICP é guiado pelo MSE mas, foi provado [36] que esta medida pode levar a uma convergência local imprecisa para o registro de imagens de profundidade, mesmo quando versões aprimoradas do ICP são utilizadas [21, 32]. Também em [36] os autores sugerem que uma medida mais apropriada, a medida de interpenetração de superfície, *Surface Interpenetration Measure* (SIM) [37], deve ser utilizada para avaliar a qualidade do alinhamento.

Uma alternativa ao ICP é a utilização de algoritmos genéticos, *Genetic Algorithm* (GA) como método de registro [5, 36, 37] já que estes não requerem um bom pré-alinhamento para conseguir um alinhamento satisfatório. No entanto, o GA é um algoritmo que possui um alto custo computacional, tornando imprópria a sua utilização para o reconhecimento facial 3D.

Para evitar todos os problemas associados aos métodos de registro (ICP e GA) e às medidas de erro utilizadas (MSE e RMSE), é proposta a utilização do SA com duas medidas robustas para conduzir e avaliar o registro: o MSAC e a SIM.

²3DIM - 3-D Digital Imaging and Modeling - <http://www.3dimconference.org/>

CAPÍTULO 2

CONCEITOS

Neste capítulo são explanados conceitos que possibilitam a correta interpretação dos resultados obtidos bem como a compreensão dos componentes dos algoritmos desenvolvidos (Capítulo 4).

O ICP é o método de registro mais utilizado nos algoritmos de reconhecimento facial 3D baseados em alinhamento e, como pode ser visto no Capítulo 3, ele é utilizado pelos que representam o estado da arte nesta área.

Embora o ICP seja amplamente utilizado como método de registro, ele apresenta algumas desvantagens: (1) tende a convergir para mínimos locais e requer um bom alinhamento inicial [38]; (2) as medidas de erro utilizadas pelo ICP (MSE e RMSE) não são robustas o suficiente para produzir alinhamentos precisos.

A proposta para minimizar e, por vezes, eliminar os problemas intrínsecos ao ICP e às medidas de erro por ele empregadas, é a utilização de um método de busca estocástico, o *Simulated Annealing*, com medidas de erro mais robustas, o MSAC e a SIM.

Inicialmente são apresentados conceitos referentes à biometria, incluindo formas de avaliação dos sistemas biométricos, Seção 2.1. Posteriormente, Seção 2.2, foca-se na utilização da face como característica biométrica, representada por imagens 3D, abordando sua aquisição, tratamento e representação. Nas seções seguintes, são contextualizados os componentes dos algoritmos desenvolvidos. Na Seção 2.3 é feita uma descrição do SA, seu funcionamento e suas particularidades. Na Seção 2.4 o MSAC e a SIM são apresentadas.

Após o entendimento destes conceitos, os algoritmos desenvolvidos e seus resultados (Capítulo 4) são facilmente compreendidos e interpretados.

2.1 BIOMETRIA

A palavra biometria tem origem no grego antigo e é composta da junção de duas palavras: *bíos*, que significa vida e; *métron* que significa medição.

A biometria é o ramo da ciência que estuda métodos que visam identificar pessoas através

de traços físicos ou comportamentais intrínsecos a elas. Dá-se também o nome de biometria às características biométricas utilizadas para a identificação.

Os sistemas biométricos tendem a substituir ou complementar os sistemas de verificação convencionais. Atualmente, os sistemas de verificação convencionais podem ser enquadrados em dois modelos:

- **Baseado em Posse:** o usuário do sistema possui um *token* (cartão magnético, cartão *smart*, etc.) que o identifica. Toda a segurança do sistema é baseada neste *token* que, se perdido, permite ao “novo dono” assumir a identidade do antigo;
- **Baseado em Conhecimento:** o usuário faz utilização de uma senha para identificar-se no sistema. Se a senha for muito pequena, pode ser descoberta facilmente e, se for muito grande, pode ser difícil lembrar-se dela e, o usuário comum poderá anotá-la em algum lugar para não esquecê-la e, assim, correr o risco de ter a mesma roubada ou perdida.

Como pode ser visto, é evidente a fragilidade dos sistemas de verificação convencionais. Esta fragilidade pode ser evitada se utilizarmos nosso próprio corpo como chave do sistema. Já que, alguns traços físicos ou comportamentais são muito mais complicados de serem forjados que uma cadeia de caracteres. Abaixo podem ser vistas as principais vantagens dos sistemas biométricos de verificação sobre os sistemas convencionais:

- características biométricas não são passíveis de perda ou esquecimento;
- características biométricas são difíceis de serem copiadas, compartilhadas e distribuídas;
- os sistemas biométricos requerem que a pessoa que está sendo autenticada esteja presente no momento e local da autenticação.

Além disso, os sistemas biométricos podem ser utilizados juntamente com senhas e *tokens* e, deste modo, aumentar a segurança de sistemas preexistentes sem a necessidade de substituí-los.

Os sistemas biométricos podem prover, basicamente, duas funções:

- **Verificação** ou **Autenticação**

- *Descrição:* neste caso, a pessoa que está utilizando o sistema declara uma identidade e é feita a comparação da característica biométrica desta pessoa com a característica biométrica da identidade declarada, que está armazenada em uma base de dados.
- *Responde a:* esta pessoa é quem afirma ser?
- *Utilização:* sistemas de autenticação, restrição de acesso, etc.

– *Comparação*: um para um (1 : 1).

- **Identificação** ou **Reconhecimento**

– *Descrição*: neste caso, é fornecido ao sistema uma característica biométrica pertencente a uma pessoa desconhecida e, comparando-se esta característica com todas as outras armazenadas na base de dados, tenta-se obter a identidade desta pessoa desconhecida.

– *Responde a*: quem é esta pessoa?

– *Utilização*: sistemas de monitoramento, em geral.

– *Comparação*: um para muitos (1 : N).

Existe também um subcaso da identificação que é chamado de classificação (em inglês *screening*), onde tenta-se verificar se uma pessoa pertence a um determinado grupo. Neste caso, o número de pessoas na base de dados é muito menor e a comparação é do tipo um para muitos (1 : n , onde $n \ll N$).

2.1.1 Características Biométricas

Característica biométrica é um traço físico ou comportamental que é mensurável e tem potencial para identificar uma pessoa. As características biométricas podem ser divididas em duas classes:

- **Fisiológica**: engloba as características relacionadas à composição do corpo humano, em especial seu formato;
- **Comportamental**: compreende as características relacionadas ao comportamento das pessoas.

Na Figura 2.1 é possível ver a classificação de algumas características biométricas, com exemplos, segundo a classe a qual pertencem.

A Tabela 2.1 compara algumas características biométricas de acordo com suas possíveis qualidades [2]:

- **Precisão**: indica se a identificação baseada nesta característica biométrica produz resultados precisos;
- **Custo**: indica se o custo do equipamento exigido para utilizar esta característica biométrica é baixo;

Características Biométricas

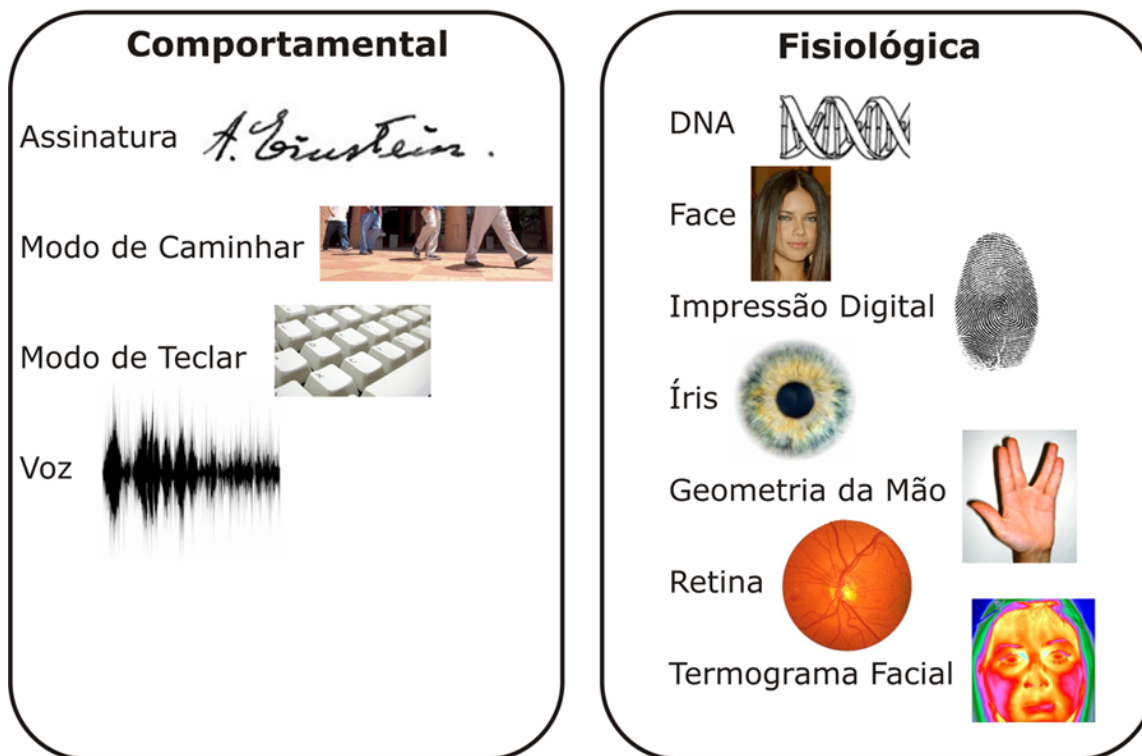


Figura 2.1: Exemplos de algumas características biométricas separadas em classes.

- Privacidade: indica que não existe a possibilidade de uso impróprio desta característica biométrica caso a base de dados onde ela está armazenada seja comprometida. Um exemplo de característica biométrica que permite uso indevido é a assinatura;
- Integrabilidade: indica se um sistema baseado nesta característica biométrica permite integração com outros sistemas;
- Facilidade de Uso: indica se é fácil lidar com esta característica biométrica, principalmente com relação à aquisição da mesma;
- Desenvolvimento: indica se é simples desenvolver um sistema de identificação que utiliza esta característica biométrica;
- Ocultação de Identidade: indica se esta biometria não é suscetível a evasão em um ambiente de reconhecimento;
- Falsificação de Identidade: indica se esta característica biométrica não é passível de ser “falsificada” em um ambiente de verificação.

	Face 2D	Assinatura	Voz	Íris	Impressão Digital	Retina	Modo de Caminhar	Impressão Palmar	Face 3D
Precisão				•	•	•			•
Custo	•		•		•				
Privacidade	•			•		•	•		•
Integrabilidade	•		•		•				•
Facilidade de Uso	•	•	•		•			•	•
Desenvolvimento	•	•	•		•			•	•
Ocultação de Identidade				•		•		•	
Falsificação de Identidade				•	•	•		•	•

Tabela 2.1: Exemplos de características biométricas, suas vantagens e desvantagens. Fonte: [2]

2.1.2 Avaliação de Sistemas Biométricos de Verificação

Um sistema de verificação baseado em biometria deve lidar com dois tipos de eventos: (1) uma pessoa é quem afirma ser (neste caso, chamada de cliente); (2) uma pessoa não é quem afirma ser (neste caso, chamada de impostor) [6].

Geralmente, os sistemas biométricos utilizam um escore para indicar a similaridade entre as amostras da biometria utilizada. Quanto maior o escore (ou menor, dependendo da medida de similaridade)¹, maior a semelhança entre as duas amostras. Em teoria, os escores dos clientes deveriam ser sempre maiores que os escores dos impostores. Se isso fosse verdade, a simples utilização de um limiar seria suficiente para separar os dois grupos de escore e, assim, diferenciar clientes e impostores. No entanto, esta suposição não é válida para sistemas biométricos do mundo real, já que alguns escores de impostores são maiores que alguns escores de clientes.

Sendo assim, a partir das duas decisões que o sistema pode tomar — aceitar uma pessoa como cliente ou rejeitá-la, classificando-a como um impostor — ele pode cometer dois tipos de erros: uma falsa aceitação, *False Acceptance* (FA), quando o sistema aceita um impostor; e uma falta rejeição, *False Rejection* (FR), quando o sistema rejeita um cliente [6].

A avaliação de sistemas biométricos de verificação, independente da biometria, normalmente, é feita em função de duas taxas de erro: *False Acceptance Rate* (FAR) e *False Rejection Rate* (FRR). Outras medidas, como a *Equal Error Rate* (EER) e a taxa de verificação, são derivadas destas duas.

¹Nos exemplos será considerado que, quanto maior o escore, maior a similaridade entre as características biométricas analisadas.

2.1.2.1 Taxa de Falsa Aceitação

A taxa de falsa aceitação, *False Acceptance Rate* (FAR), corresponde à probabilidade de um sistema biométrico, incorretamente, classificar um impostor como cliente. Ou seja, o escore do impostor está acima do limiar que separa clientes e impostores.

O cálculo feito para obter a FAR é visto na Equação 2.1:

$$FAR = \frac{N_{fa}}{N_i} \quad (2.1)$$

onde, N_{fa} corresponde ao número de falsas aceitações e N_i ao número de impostores que tentaram acessar o sistema. O valor da FAR varia no intervalo fechado $[0,1]$ (este valor também costuma ser representado em porcentagem).

Se o valor da FAR for 1, todos os impostores serão falsamente aceitos, em contrapartida, se seu valor for 0, todos os impostores serão corretamente classificados e, portanto, rejeitados. Como é possível perceber, quanto menor o valor da FAR mais seguro será o sistema. A Figura 2.2 (b) mostra como a FAR varia para uma distribuição fictícia de escore de impostores que pode ser observada na Figura 2.2 (a).

Como normalmente os sistemas biométricos de verificação são utilizados para restringir acesso/ações por parte de impostores, a FAR pode ser considerada o fator mais crítico deste tipo de sistema biométrico.

2.1.2.2 Taxa de Falsa Rejeição

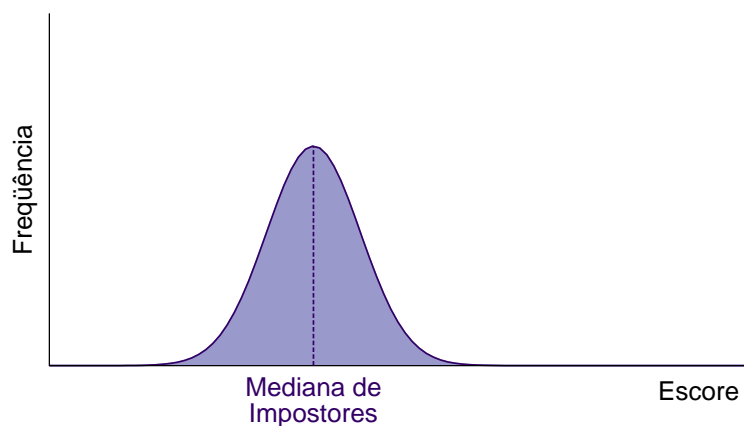
A taxa de falsa rejeição, *False Rejection Rate* (FRR), indica a probabilidade de um sistema biométrico rejeitar um cliente, classificando-o como impostor. Ou seja, o escore do cliente está abaixo do limiar que separa clientes e impostores.

O cálculo feito para obter a FRR é visto na Equação 2.2:

$$FRR = \frac{N_{fr}}{N_c} \quad (2.2)$$

onde, N_{fr} corresponde ao número de falsas rejeições e N_c ao número de clientes que tentaram acessar o sistema. O valor da FRR, assim como o da FAR, varia no intervalo fechado $[0,1]$ (este valor também costuma ser representado em porcentagem).

Se o valor da FRR for 1, todos os clientes serão falsamente rejeitados, no entanto, se seu valor for 0, todos os clientes serão corretamente classificados e, portanto, aceitos. Como é



(a) distribuição de impostores

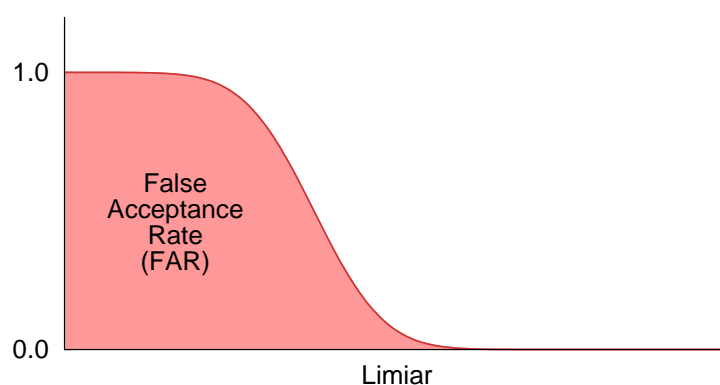
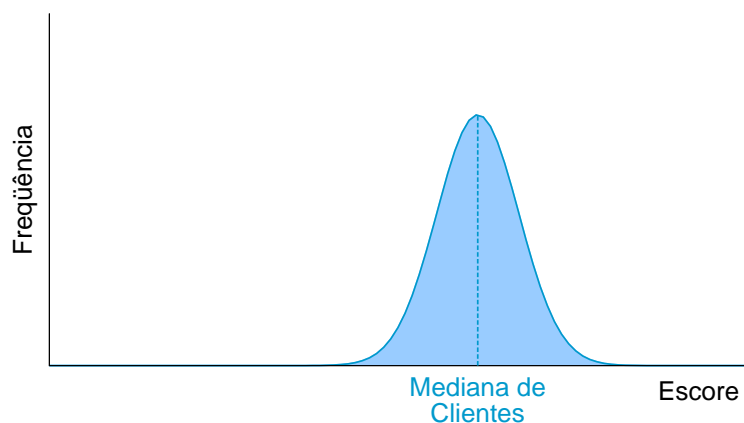
(b) *False Acceptance Rate*

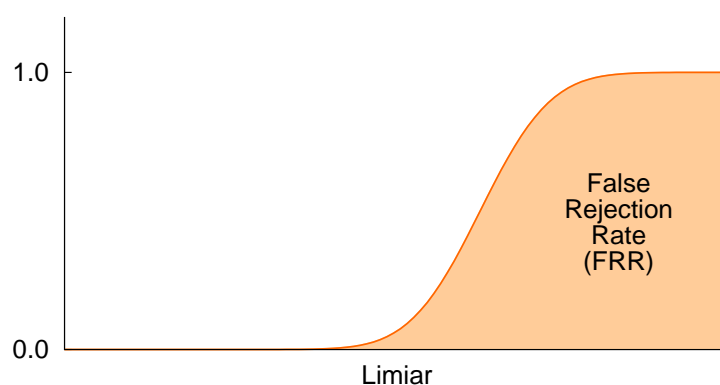
Figura 2.2: Distribuição do escore de impostores e sua FAR. (a) distribuição do escore de impostores, (b) a FAR para a distribuição de (a).

possível perceber, quanto menor o valor da FRR maior será a chance de acontecer uma falsa aceitação. A Figura 2.3 (b) mostra como a FRR varia para uma distribuição fictícia de escore de clientes que pode ser observada na Figura 2.3 (a).

Em sistemas biométricos onde a performance tem prioridade sobre a segurança, a FRR deve ser reduzida (diminuindo o limiar) para minimizar a ocorrência de falsas rejeições.



(a) distribuição de clientes



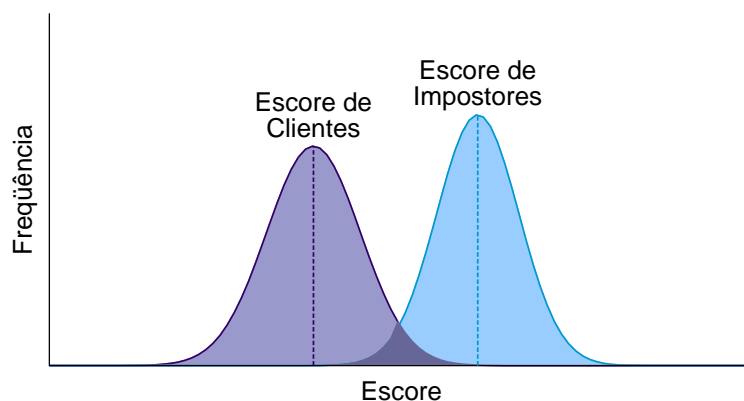
(b) False Rejection Rate

Figura 2.3: Distribuição do escore de clientes e sua FRR. (a) distribuição do escore de clientes, (b) a FRR para a distribuição de (a).

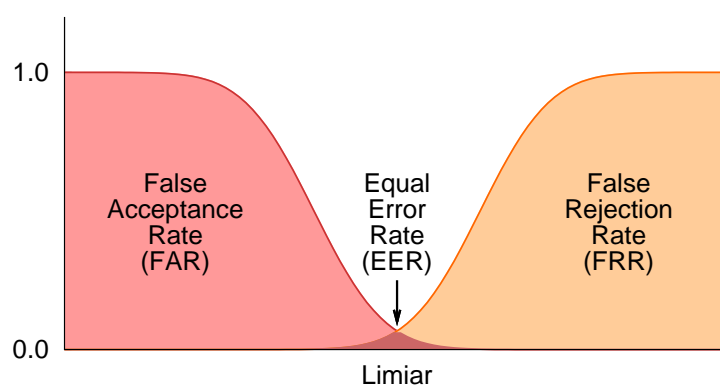
2.1.2.3 Taxa Equivalente de Erro

A taxa equivalente de erro, *Equal Error Rate* (EER), corresponde à taxa de erro na qual a FAR e a FRR possuem o mesmo valor.

Quando há sobreposição na distribuição dos escores de clientes e de impostores, Figura 2.4 (a), a FAR e a FRR se cruzam em um determinado ponto, Figura 2.4 (b). O valor da FAR e da FRR neste ponto de cruzamento que, obviamente, é o mesmo para ambos, é chamado de EER.



(a) distribuição de clientes e impostores



(b) Equal Error Rate

Figura 2.4: Distribuição dos escores de clientes e impostores e sua EER. (a) distribuição dos escores de clientes e impostores, (b) a EER para a distribuição de (a).

O valor da EER varia no intervalo fechado $[0,1]$, muitas vezes sendo representado em forma de porcentagem.

Os sistemas de verificação podem ser ajustados utilizando um limiar de decisão para conseguir uma baixa FAR ou uma baixa FRR [6]. Existe, portanto, uma troca entre as duas taxas, que pode ser explorada de acordo com a aplicação do sistema: em alguns casos pode ser mais importante possuir uma FAR baixa, enquanto que em outros, uma FRR baixa pode ser indispensável [6]. Para visualizar o desempenho de um sistema biométrico de acordo com essa troca entre a FAR e a FRR, geralmente são utilizadas curvas *Receiver Operating Characteristic* (ROC), que representam a FRR em função da FAR. É possível perceber que quanto mais

próxima da coordenada $(0,0)$ estiver esta curva, melhor será o desempenho do sistema, já que tanto a FAR quanto a FRR serão baixas. Uma variação da curva ROC é a curva *Detection Error Trade-off* (DET), que é obtida utilizando escala logarítmica nos dois eixos. A Figura 2.5 é um exemplo de curva ROC e, observando-a, é possível notar que quanto maior a FAR menor é a FRR e vice-versa. Ainda na curva da Figura 2.5, foram colocados três pontos que indicam três valores de FAR e três valores de FRR. Estes pontos representam o mesmo sistema mas, com comportamentos diferentes:

- Ponto *A*: representa um ponto de equilíbrio ($FAR_A = FRR_A$), chamado de EER;
- Ponto *B*: neste ponto a FAR é menor que a FRR ($FAR_B < FRR_B$). Em relação ao ponto *A*, temos um comportamento:
 - Mais seguro: $FAR_B < FAR_A$
 - Mais inconveniente: $FRR_B > FRR_A$
- Ponto *C*: neste ponto a FAR é maior que a FRR ($FAR_C > FRR_C$). Em relação ao ponto *A*, temos um comportamento:
 - Menos seguro: $FAR_C > FAR_A$
 - Menos inconveniente: $FRR_C < FRR_A$

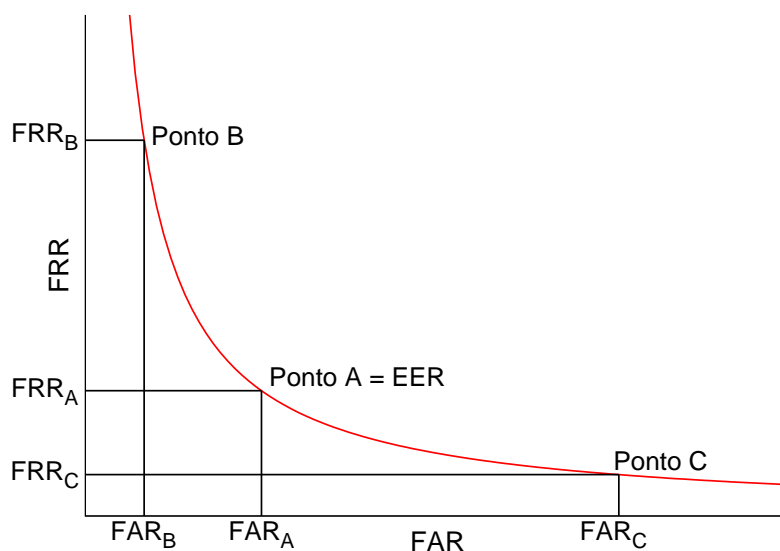


Figura 2.5: Curva ROC, relacionando a FAR e a FRR.

Como os sistemas podem ter comportamentos diferentes de acordo com sua aplicação, normalmente utiliza-se a EER para uma comparação não rigorosa entre diferentes sistemas biométricos. Quanto menor for a EER, mais preciso é considerado o sistema avaliado.

2.1.2.4 Taxa de Verificação

A taxa de verificação indica a probabilidade de um sistema biométrico aceitar um cliente. Ou seja, o seu score está acima do limiar e , portanto, o cliente é corretamente classificado.

O cálculo feito para obter a taxa de verificação é visto na Equação 2.3:

$$TV = 1.0 - FRR \quad (2.3)$$

O valor da taxa de verificação varia no intervalo fechado $[0,1]$, muitas vezes sendo representada em forma de porcentagem.

Como o valor da FRR varia de acordo com o limiar escolhido, ver Figura 2.3 (b), o mesmo acontecerá com a taxa de verificação.

Para aumentar a taxa de verificação de um sistema biométrico, é necessário diminuir a FRR. No entanto, como a FAR varia inversamente à FRR, ver Figura 2.4 (b), isto pode resultar em um sistema com altas taxas de falsa aceitação e portanto, pouco seguro. De acordo com [30], uma FAR aceitável é de cerca de 0,1%, ou seja, um impostor aceito a cada mil tentativas de acesso por parte de impostores.

O ideal é que um sistema biométrico possua uma taxa de verificação de 100% e, ainda assim, preserve uma FAR de 0%, no entanto, este patamar é difícil de ser atingido por sistemas biométricos do mundo real.

2.1.3 Avaliação de Sistemas Biométricos de Identificação

Como dito anteriormente, a identificação é um problema de comparação “um para muitos”. Ou seja, uma característica biométrica submetida ao sistema será comparada com todas as características biométricas armazenadas na base de dados do sistema. Por exemplo, supondo que um sistema possua 500 faces cadastradas em sua base de dados e, submete-se a este sistema, para identificação, uma outra face; a face submetida será comparada com cada uma das 500 faces do sistema, resultando em 500 comparações, para determinar a identidade da face submetida.

A avaliação de sistemas biométricos de identificação, basicamente, é feita através da taxa de identificação, ou de reconhecimento, *rank-one*. Esta taxa compreende a porcentagem de submissões feitas ao sistema biométrico que são corretamente identificadas [41].

Os passos do algoritmo de identificação para a obtenção da taxa de reconhecimento *rank-one* são apresentados a seguir:

1. Uma característica biométrica submetida ao sistema é comparada com todas as características biométricas pertencentes à base de dados deste sistema. O grupo ao qual as características biométricas submetidas pertencem, é chamado de *probe*;
2. Seleciona-se a combinação que apresenta maior similaridade (de acordo com a medida de similaridade utilizada). A base de dados, contra a qual as características biométricas submetidas são comparadas, é chamada de *gallery*;
3. Se as características biométricas desta combinação pertencem à mesma pessoa, então, elas são ditas correspondentes. Ou seja, o sistema, corretamente, identificou a característica biométrica submetida;
4. A taxa de reconhecimento *rank-one* é a porcentagem das características biométricas submetidas para as quais definiu-se uma correspondência correta.

A Equação 2.4 define como o cálculo da taxa de reconhecimento *rank-one* é feito:

$$RankOne = \frac{N_c}{N_s} \quad (2.4)$$

onde, N_c corresponde ao número de submissões corretamente identificadas e N_s ao número total de submissões. O valor da taxa de reconhecimento *rank-one*, varia no intervalo fechado $[0,1]$ mas, também pode ser expresso em porcentagem.

O nome *rank-one* vem do fato de que verifica-se somente se a identidade correta da imagem submetida encontra-se na imagem que apresentou maior valor de similaridade. Existe também outros *ranks*, como por exemplo, o *rank-two*, onde verifica-se se a identidade correta é uma das duas imagens que apresentaram o maior valor de similaridade, não importando se a identidade correta é a imagem de maior valor ou a segunda de maior valor. O número possível de *rank* é igual ao número de indivíduos da base analisada, sendo que, neste caso extremo, todos os algoritmos conseguiriam taxa de 100%. Em outras palavras, o que esta taxa demonstra, por exemplo para o *rank-X*, é o quão frequentemente ocorre a presença da identidade correta, da imagem submetida, nas X imagens com maior valor de similaridade. O gráfico que traça uma curva que relaciona a taxa de reconhecimento de acordo com a variação do *rank* é chamado de *Cumulative Match Characteristic* (CMC). A Figura 2.6 mostra um exemplo de curva CMC para valores fictícios.

Como pode ser percebido, quanto menor o *rank*, menor é a taxa de reconhecimento. Embora difícil de ser atingido, o ideal é que um sistema de identificação possua um índice de 100% para a taxa de reconhecimento *rank-one*.

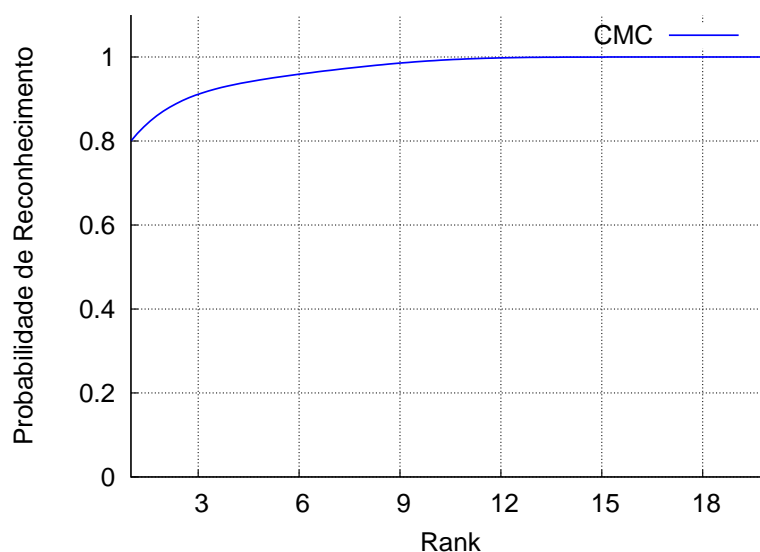


Figura 2.6: Curva CMC.

2.2 RECONHECIMENTO FACIAL 3D

Duas particularidades fazem da face uma característica biométrica atrativa [2]:

1. A aquisição da face é feita de forma fácil e não-intrusiva: esta propriedade é particularmente interessante caso o sistema seja utilizado frequentemente ou por um número grande de pessoas. Características biométricas que não possuem esta propriedade: íris e retina;
2. Baixa privacidade da informação: a face é exposta constantemente e, caso uma base de faces seja roubada, estas informações não possibilitam seu uso impróprio. Características biométricas que não possuem esta propriedade: assinatura e impressão digital.

Em alguns aspectos, o reconhecimento facial 3D representa uma melhoria do reconhecimento facial 2D. O reconhecimento baseado em imagens de intensidade é um problema de solução complexa porque as variações de iluminação, pose e expressão provocam grandes mudanças na face que são difíceis de serem tratadas. O reconhecimento facial 3D tem o potencial de melhorar a localização de características da face e superar os problemas de pose e iluminação, além de poder ser utilizado juntamente com os sistemas baseados em imagens 2D da face [2].

2.2.1 Aquisição

A precisa aquisição da superfície da face é crucial para o reconhecimento facial 3D [11]. Os dispositivos de aquisição de imagens 3D podem ser divididos em dois tipos: com contato e sem contato. Esta classificação se dá pela necessidade, ou não, do dispositivo de entrar em contato físico com o objeto que está sendo analisado.

Nesta seção serão considerados apenas os dispositivos que não necessitam de contato físico com a face a fim de obter informações tridimensionais sobre ela.

De acordo com o método de aquisição, os dispositivos que não exigem contato podem ser divididos em dois grupos: ativo e passivo. O método ativo emite algum tipo de radiação e detecta sua reflexão para fazer a sondagem do objeto no ambiente. Já o método passivo não emite qualquer tipo de radiação, ao invés disso, baseia-se na detecção da radiação do ambiente refletida, normalmente esta radiação é a luz.

Serão abordados um tipo de dispositivo passivo — baseado em visão estéreo — e dois ativos — baseados em luz estruturada e laser.

Na aquisição baseada em visão estéreo, duas, ou mais câmeras que estão posicionadas e calibradas são utilizadas para adquirir imagens simultâneas do sujeito que será analisado. Através da solução do problema da correspondência e da utilização de modelos geométricos, é possível inferir informação de profundidade para cada ponto da cena observada (sujeito que será identificado). A Figura 2.7 demonstra o esquema de aquisição baseada em visão estéreo.

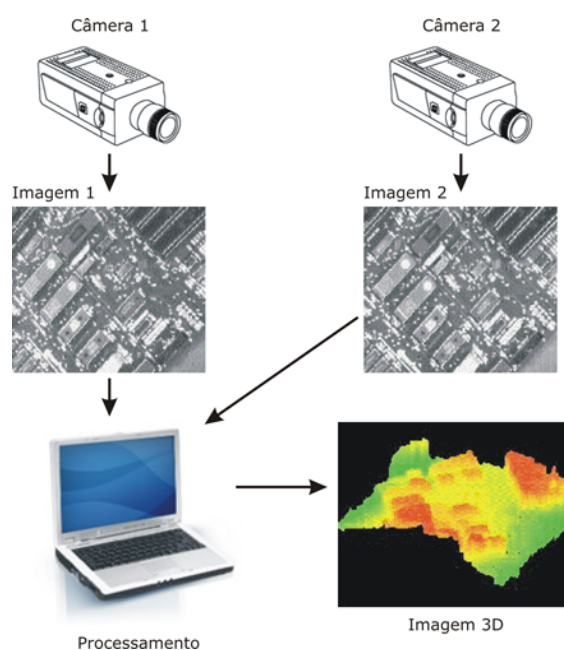


Figura 2.7: Esquema de funcionamento de aquisição baseada em visão estéreo.

Este método é o que possui o menor custo e a maior facilidade de uso, no entanto, não possui

grande precisão e variações de iluminação entre as imagens adquiridas podem comprometer a qualidade da imagem 3D.

A técnica de luz estruturada utiliza um padrão de luz (grade, listras, padrões elípticos, etc.) que é projetado na face do indivíduo que está sendo analisado. Através da distorção dos padrões, o formato da superfície examinada pode ser deduzido. Utilizando-se de informações sobre a geometria da câmera e do projetor é possível calcular informação de profundidade através de uma técnica semelhante à triangulação. A Figura 2.8 demonstra o esquema de aquisição baseada em luz estruturada.

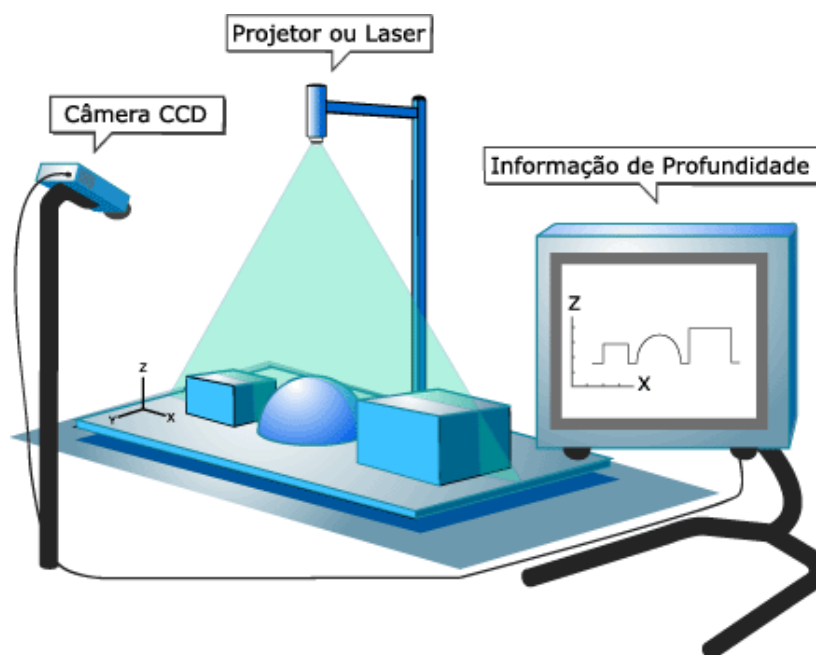


Figura 2.8: Esquema de funcionamento de aquisição baseada em luz estruturada.

Esta técnica é relativamente rápida, barata e permite que uma simples câmera produza informações de profundidade e de textura.

A aquisição de imagens tridimensionais da face também pode ser feita utilizando um sensor laser baseado em triangulação. Princípio desta técnica, ver Figura 2.9: o feixe de laser é projetado no objeto; a lente projeta o ponto do laser no sensor CCD (ou PSD); a triangulação é feita a partir do ponto projetado, da câmera e do emissor de laser. Como pode ser visto, ainda na Figura 2.9, um deslocamento do objeto (DZ) produz um deslocamento na imagem (Dz).

Este método é o que possui maior precisão, no entanto, é também o que possui o equipamento mais caro.

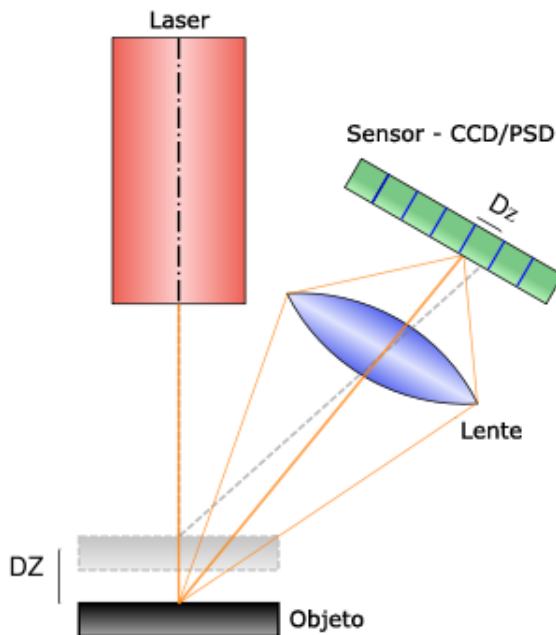


Figura 2.9: Esquema de funcionamento de aquisição baseada em sensor laser (triangulação).

2.2.2 Pré-processamento

Até mesmo sob condições ideais de iluminação para um determinado sensor, na aquisição de uma imagem 3D, é comum a ocorrência de imperfeições, também chamadas de artefatos, em regiões da face, tais como regiões oleosas que se apresentam reflexivas, dos olhos e regiões com pêlos, tais como sobrancelhas, bigode e barba [9]. As imagens 3D com estas imperfeições normalmente degradam a performance de algoritmos de reconhecimento facial que as utilizam. Ou seja, para garantir bons resultados na etapa de reconhecimento, as imagens 3D precisam ser pré-processadas após a aquisição.

A não ser que seja utilizado um *scanner* laser, a iluminação ainda afeta a aquisição de imagens 3D [2, 9]. A Figura 2.10 mostra um exemplo de como variação de iluminação pode influenciar a aquisição de uma imagem 3D da face.

As imperfeições mais comumente encontradas em imagens 3D são [8, 9, 15]:

- ausência de pontos: regiões da imagem onde há ausência de pontos (vértices), criando “buracos”. Ocorre devido à incapacidade do sensor de adquirir os dados (informação 3D) nestas regiões;
- ruído do tipo *spike*: são pontos que se destoam dos pontos que estão ao seu redor, seja por possuírem um valor muito alto ou muito baixo. Os pontos *spike* são *outliers* nos dados adquiridos, resultantes, por exemplo, de uma inter-reflexão em sensores baseados em luz estruturada ou de uma falsa correspondência em sensores baseados em visão estéreo [9].

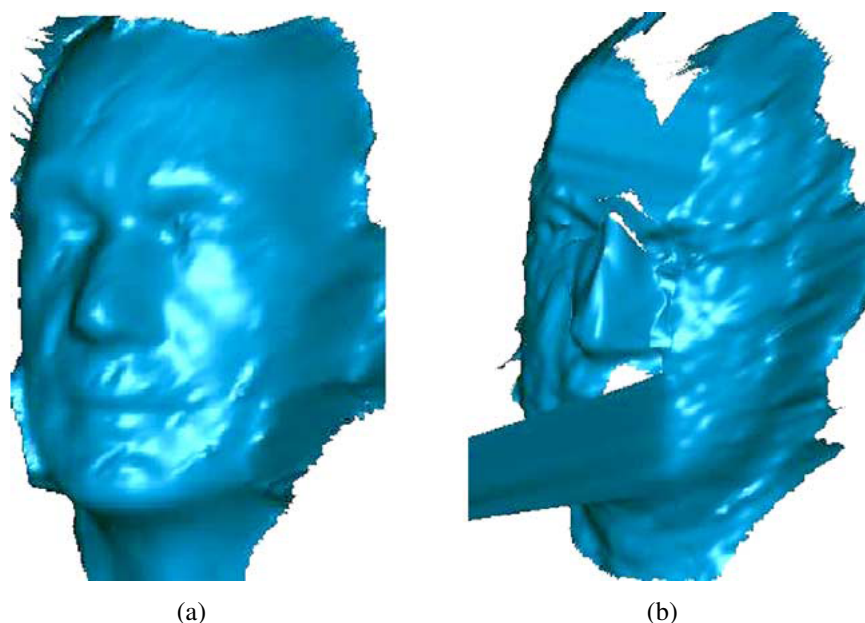


Figura 2.10: Efeito da variação de iluminação na aquisição de uma imagem 3D da face: (a) aquisição com iluminação apropriada para o sensor; (b) utilização de um refletor adicional a 1,5m de distância da face. Fonte: [9].

Para corrigir o problema da ausência de pontos, é feita uma interpolação dos pontos válidos que estão ao redor do buraco. Por exemplo, utilizando imagens de profundidade (imagens *range*), supondo que o ponto (x,y) seja um buraco, e que os seus vizinhos são pontos válidos, a Equação 2.5 mostra como redefinir este ponto:

$$P(x,y) = \frac{1}{8} \left[\sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} P(i,j) \right] \quad (2.5)$$

onde, $P(x,y)$ corresponde ao ponto 3D localizado na posição (x,y) da imagem P , do tipo *range*. Foi considerado que o ponto inválido (x,y) retorna o valor identidade para as operações de adição e subtração, ou seja, corresponde ao ponto 3D de coordenadas $(0,0,0)$.

A Figura 2.11 mostra um exemplo onde foi feita a interpolação das regiões onde havia algumas áreas com ausência de pontos.

O ruído do tipo *spike* (estaca, em inglês) pode ser visto na Figura 2.12. Esta imagem 3D foi renderizada em posição lateral para evidenciar este ruído, que se apresenta como “estacas” que se afastam da superfície da face.

Para remover o ruído deste tipo, basta aplicar um filtro da mediana no ponto que apresenta o ruído. O filtro da mediana consiste em: (1) selecionar os n pontos² contidos na janela de raio r , cujo centro encontra-se no ponto p ; (2) ordenar os pontos que pertencem a esta janela

² $n = (2r + 1)^2$, onde r é o raio da janela.

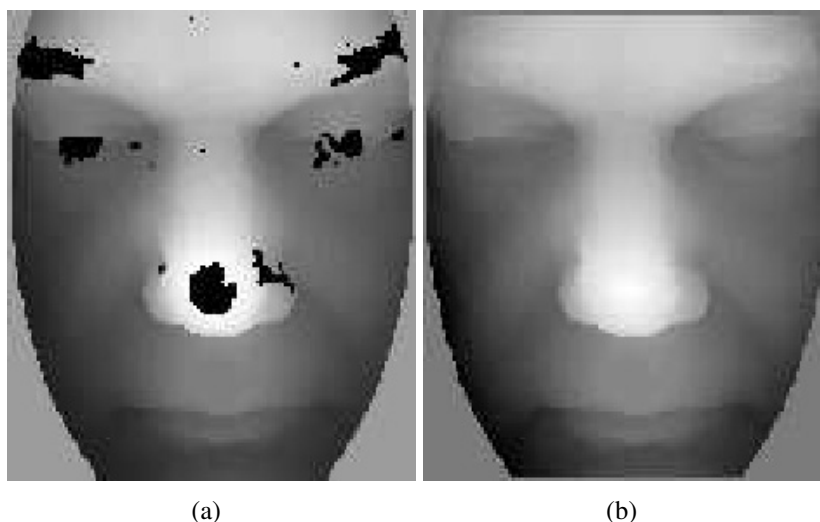


Figura 2.11: Exemplo de imagem 3D com “buracos” e que foram removidos através de pré-processamento: (a) imagem 3D com regiões sem pontos (áreas negras); (b) a imagem (a) após a aplicação de pré-processamento, no caso, a interpolação. Fonte: [15]

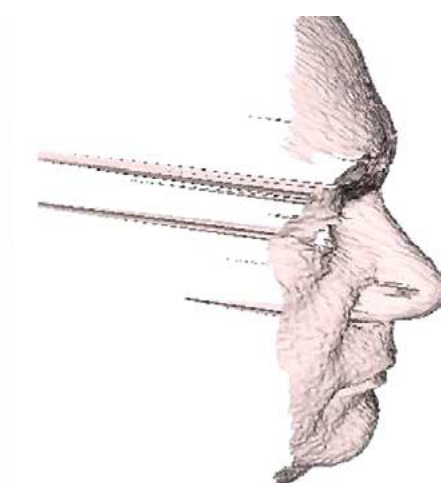


Figura 2.12: Exemplo de ruído do tipo *spike*. Fonte: [9].

(inclusive o próprio ponto p); (3) substituir o ponto p pelo valor da mediana, ou seja, o valor que encontra-se na posição $(n - 1)/2$ do vetor³ ordenado.

A Figura 2.13 mostra um exemplo onde foi aplicado o filtro da mediana para remover o referido ruído.

Como pôde ser observado, os dois tipos de ruído são tratados de maneira semelhante, que é através da geração de novos valores para os pontos ruidosos, a partir dos valores dos pontos próximos a estes.

Ao avaliar se uma imagem 3D requer ou não pré-processamento, é importante que análise seja feita sem a textura que foi adquirida juntamente com o modelo 3D. Quando o modelo

³O primeiro elemento deste vetor encontra-se na posição 0 (zero), e o último na posição $n - 1$.

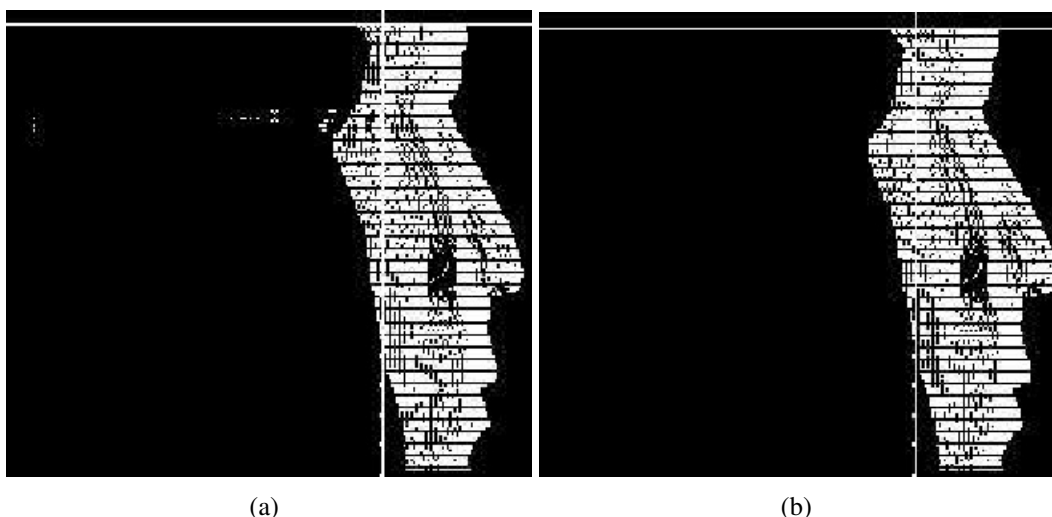


Figura 2.13: Exemplo de imagem 3D onde o ruído do tipo *spike* foi removido por um pré-processamento: (a) imagem com ruído; (b) a imagem de (a) após a aplicação do pré-processamento, no caso, o filtro da mediana. Fonte: [15]

3D é renderizado juntamente com o seu mapa de textura, este mapa pode ocultar imperfeições presentes no modelo 3D, como pode ser visto na Figura 2.14. Para fazer este tipo de análise, recomenda-se que o modelo 3D seja renderizado como um modelo sombreado [9].

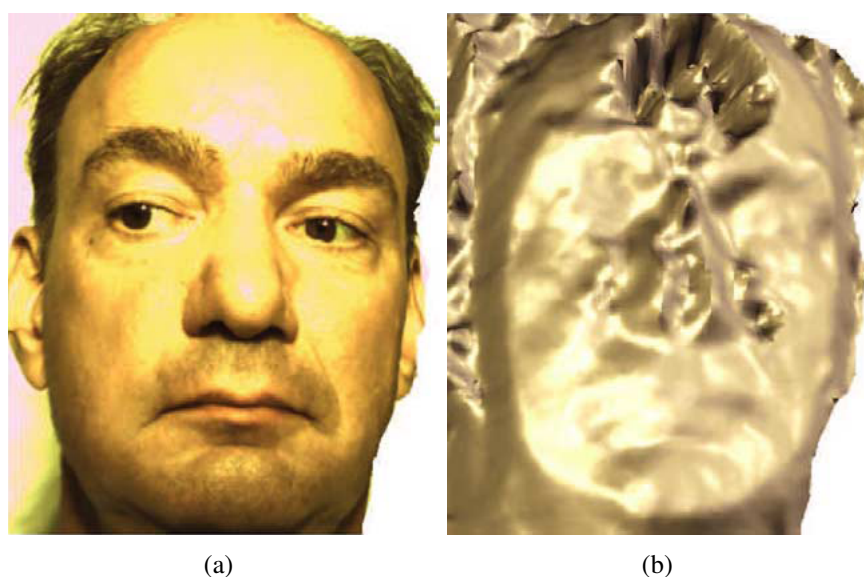


Figura 2.14: Exemplo de como o mapa de textura pode ocultar imperfeições do modelo 3D: (a) modelo 3D renderizado com seu mapa de textura; (b) o mesmo modelo 3D de (a) mas, que foi renderizado como um modelo sombreado. Fonte: [9].

Os pré-processamentos aplicados para remover as imperfeições das imagens 3D, melhoram-nas muito e não são computacionalmente custosos sendo, portanto, aconselhável sua utilização para melhorar os resultados dos algoritmos de reconhecimento facial que utilizam tais imagens.

2.2.3 Representação das Faces

Os algoritmos de reconhecimento facial 3D utilizam várias formas de representação das faces. Muitas vezes, esta representação está associada com o método de reconhecimento utilizado.

- **Curvatura e Características de Superfície:** nesta representação, a curvatura de regiões da face é utilizada em diferentes estratégias para realizar o reconhecimento facial. Por utilizar derivadas parciais de segunda ordem, o cálculo das curvaturas é fortemente afetado pela presença de ruído nos dados de entrada [2, 22];
- **Malhas e Nuvens de Pontos:** nuvem de pontos é a forma mais primitiva de representação de faces, sendo, assim, de difícil utilização. Quando as faces estão neste formato, o ICP é o método mais utilizado para alinhar as imagens [2];
- **Imagens de Profundidade:** nesta representação, os pontos tridimensionais estão dispostos em grade e, portanto, existe informação de vizinhança. Normalmente são utilizados em conjunto com métodos subespaciais, como a análise de componentes principais, *Principal Components Analysis* (PCA);
- **Contorno:** utiliza apenas um contorno para fazer o reconhecimento. O maior problema para estes algoritmos é a extração do contorno. Para se obter o contorno, as faces são cruzadas com diferentes superfícies (plano vertical, horizontal, cilindros, etc.) [2];
- **Análise por Síntese:** modelos deformáveis são utilizados para corrigir pose e iluminação em imagens 2D e aproveitar a informação de profundidade fornecida pelo modelo para fazer o reconhecimento facial [2];
- **Combinação de Representações:** utilização de várias representações, inclusive a 3D, para fazer o reconhecimento facial. A maioria dos trabalhos utiliza esta representação [2].

A escolha da forma de representação da imagem pode ser decisiva para o correto funcionamento do algoritmo. Por exemplo, a SIM utiliza o conceito de vizinhança de pontos (em uma janela $n \times n$) que pode ser facilmente obtida se a face estiver representada sob a forma de uma imagem de profundidade.

2.2.4 Desafios

Algumas questões ainda impedem que sistemas de verificação/reconhecimento baseados em imagens 3D da face sejam amplamente utilizados. Avanços na captura das faces e nos

algoritmos empregados poderiam contribuir para a disseminação desta característica biométrica, em 3D, nos sistemas baseados em biometria.

Embora em fase de barateamento, equipamentos precisos de aquisição de imagens 3D ainda possuem altos preços. Este alto custo pode favorecer a utilização de características biométricas que utilizam equipamentos mais baratos, como por exemplo, a impressão digital. Outras melhorias são apontadas em [8] para aperfeiçoamento dos sensores de imagens 3D: (1) redução na intensidade e na frequência de ocorrência dos artefatos; (2) aumentar a profundidade de campo; (3) aumentar a resolução espacial e de profundidade; (4) reduzir o tempo de captura.

Como mostrado na Seção 2.2.2, a aquisição de imagens 3D está sujeita à presença de imperfeições. Estas imperfeições possuem diversas origens, podendo ocorrer devido ao tipo de região e até mesmo iluminação imprópria, neste caso, dependendo do método de aquisição [8].

Uma outra limitação dos sistemas de aquisição atuais é a pequena profundidade de campo⁴ que possuem. O aumento desta propriedade é especialmente relevante quando o sistema será utilizado em condições onde não se espera a cooperação da pessoa que será analisada. A profundidade de campo para aquisição de imagens utilizáveis para o reconhecimento facial 3D varia de 30 centímetros, para sistemas baseados em visão estéreo, a um metro, para sistemas baseados em luz estruturada [8].

Existem evidências [15] de que algoritmos de reconhecimento facial 3D podem se beneficiar de uma resolução de profundidade com precisão inferior a um milímetro [8]. No entanto, esta informação é baseada em experimentos com uma base de dados peculiar e com um algoritmo igualmente peculiar. Como o preço dos sensores pode aumentar dramaticamente de acordo com a precisão do mesmo, é necessário um estudo maior para determinar o que é realmente necessário para aplicações de reconhecimento facial [9].

Por fim, o tempo de aquisição da imagem pode representar um problema dependendo da aplicação do sistema de reconhecimento. Este problema é mais evidente nos sistemas baseados em luz estruturada uma vez que o tempo de captura destes sistemas é muito maior que o dos baseados em visão estéreo. O problema do tempo reside no fato de que quando a captura é lenta, a movimentação da pessoa que está sendo analisada pode contribuir para a geração de imagens com distorção e que dificilmente podem ser utilizadas em sistemas de reconhecimento facial. Para sistemas de autenticação, onde presume-se a cooperação das pessoas analisadas, o tempo de aquisição não representa um fator crítico.

⁴Distância entre as posições extremas de um objeto cuja imagem formada por um sistema óptico é de boa qualidade.

2.3 SIMULATED ANNEALING

O *Simulated Annealing* (SA) [24] (Têmpera Simulada, em tradução livre) é um algoritmo de busca local que a partir de uma solução candidata, iterativamente, move-se para uma solução vizinha que melhor soluciona o problema.

A principal diferença do SA para outros algoritmos de busca local, como por exemplo o *Hill Climbing*, é a possibilidade que o SA tem de aceitar uma solução pior que a atual (a solução atual também é conhecida como solução candidata). Isto faz com que o SA não fique “preso” em mínimos locais e consiga atingir seu objetivo, que é uma solução próxima à do mínimo global. A Figura 2.15 mostra um gráfico de soluções e seus custos associados, evidenciando os mínimos locais e o mínimo global.

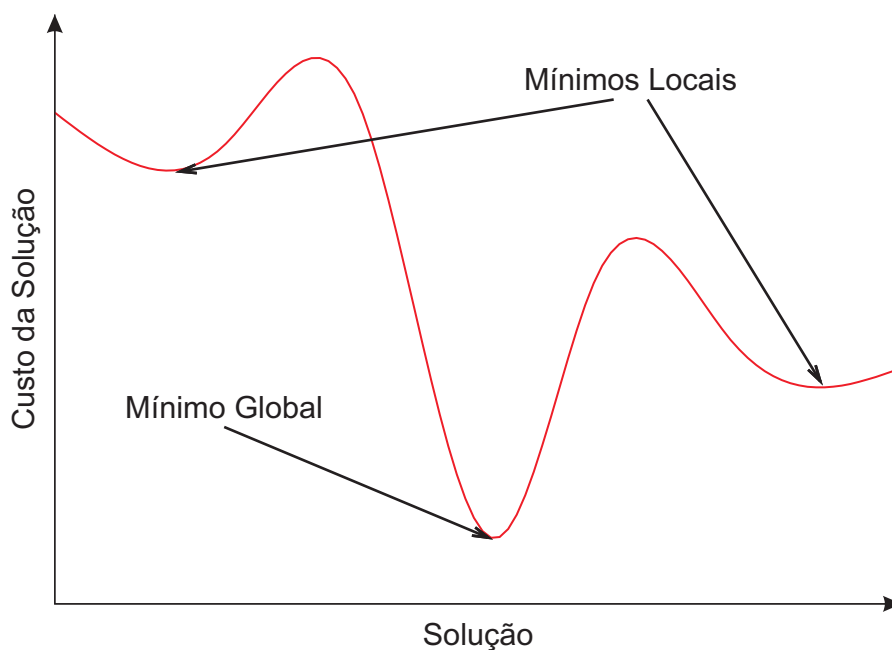


Figura 2.15: Mínimo global e mínimos locais.

O nome do algoritmo revela o fenômeno real no qual foi inspirado:

- *Simulated* – simulado; reproduzido por meio de um modelo;
- *Annealing* – têmpera; processo de aquecimento de metais, vidros, etc. seguido de resfriamento lento e gradual, cujo objetivo é tornar o material mais rígido.

Assim, o SA explora a analogia entre o modo como um metal se resfria e se congela numa estrutura cristalina de energia mínima (o processo real de *annealing*) e a busca por um mínimo num sistema qualquer. O SA foi apresentado em [24].

A Tabela 2.2 estabelece a relação entre os componentes e eventos do *annealing* real e do *Simulated Annealing*.

Simulação Termodinâmica	Otimização Combinatória
Estados do Sistema	Soluções Possíveis
Energia	Custo
Mudança de Estado	Soluções Vizinhas
Temperatura	Parâmetro de Controle
Estado Congelado	Solução Heurística

Tabela 2.2: Relação entre o *Simulated Annealing* e o processo real de *annealing*.

Os algoritmos da classe do SA, também conhecidos como “algoritmos de limiar” (*threshold algorithms*), possuem um importante papel na pesquisa local por duas razões: (1) apresentam um bom desempenho quando aplicados a um vasto conjunto de problemas práticos; (2) alguns deles, como é o caso do SA, têm um componente estocástico que facilita uma análise teórica da sua convergência assintótica [40].

Embora o SA possa ser aplicado em um grande conjunto de problemas, aconselha-se sua utilização nos seguintes casos:

- Conhecimento sobre o problema é escasso;
- Método de resolução do problema se apresenta complexo;
- Soluções próximas da ótima são aceitáveis.

Para que seja possível resolver um problema utilizando o SA, deve-se possuir alguns elementos. São eles:

- **Função de energia** - também conhecida como função de custo ou função-objetivo. Esta função avalia as soluções do problema, informando o quanto elas o satisfazem;
- **Função de incremento** - uma função que retorna uma solução vizinha a partir de uma dada solução. Esta função permite explorar a vizinhança de uma solução candidata;
- **Solução inicial** - um conjunto de parâmetros que aplicados ao problema visam solucioná-lo. Para a solução inicial, se não for possível obter uma solução aproximada, estes parâmetros podem ser totalmente aleatórios.

Quando se obtém uma solução vizinha da solução atual e calcula-se sua energia, podem ocorrer duas situações distintas tendo-se em vista a minimização a energia:

1. A energia da solução vizinha é menor ou igual que a da atual: ou seja, a energia manteve-se igual ou diminuiu. A nova solução será aceita e se tornará a solução candidata (ou atual);
2. A energia da solução vizinha é maior que a da atual: ou seja, a energia aumentou. A nova solução poderá ser aceita de acordo com uma probabilidade.

No segundo caso, a probabilidade de aceitação da nova solução é dada pela Equação 2.6. A função $h(\Delta E)$ determina, com base na temperatura, a probabilidade de aceitação de um novo candidato a partir do valor de sua energia e da energia do candidato anterior (o candidato atual, até então). Ou seja, baseia-se na chance de obtenção de um novo estado de energia E_{k+1} a partir de um prévio estado de energia E_k ,

$$\begin{aligned} h(\Delta E) &= \frac{e^{-\frac{E_{k+1}}{T}}}{e^{-\frac{E_{k+1}}{T}} + e^{-\frac{E_k}{T}}} \\ &= \frac{1}{1 + e^{\frac{\Delta E}{T}}} \\ &\approx e^{-\frac{\Delta E}{T}} \end{aligned} \quad (2.6)$$

onde, T corresponde à temperatura atual do sistema e ΔE representa a diferença entre as energias do candidato atual e do novo candidato, i. e., $\Delta E = E_{k+1} - E_k$. Essencialmente, esta é a distribuição de Boltzmann [23].

Para efetivamente determinar se a solução será aceita ou não, $h(\Delta E)$ é comparada com uma variável aleatória r como pode ser vista na Equação 2.7:

$$\begin{aligned} \text{se } r < h(\Delta E), \text{ então } &\mathbf{aceita} \text{ a nova solução;} \\ \text{se } r \geq h(\Delta E), \text{ então } &\mathbf{rejeita} \text{ a nova solução.} \end{aligned} \quad (2.7)$$

onde, r é um número real aleatório e está contido no intervalo ($0 < r < 1$).

Analisando em conjunto as Equações 2.6 e 2.7, é possível verificar que: (1) a possibilidade de aceitação de uma solução pior que a atual decresce juntamente com a temperatura; (2) quanto maior a diferença entre a energia da nova solução e a da solução atual, menor a chance de aceitação desta nova solução (ou seja, quanto pior a nova solução maior a possibilidade desta ser rejeitada).

O funcionamento do algoritmo do SA pode ser visto através do pseudo-código abaixo, Algoritmo 1, e/ou do fluxograma na Figura 2.16:

Algoritmo 1 *Simulated Annealing*

```

1: função SA( $s_0$ )
2:    $candidato \leftarrow s_0$ 
3:    $T \leftarrow T_0$ 
4:   repita
5:     para  $i \leftarrow 1$  até  $N_i$  faça
6:        $proximo \leftarrow vizinho(candidato)$            ▷ Solução vizinha de  $candidato$ 
7:        $\Delta E \leftarrow energia(proximo) - energia(candidato)$ 
8:       se  $\Delta E \leq 0$  então
9:          $candidato \leftarrow proximo$                  ▷ Sempre aceita uma solução melhor
10:      senão
11:        se  $aleatorio() < e^{-\frac{\Delta E}{T}}$  então
12:           $candidato \leftarrow proximo$              ▷ Eventualmente, aceita uma solução pior
13:        fim se
14:      fim se
15:    fim para
16:     $T \leftarrow proximaTemperatura(T)$              ▷ Diminui a temperatura
17:  até  $T < T_f$ 
18:  retorna  $candidato$ 
19: fim função

```

onde:

- s_0 é a solução inicial (candidato a mínimo);
- T_0 e T_f correspondem à temperatura inicial e final, respectivamente;
- N_i é o número de tentativas em cada temperatura;
- $vizinho$ é a função que retorna uma solução vizinha à atual;
- $energia$ é a função de energia;
- $aleatorio$ retorna um número real no intervalo $[0,1)$;
- $proximaTemperatura$ é a função que reduz a temperatura em cada iteração.

2.3.1 Esquema de Resfriamento

O modo como se reduz a temperatura é crucial para o sucesso do algoritmo. O esquema de resfriamento do *Simulated Annealing* é composto de quatro componentes, são eles:

- Temperatura inicial;

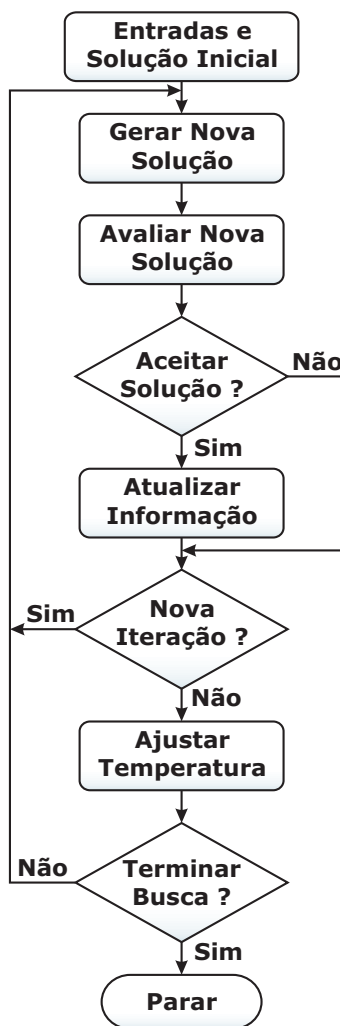


Figura 2.16: Fluxograma do *Simulated Annealing*.

- Temperatura final;
- Decremento da temperatura;
- Número de iterações em cada temperatura.

Cada um deles será melhor detalhado nas seções subseqüentes.

2.3.1.1 Temperatura Inicial

A temperatura inicial deve ser quente o suficiente para permitir a aceitação de quase todas as soluções de uma dada vizinhança. Caso isso não aconteça, a solução final pode ser a própria solução inicial, ou muito próxima desta e, nesse caso, a implementação seria praticamente idêntica a do *Hill Climbing*. Entretanto, se a temperatura for muito elevada, todo e qualquer

vizinho será aceito e, portanto, a busca será praticamente aleatória (pelo menos nos estágios iniciais)⁵.

Atualmente, não existe nenhum método para encontrar uma temperatura inicial adequada para qualquer problema. No entanto, existem algumas considerações que podem contribuir na escolha de uma temperatura inicial conveniente.

Se a maior distância (diferença na função de custo) entre um vizinho e outro é conhecida, então esta informação pode ser utilizada para determinar uma temperatura inicial adequada.

Um outro método, sugerido em [31], consiste em iniciar o sistema com uma alta temperatura e reduzi-la rapidamente, até que cerca de 60% das soluções ruins⁶ sejam aceitas. A temperatura que satisfizer este critério será a verdadeira temperatura inicial do sistema e, portanto, passa a ser reduzida normalmente (de acordo com o critério utilizado).

Uma idéia similar, sugerida em [19], é a de aquecer o sistema rapidamente até que uma certa porcentagem de soluções ruins sejam aceitas e então inicia-se o resfriamento normal. Este método se parece mais com o real processo de *annealing*, onde um material é aquecido até que se liquéfique e então o processo de resfriamento se inicia (i.e., uma vez que o material encontra-se líquido, não há a necessidade de continuar a aquecê-lo).

2.3.1.2 Temperatura Final

A temperatura final é utilizada como parâmetro de controle e estabelece quando o algoritmo deverá ser encerrado. Normalmente, a temperatura é reduzida até que atinja o valor zero (nesse caso, a temperatura final seria zero). No entanto, isto pode fazer com que o algoritmo seja executado por um tempo muito maior que o necessário, principalmente quando a temperatura é reduzida geometricamente (ver Seção 2.3.1.3).

Na prática, não é necessário permitir que a temperatura atinja zero, já que à medida que ela se aproxima de zero, a chance de se aceitar uma solução ruim é praticamente a mesma que quando a temperatura é igual a zero. Portanto, o critério de parada pode ser uma temperatura adequadamente baixa ou então quando o sistema encontra-se “congelado”⁷ na temperatura atual.

⁵A busca será aleatória até que a temperatura seja baixa o suficiente para que o algoritmo passe a agir como o SA.

⁶Neste contexto, soluções ruins são aquelas que apresentam uma energia maior que a da solução atual (solução candidata).

⁷Um sistema é considerado “congelado” quando soluções ruins não são mais aceitas e soluções melhores não são encontradas.

2.3.1.3 Decremento da Temperatura

O decréto da temperatura controla o modo como será feito o resfriamento do sistema e possibilita chegar na temperatura final (critério de parada) a partir da temperatura inicial.

O sucesso do algoritmo depende fortemente do modo como a temperatura é reduzida. Normalmente, a redução da temperatura pode ser **linear** (Equação 2.8) ou **geométrica** (Equação 2.9).

$$T_{k+1} = T_k - \Delta T \quad (2.8)$$

onde,

- T_{k+1} corresponde ao valor da temperatura na próxima iteração;
- T_k representa o valor da temperatura na iteração atual;
- ΔT é o valor que será reduzido da temperatura em cada iteração.

$$T_{k+1} = \alpha \times T_k \quad (2.9)$$

onde, α determina a velocidade de redução da temperatura e se situa no intervalo aberto $(0; 1)$.

No caso da redução geométrica, estudos mostram que α deve se situar entre 0,8 e 0,99, com melhores resultados nas proximidades do limite superior. No entanto, quanto maior o valor de α , maior será o tempo necessário para que a temperatura mínima (final) seja atingida e, portanto, maior será o esforço computacional do algoritmo. O resfriamento geométrico foi o proposto por Kirkpatrick *et al.* [24], com $\alpha = 0,95$.

2.3.1.4 Número de Iterações em Cada Temperatura

Uma última decisão que deve ser tomada, é o número de iterações do algoritmo em cada temperatura.

A teoria indica que se deve permitir iterações suficientes em cada temperatura para que o sistema se estabilize nesta temperatura. Em contrapartida, a teoria também indica que o número de iterações em cada temperatura, para alcançar tal estabilidade, pode ser exponencial ao tamanho do problema. Como isto é impraticável, se faz necessário chegar a um termo comum. Este pode ser obtido das seguintes formas:

- realizar um grande número de iterações em poucas temperaturas;
- realizar um pequeno número de iterações em várias temperaturas;
- um equilíbrio entre as opções anteriores.

Ao invés de se utilizar um número fixo de iterações, é possível alterar, dinamicamente, o número de iterações conforme o progresso do algoritmo. Em temperatura baixas, é importante que seja feito um grande número de iterações para que o mínimo onde o SA se encontra (seja local ou global) seja totalmente explorado. Em temperaturas mais altas, o número de iterações pode ser menor, já que nos estágios iniciais do algoritmo não é necessário que toda a vizinhança seja explorada.

Um outro método sugerido por Lundy em [28], é o de fazer apenas uma iteração em cada temperatura e reduzi-la muito vagarosamente. A fórmula utilizada no decremento da temperatura pode ser vista na Equação 2.10

$$T_{k+1} = \frac{T_k}{1 + (\beta \times T_k)} \quad (2.10)$$

onde, β é um valor adequadamente pequeno.

2.3.2 Decisões Específicas do Problema

O esquema de resfriamento (Seção 2.3.1) condiz, principalmente, às decisões feitas no contexto do *Simulated Annealing*. Existe um outro conjunto de decisões necessárias, que são aquelas específicas do problema que se está tentando resolver.

2.3.2.1 Função de Custo

A função de custo avalia as soluções do problema que está sendo tratado pelo SA, indicando o quão bem cada solução satisfaz o problema.

Ao se implementar uma função de custo, duas preocupações devem ser consideradas: (1) ter certeza de que a função de custo representa o problema que se está tentando resolver; (2) implementá-la de forma que seja calculada do modo mais eficiente possível, já que normalmente a função de custo é o “gargalo” do algoritmo. Algumas técnicas que ajudam no desenvolvimento de uma função de custo eficiente são:

- **Avaliação Delta** - quando se utiliza o custo da solução atual no cálculo do custo de uma solução vizinha;

- **Avaliação Parcial** - quando se utiliza uma função simplificada de avaliação que não retorna o valor exato mas fornece uma boa aproximação (na indicação da qualidade da solução).

Se possível, a função de custo deve ser desenvolvida de modo que seja capaz de “guiar” o algoritmo na busca pelo mínimo. Um modo de conseguir isto é evitar funções de custo onde vários estados diferentes possuem o mesmo custo, pois, caso contrário, o SA ficaria sem conhecimento para determinar para onde ele deve seguir. No espaço de busca, esses vários estados com o mesmo custo podem ser vistos como um platô.

Muitas funções de custo têm de ser preocupar com soluções que podem ser ilegais. Quando isto é necessário, frequentemente, são utilizadas dois tipos de restrições:

- **Restrições fortes** - estas restrições não podem ser violadas em uma solução praticável (e.g., na alocação de espaço para um campus, uma restrição forte pode ser a impossibilidade de que os professores ocupem salas menores que um certo tamanho);
- **Restrições fracas** - estas restrições não deveriam, idealmente, ser violadas, mas, se forem, ainda assim a solução é praticável (e.g., no escalonamento de atividades, estudantes não devem ter exames marcados em aulas consecutivas).

Normalmente, as restrições são ponderadas. Pesos altos são dados às restrições fortes, para que, caso sejam violadas, produzam um alto valor de custo e, assim, invalidando a solução. As restrições fracas são ponderadas de acordo com sua importância. Algumas vezes, o valor dos pesos pode mudar conforme o algoritmo progride. Isto permite, por exemplo, que restrições fortes sejam aceitas no início do algoritmo e rejeitadas no final.

2.3.2.2 Estrutura de Vizinhança

Quando se está definindo o problema, uma das primeiras considerações será como mover de um estado, ou solução, para outro. Ou seja, é necessário definir uma vizinhança. Em outras palavras, quando se está em um certo estado, é preciso definir quais outros estados são atingíveis.

Por exemplo, no problema do caixeiro-viajante, a função que retorna um estado vizinho pode ser definida como a execução de apenas uma troca na ordem de visita entre duas cidades (do estado atual).

Alguns resultados indicam que a estrutura de vizinhança deva ser simétrica, ou seja, se a partir de um estado i é possível chegar a j então, também deve ser possível chegar a i a partir de

j. Também é importante assegurar que qualquer estado do espaço de solução possa ser atingido a partir de qualquer outro estado.

2.3.2.3 Espaço de Solução

Experimentos mostraram que, para um espaço de solução (espaço de busca) pequeno, o número de iterações necessárias para que o SA convirja para uma solução ótima é menor do que para um espaço de busca maior.

Se na definição da função de custo forem permitidas soluções impraticáveis (ou seja, que violam as restrições fortes), o espaço de busca será maior e, portanto, serão necessárias mais iterações para que haja a convergência.

Além de diminuir o espaço de solução, também recomenda-se que a vizinhança seja pequena. Isto faz com que a busca seja mais rápida, no entanto, a possibilidade de haver melhorias significativamente grandes entre uma iteração e outra será reduzida.

2.3.2.4 Sumário

Na definição do problema, existem alguns conflitos de interesse. Necessita-se de uma função de custo que modele bem o problema mas que seja rápida no cálculo. Necessita-se de uma função de custo que não permita soluções impraticáveis mas, algumas vezes é necessário explorar as áreas dessas soluções para que seja possível encontrar uma boa solução.

Necessita-se de um espaço de solução o menor possível, no entanto, não se quer restringir muito a busca. Também se quer uma vizinhança pequena mas, novamente, não em detrimento da qualidade da solução.

Ou seja, o objetivo é fazer o uso mais eficiente possível de cada iteração, enquanto tentando assegurar que uma solução de boa qualidade será encontrada.

2.4 MEDIDAS DE ERRO

A utilização de medidas de erro robustas é crucial para o correto funcionamento do algoritmo proposto pois, como já foi provado em [36], o MSE e o RMSE (as medidas mais utilizadas) podem levar a convergências locais, o que ocasiona alinhamentos imprecisos.

As duas medidas utilizadas no algoritmo são: o MSAC, que é utilizado para conseguir um

alinhamento grosseiro; e a SIM, que mede a interpenetração das superfícies e é utilizada para um ajuste fino no alinhamento.

2.4.1 *M-estimator Sample Consensus*

O *M-estimator Sample Consensus* (MSAC) foi apresentado em [39] e é um avaliador robusto para a correspondência de pontos.

O MSAC é semelhante ao *Random Sample Consensus* (RANSAC) [20] e, ambos classificam os pontos como *inliers* ou *outliers*. *Inliers* são pontos cujas correspondências são consideradas válidas (segundo algum critério), já as correspondências dos *outliers* consideradas inválidas.

O RANSAC classifica os pontos de acordo com um limiar T , definido de acordo com cada problema, como pode ser visto nas Equações 2.11 e 2.12

$$C = \sum_i \rho(e_i^2) \quad (2.11)$$

onde, $\rho()$ é:

$$\rho(e^2) = \begin{cases} 0 & \text{se } e^2 < T^2 \\ \text{constante} & \text{se } e^2 \geq T^2 \end{cases} \quad (2.12)$$

onde, C corresponde à função de custo do RANSAC, ρ indica o custo de um determinado par de pontos correspondentes, e corresponde ao erro associado a um determinado par correspondente e T ao limiar escolhido.

Em outras palavras, o RANSAC não aumenta o custo para *inliers* e aplica uma penalidade constante para *outliers*. Portanto, quanto maior o valor de T^2 maior será o número de soluções com o mesmo valor de C , resultando em uma avaliação ineficiente. Por exemplo, se T for suficientemente alto, então todas as soluções possuirão o mesmo custo, já que todos os pares de pontos correspondentes seriam considerados *inliers* [39]. No MSAC, sem custo adicional de processamento, esta situação indesejável pode ser remediada. Um nova função de custo, que deve ser minimizada, é definida:

$$C_2 = \sum_i \rho_2(e_i^2) \quad (2.13)$$

onde, a função robusta $\rho_2()$ é:

$$\rho_2(e^2) = \begin{cases} e^2 & \text{se } e^2 < T^2 \\ T^2 & \text{se } e^2 \geq T^2 \end{cases} \quad (2.14)$$

Na Equação 2.14 pode ser observado que *outliers* ainda recebem uma penalidade fixa mas, a cada *inlier* é atribuído um valor de acordo com quão adequado ele é para a solução do problema. A Figura 2.17 ilustra o funcionamento do MSAC: a união das linhas tracejada e pontilhada representa o erro associado ao par de pontos correspondentes; a linha pontilhada representa o valor retornado por ρ_2 e que é agregado à função de custo total C_2 ; a região em cinza claro representa a área (definida pelo limiar T) onde os pontos correspondentes são considerados *inliers*.

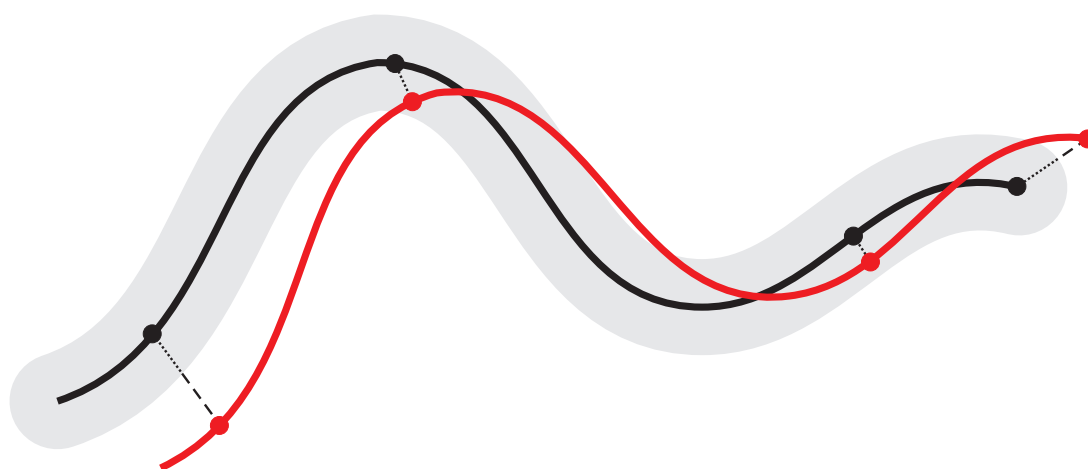


Figura 2.17: Ilustração do funcionamento do MSAC.

2.4.2 Surface Interpenetration Measure

A *Surface Interpenetration Measure* (SIM) foi desenvolvida através de uma análise visual de duas superfícies alinhadas, renderizadas em cores diferentes, que cruzavam-se repetidamente na área de sobreposição. O efeito da interpenetração ocorre devido a pequenas distorções locais causadas pelo equipamento de aquisição ou pela presença de ruído, gerando uma certa “rugosidade” nas imagens adquiridas. Sendo assim, até mesmo superfícies planas apresentam esta rugosidade e, portanto, é possível assumir que, independente do formato da superfície, a interpenetração sempre ocorrerá [5, 35, 36].

A SIM é uma medida que quantifica a interpenetração, possibilitando a avaliação precisa do registro de imagens e provendo um controle altamente robusto. A SIM baseia-se no vetor normal à superfície, calculado através do encaixe de planos dos quadrados mínimos locais, em cada ponto.

É possível definir a SIM da seguinte forma: dado o alinhamento de duas imagens, A e B , é identificado o conjunto de pontos interpenetrados de A em relação a B . Para cada ponto $p \in A$, é definida a vizinhança N_p como uma pequena janela $n \times n$ centrada em p . Baseando-se no fato de que a interpentração é um efeito local, foi escolhido $n = 5$. Sendo q um ponto na vizinhança N_p , c o ponto correspondente a p na imagem B (obtido pelo cálculo de vizinho mais próximo, utilizando a estrutura *KD-tree* para otimizar a busca), e \vec{n}_c o vetor normal à superfície no ponto c , o conjunto de pontos interpenetrados é definido como: (Equação 2.15) [35, 36]:

$$C_{(A,B)} = \{p \in A \mid [(q_i - c) \cdot \vec{n}_c][(q_j - c) \cdot \vec{n}_c] < 0\} \quad (2.15)$$

onde, $q_i, q_j \in N_p$ e $i \neq j$. Este conjunto de pontos interpenetrados engloba os pontos em A cujas vizinhanças contêm pelo menos um par de pontos separados pelo plano tangente local, calculado a partir de seus correspondentes em B . A Figura 2.18 ilustra a situação acima descrita.

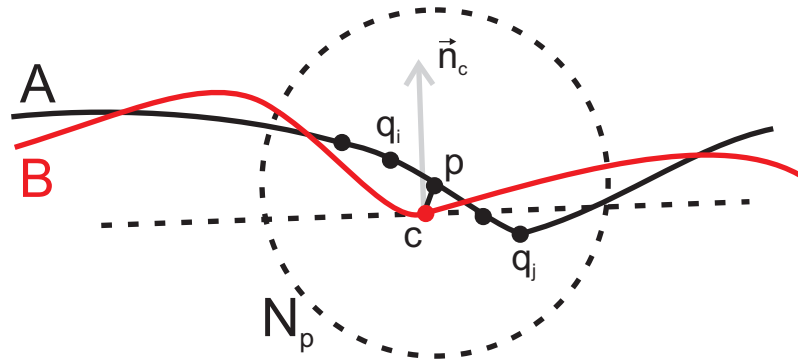


Figura 2.18: Ponto interpenetrado p de A em relação a B . Adaptado de: [36]

Desta forma, apresenta-se a SIM como a porcentagem dos pontos de A que estão interpenetrados em B , como na Equação 2.16.

$$SIM_{(A,B)} = \frac{|C_{(A,B)}|}{|A|} \quad (2.16)$$

Assim, definimos a $SIM_{(A,B)}$ como a quantidade de pontos de A que estão interpenetrados em B ($|C_{(A,B)}|$) dividida pela quantidade de pontos em A ($|A|$). Onde, $|C_{(A,B)}| \leq |A|$ e a cardinalidade de ambos conjuntos é finita e pertencente ao conjunto dos números naturais.

Comparada ao MSE, a SIM é uma medida mais sensível e mais eficiente quando utilizada para refinar bons alinhamentos e, assim, torná-los mais precisos. Por exemplo, um “correto” alinhamento com baixo valor de MSE pode não apresentar interpentração se as superfícies são paralelas e possuem um pequeno deslocamento. No entanto, maximizando a SIM, é possível atingir um alinhamento preciso e, ainda assim, manter um baixo valor de MSE [36].

A Figura 2.19 mostra como o MSE pode falhar ao indicar um bom alinhamento, nela são apresentados diferentes alinhamentos, todos com baixos e semelhantes valores de MSE, no entanto, a situação (b) é a que apresenta o alinhamento mais preciso mas, pode não ser a que apresenta o menor valor de MSE entre as três. A SIM, sozinha, não é suficiente para deter-

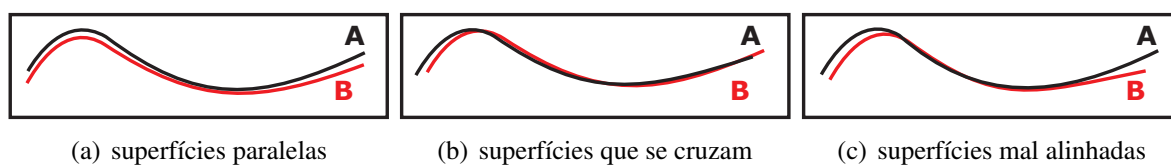


Figura 2.19: Diferentes alinhamentos de duas imagens A e B , todas com baixo MSE. (a) superfícies paralelas, (b) superfícies que se cruzam em diferentes regiões, (c) superfícies mal alinhadas com baixo MSE. Adaptado de: [36]

minar se um alinhamento é correto mas, pode confirmar um bom resultado quando analisada juntamente com o MSE [36].

Através da análise da SIM, foi possível chegar às seguintes conclusões:

- o registro de duas imagens que apresentam uma boa interpenetração, produz altos valores de SIM;
- alinhamentos incorretos apresentam baixos valores de SIM;
- pequenas diferenças no MSE podem produzir grandes variações na SIM;
- alinhamentos com altos valores de SIM apresentam uma distância muito pequena entre os pontos correspondentes.

A SIM é uma métrica robusta e sensível. Um indicador ideal para medir a qualidade de alinhamentos, quando comparando alinhamentos “razoáveis”.

CAPÍTULO 3

TRABALHOS RELACIONADOS

Em [14, 17] é feito um estudo sobre o efeito que as expressões faciais têm sobre o reconhecimento facial 3D.

As duas abordagens mais utilizadas para reconhecimento facial 3D foram utilizadas para demonstrar a degradação de performance provocada pela presença de expressões faciais: (1) algoritmo baseado em ICP; (2) algoritmo baseado em análise de componentes principais, *Principal Components Analysis* (PCA) [17]. As taxas de reconhecimento *rank-one* obtidas podem ser vistas na Tabela 3.1.

ICP		PCA	
sem expressão	com expressão	sem expressão	com expressão
93,6%	61,1%	75,5%	52,4%

Tabela 3.1: Efeito da expressão em faces 3D no desempenho das abordagens baseadas em ICP e PCA.

Após a comprovação de que a presença de expressões degrada a performance dos algoritmos de reconhecimento facial 3D em cerca de 28%, foi sugerida a utilização de regiões da face que são menos sensíveis às deformações provocadas pelas várias expressões. A região proposta foi a do nariz, incluindo a área ao redor deste [14].

Uma outra vantagem de se utilizar o nariz para reconhecimento é que, por se tratar de uma região pequena, o número de pontos utilizados é consideravelmente menor que o da face toda e, portanto a execução do algoritmo de reconhecimento é mais rápida.

Todos os algoritmos apresentados neste capítulo utilizam o ICP como método de registro. O ICP considera as faces 3D como um conjunto de pontos. E, o esquema básico do ICP para alinhar dois conjuntos de pontos é descrito abaixo [26, 27]:

1. Seleciona os pontos de controle em um conjunto de pontos;
2. Encontra os pontos mais próximos no outro conjunto de pontos (faz a correspondência);
3. Calcula a transformação ótima entre os dois conjuntos baseado na correspondência atual;
4. Aplica a transformação nos pontos;

5. Repete os passos 2, 3 e 4 até que haja convergência (ou seja, a diferença entre o alinhamento de duas iterações consecutivas é pequena ou nenhuma).

Entretanto, antes de seguir os passos supra citados, é necessário que haja um alinhamento inicial entre os dois conjuntos de pontos.

Em [17], foi descrito o algoritmo *Adaptive Rigid Multi-region Selection (ARMS)* baseado nas observações e resultados obtidos em [14], onde foram utilizadas três segmentações da região do nariz: (1) nariz propriamente dito; (2) área circular ao redor do nariz; (3) área elíptica ao redor do nariz. Estas áreas podem ser observadas na Figura 3.1.

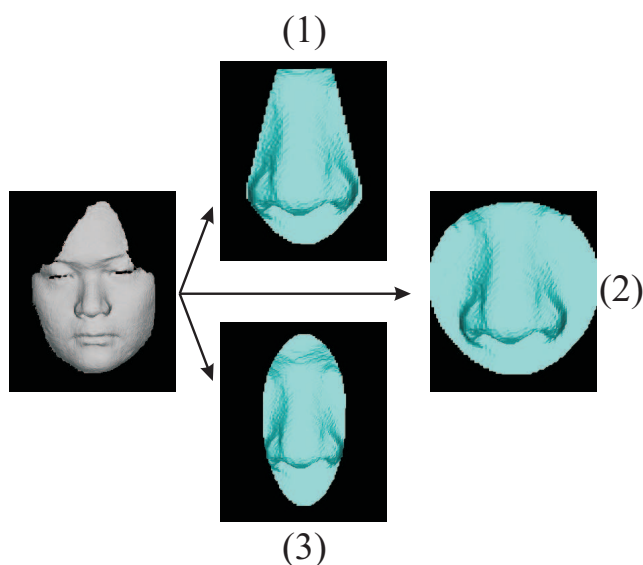


Figura 3.1: Segmentações do nariz utilizadas em [14, 16, 17].

Durante os experimentos com o ARMS, descobriu-se que a combinação da área elíptica ao redor do nariz e a do próprio nariz já produzia resultados semelhantes aos que utilizavam as três regiões e, por esse motivo, optou-se por utilizar somente estas duas regiões. Com a finalidade de combinar os resultados dos alinhamentos com cada uma das segmentações, foi utilizada a regra do produto que, nada mais é que a multiplicação do RMSE de cada um dos alinhamentos produzidos.

A Tabela 3.2 exibe as taxas de reconhecimento *rank-one* obtidas com o ARMS para as suas duas versões: (1) versão manual, que requer a localização de determinados pontos da face; (2) versão automática, que não requer qualquer supervisão.

ARMS automático		ARMS manual	
sem expressão	com expressão	sem expressão	com expressão
96,1%	87,1%	96,5%	88,6%

Tabela 3.2: Taxas de reconhecimento *rank-one* obtidas pelo algoritmo ARMS.

O ARMS produziu bons resultados utilizando apenas o nariz como região de interesse, *Region Of Interest* (ROI), no entanto, utilizando o ICP como método de registro e o RMSE como medida de erro, o algoritmo fica suscetível aos problemas anteriormente mencionados. Primeiro, requer um bom pré-alinhamento que, neste caso, é proporcionado por um método que localiza determinados pontos da face para fazer o alinhamento inicial, no entanto, se este método falhar, é pouco provável que o alinhamento seja bem sucedido. Segundo, mesmo com um bom pré-alinhamento, o ICP é guiado pelo RMSE que, por si só, não é uma boa medida para produzir e avaliar alinhamentos precisos.

Em [16] foi apresentado o trabalho oriundo do ARMS [17]. Diferentemente do ARMS, neste trabalho as três segmentações do nariz são utilizadas na composição de seu melhor resultado. Os resultados obtidos individualmente para a região do nariz (N), região circular do nariz (C) e região elíptica do nariz (I), bem como os resultados advindos de suas combinações através da regra do produto, tanto para casos com e sem expressão facial podem ser vistos na Tabela 3.3.

Região	Taxa de Reconhecimento <i>Rank-One</i>	
	Com Expressão	Sem Expressão
Nariz Região Circular (C)	77,1%	95,9%
Nariz Região (N)	83,5%	95,4%
Nariz Região Elíptica (I)	82,7%	96,6%
Regiões C + N	85,1%	96,8%
Regiões C + I	83,0%	96,9%
Regiões N + I	87,1%	96,6%
Regiões C + N + I	86,1%	97,1%

Tabela 3.3: Taxas de reconhecimento *rank-one* obtidas pelo algoritmo descrito em [16].

Como pode ser visto, o melhor resultado para os casos onde não há expressão facial foi obtido pela combinação das três regiões do nariz, alcançando uma taxa de reconhecimento *rank-one* de 97,1%, já para os casos onde há expressão facial, o melhor resultado foi obtido através da combinação da região do nariz com a região elíptica ao redor do nariz, atingindo uma taxa de reconhecimento *rank-one* de 87,1%.

Em [16], é feita uma comparação dos resultados obtidos utilizando as segmentações do nariz, com um método de referência que utiliza o ICP para a região da face. Para este método de referência, foram marcados, manualmente, pontos de referência para prover um alinhamento inicial para o ICP. Ou seja, este método de referência representa um nível idealizado de desempenho para a abordagem [16]. Ainda assim, os resultados ficaram bem abaixo dos obtidos pelo algoritmo de [16], atingindo 61,5% e 91,0% de taxa de reconhecimento *rank-one* para casos

com e sem expressão facial, respectivamente.

Assim como em [17], os resultados obtidos em [16] são satisfatórios, no entanto carecem dos mesmos problemas apresentados pelo ARMS.

A mesma base de dados foi utilizada em [17] e em [16]. As imagens da base foram obtidas pelo sensor laser Minolta Vivid 900/910 a partir de 546 indivíduos, sendo que 2.798 são imagens onde não há expressão facial, divididas entre uma *gallery* com 449 imagens e 9 *probes* com 2.349 no total, e 1.590 com expressão facial, divididos em 8 *probes*. A *gallery* das faces com expressão facial possui 355 imagens e é um subconjunto da *gallery* das faces sem expressão facial.

Em [26, 27], é proposto um algoritmo de reconhecimento facial que utiliza tanto as imagens de profundidade quanto as imagens de textura (capturadas juntamente com as de profundidade). Este algoritmo possui dois módulos principais, cada um lidando com um tipo de imagem.

O primeiro módulo trabalha com as imagens de profundidade. Para que possa ser utilizado, é preciso que para cada indivíduo a ser identificado exista um modelo tridimensional de sua face na base de faces. Este modelo é gerado a partir de várias imagens de profundidade combinadas (neste experimento, foram utilizadas 5 imagens). Uma vez que já existam os modelos tridimensionais, uma imagem de profundidade com sua respectiva textura (imagem 2D simples) são submetidas ao algoritmo para que seja feito o reconhecimento, seguindo as seguintes etapas:

- alinhamento grosseiro: é feito um alinhamento grosseiro baseado nas características da face submetida. Ao menos três pontos devem ser usados para garantir um alinhamento razoável. Os pontos utilizados encontram-se nas proximidades do nariz e dos olhos, ver Figura 3.2. Os pontos foram informados manualmente;
- alinhamento fino: este alinhamento é feito através de um ICP híbrido. Este ICP é uma integração do ICP clássico proposto por Besl e McKay [7], com o ICP proposto por Chen e Medioni [18]. A diferença entre os dois é que, enquanto o ICP de Besl e McKay utiliza a distância ponto-a-ponto (Figura 3.3 (a)), o ICP de Chen e Medioni utiliza a distância ponto-a-plano. Este último requer menos iterações para convergir e é menos suscetível a mínimos locais, no entanto, possui um custo computacional superior. O ICP híbrido consiste em executar iterações alternando os dois métodos, começando com o ICP de Besl e McKay;
- distância de alinhamento de superfície: o RMSE minimizado pelo algoritmo do ICP é utilizado como a principal medida de erro do alinhamento. Mais especificamente, utiliza-se o RMSE ponto-a-plano.



Figura 3.2: Pontos selecionados para alinhamento grosseiro, segundo o método descrito em [26, 27]. Fonte: [26]

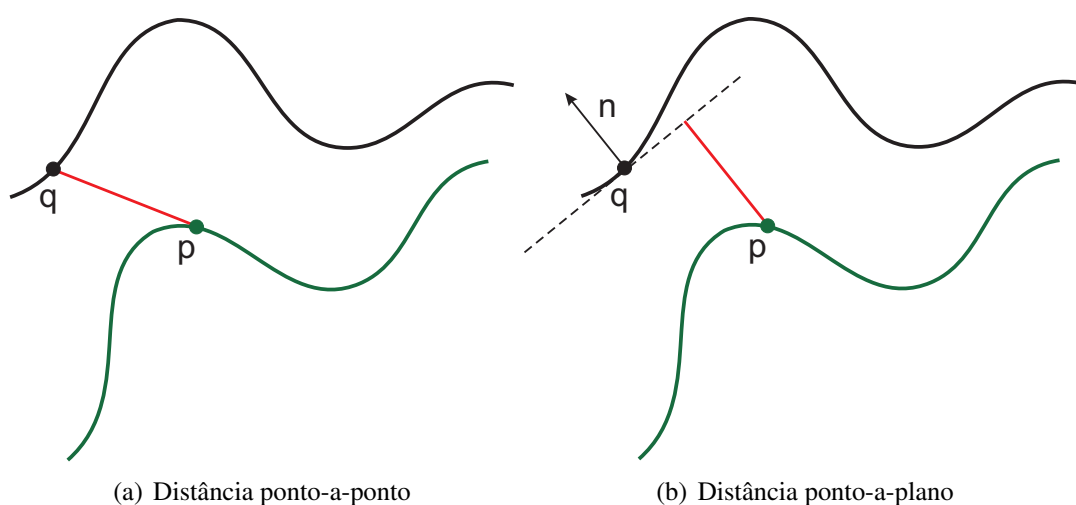


Figura 3.3: Distâncias entre pontos correspondentes. (a) ponto-a-ponto; (b) ponto-a-plano.

O segundo módulo utiliza análise linear discriminante, *Linear Discriminant Analysis* (LDA), [3, 29] para verificar a similaridade de aparência da face submetida para avaliação em um subespaço gerado dinamicamente pelas faces que obtiveram melhores valores no módulo anterior (ou seja, as que foram mais parecidas tridimensionalmente à face submetida). Este subespaço ainda inclui variações de deslocamento e iluminação das faces selecionadas.

O resultado final do algoritmo é a combinação do resultado de cada um dos módulos através de uma soma ponderada. A Tabela 3.4 apresenta as taxas de reconhecimento *rank-one* obtidas em [26]. Os resultados obtidos em [27] podem ser vistos na Tabela 3.5

Esquema Considerado	Toda Base	Sem Expressão
ICP	87%	97%
LDA	77%	84%
ICP + LDA	91%	98%

Tabela 3.4: Taxas de reconhecimento *rank-one* obtidas pelo algoritmo descrito em [26].

Este algoritmo, embora produza resultados satisfatórios, depende fortemente de: (1) um ex-

Esquema Considerado	Com Expressão	Sem Expressão
ICP	68%	98%
LDA	71%	86%
ICP + LDA	77%	99%

Tabela 3.5: Taxas de reconhecimento *rank-one* obtidas pelo algoritmo descrito em [27].

trator de características e (2) um bom alinhamento com o ICP. Esta dependência pode acarretar problemas no desempenho do algoritmo, por exemplo: se o extrator de características não localizar corretamente os pontos para o alinhamento inicial, provavelmente o ICP não irá convergir para um bom registro e as faces selecionadas para gerar o subespaço de busca do LDA não corresponderão às melhores candidatas, gerando um resultado incorreto. Uma outra situação, é a possibilidade do registro do ICP não convergir pois, embora seja utilizada uma versão aprimorada, este ICP ainda se baseia no resultado do RMSE, que não é boa medida para realizar registros precisos, e, se o registro com o ICP falhar, este erro irá se propagar para o LDA, produzindo um novo erro e, por fim, gerando um resultado incorreto.

As imagens utilizadas foram obtidas através de um sensor laser Minolta Vivid 910. Em [26], a *gallery* é composta de 100 modelos 3D da face, já em [27] a *gallery* contém 200 modelos 3D da face. Em ambos trabalhos, a conjunto *probe* consiste de 598 imagens, incluindo pose frontal e perfil. Os resultados da Tabela 3.5, apresenta somente os valores para as imagens com pose frontal.

Os algoritmos apresentados [14, 16, 17, 26, 27] neste capítulo representam o estado da arte em reconhecimento facial 3D. No entanto, todos eles possuem as mesmas deficiências:

- Necessitam de um bom pré-alinhamento;
- Utilizam o ICP, cuja possibilidade de ficar “preso” em mínimos locais é alta, como método de registro;
- Utilizam o RMSE, que dificilmente produz alinhamentos precisos, como medida de erro.

Por esses motivos, se faz necessária a utilização de uma nova abordagem que utilize medidas de erro mais robustas e que não necessite de supervisão.

CAPÍTULO 4

ALGORITMOS DESENVOLVIDOS

Foram desenvolvidas duas versões para o sistema de reconhecimento facial. Em ambas versões, o componente fundamental é o *Simulated Annealing* pois, é através dele que o espaço de busca de seis dimensões é explorado e o melhor alinhamento entre as faces encontrado.

A implementação do *Simulated Annealing* foi baseada naquela encontrada na biblioteca *GNU Scientific Library* (GSL)¹. No entanto, algumas melhorias foram feitas.

As principais diferenças entre a nova implementação e aquela disponível na GSL são: (1) mudança para o paradigma de orientação a objetos (C++); (2) a possibilidade de minimizar e maximizar a função de custo; (3) definir vários esquemas de resfriamento (e.g., esquema definido em [28]); (4) método para encontrar uma temperatura inicial adequada; (5) possibilidade da função de custo acessar e modificar os atributos do SA, o que, na prática, permite mudar seu comportamento durante sua execução; (6) a possibilidade de iniciar o algoritmo com várias soluções iniciais e, prosseguir naquela que apresentar melhor resultado.

O reconhecimento facial realizado pelo sistema desenvolvido, de codinome *Simulated Annealing Matcher* (SAM), é baseado no registro das imagens da face. As transformações geométricas que alinham as duas imagens são representadas por um vetor de seis elementos: T_x , T_y e T_z indicando as translações nos eixos X , Y e Z , respectivamente; R_x , R_y e R_z indicando as rotações nos eixos X , Y e Z , respectivamente. Este vetor pode ser visualizado na Equação 4.1 e é chamado de vetor de transformação:

$$\left[T_x \quad T_y \quad T_z \quad R_x \quad R_y \quad R_z \right] \quad (4.1)$$

Uma vez que a implementação do SA utilizada foi descrita e os componentes do sistema foram contextualizados, Seção 2, seus estágios e suas duas versões podem ser apresentadas.

¹Disponível em: <http://www.gnu.org/software/gsl/>

Embora diferenças significativas existam entre as duas versões, os principais estágios do SAM continuam o mesmo. De modo geral, o SAM, pode ser dividido em três estágios principais:

1. Obtenção e avaliação de pré-alinhamentos;
2. Execução do SA com o MSAC;
3. Execução do SA com a SIM.

Cada um desses estágios é explicado em detalhes nas seções que se seguem.

4.1 ESTÁGIOS

Cada uma das principais etapas que compõem o SAM é descrita nas próximas seções. Iniciando com os pré-alinhamentos e, depois com a execução do SA com as duas medidas de erro.

4.1.1 Pré-alinhamentos

O pré-alinhamento representa a solução inicial do SA. Embora o SAM seja um algoritmo robusto, um pré-alinhamento ruim pode impedir a convergência para o mínimo global. Além disso, um bom pré-alinhamento pode garantir uma convergência mais rápida.

Para assegurar que o método convirja rapidamente, foram utilizados três pré-alinhamentos e, cada um deles será descrito nas seções subsequentes. Cada pré-alinhamento é avaliado com algumas poucas iterações do SA utilizando o MSAC como medida de erro.

Os pré-alinhamentos empregados só utilizam os elementos de translação do vetor de transformação; aos elementos de rotação é atribuído o valor zero, ou seja, sem rotação.

Exemplos de todos os pré-alinhamentos podem ser vistos na Seção 4.2.1 (Resultados).

4.1.1.1 Centro de Massa

O centro de massa (ou centróide) de uma imagem de profundidade corresponde ao valor médio (em cada coordenada) assumido pelos pontos que a compõem.

Seja I uma imagem de n pontos e p_i o i -ésimo ponto desta imagem. Então, seu centro de massa (C) pode ser calculado pela Equação 4.2:

$$C = \frac{1}{n} \sum_{i \in I} p_i \quad (4.2)$$

onde, C possui valor para as coordenadas X , Y e Z , que podem ser referenciadas por $C_{(x)}$, $C_{(y)}$ e $C_{(z)}$, respectivamente.

O alinhamento pelo centro de massa é feito pela subtração dos centros de massa das duas imagens que serão alinhadas. Seja C^1 e C^2 os centros de massa das imagens I^1 e I^2 , respectivamente, e deseja-se alinhar I^2 em relação a I^1 então, o vetor de transformação para o alinhamento pelo centro massa seria (Equação 4.3):

$$\left[\begin{array}{cccccc} (C_{(x)}^1 - C_{(x)}^2) & (C_{(y)}^1 - C_{(y)}^2) & (C_{(z)}^1 - C_{(z)}^2) & 0 & 0 & 0 \end{array} \right] \quad (4.3)$$

Este alinhamento produz resultados satisfatórios e é pouco sensível à presença de ruídos.

4.1.1.2 Ponta do Nariz

Este é um alinhamento heurístico e que utiliza características da face para ser feito.

Baseando-se no fato de que, normalmente, o nariz é a parte da face mais próxima do equipamento de aquisição, faz-se uma busca pelo ponto que possui o maior valor na coordenada Z . No entanto, com a intenção de evitar que ruídos contribuam para a escolha de um ponto errado, aconselha-se a utilização de uma vizinhança.

Seja p_i o i -ésimo ponto da imagem I e w_i a janela de tamanho $n \times n$ centrada em p_i , é considerada a ponta do nariz o ponto cuja janela nele centrada apresenta os pontos de maior valor no eixo Z , como visto na Equação 4.4:

$$P = \max_{i \in I} \left\| \sum_{j \in w_i} p_{j(z)} \right\| \quad (4.4)$$

O valor de n utilizado deve ser pequeno e, atualmente é utilizada uma janela 3×3 . Este

alinhamento, normalmente, é o que produz melhores resultados.

Além do ruído, uma rotação da face no eixo Y pode comprometer este alinhamento pois, neste caso, outras partes da face (normalmente, testa ou queixo) podem ser escolhidas como o ponto mais próximo ao sensor e, portanto, a “ponta do nariz”.

Uma vez encontradas as supostas pontas do nariz das duas faces, o alinhamento seguindo esta técnica é análogo ao apresentado na Seção 4.1.1.1.

4.1.1.3 *Bounding Box*

Bounding Box corresponde ao menor cubóide que engloba, totalmente, um determinado objeto.

Este alinhamento é feito pelo centro da *bounding box* do objeto, que também é chamado de centro geométrico.

Seja p_i o i -ésimo ponto da imagem I , o centro B da *bounding box* da imagem I pode ser definido como (Equação 4.5):

$$\begin{aligned} B_{(x)} &= \frac{\min_{i \in I}(p_{i(x)}) + \max_{i \in I}(p_{i(x)})}{2} \\ B_{(y)} &= \frac{\min_{i \in I}(p_{i(y)}) + \max_{i \in I}(p_{i(y)})}{2} \\ B_{(z)} &= \frac{\min_{i \in I}(p_{i(z)}) + \max_{i \in I}(p_{i(z)})}{2} \end{aligned} \quad (4.5)$$

O alinhamento através desta técnica, uma vez encontrados os centros geométricos das duas faces, é análogo ao apresentado na Seção 4.1.1.1.

Este alinhamento é extremamente sensível a ruídos e, raramente, produz resultados superiores aos dos outros dois alinhamentos já apresentados.

4.1.2 *Simulated Annealing com o MSAC*

Neste estágio do SAM, é feito um alinhamento grosseiro utilizando o MSAC como medida de erro. As definições relacionadas ao SA são dadas a seguir.

A **solução inicial** corresponde ao vetor de transformação pertencente ao melhor dos três pré-alinhamentos apresentados na Seção 4.1.1. Como para avaliar os pré-alinhamentos foram executadas algumas iterações do SA, é considerado o melhor pré-alinhamento aquele que for-

nece o menor valor de MSAC nessas iterações.

A **função de incremento** consiste em adicionar a cada elemento do vetor de transformação um valor aleatório que varia no intervalo fechado $[-1; 1]$. A cada chamada da função de incremento todos os elementos do vetor são modificados, cada um por um valor aleatório diferente.

Para que o esforço computacional necessário para calcular a **função de custo** fosse reduzido, apenas uma amostragem dos pontos da imagem que será alinhada é utilizada no cálculo (ver avaliação parcial, na Seção 2.3.2.1). Considerando B a imagem que será alinhada com a imagem A , os passos para se calcular o valor da função de custo, são dados a seguir:

1. para cada ponto amostrado da imagem B (com as devidas transformações geométricas aplicadas):
 - (a) encontrar o ponto mais próximo (ponto correspondente) na imagem A , através da *k-d tree*;
 - (b) calcular a distância Euclidiana entre eles;
 - (c) a esta distância, aplicar o avaliador MSAC;
2. a saída da função de custo é o somatório dos valores retornados pelo MSAC dividido pelo número de pontos avaliados.

Os parâmetros relacionados ao esquema de resfriamento do SA foram definidos da seguinte maneira:

- Esquema de resfriamento: proposto por Lundy [28];
- Número de iterações em cada temperatura: 1 (conforme o esquema de resfriamento adotado);
- Temperatura inicial: 0,002 (2×10^{-3});
- Temperatura final: <indefinida>;
- Fator de resfriamento: 0,001 (1×10^{-3}).

A condição de parada, ao invés de ser determinada pela temperatura final, é atingida quando o sistema permanece congelado (ou seja, não consegue uma solução melhor que a atual) por um determinado número de iterações.

4.1.3 *Simulated Annealing com a SIM*

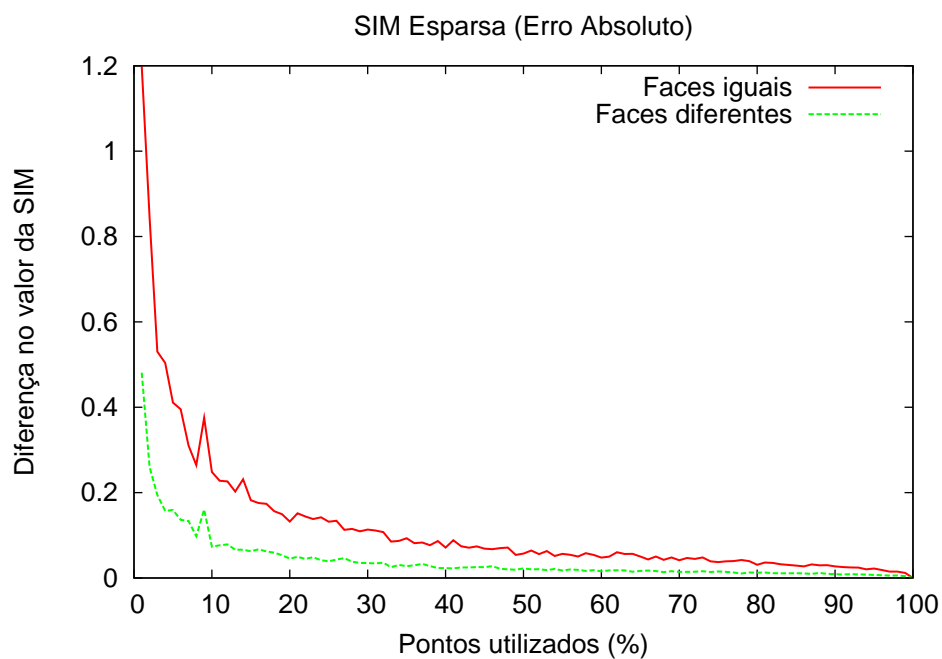
Com a finalidade de refinar o alinhamento, neste estágio do algoritmo, a SIM é utilizada como medida de erro. As definições relativas ao SA são dadas a seguir.

A **solução inicial** deste estágio, corresponde ao vetor de transformação da melhor solução encontrada no estágio anterior (ou seja, a que possui menor valor de MSAC).

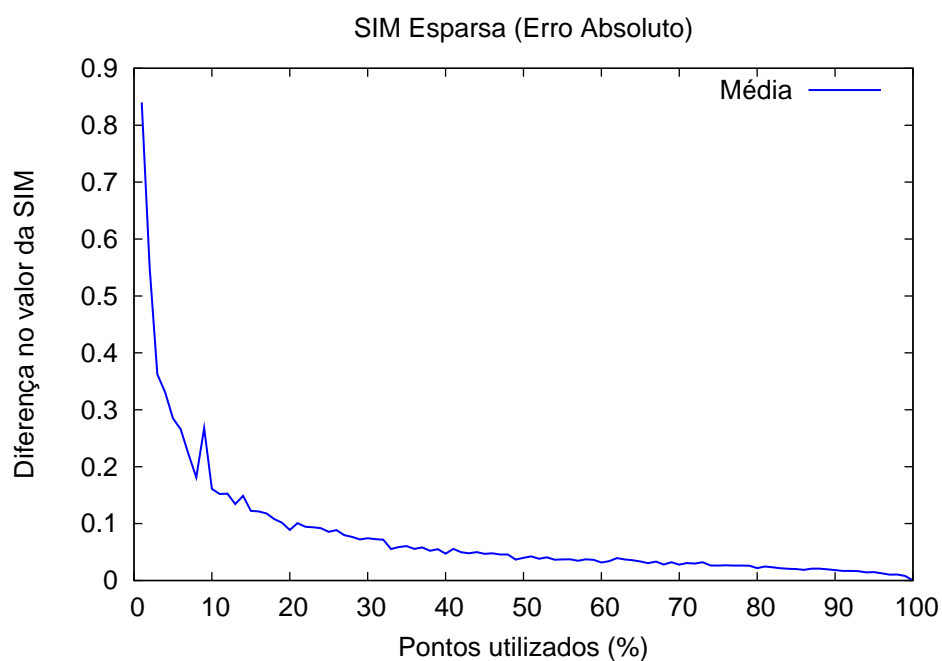
A **função de incremento** é exatamente igual à do estágio anterior, no entanto os valores aleatórios agora variam no intervalo $[-0,06667; 0,06667]$ (equivalente a $[-1/15; 1/15]$). A diminuição neste valor se deve a dois fatores: (1) o alinhamento já encontra-se razoável e, a idéia agora é refiná-lo lentamente; (2) a SIM é uma medida muito sensível e pequenas variações no alinhamento podem provocar mudanças bruscas em seu valor.

A **função de custo** corresponde à porcentagem de pontos interpenetrados (ou seja, a SIM). No entanto, com o intuito de reduzir o custo necessário para fazer este cálculo, é utilizada a SIM esparsa que, nada mais é que a SIM aplicada a uma determinada amostragem de pontos.

Para que fosse possível determinar o erro médio provocado pela SIM esparsa em relação à SIM normal (que utiliza todos os pontos) quando aplicada ao reconhecimento facial, foi feito o seguinte experimento: (a) foram feitos 100 alinhamentos entre faces de mesmos indivíduos; (b) também foram feitos 100 alinhamentos entre faces de indivíduos diferentes. Para estes 200 alinhamentos, foram calculados o erro absoluto e o erro relativo da SIM esparsa (variando de 1% a 100% de amostragem) em relação à SIM normal. As Figuras 4.1 (a) e 4.2 (a) mostram o impacto da porcentagem de pontos utilizadas nos erros absoluto e relativo, respectivamente, para o cálculo da SIM para faces iguais e diferentes, já as Figuras 4.1 (b) e 4.2 (b) mostram a média deste impacto nos erros absoluto e relativo, respectivamente.

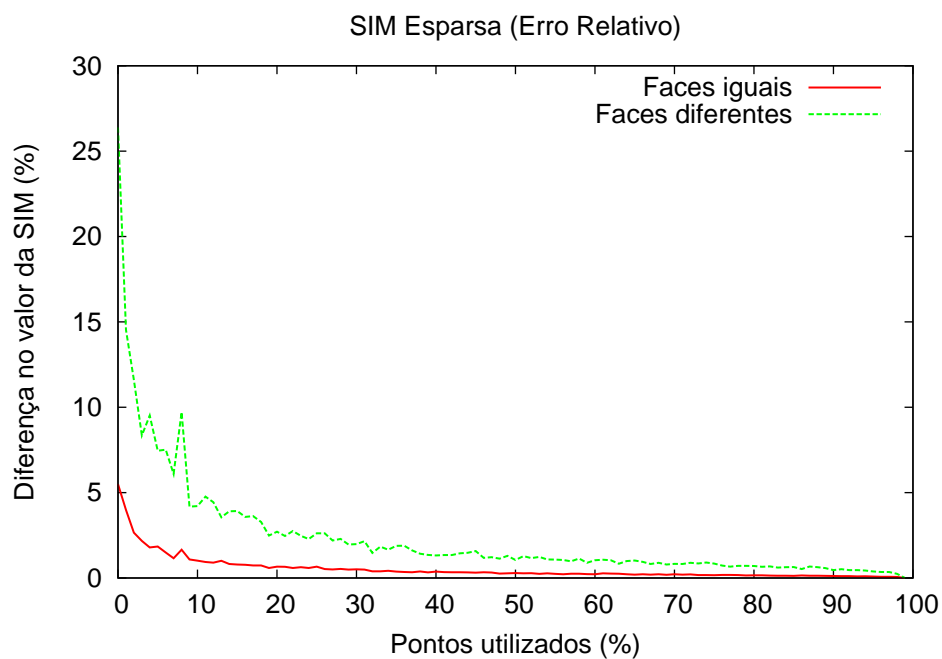


(a) Entre faces iguais e diferentes

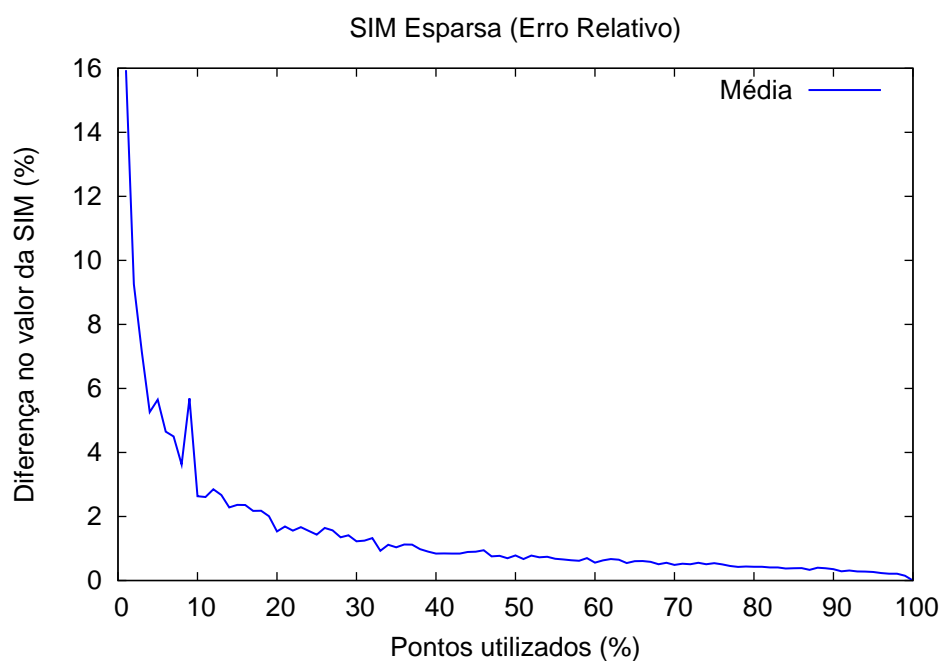


(b) Média

Figura 4.1: Erro absoluto da SIM esparsa em relação à SIM normal: (a) para mesmos indivíduos e indivíduos diferentes; (b) média.



(a) Entre faces iguais e diferentes



(b) Média

Figura 4.2: Erro relativo da SIM esparsa em relação à SIM normal: (a) para mesmos indivíduos e indivíduos diferentes; (b) média.

Os parâmetros relacionados ao esquema de resfriamento do SA foram definidos da seguinte maneira:

- Esquema de resfriamento: proposto por Lundy [28];
- Número de iterações em cada temperatura: 1 (conforme o esquema de resfriamento adotado);
- Temperatura inicial: 0,15 ($1,5 \times 10^{-1}$);
- Temperatura final: <indefinida>;
- Fator de resfriamento: 0,001 (1×10^{-3}).

A condição de parada, assim como no estágio anterior, é atingida quando o sistema permanece congelado por um determinado número de iterações. A temperatura inicial é maior devido à diferença nas ordens de grandeza do MSAC e da SIM, uma vez que, para aceitar soluções piores com uma maior diferença entre as energias, é necessária uma maior temperatura. Por exemplo, uma solução pior que a atual e que gera uma diferença de energia de 0,5 para a SIM é pouco significativa e, pode ser aceita, já pra o MSAC, com o limiar empregado no algoritmo, tal solução representaria um grande retrocesso no alinhamento e portanto, deveria ser rejeitada.

4.2 VERSÃO PRELIMINAR

O funcionamento da versão preliminar [4] do SAM pode ser visto no diagrama da Figura 4.3. Seu caminho de execução, da entrada à saída, é visto abaixo:

1. Recebe duas regiões biométricas (face, região do nariz ou dos olhos), sob a forma de imagens de profundidade, como entrada;
2. Utilizando o MSAC como medida de erro, testa a convergência de três pré-alinhamentos;
3. Executa o SA para o melhor dos três pré-alinhamentos, ainda utilizando o MSAC como medida de erro;
4. Executa o SA para o melhor resultado obtido no passo anterior, mas, desta vez, utilizando a SIM como medida de erro;
5. Devolve, como saída, o melhor valor da SIM obtido no passo anterior. Neste valor, aplica-se um limiar que decide se as regiões biométricas pertencem ou não ao mesmo indivíduo.

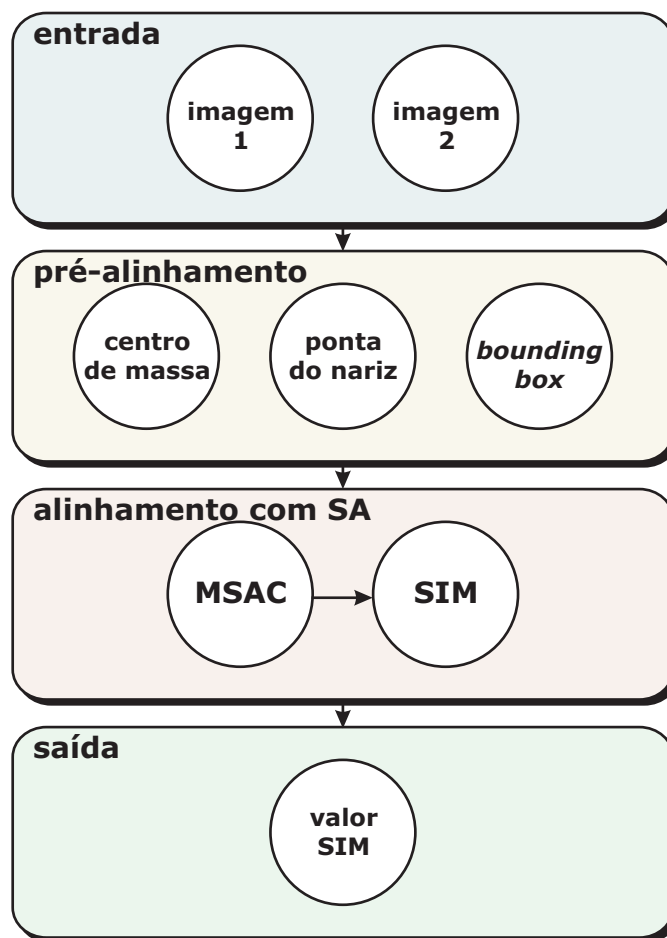


Figura 4.3: Diagrama do SAM na sua versão preliminar.

4.2.1 Resultados Experimentais

Nos experimentos realizados foram utilizadas três diferentes abordagens de segmentação, porque as imagens apresentam áreas distintas que podem interferir no alinhamento, e.g., porções de cabelo, pescoço, ombro, como mostra a Figura 4.4. Neste caso, o alinhamento foi realizado independentemente para cada uma das abordagens de segmentação e em seguida a SIM foi calculada para cada alinhamento.

4.2.1.1 Base de Dados

Os experimentos foram realizados usando 778 imagens da base de dados *Biometric Experimentation Environment (BEE) 3D Face Database*². As imagens possuem dimensão de 640×480 e foram adquiridas utilizando o sensor laser Minolta Vivid 900³.

²<http://www.bee-biometrics.org>

³<http://www.minoltausa.com/vivid>

Esta base de dados foi utilizada por se tratar de uma base pública e ser amplamente utilizada por algoritmos de reconhecimento facial 3D e, portanto, possibilita uma comparação mais rigorosa entre eles.

A base de dados possui cerca de 250 indivíduos, todos com vista frontal e a maioria com expressão facial neutra. Na Figura 4.4 são apresentadas algumas das imagens presentes na base onde pode-se observar características diferentes, como por exemplo áreas sem informação (em preto), partes do cabelo, imagens onde os indivíduos aparecem sem ombros, entre outras.

Embora a maior parte das faces da base de dados apresentem expressão facial neutra (e.g., Figura 4.4(a)), foram observadas algumas faces com expressões faciais mais significativas como indivíduos sorrindo (e.g., Figura 4.4(k) e Figura 4.4(l)) e distorções ou inconsistências nas superfícies (Figura 4.4(g) e Figura 4.4(i)).

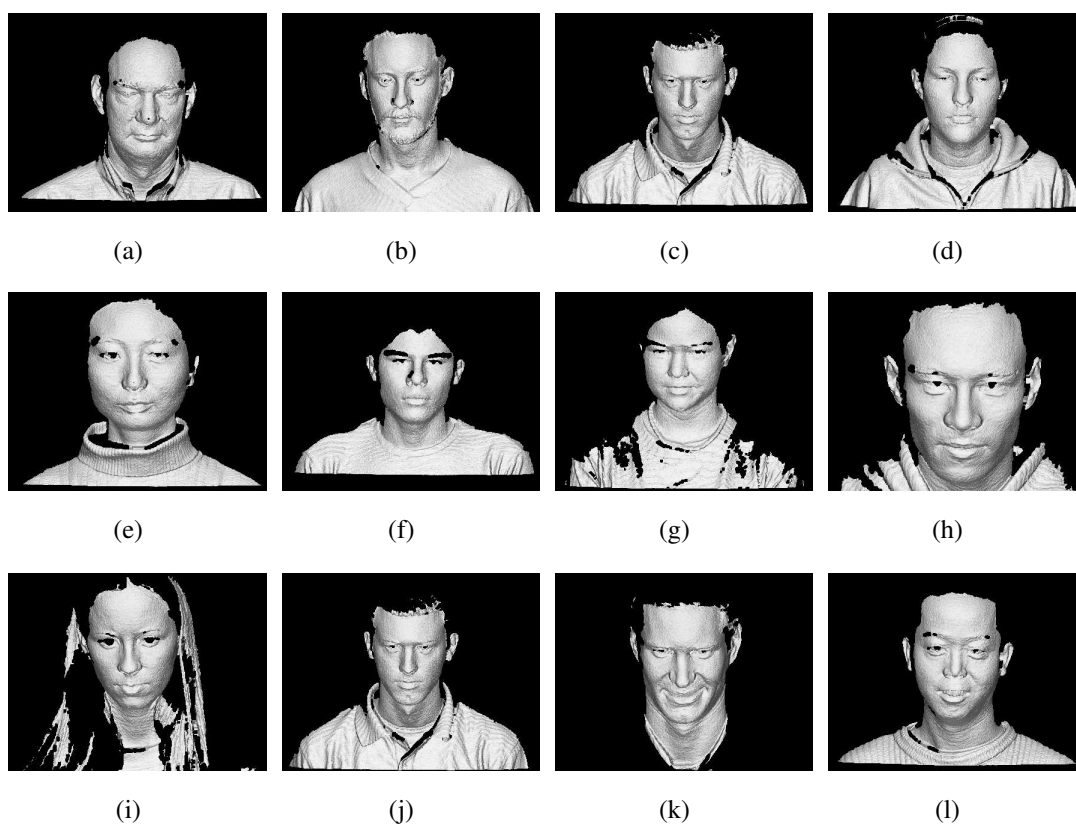
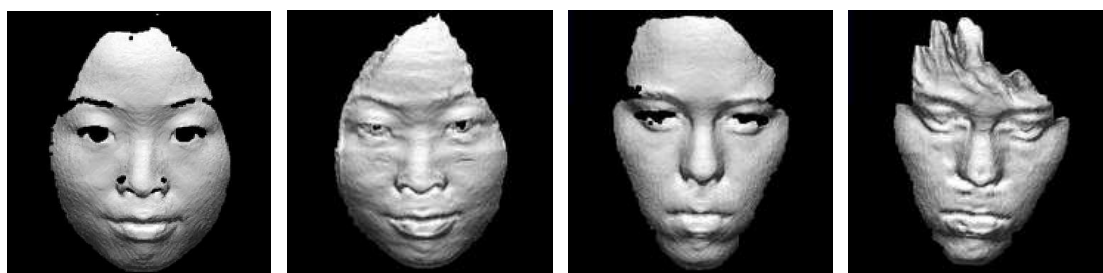


Figura 4.4: Exemplos de imagens da base de dados BEE.

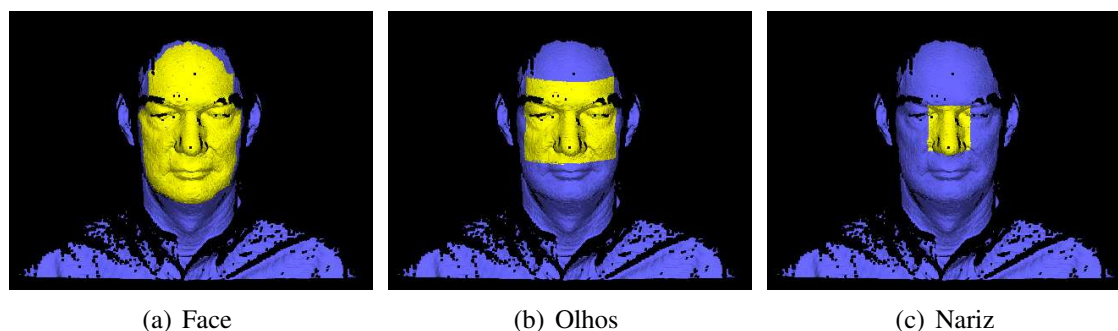
Como a quantidade dessas imagens é baixa, não foi possível realizar nenhuma análise precisa e detalhada com relação à presença de expressão facial. Além disso, algumas das imagens apresentam um alto nível de ruído, como pode ser visto na Figura 4.5. O método de segmentação utilizado, descrito em [33, 34], além de produzir as imagens de cada região, também emprega técnicas de redução de ruído sendo, estas técnicas, suficientes para o objetivo de suavização das imagens, embora não tenha resolvido alguns casos como os apresentados na Figura 4.5.



(a) Indivíduo A - Normal (b) Indivíduo A - Ruído (c) Indivíduo B - Normal (d) Indivíduo B - Ruído

Figura 4.5: Exemplos de imagens com ruído: vistas (a) e (c) são imagens normais; vistas (b) e (d) apresentam ruído. (a)-(b) e (c)-(d) pertencem aos mesmos indivíduos.

Foram utilizadas três segmentações: (a) região da face, (b) região dos olhos e nariz e (c) região do nariz (ver Figura 4.6). Para as regiões (b) e (c) foi utilizada uma abordagem similar à proposta em [14, 17, 26, 27], onde ambas as regiões selecionadas consistem em áreas retangulares (Figura 4.6(b)-4.6(c)).



(a) Face

(b) Olhos

(c) Nariz

Figura 4.6: Regiões utilizadas das imagens: (a) região da face; (b) região retangular em torno dos olhos e nariz; (c) região retangular em torno do nariz.

4.2.1.2 Resultados

Os experimentos foram realizados combinando cada imagem da base de dados com todas as outras, totalizando 302.253 combinações (300.988 entre indivíduos diferentes e 1.265 entre mesmos indivíduos) para cada uma das regiões segmentadas. Uma vez obtidos os resultados para todas as combinações, foram estabelecidos, automaticamente, limiares discriminatórios que permitissem uma taxa de falsa aceitação, FAR, de 0%, para cada região.

Os experimentos também foram realizados utilizando o ICP como método de registro. Para esta abordagem, uma vez que as imagens estavam alinhadas, era calculada a SIM para medir a similaridade das faces analisadas.

A utilização de duas abordagens visa compará-las e validar a SIM como uma medida discriminatória para o reconhecimento facial 3D.

A Tabela 4.1 apresenta a taxa de verificação e o limiar encontrado para o método baseado no ICP.

Região	Taxa de Verificação	Limiar
Face	84,66%	14,5%
Olhos	91,86%	11,5%
Nariz	93,83%	18,5%

Tabela 4.1: Taxa de verificação, usando o registro com ICP e FAR de 0%.

A área em torno dos olhos e do nariz é considerada uma região rígida [14, 26], pois sofre menor variação com a presença de expressões faciais quando comparada com outras regiões da face ou mesmo em relação a face inteira. Na Tabela 4.1 pode ser observado que a taxa de verificação com a região do nariz (93,83%) é maior do que usando a face inteira (84,66%). Analisando as combinações entre mesmos indivíduos que ficaram abaixo do limiar estabelecido, foi observado que o conjunto das imagens não reconhecidas com a região da face, não era necessariamente o mesmo conjunto das imagens não reconhecidas com a região do nariz. Por essa razão foram utilizadas regras para combinar as medidas obtidas para cada região, na tentativa de fazer com que as regiões complementem-se umas às outras.

Dentre as regras de fusão (regra da soma, do produto, da média), a regra da soma apresentou os melhores resultados [25]. A idéia básica da fusão das medidas é alinhar as duas vistas e somar o valor da SIM obtido para cada uma das três regiões segmentadas. Uma vez adquiridos os resultados é estabelecido um limiar discriminatório que permita uma FAR de 0%. Os resultados obtidos pela abordagem do ICP na combinação dos valores de SIM de cada segmentação são apresentados na Tabela 4.2.

Combinação	Taxa de Verificação
Face + Olhos	93,44%
Face + Nariz	96,78%
Olhos + Nariz	98,65%
Face + Olhos + Nariz	98,85%

Tabela 4.2: Taxa de verificação para combinação das medidas, usando o registro com ICP e FAR de 0%.

Utilizando um FAR de 0% elimina-se a possibilidade de identificar um indivíduo não autorizado em um sistema de autenticação, obtendo uma taxa de verificação de 98,85% usando a combinação das três segmentações.

A Tabela 4.3 mostra os resultados obtidos com o SA para cada uma das regiões segmentadas, e a Tabela 4.4 mostra os resultados obtidos utilizando a combinação das regiões. Também para os experimentos realizados com o SA, considerou-se uma FAR de 0%.

Região	Taxa de Verificação	Limiar
Face	97,23%	16,75%
Olhos	94,31%	14,75%
Nariz	93,75%	22,5%

Tabela 4.3: Taxa de verificação usando o registro com SA e FAR de 0%.

Combinação	Taxa de Verificação
Face + Olhos	97,79%
Face + Nariz	98,26%
Olhos + Nariz	96,28%
Face + Olhos + Nariz	99,05%

Tabela 4.4: Taxa de verificação para combinação das medidas, usando o registro com SA e FAR de 0%.

Os resultados obtidos com o SA foram mais discriminatórios do que os resultados obtidos com o ICP. Por ser guiado pela SIM em suas etapas finais, o SA tende a gerar valores mais altos da SIM tanto para combinações entre mesmos indivíduos quanto para indivíduos distintos. Por esse motivo o limiar que permite FAR de 0% aumentou em relação ao do ICP. Foi observado que para indivíduos diferentes a taxa de aumento da SIM é muito menor do que para mesmos indivíduos. Sendo alinhada apenas a região da face, com o SA obteve-se uma taxa de verificação de 97,31% contra os 84,66% obtidos com o ICP.

Utilizando a combinação das três abordagens de segmentação, com o SA obtivemos uma taxa de verificação de 99,05%. Dos casos que deram errado e não foram reconhecidos, observamos que pelo menos uma das imagens de cada combinação apresentava ruído (e.g., Figura 4.5(b) e Figura 4.5(d)). Os experimentos mostraram resultados promissores e confirmam que a SIM pode ser muito eficiente para o problema do reconhecimento de faces em imagens 3D. De fato, de acordo com [30] o desempenho esperado para o *Face Recognition Grand Challenge* (FRGC) 2006 é uma taxa de verificação de 98% com um FAR de 0,1%.

4.2.1.3 Exemplos de Resultados

Nesta seção serão apresentados exemplos de resultados obtidos com a utilização do SAM. Os exemplos estão divididos em duas classes: (a) pré-alinhamentos e resultado final; (b) pré-

alinhamentos, resultado final e gráficos de convergência da SIM e do MSAC.

As imagens binárias devem ser interpretadas da seguinte maneira: os pontos brancos representam os pontos onde há interpenetração e os negros, onde não há.

Para obter os resultados desta seção, o SAM foi configurado da seguinte maneira:

- Encerrar o estágio do MSAC após o sistema permanecer congelado por 350 iterações e, o estágio da SIM, após 40 iterações. Estes valores foram obtidos empiricamente;
- Utilizar amostragem de pontos de 3,24% para o estágio do MSAC. Como neste estágio o objetivo é conseguir um alinhamento grosseiro, uma pequena amostragem de pontos, como a empregada, é suficiente para atingi-lo. Para o estágio da SIM, a amostragem utilizada foi de 10,89%, este valor foi escolhido após estudo da SIM esparsa, Seção 4.1.3, onde foi demonstrado que o valor de amostragem utilizado gera uma diferença de apenas cerca de 2% em relação ao valor da SIM para todos os pontos;
- Avaliar cada pré-alinhamento por 50 iterações do SA, utilizando o MSAC.

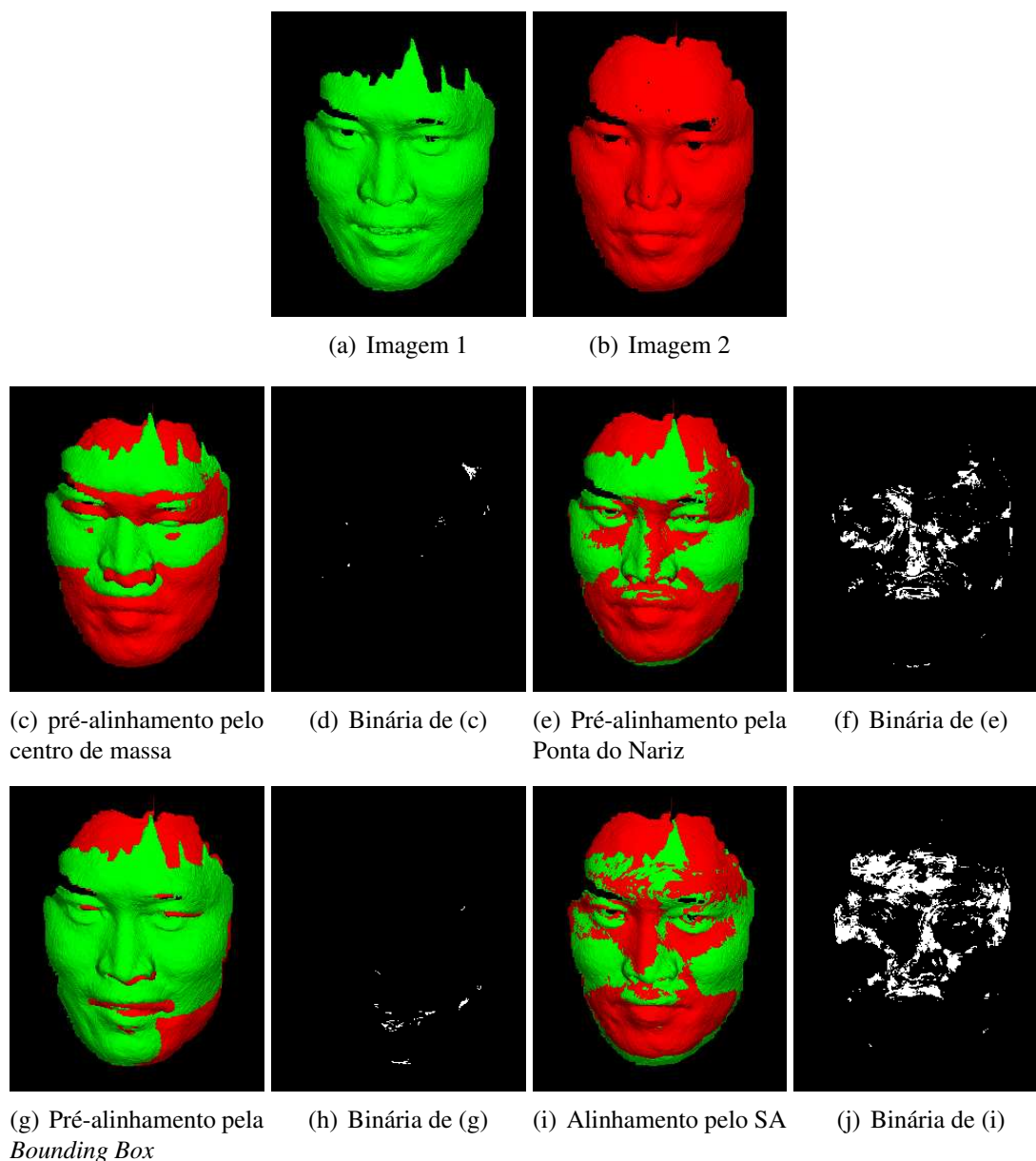


Figura 4.7: Exemplo de registro e pré-alinhamento para indivíduos iguais (com expressão): (a) imagem 1; (b) imagem 2; (c) $MSE = 24,887455 / SIM = 0,311710$ (111 pontos); (d) binária de (c); (e) $MSE = 1,570920 / SIM = 8,385285$ (2.986 pontos); (f) binária de (e); (g) $MSE = 35,390126 / SIM = 0,508284$ (181 pontos); (h) binária de (g); (i) $MSE = 0,890670 / SIM = 17,217074$ (6.131 pontos); (j) binária de (i).

A Figura 4.7 mostra um exemplo de alinhamento entre faces de um mesmo indivíduo mas, com diferentes expressões. Pode-se perceber na imagem binária (j) que a interpenetração se concentrou em áreas que não foram afetadas pelas diferentes expressões faciais.

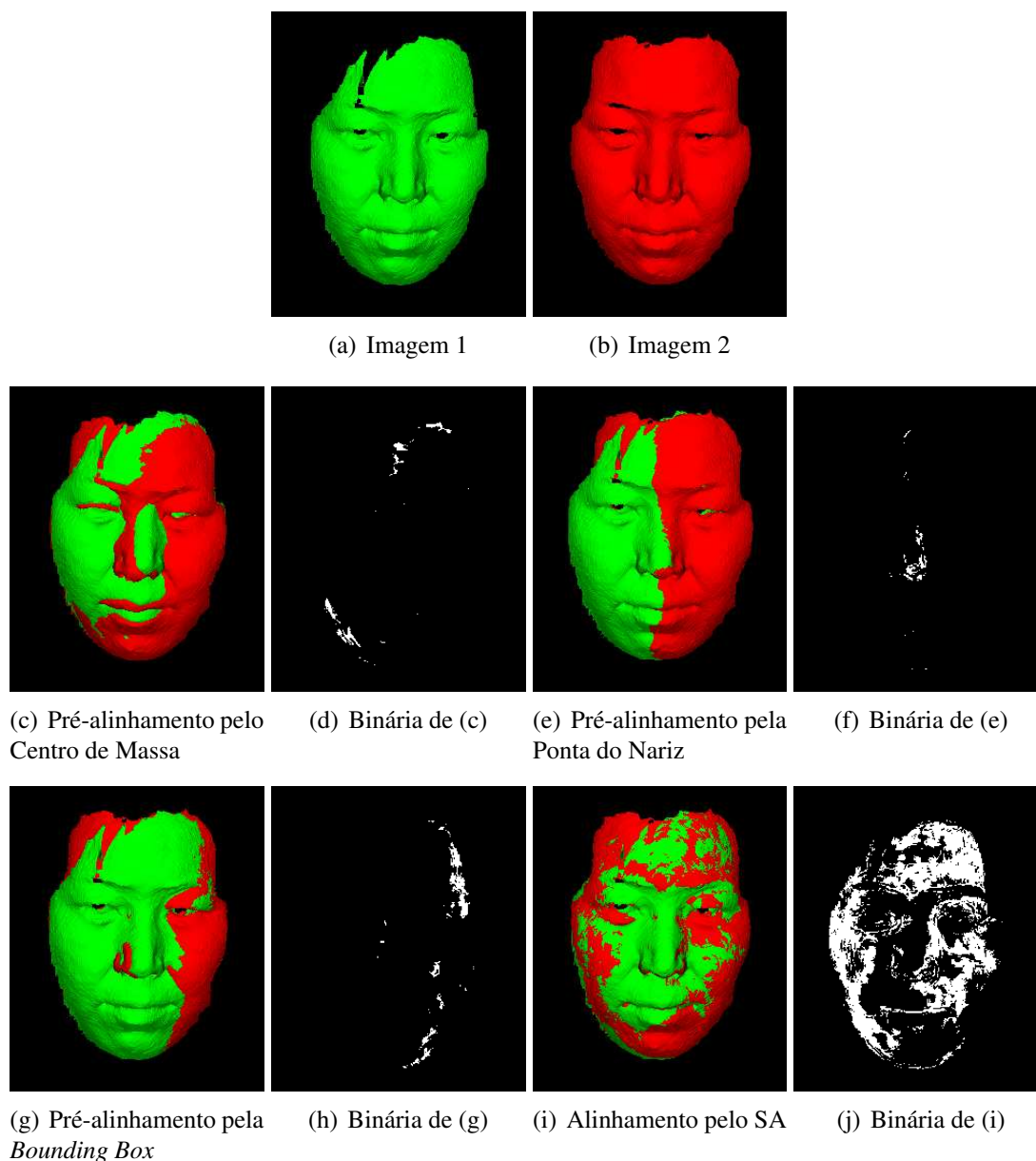


Figura 4.8: Exemplo de registro e pré-alinhamento para indivíduos iguais (sem expressão): (a) imagem 1; (b) imagem 2; (c) $MSE = 8,910562$ / $SIM = 1,288401$ (437 pontos); (d) binária de (c); (e) $MSE = 23,350791$ / $SIM = 0,742968$ (252 pontos); (f) binária de (e); (g) $MSE = 9,328597$ / $SIM = 2,461820$ (835 pontos); (h) binária de (g); (i) $MSE = 0,301815$ / $SIM = 33,754938$ (11.449 pontos); (j) binária de (i).

A Figura 4.8 mostra um alinhamento entre faces, ambas com expressão neutra, de um mesmo indivíduo. Neste exemplo, o pré-alinhamento pela *bounding box* foi o que apresentou melhor valor SIM e o pela ponta do nariz o pior. O alinhamento final pode ser considerado bom.

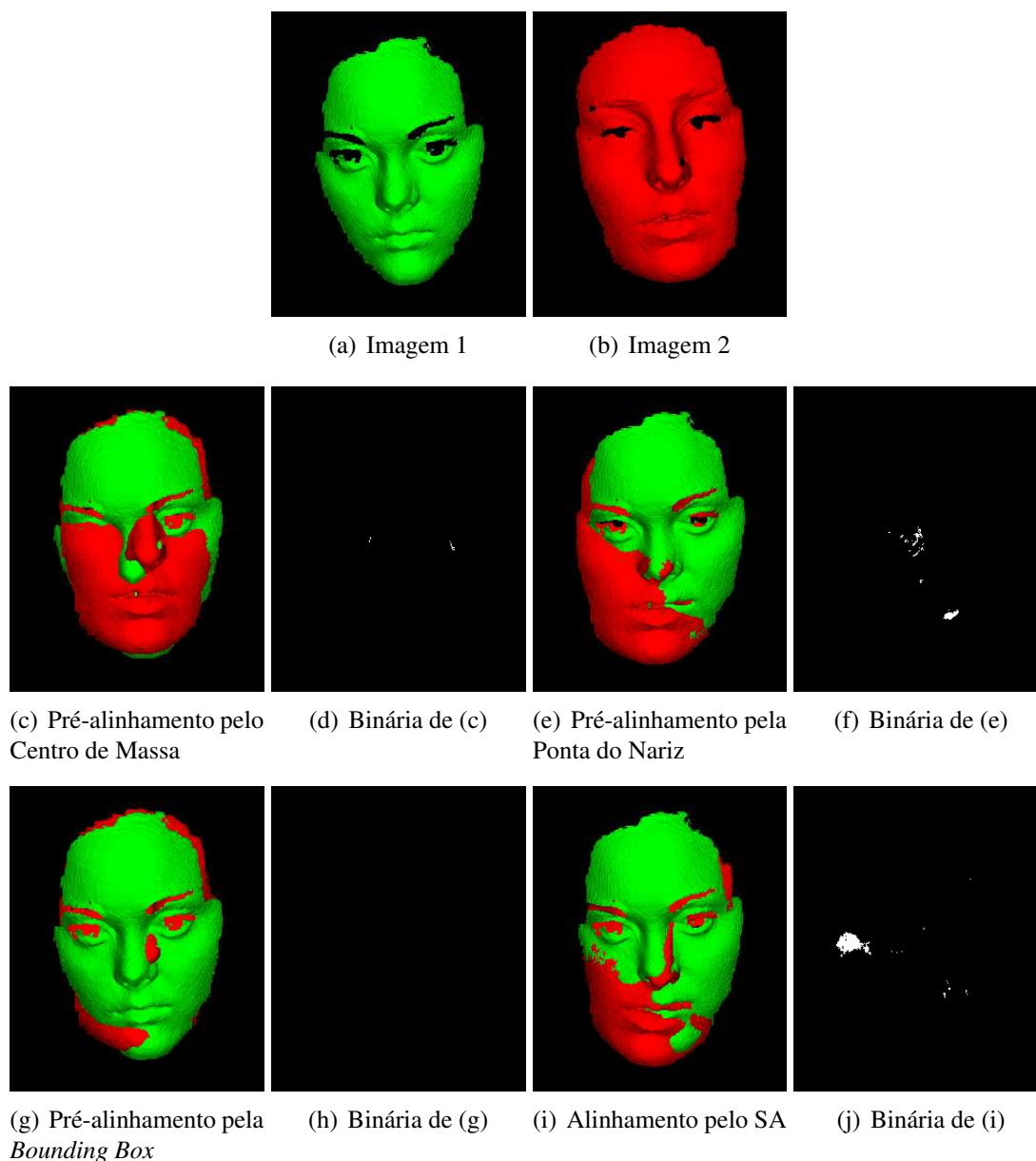


Figura 4.9: Exemplo de registro e pré-alinhamento para indivíduos diferentes: (a) imagem 1; (b) imagem 2; (c) $MSE = 157,944532 / SIM = 0,055173$ (15 pontos); (d) binária de (c); (e) $MSE = 97,789687 / SIM = 0,592195$ (161 pontos); (f) binária de (e); (g) $MSE = 190,198022 / SIM = 0,000000$ (0 ponto); (h) binária de (g); (i) $MSE = 46,830986 / SIM = 1,449222$ (394 pontos); (j) binária de (i).

A Figura 4.9 mostra um alinhamento entre faces de diferentes indivíduos. Como pode ser observado, o alinhamento final possui pouca interpenetração entre as duas faces.

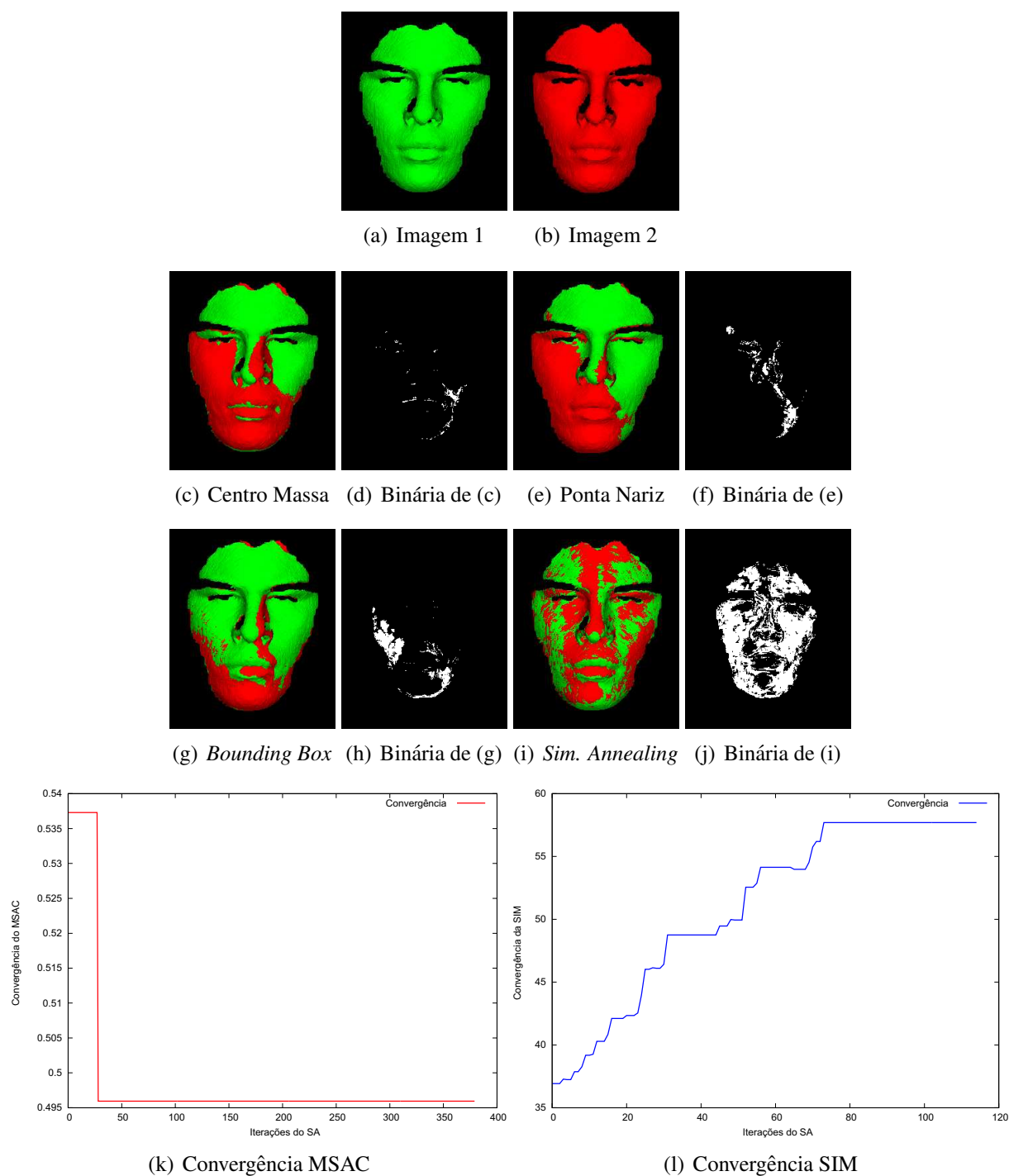


Figura 4.10: Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos iguais: (a) imagem 1; (b) imagem 2; (c) MSE = 3,069447 / SIM = 1,898326 (441 pontos); (d) binária de (c); (e) MSE = 5,098784 / SIM = 5,574448 (1.295 pontos); (f) binária de (e); (g) MSE = 5,124627 / SIM = 9,194611 (2.136 pontos); (h) binária de (g); (i) MSE = 0,230891 / SIM = 57,104731 (13.266 pontos); (j) binária de (i); (k) convergência do MSAC; (l) convergência da SIM.

A Figura 4.10 mostra um ótimo alinhamento entre faces pertencentes a um mesmo indivíduo. Como pode ser observado, as duas imagens são extremamente parecidas e, a ausência de expressão facial faz com que as regiões interpenetradas estejam uniformemente distribuídas.

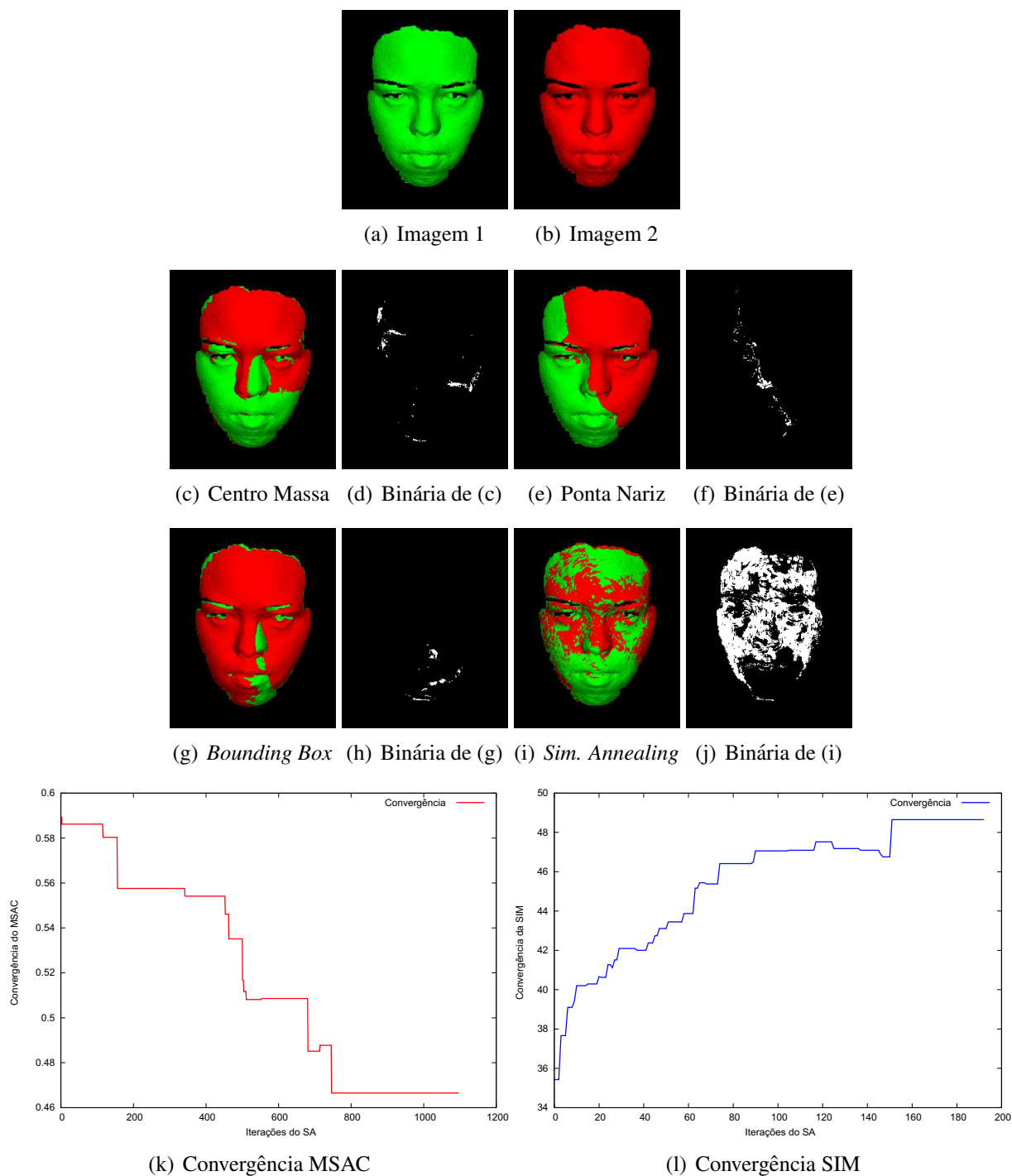


Figura 4.11: Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos iguais: (a) imagem 1; (b) imagem 2; (c) $MSE = 3,758885$ / $SIM = 1,707008$ (512 pontos); (d) binária de (c); (e) $MSE = 18,788218$ / $SIM = 1,523638$ (457 pontos); (f) binária de (e); (g) $MSE = 19,523556$ / $SIM = 0,976862$ (293 pontos); (h) binária de (g); (i) $MSE = 0,248041$ / $SIM = 48,232980$ (14.467 pontos); (j) binária de (i); (k) convergência do MSAC; (l) convergência da SIM.

A Figura 4.11 também mostra um ótimo alinhamento entre faces de um mesmo indivíduo. Como se pode observar, os pré-alinhamentos não produziram bons valores de SIM mas, com a utilização do SA, foi possível aumentar consideravelmente o valor da SIM no alinhamento final

em relação aos pré-alinhamentos (cerca de 28 vezes).

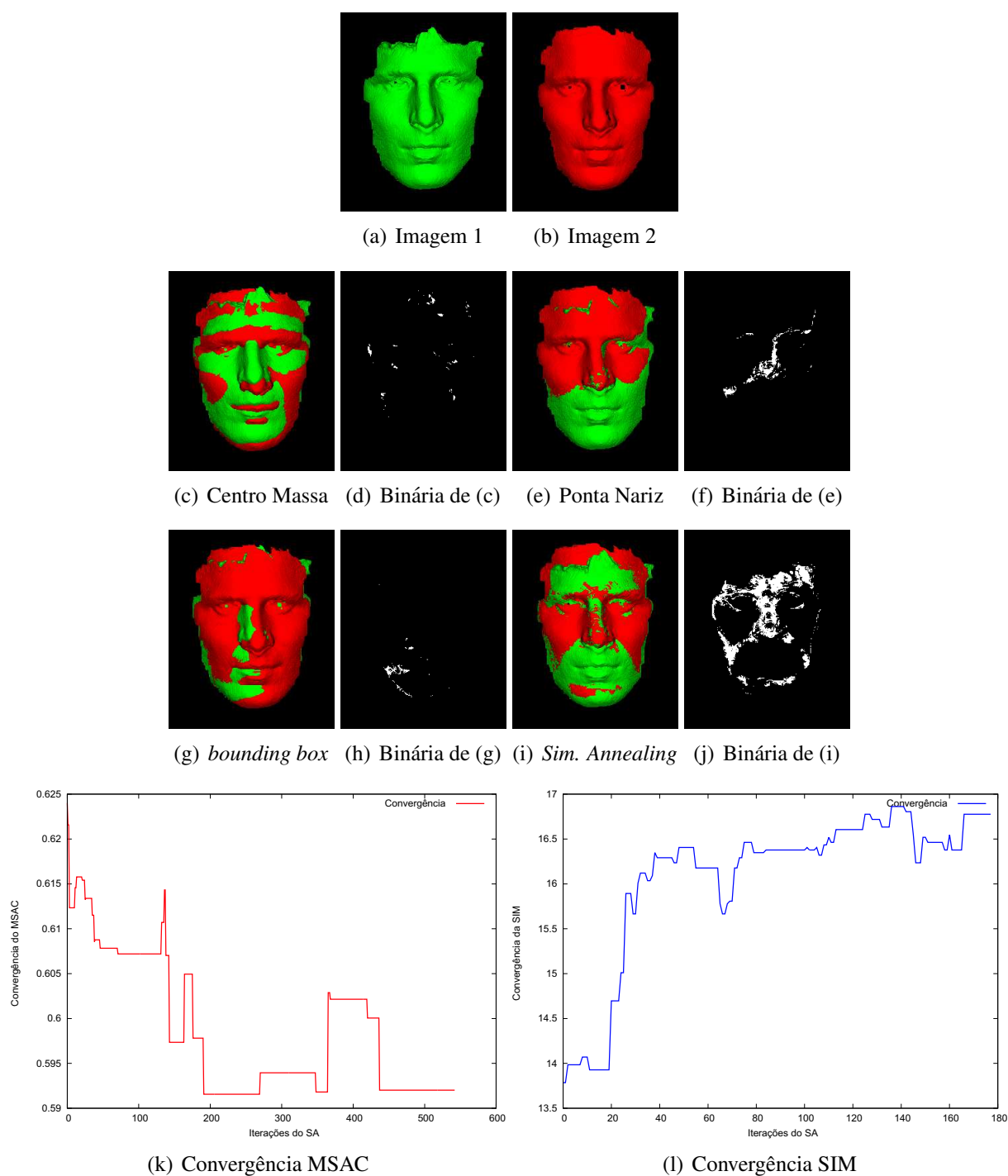


Figura 4.12: Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos iguais: (a) imagem 1; (b) imagem 2; (c) MSE = 5,057049 / SIM = 0,888127 (285 pontos); (d) binária de (c); (e) MSE = 5,740972 / SIM = 2,963540 (951 pontos); (f) binária de (e); (g) MSE = 14,643231 / SIM = 0,875662 (281 pontos); (h) binária de (g); (i) MSE = 0,820948 / SIM = 16,631349 (5.337 pontos); (j) binária de (i); (k) convergência do MSAC; (l) convergência da SIM.

A Figura 4.12 mostra um alinhamento razoável entre faces pertencentes a um mesmo indivíduo. É possível observar que as faces possuem recortes diferentes. Estes recortes diferentes

tendem a produzir pré-alinhamentos ruins para o centro de massa.

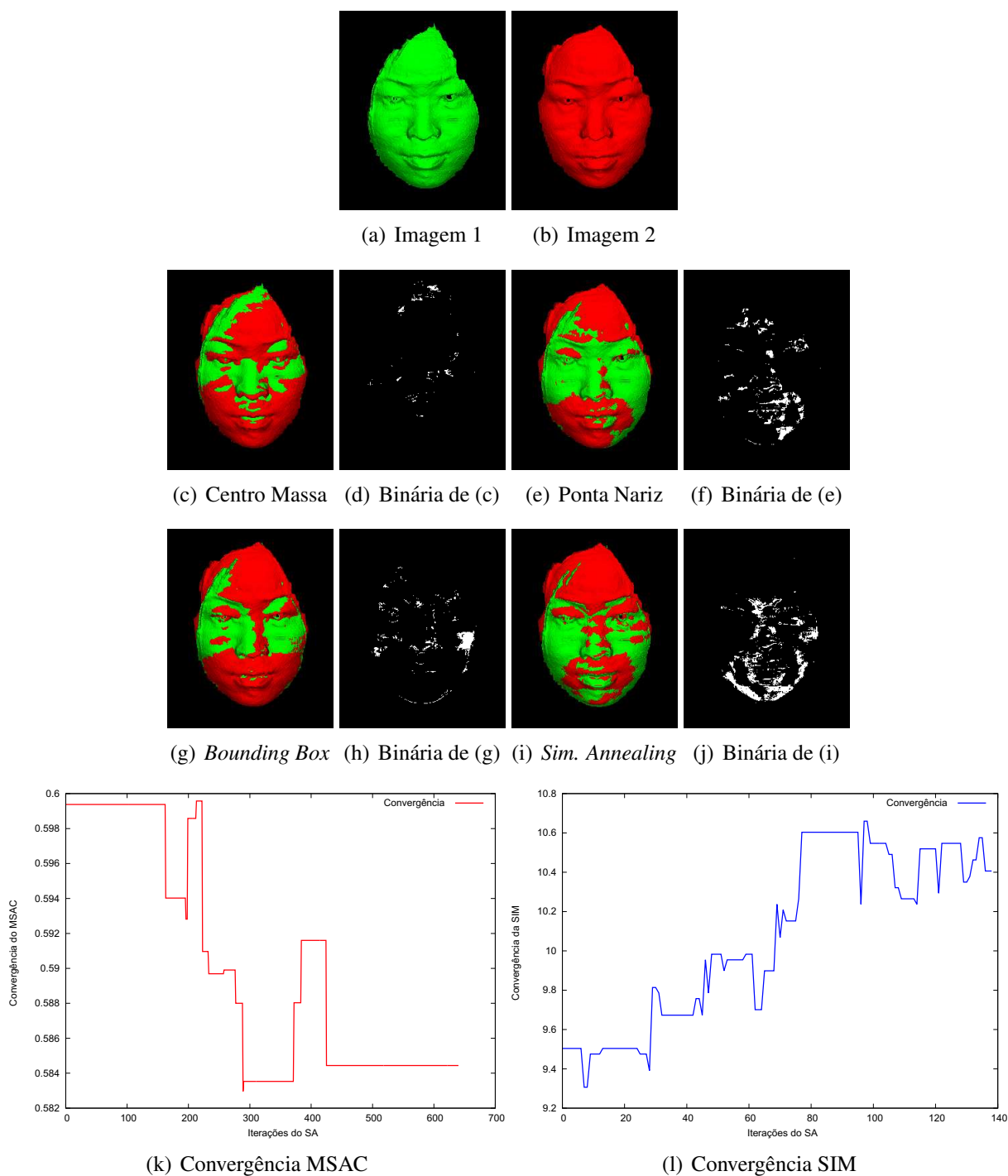


Figura 4.13: Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos iguais: (a) imagem 1; (b) imagem 2; (c) MSE = 6,482668 / SIM = 1,157045 (377 pontos); (d) binária de (c); (e) MSE = 1,464354 / SIM = 5,487524 (1.788 pontos); (f) binária de (e); (g) MSE = 2,034887 / SIM = 3,676764 (1.198 pontos); (h) binária de (g); (i) MSE = 2,378580 / SIM = 10,330540 (3.366 pontos); (j) binária de (i); (k) convergência do MSAC; (l) convergência da SIM.

A Figura 4.13 mostra um alinhamento ruim (apenas 10,0% de SIM no alinhamento final) para faces de um mesmo indivíduo. O motivo para um alinhamento tão ruim, é presença de

ruído nas duas imagens.

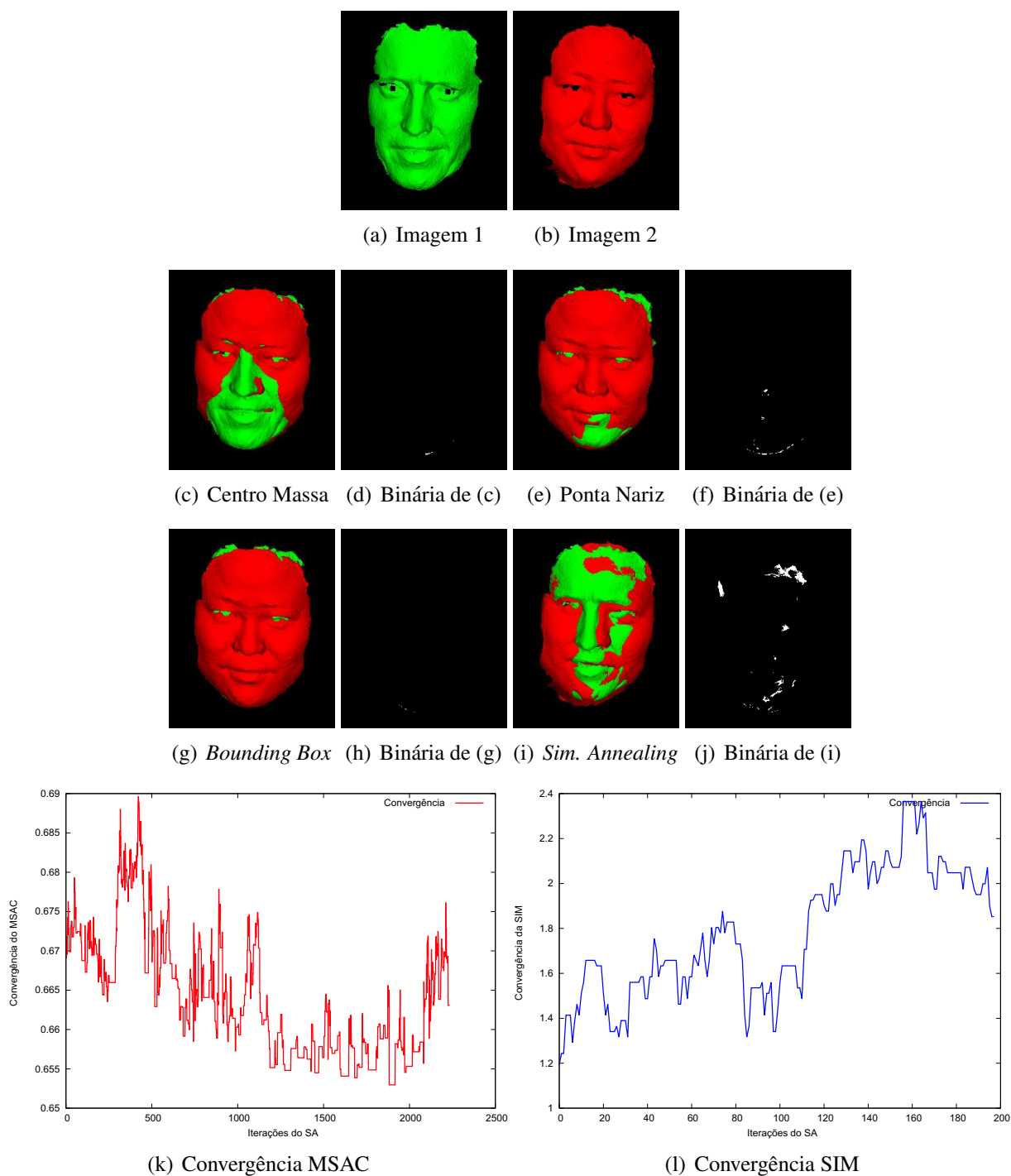


Figura 4.14: Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos diferentes: (a) imagem 1; (b) imagem 2; (c) MSE = 63,023682 / SIM = 0,042418 (16 pontos); (d) binária de (c); (e) MSE = 176,689948 / SIM = 0,243902 (92 pontos); (f) binária de (e); (g) MSE = 623,861727 / SIM = 0,015907 (6 pontos); (h) binária de (g); (i) MSE = 15,943950 / SIM = 2,218982 (837 pontos); (j) binária de (i); (k) convergência do MSAC; (l) convergência da SIM.

A Figura 4.14 mostra um alinhamento entre faces de diferentes indivíduos. Como é possível observar, o alinhamento final possui um baixo valor de SIM. Um outro detalhe interessante, é

a convergência do MSAC e da SIM, que se comportam de maneira diferente (se comparadas às convergências de alinhamentos entre faces de um mesmo indivíduo).

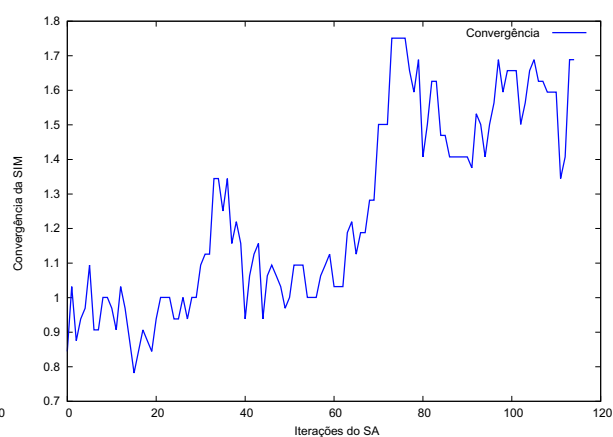
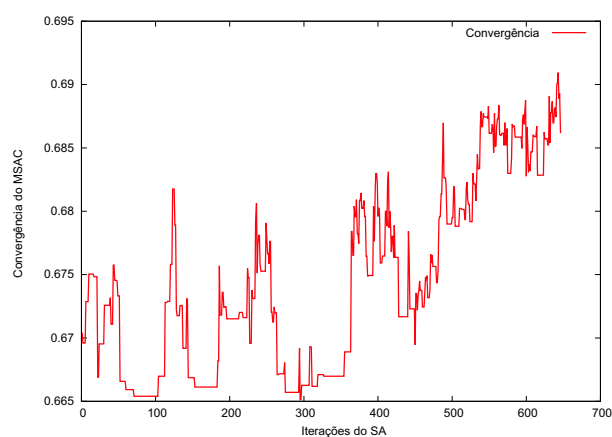
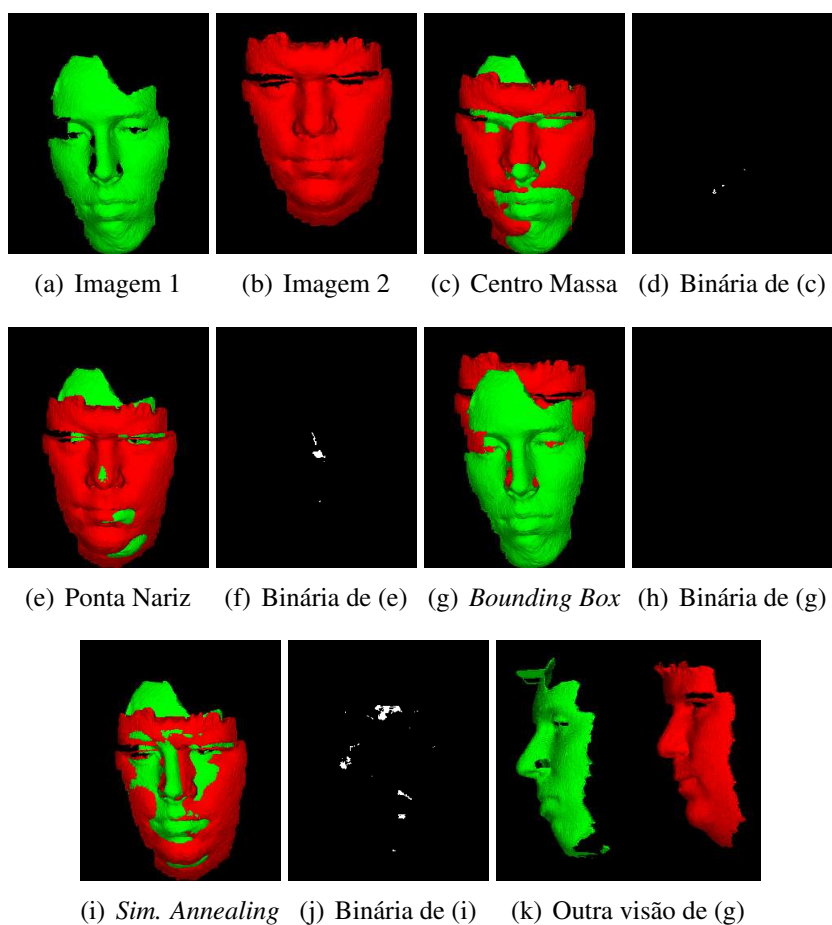


Figura 4.15: Exemplo de registro, pré-alinhamento e convergência do MSAC e da SIM para indivíduos diferentes: (a) imagem 1; (b) imagem 2; (c) $MSE = 18,833433$ / $SIM = 0,071628$ (21 pontos); (d) binária de (c); (e) $MSE = 43,658038$ / $SIM = 0,416127$ (122 pontos); (f) binária de (e); (g) $MSE = <indefinido>$ / $SIM = 0,000000$ (0 ponto); (h) binária de (g); (i) $MSE = 5,036215$ / $SIM = 1,633809$ (479 pontos); (j) binária de (i); (k) outra visão de (g); (l) convergência do MSAC; (m) convergência da SIM.

A Figura 4.15 mostra um alinhamento entre faces de indivíduos diferentes. Também é possível observar um comportamento de convergência bem diferente do usual. Observando as Figuras (g) (k), é possível observar o efeito que a presença de ruído tem sobre o pré-alinhamento pela *bounding box*.

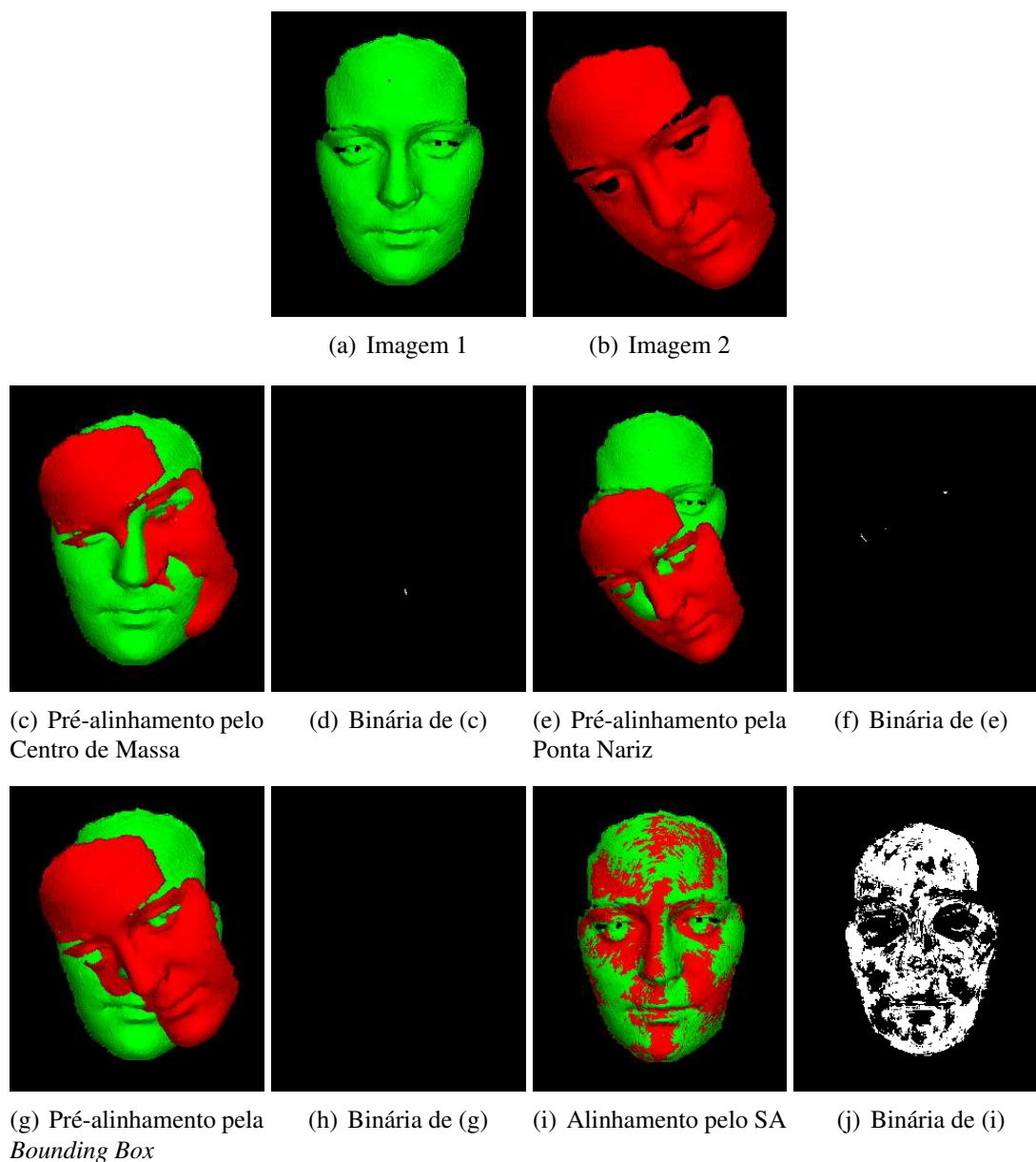


Figura 4.16: Exemplo de registro, pré-alinhamento para indivíduos iguais e com faces muito desalinhadas: (a) imagem 1; (b) imagem 2; (c) $MSE = 123,260922 / SIM = 0,026320$ (7 pontos); (d) binária de (c); (e) $MSE = 227,739512 / SIM = 0,056399$ (15 pontos); (f) binária de (e); (g) $MSE = 788,741215 / SIM = 0,000000$ (0 ponto); (h) binária de (g); (i) $MSE = 0,166690 / SIM = 56,241540$ (14.958 pontos); (j) binária de (i).

A Figura 4.16 mostra um exemplo de alinhamento entre faces pertencentes à mesma pessoa mas que estão totalmente desalinhadas. Este desalinhamento foi feito intencionalmente aplicando as seguintes transformações geométricas: 20, 37 e 11 unidades de deslocamento nos

eixos X , Y e Z , respectivamente; e, 14° , 13° e 30° de rotação nos eixos X , Y e Z , respectivamente. Como pode ser observado, todos os pré-alinhamentos estão deficientes, no entanto, o SAM foi capaz de atingir um registro preciso. A Figura 4.17 mostra as imagens de entrada do algoritmo sob diferentes ângulos e , sem qualquer pré-alinhamento; mostra também os gráficos de convergência do MSAC e da SIM. Pode-se perceber que nas primeiras 850 iterações do MSAC, o resultado era precário pois o SA estava longe do mínimo global. No entanto, o SA encontra o “caminho certo” e o segue até atingir uma solução muito próxima do mínimo global e , como resultado, foi produzido um ótimo alinhamento.

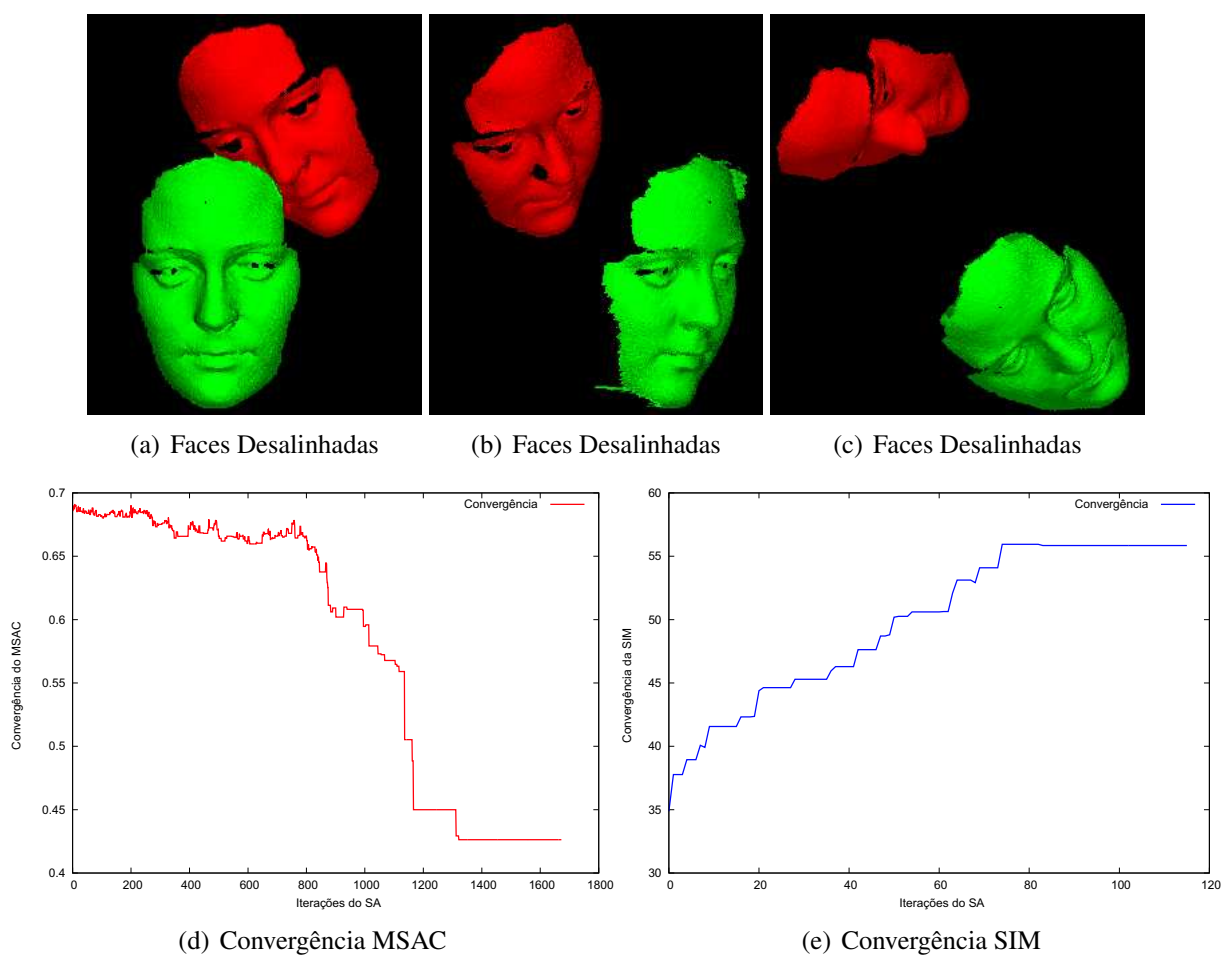


Figura 4.17: Faces desalinhadas da Figura 4.16 e gráficos de convergência do MSAC e da SIM: (a), (b) e (c) imagens desalinhadas; (d) convergência do MSAC; (e) convergência da SIM.

4.3 VERSÃO ATUAL

A versão atual do SAM segue, basicamente, o mesmo funcionamento de sua versão anterior. As mudanças feitas têm basicamente dois propósitos: (1) melhorar os resultados do algoritmo; (2) diminuir seu tempo de execução.

Para atingir o propósito de melhoria nos resultados do algoritmo, foram utilizadas três regiões do nariz, conforme proposto em [17], ao invés de uma única região, como na versão preliminar. As regiões do nariz podem ser vistas na Figura 3.1 e são:

- nariz propriamente dito;
- área circular ao redor do nariz;
- área elíptica ao redor do nariz.

A versão preliminar do algoritmo, para obter sua melhor taxa de verificação, que é de 98,85%, precisa combinar os resultados obtidos com o registro da região do nariz, dos olhos e da face como um todo. No entanto, executar o SAM para todas estas regiões é uma atividade muito custosa e que, por consequência, consome muito tempo. Na Figura 4.18 é possível verificar o tempo médio de execução⁴, em segundos, para cada uma das regiões da face.

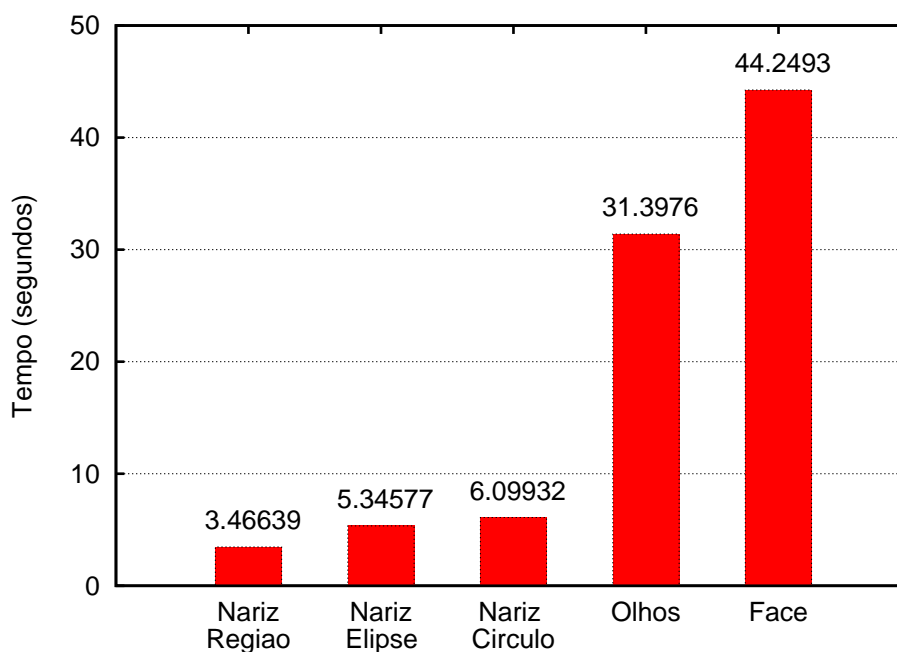


Figura 4.18: Tempo de execução do SAM para diferentes regiões da face.

⁴Valores obtidos em um computador AMD Sempron™ 3100+ com 1GB de memória RAM.

Na versão preliminar do SAM, a região do nariz que foi utilizada difere das apresentadas na Figura 4.18, no entanto, seu tempo execução assemelha-se ao da região circular ao redor do nariz. Utilizando estes valores de tempo de execução, o SAM em sua versão preliminar levaria cerca de 81,72 segundos por combinação para atingir sua taxa de verificação máxima.

A utilização de três regiões do nariz na versão atual do SAM teria como efeito imediato um aumento no tempo execução, se comparada à versão anterior. Entretanto, foi adotado um esquema de avaliação em cascata que reduz sensivelmente o tempo médio para verificação de um indivíduo além de melhorar a taxa de verificação do algoritmo.

Considerando que, na versão preliminar, utilizando apenas uma das regiões biométricas é possível obter resultados razoáveis, a nova versão do SA verifica se o resultado obtido com uma dada região biométrica é suficiente para determinar se as imagens de entrada pertencem ao mesmo indivíduo ou não. Caso não seja possível determinar com exatidão se se tratam de imagens do mesmo indivíduo, utiliza-se outra região biométrica. A ordem em que as regiões biométricas devem ser utilizadas é da menor para a maior, já que as menores regiões, por possuírem menos pontos, são executadas mais rapidamente. Desta forma, garante-se que o SAM será executado no menor tempo possível. O fluxo de execução pode ser acompanhado pelo diagrama da Figura 4.19 e pela descrição abaixo:

1. Seleciona-se a segmentação que compreende a região do nariz;
2. Utilizando a região biométrica selecionada e o MSAC como medida de erro, é feita a análise de três pré-alinhamentos;
3. A partir do pré-alinhamento que obteve o melhor resultado, executa-se o SA, ainda utilizando o MSAC como medida de erro;
4. Executa o SA para o melhor resultado obtido no passo anterior, mas, desta vez, utilizando a SIM como medida de erro;
5. Utilizando o melhor valor de SIM obtido no passo anterior, verifica-se se este é suficientemente alto (acima do limiar de reconhecimento) e, portanto, confirmando que as duas regiões biométricas pertencem ao mesmo indivíduo ou que é muito baixo (abaixo do limiar de rejeição) e, portanto, rotulando as duas regiões biométricas como pertencentes a indivíduos diferentes. Caso alguma dessas duas condições seja satisfeita, exhibe-se o resultado e o algoritmo é encerrado. Se nenhuma das condições é atendida (ou seja, o valor da SIM encontra-se entre os dois limiares), utiliza-se uma outra região biométrica como entrada do algoritmo de reconhecimento;
6. Seleciona-se a área elíptica que engloba o nariz como região biométrica e executam-se os passos de 2 a 5;

7. Seleciona-se a área circular ao redor do nariz como região biométrica e executam-se os passos de 2 a 5;
8. Seleciona-se a região dos olhos como região biométrica e executam-se os passos de 2 a 5;
9. Seleciona-se a face como região biométrica e executam-se os passos de 2 a 5;
10. Por fim, os resultados obtidos para todas as regiões biométricas são combinados em um único valor. E, a este valor é aplicado um limiar que determina se as imagens pertencem ou não ao mesmo indivíduo.

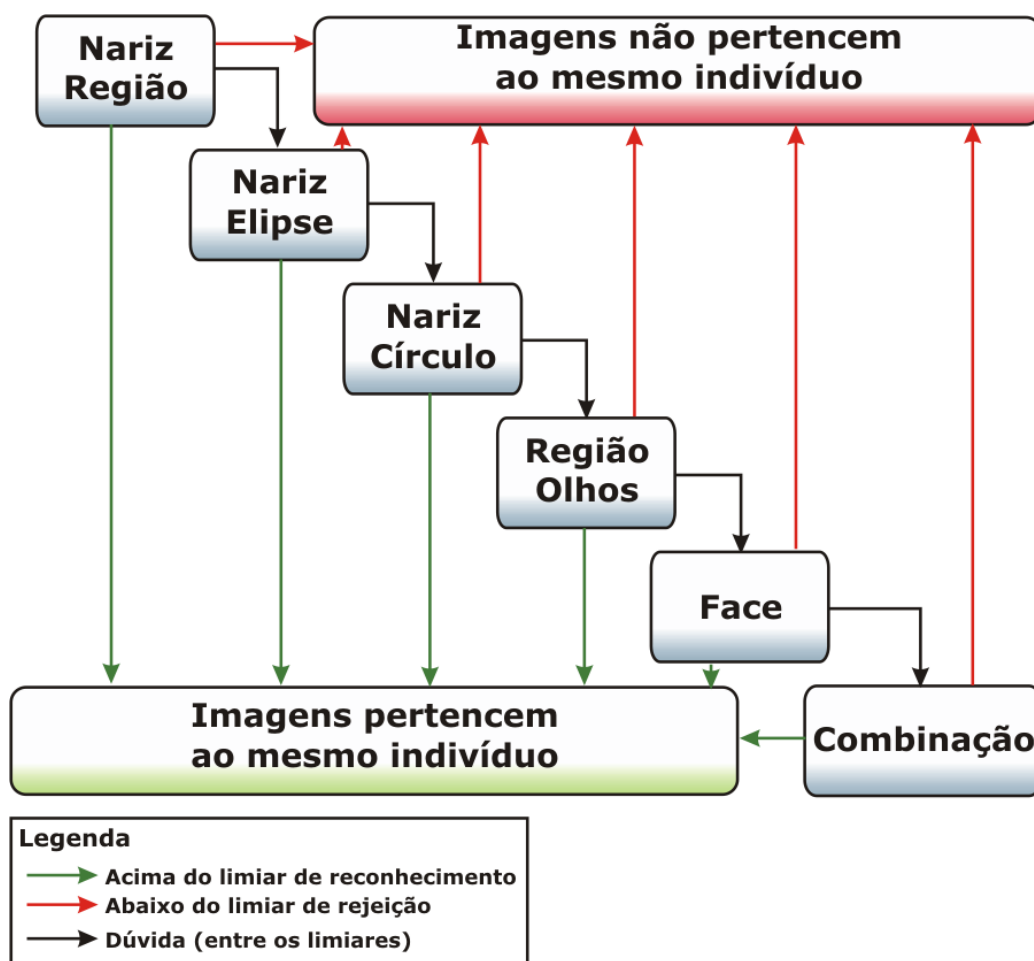


Figura 4.19: Diagrama do SAM na versão atual.

No diagrama da Figura 4.19, cada retângulo rotulado com o nome de uma região biométrica corresponde, praticamente, à execução da versão preliminar do sistema (ver Figura 4.3) para a região biométrica em questão. Os limiares utilizados foram definidos através da análise de todas as combinações e foram definidos de forma que: (a) nenhuma combinação entre imagens de indivíduos diferentes fique acima do limiar de aceitação; (b) nenhuma combinação entre imagens de um mesmo indivíduo fique abaixo do limiar de aceitação.

Esta modificação, além de melhorar o tempo de execução do algoritmo, também garante um resultado melhor ou, na pior das hipóteses, igual ao da versão preliminar. Isto se deve ao fato de que, em alguns casos, conclusões diferentes são obtidos acerca de dois alinhamentos de diferentes regiões biométricas. Por exemplo, supondo que seja feita a comparação, nas duas versões do sistema, de imagens pertencentes ao mesmo indivíduo:

- **Versão preliminar:**

- *nariz*: obtém um resultado de SIM acima do limiar e, portanto, se fosse analisado individualmente, chegar-se-ia à conclusão de que as imagens pertencem ao mesmo indivíduo;
- *região dos olhos*: obtém um resultado de SIM ligeiramente abaixo do limiar e, portanto, se fosse analisado individualmente, chegar-se-ia à conclusão de que as imagens não pertencem ao mesmo indivíduo;
- *face*: obtém um resultado de SIM abaixo do limiar e, portanto, se fosse analisado individualmente, chegaria-se à conclusão de que as imagens não pertencem ao mesmo indivíduo;
- *valor combinado*: a combinação de todos os resultados num único valor, que é o caso de melhor resultado obtido por esta versão do sistema, ficaria abaixo do limiar para o valor combinado e, portanto, chega-se à conclusão de que as imagens não pertencem ao mesmo indivíduo;
- **Sumário de execução**: *Tempo de execução*: 81,72 segundos; *Conclusão*: errada (imagens não pertencem ao mesmo indivíduo).

- **Versão atual:**

- *região do nariz*: obtém um resultado de SIM acima do limiar de reconhecimento e, portanto, conclui-se que as imagens pertencem ao mesmo indivíduo;
- **Sumário de execução**: *Tempo de execução*: 3,46 segundos; *Conclusão*: correta (imagens pertencem ao mesmo indivíduo).

O tipo de situação ilustrada neste exemplo, pode acontecer principalmente quando em uma das imagens verificadas existe a presença de expressão facial. Assim, regiões como o nariz, que possuem menor sensibilidade à presença expressões faciais, conseguem chegar ao resultado correto; em contrapartida, o alinhamento da região dos olhos ou da face toda, é comprometido pela expressão facial, levando a um resultado errado.

Na versão atual do algoritmo, quando nenhuma região, sozinha, é suficiente para determinar se as duas imagens pertencem ou não ao mesmo indivíduo, é feita a combinação de todas as

regiões. Esta combinação baseia-se em uma média ponderada dos valores de SIM obtidos para cada uma das regiões. Os pesos utilizados para obter os resultados apresentados na seção 4.3.1.2 podem ser vistos na Figura 4.20 e foram obtidos utilizando-se um algoritmo, baseado em SA, que buscava encontrar os pesos que aplicados aos valores de SIM obtidos produziam os melhores resultados.

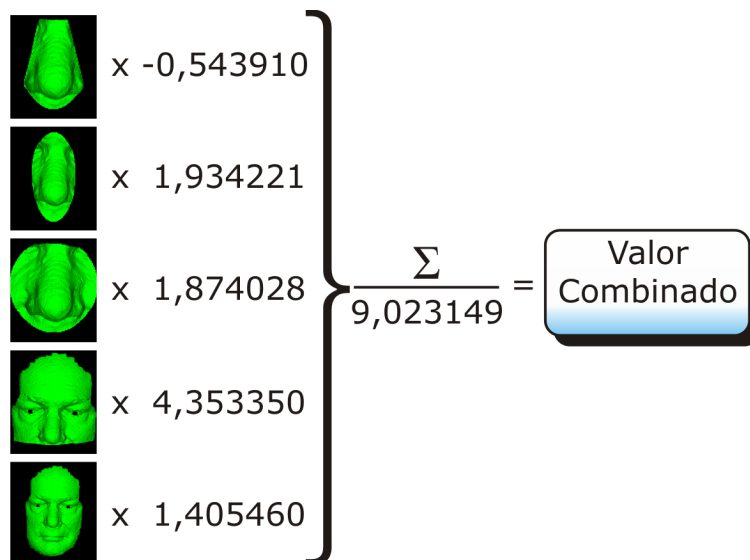


Figura 4.20: Pesos utilizados para combinar os valores obtidos em cada segmentação da face.

4.3.1 Resultados Experimentais

Nesta série de experimentos, para cada imagem da base de dados foram utilizadas cinco segmentações. Em comparação à versão preliminar, a diferença é que, ao invés de utilizar uma única segmentação do nariz, de formato retangular, foram utilizadas três segmentações: circular, elíptica e nariz propriamente dito (conforme sugerido em [17]).

4.3.1.1 Base de Dados

Os experimentos foram realizados utilizando 978 imagens da base de dados FRGC 3D *Face Database* v2.0. Esta base de dados possui 4.007 faces, entretanto, antes do início dos experimentos, foram removidas as imagens que apresentavam muito ruído ou que apresentavam expressão facial acentuada. As imagens possuem dimensão de 640×480 e foram adquiridas utilizando o sensor laser Minolta Vivid 900. Embora estas imagens possuam o mesmo número de pontos que aquelas utilizadas com a versão preliminar do sistema (Seção 4.2.1.1), a distância entre os indivíduos e o equipamento de aquisição, nesta base, diminuiu, o que resultou em faces com um número maior de pontos.

A base de dados possui um total de 256 indivíduos, todos com vista frontal e com expressão facial neutra. O número de imagens por indivíduo varia muito, chegando a até 15 imagens por indivíduo.

Assim como na versão anterior, as segmentações utilizadas foram obtidas através do método descrito em [33, 34], que também utiliza técnicas de redução de ruído.

4.3.1.2 Resultados

Da mesma forma que na versão anterior do sistema, os experimentos foram realizados combinando cada imagem da base de dados com todas as outras, gerando um total de 477.753 combinações, sendo que, destas, 2.662 são combinações entre imagens de mesmos indivíduos e 475.091 entre imagens de indivíduos diferentes. Assim como na versão anterior do sistema, os limiares empregados são definidos de forma a garantir uma FAR de 0%.

A Tabela 4.5 apresenta o limiar, taxa de verificação e número de classificações incorretas para cada segmentação, com FAR de 0%, sem utilizar a avaliação em cascata. Já que estes resultados utilizam FAR de 0%, o número de combinações incorretas representa o número de combinações entre faces pertencentes à mesma pessoa abaixo do limiar apresentado.

Região	Limiar	Taxa de Verificação	No. de Falsas Rejeições
Nariz Região	29,606	74,605%	676
Nariz Região Elíptica	24,436	85,537%	385
Nariz Região Circular	23,314	83,132%	449
Região dos Olhos	14,106	94,327%	151
Face	10,096	92,036%	212
<i>Combinação</i>	11,718	99,737%	7

Tabela 4.5: Taxa de verificação e limiar para as segmentações com FAR de 0% da base FRGC v2.0.

As Figuras 4.21–4.26 mostram a distribuição dos alinhamentos, de imagens de mesmos indivíduos e de indivíduos diferentes, em relação ao valor final da SIM para cada uma das segmentações. Como o número de combinações entre imagens de indivíduos diferentes é várias vezes maior que entre imagens de mesmos indivíduos, optou-se por utilizar a porcentagem dos dois tipos de combinação para gerar gráficos mais fáceis de serem compreendidos. Os valores de SIM obtidos foram truncados para agregar as combinações, assim, por exemplo, 12,02% e 12,96% estão agregados na mesma “barra”. Exemplo de como interpretar o gráfico: para o alinhamento da região do nariz (Figura 4.21) cerca de 6% das combinações entre indivíduos

diferentes obtiveram um valor de SIM maior ou igual a 0% e menor que 1%; e, pouco mais que 14% das combinações entre imagens de indivíduos diferentes obtiveram um valor de SIM maior ou igual a 1% e menor que 2%.

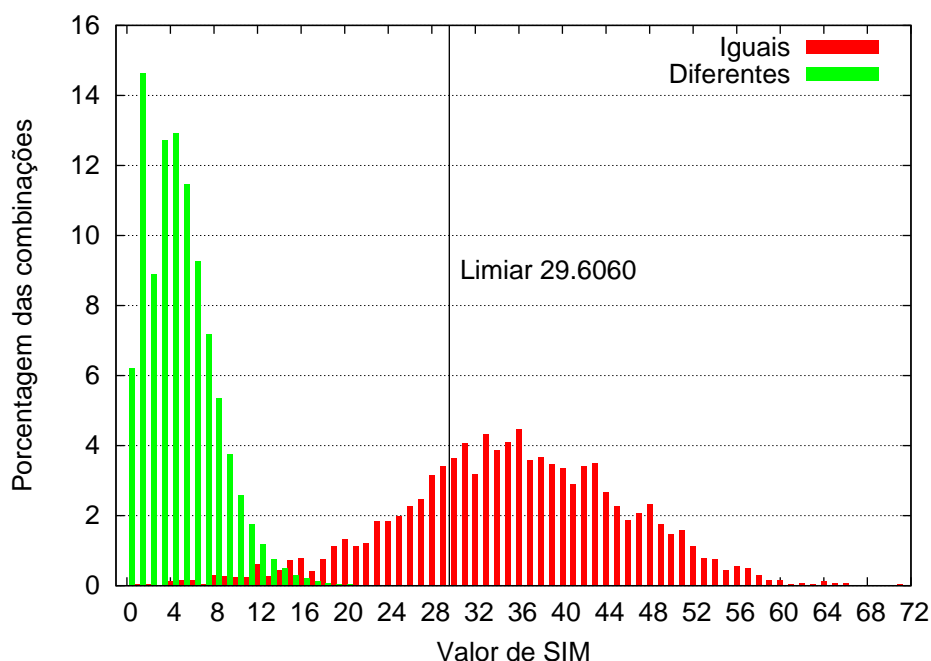


Figura 4.21: Distribuição das combinações para a região do nariz.

A segmentação que compreende a região do nariz foi a que obteve o pior resultado dentre todas as outras. Isto se deve ao fato de que estas regiões são muito pequenas e podem possuir grandes semelhanças mesmo em se tratando de indivíduos diferentes. Chegou-se a obter 29,60% de pontos interpenetrados (SIM) combinando imagens de indivíduos diferentes.

A segmentação elíptica ao redor do nariz obteve um resultado sensivelmente melhor que a segmentação da região do nariz. O motivo para esta melhoria é a utilização de uma região um pouco maior, o que dificulta a ocorrência de alinhamentos com alto valor de SIM entre imagens de indivíduos diferentes e ainda mantendo esta região pouco sensível às contrações da musculatura facial.

A segmentação circular do nariz obteve resultados semelhantes aos da região elíptica, tanto para a taxa de verificação quanto para o limiar. Isto se deve ao fato de que as duas regiões são parecidas. A maioria das combinações que foram classificadas incorretamente utilizando a região elíptica do nariz, também o foram para a região circular do nariz, cerca de 65,19%, o que equivale a 251 combinações.

A região dos olhos foi a segmentação que, analisada individualmente, apresentou o melhor resultado, atingindo uma taxa de verificação de 94,327%. O principal motivo para um resultado tão expressivo se deve ao fato de que, por se tratar de uma região relativamente grande da face,

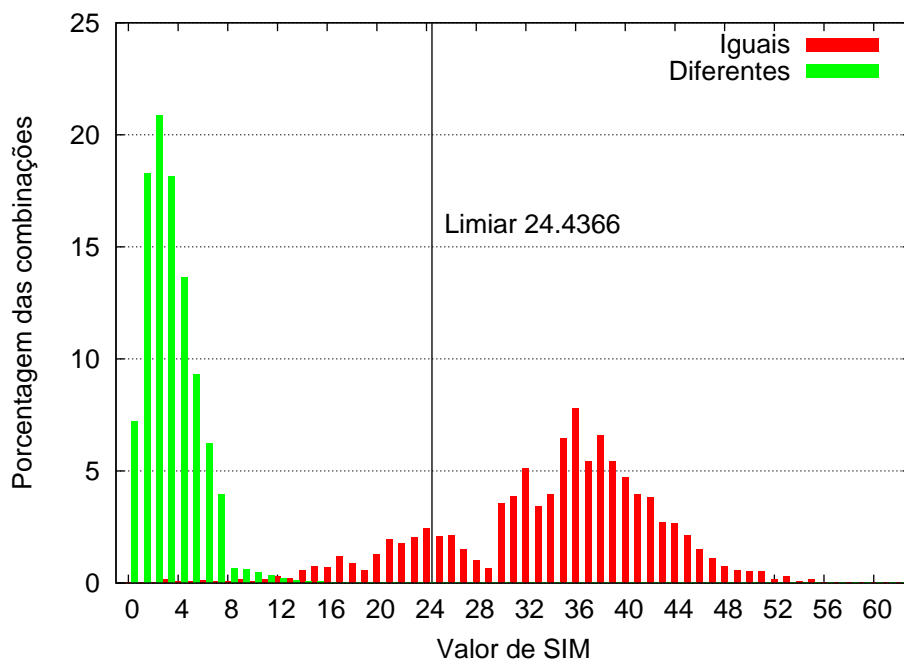


Figura 4.22: Distribuição das combinações para a região elíptica ao redor do nariz.

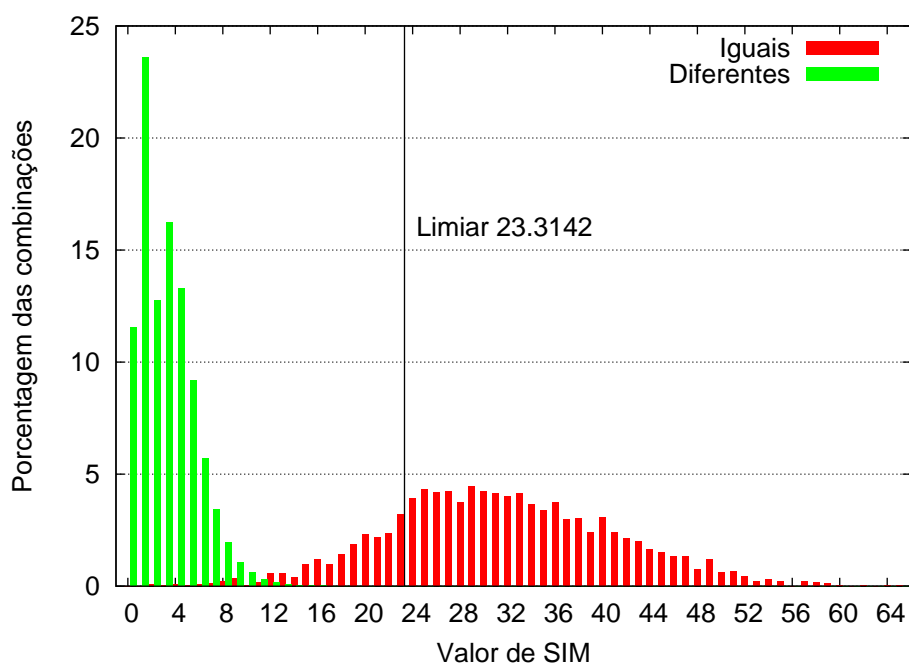


Figura 4.23: Distribuição das combinações para a região circular ao redor do nariz.

a chance de imagens pertencentes a pessoas diferentes produzirem altos valores de SIM é muito baixa. Como está sendo considerado uma FAR de 0%, uma única combinação entre imagens de indivíduos diferentes com elevado valor de SIM é suficiente para comprometer a taxa de verificação do algoritmo para a segmentação utilizada.

A utilização de toda a face apresentou resultados semelhantes aos da região dos olhos. No-

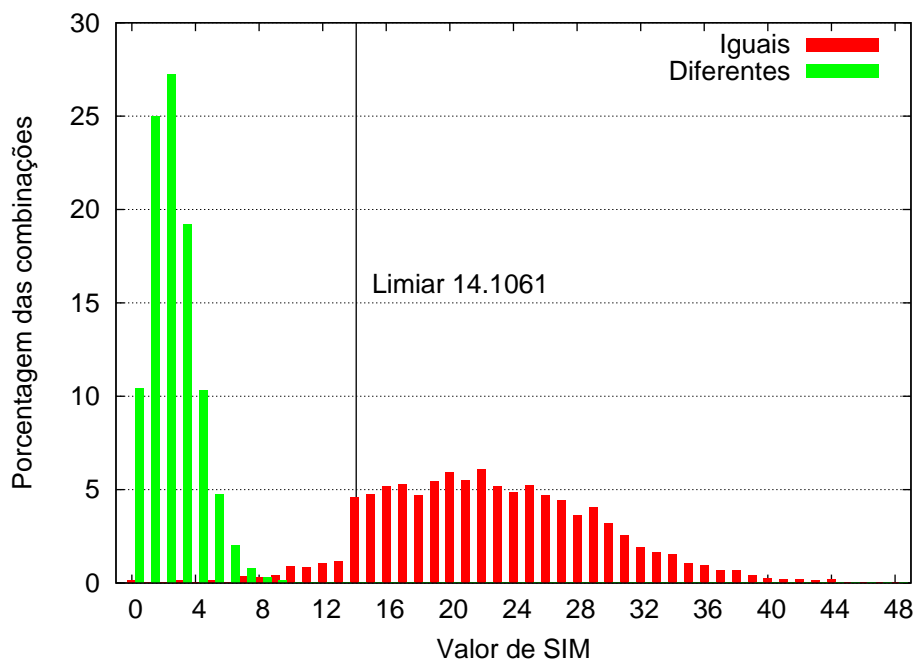


Figura 4.24: Distribuição das combinações para a região dos olhos.

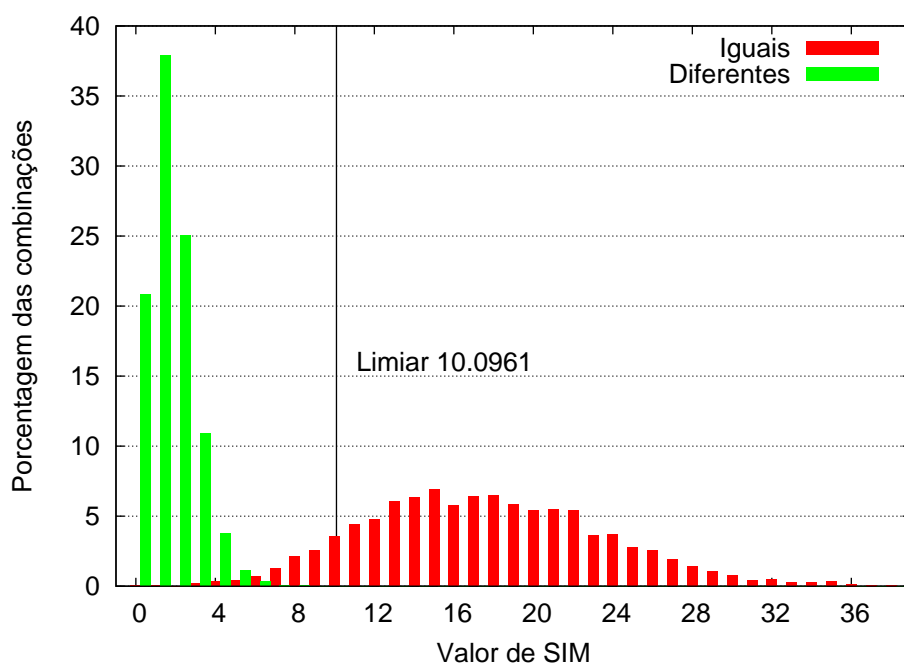


Figura 4.25: Distribuição das combinações para a região da face.

vamente, o motivo para tal semelhança é a área em comum que estas duas regiões possuem. A grande variação que ocorre na região da boca na presença de expressões faciais, mesmo quando sutis, pode ter contribuído para uma taxa de verificação abaixo daquela obtida quando utilizou-se somente a região dos olhos.

A avaliação em cascata do algoritmo rejeita, falsamente, apenas 6 combinações. O ganho

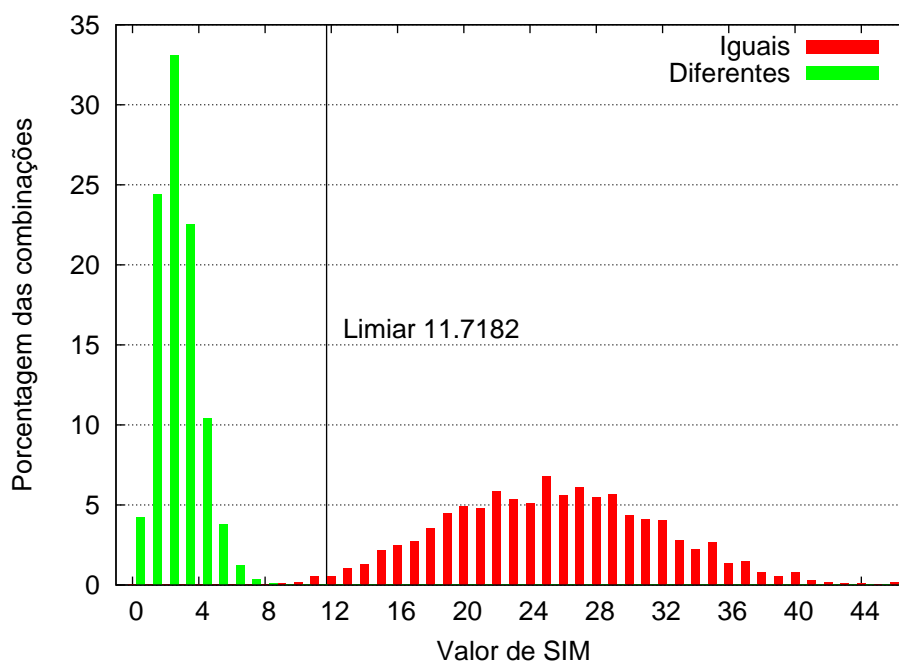


Figura 4.26: Distribuição das combinações para a combinação de todas as regiões.

final, em relação à utilização do valor combinado das características pode parecer pequeno, já que utilizando o valor combinado, o número de falsas rejeições é de 7. No entanto, o principal diferencial para esta abordagem é que, além de melhorar o resultado final, o tempo de execução sofre uma sensível redução, já que, não é necessário executar o algoritmo para todas as segmentações. A taxa de verificação para o algoritmo em cascata é de 99,774%.

Na Figura 4.27 (a) e (b) é possível observar as curvas de FAR e FRR, respectivamente, para o algoritmo em cascata. Na curva da FAR nota-se que sua queda é muito acentuada, devido aos baixos valores de SIM produzidos para os alinhamentos entre pessoas diferentes.

Na Figura 4.28 (a) é possível ver a junção das curvas de FAR e de FRR. Para que fique mais clara sua visualização, na Figura 4.28 (b) está representado o ponto de cruzamento das curvas, de onde é possível obter o EER do algoritmo. No caso, o EER do algoritmo é de aproximadamente 0,033% (ou $3,3 \times 10^{-4}$).

Na Figura 4.29 (a) pode-se observar a área da curva ROC próxima à coordenada (0,0). Na Figura 4.29 (b) é apresentada a curva DET da versão atual do algoritmo.

Além da avaliação de performance como sistema de verificação, para esta versão do algoritmo, também foi avaliado seu desempenho como sistema de identificação (ou reconhecimento).

Esta avaliação consiste em, dada uma imagem de entrada, descobrir a identidade da pessoa desta imagem, através de uma busca na base dados. Para tanto, as imagens de cada segmentação

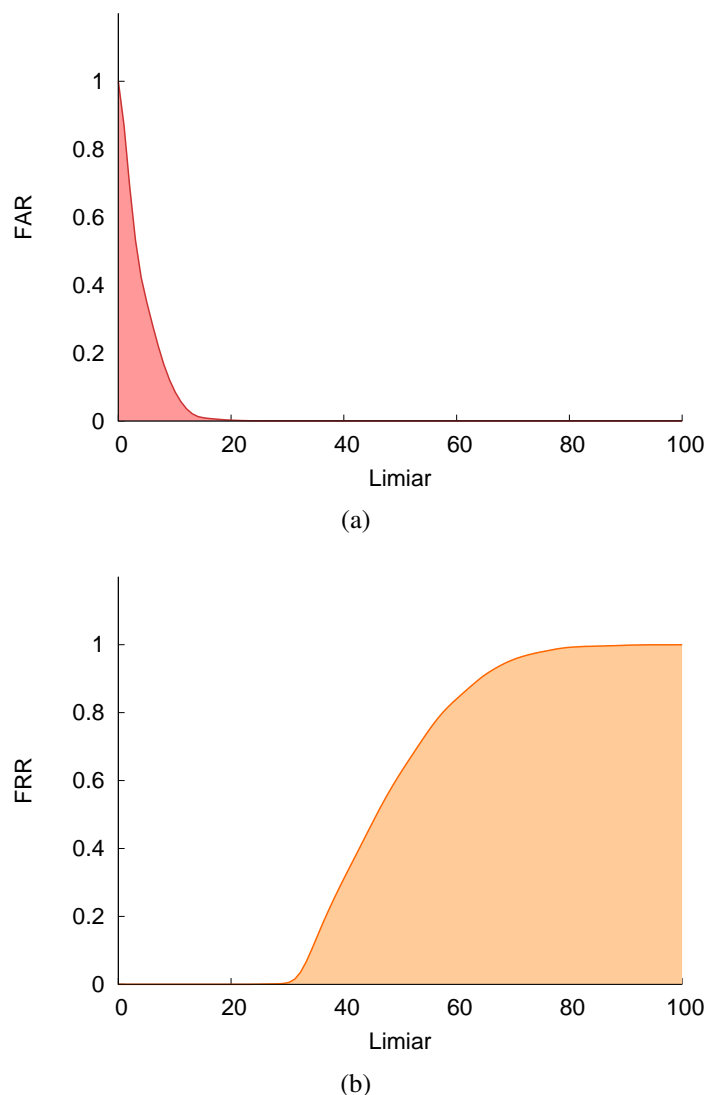


Figura 4.27: Curvas de FAR e de FRR de acordo com o limiar, para a versão atual do algoritmo: (a) FAR; (b) FRR.

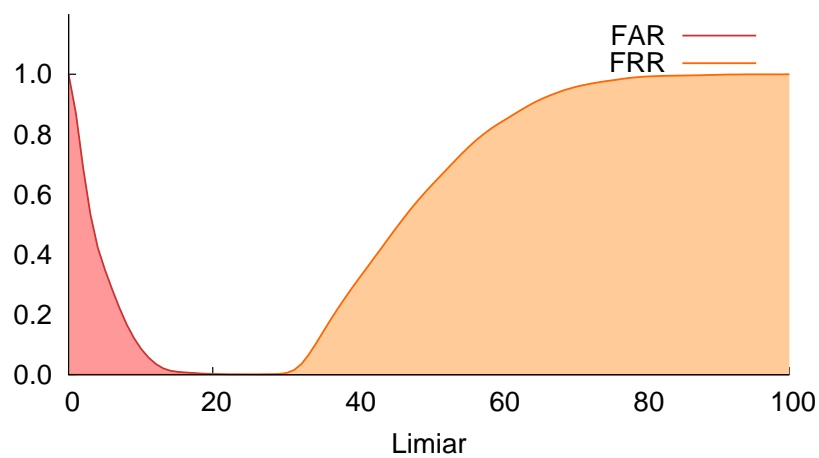
foram divididas em grupos:

Base de Dados composto de 256 imagens, este grupo contém exatamente uma imagem para cada um dos 256 indivíduos. Cada imagem submetida será comparada com todas as imagens deste grupo. Esta é a *gallery* do sistema;

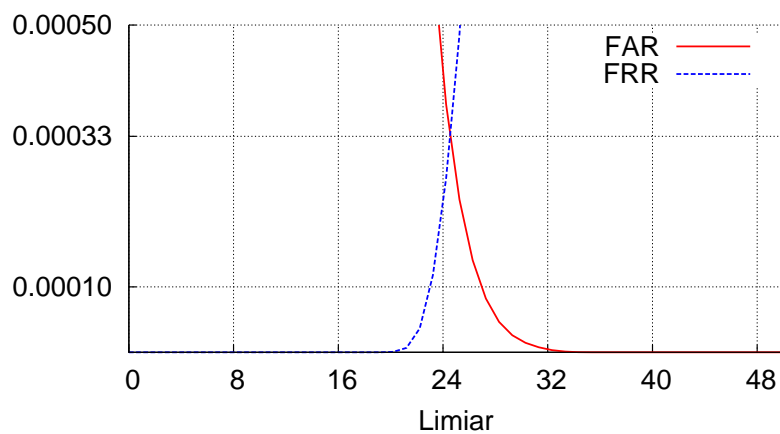
Provação 1 composto de 361 imagens, as imagens deste grupo são submetidas ao algoritmo desenvolvido para que este tente identificá-las. Este é um dos grupos de *probe* do sistema;

Provação 2 assim como o grupo anterior, este é composto de 361 imagens que são submetidas ao algoritmo para identificação.

A Tabela 4.6 apresenta a taxa de reconhecimento *rank-one* e o número de identificações



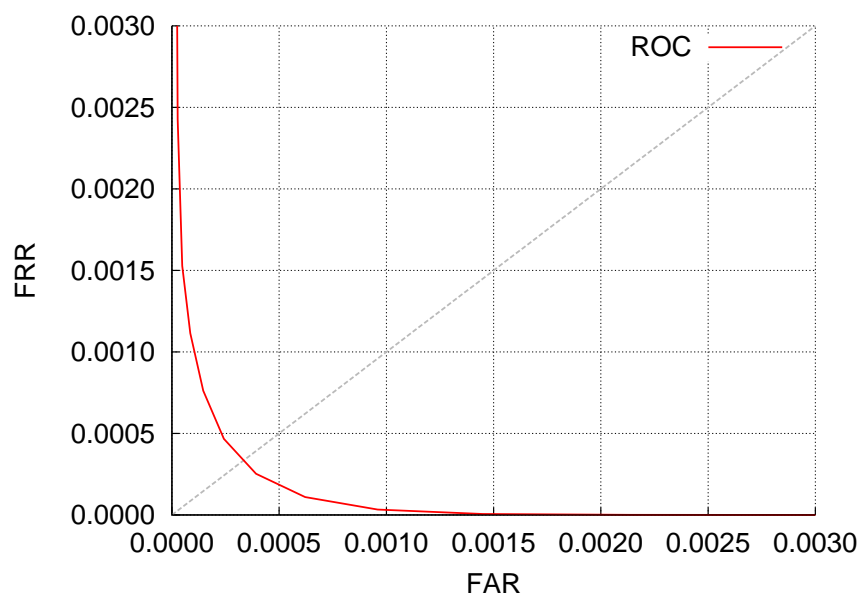
(a)



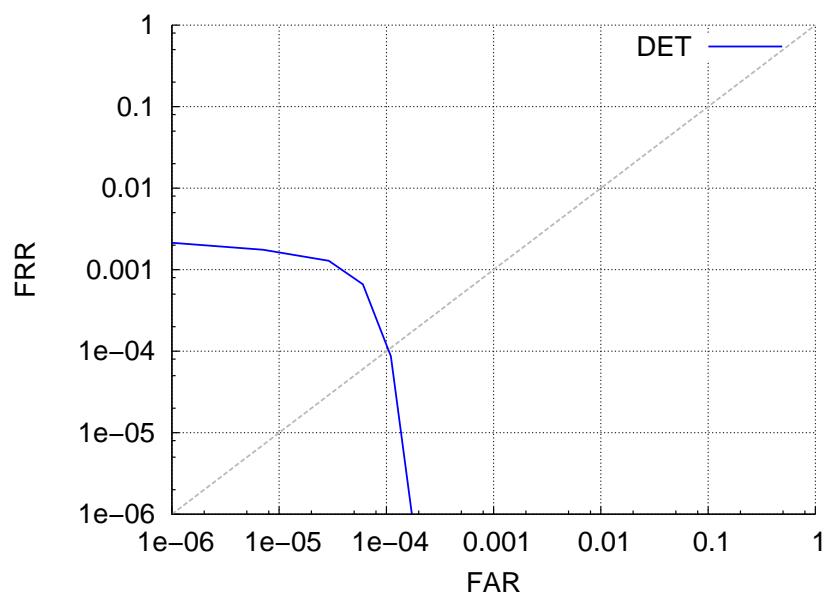
(b)

Figura 4.28: Cruzamento das curvas de FAR e de FRR, para a versão atual do algoritmo: (a) curvas de FAR e de FRR; (b) região do cruzamento das curvas.

incorretas, entre parênteses, obtidas pelo algoritmo, para cada uma das segmentações por ele utilizadas.



(a)



(b)

Figura 4.29: Curvas ROC e DET, para a versão atual do algoritmo: (a) curva ROC; (b) curva DET.

Região	Taxa de Reconhecimento <i>Rank-One</i> / No. de Ident. Incorretas		
	Grupo de Probação 1	Grupo de Probação 2	Média
Nariz Região	95,291% (17)	97,784% (8)	96,537% (12,5)
Nariz Região Elíptica	98,338% (6)	98,615% (5)	98,476% (5,5)
Nariz Região Circular	98,338% (6)	98,892% (4)	98,615% (5)
Região dos Olhos	99,723% (1)	99,446% (2)	99,584% (1,5)
Face	98,892% (4)	99,169% (3)	99,030% (3,5)
<i>Combinação</i>	100,000% (0)	100,000% (0)	100,000% (0)

Tabela 4.6: Taxa de reconhecimento *rank-one* e número de identificações incorretas da base FRGC v2.0.

Como pôde ser visto, para a combinação das regiões, o valor obtido para a taxa de reconhecimento *rank-one* foi de 100%. Isto significa que o algoritmo foi capaz de identificar, corretamente, todas as imagens que compõem os grupos de probação 1 e 2. A Figura 4.30 apresenta a curva CMC do algoritmo.

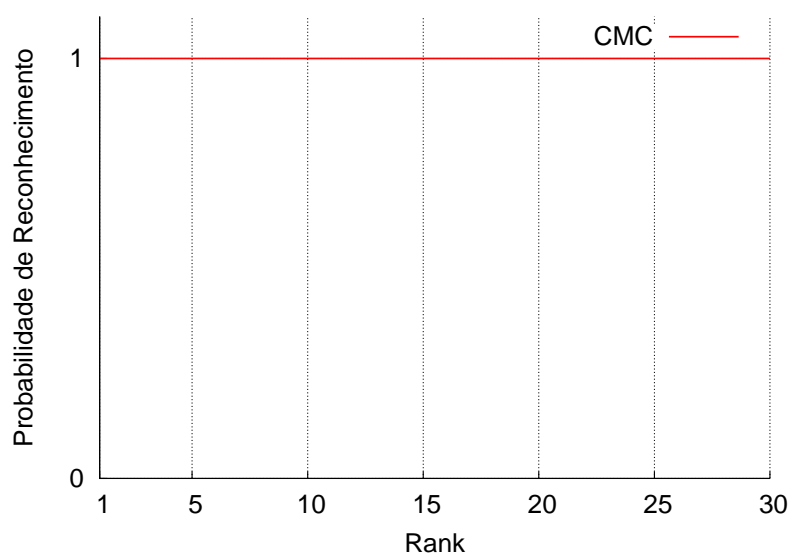


Figura 4.30: Curva CMC, para a versão atual do algoritmo.

4.3.1.2.1 Exemplos de Resultados Nesta seção serão apresentados alguns exemplos de resultados obtidos com o a versão atual do SAM. Para cada par de imagens será exibido: as imagens de entrada da face, os alinhamentos para cada segmentação e a imagem binária para a região da face. Além das imagens, será informado o valor de SIM obtido para cada alinhamento de cada segmentação.

As Figuras de 4.31 a 4.36 mostram as combinações entre imagens de mesmos indivíduos que foram falsamente rejeitadas.

As Figuras de 4.37 a 4.40 mostram algumas combinações entre imagens de mesmos indivíduos que obtiveram altos valores de SIM.

A Figura 4.41 mostra um exemplo de combinação de imagens de indivíduos diferentes. Já a Figura 4.42 mostra a combinação entre imagens de pessoas diferentes que obteve o maior valor de SIM (para a combinação das segmentações).

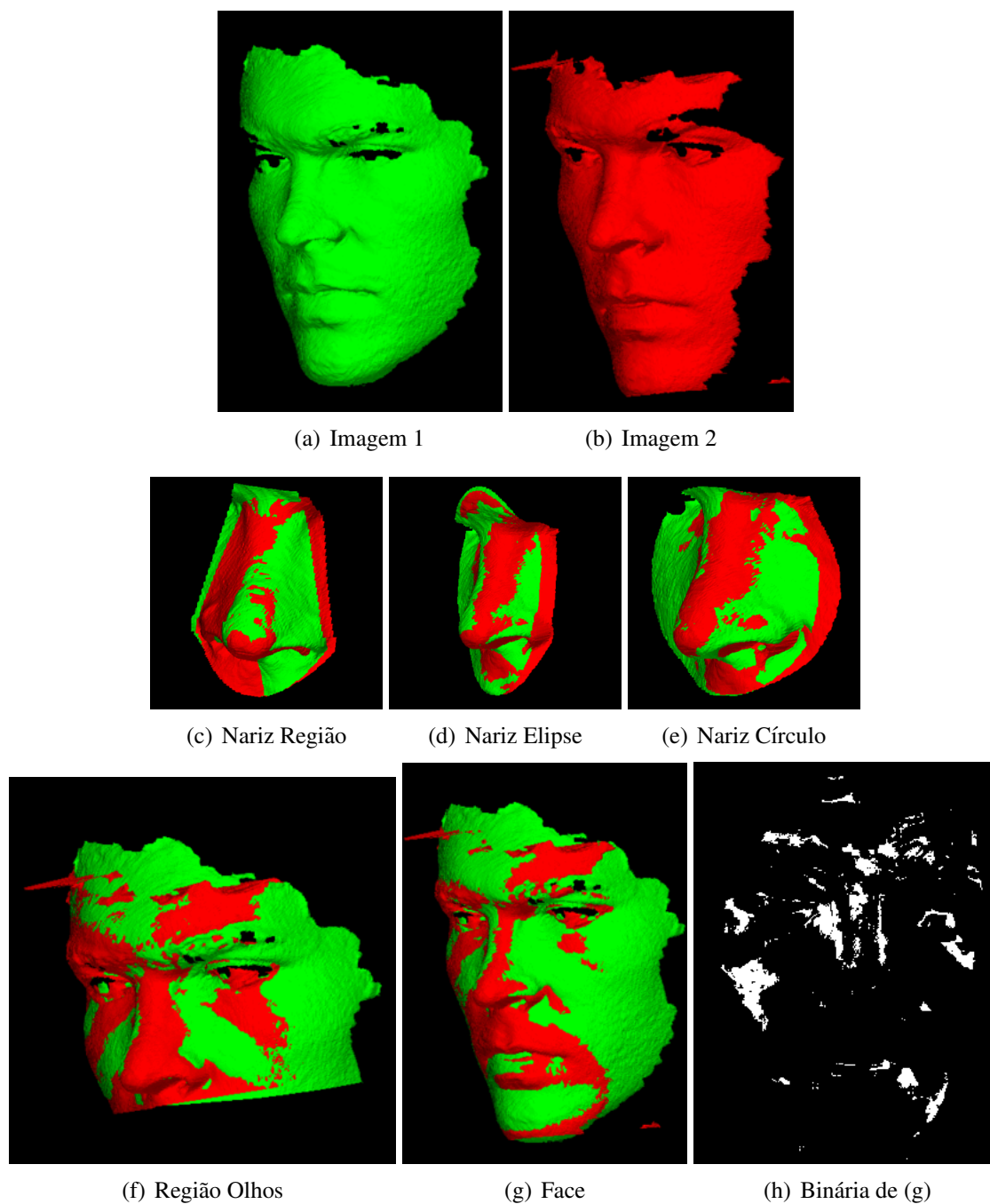


Figura 4.31: Exemplo de alinhamento das segmentações utilizadas pelo algoritmo para indivíduos iguais: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 13,33\%$; (d) região elíptica do nariz, $SIM = 16,50\%$; (e) região circular do nariz, $SIM = 11,62\%$; (f) região dos olhos, $SIM = 15,33\%$; (g) face, $SIM = 7,98\%$; (h) imagem binária de (g).

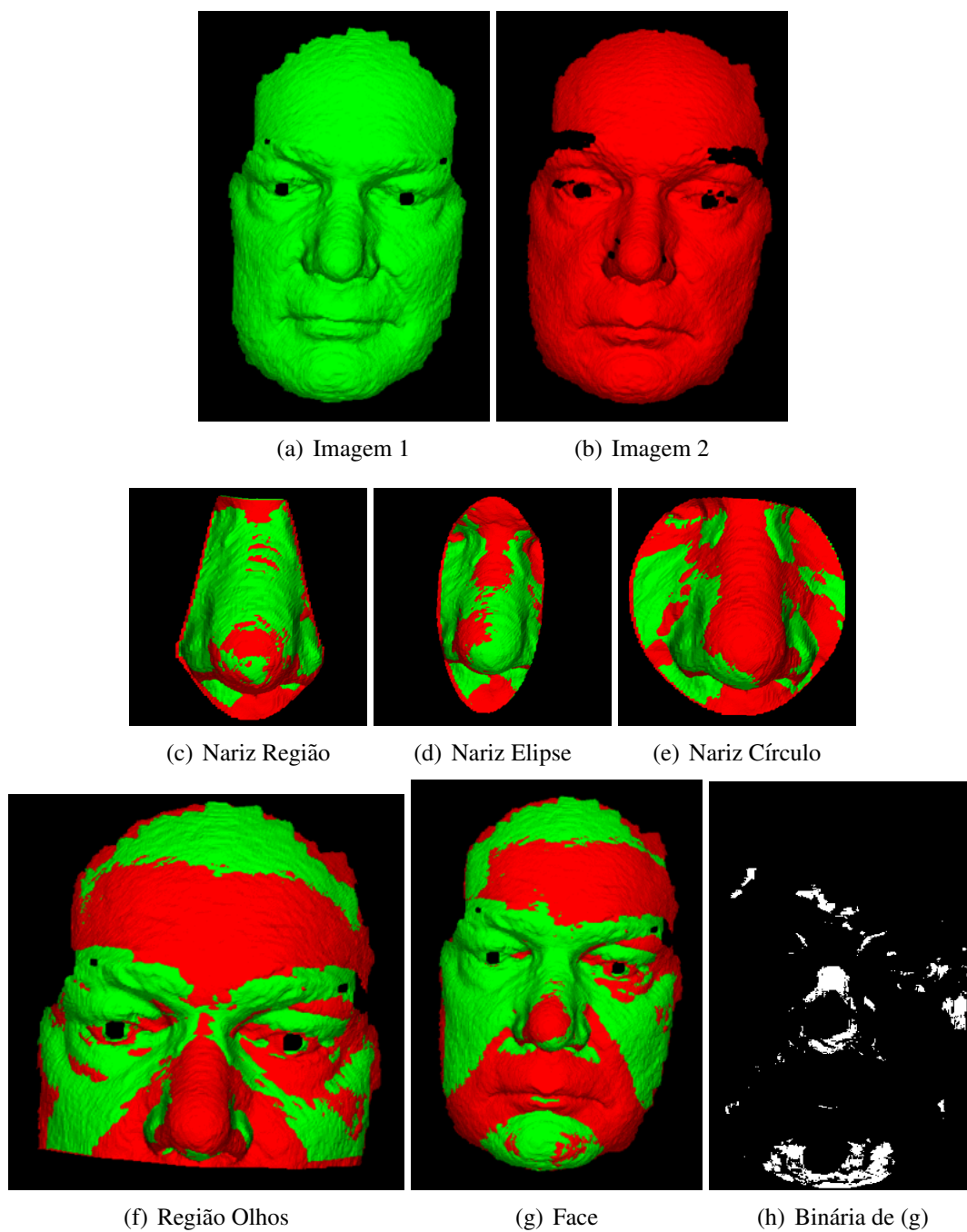


Figura 4.32: Exemplo de alinhamento das segmentações utilizadas pelo algoritmo para indivíduos iguais: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 16,80\%$; (d) região elíptica do nariz, $SIM = 11,98\%$; (e) região circular do nariz, $SIM = 13,20\%$; (f) região dos olhos, $SIM = 7,75\%$; (g) face, $SIM = 6,58\%$; (h) imagem binária de (g).

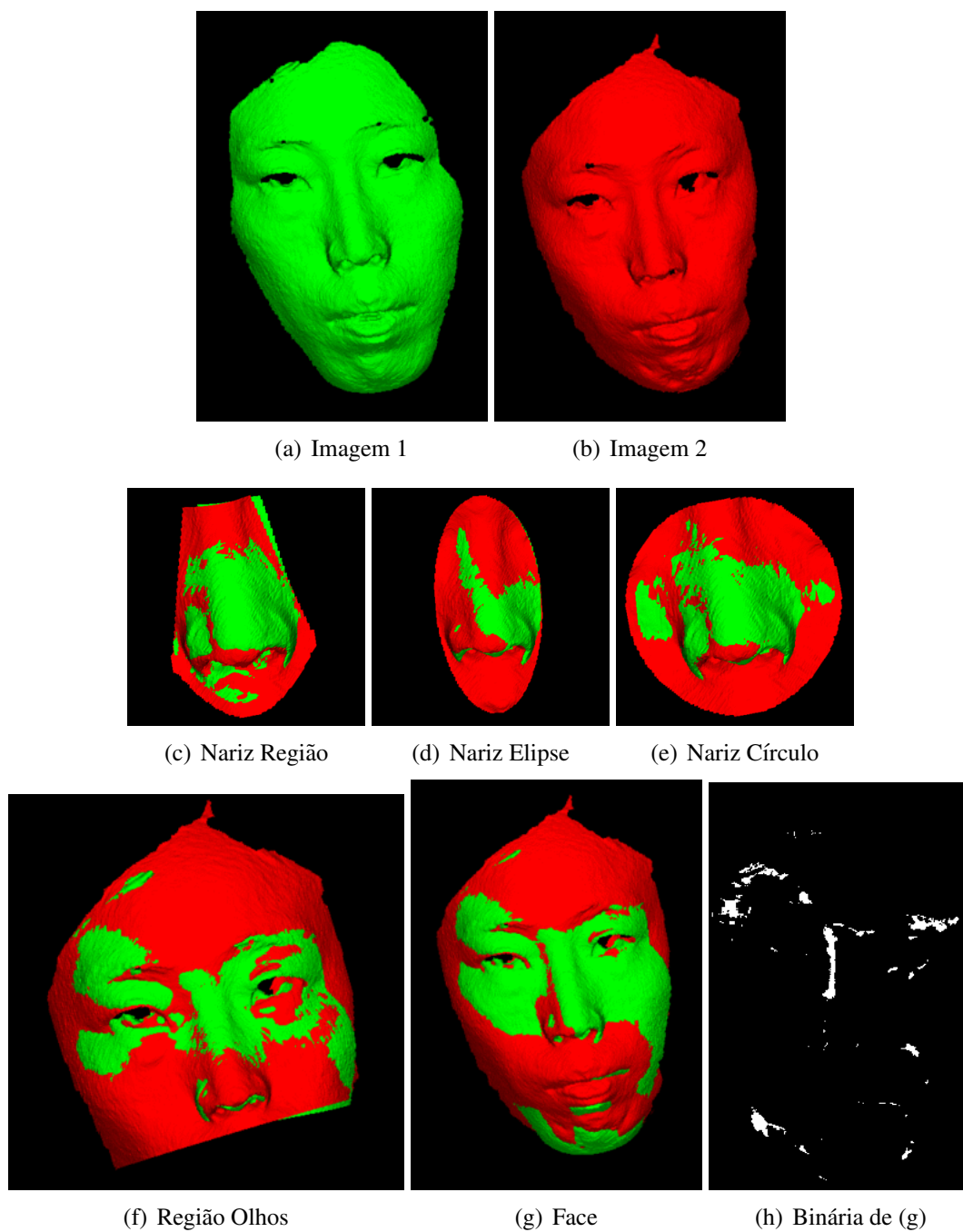


Figura 4.33: Exemplo de alinhamento das segmentações de um mesmo indivíduo que foi falsamente rejeitado: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 20,13\%$; (d) região elíptica do nariz, $SIM = 10,28\%$; (e) região circular do nariz, $SIM = 22,52\%$; (f) região dos olhos, $SIM = 8,64\%$; (g) face, $SIM = 3,00\%$; (h) imagem binária de (g).

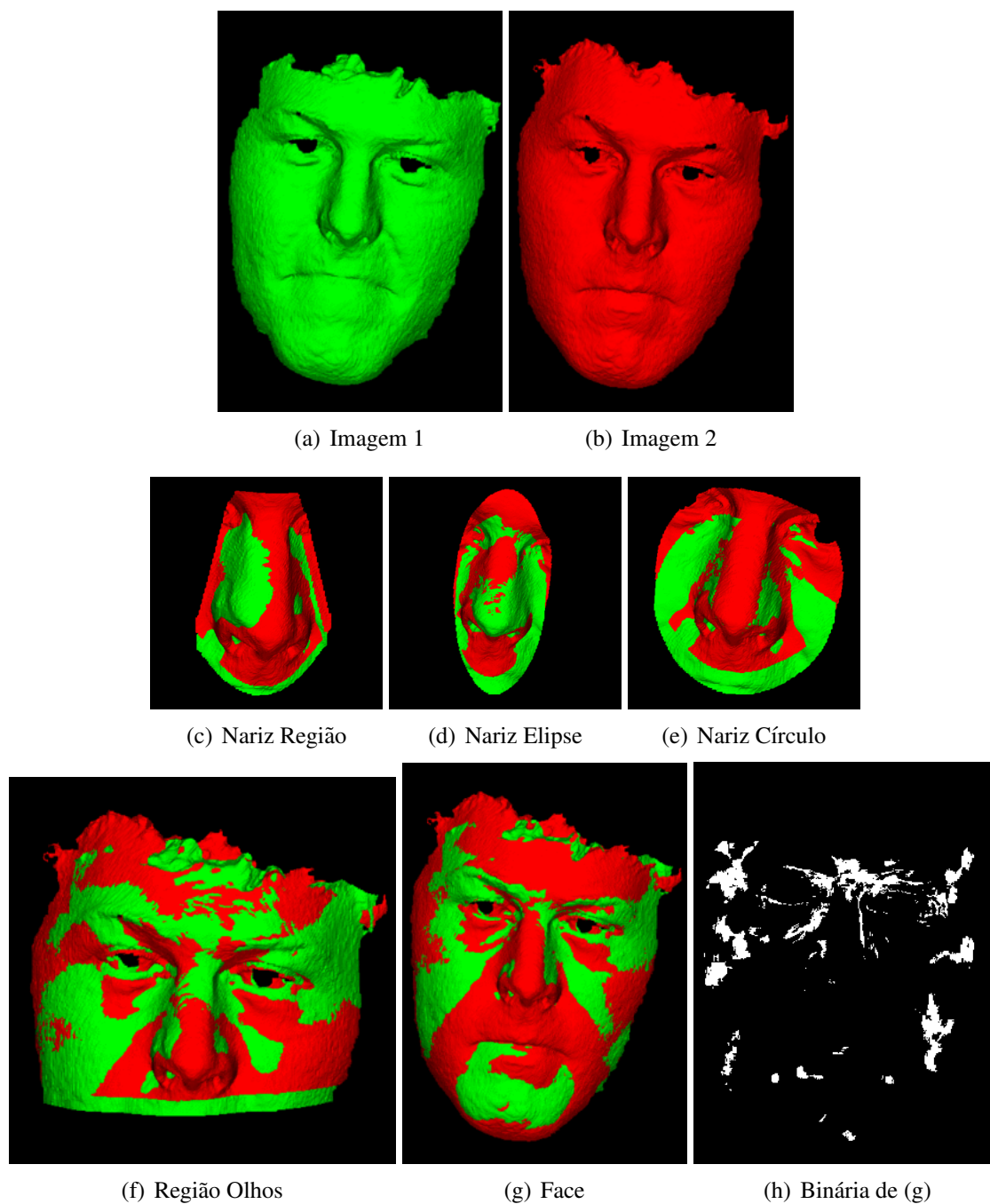


Figura 4.34: Exemplo de alinhamento das segmentações de um mesmo indivíduo que foi falsamente rejeitado: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 0,83\%$; (d) região elíptica do nariz, $SIM = 16,70\%$; (e) região circular do nariz, $SIM = 7,86\%$; (f) região dos olhos, $SIM = 11,75\%$; (g) face, $SIM = 8,54\%$; (h) imagem binária de (g).

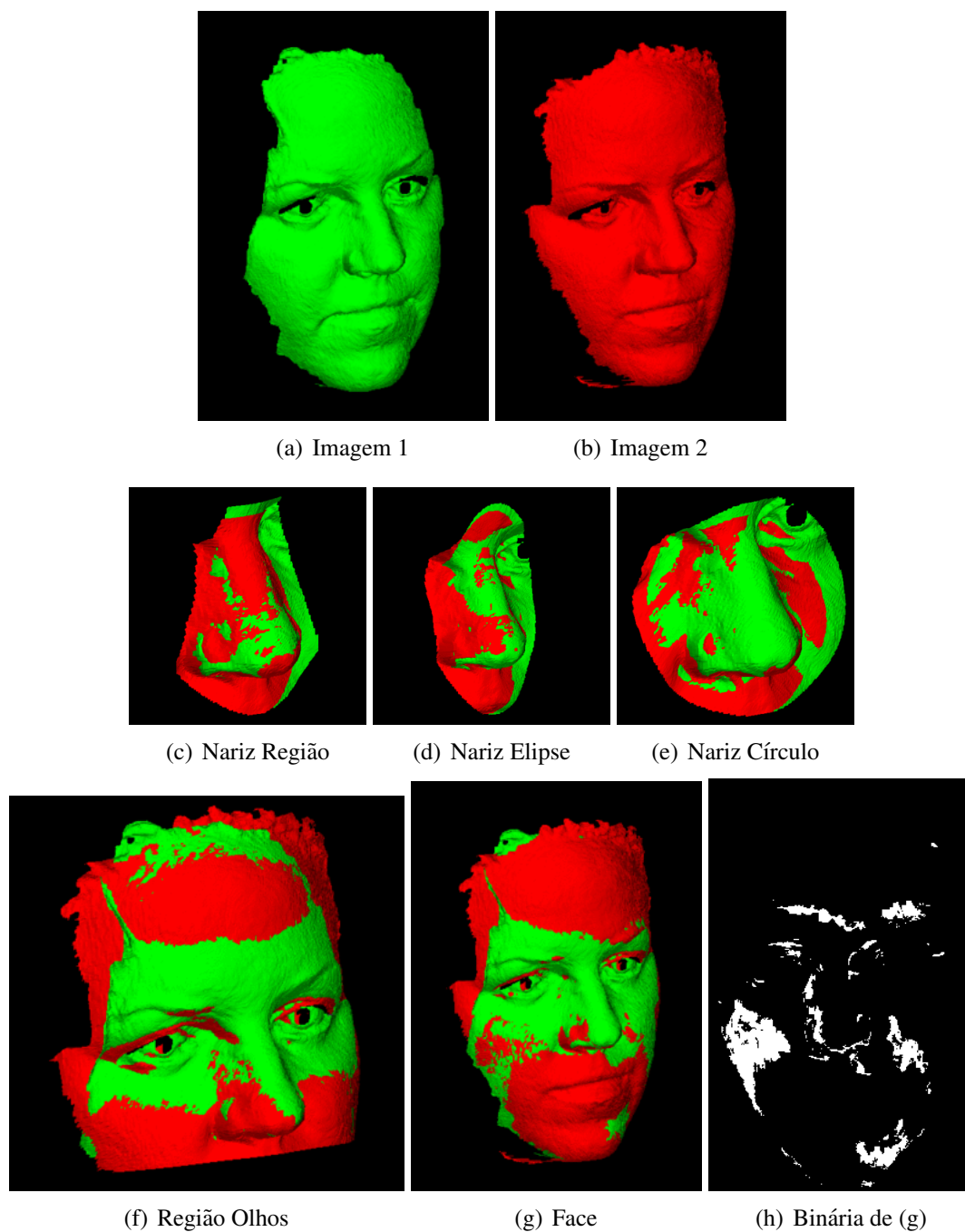


Figura 4.35: Exemplo de alinhamento das segmentações de um mesmo indivíduo que foi falsamente rejeitado: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 21,54\%$; (d) região elíptica do nariz, $SIM = 19,55\%$; (e) região circular do nariz, $SIM = 14,75\%$; (f) região dos olhos, $SIM = 7,75\%$; (g) face, $SIM = 8,29\%$; (h) imagem binária de (g).

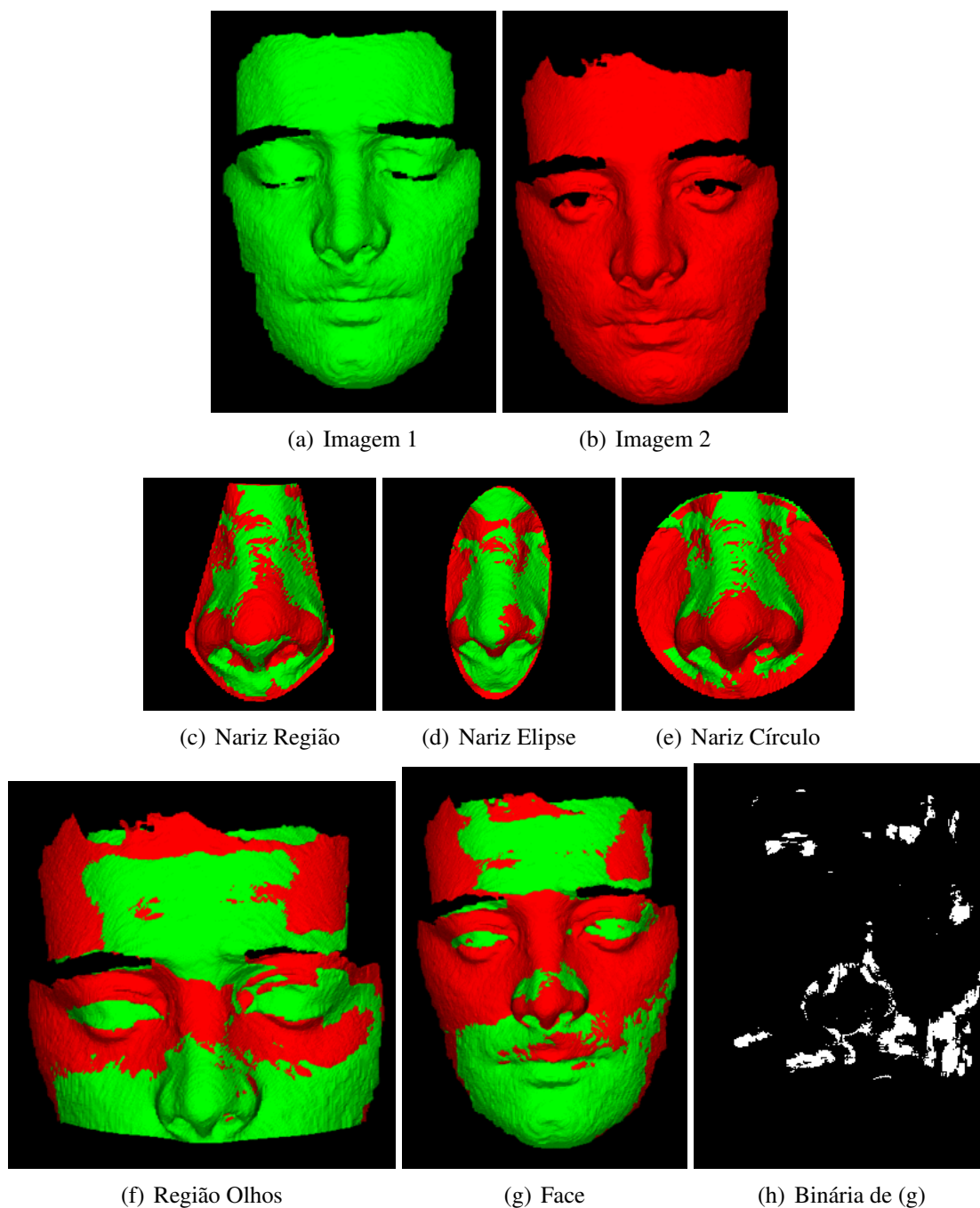


Figura 4.36: Exemplo de alinhamento das segmentações de um mesmo indivíduo que foi falsamente rejeitado: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 34,48\%$; (d) região elíptica do nariz, $SIM = 18,07\%$; (e) região circular do nariz, $SIM = 23,64\%$; (f) região dos olhos, $SIM = 6,55\%$; (g) face, $SIM = 7,27\%$; (h) imagem binária de (g).

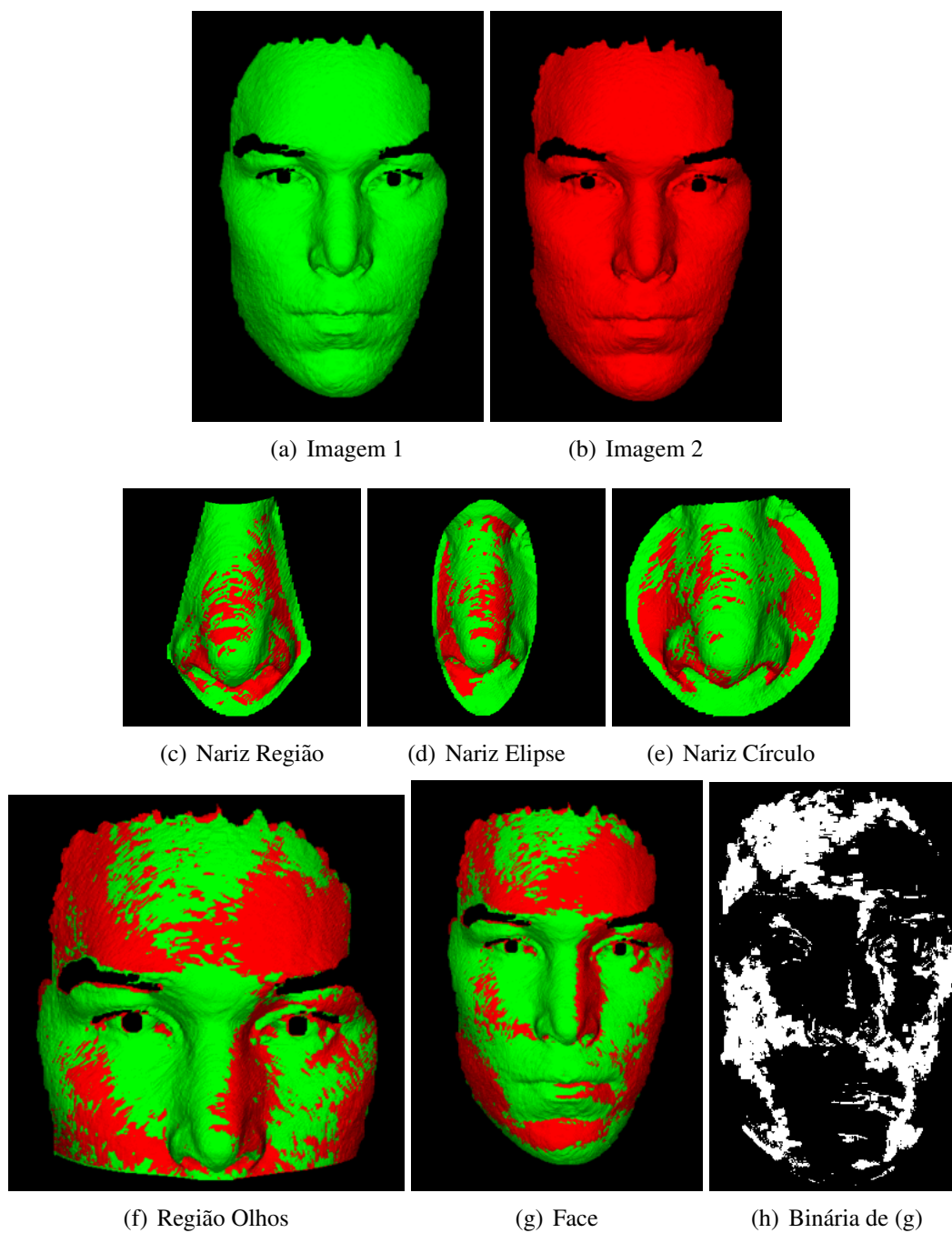


Figura 4.37: Exemplo de alinhamento das segmentações de um mesmo indivíduo: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 71,83\%$; (d) região elíptica do nariz, $SIM = 59,06\%$; (e) região circular do nariz, $SIM = 68,20\%$; (f) região dos olhos, $SIM = 46,80\%$; (g) face, $SIM = 33,36\%$; (h) imagem binária de (g).

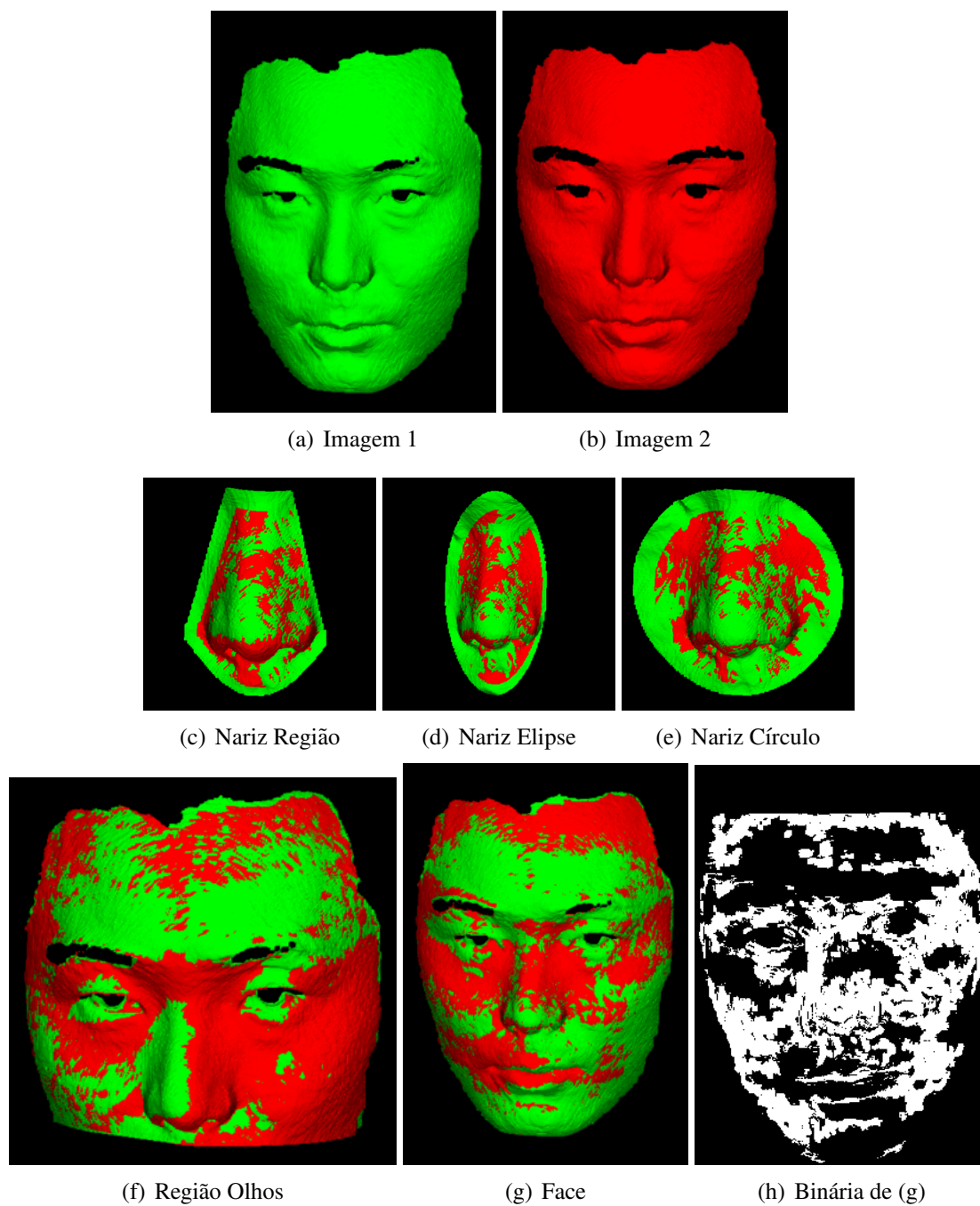


Figura 4.38: Exemplo de alinhamento das segmentações de um mesmo indivíduo: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 68,82\%$; (d) região elíptica do nariz, $SIM = 68,50\%$; (e) região circular do nariz, $SIM = 79,55\%$; (f) região dos olhos, $SIM = 39,33\%$; (g) face, $SIM = 45,34\%$; (h) imagem binária de (g).

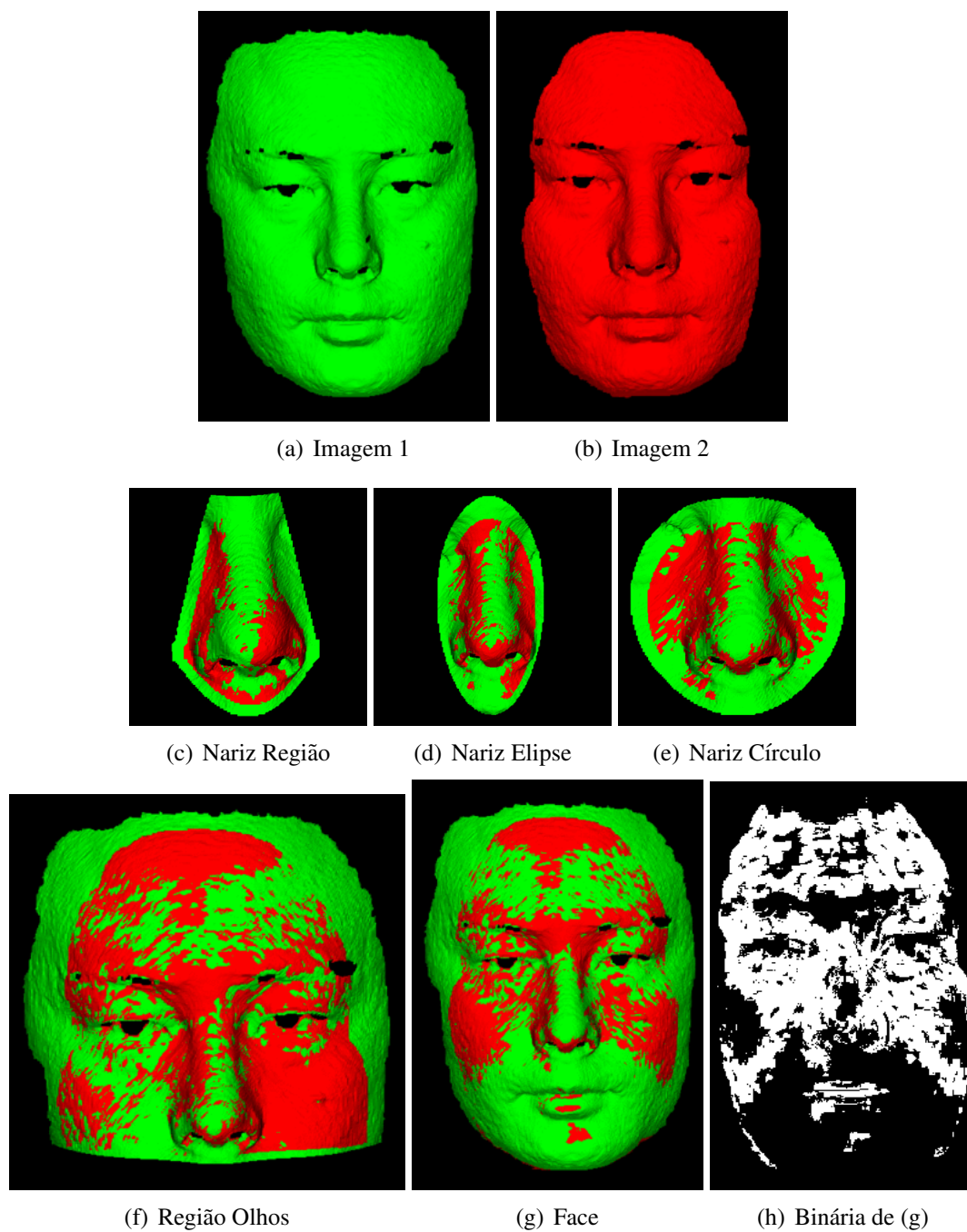


Figura 4.39: Exemplo de alinhamento das segmentações de um mesmo indivíduo: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 49,08\%$; (d) região elíptica do nariz, $SIM = 59,84\%$; (e) região circular do nariz, $SIM = 70,84\%$; (f) região dos olhos, $SIM = 55,91\%$; (g) face, $SIM = 47,11\%$; (h) imagem binária de (g).

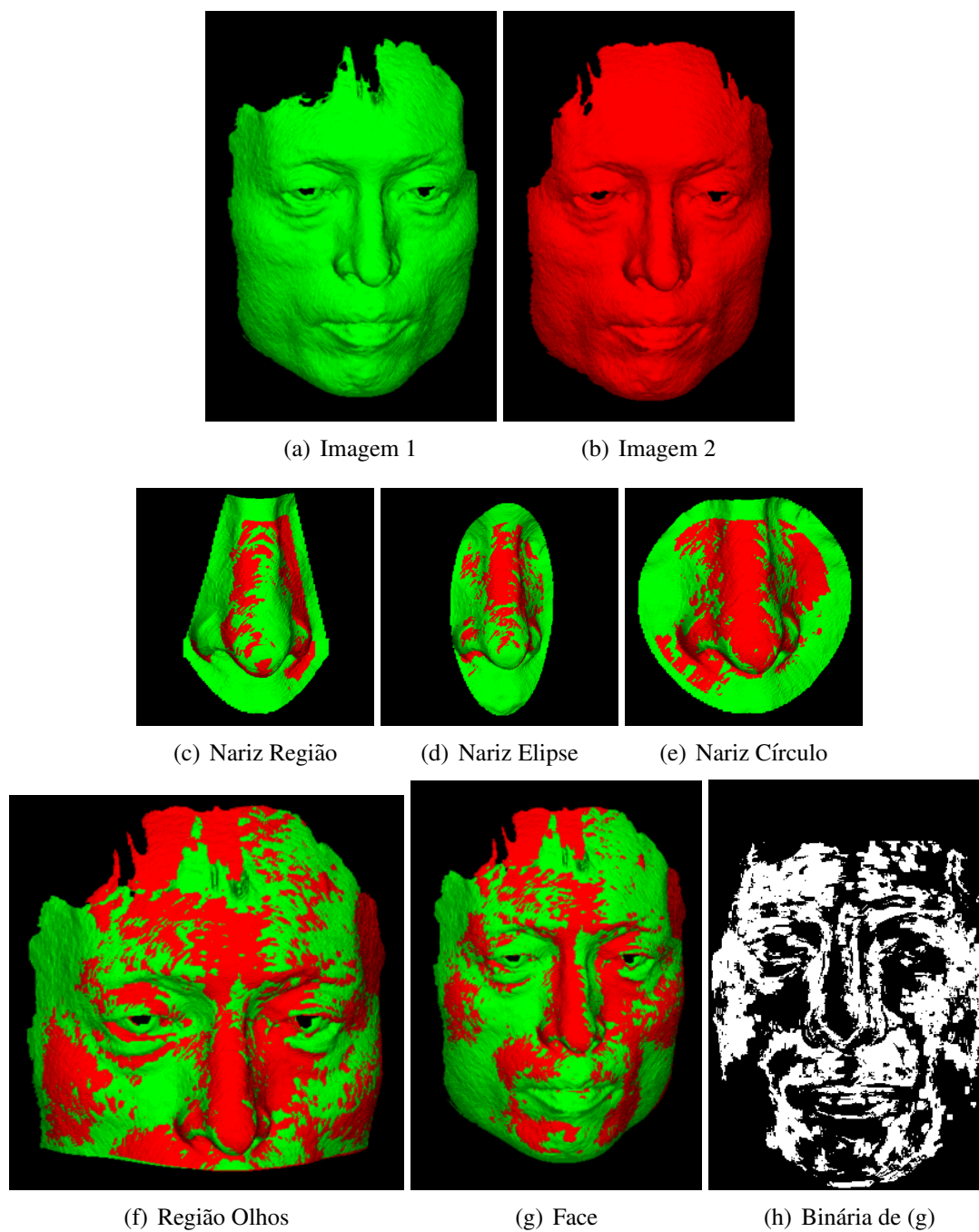


Figura 4.40: Exemplo de alinhamento das segmentações de um mesmo indivíduo: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 51,59\%$; (d) região elíptica do nariz, $SIM = 48,91\%$; (e) região circular do nariz, $SIM = 45,60\%$; (f) região dos olhos, $SIM = 51,84\%$; (g) face, $SIM = 41,38\%$; (h) imagem binária de (g).

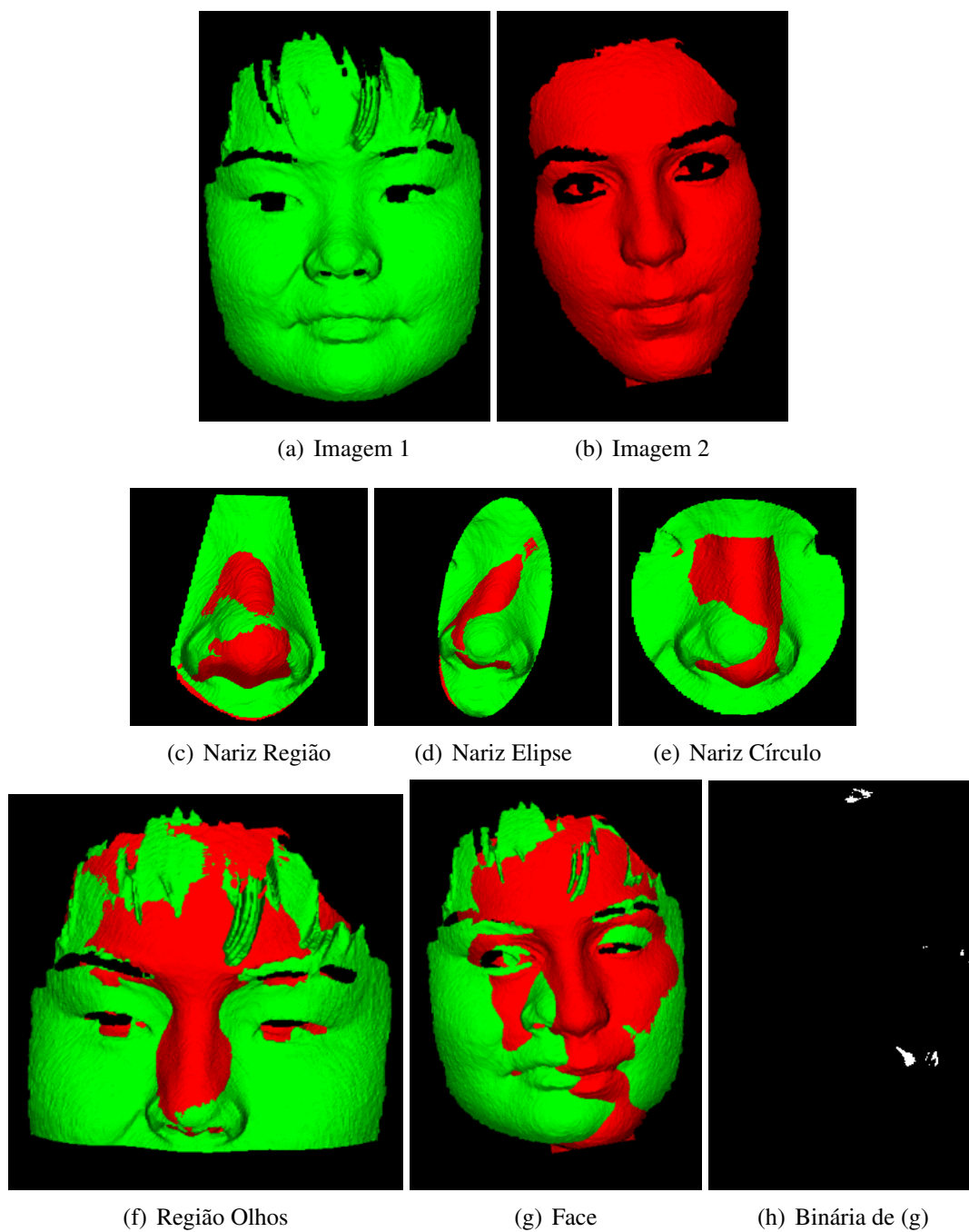


Figura 4.41: Exemplo de alinhamento das segmentações de indivíduos diferentes: (a) imagem 1; (b) imagem 2; (c) região do nariz, $SIM = 4,54\%$; (d) região elíptica do nariz, $SIM = 1,49\%$; (e) região circular do nariz, $SIM = 0,01\%$; (f) região dos olhos, $SIM = 2,17\%$; (g) face, $SIM = 0,47\%$; (h) imagem binária de (g).

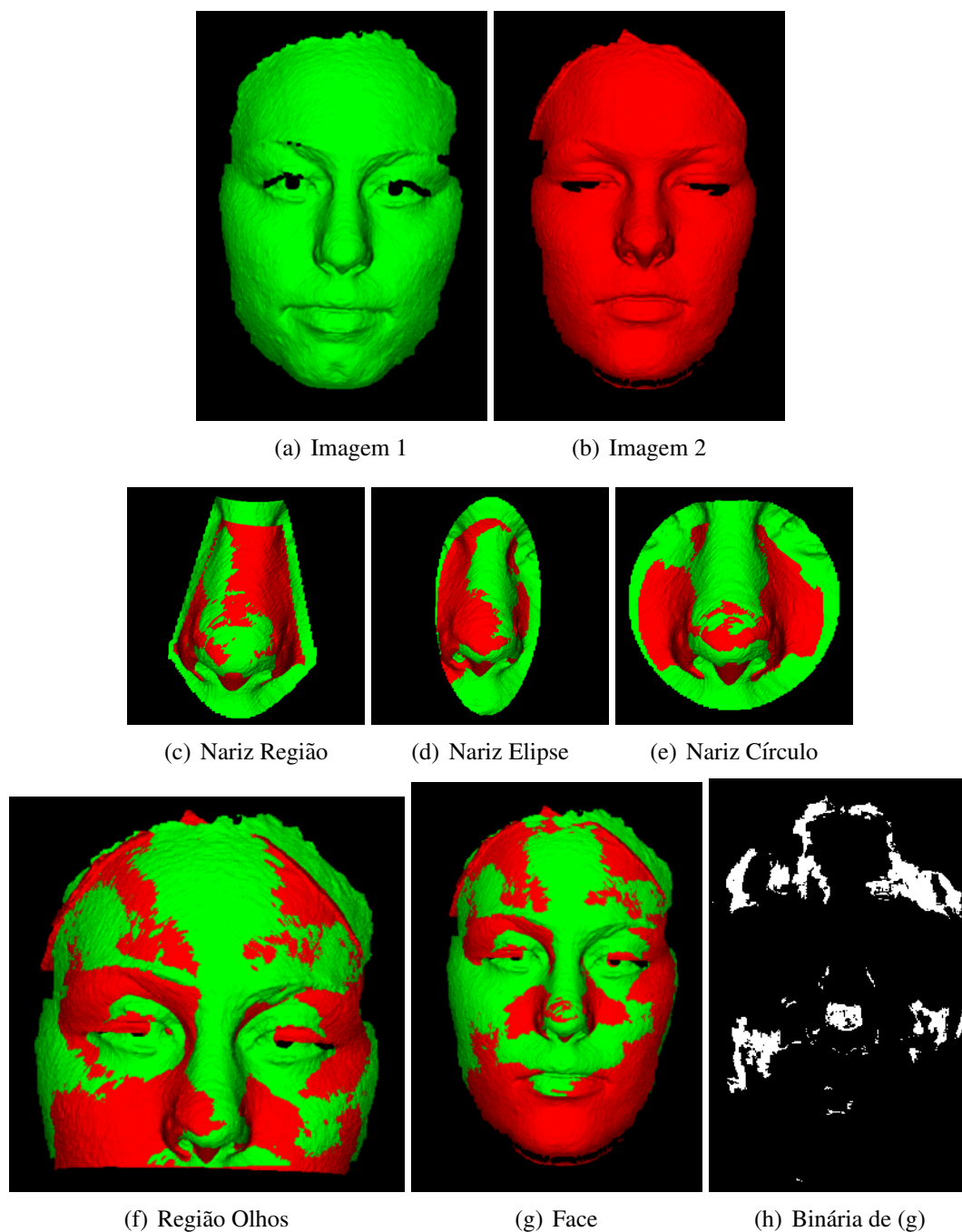


Figura 4.42: Alinhamento das segmentações de indivíduos diferentes, que produziram o maior valor de SIM: (a) imagem 1; (b) imagem 2; (c) região do nariz, SIM = 18,47%; (d) região elíptica do nariz, SIM = 7,85%; (e) região circular do nariz, SIM = 9,46%; (f) região dos olhos, SIM = 14,80%; (g) face, SIM = 11,23%; (h) imagem binária de (g).

4.4 VANTAGENS VS. DESVANTAGENS E LIMITAÇÕES

Se compararmos o SAM com outros algoritmos, apresentados no Capítulo 3, podemos ver que seus resultados são superiores. A Tabela 4.7 mostra a comparação do SAM segundo alguns critérios.

	ARMS[17]	Alg. [16]	Alg. [26]	Alg. [27]	SAM
Rank-One	96,5%	97,1%	98,0%	99,0%	100%
Tax. Verif.	-	-	-	-	99,77%
EER	-	0,12%	-	-	0,033%
Tempo	-	1s	30s	20s ¹	3s – 81s ²

Tabela 4.7: Comparação do SAM com os algoritmos [16, 17, 26, 27]

¹ Intel Pentium 4™ 2.8GHz

² AMD Sempron™ 3100+

Como pode ser visto, a única desvantagem do SAM em relação aos demais algoritmos, é o seu tempo de processamento. No entanto, algumas modificações foram feitas para melhorar sua velocidade, entre elas destacam-se:

- Simplificação de algumas estruturas de dados e tradução dessas estruturas para C;
- Utilização de apenas um pré-alinhamento (centro de massa);
- Utilização de uma amostragem menor de pontos, apenas 0,6%, no alinhamento com o MSAC.

Com estas modificações o tempo de execução do algoritmo tem uma sensível redução. Ao invés do tempo de 3s a 81s, o novo tempo varia de 0,5s a 18,3s⁵. Se considerarmos que o algoritmo utiliza avaliação em cascata, o tempo de execução do algoritmo tende a ficar mais próximo do limite inferior. Maiores detalhes sobre a execução do SA utilizando as modificações supracitadas, podem ser vistas no Apêndice A.

A utilização de GA com a SIM pode produzir resultados semelhantes e, até mesmo, melhores que os do SAM mas, a um custo computacional muito mais elevado.

A comparação entre o ICP, GA com SIM e o SA, todos utilizando a SIM como medida de similaridade entre as imagens, pode ser visto no Apêndice A.

As limitações da utilização do SAM se devem justamente ao seu tempo de execução que, embora não seja tão elevado na versão modificada, pode ser considerado inadequado por uma aplicação de reconhecimento que possua uma grande base de dados e com restrições em relação

⁵Valores obtidos em um AMD Opteron™ 280 2.4GHz

ao tempo de resposta. Neste caso, uma possível solução seria a utilização de técnicas de filtragem ou de indexação das imagens [27]. Mas, se a restrição for em relação à precisão dos resultados, o SAM se apresenta como uma ótima solução.

CAPÍTULO 5

CONCLUSÕES

Sendo o reconhecimento facial 3D uma área de pesquisa recente e, que promete revolucionar o reconhecimento facial, métodos robustos que explorem diferentes técnicas e apresentem bons resultados podem dar uma grande contribuição para este campo.

O *Simulated Annealing* aplicado ao reconhecimento facial 3D representa uma linha de pesquisa diferente da maior parte dos trabalhos da área. Esta abordagem apresentou bons resultados e mostrou-se como uma solução praticável ao problema, desde que sejam utilizadas medidas de erro apropriadas. Além disso, o SA é mais rápido que outros métodos que produzem resultados tão bons quanto os seus (e.g., algoritmo genético).

O ICP, embora produza resultados razoáveis, baseia-se em medidas pouco robustas e, portanto, pode ficar aquém dos resultados esperados. A principal vantagem do SA sobre o ICP, reside no fato de não precisar de supervisão e de um bom pré-alinhamento.

Como trabalho futuro para comprovar sua eficácia e aumentar sua aplicabilidade, sugere-se um estudo de técnicas para melhorar o desempenho computacional do SAM e uma análise de seu comportamento quando, ao menos uma das faces analisadas possuir expressão facial.

Para o problema de imagens com expressão facial, segundo [16, 17], existem ao menos quatro possíveis abordagens para lidar com o problema:

1. Concentrar o reconhecimento em áreas da face que são menos sujeitas às variações provocadas pela presença de expressão facial;
2. Construir uma *gallery* onde cada indivíduo nela cadastrado, possui várias imagens, cada uma delas contemplando uma expressão facial diferente;
3. Mapear as expressões faciais em modelos que podem ser aplicados nas imagens da *gallery* para simular expressões faciais;
4. Utilizar representações de face que sejam invariantes quanto à presença de expressões faciais, como em [10–12].

Para melhorias no desempenho computacional, a utilização de técnicas de paralelização em partes do algoritmo, sobretudo na avaliação de soluções e na busca pelo ponto mais próximo

(*KD-tree*), pode reduzir significativamente o seu tempo de execução. Além disso, a classificação das faces de acordo com suas características, pode reduzir o número de comparações necessárias para identificar uma face submetida. Ao invés de comparar a face submetida com todas as faces da *gallery*, a comparação é feita somente com aquelas que possuem as mesmas características (de acordo com a classificação) da face submetida.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Andrea F. Abate, Michele Nappi, Daniel Riccio, e Gabriele Sabatino. 3d face recognition using normal sphere and general fourier descriptor. *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pp. 1183–1186, Washington, DC, USA, 2006. IEEE Computer Society.
- [2] Lale Akarun, Berk Gökberk, e Albert Ali Salah. 3d face recognition for biometric applications. *13th European Signal Processing Conference (EUSIPCO 2005)*, Setembro de 2005. Disponível em citeseer.ist.psu.edu/742997.html.
- [3] Peter N. Belhumeur, Joao Hespanha, e David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.
- [4] O. R. P. Bellon, Silva L., C. C. Queirolo, e M. Drovetto Jr., S. A. Pamplona Segundo. 3d face image registration for face matching guided by the surface interpenetration measure. *ICIP*, pp. 2661–2664. IEEE, 2006. Disponível em <http://dblp.uni-trier.de/db/conf/icip/icip2006.html#BellonSQDP06>.
- [5] O. R. P. Bellon, L. Silva, e C. C. Queirolo. 3d face matching using the surface interpenetration measure. *ICIAP 2005, 13th International Conference on Image Analysis and Processing, 2005, Cagliari, Italy. Lecture Notes in Computer Science. Berlin : Springer-Verlag, 2005, v. 3617, pp. 1051-1058*, Setembro de 2005.
- [6] Samy Bengio e Johnny Mariéthoz. The expected performance curve: a new assessment measure for person authentication. *Proceedings of Odyssey 2004: The Speaker and Language Recognition Workshop*, 2004.
- [7] Paul J. Besl e Neil D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
- [8] Kevin W. Bowyer, Kyong Chang, e Patrick Flynn. A survey of 3d and multi-modal 3d+2d face recognition. Relatório técnico, Notre Dame Department of Computer Science and Engineering, Janeiro de 2004.
- [9] Kevin W. Bowyer, Kyong Chang, e Patrick Flynn. A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition. *Computer Vision and Image Understanding*, 101(1):1–15, 2006.
- [10] A. M. Bronstein, M. M. Bronstein, e R. Kimmel. Expression-invariant representations of faces. *IEEE Transactions on Image Processing*, 16(1):188–197, Janeiro de 2007.

- [11] A. M. Bronstein, M. M. Bronstein, e Ron Kimmel. Expression-invariant 3d face recognition. Joseph Kittler e Mark S. Nixon, editors, *Proc. Audio & Video-based Biometric Person Authentication (AVBPA)*, number 2688 in Lecture Notes in Comp. Science, pp. 62–69. Springer Verlag, 2003.
- [12] Alexander M. Bronstein, Michael M. Bronstein, e Ron Kimmel. Three-dimensional face recognition. *International Journal of Computer Vision*, 64(1):5–30, Agosto de 2005.
- [13] Lisa G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.
- [14] Kyong I. Chang, K. W. Bowyer, e P. J. Flynn. Effects on facial expression in 3D face recognition. *Proc. SPIE - Biometric Tech. for Human Identification*, volume 5779, pp. 132–143, 2005.
- [15] Kyong I. Chang, Kevin W. Bowyer, e Patrick J. Flynn. Face recognition using 2d and 3d facial data. *Proceedings of ACM Workshop in Multimodal User Authentication*, pp. 25–32, Santa Barbara, California, USA, Dezembro de 2003.
- [16] Kyong I. Chang, Kevin W. Bowyer, e Patrick J. Flynn. Multiple nose region matching for 3d face recognition under varying facial expression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10):1695–1700, Outubro de 2006.
- [17] Kyong I. Chang, Kevin W. Bowyer, e Patrick J. Flynn. Adaptive rigid multi-region selection for handling expression variation in 3d face recognition. *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pp. 157, Washington, DC, USA, 2005. IEEE Computer Society.
- [18] Yang Chen e Gérard Medioni. Object modelling by registration of multiple range images. *Image Vision Computing*, 10(3):145–155, 1992.
- [19] Kathryn A. Dowsland. Simulated annealing. *Modern heuristic techniques for combinatorial problems*, pp. 20–69. John Wiley & Sons, Inc., New York, NY, USA, 1993.
- [20] Martin A. Fischler e Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Readings in Computer Vision: Issues, Problems, Principles, and Paradigms*, pp. 726–740. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.
- [21] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, e M. Levoy. Geometrically stable sampling for the ICP algorithm. *3DIM '03: Proceedings of 4th International Conference on 3D Digital Imaging and Modeling*, pp. 260–267, Banff, Canada, 2003.

- [22] Gaile G. Gordon. Face recognition based on depth maps and surface curvature. *Geometric Methods in Computer Vision*, volume 1570 of *SPIE Proceedings*, pp. 234–247, Julho de 1991.
- [23] Lester Ingber. Simulated annealing: Practice versus theory. *Mathl. Comput. Modelling*, 18(11):29–57, 1993.
- [24] S. Kirkpatrick, C.D. Gelatt, Jr., e M.P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, Maio de 1983.
- [25] J. Kitler, M. Hatef, R. Duin, e J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, 1998.
- [26] X. Lu e A. K. Jain. Integrating range and texture information for 3D face recognition. *Proc. 7th IEEE Workshop on Applications of Computer Vision*, pp. 156–163, Breckenridge, CO, 2005.
- [27] X. Lu, A. K. Jain, e D. Colbry. Matching 2.5D face scans to 3D models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):31–43, 2006.
- [28] M Lundy e A Mees. Convergence of an annealing algorithm. *Math. Program.*, 34(1):111–124, 1986.
- [29] Aleix M. Martínez e Avinash C. Kak. Pca versus Ida. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):228–233, 2001.
- [30] P. Jonathon Phillips, Patrick J. Flynn, Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, e William Worek. Overview of the face recognition grand challenge. *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pp. 947–954, Washington, DC, USA, 2005. IEEE Computer Society.
- [31] V. J. Rayward-Smith, I. H. Osman, C. R. Reeves, e G. D. Smith. *Modern Heuristic Search Methods*. John Wiley & Sons Ltd., 1996.
- [32] S. Rusinkiewicz e M. Levoy. Efficient variants of the ICP algorithm. *Proceedings of 3rd International Conference on 3D Digital Imaging and Modeling*, pp. 145–152, 2001.
- [33] M. P. Segundo, O. R. P. Bellon, e L. Silva. Segmentação facial em imagens de profundidade. *XIX Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens, IV Workshop de Iniciação Científica*, 2006.
- [34] M. P. Segundo, C. C. Queirolo, O. R. P. Bellon, e L. Silva. Automatic 3d facial segmentation and landmark detection. *XIV International Conference on Image Analysis and Processing*, pp. 431–436, 2007.

- [35] Luciano Silva, Olga R.P. Bellon, e Kim L. Boyer. Computer vision and graphics for heritage preservation and digital archaeology. *Revista de Informática Teórica e Aplicada*, 11(1):9–32, Setembro de 2004.
- [36] Luciano Silva, Olga R.P. Bellon, e Kim L. Boyer. Precision range image registration using a robust surface interpenetration measure and enhanced genetic algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5):762–776, Maio de 2005.
- [37] Luciano Silva, Olga R.P. Bellon, e Kim L. Boyer. *Robust range image registration using genetic algorithms and the surface interpenetration measure*, volume 60. World Scientific Publishing Co. Pte. Ltd., 2005. Series in Machine Perception and Artificial Intelligence.
- [38] A. Thayananthan, B. Stenger, P. Torr, e R. Cipolla. Shape context and chamfer matching in cluttered scenes. *CVPR '03: Proceedings of IEEE Conference of Computer Vision and Pattern Recognition 2003*, volume 1, pp. 127–133, Madison, WI, USA, 2003.
- [39] P. H. S. Torr e A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78:138–156, 2000.
- [40] Maria Leonilde Rocha Varela e Rita Almeida Ribeiro. Utilização de simulated annealing em otimização difusa. *Revista de Investigação Operacional*, 21(2):205–231, Dezembro de 2001.
- [41] W. Zhao, R. Chellappa, P. J. Phillips, e A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003.

APÊNDICE A

COMPARAÇÃO DAS ABORDAGENS: ICP, GA E SA

A comparação aqui apresentada foi realizada em um pequeno número de imagens da base FRGC 3D *Face Database* v2.0. Foram utilizadas 50 imagens, todas sem expressão, de 25 indivíduos (2 imagens por indivíduo). O total de combinações foi de 1.225, sendo que 1.200 entre imagens de mesmos indivíduos e 25 entre imagens de indivíduos diferentes.

O *Simulated Annealing* (SA) aqui utilizado possui algumas modificações, são elas:

- Simplificação de algumas estruturas de dados e tradução dessas estruturas para C;
- Utilização de apenas um pré-alinhamento (centro de massa);
- Utilização de uma amostragem menor de pontos, apenas 0,6%, no alinhamento com o MSAC;
- MSAC com valor limiar de 3,0.

O *Iterative Closest Point* (ICP) utilizado corresponde a uma versão aprimorada, descrita em [18].

O algoritmo genético, *Genetic Algorithm* (GA), aqui utilizado, provido pela biblioteca GA-Lib¹, emprega o modelo *steady-state*. O algoritmo foi configurado com uma taxa de cruzamento de 90% e uma taxa de mutação de 15%, a população inicial é de 60 indivíduos e o número de gerações é de 100.

A Tabela A.1 faz uma comparação de tempo de execução entre as abordagens, para cada uma das possíveis segmentações.

Abordagem	Tempo Médio de Execução				
	Nariz	N. Elipse	N. Círculo	Olhos	Face
<i>Iterative Closest Point</i>	0,85s	1,33s	1,58s	9,02s	13,94s
<i>Genetic Algorithm</i>	9,63s	15,48s	18,50s	111,68s	171,45s
<i>Simulated Annealing</i>	0,52s	0,82s	1,00s	6,19s	9,84s

Tabela A.1: Comparação do tempo de alinhamento para os métodos ICP, GA e SA.

¹Disponível em: <http://lancet.mit.edu/ga/>

As Tabelas de A.2 a A.7 mostram os valores médios obtidos por cada abordagem, em cada segmentação para clientes e impostores.

Abordagem	Valor Médio de SIM para Clientes				
	Nariz	N. Elipse	N. Círculo	Olhos	Face
<i>Iterative Closest Point</i>	48,38%	44,63%	41,56%	33,39%	21,68%
<i>Genetic Algorithm</i>	51,33%	46,95%	45,51%	31,09%	25,25%
<i>Simulated Annealing</i>	48,07%	41,44%	40,39%	34,90%	24,53%

Tabela A.2: Comparação do valor médio de SIM para clientes dos métodos ICP, GA e SA.

Abordagem	Valor Médio de SIM para Impostores				
	Nariz	N. Elipse	N. Círculo	Olhos	Face
<i>Iterative Closest Point</i>	2,85%	1,81%	1,41%	1,51%	1,11%
<i>Genetic Algorithm</i>	9,42%	7,44%	6,88%	5,58%	3,95%
<i>Simulated Annealing</i>	5,72%	4,23%	3,73%	3,67%	2,62%

Tabela A.3: Comparação do valor médio de SIM para impostores dos métodos ICP, GA e SA.

Abordagem	Valor Médio de MSAC para Clientes				
	Nariz	N. Elipse	N. Círculo	Olhos	Face
<i>Iterative Closest Point</i>	0,52	0,52	0,52	0,56	0,66
<i>Genetic Algorithm</i>	0,58	0,61	0,58	0,70	0,82
<i>Simulated Annealing</i>	0,55	0,59	0,55	0,58	0,68

Tabela A.4: Comparação do valor médio de MSAC para clientes dos métodos ICP, GA e SA.

Abordagem	Valor Médio de MSE para Impostores				
	Nariz	N. Elipse	N. Círculo	Olhos	Face
<i>Iterative Closest Point</i>	1,41	1,54	1,72	1,71	1,86
<i>Genetic Algorithm</i>	1,66	1,75	1,78	1,90	2,02
<i>Simulated Annealing</i>	1,45	1,52	1,57	1,65	1,78

Tabela A.5: Comparação do valor médio de MSE para impostores dos métodos ICP, GA e SA.

Abordagem	Valor Médio de MSE para Clientes				
	Nariz	N. Elipse	N. Círculo	Olhos	Face
<i>Iterative Closest Point</i>	0,58	0,58	0,58	0,73	0,79
<i>Genetic Algorithm</i>	0,65	0,70	0,63	0,88	0,98
<i>Simulated Annealing</i>	0,61	0,68	0,61	0,74	0,81

Tabela A.6: Comparação do valor médio de MSE para clientes dos métodos ICP, GA e SA.

Abordagem	Valor Médio de MSE para Impostores				
	Nariz	N. Elipse	N. Círculo	Olhos	Face
<i>Iterative Closest Point</i>	1,99	2,16	2,37	2,76	3,01
<i>Genetic Algorithm</i>	2,56	2,77	2,90	3,51	3,90
<i>Simulated Annealing</i>	1,70	1,93	2,21	2,48	2,69

Tabela A.7: Comparação do valor médio de MSE para impostores dos métodos ICP, GA e SA.

A Tabela A.8 exhibe o valor de SIM para algumas combinações “especiais” das várias abordagens, para a segmentação da face. Estas combinações são: (1) o cliente que conseguiu o maior valor de SIM, ou seja, o alinhamento mais preciso; (2) o cliente com menor valor de SIM, normalmente, esta combinação seria falsamente rejeitada; (3) o impostor com maior valor de SIM, é com base neste valor que se devem estabelecer os limiares de sistemas biométricos que exigem FAR de 0%. Esta tabela também relaciona cada um destes alinhamentos à sua respectiva imagem, Figuras de A.1 a A.6. Para cada caso apresentado na tabela, o valor da SIM e o alinhamento de todas as abordagens são apresentados na figura correspondente.

Abordagem	Cliente maior SIM		Cliente menor SIM		Impostor maior SIM	
	SIM	Figura	SIM	Figura	SIM	Figura
ICP	46,15%	A.1	7,70%	A.2	6,25%	A.3
GA	46,16%	A.1	11,22%	A.4	11,46%	A.5
SA	47,21%	A.1	12,17%	A.6	10,83%	A.5

Tabela A.8: Comparação de alinhamentos para os métodos ICP, GA e SA.

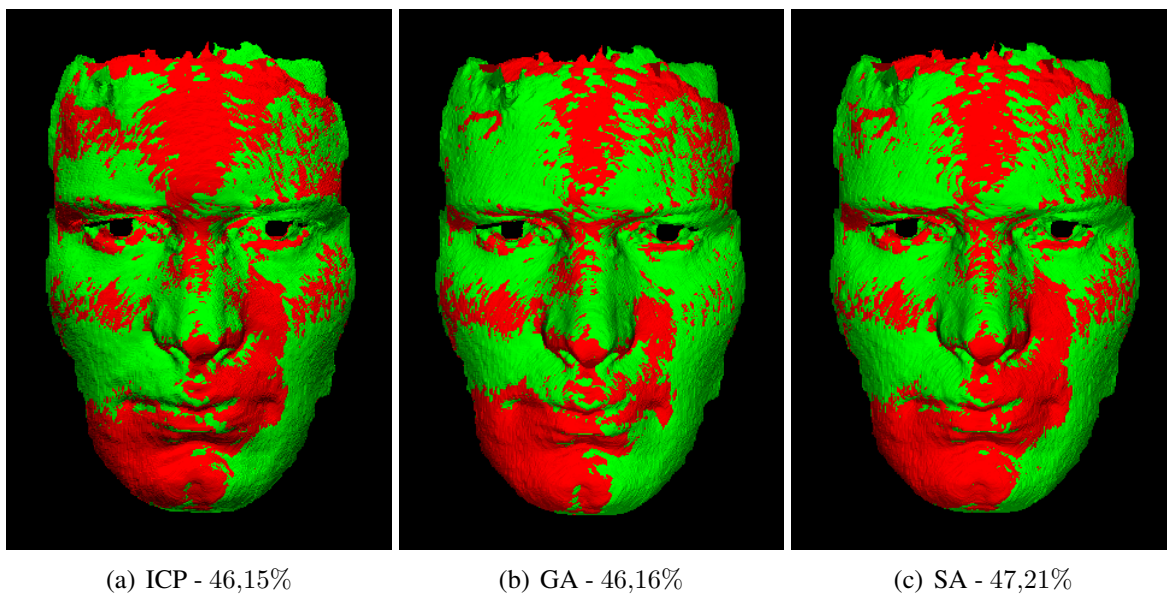


Figura A.1: Exemplo de alinhamento 1: (a) ICP, SIM 46,15%; (b) GA, SIM 46,16%; (c) SA, SIM 47,21%

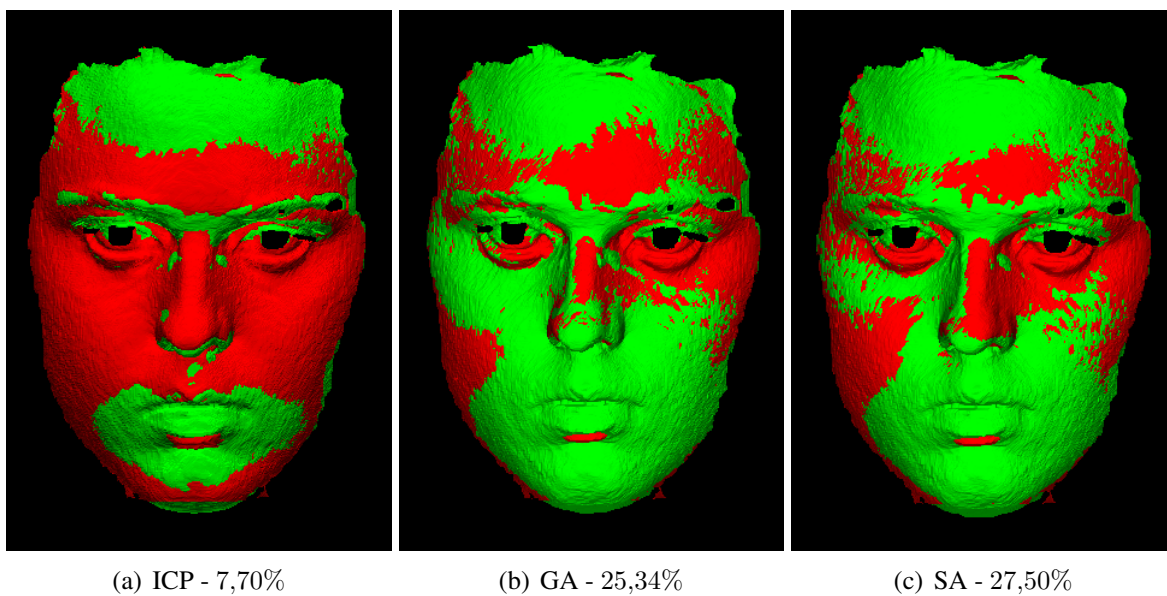


Figura A.2: Exemplo de alinhamento 2: (a) ICP, SIM 7,70%; (b) GA, SIM 25,34%; (c) SA, SIM 27,50%

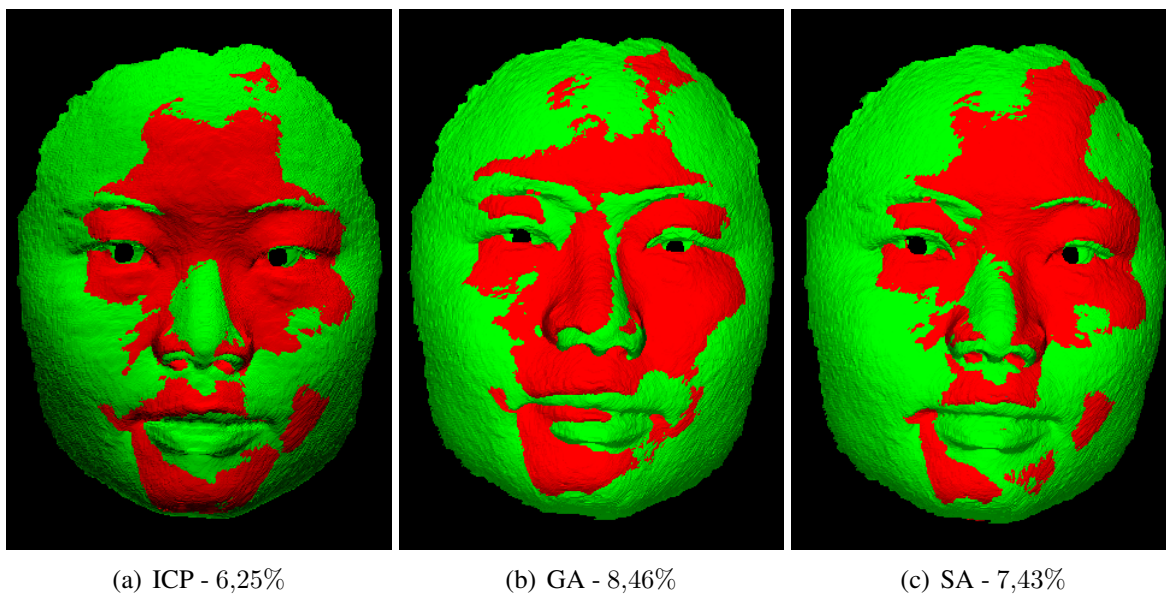


Figura A.3: Exemplo de alinhamento 3: (a) ICP, SIM 6,25%; (b) GA, SIM 8,46%; (c) SA, SIM 7,43%

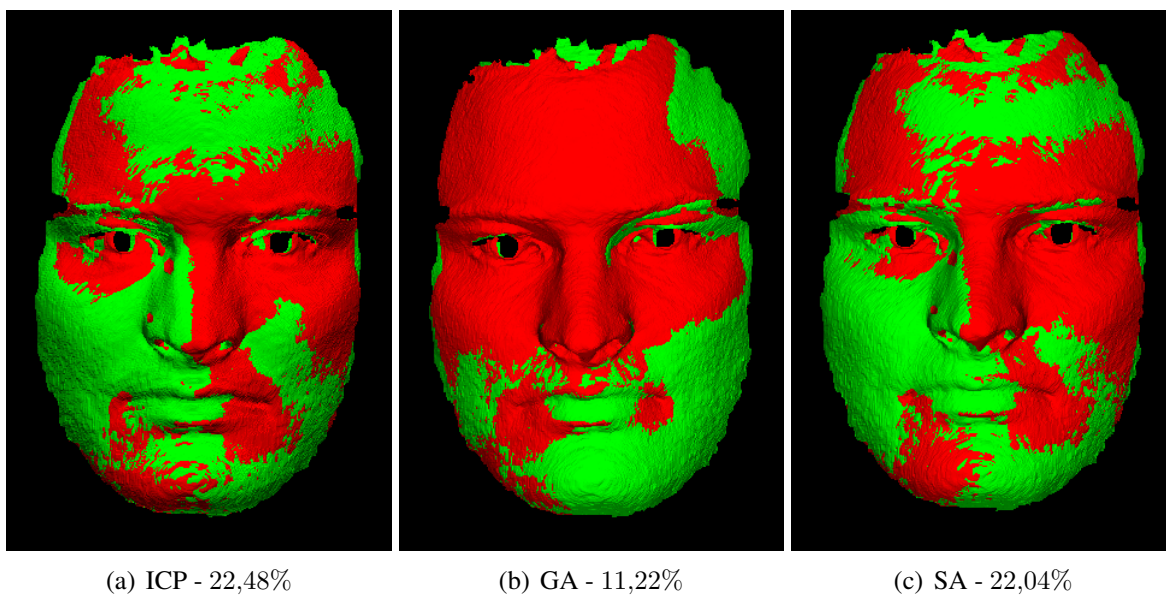


Figura A.4: Exemplo de alinhamento 4: (a) ICP, SIM 22,48%; (b) GA, SIM 11,22%; (c) SA, SIM 22,04%

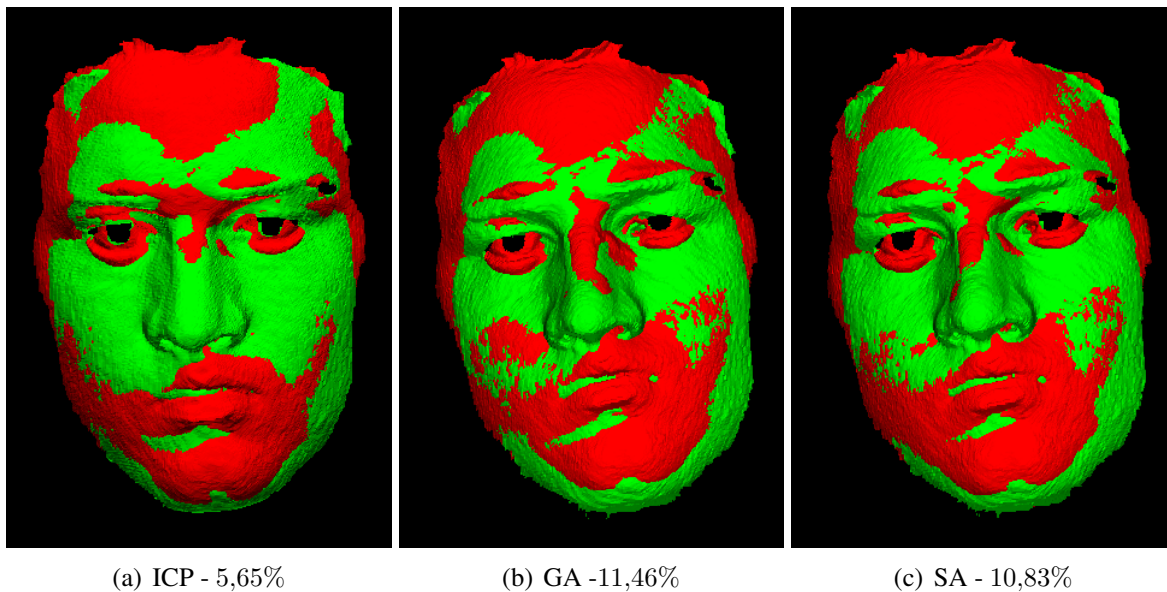


Figura A.5: Exemplo de alinhamento 5: (a) ICP, SIM 5,65%; (b) GA, SIM 11,46%; (c) SA, SIM 10,83%

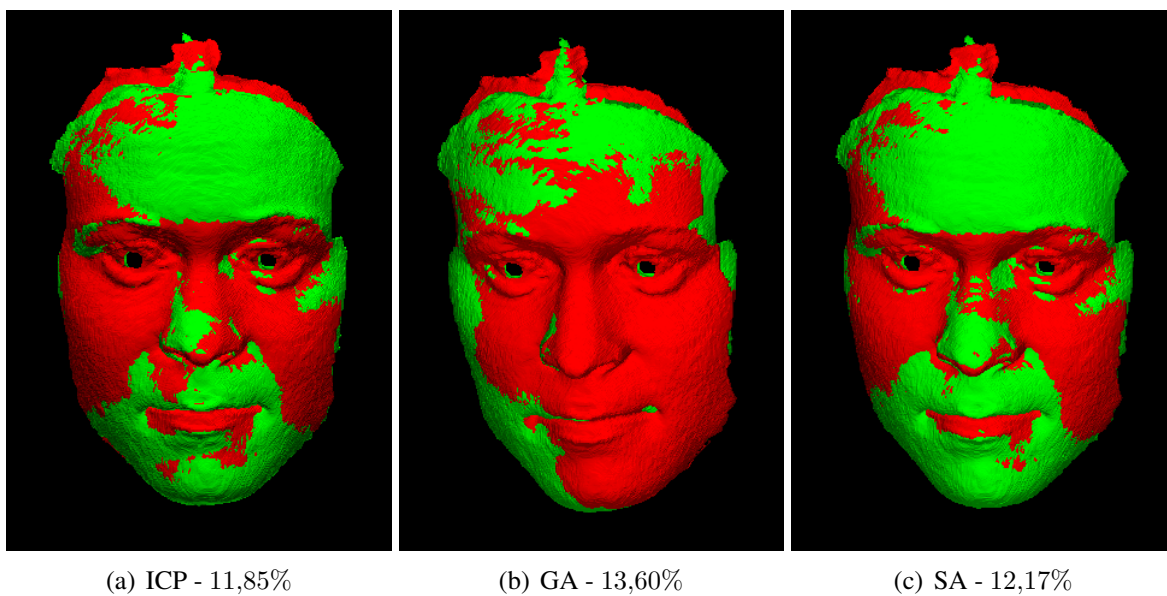


Figura A.6: Exemplo de alinhamento 6: (a) ICP, SIM 11,85%; (b) GA, SIM 13,60%; (c) SA, SIM 12,17%