

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
ELÉTRICA

DETECÇÃO FACIAL A PARTIR DE DIFERENÇAS DE COR

KELLY VINENTE DOS SANTOS

MANAUS
2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

**UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
ELÉTRICA**

KELLY VINENTE DOS SANTOS

DETECÇÃO FACIAL A PARTIR DE DIFERENÇAS DE COR

Dissertação apresentada ao
Programa de Pós-Graduação em
Engenharia Elétrica da
Universidade Federal do
Amazonas, como requisito parcial
para a obtenção do título de
Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Rogério Caetano – UFAM

**MANAUS
2008**

Ficha Catalográfica
(Catalogação na fonte realizada pela Biblioteca Central – UFAM)

Santos, Kelly Vinente dos

S237d

Detecção facial a partir de diferenças de cor / Kelly Vinente dos Santos. - Manaus: UFAM, 2008.

111 f.; il. color.

Dissertação (Mestrado em Engenharia Elétrica) — Universidade Federal do Amazonas, 2008.

Orientador: Prof. Dr. Rogério Caetano

1. Processamento de imagens 2. Informática – Método *skin locus* 3. Técnicas digitais 4. Morfologia matemática I. Caetano, Rogério II. Universidade Federal do Amazonas III. Título

CDU 004.932(043.3)

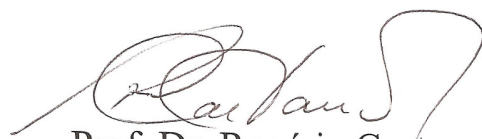
KELLY VINENTE DOS SANTOS

DETECÇÃO FACIAL A PARTIR DE DIFERENÇAS DE COR

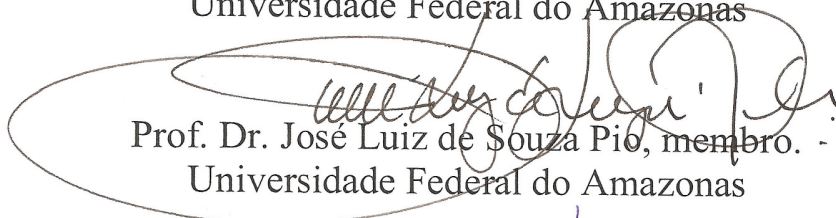
Dissertação apresentada ao
Programa de Pós-Graduação em
Engenharia Elétrica da
Universidade Federal do
Amazonas, como requisito parcial
para a obtenção do título de
Mestre em Engenharia Elétrica.

Aprovado em 28 de fevereiro de 2008.

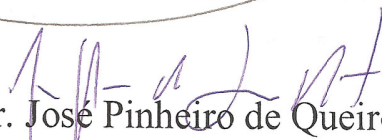
BANCA EXAMINADORA



Prof. Dr. Rogério Caetano, presidente.
Universidade Federal do Amazonas



Prof. Dr. José Luiz de Souza Pio, membro.
Universidade Federal do Amazonas



Prof. Dr. José Pinheiro de Queiroz Neto, membro.
Centro Federal de Educação Tecnológica do Amazonas

A minha família

AGRADECIMENTOS

A Deus acima de tudo.

A minha família, minha mãe e irmãos pelo apoio sincero durante todo este caminho.

Ao meu namorado Daniel Oliveira, pela motivação e ajuda.

Ao amigo Waldir Sabino pelas palavras de sabedoria.

Aos colegas que estiveram sempre presentes contribuindo para este trabalho.

A SUFRAMA e ao CETELI pelo suporte financeiro.

Ao coordenador e professores do curso de pós-graduação em engenharia elétrica da UFAM.

Aos professores que gentilmente aceitaram participar da banca.

A todos de uma forma geral que sempre me incentivaram a prosseguir.

A face é a alma do corpo.
Ludwig Wittgenstein

RESUMO

O ato de reconhecer é uma atividade rotineira realizada pelo cérebro humano ao longo de toda a existência do ser. É por meio de um sistema de reconhecimento natural que cada indivíduo identifica seus entes queridos, amigos e conhecidos em um grupo de pessoas. Esta habilidade inerente ao ser humano vem sendo, nas últimas décadas, fonte de inspiração para o desenvolvimento de sistemas computacionais. Tais tarefas são desenvolvidas para processar automaticamente imagens e vídeos digitais como a detecção de pele utilizada em grande escala em aplicações rotineiras das câmeras de vigilância, autenticação de faces e segurança de informação. Dentre as várias utilidades dessas aplicações, destaca-se nesta dissertação a detecção facial em imagens sob várias condições de iluminação e plano de fundo complexo por um algoritmo que usa o espaço de cor RGB. A intenção desta dissertação é investigar as regiões de uma dada imagem para identificar quais partes dela podem ser classificadas como pele e desprezar os outros elementos que aparecem nela. Foi realizado um levantamento do perfil comportamental dos pixels nos canais R, G e B das regiões de pele e comparados ao perfil dos outros elementos da imagem. A partir dessas análises foi proposto um método de classificação dos pixels de pele e usando técnicas baseadas em morfologia matemática foi realizada a extração das características faciais para validar a região de pele pertencente à face. Os resultados alcançados correspondem a uma média de acertos de 95% na classificação dos pixels de pele e de 90% na verificação de face nos bancos de imagens utilizados. A pesquisa revelou que o algoritmo proposto para classificar os pixels de pele é eficaz tanto em imagens de pessoas com características negras quanto caucasianas.

Palavras-chave: detecção de face, segmentação de cor de pele, morfologia matemática.

ABSTRACT

Recognition is a routine activity carried through by the human brain throughout all the existence of the being. It is through the natural recognition system that each one identifies his dear relativities, friends and known people in a group. This inherent ability to the human being comes being, in the last times, source of inspiration for the development of computational systems. Such tasks are developed to process digital images and videos automatically as detection of skin color used in wide scale in routine applications of the video surveillance, authentication of faces and security of information. Amongst several utilities of those applications, it is distinguished in this dissertation work the face detention in images under several conditions of illumination and complex background through an algorithm that uses the space of color RGB. The intention in this present work is to investigate the regions of a given image to identify which parts of it can be classified like skin and ignore the others elements that appear in it. It was realized a survey of the behavior profile of pixels in canals R, G and B of the regions of skin and it was compared with the profile of the others elements of the image. From these analyses, it was proposed a method to identify the region of skin color and through of morphology-based techniques it was extracted the face characteristics to become valid the region of skin color that belongs to the face. The reached results correspond to an average of rightness of 95% in classification of skin color pixels and 90% in verification of face of the images of the banks that were used in these applications. The research showed that the algorithm developed to classify the skin color pixels is efficient with images of Black and Caucasians people.

Key-words: detection of face, segmentation of skin color, morphology-based techniques

LISTA DE FIGURAS

Figura 1 - Exemplo de pré-processamento: <i>deblurring</i> – quando aplicado na imagem retira ruídos	19
Figura 2 - Segmentação: remoção das regiões que não são pele.....	19
Figura 3 - Rastreamento: identificação da face em uma sequência de quadros	20
Figura 4 - Extração de Atributos: identificação das regiões dos olhos, nariz e boca	20
Figura 5 - Classificação: comparação da face com outras faces armazenadas num banco de dados.....	21
Figura 6 – Resultado proposto pelo algoritmo de SINGH et al (2003).....	26
Figura 7 – Representação cônica do sistema HSI.....	29
Figura 8 – Relação entre os sistemas RGB e HSI. Triângulo de cores HSI dentro do cubo de cores RGB	30
Figura 9 – Esquema de classificação de pele usado por LIN (2007)	30
Figura 10 – Resultado apresentado por LIN (2007) para classificação dos pixels de pele.	31
Figura 11 – Resultado da detecção de face de LIN (2007)	31
Figura 12 – <i>Skin Locus</i> de Störring (1999).....	33
Figura 13 – Exemplo de resultado obtido por Störring (1999)	34
Figura 14 – <i>Skin locus</i> obtido por Martinkauppi (2000)	36
Figura 15 - <i>Skin locus</i> obtidos com duas câmeras diferentes: Sony e Winnov	37
Figura 16 – <i>Skin locus</i> de 16 amostras de pele do banco de imagens A	39
Figura 17 – Primeiro exemplo de <i>skin locus</i> com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	39
Figura 18 – Segundo exemplo de <i>skin locus</i> com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	40
Figura 19 – Terceiro exemplo de <i>skin locus</i> com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	41
Figura 20 – Primeiro exemplo de <i>skin locus</i> com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	42
Figura 21 – Segundo exemplo de <i>skin locus</i> com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	42
Figura 22 – Terceiro exemplo de <i>skin locus</i> com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	43
Figura 24 – <i>Skin locus</i> de 16 amostras de pele do banco de imagens B.....	44
Figura 25 – Primeiro exemplo de <i>skin locus</i> com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	45
Figura 26 – Segundo exemplo de <i>skin locus</i> com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	45

Figura 27 – Terceiro exemplo de <i>skin locus</i> com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	46
Figura 28 – Primeiro exemplo de <i>skin locus</i> com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	46
Figura 29 – Segundo exemplo de <i>skin locus</i> com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	47
Figura 30 – Terceiro exemplo de <i>skin locus</i> com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	47
Figura 31 - <i>Skin locus</i> de 16 amostras de pele a partir de imagens de <i>webcam</i>	48
Figura 32 - Primeiro exemplo de <i>skin locus</i> sem perda de informação de pele para <i>webcam</i> . (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	49
Figura 33 - Segundo exemplo de <i>skin locus</i> sem perda de informação de pele para <i>webcam</i> . (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	49
Figura 34 - Primeiro exemplo de <i>skin locus</i> com perda de informação de pele para <i>webcam</i> . (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	50
Figura 35 - Segundo exemplo de <i>skin locus</i> com perda de informação de pele para <i>webcam</i> . (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do <i>skin locus</i> e (c) gráfico do <i>skin locus</i> da imagem original	50
Figura 36 – Resultado da aplicação do <i>skin locus</i> em uma imagem do banco A	51
Figura 37 – Valores de pixels de fundo classificados como pele	52
Figura 38 – <i>Skin locus</i> da Figura 37. Localização dos pixels de fundo classificados como pele no <i>skin locus</i>	52
Figura 39 – Resultado da aplicação do <i>skin locus</i> em uma imagem do banco B	53
Figura 40 - Resultado da aplicação do <i>skin locus</i> em uma imagem do banco B com indivíduo negro	53
Figura 41 – Imagens faciais apresentando valores de R, G e B para a pele. (a) Pele negra; (b) Pele caucasiana	60
Figura 42 – Seleção da linha da imagem para verificar o perfil dos valores R, G e B (indivíduo de pele caucasiana)	62
Figura 43 – Seleção da linha da imagem para verificar o perfil dos valores R, G e B (indivíduo de pele negra)	63
Figura 44 – Imagens originais do (a) Banco A e do (b) Banco B	64
Figura 45 – Máscaras contendo as regiões de pele utilizada como referência para comparação dos resultados. (a) Máscara da imagem do Banco A e (b) máscara da imagem do Banco B	65
Figura 46 - Resultado após utilização do algoritmo de modelagem do <i>skin locus</i> (MARTINKAUPPI, 2000)	65
Figura 47 – Resultado utilizando a diferença entre os canais R e G	65
Figura 48 – Classificação dos pixels da imagem utilizando a diferença RB	66
Figura 49 - Classificação dos pixels da imagem utilizando a diferença GB	67
Figura 50 – Classificação dos pixels da imagem utilizando a diferença BG ((a) indivíduo de pele caucasiana e (b) indivíduo negro)	67
Figura 51 - Etapas da segmentação proposta	68
Figura 52 - Exemplo de histograma da diferença RG	69

Figura 53 - Resultados da aplicação do <i>Skin Locus</i> e da diferença RGB nas imagens do Banco A	71
Figura 54 - Resultados da aplicação do <i>Skin Locus</i> e da diferença RGB nas imagens do Banco B	71
Figura 55 – Metodologia aplicada para eliminação de objetos não pertencentes à face	72
Figura 56 - Somas horizontal e vertical da imagem binária obtida do resultado da diferença de canais R, G e B	73
Figura 57 – Estabelecimento dos limites vertical e horizontal a partir das respectivas somas	74
Figura 58 – Extração da região candidata a face	74
Figura 59 – Somas horizontal e vertical da imagem binária obtida do resultado da diferença de canais R, G e B, e a imagem segmentada após o uso dos limites obtidos a partir dessas somas. Figura do banco de imagens B.	75
Figura 60 – Etapas de um sistema genérico de identificação/processamento de face.....	77
Figura 61 – A busca por vales na imagem. (a) Imagem original e (b) o perfil dos níveis de cinza da imagem na coluna selecionada	79
Figura 62 – Símbolo da operação lógica XOR, da álgebra booleana, seguida da operação de negação.	80
Figura 63 – Resultado das operações morfológicas. (a) Imagem original e (f) perfil do nível de cinza da imagem vale final, com os pixels de íris destacados	80
Figura 64 – Detecção facial de Han et al (2000).	81
Figura 65 - Detecção facial de Perlibakas (2003).....	82
Figura 66 - Primeira fase da metodologia proposta: detecção da região dos olhos	83
Figura 67 – Resultado da diferença RG com limiar global (banco de imagens A).....	84
Figura 68 – Resultado da diferença RG com limiar global (banco de imagens A).....	84
Figura 69 – Resultado da diferença RG com limiar global (banco de imagens B)	85
Figura 70 – Resultado da diferença RG com limiar global (banco de imagens B)	85
Figura 71 – Relação entre os objetos presentes na imagem	86
Figura 72 – Comportamento dos pixels na região dos olhos.....	87
Figura 73 – Estrutura do olho humano	87
Figura 74 – Diagrama de blocos da utilização de operadores morfológicos.....	88
Figura 75 – Contorno do nariz conectado aos olhos. O limiar aplicado foi calculado pelo método de Otsu pelo MatLab	89
Figura 76 – Resultado final a aplicação morfológica de fechamento usando a imagem de entrada em (a) níveis de cinza e (b) a soma dos resultados nos canais R, G e B	89
Figura 77 – Imagem binária resultado da aplicação de limiar maior que 100.....	90
Figura 78 – Classificação dos objetos para obtenção das sobrancelhas, nariz e boca.....	91
Figura 79– Resultados dos processos de detecção dos olhos, sobrancelhas e boca para o banco de imagens A	92
Figura 80 – Resultados dos processos de detecção dos olhos, sobrancelhas e boca para o banco de imagens A.....	93
Figura 81 – Resultados dos processos de detecção dos olhos, nariz e boca para o banco de imagens B	94
Figura 82 – Resultados dos processos de detecção dos olhos, sobrancelhas, nariz e boca para o banco de imagens B	95
Figura 83 – Exemplos de imagens do Banco de Imagens A com diferentes condições de iluminação e complexidade de fundo: (a) e (b) sob condições de iluminação em ambiente fechado e (c) sob condições de iluminação de ambiente aberto.....	102
Figura 84 - Exemplos de imagens do Banco de Imagens A que apresentam oclusões: (a) Barba; (b) Olhos fechados e (c) oclusão do olho direito	102
Figura 85 – Desenho de pessoas do Banco de Imagens A	102

Figura 86 – Exemplos de imagens do Banco de Imagens B: (a) Face inclinada; (b) Indivíduo caucasiano e (c) indivíduo com tom de pele semelhante ao objeto de fundo da cena.....	103
Figura 87 – Exemplos de imagens de <i>webcam</i> utilizadas nos experimentos.	104

LISTA DE TABELAS

Tabela 1 – Taxas de reconhecimento obtidas utilizando-se o algoritmo de detecção de pixels de pele proposto por Martinkauppi (2000)	51
Tabela 2 – Taxas de reconhecimento de pele e fundo utilizando-se a máscara binária da diferença entre os canais RG	64
Tabela 3 – Comparação das taxas de reconhecimento para as Figura 46 e Figura 47	66
Tabela 4 – Taxas de reconhecimento utilizando o <i>skin locus</i> e a máscara baseada nas diferenças dos canais R, G e B	70

LISTA DE QUADROS

Quadro 1 - Temperaturas de cor da fonte de iluminação simuladas para obtenção do <i>skin locus</i>	34
Quadro 2 - Temperaturas de cor da fonte de iluminação simuladas por Martinkauppi (2000)	35
Quadro 3 - Valores de normalização dos pixels no plano RGB.	37
Quadro 4 - Valores da mediana para as distribuições r e g dos bancos de imagens A e B	55
Quadro 5 - Faixas de valores propostas e suas respectivas porcentagens de reconhecimento (banco de imagens A)	56
Quadro 6 – Faixas de valores propostas e suas respectivas porcentagens de reconhecimento (banco de imagens B)	57

LISTA DE GRÁFICOS

Gráfico 1 – Histogramas (a) e (b) das máscaras de pele do banco de imagens A.....	54
Gráfico 2 - Histogramas (a) e (b) das máscaras de pele do banco de imagens B.....	54
Gráfico 3 - Desempenho de cada faixa de pixels vermelho e verde experimentadas para classificação dos pixels de pele (banco de imagens A)	56
Gráfico 4 – Desempenho de cada faixa de pixels vermelho e verde experimentadas para classificação dos pixels de pele (banco de imagens B)	57
Gráfico 5 - Perfil do comportamento R, G e B na face do indivíduo da Figura 41(a)	60
Gráfico 6 – Perfil do comportamento R, G e B na face do indivíduo da Figura 41 (b)	61
Gráfico 7 – Comportamento dos níveis R, G e B da linha selecionada na Figura 42	62
Gráfico 8 – Comportamento dos níveis R, G e B da linha selecionada na Figura 43	63

LISTA DE SIGLAS

CCD	Dispositivo de carga acoplada
FP	Falsos positivos
HSI	Modelo de cores formado pelos componentes <i>Hue</i> (Tonalidade), <i>Saturation</i> (Saturação) e <i>Intensity</i> (Intensidade)
HSV	Modelo de cores formada pelos componentes <i>Hue</i> (Tonalidade), <i>Saturation</i> (Saturação) e <i>Value</i> (Valor)
JPEG	<i>Joint Picture Expert Group</i> - algoritmo criado para comprimir imagens
KODAK	Empresa multinacional dedicada ao design, produção e comercialização de equipamentos fotográficos
NCC	Coordenadas de cor normalizadas
RGB	Modelo de cores aditivas formado pelo vermelho (<i>Red</i>), verde (<i>Green</i>) e azul (<i>Blue</i>)
USB	<i>Universal Serial Bus</i>
VP	Verdadeiro Positivo
YCbCr	Modelo de cores que considera luminância (Y), croma azul (Cb) e croma vermelho (Cr)
<i>Webcam</i>	Câmera de vídeo projetada para PCs

SUMÁRIO

INTRODUÇÃO.....	18
1 DETECÇÃO DE PELE UTILIZANDO O MÉTODO DO <i>SKIN LOCUS</i>	24
1.1 Introdução	24
1.2 Espaço de cor YCbCr	25
1.3 Espaço de cor HSI	28
1.4 Espaço de cor RGB	32
1.5 Classificação de pixels de pele usando <i>skin locus</i>	38
1.5.1 <i>Skin locus</i> do Banco de imagens A.....	38
1.5.2 <i>Skin locus</i> do Banco de imagens B	44
1.5.3 <i>Skin locus</i> das imagens de webcam	48
1.5.4 Resultados da classificação de pele por <i>skin locus</i>	50
1.5.5 Novas faixas de valores <i>r</i> e <i>g</i> e resultados	54
2 DETECÇÃO FACIAL A PARTIR DE DIFERENÇAS EM RGB	59
2.1 Introdução	59
2.2 Comportamento dos pixels no espaço de cor RGB	60
2.3 Detecção por diferenças em RGB	67
2.4 Resultados da detecção de pixels de pele	70
3 VERIFICAÇÃO DA FACE	76
3.1 Introdução	76
3.2 Extração das características da face por morfologia matemática.....	78
3.3 Extração das características da face por diferenças em RGB.....	82
3.4 Resultados da detecção de face	91
4 CONCLUSÃO E TRABALHOS FUTUROS	96
REFERÊNCIAS	98
ANEXO A	101
ANEXO B	110

INTRODUÇÃO

O uso das características da face para detecção automática é uma tarefa difícil, pois a aparência facial tende a mudar a todo tempo e ser bastante suscetível a fatores emocionais.

A partir de imagens estáticas e vídeos, a detecção e identificação da face vêm emergindo há décadas como uma área de pesquisa que possibilita numerosas aplicações. Segundo Zhao e Chellappa (2006) essas aplicações são: entretenimentos (vídeo games, realidade virtual, programas de treinamento, interação entre homem e máquina, álbuns de foto da família, etc.), industriais (*smart cards*, passaportes, licença para dirigir, etc.), segurança da informação (acesso a telefones celulares, *login* em computadores pessoais, acesso a internet, etc.), de imposição de lei (mecanismos de segurança, investigações de imagens pornográficas na internet, anti-furto, etc.) e em áreas emergentes (ferramentas de estudo psicológico).

Embora, os seres humanos consigam reconhecer faces de forma natural por meio de um aprendizado próprio (OLIVEIRA, 1989), fazer com que um computador reconheça faces não é uma tarefa trivial por diversas razões.

Umas das principais dificuldades é a localização da face sem nenhum conhecimento prévio de suas escalas, orientações (ereta, inclinada), posicionamento (perfil, frontal) e existência de oclusões.

As condições de iluminação também tendem a somar dificuldades no processo de detecção facial, assim como a aquisição da imagem que não depende somente da pessoa, mas, também, do tipo de dispositivo utilizado e do seu campo de visão.

Devido às dificuldades envolvidas, o reconhecimento facial tem sido estudado por meio de diversas abordagens, variando desde sistemas de reconhecimento de padrões por

redes neurais (ROWLEY, 1998; SUNG e POGGIO, 1998) até varreduras infravermelhas de pontos da face (WILDER, 1996).

Um sistema genérico de reconhecimento computacional de faces pode ser dividido nas seguintes etapas (ZHAO e CHELLAPPA, 2006):

- Pré-Processamento: utilizam-se algoritmos de processamento de imagens para minimizar as interferências (Figura 1) que podem ocorrer na imagem (TURK e PENTLAND, 1991).



Figura 1 - Exemplo de pré-processamento: *deblurring*¹ – quando aplicado na imagem retira ruídos

- Segmentação e Detecção: etapa que se identificam as regiões de pele (Figura 2) para separá-las dos demais elementos na imagem (LI e JAIN, 2004).



Figura 2 - Segmentação: remoção das regiões que não são pele

- Rastreamento: (usado apenas na abordagem dinâmica, ou seja, com vídeos e seqüências de imagens) fase onde se rastreia a face (Figura 3) em uma seqüência de quadros (SAXE e FOULDS, 1996; ZHOU, 2003; DORNAIKA e AHLBERG, 2004).

¹ Retira os borrões das imagens degradadas.



Figura 3 - Rastreamento: identificação da face em uma sequência de quadros

Fonte: Retirado de <<http://herkules.oulu.fi/isbn9514267885/html/i1030753.html>>. Acessado em: 15 maio 2007.

- Extração de Atributos: muitos atributos podem ser extraídos (Figura 4), desde distâncias entre regiões faciais importantes até medidas obtidas por ferramentas matemáticas poderosas como as transformações de Fourier e as *wavelets*.

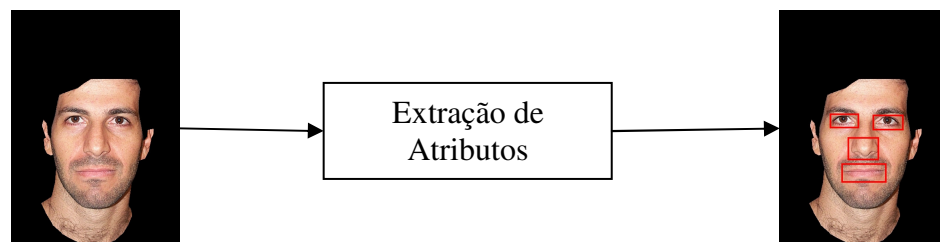


Figura 4 - Extração de Atributos: identificação das regiões dos olhos, nariz e boca

- Classificação: geralmente é feita por algoritmos que comparam o vetor de atributos da face a ser reconhecida com os vetores de atributos de faces já conhecidas (Figura 5), armazenadas em uma base de dados (JOLLIFFE, 1996).

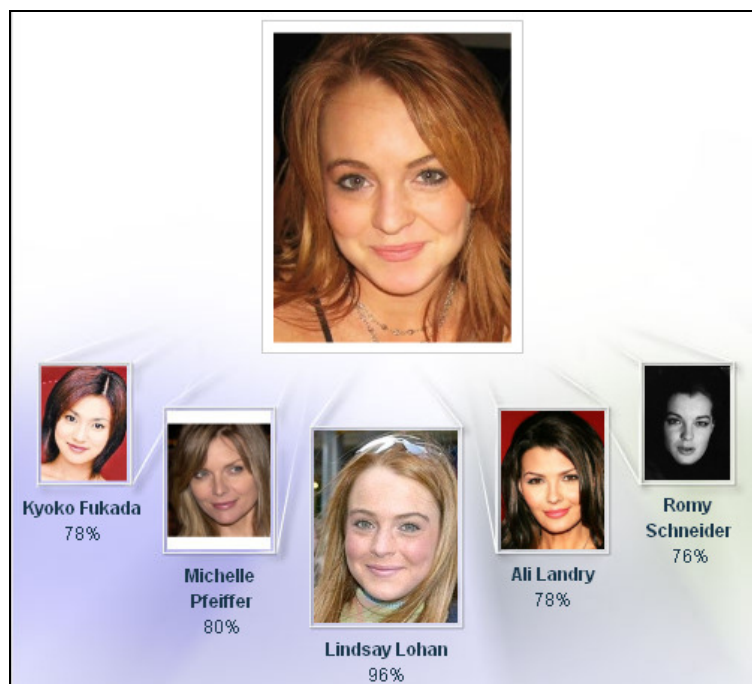


Figura 5 - Classificação: comparação da face com outras faces armazenadas num banco de dados

Para construir um sistema, totalmente automatizado e eficiente, que analise a informação de face contida em imagens, é necessário o desenvolvimento de um algoritmo robusto de detecção facial. Para isso muitas técnicas têm sido desenvolvidas com o objetivo de detectar faces em uma imagem simples. Os métodos de detecção analisados estão baseados em:

- **Conhecimento:** representam técnicas que codificam algum conhecimento sobre a face (YANG, 2002). As técnicas baseadas em *templates* (SCLAROFF e ISIDORO, 1998; COOTES e TAYLOR, 2001; HOU *et al*, 2001) encontram-se nesse grupo, pois necessitam de conhecimento do objeto para serem gerados;
- **Características invariantes:** são técnicas que identificam características da face independente da orientação (BELHUMEUR, 1997; MARTINKAUPPI, 2000);
- **Aparência:** não necessitam de conhecimento prévio da face a ser detectada. No treinamento do sistema são utilizadas várias faces para que o sistema aprenda ou codifique o que é necessário para detectá-las. Exemplo dessa abordagem são redes neurais e *eigenfaces* (TURK e PENTLAND, 1991);

Entre eles, os métodos fundamentados em características apresentam uma boa solução para detecção de faces em diferentes posições e tamanhos.

Dentre as características mais utilizadas para auxiliar a identificação de faces está a cor, a qual é muito utilizada devido sua eficiência computacional e ao seu alto poder discriminatório.

Esta dissertação propõe um algoritmo de detecção facial baseado na identificação de tons de pele e de pontos característicos como olhos e boca de uma imagem digital de alta qualidade. Essa detecção visa identificar todas as regiões da imagem que contenham face, independente da posição, orientação e condições de iluminação, mas levando em consideração a classificação dos pixels coloridos como pixels da face em uma imagem com fundo complexo e, mesmo assim, distinguir a face do fundo.

O estudo dessa dissertação baseia-se em algoritmos de região em espaço de cor para detecção da pele, portanto são analisados os espaços de cor mais utilizados na literatura: YCbCr, HSV e RGB.

O método proposto nesta dissertação foi implementado em MatLab e utiliza a análise de cor por ser muito eficiente na detecção de faces inclinadas ou em diferentes escalas, pois essas mudanças dificilmente afetam a informação de cor, quando observadas sob uma iluminação uniforme e porque é uma característica que tende a reduzir significativamente o número de candidatos a face, diminuindo os falsos positivos.

Contudo, somente a informação de cor não é suficiente para se realizar a detecção de faces, pois em algumas imagens é difícil distinguir faces de outros objetos da cena, como braços, madeira e roupas com cores similares (MARTINKAUPPI, 2000). Assim, faz-se uso, também, de outras técnicas para verificação da região selecionada como face, por exemplo, a detecção dos olhos por morfologia matemática.

A avaliação do método proposto por esta dissertação utilizou dois bancos de imagens obtidos na internet, denominados Banco de Imagens A e Banco de Imagens B (Anexo A). As imagens de dos bancos apresentam resolução espacial de 896 x 592 pixels e 640 x 480 pixels, respectivamente. Foram utilizadas 160 imagens do Banco de Imagens A e 150 do Banco de Imagens B.

Do ponto de vista estrutural, o trabalho apresenta seis capítulos. O Capítulo I apresenta os métodos de classificação de pele utilizando o espaço de cor RGB a partir do seu *skin locus*² e os passos fundamentais para criação do classificador descritos na literatura, além da aplicação deste em imagens de alta qualidade.

O Capítulo II aborda os espaços de cor YCbCr e HSV, considerados para a proposição de um classificador de pele, destacando outras técnicas de classificação na detecção pele, necessárias para a construção de um algoritmo de detecção de face em uma imagem.

No terceiro Capítulo é proposto um novo método de detecção dos pixels de pele baseado na combinação das diferenças das informações dos canais do espaço de cor RGB.

No Capítulo IV é proposta uma metodologia para validação da região de face baseada no método de diferenças de canais RG e identificação de características presentes na face (olhos, nariz e boca).

O último Capítulo (VI) está destinado às conclusões obtidas após a concepção, prototipação e avaliação da abordagem proposta, onde apresenta-se, também, possíveis sugestões, possibilidades de extensão e trabalhos futuros.

² Local onde os pixels de pele tendem a se agrupar no plano de coordenadas de cor normalizadas.

1 DETECÇÃO DE PELE UTILIZANDO O MÉTODO DO *SKIN LOCUS*

1.1 Introdução

A localização e detecção da face humana são sempre os primeiros passos em aplicações como vídeo segurança, interface entre humanos e computadores, reconhecimento de faces e gerenciamento de dados de imagens, são também pré-requisitos para as análises de expressões faciais.

Para detecção de faces existem vários algoritmos incluindo os baseados em cor de pele. A cor é uma característica importante das faces e seu uso tem grandes vantagens, como rapidez de processamento. A cor é invariável sob certas condições de iluminação.

O estudo de classificação de cor de pele ganhou crescente atenção nas últimas décadas em pesquisas com imagens baseadas em conteúdo. Por exemplo, a habilidade de localizar objetos na imagem como a face, pode ser explorada por codificação de imagem, edição, indexação ou outra proposta de interatividade.

Há três maneiras de classificar um algoritmo de detecção de pele em imagens estáticas: baseado em estatísticas, em regiões no espaço de cor ou na correção de cor dessa região.

Os algoritmos baseados em regiões no espaço de cor separam os pixels de pele dos outros objetos na imagem. A região onde se encontram os valores dos pixels de pele é determinada a partir de um conjunto de imagens, contendo apenas regiões de pele sob diferentes condições de iluminações.

A região, no espaço de cor, que contém os valores de pixels de pele pode ser modelada de diferentes maneiras como: determinação de valores limiares ou tabelas de consultas. O formato dessa região pode variar de simples para complexa.

Logo, para esse tipo de algoritmo é preciso ter primeiro um conjunto de treinamento para depois obter a faixa de valores pertencentes às regiões de pele.

Um dos principais problemas dessa abordagem é a determinação da faixa de valores dos pixels de pele, porque muitas vezes classifica-se como pixels de pele, pixels que apresentam cor semelhante à pele.

A correção de cor é utilizada por muitos algoritmos de detecção de pele que visam tratar cores deslocadas devido a reflexão da luz ou variação de iluminação. Nesse tipo de algoritmo é feito a compensação de branco, mas para isso é necessário conhecer as características do dispositivo de aquisição de imagens.

Neste capítulo é apresentado o estudo sobre o *skin locus* para os espaços de cor RGB, YCbCr e HSI encontrados na literatura. Para cada espaço de cor são mostradas metodologias utilizadas para detecção facial, bem como a definição de cada espaço de cor.

Assim, este capítulo aborda os tipos de espaços utilizados na literatura e na pesquisa para a detecção facial a partir do reconhecimento de tons de pele, tais como o YCbCr, espaço de cor que utiliza crominâncias azul e vermelha, bem como a luminância, o HSI, espaço de cor que utiliza a intensidade de cor, matiz e saturação, e o RGB, espaço de cor que separa os plano de cor vermelho, verde e azul. Cada espaço de cor apresenta suas vantagens em relação à separação das características da pele e são utilizados sob diferentes técnicas pelos autores.

Foi dado enfoque ao espaço de cor RGB, utilizando-se o *skin locus* e avaliando-se seus resultados para os bancos de imagens utilizados.

1.2 Espaço de cor YCbCr

O espaço de cor YCbCr é muito utilizado em vídeo digital e em compressão de imagens no formato JPEG. Neste formato, a informação de cor é representada pela luminância

(Y), e as componentes de cromaticidade Cb e Cr. Onde Cb representa a cromaticidade de azul e Cr, a de vermelho.

A conversão do espaço de cor RGB para YCbCr é dada por:

$$Y = 0,299R + 0,587G + 0,114B \quad (1)$$

$$Cb = -0,1687R - 0,3313G + 0,5B + 128 \quad (2)$$

$$Cr = 0,5R - 0,4187G - 0,0813B + 128 \quad (3)$$

A simplicidade de conversão e facilidade de separação da luminância torna esse espaço de cor muito atrativo para a modelagem dos pixels de pele.

Utilizando os espaços de cor RGB, YCbCr e HSI, SINGH *et al* (2003) propôs um algoritmo que identifica as regiões de pele em cada um desses espaços e unifica todos os resultados em uma imagem final, aplicando operadores morfológicos para obter toda a região da face (Figura 6).

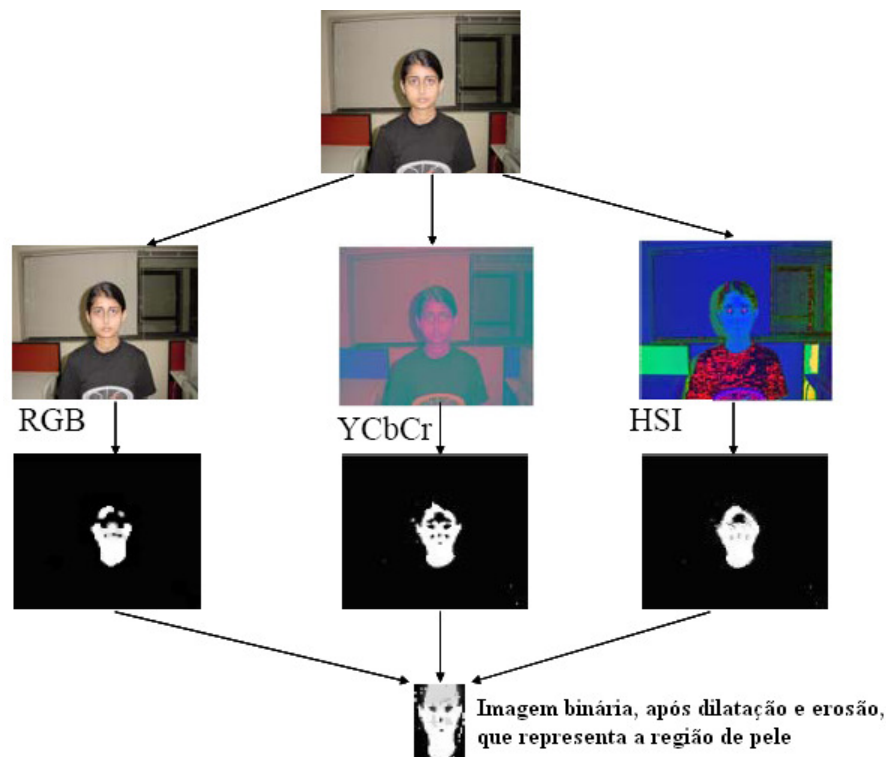


Figura 6 – Resultado proposto pelo algoritmo de SINGH *et al* (2003)

Fonte: SINGH *et al* (2003)

Em cada um desses espaços de cor foram estabelecidos limiares para determinar se um pixel é classificado como pixel de pele ou não.

Para o espaço de cor RGB, fez-se a transformação em NCC para estabelecer os limites de pele de acordo com Sanjay Kr. Singh *et al* (1997). Utilizando histogramas, também, estabeleceu-se limites para os espaços de cor YCbCr e HSI (H – matiz, S – saturação e I – intensidade).

No espaço YCbCr levou-se em consideração apenas as componentes Cb e Cr como Cah e Ngan (1999). Da mesma forma, no espaço HSI considerou-se apenas as componentes H e S (KJELDSOON e KENDER, 1996).

Utilizou-se a teoria de conjuntos para a união dos resultados de cada espaço de cor, aplicando-lhes a idéia básica do diagrama de Venn (RUSKEY e WESTON, 2005) da teoria de conjuntos.

A resposta individual em cada espaço delimita a região da face, porém, a união das respostas promove um resultado mais preciso quando combinado com as operações morfológicas de abertura e fechamento.

O processo de verificação da face utilizado foi a triangulação dos objetos mais escuros da imagem. Essa metodologia prevê um triângulo entre os olhos e a boca. De acordo com SINGH *et al* (2003), a distância Euclidiana entre os olhos é cerca de 90-110% a distância Euclidiana entre os olhos e a boca.

Após a obtenção do triângulo que pode ser isósceles (face frontal) ou não (face em perfil), são aplicadas regras de obtenção das bordas da potencial face.

De acordo com os resultados apresentados por SINGH *et al* (2003), não foram detectados 3% das faces quando o método foi executado em 1100 imagens.

Da mesma forma, BAO (2005) realizou a segmentação da cor de pele em imagens com fundo complexo determinando limiares, utilizou também a componente Y. Os limiares

foram obtidos nos histogramas das imagens de pessoas, de diferentes raças, obtidas na internet. Por meio desses limiares, um pixel será classificado como pele se:

$$Y1 < Y < Y2 \quad (4)$$

$$Cb1 < Cb < Cb2 \quad (5)$$

$$Cr1 < Cr < Cr2 \quad (6)$$

Os ruídos são removidos com o uso de operadores morfológicos de abertura. Em seguida são obtidas as bordas de cada região de pele detectada na imagem e, por meio de lógica Fuzzy, se valida a região como face se forem encontrados olhos e boca.

1.3 Espaço de cor HSI

O espaço de cor HSI é muito utilizado porque apresenta a intensidade de brilho (I) independente, o que facilita na resolução do problema da variação de iluminação do espaço RGB.

O espaço HSI é similar ao modo como os humanos tendem a perceber as cores. Ele define o espaço de cor em termos de 3 componentes:

- Matiz (Hue) - o tipo de cor. Exemplo: vermelho, azul ou amarelo.

Faixa: 0 a 360.

Faixa normalizada: 0 a 100%.

- Saturação (Saturation) - a pureza da cor, o nível de diluição da cor pura pela luz branca.

Faixa: 0 a 100%.

Uma saturação baixa da cor significa maior acinzentamento na imagem e maior desbotamento da cor.

- Intensidade (Intensity) ou Valor (Value) – brilho da cor, nível de luminância.

Faixa: 0 a 100%.

O modelo HSI foi criado em 1978 por Alvy Ray Smith. Esse modelo pode ser representado por um cone (Figura 7). Nessa representação, a matiz (H) é definida no sentido da seção circular do cone. A saturação (S) é representada pela distância do centro da seção circular do cone em direção a borda e a Intensidade (I) é a distância do ponto final do cone em direção ao centro da seção circular.

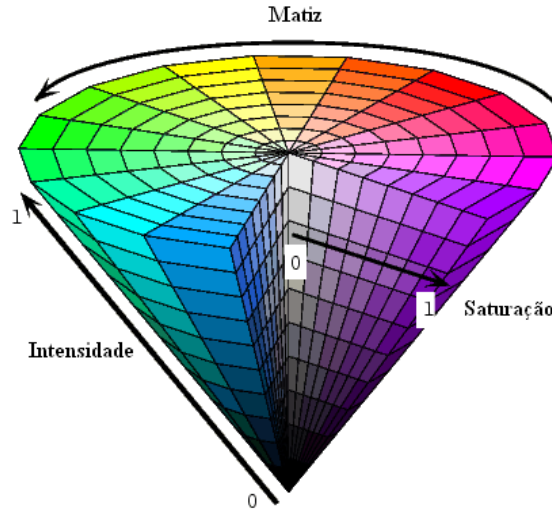


Figura 7 – Representação cônica do sistema HSI

O espaço de cor HSI é uma transformação não-linear do espaço RGB, e pode ser transformado a partir das coordenadas RGB.

Dado um pixel de cor definido por RGB, onde R, G e B são valores entre 0.0 e 1.0. Um equivalente de cor HSI pode ser determinado por uma série de fórmulas.

Fazendo MAX igual ao máximo dos valores de RGB e MIN, o menor dos valores, as fórmulas podem ser escritas como:

$$H = \begin{cases} \left(0 + \frac{G - B}{MAX - MIN}\right) \times 60, & \text{se } R = MAX \\ \left(2 + \frac{B - R}{MAX - MIN}\right) \times 60, & \text{se } G = MAX \\ \left(4 + \frac{R - G}{MAX - MIN}\right) \times 60, & \text{se } B = MAX \end{cases} \quad (7)$$

$$S = \frac{MAX - MIN}{MAX} \quad (8)$$

$$I = \text{MAX} \quad (9)$$

Se $S=0$, então a cor encontra-se ao longo da linha central de cinzas, e ela não tem matiz, a coordenada angular não tem sentido.

Se $\text{MAX}=0$ (isto é, $I=0$), então S é indefinido.

Se $I=0$, então a cor é puramente preto, não tendo matiz ou saturação. Assim, o diagrama cônico converge a um único ponto onde as coordenadas do ângulo e do raio não têm sentido.

Este modelo pode ser representado pela forma triangular bidimensional dentro do cubo formado pelo modelo RGB, Figura 8.

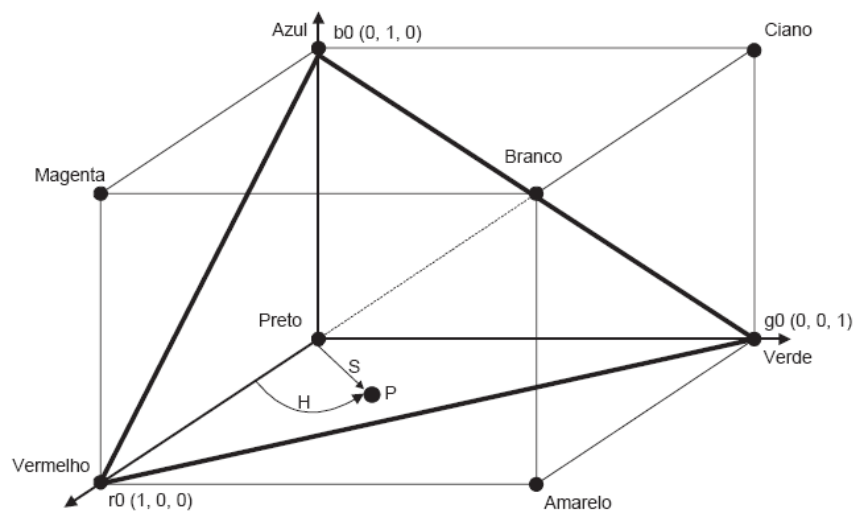


Figura 8 – Relação entre os sistemas RGB e HSI. Triângulo de cores HSI dentro do cubo de cores RGB

A classificação dos pixels de pele proposta por LIN (2007) segue o esquema da Figura 9.

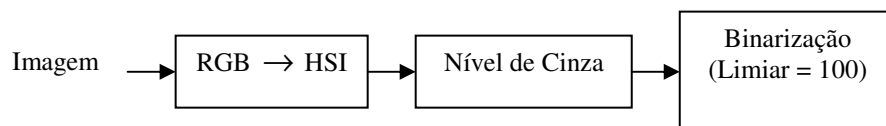


Figura 9 – Esquema de classificação de pele usado por LIN (2007)

Logo após a binarização da imagem, ela é rastreada para se obter os pixels com vizinhança N4. De acordo com Gonzalez (2007), um pixel, na coordenada (x,y) , tem vizinhança de 4 quando os pixels das coordenadas $(x+1,y)$, $(x-1,y)$, $(x,y+1)$ e $(x,y-1)$ também

possuem o mesmo valor. Esses objetos são rotulados e depois é identificado o centro de cada um deles.

Para verificar se as regiões rotuladas são faces, é utilizada a heurística de extração do triângulo isósceles, Figura 11, correspondente às distâncias entre os olhos e a boca. Faz-se uso de uma rede neural para a verificação da face.

Esse algoritmo foi aplicado por LIN (2007) em 1000 imagens de faces sob 200 condições de iluminação e o resultado apresentado foi de 98% de sucesso na detecção de face.

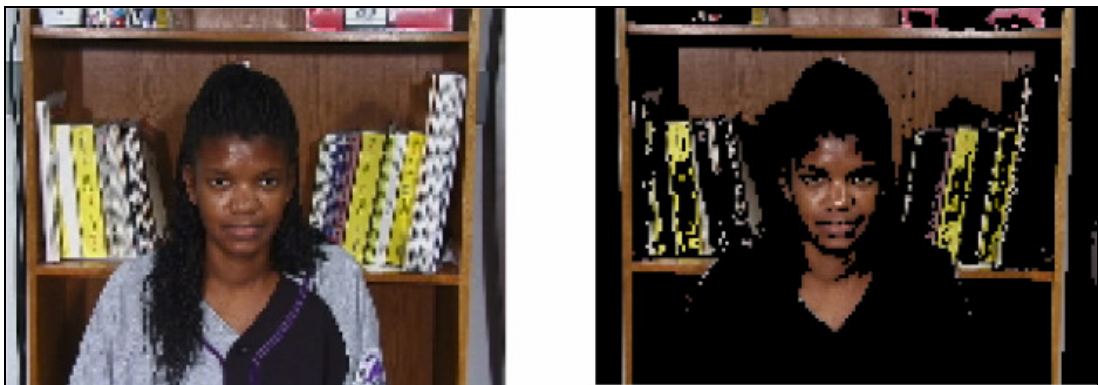


Figura 10 – Resultado apresentado por LIN (2007) para classificação dos pixels de pele
Fonte: LIN (2007)



Figura 11 – Resultado da detecção de face de LIN (2007)
Fonte: LIN (2007)

1.4 Espaço de cor RGB

Dentre os estudos feitos sobre a modelagem de cor de pele em imagem no espaço de cor RGB, Störring (1999) destaca-se pela análise da concentração dos pixels de pele sob diversas condições de iluminação, no plano de cromaticidade ³*rg*, segundo a fórmula, na qual:

$$r = \frac{R}{R + G + B} \quad (10)$$

e

$$g = \frac{G}{R + G + B} \quad (11)$$

O cálculo da componente *b* se dá pela equação 12:

$$b = 1 - r - g \quad (12)$$

De acordo com este método, para saber onde se concentram os pixels de pele no plano de cor RGB, é necessário o conhecimento de parâmetros da câmera como: espectro de luz de incidência, sensibilidade espectral do sensor da câmera e a transmitância espectral dos filtros em frente ao sensor.

A pele humana é tratada como um material refletor e pela análise de sua refletância pode-se propor um método para modelá-la.

Em sua pesquisa, Störring (1999) analisou o espectro de refletância de luz da pele desde o tipo caucasiano até o negro. Seu principal objetivo foi analisar o processo de formação da imagem sob diferentes condições de iluminação. Ele observou a aparência da cor de pele quando variava a temperatura de cor da fonte de iluminação.

Segundo este modelo, é possível estimar a corrente temperatura de cor da fonte de iluminação e ajustar corretamente o balanço de branco da câmera.

³ Refere-se à qualidade de uma cor em relação ao seu matiz dominante.

Os resultados mostraram que, para cada tipo de temperatura de cor da fonte de iluminação, os pixels de pele se concentravam em determinada região do espaço de cromaticidade rg como apresentado na Figura 12. Para essas concentrações de pixels de pele ele atribuiu o termo *skin locus*.

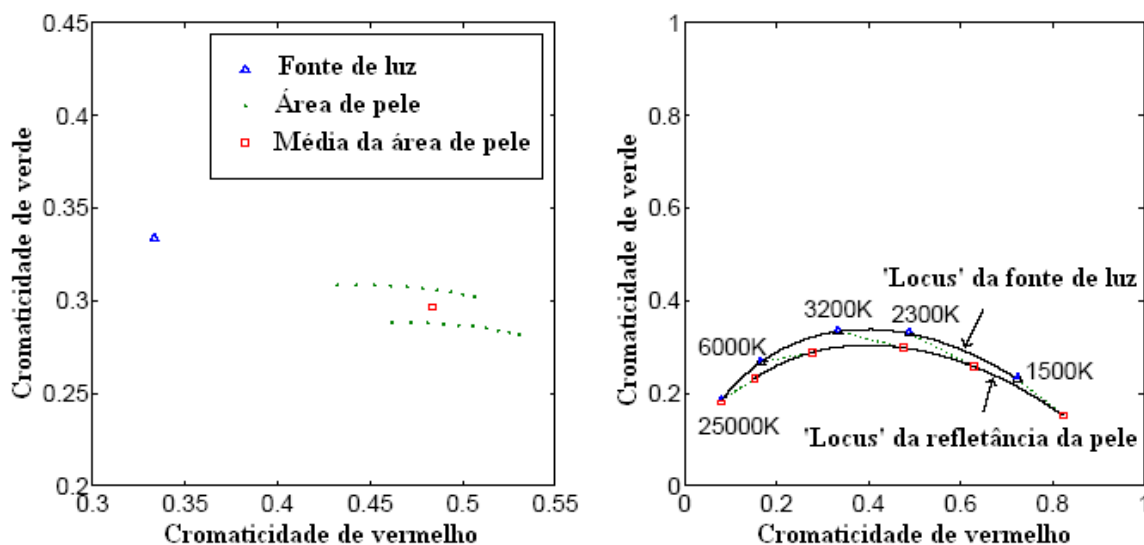


Figura 12 – Skin Locus de Störring (1999)
Fonte: Störring (1999)

A unidade de temperatura de cor da fonte de iluminação é Kelvin (K). Nos experimentos de Störring (1999) foram utilizadas quatro lâmpadas fluorescentes como fonte de iluminação: Philips TLD 927 (2760K), TLD 940 (3870K), TLD 950 (5190K) e TLD 965 (6900K).

Essas lâmpadas foram utilizadas para produzir as temperaturas de cor correspondentes ao Quadro 1.

Temperatura de cor	Tipo de iluminação
1500 K	Luz de vela
2300 K	Lâmpada incandescente
3200 K	Pôr-do-sol
6000 K	Luz do dia, ao meio-dia
25000 K	Luz do dia

Quadro 1 - Temperaturas de cor da fonte de iluminação simuladas para obtenção do *skin locus*

As imagens foram adquiridas em um ambiente fechado, com fundo preto, de pessoas características da Ásia, Camarões, Dinamarca, Grécia, Índia, Irã e Letônia.

Classificando-se os pixels por meio do comportamento característico dos pixels de pele, Störring (1999) obteve resultados como o mostrado na Figura 13.

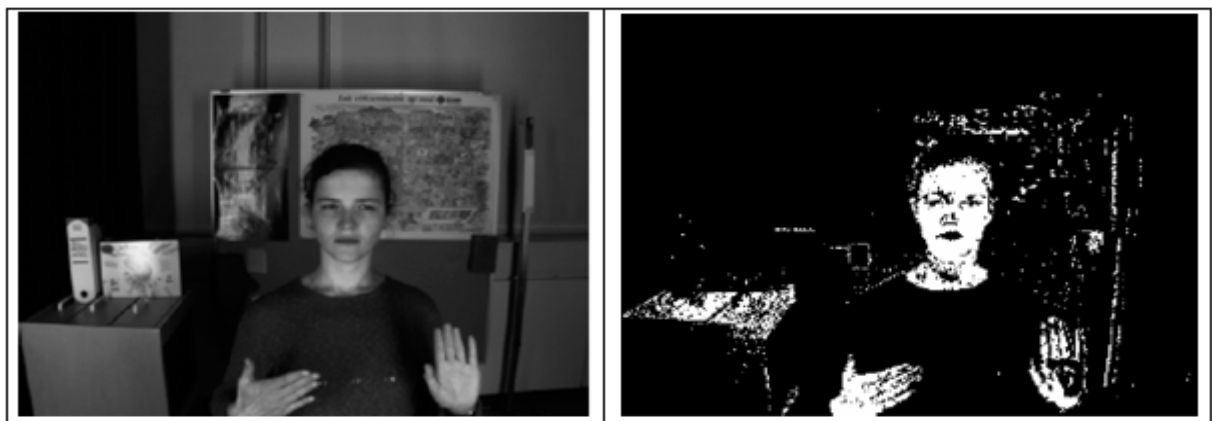


Figura 13 – Exemplo de resultado obtido por Störring (1999)
Fonte: Störring (1999)

Em seu trabalho, Störring afirma que o modelo pode ser usado com sucesso, pois quando utilizado com sete pessoas de diferentes cores de pele, sob quatro diferentes fontes de iluminação, pode-se identificar as regiões de pele nas imagens da câmera utilizada. Ressalta também que é preciso conhecer o espectro da fonte de iluminação e suas características.

Baseado nos resultados do experimento de Störring (1999), Martinkauppi (2000) propôs uma forma de modelar o *skin locus* por meio de polinômios.

Para obter todas as possibilidades de iluminação sobre a pele, ele simulou as seguintes iluminações em uma sala escura:

Temperatura de cor	Tipo de iluminação
2300 K	Luz do dia no horizonte
2866 K	Lâmpada incandescente
4000 K	Pôr-do-sol
6500 K	Luz do dia

Quadro 2 - Temperaturas de cor da fonte de iluminação simuladas por Martinkauppi (2000)

Para cada iluminação, foi feito o balanceamento de branco na câmera. Foram utilizadas 16 condições (4 de iluminação e 4 de calibração) para cada imagem de face capturada, sendo essas faces caucasianas e asiáticas.

As imagens foram recortadas, deixando apenas as regiões de pele. Foi realizada a normalização da imagem, ou seja, transformou-se o espaço RGB em coordenadas de cor normalizadas (NCC). O autor justifica o uso apenas das componentes de cromaticidade r e g afirmando que, a componente b de muitas câmeras fotográficas apresenta pouca sensibilidade ou pode conter muitos ruídos.

O *skin locus* das regiões de pele apresentado por Martinkauppi (2000), Figura 14, apresentou resultado semelhante ao de Störring (1999).

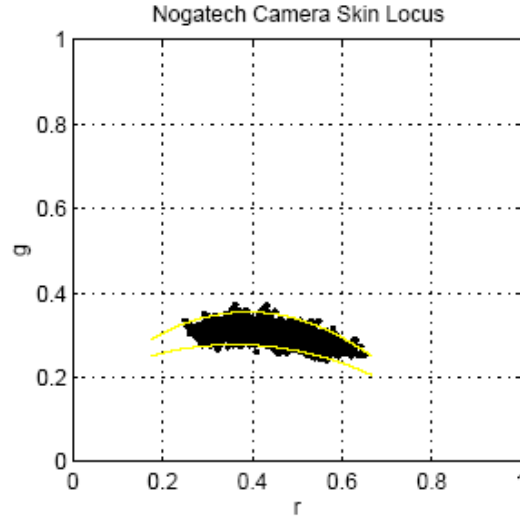


Figura 14 – Skin locus obtido por Martinkauppi (2000)
Fonte: Martinkauppi (2000)

O *skin locus* da Figura 14 foi modelado por um polinômio superior e um inferior (curvas em amarelo na figura) para classificação dos pixels que se encontram nessa região como pixels de pele.

A estimação dos mínimos quadrados foi utilizada para definir os valores dos coeficientes dos polinômios superior e inferior do *skin locus*.

Coeficientes da borda superior: $a_s = -1,3767$, $b_s = 1,0743$ e $c_s = 0,1452$.

Coeficientes da borda inferior: $a_i = -0,776$, $b_i = 0,5601$ e $c_i = 0,1766$.

Os pixels são classificados como pele ou não, de acordo com:

$$S = \begin{cases} 1 & \text{se } (g < Q_+) \& (g > Q_-) \& (W > 0,0004) \\ 0 & \text{caso contrário} \end{cases} \quad (13)$$

Onde:

$$Q_+ = a_s r^2 + b_s r + c_s \quad (14)$$

$$Q_- = a_i r^2 + b_i r + c_i \quad (15)$$

$$W = (r - 0,33)^2 + (g - 0,33)^2 \quad (16)$$

O conjunto de valores de r utilizados é $[0,2, 0,6]$.

Nota-se que o uso de coordenadas de cor normalizadas transforma pixels com a mesma intensidade em ambos os canais de RGB, para um mesmo valor, igual a 0,33, nas coordenadas de cromaticidade (Quadro 3).

R	G	B	R+G+B	$\frac{R}{R+G+B}$	$\frac{G}{R+G+B}$	$\frac{B}{R+G+B}$
255	255	255	765	0,33	0,33	0,33
10	10	10	30	0,33	0,33	0,33

Quadro 3 - Valores de normalização dos pixels no plano RGB.

Para evitar classificar pixels brancos ou escuros como pele, Martinkauppi (2000) propõe a exclusão dos pixels pertencentes ao círculo (W) de centro 0,33 e raio 0,02.

O algoritmo de Martinkauppi (2000) é usado para identificar, inicialmente, pixels pertencentes à pele de uma pessoa e na sequência é utilizado outro, de rastreamento, para identificar a região da face.

As imagens utilizadas nos experimentos de Martinkauppi (2000) foram obtidas por uma CCD USB webcam Nogatech. Segundo o autor, o uso de diferentes câmeras produz *skin locus* diferente (Figura 15) do mostrado na Figura 14. “A câmera e suas propriedades determinam o *skin locus*, como indicado pelo *loci* de quatro câmeras. Entretanto, algumas regiões são comuns para todas, mais notavelmente a região do tom de pele.” (MARTINKAUPPI, 2000).

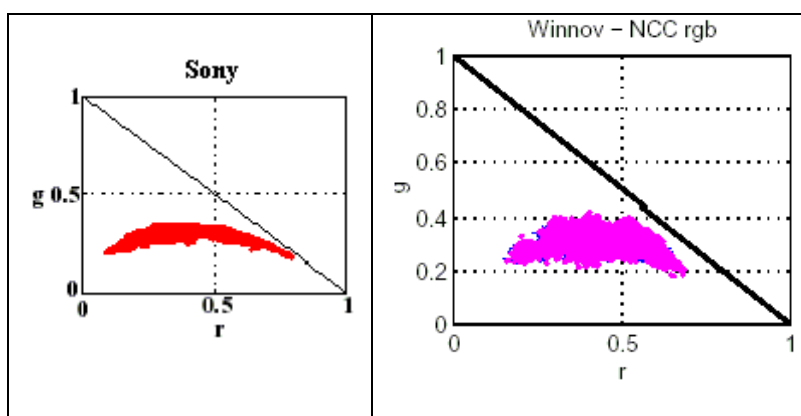


Figura 15 - *Skin locus* obtidos com duas câmeras diferentes: Sony e Winnov

Fonte: Retirado de <<http://herkules.oulu.fi/isbn9514267885/html/i1030753.html>>. Acessado em: 15 maio 2007.

1.5 Classificação de pixels de pele usando *skin locus*

Na etapa de classificação dos pixels de pele fez-se uso dos passos fundamentais apresentados na literatura (MARTINKAUPPI, 2000), para a criação de um modelo de classificação de cor de pele. Enumerados como:

- a) Escolha de um conjunto de imagens;
- b) Obtenção das regiões de pele das imagens escolhidas;
- c) Escolha da faixa de valores de r e g à qual a pele pertence.

A descrição dos passos a e b, bem como as ferramentas utilizadas, foram abordadas no capítulo 1. O resultado do passo b, máscaras de pele, serve para a etapa de treinamento do algoritmo, pois serão detectados os valores r e g que mais ocorrem nas regiões de pele.

Em seguida são construídos os histogramas de r e g . Não foi utilizada a componente b por ser considerada com pouca informação de pele na literatura.

1.5.1 *Skin locus* do Banco de imagens A

Utilizando-se a modelagem proposta por Martinkauppi (2000) nas máscaras de pele para o banco de imagens A (Anexo A), observou-se que o *skin locus* produzido (Figura 16) era diferente do que esse autor apresentou. Em sua proposta, Martinkauppi (2000) apresenta o formato que deveria ter o *skin locus* (Figura 14) quando submetido a 16 condições de iluminação diferentes.

Usando a mesma quantidade de amostras, 16 máscaras de pele, construiu-se o *skin locus* para o banco de imagens A (Figura 16). De acordo com sua disposição no plano de cromaticidade, concluiu-se que o ajuste dos polinômios delimitadores proposto por Martinkauppi (2000) não seria apropriado devido a ampla distribuição dos pixels de pele.

Para a Figura 17 houve perda apenas dos pixels que se localizavam no círculo, representando uma porcentagem muito pequena.

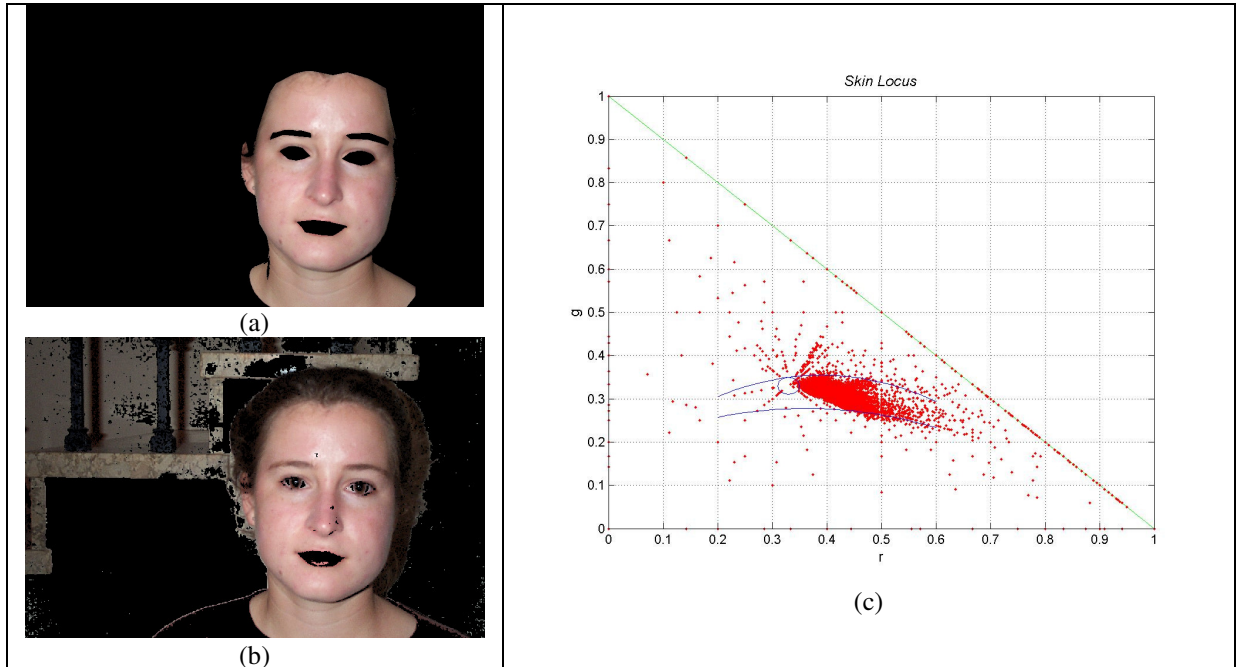


Figura 18 – Segundo exemplo de *skin locus* com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

Apesar de mostrar um espalhamento dos pixels de pele no plano de cromaticidade, a Figura 18 apresentou ínfimas perdas de informação.

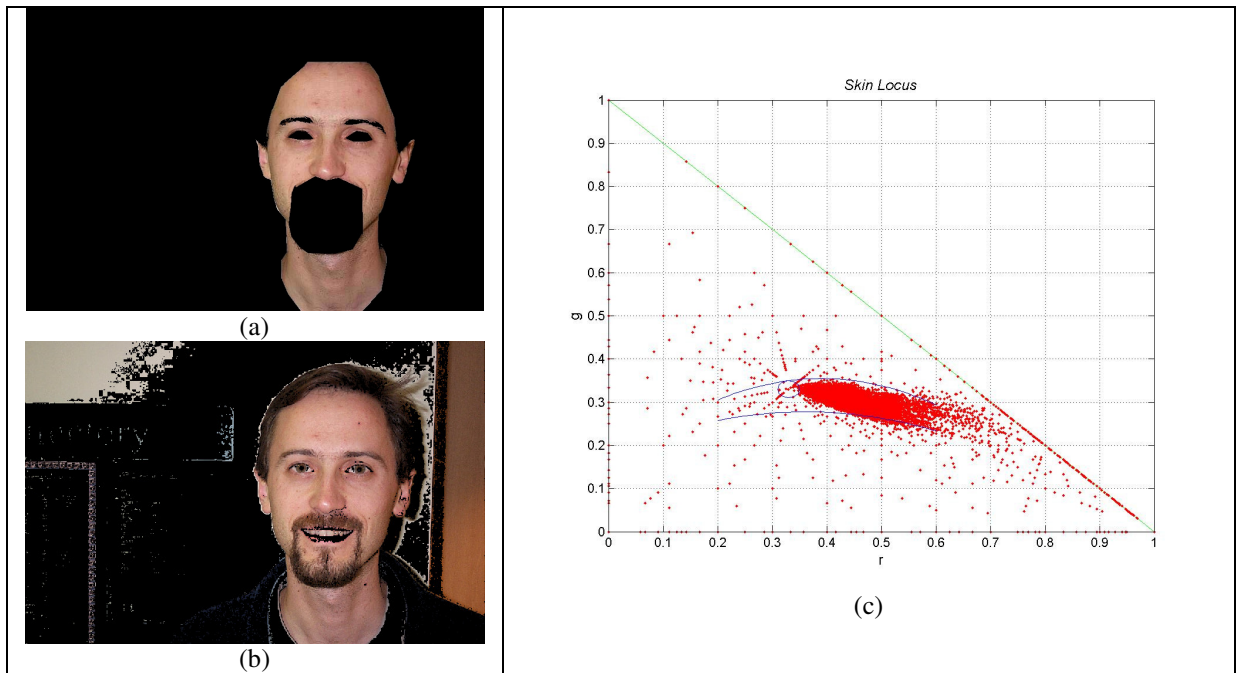


Figura 19 – Terceiro exemplo de *skin locus* com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

Semelhante à Figura 18, a Figura 19 apresentou ligeira dispersão dos pixels sendo que a grande maioria situou-se entre os polinômios de modelagem.

- Resultados com perda de informação de pele

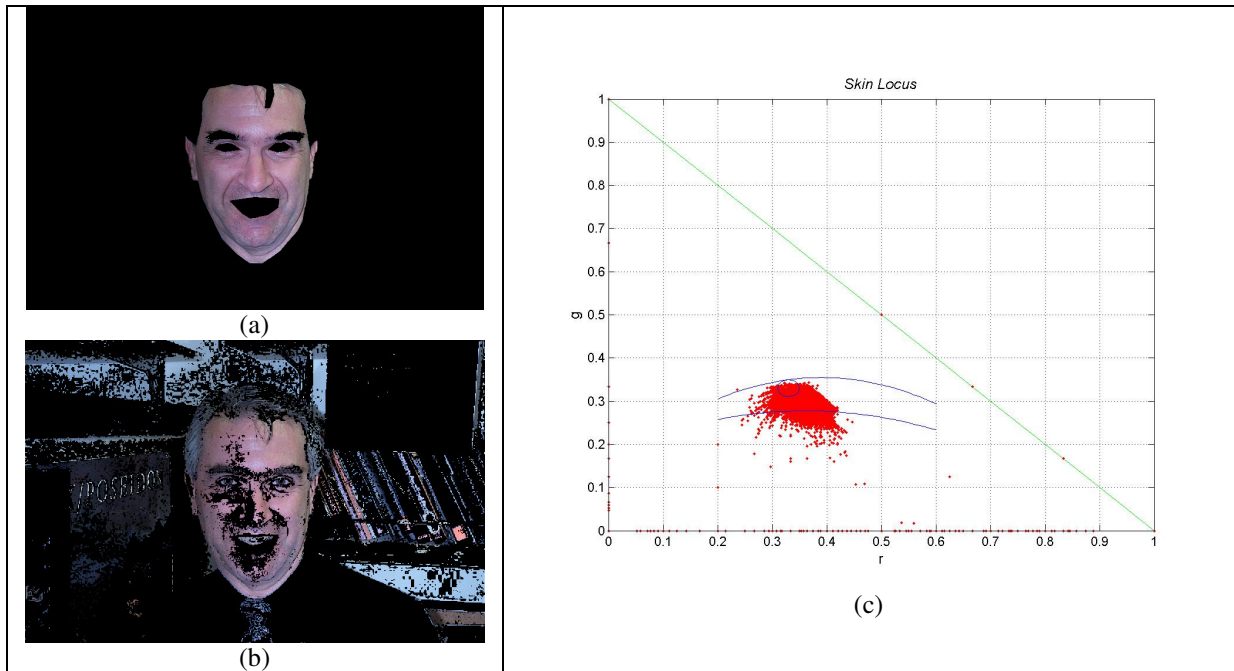


Figura 20 – Primeiro exemplo de *skin locus* com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

Na Figura 20 pode-se verificar que praticamente metade dos pixels de pele é classificada como fundo por essa modelagem. O mesmo acontece nas Figura 21 e Figura 22, nas quais apresentam pixels fora da região modelada.

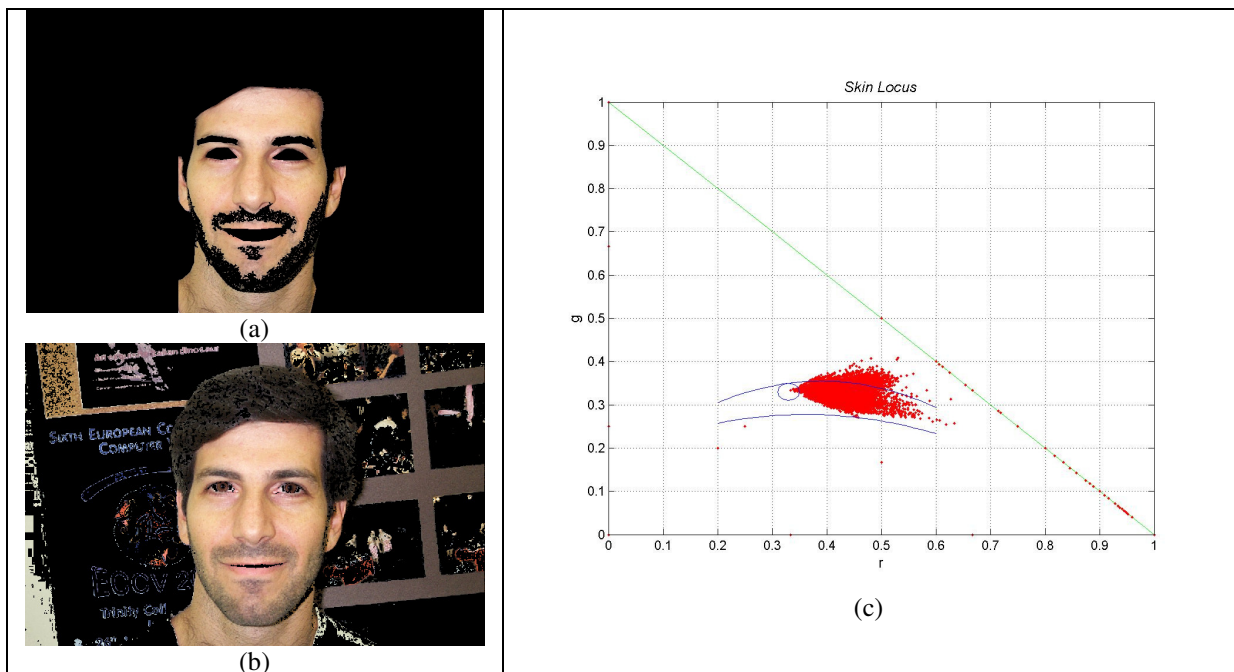


Figura 21 – Segundo exemplo de *skin locus* com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

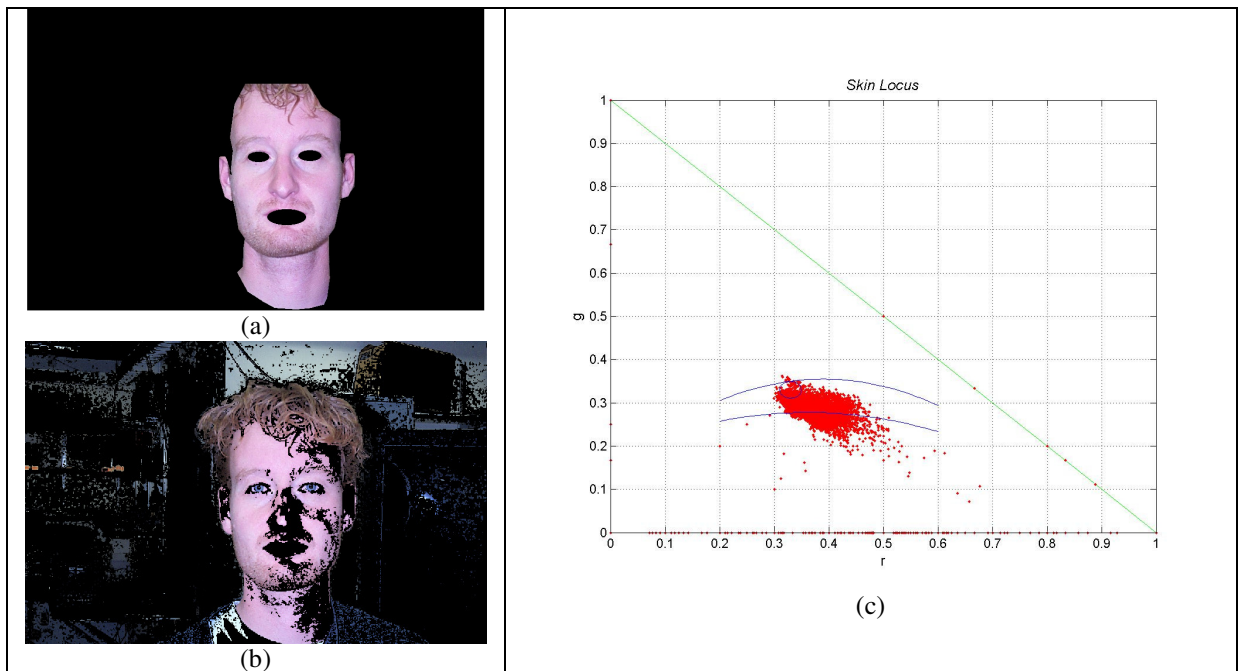


Figura 22 – Terceiro exemplo de *skin locus* com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

Em ambos os tipos de resultados, os polinômios delimitadores classificam muitos pixels do fundo da cena como pixels de pele, pois estes se encontram dentro da área de validação.

1.5.2 *Skin locus* do Banco de imagens B

A modelagem do *skin locus* quando aplicada no banco de imagens B (Anexo A) apresentou resultados muito semelhantes aos do banco de imagens A. Os resultados foram classificados com e sem perda de informação, além da classificação dos pixels de fundo como pele.

Usou-se 16 amostras do banco de imagens B, máscaras de pele, para a construção do *skin locus* da Figura 23. Observou-se que esse *skin locus* manteve-se mais concentrado em uma região do plano.

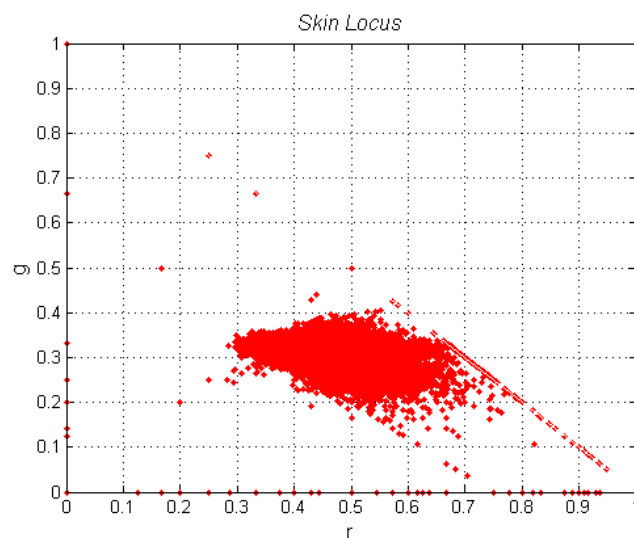


Figura 23 – *Skin locus* de 16 amostras de pele do banco de imagens B

- Resultados com pouca perda de informação de pele

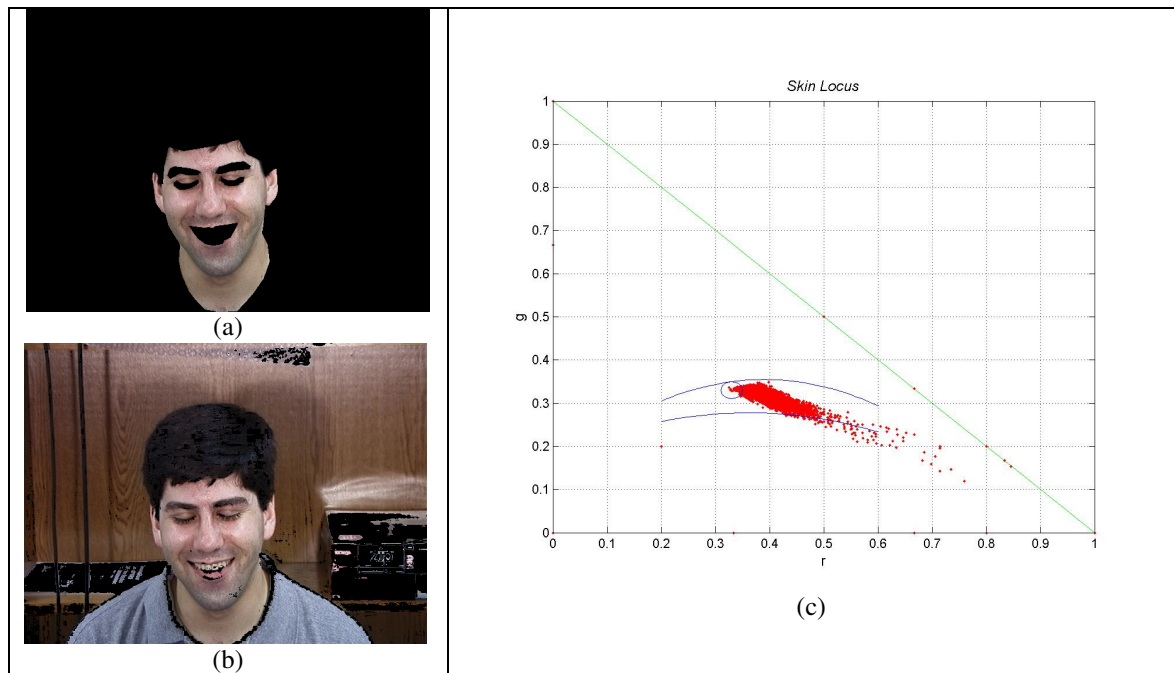


Figura 24 – Primeiro exemplo de *skin locus* com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

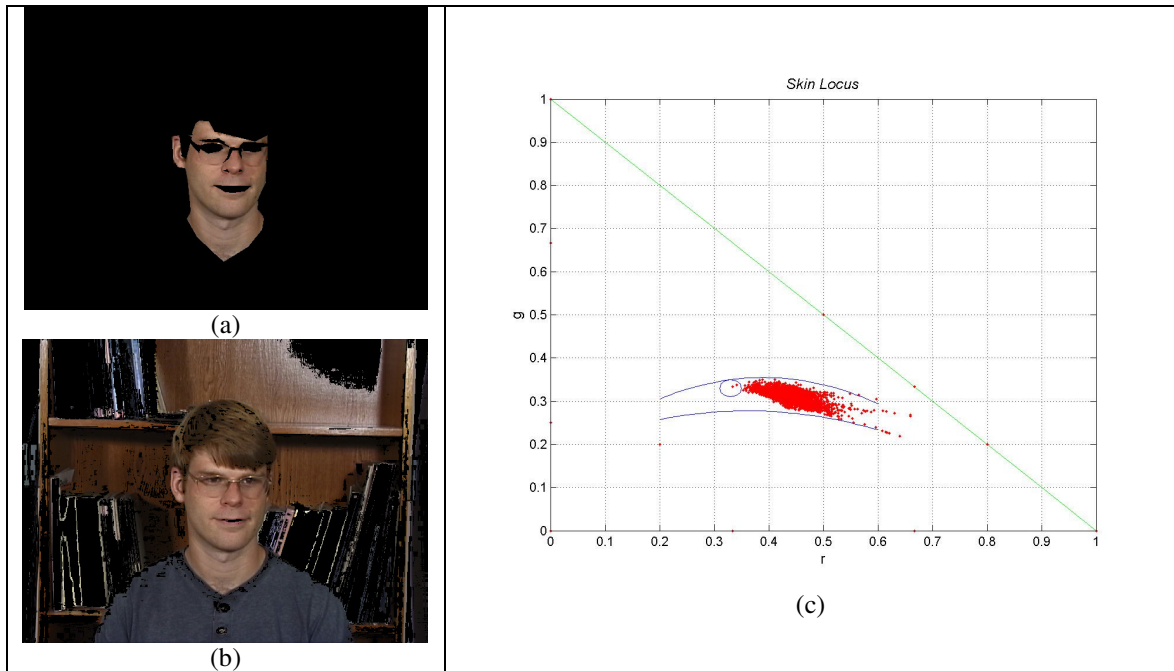


Figura 25 – Segundo exemplo de *skin locus* com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

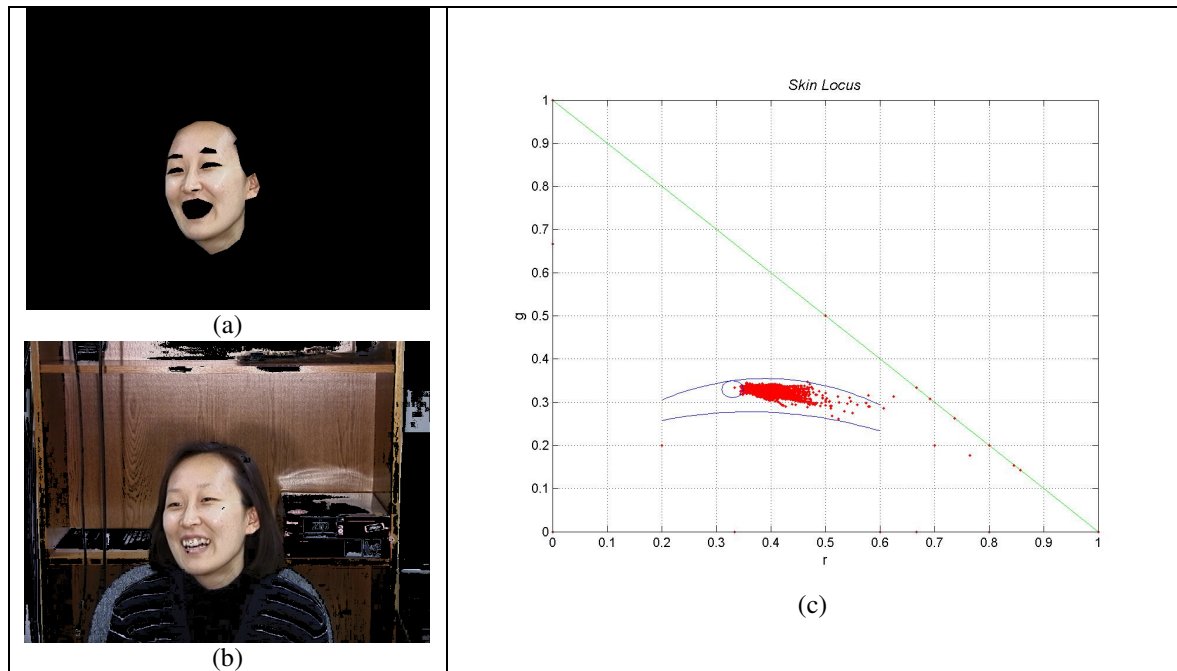


Figura 26 – Terceiro exemplo de *skin locus* com pouca perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

As Figura 24 à Figura 26 mostram exemplos de *skin locus* que é classificado corretamente, com pouco ou praticamente nenhuma perda.

- Resultados com perda de informação de pele

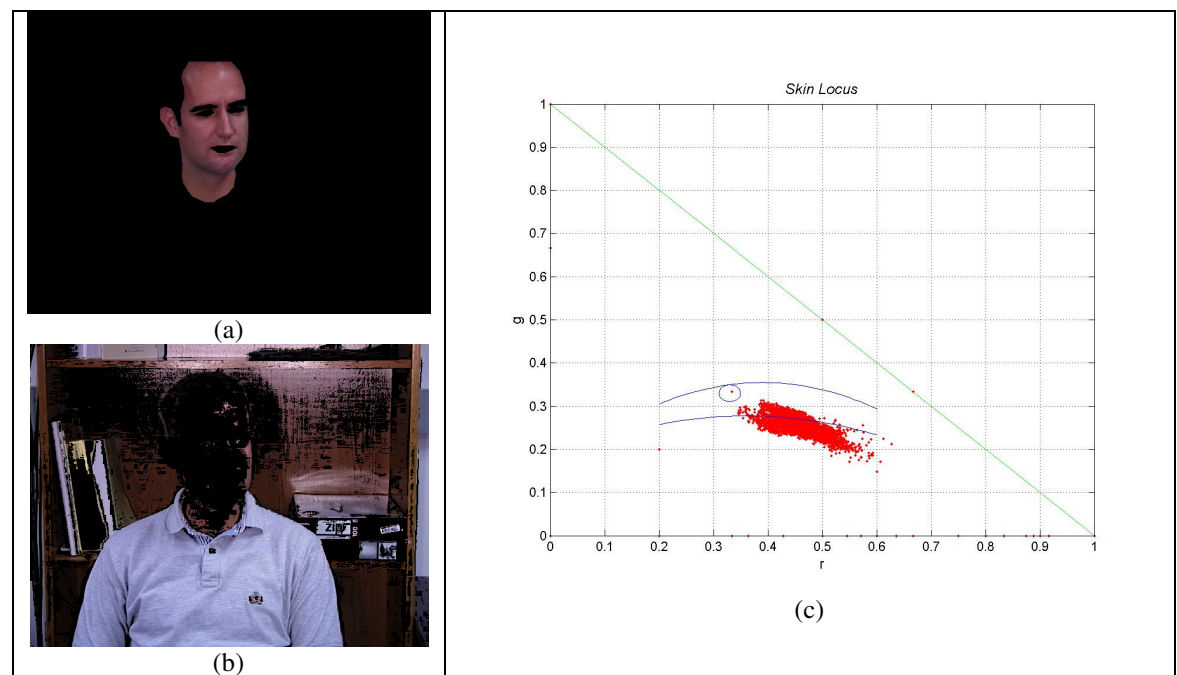


Figura 27 – Primeiro exemplo de *skin locus* com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

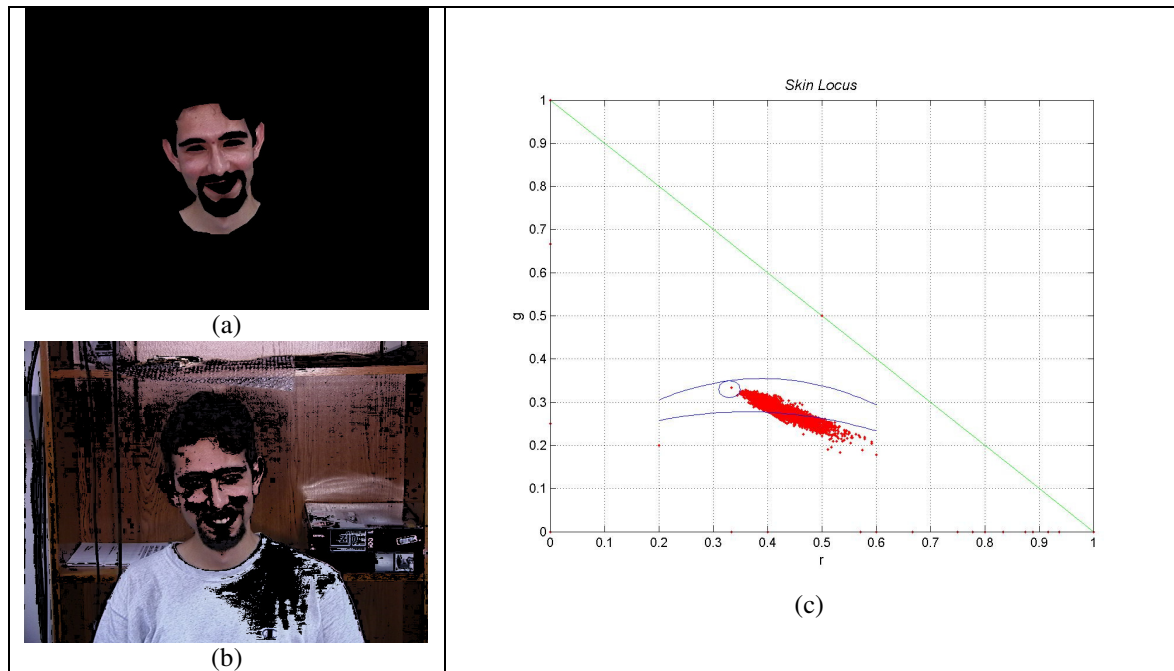


Figura 28 – Segundo exemplo de *skin locus* com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

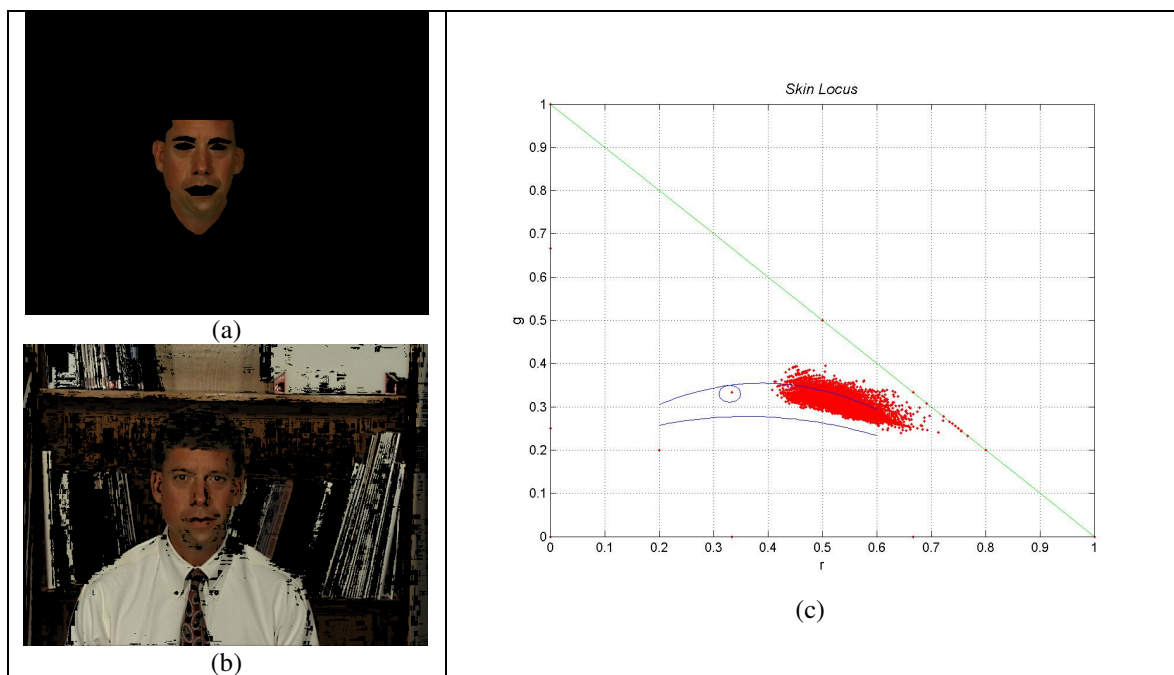


Figura 29 – Terceiro exemplo de *skin locus* com perda de informação de pele. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

Para as Figura 27 até a Figura 29 houve muita perda de informação de pele na modelagem, caracterizando em um resultado ruim, pois classificou poucos pixels e pele e muitos pixels do fundo da cena como se fosse pele.

1.5.3 *Skin locus* das imagens de webcam

Realizou-se a mesma modelagem com as imagens de *webcam* descritas no anexo A. Esses testes foram realizados, pois Martinkauppi (2000) enfatiza essa modelagem feita principalmente para as imagens de uma *webcam*.

Quando utilizadas 16 máscaras de pele das imagens de *webcam*, produziu-se um *skin locus* mais concentrado, Figura 30 semelhante o de Martinkauppi (2000).

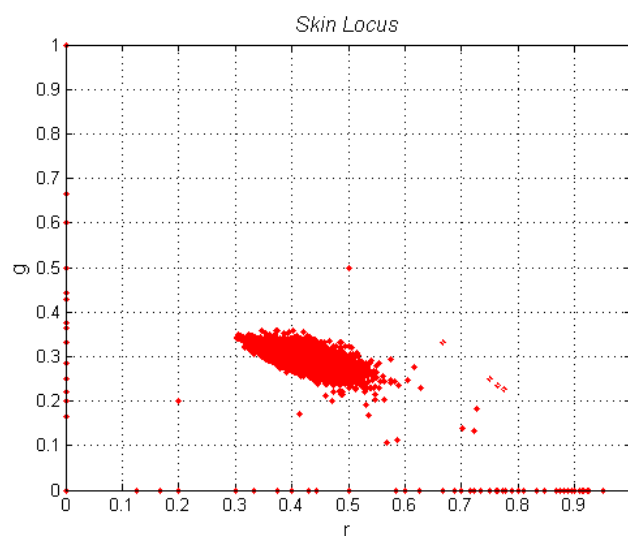


Figura 30 - *Skin locus* de 16 amostras de pele a partir de imagens de *webcam*

Nas Figura 31 e Figura 32 são mostrados *skin locus* de imagens de *webcam* que puderam ser corretamente modelados pelas regras de Martinkauppi (2000) e nas Figura 33 e Figura 34 as que apresentaram perda de informação durante a modelagem.

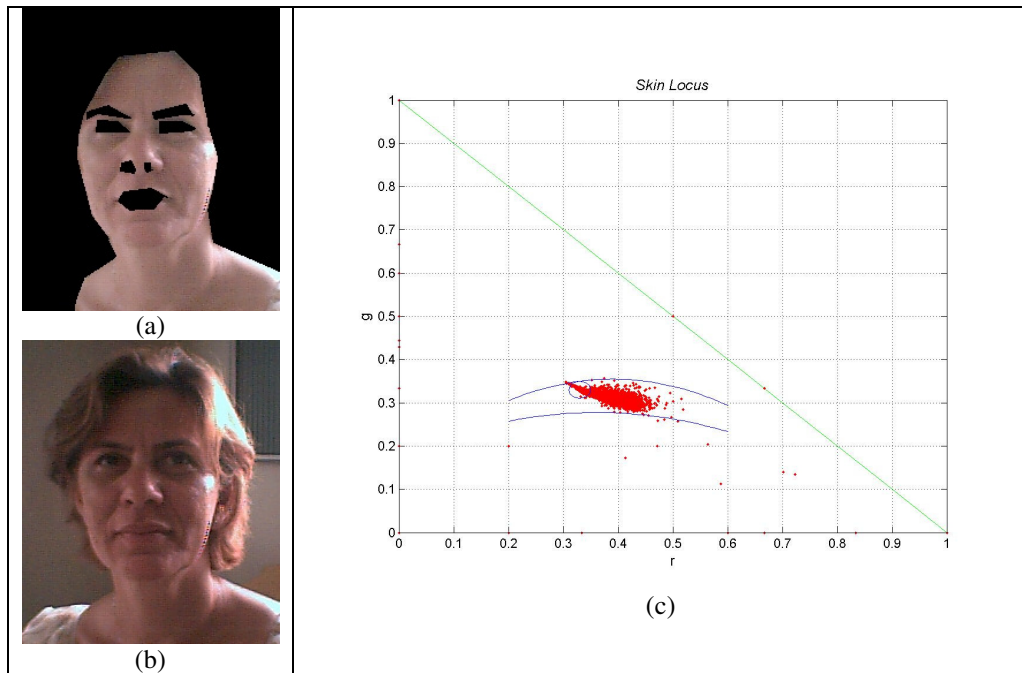


Figura 31 - Primeiro exemplo de *skin locus* sem perda de informação de pele para *webcam*. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

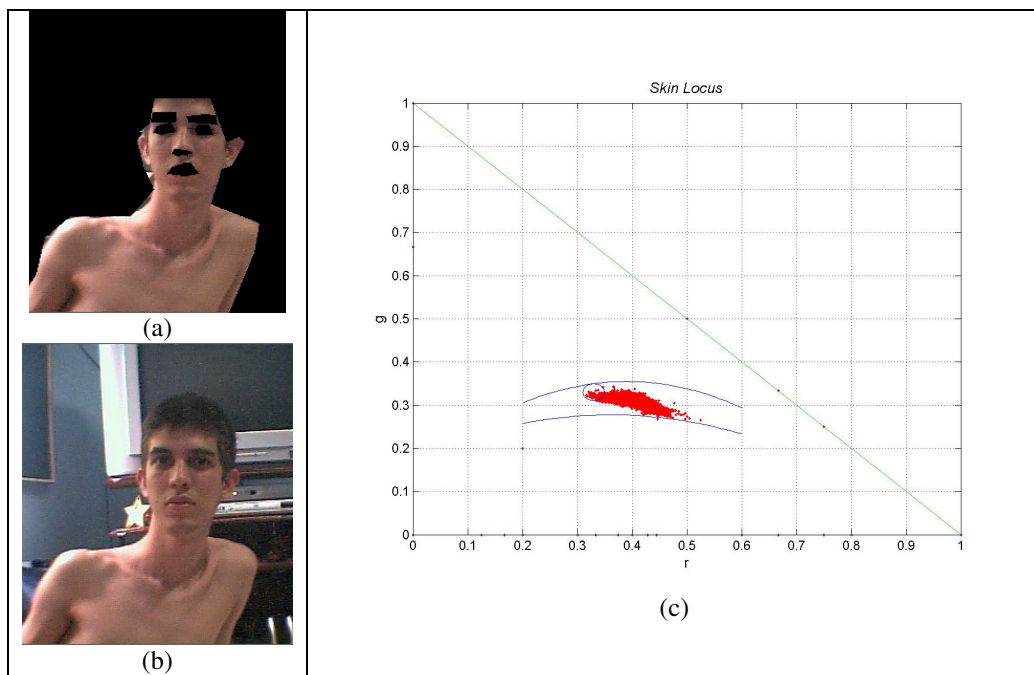


Figura 32 - Segundo exemplo de *skin locus* sem perda de informação de pele para *webcam*. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

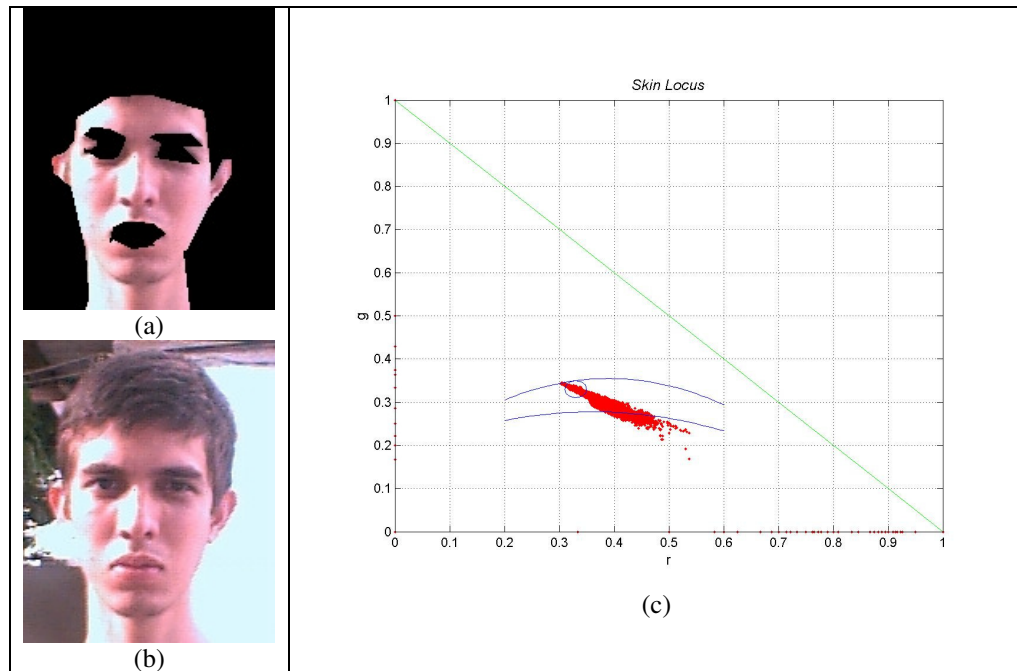


Figura 33 - Primeiro exemplo de *skin locus* com perda de informação de pele para *webcam*. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

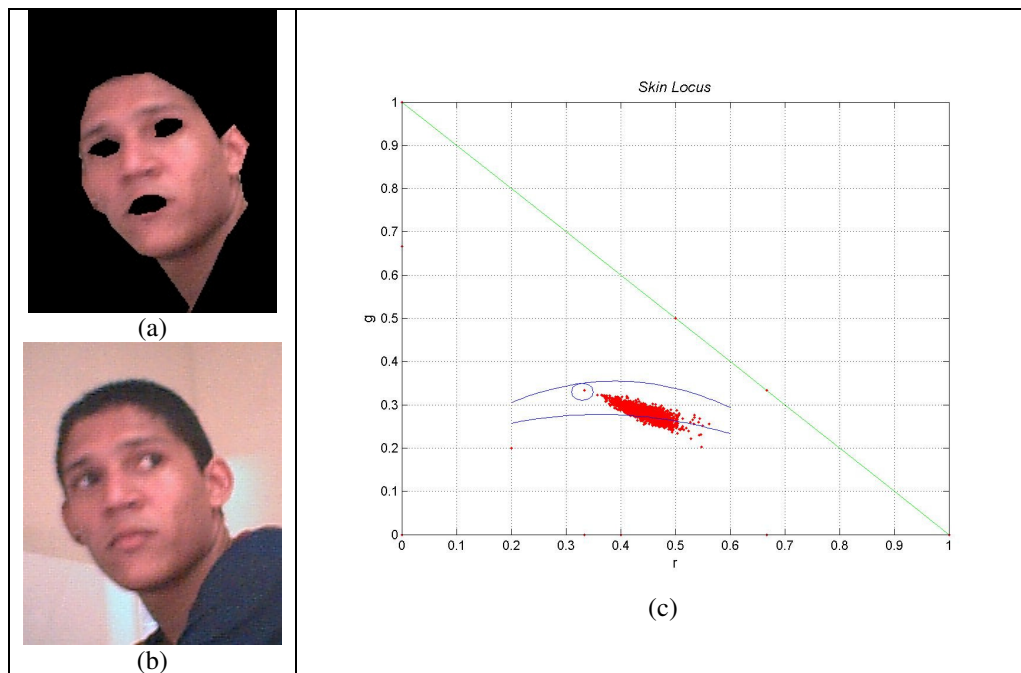


Figura 34 - Segundo exemplo de *skin locus* com perda de informação de pele para *webcam*. (a) Máscara de pele da imagem original, (b) Imagem após a aplicação do *skin locus* e (c) gráfico do *skin locus* da imagem original

1.5.4 Resultados da classificação de pele por *skin locus*

Aplicou-se a modelagem do *skin locus* para todas as imagens selecionadas do banco de imagens A e B, e obteve-se resultado melhor no banco de imagens A (Tabela 1). No banco de

imagens B classificou-se mais da metade do fundo como pele. Uma justificativa para esse pior desempenho são as características dos objetos de fundo da imagem, descritas no anexo A.

	Banco de imagens A	Banco de imagens B
VP (%)	93,44	94,60
FP (%)	26,67	68,62

Tabela 1 – Taxas de reconhecimento obtidas utilizando-se o algoritmo de detecção de pixels de pele proposto por Martinkauppi (2000)

Os valores de verdadeiro positivo (VP) foram obtidos comparando-se os pixels, da imagem de entrada, classificados como pele com os pixels da máscara. Os falsos positivos (FP) foram obtidos a partir da comparação dos pixels de fundo classificados como pele com a máscara.

A Figura 35 mostra um exemplo de resultado da aplicação do *skin locus* em uma imagem do banco A. Neste caso 98% dos pixels de pele foram identificados corretamente (VP) e 37% do fundo foram classificados como pele (FP).



Figura 35 – Resultado da aplicação do *skin locus* em uma imagem do banco A

Na Figura 35 podemos ver que objetos de cor azul e laranja foram classificados como pele, sendo que nunca ocorrerão casos em que a pele seja representada por essas cores.

A Figura 36 mostra os valores R, G e B de dois pixels de fundo, classificados como pele devido a sua localização no *skin locus*.

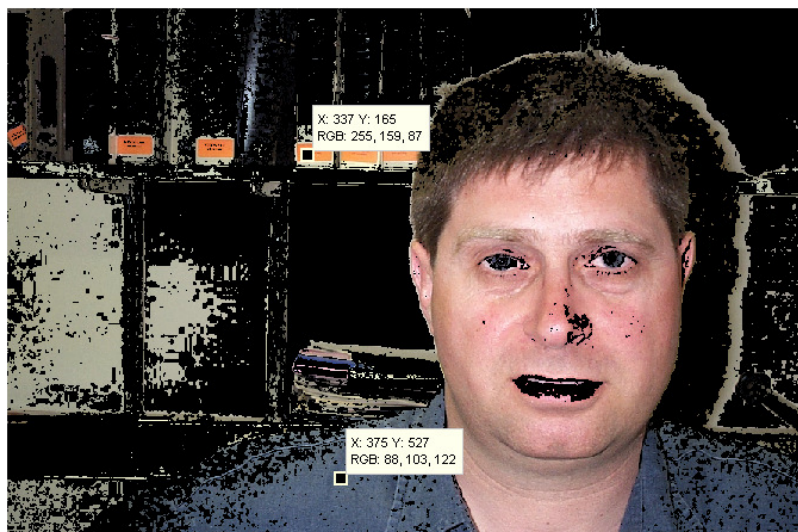


Figura 36 – Valores de pixels de fundo classificados como pele

O fato de objetos de cor azul e laranja serem classificados como pele é porque ao se transformar para NCC, o pixel com tonalidade laranja apresenta os valores $r = 0,508$ e $g = 0,317$, e o pixel com tonalidade azul é o equivalente a $r = 0,281$ e $g = 0,329$. Ambos os pixels encontram-se na região entre os polinômios delimitadores e fora do círculo, Figura 37.

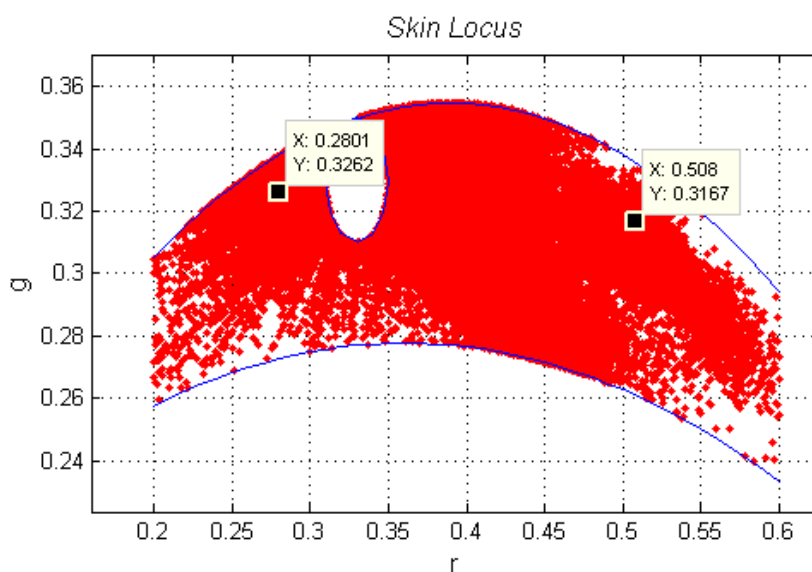


Figura 37 – Skin locus da Figura 36. Localização dos pixels de fundo classificados como pele no skin locus

Nas imagens do banco B os resultados apresentaram-se piores do que os do banco A devido o elemento de madeira no fundo da cena. A Figura 38 apresenta um típico resultado para esse banco onde $VP = 100\%$ e $FP = 60\%$.



Figura 38 – Resultado da aplicação do *skin locus* em uma imagem do banco B

O objeto de fundo da Figura 38 apresenta tonalidade semelhante à pele de indivíduos negros, Figura 39. Para a Figura 39, a taxa de PV = 99,51% e FP = 57,78%.



Figura 39 - Resultado da aplicação do *skin locus* em uma imagem do banco B com indivíduo negro

Visando realizar a classificação dos pixels de pele com taxa de falsos positivos (FP) menores em ambos os bancos de imagens

1.5.5 Novas faixas de valores r e g e resultados

Por conseguinte, analisou-se os histogramas dos valores r e g (Gráfico 1 e Gráfico 2) para determinação de uma nova faixa de valores da pele para imagens de resolução dos bancos A e B, visto que a proposta de Martinkauppi (2000) mostrou-se mais adequada para imagens de baixa resolução (*webcam*), concluindo assim o passo b da criação do classificador de pixels de pele.

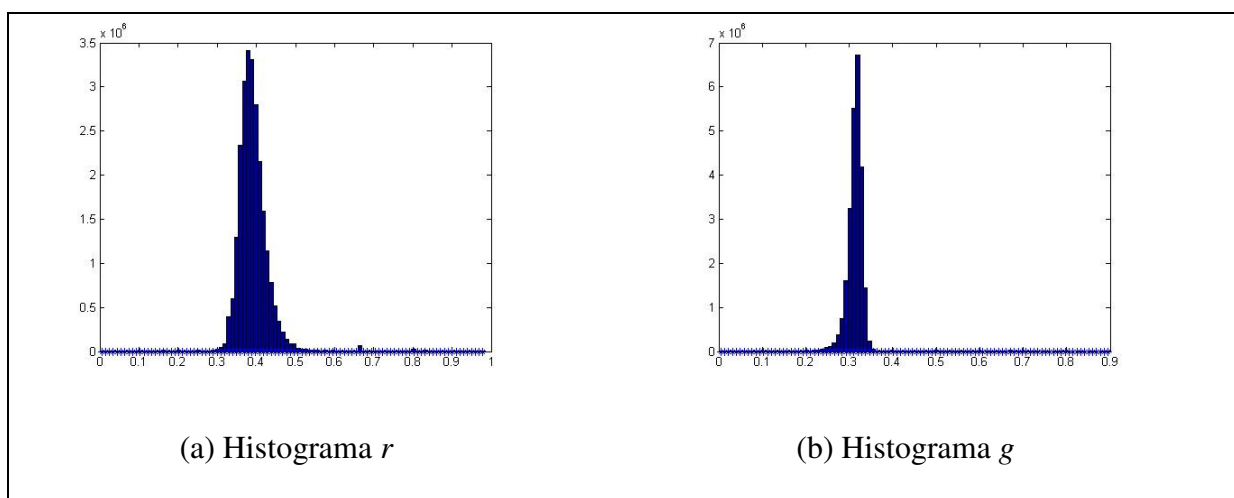


Gráfico 1 – Histogramas (a) e (b) das máscaras de pele do banco de imagens A

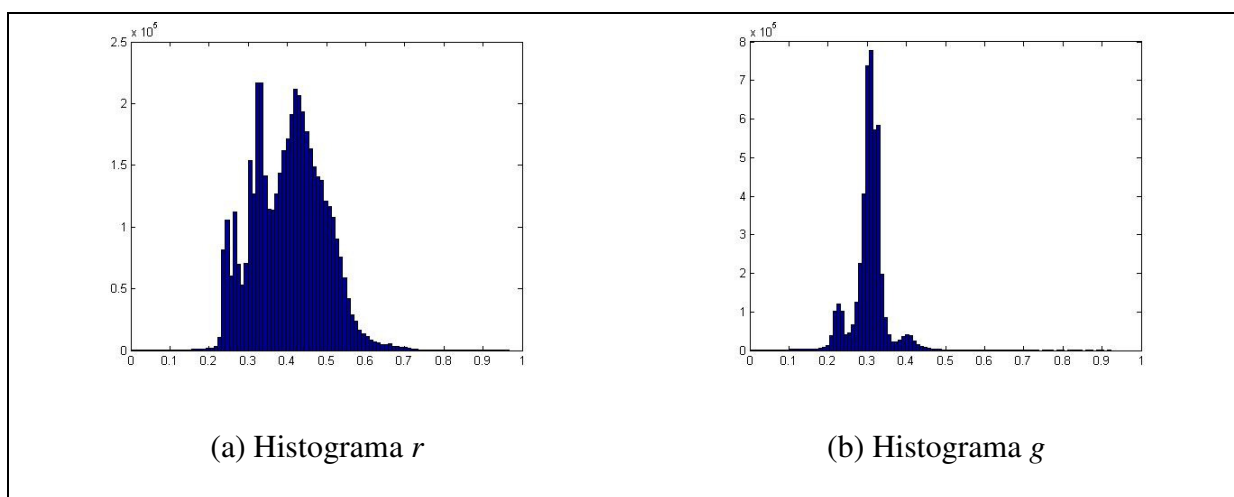


Gráfico 2 - Histogramas (a) e (b) das máscaras de pele do banco de imagens B

A partir dos valores de mediana da distribuição gaussiana dos histogramas de r e g , foram estimadas faixas de valores para validação dos pixels.

Os valores de mediana para cada gráfico são apresentados no Quadro 4.

	Banco de imagens A		Banco de Imagens B	
	R	G	R	G
Mediana	0,38	0,31	0,41	0,31

Quadro 4 - Valores da mediana para as distribuições r e g dos bancos de imagens A e B

Por experimentação, estimou-se faixas de valores r e g. As taxas de reconhecimentos para os bancos de imagens A e B são apresentadas nos Quadro 5 e Quadro 6. Os desempenhos para essas faixas podem ser visualizados nos Gráfico 3 e Gráfico 4.

	Faixa de valores r	Faixa de valores g	VP (%)	FP (%)
1	$0,360 < r < 0,410$	$0,290 < g < 0,330$	91,24	24,45
2	$0,355 < r < 0,405$	$0,305 < g < 0,330$	62,77	32,92
3	$0,360 < r < 0,410$	$0,295 < g < 0,330$	62,43	34,52
4	$0,360 < r < 0,410$	$0,305 < g < 0,330$	62,27	31,19
5	$0,360 < r < 0,410$	$0,300 < g < 0,330$	62,36	31,19
6	$0,355 < r < 0,400$	$0,290 < g < 0,330$	58,31	36,84
7	$0,360 < r < 0,405$	$0,305 < g < 0,330$	57,93	30,33
8	$0,360 < r < 0,410$	$0,290 < g < 0,335$	54,18	41,00
9	$0,355 < r < 0,395$	$0,290 < g < 0,330$	52,60	35,76
10	$0,355 < r < 0,395$	$0,305 < g < 0,330$	52,40	31,05
11	$0,355 < r < 0,390$	$0,305 < g < 0,330$	46,31	30,08

Quadro 5 - Faixas de valores propostas e suas respectivas percentagens de reconhecimento (banco de imagens A)

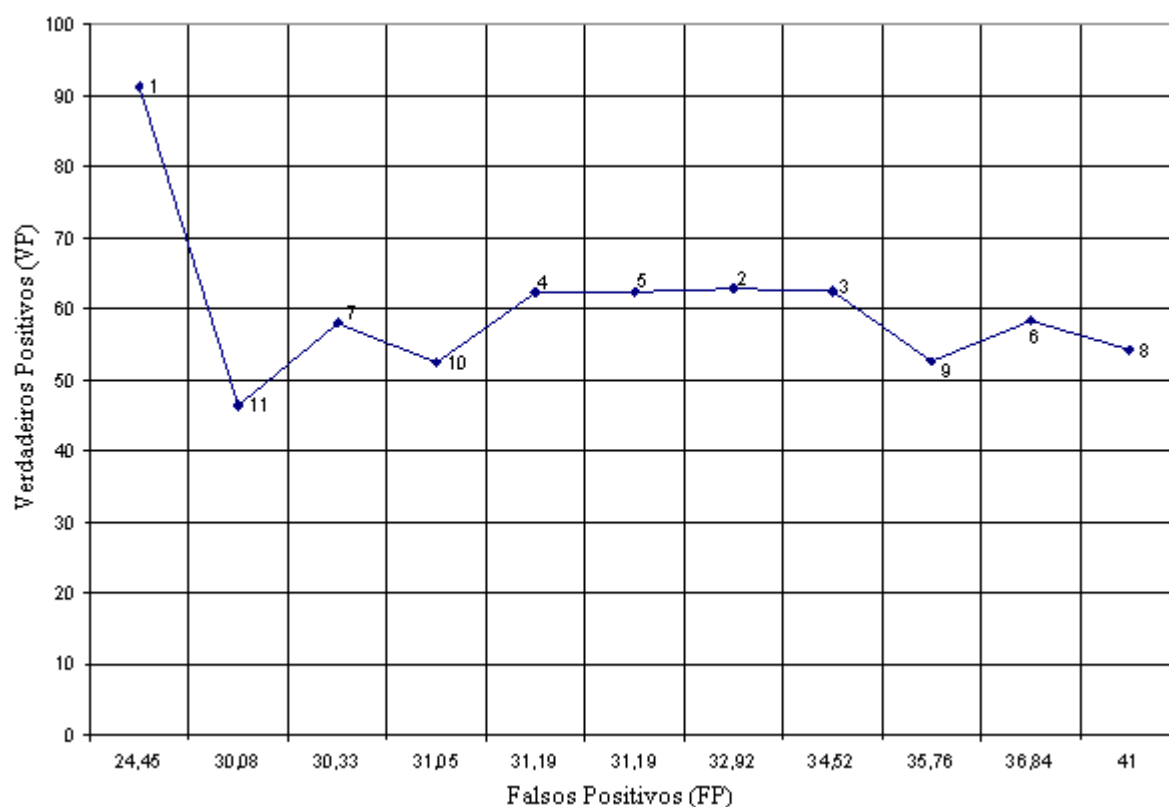


Gráfico 3 - Desempenho de cada faixa de pixels vermelho e verde experimentadas para classificação dos pixels de pele (banco de imagens A)

	Faixa de valores r	Faixa de valores g	VP (%)	FP (%)
1	0,360<r<0,410	0,290<g<0,330	83,95	67,76
2	0,355<r<0,405	0,305<g<0,330	58,75	52,47
3	0,360<r<0,410	0,295<g<0,330	77,90	63,65
4	0,360<r<0,410	0,305<g<0,330	61,28	52,53
5	0,360<r<0,410	0,300<g<0,330	69,67	58,09
6	0,355<r<0,400	0,290<g<0,330	82,99	67,58
7	0,360<r<0,405	0,305<g<0,330	58,73	51,74
8	0,360<r<0,410	0,290<g<0,335	84,90	72,79
9	0,355<r<0,395	0,290<g<0,330	82,60	67,13
10	0,355<r<0,395	0,305<g<0,330	53,99	51,09
11	0,355<r<0,390	0,305<g<0,330	52,21	50,42

Quadro 6 – Faixas de valores propostas e suas respectivas porcentagens de reconhecimento (banco de imagens B)

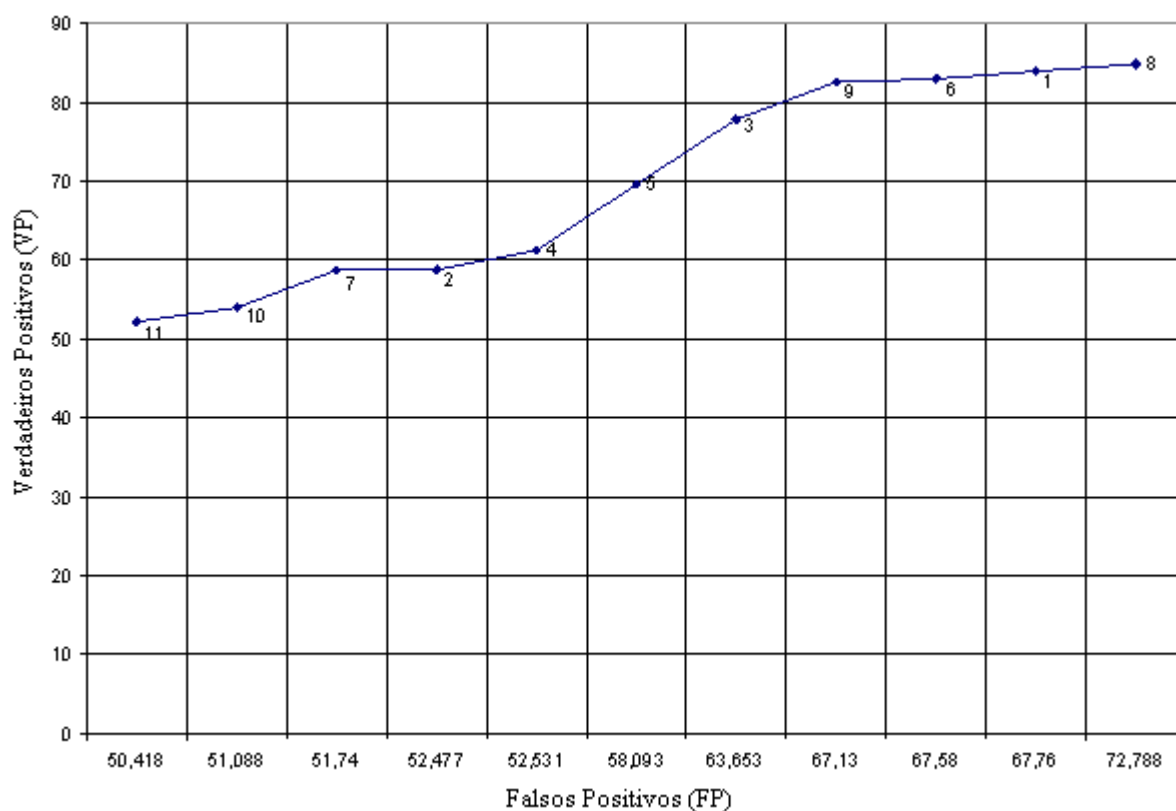


Gráfico 4 – Desempenho de cada faixa de pixels vermelho e verde experimentadas para classificação dos pixels de pele (banco de imagens B)

Dentre as faixas estimadas no reconhecimento de pixels de pele, a faixa 1 foi a que apresentou melhor taxa de verdadeiros positivos (VP) e menor de falsos positivos (FP)

quando utilizada no banco de imagens A. A mesma resposta não foi obtida quando utilizou-se esta faixa no banco de imagens B.

Para o banco de imagens B, a faixa 11 apresentou menor taxa de falsos positivos, mas, também apresentou a menor taxa de positivos verdadeiros, apesar dos valores elevados. Essa faixa foi a que a obteve pior resultado quando aplicada no banco de imagens A.

Comparando-se esses resultados com os resultados da aplicação dos polinômios delimitadores do *skin locus*, observou-se que:

- Os resultados dos polinômios delimitadores do *skin locus* foram superiores na taxa de positivos verdadeiros;
- Os resultados dos dois foram equivalentes em falsos positivos.

A partir desses resultados observou-se que os passos, apresentados pela literatura, para criação de um classificador de cor de pele torna-o muito específico para o banco de imagens que se está utilizando.

2 DETECÇÃO FACIAL A PARTIR DE DIFERENÇAS EM RGB

2.1 Introdução

A detecção de pixels de pele utilizando o *skin locus* é largamente utilizada na literatura em diversos modelos de cor, porém seus resultados mostraram-se pouco satisfatórios quando aplicado a imagens de alta resolução. Devido à considerável taxa de reconhecimento de falsos positivos obtida pelo *skin locus*, buscou-se outra abordagem para criar um classificador de pixels de pele que apresentasse uma taxa de falsos positivos menor.

Realizou-se um estudo sobre o comportamento dos pixels de pele no espaço de cor RGB e propôs-se um classificador baseado em limiares que pode ser utilizado nos dois bancos de imagens (Banco de Imagens A e Banco de Imagens B).

O objetivo deste capítulo é mostrar como se comportam os pixels de pele no espaço de cor RGB e a partir desse comportamento, como foi construído o classificador utilizando-se diferenças entre os planos do espaço de cor RGB.

O estudo do comportamento dos pixels no espaço de cor RGB possibilitou o estabelecimento de faixas de valores para a extração dos pixels de pele da imagem. Mesmo utilizando-se essas faixas de valores, alguns pixels com tonalidades semelhantes eram confundidos com os de pele aumentando a taxa de falsos positivos.

Para reduzir essa taxa, utilizou-se limites horizontais e verticais para a região de pele da face, que sempre seria a região com maior quantidade de pixels classificados como pele. O uso desses limites reduziram o espaço de busca por características da face, agilizando-se o processo de identificação da mesma.

Assim, este capítulo faz uma análise do comportamento dos pixels de pele no espaço de cor RGB, a obtenção dos limites para classificá-los como pixels de pele e o uso de limites verticais e horizontais para restringir o espaço de busca pela face.

2.2 Comportamento dos pixels no espaço de cor RGB

Nessa investigação, observou-se o comportamento dos valores dos pixels de pele nas imagens do Banco de Imagens A e no Banco de Imagens B como mostra a Figura 40.

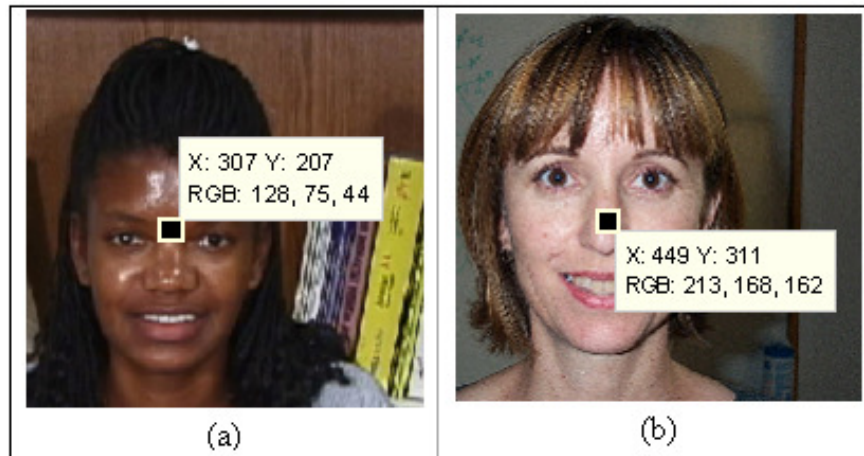


Figura 40 – Imagens faciais apresentando valores de R, G e B para a pele. (a) Pele negra; (b) Pele caucasiana

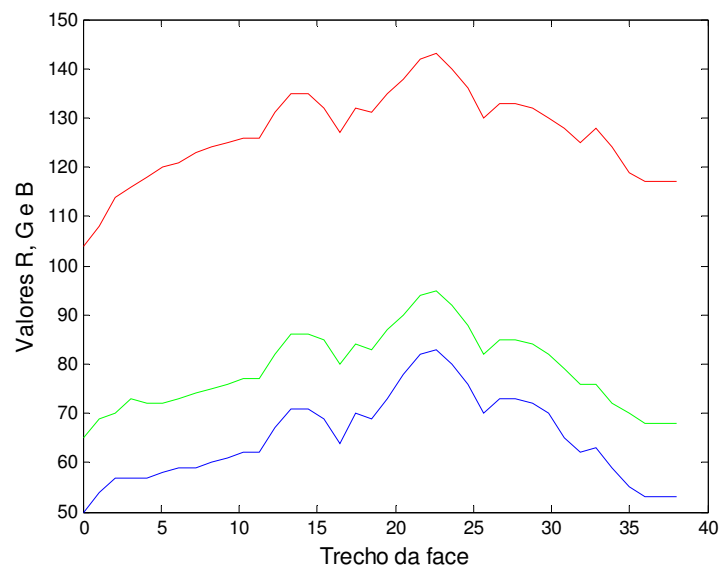


Gráfico 5 - Perfil do comportamento R, G e B na face do indivíduo da Figura 40(a)

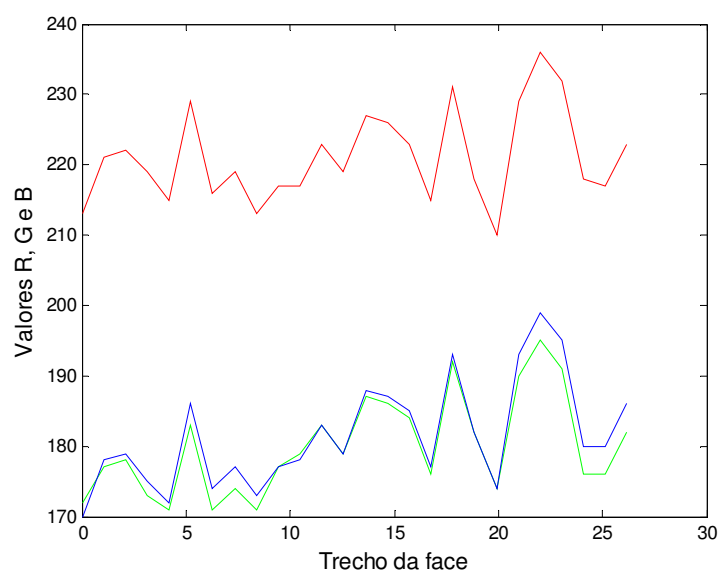


Gráfico 6 – Perfil do comportamento R, G e B na face do indivíduo da Figura 40 (b)

De acordo com essa investigação foram feitas três observações, sendo a primeira: os pixels vermelhos possuem sempre valores superiores aos pixels verde e azul. A segunda foi que os pixels verdes são visivelmente superiores aos pixels azuis, em indivíduos de pele negra (Gráfico 5). E a terceira, em indivíduos de pele caucasiana, os pixels azuis e verdes apresentam praticamente o mesmo valor ou os pixels verdes são ligeiramente inferiores aos azuis (Gráfico 6).

Traçando-se o perfil no plano de cor RGB, percebeu-se que os pixels vermelho, verde e azul da região de pele possuem comportamento diferente de outros objetos na imagem, podendo a região de pele ser claramente distinguida. Os Gráfico 7 e 8 apresentam o comportamento das R, G e B para as linhas selecionadas nas imagens das Figura 41 e Figura 42, respectivamente.



Figura 41 – Seleção da linha da imagem para verificar o perfil dos valores R, G e B (indivíduo de pele caucasiana)

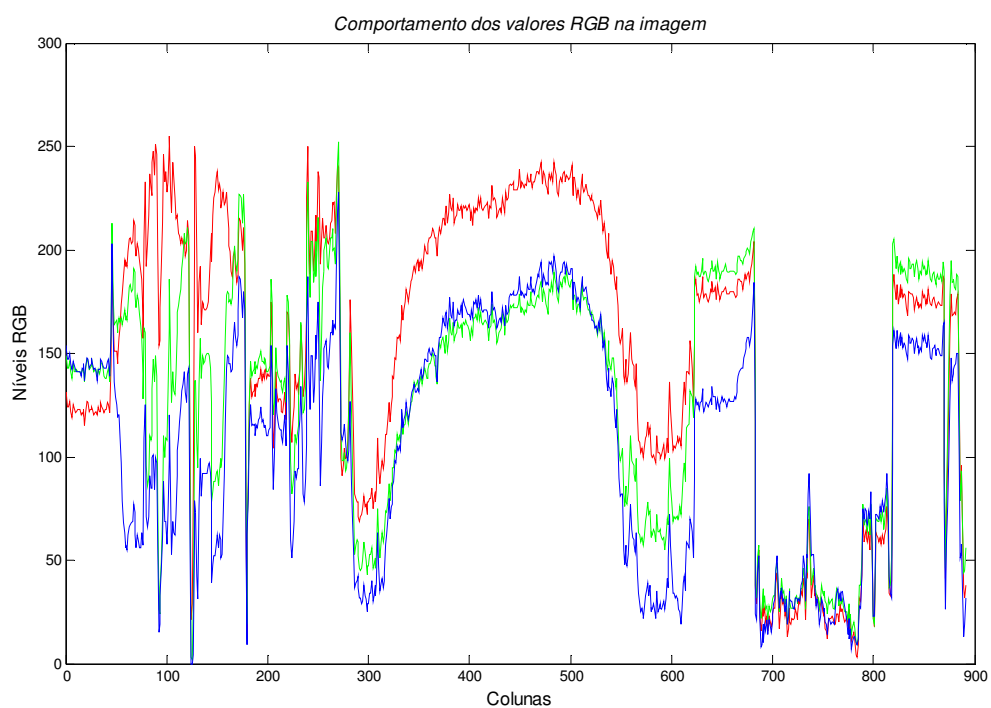


Gráfico 7 – Comportamento dos níveis R, G e B da linha selecionada na Figura 41



Figura 42 – Seleção da linha da imagem para verificar o perfil dos valores R, G e B (indivíduo de pele negra)

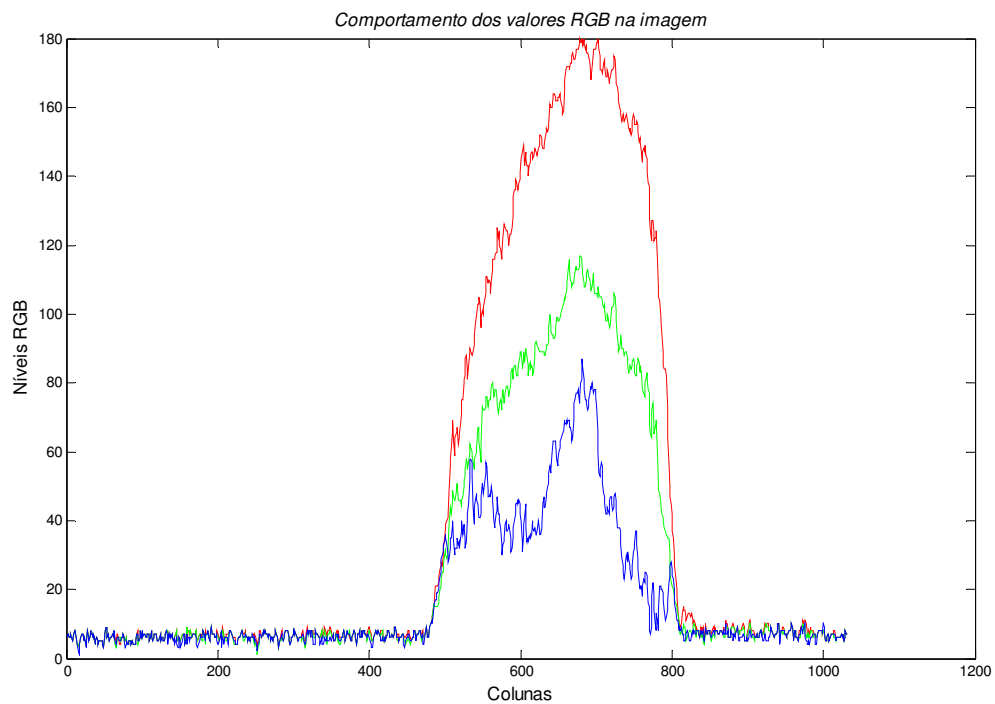


Gráfico 8 – Comportamento dos níveis R, G e B da linha selecionada na Figura 42

Tomando-se como base a afirmação de Störring (1999) de que os canais com maiores informações sobre a cor de pele são os canais vermelho (R) e verde (G), realizou-se uma operação de diferença entre esses canais e aplicou-se um limiar adaptativo para construção de uma máscara binária.

O limiar adaptativo foi obtido pela divisão horizontal da imagem em 3 partes, para as quais foram calculados os limiares globais.

Utilizando-se este processo na criação da máscara binária para classificação de pixels de pele, as taxas de reconhecimento obtidas foram:

	Banco de Imagens A	Banco de Imagens B
PV	89,61 %	91,19 %
FP	10,73 %	34,72 %

Tabela 2 – Taxas de reconhecimento de pele e fundo utilizando-se a máscara binária da diferença entre os canais RG

Estes resultados, quando comparados com os do algoritmo de Martinkauppi (2000), apresentam uma taxa de positivos verdadeiros (VP) 5% menor e uma redução em 50% na taxa de falsos positivos (FP) para ambos os bancos de imagens.

Dois exemplos dos resultados obtidos com a diferença entre R e G são apresentados na Figura 46, sendo uma imagem do Banco A e a outra do Banco B. A Figura 43 apresenta as imagens originais, antes da aplicação do método, e a Figura 44 mostra as máscaras obtidas dessas imagens contendo apenas as regiões de pele. Essas máscaras foram utilizadas para comparação do resultado de classificação dos pixels de pele para os métodos avaliados nesse trabalho.



Figura 43 – Imagens originais do (a) Banco A e do (b) Banco B

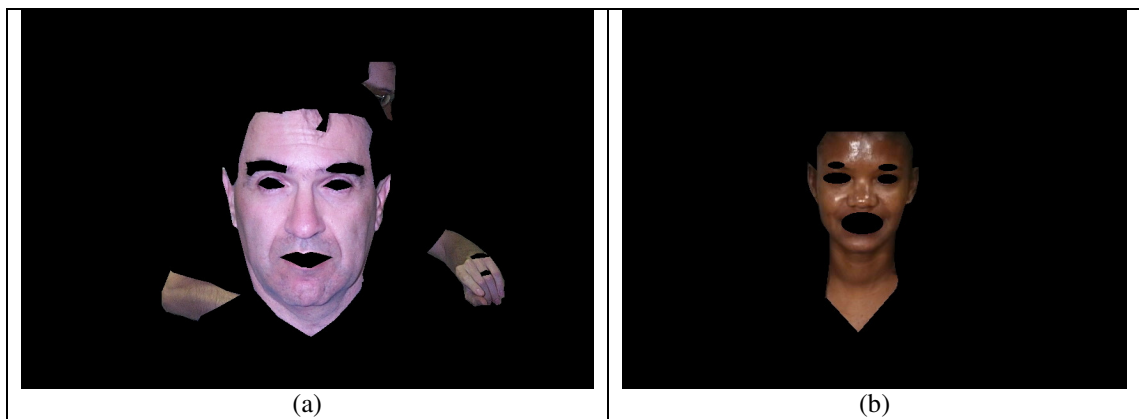


Figura 44 – Máscaras contendo as regiões de pele utilizada como referência para comparação dos resultados. (a) Máscara da imagem do Banco A e (b) máscara da imagem do Banco B



Figura 45 - Resultado após utilização do algoritmo de modelagem do *skin locus* (MARTINKAUPPI, 2000)



Figura 46 – Resultado utilizando a diferença entre os canais R e G

A diferença nos resultados da classificação de pixels pela utilização dos dois métodos, *Skin Locus* (Figura 45) e diferença RG (Figura 46), são claramente visíveis. A Tabela 3 apresenta as taxas de reconhecimento para essas duas figuras.

	<i>SkinLlocus</i> (Figura 45)	R-G (Figura 46)
VP (%)	77,83	90,82
PF (%)	32,92	1,69

Tabela 3 – Comparação das taxas de reconhecimento para as Figura 45 e Figura 46

Prosseguindo na investigação comportamental dos canais R, G e B para pele, analisou-se a classificação de pixels de pele quando utilizadas as diferenças RB, GB e BG com limiar adaptativo.

Quando foram construídas as máscaras para classificação de pixels de pele a partir das diferenças entre os canais RB, GB e BG verificou-se que em seus resultados:

- RB: Em caucasianos, enfatiza regiões de pele, captando baixo percentual de outros objetos contidos na imagem (Figura 47(a)). Em negros, poucos pixels de pele são classificados, porém outros objetos são destacados (Figura 47(b)).

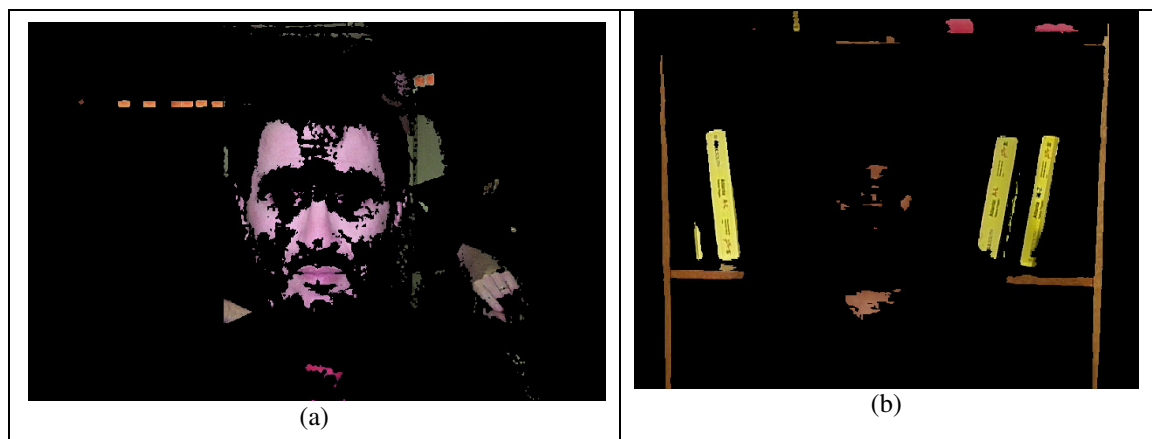


Figura 47 – Classificação dos pixels da imagem utilizando a diferença RB

- GB: Enfatiza os objetos contidos na imagem. Não apresenta regiões de pele (Figura 48).

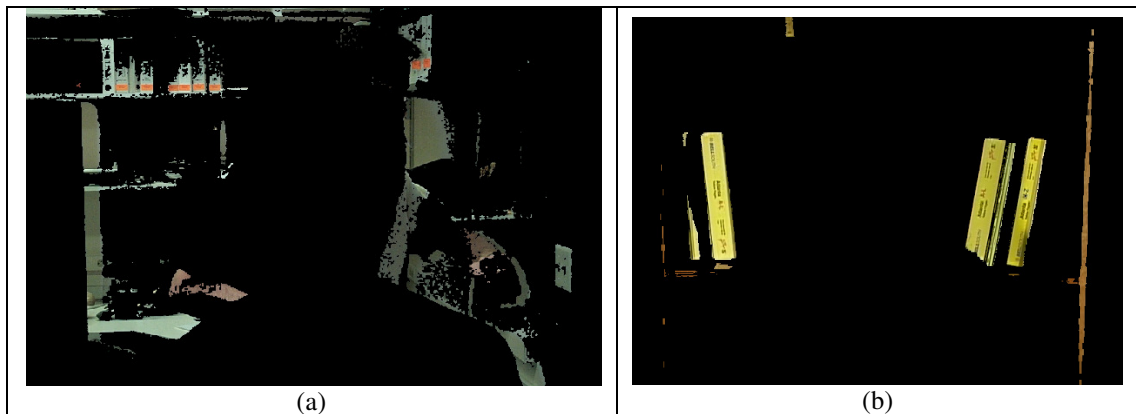


Figura 48 - Classificação dos pixels da imagem utilizando a diferença GB

Os pixels de pele apresentaram diferença entre G e B menor que 30 em indivíduos caucasianos e menor que 50 em negros.

- BG: Para indivíduos de pele com características caucasianas, enfatizou-se regiões de pele e objetos de fundo (Figura 49 (a)). Para indivíduos de pele com características negras, enfatizaram-se objetos do fundo com níveis de azul elevado (Figura 49 (b)).



Figura 49 – Classificação dos pixels da imagem utilizando a diferença BG ((a) indivíduo de pele caucasiana e (b) indivíduo negro)

2.3 Detecção por diferenças em RGB

O método proposto para classificação dos pixels de cor de pele baseia-se nas respostas obtidas nas imagens de diferenças entre os canais do plano RGB. As etapas dessa

segmentação são mostradas no diagrama de blocos da Figura 50. A seguir, são relatadas essas etapas.

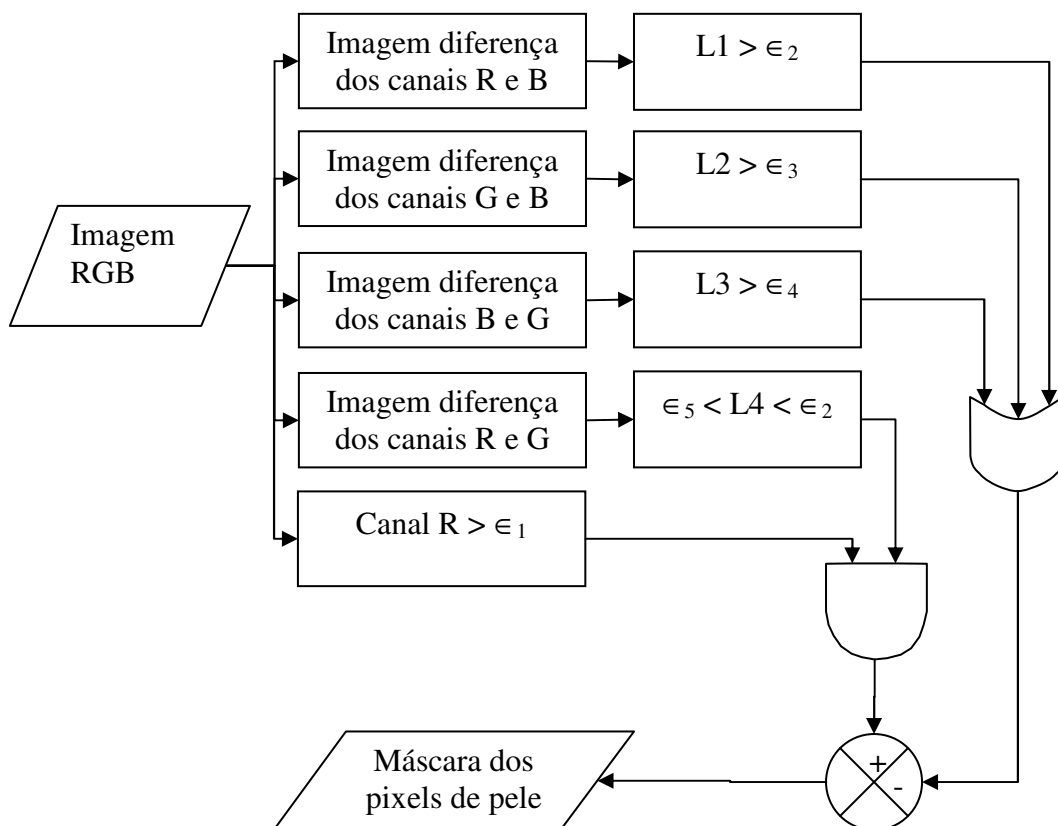


Figura 50 - Etapas da segmentação proposta

Onde:

$L1, L2, L3$ e $L4$ – Limiar global

$\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5$ e ϵ_6 – Constantes

O primeiro passo foi a obtenção das imagens diferença entre os canais R, G e B. As imagens diferença RB, GB e BG foram utilizadas na criação de uma máscara de objetos. Essa máscara tem como foco os objetos presentes na imagem que não são pele. Assim, com essa máscara de objetos construída é possível saber quem é objeto ou não na imagem original e subtrair dela os objetos que não são o alvo principal da análise.

As constantes utilizadas pelos limiares foram obtidas através da análise dos valores referente a diferença entre os canais das imagens do Banco de Imagens A e Banco de Imagens B.

Após a obtenção dos valores diferenças de todas as imagens, construiu-se um histograma com esses valores. Verificou-se quais eram seus limites para utilizá-los como limiares. Ou seja, construiu-se um histograma com os valores da diferença RG de todas as imagens do Banco A e do Banco B. Nesse histograma, identificou-se qual era o maior e o menor valor que ocorria e o selecionou como limiar superior e inferior (Figura 51). O mesmo processo foi realizado para a diferença RB, BG e GB.

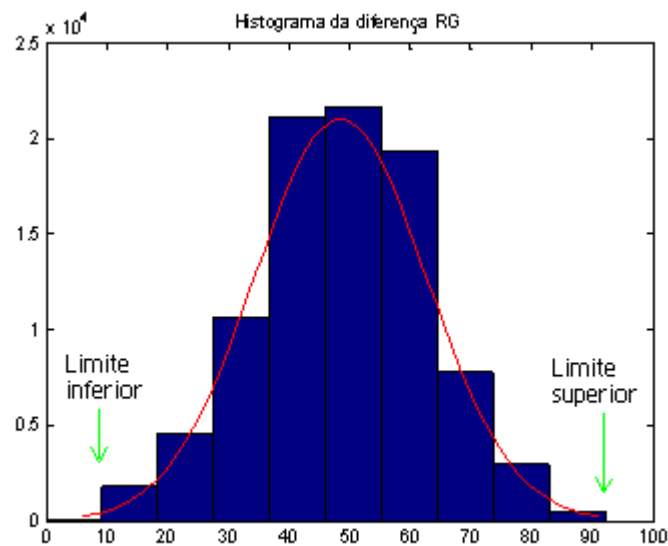


Figura 51 - Exemplo de histograma da diferença RG

A imagem diferença RG, com limiar global (L4) maior que 10 (\in_5) e menor que 80 (\in_2), apresenta apenas regiões que provavelmente são pele. Assim como a máscara resultante do canal R (L1) para valores maiores que 70 (\in_1).

A operação lógica “E” é aplicada na imagem proveniente do canal R e na imagem diferença RG. Dessa imagem são excluídos os objetos destacados na imagem resultante da

operação lógica “OU” aplicada nas imagens diferenças RB, GB e BG que utilizam as constantes $\epsilon_2 = 80$, $\epsilon_3 = 50$ e $\epsilon_4 = 60$, respectivamente.

A utilização dessa sequência promove uma imagem segmentada com ênfase nas regiões de pele.

2.4 Resultados da detecção de pixels de pele

Os percentuais da classificação dos pixels de pele obtidos na segmentação da imagem, para os dois bancos de imagens utilizados, são mostrados na Tabela 4.

	<i>Skin Locus</i>		Diferença RGB	
	Banco de imagens A	Banco de imagens B	Banco de imagens A	Banco de imagens B
PV (%)	93,44	94,60	94,99	96,42
FP (%)	26,67	68,62	10,87	35,90

Tabela 4 – Taxas de reconhecimento utilizando o *skin locus* e a máscara baseada nas diferenças dos canais R, G e B

Comparando-se os resultados obtidos pela diferença entre os canais RGB e o *Skin Locus*, verifica-se que a taxa de positivos verdadeiros foi praticamente a mesma, porém a taxa de falsos positivos foi reduzida pela metade utilizando-se a diferença RGB. Alguns exemplos desses resultados podem ser visualizados nas Figura 52 e Figura 53.



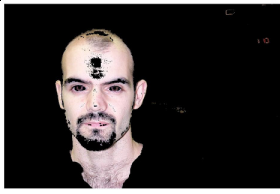


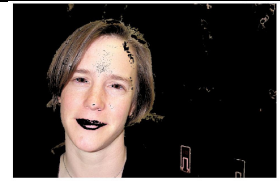


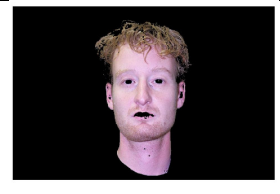
Imagem Original	Skin Locus	Diferença RGB
		
		
		

Figura 52 - Resultados da aplicação do *Skin Locus* e da diferença RGB nas imagens do Banco A

Imagem Original	Skin Locus	Diferença RGB
		
		
		

Figura 53 - Resultados da aplicação do *Skin Locus* e da diferença RGB nas imagens do Banco B

Ainda assim, após a aplicação deste procedimento de segmentação na imagem, outros objetos que não são pele podem estar presentes na imagem resultante.

Como o objetivo final é obter apenas a região de pele da face, adotou-se também a utilização de um método baseado nas somas verticais e horizontais das imagens para excluir os demais objetos (Figura 54).

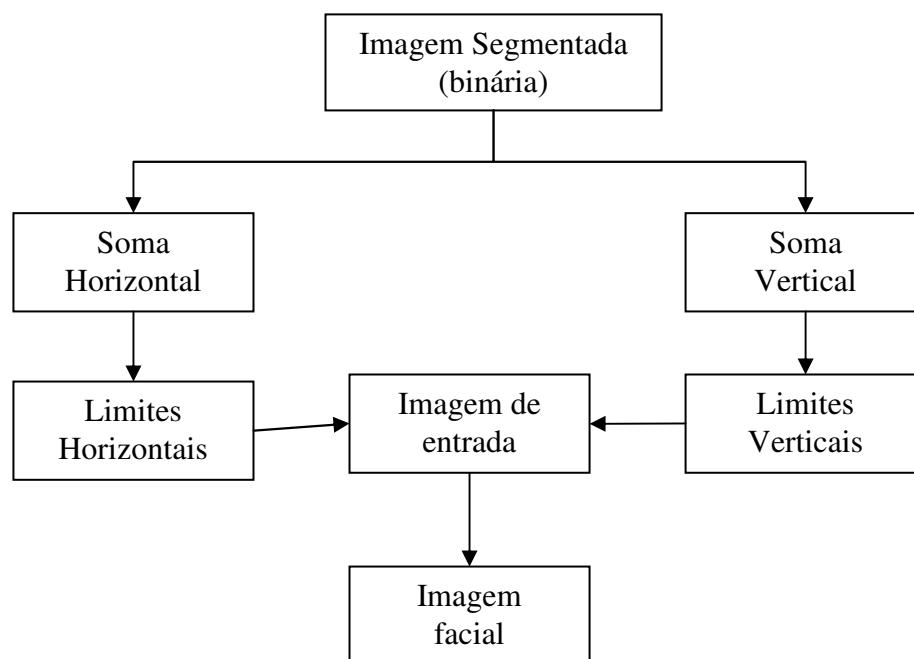


Figura 54 – Metodologia aplicada para eliminação de objetos não pertencentes à face

Extraíu-se os limites faciais das somas horizontais e verticais (Figura 55) da imagem binária resultante da diferença de canais em RGB.

Determinou-se um valor a partir do qual deveria ser pego o limite vertical e horizontal. Esse valor corresponde a 20% do máximo valor ocorrido nas projeções. Tal procedimento elimina objetos que não são pele de forma simples e rápida. Às vezes, a sombra criada pelo indivíduo em foco apresenta valores de R, G e B semelhantes às de pele, nesses casos, dependendo da área da sombra, torna-se impossível sua eliminação por meio dos limites (Figura 56).

Em seguida, fez-se o preenchimento das lacunas do objeto em foco e construiu-se uma imagem contendo apenas a região candidata à face (Figura 57).

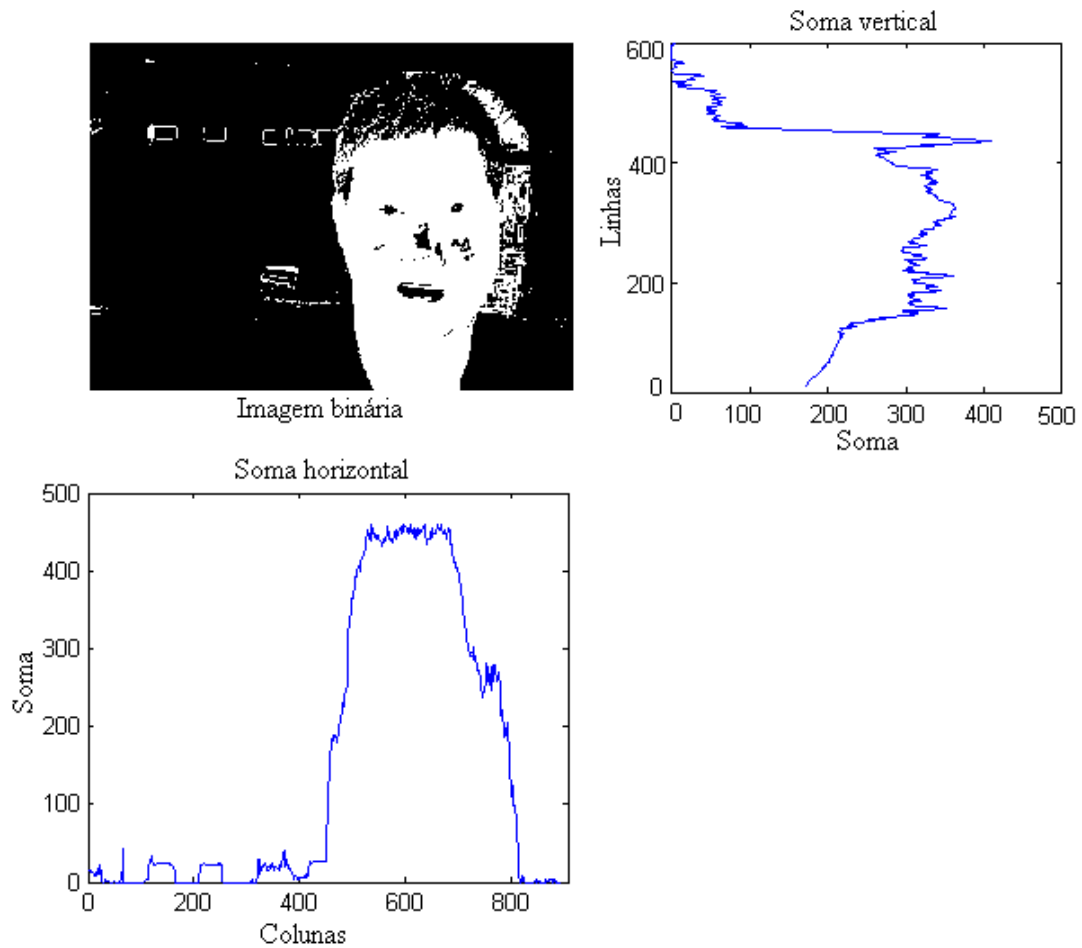


Figura 55 - Somas horizontal e vertical da imagem binária obtida do resultado da diferença de canais R, G e B



Figura 56 – Estabelecimento dos limites vertical e horizontal a partir das respectivas somas



Figura 57 – Extração da região candidata a face

No banco de imagens B, devido o elemento de fundo possuir tonalidade semelhante à pele, o uso de limites vertical e horizontal reduziu poucos elementos de fundo como pode ser visto na Figura 58. No momento em que se faz a varredura horizontal encontra-se os picos no início da imagem, por isso o limite não é aplicado, pois pode ser uma face.

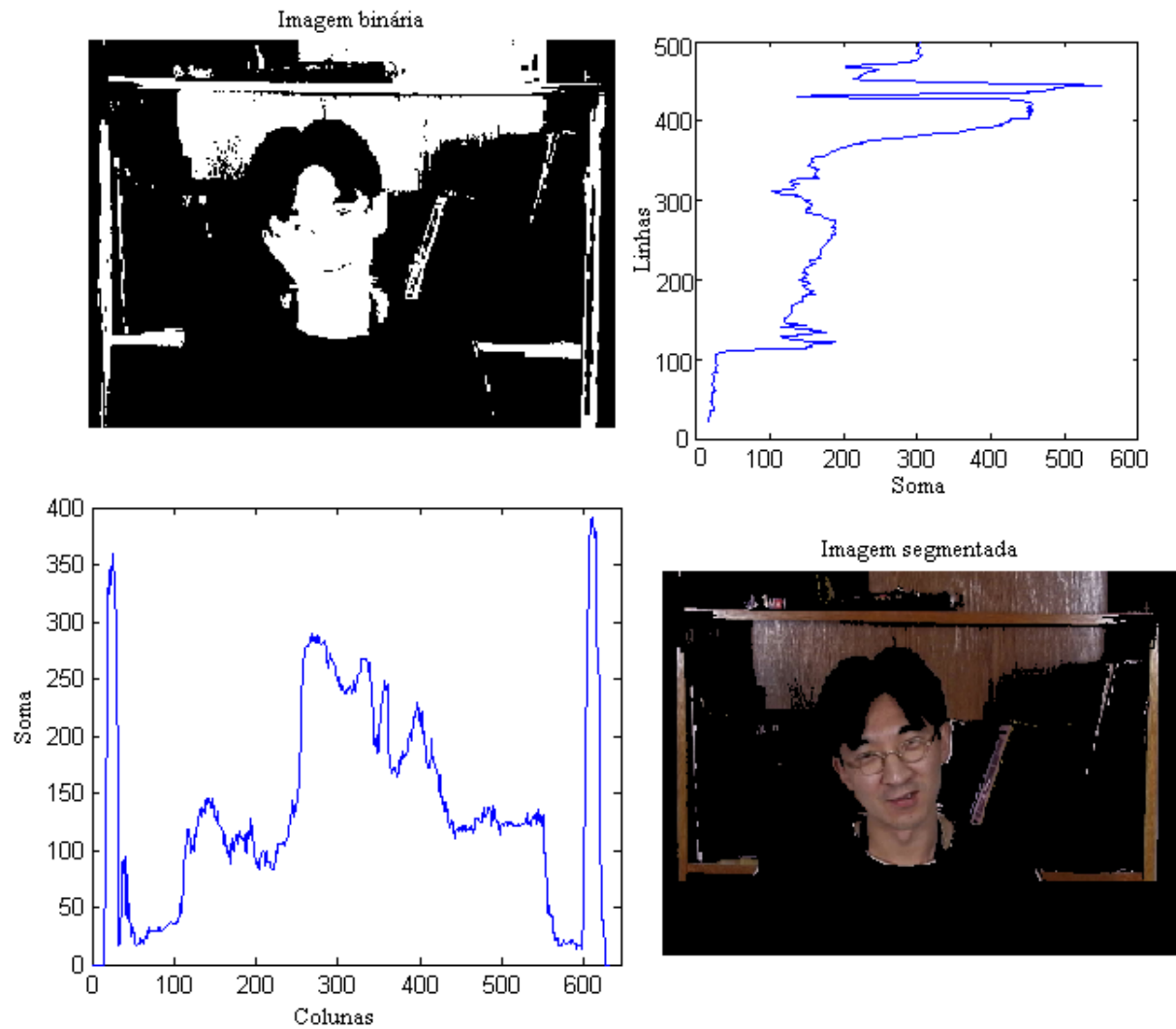


Figura 58 – Somas horizontal e vertical da imagem binária obtida do resultado da diferença de canais R, G e B, e a imagem segmentada após o uso dos limites obtidos a partir dessas somas. Figura do banco de imagens B.

Após o término da classificação dos pixels e extração das regiões de pele, buscou-se verificar quais regiões de pele detectadas correspondiam a face.

3 VERIFICAÇÃO DA FACE

3.1 Introdução

O cérebro humano possui um sistema de reconhecimento natural, o qual nos permite identificar pessoas conhecidas entre tantas outras pessoas. No momento em que entramos e contato com alguém que conhecemos, automaticamente, procuramos identificar particularidades dessa pessoa que nos façam saber quem ela é. Segundo Oliveira (1989, p. 48-49):

1 Uma pessoa vê um rosto que lhe parece familiar, mas por algum motivo não identifica imediatamente de quem se trata. O cérebro então registra os traços essenciais daquela imagem – o bigode, o formato da face e do nariz.

2 Com essas pistas, a memória busca retratos aparentados. Assim o cérebro compara imagens que vêm com as lembranças de um ex-chefe, de um antigo médico da família, de um primo distante, de um professor dos tempos de colégio. Este último possui o mesmo formato de rosto e tem nariz e cabelos iguais. Mas na imagem gravada na memória o seu rosto aparece de barba.

3 Sem ter certeza absoluta, o cérebro se decide pelo professor, cujo rosto é o mais parecido. A partir daí, surgem lembranças [...] e tudo vai reforçar a decisão de que é de fato o professor, só que sem barba.

De acordo com o exposto, cada ser humano possui características próprias que os diferem um dos outros. Sejam essas características físicas ou psicológicas, o fato é que até mesmo os gêmeos idênticos são distinguíveis.

Desta forma, no processo de reconhecimento natural podemos enumerar algumas características que usamos para classificar as pessoas, como:

- 1 Cor de pele;
- 2 Formato do rosto;
- 3 Características dos olhos;
- 4 Características do nariz;

5 Características da boca;

6 Barba, óculos, cabelos, outros.

Nos capítulos anteriores foi abordada a dificuldade de identificação automática da primeira característica, a cor de pele. As dificuldades na identificação das outras características serão abordadas nessa seção, sendo enfatizadas as técnicas que utilizam operadores morfológicos devido a sua simplicidade e rapidez de execução.

As etapas de um sistema genérico de identificação da face em imagens digitais podem ser demonstradas no esquema da Figura 59.

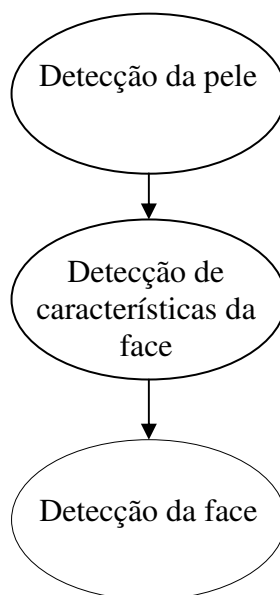


Figura 59 – Etapas de um sistema genérico de identificação/processamento de face

Em um sistema de verificação de faces, para regiões candidatas, procuram-se identificar a existência de características como olhos, nariz e boca.

Os olhos são informações cruciais para validação de uma região como face. Sua aparência tem menor variação em relação às outras características da face, ao longo dos anos.

Conhecendo-se a localização dos olhos é possível estimar a escala e a inclinação da face; além disso, facilita a localização de outras características da face como o nariz e a boca.

Dessa forma, os olhos tornam-se a base fundamental da etapa de detecção das características da face.

A detecção da face pode ser feita usando a relação geométrica entre os componentes da face (olhos, nariz e boca).

Dentre os métodos propostos para detecção de olhos em imagem de faces humanas, pode-se destacar as técnicas que utilizam modelos deformáveis (KASS *et al*, 1988), algoritmos de análise do componente principal (JOLLIFE, 1986), operadores morfológicos (HAN *et al*, 2000), algoritmos usando variância (FENG e YUEN, 2001) ou aplicação da transformada de Hough (CHIANG *et al*, 2004).

Os algoritmos que usam operadores morfológicos apresentam grande popularidade devido a sua simplicidade de execução. Os resultados obtidos por esses algoritmos mostram-se muito eficientes quando aplicados em regiões específicas para localização das características da face humana.

Este capítulo está apresenta as técnicas de extração das características faciais por morfologia matemática mais comuns na literatura, bem como a técnica de extração das características faciais a partir de diferenças no espaço RGB. Enfatiza também a técnica de rotulação de objetos na imagem para classificação dos mesmos como características faciais.

Após a apresentação das técnicas utilizadas, mostra-se como elas foram utilizadas em conjunto para a determinação dos olhos, sobrancelhas, nariz e boca.

3.2 Extração das características da face por morfologia matemática

Segundo Han *et al* (2000), as sobrancelhas, olhos, narinas e boca estão sempre nas regiões mais escuras de uma imagem em nível de cinza.

Para se obter a localização dos olhos, é feita uma busca por vales na imagem, ou seja, busca-se as regiões onde a imagem apresenta os menores níveis de cinza, pois nessas regiões encontram-se as íris (Figura 60).

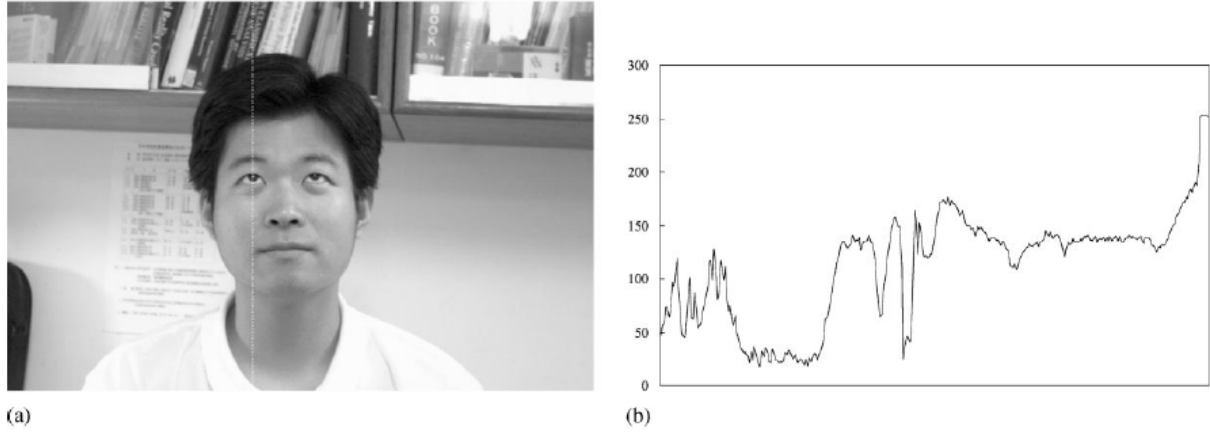


Figura 60 – A busca por vales na imagem. (a) Imagem original e (b) o perfil dos níveis de cinza da imagem na coluna selecionada
Fonte: Han et al (2000)

Os vales da imagem podem ser obtidos por meio da operação morfológica de fechamento (Figura 62). A imagem resultante da aplicação da operação morfológica é descrita por:

$$E_1 = T_1(X \ominus (X \bullet Sh)) \quad (17)$$

$$E_2 = T_2(X^{1/2} \ominus (X^{1/2} \bullet Sh)) \quad (18)$$

$$E_3 = T_3(X \ominus (X \bullet Sv)) \quad (19)$$

$$E_4 = T_4(X^{1/2} \ominus (X^{1/2} \bullet Sv)) \quad (20)$$

$$E = E_1 \vee E_2^2 \vee E_3 \vee E_4^2 \quad (21)$$

Onde:

X é a imagem de entrada, em nível de cinza.

S_v é um elemento estruturante vertical com tamanho 7×1 .

S_h é um elemento estruturante horizontal com tamanho 1×7 .

\ominus é a operação XOR (OU-exclusivo) seguida da operação de negação⁴.

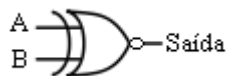


Figura 61 – Símbolo da operação lógica XOR, da álgebra booleana, seguida da operação de negação.

• é a operação morfológica de fechamento.

\vee é a operação OU.

T é o limiar, corresponde à média de cada imagem.

E_1 a E_4 são imagens resultantes das operações morfológicas.

E é a imagem resultante final dos vales.

Os subscritos 2 e $\frac{1}{2}$ significam interpolado por 2 e subamostrado por 2, respectivamente.

A utilização dessa sequência de operações resulta em uma imagem com os pixels da íris destacados, porém outros artefatos também aparecerão como: letras, janelas, etc.

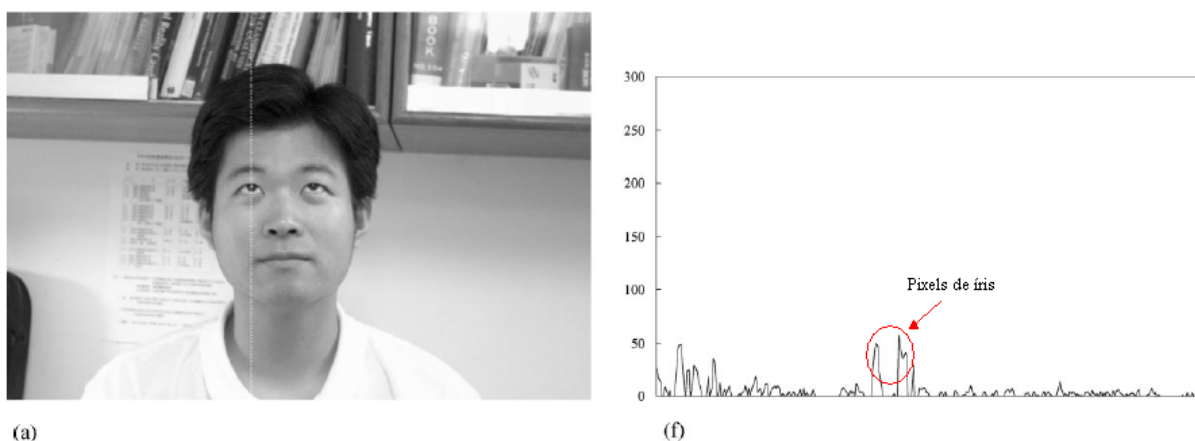


Figura 62 – Resultado das operações morfológicas. (a) Imagem original e (f) perfil do nível de cinza da imagem vale final, com os pixels de íris destacados

Fonte: Han *et al* (2000) com adaptações

Para eliminação de pixels indesejados, Han *et al* (2000) propôs a utilização de regras de exclusão considerando as dimensões dos artefatos e utilização da operação morfológica de

⁴ Também pode ser encontrada na literatura como XNOR.

dilatação, com diferentes elementos estruturantes, de acordo com a orientação dos segmentos da imagem.

Ademais, Han *et al* (2000) propôs regras para identificação dos segmentos correspondentes aos olhos. Tais regras consideram a geometria da face como a distância entre os olhos, narinas e boca.

Os resultados apresentaram desempenho satisfatório na identificação da face para diferentes escalas e inclinações. Porém, não é feito um pré-processamento para redução de falsos candidatos, contribuindo para que outros artefatos sejam considerados face, como mostra a Figura 63.



Figura 63 – Detecção facial de Han *et al* (2000)

Fonte: Han *et al* (2000)

Na mesma linha de raciocínio de Han *et al* (2000), Perlibakas (2003) sugeriu algumas diferenças na obtenção da imagem com vales. Estes trabalhos apresentaram diferentes regras para o cálculo do limiar e classificação dos segmentos. A Equação 22 mostra o cálculo efetuado para a obtenção da imagem resultante final dos vales.

$$E = (E_1 + E_2^2 + E_3 + E_4^2)/4 \quad (22)$$

Segundo essas regras, o menor tamanho de face detectável é de 50 pixels e o maior tamanho é de 150. A partir desses tamanhos, são definidas outras 25 regras para determinar quais tamanhos podem ter os olhos e a boca.

Depois de localizar as características da face, Perlibakas (2003) fez uso de contornos ativos ⁵, Figura 64, para determinar o contorno exato da face.



Figura 64 - Detecção facial de Perlibakas (2003)

Fonte: Perlibakas (2003)

Ambos os algoritmos foram testados em bancos de imagens em níveis de cinza como o BioID (2003) e FERET (PHILIPS e RAUSS, 1996).

3.3 Extração das características da face por diferenças em RGB

Após a etapa de segmentação da imagem de entrada, utilizando o método de diferença dos canais RGB proposto no capítulo II e visando a validação da região facial é apresentada uma metodologia para extração das características da face. Essa metodologia é apresentada em duas fases: a primeira fase extrai a região dos olhos e a segunda, a região das sobrancelhas, nariz e boca.

No diagrama de blocos da Figura 65 são ilustradas as fases da metodologia proposta para extração de características faciais a partir da diferença entre R e G.

⁵ Também conhecido como *snakes* (KASS, 1988).

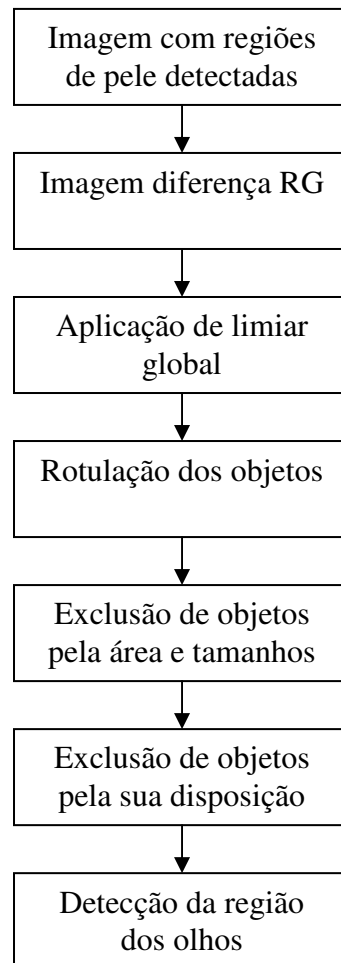


Figura 65 - Primeira fase da metodologia proposta: detecção da região dos olhos

Na primeira fase dessa metodologia utilizou-se a imagem diferença RG e aplicou-se um limiar global de Otsu⁶. O resultado apresentou uma imagem binária com destaque das regiões com baixos níveis de pixels vermelhos, Figura 66 à Figura 69.

⁶ Método padrão para o cálculo de limiar efetuado pelo MatLab.



Figura 66 – Resultado da diferença RG com limiar global (banco de imagens A)

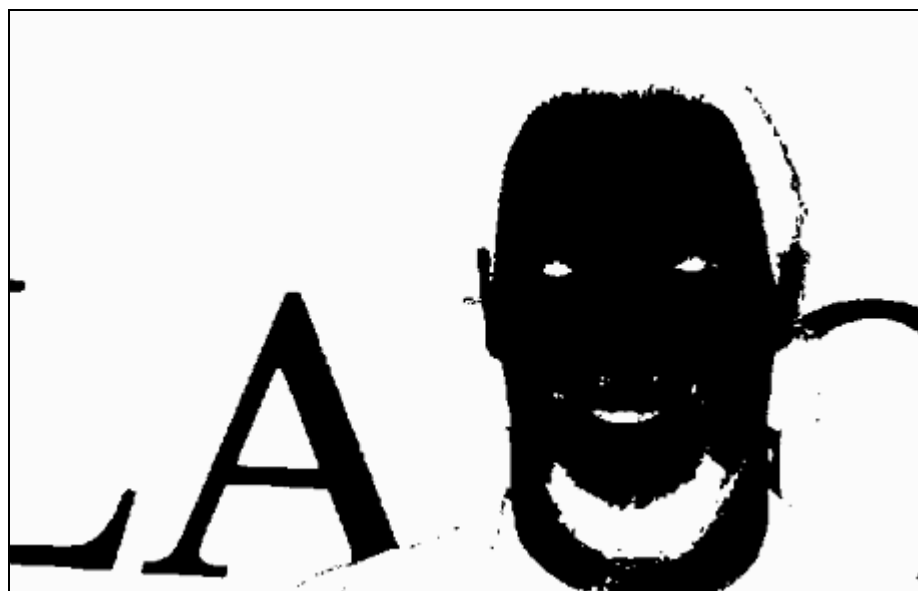


Figura 67 – Resultado da diferença RG com limiar global (banco de imagens A).



Figura 68 – Resultado da diferença RG com limiar global (banco de imagens B)



Figura 69 – Resultado da diferença RG com limiar global (banco de imagens B)

Os objetos presentes na imagem binária resultante foram rotulados e aplicaram-se regras para exclusão de objetos de acordo com sua área, dimensões e disposição na imagem.

As regras de exclusão, baseadas nas áreas e dimensões dos objetos, utilizadas foram:

$$\text{Área} < 15 \quad (23)$$

$$\text{Dimensões} < 10 \ \& \ \text{Dimensões} > (\text{Dimensões da imagem})/2 \quad (24)$$

$$\text{Dimensão vertical} > 2x (\text{Dimensão horizontal}) \quad (25)$$

$$\text{Dimensão horizontal} > 3x (\text{Dimensão vertical}) \quad (26)$$

$$\text{Dimensão vertical} > 50 \quad (27)$$

$$\text{Dimensão horizontal} > 80 \quad (28)$$

As disposições dos objetos presentes na imagem foram analisadas para definir o potencial par de olhos, como na Figura 70.

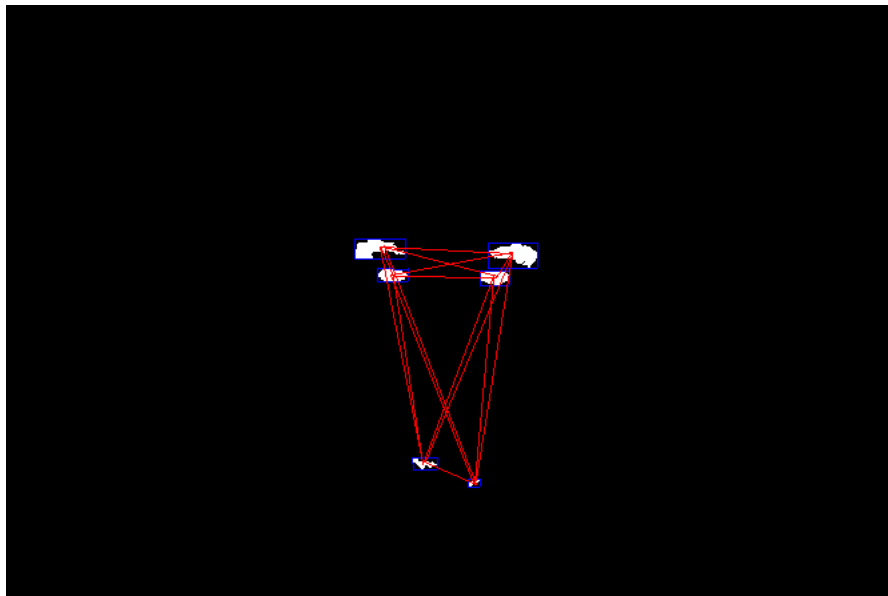


Figura 70 – Relação entre os objetos presentes na imagem

A relação entre os objetos presentes na imagem é analisada a partir das distâncias entre seus centróides. A regra utilizada para os centróides dos candidatos ao par de olhos foi:

$$i_{\text{Objeto1}} - i_{\text{Objeto2}} > 50 \quad (29)$$

$$j_{\text{Objeto1}} - j_{\text{Objeto2}} < 50 \quad (30)$$

Onde i é a coordenada horizontal dos objetos e j é a vertical. O valor 50 foi obtido a partir da análise das imagens dos bancos de imagens, onde o foco não são os olhos e sim a face da pessoa.

Nessa abordagem, os olhos sempre são perceptíveis na imagem binária. Para indivíduos que apresentam tonalidade de cabelos escuros, as sobrancelhas também aparecem

na imagem binária. Em indivíduos que apresentam expressões faciais como o sorriso, a abertura da boca, também, estará presente na imagem binária.

Após o processo de exclusão, é analisada a área dos objetos candidatos a olhos. Na imagem em nível de cinza, extraí-se a região respectiva a área de cada objeto e verifica-se o comportamento desses pixels. Espera-se encontrar um comportamento semelhante ao da Figura 71.

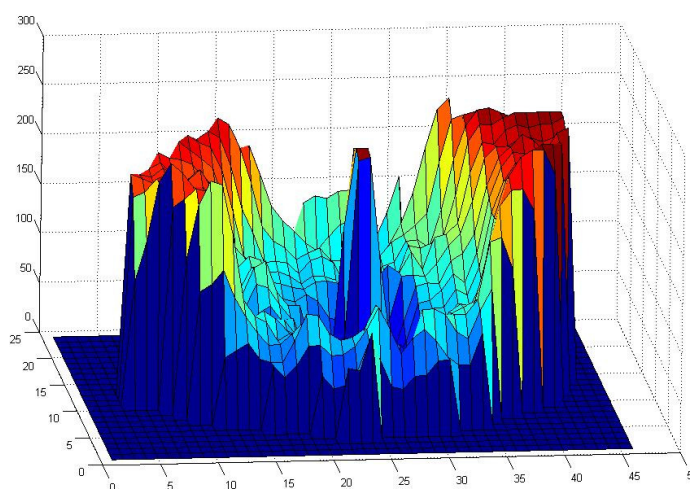


Figura 71 – Comportamento dos pixels na região dos olhos

A Figura 71 apresenta o comportamento dos pixels da região dos olhos onde ocorre a sequência de pico → vale → pico → vale → pico. Essa não é a principal sequência para essa região, mas é a mais comum, visto que muitas vezes a região da pupila tende a apresentar níveis de branco devido a iluminação do ambiente ou do *flash* da câmera. Logo, essa sequência corresponde a: esclera → íris → pupila → íris → esclera (Figura 72).

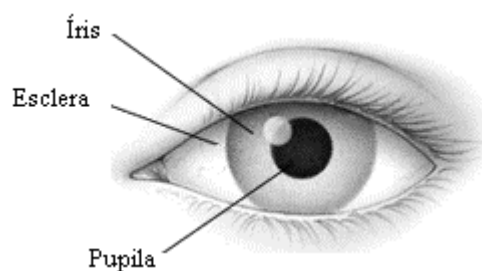


Figura 72 – Estrutura do olho humano

Fonte: <<http://www.compuland.com.br/anatomia/olho.htm>>. Com adaptações.

Essa metodologia reduz a quantidade de candidatos a olhos, facilitando sua localização. Porém, apresenta-se insuficiente para a determinação da boca.

Segundo a pesquisa de LIN (2007) na face existe uma relação de distâncias entre os olhos e a boca. Em uma imagem facial frontal, as distâncias entre os olhos e a boca formaram um triângulo isósceles. Procedeu-se a investigação para localização da região da boca seguindo essa afirmação.

Contudo, para os dois bancos de imagens utilizados nesse trabalho, mesmo normalizando somente as regiões da face, não se obteve as dimensões do triângulo isósceles entre os olhos e boca.

Lançou-se mão de métodos que utilizam operações morfológicas em conjunto com a metodologia proposta para detectar as demais características da face.

O emprego da operação morfológica na imagem é demonstrado no diagrama de blocos da Figura 73.

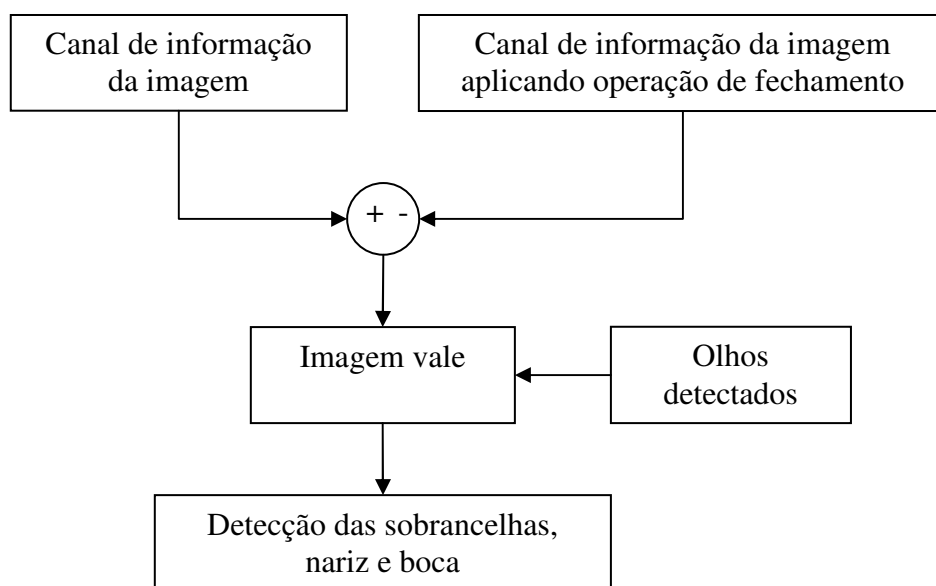


Figura 73 – Diagrama de blocos da utilização de operadores morfológicos

A operação morfológica de fechamento foi aplicada em cada canal da imagem (R, G e B) para se obter as imagens diferenças entre esses resultados e seus respectivos canais. Somaram-se essas diferenças para construção de uma imagem vale final.

Optou-se por essa metodologia, pois se verificou que dessa forma, na imagem diferença final, os olhos apresentam mais pixels do contorno inferior. Quando se utilizou a imagem em nível de cinza da entrada, ao tentar estabelecer um limiar global para todas as imagens, na maioria dos casos o contorno inferior ou superior não era evidenciado. Se diminuísse o valor do limiar, para evidenciar o contorno dos olhos, o contorno do nariz se conectava aos olhos, Figura 74.



Figura 74 – Contorno do nariz conectado aos olhos. O limiar aplicado foi calculado pelo método de Otsu pelo MatLab

Utilizando-se a soma da resposta de cada canal, os pixels dos olhos assumiram valores superiores ao contorno do nariz e sua forma elíptica ficou mais evidente (Figura 75 (b)).



Figura 75 – Resultado final a aplicação morfológica de fechamento usando a imagem de entrada em (a) níveis de cinza e (b) a soma dos resultados nos canais R, G e B

Ao aplicar um limiar maior que 100 em todas as imagens, a resposta foi semelhante a Figura 76, onde as características da face não interferiam uma nas outras.



Figura 76 – Imagem binária resultado da aplicação de limiar maior que 100

Com base nas informações das regiões de olhos detectadas previamente pela imagem diferença RG, pode-se identificar a região da boca, nariz e sobrancelhas, examinando-se suas disposições em relação aos olhos (Figura 77) na imagem vale final.

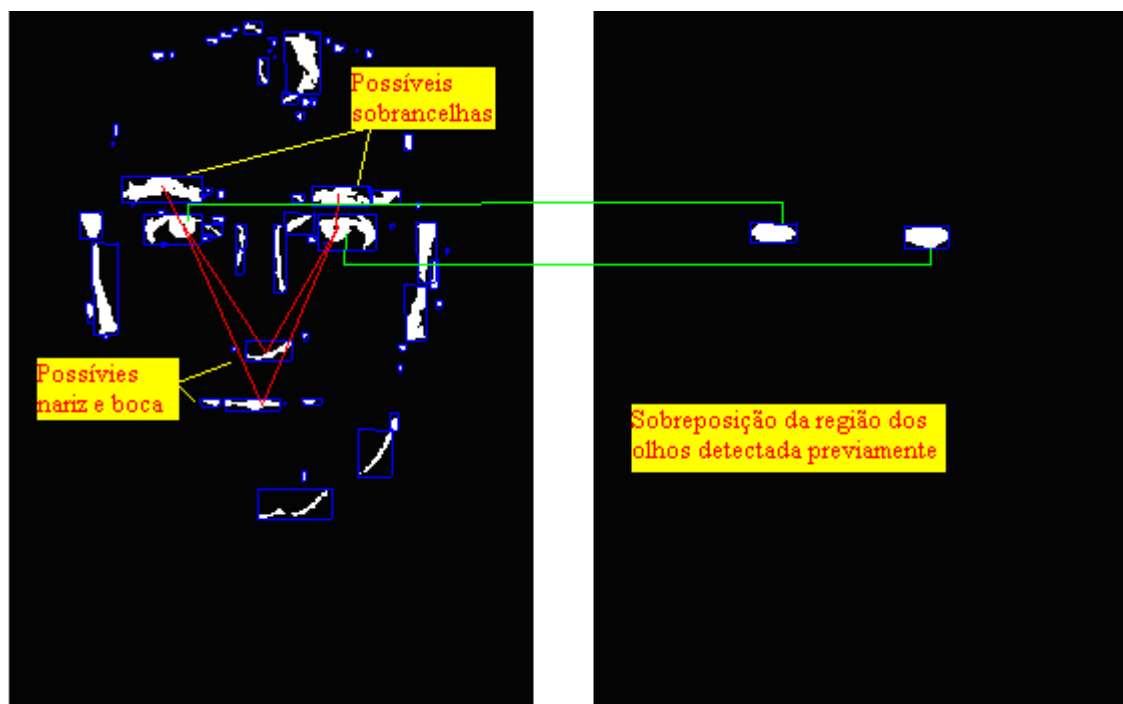


Figura 77 – Classificação dos objetos para obtenção das sobrancelhas, nariz e boca

A classificação das outras características (sobrancelhas, nariz e boca) em relação aos olhos é feita considerando-se uma hierarquia de disposição na face:

- As sobrancelhas apresentam-se como objetos acima dos olhos.
- Os traços do nariz estarão abaixo dos olhos, na posição mediana entre eles.
- Os traços da boca localizam-se na mesma mediana, logo abaixo dos traços do nariz.

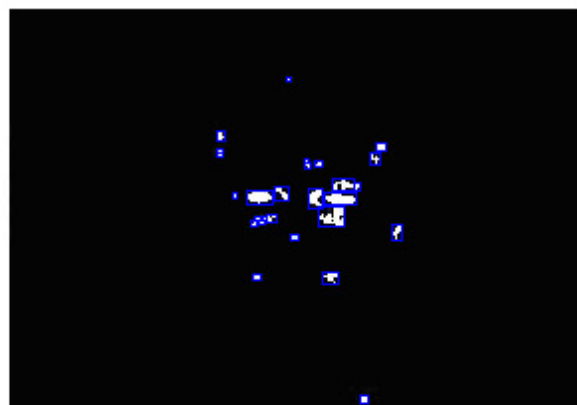
3.4 Resultados da detecção de face

Aplicando-se essa metodologia no banco de imagens A não foi possível a verificação da face em 5% das imagens, pois elas apresentavam oclusões dos olhos. Para o banco de imagens B, esse resultado foi de 13%. Nesse banco, além das oclusões de olhos, houve imagens em que a face encontrava-se em perfil.

Exemplos de resultados dos casos de acerto, para o banco de imagens A, são ilustrados nas Figura 78 e Figura 79.



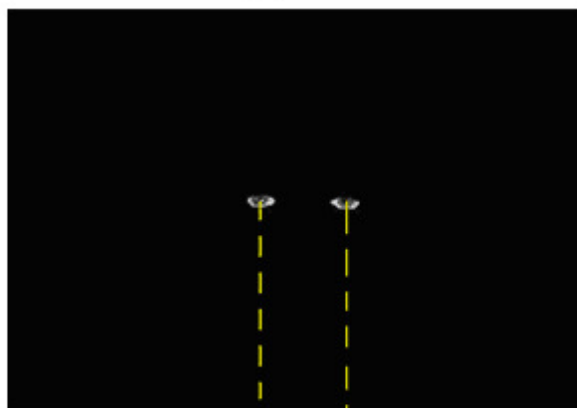
Imagem original
(a)



Deteção dos olhos por diferença RG
(d)



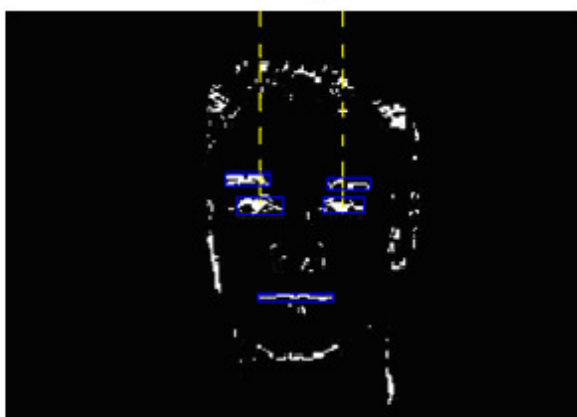
Pixel classificados como pele
(b)



Olhos detectados
(e)



Imagem após aplicação dos limites vertical e horizontal
(c)



Deteção das características da face
(f)

Figura 78– Resultados dos processos de detecção dos olhos, sobrancelhas e boca para o banco de imagens A

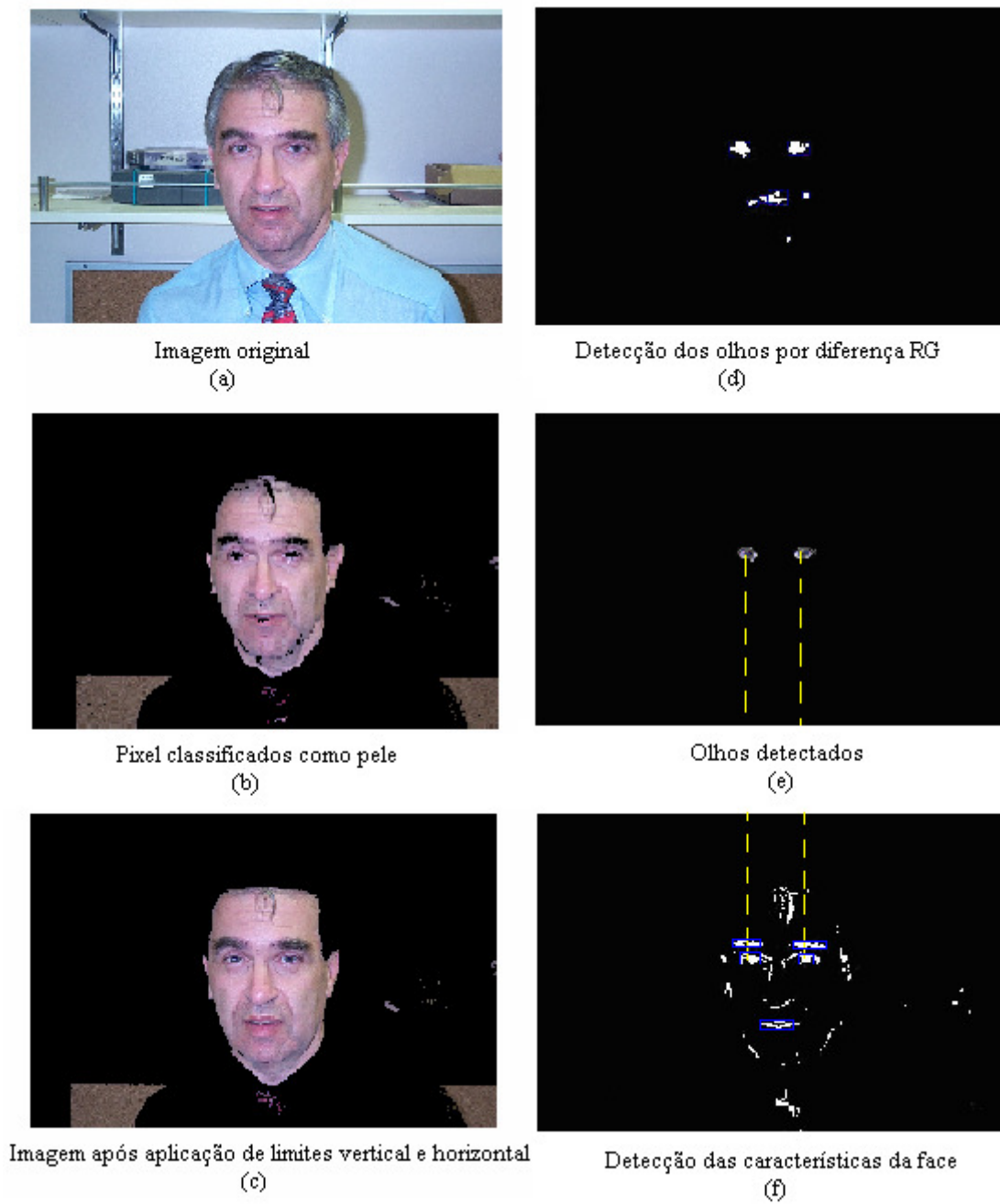
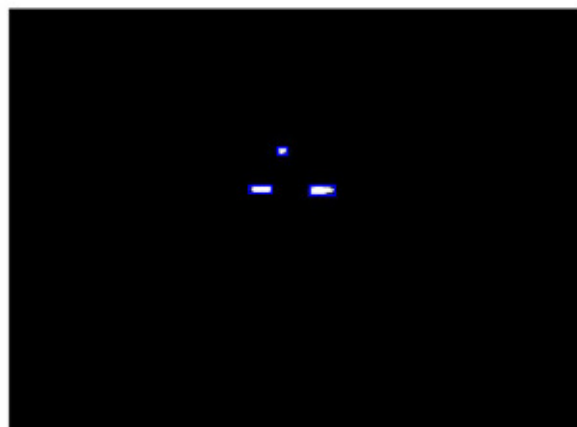


Figura 79 – Resultados dos processos de detecção dos olhos, sobrancelhas e boca para o banco de imagens A

Para o banco de imagens B, as Figura 80 e Figura 81 apresentam exemplos dos resultados obtidos.



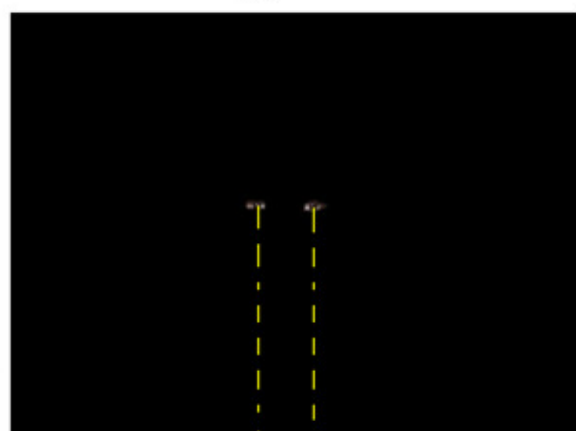
Imagem original
(a)



Deteção dos olhos por diferença RG
(d)



Pixels classificados como pele
(b)



Olhos detectados
(e)

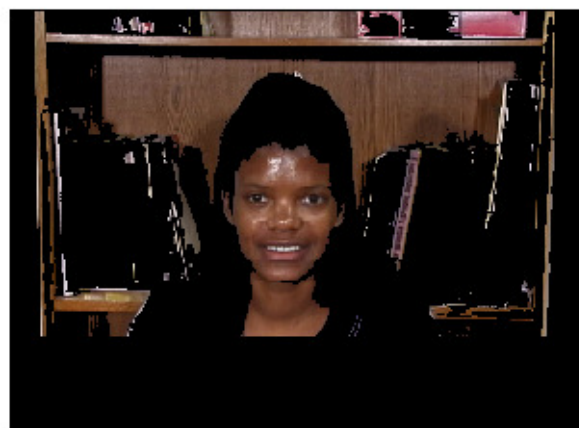
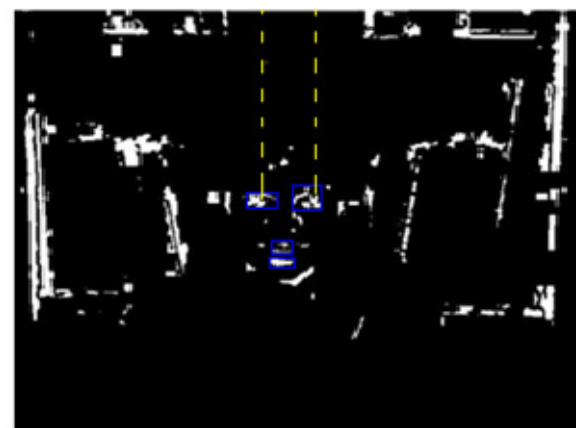


Imagem após aplicação de limites vertical e horizontal
(c)

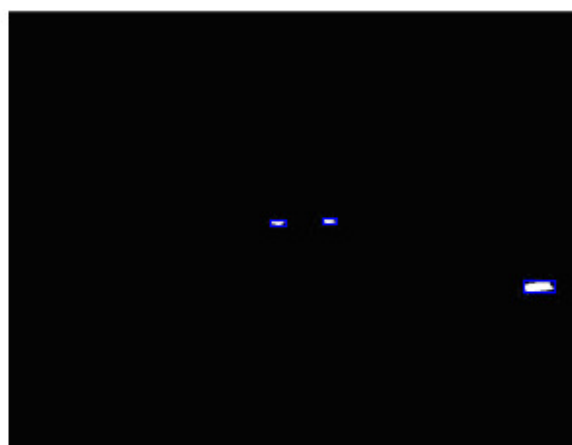


Deteção das características da face
(f)

Figura 80 – Resultados dos processos de detecção dos olhos, nariz e boca para o banco de imagens B



Imagem original
(a)



Deteção dos olhos por diferença RG
(d)



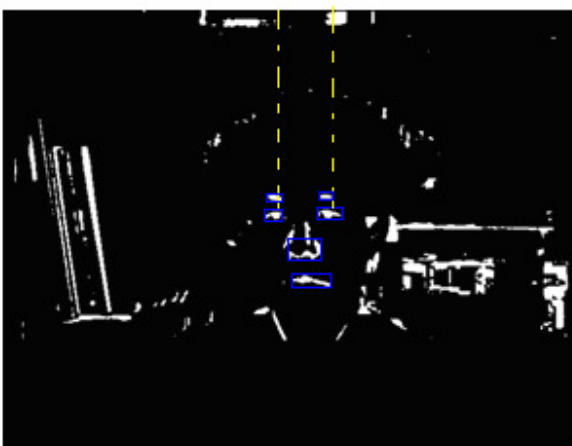
Pixels classificados como pele
(b)



Olhos detectados
(e)



Imagem após aplicação de limites vertical e horizontal
(c)



Deteção das características da face
(f)

Figura 81 – Resultados dos processos de detecção dos olhos, sobrancelhas, nariz e boca para o banco de imagens B

4 CONCLUSÃO E TRABALHOS FUTUROS

Nesta dissertação foi proposto um novo algoritmo que faz a detecção de pixels de cor de pele baseada na combinação das diferenças das informações contidas nos canais R, G e B de uma imagem digital de alta qualidade. O algoritmo em questão apresentou um desempenho superior ao método largamente referenciado na literatura, o *skin locus*. Na análise, o *skin locus* mostrou-se eficiente somente quando utilizado em imagens de baixa qualidade, do tipo *webcam*, mas quando aplicado em imagens digitais de alta qualidade, típicas de câmeras digitais, apresentou um desempenho insatisfatório, mesmo após os ajustes dos limiares de decisão.

O método de detecção de cor de pele proposto mostrou-se eficaz na detecção de pele de pessoas tanto com características negras quanto caucasianas. A principal dificuldade encontrada ocorreu quando o mesmo foi utilizado com imagens onde no fundo existiam objetos com a mesma cor da pele da pessoa, tais como: madeira e sombras. Mesmo tendo essa limitação, o algoritmo de detecção desenvolvido reduziu significativamente os falsos positivos (fundo reconhecido como pele) se comparado ao *skin locus*.

Por trabalhar somente com as diferenças entre as informações dos canais dos R, G e B e utilizar-se de limiares de decisão para identificar os pixels de cor de pele, o método apresenta baixa complexidade computacional, podendo, assim, ser utilizado como parte da grande maioria dos sistemas de reconhecimento facial expostos na literatura e que se baseiam em identificação dos pixels de cor de pele.

Também foi apresentado, nesta dissertação, um algoritmo de validação da máscara facial de cor de pele gerada a partir do algoritmo de detecção de cor de pele baseado em diferenças das informações contidas nos canais R, G e B, o qual restringiu o espaço de busca da face. Tal procedimento de validação foi baseado na identificação de características que só

estão presentes na face (olhos, boca e nariz). Neste, primeiramente buscou-se o par de olhos, depois o nariz e, por último a boca. Com a utilização deste método de validação pode-se definir com maior precisão qual região de cor de pele corresponde à região da face, a qual poderá, posteriormente, ser utilizada para identificação da face.

A verificação da face por um algoritmo baseado na diferença dos canais R e G destaca as regiões de vales, facilitando a identificação dos olhos. Porém, a identificação dos olhos torna-se mais difícil quando estes apresentam interferência de outros objetos, como o cabelo. Casos em que os olhos têm conexão com o fundo da imagem ou com o cabelo, quando se realiza a binarização da imagem, tornam-se um único objeto. Ainda que seja utilizado o método da imagem vale por morfologia matemática, se os olhos estiverem conectados ao fundo, torna-se difícil sua identificação.

Propõe-se, como trabalhos futuros, o estudo dos métodos propostos nesta pesquisa para imagens que contenham mais de uma face. Embora, o algoritmo de detecção de pele seja indiferente à quantidade de pessoas presentes na imagem. O método de verificação da área da face não é tão generalista.

Considerando que os métodos propostos compõem partes de um sistema de reconhecimento computacional, sugere-se que seja dada continuidade nesta linha de pesquisa, investigando métodos de classificação das faces em conjunto com os apresentados nesta dissertação.

REFERÊNCIAS

BAO, Pham; KIM, Jin; NA, Seung. Fast multi-face detection in color images using Fuzzy logic. Proceedings of 2005 International Symposium on Intelligent Signal Processing and Communication Systems. Hong Kong, 2005.

BELHUMEUR, V; HESPANHA, J e KRIEGMAN, D. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Transaction on Pattern Analysis and Machine Intelligence, pag 711-720, 1997.

BioID. The BioID Face Database. 2003. Disponível em <<http://www.humanscan.de/support/downloads/facedb.php>>. Acessado em Nov 2007.

CAHI, D. e Ngan, K. N. Face segmentation using skin-color map in videophone applications. IEEE Transaction on circuit and systems for video technology, vol. 9, pag 551-564, 1999.

CHIANG, Pei-Ying; LIAO, Wen-Hung; LI Tsai-Yen. Automatic Caricature Generation by Analyzing Facial Features. Asian Conference on Computer Vision, Jeju Island, 2004. pag 27 – 30.

COOTES, T. F. e TAYLOR, C. J. On representing edge structure for model matching. Proceedings Computer Vision on Pattern Recognition, vol 1, pag 1114-1119, 2001.

DORNAIKA, F. e AHLBERG, J. Face and facial feature tracking using deformable models. International Journal of Image and Graphics, pag 499-532, 2004.

FENG, G. C.; YUEN, P. C. Multi-cues eye detection on gray intensity image. Pattern Recognition, Vol 34, 2001.

GONZALEZ, R. C.; WOODS, R.E.; Digital Image Processing, 3 ed. Upper Saddle River, New Jersey: Prentice-Hall, Inc., 2007.

HAN, C. C; LIAO, H; YU, L; HUA, L. Fast face detection via morphology-based pre-processing. Pattern Recognition, Elsevier, 2000.

HOU, X; LI, S; ZHANG, H. e CHENG, Q. Direct appearance model. Proceeding Computer Vision and Pattern Recognition, vol 1, pag 828-833, 2001.

JOLLIFFE, I. T; Principal component analysis. Springer, Nova York, 1986.

KASS, Michael; WITKIN, Andrew; e TERZOPOULOS, Demetri. Snakes: active contour models. International Journal of Computer Vision, 1988.

KOVAC, J; PEER, P; SOLINA, F. Human Skin Colour Clustering for Face Detection. EUROCON – 2003 – International Conference on Computer as a Tool, 2003.

LIN, Chiunhsiun. Using HSV Color Space and Neural Network for Face Detection with Various Illuminations. Proceedings of the 2007 WSEAS International Conference on Computer Engineering and Applications. Australia, 2007.

MARTINKAUPPI, B. Face Colour under Varying Illumination – Analysis and Applications. Ph. D. Thesis, University of Oulu, 2002. Disponível em: <<http://herkules.oulu.fi/isbn9514267885/>>. Acessado em 20 Mar 2007.

OHTSUKI, T; HEALEY, G. Using color and geometric models for extracting facial features. Journal of Imaging Science and Technology, 42(6):554-561, 1998.

OLIVEIRA, Lúcia Helena. O endereço da inteligência. Superinteressante, p. 48-49, dez. 1989. 1 CD-ROM.

PERLIBAKAS, Vytutas. Automatical detection of faces features and exact face contour. Elsevier. Lituânia, 2003.

PHILLIPS, P.J.; RAUSS, P. FERET recognition algorithm development and test report. 1996. Disponível em <<http://www.frvt.org/FERET/default.htm>>.

RUSKEY, Frank; WESTON, Mark. A survey of Venn diagrams. The Electronic Journal of Combinatorics. Ed. julho 2005, DS#5. Disponível em: <<http://www.combinatorics.org/Surveys/ds5/VennEJC.html>>.

SAXE, D e FOULDS, R. Toward robust skin identification in video images. Proceedings of 2nd International Conference on Automatic Face and Gesture Recognition, Killington, pag 379-384, 1996.

SCLAROFF, S. e ISIDORO, J. Actives blobs. International Journal of Computer Vision, pag 1146-1153, 1998.

SINGH, S; CHAUHAN, D. S; VATSA, M; SINGH, R. A Robust Skin Color Based Face Detection Algorithm. Tamkang Journal of Science and Engineering. Vol. 6, No. 4, pag 227 – 23, 2003.

TURK, M.A; PENTLAND, A. P. Face Recognition Using Eigenfaces. Proceedings of International Conference on Pattern Recognition, pag 586-591, 1991.

WILDER, J; PHILLIPS, P.; JIANG,C.; WIENER, S. Comparison of visible and infra-red imagery of face recognition. In: Proc. of International Conference on Automatic Face and Gesture Recognition, pag 215 -220, 2002.

YANG, Ming-Hsuan; KRIEGMAN, David J; AHUJA, Narendra. Detecting Faces in Images: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 24, N° 1, Janeiro 2002.

ZHANG, J; HE, M; CHEN, J. RGB Level face detection. Relatório de projeto, 2002. Disponível em: <<http://scien.stanford.edu/2002projects/ee368/Project/reports/ee368group13.pdf>>. Acessado em: 23 Jun 2007.

ZHAO, W.; CHELLAPPA, R.; PHILLIPS, P. Subspace linear discriminant analysis for face recognition. Technical Report CAR-TR 914, University of Maryland, 1999.

ZHAO, Wenyi e CHELLAPPA, Rama. Face processing. Elsevier Inc. 2006

ZHOU, S; KRUEGER, V. e CHELAPPA, R. Probabilistic recognition of human faces from vídeo. Computer Vision and Image Understanding, pag 214-245, 2003.

ANEXO A

AQUISIÇÃO DE IMAGENS

As imagens utilizadas para o desenvolvimento desse trabalho foram capturadas a partir de bancos de dados disponibilizados na internet.

Constatou-se que a grande maioria dos bancos de imagens disponibilizadas na internet corresponde a imagens em níveis de cinza ou apresentam imagens coloridas provenientes de dispositivos de captura antigos, resultando em imagens de baixa qualidade.

1 Bancos de Imagens

Utilizou-se dois bancos de imagens sendo estes nomeados: Banco de Imagens A e Banco de Imagens B.

1.1 Banco de imagem A

Este banco de imagens está disponível no endereço eletrônico: <http://www.vision.caltech.edu/html-files/archive.html>. É composto de 450 imagens faciais, com uma única pessoa por imagem. As imagens têm predominância de indivíduos caucasianos.

As escalas do tamanho das faces são variadas, apresentando estas, também, inclinações em alguns casos.

Não foram mencionadas as condições de iluminação na aquisição das imagens que se apresentam diferentes para cada imagem, Figura 82.

O fundo das imagens não é controlado e apresenta, na sua maioria, um fundo complexo. Por complexidade do fundo compreende-se a presença de artefatos com cores muito semelhantes aos de pele.

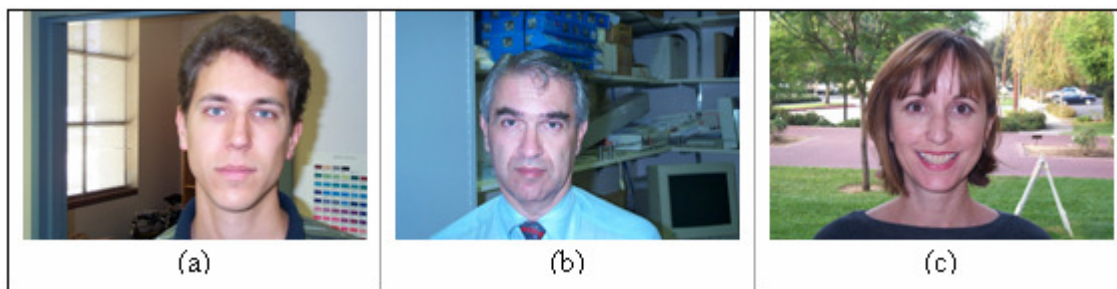


Figura 82 – Exemplos de imagens do Banco de Imagens A com diferentes condições de iluminação e complexidade de fundo: (a) e (b) sob condições de iluminação em ambiente fechado e (c) sob condições de iluminação de ambiente aberto

Neste banco de imagens, há também a ocorrência de oclusões como barba, óculos ou gestos que ocultam parcialmente alguma característica da face, Figura 83.

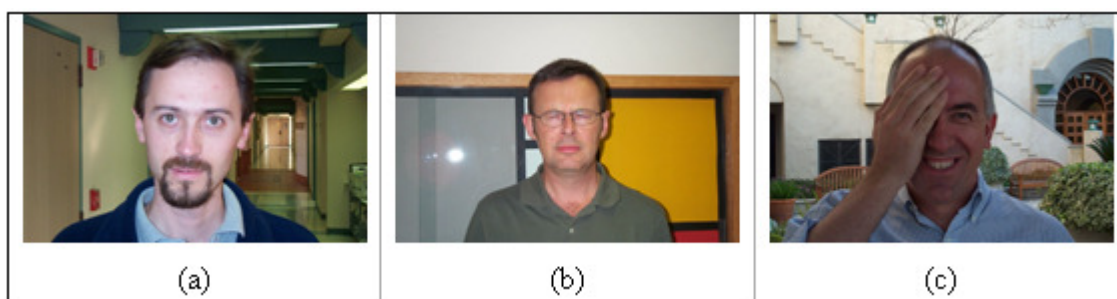


Figura 83 - Exemplos de imagens do Banco de Imagens A que apresentam oclusões: (a) Barba; (b) Olhos fechados e (c) oclusão do olho direito

O banco de imagens A é composto por grupos de imagens de 28 pessoas, somando um total de 447 imagens. Três imagens deste banco não foram utilizadas por se tratarem de desenho de pessoas em escala de cinza (Figura 84), não podendo assim serem utilizadas no método de classificação de pixels de pele.



Figura 84 – Desenho de pessoas do Banco de Imagens A

As imagens deste banco são provenientes de uma câmera Eastman Kodak Company, KODAK DC280. O formato é JPEG e a resolução espacial é de 896 x 592 pixels com 24 bits de profundidade.

No treinamento do algoritmo de classificação dos pixels de pele desenvolvido foram utilizadas 160 imagens deste banco para construção de máscaras, por meio de um processo manual, onde os pixels que não pertenciam à pele foram levados ao nível zero.

1.2 Banco de imagem B

Este banco de imagens está disponível para no endereço eletrônico: <http://www.anefian.com/face_reco.htm>. É composto de 50 diretórios, cada um contendo 15 imagens faciais de uma única pessoa. Os indivíduos apresentados nas imagens variam de caucasianos a negros (Figura 85 (b) e (c)).

A média do tamanho das faces é de 150 x 150 pixels. As imagens são frontais, apresentando também inclinações em alguns casos (Figura 85 (a)), com diferentes condições de iluminação.

O fundo das imagens não é controlado, mas apresenta sempre a mesma característica, sendo esse fundo classificado como complexo. Sua complexidade se define pela presença de elemento de madeira, uma estante, dificultando a separação entre a pele e o fundo (Figura 85 (c)).



Figura 85 – Exemplos de imagens do Banco de Imagens B: (a) Face inclinada; (b) Indivíduo caucasiano e (c) indivíduo com tom de pele semelhante ao objeto de fundo da cena

As imagens são provenientes de uma câmera modelo HANDCAM. O formato é JPEG e a resolução espacial é de 640 x 480 pixels com 24 bits de profundidade.

Para o treinamento de detecção de pixels de pele, foram utilizadas 150 imagens deste banco correspondendo a três imagens de cada diretório.

1.3 Imagens de webcam

Não foram encontrados bancos de imagens construídos a partir de câmeras tipo *webcam* disponíveis na internet para utilização em experimentos. Quando necessário, utilizou-se imagens como as da Figura 86.

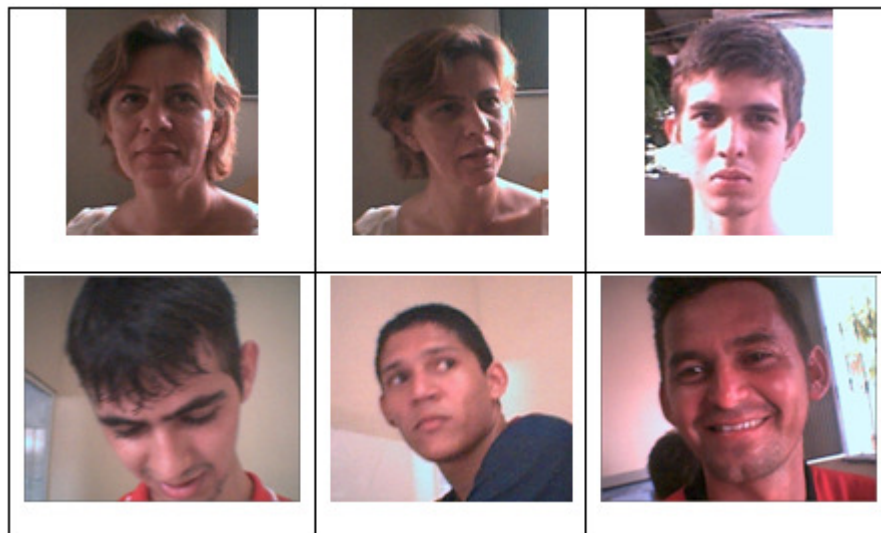
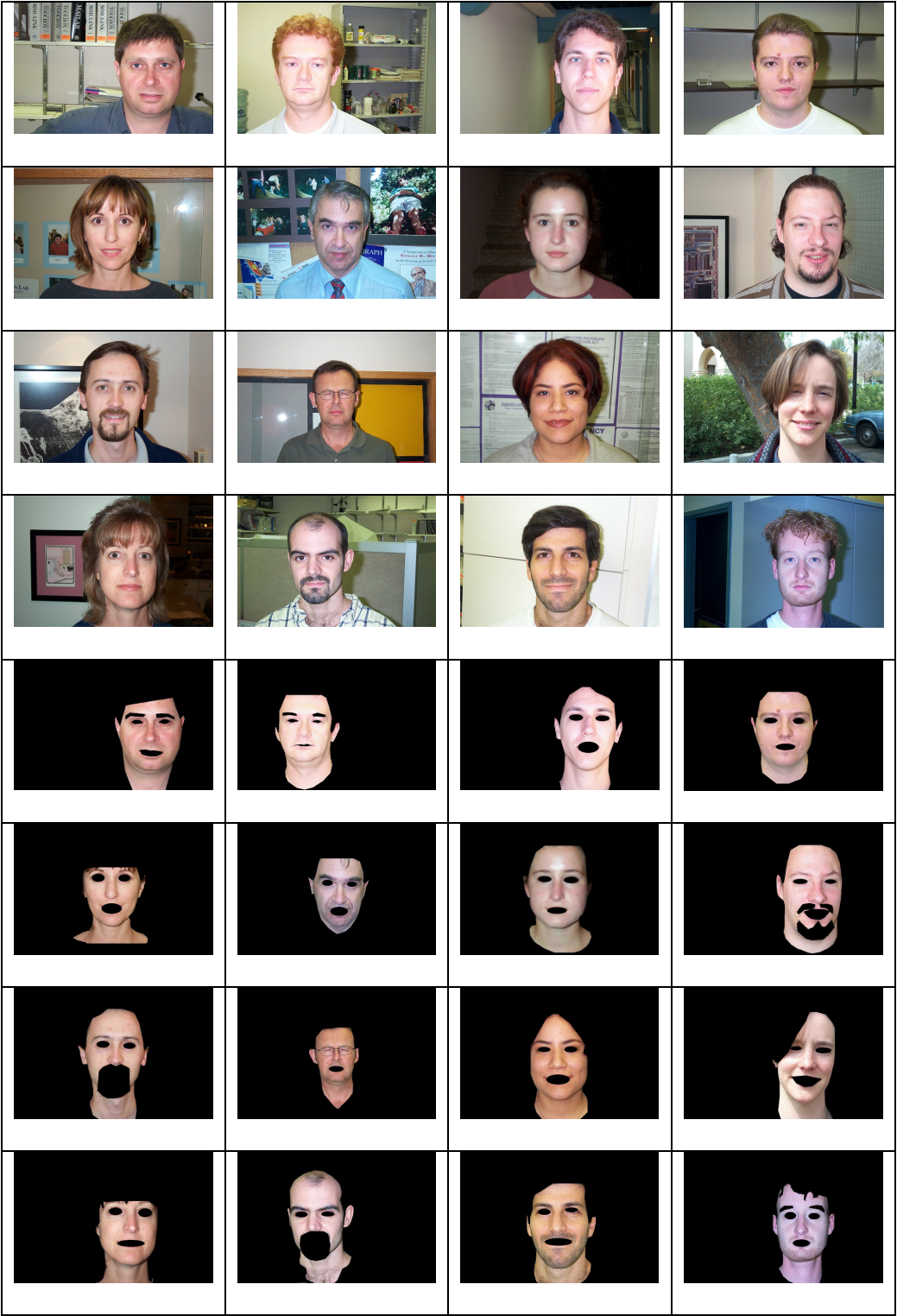


Figura 86 – Exemplos de imagens de *webcam* utilizadas nos experimentos.

O total de imagens utilizadas foi 16, provenientes de uma *webcam* modelo Genius. O formato é JPEG e a resolução espacial é de 352 x 288 pixels com 24 bits de profundidade.

Não houve a necessidade de construção de um banco de imagens de *webcam* para esse trabalho uma vez que foi explorado outros trabalhos voltado para esse fim e essa dissertação é voltada para análise de imagens com qualidade superior, como imagens de mais de 2 megapixels. As imagens de *webcam* foram utilizadas somente para efeito de verificação de métodos relacionados a esse tipo de imagem.

Exemplos de imagens do Banco de imagens A e Máscaras

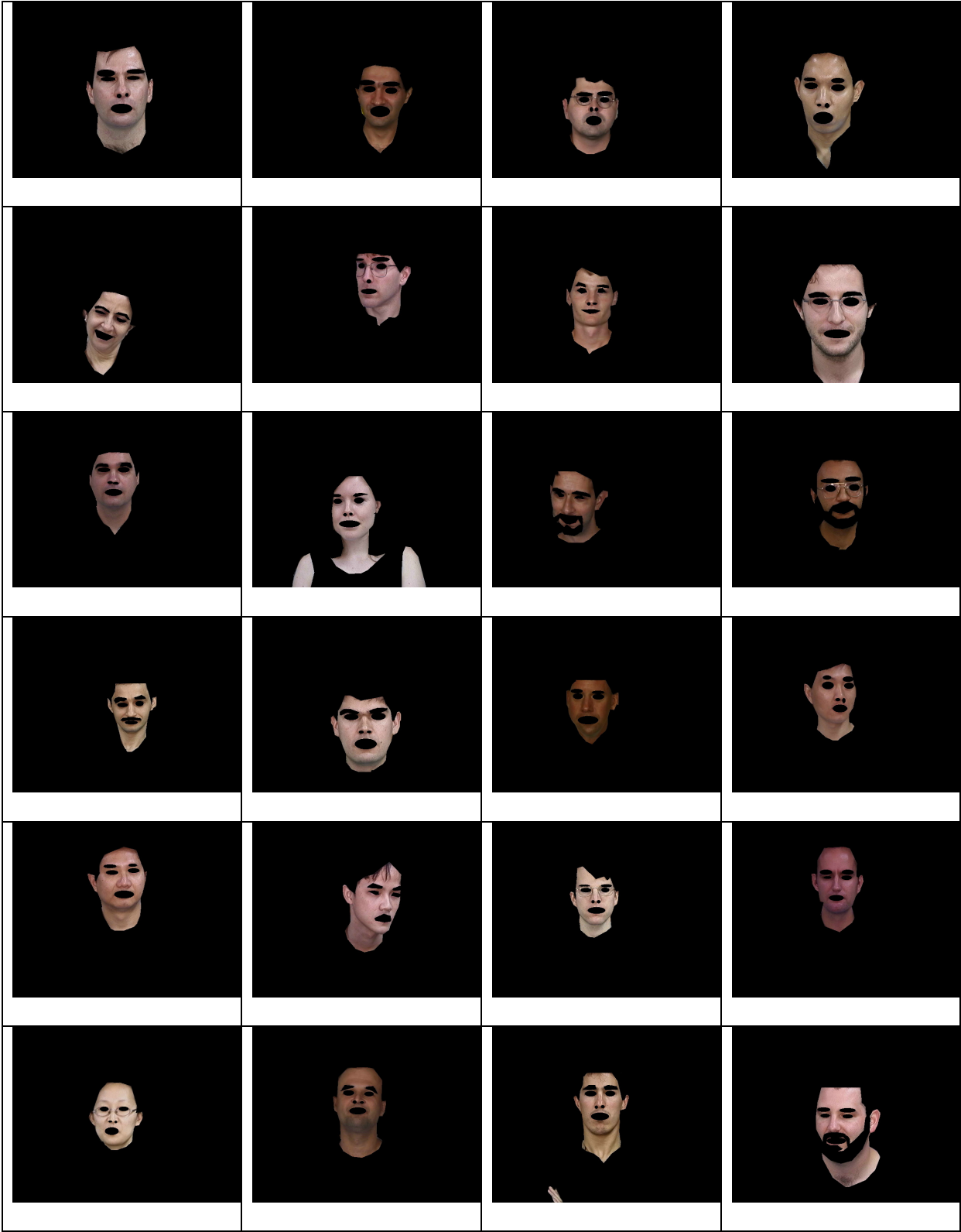


Exemplos de imagens do Banco de imagens B e Máscaras





Máscaras do Banco de Imagens B





ANEXO B

MATERIAIS E MÉTODOS

No desenvolvimento deste trabalho foi fundamental o uso de duas ferramentas, uma para geração das máscaras das imagens e outra para a codificação do método de detecção facial.

A geração das máscaras correspondentes a cada imagem do banco de imagens utilizou-se o software PhotoImpact (software proprietário) e sua ferramenta *Lasso Tool* para selecionar manualmente e excluir todas as regiões da imagem que não fossem pele. O mesmo resultado pode ser obtido utilizando outra ferramenta de edição de imagens como o Gimp (ferramenta de código livre).

A codificação foi desenvolvida em Matlab, pois de acordo com o site da MathWorks: “as ferramentas para processamento de imagens provêm um conjunto de algoritmos de referência padronizados para pré- e pós-processamento, além de disponibilizar ferramentas gráficas, análise, visualização e desenvolvimento de algoritmos. Pode-se restaurar ou degradar imagens, realçá-las para promover inteligibilidade, extrair características, analisar formas e texturas.”

As principais funções do Matlab utilizadas neste trabalho foram:

- `bwlabeln` – rotula os objetos na imagem;
- `graythresh` – calcula o limiar global da imagem;
- `hist` – plota o histograma da imagem;
- `im2bw` – converte uma imagem em escala de cinza para binária;
- `imclose` – efetua a operação morfológica de fechamento na imagem;

- `imfilter` – executa uma filtragem multidimensional na imagem;
- `regionprops` – obtém as medidas de um conjunto de propriedades das regiões rotuladas como bordas, centróides, área, etc.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)