

OSVALDO SEVERINO JUNIOR

Mistura de cores: Uma nova abordagem para processamento de cores e sua aplicação na segmentação de imagens

Tese apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de doutor em Engenharia Elétrica.

Área de Concentração: Processamento de Sinais e Instrumentação.

Orientador: Prof. Dr. Adilson Gonzaga

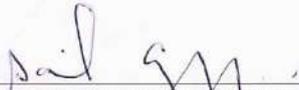
São Carlos

2009

FOLHA DE JULGAMENTO

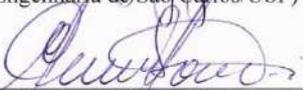
Candidato: Bacharel OSVALDO SEVERINO JUNIOR

Tese defendida e julgada em 28/05/2009 perante a Comissão Julgadora:



Prof. Associado **ADILSON GONZAGA (Orientador)**
(Escola de Engenharia de São Carlos/USP)

APROVADO



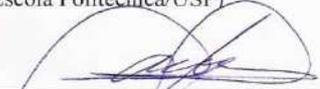
Prof. Dr. **EVANDRO LUÍS LINHARI RODRIGUES**
(Escola de Engenharia de São Carlos/USP)

APROVADO



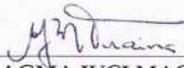
Prof^a. Associada **ROSELI DE DEUS LOPES**
(Escola Politécnica/USP)

APROVADO



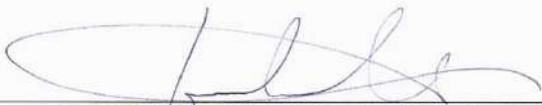
Prof. Dr. **JANDER MOREIRA**
(Universidade Federal de São Carlos/UFSCar)

APROVADO



Prof^a. Titular **AGMA JUCI MACHADO TRAINA**
(Instituto de Ciências Matemáticas e de Computação/USP)

APROVADO



Prof. Associado **GERALDO ROBERTO MARTINS DA COSTA**
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica e
Presidente da Comissão de Pós-Graduação

A Ciomara, minha esposa, Matheus e Gabriel, meus filhos, com amor, admiração e gratidão pela compreensão, carinho, presença e incansável apoio ao longo do período de elaboração deste trabalho.

Agradecimentos

A Deus, pelos dons da vida, da sabedoria e da perseverança.

Ao Prof. Dr. Adilson Gonzaga, pela amizade e que nos anos de convivência, muito me ensinou, contribuindo para meu crescimento científico e intelectual.

À Escola de Engenharia de São Carlos, pela oportunidade de realização do curso de pós-graduação à nível de doutorado.

Aos profs. e funcionários do Departamento de Engenharia Elétrica pelo ensino e trabalho prestados durante estes anos na USP.

Ao IMES/FAFICA de Catanduva, pelo apoio para a concretização deste trabalho.

*“Bastam algumas cores para despertar e
aguçar nossa imaginação”*

Resumo

SEVERINO JR, O. **Mistura de cores: Uma nova abordagem para processamento de cores e sua aplicação na segmentação de imagens.** 2009. 121 f. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2009.

Inspirado nas técnicas utilizadas por pintores que sobrepõem camadas de tintas de diversos matizes na geração de uma tela artística e também observando-se a distribuição da quantidade dos cones na retina do olho humano na interpretação destas cores, este trabalho propõe uma técnica de processamento de imagens baseada na mistura de cores. Trata-se de um método de quantização de cores estático que expressa a proporção das cores preto, azul, verde, ciano, vermelho, magenta, amarelo e branco obtida pela representação binária da cor que compõe os pixels de uma imagem RGB com 8 *bits* por canal. O histograma da mistura é denominado de Misturograma e gera planos que interceptam o espaço RGB, definindo o espaço de cor HSM (*Hue, Saturation and Mixture*). A posição destes planos dentro do cubo RGB é modelada por meio da distribuição dos cones sensíveis aos comprimentos de onda curta (Short), média (Middle) e longa (Long) consideradas para a retina humana. Para demonstrar a aplicabilidade do espaço de cor HSM, é proposta, neste trabalho, a segmentação dos *pixels* de uma imagem digital em “pele humana” ou “não pele” com o uso dessa nova abordagem. Para análise de desempenho da mistura de cores foi implementado um método tradicional no espaço de cor RGB e também usando uma distribuição gaussiana nos espaços de cores HSV e HSM. Os resultados obtidos demonstram o potencial da técnica que emprega a mistura de cores para a segmentação de imagens digitais coloridas. Verificou-se também que, baseando-se apenas na camada mais significativa da mistura de cores, gera-se a imagem esboço de uma imagem facial denominada Esboço da Face. Os resultados obtidos comprovam o bom desempenho do Esboço da Face em aplicações CBIR.

Palavras chaves: Misturograma, espaço de cor HSM, quantização de cor, segmentação de pele humana, processamento de imagens coloridas.

Abstract

SEVERINO JR, O. **Colors Mixture: A new approach for color processing and its application in image segmentation.** 2009. 121 p. Dr. Thesis – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2009.

Inspired on the techniques used by painters to overlap layers of various hues of paint to create oil paintings, and also on observations of the distribution of cones in human retina for the interpretation of these colors, this thesis proposes an image processing technique based on color mixing. This is a static color quantization method that expresses the mixture of black, blue, green, cyan, red, magenta, yellow and white colors quantified by the binary weight of the color that makes up the pixels of an RGB image with 8 bits per channel. The mixture histogram, called a Mixturegram, generates planes that intersect the RGB color space, defining the HSM (Hue, Saturation and Mixture) color space. The position of these planes inside the RGB cube is modeled by the distribution of cones sensitive to the short (S), middle (M) and long (L) wave lengths of the human retina. To demonstrate the applicability of the HSM color space, this thesis proposes the segmentation of the pixels of a digital image of “human skin” or “non-skin” using this new approach. The performance of the color mixture is analyzed by implementing a traditional method in the RGB color space and by a Gaussian distribution in the HSV and HSM color spaces. The results demonstrate the potential of the proposed technique for color image segmentation. It was also noted that, based only on the most significant layer of the colors mixture, it is possible to generate the face sketch image. The results show the performance of the face sketch image in CBIR applications.

Keywords: Mixturegram, HSM color space, color quantization, human skin segmentation, color images processing.

Lista de Figuras

Figura 2.1: Espectro Eletromagnético.	30
Figura 2.2: Esquema da percepção de cor pelo olho humano.	30
Figura 2.3: Sensibilidade dos cones da retina em função do comprimento de onda.	31
Figura 2.4: Exemplo do processo de formação de cores aditivo.	32
Figura 2.5: Exemplo do processo de formação de cores subtrativo.	33
Figura 2.6: Espaço de cor do modelo RGB.	35
Figura 2.7: Espaço de cor do modelo CIE.	37
Figura 2.8: Espaço de cor do modelo CMY.	38
Figura 2.9: Espaço de cor do modelo HLS.	40
Figura 2.10: Espaços de cor dos modelos YIQ e YCbCr.	42
Figura 2.11: Tipos de erros em sistemas de classificação de pele humana.	51
Figura 2.12: FMR e $FNMR$ calculados para um $threshold$ t sobre os valores das distribuições de <i>genuínos</i> e <i>impostores</i>	53
Figura 2.13: Variação do $threshold$ t em um sistema de reconhecimento biométrico. a) Distribuição de impostores e usuários genuínos. b) $FMR(t)$ e $FNMR(t)$ calculados pelos valores em a).	54

Figura 2.14: Exemplos das curvas $FMR(t)$ e $FNMR(t)$ e os pontos de erro: ERR , $Zero FNMR$ e $Zero FMR$	55
Figura 2.15: Curva ROC segundo a abordagem utilizada pela Biometria.....	55
Figura 2.16: Representação de uma matriz de confusão.....	56
Figura 3.1: Evolução dos passos realizados para a pintura de uma tela artística. (RIBEIRO, 2005).....	60
Figura 3.2: Decomposição do quadro do druida nas camadas K_j	63
Figura 3.3: Representação de uma cor $C = (109,146,137)$ composta pela mistura das cores: verde, ciano, vermelho e magenta.	68
Figura 3.4: Aplicação da mistura de cores. a) imagem Original. b) imagem quantizada pelo valor da mistura de cores.	70
Figura 3.5: Misturograma (Histograma normalizado) das cores quantizadas pelo valor da mistura da Figura	71
Figura 3.6: Intersecção do plano m com o vértice azul do cubo RGB.....	74
Figura 3.7: Projeção de MR_m no plano BG.	76
Figura 3.8: Intersecção do plano m com o vértice verde do cubo RGB.	76
Figura 3.9: Intersecção do plano m com o vértice ciano do cubo RGB.....	78
Figura 3.10: Intersecção do plano m com o vértice vermelho do cubo RGB.	79
Figura 3.11: Intersecção do plano m com o vértice magenta do cubo RGB.....	80
Figura 3.12: Intersecção do plano m com o vértice amarelo do cubo RGB.	82
Figura 3.13: Planos de intersecção do espaço HSM nos vértices azul, verde, ciano, vermelho, magenta e amarelo do cubo RGB.	84
Figura 3.14: Definição do plano da mistura de cores que intercepta o cubo RGB.....	84
Figura 4.1: Definição da região de pele no espaço RGB. (BRAND; MASON, 2000).....	91
Figura 4.2: Amostras de pele segmentadas manualmente dos bancos AR e FERET.	92

Figura 4.3: Representação da distribuição dos pontos de pele nos espaços HSM e HSV. a) $m = v = 0,39$. b) $m = v = 0,49$. c) $m = v = 0,59$. d) $m = v = 0,69$. e) $m = v = 0,79$	93
Figura 4.4: Representação da distribuição dos pontos de pele obtidos pela distância de Mahalanobis (cor branca) e por amostragem (cor preta) nos espaços HSM e HSV. a) $m = v = 0,39$. b) $m = v = 0,49$. c) $m = v = 0,59$. d) $m = v = 0,69$. e) $m = v = 0,79$	96
Figura 4.5: Imagem obtida pela mistura de cores utilizando-se apenas a camada K_7 . a) imagem original.	97
Figura 4.6: Exemplo da imagem esboço da face. a) Imagem original. b) Imagem do esboço da face.	98
Figura 4.7: Exemplos de imagens esboço da faces segmentadas a partir do banco AR pela mistura de cores.....	99
Figura 4.8: Segmentação da região de pele humana. a) Imagem original. b) Imagem segmentada pelo classificador não paramétrico de Peer no RGB. c) Imagem segmentada pelo classificador paramétrico gaussiano no HSV. d) Imagem segmentada pelo classificador paramétrico gaussiano no HSM. e) Imagem segmentada manualmente para criar o “ground truth”.	102
Figura 4.9: As imagens m-001-1, m-001-2, m-001-3, m-001-4, m-001-5, m-001-6 e m-001-7 referentes a classe masculina (m) 001 do banco AR.....	104
Figura 4.10: Etapas para a construção da matriz de projeção do subespaço PCA.	105
Figura 4.11: Obtenção do vetor característica da imagem m-001-1 do banco de imagens originais.....	106
Figura 4.12: Gráfico de desempenho do banco com 840 imagens apresentando diferentes expressões faciais e mudanças de iluminação.....	107

Lista de Tabelas

Tabela 1: Métodos de Quantização de cores	50
Tabela 2: Representação binária da cor $C = (109, 146, 137)$	67
Tabela 3: Representação dos cálculos K_i , NK_i e N_i da cor $C = (109, 146, 137)$	67
Tabela 4: Resultados obtidos pelos classificadores paramétricos e não paramétricos	100
Tabela 5: Índices de desempenho dos classificadores paramétricos e não paramétricos.	101
Tabela 6: Interpretação do índice de concordância Kappa (BISHOP; GREEN, 1993)	103
Tabela 7: Dados obtidos nos sistemas CBIR imagem original e imagem esboço da face.	107

Lista de Abreviaturas

- AR:** The AR Face Database – banco de imagens faciais denominado AR.
- CBIR:** *Content-Based Image Retrieval* – Recuperação de imagens por Conteúdo.
- CIE:** *Commission Internationale de l'Éclairage* – modelo de cor criado pela Comissão Internacional de l'Éclairage.
- CMY:** *Cyan Magenta Yellow* - modelo de cor baseado nas cores complementares: ciano, magenta e amarelo.
- FERET:** The *FERET Face Database* – banco de imagens faciais denominado FERET.
- FMNR:** *False Non-Match Rate* – Probabilidade de falsa rejeição.
- FMR:** *False Match Rate* – Probabilidade de falsa aceitação.
- HLS:** *Hue Lightness Saturation* : Matiz Luminosidade Saturação - modelo de cor que emprega os conceitos qualitativos de matiz, luminosidade e tonalidade.
- HSM:** *Hue Saturation Mixture* – Matiz Saturação e Mistura – modelo de cor que emprega os conceitos qualitativos de matiz, luz e o valor da mistura.
- HSV:** *Hue Saturation Value* : Matiz Saturação Valor - modelo de cor que emprega os conceitos qualitativos de matiz, luz e tonalidade.
- PCA:** *Principal Components Analysis* – Análise dos Principais Componentes.

RGB: *Red Green Blue* - modelo de cor concebido com base nos dispositivos gráficos com três cores primárias: vermelho, verde e azul

ROC: *Receiver Operating Characteristic*

RP: *Recall Precision* – curva de Refurcação x Precisão.

YCbCr: Luminância Cromaticidade Azul Cromaticidade Vermelha – modelo de cor baseado na separação dos sinais de cor RGB em um sinal de luminosidade, ou luminância, e dois sinais de cromaticidade ou diferença de cor correspondentes as componentes azul e vermelha.

Sumário

1	Introdução	25
2	Processamento de Imagens Coloridas.....	29
2.1	Introdução	29
2.2	Modelos de cor.....	34
2.2.2	Modelo RGB	34
2.2.2	Modelo CIE	36
2.2.3	Modelo CMY.....	38
2.2.4	Modelo HSL	39
2.2.6	Modelo YIQ e YCbCr	42
2.3	Abordagens tradicionais de quantização de cores	43
2.4	Métricas de Avaliação de Desempenho.....	49
2.4.1	Curva <i>ROC</i> (<i>Receiver Operating Characteristic</i>)	50
2.4.2	Matriz de Confusão	56
2.5	Considerações Finais	58
3	Mistura de Cores	59
3.1	Introdução	59
3.2	Hipótese da Mistura de Cores.....	61
3.2.1	Camadas Binárias das Cores	61

3.2.2	Mistura de cores	64
3.3	Quantização de cores	69
3.4	Espaço de cor HSM definido pelo Valor da mistura.....	72
3.4.1	Conversão do RGB para HSM.....	73
3.4.1	Conversão do HSM para RGB	84
3.5	Considerações Finais	86
4	Uso da Mistura de Cores na Segmentação de Imagens de Pele Humana e o Esboço da Face.....	89
4.1	Introdução.....	89
4.3.1	Amostragem da Região de Pele no RGB, no HSM e no HSV.....	91
4.3.2	Equacionamento da distribuição dos pontos de Pele da Face Humana amostrados no RGB, no HSM e HSV	94
4.3.3	Detecção do Esboço da Face da Face Humana pela Mistura de cores	97
4.4	Resultados	98
4.4.1	Desempenho dos Classificadores da Região de Pele Humana	100
4.4.2	Desempenho das Imagens Esboço da Face.....	103
4.5	Considerações Finais	108
5	Conclusões.....	109
5.1	Contribuições.....	110
5.2	Trabalhos Futuros.....	111
	Referências Bibliográficas	113

Capítulo 1

Introdução

A cor transmite informações sobre as condições de iluminação e é uma das principais características dos objetos visualizados, o que permite acentuar os contornos das formas e dos sombreamentos.

Na retina existem células sensíveis à luz, os bastonetes e os cones, que transformam a energia que absorvem, na forma de *fótons*, em impulsos nervosos. Os bastonetes são sensíveis à intensidade luminosa em toda a gama de comprimentos de ondas ao qual o olho humano responde. Ao contrário dos bastonetes, os cones são mais sensíveis à luz em certos intervalos de comprimentos de onda.

Há cones sensíveis aos comprimentos de onda curta (Short), média (Middle) e longa (Long) na retina humana, tal que para dois cones sensíveis ao comprimento de onda longa existe um cone sensível ao comprimento de onda médio. Assim, existem cones sensíveis à luz na região do vermelho (faixa de comprimento de onda longo), do verde (faixa de comprimento de onda médio) e do azul (faixa de comprimento de onda curto). Observa-se que existe uma distribuição na proporção (2:1) dos cones sensíveis à luz na região do vermelho

em relação aos cones sensíveis à luz na região do verde, tal que, para dois cones sensíveis à luz na região do vermelho existe um cone sensível à luz na região do verde.

A diferença entre as respostas dos três tipos de cones permite interpretar diferentes comprimentos de onda que corresponde a diferentes cores. Esta interpretação é conhecida como *Teoria dos Três Estímulos*.

Tomando por base a Teoria dos Três Estímulos supõe-se que qualquer modelo de cor deverá possuir três parâmetros. O problema essencial da modelagem da cor está na definição de que grandezas devem ser associadas a cada um dos três parâmetros.

A pintura de uma tela artística, no entanto, não é realizada em uma única etapa. Durante a evolução de uma pintura artística, uma nova camada de pigmentos de cores é misturada às camadas anteriores e, como consequência, novos tons vão surgindo. Esse processo se repete até que as tonalidades desejadas sejam alcançadas.

1.1 Objetivos

Observando o processo efetuado por artistas na pintura de uma tela e, na distribuição dos cones sensíveis à luz na região do vermelho e do verde e, na teoria dos três estímulos que define um modelo de cor com três parâmetros, propõe-se neste trabalho: i) o desenvolvimento da mistura de cores que define o espaço de cor HSM, ii) a proposta de um método para segmentar a região de pele humana através do valor da mistura de cores e iii) a obtenção da imagem esboço da face para a descrição global de uma imagem facial.

1.2 Materiais e Métodos

A proposta precípua desta tese de doutoramento é a apresentação de um novo espaço de cor definido a partir do espaço de cor RGB, aqui denominado HSM. Para demonstrar sua aplicabilidade é implementada a segmentação de *pixels* da cor da pele humana. Esta segmentação é necessária para algoritmos de detecção e de reconhecimento de faces humanas em Biometria. Um conjunto de treinamento é obtido a partir de amostras de cor de pele humana retirada de bases de dados amplamente utilizadas na literatura. Por meio da distribuição da cor dos *pixels* dessas amostras no espaço de cor HSM, propõe-se o uso de um método paramétrico com o uso de uma função gaussiana para a classificação da cor dos *pixels* de imagens faciais como “pele” ou “não pele”. Verificando a cor *dos pixels* presentes na camada sete da mistura de cores é proposta a imagem de esboço da face. A imagem esboço da face é aplicada num sistema CBIR demonstrando o seu potencial quanto à recuperação de imagens baseadas por conteúdo.

1.3 Organização do Trabalho

Esta tese de doutoramento está dividida nos seguintes capítulos:

O capítulo 2 apresenta o conceito de cor e as abordagens tradicionais de processamento de imagens coloridas baseadas nos modelos de cor.

No capítulo 3 é proposto o valor da mistura baseado no conceito de mistura de cor empregado nas artes visuais e a distribuição dos cones na retina humana.

O capítulo 4 aborda a proposta do trabalho de doutoramento e são avaliados os métodos de segmentação da região de pele: não paramétrico, implementado no espaço de cor RGB e,

paramétrico gaussiano implementados nos espaços de cores HSV e HSM. Neste capítulo também é abordada a obtenção da imagem Esboço da Face.

No capítulo 5 são apresentados as conclusões e trabalhos futuros.

Capítulo 2

Processamento de Imagens Coloridas

2.1 Introdução

Segundo Sharma (2003), a cor é essencialmente um conceito subjetivo do ser humano, pois consiste na interpretação que o sistema sensorial e o cérebro dão aos diferentes comprimentos de onda da luz recebida.

O estudo da luz e da cor deve ser iniciado pela Física elementar, uma vez que a luz é uma onda eletromagnética. A Física define que todas as ondas eletromagnéticas se propagam no vácuo com a mesma velocidade c com o valor de 3×10^8 m/s (velocidade da luz) e, sabendo-se a frequência de uma onda eletromagnética (f), pode-se determinar o comprimento de onda (λ) desta radiação, usando a equação $\lambda = c/f$.

Desta forma, podem-se exemplificar as ondas eletromagnéticas de maior importância nas pesquisas e nas aplicações práticas, em função do comprimento de onda: Raios-X (faixa de 5 pm até 1 nm), raios ultravioletas (faixa de 1 até 400 nm), o espectro de luz visível (faixa de 400 até 700 nm), raios infravermelhos (faixa de 700 nm até 1 mm) e faixas de radiofrequência que variam de 20 cm até 10^5 m, Figura 2.1, (GOMES; VELHO, 1998).

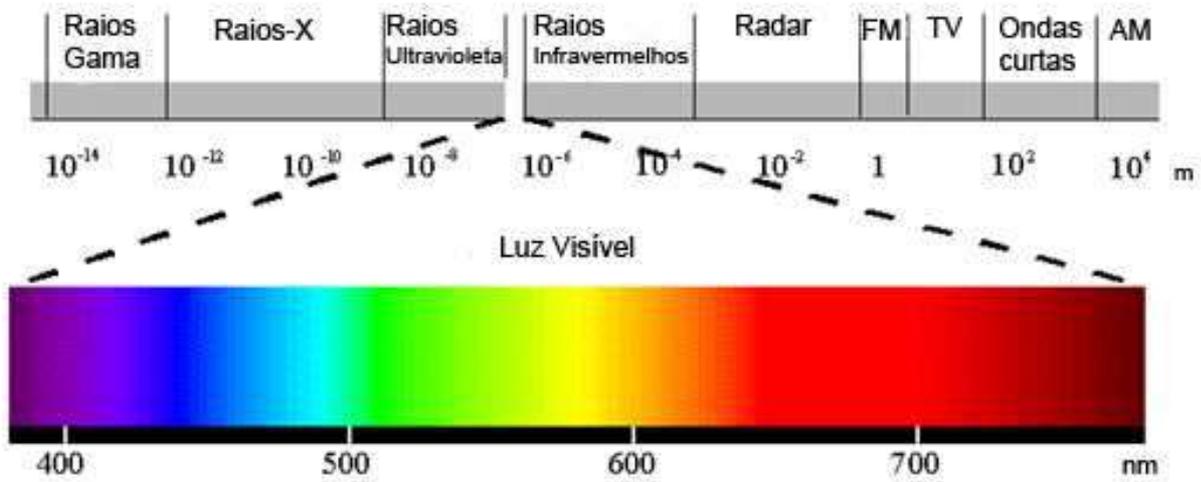


Figura 2.1: Espectro Eletromagnético.

O espectro de luz visível pode assumir diversas cores (desde o violeta até o vermelho), em função do comprimento de onda.

A percepção de cor pelo olho humano pode ser esquematizada pela Figura 2.2. De forma resumida, os raios luminosos incidem na córnea, sendo então refratados. Os raios refratados incidem sobre o cristalino que tem por objetivo projetá-los na retina. Na retina encontram-se dois tipos de fotorreceptores, os cones e os bastonetes, que convertem a intensidade e a cor da luz recebida em impulsos nervosos. Estes impulsos são enviados ao cérebro através do nervo óptico e então se tem a percepção de uma imagem.

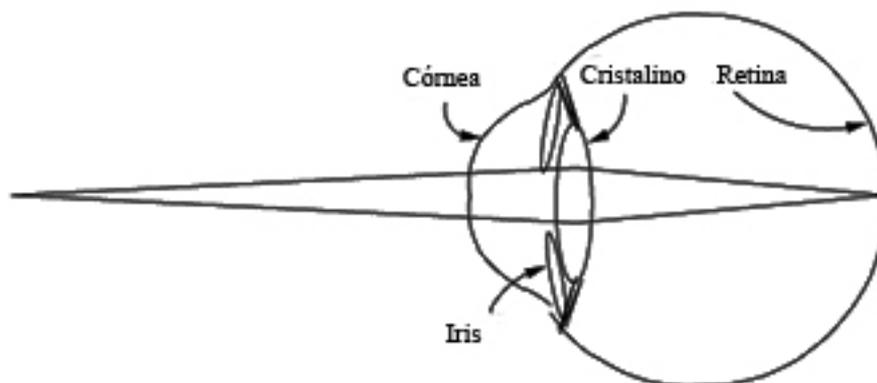


Figura 2.2: Esquema da percepção de cor pelo olho humano.

Os bastonetes, embora sejam maioria absoluta, só conseguem captar a intensidade (luminosidade) da cor, ou seja, respondem igualmente a qualquer espectro e desta forma não diferenciam cores. Os bastonetes são responsáveis pela detecção das formas dos objetos, pela visão noturna e pela informação necessária à orientação.

Diferentemente dos bastonetes, os cones respondem a espectro de cores distintos (vermelho, verde e azul). Assim, existem cones sensíveis à luz na região do vermelho (ρ), do verde (γ) e do azul (β). Portanto, diz-se que o sistema visual humano distingue as cores pelo processo da **tricromacia**. Na Figura 2.3 é apresentado os comprimentos de onda dos cones que respondem na faixa do azul, do vermelho e do verde, nota-se que a eficiência do cone que responde na faixa do azul possui uma sensibilidade bem menor do que os outros dois tipos de cones (GOMES; VELHO, 1998).

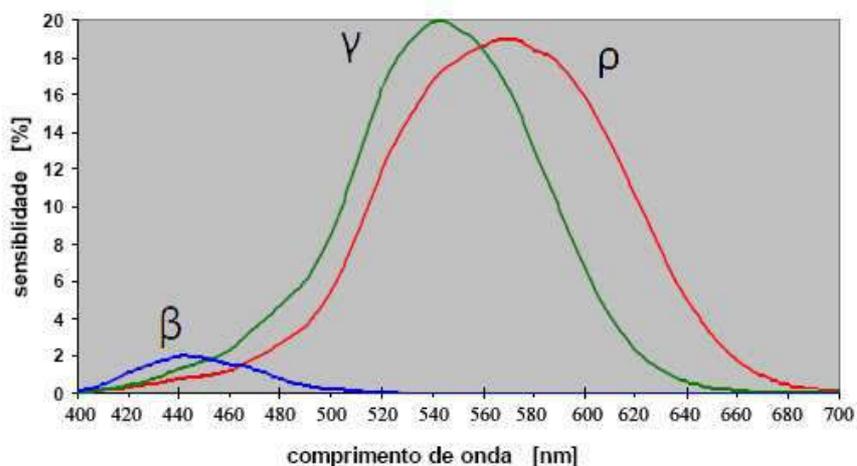


Figura 2.3: Sensibilidade dos cones da retina em função do comprimento de onda.

Em Morris (2000), verifica-se que existe uma proporção (2:1) para distribuição dos cones sensíveis ao comprimento de onda longo (Long) em relação aos cones sensíveis ao comprimento de onda médio (Meddle) na retina, ou seja, para cada dois cones sensíveis a faixa do vermelho existe um cone sensível a faixa do verde.

Visualizando o espectro eletromagnético (Figura 2.1) a literatura define o termo cores espectrais como sendo as cores que correspondem a comprimentos de onda bem determinados do espectro da luz visível. Mas a observação do espectro da luz visível mostra que nele não se encontram todas as cores visíveis, porque existe uma gama de cores que habitualmente definimos entre o azul e o vermelho (por ex. o magenta) que estão ausentes do espectro da luz visível, pois equivaleria a ligar comprimentos de onda menores (azuis) aos maiores (vermelhos) sem passar pelos comprimentos de onda intermediários (verdes e amarelos). Também a cor branca não é uma cor do espectro da luz visível porque não corresponde a nenhum comprimento de onda preciso, pois resulta da detecção simultânea da radiação visível em muitos comprimentos de onda, com uma intensidade mais ou menos uniforme. Portanto, esses exemplos mostram que existem mais cores visíveis do que cores espectrais.

A existência de uma quantidade maior de cores visíveis do que espectrais é explicada pelo processo de formação de cores. Existem dois processos de formação de cores: o aditivo e o subtrativo.

No processo aditivo, duas fontes luminosas de cores diferentes são projetadas em duas regiões, como ilustrado na Figura 2.4. Na área de interseção destas regiões há a formação de uma nova cor constituída pela adição dessas duas cores. O processo aditivo é usado largamente nas televisões comerciais, monitores de vídeo e câmeras digitais.

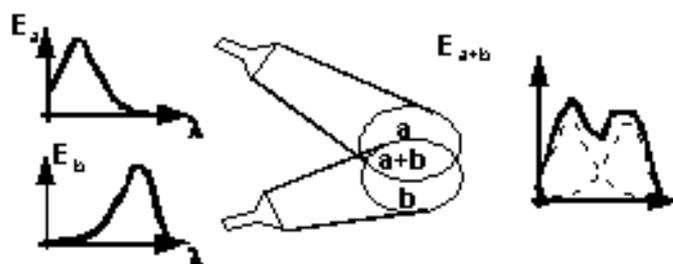


Figura 2.4: Exemplo do processo de formação de cores aditivo.

O processo de formação de cores subtrativo baseia-se no uso de filtros ou corantes que têm por objetivo filtrar determinados comprimentos de onda. A Figura 2.5 exemplifica o processo. Ao emitir uma luz branca (que possui todos os comprimentos de onda) sobre um filtro verde que filtra todos os comprimentos de onda deixando só passar o comprimento de onda relativo a cor verde, produzindo assim o verde. Na utilização de corantes o processo é o mesmo, só que são usados pigmentos que absorvem e refletem alguns comprimentos de onda. Esse processo é utilizado em eslaides (SHARMA, 2003).

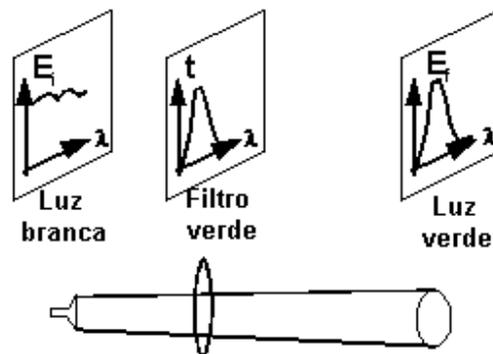


Figura 2.5: Exemplo do processo de formação de cores subtrativo.

Quanto ao processo de manipulação de cor em sistemas computacionais, a literatura emprega o termo *true color* para a representação de uma cor por três ou mais *bytes*. Como o sistema humano utiliza o processo da tricromacia para distinguir cores, utilizando um *byte* (8 *bits*) para cada componenete do modelo de cor, representam-se 3×2^8 intensidades de cor, o que resulta em 16.777.216 cores. Acredita-se que o olho humano consegue distinguir algo em torno de 10 milhões de cores. Portanto, o termo *true color* representa mais cores que o olho humano consegue distinguir e, em consequencia, há a ilusão de cores reais (SHARMA, 2003).

2.2 Modelos de cor

A modelagem da cor assume extrema importância em sistemas de Visão Computacional, pois, embora a captura e apresentação de imagens sejam feitas por dispositivos gráficos que utilizam o modelo de cor RGB, o processamento de imagens poderá depender de características e formatos específicos que consigam manipular a informação de uma forma coerente e precisa e que introduzam o menor número de erros ou falhas quanto à precisão. Como visto, o sistema visual humano distingue as cores pelo processo da tricromacia e, portanto, a especificidade da modelagem da cor está na definição de que grandezas devem ser associadas a cada um dos três parâmetros.

2.2.2 Modelo RGB

O modelo RGB (*Red*, *Green* e *Blue*) é um modelo de cor desenvolvido com base nos dispositivos gráficos que apresentam as três cores primárias: vermelho, verde e azul. Este modelo descreve as cores como o resultado da adição das três cores primárias, com uma intensidade que pode variar entre 0 e 1. O valor 1 corresponde à intensidade máxima com que a cor pode ser apresentada no dispositivo gráfico e o valor 0 à intensidade mínima. A cor branca é resultante da adição simultânea das três cores primárias com intensidade máxima. A cor preta é obtida quando todas as cores primárias apresentam intensidade mínima

Como o modelo RGB está associado às superfícies emissoras de luz, este modelo é universalmente empregado nos equipamentos que manipulam a emissão de luz, tais como os monitores de vídeo, os televisores em cores, as câmeras fotográficas e as filmadoras.

As cores deste modelo definem um subespaço de cor em que $0 \leq (R,G,B) \leq 1$ representado por um cubo de aresta unitária. Na Figura 2.6 é apresentado o espaço de cor RGB, onde:

- Vermelho = (1,0,0)
- Verde = (0,1,0)
- Azul = (0,0,1)
- Preto = (0,0,0)
- Branco = Vermelho (1,0,0) + Verde (0,1,0) + Azul (0,0,1) = (1,1,1)
- Amarelo = Vermelho (1,0,0) + Verde (0,1,0) = (1,1,0)
- Ciano = Verde (0,1,0) + Azul (0,0,1) = (0,1,1)
- Magenta = Vermelho (1,0,0) + Azul (0,0,1) = (1,0,1)

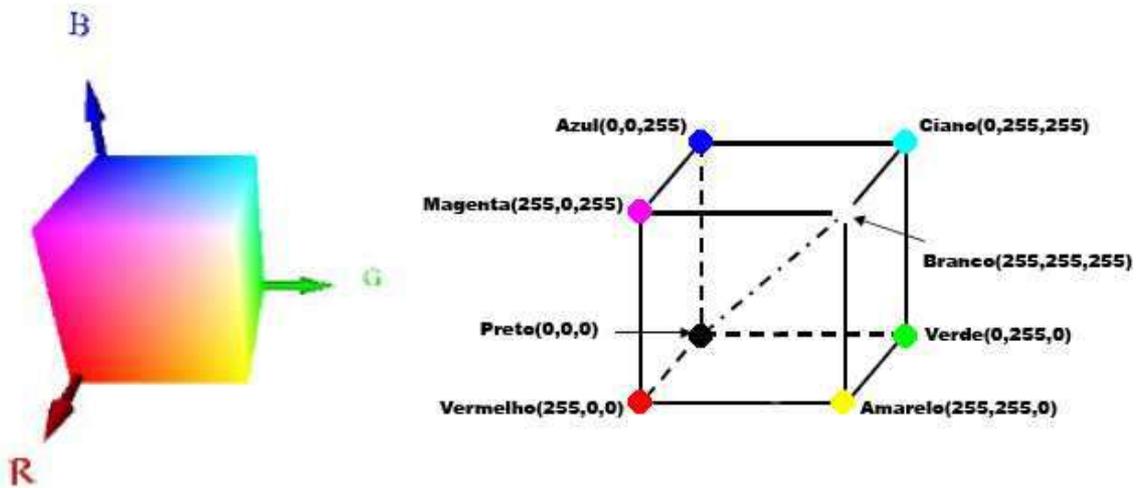


Figura 2.6: Espaço de cor do modelo RGB.

A denominação de cor complementar atribuída ao amarelo, ciano e magenta pela literatura, deve-se ao fato da localização destas, no cubo do espaço de cor RGB, ser oposta aos vértices das cores primárias (vermelho para o ciano, verde para o magenta e azul para o

amarelo) e, a adição da cor complementar à respectiva cor primária resultar sempre na cor branca (GOMES; VELHO, 1998), ou seja:

- Vermelho (1,0,0) + Ciano (0,1,1) = Branco (1,1,1)
- Verde (0,1,0) + Magenta (1,0,1) = Branco (1,1,1)
- Azul (0,0,1) + Amarelo(1,1,0) = Branco (1,1,1)

Os tons de cinza correspondem aos pontos situados sobre a diagonal principal do cubo, no qual as três componentes apresentam a mesma intensidade (x, x, x) com $0 \leq x \leq 1$, no espaço RGB.

Tradicionalmente, as implementações do modelo RGB nos sistemas gráficos empregam valores inteiros entre 0 e 255 para exprimir o valor da intensidade de cada componente em vez de valores reais normalizados entre 0 e 1. Esta tradição ainda se deve ao fato que o processamento de valores inteiros era muito mais rápido do que o processamento de valores reais nos primeiros sistemas gráficos, além do emprego de valores inteiros ser mais simples de escrever e apreender do que a representação com valores reais fracionários.

2.2.2 Modelo CIE

Modelo de cor criado pela CIE (*Commission Internationale de l'Éclairage*) em 1931 é um modelo com três cores primárias denominadas X, Y e Z, que substituem as cores primárias vermelho, verde e azul, e alcançam a representação de todas as cores do espectro visível (Figura 2.7).

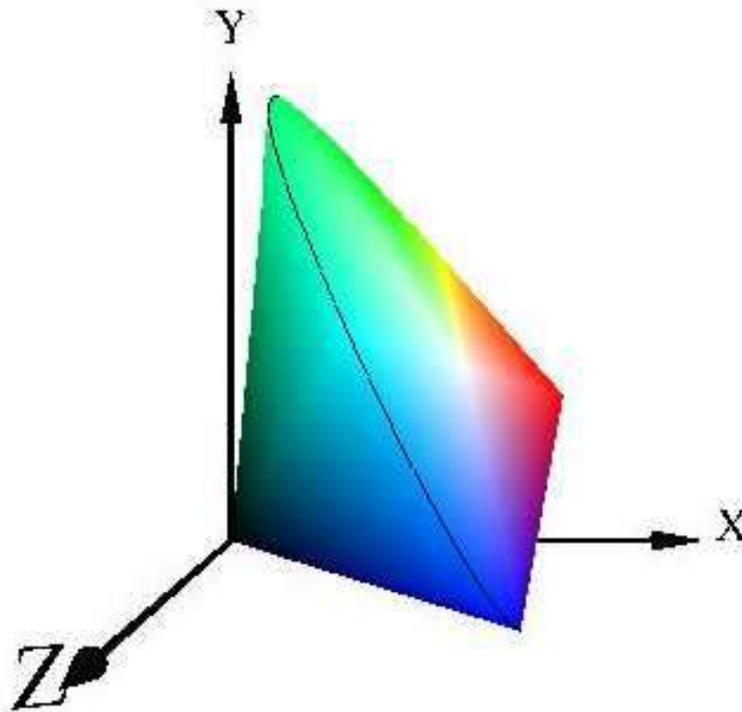


Figura 2.7: Espaço de cor do modelo CIE.

A transformação entre o modelo CIE XYZ e o modelo RGB (HUNT, 1991) é dado pelas equações 1 e 2:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M^{-1} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2)$$

e o valor de M é calculado conforme a equação 3, segundo CIE XYZ:

$$M = \begin{bmatrix} 0,430574 & 0,341550 & 0,178325 \\ 0,222015 & 0,706655 & 0,071330 \\ 0,020183 & 0,129553 & 0,939180 \end{bmatrix} \quad (3)$$

2.2.3 Modelo CMY

O modelo CMY (*Cyan Magenta Yellow*) é um modelo baseado nas cores complementares: ciano, magenta e amarelo (FIELD, 1988). Este modelo baseia-se no fenômeno que ocorre quando a luz branca incide em superfícies que podem absorver, refletir ou refratar a luz de forma desigual. A luz correspondente aos comprimentos de onda não absorvidos é, em geral, refletida. Assim, quando a luz natural incide numa superfície que absorve os comprimentos de onda na zona do vermelho, a luz refletida não possuirá quaisquer componentes nessa gama e será constituída por apenas as cores verdes e cores azuis, ou seja, o olho humano detectará a superfície percebendo a cor ciano. Esta subtração (por absorção) da luz em determinados comprimentos de onda é a razão pela qual o modelo CMY é também designado por modelo subtrativo da cor, em oposição ao modelo RGB que é um modelo aditivo da cor (Figura 2.8).

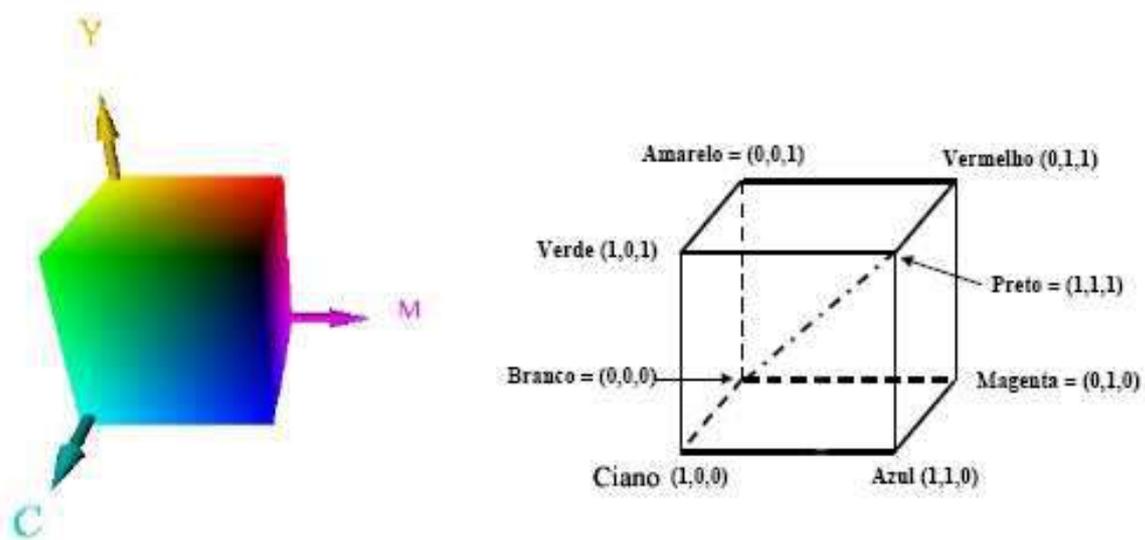


Figura 2.8: Espaço de cor do modelo CMY.

A transformação entre o modelo CMY e o modelo RGB (FIELD, 1988) é dado pelas equações 4 e 5.

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix} \quad (5)$$

2.2.4 Modelo HSL

O modelo HSL (*Hue, Saturation e Lightness*) é um modelo mais intuitivo, se comparado aos modelos RGB e CMY. Com estes parâmetros, o modelo HSL aproxima-se muito do modelo que emprega os conceitos qualitativos de matiz, luz e tonalidade. O parâmetro H indica a matiz ou a cor. O parâmetro S indica a saturação de uma cor. O parâmetro L corresponde à intensidade da cor e varia entre intensidade nula, ou seja, preto e intensidade máxima, o branco (Figura 2.9).

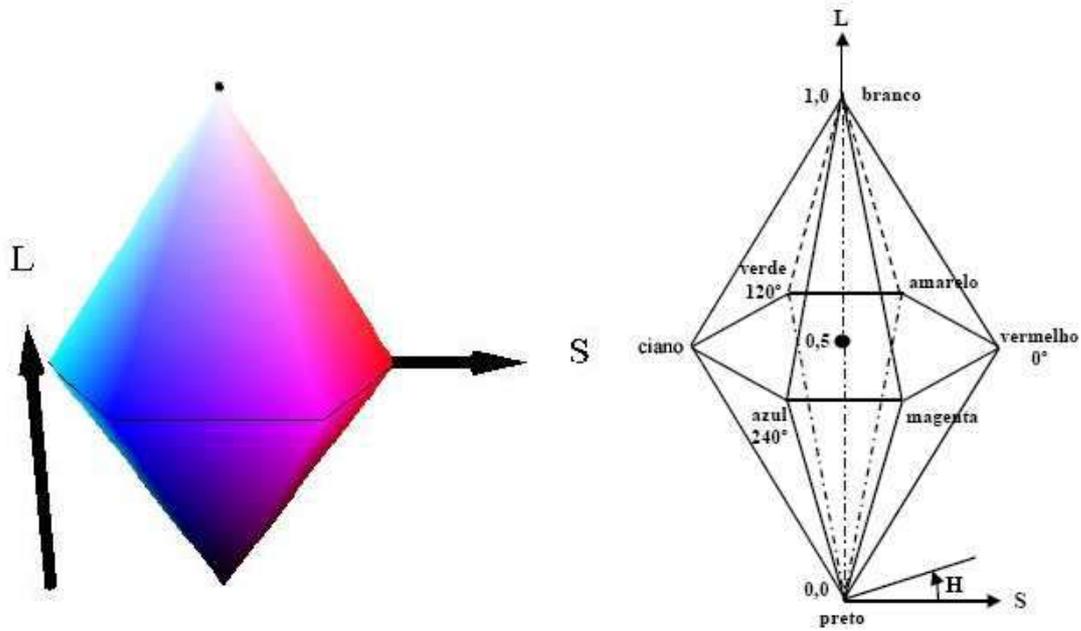


Figura 2.9: Espaço de cor do modelo HLS.

A transformação do modelo RGB para o modelo HSL (GONZALEZ; WOODS; EDDINS, 2004) é dada por:

$$H = \begin{cases} \theta & \rightarrow B \leq G \\ 360 - \theta & \rightarrow B > G \end{cases} \quad (6)$$

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G) + (R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\} \quad (7)$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)] \quad (8)$$

$$L = \frac{1}{3}(R+G+B) \quad (9)$$

A transformação do modelo HSL e para o modelo RGB é dada por:

$$\begin{aligned} (0^\circ \leq H \leq 120^\circ) \\ B = I(1 - S) \end{aligned} \quad (10)$$

$$R = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] \quad (11)$$

$$G = 3I - (R + B) \quad (12)$$

$$(120^\circ \leq H \leq 240^\circ) \quad (13)$$

$$H = H - 120^\circ$$

$$R = I(1 - S) \quad (14)$$

$$G = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] \quad (15)$$

$$B = 3I - (R + G) \quad (16)$$

$$(240^\circ \leq H \leq 360^\circ) \quad (17)$$

$$H = H - 240^\circ$$

$$G = I(1 - S) \quad (18)$$

$$B = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] \quad (19)$$

$$R = 3I - (G + B) \quad (20)$$

2.2.6 Modelo YIQ e YCbCr

Os modelos de cor YIQ e YCbCr (Figura 2.10) foram criados para permitir que as emissões dos sistemas de televisão em cores fossem compatíveis com os receptores preto e branco. O sistema NTSC foi criado em 1953 e emprega o modelo YIQ. Os sistemas SECAM (*Séquence Electronique Couleur Avec Mémoire*) e PAL (*Phase Alternating Line*), de 1961, empregam o modelo YCbCr. Os dois modelos são muito semelhantes e são baseados na separação dos sinais de cor RGB em um sinal de luminosidade, ou luminância (Y), e dois sinais de cromaticidade ou diferença de cor correspondentes às componentes: azul e vermelha. Desta forma, o sinal da televisão em cores transmite a luminância da mesma forma que o sinal da televisão em preto e branco e, portanto, os receptores preto e branco podem receber as emissões da televisão em cores (SPROSON, 1983).

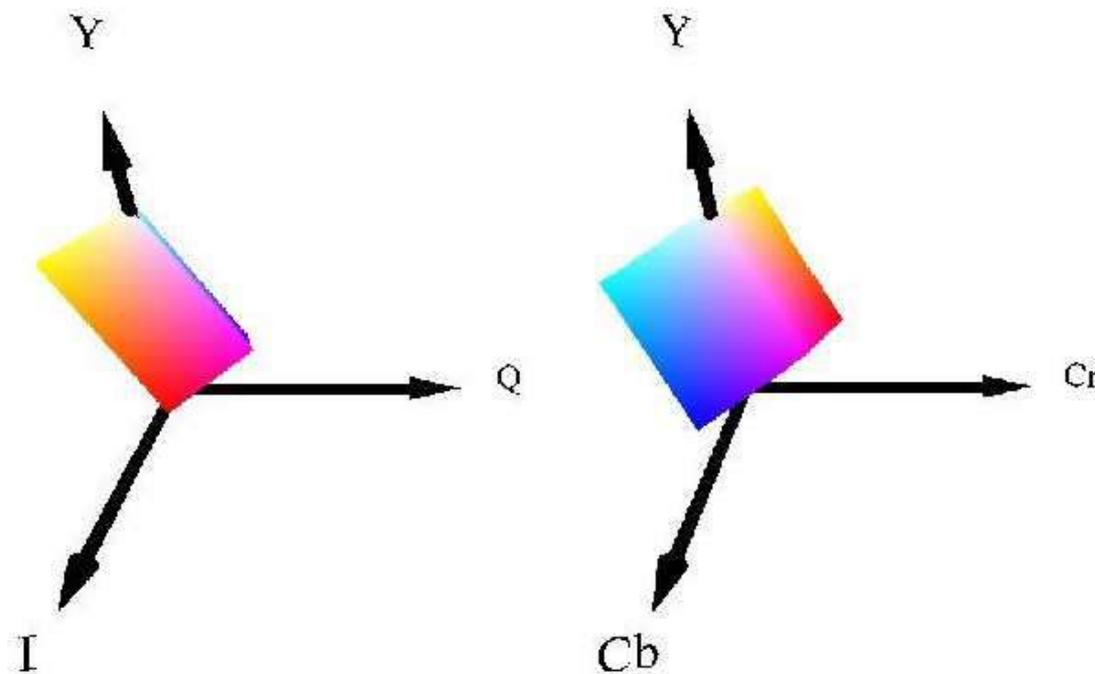


Figura 2.10: Espaços de cor dos modelos YIQ e YCbCr.

A transformação entre do RGB para o YIQ é mostrada nas equações 21-23:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = M \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (21)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M^{-1} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \quad (22)$$

$$M = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,275 & -0,321 \\ 0,212 & -0,523 & 0,311 \end{bmatrix} \quad (23)$$

A transformação do RGB para o YCbCr e é dada pelas equações 24-26.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = M \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (24)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M^{-1} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} \quad (25)$$

$$M = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,169 & -0,331 & 0,5 \\ 0,5 & -0,419 & -0,0813 \end{bmatrix} \quad (26)$$

2.3 Abordagens tradicionais de quantização de cores

Em uma imagem digital colorida definida no formato *True Color*, a cor de cada pixel é formada pelo menos por 24 bits gerando, mais de 16 milhões (2^{24}) de diferentes cores no espaço de cor RGB, enquanto em uma imagem em níveis de cinza, a cor de cada pixel da

imagem é formada por 8 bits, ou seja, 256 (2^8) diferentes cores, o que indica uma redução de 65.536 vezes o número de cores.

A redução do número de cores de uma imagem é uma tarefa importante para a representação, segmentação e compressão de imagens coloridas. O processo de redução do número de cores representados em um espaço de cor denomina-se quantização do espaço de cor. Há dois tipos de quantização: a estática, que faz uso de paletas de cores pré-definidas e a dinâmica, que utiliza agrupamentos e ou segmentação espacial. Formalmente, seja C um espaço de cores então um subespaço de C é denominado espaço de quantização. Uma função Q que mapeia cada cor de C para um elemento em P é denominada quantizadora ($Q : C \rightarrow P$), e é definida pela equação 27.

$$P = \left\{ c_1, c_2, \dots, c_n \mid c_i \in C, n \ll \|C\| \right\} \quad (27)$$

Diferentes métodos de quantização de cores são propostos na literatura de processamento de imagens. Dentre eles citam-se:

- Syeda-Mahmood (1992) particionou o espaço de cor RGB em 220 subespaços (categorias) nos quais a cor é perceptualmente a mesma e distintamente diferente dos subespaços vizinhos. Essa partição foi obtida por um experimento extensivamente psicofísico, o qual se examinou sistematicamente um dispositivo dependente do espaço de cor HSV. Utilizou-se uma tabela de índice para mapear os valores do RGB e as cores de cada categoria;
- Liu e Yang (1994) propuseram uma função para efetuar a segmentação de imagens coloridas que não requeria nenhum conjunto de parâmetros definidos pelo usuário, ou valores de *threshold* para definir características como tamanho, forma ou homogeneidade de regiões. Essa função incorporaria, de forma direta ou indireta, três

dos quatro critérios heurísticos sugeridos por Haralick e Shapiro (1985): i) a região deve ser uniforme e homogênea, ii) o interior da região deve ser simples, ou seja, possuir a menor quantidade possível de buracos e iii) regiões adjacentes devem apresentar valores significativamente diferentes para as características de uniformidade;

- Smith e Chang (1995) particionaram a componente *Hue* (matiz) do espaço HSV em 18 níveis, a componente *Value* (valor) em três níveis e a componente *Saturation* (saturação) em três níveis, num total de 162 níveis. Antes de se verificar o índice de cada pixel da imagem, um filtro da mediana é aplicado em cada componente HSV da imagem para enfatizar as regiões coloridas proeminentes;
- Mehtre et al. (1995) definiram de forma heurística uma pequena tabela de referência de cores do espaço RGB. Essa tabela referencia aproximadamente quase todas as cores das imagens que constituíam a aplicação, de forma que a cor de cada pixel de cada imagem é definida pela referência da tabela mais próxima da cor. Entretanto, quando o conjunto de imagens da aplicação crescia demasiadamente, era utilizado o algoritmo de Jaynes (1982) para maximizar o número de cores da tabela de referência. Essa maximização estabelecia que cada cor referenciada pela tabela deveria possuir a mesma probabilidade de ocorrência em toda sua aplicação;
- Ito et al. (1995) particionaram o espaço de cor HLS com auxílio da lógica *Fuzzy*, de tal modo que a cor era representada pelas funções de pertinências: *Lightness* baseada no histograma de cor do canal L entre as cores preto, cinza e branco e, *Hue* baseado no histograma de cor do canal H entre as cores azul, ciano, verde, amarelo, vermelho e magenta. Se, a saturação de um pixel da imagem fosse abaixo de um valor de *threshold*, o valor do *Hue* do pixel seria ignorado;

- O sistema QBIC (HAFNER et al, 1995) permite calcular um histograma com K cores do espaço de cor RGB, onde K possui valor inicial igual a 64, mas pode ser redefinido pelo usuário. Inicialmente, cada eixo do espaço de cor RGB é particionado em 16 níveis obtendo 4096 células. Todas as cores que pertencessem a cada célula são convertidas no espaço de cor modificado de Munsell (MIYAHARA; YOSHIDA, 1988). A partir daí as coordenadas do centro de cada célula são calculadas. Depois, são calculadas as K mínimas somas dos quadrados das coordenadas das células. Vailaya et al. (1998) aplicou uma partição similar no espaço de cores HSV;
- Gong et al. (1996) particionaram o espaço de cor modificado de Munsell (MIYAHARA; YOSHIDA, 1988) em 11 níveis de cores definidas e validadas empiricamente por diferentes grupos de examinadores. Cox et al. (1996) realizaram um particionamento semelhante no espaço de cor HSV no sistema *PicHunter* ;
- Pass et al. (1996) e Mitra et al. (1997) utilizaram somente os bits mais significativos (geralmente os dois primeiros) de cada canal do espaço de cor RGB para reduzir severamente o número de cores da imagem;
- Gagliardi e Schettini (1997) particionaram a gama de cores existentes em classes de equivalência de acordo com as denominações linguísticas para cor. Inicialmente, o espaço de cor CIELAB foi dividido em 256 subespaços, de modo que a cor de cada subespaço era perceptualmente a mesma, distintamente diferente dos subespaços vizinhos e denominada de acordo com as nomenclaturas definidas pelo Conselho da Sociedade Internacional de Cor e pela agência Nacional de Padrões – ISCC/NBS (*Inter-Society Color Council / National Bureau of Standards*) propostas em 1955. A partir daí, os subespaços foram agrupados em 13 classes de cores (preto, cinza, branco, vermelho, laranja, amarelo, verde, azul, violeta, roxo, rosa, marrom, verde oliva) e produziram uma segmentação não supervisionada da imagem;

- No sistema *ImageRover* (SCLAROFF et al.,1997) o espaço de cor RGB foi convertido no espaço de cor CIELUV e, a partir daí, cada eixo de cor (L, U e V) do espaço CIELUV é subdividido em 4 *bins* do mesmo tamanho, resultando em um número total de 64 *bins*, de modo que todas as cores muito escuras e muito claras não possuíssem uma saturação muito alta;
- Xiang (1997) preocupado com a questão de minimizar a discrepância entre a cor original de um pixel da imagem e a correspondente cor definida pelo método de quantização de cores, propôs um método de quantização que, inicialmente, agrupava todas as cores de uma imagem em *clusters* (agrupamentos) tão pequenos quanto possível utilizando o algoritmo de *clustering nonhierarchial* (agrupamentos não hierárquicos) de Gonzales (1985) e depois atribuía uma cor ao *cluster*, representada pelo seu centróide. Esse método utilizava vetores de cor com 24 *bits* e minimizava a distorção entre a imagem original e a imagem quantizada;
- Ciocca e Schettini (1999), inicialmente, geraram amostras randômicas contendo milhões de cores do espaço de cor RGB. Após esse procedimento, as amostras foram mapeadas no espaço de cor CIELAB, em que o algoritmo de *cluster* competitivo de Uchiyama e Arbib (1994) foi aplicado para localizar as 64 cores mais significativas. A cor de cada pixel da imagem que fosse indexada era mapeada no espaço de cor CIELAB e, então, atribuída a um dos 64 centróides que mais se aproximam da cor;
- Papamarkos et al. (2002) propuseram um novo método de quantização de cores que utilizava as características espaciais locais do pixel da imagem (valores mínimo, máximo e de entropia), além das cores. Utilizando as características espaciais e cor de cada pixel da imagem foi definida uma estrutura do tipo árvore composta por m classes de características, de modo que, as condições predefinidas estabelecessem uma classe de características que deveriam ser subdivididas ou incorporadas a outras. Em

sequência, utilizaram *PCA* e a rede neural de *Kohonen* com a finalidade de definirem a cor de cada classe de características;

- Liapis e Tziritas (2004) utilizam o espaço de cor CIELAB, por ser considerado perceptualmente uniforme. Descartam o *lightness* representado pela componente L e calculam os histogramas (1-D) com 232 *bins* da componente A e com 233 *bins* da componente B. Por meio da utilização da distribuição gaussiana, calcula-se a média e a variância das coordenadas das cores da imagem;
- Sirisathitkul et al. (2004) utilizaram o espaço de cor RGB para efetuar uma divisão hierárquica do mapa de cores. Em princípio, criaram um histograma (3D) com os 24 *bits* da cor de cada pixel de uma imagem que serviria como espaço de cor a ser quantizado. A partir daí, ordenaram as cores do histograma de forma que, a primeira cor possuísse a maior variância quanto à distribuição de cores. Os autores definiram essas cores ordenadas como a primeira célula do método de quantização, em seguida dividiram a célula em duas, com a finalidade de balancear a soma da distância entre as cores adjacentes e a frequência de distribuição das mesmas, que pertenciam à célula. Essa divisão foi realizada em todas as células existentes e criadas até o momento da obtenção do balanceamento. A cor representante de cada célula foi calculada pelo centróide da célula;
- Qiu et al. (2004) investigaram a compreensão, ou seja, a redução da dimensionalidade de descritores baseados em histogramas de cores por meio do *PCA* e utilizaram em um *support vector machine* para treinar e classificar as imagens dos auto-histogramas gerados pelo *PCA*;
- Kim e Kehtarnavaz (2005) utilizaram a transformada de *wavelet* discreta (DWT) para determinar automaticamente o número de representantes de cada partição (*cluster*) no

algoritmo multi-escala *clustering*, entretanto tentavam manter a distorção das cores quantizadas em um nível aceitável;

- Papamarkos e Atsalakis (2006) desenvolveram um novo classificador em rede neural que combinava as características de auto-crescimento (*Growing Neural Gas*) e de auto-organização (*Kohonen Feature Map*). Essa combinação produziu uma nova rede neural com características proeminentes que introduzia uma série de critérios que tinha efeito sobre a introdução ou remoção de neurônios, além de permitir o número de neurônios e sua topologia;
- Suzuki e Furukawa (2007) investigaram a redução do número necessário de *bits* para a quantização de cor em imagens definidas no espaço CIE XYZ pela aplicação do método de Taylor para calcular os valores de uma região segmentada;
- Fun e Wong (2008), verificaram que a abordagem de quantização de cor no espaço HSV é perceptivamente mais uniforme na forma cônica que forma cilíndrica.

Um resumo dos diferentes métodos de quantização de cores pode ser visto na Tabela 1.

2.4 Métricas de Avaliação de Desempenho

Em um processo de classificação podem ocorrer erros. Nesse sentido, a literatura apresenta algumas métricas de avaliação de desempenho quanto ao processo de classificação, entre elas pode-se citar a curva ROC e a matriz de confusão.

Tabela 1: Métodos de Quantização de cores

Escopo	Método	Trabalhos
Estático	<i>Bits</i> mais significativos	<ul style="list-style-type: none"> ▪ Pass et al. (1996) ▪ Mitra et al. (1997)
	Partição do espaço de Cor	<ul style="list-style-type: none"> ▪ Syeda-Mahmood (1992) ▪ Hafner et al. (1995) ▪ Ito et al. (1995) ▪ Smith e Chang (1995) ▪ Gong et al. (1996) ▪ Gagliardi e Schetini (1997) ▪ Sclaroff et al. (1997) ▪ Vailaya et al. (1998) ▪ Liapis e Tziritas (2004) ▪ Qiu et al. (2004) ▪ Fun e Wong (2008)
	Agrupamento (<i>clusters</i>) do espaço de cor	<ul style="list-style-type: none"> ▪ Xiang (1997) ▪ Ciocca e Schettini (1999) ▪ Papamarkos et al. (2002) ▪ Sirisathitkul et al. (2004) ▪ Kim e Kehtarnavaz (2005) ▪ Papamarkos e Atsalakis (2006)
	Referência de cores	<ul style="list-style-type: none"> ▪ Mehtre et al. (1995)
Dinâmico	Segmentação da imagem	<ul style="list-style-type: none"> ▪ Liu e Yang (1994) ▪ Suzuki e Furukawa (2007)

2.4.1 Curva ROC (*Receiver Operating Characteristic*)

Em um processo de classificação podem ocorrer dois tipos de erros (MALTONI, 2003) na comparação do modelo com a amostra: erro da falsa aceitação (*false match*), no qual mesmo que a amostra seja diferente do modelo, o sistema classifica-a como igual, aceitando-a; erro da falsa rejeição (*false no match*), no qual mesmo que a amostra se mostre similar ao modelo, o sistema não a classifica corretamente, rejeitando-a (Figura 2.11). Portanto, a falsa aceitação implica na aceitação de um impostor, enquanto a falsa rejeição implica na rejeição de um genuíno.

A formulação do problema pode ser feita da seguinte maneira (MALTONI, 2003): seja M o modelo da característica armazenada no sistema e A a amostra a ser observada. Apresentam-se, então, as seguintes hipóteses:

- $H_0 : A \neq M$ significa que a amostra A não pertence ao modelo M ;
- $H_1 : A = M$ significa que a amostra A pertence ao modelo M .

e as seguintes decisões:

- D_0 : impostor;
- D_1 : genuíno.



Figura 2.11: Tipos de erros em sistemas de classificação de pele humana.

A verificação de igualdade ou não entre a amostra A e o modelo M é efetuada pelo sistema de classificação usando métricas de similaridade, representadas por $s(A, M)$. Essas

técnicas obtêm valores que são analisados pelo sistema de decisão, por meio de um valor de um limiar (*threshold*) t , para que se possa optar por D_0 ou D_1 , dado pela equação 28.

$$s(A, M) \begin{cases} \geq t \rightarrow D_1 \\ < t \rightarrow D_0 \end{cases} \quad (28)$$

O sistema de decisão pode cometer dois tipos de erros:

- *Tipo I*: falsa aceitação ocorre quando o sistema de classificação decide por D_1 mesmo que H_0 seja verdadeiro;
- *Tipo II*: falsa rejeição ocorre quando o sistema de classificação decide por D_0 mesmo que H_1 seja verdadeiro.

Seja *FMR* (*False Match Rate*) a probabilidade de ocorrer o erro do *Tipo I* e *FNMR* (*False Non-Match Rate*) a probabilidade de ocorrer o erro do *Tipo II*, calcula-se:

- $FMR = P(D_1 | H_0 = true)$;
- $FNMR = P(D_0 | H_1 = true)$.

Sabendo-se que os valores de decisão gerados pelo sistema de classificação das amostras genuínas são denominados *distribuição de genuínos* e os das amostras de impostores são denominados *distribuição de impostores*, a Figura 2.12 ilustra o cálculo da FMR e FNMR sobre estas distribuições.

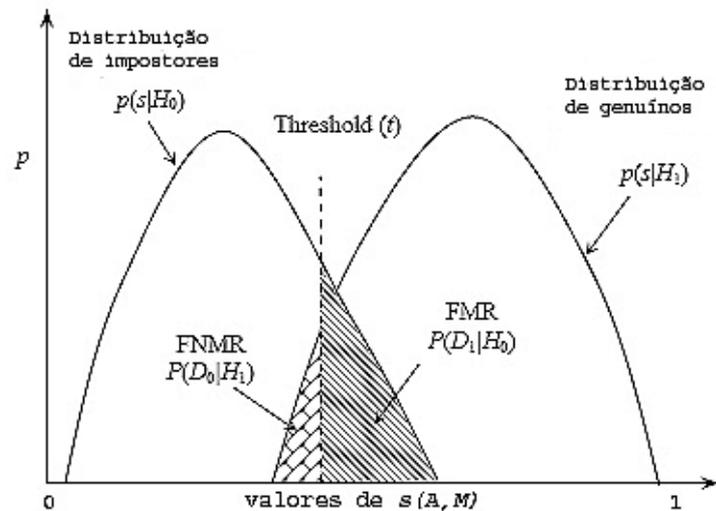


Figura 2.12: *FMR* e *FNMR* calculados para um *threshold* t sobre os valores das distribuições de *genuínos* e *impostores*.

Pelo fato das funções *FMR* e *FNMR* serem calculadas sobre um *threshold* t , devem ser referenciadas como $FMR(t)$ e $FNMR(t)$. Na Figura 2.13, verifica-se que, ao se diminuir o valor de t , o sistema de classificação torna-se mais tolerante às variações e ruídos aumentando, dessa forma, o valor da $FMR(t)$ e permitindo a aceitação de impostores, enquanto ao se aumentar o valor de t , o classificador torna-se mais rígido às variações e ruídos, assim, aumenta o valor da $FNMR(t)$ e pode ocorrer a rejeição de genuínos.

De acordo com as distribuições expressas nas curvas da Figura 2.13-b, podem-se verificar três fatores de erros relevantes para os classificadores (ver Figura 2.14):

- *EER (Equal-ErrorRate)*, que denota a mesma taxa de erro para $FMR(t)$ e $FNMR(t)$;
- *Zero FNMR*, que denota o ponto no qual não ocorre falsa rejeição (*FNMR*);
- *Zero FMR*, que denota o ponto no qual não ocorre falsa aceitação (*FMR*).

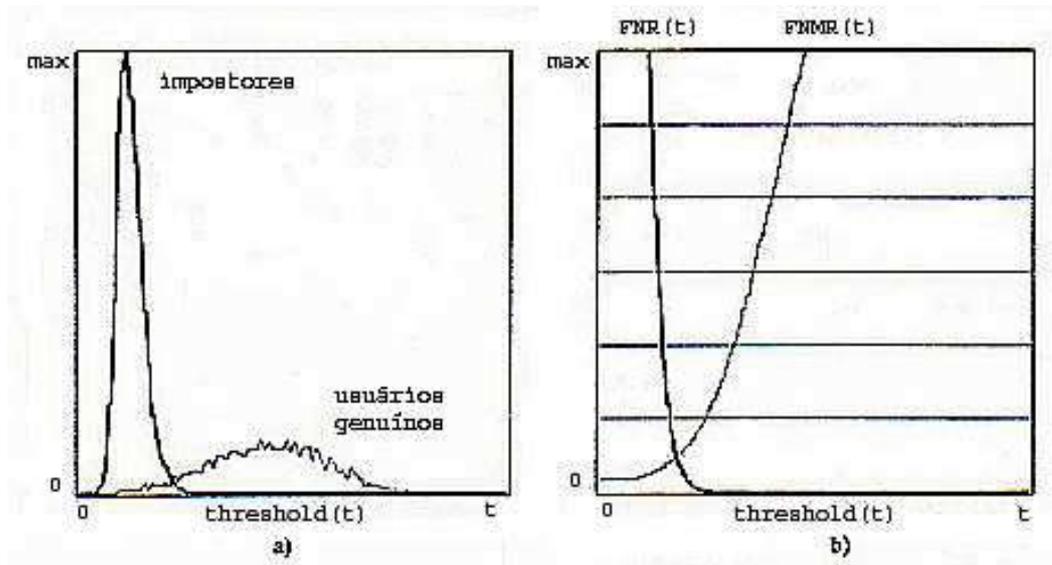


Figura 2.13: Variação do *threshold* t em um sistema de reconhecimento biométrico. a) Distribuição de impostores e usuários genuínos. b) $FMR(t)$ e $FNMR(t)$ calculados pelos valores em a).

Ao se analisar os três erros ilustrados na Figura 14, obtém-se a curva *ROC* que expressa a curva dada pelos pontos de $FMR(t)$ e $FNMR(t)$, para vários valores do *threshold* t . A Figura 2.15 ilustra uma Curva *ROC* onde se destacam:

- As aplicações forenses denotadas pelo ponto *Zero FNMR*, que torna o classificador mais tolerante às variações e ruídos e, desta forma, permite a aceitação de impostores;
- As aplicações comerciais denotadas pelo ponto ERR, que possuem valores iguais para $FMR(t)$ e $FNMR(t)$;
- As aplicações de alta segurança denotadas pelo ponto *Zero FMR*, que torna o classificador mais rígido às variações e ruídos aumentando e, desta forma, possibilita a ocorrência da rejeição de genuínos.

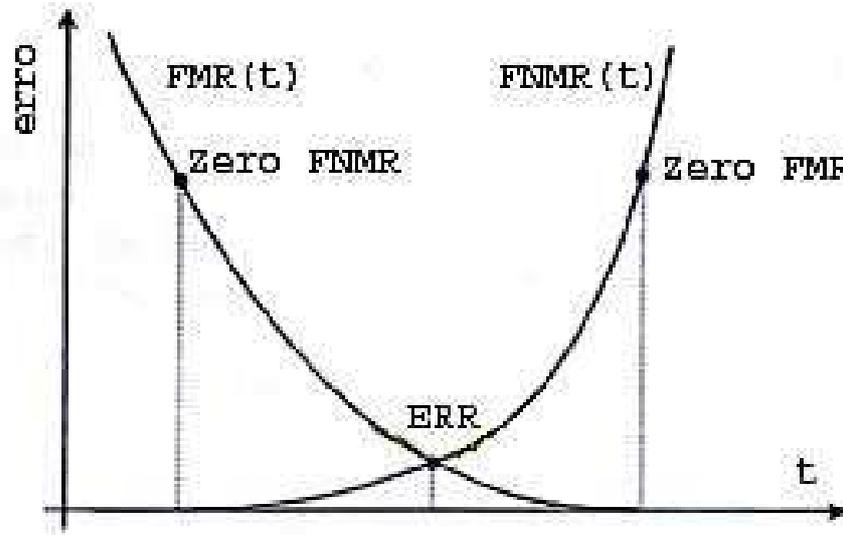


Figura 2.14: Exemplos das curvas $FMR(t)$ e $FNMR(t)$ e os pontos de erro: ERR , $Zero FNMR$ e $Zero FMR$.

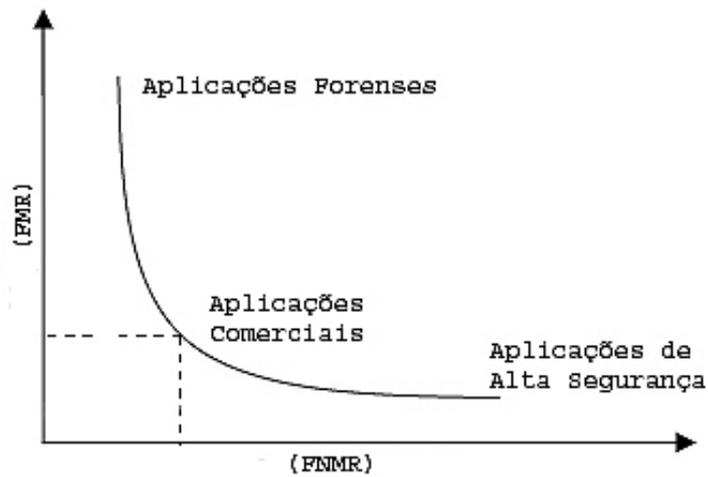


Figura 2.15: Curva ROC segundo a abordagem utilizada pela Biometria.

2.4.2 Matriz de Confusão

A matriz de confusão, ou matriz de erro, é uma matriz quadrada de números que expressam a quantidade de unidades amostrais associada a uma dada categoria durante o processo de classificação efetuado e à categoria real às quais pertencem essas unidades (ROSENFELD; FITZPATRICK, 1986). Desta forma, a matriz de confusão de uma hipótese h , oferece uma medida efetiva do modelo de classificação ao mostrar o número de classificações corretas *versus* as classificações preditas para cada classe, sobre um conjunto de teste T , dada pela equação 29.

$$M(C_i, C_j) = \sum \left\| h(x) = C_j \right\|, \forall (x, y) \in T, y = C_i \quad (29)$$

Classe	classificada como C_1	classificada como C_2	...	classificada como C_r
C_1	$M(C_1, C_1)$	$M(C_1, C_2)$...	$M(C_1, C_r)$
C_2	$M(C_2, C_1)$	$M(C_2, C_2)$...	$M(C_2, C_r)$
\vdots	\vdots	\vdots	\vdots	\vdots
C_r	$M(C_r, C_1)$	$M(C_r, C_2)$...	$M(C_r, C_r)$

Figura 2.16: Representação de uma matriz de confusão.

A Figura 2.16 ilustra que o número de acertos para cada classe se localiza na diagonal principal $M(C_i, C_i)$ da matriz, enquanto os demais elementos $M(C_i, C_j)$, para $i \neq j$, representam erros na classificação.

Um dos parâmetros mais utilizados e mais eficientes na determinação da exatidão da classificação é o coeficiente de concordância *Kappa*, que é obtido por meio da matriz de confusão (CONGALTON; GREEN, 1993). O coeficiente *Kappa* (K), é calculado a partir da

equação 30 (BISHOP; et. al., 1975), sendo que N representa o número de observações (pontos amostrais), enquanto r representa o número de linhas da matriz confusão.

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_i \times x_j)}{N^2 - \sum_{i=1}^r (x_i \times x_j)} \quad (30)$$

$$x_{ii} = M(C_i, C_j) \quad (31)$$

$$x_i = \sum_{j=1}^r M(C_i, C_j) \quad (32)$$

$$x_j = \sum_{i=1}^r M(C_i, C_j) \quad (33)$$

Entre outras métricas derivadas da matriz de confusão, cita-se:

- a Taxa de Erro da Classe (TEC) da linha i :

$$TEC_i = \frac{\sum_{j=1}^r M(C_i, C_j) - M(C_i, C_i)}{\sum_{j=1}^r M(C_i, C_j)} \quad (34)$$

- a Exatidão Global (EG):

$$EG = \frac{\sum_{i=1}^r x_{ii}}{N} \quad (35)$$

2.5 Considerações Finais

Neste capítulo foi apresentado que na retina existem dois tipos de fotorreceptores: os cones e os bastonetes. Os bastonetes proporcionam a detecção das formas dos objetos, enquanto os cones são sensíveis à luz na região do vermelho, do verde e do azul. A diferença entre as respostas dos três tipos de cones (vermelho, verde e azul) permite interpretar diferentes comprimentos de onda que correspondem a cores diferentes. Essa interpretação é conhecida como Teoria dos Três Estímulos. Tomando por base a Teoria dos Três Estímulos, a característica essencial na modelagem da cor está na definição de como essas três grandezas devem ser associadas. Essa associação origina a existência de vários modelos de cor.

Capítulo 3

Mistura de Cores

3.1 Introdução

Segundo Morris (2000), a noção de cor, nas artes visuais, encontra-se ligada aos pigmentos das tintas utilizadas e aos termos *cor pura*, *pastel*, *sombreado* e *tom*. O sombreado é o resultado da adição de um pigmento negro a um pigmento de cor pura, sendo que o pigmento negro retira a luminosidade do pigmento puro. A cor pastel é o resultado da adição de um pigmento branco a um pigmento puro que leva à diminuição da saturação do pigmento puro. O tom resulta da adição simultânea de um pigmento negro mais um pigmento branco mais um pigmento puro, o que resulta na redução da saturação e da luminosidade do pigmento puro. Sendo assim, a definição de cor num sistema de artes visuais é dada pela especificação dos pigmentos de cores puras, dos brancos e dos negros, para que os tons pretendidos sejam obtidos.

A pintura de uma tela artística não é realizada numa única etapa. Na Figura 3.1 verificam-se os passos efetuados para a pintura de uma tela artística de um druida (RIBEIRO, 2005).

No primeiro passo é esboçado o druida com um lápis *dermograph* azul direto na tela e, após o esboço, aplica-se uma camada fina e transparente de um fixador, no segundo passo foi aplicada a cor sombra queimada e reforçado os contornos. Com a finalidade de estabelecer as regiões claras e escuras, aplica-se uma camada transparente de carmim, misturado com sombra queimada e diluídos numa mistura de terebintina, secante de cobalto e óleo de linhaça. No terceiro passo, inicia-se a pintura do rosto. No quarto passo, pinta-se o fundo do quadro e no quinto define-se o cabelo, a barba e a roupa. No último passo, aplica-se uma película da mistura de carmim e verniz para alterar a matiz das cores e dar um efeito à pintura final da tela. Durante a evolução dos passos, uma nova camada de pigmentos de cores é misturada às camadas anteriores e, como consequência, novos tons vão surgindo. Esse processo se repete até que as tonalidades desejadas sejam alcançadas.



Figura 3.1: Evolução dos passos realizados para a pintura de uma tela artística. (RIBEIRO, 2005)

3.2 Hipótese da Mistura de Cores

Baseando-se na decomposição em camadas de uma pintura artística, na mistura de pigmentos de cor pura, de cor branco e de cor negro para obtenção de tons na pintura artística, na proporção média de distribuição (2:1) dos cones sensíveis na faixa do vermelho em relação aos sensíveis na faixa do verde e, no fato da visualização de imagens digitais utilizarem o espaço de cor RGB, este trabalho propõe a mistura de cores, que utiliza a decomposição em camadas da representação binária dos valores R, G e B do espaço de cor RGB para definição de cor.

3.2.1 Camadas Binárias das Cores

Estendendo-se o conceito de decomposição em camadas e, a mistura de pigmentos de cor pura, branco e negro para obtenção de novos tons das artes visuais para o conceito de cor no espaço RGB, pode-se definir $C = (R, G, B)$ como uma cor que, representada na forma binária em um sistema computacional com 24 *bits* (8 *bits* para cada componente), é igual a:

$$C = (R_7 R_6 R_5 R_4 R_3 R_2 R_1 R_0, \\ G_7 G_6 G_5 G_4 G_3 G_2 G_1 G_0, \\ B_7 B_6 B_5 B_4 B_3 B_2 B_1 B_0) \quad (36)$$

Os *bits* R_0 , G_0 e B_0 correspondem aos *bits* menos significativos de R, G e B, e os *bits* R_7 , G_7 e B_7 correspondem aos mais significativos. Desta forma, os valores de R, G e B variam

de 0 (00000000) a 255 (11111111). Dado que cada componente possui 8 *bits*, pode-se também definir oito camadas K_i , dada pela equação 37.

$$K_i = (R_i, G_i, B_i) \text{ para } i = 0 \dots 7 \quad (37)$$

K_i define a *i-ésima* camada binária da cor do pixel. Nota-se que cada camada K_i define as cores dos pigmentos puros (vermelho, verde e azul), dos pigmentos complementares (ciano, magenta e amarelo), e do branco e do negro, de forma que:

- $K_i = (0, 0, 0)$ expressa a ausência das cores vermelho, verde e azul e define o preto;
- $K_i = (0, 0, 1)$ define a cor azul;
- $K_i = (0, 1, 0)$ define a cor verde;
- $K_i = (0, 1, 1)$ expressa a mistura das cores: verde e azul e define a cor ciano;
- $K_i = (1, 0, 0)$ define a cor vermelho;
- $K_i = (1, 0, 1)$ expressa a mistura das cores: vermelho e azul e define a cor magenta;
- $K_i = (1, 1, 0)$ expressa a mistura das cores: vermelho e verde e define a cor amarelo;
- $K_i = (1, 1, 1)$ expressa a mistura das cores: vermelho, verde e azul e define a cor branco.

A Figura 3.2 exemplifica o quadro do druida nas camadas K_i que compõe a representação binária de cada *pixel* do quadro.

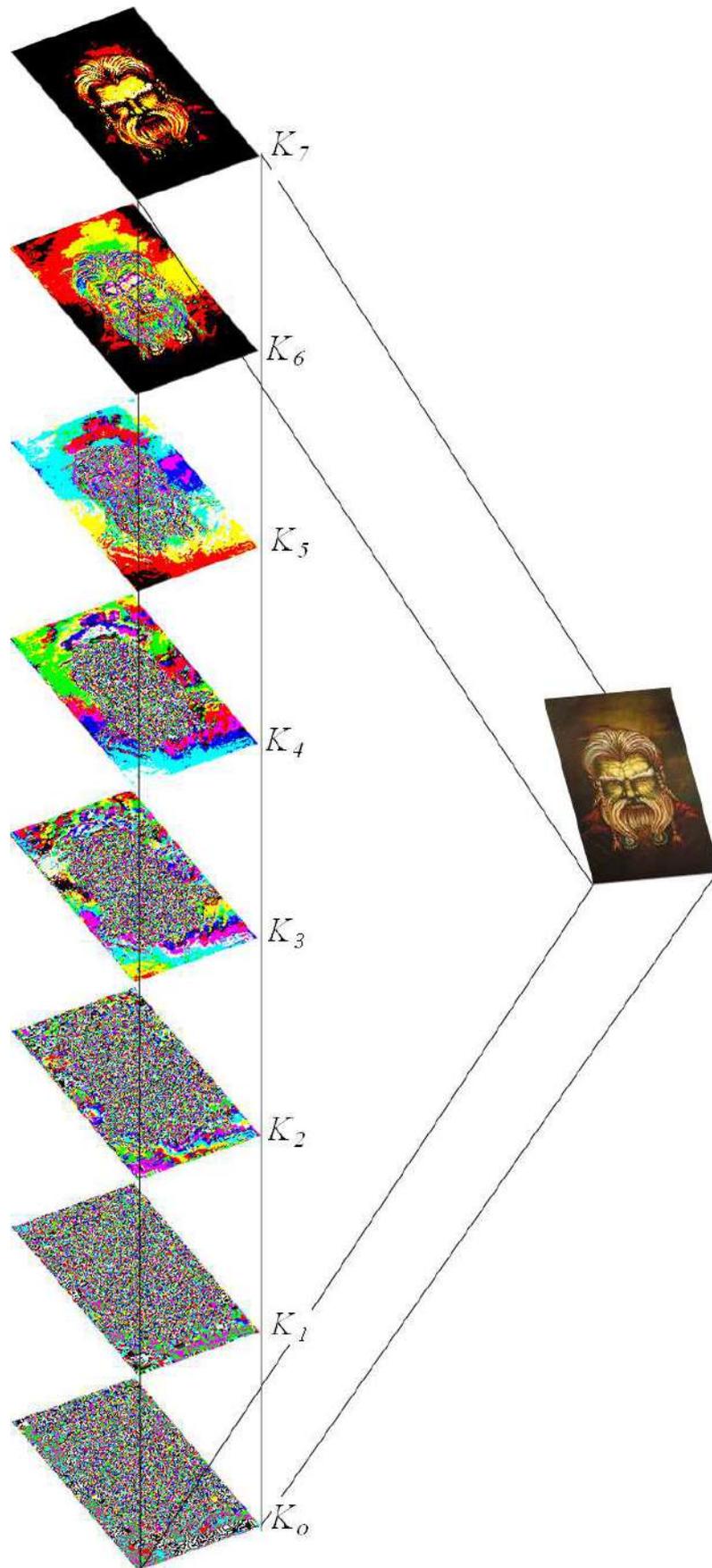


Figura 3.2: Decomposição do quadro do druida nas camadas K_i .

3.2.2 Mistura de cores

DeVries (1946), Rushton e Baker (1964), Cicerone e Nерger (1989), Pokorny et al. (1991), Yamaguchi et al. (1997), Hagstrom et al. (1998), Brainard et al. (1999), Deeb et al. (2000), Kremeres et al. (2000), Otake e Cicerone (2000), Carroll et al. (2002) e Gunther e Dobkins (2002) têm, por meio de medidas indiretas, sugerido que em média existem mais cones do tipo R do que cones do tipo G na retina humana.

Hofer et al. (2005) caracterizaram arranjos de cones sensíveis aos comprimentos de onda curto, médio e longo em oito mosaicos da fóvea humana, determinando que os seres humanos com visão normal apresentam a proporção 1,1:1 até 16,5:1 de cones de comprimento longo (faixa sensível à luz na região do vermelho) em relação aos de comprimento médio (faixa sensível à luz na região do verde).

Baseado nos estudos de Hofer et. al., 2005, este trabalho propõe o modelo de mistura da quantidade das cores primárias em cada camada de acordo com a equação 38, considerando-se que a retina humana, na captura das cores, possui maior quantidade de cones para a cor vermelha, seguida pela cor verde e menos cones para a cor azul. O modelo da equação 38 adota uma distribuição de pesos binários visando a implementação digital. Define-se, assim, o valor numérico NK da cor de cada camada, obtido pelo valor decimal dos *bits* que a compõe.

Pode-se dizer que o valor de NK modela uma “retina artificial” no que se refere à distribuição de sensores para R, G e B.

$$NK_i = 2^2R_i + 2^1G_i + 2^0B_i, \quad \text{para } i = 0, \dots, 7 \quad (38)$$

e assim,

- $NK_i = 0$ é representado pelo preto;
- $NK_i = 1$ é representado pelo azul;
- $NK_i = 2$ é representado pelo verde;
- $NK_i = 3$ é representado pelo ciano;
- $NK_i = 4$ é representado pelo vermelho;
- $NK_i = 5$ é representado pelo magenta;
- $NK_i = 6$ é representado pelo amarelo;
- $NK_i = 7$ é representado pelo branco.

Como cada elemento de cor considerado é um número de 8 bits, seu valor binário convertido para seu valor decimal é obtido multiplicando-se dígito a dígito por um peso P_i na base 2, tal que:

$$P_i = 2^i, i = 0, \dots, 7 \quad (39)$$

$$SP = \sum_{i=0}^7 P_i = 2^8 - 1 \quad (40)$$

SP define a soma de todos os pesos. Logo, N_i é a normalização de P_i e, é dada pelas equações 41-42:

$$N_i = \frac{2^i}{SP} \quad (41)$$

$$NP = \sum N_i = 1 \quad (42)$$

Dessa forma, à cada camada K_i é associado um peso normalizado N_i que correspondente à influência que a cor expressa na camada K_i exerce na representação final da cor.

Considerando-se o uso de cor nas artes visuais, pode-se afirmar que as camadas de menor peso equivalem às camadas inferiores nas telas artísticas e por isso têm menor influência na visualização da obra, ou seja, as noções de cores finais, formas, sombras e tonalidades de tinta são mais afetadas pelas camadas superiores de tinta, ou seja, pelas camadas de maior peso N_i .

A Tabela 2 mostra o exemplo de uma cor digital $C = (109, 146, 137)$ no espaço RGB, que na representação binária com 8 bits/cor é $C = (01101101, 10010010, 10001001)$. Aplicando-se a hipótese da Mistura de cores, pode-se dizer, que esta cor foi obtida através da mistura de 50,20% do pigmento ciano, 39,22% do pigmento vermelho, 7,05% do pigmento verde e 3,53% do pigmento magenta. Estas porcentagens foram calculadas da seguinte maneira (Tabela 3):

- $k_7 = (0,1,1)$ corresponde à cor ciano com um peso $N_7 = \frac{2^7}{2^8 - 1} = 0,5020$;
- $k_6 = (1,0,0)$ corresponde à cor vermelha com um peso $N_6 = \frac{2^6}{2^8 - 1} = 0,2510$;
- $k_5 = (1,0,0)$ corresponde à cor vermelha com um peso $N_5 = \frac{2^5}{2^8 - 1} = 0,1255$;
- $k_4 = (0,1,0)$ corresponde à cor verde com um peso $N_4 = \frac{2^4}{2^8 - 1} = 0,0627$;
- $k_3 = (1,0,1)$ corresponde à cor magenta com um peso $N_3 = \frac{2^3}{2^8 - 1} = 0,0314$;
- $k_2 = (1,0,0)$ corresponde à cor vermelha com um peso $N_2 = \frac{2^2}{2^8 - 1} = 0,0157$;
- $k_1 = (0,1,0)$ corresponde à cor verde com um peso $N_1 = \frac{2^1}{2^8 - 1} = 0,0078$;

- $k_0 = (1,0,1)$ corresponde à cor magenta com um peso $N_0 = \frac{2^0}{2^8 - 1} = 0,0039$.

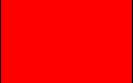
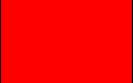
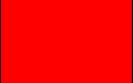
De forma, a cor $C = (109, 146, 137)$ é constituída por (Figura 3.3):

- Ciano presente em (K_7) totalizando 50,2%;
- Vermelho presente em $(K_6 + K_5 + K_2)$ totalizando 39,22%;
- Verde presente em $(K_4 + K_1)$ totalizando 7,05%;
- Magenta presente em $(K_3 + K_0)$ totalizando 3,53%.

Tabela 2: Representação binária da cor $C = (109, 146, 137)$.

R								G								B							
109								146								137							
0	1	1	0	1	1	0	1	1	0	0	1	0	0	1	0	1	0	0	0	1	0	0	1
R_7	R_6	R_5	R_4	R_3	R_2	R_1	R_0	G_7	G_6	G_5	G_4	G_3	G_2	G_1	G_0	B_7	B_6	B_5	B_4	B_3	B_2	B_1	B_0

Tabela 3: Representação dos cálculos K_i , NK_i e N_i da cor $C = (109, 146, 137)$.

i	K_i	R_i	G_i	B_i	Cor	NK_i	N_i
7	(R_7, G_7, B_7)	0	1	1		3	0,5020
6	(R_6, G_6, B_6)	1	0	0		4	0,2510
5	(R_5, G_5, B_5)	1	0	0		4	0,1255
4	(R_4, G_4, B_4)	0	1	0		2	0,0627
3	(R_3, G_3, B_3)	1	0	1		5	0,0314
2	(R_2, G_2, B_2)	1	0	0		4	0,0157
1	(R_1, G_1, B_1)	0	1	0		2	0,0078
0	(R_0, G_0, B_0)	1	0	1		5	0,0039

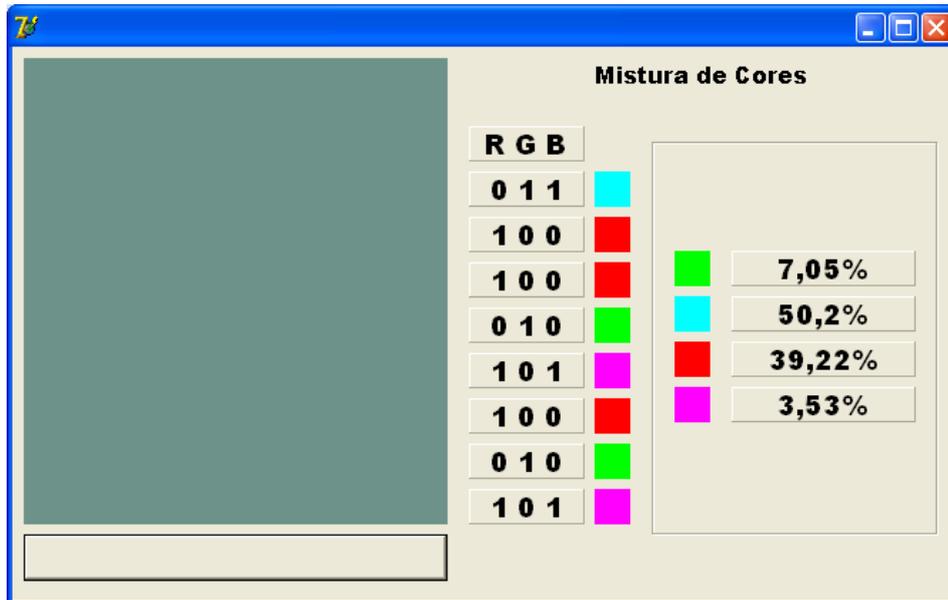


Figura 3.3: Representação de uma cor $C = (109,146,137)$ composta pela mistura das cores: verde, ciano, vermelho e magenta.

Pode-se abstrair que para se obter a cor mostrada na Figura 3.3, o artista deve utilizar na primeira camada, 0,39% de tinta magenta, na segunda 0,78% de verde, na terceira 1,57% de vermelho, na quarta deve aplicar mais 3,14% de magenta, na quinta camada mais 6,27% de pigmento verde, na sexta 12,55% de vermelho, na sétima aumentar o vermelho para mais 25,10% de pigmento e finalmente passar 50,20% de ciano na última camada.

É claro que o raciocínio indutivo não equivale à prática, mas apenas a uma maneira de comparar o valor da mistura de cores à sua inspiração.

Como o valor da mistura é diretamente derivado do RGB, também é aditivo. A cor final da mistura pode ser dada pela soma da contribuição da cor representada em cada camada, ou seja, pela soma do valor numérico de cada camada NK multiplicado pelo seu respectivo peso normalizado N . Assim, define-se o valor da mistura v como mostra a equação 43.

$$v = \sum_{i=0}^7 (4R_i + 2G_i + 1B_i) \frac{2^i}{2^8 - 1} \quad (43)$$

Ou seja,

$$v = \sum_{i=0}^7 NK_i N_i \quad (44)$$

Portanto, a cor $C = (109, 146, 137)$ possui:

- Ciano $NK = 3$ na proporção de 0,5020;
- Vermelho $NK = 4$ na proporção de 0,3922;
- Verde $NK = 2$ na proporção de 0,0705;
- Magenta $NK = 5$ na proporção de 0,0353.

ou seja, o valor da mistura para o exemplo é $v = 3 \times 0,5020 + 4 \times 0,3922 + 2 \times 0,0705 + 5 \times 0,0353 = 3,3923$. Dessa forma, qualquer mistura de cores pode ser associada a valores reais de 0 a 7, onde 0 é o valor correspondente ao valor de RGB igual a (0,0,0) e, 7 ao valor RGB igual a (255,255,255).

3.3 Quantização de cores

Deriva-se do valor da mistura uma aplicação imediata na quantização de cores, (SEVERINO JR, GONZAGA; 2005) e (SEVERINO JR et al.; 2006). Como $0 \leq v \leq 7$ e a mistura pode ser obtida por 8 cores (preto, azul, verde, ciano, vermelho, magenta, amarelo e branco), valores inteiros podem ser definidos para v , tal que:

- para $0 \leq v < 1 \times \frac{7}{8}$ o valor inteiro de $v = 0$ representa a cor preta;

- para $1 \times \frac{7}{8} \leq v < 2 \times \frac{7}{8}$ o valor inteiro de $v = 1$ representa a cor azul;
- para $2 \times \frac{7}{8} \leq v < 3 \times \frac{7}{8}$ o valor inteiro de $v = 2$ representa a cor verde;
- para $3 \times \frac{7}{8} \leq v < 4 \times \frac{7}{8}$ o valor inteiro de $v = 3$ representa a cor ciano;
- para $4 \times \frac{7}{8} \leq v < 5 \times \frac{7}{8}$ o valor inteiro de $v = 4$ representa a cor vermelho;
- para $5 \times \frac{7}{8} \leq v < 6 \times \frac{7}{8}$ o valor inteiro de $v = 5$ representa a cor magenta;
- para $6 \times \frac{7}{8} \leq v < 7 \times \frac{7}{8}$ o valor inteiro de $v = 6$ representa a cor amarelo;
- para $7 \times \frac{7}{8} \leq v \leq 8 \times \frac{7}{8}$ o valor inteiro de $v = 7$ representa a cor branca.

Assim, a cor $C = (109, 146, 137)$ possui $v = 3,3923$ que equivale a $v = 3$ representando a cor ciano.

Na Figura 3.4-a é apresentada uma imagem do banco de faces AR (MARTINEZ; BENAVENTE, 1998) no formato RGB com 24 *bits* e na Figura 3.4-b, a mesma imagem, quantizada em 8 cores (preto, azul, ciano, verde, vermelho, magenta, amarelo e branco) pelo valor v da mistura de cores da imagem.

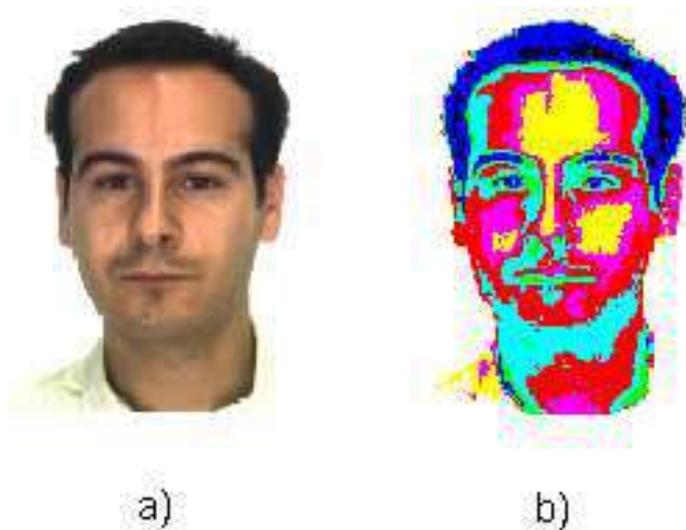


Figura 3.4: Aplicação da mistura de cores. a) imagem Original. b) imagem quantizada pelo valor da mistura de cores.

O valor da mistura pode ser utilizado como um método de quantização de cores estático gerando um histograma, definido nesse trabalho, como Misturograma, que indica a quantidade das cores preto, azul, verde, ciano, vermelho, magenta, amarelo e branco (Figura 3.5). O método de quantização é aplicado a cada *pixel* que compõe uma imagem RGB e, que é quantificado pelo valor v da mistura de cor (equação 43).

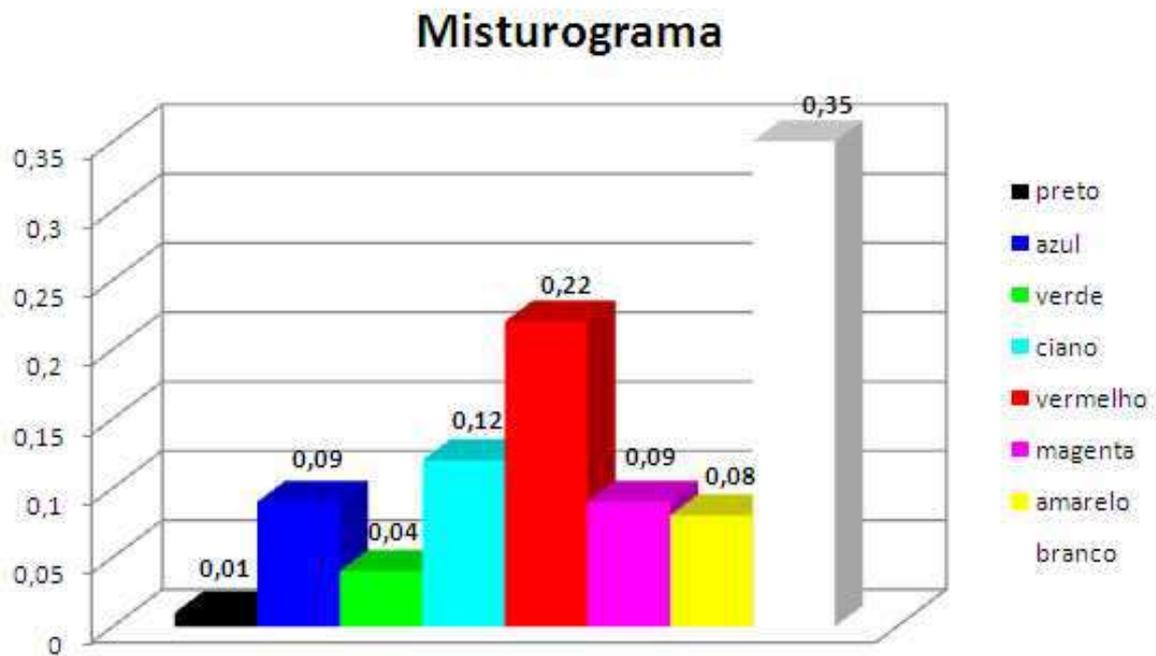


Figura 3.5: Misturograma (Histograma normalizado) das cores quantizadas pelo valor da mistura da Figura 3.4-b.

O diferencial da Mistura de Cores em relação a outros métodos de quantização de cores é possibilitar a segmentação de uma imagem pela mistura de cores, separando regiões semelhantes que tenham pouca variação de tonalidade observáveis, de outras, que diferem mais acentuadamente. Esta segmentação pode ser vista na Figura 3.4-b, onde se destacam: a pele, os olhos e o cabelo em regiões quase homogêneas. Além disso, a quantização em apenas

oito cores reduz drasticamente o tempo de processamento das imagens. Os métodos tradicionais de quantização de cores procuram manter a qualidade visual das imagens em detrimento a uma maior facilidade para o processamento visando à segmentação de partes das imagens.

3.4 Espaço de cor HSM definido pelo Valor da mistura

O valor da mistura pode ser obtido também pela da equação 45.

$$v = \frac{(4 \sum_{i=0}^7 R_i 2^i + 2 \sum_{i=0}^7 G_i 2^i + 1 \sum_{i=0}^7 B_i 2^i)}{2^8 - 1} \quad (45)$$

Ou seja

$$v = \frac{4R + 2G + 1B}{255}, \text{ para } (0 \leq R \leq 255), (0 \leq G \leq 255) \text{ e } (0 \leq B \leq 255) \quad (46)$$

Portanto, para $C = (109, 146, 137)$ tem-se que:

$$v = \frac{4 \times 109 + 2 \times 146 + 1 \times 137}{255} = 3,3923 \quad (47)$$

Isto é, o valor da mistura é a equação de um plano que intercepta o espaço RGB . As cores são números inteiros que variam de $RGB = (0, 0, 0)$ com $v = 0$ a $RGB = (255, 255, 255)$ com $v = 7$, e o valor da mistura normalizada ($0 \leq m \leq 1$) é dado pela equação 48.

$$m = \frac{v}{7} \quad (48)$$

Como o valor da mistura define planos que interceptam o espaço de cores RGB e cada plano possui um valor único para as cores nele contida, propõe-se denominar HSM (*Hue*, *Saturation* e *Mixture*) o espaço de cores do valor da mistura e utilizá-lo como ferramenta em algoritmos de processamento de imagens que levem em conta as características de cores observadas pelo sistema visual humano.

3.4.1 Conversão do RGB para HSM

Para se calcular os valores h (*hue*) e s (*saturation*) do espaço HSM deve-se analisar a forma geométrica da intersecção do plano m com o cubo RGB nos seus vértices que correspondem às cores azul, verde, ciano, vermelho, magenta e amarelo. A intersecção do plano m com os vértices correspondentes às cores “preto” e “branco” representa um ponto. Utilizam-se os valores de $(0 \leq R \leq 1)$, $(0 \leq G \leq 1)$ e $(0 \leq B \leq 1)$ para calcular h e s e, a equação 49 para calcular o valor de m .

$$m = \frac{4R + 2G + B}{7}, \text{ para } (0 \leq R \leq 1), (0 \leq G \leq 1) \text{ e } (0 \leq B \leq 1) \quad (49)$$

3.4.1.1 Vértice Azul do cubo RGB

O plano do valor da mistura definido por $m = 1/7$ intercepta o cubo RGB nos pontos $B_m = (0, 0, 1)$, $G_m = (0, 0.5, 0)$ e $R_m = (0.25, 0, 0)$ e define um triângulo (Figura 3.6). Isso ocorre para $0 \leq m \leq 1/7$.

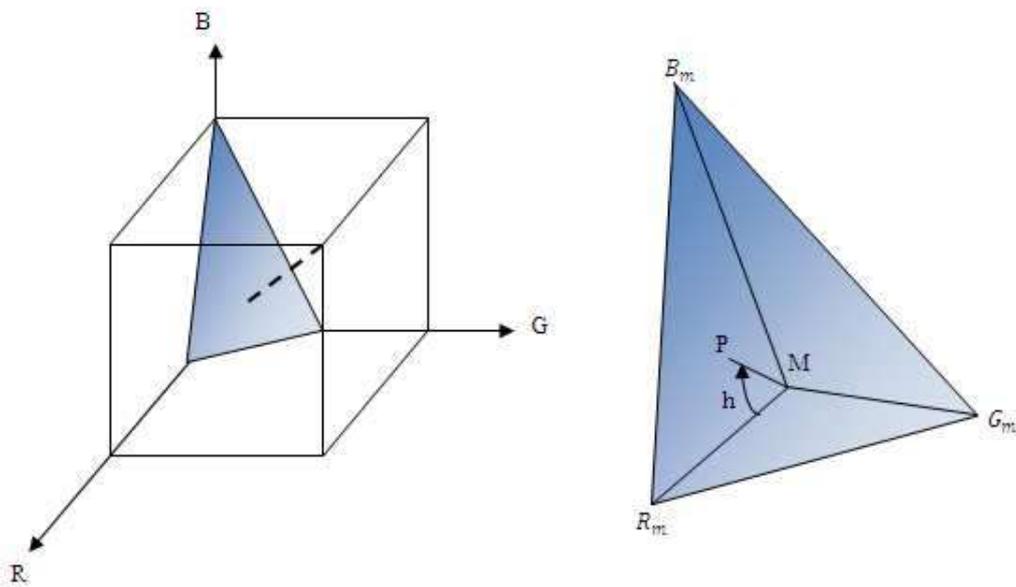


Figura 3.6: Intersecção do plano m com o vértice azul do cubo RGB.

Seja $P = (R, G, B)$ um ponto que pertence a intersecção do plano m com o cubo RGB, tem-se que:

- $M = (m, m, m)$ representa o ponto da diagonal principal do cubo onde $R = G = B = m$,
- $R_m = (7/4 m, 0, 0)$,
- $\overrightarrow{MR_m} = (3/4 m, -m, -m)$, seja $\vec{u} = (3/\sqrt{41}, -4/\sqrt{41}, -4/\sqrt{41})$ o vetor unitário com mesma direção e mesmo sentido de $\overrightarrow{MR_m}$,
- $\vec{v} = \overrightarrow{MP} = (R - m, G - m, B - m)$,

- θ é o ângulo formado pelos vetores \vec{u} e \vec{v} , tal que:

$$\cos \theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|} \quad (50)$$

Ou seja,

$$\theta = \cos^{-1} \left(\frac{\frac{3}{\sqrt{41}}(R-m) - \frac{4}{\sqrt{41}}(G-m) - \frac{4}{\sqrt{41}}(B-m)}{\sqrt{(R-m)^2 + (G-m)^2 + (B-m)^2}} \right) \quad (51)$$

Sendo que (Figura 3.7)

$$\begin{cases} \omega = \theta, & \text{se } B \leq G \\ \omega = 2\pi - \theta, & \text{se } B > G \end{cases} \quad (52)$$

Dessa forma, define-se h , como sendo ω normalizado

$$h = \frac{\omega}{2\pi} \quad (53)$$

- $B_m = (0, 0, 7m)$,
- $G_m = \left(0, \frac{7m}{2}, 0\right)$
- $\|MB_m\| > \|MG_m\|$ e $\|MB_m\| > \|MR_m\|$, S é calculado pela equação 54.

$$S = \frac{\|MP\|}{\|MB_m\|} = \frac{\sqrt{(R-m)^2 + (G-m)^2 + (B-m)^2}}{\sqrt{(0-m)^2 + (0-m)^2 + (7m-m)^2}} \quad (54)$$

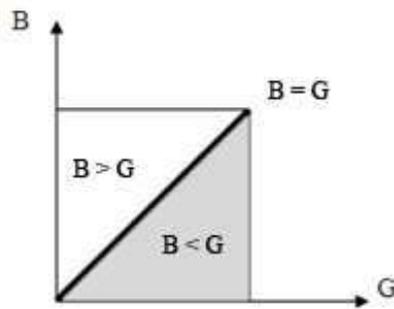


Figura 3.7: Projeção de \overrightarrow{MR}_m no plano BG.

3.4.1.2 Vértice Verde do cubo RGB

O plano do valor da mistura definido por $m = 2/7$ intercepta o cubo RGB nos pontos $G_m = (0, 1, 0)$, $BG_m = (0, 0.5, 1)$, $BR_m = (0.25, 0, 1)$ e $R_m = (0.5, 0, 0)$ e define um trapézio (Figura 3.8). Isso ocorre para $1/7 < m \leq 2/7$.

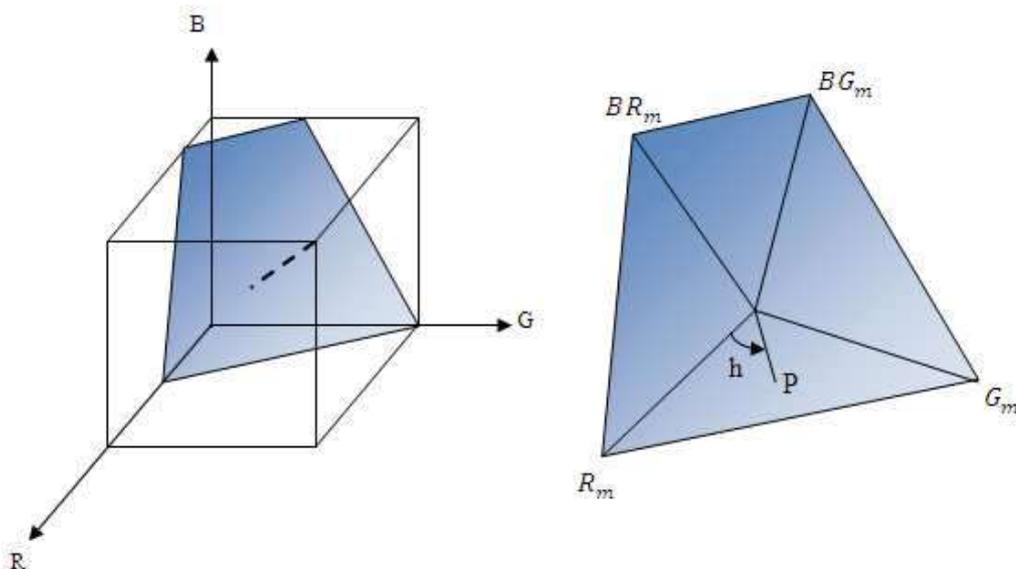


Figura 3.8: Intersecção do plano m com o vértice verde do cubo RGB.

Seja $P = (R, G, B)$ um ponto que pertence a intersecção do plano m com o cubo RGB, tem-se que:

- $M = (m, m, m)$ representa o ponto da diagonal principal do cubo onde $R = G = B = m$,
- $R_m = \left(\frac{7m}{4}, 0, 0\right)$,
- h é o ângulo formado pelos vetores \overrightarrow{MP} e $\overrightarrow{MR_m}$ (equações 51-53),
- $G_m = \left(0, \frac{7m}{2}, 0\right)$,
- $BG_m = \left(0, \frac{7m-1}{2}, 1\right)$,
- $BR_m = \left(\frac{7m-1}{4}, 0, 1\right)$,
- $\|MBG_m\| > \|MG_m\|$ e $\|MBG_m\| > \|MR_m\|$ e $\|MBG_m\| > \|MBR_m\|$, S é calculado pela equação 55.

$$S = \frac{\|MP\|}{\|MBG_m\|} = \frac{\sqrt{(R-m)^2 + (G-m)^2 + (B-m)^2}}{\sqrt{(0-m)^2 + \left(\frac{7m-1}{2} - m\right)^2 + (1-m)^2}} \quad (55)$$

3.4.1.3 Vértice Ciano do cubo RGB

O plano do valor da mistura definido por $m = 3/7$ intercepta o cubo RGB nos pontos $BG_m = (0, 1, 1)$, $BR_m = (0.5, 0, 1)$, $GR_m = (0.25, 1, 0)$ e $R_m = (0.75, 0, 0)$ e define um paralelograma (Figura 3.9). Observa-se esse fato para $(2/7 < m \leq 3/7)$.

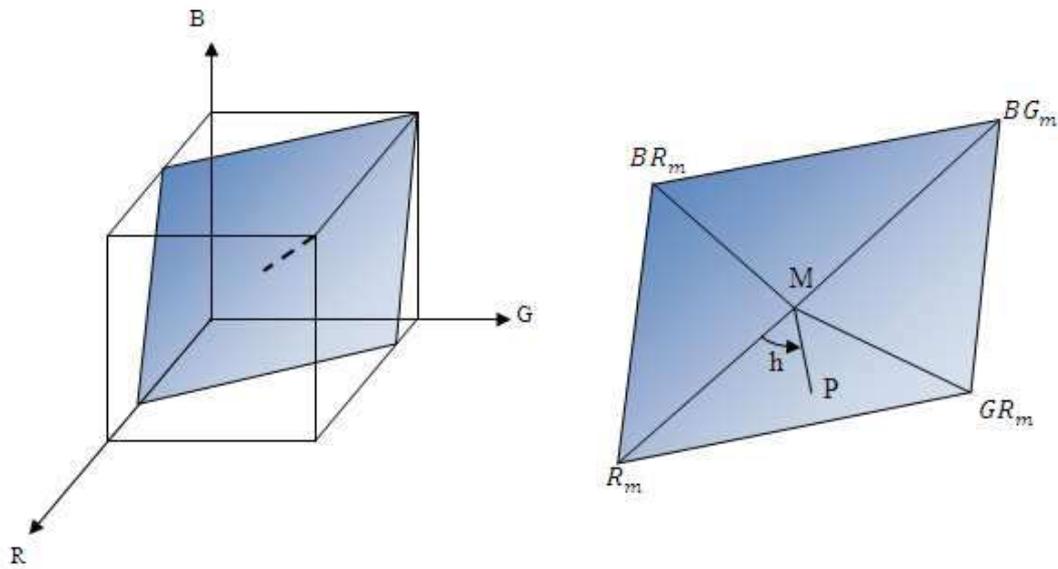


Figura 3.9: Interseção do plano m com o vértice ciano do cubo RGB.

Seja $P = (R, G, B)$ um ponto que pertence a intersecção do plano m com o cubo RGB,

tem-se que:

- M representa o ponto da diagonal principal do cubo onde $R = G = B = m$
- $R_m = \left(\frac{7m}{4}, 0, 0\right)$,
- h é o ângulo formado pelos vetores \overrightarrow{MP} e $\overrightarrow{MR_m}$ (equação 51-53),
- $BG_m = \left(0, \frac{7m-1}{2}, 1\right)$,
- $BR_m = \left(\frac{7m-1}{4}, 0, 1\right)$,
- $GR_m = \left(\frac{7m-2}{4}, 1, 0\right)$,
- $\|MBG_m\| > \|MBR_m\|$ e $\|MBG_m\| > \|MR_m\|$ e $\|MBG_m\| > \|MGR_m\|$, S é calculado pela equação 55.

3.4.1.4 Vértice Vermelho do cubo RGB

O plano do valor da mistura definido por $m = 4/7$ intercepta o cubo RGB nos pontos $BGR_m = (0.25, 1, 1)$, $BR_m = (0.75, 0, 1)$, $GR_m = (0.5, 1, 0)$ e $R_m = (1, 0, 0)$ e define um paralelograma (Figura 3.10). Verifica-se esse fato para $3/7 < m \leq 4/7$.

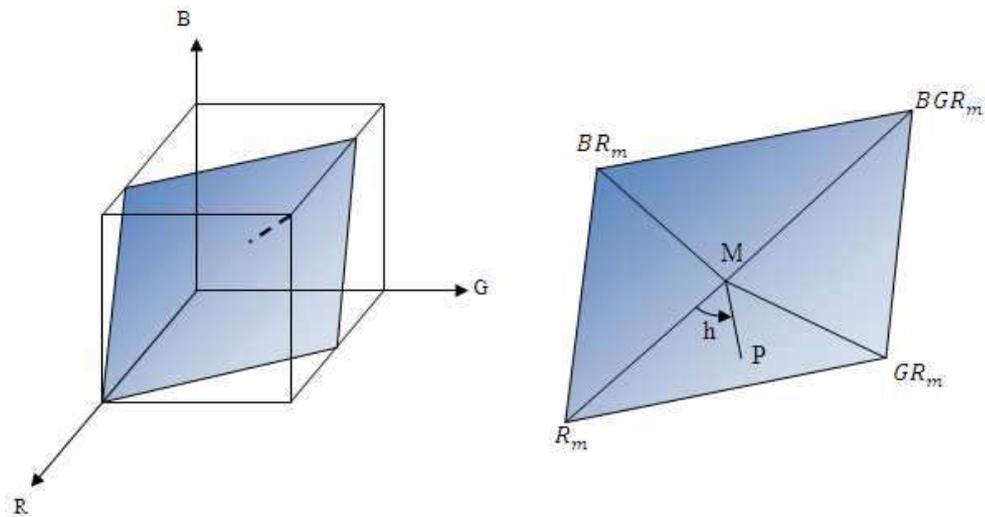


Figura 3.10: Intersecção do plano m com o vértice vermelho do cubo RGB.

Seja $P = (R, G, B)$ um ponto que pertence a intersecção do plano m com o cubo RGB, tem-se que:

- $M = (m, m, m)$ representa o ponto da diagonal principal do cubo onde $R = G = B = m$,
- $R_m = \left(\frac{7m}{4}, 0, 0\right)$,
- h é o ângulo formado pelos vetores \overrightarrow{MP} e $\overrightarrow{MR_m}$ (equação 51-53),
- $BGR_m = \left(\frac{7m-3}{4}, 1, 1\right)$,
- $BR_m = \left(\frac{7m-1}{4}, 0, 1\right)$,
- $GR_m = \left(\frac{7m-2}{4}, 1, 0\right)$,

- $\|MBGR_m\| > \|MBR_m\|$ e $\|MBGR_m\| > \|MR_m\|$ e $\|MBGR_m\| > \|MGR_m\|$, para $(3/7 < m \leq 1/2)$ então:

$$S = \frac{\|MP\|}{\|MBGR_m\|} = \frac{\sqrt{(R-m)^2 + (G-m)^2 + (B-m)^2}}{\sqrt{\left(\frac{7m-3}{2} - m\right)^2 + (1-m)^2 + (1-m)^2}} \quad (56)$$

- $\|MR_m\| > \|MBR_m\|$ e $\|MR_m\| > \|MBGR_m\|$ e $\|MR_m\| > \|MGR_m\|$, para $(1/2 < m \leq 4/7)$ então:

$$S = \frac{\|MP\|}{\|MBGR_m\|} = \frac{\sqrt{(R-m)^2 + (G-m)^2 + (B-m)^2}}{\sqrt{\left(\frac{7m}{4} - m\right)^2 + (0-m)^2 + (0-m)^2}} \quad (57)$$

3.4.1.5 Vértice Magenta do cubo RGB

O plano do valor da mistura definido por $m = 5/7$ intercepta o cubo RGB nos pontos $RB_m = (1, 0, 1)$, $BGR_m = (0.5, 1, 1)$, $RG_m = (1, 0.5, 0)$ e $GR_m = (0.75, 1, 0)$ e define um trapézio (Figura 3.11). Observa-se esse fato para $4/7 < m \leq 5/7$.

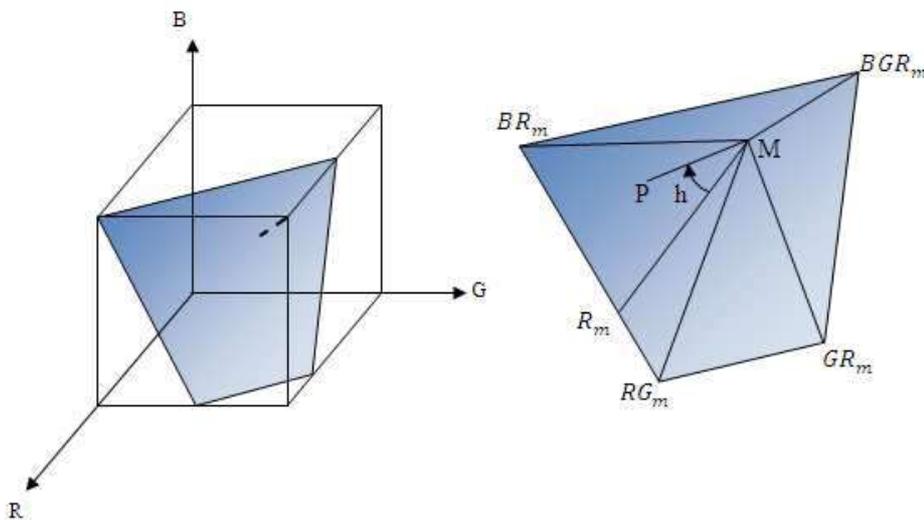


Figura 3.11: Intersecção do plano m com o vértice magenta do cubo RGB.

Seja $P = (R, G, B)$ um ponto que pertence a intersecção do plano m com o cubo RGB, tem-se que:

- $M = (m, m, m)$ representa o ponto da diagonal principal do cubo onde $R = G = B = m$,
- $R_m = \left(1, \frac{7m-4}{3}, \frac{7m-4}{3}\right)$,
- $\overrightarrow{MR_m} = (1 - m, (4m - 4)/3 - m, (4m - 4)/3 - m)$, seja
 $\vec{u} = \left(3/\sqrt{41}, -4/\sqrt{41}, -4/\sqrt{41}\right)$ o vetor unitário com mesma direção e mesmo sentido de $\overrightarrow{MR_m}$,
- h é o ângulo formado pelos vetores \overrightarrow{MP} e $\overrightarrow{MR_m}$ (equação 51-53),
- $BGR_m = \left(\frac{7m-3}{4}, 1, 1\right)$,
- $RG_m = \left(1, \frac{7m-4}{2}, 0\right)$,
- $GR_m = \left(\frac{7m-2}{4}, 1, 0\right)$
- $\|MRG_m\| > \|MBR_m\|$ e $\|MRG_m\| > \|MBGR_m\|$ e $\|MRG_m\| > \|MGR_m\|$, S é calculado pela equação 58.

$$S = \frac{\|MP\|}{\|MRG_m\|} = \frac{\sqrt{(R - m)^2 + (G - m)^2 + (B - m)^2}}{\sqrt{(1 - m)^2 + \left(\frac{7m - 4}{2} - m\right)^2 + (0 - m)^2}} \quad (58)$$

3.4.1.6 Vértice Amarelo do cubo RGB

O plano do valor da mistura definido por $m = 6/7$ intercepta o cubo RGB nos pontos $RBG_m = (1, 0.5, 1)$, $BGR_m = (0.75, 1, 1)$ e $RG_m = (1, 1, 0)$ e define um triângulo (Figura 3.12). Isso ocorre para $5/7 < m \leq 6/7$

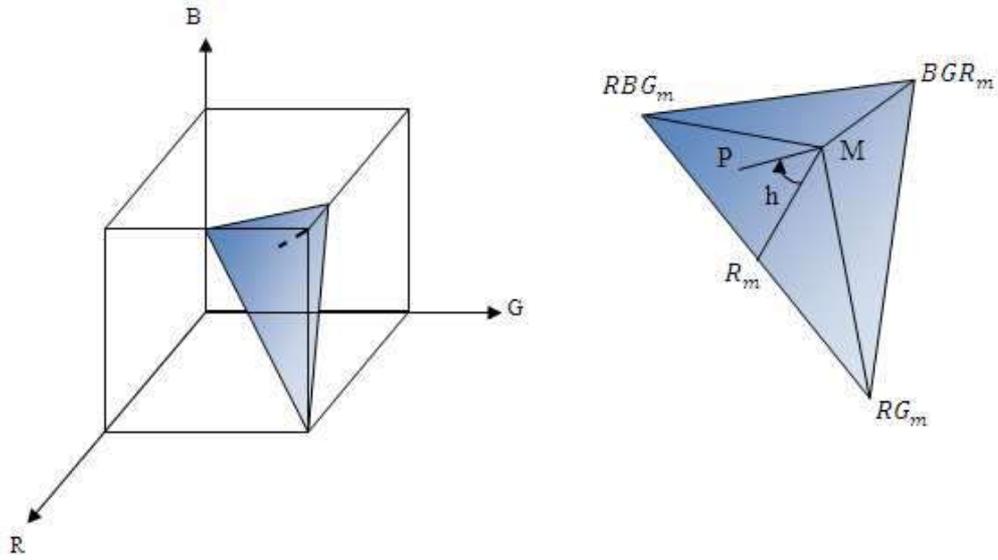


Figura 3.12: Intersecção do plano m com o vértice amarelo do cubo RGB.

Seja $P = (R, G, B)$ um ponto que pertence a intersecção do plano m com o cubo RGB, tem-se que:

- $M = (m, m, m)$ representa o ponto da diagonal principal do cubo onde $R = G = B = m$,
- $R_m = \left(1, \frac{7m-4}{3}, \frac{7m-4}{3}\right)$,
- h é o ângulo formado pelos vetores \overrightarrow{MP} e $\overrightarrow{MR_m}$ (equação 51-53),
- $BGR_m = \left(\frac{7m-3}{4}, 1, 1\right)$,
- $RBG_m = \left(1, \frac{7m-5}{2}, 1\right)$,
- $RG_m = \left(1, \frac{7m-4}{2}, 0\right)$
- $\|MRG_m\| > \|MRBG_m\|$ e $\|MRG_m\| > \|MBGR_m\|$, S é calculado pela equação 59.

$$S = \frac{\|MP\|}{\|MRG_m\|} = \frac{\sqrt{(R-m)^2 + (G-m)^2 + (B-m)^2}}{\sqrt{(1-m)^2 + \left(\frac{7m-4}{2} - m\right)^2 + (0-m)^2}} \quad (59)$$

Para os valores $6/7 < m \leq 1$, tem-se que:

- $M = (m, m, m)$ representa o ponto da diagonal principal do cubo onde $R = G = B = m$,
- $R_m = \left(1, \frac{7m-4}{3}, \frac{7m-4}{3}\right)$,
- h é o ângulo formado pelos vetores \overrightarrow{MP} e $\overrightarrow{MR_m}$ (equação 51-53),
- $BGR_m = \left(\frac{7m-3}{4}, 1, 1\right)$,
- $RBG_m = \left(1, \frac{7m-5}{2}, 1\right)$,
- $RG_m = (1, 1, 7m - 6)$
- $\|MRG_m\| > \|MRBG_m\|$ e $\|MRG_m\| > \|MBGR_m\|$, S é calculado pela equação 60.

$$S = \frac{\|MP\|}{\|MRG_m\|} = \frac{\sqrt{(R-m)^2 + (G-m)^2 + (B-m)^2}}{\sqrt{(1-m)^2 + (1-m)^2 + ((7m-6)-m)^2}} \quad (60)$$

A Figura 3.13 ilustra a forma geométrica do espaço HSM de acordo com os planos de intersecção com o espaço RGB para $m = 1/7$, $m = 2/7$, $m = 3/7$, $m = 4/7$, $m = 5/7$ e $m = 6/7$.

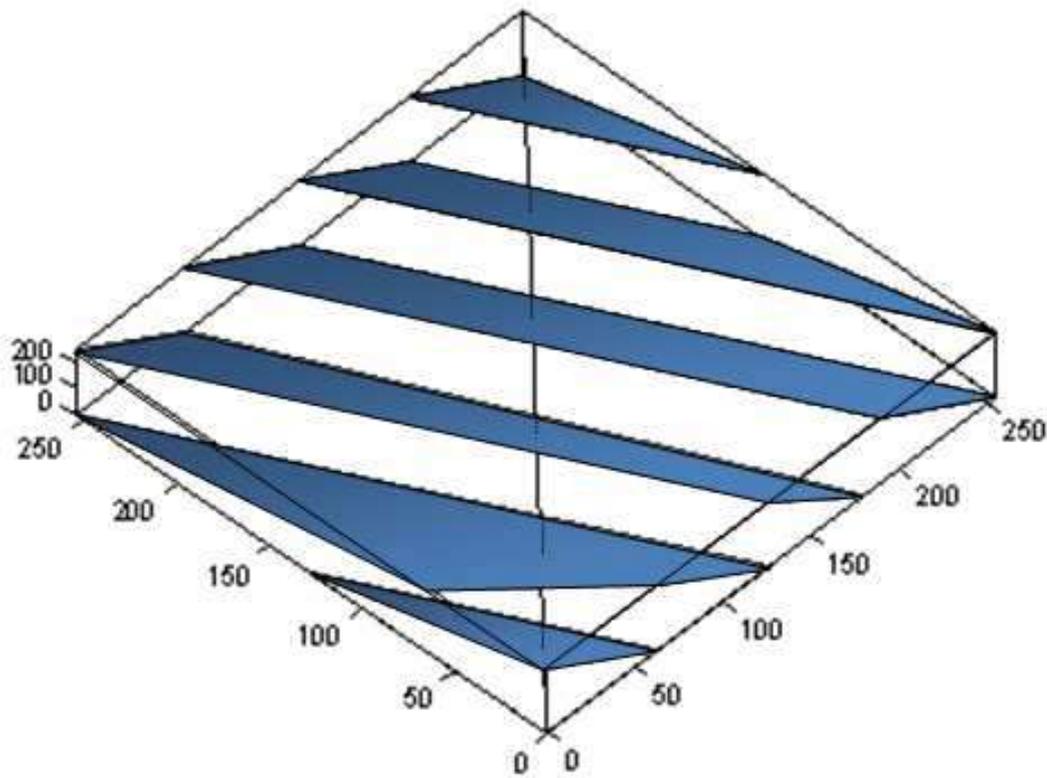


Figura 3.13: Planos de intersecção do espaço HSM nos vértices azul, verde, ciano, vermelho, magenta e amarelo do cubo RGB.

3.4.1 Conversão do HSM para RGB

A conversão de HSM para RGB utiliza os valores $0 \leq h \leq 1$, $0 \leq s \leq 1$ e $0 \leq m \leq 1$. O plano da mistura de cores que intercepta o cubo RGB é definido pelos vetores $\overline{MR_m}$ e \overline{MP} , Figura 3.14, com $M = (m, m, m)$ e $P = (R, G, B)$.

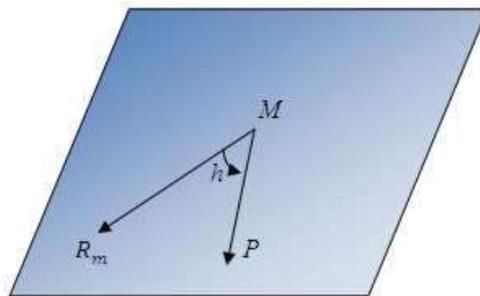


Figura 3.14: Definição do plano da mistura de cores que intercepta o cubo RGB.

Tem-se que:

- $\theta = 2\pi h$
- na conversão de RGB para HSM, verificou-se que $\vec{u} = \left(3/\sqrt{41}, -4/\sqrt{41}, -4/\sqrt{41}\right)$ é o vetor unitário com mesma direção e mesmo sentido de $\overrightarrow{MR_m}$,
- $\vec{v} = \overrightarrow{MP} = (R - m, G - m, B - m)$,
- θ é o ângulo formado pelos vetores \vec{u} e \vec{v} , tal que:

$$\cos \theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|}$$

ou seja

$$\cos \theta = \frac{\frac{3}{\sqrt{41}}(R - m) - \frac{4}{\sqrt{41}}(G - m) - \frac{4}{\sqrt{41}}(B - m)}{1s} \quad (61)$$

tal que

$$3R - 4G - 4B - \sqrt{41}s \cos \theta + 5m = 0 \quad (62)$$

- $\|\overrightarrow{MP}\| = s$, ou seja

$$(R - m)^2 + (G - m)^2 + (B - m)^2 - s^2 = 0 \quad (63)$$

- como $4R + 2G + B = 7m$, (equação 49), pode-se definir o sistema (equações 62-63):

$$\begin{cases} (R - m)^2 + (G - m)^2 + (B - m)^2 - s^2 = 0 \\ 3R - 4G - 4B - \sqrt{41}s \cos \theta + 5m = 0 \\ 4R + 2G + B - 7m = 0 \end{cases} \quad (64)$$

ao resolver o sistema definido na equação 64 em função de R , G e B , tem-se que R é calculado pela equação 65, G (equação 66) e B (equação 67).

$$R = \frac{3}{\sqrt{41}} \text{scos}(h) + m - \frac{4}{861} \sqrt{861s^2(1 - \cos^2(h))} \quad (65)$$

$$G = \frac{\sqrt{41} \text{scos}(h) + 23m - 19R}{4} \quad (66)$$

$$B = \frac{11R - 9m - \sqrt{41} \text{scos}(h)}{2} \quad (67)$$

3.5 Considerações Finais

Neste capítulo foi proposta uma nova metodologia denominada mistura de cores que é baseada na noção de cores em artes visuais. Essa nova metodologia é fundamentada na distribuição dos cones sensíveis ao comprimento de onda na faixa do “vermelho” ser maior do que a distribuição dos cones sensível ao comprimento de onda na faixa do “verde”, que por sua vez é maior do que a distribuição de cones sensível ao comprimento de onda na faixa do “azul” que foi implementada pela definição binária de cada camada que define a cor. O plano de intersecção do RGB com a mistura de cores é definido pelo valor da mistura v que define o espaço de cor HSM onde a característica essencial na modelagem da cor está na definição do matiz, da saturação e do valor da mistura. A mistura de cores pode, também, ser utilizada como um método estático de quantização de cores que permite descrever uma imagem de

maneira global e representá-la com um número reduzido de cores (preto, azul, verde, ciano, vermelho, magenta, amarelo e branco) que permite definir um histograma de cores denominado Misturograma. O algoritmo da mistura de cores como um método de quantização de cor é paralelizável, pois trabalha diretamente no nível do *bit*, todas as camadas K_i podem ser calculadas ao mesmo tempo, definindo 8 processos paralelos, um para cada $i = 0, \dots, 7$ e dessa forma pode ser implementado em hardware e, portanto constitui um recurso de grande valia na composição de vetores de características associados à descrição global de cores de uma imagem digital.

Capítulo 4

Uso da Mistura de Cores na Segmentação de Imagens de Pele Humana e o Esboço da Face

4.1 Introdução

A pele humana é constituída por diferentes pigmentos. Segundo Domenech e Bolivar (1965) levando-se em conta, como caráter fundamental, a cor da pele, pode-se definir grandes grupos, ou troncos étnicos.

Na literatura de Visão Computacional a detecção da cor da pele é frequentemente utilizada como uma fase de segmentação, em sistemas biométricos, para o reconhecimento de pessoas pela face (KAKUMANU; MAKROGIANNIS; BOURBAKIS, 2007). Verifica-se, também, que Chellapa et al. (1995), Hjelmas e Low (2001), Yang et al. (2002) e Zhao et al. (2003) acrescentam outras características como a textura.

O modelo de cor RGB tem sido usado para a modelagem e a detecção de regiões de pele humana (YANG; LU; WAIBEL, 1998) (YANG; AHUJA, 1999). Verifica-se a popularidade do seu uso nos trabalhos realizados por Kim et al. (1998), Yang et al. (1998), Wang e Sung (1999), Yang e Ahuja (1999), Bergasa et al. (2000), Schwerdt e Crowley (2000), Brown et al. (2001), Caetano e Barone (2001), Oliver et al. (2001), Soriano et al. (2003), Störring et al. (2003), Sebe et al. (2004).

Inspirado na maneira como os pintores utilizam-se de camadas de tinta para atingir a cor ideal em suas obras, foi desenvolvida a metodologia chamada mistura de cores, conforme detalhada no Capítulo 3. Esta metodologia é agora explorada quanto a sua capacidade e desempenho na segmentação da região de pele humana em imagens digitais.

Considerando-se o valor da mistura de cores v , *pixels* da região de pele são amostrados e projetados nos planos HSM e HSV.

Para a implementação do HSM utilizou-se a software MatLab versão 7 (R14) no sistema operacional Windows XP e, para a amostragem dos *pixels* da região de pele, foram utilizados os bancos de faces humanas AR (MARTINEZ; BENAVENTE, 1998) com imagens RGB com 8 *bits* por canal e resolução espacial de 576 x 768 *pixels* por imagem e a base FERET (PHILLIPS et al., 2000) com imagens RGB, também de 8 *bits* por canal e resolução espacial de 768 x 512 *pixels* por imagem.

4.3.1 Amostragem da Região de Pele no RGB, no HSM e no HSV

Brand e Mason (2000) estabeleceram a região de pele no espaço RGB pelo método de amostragem, conforme mostra a figura 4.1. Nota-se que a região é situada próxima a diagonal principal do cubo RGB que corresponde ao nível de cinza. Em nosso trabalho, a região de pele foi determinada por um banco de amostras de *pixels* de pele humana com 5264 imagens dos bancos de faces AR e FERET, que foram segmentadas, de forma manual, com a utilização do software gráfico Adobe Photoshop CS3 (versão 10.0), conforme mostra a figura 4.2.

Definido o banco de amostras, calcularam-se os valores RGB, HSM (equações 49-60) e HSV (GONZALEZ; WOODS; EDDINS, 2004) de cada *pixel* que compõe cada imagem do banco. Na figura 4.3 pode-se ver a distribuição dos pontos de pele (cor preta) amostrados.

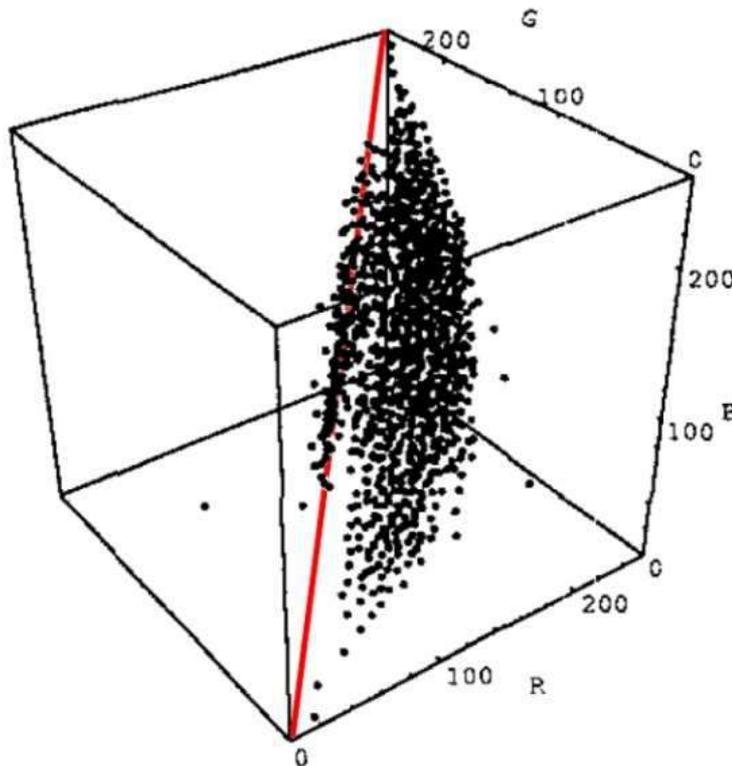


Figura 4.1: Definição da região de pele no espaço RGB. (BRAND; MASON, 2000)

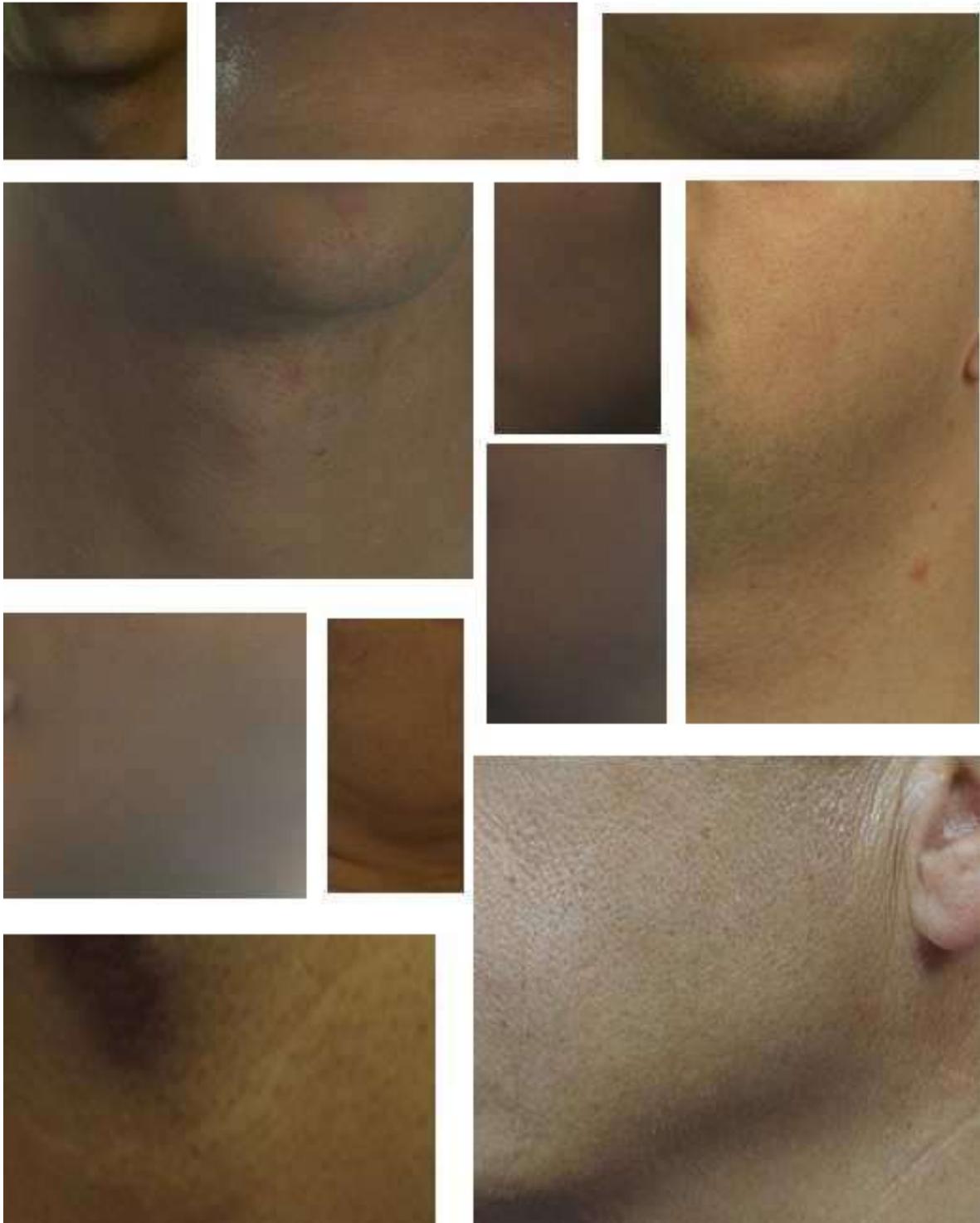


Figura 4.2: Amostras de pele segmentadas manualmente dos bancos AR e FERET.

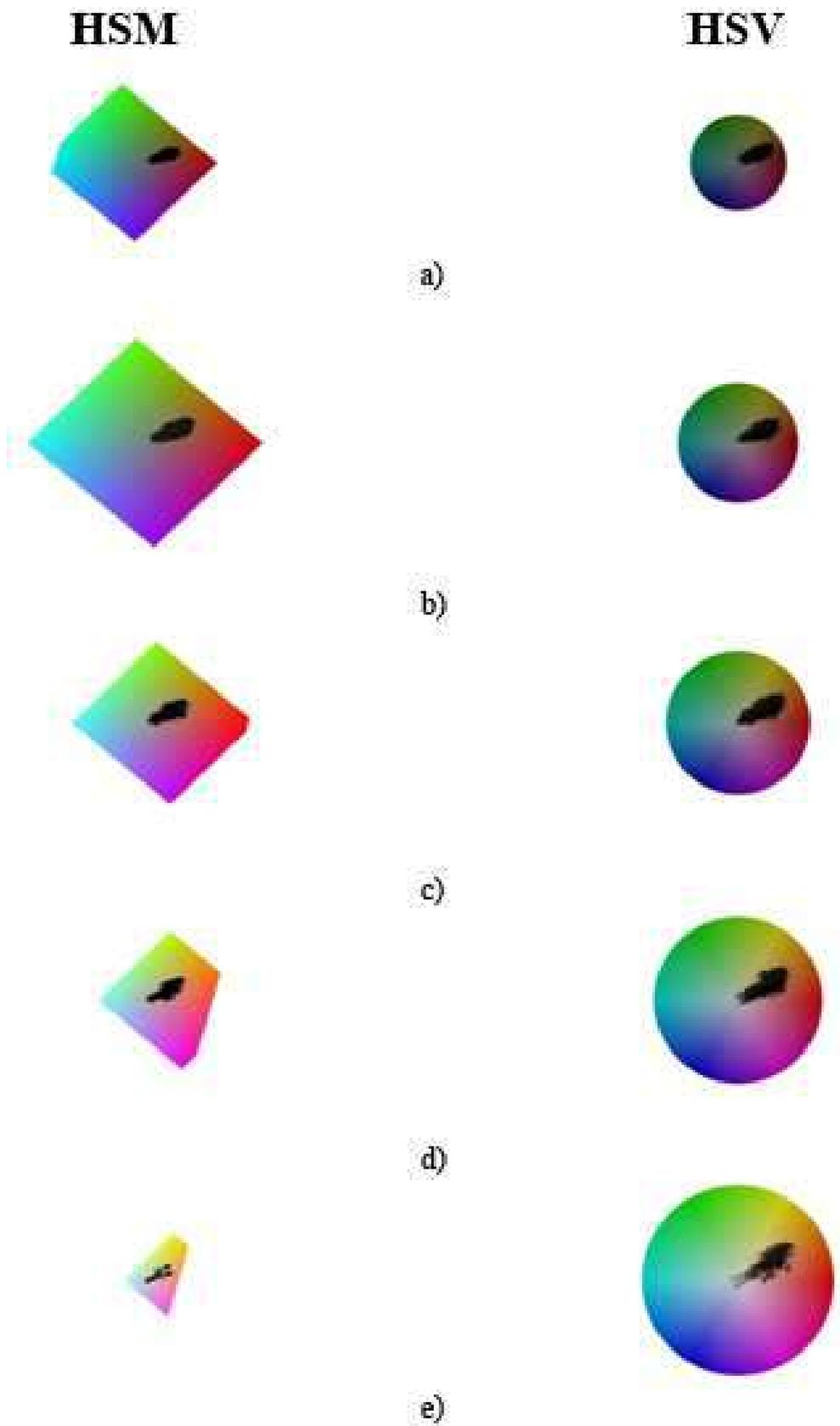


Figura 4.3: Representação da distribuição dos pontos de pele nos espaços HSM e HSV. a) $m = v = 0,39$. b) $m = v = 0,49$. c) $m = v = 0,59$. d) $m = v = 0,69$. e) $m = v = 0,79$.

4.3.2 Equacionamento da distribuição dos pontos de Pele da Face

Humana amostrados no RGB, no HSM e HSV

Peer et al. (1999) desenvolveram um classificador definindo explicitamente os valores de R, G e B no espaço RGB para pele para condições uniformes de luz (equação 68) e para condições com alta luminosidade (equação 69).

$$R > 95 \text{ e } G > 40 \text{ e } B > 20 \text{ e } \max(R, G, B) - \min(R, G, B) > 15 \text{ e} \quad (68)$$

$$|R - G| > 15 \text{ e } R > G \text{ e } R > B$$

$$R > 220 \text{ e } G > 210 \text{ e } B > 170 \text{ e } |R - G| \leq 15 \text{ e} \quad (69)$$

$$R > G \text{ e } G > B$$

A simplicidade deste classificador atraiu pesquisadores da época como, Ahlberg (1999) e Jordao et al. (1999), e ainda hoje atrai outros pesquisadores. Entretanto, a dificuldade deste método é a necessidade de adequar as regras de decisão de forma empírica para melhorar a taxa de reconhecimento. Outra característica deste método é estimar a distribuição da cor da pele a partir de um conjunto de treinamento (amostras de pele) sem derivar um modelo explícito, ou seja, constitui um método de modelagem não paramétrico da região da pele.

Ao se observar o método de modelagem não paramétrico, verificam-se duas vantagens: é rápido quanto ao uso e teoricamente independente da forma da distribuição da cor de pele no modelo utilizado. Como desvantagem não é capaz de interpolar ou generalizar o conjunto de treinamento de dados. Com o intuito de interpolar ou generalizar o conjunto de

treinamento de dados, utiliza-se o método paramétrico que é dependente da forma da distribuição da cor da pele humana na modelagem.

Menser & Wien (2000) implementaram um classificador paramétrico gaussiano para classificação de pele humana definido por:

$$p(c/pele) = \frac{1}{2\pi|\Sigma_s|^{1/2}} e^{-\frac{1}{2}(c-\mu_s)^t \Sigma_s^{-1} (c-\mu_s)} \quad (70)$$

onde, c é o vetor de cores, μ_s e Σ_s são os parâmetros da distribuição normal (vetor média e matriz de covariância calculados pelas equações 71 e 72). Esses parâmetros do modelo são estimados pelos dados de treinamento.

$$\mu_s = \frac{1}{n} \sum_{j=1}^n c_j \quad (71)$$

e

$$\Sigma_s = \frac{1}{n-1} \sum_{j=1}^n (c_j - \mu_s)(c_j - \mu_s)^t \quad (72)$$

onde, Terrillon et al. (2000) utilizaram a distância de Mahalanobis que faz uso do vetor média μ_s e da matriz de covariância Σ_s dos pontos que compõem uma distribuição para verificar se dado um ponto ele pertence ou não à distribuição, (equação 73).

$$\lambda_s(c) = (c - \mu_s)^t \Sigma_s^{-1} (c - \mu_s) \quad (73)$$

Baseado na equação 73, foram definidos nesse trabalho, dois classificadores paramétricos gaussianos para a modelagem da região de pele humana: o primeiro implementado no HSM utilizando como parâmetro a distância de Mahalanobis $\lambda_s \leq 5$ nos planos perpendiculares a m nos valores $0,03 \leq m \leq 1$ e o segundo no HSV utilizando a distância $\lambda_s \leq 4,5$ nos planos perpendiculares a v nos valores $0,03 \leq v \leq 1$. Na Figura 4.4

verificam-se a distribuição dos pontos de pele humana obtidos pela distância de Mahalanobis, apresentada na cor branca e a distribuição dos pontos de pele humana obtidos no processo de amostragem, representada pela cor preta, dos classificadores paramétricos gaussianos aplicados nos espaços HSM e HSV.

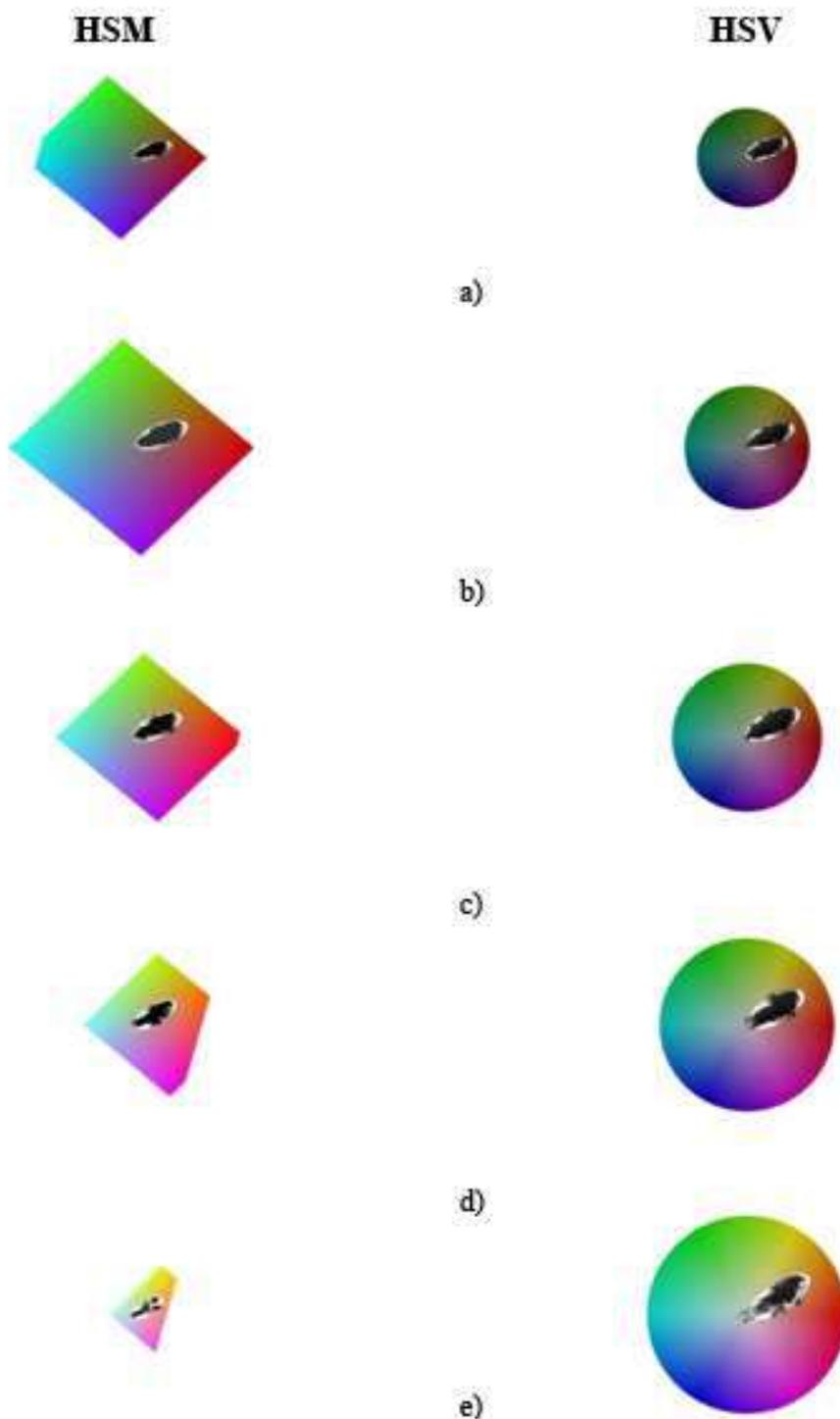


Figura 4.4: Representação da distribuição dos pontos de pele obtidos pela distância de Mahalanobis (cor branca) e por amostragem (cor preta) nos espaços HSM e HSV. a) $m = v = 0,39$. b) $m = v = 0,49$. c) $m = v = 0,59$. d) $m = v = 0,69$. e) $m = v = 0,79$.

4.3.3 Detecção do Esboço da Face da Face Humana pela Mistura de cores

No processo de quantização em oito cores (preto, azul, verde, ciano, vermelho, magenta, amarelo e branco), visto no capítulo 3, as camadas K_i da imagem equivalem às cores separadas cuja mistura gera a cor final quantificada em oito camadas e a influência de cada camada no aspecto final da imagem é proporcional a P_i , ou seja, ao peso da camada. Observando a imagem da camada K_7 na Figura 3.2 e a proposta de quantização de cores de Pass et al. (1996) e Mitra et al. (1997) que utilizaram somente os *bits* mais significativos de cada canal do espaço de cor RGB, decidiu-se, então, investigar a capacidade de segmentação da mistura de cores utilizando apenas a imagem obtida na camada mais significativa da cor, ou seja, a camada K_7 que corresponde ao *bit* mais significativo da representação binária da cor de cada *pixel* da imagem. Na Figura 4.5 é apresentado um exemplo de uma imagem obtida pela camada K_7 utilizando a mistura de cores.

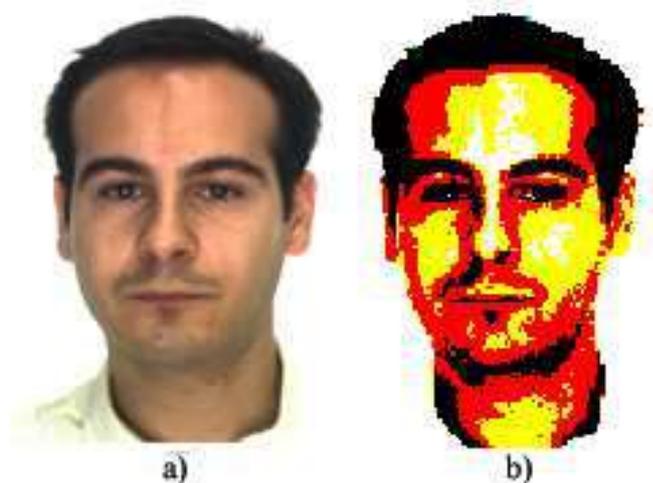


Figura 4.5: Imagem obtida pela mistura de cores utilizando-se apenas a camada K_7 . a) imagem original. b) imagem obtida pela camada K_7 .

Na Figura 4.6 pode-se notar que a região da face, correspondente ao cabelo, sobrancelhas e olhos, torna-se facilmente identificadas pela cor preta. Segmentou-se a imagem original com os *pixels* que apresentavam o valor $K_i=(0,0,0)$, que define a cor preta, na camada K_7 na mistura de cores. A imagem obtida desse processo denominou-se, neste trabalho, imagem Esboço da Face (Figura 4.6). A Figura 4.7 ilustra outros exemplos de imagens esboço da face segmentadas a partir do banco de faces AR (SEVERINO JR, GONZAGA; 2006a).

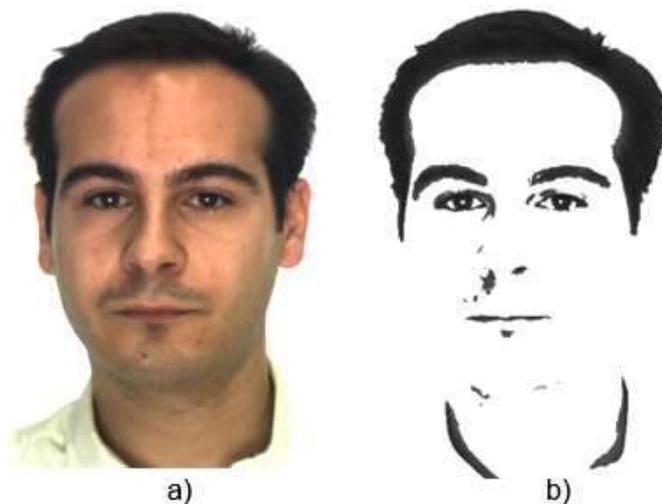


Figura 4.6: Exemplo da imagem esboço da face. a) Imagem original. b) Imagem do esboço da face.

4.4 Resultados

Para validação do método da mistura de cores na segmentação de imagens coloridas de faces humanas, utilizando-se apenas a cor como característica, foram implementados dois tipos de classificadores para a região de pele: não paramétrico (classificador RGB) e paramétrico gaussiano (nos espaços HSM e HSV). Para validar a imagem esboço da face foi implementada uma aplicação CBIR (*Content-Based Image Retrieval*) (SMEULDERS, et al.; 2000).



Figura 4.7: Exemplos de imagens esboço da face segmentadas a partir do banco AR pela mistura de cores.

4.4.1 Desempenho dos Classificadores da Região de Pele Humana

Para verificar o desempenho dos classificadores da região de pele humana foi definido um banco de teste de imagens faciais, formado por 381 imagens do banco AR e 882 imagens FERET totalizando 1263 imagens. Inicialmente, as imagens do banco de teste foram segmentadas manualmente, para criar o “ground truth”, com a utilização do software gráfico Adobe Photoshop CS3 (versão 10.0). Após, as imagens do banco de teste foram segmentadas pelo método não paramétrico de Peer (equações 68-69) no RGB, pelo método gaussiano paramétrico implementado no HSM utilizando a distância de Mahalanobis $\lambda_s \leq 5$ (equação 73) como parâmetro e pelo método gaussiano paramétrico implementado no HSV utilizando, também, a distância de Mahalanobis $\lambda_s \leq 4,5$ (equação 73) como parâmetro. As imagens segmentadas pelos classificadores foram comparadas com as imagens obtidas de forma manual de acordo com os critérios definidos na tabela 4 abordados por Congalton e Green (1993).

Tabela 4: Resultados obtidos pelos classificadores paramétricos e não paramétricos

Classificador	Taxa de acerto da classe pele	Taxa de acerto da classe não pele	Taxa de Falsa Rejeição (false negative)	Taxa de Falsa Aceitação (false positive)
paramétrico no HSM	98,1%	93,5%	1,9%	6,5%
paramétrico no HSV	97,6%	65,1%	2,4%	34,9%
não paramétrico no RGB	70,7%	91,7%	29,3%	8,3%

Na Tabela 4 pode ser visto o desempenho dos classificadores quanto a suas exatidões globais e o índice de concordância Kappa (BISHOP; et. al., 1975), enquanto na Figura 4.8 são ilustrados exemplos das imagens segmentadas por três classificadores.

Tabela 5: Índices de desempenho dos classificadores paramétricos e não paramétricos.

Classificador	Taxa de acerto da classe pele	Taxa de acerto da classe não pele	Exatidão Global	Índice concordância Kappa
paramétrico no HSM	98,1%	93,5%	94,8%	0,87
paramétrico no HSV	97,6%	65,1%	74,1%	0,49
não paramétrico no RGB	70,7%	91,7%	86,2%	0,64

Analisando-se os dados obtidos e apresentados nas tabelas 4 e 5, considerando-se a aplicação que utiliza a metodologia de mistura de cores na classificação de *pixels* de uma imagem digital como “pele humana” ou “não pele”, observa-se que:

- O desempenho da classificação dos pixels de pele humana dos classificadores paramétricos gaussianos no HSM e no HSV são superiores a 27,4 e 26,9 respectivamente, em relação ao classificador não paramétrico no RGB, obtendo a taxa de falsa rejeição de 1,9% e 2,4% para os classificadores paramétricos gaussianos no HSM e no HSV, respectivamente, e a taxa de 29,3% para o classificador não paramétrico no RGB;
- O desempenho da classificação, dos pixels de não pele humana dos classificadores paramétrico gaussiano no HSM e não paramétrico no RGB é superior a 28,4 e 26,6 respectivamente, em relação ao classificador paramétrico gaussiano no HSV, refletindo na taxa de falsa aceitação de 6,5% e 8,3% para os classificadores paramétrico gaussiano no HSM e não paramétrico no RGB, respectivamente, e na taxa de 34,9% para o classificador paramétrico gaussiano no HSV;
- O índice de concordância Kappa varia de 0 a 1, sendo que 0 representa não haver concordância, e 1 representa a concordância perfeita (tabela 6). Nesse trabalho a concordância é estabelecida entre a classificação da cor de *pixel* como “pele” e “não pele”. A exatidão global dos classificadores paramétrico gaussiano no HSM e não paramétrico

no RGB, é superior a 20,7 e 12,1 respectivamente, em relação ao classificador paramétrico gaussiano no HSV, obtendo um índice de concordância Kappa de 0,87 e 0,64 para os classificadores paramétrico gaussiano no HSM e não paramétrico no RGB, respectivamente, e índice de 0,49 para o classificador paramétrico gaussiano no HSV.

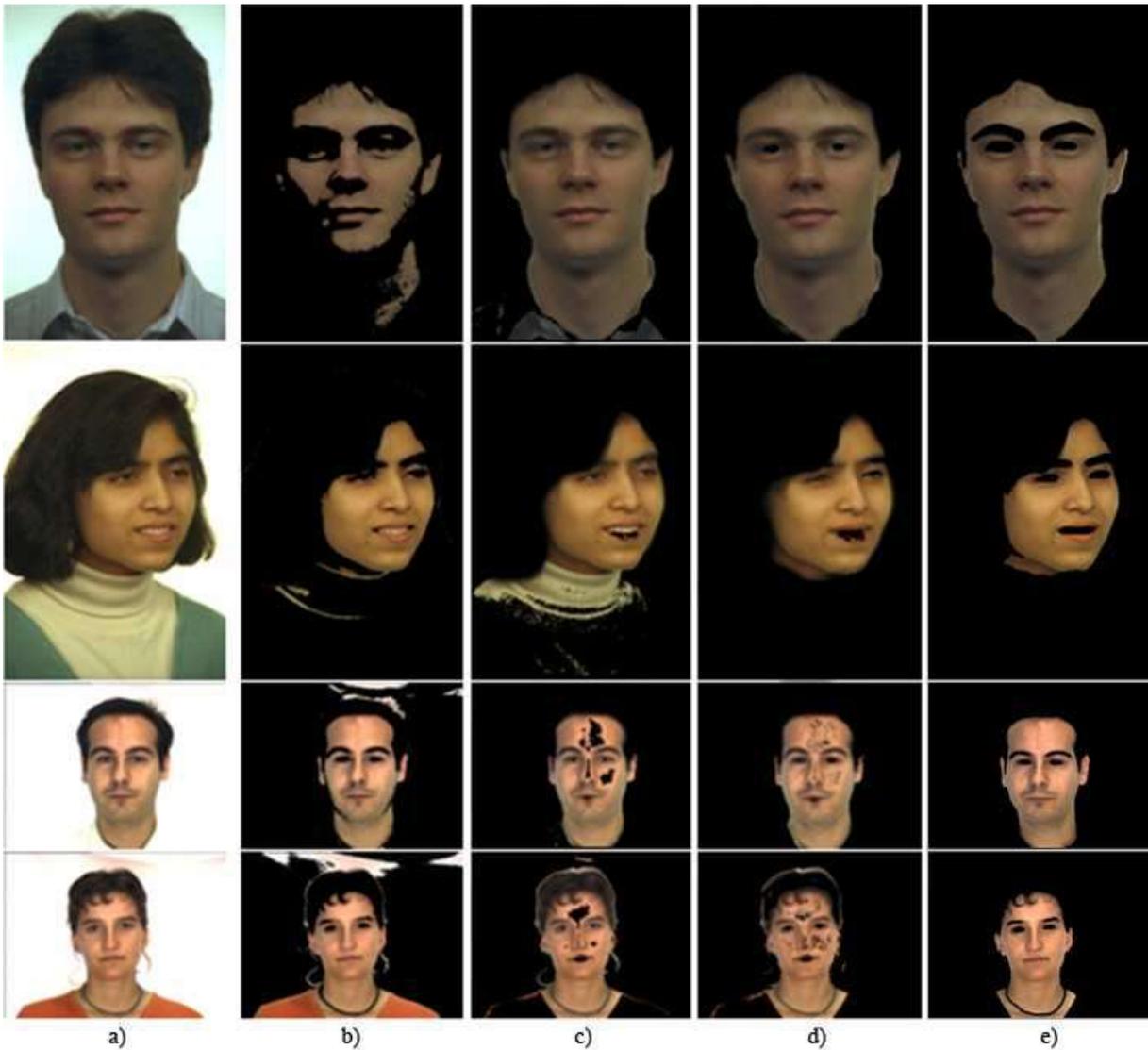


Figura 4.8: Segmentação da região de pele humana. a) Imagem original. b) Imagem segmentada pelo classificador não paramétrico de Peer no RGB. c) Imagem segmentada pelo classificador paramétrico gaussiano no HSV. d) Imagem segmentada pelo classificador paramétrico gaussiano no HSM. e) Imagem segmentada manualmente para criar o “ground truth”.

Tabela 6: Interpretação do índice de concordância Kappa (BISHOP; GREEN, 1993)

Índice de concordância kappa	Concordância
0	Pobre
0 – 0,20	Ligeira
0,21 – 0,40	Considerável
0,41 – 0,60	Moderada
0,61 – 0,80	Substancial
0,81 – 1	Excelente

4.4.2 Desempenho das Imagens Esboço da Face

Com a finalidade de validar o desempenho das imagens Esboço da Face, criou-se um banco das imagens esboço da face obtida pelos *pixels* com o valor $K_i=(0,0,0)$ na camada K_7 da mistura de cores (equação 37). O banco de imagens foi composto por três imagens referentes às diferenças quanto à iluminação e as quatro imagens referentes às diferenças quanto à expressão facial para cada uma das 120 classes do banco AR (65 classes de homens identificados pela letra inicial *m* e 55 de mulheres identificadas pela letra inicial *w*), totalizando 840 imagens (SEVERINO JR, GONZAGA; 2006b).

As classes do banco AR obedecem à classificação (Figura 4.9):

- uma imagem da pessoa, com a identificação -1;
- três imagens da pessoa com diferentes expressões faciais, com as identificações -2,-3 e -4;
- três imagens da pessoa com diferentes focos de iluminações (do lado direito da face, do lado esquerdo da face e em toda face), com as identificações -5,-6 e -7.

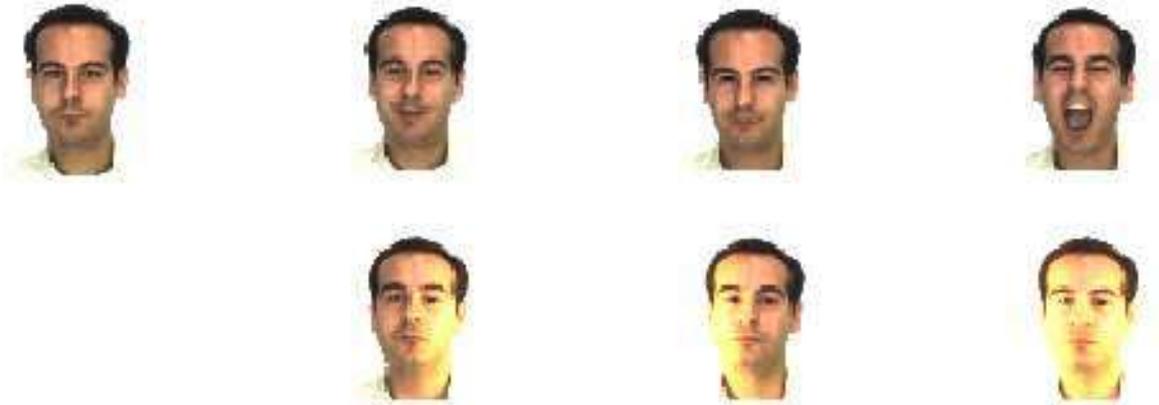


Figura 4.9: As imagens m-001-1, m-001-2, m-001-3, m-001-4, m-001-5, m-001-6 e m-001-7 referentes a classe masculina (m) 001 do banco AR.

Implementaram-se duas aplicações CBIR, a primeira com o banco do esboço da face e a segunda e das imagens originais, utilizando a técnica PCA (*Principal Components Analysis*) (BAEK; BARTLETT; DRAPER, 2003) para compor o vetor característica da descrição global das imagens.

Para a definição da matriz de projeção do subespaço PCA, cada imagem de cada banco com tamanho de 576×768 pixels no formato “bmp”, aqui representadas por a^i para $i = 1, \dots, 840$, foi convertida em nível de cinza e, posteriormente, reduzida para 32×36 pixels e transformadas em um vetor 1152 (32×36) valores (1ª etapa, Figura 4.10). Em seguida, todos os vetores foram agrupados formando uma matriz com dimensão 1152×840 e, então, foi calculada a matriz de covariância c (2ª etapa, Figura 4.10). Após, foram calculadas as matrizes de autovetores λ e autovalores v . Formou-se a matriz do subespaço PCA p com as colunas da matriz v que correspondem às colunas dos cem maiores autovetores da matriz λ (3ª etapa, Figura 4.10).

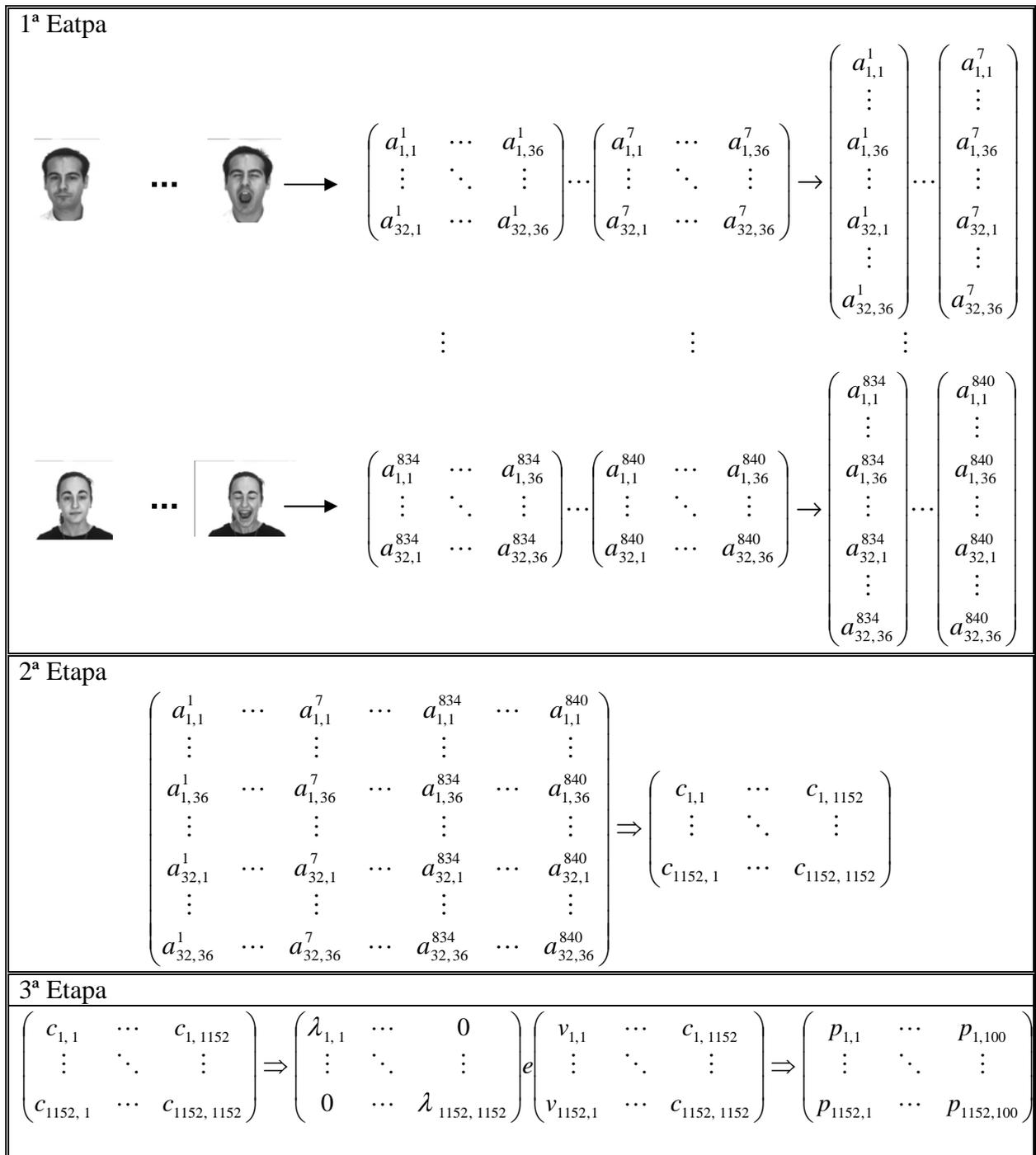


Figura 4.10: Etapas para a construção da matriz de projeção do subespaço PCA.

Para obter o vetor característica vc^i de cada imagem de cada aplicação CBIR, multiplicou-se a matriz de projeção transposta do subespaço PCA p associado ao banco da aplicação pela transformação da imagem a^i do banco em um vetor (Figura 4.11).

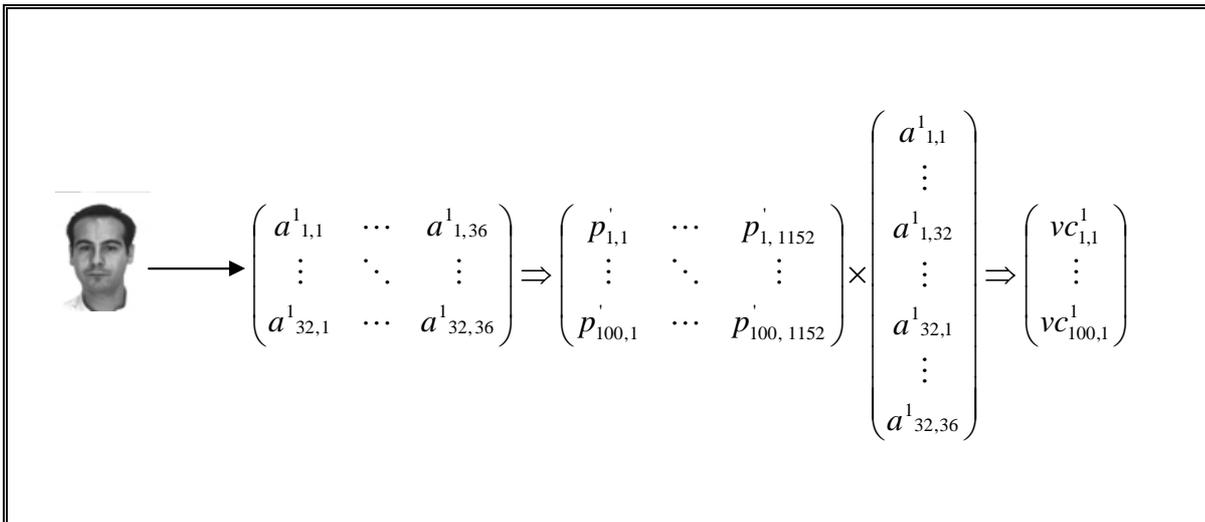


Figura 4.11: Obtenção do vetor característica da imagem m-001-1 do banco de imagens originais.

A Figura 4.12 apresenta as curvas *RP* (*Recall x Precision*) (BAEZA-YATES; RIBEIRO NETO, 1999), dados descritos na tabela 7, de cada aplicação CBIR utilizando acesso seqüencial e a distância euclidiana como critério de similaridade. As curvas comprovam o desempenho da metodologia da mistura de cores quanto à recuperação das imagens esboço da face em relação às imagens originais envolvendo diferenças de iluminação e expressão facial. O desempenho do sistema CBIR composto pelas imagens esboço da face em relação ao sistema CBIR composto pelas imagens originais foi superior:

- 0,0123 para a busca igual a três;
- 0,0259 para a busca igual a quatro,
- 0,0488 para a busca igual a cinco,
- 0,0726 para busca igual a seis e
- 0,0877 para a busca igual a sete.

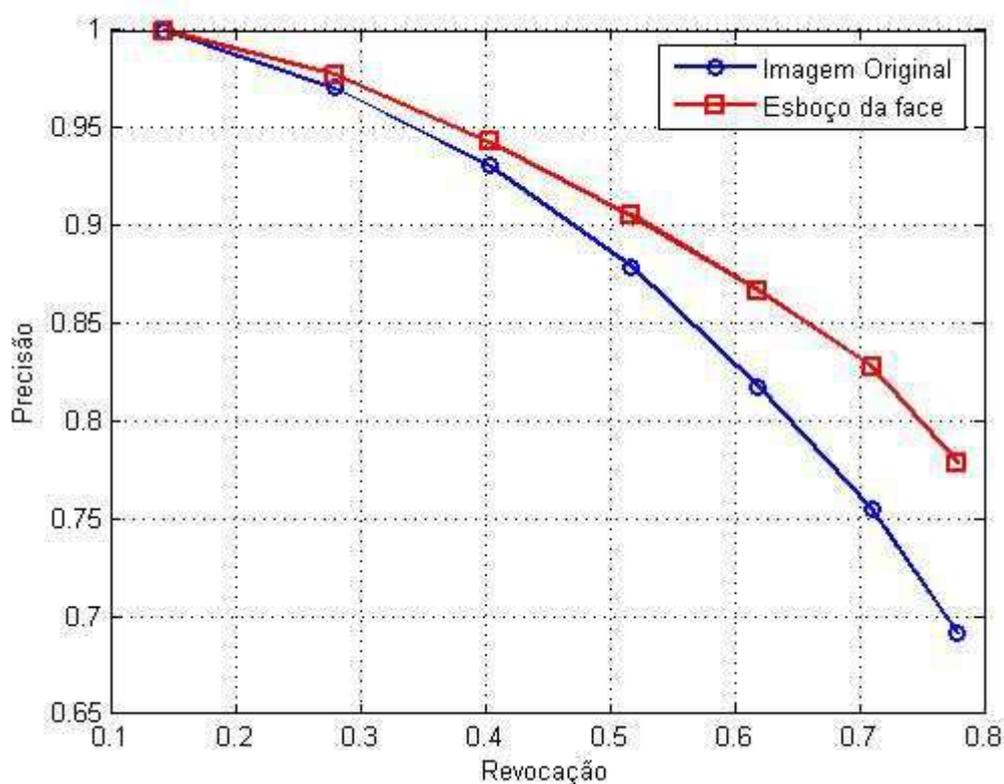


Figura 4.12: Gráfico de desempenho do banco com 840 imagens apresentando diferentes expressões faciais e mudanças de iluminação.

Tabela 7: Dados obtidos nos sistemas CBIR imagem original e imagem esboço da face.

Busca	Revocação	Precisão (imagem original)	Precisão (imagem esboço da face)
primeira mais próxima	0,1429	1,0	1,0
duas mais próximas	0,2793	0,9708	0,9774
três mais próximas	0,4041	0,9306	0,9429
quatro mais próximas	0,5172	0,8792	0,9051
cinco mais próximas	0,6189	0,8176	0,8664
seis mais próximas	0,7097	0,7554	0,8280
sete mais próximas	0,7787	0,6910	0,7787

4.5 Considerações Finais

Neste capítulo foi apresentada a modelagem de três métodos para segmentação da região de pele humana. O primeiro, modelado no espaço RGB, foi definido de forma não paramétrica que apresenta como vantagens ser rápido quanto ao uso e independente da forma da distribuição da cor de pele no modelo utilizado; entretanto, não é capaz de interpolar ou generalizar o conjunto de treinamento de dados. O segundo, desenvolvido no espaço HSM, faz uso de uma distribuição gaussiana e da distância de Mahalanobis. O terceiro, desenvolvido no espaço HSV, também, faz uso de uma distribuição gaussiana e da distância de Mahalanobis. Foi mostrada, também, a obtenção da imagem do esboço da face obtida pelos *pixels* com o valor $K_i=(0,0,0)$ na camada K_7 da mistura de cores. Demonstrou-se o potencial da mistura de cores aplicada na classificação de um *pixel* como “pele humana” e “não pele humana” em imagens faciais digitais e verificou-se que a imagem do esboço da face constitui uma nova ferramenta na descrição global de uma imagem digital e na constituição de vetores características.

Capítulo 5

Conclusões

Este trabalho propõe uma nova abordagem para o processamento de imagens coloridas. A maneira como os pintores executam suas obras inspiram a proposta da mistura de cores. Considerando que a retina humana para a percepção das cores possui uma distribuição maior de cones sensíveis a cor vermelha, uma distribuição menor para a cor verde e possuindo uma pequena distribuição para a cor azul, a mistura de cores foi desenvolvida utilizando os *bits*, do mais significativo ao menos significativo dos três canais de cores do espaço de cor RGB, como formantes de oito camadas de cores, que ao serem misturados em proporções que dependem da posição do *bit*, geram a cor final de cada *pixel*.

Com esta nova abordagem foi possível demonstrar que a mistura de cores pode ser aplicada no processamento de imagens coloridas. A aplicação imediata, devido ao aspecto da geração por camadas, é na quantização das cores de uma imagem. A quantização permite descrever uma imagem de maneira global representando-a com um número de oito cores (preto, azul, verde, ciano, vermelho, magenta, amarelo e branco).

Baseando-se na definição do matiz e da saturação da mistura de cores foi proposta a modelagem de um novo espaço de cor denominado HSM que apresenta como característica a percepção de cor pelo ser humano.

Na aplicação da segmentação da região de pele de imagens faciais os resultados obtidos comprovam o ganho de desempenho do uso do modelo de cor proposto HSM. Este desempenho foi comprovado pela implementação de um método paramétrico gaussiano no

HSM utilizando a distância de Mahalanobis como parâmetro, quando comparado ao mesmo método implementado no HSV e a outro método não paramétrico implementado no RGB.

Com o uso pelos *pixels* com o valor $K_i=(0,0,0)$ na camada K_7 da mistura de cores foi proposta a obtenção da imagem denominada de esboço da face. A utilização da imagem esboço da face em aplicações CBIR demonstrou o ganho de desempenho na recuperação de imagens. Dessa forma, o uso da imagem esboço da face como característica apresenta-se como uma ferramenta de grande valia em aplicações CBIR.

5.1 Contribuições

A principal contribuição deste trabalho é a proposta de uma nova abordagem para o processamento de cor baseada na execução de uma pintura artística e na percepção de cores pela retina humana.

Essa nova abordagem de cor contribui para:

- a proposta de um novo método de quantização de cores descrevendo uma imagem digital de forma global em apenas oito cores e
- a definição do novo espaço de cor denominado HSM, baseado na percepção de cor pela retina humana,
- a proposta de um novo método de segmentação da região de pele em imagens faciais implementado por um método paramétrico gaussiano utilizando a distância de Mahalanobis como parâmetro e
- a obtenção da imagem esboço da face como uma nova característica a ser abordada em aplicações CBIR.

5.2 Trabalhos Futuros

Os direcionamentos futuros relacionados a este trabalho envolvem:

- o projeto e o desenvolvimento de um hardware dedicado à captura da mistura de cores dos *pixels* que compõem uma imagem digital, justificado pelo fato da mistura de cores basear-se na definição binária da cor;
- a exploração do espaço de cor HSM quanto ao potencial na segmentação de outros tipos de imagens (paisagens, mata, etc);
- a obtenção do desempenho do espaço HSM quando for aplicado à outras técnicas de processamento de imagens digitais coloridas e
- a elaboração de um estudo comparativo sobre o desempenho das imagens esboço da face em outras aplicações CBIR envolvendo imagens faciais.

Referências Bibliográficas

AHLBERG, J. A system for face localization and facial feature extraction. **Tech. Rep. LiTH-ISY-R-2172**, Linkoping University, 1999.

BAEK, K.; BARTLETT, M. S.; DRAPER, B. A. Recognizing faces with PCA and ICA. **Computer Vision and Image Understand**, v. 91, p. 115-137, 2003.

BAEZA-YATES, R.; RIBEIRO NETO, B. **Modern Information Retrieval**. New York, ACM Press, 1999.

BERGASA, L. M.; MAZO, M.; GARDEL, A.; SOTELO, M. A.; BOQUETE, L. Unsupervised and adaptive Gaussian skin-color model. **Image Vision Computer**, v. 18, n. 12, p. 987–1003, 2000.

BISHOP, Y.; FIENBERG, S.; HOLLAND, P. **Discrete multivariate analysis: theory and practice**. Cambridge, Estados Unidos: MIT, 1975.

BRAINARD, D. H.; CALDERONE, J. B.; NUGENT, A. K.; JACOBS, G. H. Flicker ERG responses to stimuli parametrically modulated in color space. **Investigative Ophthalmology & Visual Science**, n. 40, 1999, p. 2840 –2847.

BRAND, J.; MASON, J. S. A comparative assessment of three approaches to pixel-level Human Skin-Detection. In: **PROCEEDINGS OF IEEE**, p. 1056-1059, 2000.

BROWN, D.; CRAW, I.; LEWTHWAITE, J. A SOM based approach to skin detection with application in real time systems. In: **BRITISH MACHINE VISION CONFERENCE (BMVC01)**, 2001.

CAETANO, T. S.; BARONE, D. A. C. A probabilistic model for the human skin-color. In: **PROCEEDINGS OF 11TH INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING (ICIAP01)**, p. 279–283, 2001.

CARROLL, J.; NEITZ M.; NEITZ, J. Estimates of L:M cone ratio from ERG flicker photometry and genetics. **Journal Vision**, n. 2, p. 531–542, 2002.

CHELLAPPA, R.; WILSON, C.; SIROHEY, S. Human and machine recognition of faces: a survey. In: **PROCEEDINGS OF IEEE**, v. 8, n. 5. p. 705–740, 1995.

CIOCCA, G.; SCHETTINI, R. A Relevance Feedback Mechanism for Content-Based Image Retrieval. In: **INFORMATION PROCESSING AND MANAGEMENT**, v. 35, p. 605-632, 1999.

CICERONE, C. M.; NERGER, J. L. The relative numbers of long-wavelength-sensitive to middle-wavelength-sensitive cones in the human fovea centralis. **Vision Reserach**, n. 29, 1989, p.115–128.

CONGALTON, R. G.; GREEN, K. A practical look at sources of confusion in error matrix generation. **Photogrammetric Engineering and Remote Sensing**, v. 59, n. 5. p. 641-644, 1993.

COX, I. J., et al. PicHunter: Bayesian Relevance Feedback for Image Retrieval. In: **PROCEEDINGS OF 13TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR'96)**, p. 361-369, 1996.

DEEB, S. S.; DILLER, L. C.; WILLIAMS, D. R.; DACEY, D. M. Interindividual and topographical variation of L:M cone ratios in monkey retinas. In: **JOURNAL OF THE OPTICAL SOCIETY OF AMERICA A: OPTICS, IMAGE SCIENCE, AND VISION**, n. 17, p. 538 –544, 2000.

DeVRIES, H. L. Luminosity curve of trichromats. **Nature**, n. 157, p.736 –737, 1946.

DOMENECH, J. M. T., BOLIVAR, A. P. **Atlas das Raças Humanas**. Rio de Janeiro, Ibero-Americano Ltda, 1965.

FIELD, G. G. **Color and Its Reproduction**. Graphics Arts Technical Foundation: 1988.

FUN, C. M.; WONG, C. F. Image retrieval using a novel color quantization approach. In: **9TH INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING (ICSP 2008)**, p. 773-776, 2008.

GAGLIARDI, I.; SCHETTINI, R. A Method for the Automatic Indexing of Color Images for Effective Image Retrieval.. **The New Review of Hypermedia and Multimedia**. v. 3, p. 201-224, 1997.

GOMES, J.; VELHO, L. **Computação Gráfica**, Volume 1. IMPA, 1998.

GONG, Y.; CHUAN, C. H.; XIAOYI, G. Image Indexing and Retrieval using Color Histograms. **Multimedia Tools and Applications**, v. 2, p. 133-156, 1996.

GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. L. **Digital Image Processing using Matlab**. Pearson Education, 2004.

GONZALES, T. F. Clustering to Minimize the Maximum Intercluster Distance. **Theoretical Computer Science**, v. 38, p. 293-306, 1985.

GUNTHER, K. L.; DOBKINS, K. R. Individual differences in chromatic (red/green) contrast sensitivity are constrained by the relative number of L- versus M-cones in the eye, **Vision Research**, v. 42, n. 11, May 2002, pp. 1367-1378, 2002.

HAFNER, J., et al. Efficient Color Histogram Indexing for Quadratic from Distance Functions. **IEEE Transactions Pattern Analysis and Machine Intelligence**, v. 17, p. 729-736, 1995.

HAGSTROM, S. A.; NEITZ, J.; NEITZ, M. Variations in cone populations for red-green color vision examined by analysis of mRNA. *Neuro Report*, n. 9, p. 1963–1967, 1998.

HARALICK, R. M.; SHAPIRO, L. G. Image Segmentation Techniques. **Computer Vision Graphics and Image Processing**, v. 29, p. 100-132, 1985.

HJELMAS, E.; LOW, B. K. Face detection: a survey. *J. Computer Vision Image Understanding*, v. 83, p. 236–274, 2001.

HOFER, H.; CARROLL, J.; NEITZ, J.; NEITZ, M.; WILLIAMS, D. R. Organization of the Human Trichromatic Cone Mosaic. In: *THE JOURNAL OF NEUROSCIENCE*, October 19, v. 25, n. 42, p. 9669 –9679, 2005.

HUNT, R. W. G. **Measuring Color**. Ellis Horwood : 1991.

ITO, N., et al. Fuzzy Logic Based Non-Parametric Color Image Segmentation with Optional Block Processing. **Association of Computing Machinery**, p. 119-126, 1995.

JAYNES, E. T. On the Rationale of Maximum-Entropy Methods. In: PROCEEDINGS OF IEEE, v. 70. p. 939-952, 1982.

JORDAO, L.; PERRONE, M.; COSTEIRA, J.; SANTOS-VICTOR, J. Active face and feature tracking. In PROCEEDINGS OF THE 10TH INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING, p. 572–577, 1999.

KAKUMANU, P.; MAKROGIANNIS, S.; BOURBAKIS, N. A survey of skin-color modeling and detection methods. **Pattern Recognition**, v. 40, p. 1106 –1122, 2007.

KIM, S. H.; KIM, N. K.; AHN, S. C.; KIM, H. G. Object oriented face detection using range and color information. In: PROCEEDINGS OF THE 3RD INTERNATIONAL CONFERENCE ON AUTOMATIC FACE AND GESTURE RECOGNITION (AFGR 98), 1998.

KIM, N.; KEHTARNAVAZ, N. DWT-based scene-adaptive color quantization. **Real-Time Imaging**, v. 11, p. 443-453, 2005.

KREMERS, J.; SCHOLL, H. P. N.; KNAU, H.; BERENDSCHOT, T. T. J. M.; USUI, T.; SHARPE, L. T. L/M cone ratios in human trichromats assessed by psychophysics, electroretinography, and retinal densitometry, In: JOURNAL OF THE OPTICAL SOCIETY OF AMERICA A: OPTICS, IMAGE SCIENCE, AND VISION, n. 17, p. 517–526, 2000.

LIAPIS, S.; TZIRITAS, G. Color and Texture Image Retrieval using Chromaticity Histograms and Wavelet Frames. **IEEE Transactions on Multimedia**, v. 6, n. 5, p. 676-686, 2004.

LIU, J.; YANG, Y. H. Multiresolution Color Image Segmentation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 16, p. 689-700, 1994.

MARTINEZ, A. M.; BENAVENTE, R. The AR Face Database. Computer Vision Center Technical Report, n. 24, 1998.

MALTONI, D.; et al. **Handbook of Fingerprint Recognition**. New York: Springer, 2003.

MEHTRE, B. M., et al. Color Matching for Image Retrieval. **Pattern Recognition Letters**, v. 16, p. 325-311, 1995.

MENSER, B.; WIEN, M. Segmentation and tracking of facial regions in color image sequences. In PROCEEDINGS OF SPIE VISUAL COMMUNICATIONS AND IMAGE PROCESSING, p. 731–740, 2000.

MITRA, M.; HUANG J.; KUMAR, S. R. Combining Supervised Learning with Color Correlograms for Content-Based Image Retrieval. In: PROCEEDINGS OF FIFTH ACM MULTIMEDIA CONFERENCE, 1997.

MIYAHARA, M.; YOSHIDA, Y. Mathematical transform of (RGB) Color Data to Munsell (HVC) Color Data. In: PROCEEDINGS OF SPIE VISUAL COMMUNICATIONS AND IMAGE PROCESSING, v. 1001. p. 650-657, 1988.

MORRIS, T. **Multimedia Systems: Delivering, Generating, and Interacting with Multimedia**. Springer, 2000

OLIVER, N.; PENTLAND, A.; BERARD, F. Lafter: lips and face real time tracker. In: IEEE COMPUTER VISION AND PATTERN RECOGNITION (CVPR 97), 1997.

OTAKE, S.; CICERONE, C. M. L and M cone relative numerosity and redgreen opponency from fovea to midperiphery in the human retina. In: JOURNAL OF THE OPTICAL SOCIETY OF AMERICA A: OPTICS, IMAGE SCIENCE, AND VISION, n. 17, 2000, pp. 615– 627.

PAPAMARKOS, N.; ATSALAKIS, A. E.; STROUTHOPOULOS, C. P. Adaptive Color Reduction. **IEEE Transactions on Systems, Man and Cybernetics**, v. 32, n. 1, p. 44-56, 2002.

PAPAMARKOS, N.; ATSALAKIS, A. E. Color reduction and estimation of the number of dominant colors by using a self-growing and self-organized neural gas. **Engineering Applications of Artificial Intelligence**, v. 19, p. 769-786, 2006.

POKORNY, J.; SMITH, V. C.; WESNER, M. F. Variability in cone populations and implications. In: FROM PIGMENTS TO PERCEPTION: ADVANCES IN UNDERSTANDING VISUAL PROCESSES (Valberg, A., Lee, B. B., eds), 1991, pp 23–34. New York: Plenum.

PASS, G.; ZABIH, R.; MILLER, J. Comparing Images Using Color Coherence Vectors. In: PROCEEDINGS OF FOURTH ACM MULTIMEDIA CONFERENCE, 1996.

PEER, P.; SOLINA, F. An Automatic Human Face Detection Method. In: PROCEEDINGS OF THE 4TH COMPUTER VISION WINTER WORKSHOP (CVWW'99). Raststfeld, Austria, p. 122-130 1999.

PHILLIPS, P. J.; MOON, H.; RAUSS, P. J.; RIZVI, S. The FERET evaluation methodology for face recognition algorithms. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 10, 2000.

QIU, G.; FENG, X.; FANG, J. Compressing Histogram Representations for Automatic Colour Photo Categorization. **Pattern Recognition**, v. 37, p. 2177-2193, 2004.

ROSENFELD, G. H.; FITZPATRICK, K. A coefficient of agreement as a measure of thematic classification accuracy. **Photogrammetric Engineering & Remote Sensing**, v. 52, n. 2, p. 227-233, 1986.

RIBEIRO, S. E. M. **Ilustrações lápis, nanquim e cores**. Silvio Ribeiro, 2005.

RUSHTON, W. A. H.; BAKER, H. D. Red/green sensitivity in normal vision. **Vision Research**, n. 4, 1964, p.75– 85.

SCHWERDT, K.; CROWLEY, J. L. Robust face tracking using color. In: PROCEEDINGS OF THE 3RD INTERNATIONAL CONFERENCE ON AUTOMATIC FACE AND GESTURE RECOGNITION (AFGR 00), 2000.

SCLAROFF, S.; TAYCHER, L.; LA CASCIA, M. Image-Rover: A Content-Based Image Browser for the World Wide. In: PROCEEDINGS OF IEEE WORKSHOP ON CONTENT-BASED ACCESS IMAGE AND VIDEO LIBRARIES, 1997.

SEBE, N.; COHEN, T.; HUANG, T. S.; GEVERS, T. Skin detection, a Bayesian network approach. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR), 2004.

SEVERINO JR, Osvaldo, GONZAGA, A. Misturograma – Uma proposta de Quantização do Histograma através da Mistura de Cores In: I Workshop de Visão Computacional (WVC'2005). Piracicaba, 2005, p.76 – 79.

SEVERINO JR, Osvaldo, GONZAGA, A. Auto-esboço e Auto-pele pelo Misturograma. In: II Workshop de Visão Computacional (WVC'2006). São Carlos, 2006a, p. 81 – 86.

SEVERINO JR, Osvaldo, GONZAGA, A. Auto-Fusão na Recuperação de Imagens Baseadas em Conteúdo. In: II Workshop de Visão Computacional (WVC'2006). São Carlos, 2006b, p. 305 – 310.

SEVERINO JR, Osvaldo, GONZAGA, A., BARCELLOS, R., SILVA, K. V. Performance comparison of content-based image retrieval systems using color autocorrelograms, mixturegrams and histograms of perceptual colors as feature Vectors In: II Workshop de Visão Computacional (WVC'2006). São Carlos, 2006, p.117 – 122.

SEVERINO JR, Osvaldo, GONZAGA, A. Auto-Fusão da Auto-Face, do Auto-Esboço e da Auto-Pele pelo Misturograma em imagens em nível de cinza. In: III Workshop de Visão Computacional (WVC'2007). São José do Rio Preto, 2007, p. 64 – 68.

SHARMA, G. **Digital Color Image Handbook**. New York: CRC Press LLC, 2003.

SIRISATHITKUL, Y.; AUWATANAMONGKOL, S.; UYYANONVARA, B. Color Image Quantization using Distances Between Adjacent Colors Along the Color Axis With Highest Color Variance. **Pattern Recognition Letters**, v. 25, p. 1025-1043, 2004.

SMITH, J. R.; CHANG, S. F. Single Color Extraction and Image Query. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 1995.

SMEULDERS, A. W. M. et al. Content-Based Image Retrieval at the End of the Early Years. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 12, p. 1349-1380, 2000.

SORIANO, M.; MARTINKAUPPI, J. B.; HUOVINEN, S.; LÄÄKSONEN, M. Adaptive skin color modeling using the skin locus for selecting training pixels. **Pattern Recognition**, v. 36, n. 3, p. 681–690, 2003.

SPROSON, W. N. **Color Science in Television and Display Systems**. Adam Hilger Ltda: 1983.

STÖRRING, M.; KOÈKA, T.; ANDERSON, H. J.; GRANUM, E. Tracking regions of human skin through illumination changes. **Pattern Recognition Letters**, v. 24, n. 11, 2003.

SUZUKI, J.; FURUKAWA, I. Required Number of Quantization Bits for CIE XYZ Signals Applied to Various Transforms in Digital Cinema Systems. In: **IEICE TRANSACTIONS ON FUNDAMENTALS OF ELECTRONICS, COMMUNICATIONS AND COMPUTER SCIENCES**, p. 1072-1084, 2007.

SYEDA-MAHMOOD, T. F. Data and Model-Driven Selection Using Color Regions. Massachusetts Institute of Technology, n, 1270, 1992

TERRILLON, J. C.; SHIRAZI, M. N.; FUKAMACHI, H.; AKAMATSU, S. Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. IN **PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON FACE AND GESTURE RECOGNITION**, p. 54–61, 2000.

UCHIYAMA, T.; ARBIB, M. A. Color Image Segmentation Using Competitive Learning. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 16, p. 1197-1206, 1994.

VAILAYA, A.; JAIN, A.; ZHANG, H. J. On Image Classification: Cities Images vs. Landscapes. **Pattern Recognition**, v. 31, p. 1921-1935, 1998.

XIANG, Z. Color Image Quantization by Minimizing the Maximum Intercluster Distance. **ACM Transaction on Graphics**, v. 16, n. 3, p. 260-279, 1997.

WANG, J. G.; SUNG, E. Frontal-view face detection and facial feature extraction using color and morphological operations. **Pattern Recognition Letters**, v. 20, p. 1053–1068, 1999.

YANG, J.; LU, W.; WAIBEL, A. Skin-color modeling and adaptation. In: **3RD ASIAN CONFERENCE ON COMPUTER VISION (ACCV'98)** , 1998.

YANG, M. H.; AHUJA, N. Gaussian Mixture model for human skin color and its application in image and video databases. In: **PROCEEDINGS OF SPIE: CONFERENCE ON STORAGE AND RETRIEVAL FOR IMAGE AND VIDEO DATABASES**, v. 36, n. 56, p. 458–466, 1999.

YANG, M. H.; KRIEGMAN, D. J.; AHUJA, N. Detecting faces in images: a survey. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 1, p. 34–58, Mach Intell, 2002.

YAMAGUCHI, T.; MOTULSKY, A. G.; DEEB, S. S. Visual pigment gene structure and expression in the human retinae, **Human Molecular Genetics.**, n. 6, 1997, pp. 981–990.

ZHAO, W.; CHELLAPPA, R.; PHILIPS, P.J.; ROSENFELD, A. Face recognition: a literature survey. **ACM Computing Surveys**, v. 85, n. 4, 2003.