

Capítulo 5 – Metodologia Proposta

5.1 Introdução

O histograma de uma imagem em nível de cinza fornece informação da distribuição dos valores dos *pixels* e é útil para a análise da imagem. É representado por um gráfico que dá o número de *pixels* na imagem para cada nível de cinza.

O histograma de uma imagem em nível de cinza é uma função de distribuição de probabilidades, onde cada elemento da imagem é calculado por:

$$\Pr(r_k) = \frac{n_k}{n} \quad (5.1)$$

Onde:

- $k=0,1,\dots L-1$ e L é a quantidade de nível de cinza da Imagem
- n é o número total de *pixels* na Imagem
- n_k é o número de *pixels* cujo nível de cinza corresponde a k
- $\Pr(r_k)$ é a probabilidade do k -ésimo nível de cinza

Quando a imagem é colorida, o valor de cada *pixel* pode ser representado por 3 componentes (R,G,B) ou (H,S,V) ou (X,Y,Z).

A quantidade de níveis de cinza em uma imagem tem significado visual (perceptual), ou seja, se a imagem é mais escura significa que há uma grande

concentração do histograma próximo ao eixo vertical e se é mais clara os valores concentram-se no final da escala de cinza.

No caso de imagens coloridas, o histograma deveria ser tridimensional mostrando a distribuição das cores no espaço. Não tem significado perceptual a subdivisão em 3 histogramas separados, um para cada elemento formante da cor. Logo, uma cor com seus três elementos deveriam ocupar apenas uma célula (*bin*) em uma distribuição de probabilidades de cores unidimensional. Isso recai no seguinte problema: como projetar no espaço bidimensional a distribuição tridimensional dos agrupamentos de cores em uma imagem?

Além disso, as aproximadamente 16 milhões de cores representadas em um monitor de computador para imagens digitalizadas seriam significativas para a interpretação humana tal que, ao se fazer uma busca por conteúdo de uma imagem colorida, seja possível a um usuário detalhar a quantidade de cada uma das 16 milhões de possibilidades?

Com estas duas questões em mente, este trabalho propõe o uso de um Histograma Perceptual baseado nas onze cores culturais já amplamente estudadas por pesquisadores da área de psicologia humana (BERLIN; KAY, 1991), realizando aplicações na recuperação de Imagens Baseada em Conteúdo.

A metodologia proposta para agrupamento de todas as cores em apenas onze cores culturais é baseada em um modelo "*Fuzzy*" onde as regras são estabelecidas através do conhecimento humano obtido por meio de uma pesquisa de classificação das cores disponibilizada na *Internet*.

Alguns trabalhos com classificação de imagens coloridas utilizam o conceito de cores culturais. No entanto, não se descreve a metodologia para a classificação das diversas cores de uma imagem em cores culturais (como por exemplo no trabalho de Meng, Chang e Li (2004)).

De uma maneira mais ampla, a metodologia como as cores são agrupadas é significativa para o sucesso de um sistema de CBIR baseado nesse tipo de característica.

A solução adotada para que se pudesse obter a categorização das cores de uma imagem em cores culturais foi a de associar a percepção humana das

cores através de um mecanismo de votação. Para isso foi desenvolvido um "Site" na *Internet* com o objetivo de se "extrair" o conhecimento sobre cores de um maior número de usuários possível e com isso estabelecer as regras "Fuzzy" que classificassem as cores em cores culturais gerando assim o que chamamos de Histograma Perceptual de Cores.

A idéia é que pessoas de qualquer parte do mundo pudessem acessar e dar sua opinião individualmente através de votação, de acordo com a sua percepção.

Os dados levantados através desta metodologia devem "conviver" com os erros oriundos de diversas fontes, a saber:

- a) Erros intencionais do usuário na classificação de cores.
- b) Diferenças de cores entre diferentes monitores de vídeo.
- c) Problemas de "saúde visual" do usuário.
- d) Outros erros.

A distribuição de todos os erros foi absorvida pela Lógica Nebulosa constituindo parte inerente do próprio sistema.

As páginas foram construídas em PHP (*Hypertext Preprocessor*), uma linguagem de programação utilizada para desenvolvimento de *Web-Sites*. Para o armazenamento das informações foi utilizado o banco de dados MySQL (*My Structured Query Language*), pela sua facilidade e segurança em se trabalhar com os dados.

Na figura 5.1 é mostrado o modelo do banco de dados, que foi elaborado para o desenvolvimento do *Site*.

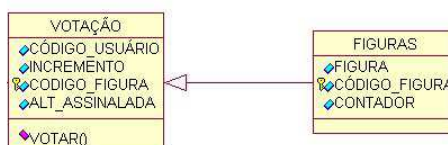


Figura 5.1: Modelo do Banco de Dados

Foram criadas duas tabelas (IMAGENS E VOTAÇÃO), especificadas como a seguir:

IMAGENS \Rightarrow Nesta tabela são armazenadas todas as informações das figuras mostradas no *Site* e do(s) usuário(s) que está(ão) acessando o *Site* naquele momento. No campo "FIGURA" é armazenado o nome da figura para referência para poder ser exibida na tela. O campo "CÓDIGO FIGURA" é um código seqüencial para cada uma das figuras, apenas para não ter duplicidade de dados e por último, o campo "CONTADOR" destina-se a armazenar quantas vezes foi votada aquela figura e para referência na hora da exibição, visto que o *Site* alterna as figuras de acordo com a angulação do (H), ou seja, as figuras são embaralhadas no momento da exibição, reduzindo a possibilidade de uma figura ficar sem votação ou ter mais votações que outras.

VOTAÇÃO \Rightarrow São armazenados os dados da votação do usuário de cada figura acessada pelo mesmo.

Todos os dados armazenados da votação permitirão estabelecer regras para a classificação da cor do *pixel*, de acordo com as onze cores culturais. Por exemplo, dado um *pixel* de entrada, será possível avaliar em qual das cores culturais o mesmo se enquadra.

Algumas cores geram mais incertezas do que outras na hora da classificação. Normalmente essas cores duvidosas em relação ao seu reconhecimento são aquelas encontradas no limite final de uma cor e início da próxima, pelo fato dessa pequena distância confundir ou misturar a percepção de quem vê.

5.2 – Resultados do *Site*

Com o intuito de simplificar o processo de aquisição do conhecimento, antes de serem lançadas no *Site*, as cores foram classificadas preliminarmente. Algumas pessoas foram contactadas pessoalmente a dar sua opinião e, de acordo com um consenso geral, essas imagens não foram inseridas nas imagens do *Site*, pois tiveram cem por cento de votação em uma única cor, não

havendo necessidade das mesmas serem inseridas, pois o universo de imagens seria muito grande.

O modelo utilizado foi o HSV. Esse modelo traz uma visualização melhor que o modelo RGB e é compreendido melhor pelos monitores do que o modelo CIE, pelo fato do modelo CIE possuir mais cores do que o monitor de vídeo consegue interpretar. Durante os testes, o modelo HSV mostrou-se mais nítido, mais claro perceptualmente, apresentou resultados melhores para se trabalhar com imagens coloridas e é o modelo que mais se aproxima a percepção humana.

Após a definição de qual modelo utilizar, algumas paletas de cores foram geradas para a classificação dos intervalos, como pode ser observado na figura 5.2.

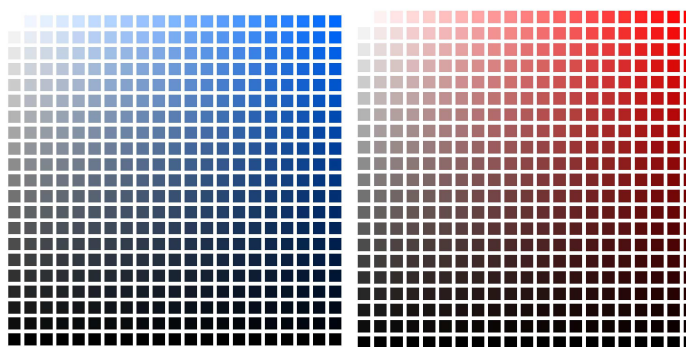


Figura 5.2: Paleta de Cores

Com essas paletas de cores, primeiramente foi feita a classificação com algumas pessoas escolhidas aleatoriamente e, após um consenso geral foram retiradas aquelas onde houve alguma divergência de reconhecimento. Cada célula da paleta onde se constatou a divergência, corresponde a uma figura do *Site*.

As paletas foram desenvolvidas em MATLAB, numa variação de trinta em trinta graus em (H), pois uma variação menor, quase não é perceptível pelo olho humano, não havendo a necessidade de incluí-las.

Alguns cuidados e estratégias foram adotados para se montar a sequência das figuras. A mesma figura foi disponibilizada no *Site* três vezes, de acordo com os três contextos de fundo definidos (branco, preto e cinza). O contexto pode influenciar na percepção da cor, e isso pode ser importante na aquisição de conhecimento.

As figuras foram ordenadas de acordo com o grau da *Matiz* (H).

O intuito dessa estratégia foi embaralhar as figuras e verificar se o usuário votaria na mesma opção de cor para as três figuras disponibilizadas no *Site* e até que ponto o contexto de fundo influencia a percepção.

Na disposição das figuras, foi inserido um fundo de tela branco, na tentativa de tratar a cor, visto que qualquer outra cor como fundo poderia influenciar a percepção.

Todos os dados adquiridos no *Site* foram armazenados em um banco de dados e foram utilizados para a classificação dessas cores, dentro das onze cores culturais.

A tela com as figuras é mostrada na figura 5.3. Nessa figura pode-se observar um quadrado central e outro em volta. O quadrado em volta é o contexto, que poderá influenciar na percepção da cor, e o quadrado central é a cor que se deseja obter a classificação.

As figuras foram elaboradas com apenas três cores de contexto de fundo, conforme pode-se observar na figura 5.4 pelo fato da necessidade de se obter um rápido retorno da pesquisa e para não ficar cansativo para a pessoa no ato da votação. O usuário não é obrigado a votar em todas as figuras, fica a critério de cada um a quantia a se responder. Apenas é salientado no *Site* para que a pessoa colabore e responda ao maior número de figuras possível.

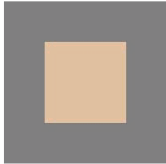
O *Site* possui um total de 417 figuras e foi disponibilizado do servidor da USP, através do *link*: (<http://iris.sel.eesc.sc.usp.br/tutorial>).

Os totais de dados colhidos da votação até a data de redação desta dissertação são mostrados na tabela 5.1.

Tabela 5.1: Classificação das cores pelos usuários do Site

Cores	Votos
Vermelho	164
Amarelo	131
Violeta	1375
Azul	583
Verde	711
Rosa	462
Marrom	439
Preto	56
Branco	24
Laranja	74
Cinza	510
Total	4529

Which color is similar to the square center color?




- ☐ Red
- ☐ Yellow
- ☐ Purple
- ☐ Blue
- ☐ Green
- ☐ Pink
- ☐ Brown
- ☐ Black
- ☐ White
- ☐ Orange
- ☐ Gray

<< Previous Next >> Quit

3 of 417

Please, vote in any color you want. You can stop any time. Your help is very important for our project. Thank you.

Qual a cor que mais se aproxima a do centro do quadrado ?



- ☐ Vermelho
- ☐ Amarelo
- ☐ Violeta
- ☐ Azul
- ☐ Verde
- ☐ Rosa
- ☐ Marrom
- ☐ Preto
- ☐ Branco
- ☐ Laranja
- ☐ Cinza

Avançar >> Sair

1 de 417

Por favor, vote no maior número de cores possível. Você pode parar a qualquer instante que quiser. Sua colaboração é muito importante para nosso projeto !

Figura 5.3: Figura para votação

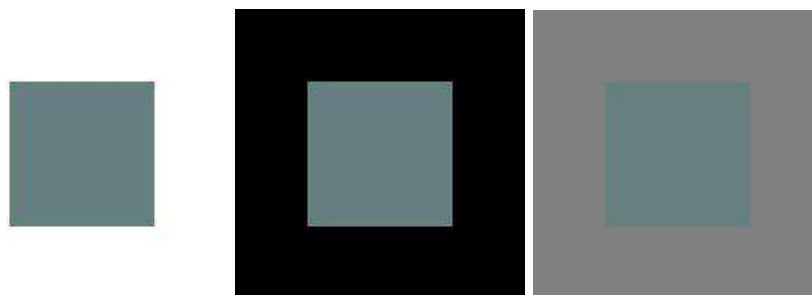


Figura 5.4: Exemplos de diferentes contextos

5.3 – Materiais e Métodos

Durante todo o trabalho, o enfoque principal abordado foi a importância e a influência das cores na recuperação de imagens e o quanto elas representam para esta área de pesquisa.

Para efeito de comparação, desempenho e visualização, foram construídos dois sistemas que são abordados com maiores detalhes no Apêndice B, ambos tendo como característica principal as cores.

5.3.1 – A Base Computacional

Ao longo dos testes trabalhou-se com quatro bases de imagens. Duas bases de imagens genéricas, uma base de imagens encontrada na Internet e outra base de imagens de faces humanas. As imagens destas duas bases (genéricas e faces humanas) geraram duas novas onde a iluminação foi alterada através de pré-processamento de brilho.

A primeira base compreende 540 imagens, sendo 135 imagens originais e 3 modificadas, conforme mostra a figura 5.5. A imagem 1 da figura 5.5 é a principal, a imagem 2 é a inversão horizontal da imagem 1, a imagem 3 é a inversão vertical da imagem 1 e a imagem 4 é a inversão horizontal da imagem 2. Nesta primeira base de imagens trabalhou-se apenas com rotação.



Figura 5.5: Base 1 – Imagens Genéricas refletidas em dois eixos

As imagens genéricas têm a dimensão de 82 linhas por 110 colunas. Todas as imagens encontram-se no formato RGB, com 24 *bits*.

Para os testes converte-se a imagem para o modelo de cores HSV.

Após o processamento das informações, as imagens são mostradas ao usuário no modelo RGB, padrão dos monitores de vídeo.

Na segunda base de dados, trabalhou-se com faces humanas (MARTINEZ; BERRAVENTE, 1998), sendo 76 imagens originais de faces masculinas e 59 de faces femininas. Estas imagens possuem 96 linhas por 128 colunas. A figura 5.6 traz amostras das imagens que fazem parte da Base 2, sendo os testes realizados com as imagens originais (imagens 1 e 5) e com algumas modificações nas expressões faciais (imagens 2, 3, 4, 6, 7 e 8).

O intuito de se testar imagens genéricas e faciais foi exatamente para averiguar a eficiência do Histograma Perceptual de Cores proposto. Ao longo dos testes, devido às imagens de faces terem um número menor de tonalidades de cores, não se conseguiu uma boa performance quanto em comparação com as imagens genéricas. Essa baixa performance deve-se ao fato de que os histogramas das diferentes classes que compõem as imagens de faces serem próximos uns dos outros. Já nas imagens genéricas entre uma classe e outra essa diferença não é tão mínima, favorecendo os resultados.

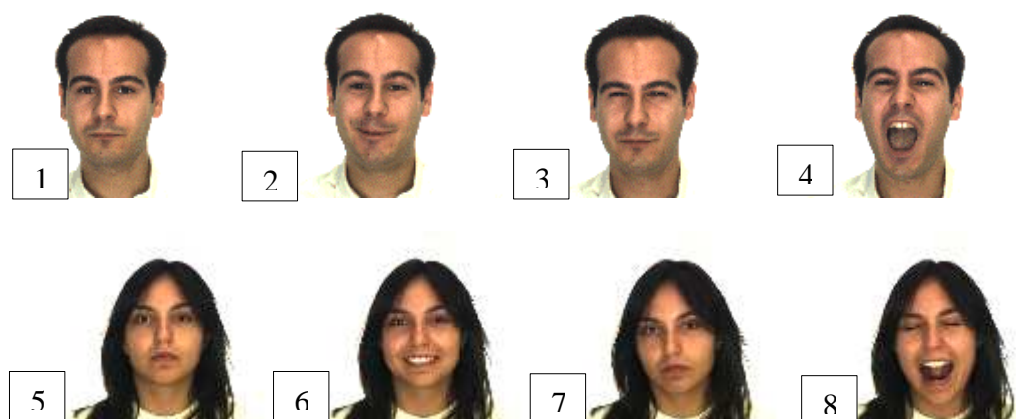


Figura 5.6: Base 2 - Faces masculinas e femininas

Na terceira base de dados aumentou-se duas imagens para cada uma das classes, ou seja, uma imagem com aumento da intensidade do brilho e outra com diminuição do brilho, conforme se pode observar nas imagens 5 e 6 da figura 5.7. A terceira base ficou com um conjunto de 810 imagens.

A quarta base de dados com faces de homens e mulheres, conforme figura 5.8, corresponde às mesmas imagens da segunda base de dados acrescidas das imagens com o brilho alterado.



Figura 5.7: Base 3 - Imagens genéricas com alteração da iluminação

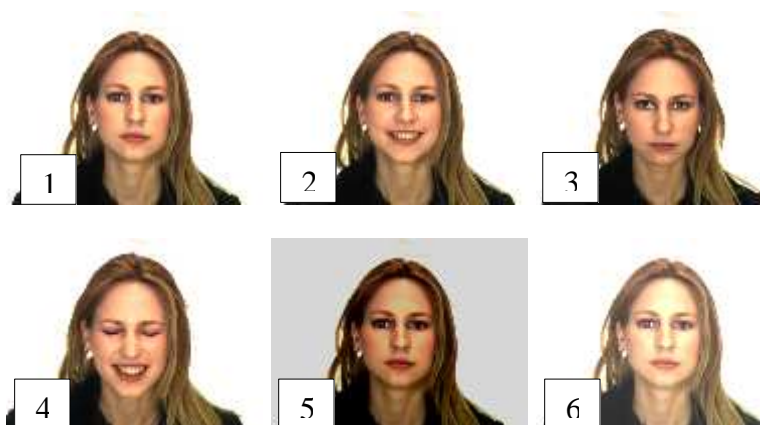


Figura 5.8: Base 4 - Imagens de faces humanas com alteração da iluminação

Como método de comparação de desempenho, utilizou-se gráficos de *Precision* (precisão) x *Recall* (revocação). Optou-se por manter nos gráficos a nomeação dos eixos em inglês.

Para a primeira e segunda base de imagens usou-se um *search* de 4, ou seja, buscou-se as imagens cujo histograma perceptual tem as 4 menores distâncias entre os vetores, da imagem base (*query image*). A distância utilizada para comparar os vetores foi a Euclidiana. Já para a terceira e a quarta base de imagens utilizou-se um *search* de 6.

5.4 – Classificador *Fuzzy*: Histograma Perceptual de Cores (HPC)

De acordo com a votação através do *Site*, montou-se a base de dados do classificador *Fuzzy*. Todas as informações da votação foram transferidas para uma matriz que foi treinada e aprimorada para poder classificar corretamente os *pixels*. A base de dados é constituída de 936 registros, onde cada registro corresponde a uma especificação de cor. Uma amostra desta base pode ser visualizada na tabela 5.2. A tabela completa pode ser melhor analisada na tabela C.7 do Apêndice C.

Analisando a tabela 5.2, tem-se as três entradas de dados (H, S e V) e um índice para se estabelecer a representação da cor. Esta tabela também é utilizada como conjunto de dados para o treinamento do sistema *Fuzzy*, como pode ser observado na figura 5.12.

Exemplo: a cor amarela é representada na base de dados com o índice 1, já o azul é representado pelo índice 2, e assim sucessivamente para todas as cores.

Tabela 5.2 – Representação parcial da Base de Dados *Fuzzy*

H	S	V	ÍNDICE
0.8889	0	1.0	1
0.9167	0	1.0	1
0.9444	0	1.0	1
0.6667	0.1	0.9	2
0.7500	0.1	0.9	2
0.8333	0.1	0.9	2
1.0	0	0	3
0	0.1	0.2	3
0.0833	0.1	0.2	3

Devido ao grande número de imagens as cores foram apresentadas em intervalos de 30 em 30 graus, de acordo com o modelo de cores HSV. Nos testes para classificação das cores, verificou-se que esse intervalo estabelecido gerava uma taxa de erros muito grande, comprometendo a eficiência do Classificador *Fuzzy*.

Dividiu-se, então, novamente o modelo HSV de 10 em 10 graus redefinindo o intervalo das cores. Após esta segunda definição notou-se uma melhora nos resultados.

Com a taxa de erro muito alta, um *pixel* vermelho, com H=1, S=0, V=0, por exemplo, pode ser classificado como rosa, afetando negativamente o Classificador.

A figura 5.9 mostra o funcionamento do Classificador *Fuzzy*, apresentando a forma como o mesmo processa as informações dadas como

entrada (H, S e V) e retorna o *pixel* classificado (R), de acordo com o treinamento na base de dados.

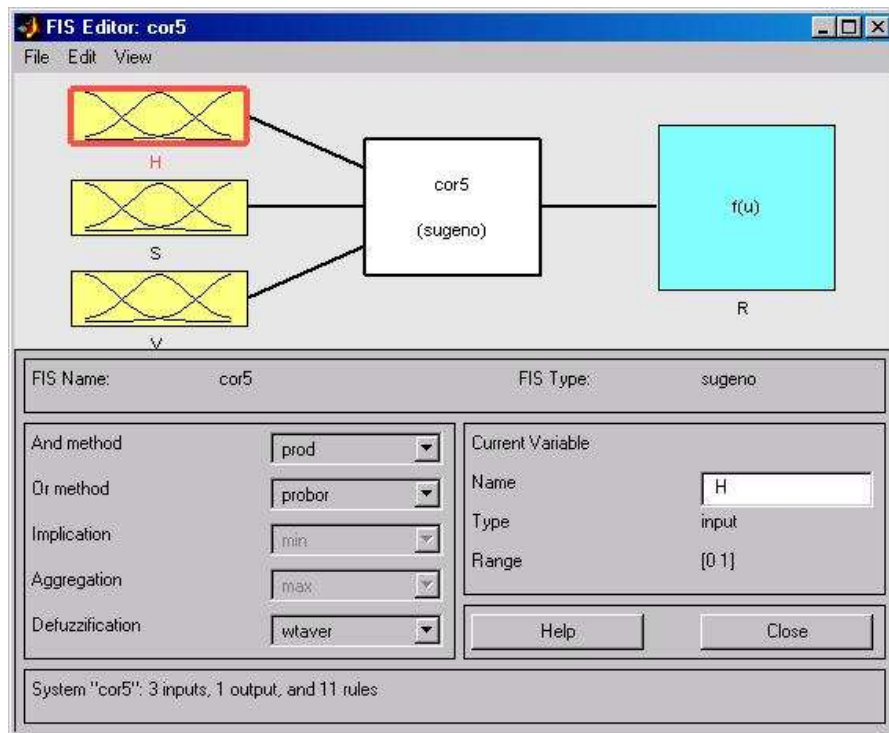


Figura 5.9 – Classificador *Fuzzy*

As informações de H, S e V de cada *pixel* são enviadas para cada uma das entradas do Classificador *Fuzzy*, conforme pode-se observar na figura 5.9. De acordo com a base de dados treinada, o Classificador *Fuzzy* retorna um grau de pertinência daquele *pixel* a uma das onze cores culturais.

A figura 5.10 mostra as variáveis de saída (R). Após o processamento, o Sistema *Fuzzy* retornará apenas um resultado, ou seja, uma das 11 cores culturais perceptuais.

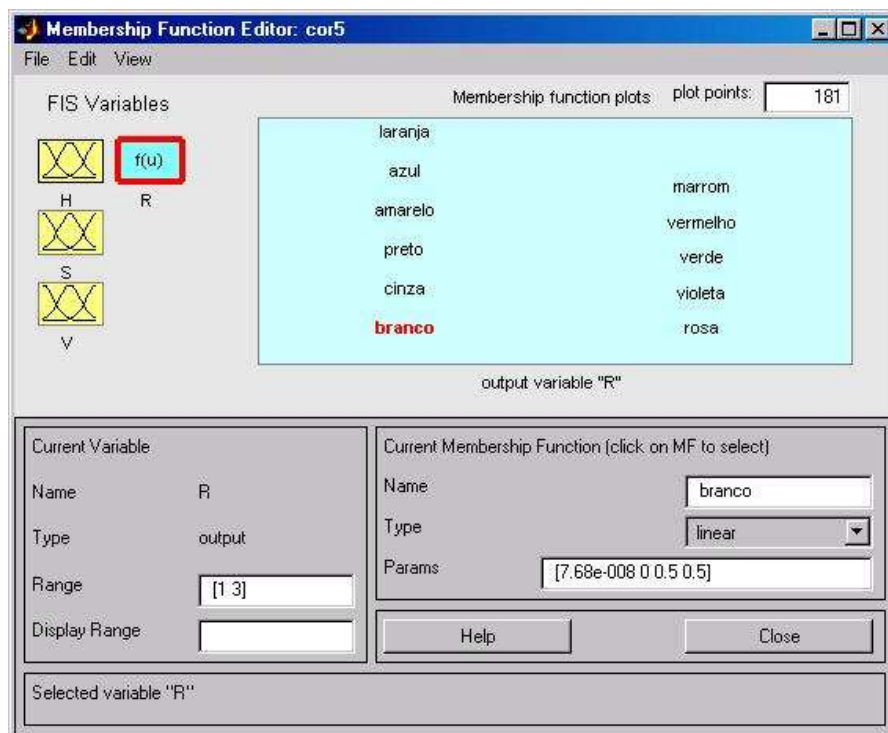


Figura 5.10 – Resultados do Classificador *Fuzzy*

O resultado é retornado em uma pertinência de que aquele *pixel* possa ser uma cor específica. O Classificador *Fuzzy* faz referências às cores através dos índices definidos na base de dados para treinamento, ou seja, para cada cor representada no resultado (R), o Classificador *Fuzzy* a interpreta como um número.

Para que o Classificador *Fuzzy* possa gerar o HPC é necessário estabelecer algumas regras, conforme pode ser visualizado na figura 5.11. Estas regras foram criadas de acordo com o intervalo estabelecido para cada uma das cores culturais. Considere como exemplo a primeira linha, tendo como representação a cor branca, destacada em azul na figura 5.11.

O intervalo do branco compreende H variando de 0 a 1, S = 0 e V = 1. Caso as informações do *pixel* que está sendo classificado se enquadre nesse intervalo do branco, essa regra se tornará ativa e o sistema retornará como resultado o índice 3, ou seja, referenciado como branco.

Do mesmo modo que se estabeleceu esta regra para a cor branca, o padrão se estende para as demais cores. É preciso saber qual o intervalo que

compreende a respectiva cor e inserir a regra, que o Classificador *Fuzzy* a utiliza como método para se estabelecer o resultado.

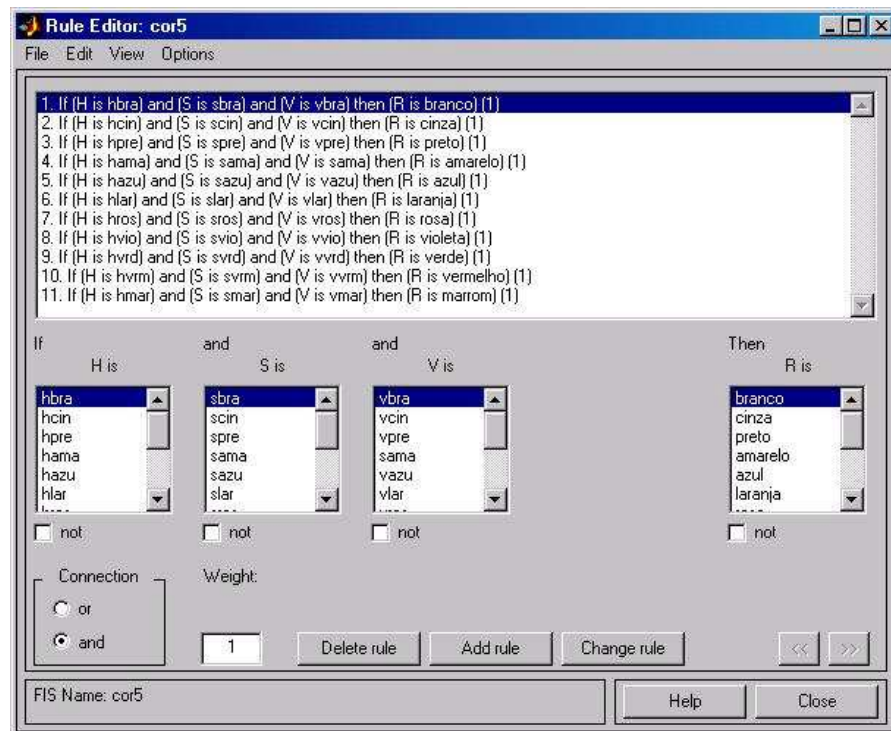


Figura 5.11 – Regras *Fuzzy*

A figura 5.12 mostra como o Classificador *Fuzzy* faz o treinamento na base de dados. Na opção para localização dos dados (*Load Data*), é selecionado o arquivo que contém todos os registros com as definições das cores. Essa base de dados é um arquivo com extensão ".dat", sendo padronizada conforme mostrado na tabela 5.2.

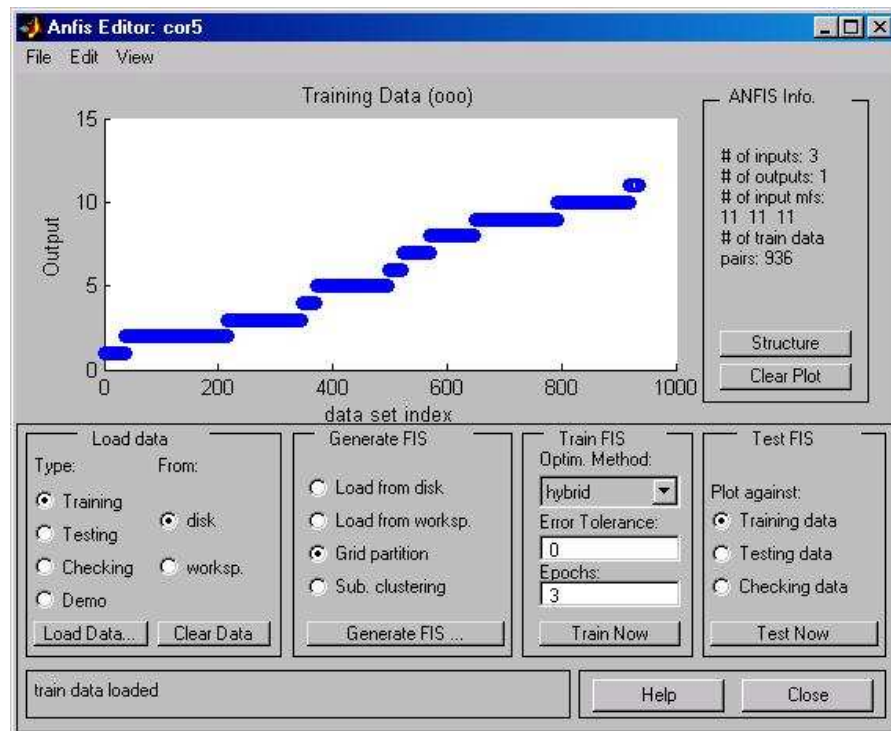


Figura 5.12– Treinamento da Base de Dados *Fuzzy*

Ao selecionar a base de dados(a tabela completa destes valores pode ser observada no apêndice C, tabela c.7), o próprio sistema gera a representação dos dados no formato de um gráfico, conforme mostra a figura 5.12. Esses dados serão treinados (*training data*) para que o sistema possa classificar cada *pixel* de acordo com os parâmetros pré-estabelecidos.

A opção de teste (*test now*) deve ser ativada após o treinamento dos dados, conforme pode ser visualizado na figura 5.13. Através desta opção pode-se observar se cada um dos *pixels* representados na tabela da base de dados *Fuzzy* está devidamente classificado no seu intervalo específico.

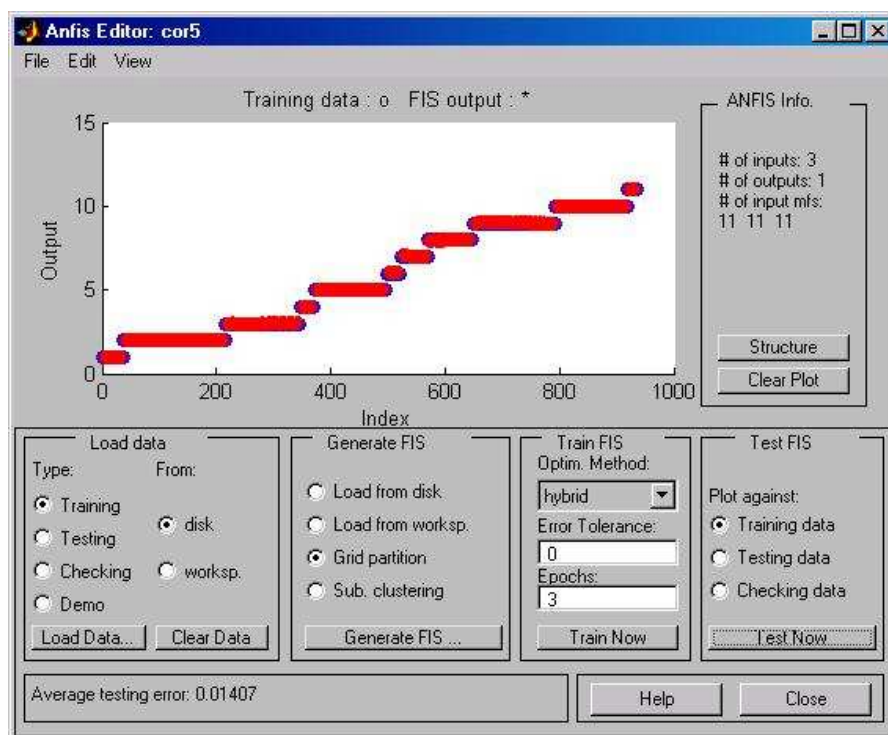


Figura 5.13 – Resultado após o treinamento na Base de Dados *Fuzzy*

A cor em vermelho representada no gráfico da figura 5.13 significa que a base de dados está devidamente treinada e que todos os *pixels* se enquadram em seu intervalo respectivo.

A figura 5.13 mostra que após o treinamento o classificador *Fuzzy* gera a informação da taxa de erros (*average testing error*). Essa taxa de erros tem grande influência na classificação dos *pixels*. Quanto menor essa taxa, mais confiável é o resultado fornecido pelo Classificador *Fuzzy*.

Uma outra forma de visualização da operação do Classificador *Fuzzy* é dada na figura 5.14, onde pode ser observadas as entradas, as regras, as saídas e o resultado. O Classificador *Fuzzy* não retornará resultados diferentes daqueles já pré-estabelecidos.

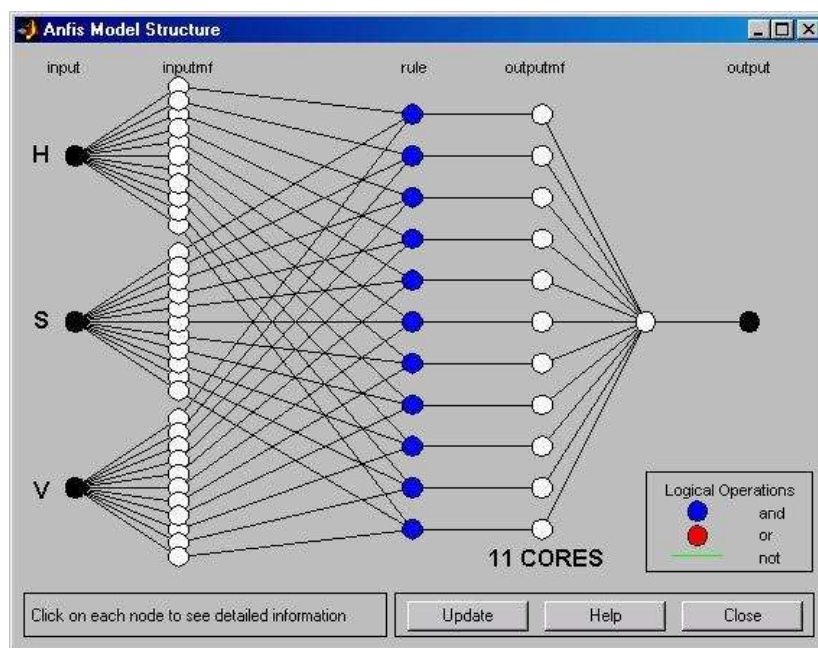


Figure 5.14 – Visualização regras *Fuzzy*

As figuras 5.15, 5.16 e 5.17 são as representações dos intervalos de cada uma das cores para os canais de H, S e V.

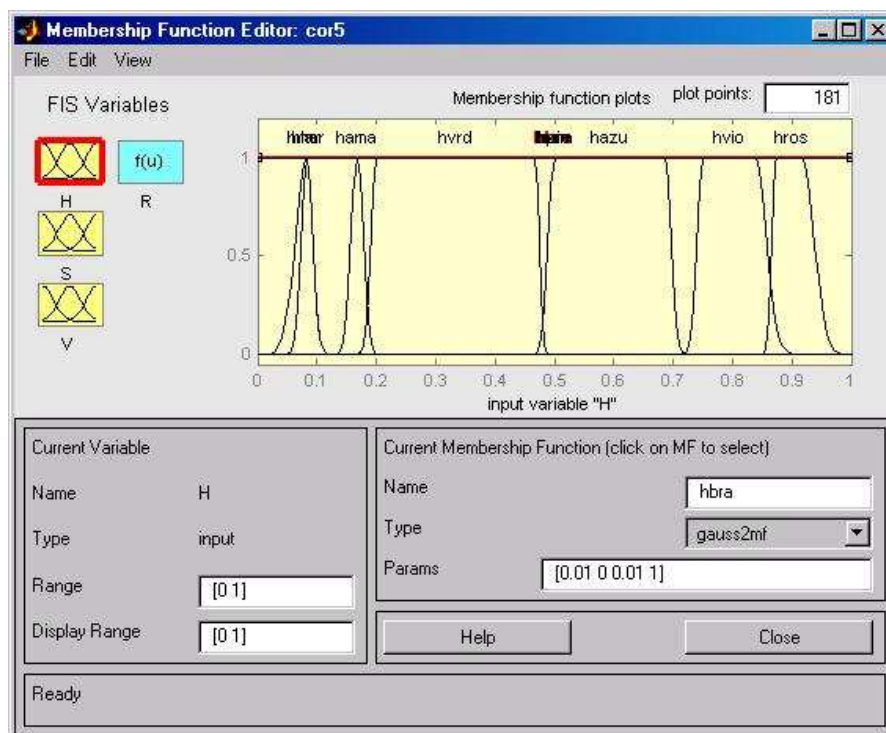


Figure 5.15 – Gráfico das Gaussianas (Entradas de H)

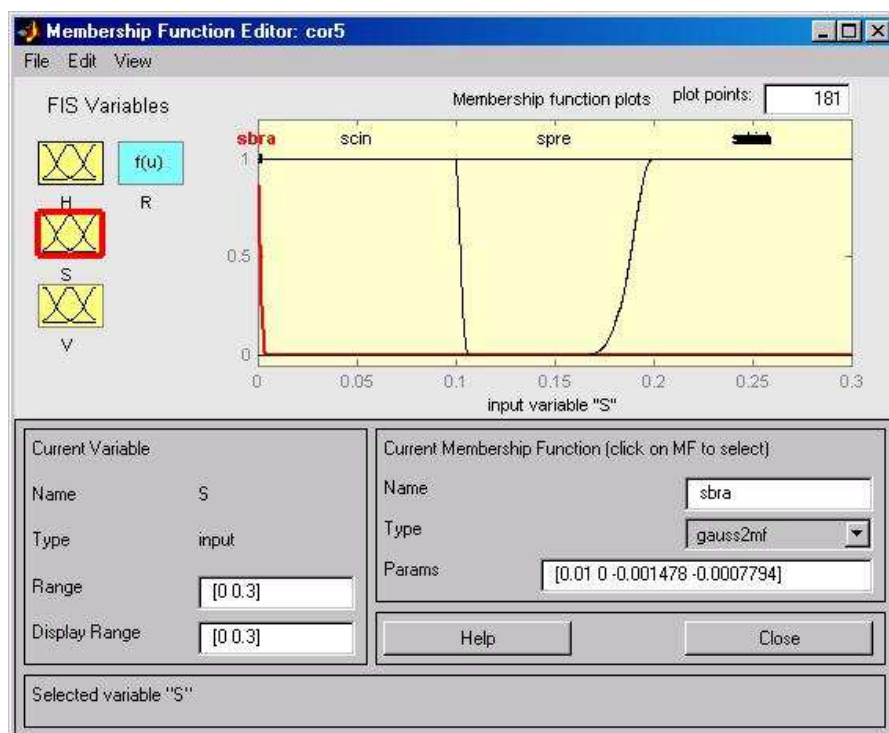


Figura 5.16 – Gráfico das Gaussianas (Entradas de S)

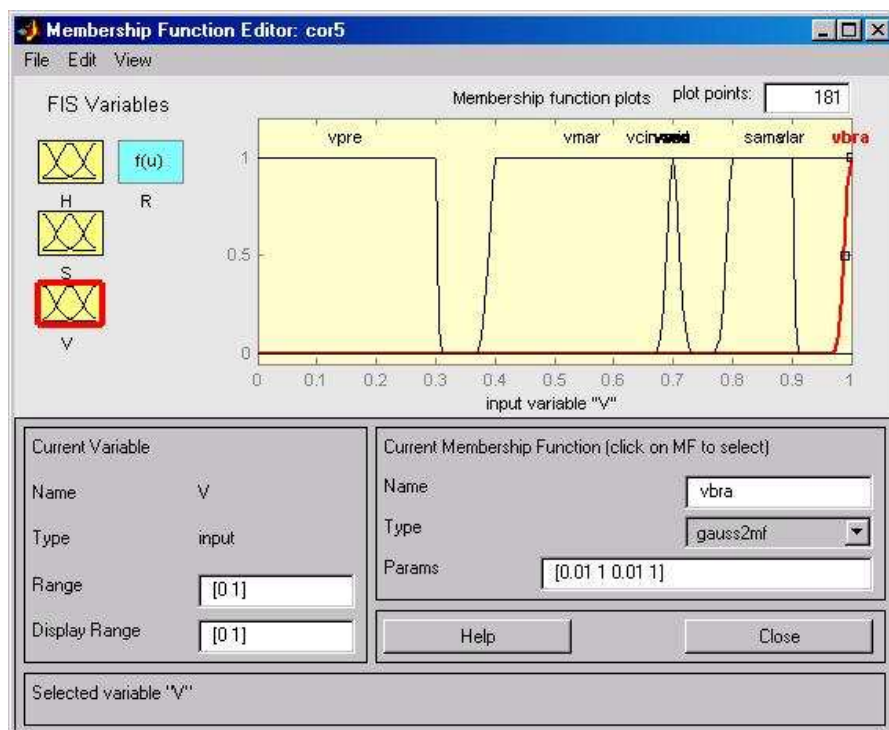


Figura 5.17 – Gráfico das Gaussianas (Entrada de V)

Estas são as curvas das cores moldadas pelo Classificador *Fuzzy* (figuras 5.15, 5.16 e 5.17). Quanto mais intercaladas forem estas curvas gaussianas, maior é a taxa de erros. De acordo com a redefinição das cores, foi reduzida a intercalação entre essas curvas diminuindo essa taxa de erros.

Uma vez classificado cada *pixel* da imagem, de acordo com as 11 cores culturais, o resultado final é normalizado para um intervalo de valores de 0 a 1 dividindo-se o número de *pixels* de uma determinada cor pela quantidade total de *pixels* da imagem.

Após serem classificados todos os *pixels* para o HPC é gerado o vetor de características, constituído do nome da imagem e dos números totais de *pixels* normalizados das cores amarelo, azul, branco, cinza, laranja, marrom, preto, rosa, violeta, verde e vermelho.

Os valores das coordenadas de H, S e V de cada *pixel* foram armazenados em uma matriz.

Com o classificador *Fuzzy* operando pixel a pixel da Imagem, o tempo de processamento mostrou-se muito lento. Em uma máquina Athlon 1.2 Ghz com 256 MB de memória, o tempo de classificação de todos os pixels de uma imagem ficou em torno de 12 minutos. Armazenando-se todos os valores em uma única matriz, o tempo de processamento foi de aproximadamente 3 minutos.