

Universidade Federal do Maranhão
Centro de Ciências Exatas e Tecnologia
Programa de Pós-Graduação em Engenharia de Eletricidade

ESTUDO DE ÍNDICES DE DIVERSIDADE COMO
DESCRITORES DE TEXTURA PARA CLASSIFICAÇÃO DE
IMAGENS DE FACES

Artur Bernardo Silva Reis

São Luís
2012

Universidade Federal do Maranhão
Centro de Ciências Exatas e Tecnologia
Programa de Pós-Graduação em Engenharia de Eletricidade

ESTUDO DE ÍNDICES DE DIVERSIDADE COMO
DESCRITORES DE TEXTURA PARA CLASSIFICAÇÃO DE
IMAGENS DE FACES

Artur Bernardo Silva Reis

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão, para a obtenção de Título de Mestre em Engenharia de Eletricidade.

Orientador: Prof^o Aristófanês Corrêa Silva

Co-orientador: Prof^o Anselmo Cardoso de Paiva

São Luís
2012

Reis, Artur Bernardo Silva.

Estudo de Índices de Diversidade como Descritores de Textura para Classificação de Imagens de Faces | Artur Bernardo Silva Reis. - 2012.

81 f.

Impresso por computador.

Orientador: Aristófanês Corrêa Silva.

Co-Orientador: Anselmo Cardoso de Paiva.

Dissertação (Mestrado) - Universidade Federal do Maranhão, Programa de Pós Graduação em Engenharia de Eletricidade, 2012.

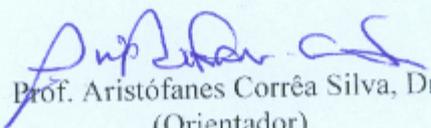
1. Imagens - Processamento 2. Padrões - reconhecimento 3. Índices de diversidade 4. Faces 5. Máquina de Vetores Suporte I. Título

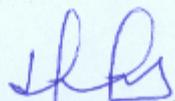
CDU 004.932

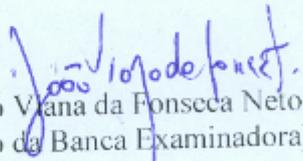
**ESTUDOS DE ÍNDICES DE DIVERSIDADE COMO
DESCRITORES DE TEXTURA PARA CLASSIFICAÇÃO DE
IMAGENS DE FACES**

Artur Bernardo Silva Reis

Dissertação aprovada em 13 de abril de 2012.


Prof. Aristófanes Corrêa Silva, Dr.
(Orientador)


Prof. Dimas Martínez Morera, Dr.
(Membro da Banca Examinadora)


Prof. João Viana da Fonseca Neto, Dr.
(Membro da Banca Examinadora)

À minha família, à Adrianne e amigos pelo apoio e companheirismo.

"Construímos muitos muros e poucas pontes."
Isaac newton

Agradecimentos

Agradeço a Deus por tudo.

À minha família por toda ajuda, incentivo, carinho e fé;

À minha namorada Adrienne, pelo carinho, apoio, compreensão e paciência;

Aos meus dois orientadores: Aristófanês Corrêa Silva e Anselmo Cardoso de Paiva. O primeiro pela paciência, incentivo, contribuições, ensinamentos, conselhos, oportunidades e toda ajuda. O segundo pelos conselhos, contribuições, incentivos e por ter aberto várias portas do mundo acadêmico.

Aos colaboradores deste trabalho: Eduardo Batista, Eduardo Reis, Tarsísio Almeida;

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES que financiou minhas pesquisas;

Aos colegas do Laboratório de Análise e Processamento de Imagens - LABPAI;

Aos colegas do Instituto Federal do Maranhão - IFMA pela ajuda prestada e compreensão;

A todos os professores e funcionários do Departamento de Pós-Graduação de Engenharia de Eletricidade da Universidade Federal do Maranhão - UFMA, pela competência e profissionalismo;

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

RESUMO

O objeto de estudo deste trabalho é uma investigação sobre o potencial grau de caracterização de texturas de faces humanas em imagens digitais usando índices de diversidade. Para isto é desenvolvida uma metodologia que utiliza o retinex como técnica de normalização de iluminação, índices de diversidade para a extração de características e a Máquina de Vetores Suporte como classificador. Para isto são realizados vários testes utilizando quatro bases de imagens de faces, com várias características peculiares para que possa ser verificada a usabilidade dos índices utilizados. Os resultados obtidos são muito promissores, uma vez que em dezoito testes realizados, foram atingidas acurácias acima de 95%. Além disso, obtivemos sensibilidade de 100% em um caso de teste e taxa de falsos positivos de 0.2%, indicando que os índices de diversidade podem ser uma boa ferramenta para a descrição de texturas de faces humanas.

Palavras-chave: Processamento de imagens. Reconhecimento de padrões. Índice de diversidade. Faces. Máquinas de vetores suporte.

ABSTRACT

This work purposes an investigation into the potential of characterization of textures of human faces in digital images using the diversity index. To this we developed a methodology that uses the retinex to normalization of lighting conditions, diversity indexes for feature extraction and Support Vector Machine as classifier. Several tests are performed using four bases of face images, each base with unques characteristics that can be verified the usability of the indexes used. The results are very promising since in eighteen tests were reached accuracies over 95%. In addition, a sensitivity of 100% in a test case and false positive rate of 0.2%, indicating that the diversity index can be a good tool for describing textures of human faces.

Keywords: Image processing. Pattern recognition. Diversity index. Faces. Support vector machine.

Lista de Tabelas

3.1	Quantidades de imagens por Base de Face	52
4.1	Resultados obtidos para o índice de Magalef	60
4.2	Resultados obtidos para o índice de Menhinick	60
4.3	Resultados obtidos para o índice de Shannon	61
4.4	Resultados obtidos para o índice de Pielou	62
4.5	Resultados obtidos para o índice de Simpson	63
4.6	Resultados obtidos para o índice de Simpson para equabilidade . .	64
4.7	Resultados obtidos para o índice de MacIntosh	64
4.8	Resultados obtidos para o índice de MacIntosh para equabilidade .	64
4.9	Resultados obtidos para o índice de brillouin	65
4.10	Resultados obtidos para o índice de Berger-Parker	66
4.11	Resultados obtidos de todos os índices para todas as bases juntas .	67
4.12	Resultados obtidos de todos os índices para todas as bases juntas .	67
4.13	Resultados para todos os índices aplicados a todas as ROIs	70
4.14	Comparativo com métodos utilizados na literatura	72

Lista de Figuras

2.1	Distribuição de indivíduos nas comunidades 1 e 2	26
2.2	Índice de Shannon para comunidades A e B	29
2.3	Etapas do Processamento de Imagens Digitais. Adaptado de (GONZALEZ and WOODS 1992)	34
2.4	Aplicação do Retinex	38
2.5	Hiperplano separando o espaço em duas regiões.	40
2.6	Hiperplano separando o espaço em duas regiões.	41
2.7	Distância entre hiperplanos e vetores suporte.	43
2.8	Conjuntos linearmente separáveis e não linearmente separáveis.	44
2.9	Formato da matriz de confusão.	46
3.1	Exemplos de imagens de não faces utilizados.	50
3.2	Exemplos de imagens da base ORL.	51
3.3	Exemplos de imagens de face da base CBCL.	51
3.4	Exemplos de imagens de face da base Libor Spacek.	52
3.5	Exemplos de imagens de faces da base Faces Expression.	52
3.6	Fuxo de atividades da metodologia proposta.	53
3.7	Fuxo de atividades durante o pré-processamento.	54

3.8	Extração do vetor de características utilizando a primeira abordagem (1 índice por ROI).	55
3.9	Extração do vetor de características utilizando a segunda abordagem (10 índices por ROI).	56
3.10	Fluxo de atividades durante a fase de reconhecimento de padrões.	56
4.1	Imagens classificadas como não face	68
4.2	Imagens classificadas como face	68
4.3	Imagens classificadas corretamente como face utilizando todos os índices	71
4.4	Imagens classificadas como incorretamente como não face utilizando todos os índices por ROI	71

Sumário

1	Introdução	15
1.1	Objetivos	16
1.1.1	Objetivos Específicos	17
1.2	Contribuição	18
1.3	Trabalhos Relacionados	18
1.4	Organização do Trabalho	22
2	Fundamentos Teóricos	23
2.1	Índices de Diversidade	23
2.1.1	Índice de Shannon	27
2.1.2	Índice de Pielou	29
2.1.3	Índice de Simpson	30
2.1.4	Índice de Brillouin	32
2.1.5	Índice de MacIntosh	33
2.1.6	Índice de Berger-Parker	33
2.2	Processamento de Imagens Digitais	34
2.2.1	Retinex	35
2.3	Reconhecimento de Padrões	37

Sumário	13
2.3.1 Máquinas de Vetores Suporte	39
2.4 Validação de Resultados	45
3 Materiais e Métodos	48
3.1 Software Utilizados	48
3.2 Base de Imagens Utilizadas	50
3.3 Metodologia Proposta	52
3.3.1 Pré-processamento	53
3.3.2 Extração de Características	54
3.3.3 Reconhecimento de Padrões	55
4 Resultados e Discussão	58
4.1 Aplicação de cada Índice Individualmente	58
4.1.1 Índice de Magalef	59
4.1.2 Índice de Menhinick	60
4.1.3 Índice de Shannon	61
4.1.4 Índice de Pielou	61
4.1.5 Índice de Simpson	62
4.1.6 Índice de Simpson para Equabilidade	63
4.1.7 Índice de MacIntosh	63
4.1.8 Índice de Brillouin	64
4.1.9 Índice de Berker-Parker	65
4.1.10 Combinação de Todas as Bases	66
4.1.11 Análise geral do Teste	67
4.2 Combinação de Todos os Índices	69
4.3 Comparação com métodos da Literatura	72

Sumário	14
5 Considerações Finais	73
Referências	76

Capítulo 1

Introdução

Com o avanço da tecnologia da informática surgiram muitas possibilidades de automatizar tarefas que só poderiam ser realizadas por seres humanos. A idéia de simular características humanas torna-se mais próxima da realidade, a medida que computadores recriam processos que apenas seres humanos são capazes de realizar. O sentido da visão tem recebido grande destaque na área da pesquisa e aplicações comerciais. O ramo da ciência da computação que estuda os processos que reproduzem a capacidade humana de enxergar usando máquinas é conhecido como visão computacional (MA, 2007).

O rápido crescimento na área de pesquisa relacionada à visão computacional é baseado na premissa da recuperação de informações de identificação humana, de tal forma, que os computadores sejam capazes de observar características de faces humanas (RANGEL, 2009). Nos últimos anos o reconhecimento de face tem atraído grande atenção de pesquisadores e da indústria. Algumas das aplicações que utilizam o processamento de uma face humana estão na área de segurança, biometria ou coleções de fotografias pessoais ou institucionais, etc. Uma das

etapas da identificação de pessoas em uma imagem, é encontrar uma maneira de descrever e reconhecer uma imagem como sendo uma face humana.

A ideia por trás do reconhecimento de um objeto em uma imagem, no caso uma face, consiste basicamente em duas etapas. Na primeira etapa, a imagem é representada de uma maneira que a descrição da informação deve ser suficiente para que haja uma distinção entre os objetos de interesse, este processo é chamado de extração de características de textura. Após termos a descrição da textura partimos para segunda etapa que é a classificação destas informações. A classificação é a atribuição de rótulos a um determinado conjunto de características de textura da imagem, este processo é realizado por uma técnica de aprendizado de máquina.

A importância da extração de características está no fato de poder fornecer informações de como os pixels da imagem estão distribuídos espacialmente, variações de tonalidade de um dado objeto, repetição de tonalidade. Existem várias técnicas que permitem a descrição da textura, neste trabalho usaremos estatísticas de distribuição de espécies em comunidades bióticas para esta tarefa.

1.1 Objetivos

O presente trabalho visa investigar e analisar em detalhes a relevância que os índices de diversidade têm ao obter informação textural presente em imagens digitais de faces. Para isso é proposta uma metodologia computacional que realiza tarefas em etapas: pré-processamento que visa normalizar condições de iluminação das imagens utilizadas; extração dos índices de diversidade das imagens; e por fim, classificação destas características em face e não face.

1.1.1 Objetivos Específicos

No sentido de alcançar o objetivo geral pretendido, buscar-se-á atingir os seguintes objetivos específicos:

- Analisar o desempenho dos seguintes índices de diversidade como descritores de textura de imagens de face e não face:
 - Índice de Magalef;
 - Índice de Menhinick;
 - Índice de Shannon;
 - Índice de Simpson;
 - Índice de Simpson de Equabilidade;
 - Índice de Pielou;
 - Índice de MacIntosh;
 - Índice de MacIntosh de Equabilidade;
 - Índice de Brillouin;
 - Índice de Berger-Parker;
- Analisar técnicas de reconhecimento de padrões, em especial a Máquina de Vetores de Suporte, para classificação de padrões obtidos;
- Analisar os resultados obtidos para cada índice de diversidade aplicado a diversas bases de imagens de face;
- Analisar os resultados obtidos para todos índices de diversidade aplicados de maneira conjunta a diversas bases de imagens de face.

1.2 Contribuição

As principais contribuições deste trabalho são:

- a) A utilização de índices de diversidade (Índice de Magalef, Índice de Menhinick, Índice de Shannon, Índice de Simpson, Índice de Pielou, Índice de MacIntosh, Índice de Brillouin, Índice de Berger-Parker) como descritores de características de textura de faces, pois são medidas estatísticas simples que apresentaram um bom desempenho mesmo gerando poucas características se comparados a outros trabalhos na literatura, implicando em baixo tempo de treinamento e classificação das imagens;
- b) Criação de uma metodologia computacional para classificação de imagens de faces humanas.

1.3 Trabalhos Relacionados

Na literatura existem várias abordagens para a caracterização e classificação de faces humanas.

No trabalho proposto por (TURK, 1991) as imagens correspondentes a faces são representadas por um método chamado *eigenfaces* que é baseado na transformada de Karhunen-Loève (KLT), ou PCA (Principal Component Analysis). Esta representação foi utilizada, pois segundo os autores, as imagens de faces não se encontram distribuídas de forma aleatória em um espaço vetorial de alta dimensionalidade, portanto, elas podem ser descritas de alguma forma em um espaço de dimensão menor, utilizando noções de base vetorial. De acordo com esta ideia, faz-se o uso da transformada KLT com o objetivo de achar os

autovetores que melhor descrevem a distribuição do conjunto original de imagens dentro do espaço. Tais autovetores são denominados *eigenfaces*, pois as imagens de face são tratadas como os autovetores.

Um método bastante conhecido na literatura foi proposto em (ROWLEY et al., 1998), que apresenta um sistema de detecção de faces baseado em redes neurais. Neste trabalho cada região da imagem candidata à face é submetida à equalização de histograma e a normalização da luz da cena, em seguida todos os *pixels* da região em análise são enviados para uma rede neural que decidirá se a região é uma face ou não. Atingindo uma taxa de acertos de 92%.

A metodologia proposta por (VIOLA, 2004) permite a caracterização da textura de face de forma robusta e rápida, usando características de Haar. Este nome se deve pelo fato do método ser inspirado nas *Wavelets de Haar*. Inicialmente a imagem é transformada para uma representação intermediária, chamada imagem integral e em seguida, é realizado o cálculo das características de textura usando a diferenças de intensidade de pixels dentro de regiões da imagem, atingindo uma taxa de acertos de 94%. Em (PAPAGEORIOU et al., 1998) é proposto um método para detecção de faces utilizando *Wavelets de Haar* e redes neurais obtendo taxas de acertos de 90%.

Outro tipo de técnica amplamente utilizada para extração de características, é o uso dos filtros de Gabor (Wu et al., 2010) (KAUSHAL, 2010) (ZHAO et al., 2011), que consiste em utilizar os coeficientes do resultado de vários filtros de Gabor para destacar características numa imagem, pois estes filtros conseguem uma boa descrição da estrutura de frequência espacial (repetições de níveis de cinza) e relações espaciais. Em (KAUSHAL, 2010) é descrito um método automático para detecção de faces utilizando um banco de 40 filtros de Gabor, sendo 4

frequências espaciais e 8 orientações. Em (ZHAO et al., 2011) é utilizada a extração de características faciais utilizando filtros de Gabor em conjunto da análise de componentes principais (PCA).

Podemos destacar ainda as abordagens baseadas em histogramas como podemos ver em (WARING, 2005), que utiliza histogramas espectrais para descrição de textura de faces em conjunto com o classificador Máquina de Vetores Suporte, o histograma espectral é a concatenação de vários histogramas obtidos da mesma imagem quando esta é aplicada a diferentes tipos de filtros. Em (DALAL, 2005) e (LOWE, 1999) utilizam com sucesso descritores baseados em histogramas obtidos do mapa de orientações do gradiente das imagens para o reconhecimento de objetos.

Quanto aos trabalhos que utilizam abordagens puramente estatísticas podemos citar o trabalho mais famoso proposto por Haralick (HARALICK et al., 1973). Neste trabalho é definido um método de descrição de textura utilizando matrizes de co-ocorrência. O método consiste em calcular medidas estatísticas para uma representação da imagem, a matriz de co-ocorrência, que é uma matriz onde cada elemento representa quantas vezes ocorre o par de níveis de cinza representado pela linha i e coluna j em uma determinada distância e orientação.

Em (SILVA, 2004) são utilizadas funções geoestatísticas e redes neurais de múltiplas camadas para a classificação para a classificação de nódulos pulmonares, atingindo valores acima 0.8 da área da curva ROC. Em (JÚNIOR, 2008) utiliza funções geoestatísticas e Máquinas de Vetores Suporte para a classificação de massas e não massas em imagens de mamografia, e assim avaliar quão bom é o conjunto de características geradas através de uma seleção de características utilizando o método da análise de discriminante

linear. Já (ALMEIDA, 2010) são utilizadas funções geoestatísticas para extrair características de textura de imagens de olhos e posteriormente é feita a classificação utilizando as Máquinas de Vetores Suporte, atingindo taxa de acerto de 94%.

Em (ZHENG et al., 2004) são propostas duas metodologias para a classificação de texturas de faces. Ambas metodologias são baseadas em regras de vizinhança, o primeiro caso cada imagem é amostrada num espaço de características, onde cada *pixel* é uma característica e é calculada a distância linear para cada ponto, se a distância do ponto em questão (face a ser classificada) estiver dentro de um determinado limiar, então o ponto pertence a classe face, obtendo 95.18% de acurácia. No segundo caso são utilizadas projeções para o cálculo da distância euclidiana, utilizando a relação de distância de um ponto a um plano obtendo 95.75 de acurácia.

Em (Lu et al., 2007) é apresentado um método de extração de características faciais baseado em análise de componentes principais simétricos (SPCA). Para isso são realizados testes de acordo com a quantidade de componentes principais, atingindo a melhor taxa de acertos igual a 94.75%.

Em (CORRÊA et al., 2008) é proposta uma metodologia para a classificação de faces. As imagens de faces são representadas através da análise de componentes principais (PCA). A classificação das faces é feita por um tipo de rede neural chamado de LSTM (Long-Short Term Memory). As redes LSTM seguem uma arquitetura semelhante à redes neurais de múltiplas camadas, mas ao invés de utilizarem neurônios como unidades básicas, são utilizados blocos de memória. Os blocos de memória decidem quais informações (pesos) devem ser

armazenados na estrutura. São utilizadas duas bases de imagens de faces, ORL e CBCL, os resultados obtidos para cada base são respectivamente 96% e 85%.

1.4 Organização do Trabalho

O restante da dissertação é organizada como segue: O capítulo 2 apresenta a fundamentação teórica necessária para a compreensão dos conceitos e técnicas utilizadas no presente trabalho. Serão descritos conceitos como o de textura, biodiversidade, índices de diversidade. Além destes conceitos são descritas técnicas de processamento de imagens (retinex), a técnica de aprendizado de máquina (Máquina de Vetores Suporte) e as medidas utilizadas na avaliação dos resultados obtidos nos testes realizados. O capítulo 3 descreve os passos executados pela metodologia para a classificação das texturas (pré-processamento, extração de características e classificação). No capítulo 4 são apresentados e discutidos os resultados obtidos através da metodologia proposta. O capítulo 5 apresenta a conclusão sobre este trabalho, mostrando a eficiência dos índices de diversidade aplicados na extração de características de textura e apresentando sugestões para trabalhos futuros.

Capítulo 2

Fundamentos Teóricos

Este capítulo apresenta a fundamentação teórica utilizada no desenvolvimento deste trabalho e necessária para compreensão das técnicas utilizadas para alcançar os objetivos esperados pelo mesmo. Aborda-se os seguintes índices de diversidade: índices de Magalef, Menhinick, Shannon, Pielou, Simpson e o índice de Simpson para equabilidade, MacIntosh e sua versão para equabilidade, Brillouin e Berger-Parker. Trata-se também de técnicas de processamento de imagens e finaliza com a técnica de reconhecimento de padrões, máquina de vetores suporte.

2.1 Índices de Diversidade

Diversidade de espécies no contexto da ecologia indica variedade de espécies, podendo ou não abordar informações sobre como estão distribuídas as espécies em uma ou mais comunidades de seres vivos. As informações que podem ser abordadas são: a riqueza de espécies e a equabilidade de espécies. A riqueza de

espécies refere-se ao número de espécies em uma determinada área geográfica, região ou comunidade. Quanto maior o número de espécies maior a riqueza de uma comunidade. Já a equabilidade é a abundância relativa em uma determinada comunidade, isto significa que quanto mais próximas as abundâncias das espécies dentro de uma comunidade, maior a equabilidade.

A importância em conhecer a diversidade de espécies em uma comunidade é poder compreender melhor padrões na distribuição de seres vivos em comunidades na natureza (MELO, 2008).

Existem várias ferramentas que auxiliam pesquisadores a mensurar a diversidade, em particular destacam-se a riqueza de espécies, estimadores de riqueza e índices de diversidade (MAGURRAN, 2004). A riqueza de espécies, conforme fora dito anteriormente, consiste na contagem do número de espécies que temos numa determinada comunidade ou área de interesse. Trata-se da maneira mais intuitiva de se mensurar a diversidade, e apesar de muito popular, apresenta alguns problemas, pois tem a obrigatoriedade de um levantamento completo de todas as espécies, e a medida que o número de indivíduos em uma comunidade aumenta, temos um aumento considerável no levantamento de espécies o que inviabiliza o cálculo da riqueza. Outro problema em se utilizar a riqueza de espécies são os movimentos migratórios o que faz com que a riqueza varie sazonalmente (MELO, 2008). Esta relação de dependência da quantidade de espécies levantadas dificulta a comparação de riqueza entre comunidades estudadas.

Uma solução utilizada para contornar os problemas da riqueza de espécies é usar estimadores de riqueza, também conhecidos com índices de riqueza (MAGURRAN, 2004). Estes índices consistem na relação entre o número de

espécies levantadas em uma amostra e a quantidade de indivíduos presentes na amostra. A vantagem destes índices é que, para uma mesma comunidade, eles tendem a assumir um valor constante em relação a aumento da quantidade de indivíduos na amostra. Os principais índices de riqueza são o índice de Magalef e índice de Menhinick. O índice de Magalef estima a riqueza de espécies baseando-se no logaritmo da quantidade de indivíduos, é calculado pela Equação 2.1. O índice de Menhinick utiliza a raiz quadrada da quantidade de indivíduos para estimar a riqueza (Equação 2.2) (MAGURRAN, 2004)

$$ml = \frac{S - 1}{\ln N} \quad (2.1)$$

$$D_{mn} = \frac{S - 1}{\sqrt{N}} \quad (2.2)$$

sendo S a quantidade de espécies na amostra e N é o tamanho da população.

Outra maneira de se mensurar diversidade é utilizando índices de diversidade. Estes índices consistem em expressar uma relação entre o número de espécies e a equabilidade de uma comunidade em um número (RICOTTA, 2005). Para entendimento melhor sobre os componentes de um índice de diversidade observe os seguintes exemplos. Imagine uma floresta com 4 espécies de árvores e em seguida uma outra floresta com 40 espécies. Na primeira floresta podemos facilmente encontrar e memorizar todas as espécies de árvores, enquanto na segunda floresta teríamos uma dificuldade maior. Quanto a equabilidade considere um outro exemplo com duas novas comunidades. A comunidade 1 contendo 100 espécies cada uma com 10 representantes e a comunidade 2 contendo o mesmo número de representantes, mas divididos da seguinte maneira: uma espécie tem 91

representantes enquanto que as outras espécies possuem apenas 1 representante. Podemos notar facilmente que na comunidade 1 a distribuição das espécies é igual, na comunidade 2 há um grande desbalanceamento, haja vista que a grande maioria dos indivíduos pertencem a uma espécie. Note que em ambas comunidades a riqueza de espécies é igual 10 mas a equabilidade é diferente. A Figura 2.1 ilustra bem este exemplo.

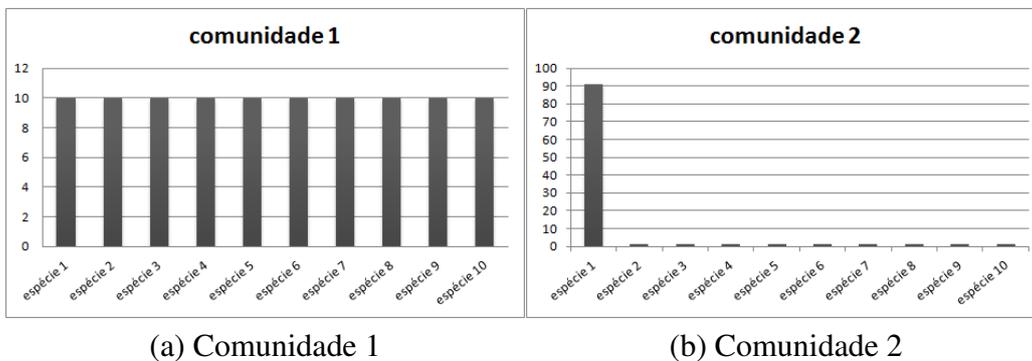


Figura 2.1: Distribuição de indivíduos nas comunidades 1 e 2

A grande vantagem em usar um índice de diversidade é concentrar numa mesma medida os atributos riqueza e equabilidade da comunidade, embora os dois componentes consistam de atributos distintos da comunidade. Outra vantagem é que os índices são estatísticas, isto significa que podemos fazer inferências sobre a comunidade baseando-se em amostras. Inclusive, alguns índices têm a propriedade de obter resultados muito próximos para diferentes esforços amostrais (MAGURRAN, 2004). Antes de usar índices de diversidade é importante ter em mente as suas desvantagens.

Uma desvantagem é que para alguns índices, apenas o seu valor não quer dizer muito sobre a comunidade em estudo, ou seja, o índice não é capaz de informar se uma comunidade é diversa ou não. Este tipo de interpretação só faz sentido

se tivermos o índice de outra comunidade para que possa haver uma comparação. Outra desvantagem é que cada índice tem uma compreensão própria dos seus valores, e que nem sempre essa compreensão é semelhante para dois índices. Por exemplo, existem situações em que um índice mostra que uma comunidade A é mais diversa que uma comunidade B, por outro lado outro índice pode mostrar o contrário, isso ocorre porque alguns índices dão maior importância a riqueza de espécies, outros dão mais importância a equabilidade (MELO, 2008). Outra desvantagem é que a maioria dos índices foram idealizados para amostras grandes, de tal forma que quando há uma amostragem muito pequena, os valores de alguns índices podem não ser confiáveis (MAGURRAN, 2004). A seguir entraremos em detalhes sobre os índices de diversidade usados neste trabalho.

2.1.1 Índice de Shannon

O índice de Shannon, também conhecido como de índice de Shannon-Wiener, foi desenvolvido pelo matemático Claude Shannon para medir a incerteza sobre espaços desordenados (SANTOS, 2009). Parte do princípio que a informação pode ser representada por símbolos e estima a entropia de uma amostra. Isto é feito multiplicando a probabilidade de se selecionar um elemento da amostra pelo logaritmo desta probabilidade, como pode ser visto na Equação 2.3. Podemos mapear estes conceitos para medir a diversidade, uma vez que esta está relacionada com a organização de indivíduos em espécies. Se tratarmos as espécies como símbolos e considerarmos que um elemento da amostra passa a ser um indivíduo de uma comunidade, podemos utilizar este índice para mensurar

diversidade.

$$H' = - \sum_{i=1}^S p_i \ln p_i \quad (2.3)$$

onde H' é o índice de Shannon, S é o número de espécies presentes na amostra, p_i é a probabilidade de um indivíduo pertencer à espécie i , e é calculada pela Equação 2.4:

$$p_i = \frac{n_i}{N} \quad (2.4)$$

onde p_i é a probabilidade de selecionar um indivíduo da espécie i , n_i é o número de indivíduos da espécie i e N é o total de indivíduos na amostra.

O índice de Shannon é considerado o índice de diversidade mais completo (SANTOS, 2009), pois dá o mesmo peso a espécies raras como espécies de maior abundância. Em geral o valor do índice não ultrapassa 5.0, e quando ocorre, somente em situações que o número de indivíduos na amostra é superior a 10^5 (MAGURRAN, 2004). Apenas com o valor do índice calculado para uma comunidade não é possível concluir que a comunidade tem alta diversidade ou não. Somente podemos ter esse tipo de conclusão quando comparamos o valor do índice de Shannon com o de outra comunidade, ambas com mesmo tamanho da amostra. Só assim podemos afirmar que a comunidade mais diversa é a que apresentar o maior valor para o índice de Shannon. O seguinte exemplo ilustra esta situação. Considere duas comunidades A e B, ambas com mesma quantidade de indivíduos. Na comunidade A os indivíduos estão divididos igualmente em 5 espécies, já a comunidade B os indivíduos estão divididos da seguinte maneira: a espécie 1 possui 80 indivíduos, a espécie 2 possui 10 indivíduos, a espécie 3 possui 5 indivíduos, a espécie 4 possui 4 indivíduos e a espécie 5 possui apenas 1 indivíduo. Nestas condições os valores encontrados para o índice de Shannon nas

comunidades A e B são respectivamente: 1.609437912 e 0.733366699. Podemos notar que para comunidades com mesmo tamanho, quanto maior o valor do índice, mais semelhante será a distribuição de indivíduos por espécie. Podemos ainda afirmar que quando a distribuição de indivíduos é igual para todas as espécies (equabilidade máxima) o índice se torna igual a $\ln S$. Voltando ao exemplo, verificamos que $\ln 5 = 1.609437912$, portanto, confirmando a observação de que na comunidade A a equabilidade é máxima. A Figura 2.2 nos dá uma boa ilustração do exemplo dado.

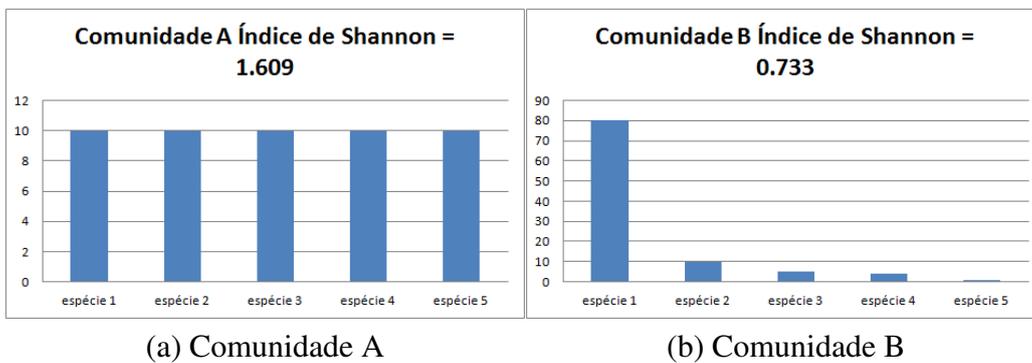


Figura 2.2: Índice de Shannon para comunidades A e B

2.1.2 Índice de Pielou

O índice de Pielou é uma medida de equabilidade que baseia-se no índice de diversidade de Shannon. Consiste em dividir o resultado do índice de diversidade de Shannon pelo logaritmo natural do número de espécies, que corresponde ao índice de Shannon quando a equabilidade é máxima. Desta forma, temos uma medida que independe da riqueza de espécies, pois o valor deste índice não se altera se incluirmos ou removermos espécies da amostra, desde que as relações

entre espécies e indivíduos sejam mantidas. Outra característica interessante do índice de Pielou, é que o seu valor por si só exprime o grau de equabilidade de uma comunidade, uma vez que 1.0 corresponde a equabilidade máxima. O índice de Pielou que é dado por:

$$J' = \frac{H'}{\ln S} \quad (2.5)$$

onde J' é o índice de Pielou, H' é o índice de Shannon e S é o número de espécies na amostra. Quando calculamos o índice de Pielou para as comunidades A e B do exemplo da Subseção 2.1.1 temos 1.0 para comunidade A e 0.455 para comunidade B.

2.1.3 Índice de Simpson

É um índice que estima a probabilidade de dois indivíduos de uma mesma espécie serem sorteados a partir da mesma amostra. Foi o primeiro índice a ser utilizado em estudos ecológicos e exprime dominância de alguma espécie na comunidade. Dominância é quando há uma ou poucas espécies que detém a maior parte de indivíduos na comunidade. O índice de Simpson é dado por:

$$D = \sum_{i=1}^S p_i^2 \quad (2.6)$$

onde p_i é obtido pela Equação 2.4.

Nesta forma o índice de Simpson é considerado um estimador tendencioso, pois o cálculo da probabilidade de sortear dois indivíduos da mesma espécie é feito considerando reposição na amostra, e é indicado para comunidades infinitamente grandes (MAGURRAN, 2004).

A relação mais adequada para comunidades finitas consiste em calcular a probabilidade de sortear dois indivíduos da mesma espécie é feito sem reposição na amostra que é dado por:

$$D = \sum_{i=1}^S \frac{n_i(n_i - 1)}{N(N - 1)} \quad (2.7)$$

sendo n_i a quantidade de indivíduos da espécie i e N é o tamanho da amostra. Ao contrário do índice de Shannon, o índice de Simpson dá maior peso às espécies comuns (MELO, 2008). Isto significa que um pequeno aumento na quantidade de indivíduos de uma espécie rara provoca pequenas alterações no valor do índice, no caso do índice de Shannon qualquer alteração na relação entre indivíduos e espécies ocorre uma variação considerável no valor obtido pelo índice. Voltando ao exemplo da Subseção 2.1.1, o cálculo do índice de simpson para ambas comunidades nos dá os seguintes resultados. A comunidade A obteve 0.2 e a comunidade B obteve 0.6542. Vemos que ao contrário do índice de Shannon que indica o aumento da diversidade conforme o aumento do valor do índice, o índice de Simpson indica aumento de diversidade conforme o valor do índice diminui (a comunidade A é mais diversa que a B), por esta razão alguns pesquisadores utilizam as versões: complemento ($D - 1$) e recíproco ($1/D$) (KANIESKI, 2010), para facilitar a interpretação do índice.

Também existe uma variação do índice de Simpson para mensurar a equabilidade. Isto é feito dividindo a versão recíproca do índice de Simpson pela quantidade de espécies na amostra, conforme é visto na Equação 2.8.

$$E_{1/D} = \frac{S}{D} \quad (2.8)$$

onde D é o índice de Simpson e S é o número de espécies na amostra. A faixa de valores compreendidos por este índice é de 0 a 1, sendo que 1 indica equabilidade máxima e 0 equabilidade mínima, embora em situações reais o valor 0 não é atingido (MAGURRAN, 2004). Considerando as comunidades descritas no exemplo da Subseção 2.1.1, os valores obtidos para o índice de Simpson para as comunidades A e B são respectivamente 1.0 e 0.305716906. Lembrando que na comunidade A a equabilidade é máxima.

2.1.4 Índice de Brillouin

O índice de diversidade de Brillouin é o índice mais apropriado para utilizar quando temos uma comunidade completamente inventariada, ou seja, todos indivíduos da comunidade estão identificados (MAGURRAN, 2004). Por esse motivo este índice se mostra muito sensível a variações do tamanho da população. O índice de Brillouin é calculado pela Equação 2.9:

$$HB = \frac{\ln N! - \sum_{i=1}^S \ln n_i!}{N} \quad (2.9)$$

onde HB é o índice de Brillouin, N é o tamanho da população e n_i é o número de indivíduos da espécie i . Quando utilizado para comunidades com mais de 69 indivíduos o cálculo do fatorial se torna intratável para calculadoras e microcomputadores (Dutra, 1995). Para isso utilizamos a aproximação de Stirling para $n!$ que é calculada da seguinte maneira:

$$n! = \sqrt{2\pi n} n^n e^{-n} \quad (2.10)$$

2.1.5 Índice de MacIntosh

O índice de McIntosh é um índice que parte do princípio que uma comunidade pode ser encarada como um ponto em um hipervolume S-dimensional, e que a distância euclidiana do ponto em questão em relação à origem pode ser usada como medida de diversidade (KANIESKI, 2010). A distância é representada por U (Equação 2.11).

$$U = \sqrt{\sum_{i=1}^S n_i^2} \quad (2.11)$$

onde, S é o número de espécies, n_i é o número de indivíduos da espécie i . Para calcular o índice usamos a Equação 2.12. Nesta forma o índice de MacIntosh mensura se há dominância em uma comunidade. A Equação 2.13 apresenta a sua variação para cálculo de equabilidade.

$$U_D = \frac{N - U}{N - \frac{N}{\sqrt{S}}} \quad (2.12)$$

$$U_E = \frac{N - U}{N - \sqrt{N}} \quad (2.13)$$

2.1.6 Índice de Berger-Parker

Trata-se do índice mais simples e de fácil interpretação, pois verifica se há dominância na comunidade apenas pela probabilidade de selecionar o indivíduo da espécie com mais representantes na amostra, sendo que 1 indica dominância total (apenas uma espécie na comunidade). O cálculo é feito conforme a Equação 2.14:

$$d = \frac{N_{max}}{N} \quad (2.14)$$

onde d é o índice de Berger-Parker, N_{max} é o número de indivíduos da espécie mais abundante e N é o tamanho da população.

2.2 Processamento de Imagens Digitais

O processamento de imagens é a atividade de tratar dados visuais no computador, com a finalidade de melhorá-los ou apenas resolver algum problema de modo a permitir que o usuário possa ter acesso a algum tipo de informação que, a princípio, ele não tinha.

O processamento de imagens para fins de didática pode ser dividido em cinco etapas principais: aquisição, melhoramento, segmentação, representação e descrição, reconhecimento e interpretação (WOODS, 1992). A Figura 2.3 ilustra estas etapas e como elas estão relacionadas. Na aquisição utiliza-se algum aparato

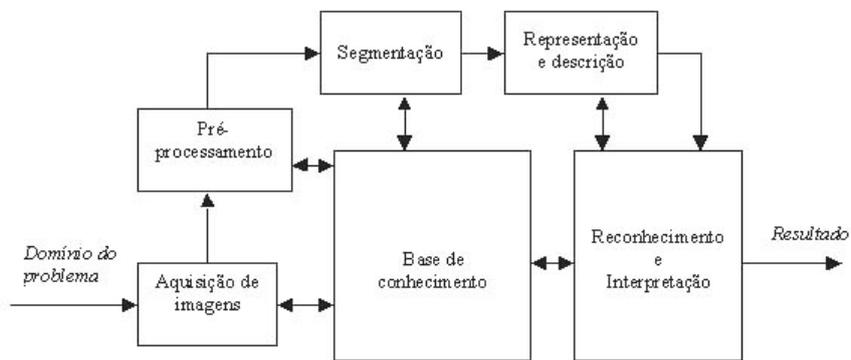


Figura 2.3: Etapas do Processamento de Imagens Digitais. Adaptado de (GONZALEZ and WOODS 1992)

tecnológico para gerar as imagens que se deseja processar. As imagens podem ser obtidas tanto através de equipamentos de captura como câmeras e radares quanto através de simulações por computador.

O melhoramento tem a finalidade de aumentar a qualidade da imagem através da redução de ruídos. Esta etapa é importante porque aumenta a eficiência das etapas posteriores. A segmentação é um passo extremamente importante, pois permite isolar o objeto de estudo, onde os seus resultados são muito importantes na determinação de eventual sucesso ou falha na análise da imagem.

A representação tem a finalidade de extrair da região segmentada um conjunto descritivo de características mensuráveis. Estas características variam muito de acordo com o que se pretende extrair da imagem, mas podem incluir perímetro, cor dos *pixels*, geometria, etc.

Por último, na classificação, as características obtidas na etapa de representação são analisadas para se chegar a alguma conclusão sobre a imagem. Nesta etapa pode-se identificar objetos, agrupá-los em categorias. A seguir veremos uma técnica de melhoramento de imagens com o objetivo de normalizar condições de iluminação de imagens, pois em algumas imagens existem grandes variações nas condições de iluminação.

2.2.1 Retinex

As diferentes condições de iluminação de uma cena podem implicar o surgimento de problemas no processo de imageamento da cena, tais como falsas bordas geradas por sombras, produção de imagens muito escuras ou muito brilhantes. O sistema visual humano consegue distinguir objetos com certa facilidade para estes tipos de cenários (GILCHRIST et al., 1999), mas estas condições de iluminação da cena causam grandes influências nos algoritmos de visão computacional (RANGEL, 2009). Tal fato inspirou o desenvolvimento

de algoritmos que visam dar uma constância cromática a imagens, uma técnica amplamente utilizada na literatura que visa melhorar as condições de iluminação é o retinex, conforme veremos a seguir.

O retinex é uma representação do modelo da visão humana que realiza a normalização da iluminação de uma cena, para isso assume que as variações espaciais pequenas nas imagens são consequências da iluminação do objeto, enquanto que variações grandes são decorrentes da própria superfície do objeto. Inicialmente o retinex foi proposto por (LAND, 1971), onde a modelagem do sistema visual humano é composto de sistemas retina-córtex, daí o nome retinex. Cada sistema capta informações de baixa, alta e média frequência da faixa visível do espectro eletromagnético. A partir desta modelagem foram construídos vários algoritmos para implementar o retinex (HEUSCH et al., 2005). Uma versão muito utilizada na literatura foi proposta por (JOBSON et al., 1997) por ser simples e eficiente.

O algoritmo proposto em (JOBSON et al., 1997) consiste em estimar a iluminação baseando-se na vizinhança e em seguida realizar correções nas imagens. A correção da imagem, para um determinado *pixel*, é então, feita com base no intervalo de valor máximo e mínimo em uma escala logarítmica. Existem diferentes versões para o retinex, a que será apresentada neste trabalho é a versão para apenas uma escala do retinex chamado *Single Scale Retinex (SSR)*.

O SSR proposto por (JOBSON et al., 1997) define uma implementação prática do retinex que é realizado em duas etapas, o cálculo da contribuição da iluminação por *pixel* (apresentado na Equação 2.15) e em seguida realiza o cálculo da correção

da iluminação chamada *gain/offset* (Equação 2.16).

$$R(x,y) = \ln I(x,y) - \ln(G(x,y) \otimes I(x,y)) \quad (2.15)$$

sendo $R(x,y)$ o retinex em única escala (SSR), $I(x,y)$ é a imagem original e $G(x,y)$ é uma função chamada *surround*. Esta função considera a contribuição dos vizinhos na iluminação de um *pixel* da imagem. Segundo (RAHMAN et al., 2011) a função gaussiana é mais adequada. Em seguida é calculada a correção *gain/offset*, este é um método que visa ajustar o brilho da imagem para que ele fique na faixa de $R(x,y)_{min}$ e $R(x,y)_{max}$ conforme é mostrado na Equação 2.16.

$$I'(x,y) = \frac{d_{max}}{R(x,y)_{max} - R(x,y)_{min}} \cdot (I(x,y) - R(x,y)_{min}) \quad (2.16)$$

onde $I'(x,y)$ é a imagem com o brilho ajustado, d_{max} é a faixa dinâmica de visualização, normalmente seu valor é de 255. A Figura 2.4 mostra um exemplo da aplicação do retinex em uma imagem.

2.3 Reconhecimento de Padrões

As técnicas de Reconhecimento de Padrões (RP) classificam informações (padrões) baseado ou em conhecimento prévio ou em informações estatísticas extraídas dos padrões. Um padrão é tudo aquilo para o qual existe uma entidade nomeável representante, geralmente, criada através do conhecimento cultural humano (ALMEIDA, 2010). O objetivo do RP é a classificação (agrupamento das amostras em classes) e o reconhecimento de objetos desconhecidos como pertencente a uma das classes criadas.

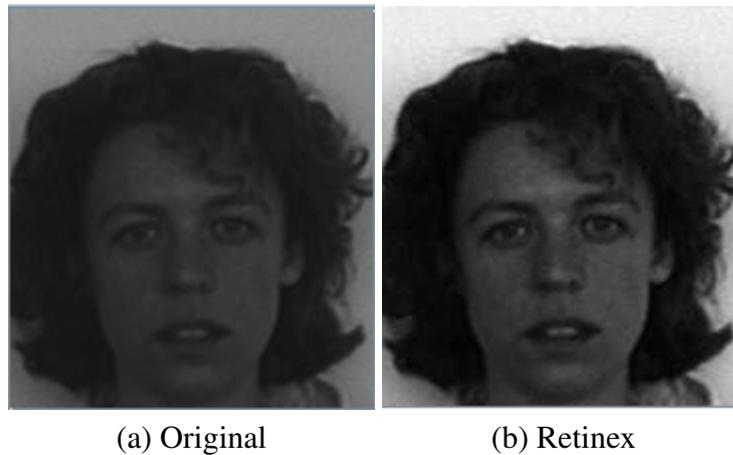


Figura 2.4: Aplicação do Retinex

Os sistemas de RP são formados por um sensor que captam observações a serem classificadas ou descritas; um mecanismo de extração de características que discretiza as observações captadas pelo sensor para formar o vetor de características; e por último um mecanismo que rotula as amostras observadas, essa rotulagem pode ser feita de duas maneiras: supervisionada e não supervisionada. Na abordagem supervisionada um supervisor apresenta um conjunto de dados já rotulados, chamado conjunto de treinamento, para que o padrão seja aprendido. Na abordagem não supervisionada o próprio mecanismo de classificação agrupa os dados e consecutivamente rotula os agrupamentos.

Um dos pontos cruciais de um sistema RP é a escolha de características. As características devem conter a maior parte da informação relevante e a dimensão do conjunto de características não deve ser muito grande. Para isso, é importante utilizar um método que diminua a quantidade de características, principalmente as que não acrescentam muita informação a cerca de um padrão.

A seguir veremos um método de classificação supervisionado amplamente utilizado na literatura pela sua robustez e rapidez, chamado Máquina de Vetores Suporte.

2.3.1 Máquinas de Vetores Suporte

A Máquina de Vetor Suporte (MVS) é um método de aprendizagem de máquina supervisionado usado para estimar uma função com o objetivo de classificar um conjunto de dados em classes. Um método de aprendizado supervisionado significa que a partir de um conjunto de casos previamente rotulados, chamado conjunto de treinamento, o método é capaz de aprender a reconhecer um padrão (LORENA, 2003). A MVS é fortemente baseada na teoria da otimização matemática, já que o seu funcionamento consiste em maximizar a distância entre pontos que representam as duas classes e uma superfície de separação, chamada de hiperplano. As restrições do problema são representadas pelos casos do conjunto de treinamento. A MVS foi originalmente desenvolvida para lidar com conjuntos de dados linearmente separáveis. A classificação de conjuntos linearmente separáveis consiste em determinar uma função $f : x \subseteq \mathfrak{R}^n \rightarrow \mathfrak{R}$, onde uma dada entrada $x = \{x_1, x_2, \dots, x_n\}^T$ é considerada parte da classe positiva se $f(x) > 0$ e negativa em caso contrário, se $f(x)$ é uma função linear de $x \in X$ então $f(x)$ pode ser escrita como:

$$f(x) = \langle w \cdot x \rangle + b = \sum_{i=1}^n w_i x_i + b \quad (2.17)$$

sendo $(w, b) \in \mathfrak{R}^n \times \mathfrak{R}$ os parâmetros que controlam a função e a regra de decisão que é geralmente dada pelo sinal de $f(x)$, convencionamos que o $sign(0) = 1$. Os

parâmetros w_i são aprendidos por um algoritmo de aprendizado a partir dos x_i . Geometricamente esta hipótese pode ser interpretada como um hiperplano, dado pela Equação $\langle w \cdot x \rangle + b = 0$ que separa o espaço X , as entradas, em duas partes conforme pode ser visto na Figura 2.5. Um hiperplano é um subespaço de dimensão $n-1$ que divide o espaço de entradas X em duas metades. No espaço de dimensão 2 o hiperplano é uma reta

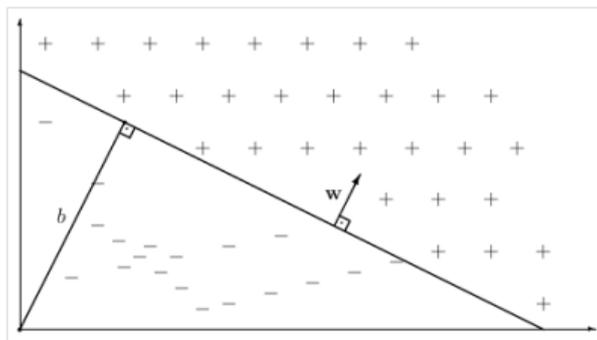


Figura 2.5: Hiperplano separando o espaço em duas regiões.

Como exemplo pode se observar na Figura 2.5, que a linha diagonal corresponde ao hiperplano com a classe positiva acima representada por uma cruz e a classe negativa abaixo denotada por um traço. O vetor w define a direção do hiperplano, enquanto que a variação do valor de b faz com que o hiperplano mova-se paralelamente a ele mesmo. Os valores w e b são chamados de peso e tendência, respectivamente. Basicamente a MVS utiliza um algoritmo que busca encontrar o hiperplano h com uma margem máxima de separação entre os dados de classes distintas, representados no espaço vetorial $X \in \mathfrak{R}^n$. Um hiperplano é denominado de margem máxima ou separação ótima caso separe um conjunto de vetores de classes distintas e a distância entre os vetores mais próximos ao hiperplano seja máxima (SMOLA, 2000). A margem máxima é definida como a distância entre o

hiperplano e os vetores que estão mais próximos a ele (Figura 2.6), esses vetores são chamados de vetores suporte, de acordo com (LIMA, 2004) os vetores suporte são padrões, que sozinhos determinam o hiperplano ótimo.

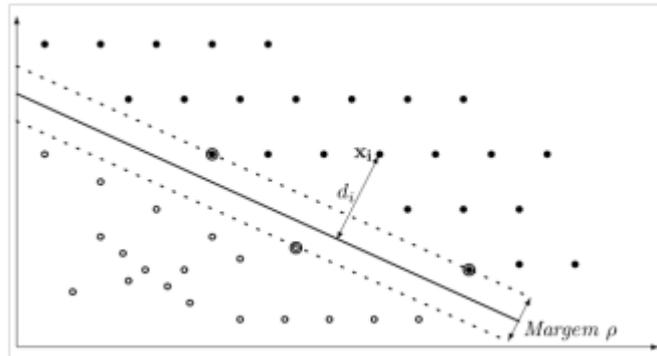


Figura 2.6: Hiperplano separando o espaço em duas regiões.

Podemos definir a margem máxima ρ pela Equação:

$$\rho = \min_{1, \dots, l} y_i f(x_i) \tag{2.18}$$

A MVS utiliza um algoritmo de aprendizado supervisionado, chamado de vetores suporte para encontrar o hiperplano, isto significa que é necessária a existência de uma seqüência de casos de treino os quais contém um rótulo indicando a classe da amostra pré-classificada. O objetivo após o treinamento é classificar novas amostras ainda não rotuladas. No caso linearmente separável, o algoritmo de vetor suporte que procura o hiperplano de separação com a maior margem, para isso os casos de treino devem satisfazer as restrições:

$$\langle w \cdot x_i \rangle + b \geq 1, \text{ para } y_i = 1 \tag{2.19}$$

$$\langle w \cdot x_i \rangle + b \leq -1, \text{ para } y_i = -1 \tag{2.20}$$

Estas restrições indicam as bordas da margem de separação para cada classe, para classe positivo ($y_i = 1$) e para classe negativa temos ($y_i = -1$). Combinando as duas restrições acima para o cálculo da margem temos a seguinte inequação:

$$y_i(\langle w \cdot x_i \rangle + b) \geq 1, \text{ para } i = 1, \dots, l \quad (2.21)$$

Seja $d_+(d_-)$ a distância euclidiana entre os vetores suporte positivos (negativos) e o hiperplano, definimos como margem ρ de um hiperplano de separação como sendo a maior margem geométrica entre todos os hiperplanos, podemos representar por $\rho = d_+ + (d_-)$. Denotaremos por $d_i(w, b; x_i)$, a distância de um dado x_i ao hiperplano (w, b) e é calculada pela Equação 2.22 (LIMA, 2004).

$$d_i(w, b; x_i) = \frac{|(\langle w \cdot x_i \rangle + b)|}{\|w\|} = \frac{y_i(\langle w \cdot x_i \rangle + b)}{\|w\|} \quad (2.22)$$

Levando em consideração a restrição imposta pela Equação 2.19, podemos escrever:

$$d_i(w, b; x_i) \geq \frac{1}{\|w\|} \quad (2.23)$$

Com isso podemos identificar $\frac{1}{\|w\|}$ como o limite inferior da distância entre os vetores suporte x_i e o hiperplano de separação (w, b) (Figura 2.7), as distâncias d_+ e d_- tornam-se:

$$d_+ = d_- = \frac{1}{w} \quad (2.24)$$

A minimização de $\|w\|$ leva a maximização da margem ρ . A partir disto podemos definir a margem ρ através da Equação 2.25.

$$\rho = d_+ + (d_-) = \frac{2}{w} \tag{2.25}$$

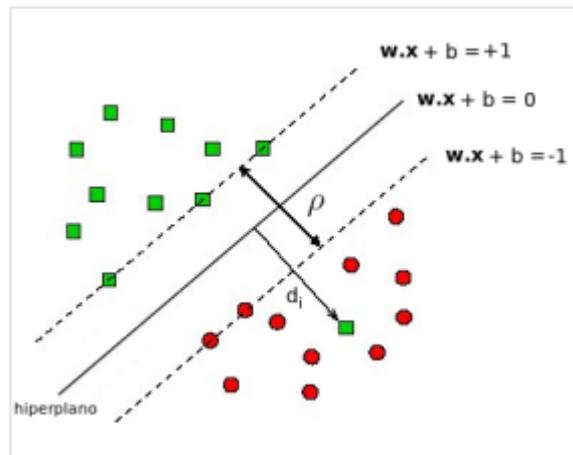


Figura 2.7: Distância entre hiperplanos e vetores suporte.

O hiperplano ótimo é dado pela minimização da norma $\|w\|$, considerando a restrição da Equação 2.19. Como pode ser observado, o problema acima se trata de um problema de otimização, este problema pode ser resolvido com o método de multiplicadores de Lagrange (LORENA, 2003). Através da teoria dos multiplicadores de Lagrange, chega-se à Equação 2.26. O objetivo então passa a ser encontrar os multiplicadores de Lagrange α_i ótimos que satisfaçam a Equação 2.26 (BURGES, 1998).

$$w(\alpha) = \sum_{i=1}^N \alpha_i \alpha_j y_i y_j (x_i x_j) \tag{2.26}$$

Apenas os pontos em que a Equação 2.19 é exatamente igual à unidade têm correspondentes $\alpha_i \neq 0$. Esses pontos são exatamente os vetores de suporte. Da forma que foi descrita acima, a MVS trabalhava apenas com conjuntos de

dados linearmente separáveis, para que a MVS possa classificar amostras que não são linearmente separáveis, é necessário aplicar uma transformação não-linear que transforme o espaço entrada para um novo espaço chamado espaço de características. Um conjunto de dados é dito não-linearmente separável, caso não seja possível separar os dados com um hiperplano. A Figura 2.8 mostra um conjunto linearmente separável e outro não - linearmente separável. O espaço de

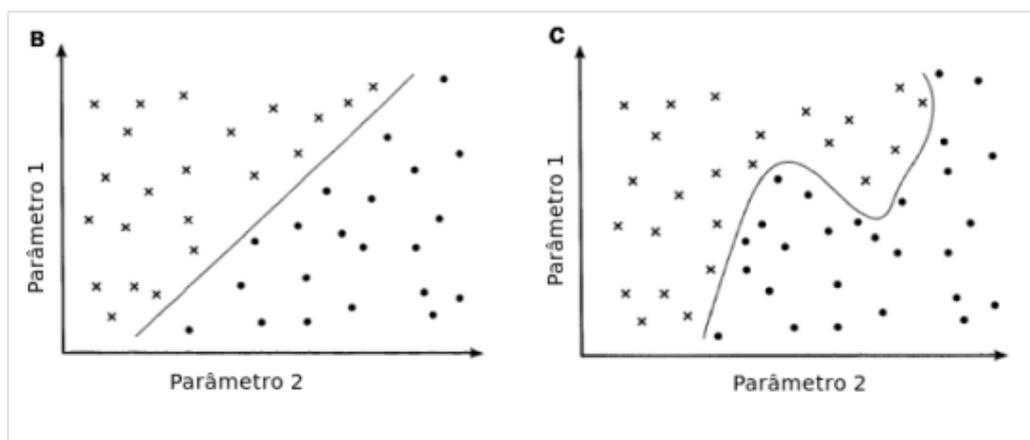


Figura 2.8: Conjuntos linearmente separáveis e não linearmente separáveis.

características deve apresentar dimensão suficientemente grande (SMOLA, 2000), e através dele, a amostra pode ser linearmente separável. Uma vez realizada a mudança do espaço de entrada para o espaço de características é possível definir um hiperplano de separação através de uma função linear. Para a construção do hiperplano no espaço de características é utilizada uma função K chamada de núcleo (*kernel*). Uma função núcleo calcula o produto interno de dois vetores do espaço de entrada (x_i e x_j) para o espaço de características. A função núcleo é dada por:

$$K(x_i, x_j) = (\langle \Phi(x_i) \cdot \Phi(x_j) \rangle) \quad (2.27)$$

É necessário que a função $\Phi(\cdot)$ pertença a um domínio que seja possível o cálculo do produto interno. A função K pode realizar o mapeamento das amostras para um espaço de dimensão muito elevada sem aumentar a complexidade dos cálculos. Uma importante família de funções de núcleo é a função de base radial, muito utilizada em problemas de reconhecimento de padrões e também utilizada neste trabalho. A função de base radial é definida por:

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|}{2\gamma}} \quad (2.28)$$

Neste trabalho a MVS será utilizada para a classificação de objetos faciais a partir das características de textura de imagens obtidas pelos índices de diversidade.

2.4 Validação de Resultados

A avaliação de um algoritmo de aprendizado de máquina supervisionado é normalmente realizada por meio da análise do desempenho do classificador durante a predição de novos dados, não apresentados durante a fase de treinamento (LORENA, 2003). Uma técnica muito utilizada para análise da precisão de um classificador é a matriz de confusão ou matriz de erro. Nela os dados são dispostos de forma tabular, onde as linhas representam a quantidade de amostras de uma categoria que são inferidas por um classificador e as colunas expressam dados de referência para que possa ser comparada com os valores obtidos na classificação. A diagonal principal indica a quantidade de amostras que foram classificadas corretamente conforme pode ser visto na Figura 2.9.

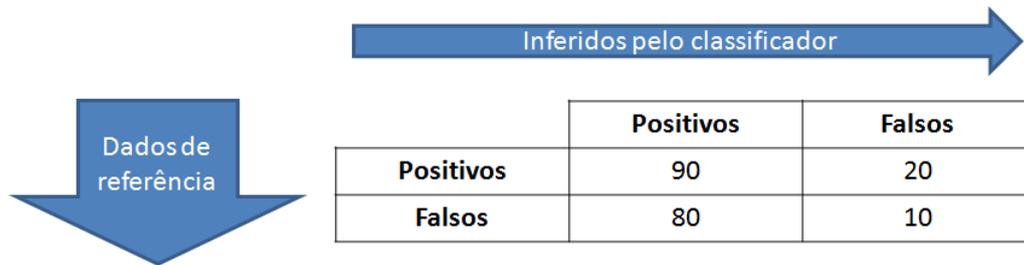


Figura 2.9: Formato da matriz de confusão.

A partir da matriz de confusão podemos extrair medidas para quantificar a precisão do classificador. Se considerarmos apenas duas categorias para os dados (positivo e negativo) temos as seguintes medidas: a quantidade de verdadeiros positivos TP, ou seja, a quantidade de amostras pertencentes a classe dos positivos e foram classificadas como positivas. A quantidade de verdadeiros negativos TN são as amostras pertencentes a classe dos negativos que foram classificados corretamente, FP é a quantidade de amostras pertencentes a classe dos negativos que foram classificadas como positivos e FN é a quantidade de falsos negativos, casos positivos que foram classificados como negativos. Com base nesses valores podemos extrair as seguintes medidas: acurácia, sensibilidade e taxa de falsos positivos. A acurácia mede a quantidade de acertos global, levando-se em conta tanto os acertos de verdadeiros negativos e os verdadeiros positivos e é dada pela Equação 2.29.

$$ac = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{2.29}$$

A sensibilidade é uma medida que quantifica os acertos para a classe dos positivos e é calculada pela Equação 2.30.

$$sen = \frac{TP}{(TP + FN)} \tag{2.30}$$

A taxa de falsos positivos mede quantas amostras contidas no teste foram classificadas erroneamente como pertencentes à classe dos positivos que é dada por:

$$tfp = \frac{FP}{(TN + FN)} \quad (2.31)$$

Neste trabalho a matriz de confusão será utilizada para mensurar o desempenho da MVS treinada a partir dos índices de diversidade. Para isto será adotada a seguinte convenção: imagens de faces serão consideradas representantes da classe dos positivos e não-faces serão consideradas representantes da classe dos negativos.

Capítulo 3

Materiais e Métodos

Este capítulo descreve os procedimentos e materiais utilizados para a avaliação dos índices de diversidade, apresentados no capítulo 2, como descritores de textura de imagens de faces. Inicialmente é apresentada a infraestrutura computacional utilizada nos experimentos, em seguida apresenta-se a base de imagens utilizadas nos testes e finalmente, é descrita a sequência de etapas realizadas para alcançar os objetivos da metodologia.

3.1 Software Utilizados

A linguagem JAVA é atualmente uma das linguagens orientadas a objetos mais utilizadas atualmente (JAVA, 2011). A sua importância no cenário do desenvolvimento de *software* é justificada por várias características, dentre elas estão: código livre, padronização, portabilidade e uma série de recursos disponíveis através de bibliotecas.

Todos os índices utilizados neste trabalho foram implementados seguindo a especificação da linguagem JAVA versão 1.6.0 (JAVA, 2011) e a IDE (*Integrated Development Environment*) *freeware* de desenvolvimento denominada Eclipse versão 3.6 (ECLIPSE, 2011).

Os métodos de manipulação das imagens foram implementados por meio da biblioteca de código *Java Advanced Imaging* (JAI) (JAI, 2011), disponível gratuitamente na *web*. A biblioteca é totalmente baseada nos princípios da programação orientada a objetos e é compatível com a tecnologia Java. A JAI fornece uma gama de funções que vão desde operações básicas como carregar uma imagem do sistema de arquivos do sistema operacional, salvar e visualizar imagens, até funções mais sofisticadas como filtragens, operações no domínio da frequência etc.

O classificador MVS foi utilizado através da suíte LIBSVM (CHANG et al., 2011). Esta suíte disponibiliza uma interface para plataforma cruzada, isto é, pode ser utilizado para mais de uma linguagem de programação e mais alguns utilitários que auxiliam no uso da MVS, entre estes utilitários destacamos o *grid*, que é uma aplicação que estima valores ótimos para os parâmetros da MVS.

A análise dos resultados obtidos da metodologia ficou por conta do ambiente WEKA (HALL et al., 2009). O WEKA é um ambiente que disponibiliza várias ferramentas de reconhecimento de padrões como classificadores, algoritmos de agrupamento, métodos de seleção de características.

3.2 Base de Imagens Utilizadas

Foram utilizadas 4 bases de faces e 1 base de não faces para a validação da metodologia proposta neste trabalho. Cada base de faces tem suas peculiaridades como tamanho, condições de iluminação, presença de óculos, barba, etc. Os exemplos de imagens correspondentes a não faces são necessárias para que a MVS possa aprender os padrões da classe positivos (faces) e negativos (não faces). As imagens de exemplos de não faces foram obtidas da web totalizando 2000 imagens que não continham faces. A Figura 3.1 mostra alguns exemplos de imagens de não faces.

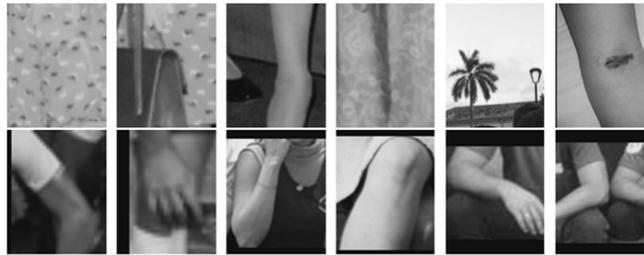


Figura 3.1: Exemplos de imagens de não faces utilizados.

As bases de faces foram as seguintes: CBCL, ORL, Libor Spacek e CMU AMP Face Expression. A base ORL (CAMBRIDGE, 2011) contém imagens de quarenta indivíduos, sendo dez imagens para cada indivíduo. Essas imagens foram obtidas sob diferentes expressões faciais (olhos fechados ou abertos, sorrindo ou sério) e presença de detalhes faciais como óculos, barba e bigode. As imagens foram tiradas em variações de inclinação e rotação da face com uma tolerância de 20 graus, também há uma variação de 10% da distância do indivíduo em relação a câmera e todas as imagens tem uma resolução de 92×112 pixels. A Figura 3.2 mostra exemplos de imagens da base ORL.

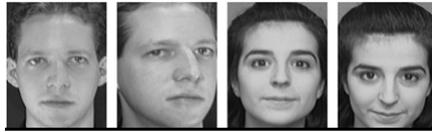


Figura 3.2: Exemplos de imagens da base ORL.

A base CBCL original é formada de 2429 faces (SUNG, 1996). As imagens foram obtidas para faces frontais a câmera, sob diferentes expressões faciais e com presença ou não de óculos, barba, bigode, etc. Um detalhe a se ressaltar é que as imagens foram obtidas a diferentes condições de iluminação e todas estão numa resolução de 20×20 *pixels*. Neste trabalho foram utilizadas apenas as imagens de faces do treinamento (SUNG, 1996), pois todas as outras bases foram testadas com o mesmo conjunto de não faces. A Figura 3.3 mostra alguns exemplos da base CBCL.

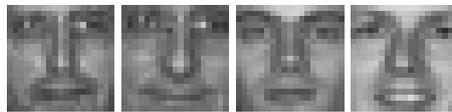


Figura 3.3: Exemplos de imagens de face da base CBCL.

A base Libor Spacek (SPACEK, 2011) consiste de 20 imagens de cada um dos 150 indivíduos, sob diferentes condições de iluminação, variações de expressões faciais e presença de barba e óculos. As imagens estão em uma resolução de 128×128 *pixels* e foram obtidas com os indivíduos na pose frontal. A Figura 3.4 mostra alguns exemplos de imagens da base Libor Spacek.

A base CMU AMP Face Expression (NAGESH, 2009) contém 975 imagens (13 indivíduos cada um com 75 imagens). Obtidas sob mesmas condições de iluminação, escala e pose (frontal). Todas as imagens estão em uma resolução



Figura 3.4: Exemplos de imagens de face da base Libor Spacek.

de 64x64 *pixels*, o único detalhe facial em alguns indivíduos é a presença de óculos. A Figura 3.5 mostra alguns exemplos de imagens da base CMU AMP Face Expression. A Tabela 3.1 resume a quantidade de imagens utilizadas em cada base de faces.



Figura 3.5: Exemplos de imagens de faces da base Faces Expression.

Base de Imagens	número de imagens
CBCL	2429
Libor Spacek	3000
ORL	400
CMU AMP Face Expression	975

Tabela 3.1: Quantidade de imagens por base

3.3 Metodologia Proposta

Esta seção descreve os passos realizados para a avaliação dos índices de diversidade como descritores de textura de faces. Os passos adotados foram: o pré-processamento, extração de características e a classificação dos dados. A Figura 3.6 de etapas realizadas na metodologia proposta, nas subseções a seguir será feito um detalhamento de cada uma dessas etapas.

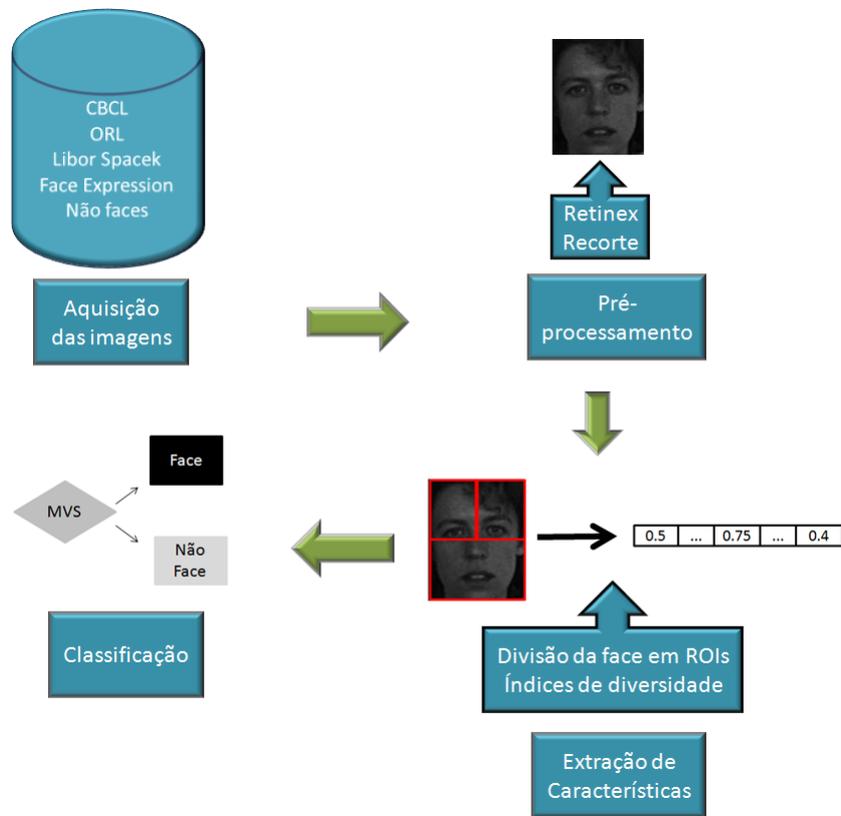


Figura 3.6: Fluxo de atividades da metodologia proposta.

3.3.1 Pré-processamento

A metodologia se inicia com a aquisição de uma imagem que pode ser uma face ou não face. Em seguida a imagem é convertida para níveis de cinza, caso a mesma seja colorida, pois é mais fácil e eficiente trabalhar com apenas um canal de cor (ALMEIDA, 2010). A conversão é feita segundo o cálculo da componente Y do modelo YCbCr (SOUTO, 2003), representado pela Equação 3.1:

$$Y = 0.299R + 0.587G + 0.114B \quad (3.1)$$

onde Y é a luminância do sistema YCbCr e R , G , B são respectivamente as componentes vermelho, verde, azul do modelo RGB. Após a conversão para níveis de cinza é aplicada a técnica retinex (Seção 2.2.1) em cada imagem das bases. Em seguida para garantir a uniformidade no tamanho de cada imagem é feito um redimensionamento na imagem para um tamanho fixo de 30×30 pixels, obtido empiricamente após várias definições de tamanho de imagens, em seguida é feito um recorte manual em cada imagem redimensionada com o intuito de remover regiões de fundo da imagem, para que na imagem contenha somente a face, conforme pode ser visto na Figura 3.7.

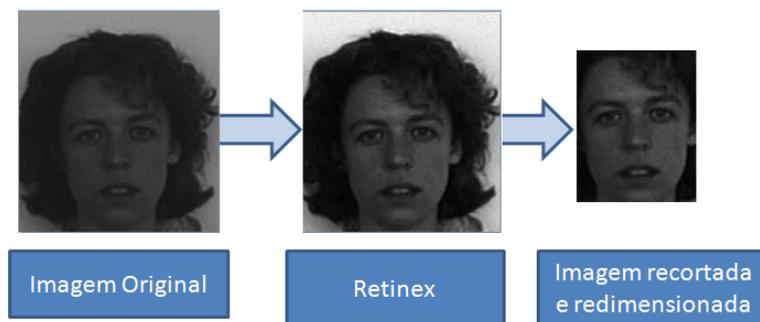


Figura 3.7: Fluxo de atividades durante o pré-processamento.

Concluída essa fase de pré-processamento serão obtidos valores que caracterizam a textura da imagem que será realizada através dos índices de diversidade.

3.3.2 Extração de Características

Nesta etapa serão obtidos valores que caracterizam a textura da face humana, os índices de diversidade apresentados na Seção 2.1. Mas para a utilização destes índices em imagens, é necessário um mapeamento de conceitos de tal

forma, que uma comunidade passa a ser interpretada como uma imagem e as espécies são as intensidades de *pixels* presentes na imagem. De posse desse mapeamento de conceitos, dividiu-se a imagem em três regiões de interesse (ROI) e para cada região foram obtidos os índices de diversidade, criando um vetor de características. Cada ROI cobre regiões significativas da face humana como os dois olhos e a região da boca e nariz. Foram utilizadas duas abordagens na geração deste vetor de características. O objetivo dessas duas abordagens é verificar quão bom um índice pode ser utilizado isoladamente para a descrição de texturas e qual o desempenho de todos os índices juntos na classificação. Na primeira abordagem foi calculado um índice por ROI gerando um vetor de 3 elementos. Na segunda abordagem foram calculados todos os índices para cada ROI gerando um vetor de 30 elementos sendo 10 índices de diversidade para 3 ROI. A Figura 3.8 ilustra como é obtido o vetor de características na primeira abordagem e a Figura 3.9 mostra a segunda abordagem na obtenção do vetor de características.

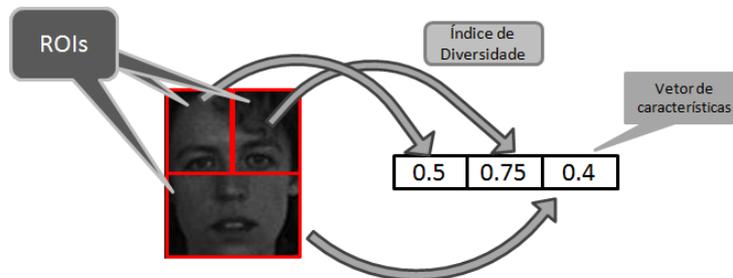


Figura 3.8: Extração do vetor de características utilizando a primeira abordagem (1 índice por ROI).

3.3.3 Reconhecimento de Padrões

Após a extração de características, é então, feita a etapa de classificação das imagens. Neste trabalho é utilizada a MVS (Seção 2.3.1) para classificar as

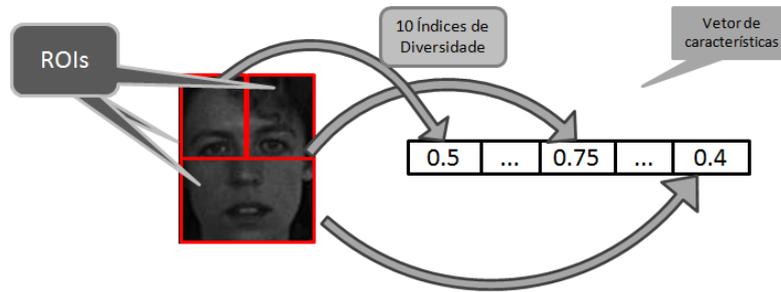


Figura 3.9: Extração do vetor de características utilizando a segunda abordagem (10 índices por ROI).

imagens em face ou não face de acordo com a informação de de textura obtida nas etapas anteriores. A Figura 3.10 ilustra o fluxo de atividades realizadas durante a fase de reconhecimento de padrões. As bases de imagens usadas no treinamento

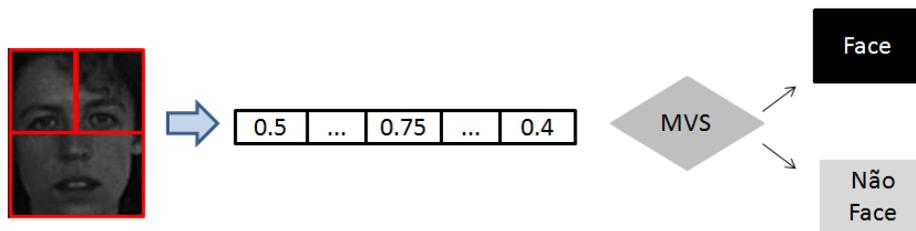


Figura 3.10: Fluxo de atividades durante a fase de reconhecimento de padrões.

e teste do classificador MVS, descritas na Seção 3.2, passaram pelo procedimento de pré-processamento (Subseção 3.3.1), em seguida cada imagem teve suas medidas de textura extraídas (Subseção 3.3.2) e foram devidamente rotuladas como face ou não face. Nesta etapa foi utilizado o pacote LIBSVM (CHANG et al., 2011), com intuito de tornar mais fácil a tarefa de se utilizar uma MVS, em particular fazer ajustes de parâmetros da MVS de forma mais prática, utilizando a ferramenta *grid.py*. Esta ferramenta trata-se de um *script* python que estima os parâmetros ótimos para MVS que utilizam núcleo radial (RBF), pois calcula valores para C (custo da função) e γ (largura da função gaussiana) em uma

grade e realiza o treinamento e classificação por meio do procedimento chamado validação cruzada. Para cada valor de C e γ é obtida uma acurácia (Seção 3.3.2) da MVS para o conjunto de treinamento, no final do processo é mostrado a acurácia máxima obtida e os valores de C e γ .

A validação cruzada é um procedimento que consiste em dividir os dados v vezes em dois subconjuntos de tamanhos iguais: treinamento e teste. Em seguida o subconjunto de teste é classificado de acordo com o conjunto de treinamento. Este processo é repetido v vezes, de tal forma que durante a divisão os elementos de cada conjunto são escolhidos aleatoriamente, dessa forma cada exemplo do conjunto de dados é classificado. A precisão da validação cruzada é a porcentagem de dados que são classificados corretamente (ALMEIDA, 2010). Neste trabalho utilizou-se $v = 10$, ou seja, organizamos os dados de treinamento e classificação de 10 maneiras diferentes. Este valor de v foi escolhido por ser amplamente utilizado na literatura.

Para construir o modelo do classificador, descrevemos um conjunto de classes pré-determinadas, face e não face. Para isto utilizou-se nesta etapa a ferramenta WEKA, que é um ambiente para a utilização de técnicas de reconhecimento de padrões (HALL et al., 2009), além dos algoritmos de reconhecimento de padrões o WEKA permite fazer uma avaliação do desempenho da classificação, pois realiza a construção da matriz de confusão (Seção 2.3) e seus indicadores. Outra característica do WEKA é realizar os testes utilizando o procedimento de validação cruzada.

Capítulo 4

Resultados e Discussão

Este capítulo descreve os experimentos realizados com base na metodologia descrita anteriormente.

Foram realizados 44 testes, divididos da seguinte maneira: um teste para cada índice de diversidade extraído por ROI (Subseção 3.3.2) totalizando dez testes aplicados a um conjunto formado por uma base de faces e uma base de não faces. O outro teste foi feito considerando o cálculo de todos os índices por ROI, aplicando aos quatro conjuntos formados pelas bases de faces e a base de não faces. A seguir será feito um detalhamento dos resultados destes cenários de testes.

4.1 Aplicação de cada Índice Individualmente

Esta seção detalha os resultados dos testes realizados considerando que o vetor de características é formado por um índice de diversidade (Subseção 3.3.2), isto é, para as quatro bases serão aplicados cada índice de diversidade.

4.1.1 Índice de Magalef

O primeiro índice utilizado nos testes foi o índice de Magalef. Esse índice mostrou bons resultados para as bases ORL e Face Expression, onde ambas obtiveram acurácia acima dos 95%, além de apresentarem uma baixa taxa de falsos positivos (abaixo de 2%). A sensibilidade em ambas as bases foi muito boa pois obtiveram valores acima de 90%. Para as bases CBCL e Libor Spacek não foram obtidos resultados parecidos, já que a acurácia para ambas as bases ficou abaixo de 88%. A Tabela 4.1 nos dá um detalhamento do desempenho do índice de Magalef para cada base. Estes resultados indicam que o índice de Magalef obteve bons resultados para imagens em que as condições de iluminação eram muito semelhantes. Na maioria dos casos não haviam faces que tinham elementos como barba e bigode, como é o caso das bases ORL e Face Expression. O fato da base Libor Spacek obter resultados mais baixos, se deve a casos de variações de iluminação e casos em que indivíduos apresentavam barba e bigode (a maior parte destes casos foram considerados não faces). Casos de faces com óculos não causaram grande influência nas medidas de avaliação. A baixa acurácia da base CBCL se deve principalmente a fatores de iluminação, pois nessa base a variação de iluminação entre as imagens é grande. Como o índice de Magalef basicamente mede a relação de espécies (níveis de cinza) e quantidade de indivíduos (número de *pixels*), se há uma grande variação dos níveis de cinza e ruído, isto afeta diretamente os valores obtidos pelo índice.

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	88.20	16.2	88.8
Libor Spacek	79.63	15	78.09
ORL	96.42	1.62	91.87
Face Expression	95.64	1.07	92.70

Tabela 4.1: Resultados para o índice de Magalef

4.1.2 Índice de Menhinick

Os resultados para o índice de Menhinick foram os que atingiram os valores mais baixos para as medidas de avaliação. Para este índice, o classificador não conseguiu fazer uma boa distinção entre faces de não faces, uma vez que todas as bases obtiveram altas taxas de falsos positivos. Isto indica que o classificador identificou muitas não faces como sendo faces, e houve uma baixa sensibilidade, principalmente na base CBCL, onde menos da metade das faces foram classificadas corretamente (sensibilidade de 48%). A Tabela 4.2 detalha os resultados obtidos para o índice de Menhinick. Um dos fatores que ajuda a compreender melhor estes resultados é do fato do índice de Menhinick utilizar uma faixa de valores para mensurar a riqueza em uma comunidade bem menores que o índice de Magalef, isto interfere diretamente no processo de aprendizado da MVS.

Base de Teste	ac(%)	tfp (%)	sen (%)
CBCL	66.30	21.10	48
Libor Spacek	79.27	31.43	86.74
ORL	80.57	17.95	74.24
Face Expression	83.59	13.49	81.03

Tabela 4.2: Resultados para o índice de Menhinick

4.1.3 Índice de Shannon

O índice de diversidade de Shannon obteve bons resultados para todas as bases. Apresentou valores altos de acurácia para as bases ORL e Face Expression. A base Libor Spacek foi a que obteve menor acurácia e sensibilidade para este índice. É notório a melhoria dos indicadores de avaliação de todas as bases para o índice de Shannon (Tabela 4.3). Isto está diretamente ligado a característica que o índice de Shannon tem de considerar, não só a quantidade de espécies (níveis de cinza) como fazem os índices de Magalef e Menhinick, mas também quão distribuídos estão as espécies. Outra característica a ser considerada é o fato do índice de Shannon dar o mesmo peso a espécies comuns, quanto as espécies raras. Isso dá uma boa relação ao se considerar possíveis indivíduos que possam ser considerados como ruído.

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	90.56	9.43	90.4
Libor Spacek	84.41	7.74	81.93
ORL	97.57	0.81	93.78
Face Expression	98.48	0.20	97.20

Tabela 4.3: Resultados para o índice de Shannon

4.1.4 Índice de Pielou

A versão do índice de Shannon para mensurar equabilidade, índice de Pielou, obteve acurácias e sensibilidades mais baixas que o índice de Shannon para todas as bases (Tabela 4.4). Apesar deste índice ser uma medida normalizada, ou seja, seus valores variam numa faixa bem definida, no caso de 0 a 1, os

valores indicadores de avaliação apontam maus resultados. Uma razão para isto pode ser o fato deste índice dar muito mais peso para espécies raras, isto implica que eventuais ruídos terão uma influência muito grande no valor do índice. A evidência é a baixa sensibilidade obtida da base CBCL (62.36%), a menor acurácia encontrada em todos os testes. Vale ressaltar que nesta base as imagens são mais ruidosas.

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	62.36	33.95	60
Libor Spacek	83.80	20.10	85.75
ORL	79.14	14.02	62.98
Face Expression	85.77	6.42	80.17

Tabela 4.4: Resultados para o índice de Pielou

4.1.5 Índice de Simpson

O índice de Simpson seguiu a tendência que os índices anteriores, classificando melhor as imagens das bases ORL e Face Expression (Tabela 4.5). Os resultados da base Libor Spacek se devem principalmente pela presença de elementos estruturais (barba, bigode, etc). A base CBCL obteve uma sensibilidade de 89% e uma taxa de falsos positivos relativamente baixa, o mesmo ocorrendo para as outras bases. Isto indica que para o índice de Simpson a MVS identificou melhor as amostras de não faces que as faces. Outra característica que podemos destacar, é que os indicadores de avaliação para o índice de Simpson atingiram uma faixa entre os valores alcançados pelo índice de Shannon e o índice de Pielou. Este fato é por que o índice de Simpson praticamente dá quase nenhum peso às

espécies raras, o que significa que desconsiderou todos os ruídos, e desconsiderou também espécies com poucos representantes que continham informação útil.

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	88.92	11.22	89.6
Libor Spacek	83.41	8.96	81.06
ORL	95.92	0.93	89.07
Face Expression	98.07	0.41	96.62

Tabela 4.5: Resultados para o índice de Simpson

4.1.6 Índice de Simpson para Equabilidade

A versão para equabilidade do índice de Simpson obteve acurácias pouco acima das obtidas pela versão original do índice de Simpson (Tabela 4.6). Apenas para base Libor Spacek não apresentou uma melhoria da acurácia, tendo inclusive um aumento na taxa de falsos positivos. Estes resultados mostram que o índice de Simpson para equabilidade não conseguiu uma boa distinção entre faces e não faces para as imagens da base Libor Spacek, provavelmente pelo número considerável de imagens de face contendo componentes faciais (barba e bigode), já que nas outras bases a quantidade de casos de faces contendo componentes faciais é menor, e em todas as outras bases foi verificada uma melhoria nos indicadores de avaliação.

4.1.7 Índice de MacIntosh

Os testes realizados com índice de MacIntosh e a sua versão para equabilidade podem ser vistos nas Tabelas 4.7 e 4.8. É interessante destacar que a versão para

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	92.7	8.0	92.6
Libor Spacek	79.27	31.43	86.74
ORL	97.28	0.91	93.10
Face Expression	98.58	0.30	97.49

Tabela 4.6: Resultados para o índice de Simpson para equabilidade

equabilidade do índice de MacIntosh obteve os melhores indicadores entre todos os índices. Para base CBCL, a acurácia é de 96.6%, taxa de falsos positivos é de 6.6%, e a sensibilidade é de 96.6%.

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	96.6	6.6	96.6
Libor Spacek	84,02	7.62	81.44
ORL	97.28	0,71	92.69
Face Expression	98.58	0.60	97.78

Tabela 4.7: Resultados para o índice de MacIntosh

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	94.4	11.8	94.6
Libor Spacek	79.16	13.64	77.27
ORL	96.57	0.92	90.93
Face Expression	95.54	0.64	92.20

Tabela 4.8: Resultados para o índice de MacIntosh para equabilidade

4.1.8 Índice de Brillouin

O destaque dos resultados obtidos para o índice de Brillouin (Tabela 4.9) é o fato das bases ORL e Face Expression obterem acurácias acima dos 97%.

A base Libor Spacek obteve a menor acurácia de 83% e a base CBCL obteve 88% de acurácia. Este índice tem o funcionamento muito semelhante ao índice de Shannon, pois também dá o mesmo peso a espécies raras e comuns, mas ao contrário do índice de Shannon o índice de Brillouin é muito sensível a amostragem e ruídos e outros fatores como a presença de barba, o que indica a baixa acurácia para a base Libor Spacek.

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	88.23	11.76	88.4
Libor Spacek	83.91	9.75	81.84
ORL	97.5	0.71	93.35
Face Expression	98.48	0.61	97.58

Tabela 4.9: Resultados para o índice de brillouin

4.1.9 Índice de Berker-Parker

O índice de Berger-Parker seguiu a tendência de obter melhores indicadores para as bases Face Expression e ORL. A base Libor Spacek obteve a maior acurácia para este índice dentre todos os outros índices utilizados, e a base CBCL obteve uma baixa sensibilidade. A principal razão para estes indicadores é o fato deste índice apresentar grande dependência em relação a iluminação, pois este índice mede a probabilidade de escolher o indivíduo de maior frequência na amostra, isto implica que sob as mesmas condições de iluminação este índice tende a apresentar valores mais próximos. Enquanto que diferenças de iluminação na cena causam discrepâncias nos valores obtidos pelo índice de Berger-Parker. Mais detalhes veja a Tabela 4.10.

Base de Teste	ac (%)	tfp (%)	sen (%)
CBCL	71.87	19.02	55.51
Libor Spacek	86.21	6.85	83.82
ORL	97.28	0.91	93.10
Face Expression	98.58	0.30	97.49

Tabela 4.10: Resultados para o índice de Berger-Parker

4.1.10 Combinação de Todas as Bases

O último teste foi realizado a partir da combinação de todas as bases de imagens. Nota-se que o índice que apresentou a maior acurácia foi o índice de Shannon (88.88%) e a menor acurácia foi verificada com o índice de Simpson (78.68%). Os demais índices obtiveram acurácias mais próximas da obtida pelo teste realizado com o índice de Shannon. Oito índices obtiveram acurácias acima de 86% e dois índices apresentaram acurácias em torno dos 78%. Quanto a taxa de falsos positivos, foi verificado que o índice de Simpson para equabilidade obteve 19.85%, a maior taxa de falsos positivos. A menor taxa de falsos positivos foi encontrada para o índice de Berger-Parker (6.04%), enquanto que a maior parte dos índices obtiveram taxas de falsos positivos abaixo dos 10%. A maior sensibilidade foi encontrada no teste realizado com o índice de Brillouin (82.5%) e o menor valor encontrado foi 53% (índice de Pielou). A Tabela 4.12 detalha todos os indicadores de desempenho dos índices de diversidade utilizados para todas as bases de imagens combinadas.

Índice	ac (%)	tfp (%)	sen (%)
Magalef	86.66	9.61	78.42
Menhinick	87.40	11.66	81.45
Shannon	88.88	7.29	81.38
Pielou	78.44	18.6	53.77
Simpson	87.57	7.76	79.05
Simpson Equabilidade	78.68	19.85	71.67
MacIntosh	88.82	7.94	81.74
MacIntosh Equabilidade	87.20	8.03	78.15
Brillouin	87.94	9.54	82.50
Berger-parker	88.79	6.74	80.86

Tabela 4.11: Resultados de todos os índices para todas as bases juntas

Tabela 4.12: Resultados de todos os índices para todas as bases juntas

4.1.11 Análise geral do Teste

No geral nota-se que a base CBCL obteve as menores acurácias que variou de 62% a 96%, enquanto que as bases ORL e Face Expression obtiveram as maiores acurácias. A base Libor Spacek apresentou valores de acurácia intermediários. Quanto ao índice que alcançou a maior acurácia (98.58%), foi o índice de MacIntosh, aplicado à base Face Expression, mas os outros índices obtiveram valores de acurácia muito próximos. Os índices que apresentaram os piores indicadores foram os índices de Menhinick (66.30% de acurácia) e Pielou (62.36% de acurácia), ambos obtidos para base CBCL. A Figura 4.1 mostra exemplos de imagens classificadas incorretamente, enquanto a Figura 4.2 mostra alguns exemplos de face classificadas corretamente.

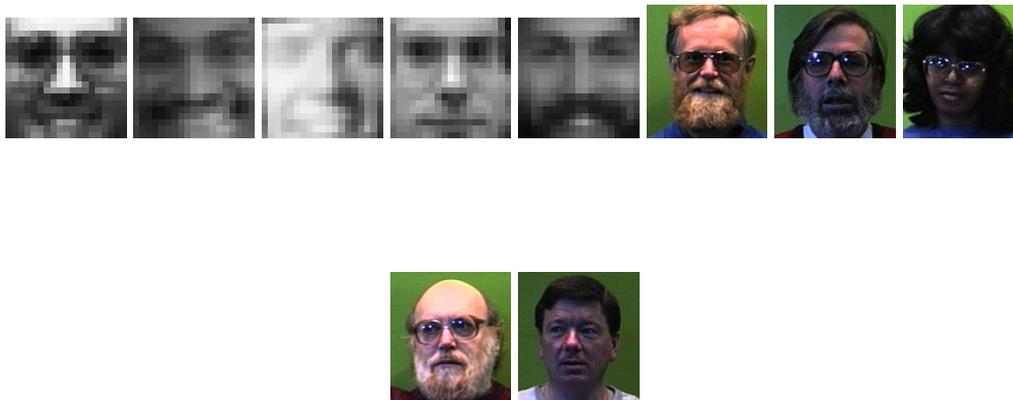


Figura 4.1: Imagens classificadas como não face

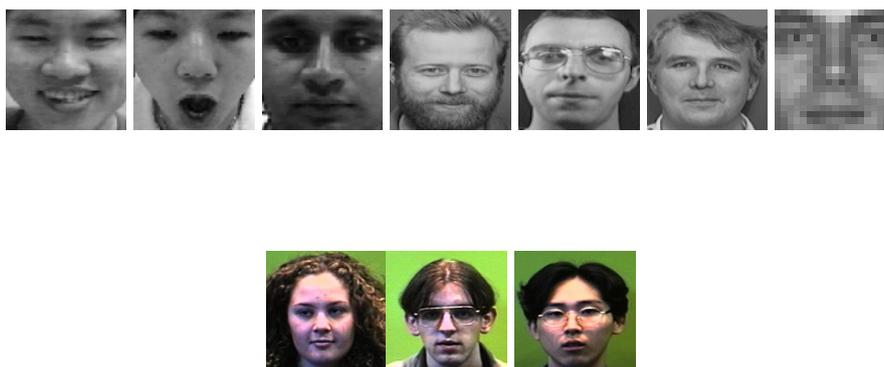


Figura 4.2: Imagens classificadas como face

4.2 Combinação de Todos os Índices

Esta seção mostra os resultados dos testes realizados considerando que o vetor de características é formado pelos 10 índices de diversidade para cada base (Subseção 3.3.2). Foram realizados 5 tipos de testes: 4 testes para cada base isoladamente e 1 teste com todas as bases juntas.

Os resultados obtidos nos quatro primeiros cenários, é mostrado na Tabela 4.13. É possível verificar que a acurácia aumentou de forma considerável em todas as bases se compararmos com os resultados obtidos no teste da Seção 4.1. A base Libor Spacek apresentou a menor sensibilidade (89.4%) e a maior taxa de falsos positivos (17%), o que explica a menor acurácia. A base CBCL obteve uma acurácia de 94% apesar da taxa de falsos positivos alta (9.4%). A base Face Expression obteve uma acurácia de 99.2% destacando-se a sensibilidade de 100%, ou seja, todas as faces foram classificadas corretamente. A taxa de falsos positivos foi a menor encontrada entre todas as bases testadas (0.8%), o que também contribuiu para a alta acurácia. A base ORL também obteve bons indicadores, já que sua acurácia foi de 97%, sensibilidade de 98% e uma baixa taxa de falsos positivos (1.4%). Esses resultados indicam que a combinação dos índices de diversidade gerou resultados melhores do que cada índice isolado, muito disso se deve ao fato de que cada índice mensura características diferentes de uma amostra, por exemplo, o índice de Shannon representa muito bem situações em que há muitas espécies (níveis de cinza) distribuídas na amostra, enquanto que os índices de Berger-Parker, Simpson conseguem uma boa representação de comunidades com poucas espécies (pouca variação de níveis de cinza).

Base de Teste	Ac(%)	tfp (%)	Sen (%)
CBCL	94.5	9.4	94.4
Libor Spacek	89	17	89.4
ORL	97.9	1.4	98
Face Expression	99.2	0.8	100

Tabela 4.13: Resultados para todos os índices aplicados a todas as ROIs

O último teste realizado foi feito a partir da combinação de todas as bases. Neste teste verificamos bons resultados, pois a acurácia foi de 97%, taxa de falsos positivos igual 6.39% e sensibilidade obteve 93.64%. Isto indica que a combinação de todas as bases de imagens juntamente com a combinação de todos os índices implicou num aumento no poder de generalização do classificador, pois mesmo as imagens da base CBCL que isoladamente obteve acurácia de 79%, isto não fez com que diminuísse a acurácia quando todas as bases foram testadas em conjunto. A figura 4.3 mostra exemplos de casos de imagens classificadas corretamente em todos os casos e a Figura 4.4.



Figura 4.3: Imagens classificadas corretamente como face utilizando todos os índices

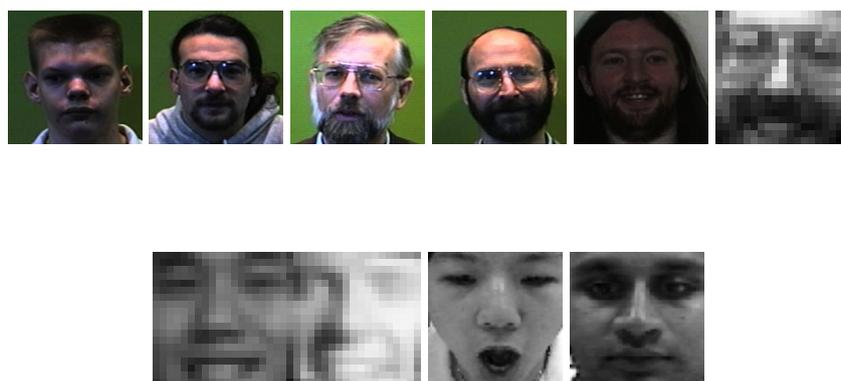


Figura 4.4: Imagens classificadas como incorretamente como não face utilizando todos os índices por ROI

4.3 Comparação com métodos da Literatura

A seguir é feita uma comparação com os métodos utilizados na literatura (Seção 1.3). A tabela 4.14 mostra os trabalhos, a técnica utilizada, a base de imagens que foram obtidos estes resultados e as acurácias obtidas em trabalhos.

Trabalho	Técnica	Base de Teste	Acurácia(%)
(ZHENG et al., 2004)	NNL	ORL	95.18
(ZHENG et al., 2004)	NNP	ORL	95.75
(Lu et al., 2007)	SPCA	CBCL	94
(CORRÊA et al., 2008)	LSTM	CBCL	85
(CORRÊA et al., 2008)	LSTM	ORL	96
(MEHTA et al., 2011)	LBP	ORL	97
(MEHTA et al., 2011)	LPA-LBP	ORL	99
Método proposto	Índice de MacInstosh	ORL	97.28
Método proposto	Todos Índices	Face Expression	99.2

Tabela 4.14: Comparação com métodos utilizados na literatura

Capítulo 5

Considerações Finais

O presente trabalho apresentou um estudo sobre a análise e classificação de imagens de face utilizando índices de diversidade como descritores de textura. Foram abordados 10 índices de diversidade e a máquina de vetores suporte para a classificação das imagens de faces.

Foi utilizada a técnica retinex nas imagens para atenuar efeitos de iluminação da cena que elas foram obtidas. Foram utilizadas duas estratégias para a obtenção de características de textura das imagens, na primeira estratégia dividiu-se a imagem em três regiões de interesse e foi extraído um índice de diversidade para cada região. A segunda estratégia também foi feita a mesma divisão da imagem, mas foram obtidos todos os 10 índices para cada região. Foram utilizadas quatro bases de imagens de face e uma base de imagens de não face para a validação dos índices apresentados. Os testes realizados consistiram de dois tipos. Em ambos os tipos foram aplicadas as estratégias de extração de características de textura para cada imagem das quatro bases utilizadas e um teste combinando todas as bases. No primeiro tipo de teste aplicou-se a primeira estratégia para a extração

de características, enquanto o segundo, utilizou-se a segunda estratégia para a obtenção de características de textura.

Observou-se que para bases de imagens que o ambiente de iluminação era bem controlado (ORL e CMU AMP Face Expression) as acurácias para oito índices foram superiores a 90% enquanto que os índices de Menhinick e Pielou apresentaram acurácias abaixo dos 85%. Na base Libor Spacek as acurácias estiveram abaixo das obtidas nas duas bases mencionadas, a maior acurácia obtida foi de 86% a pouco uma provável razão é a existência de muitas amostras de faces conterem elementos como barba e bigode, outros elementos como óculos não causaram diminuição na acurácia. A base CBCL foi a que obteve os piores resultados, sendo que o melhor resultado obtido foi de 78%. Isto se deve em muito ao fato das imagens desta base possuir ruídos e grandes variações de iluminação. Nos testes utilizando todos os índices para todas as bases apresentaram resultados melhores, sendo que as bases ORL (97%) e CMU AMP Face Expression (99%) obtiveram as melhores acurácias, já as bases Libor Spacek (89%) e CBCL (80%) obtiveram acurácias bem acima das obtidas pela quando testadas apenas para um índice.

Os experimentos demonstraram que os índices de diversidade podendo ser usados como descritores de textura de faces humanas. A partir da pesquisa realizada, novas propostas surgem para trabalhos futuros. Uma proposta seria pesquisar outros índices de diversidade (beta e gama) (MAGURRAN, 2004). Uma outra proposta é pesquisar métodos de normalização da iluminação, pois o retinex não mostrou grande eficiência na normalização de imagens com grandes variações de iluminação e também métodos para a redução de ruídos como o filtro de Wiener. Outra proposta é alterar a quantidade de ROIs na extração de

características, inclusive variando o tamanho das ROIs e finalmente utilizar esta metodologia em um sistema de detecção de faces.

Referências

- ALMEIDA, DE J.D.S. ; LUÍS, S. (2010). Metodologia computacional para detecção automática de estrabismo em imagens digitais através do teste de hirschberg. Master's thesis, UFMA.
- BURGES, C. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167.
- CAMBRIDGE (2011). Orl face database:. available at: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>. *Acesso em 2012*.
- CHANG, CHIH-CHUNG, LIN, and CHIH-JEN (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- CORRÊA, D., SALVADEO, D., LEVADA, A., and J.H., S. (2008). Using lstm network in face classification problems.

- DALAL, N. ; TRIGGS, B. (2005). Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893.
- Dutra, R. (1995). Um procedimento para o cálculo do índice de diversidade de brillouin. *Revista Brasileira de Zoologia*, 12(4):1025–1027.
- ECLIPSE, I. (2011). Disponível em: <http://www.eclipse.org>. Acesso em, 12.
- GILCHRIST, A., KOSSYFIDIS, C., BONATO, F., AGOSTINI, T., CATALIOTTI, J., LI, X., SPCHAR, B., ANNAN, V., and ECONOMOU, E. (1999). An anchoring theory of lightness perception. *Psychological review*, 106(4):795.
- HALL, M., FRANK., and HOLMES, G. (2009). The weka data mining software: An update. *SIGKDD Explorations*, 11.
- HARALICK, R., SHANMUGAM, K., and DINSTEN, I. (1973). Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, 3(6):610–621.
- HEUSCH, G., CARDUNAUX, F., and MARCEL, S. (2005). Lighting normalization algorithms for face verification. *IDIAP-com 05*, 3.
- JAI, S. (2011). Disponível em: <http://java.sun.com/javase/technologies/desktop/media/jai/>. Acesso em 2012.
- JAVA, S. (2011). Disponível em: <http://www.oracle.com/technetwork/java/>. Acesso em 2012.

- JÚNIOR, G. B. (2008). Classificação de regiões de mamografias em massa e não massa usando estatística espacial e máquina de vetores de suporte. Master's thesis, UFMA.
- JOBSON, D., RAHMAN, Z., and WOODSELL, G. (1997). A multiscale retinex for bridging the gap between color images and the human observation of scenes. *Image Processing, IEEE Transactions on*, 6(7):965–976.
- KANIESKI, M. R. (2010). Caracterização florística, diversidade de correlação ambiental da floresta nacional de são francisco de paula, rs. Master's thesis, UFSM.
- KAUSHAL, A. ; RAINA, J. (2010). Face detection using neural network and gabor wavelet transform. *International Journal of Computer Science and Technology*, 1(1).
- LAND, E.H. ; McCANN, J. (1971). Lightness and retinex theory. *Journal of the Optical society of America*, 61(1):1–11.
- LIMA, C. A. d. M. (2004). *Comitê de Máquinas: Uma abordagem Unificada Empregando Máquinas de Vetores-Suporte*. PhD thesis, Universidade Estadual de Campinas, SP.
- LORENA, A.C. ; CARVALHO, A. (2003). Introdução aos classificadores de margens largas. Technical report, ICMC.
- LOWE, D. (1999). Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150–1157.

- Lu, C., Zhang, C., Zhang, T., and Zhang, W. (2007). Kernel based symmetrical principal component analysis for face classification. *Neurocomputing*, 70(4):904–911.
- MA, E. (2007). Avaliação de características haar em um modelo de detecção de face. *CEP*, 70910:900.
- MAGURRAN, A. (2004). *Measuring biological diversity*. Blackwell Pub.
- MEHTA, R., YUAN, J., and EGIAZARIAN, K. (2011). Local polynomial approximation-local binary pattern (lpa-lbp) based face classification. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7881, page 19.
- MELO, A. (2008). O que ganhamos 'confundindo' riqueza de espécies e equabilidade em um índice de diversidade? *Biota Neotrop. Jul/Sep 2008 vol. 8, no. 3*
<http://www.biotaneotropica.org.br/v8n3/pt/abstract?article+bn00108032008>
ISSN 1676-0603.
- NAGESH, P.; LI, B. (2009). A compressive sensing approach for expression-invariant face recognition. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1518–1525.
- PAPAGEORIOU, C., OREN, M., and POGGIO, T. (1998). A general framework for object detection. In *Computer Vision, 1998. Sixth International Conference on*, pages 555–562. IEEE.

- RAHMAN, Z., JOBSON, D., and WOODSELL, G. (2011). Investigating the relationship between image enhancement and image compression in the context of the multi-scale retinex. *Journal of Visual Communication and Image Representation*.
- RANGEL, J.E.; GOMES, Q. (2009). Universidade federal de campina grande.
- RICOTTA, C. (2005). Through the jungle of biological diversity. *Acta Biotheoretica*, 53(1):29–38.
- ROWLEY, H. A., BALUJA, S., and KANADE, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and MachineIntelligence*, 20(1):23–38.
- SANTOS, D. V. (2009). Uma generalização da distribuição do índice de diversidade generalizado por good com aplicação em ciências agrárias. Master's thesis, UFRPE.
- SILVA, A. (2004). *Algoritmos para Diagnostico Assistido de Nodulos Pulmonares Solitarios em Imagens de Tomografia Computadorizada*. PhD thesis, PUC-Rio.
- SMOLA, A. (2000). *Advances in large margin classifiers*. MIT Press.
- SOUTO, R. (2003). *Segmentação de imagem multiespectral utilizando-se o atributo matiz*. PhD thesis, Tese. INPE, São José dos Campos.
- SPACEK, L. (2011). The essex faces94 database.
- SUNG, K. (1996). *Learning and Example Selection for Object and Pattern Recognition*. PhD thesis, MIT, Artificial Intelligence Laboratory and Center for Biologicaland Computational Learning, Cambridge. MA.

- TURK, M.; PENTLAND, A. (1991). Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86.
- VIOLA, P.; JONES, M. (2004). Robust real-time face detection. *International journal of computer vision*, 57(2):137–154.
- WARING, C.A.; LIU, X. (2005). Face detection using spectral histograms and svms. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 35(3):467–476.
- WOODS, R. G. R. (1992). *Digital image processing*. Addison - Wesley Publishing Company, Reading, Massachusetts.
- Wu, T., BARTLETT, M., and MOVELLAN, J. (2010). Facial expression recognition using gabor motion energy filters. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 42–47.
- ZHAO, W., PARK, J., and LEE, S. (2011). Fully automatic face detection and facial feature points extraction using local gabor filter bank and pca. In *Machine Learning and Cybernetics (ICMLC), 2011 International Conference on*, volume 4, pages 1789–1792.
- ZHENG, W., ZHAO, L., and ZOU, C. (2004). Locally nearest neighbor classifiers for pattern classification. *Pattern recognition*, 37(6):1307–1309.