

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO E SISTEMAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**RECONHECIMENTO DE PADRÕES USANDO
INDEXAÇÃO RECURSIVA**

Tese submetida à Universidade Federal de Santa Catarina para a obtenção do
título de Doutor em Engenharia de Produção

JOÃO ARTUR DE SOUZA

**FLORIANÓPOLIS
1999**

JOÃO ARTUR DE SOUZA

**RECONHECIMENTO DE PADRÕES
USANDO INDEXAÇÃO RECURSIVA**

Esta tese foi julgada adequada para a obtenção do título de DOUTOR EM ENGENHARIA DE PRODUÇÃO e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia de Produção.

Prof. Ricardo Miranda Barcia, Ph.D.
Coordenador do Curso

BANCA EXAMINADORA:

**PROF. ROBERTO C. S. PACHECO, DR
ORIENTADOR**

**PROF^a. EDIS MAFRA LAPOLLI, DR
CO-ORIENTADORA**

**PROF. FLÁVIO R. LAPOLLI, DR
MODERADOR**

RICARDO M. BARCIA, PHD

**VINÍCIUS M. KERN, DR
EXAMINADOR EXTERNO**

**MARCO A. B. CÂNDIDO, DR
EXAMINADOR EXTERNO**

Florianópolis, 16 de Setembro de 1999

É preciso recompor o todo.

Marcel Mauss

À GERTRUDES

AO HEITOR

AGRADECIMENTOS

Agradecer a todos que contribuíram para realização deste trabalho, constitui-se numa tarefa muito difícil. No entanto não posso deixar de agradecer a algumas pessoas e instituições que contribuíram diretamente para a realização desta tese.

À CAPES pelo apoio financeiro.

À Universidade Federal de Pelotas, pela liberação de minhas atividades acadêmicas, principalmente aos colegas de trabalho que me apoiaram nesta liberação.

À Universidade Federal de Santa Catarina, em especial ao Programa de Pós-Graduação em Engenharia de Produção em nome do Prof. Ricardo M. Barcia, por ter me recebido de braços abertos.

À empresa EMSA pelo fornecimento dos registros polissonográficos.

À Engenharia Biomédica, em nome do professor Fernando M. de Azevedo, por colocar a nossa disposição o material de pesquisa do grupo.

A todos os membros da banca examinadora, principalmente aos examinadores externos, por aceitarem o convite, honrando-nos com vossa presença.

Aos Profs. Ricardo M. Barcia, Alejandro Martins, José Leomar e Fernando Gauthier pelo incentivo, críticas e orientação durante alguns estágios deste aprendizado.

À professora Edis Lapolli pela sua Co-Orientação, amizade, apoio e contextualizações em várias etapas na elaboração desta tese.

Principalmente ao professor e Orientador Roberto C. S. Pacheco, pela amizade, paciência, orientação, apoio e contribuições valiosas desde os passos iniciais até a conclusão deste trabalho.

A todos os amigos, em especial a Roberto Raitz, Malcon Tafner, Paulo Goulart e Paulo Afonso B. da Costa, pela colaboração, críticas construtivas, apoio e amizade.

Aos meus familiares, dos quais nunca faltaram muito amor e incentivo, durante todo o meu desenvolvimento pessoal e profissional, principalmente a meu irmão Henor, que sempre se faz presente.

E um agradecimento todo especial as duas pessoas mais importantes de minha vida, Gertrudes e Heitor. Que sempre souberam estar do meu lado.

E finalmente à luz que nos guia... Obrigado!

SUMÁRIO

LISTA DE FIGURAS.....	ix
LISTA DE TABELAS.....	x
LISTA DE QUADROS.....	xi
RESUMO.....	xii
ABSTRACT.....	xiii
1 INTRODUÇÃO.....	1
1.1 ORIGEM DO TRABALHO	1
1.2 OBJETIVOS	4
1.3 JUSTIFICATIVA E IMPORTÂNCIA DO TRABALHO	4
1.4 METODOLOGIA DE DESENVOLVIMENTO DA PESQUISA.....	5
1.5 ESTRUTURA DO TRABALHO	6
2 RECONHECIMENTO DE PADRÕES	8
2.1 INTRODUÇÃO	8
2.2 APLICAÇÕES DE RECONHECIMENTO DE PADRÕES	9
2.3 ESTRUTURA TÍPICA DE UM SISTEMA DE RP	9
2.3.1 <i>Redução de Dimensionalidade</i>	11
2.3.2 <i>Abordagens para classificação</i>	14
2.3.2.1 Abordagens estatísticas	15
2.3.2.2 Abordagem Sintática	15
2.3.2.3 Abordagem Neural	16
2.3.2.4 Abordagem Difusa	17
2.4 AS TENDÊNCIAS DE RP	17
2.5 CONCLUSÃO.....	19
3 INDEXAÇÃO E RECUPERAÇÃO DE IMAGENS.....	20
3.1 INTRODUÇÃO	20
3.2 INDEXAÇÃO E RECUPERAÇÃO DE IMAGENS: UMA VISÃO GERAL.....	22
3.2.1 <i>Indexação de imagem Baseada em conteúdo</i>	24
3.2.2 <i>Evolução</i>	26
3.3 MÉTODOS BÁSICOS UTILIZADOS NO PROCESSO DE INDEXAÇÃO	29
3.3.1 <i>PCA (Principal Components Analysis)</i>	30

3.3.2	<i>Histogramas</i>	31
3.3.3	<i>Filtros Gaussianos</i>	31
3.3.4	<i>Wavelet</i>	33
3.3.5	<i>Transformada de Fourier</i>	34
3.3.6	<i>Redes Neurais</i>	36
3.4	DESCRIÇÃO DE ALGUNS SISTEMAS DE RECUPERAÇÃO DE IMAGENS.....	36
3.4.1	<i>QBIC</i>	37
3.4.2	<i>IRIS - Image Retrieval for Information Systems</i>	39
3.4.3	<i>Virage</i>	40
3.4.4	<i>VisualSEEk</i>	41
3.4.5	<i>Synapse</i>	42
3.5	CONCLUSÕES	43
4	INREC - INDEXAÇÃO RECURSIVA: MODELO PROPOSTO PARA REDUÇÃO DE DIMENSIONALIDADE	45
4.1	INTRODUÇÃO	45
4.2	MODELO PROPOSTO – INDEXAÇÃO RECURSIVA (INREC).....	46
4.2.1	<i>INREC</i>	46
4.2.1.1	Função f	48
4.2.2	<i>Originalidade</i>	50
4.2.3	<i>Conclusão</i>	51
5	APLICAÇÕES DO MODELO PROPOSTO (INREC).....	52
5.1	APLICAÇÃO DO INREC PARA INDEXAÇÃO DE IMAGENS.....	52
5.2	INREC PARA CLASSIFICAÇÃO DE PADRÕES COM POUCAS CARACTERÍSTICAS	54
5.2.1	<i>Aplicação da INREC aos dados da Iris de Fisher</i>	55
5.2.2	<i>Aplicação da INREC nos dados dos cromossomas</i>	56
5.3	APLICAÇÃO DA INREC PARA ESTAGIAMENTO AUTOMÁTICO DO SONO	58
5.3.1	<i>Introdução</i>	58
5.3.2	<i>O Sono e Sua Estrutura</i>	58
5.3.2.1	O Sono.....	58
5.3.2.2	A Monitorização.....	59
5.3.2.3	Considerações Sobre o Estagiamento.....	59
5.3.2.4	O Padrão de Estagiamento.....	60
5.3.2.5	Progresso do Sono Durante a Noite.....	61
5.3.3	<i>O Estagiamento Automático</i>	61
5.3.3.1	Pré-processamento dos Dados.....	61
5.3.3.1.1	Ajuste do Sinal Digital.....	62
5.3.3.1.2	Geração dos Índices Recursivos.....	62
5.3.3.1.3	Extração das Características.....	63

5.3.4	<i>Análise dos Resultados</i>	64
5.3.4.1	Considerações Iniciais	64
5.3.4.2	Resultados	65
5.4	ANÁLISE DOS RESULTADOS DAS APLICAÇÕES	66
6	CONCLUSÕES E RECOMENDAÇÕES	67
6.1	CONCLUSÕES	67
6.2	RECOMENDAÇÕES	68
7	BIBLIOGRAFIA	69

LISTA DE FIGURAS

FIGURA 1: DESENVOLVIMENTO DA PESQUISA	6
FIGURA 2 : O PROBLEMA DE RECONHECIMENTO DE PADRÕES	10
FIGURA 3: UM ESQUEMA DE SISTEMA DE RECUPERAÇÃO E ARMAZENAGEM DE IMAGEM.....	23
FIGURA 4: <i>QUERY</i> DE UMA IMAGEM BASEADA EM CONTEÚDO	26
FIGURA 5: REPRESENTAÇÃO GRÁFICA DA PRIMEIRA (v_1) E SEGUNDA COMPONENTE PRINCIPAL.....	30
FIGURA 6: UM EXEMPLO DE UMA PEQUENA IMAGEM (ESQUERDA) E O NÚCLEO (DIREITA)	32
FIGURA 7: IMAGEM POR TRANSFORMADA WAVELET	34
FIGURA 8: ARQUITETURA DO SISTEMA QBIC	38
FIGURA 9: ARQUITETURA DO SISTEMA IRIS	39
FIGURA 10: ARQUITETURA DO VIRAGE.....	41
FIGURA 11: ARQUITETURA DO VISUALSEEK.....	42
FIGURA 12: GRÁFICO DAS PRIMEIRAS 12 COMPOSIÇÕES DA FUNÇÃO $F = x^2$: (A) COM $x_0 = 0.99$ A RECURSÃO CONVERGE PARA 0; (B) COM $x_0 = 1.001$ A RECURSÃO DIVERGE	49
FIGURA 13 : GRÁFICO DE ALGUMAS FUNÇÕES UTILIZADAS PARA TESTES	49
FIGURA 14: LADO ESQUERDO SÃO AS FOTOS ORIGINAIS; LADO DIREITO AS ALTERADAS	53
FIGURA 15: APRESENTAÇÃO DAS FOTOS NA ORDEM GERADA PELO ÍNDICE	54
FIGURA 16 - DIAGRAMA DOS PRINCIPAIS REGISTROS DA POLISSONOGRRAFIA (GUITOT, 1996).....	59
FIGURA 17 - TRECHO DO SINAL BIOLÓGICO ORIGINAL COM OS LIMITES SUPERIOR E INFERIOR	62
FIGURA 18 - VETOR DE ÍNDICES	63
FIGURA 19 - FAIXA DE AMPLITUDES E ELEMENTO DE PARTIÇÃO PARA O VETOR DE ÍNDICES	63

LISTA DE TABELAS

TABELA 1 - AVALIAÇÃO DA PERFORMANCE DO INREC PARA OS DADOS DA ÍRIS	56
TABELA 2 - AVALIAÇÃO DA PERFORMANCE DA INREC PARA OS DADOS DOS CROMOSSOMOS	57
TABELA 3 - PACIENTES USADOS PARA O TRABALHO	64
TABELA 4 - DISTRIBUIÇÃO DO ACERTO PARA O PACIENTE 7706.....	65
TABELA 5 - DISTRIBUIÇÃO DO ACERTO PARA OS PACIENTES 6740 E 9300	65

LISTA DE QUADROS

QUADRO 1 : DESCRIÇÃO SUCINTA DAS CLASSES DO SONO	60
QUADRO 2 : APARÊNCIA DOS SINAIS DAS CLASSES DO SONO.....	60

RESUMO

O principal objetivo desta tese é desenvolver um modelo matemático para realizar a redução de dimensionalidade. O modelo é uma forma alternativa de redução de dimensionalidade e reconhecimento de padrões através do uso da recursividade da função sujeita aos parâmetros que representam as características dos padrões. O modelo proposto, chamado de *indexação recursiva*, utiliza o processo de recursão, ou composição, de uma função matemática para encapsular, em um número apenas, a informação contida em um padrão. A meta é gerar índices próximos quando são indexadas variações de um padrão existente.

Com a finalidade de avaliar a eficiência do modelo proposto, a indexação recursiva é utilizada em três problemas de reconhecimento de padrões distintos: indexação de imagens, classificação de padrões com poucas características (iris de Fisher e os cromossomos para os grupos de Denver) e processamento de sinal digital. Para o processamento de sinal (neste caso, reconhecimento automático das fases do sono), a indexação recursiva é utilizada como parte do sistema de pré-processamento do sinal de eletroencefalograma.

ABSTRACT

The main purpose of this dissertation is to develop a mathematical model to accomplish the dimensionality reduction. The model is an alternative form to dimensionality reduction and pattern recognition through of the use of recursivity of functions subjected to the parameters that represent the features of the patterns. The proposed model, called *recursive indexing*, uses a recursive process, or composition, of a mathematical function to encapsulate, in a single number, the information that describe a pattern. The goal is that, as indexing variations of an existing pattern, the indexes generated be close.

In order to evaluate the efficiency of the proposed model, the recursive indexing is applied in three different problems of pattern recognition: image indexing, pattern classification with small features set (iris of Fisher and the chromosomes at Denver's group) and digital signal processing. In the signal processing applications (i.e., automatic recognition of the sleep stages), the recursive indexing is used in the preprocessing of the electroencephalogram signal.

1

INTRODUÇÃO

1.1 Origem do Trabalho

Reconhecimento de Padrões é provavelmente uma das áreas da ciência de maior abrangência, estudada em diversos campos de pesquisa como biologia, psicologia, medicina, marketing, finanças, meteorologia, sensoriamento remoto, processamento de imagens, entre outros. Nos últimos anos, mesmo com o crescente desenvolvimento de técnicas inteligentes, como redes neurais, conjuntos difusos, o Reconhecimento de Padrões ainda permanece como um dos grandes desafios a ser superado.

Reconhecimento de padrões (RP) é, por natureza, uma ciência inexata, e assim admite muitas abordagens, algumas vezes complementares, outras competitivas, para aproximar, indicar ou encontrar soluções a um dado problema. A primeira explicação de um fato ou entendimento de um acontecimento novo está sempre ligada aos padrões do mesmo.

Em problemas complexos de RP, a abordagem comum utilizada é dividir o problema em dois módulos subsequentes: a extração de características e o módulo de classificação. Tanto para o primeiro como para o segundo módulo já existem várias técnicas desenvolvidas que vão desde técnicas estatísticas e matemáticas até técnicas de Inteligência Artificial (IA) (Bishop, 1995). Estas técnicas variam dependendo dos tipos de padrões utilizados.

O objetivo principal na teoria de reconhecimento de padrões é a classificação, sendo que a redução da dimensão do espaço dos padrões (extração de características) é um passo intermediário, que nem sempre é necessário. A redução de dimensionalidade pode ser efetuada com três objetivos distintos: reduzir ruídos, extrair informações redundantes, e/ou simplesmente reduzir a dimensão (para os casos onde os padrões são complexos e a dimensão

dos padrões é muito alta, como em imagens) para tornar viável computacionalmente o reconhecimento.

No último caso, a redução da dimensão dos padrões de entradas é, provavelmente, o aspecto mais importante do processo de RP. Com uma seleção de características apropriada e eficiente, tanto um algoritmo de classificação simples como um sofisticado pode ser implementado devido à redução da dimensão do espaço. Além disto, a seleção das características no espaço reduzido pode classificar melhor que no espaço dos padrões e possivelmente permitirá superfícies de decisão mais simples (função discriminante linear *versus* de alta ordem). Com uma extração de características inadequada ou ineficiente, os algoritmos de classificação necessariamente serão menos eficientes e os erros de classificação aumentarão. A grande questão na redução da dimensão é transformar um problema de classificação de um espaço em outro de dimensão menor, que ainda mantenha poder discriminatório.

Na literatura, existem duas abordagens diferentes para problemas de extração de características: (a) transformações lineares ou não do espaço de característica, a qual para cada conjunto original de características de entrada leva para um vetor de características de dimensão menor, como por exemplo, análise discriminante e componentes principais; (b) categorização do espaço de vetores original, no qual o código de cada categoria de um vetor de entrada pode ser usado como a representação de menor dimensão como, por exemplo, a rede de Kohonen e *K-means*.

Muitas questões sobre a redução de dimensionalidade continuam abertas: quais características dos padrões devem ser consideradas? Como a categorização deve ser realizada (supervisionada ou não)? Porém, quando se está trabalhando com padrões de dimensão muito alta, tem-se, além dos problemas de eficiência, os problemas do tempo computacional e da complexidade algorítmica.

Dentre os problemas de RP, o processamento de imagens é um dos grandes desafios para os pesquisadores. As imagens são padrões complexos, apresentam alta dimensão (na ordem de centenas ou milhares de características) e conseqüentemente são difíceis de processar. Por outro lado, o processo de indexação e recuperação de imagens vem se tornando cada vez mais importante e necessário devido ao advento da tecnologia de multimídia.

Um dos desafios encontrados no trabalho de projetar um sistema de indexação e recuperação de imagens é identificar um conjunto mínimo de características e operadores que realizam esta tarefa mais adequadamente. Muitas técnicas são usadas na tentativa de obter uma representação adequada das imagens, que viabilize o processo de recuperação. As tecnologias de visão computacional (*computer vision*) e técnicas de processamento de imagens estão sendo usadas para extrair automaticamente as características da imagem em um processo automático de geração de descrições. Mas estes sistemas ainda apresentam resultados, muitas vezes, insatisfatórios.

Outro problema bastante complexo em RP é o processamento de sinais digitais. As razões específicas para o processamento de um sinal digital podem ser, por exemplo, remover interferências ou ruídos do sinal, obter o espectro dos dados, transformar o sinal em uma forma mais desejada ou extrair informações destes sinais. A sua importância está na crescente aplicação em campos bastante variados como processamento da voz, telecomunicações, sistemas de radar, sinais biomédicos, entre outros.

Um deste sinais, e que será tratado neste trabalho, provém do eletroencefalograma (EEG). Esses sinais possuem uma quantidade de informação maior do que o necessário para a sua utilização. A extração das informações necessárias para uma determinada aplicação é um processo subjetivo e complicado. Este processo é dependente da relação entre informação e redundância.

Existem na literatura várias técnicas que trabalham estes sinais (Schowengerdt, 1997; Gomes e Velho, 1994; Pandya e Macy, 1995): convolução, correlação, filtros (lineares, estatísticos, gradientes), transformadas (Fourier, wavelet, Laplace), técnicas de extração de característica e de classificação. Porém, existem sinais que apresentam uma alta complexidade e que necessitam de novas abordagens de processamento para serem entendidos e utilizados com maior eficiência.

Sendo a redução de dimensionalidade um problema ainda não totalmente resolvido em RP (Carreira-Perpiñán, 1997), propõe-se neste trabalho uma forma alternativa de abordá-lo que possa na sua simplicidade conter informação discriminatória o suficiente.

Com a idéia básica de fornecer uma forma alternativa de redução de dimensionalidade que melhore o poder discriminatório dos classificadores, de forma a aumentar sua eficiência, desenvolveu-se este trabalho.

1.2 Objetivos

Este trabalho tem como objetivo principal o desenvolvimento de uma metodologia para a redução de dimensionalidade de um conjunto de características. Esta redução de dimensão é fundamentada na construção de um índice (número real) através da recursividade de uma função matemática sujeita aos parâmetros (características) que compõem o padrão. A finalidade desta alternativa é aumentar a eficiência na tomada de decisão numa situação de problema.

Como objetivos específicos têm-se:

- Manipular funções matemáticas e sua composição para fornecer uma redução de dimensionalidade;
- Analisar o desempenho de índices quando são aplicados a problemas de classificação de padrões com poucas características: Íris de Fisher e Cromossomos nos grupos de Denver;
- Analisar o desempenho de índices quando são aplicados à indexação e recuperação de imagens;
- Analisar o desempenho de índices quando aplicados aos sinais de EEG.

1.3 Justificativa e Importância do Trabalho

A redução de dimensionalidade, estudada em quase todo sistema de reconhecimento de padrões, com dados em espaços de alta dimensão, exerce um papel fundamental e em parte desconhecido para muitos tipos de dados. Pode-se citar a indexação e recuperação de imagens como um exemplo de sistema de reconhecimento de padrões que tem em sua redução de dimensionalidade o problema fundamental.

A importância de realizar um pré-processamento no sentido de redução da dimensão de espaços está intimamente ligada ao poder discriminatório em ambos os espaços: o original e o reduzido. O espaço reduzido muitas vezes é de dimensão maior que um. Uma pergunta é natural neste contexto: Qual a dimensão necessária para representar os “padrões originais”? Escolher um espaço onde as relações intrínsecas entre as características dos padrões possam

ser representadas é o principal objetivo de uma técnica de redução e reconhecimento de padrões.

O que se propõe é um método de redução de dimensionalidade que possa, na sua simplicidade, conter informação discriminatória suficiente. Esta redução não elimina ou considera uma característica mais relevante que a outra, pelo contrário, utiliza todas as características disponíveis na representação do padrão complexo. Não é uma teoria do simplismo, mas um estudo algorítmico das funções e suas recursões na formação de índices (números reais) com poder discriminatório.

De acordo com Pao (1989, prefácio) *“Nenhum aspecto ou seqüência de aspectos, de um padrão determina o significado daquele padrão. Ao invés disso, existe uma grande compensação entre todos os aspectos do padrão de modo que o significado de um padrão pode ser entendido somente se todos os aspectos estão disponíveis simultaneamente para a consideração”*.

O objetivo, então, é construir uma forma simples de representar a informação, mas que contenha poder discriminatório. E que mantenha a informação com significado. Pretende-se construir um processo que possa reduzir a dimensão e que mantenha os semelhantes, em ambos os espaços, próximos.

O foco deste estudo está atrelado a várias considerações relativas ao reconhecimento de padrões: primeiro, qual o conjunto mínimo de características necessárias para representar classes de padrões? segundo, qual a qualidade na dimensão reduzida *versus* a velocidade de processamento? e finalmente, qual o significado da informação contida em apenas um índice?

A importância deste trabalho consiste no desenvolvimento de uma nova forma de redução de dimensionalidade não supervisionada que poderá ser aplicada para indexação de imagens simples, para classificação de padrões com poucas características ou como uma forma alternativa de redução de dimensionalidade.

1.4 Metodologia de Desenvolvimento da Pesquisa

O estudo em questão foi dividido em diversas etapas, que juntas, integram o trabalho desenvolvido. Assim, o estudo teve início partindo de questões objetivas, já apresentadas na seção 1.2 (reconhecimento de padrões e redução de dimensionalidade). O primeiro passo foi

realizar um levantamento bibliográfico (livros, artigos e Internet) acerca dos conceitos, dos desenvolvimentos e das aplicações em reconhecimento de padrões, indexação e recuperação de imagens, mais especificamente sobre o problema de redução de dimensionalidade.

Uma vez conhecidas as técnicas desenvolvidas e mais utilizadas, partiu-se para o desenvolvimento de um modelo alternativo que objetivasse a redução de dimensionalidade baseado na recursividade de uma função contração. O modelo utiliza um processo recursivo sobre as características que representam cada padrão como forma de reduzir a sua dimensão mantendo o poder discriminatório.

Tendo desenvolvido o modelo, o próximo passo era a sua validação, e para tanto, o modelo proposto foi aplicado a dois problemas clássicos de reconhecimento de padrões: classificação da Íris de Fisher e a classificação dos cromossomos nos grupos de Denver. O modelo também foi utilizado em outras duas aplicações: classificação dos estágios do sono utilizando um canal do eletroencefalograma e Indexação de Imagens (especificamente indexação de faces humanas).

O trabalho é encerrado apresentando as conclusões obtidas dos experimentos realizados nas aplicações do modelo.

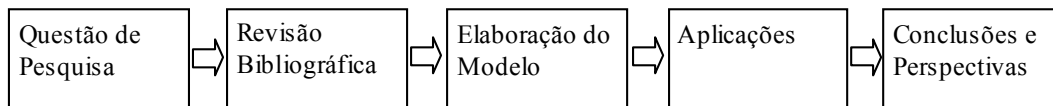


Figura 1: Desenvolvimento da Pesquisa

1.5 Estrutura do Trabalho

Este trabalho está estruturado em 7 capítulos:

No primeiro capítulo são descritos a origem, os objetivos e as justificativas, assim como a estrutura do trabalho realizado.

O segundo e terceiro capítulos apresentam uma revisão teórica dos assuntos estudados cujas teorias auxiliaram no desenvolvimento do novo método. O segundo trata da revisão teórica sobre reconhecimento de padrões, com ênfase na redução de dimensionalidade, já o terceiro capítulo refere-se à indexação e recuperação de imagens num problema de reconhecimento de padrões.

Uma proposta alternativa de redução de dimensionalidade, baseada em indexação recursiva (INREC), é apresentada no capítulo quatro. No capítulo 5, são apresentadas 3 aplicações da INREC, são elas: uma avaliação do desempenho da INREC para classificação de padrões, indexação e recuperação de imagens e a INREC em uma aplicação prática como parte do pré-processamento para estagiamento automático das fases do sono.

Finalmente, o capítulo 6 traz conclusões e recomendações com relação a esse estudo e o último capítulo apresenta a bibliografia utilizada nesse trabalho.

2

RECONHECIMENTO DE PADRÕES

2.1 Introdução

A abrangência do reconhecimento de padrões está na sua efetiva realização para as diversas áreas de pesquisas. A área abrange desde a detecção de padrões à escolha mais simples entre dois objetos, como a complexa realização da aprendizagem.

O reconhecimento de padrões está ligado aos aspectos associados a problemas específicos, como por exemplo, redução de dimensionalidade, extração de característica e escolha de um classificador adequado. Com o objetivo de analisar tais problemas, realiza-se uma pequena revisão dos principais aspectos relacionados às características do que é reconhecimento de padrões.

Existem muitas definições e muitas abordagens sobre Reconhecimento de Padrões (RP). Duda e Hart, (1973), caracterizaram RP como: “*campo que consiste no reconhecimento de regularidades significativas em meios ruidosos e complexos.*” Já Bezdek e Pal (1992) definem RP como: “a busca por estruturas em dados”.

Segundo Pao (1989), o conhecimento de reconhecimento de padrões é importante devido às ocorrências na vida humana tomarem forma de padrões. A formação da linguagem, o modo de falar, o desenho das figuras, o entendimento das imagens, tudo envolve padrões. RP é uma tarefa complexa, onde o homem busca, sempre, avaliar as situações em termos dos padrões das circunstâncias que as constituem, descobrir relações existentes no meio, para melhor entendê-lo e adaptar-se.

O conhecimento de RP é freqüentemente altamente especializado. As contribuições têm vindo de muitas áreas de pesquisas distintas (como: sistemas e processamento de sinais,

inteligência artificial, modelagem conexionista, teoria de estimação/otimização, conjuntos difusos, modelagem estrutural, linguagem formal, etc) (Kasabov, 1996). Isto é, por si só, uma indicação do sucesso da extensão e profundidade de interesse no tópico e do vigor das pesquisas associadas.

2.2 Aplicações de Reconhecimento de Padrões

O termo reconhecimento de padrões envolve uma gama significativa de problemas de processamento de informação, com grande relevância prática, desde o reconhecimento de voz e de caracteres feitos manualmente, até à detecção de erros em equipamentos ou diagnósticos médicos. Não é possível abordar todas as áreas onde seria útil e, por vezes mesmo necessário, aplicar RP. Deve-se levar em conta que as pessoas aplicam seus próprios métodos de reconhecimento de padrões em praticamente todas as áreas da atividade humana. Na prática computacional, pode-se enumerar algumas áreas onde vem sendo usado RP (Kasabov, 1996; Smith, 1998; Tafner, 1998):

- Análise, segmentação e pré-processamento de imagens;
- Reconhecimento de faces;
- Identificação de impressões digitais;
- Reconhecimento de caracteres;
- Análise de manuscritos;
- Computer Vision;
- Entendimento e reconhecimento de voz;
- Diagnóstico médico.

2.3 Estrutura Típica de um Sistema de RP

Existem, hoje, muitas técnicas de RP desenvolvidas, que se baseiam em técnicas matemáticas, estatísticas e/ou técnicas de Inteligência Artificial (redes Neurais, conjuntos difusos, etc.). Cada uma destas tenta simular o RP de forma distinta.

Independente da técnica usada no RP, Bezdek e Pal (1992) conceitualizam o problema em três estágios ou espaços: o espaço do padrão, o espaço das características e o espaço da classificação (Figura 2). O mundo físico é representado por um contínuo de parâmetros que é essencialmente de dimensionalidade infinita (Andrews, 1972). Porém, representa-se o problema do mundo real por R características. Esta é a dimensionalidade do espaço dos padrões. Como R é, geralmente, muito grande, é desejável reduzi-lo (redução de dimensionalidade) de forma que os dados resultantes ainda mantenham o poder discriminatório dos padrões que estão inerentes aos dados. O espaço das características é postulado de dimensão N ($N < R$) no qual as regras de classificação podem ser executadas em tempo razoável. O espaço de classificação, então, simplesmente é o espaço de decisão no qual K classes podem ser selecionadas e, portanto, de dimensão K.

Conceitualmente, o problema de RP pode, então, ser descrito como uma transferência do espaço de padrões P (dimensão R), para o espaço de características F (dimensão N) e finalmente para o espaço de classificação C (dimensão K):

$$P \rightarrow F \rightarrow C \quad (2.1)$$

Estas transformações são altamente não lineares e não inversíveis, mas o único objetivo é manter o poder discriminatório dos padrões. A Figura 2 ilustra um caminho razoável para um sistema de classificação de padrões.

Neste processo encontram-se duas transformações:

- 1º) Redução de Dimensionalidade;
- 2º) Classificação dos dados em K classes

Tanto para a primeira, como para a segunda transformação, existem muitas ferramentas desenvolvidas que são utilizadas para resolver problemas de RP reais. Porém, ainda não existe uma técnica universal.

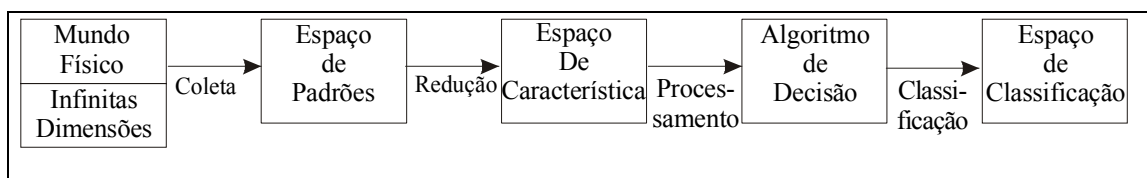


Figura 2 : O problema de reconhecimento de padrões

Fonte: Adaptado de Andrews, (1972)

Dois elementos essenciais dentro de reconhecimento de padrões devem ser definidos:

- **Padrão:** É um conjunto de características que definem um objeto ou um grupo de objetos. É essencialmente um arranjo ou uma ordenação em que alguma organização de estrutura pode ser dita existir (Pandya, 1996). Um padrão pode ser referenciado como uma quantidade ou descrição estrutural de um objeto ou algum outro item de interesse. Um padrão pode ser tão básico quanto um conjunto de medidas ou observações, geralmente sendo representado na forma de vetor ou matriz. O mundo pode ser visto como feito de padrões.
- **Características:** Podem ser entendidas como qualquer medição útil extraída no processo de identificação do padrão. Intensidade de sinais são exemplos de *características*. As *características* podem ser simbólicas, numéricas ou ambas. Podem ser variáveis contínuas ou discretas.

Sendo o objetivo principal do RP a distinção entre os diferentes tipos de padrões, a capacidade de discriminação é colocada em questão. Diferentes padrões são compostos de características diferentes, com valores numéricos diferentes ou relação entre as próprias características diferentes. A maioria dos classificadores (Estatísticos, Sintáticos ou Neurais) são baseados no conceito de similaridade. Por exemplo, se um padrão x é muito similar a outro padrão que é conhecido pertencer a uma classe c_1 , intuitivamente a tendência é classificar x como pertencente à classe c_1 .

O problema em reconhecimento de padrões está na sua definição ou composição, já que definir um conjunto de características que o representa pode não ser uma tarefa trivial. A chave é escolher e extrair um conjunto finito de características que o represente totalmente e que possa ser passível de ser manuseado (Bezdek e Pal, 1992).

2.3.1 Redução de Dimensionalidade

O termo "*The course of dimensionality*" foi criado por Bellman em 1961 (Bellman, 1961). Este termo inclui um número de problemas que surgem com o aumento da dimensão dos dados. Este problema compromete, muitas vezes, a tarefa de reconhecimento de padrões. Por exemplo, os humanos tem uma capacidade excepcional de identificar padrões em espaços

de 1, 2 ou 3 dimensões. Contudo, em espaços de dimensão maior eles não conseguem identificar nem padrões simples. A necessidade de visualização dos dados (pelos humanos) de alta dimensão leva à utilização de algumas técnicas de redução de dimensão.

A redução de dimensão é, também, uma parte importante do processo de reconhecimento de padrões (Poston e Marchette, 1998). Em muitas aplicações, numerosas características são obtidas afim de assegurar uma classificação acurada das classes desconhecidas. Algumas vezes, o número de características deve ser reduzida antes de usar um classificador pois o método de classificação utilizado pode se tornar computacionalmente e/ou analiticamente intratável (Poston e Marchette, 1998).

O problema de *redução de dimensionalidade* acentua-se em espaços de alta dimensão e como uma ferramenta de modelagem para tais dados (Carreira-Perpiñán, 1997). Busca-se um espaço de dimensão menor que ainda represente a informação contida nestes dados. É razoável pensar assim, pois a quantidade de dados muitas vezes é confundida com a qualidade das relações existentes entre os mesmos.

Na literatura existente, a redução de dimensionalidade está intimamente ligada à capacidade dos algoritmos em utilizar o espaço das características para promover uma boa classificação. Portanto, é útil ou mesmo necessário primeiro fazer um pré-processamento no sentido de reduzir a dimensão das características (dimensão dos dados) para um tamanho manuseável, ajustando tanto as informações originais quanto possível. Pode-se pensar que um fenômeno representado por um espaço de padrões de alta dimensão, e assim complexo, pode realmente ser governado por umas poucas variáveis (espaço de características).

Os métodos de redução de dimensionalidade encontrados na literatura (Carreira-Perpiñán, 1997; Vijayakumar e Schaal, 1997) estudam diversas técnicas que permitem reduzir espaços de altas dimensões para espaços de dimensões menores. A aplicação de tais técnicas ocorre freqüentemente em problemas de reconhecimento de imagem, voz, etc.. Dentre as diversas técnicas, tem-se: análise de componentes principais, análise fatorial, análise discriminante, curvas e superfícies principais e a rede de Kohonen. Estas técnicas podem ser encontradas em detalhes nos trabalhos referenciados com suas respectivas justificativas matemáticas detalhadas.

Algumas das técnicas mais utilizadas para a redução de dimensionalidade estão descritas a seguir:

- **Análise de Componentes Principais (ACP):** Também conhecida como transformação de Karhunen-Loeve, busca uma representação de dimensão menor em vetores ortogonais, onde os vetores de entrada possam ser projetados sem perda de generalidade. O objetivo das primeiras componentes é ter maior representatividade com respeito à matriz de variância-covariâncias obtida dos dados originais. Provinda de um processo de ortogonalização, fornece uma medida quantitativa do erro introduzido por omitir alguns dos vetores desta nova base. Este método é um dos mais conhecidos e usados para redução de dimensionalidade devido à sua simplicidade conceitual e ao fato de existirem algoritmos relativamente eficientes com complexidade polinomial para seu cálculo. Porém, a ACP é somente capaz de encontrar um subespaço linear e então não pode tratar apropriadamente com dados que estão em superfícies não lineares. Além disto, não se sabe com quantas componentes principais deve-se trabalhar, embora algumas regras são aplicadas na prática. Por exemplo, eliminar as componentes cujos os autovalores são menores que uma fração da média (Johnson e Wichern, 1998).
- **Análise fatorial:** A proposta essencial da análise fatorial é descrever, se possível, a relação de covariância entre muitas variáveis em termos de umas poucas fundamentais, mas não observáveis, quantidades randômicas chamadas fatores. Basicamente, o modelo fatorial é motivado pelo seguinte argumento: suponha que as variáveis possam se agrupar por suas correlações. Isto é, todas as variáveis em um grupo particular são altamente correlacionadas entre elas mesmas, mas têm relativamente pouca correlação com variáveis em um grupo diferente. É concebível que cada grupo de variáveis represente uma construção básica simples, ou fator, que é responsável pela correlação observável. A análise fatorial pode ser considerada como uma extensão da análise de componentes principais. Ambas podem ser vistas como tentativas para aproximar a matriz de covariância. A questão em análise fatorial é se os dados são consistentes com uma estrutura prescrita (Johnson e Wichern, 1998; Ghahramani e Hinton, 1996).
- **Rede de Kohonen (SOMs – Self-Organizing Maps):** Os mapas auto-organizáveis de Kohonen são uma técnica principalmente usada para visualização dos dados de alta dimensão. Esta rede neural é também considerada uma forma de redução de dimensionalidade no sentido que ela aprende, de uma forma não supervisionada, um mapeamento entre um espaço l -dimensional e o espaço dos padrões (Carreira-Perpiñán, 1997). O mapeamento do arranjo l -dimensional no espaço dos dados é topologicamente

contínuo. Pode-se visualizar o processo de aprendizagem como um subplano que se torce em torno dele mesmo no espaço dos padrões d-dimensional para assemelhar-se o quanto possível da distribuição dos vetores de dados. A aprendizagem é realizada de forma competitiva sobre todos os padrões de entrada até a rede convergir. Eles têm sido utilizados com sucesso em muitas aplicações práticas. Particularmente na visualização. Entretanto, devido à sua natureza heurística, eles têm algumas desvantagens, no sentido que nenhuma função de custo para otimizar pode ser definida, e não existe prova e nem garantia da convergência (Haykin, 1994).

A análise de conjunto de dados de alta dimensão é freqüentemente uma tarefa complexa. Os sistemas de RP sofrem do problema da alta dimensionalidade. O processo de extração de características para um estágio no qual os objetos são classificados pode ser considerado como uma redução em dimensionalidade utilizando uma transformação de variáveis onde algumas variáveis são mais significativas que outras. Estas variáveis mais importantes podem ser pensadas como características. Se aquelas variáveis de menor importância podem ser ignoradas, então a redução desejada em dimensionalidade é alcançada.

O processo de extração de características transforma o espaço dos dados (espaço do padrão) em um espaço de características que é de dimensão muito menor comparado com o espaço de dados original, ainda retendo a maioria do conteúdo de informação intrínseca dos dados.

2.3.2 Abordagens para classificação

O Reconhecimento de Padrões, assim como a Inteligência Artificial, teve seu início nos anos 50. O RP inicialmente utilizou-se de técnicas probabilísticas, mais precisamente, da Estimação e Teoria da Decisão para sua fundamentação. Assim, usou a linguagem da Probabilidade em sua origem, usando a abordagem Bayesiana (Schalkoff, 1992). Uma outra corrente apareceu depois, utilizando a abordagem estrutural, sendo usado o modelo sintático. Mais recentemente, a Modelagem Neural e Difusa apareceram como abordagens promissoras para esta área. Assim pode-se dividir os métodos de classificação em quatro grupos: RP estatístico, RP sintático, RP neural e RP difuso (Pao, 1989; Fu, 1994b ; Bishop, 1995).

2.3.2.1 *Abordagens estatísticas*

Para a abordagem de tomada de decisão, um conjunto de medidas de características é extraído dos padrões. O reconhecimento de cada padrão, ou seja, a designação de uma classe padrão, é normalmente feita particionando o espaço de características. A informação estrutural que descreve cada padrão é importante e o processo inclui, não somente a capacidade de associar o padrão a uma classe particular e classificá-lo, mas também a capacidade de descrever aspectos do padrão que o tornam inelegível para designá-lo para a outra classe.

Robert Schalkoff (1992) diz que o reconhecimento de padrões estatístico, como sugere o próprio nome, assume uma base estatística para os algoritmos de classificação. Um conjunto de medidas, que denotam as características, é extraído dos dados de entrada e usado para associar cada vetor de características a uma dentre c classes. Presume-se que as características são geradas por um estado da natureza, e, portanto, existe um modelo subordinado a um estado com um conjunto de probabilidades e/ou funções de densidade de probabilidade correspondente, passível de ser usado para representá-lo.

Segundo Bishop (1995), a forma mais geral e natural de formular soluções para o reconhecimento de padrões é o RP estatístico, através do qual é reconhecida a natureza estatística tanto da informação que se quer representar quanto dos resultados que devem ser expressos.

2.3.2.2 *Abordagem Sintática*

Muitas vezes, as informações significativas em um padrão não consistem apenas na presença ou ausência de valores numéricos de um conjunto de características. Ao invés disto, a inter-relação ou inter-conexão das características produzem informações estruturais importantes, que facilitam a descrição ou classificação. Esta é a base do RP sintático. Portanto é necessário, ao se usar a abordagem de RP sintático, quantificar e extrair as informações estruturais importantes para poder avaliar a similaridade estrutural entre os padrões.

Para a abordagem sintática, ao enfatizar a descrição estrutural dos padrões, tenta-se fazer uma analogia entre as estruturas do padrão e a sintaxe de uma linguagem. Tipicamente, o RP sintático formula uma descrição hierárquica de padrões complexos, construída a partir de sub-padrões mais simples, sendo que no nível mais baixo se encontram os elementos mais

simples, extraídos dos dados de entrada que são chamados de primitivas (Fu, 1982). Por exemplo, as frases e as sentenças são construídas juntando-se palavras, e palavras são construídas juntando-se letras.

2.3.2.3 Abordagem Neural

Com a capacidade intrínseca de aprender, e o fato de serem aproximadores universais, as redes neurais surgiram como uma ferramenta poderosa na área de reconhecimento de padrões. Sua capacidade de fazer suposições mais delicadas a respeito da distribuição dos dados de entrada do que métodos estatísticos tradicionais e a capacidade de formar fronteiras de decisão altamente não-lineares no espaço de características, levou ao seu crescente uso. Além disto, as redes neurais são usadas para a implementação de alguns modelos estatísticos, como a ACP.

Uma rede neural, também conhecida como rede conexionista, é definida em Cichoki e Unbehauen (1993), como um sistema de processamento de sinal ou informação, composto por um grande número de elementos simples de processamento, chamados neurônios artificiais, ou simplesmente nós, que são interconectados por elos diretos chamados de conexões, que colaboram para realizar um processamento paralelo distribuído (PPD) para realizar uma tarefa computacional desejada.

A tarefa computacional citada no parágrafo anterior pode ser, por exemplo, a classificação de padrões, ou seja: dado um vetor de características, deve-se associá-lo a uma classe em um determinado problema (*i. e.*, reconhecimento de padrões). Na maioria das aplicações de RP usando-se redes neurais, são estabelecidas conexões entre os valores de todas as características que definem os padrões com camadas intermediárias de neurônios e a todas as classes objetivos que são consideradas como a camada de saída. O treinamento da rede é realizado corrigindo os pesos nas conexões para estabelecer-se as relações entre as características e classes que promovam a melhor discriminação possível entre os padrões de classes diferentes. Assim, ao ser apresentado à rede um novo padrão, esta indicará a classe que o representa melhor na camada de saída.

Embora os primeiros desenvolvimentos de aplicações com sucesso datem dos anos cinquenta, a introdução das rede neurais no comércio em larga escala tem seu início nos meados dos anos oitenta com as redes MLP (Multi Layer Perceptron). Hoje, as redes mais

usadas em RP são: perceptron multicamadas (MLP) com algoritmo de aprendizagem Backpropagation, a rede de base radial (RBF), a rede LVQ (Learning Vector Quantization), e os mapas topológicos de Kohonen (SOM) (Bishop, 1995; Haykin, 1994; Omidvar e Dayhoff, 1998).

2.3.2.4 Abordagem Difusa

A utilidade da teoria dos conjuntos difusos no reconhecimento de padrões já foi reconhecida em meados da década de 60, e hoje a literatura a respeito de RP difuso é bastante extensa, segundo Klir (1995). Existem muitas maneiras de se aplicar a teoria dos conjuntos difusos em RP. Existem duas formas clássicas de RP difuso, são elas: os métodos de reconhecimento de padrões através de listas de pertinências e os métodos difusos sintáticos.

No método clássico da lista de pertinência, cada classe é caracterizada por um conjunto de padrões que é armazenado no sistema de reconhecimento de padrões. Um padrão desconhecido a ser classificado é comparado com os padrões armazenados um a um. O padrão é classificado como um membro de uma classe se ele combinar com um dos padrões pertencentes a esta classe.

No método sintático difuso, um padrão é representado por uma cadeia de sub-padrões concatenados chamados de primitivas. Estas primitivas são vistas como o alfabeto de uma linguagem formal. Um padrão, então, é uma sentença gerada por alguma gramática. Todos os padrões, cujas sentenças são geradas pela mesma gramática, pertencem a uma mesma classe. Um padrão desconhecido é classificado como pertencente a uma classe particular se ele puder ser gerado pela gramática correspondente a ela (Klir,1995).

2.4 As tendências de RP

Estimulada pelo desenvolvimento do computador digital, a área de Reconhecimento de Padrões floresceu no início dos anos 60. As contribuições vêm de muitas áreas incluindo estatística, teoria da comunicação, teoria de controle, pesquisa operacional, biologia, psicologia, lingüística, ciência da computação e inteligência artificial. Esta diversidade apresenta sérios problemas para as pessoas que querem escrever sobre o assunto (Duda e Hart, 1973). Nenhuma teoria simples de RP envolve todos os tópicos, pois cada domínio de aplicação tem características únicas que modelam a abordagem apropriada. Pode-se dizer que

a área de Reconhecimento de Padrões é uma conseqüência direta de suas aplicações. Cada área de aplicação contribui para o seu desenvolvimento.

A teoria mais proeminente independente do domínio de aplicação é a teoria da classificação. Baseada na teoria de decisão estatística, fornece procedimentos matemáticos formais para a classificação de padrões uma vez que eles são representados abstratamente como vetores (Duda e Hart, 1973).

Nas últimas décadas e principalmente nos últimos anos, as Redes Neurais Artificiais estão sendo cada vez mais usadas como classificadores e também estão sendo desenvolvidas novas redes neurais que simulam os métodos estatísticos (como a rede bayesiana e a probabilística) (Pao, 1989; Bishop, 1995; Pandya e Macy, 1995; Omidvar e Dayhoff, 1998).

Assim, como é muito difícil descrever a sua história devido a sua dependência das aplicações, também é muito difícil dizer quais são as tendências futuras de RP. Mas se pode dizer que uma das áreas que tem recebido mais atenção e, que continuará recebendo, é o domínio de imagens. Além disso, os trabalhos neste domínio tem progredido de classificação de imagens para análise e descrição de imagens (Hermes et al., 1995; Ravela e Manmatha, 1998). O desenvolvimento da tecnologia computacional em termos de hardware e software tem tornado possível a captura e o armazenamento de imagens (como na área médica e em sensoriamento remoto). O mundo WEB também reflete este desenvolvimento, onde cada vez mais imagens são utilizadas. Porém, a complexidade, a dimensão e a grande quantidade de imagens utilizadas requerem técnicas de RP cada vez mais eficientes (rapidez e acurácia) para o seu processamento (classificação, análise, descrição, indexação e recuperação).

Uma outra área proeminente, que também requer técnicas de RP eficientes é a área de Data Mining¹ (Berry e Linoff, 1997; Groth, 1998).

A presença da complexidade nos padrões refere-se normalmente a uma representação em espaços de alta dimensão, levando ao problema da dimensionalidade. A redução de dimensionalidade torna-se uma parte do processo de reconhecimento de padrões que precisa ser melhor explorado. Trabalhos recentes nesta linha mostram a preocupação e a necessidade

¹ Data Mining é a exploração e a análise, por meios automáticos e semi-automáticos, de grandes quantidades de dados com o objetivo de descobrir padrões significativos e regra (Berry e Linoff, 1997).

de novas abordagens para o tratamento da alta dimensionalidade no processo de reconhecimento de padrões (Carreira –Perpiñán, 1997; Poston e Marchette, 1998).

2.5 Conclusão

Apesar das muitas técnicas desenvolvidas e utilizadas na prática, tanto no que tange à redução de dimensão quanto à classificação propriamente dita, ainda não existe uma abordagem universal. Os modelos são, de modo geral, desenvolvidos para resolver problemas específicos.

Em Andrews (1972) e Fu (1982) pode-se encontrar técnicas matemáticas de RP, em Johnson e Wichern (1998) tem-se algumas técnicas estatísticas, em Bishop (1996) encontram-se as técnicas de Redes Neurais, já em Schalkoff (1992) e Pao (1989) tem-se técnicas matemáticas, estatísticas, Redes Neurais e conjuntos difusos.

O grande problema do RP reside nos padrões de alta dimensão, como é o caso do reconhecimento de imagens e vídeos. Nestes casos, a aplicação de uma técnica de redução da dimensão é extremamente importante. O capítulo seguinte apresenta uma revisão sobre alguns métodos usados no processo de indexação e recuperação de imagens, onde se pode concluir que os métodos utilizados em imagens ainda são limitados.

Existem diversas formas de representação e um espaço que se julga necessário para discriminá-las. Matematicamente sempre é possível encontrar um espaço de dimensão alta o suficiente onde os padrões (dados) são totalmente discriminados (separáveis linearmente), porém, este tipo de espaço não interessa, principalmente devido às limitações computacionais. Procura-se o inverso, reduzir o espaço o máximo possível e ainda manter o poder classificatório.

3

INDEXAÇÃO E RECUPERAÇÃO DE IMAGENS

3.1 Introdução

A extração de característica e, principalmente, a redução de dimensionalidade tornam-se ainda mais críticas em problemas de indexação e recuperação de imagens. Nestes problemas, as dimensões do espaço e a natureza dos padrões potencializam a necessidade de métodos eficientes na redução de dimensionalidade. Este capítulo apresenta uma revisão sobre indexação e recuperação de imagens.

O processamento de imagens é um problema genérico de RP. Existem diferentes tarefas de processamento de imagens como: reconhecimento de imagens, compressão de imagens e análise de imagens.

O reconhecimento de imagens associa uma nova imagem com uma já existente ou com uma classe de imagens. O processo de reconhecimento é difícil, pois as imagens apresentam ruídos, ou são corrompidas, e/ou não são nítidas. Além disto, o espaço das imagens é de dimensão muito alta. As imagens são comprimidas com o objetivo de codificar a imagem com um número menor de bits por pixel de tal modo que um processo de decodificação reconstrua a imagem de modo satisfatório. A extração de características e a segmentação de imagens são áreas que fazem parte da análise de imagens.

O processamento de imagens é usado com dois objetivos distintos: melhorar a aparência visual das imagens e preparar as imagens para avaliar suas características e estruturas presentes. O enfoque neste trabalho refere-se ao segundo, ou seja, analisar as informações que a imagem contém.

O significativo crescimento na quantidade e tamanhos de coleções de imagens digitais faz com que o desenvolvimento de ferramentas para a indexação de imagens torne-se uma necessidade. A fim de fornecer um sistema com capacidade de buscas de imagens e vídeos, técnicas de indexação e recuperação de imagens baseado em conteúdo vem sendo desenvolvidas. Pesquisas e discussões sobre abordagens e sistemas para recuperação de imagem foram publicadas em Nagy (1985), Cawkell (1992), Jain (1992), Schmid e Mohr (1997), Ravela e Manmatha (1997). As aplicações potenciais incluem, base de dados de imagem/vídeo, livrarias digitais, TV digital, sensoriamento remoto, medicina, etc. (Mandal et al., 1996).

Os sistemas de indexação e recuperação de imagens geralmente envolvem técnicas de compressão, análise e recuperação de imagens. A análise é realizada para extrair características como cor, textura e forma das imagens. Uma das técnicas usadas é o histograma de cor (que será descrito mais tarde, neste capítulo) (Valkealahti e Oja, 1998).

Os sistemas de informação que contêm funções de indexação e recuperação de imagens, apesar de ocorrerem cada vez com maior frequência, ainda não conseguem obter resultados completamente satisfatórios (Ravela e Manmatha, 1998). Devido às necessidades únicas e peculiares, as bases de dados de imagens não podem ser tratadas em um estilo similar a outros tipos de livrarias digitais. A dependência contextual apresentada pelas imagens, e a natureza complexa das imagens bidimensionais tornam a questão da representação mais difícil para bases de dados de imagens. Uma representação invariante de uma imagem é ainda uma questão aberta de pesquisa. Por esta razão, é difícil encontrar uma técnica universal de recuperação baseada em conteúdo.

Existe hoje na comunidade científica, um interesse crescente no desenvolvimento de sistemas de recuperação de imagens baseados em suas características. As tecnologias de visão computacional (*computer vision*) e as técnicas de processamento de imagens são usadas para extrair automaticamente as características da imagem em um processo de geração de descrições. Este capítulo visa apresentar alguns estudos desenvolvidos nesta área. Para tal procura-se introduzir algumas técnicas e conceitos gerais e em seguida apresenta-se alguns trabalhos recentes desta área.

3.2 Indexação e Recuperação de Imagens: Uma visão geral

Devido ao advento da tecnologia de multimídia, os estudos mais recentes estão preocupados em encontrar uma forma adequada para a indexação e recuperação de imagens semelhante ao que hoje é feito com dados alfanuméricos.

De um sistema multimídia, espera-se o manuseio eficiente de vários tipos de dados, como textos, imagens, áudios e vídeos. A tecnologia atual de banco de dados pode, facilmente, gerenciar dados alfanuméricos. Contudo, as técnicas orientadas aos textos não são suficientes para as aplicações relacionadas à recuperação de imagens.

No início, a idéia corrente para descrição de imagens era a de fazer a anotação de textos gerados pelos usuários. Por exemplo, o título e a legenda das imagens mais alguma descrição adicional, conforme os trabalhos descritos em Leung (1990), Bordogna (1990) e Chakravarthy (1994). Esta abordagem é restrita, devido ao esforço das anotações manuais e as diferentes interpretações dos usuários. Isto conduz a inconsistências nas anotações e uma exigência indesejada: a pessoa que formula as questões tem que adquirir conhecimento relativo aos critérios de geração de anotação. Por exemplo, alguém pode fornecer um exemplo da árvore de Natal e tentar encontrar outro através de textura. Ou, um usuário pode fornecer uma figura de papai Noel e encontrar outros por cor, formato, ou por ambos.

A indexação de imagens significa identificar objetos e calcular as características das imagens. Este é o passo chave para todos os sistemas de indexação de imagens. Até o momento, a maioria dos pesquisadores neste campo estão focalizados sobre a indexação automática do conteúdo, tais como: cor, textura e forma de uma imagem (Hermes et al., 1995; Bach et al., 1996; Ravela e Manmatha, 1998; Pala e Santini, 1999).

O objetivo de um sistema de recuperação de imagem é operar sobre uma coleção de imagens (indexadas) e, em resposta a uma solicitação (*query*), apresentar "imagens relevantes" segundo os critérios estabelecidos. A "*query*" é definida pelo usuário. Ele pode fornecer, por exemplo, as características que as imagens recuperadas deveriam conter ou mesmo um pedaço de uma imagem ou, ainda, uma imagem inteira e o sistema deve então encontrar imagens que podem ser visualmente semelhantes.

Algumas considerações devem ser levadas em conta para que se possa compreender a recuperação de imagens. Inicialmente, é preciso definir a expressão "recuperação de imagens

relevantes". Relevância, para os *usuários* de um sistema de recuperação, está associada à semântica. Por exemplo, um usuário pode querer encontrar todas as figuras de cartões de Natal em uma determinada base de dados. A um cartão de Natal está associado semanticamente, Papai Noel, árvore de Natal, rena, neve e assim por diante. A codificação de informações semânticas em um sistema de recuperação de imagens está associada a solução de problemas tais como, segmentação, reconhecimento e extração automática de característica. Estes são problemas difíceis que, segundo Ravela e Manmatha (1998), ainda são insolúveis.

Uma abordagem proposta freqüentemente para solucionar o problema da relevância vem da observação que, em muitos casos, os atributos associados (conteúdo) a uma imagem, quando usados junto com alguma informação fornecida pelo usuário, captura algum tipo de informação semântica que é desejável (Ravela e Manmatha, 1997).

Em situações mais simples, atributos textuais estão sendo associados a imagens. Porém os trabalhos mais recentes têm focado diretamente sobre o conteúdo da imagem (cor, forma, textura, etc.). O principal objetivo dos pesquisadores têm sido o desenvolvimento de técnicas que permitam armazenar e recuperar imagens baseadas em seu conteúdo. Um diagrama de um sistema de indexação e recuperação de imagens é mostrado na Figura 3.

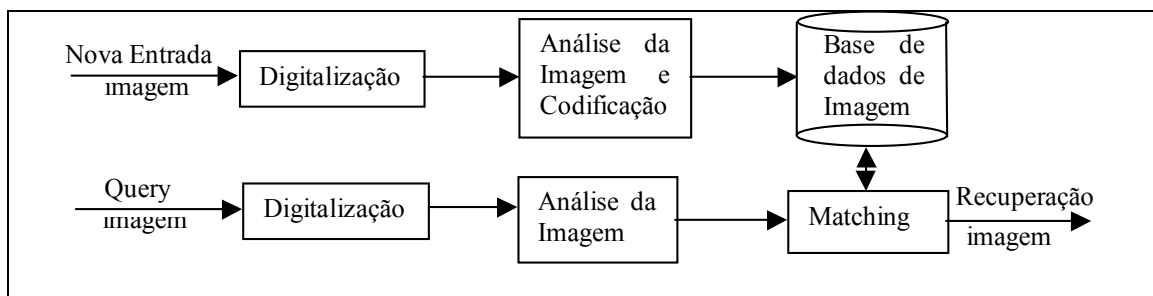


Figura 3: Um esquema de sistema de recuperação e armazenagem de imagem

Fonte: Mandal *et al.* (1996)

Neste contexto, existem dois grupos. O primeiro usa os algoritmos de *computer vision* para gerar a descrição do objeto e armazená-lo em uma forma textual. Essas anotações são transferidas, então, a um sistema de recuperação. O sistema IRIS desenvolvido por Hermes *et al.* (1995) pertence a este grupo e será apresentado nos próximos itens.

O segundo tipo dá ênfase a objetos semelhantes baseado em forma, cor e textura (Brolio *et al.*, 1989; Swain e Ballard, 1990; Chen e Chang, 1993; Niblack *et al.*, 1993;

Binaghi *et al.*, 1994; Pentland *et al.*, 1994). Picard e Minka (1995) usam a textura como uma característica no método de recuperação baseado em conteúdo. São definidas etiquetas baseadas em semânticas para regiões diferentes de uma imagem e somente então calculados os parâmetros dos modelos de textura dentro daquela região. Ao invés de usar apenas um modelo de textura, são empregados vários modelos, e o melhor deles é selecionado para discriminar uma região do resto da imagem. Outro exemplo de recuperação de imagem, por comparação (*matching*) com imagens exemplos, é o projeto QBIC (Niblack *et al.*, 1993). As características das imagens usadas são: cor, textura, e forma, juntos com alguns parâmetros geométricos das regiões.

3.2.1 Indexação de imagem Baseada em conteúdo

Indexar imagens significa identificá-las e calcular parâmetros baseados em suas características para realizar futuras comparações. Este é o passo fundamental para todo sistema de reconhecimento de imagens. Existem duas classes distintas de indexação: a indexação global e a regional (local). A global habilita comparações de imagens inteiras, enquanto que a regional permite comparações por partes das imagens. Ambas as técnicas são muito poderosas para recuperação de imagens, mas são convenientes para tipos diferentes de “*queries*”.

O crescimento gigantesco nos números e tamanhos de imagens digitais torna necessário o desenvolvimento de ferramentas para indexar imagens com o mínimo de restrições possíveis ou mesmo sem restrições (Smith, 1997). Para possibilitar a construção de sistemas eficientes com capacidade de busca de imagens, técnicas de *query* visual baseadas em conteúdo estão sendo desenvolvidas e podem ser encontradas ou descritas na literatura técnica (Hermes *et al.*, 1995; Smith, 1997; Ravela e Manmatha, 1998; Niblack *et al.*, 1993).

Consideram-se imagens sem restrições as que não são restritas a um determinado domínio particular de imagens e são freqüentemente componentes de soluções em reconhecimento de padrões tradicionais (Duda e Hart, 1973). O desenvolvimento de sistemas que analisam, comprimem e recuperam imagens sem restrições é extremamente pertinente devido à quantidade enorme deste tipo de imagem que é acessível e acumulada.

O objetivo das técnicas de indexação de imagens baseadas em conteúdo é que, mesmo na ausência do entendimento da imagem, pode-se obter e utilizar discriminantes que são úteis no processamento das "queries".

Por conteúdo entende-se todos os atributos/características visuais de uma imagem. Uma característica visual, por sua vez, é qualquer propriedade de uma imagem que pode ser calculada usando algoritmos de *computer vision* ou técnicas de processamento de imagens (cor, textura, forma, luminosidade, etc.).

A idéia básica na recuperação de imagens baseada em conteúdo é que o sistema pode procurar e encontrar em um arquivo as imagens que melhor correspondem a uma dada descrição (das características visuais de uma imagem) fornecida pelo usuário.

Com o propósito de fornecer uma rápida comparação de imagens, como ilustra a Figura 4, um sistema de "query" visual baseado em conteúdo, tipicamente extrai, armazena e indexa as características das imagens off-line. Por exemplo, histogramas de cor, que descrevem a distribuição das cores de uma imagem são pré-calculados, armazenados e usados como base para medir a similaridade das imagens em Niblack *et al.* (1993) e Smith e Chang (1996).

A métrica de dissimilaridade (D_f) é usada para produzir valores pequenos para imagens similares e valores grandes para imagens dissimilares. Contudo, a performance deste método é limitada pela qualidade da extração das características. Histogramas de cor e outras medidas de similaridades globais do conteúdo da imagem fornecem informações insuficientes para a comparação das imagens.

O processo de recuperação de imagens baseado em conteúdo tem sido investigado em diversos estudos recentes: o projeto QBIC da IBM (Niblack *et al.*, 1993), o sistema Photobook do MIT (Pentland *et al.*, 1994) e o sistema de recuperação de imagens comercial Virage (Bach *et al.*, 1996). Estes sistemas usam a abordagem baseada em características para indexação de informações visuais.

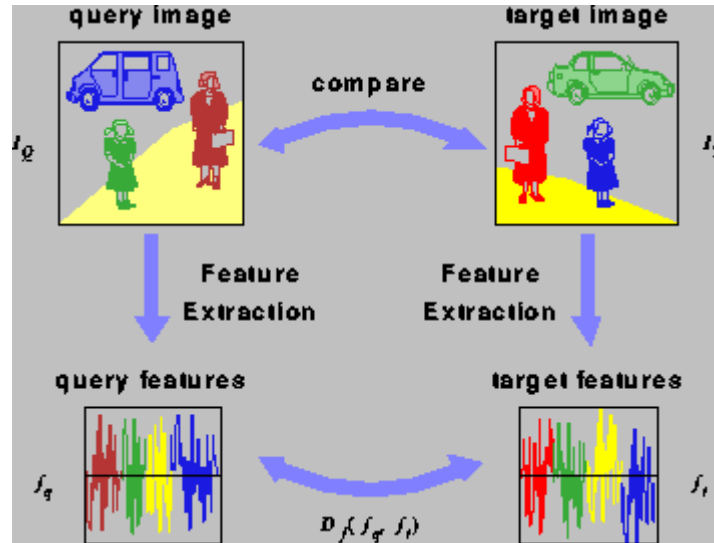


Figura 4: *Query* de uma imagem baseada em conteúdo
 Fonte: Smith, 1997

Quando a base de dados é muito grande e a representação das características das imagens é complexa, a busca exaustiva na base de dados e o cálculo da similaridade das imagens não são eficazes. Diversas técnicas recentes têm sido propostas para aumentar a velocidade da recuperação. Swain e Ballard (1990), Stricker e Orengo (1995) e QBIC (Niblack *et al.*, 1993) pré-calculam e utilizam somente características de imagens compactadas e simples. Petrakis e Faloutsos (1995) e o QBIC usam a combinação de técnicas de redução de dados e estruturas de indexação eficientes para aumentar a velocidade da recuperação. Por outro lado, o VisualSEEK (Smith e Chang, 1996) usa métodos para integrar as informações espaciais das características, a fim de melhorar a recuperação de imagens.

A seguir tem-se uma descrição sobre a evolução dos processos de indexação e recuperação de imagens.

3.2.2 Evolução

Os pesquisadores no passado mostraram muito interesse em banco de dados de imagens. Porém, o hardware disponível na época possuía sérias limitações. Com os avanços na tecnologia de hardware, os bancos de dados não textuais têm recebido mais atenção. Para dados textuais, as questões de representação e recuperação foram extensivamente desenvolvidas. A mesma análise não pode ser estendida para imagens, vídeos e áudios, devido

às dificuldades em se obter uma representação invariante dos dados não textuais. Por exemplo, um objeto é representado em uma base de dados textual pelo seu nome, descrições literais, e outros atributos que o definem com regras sintáticas simples. A *imagem* de um mesmo objeto, por outro lado, pode variar dependendo de sua orientação, luz do ambiente e do sensor, etc. Ainda, as informações contidas em uma imagem, normalmente possuem uma dimensionalidade muito maior que a de informações textuais.

Os primeiros bancos de dados de imagens eram principalmente, usados para dados de sensores remotos (Chang e Liu, 1984; Chock *et al.*, 1984). Recentemente, muitos pesquisadores têm examinado várias questões de projetos para base de dados de imagens, como: Niblack *et al.* (1993), Pentland *et al.* (1994), Smoliar e Zhang (1994), Yoshitaka *et al.* (1994), Wu e Narasimhalu (1994) e Wu *et al.* (1995). As técnicas de recuperação de imagens baseadas em conteúdo usadas nestes sistemas consideram muitas sugestões visuais, tais como: cor (Swain e Ballard, 1991), textura (Picard e Minka, 1995), e forma (Califano e Mohan, 1994). Na seqüência, são comentados alguns desdobramentos importantes no desenvolvimento das técnicas de indexação e recuperação de imagens.

Orenstein e Manola (1988) descrevem uma base de dados de imagens para modelagem de dados espaciais e temporal, em que uma estrutura de dados chamada "*points sets*" é usada para representar imagens e um operador chamado "*geometric filter*" para otimizar as "*queries*" espaciais.

Joseph e Cardenas (1988) propuseram uma linguagem *query* de alto nível, chamada PICQUERY para o gerenciamento de base de dados de imagens. Esta foi mais tarde estendida para PICQUERY⁺ por Cardenas *et al.* (1993) em que aparece uma linguagem "*query*" independente do domínio de alto nível que suporta descrições difusas e imprecisas.

Roussopolous *et al.* (1988) usam árvores R e R⁺ e uma linguagem *query* PSQL para aplicações em base de dados de cartografia. Reconhecimento de objetos baseado em índice tem tornado-se popular na comunidade de *Computer Vision* (Gorsky e Mehrotra, 1990; Califano e Mohan, 1994 e Rigoutsos e Hummel, 1992). As características extraídas de uma imagem são comparadas com características na base de dados. *Geometric hashing* (Lamdan e Wolfson, 1988) tem sido usado para este objetivo. Rigoutsos e Hummel (1992) implementaram *geometric hashing* em uma máquina paralela para reconhecimento de objetos. Califano e Mohan (1994) descreveram uma rede analítica para indexação multidimensional.

Histogramas de cor de objetos multicores propostos por Swain e Ballard (1990) fornecem outro método de indexação dentro de uma base de dados grande de modelos.

O projeto QBIC (Smith e Chang, 1996) integra muitas sugestões tais como: cor, textura e forma para recuperar os conteúdos da base de imagens. Photobook do MIT (Pentland *et al.*, 1994) descreve um conjunto de ferramentas interativas para procurar e buscar imagens e seqüências de imagens. Um conjunto de coeficientes significantes é usado para o matching de imagens. Abordagens similares têm sido estudadas por Picard e Minka (1995) para registrar uma livraria digital de imagens baseadas na textura da imagem. Para indexar e recuperar dados de vídeos, Smoliar e Zhang (1994) usam cor, textura e forma como características e constróem tabelas de índices usando técnicas de representação do conhecimento. Chang e Hsu (1993) discutem a pesquisa em um sistema de informação de imagem e sugerem o uso de ícones generalizados e índices ativos para recuperação baseado em conteúdo. Jain e Vailaya (1995) descrevem uma técnica de recuperação de imagens baseada em cor onde obtiveram resultados excelentes sobre base de dados de imagens *trademark*.

Existe um grande interesse em base de dados de faces. Bach *et al.* (1993) usam as extremidades dos mapas e outras características topológicas das faces (e.g., comprimento do nariz) como características para recuperação de imagens de faces. Wu e Narasimhalu (1994) descrevem métodos similares para recuperação de faces baseada em redes neurais por indexação *iconic*. Apesar de ser difícil designar um objetivo geral aos sistemas de recuperação de imagens baseado em conteúdo, Wu *et al.* (1995) tentaram integrar muitos requerimentos básicos em uma máquina de busca baseada em conteúdo, chamada CORE.

Alguns autores usam uma abordagem geométrica para modelar a imagem por propriedades tais como linhas, vértices, e elipses e tentam extrair estas características afim de reconhecer as imagens (Besl e Jain, 1985; Chin *et al.*, 1986). A maior dificuldade desta abordagem é que eles usam modelos feitos por humanos ou requerem representações parecidas com CAD. Além disto, os métodos geométricos têm dificuldades em descrever objetos "não-geométricos" e eles têm problemas diferentes entre muitos objetos.

Outra abordagem alternativa é usar informação sobre a luminosidade da imagem. A idéia é não impor o que tem de ser visto na imagem (pontos, linhas, ...) mas usar o que é realmente visto na imagem para caracterizar um objeto. A primeira idéia foi usar histograma

de cor (Swain e Ballard, 1991). A abordagem descrita por Schmid e Mohr (1997) usa cálculos de características de níveis de cinza local de pontos de interesse. Pontos de interesse são características locais com grandes quantidades de informações. As características locais usadas neste trabalho são baseadas sobre invariantes diferenciais de níveis de cinza. Estas asseguram invariâncias sob um grupo de deslocamento dentro da imagem.

A abordagem apresentada por Ravela e Manmatha (1998) não depende da segmentação da imagem (manual ou automática) ou extração de características binárias, nem do tempo para treinamento. Eles desenvolveram o sistema Synapse (Engenho de Busca por Aparência Sintática). Neste sistema as imagens são recuperadas usando uma caracterização da aparência visual. Intuitivamente, a aparência visual do objeto em uma imagem está aproximadamente relacionada com a descrição da forma de sua superfície de intensidade. A aparência não depende, somente, da forma geométrica tridimensional do objeto, mas também sobre seu albedo (fração da luz incidente que é difundida pela superfície), sua textura, o ponto de vista do qual é imaginada e vários outros fatores. Usando uma imagem exemplo e a interação do usuário para construir *queries*, Synapse recupera imagens semelhantes dentro de pequena variação de tamanho e vista na ordem de sua semelhança com a aparência sintática para uma query.

Para base de dados pequenas (na ordem de centenas de imagens) a maioria destas técnicas tem mostrado bons desempenhos. Conforme cresce o tamanho da base de dados, estas técnicas não são satisfatórias na prática. Além disto, estas técnicas têm sido testadas em meios controlados. Conseqüentemente, a robustez para ruídos e distorções no meio da vida real ainda não foram demonstrados.

3.3 Métodos Básicos Utilizados no Processo de Indexação

Um dos desafios encontrados no trabalho de projetar um sistema de indexação e recuperação de imagens é o de identificar o conjunto mínimo de características e operadores que realizam esta tarefa mais adequadamente. Muitas técnicas são usadas na tentativa de obter uma representação adequada das imagens que viabilize o processo de recuperação.

A seguir, apresentam-se alguns métodos que são utilizados nos processos de indexação e recuperação de imagens.

3.3.1 PCA (Principal Components Analysis)

A análise de componentes principais (Johnson e Wichern, 1998) é uma das técnicas mais usadas na prática para a redução de dimensionalidade, talvez devido à sua simplicidade conceitual e ao fato de existirem algoritmos relativamente eficientes (de complexidade polinomial) para seu cálculo.

No contexto de indexação de imagem, e em uma aplicação específica, é natural pensar que se pode identificar a imagem utilizando algumas de suas características (Hancock et al., 1992; Weng e Swets, 1999).

Na prática, a análise de componentes principais, primeiro calcula a média (\bar{x}) dos vetores de entrada (x_n) e então subtrai todos da média. Na seqüência, a matriz de covariância é calculada e seus autovetores e autovalores são encontrados. Os autovetores correspondentes aos M autovalores maiores são utilizados e os vetores de entrada (x_n) são projetados sobre estes autovetores para dar as componentes do vetor transformado no espaço M-dimensional. A Figura 5 ilustra os pontos bidimensionais projetados sobre o autovetor v_1 associado ao maior autovalor.

Assim, pode-se descartar as variáveis com menor variância, i. e., projetando sobre o subespaço das primeiras M componentes principais, e obtendo uma boa aproximação para as amostras originais.

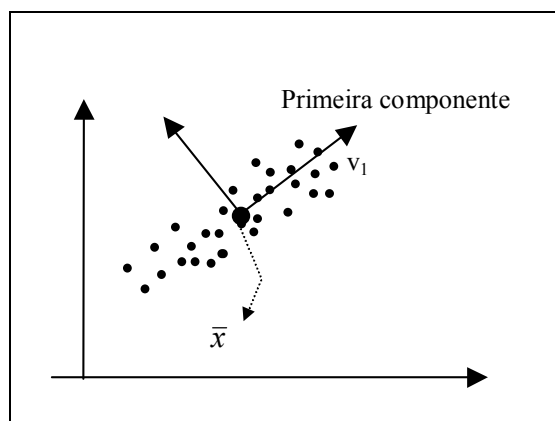


Figura 5: Representação gráfica da primeira (v_1) e segunda componente principal

3.3.2 Histogramas

A comparação de Histogramas tem recentemente tornado-se uma técnica popular para recuperação e indexação de imagens (Swain e Ballard, 1991; Stricker e Orengo, 1995; Lee e Dickinson, 1994; Colombo e Bimbo, 1999). Um histograma de cor determina a distribuição das cores em uma imagem, região ou objeto. Swain e Ballard (1991) propuseram o uso de intercessão de histogramas para o *matching* de objetos de cor, onde a hipótese básica é que as imagens similares terão as suas distribuições de cores similares. Este conceito pode ser estendido para imagens em tons de cinza. A distribuição do nível de cinza é invariante à rotação na imagem e muda suavemente com a translação. Então, a baixa sensibilidade dos histogramas de imagens à câmara e aos movimentos do objeto torna-o uma técnica viável para aplicação de indexação de imagens. Além disto, as técnicas baseadas em histogramas têm uma complexidade baixa quando comparada às técnicas clássicas de reconhecimento de padrões, facilitando a implementação em tempo real.

O histograma da imagem *query* é comparado aos histogramas de todas as imagens na base de dados. As imagens com as menores diferenças de histogramas são selecionadas para o próximo nível de busca.

A distância entre dois histogramas f e g em uma métrica L^p pode ser expressa como:

$$d_{L^k}(f, g) = \sum |f - g|^p \quad (3.1)$$

Tipicamente, as métricas do erro absoluto (L^1) ou erro quadrado (L^2) são usadas em técnicas de indexação de imagens.

3.3.3 Filtros Gaussianos

Assim como o filtro da média, usado para reduzir a quantidade de variação de intensidade entre um pixel e seu vizinho (reduzir ruído na imagem), o operador de alisamento (smoothing) Gaussiano tem a função de remover detalhes e ruídos da imagem e obscurecê-la.

A idéia do filtro da média é simplesmente substituir cada pixel na imagem com a média de seus vizinhos, incluindo ele mesmo. Tem o efeito de eliminar valores de pixels que não são representativos. O filtro da média é normalmente tratado como um filtro de convolução.

A convolução é uma operação matemática que é fundamental para muitos operadores que processam imagens. Ela fornece uma maneira de multiplicar dois *vetores* (matriz) de números, geralmente de tamanhos diferentes, mas de mesma dimensão. Pode ser usada no processamento de imagem para implementar operadores cujos valores dos pixels de saída são combinações lineares de certos valores do pixel de entrada.

No contexto de processamento de imagem, um *vetor* de entrada, geralmente, apresenta a imagem em níveis de cinza, e o segundo *vetor* é menor (conhecido como o núcleo), sendo ambos bidimensionais. A Figura 6 ilustra um exemplo de convolução.

O filtro de alisamento Gaussiano é um operador de convolução bidimensional, similar ao filtro da média, mas usa diferentes núcleos que representam a forma de uma curva gaussiana (forma de sino).

Koenderink e van Doorn (1987) argumentam que a estrutura local de uma imagem pode ser representada usando um conjunto de filtros de derivadas Gaussianas. O uso de filtros das derivativas gaussianas para realçar a aparência é motivado pelo seu uso na descrição da estrutura espacial e sua unicidade em representar o espaço escala de uma função e o fato que as componentes principais de uma imagem são melhores descritas com as Gaussianas e suas derivadas. Existe uma decomposição natural das imagens em função das gaussianas e suas derivadas (Hancock, *et al.*, 1992).

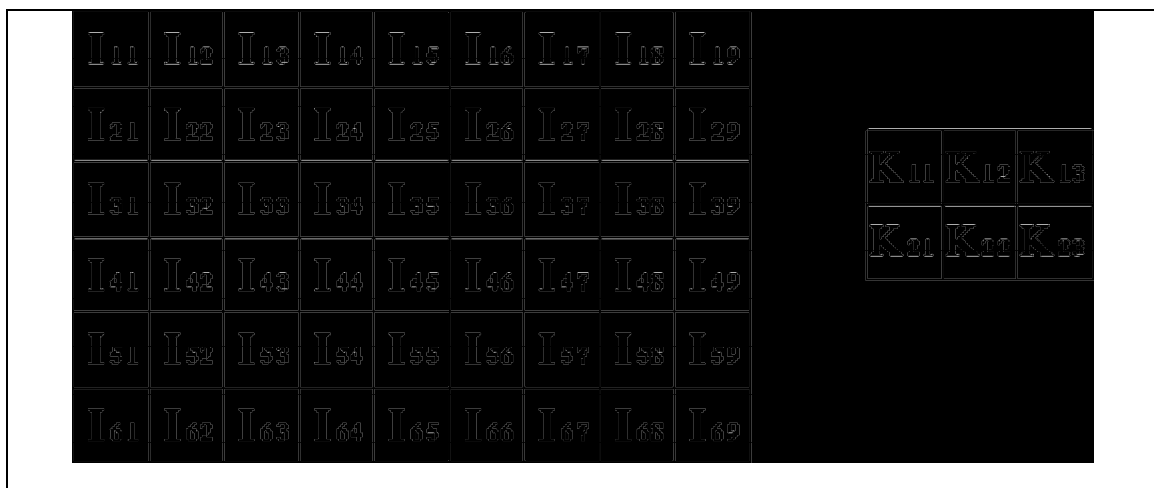


Figura 6: Um exemplo de uma pequena imagem (esquerda) e o núcleo (direita) para ilustrar convolução

Fonte: Fisher *et al.* (1996)

3.3.4 Wavelet

As imagens em uma base de dados são geralmente armazenadas de uma forma comprimida. Os resultados de uma indexação podem ser superiores se as propriedades do esquema de codificação são exploradas em uma técnica de indexação. Recentemente, as transformadas de wavelets discretas (DWT - Discret Wavelet Transform) têm sido bastante usadas em aplicações de codificação de imagens (Woods, 1991; Antonini *et al.*, 1992). As wavelets fornecem a capacidade de multi-resolução, de compactar e adaptar as características do sistema visual humano.

As transformadas de wavelets representam uma função como uma superposição de uma família de funções de base chamadas wavelets. Um conjunto de funções de base pode ser gerada pela translação e dilatação da *wavelet mãe*, que corresponde a uma base particular. O sinal é passado através de um filtro de passa-baixa (LPF - lowpass filter²) e um de passa-alta (HPF - highpass filter³) e as saídas do filtro são divididas em duas. Então, a transformada de wavelet extrai informações do sinal em diferentes escalas. Para reconstrução, os coeficientes são *upsampled* e passados através de outro conjunto de filtros de passa-baixa e passa-alta. A DWT em 2-D é geralmente calculada usando uma abordagem separável. A Figura 7 mostra uma decomposição de uma wavelet em três níveis de uma imagem S de tamanho a x b pixels. No primeiro nível de decomposição, uma subimagem de passa-baixa (S_2) e três subimagens de passa-alta com orientações seletivas (W_2^H, W_2^V, W_2^D) são criadas. No segundo nível de decomposição, a subimagem de passa baixa é decomposta em uma subimagem de passa-baixa e três de passa-alta (W_4^H, W_4^V, W_4^D). O processo é repetido sobre a subimagem de passa-baixa para formar a decomposição de wavelet de nível mais alto. Em outras palavras, a DWT decompõe uma imagem em uma estrutura de pirâmide de subimagens com várias resoluções que correspondem a diferentes escalas. A transformada de wavelet inversa é calculada de modo inverso, isto é, começando com a subimagem de resolução mais baixa, as imagens de resolução mais altas são criadas recursivamente. Note que uma decomposição com três estágios cria uma subimagem de passa-baixa e nove subimagens de passa-alta (três na direção horizontal, três na vertical e três na diagonal). A subimagem de passa-baixa é uma versão da

² Um filtro de passa-baixa atenua as frequências altas e retém as frequências baixas inalteradas.

³ Um filtro de passa-alta atenua as frequências baixas e retém as frequências altas inalteradas.

imagem original com baixa resolução. As subimagens horizontais, verticais e diagonais fornecem a informação sobre as mudanças nas direções horizontais, verticais e diagonais, respectivamente.

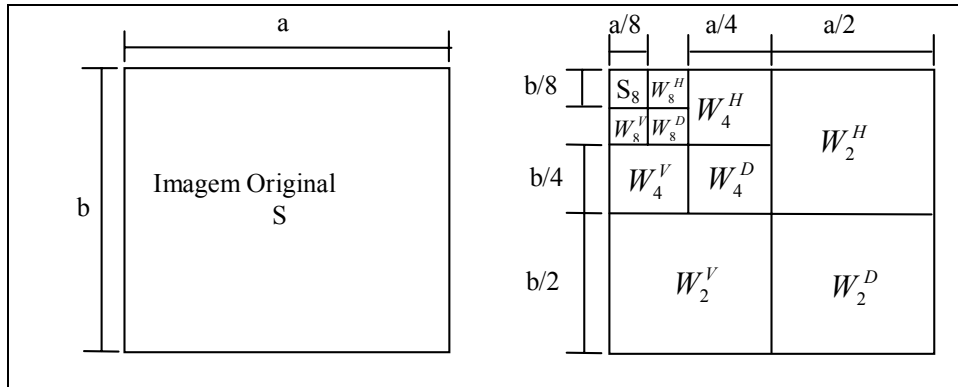


Figura 7: Imagem por transformada Wavelet
Fonte: (Mandal *et al.*, 1996)

A subimagem de passa-baixa é a mais importante entre todas as subimagens wavelet pois é uma versão *reduzida* da imagem original.

Lee e Dickinson (1994) propuseram uma técnica de indexação, no domínio de sub-bandas (Woods, 1991), onde os histogramas dos coeficientes das sub-bandas da subimagem de passa-baixa são comparadas hierarquicamente, do nível mais alto para o mais baixo da pirâmide. Esta técnica reduz a complexidade dos histogramas pois trabalha com subimagens pequenas. Idris e Panchanathan (1995) propuseram um esquema de indexação de imagens baseado em wavelet usando quantização de vetor. Mandal *et al.* (1996) e Lee (1996) também usaram as transformadas de wavelets no processo de indexação de imagens.

3.3.5 Transformada de Fourier

A transformada de Fourier é uma ferramenta importante no processamento de imagens. Ela é usada para decompor uma imagem em seus componentes de seno e co-seno. A saída da transformada representa a imagem no espaço de Fourier ou no domínio de frequência, enquanto a imagem de entrada está no domínio espacial equivalente. No domínio de Fourier cada ponto da imagem representa uma frequência particular contida no domínio espacial da imagem.

A transformada de Fourier é usada em muitas aplicações, como análise de imagens, filtros de imagens e sinais digitais, reconstrução e compressão de imagens, tomografia, análise espectral de voz, detecção de vibração, entre outras (Brigham, 1974; Gomes e Velho, 1994; Ifeachor e Jervis, 1995, Hadjiyannakis et al., 1997).

Como neste estudo objetiva-se trabalhar com imagens digitais, esta discussão restringe-se à transformada de Fourier Discreta (DFT - Discrete Fourier Transform).

O DFT são amostras da transformada de Fourier e, portanto, não contém todas as frequências que formam uma imagem, mas somente um conjunto de amostras que é grande o suficiente para descrever o domínio espacial da imagem completamente. O número de frequências corresponde ao número de pixels no domínio espacial, i.e. a imagem no domínio espacial e no domínio de Fourier são do mesmo tamanho.

Para uma imagem quadrada de tamanho $N \times N$, o DFT bi-dimensional é determinado por:

$$F(k, l) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) e^{-i2\pi \left(\frac{ki}{N} + \frac{lj}{N} \right)} \quad (3.2)$$

onde $f(i, j)$ é a imagem no domínio espacial e o termo exponencial é a função de base que corresponde a cada ponto $F(k, l)$ no espaço de Fourier. A equação (3.2) pode ser interpretada como: o valor de cada ponto $F(k, l)$ é obtido multiplicando a imagem espacial com a função de base correspondente e somando o resultado.

As funções de base são ondas de senos e co-senos com frequências crescentes, i.e., $F(0, 0)$ representa a componente discreta do cosseno da imagem que corresponde ao brilho médio e $F(N-1, N-1)$ representa a frequência mais alta.

Em um modo semelhante, a imagem de Fourier pode ser re-transformada ao domínio espacial. A transformada inversa de Fourier é determinada por:

$$F(i, j) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f(k, l) e^{i2\pi \left(\frac{ki}{N} + \frac{lj}{N} \right)} \quad (3.3)$$

Para obter-se os resultados das equações (3.2) e (3.3), uma soma dobrada tem que ser calculada para cada ponto da imagem.

A transformada de Fourier produz uma saída complexa que pode ser exibida com duas imagens, ou com uma parte real e outra imaginária ou com magnitude e fase. Neste processamento de imagens, freqüentemente só a magnitude da transformada de Fourier é exibida pois contém a maioria das informações de estrutura geométrica da imagem do domínio espacial. Porém, se o objetivo é re-transformar a imagem de Fourier no domínio espacial correto, após algum processamento no domínio de freqüência, deve-se preservar a magnitude e a fase da imagem de Fourier.

O domínio de Fourier da imagem tem um alcance maior que a imagem no domínio espacial. Consequentemente, para ser suficientemente preciso, seus valores são normalmente calculados e armazenados em valores de ponto flutuante.

A transformada de Fourier é usada quando se deseja ter acesso às características geométricas de uma imagem de um domínio espacial. Como a imagem no domínio de Fourier é decomposta em seus componentes senoidal, é fácil examinar ou processar certas freqüências da imagem que influenciam na estrutura geométrica do domínio espacial. Por exemplo, Milanese et al. (1999) usam a transformada de Fourier para a recuperação de imagens baseada em conteúdo e Joseph et al. (1999) a usam para embutir uma marca d'água em livrarias digitais com o objetivo de evitar a cópia ilícita e assim criar uma livraria digital segura.

3.3.6 Redes Neurais

As redes neurais, como a rede MLP, a rede de convolução e o SOM estão sendo usadas nos trabalhos com imagens, para classificação e redução de dimensionalidade (Lawrence et al., 1997; Tansley et al., 1998; Weng e Swets, 1999). Além disto, tem-se também as redes, como a rede bayesiana, rede de densidade, redes de análise de componentes principais, que implementam métodos estatísticos e que são usadas no processamento de imagens. Por exemplo S. Lin e Lin (1997) usaram uma rede de decisão probabilística para detecção e reconhecimento de faces.

3.4 Descrição de Alguns Sistemas de Recuperação de Imagens

Existem na literatura diversos sistemas de indexação e recuperação de imagens sendo testados e gradativamente sendo melhorados. Abaixo seguem-se alguns deles (QBIC, IRIS,

Virage, VisualSEEK, e o Synapse). Dentre os descritos, o sistema Virage é o único sendo comercializado.

3.4.1 QBIC

O sistema QBIC (Niblack et al., 1993) é um sistema de recuperação de imagens baseado nas características de cor, textura e forma. O QBIC foi desenvolvido no Centro de pesquisas Almaden IBM e seu protótipo foi escrito em linguagem de programação C, com a interface escrita em *X/Motif*. O sistema consiste de três partes principais:

1. Formar a base de dados, onde as imagens são importadas para dentro do sistema e, opcionalmente, são acrescentados objetos esboçados e textos e palavras chave.
2. Cálculo das características, onde as características das imagens e objetos são calculadas e armazenadas na base de dados.
3. Processar a "*query*", onde se cria as "*queries*", revisa-se os resultados e refina-se as "*queries*".

De acordo com estes passos, obtem-se a arquitetura do sistema QBIC, como mostra a Figura 8.

No processo de indexação do QBIC, a identificação dos objetos é o passo chave. QBIC tem devotado considerável esforço no desenvolvimento de ferramentas com objetivo de identificação de objetos. Usa métodos de segmentação não supervisionada, completamente automática para identificar objetos em uma classe restrita de imagens. Também usa ferramentas semi-automáticas para a identificação, como o método Flood-fill, que começa em um pixel do objeto simples e repetidamente adiciona pontos adjacentes cujos valores estão dentro de um dado limite do pixel original.

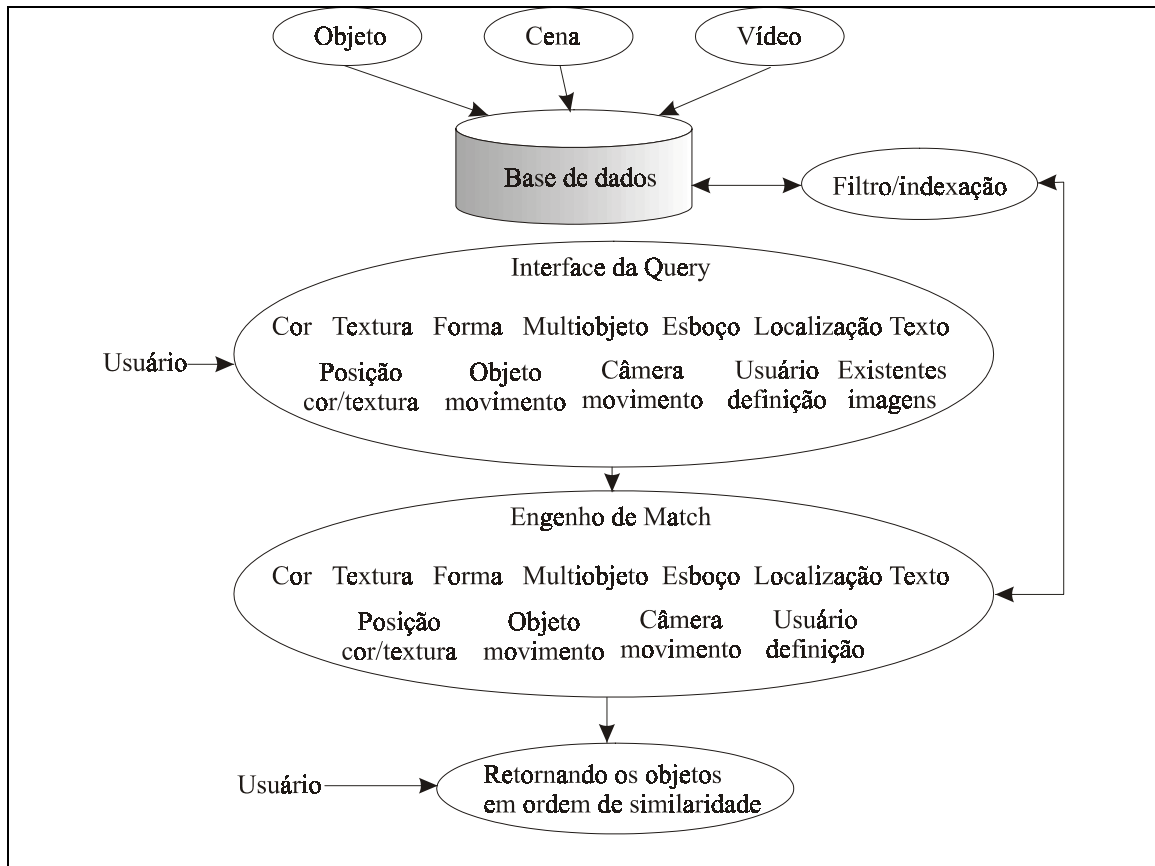


Figura 8: Arquitetura do sistema QBIC
 Fonte: He, (1996)

Após a identificação dos objetos, o QBIC calcula as características de cada objeto e imagem. No sistema QBIC, ou a cena é a imagem inteira ou um frame simples de vídeo e um objeto é uma parte da cena. Ele utiliza conjuntos de características que capturam a cor, textura e forma das imagens. Suporta duas classes do conteúdo de cor – local e global. No QBIC a extração de regiões locais (objetos) é feita manualmente, pois requer uma pessoa para desenhar os limites da região usando um *mouse*. As informações da cor local e global são representadas pela média das cores e o histograma de cor. O sistema também usa uma métrica da distância quadrada para comparação dos histogramas. Como a medida da distância quadrada é muito intensa computacionalmente, a distância de cor média é usada como um pré-filtro para as *queries* de cor.

3.4.2 IRIS - Image Retrieval for Information Systems

O sistema IRIS (Hermes et al, 1995) desenvolve e combina métodos e técnicas de *computer vision* e representação de conhecimento a fim de gerar, automaticamente, descrições dos conteúdos (cor, textura e contorno) das imagens e armazená-las em forma textual.

O sistema IRIS consiste de dois módulos principais: o módulo de análise da imagem e módulo de recuperação. A arquitetura é mostrada na Figura 9. O módulo de análise consiste de quatro sub-módulos: três módulos que extraem cada uma das características, cor, textura e contorno. O quarto módulo é responsável pelo reconhecimento da imagem. Os módulos de cor, textura e contorno extraem segmentos que são representados como textos estruturados. Estas características são extraídas independentemente umas das outras. Portanto, recebe-se informações de três fontes. Estas fontes apresentam anotações de baixo nível de uma imagem analisada.

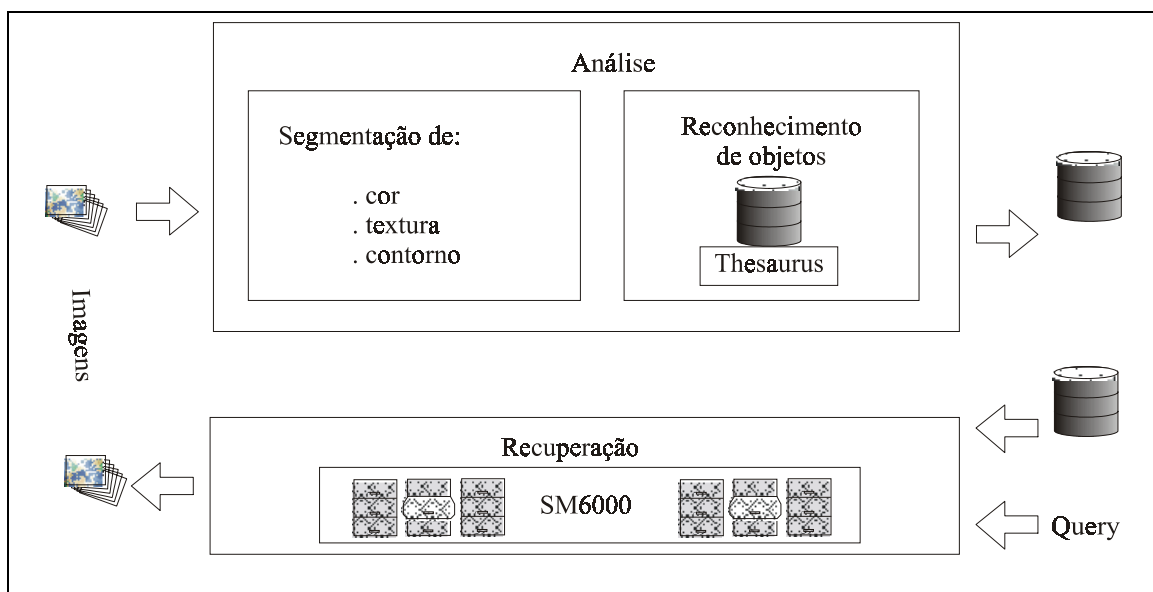


Figura 9: Arquitetura do sistema IRIS
Fonte: Hermes *et al.* (1995)

O módulo de reconhecimento do objeto é baseado nas anotações geradas dos três módulos de baixo nível e usa uma abordagem de reconhecimento de padrões sintático. Inicialmente, as relações de vizinhanças dos segmentos extraídos são calculadas. Os grafos são usados para representar estas relações de vizinhanças. Segundo, o processo de reconhecimento do objeto é tratado como o processo de transformações no grafo, isto é, o processo de reescrever o grafo. A gramática de grafos apresenta a classificação compilada,

que reflete o domínio de conhecimento. O reconhecimento do objeto fornece a informação para o quarto campo da anotação pertencente à imagem.

A geração automática da anotação textual foi completada. Portanto, a descrição pode ser indexada por técnicas de recuperação de textos bem conhecidas. A indexação e a recuperação textual no IRIS são executadas pelo produto de gerenciamento de busca da IBM. Uma *query* é exclusivamente baseada sobre anotações textuais e, portanto, o tempo de resposta é mínimo.

3.4.3 Virage

O Virage (Bach et al., 1996) é um sistema comercial para indexação e recuperação de imagens. Aqui a indexação é realizada após diversas operações de pré-processamento, tal como, suavizar e enfatizar contrastes. Cada rotina de extração primitiva toma uma imagem pré-processada e, dependendo das propriedades da imagem, calcula um conjunto específico de dados para aquela primitiva. Um vetor com os dados primitivos computados é armazenado em uma estrutura de dados própria.

No Virage a detecção das características é feita por vários métodos propostos na literatura corrente. Todas as características da imagem calculadas são chamadas primitivas. Os usuários são capazes de buscar imagens com as seguintes primitivas: Cor (a impressão de cor geral da imagem), composição (o arranjo espacial das regiões de cor na cena), textura (os "padrões" ou "texturas" de uma imagem) e estrutura (as características de forma gerais dos objetos na imagem).

Seguindo o modelo de dados para informação visual, o Virage admite quatro camadas de abstração de informação (Figura 10): a imagem linha (camada de representação da imagem), a imagem processada (a camada de objeto da imagem), as características de interesse do usuário (camada do domínio do objeto) e eventos de interesse do usuário para vídeos (a camada do evento do domínio). As três camadas do topo formam o conteúdo da imagem ou vídeo.

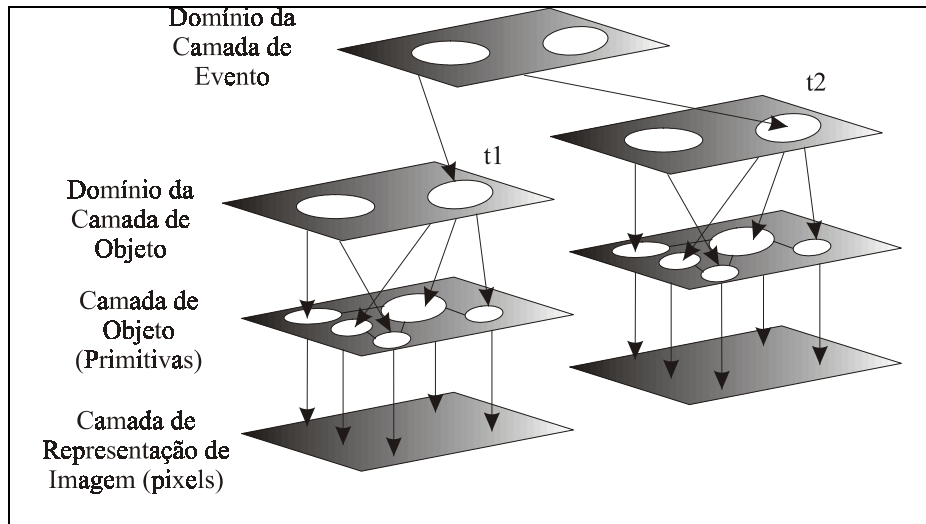


Figura 10: Arquitetura do Virage
Fonte: (Bach *et al.*, 1996)

A tecnologia do Virage é construída em torno de um módulo principal chamado o Engenho Virage e opera para um nível do objeto da imagem do modelo Virage. Existem três partes funcionais principais do engenho: análise da imagem, comparação das imagens e gerenciamento.

3.4.4 VisualSEEK

O VisualSEEK (Smith e Chang, 1996) é um sistema híbrido que integra indexação de imagem baseada em conteúdo (cor e textura) e em propriedades espaciais (como tamanho, localização e relações com outras regiões). A integração confia na representação de regiões através de um conjunto de cores. Primeiro, um conjunto de cores fornece para o sistema de extração de regiões automatizada através da retro-projeção do conjunto de cores. Segundo, os conjuntos de cores são facilmente indexados para a recuperação de imagens com um conjunto de cores semelhante. Desta forma, as imagens são comparadas através de suas regiões. Além disto, o sistema fornece ao usuário o controle da seleção das regiões e os parâmetros que são mais importantes na determinação da similaridade na query dada.

O projeto do VisualSEEK enfatiza diversos objetivos únicos a fim de melhorar a recuperação de imagens, como: (1) extração automática de características e regiões localizadas, (2) examina informações espaciais e as características, (3) extração de

características de dados compactados, (4) desenvolvimento de técnicas para indexação e recuperação rápida e (5) desenvolvimento de ferramentas altamente funcionais para o usuário.

A aplicação *server* do VisualSEEK foi desenvolvida no idioma de Java para permitir independência de plataforma de cliente e acessibilidade no *World Wide Web*. Como ilustra a Figura 11, o sistema de VisualSEEK consiste em vários componentes: o conjunto de ferramentas do usuário, o servidor de *query*, a imagem e servidor de vídeo, o arquivo de imagem e de vídeo, o banco de dados dos meta-dados e os arquivos de índice.

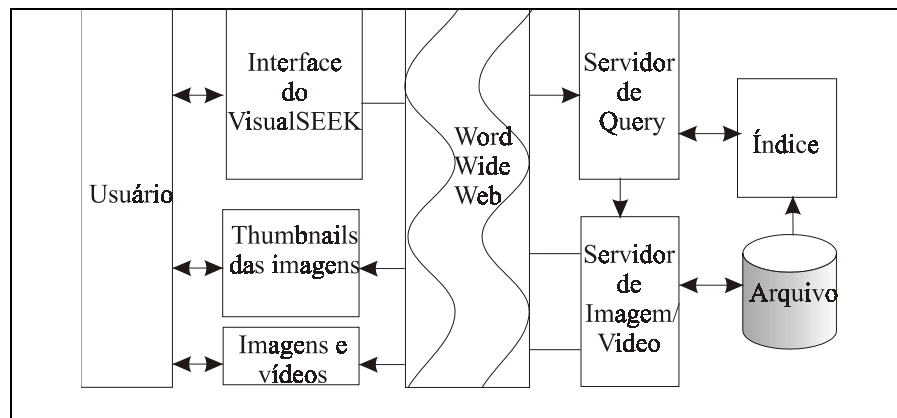


Figura 11: Arquitetura do VisualSEEK

Fonte: Smith e Chang (1996)

Atualmente, o sistema de VisualSEEK permite procurar imagens em uma base de 12000 imagens coloridas. As ferramentas de VisualSEEK também estão sendo portadas a uma aplicação para procurar em uma coleção de mais de um milhão de imagens e vídeos da *World Wide Web*.

3.4.5 Synapse

O **Synapse** (Sistema de Busca por Aparência Sintática) (Ravela *et al.*, 1996; Ravela e Manmatha, 1998) trata a recuperação de imagem pela aparência visual com uma *query*. Imagens são recuperadas usando uma caracterização da aparência visual dos objetos. A aparência visual do objeto está muito próximo da descrição da forma de sua superfície de intensidade, ou seja, os objetos que parecem ser visualmente similares podem ser recuperados pela caracterização da forma da superfície de intensidade.

A abordagem tomada por Ravela e Manmatha (1998) não depende da segmentação da imagem (manual ou automática) ou da extração de características binárias, nem do tempo para

treinamento. Usando uma imagem exemplo e a interação do usuário para construir *queries*, Synapse recupera imagens semelhantes dentro de pequena variação de tamanho e vista na ordem da similaridade com a aparência sintática para uma query. As queries (partes da imagem) são escolhidas pelo usuário para expressar sua “intenção semântica”.

Os vetores invariantes multi-escala formam uma representação local da imagem que naturalmente leva à recuperação por similaridade local por aparência. Até uma certa ordem, a aparência local da superfície de intensidade pode ser representada como resposta a um conjunto de filtros de derivadas gaussianas de escala parametrizados. Uma estratégia para recuperar imagem é, então, desenvolvida usando vetores construídos da combinação das saídas dos filtros de derivadas. Esta combinação leva a um conjunto de invariantes diferenciais, não alteráveis com as invariantes à transformações rígidas bidimensionais.

A recuperação é feita em dois passos computacionais: uma fase off-line e uma fase *run-time*. Na primeira fase, cada imagem na base de dados é filtrada para localizações amostradas e então as respostas dos filtros da base de dados são indexadas. Na Segunda fase, o sistema começa com o usuário selecionando uma imagem exemplo e marcando um conjunto de regiões salientes dentro da imagem. As respostas correspondentes destas regiões são comparadas com aquelas da base de dados e uma medida de *fitness* por imagem na base de dados é calculada em ambos espaço de características e espaço de coordenadas. Finalmente, as imagens são mostradas para o usuário em ordem de *fitness* com a *query*.

3.5 Conclusões

Analisando estes sistemas de recuperação de imagens pode-se notar a presença de duas preocupações fundamentais em sua construção: rapidez e eficiência. Com este objetivo os sistemas têm evoluído, desenvolvendo técnicas que melhor integram as características visuais das imagens. Para bases de dados pequenas (na ordem de centenas de imagens) a maioria das técnicas têm mostrado bons desempenhos. Conforme o tamanho da base de dados cresce, essas técnicas não são satisfatórias na prática. Além disto, estas mesmas técnicas têm sido testadas em meios controlados. Conseqüentemente, a robustez em relação aos ruídos e distorções nos ambientes da vida real ainda não foram demonstrados.

Problemas muito difíceis em sistemas de imagens permanecem ainda em investigação. Existe uma proliferação explosiva das imagens digitais sem restrições na forma de gráficos e vídeos devido ao melhoramento no acesso às tecnologias computacionais e na aceitabilidade do mundo Web como um meio viável para comunicação, publicação e propaganda. Isto tem criado uma necessidade de ferramentas eficientes e efetivas para catalogar, indexar, gerenciar, comprimir e buscar informações visuais sem restrições. Uma grande lacuna permanece entre a habilidade computacional em analisar imagens e vídeos em um nível de características (cor, textura, forma) e a inabilidade em um nível semântico (objetos, cenas, pessoas, valores artísticos). O aperfeiçoamento de tecnologias para preencher esta lacuna pode melhorar sistemas já existentes ou criar novas abordagens.

O maior problema de um sistema de indexação e recuperação de imagem é a determinação de um conjunto de características que possa fornecer uma boa performance na classificação das imagens para uma recuperação adequada. A este conjunto de características deve-se os diferentes tipos de sistemas encontrados na literatura. Assim, ainda existe a necessidade do desenvolvimento de ferramentas que possibilitem ou contribuam para o desenvolvimento de um sistema de indexação e recuperação de imagens.

4

INREC - INDEXAÇÃO RECURSIVA: MODELO PROPOSTO PARA REDUÇÃO DE DIMENSIONALIDADE

4.1 Introdução

O RP, apesar de ser uma área de pesquisa relativamente consolidada, ainda não consegue resolver todos os problemas de reconhecimento. Um dos grandes desafios é a análise de padrões que são caracterizados por domínios de espaço de alta dimensão. Por exemplo, o reconhecimento de caracteres manuscritos é uma tarefa difícil devido à variedade de estilos que são únicos para cada indivíduo. A tarefa se torna muito mais difícil de reconhecer quando se muda a fonte, a escala, a rotação ou quando os caracteres apresentam ruídos.

Outro problema bastante complexo em RP é o processamento de sinais digitais. As razões específicas para o processamento de um sinal digital podem ser, por exemplo, remover interferências ou ruídos do sinal, obter o espectro dos dados, transformar o sinal em uma forma mais desejada ou extrair informações destes sinais. A sua importância está na crescente aplicação em campos bastante variados como o processamento da voz, as telecomunicações, os sistemas de radar, sinais biomédicos, entre outros.

O processamento de imagens é uma área genérica do RP. O processo de reconhecimento de imagens é muito mais complexo, pois nas imagens, além dos problemas com ruído, distorções, e falta de nitidez, tem-se o problema da alta dimensão, que torna

inviável computacionalmente o uso de técnicas de classificação sem um processamento prévio.

Um dos pré-processamentos utilizados nestes e em outros casos é a redução de dimensionalidade como uma forma para auxiliar no reconhecimento. Porém, como foi visto no capítulo 2, os métodos existentes de redução de dimensionalidade ainda não são eficientes o bastante quando se deseja trabalhar com padrões de alta dimensão. Tendo em vista este problema, propõe-se aqui uma forma alternativa de redução de dimensionalidade, a INdexação RECURSiva (INREC). Esse modelo realiza uma indexação recursiva de uma função matemática para encapsular, em um único número (índice), a informação contida em um padrão. A meta é que, ao indexar as variações dos padrões da mesma classe, obtenham-se índices próximos.

4.2 Modelo Proposto – INdexação RECURSiva (INREC)

O foco deste estudo está em várias considerações relativas ao reconhecimento de padrões: qual o conjunto mínimo de características necessárias para representar classes de padrões? qual a qualidade na dimensão reduzida *versus* a velocidade de processamento? qual o significado da informação contida em apenas um índice?.

O objetivo, então, é construir uma forma simples de representar a informação, mas que contenha poder discriminatório. Construir um processo que possa reduzir a dimensão e que ainda mantenha as informações significativas para que os similares, em ambos os espaços, fiquem próximos.

4.2.1 INREC

Sendo a redução de dimensionalidade um problema ainda não totalmente resolvido em RP, propõe-se aqui uma forma alternativa de abordá-lo. Uma forma que possa na sua simplicidade conter informações discriminatórias. Esta redução não elimina ou considera uma característica mais relevante que a outra, pelo contrário, utiliza todas as características disponíveis na representação de um padrão complexo. Não é uma teoria do simplismo, mas um estudo algorítmico das funções e suas recursões na formação de índices (números reais)

com poder discriminatório. Cada índice é construído a partir das características que compõem cada padrão.

Seja um padrão X^i representado pelas suas p características, define-se um índice ρ , recursivo (INREC):

$$\rho = f(x_1^a f(x_2^a f(x_3^a \dots f(x_p^a) \dots))) \quad (4.1)$$

onde f é a função para a recursão, x_i são as características do padrão, e $a \in \{1, -1\}$ cuja função é promover uma padronização das características em questão.

Entre as funções matemáticas que podem ser utilizadas, podemos destacar as funções contrações⁴, pois as mesmas apresentam estabilidade na recursão sucessiva, ou seja, feita a recursão de um ponto do seu domínio, as mesmas convergem para um único ponto, chamado *ponto fixo* (Lima, 1977). Vale observar que o modelo proposto não realiza a recursão de um ponto apenas do domínio, é realizado sim uma recursão das características de um padrão, procurando usar a propriedade da função para encontrar relações entre as características do padrão.

A proposta deste estudo é avaliar a possibilidade de serem construídos algoritmos rápidos e eficientes baseados na equação (4.1), portanto, uma abordagem estritamente algorítmica.

A motivação para o desenvolvimento do método proposto, INREC, surgiu a partir da hipótese de se encontrar uma transformação que leve um vetor de um espaço \mathfrak{R}^n para um espaço de valores numéricos unidimensional \mathfrak{R} , que mantenha a capacidade de identificar os pontos próximos em ambos os espaços. Ou seja, a intenção é conseguir reduzir o espaço de características dos padrões de dimensão p de um determinado domínio de problema a um valor numérico (índice) $\rho \in \mathfrak{R}$, mantendo poder discriminatório. De acordo com Pao (1989, prefácio) “*Nenhum aspecto ou seqüência de aspectos, de um padrão determina o significado daquele padrão. Ao invés disso, existe uma grande compensação entre todos os aspectos do padrão de modo que o significado de um padrão pode ser entendido somente se todos os aspectos estão disponíveis simultaneamente para a consideração*”. Segue contudo que, na

⁴ Uma aplicação $f: M \rightarrow N$ chama-se uma contração do espaço métrico M no espaço métrico N , quando existe uma constante c , com $0 \leq c < 1$, tal que $d(f(x), f(y)) \leq cd(x, y)$ para quaisquer $x, y \in M$, onde $d(\cdot)$ é a distância definida.

geração do índice são consideradas todas as características e a posição das mesmas através de uma função f .

4.2.1.1 Função f

Algumas características a respeito da função f utilizada no processo recursivo, tem sua origem em alguns testes realizados com várias funções e sobre algumas considerações sobre a recursividade simples utilizada em problemas numéricos.

A recursividade simples, definida como $x_n = f^n(x_0)$ apresenta algumas características:

- 1) a recursão ou composição é aplicada a um número apenas (x_0);
- 2) dependendo da função utilizada para este processo, pode-se garantir a convergência ou não. A garantia decorre do teorema (Teorema de Banach, sobre pontos fixos e contrações): *Se M é um espaço métrico completo, toda contração $f: M \rightarrow M$ possui um único ponto fixo em M . Mais precisamente, se escolhermos um ponto $x_0 \in M$ e pusermos $x_1 = f(x_0)$, $x_2 = f(x_1)$, ..., $x_{n+1} = f(x_n)$, ... a seqüência (x_n) converge em M e $a = \lim x_n$ é o único ponto fixo⁵ de f (Lima, 1977);*
- 3) se a função não for uma contração podem acontecer fatos indesejáveis, como por exemplo, a sensibilidade em relação ao ponto inicial, ou seja, tomando pontos iniciais próximos à recursão pode levar a pontos muito distintos (Figura 12).

Baseado nestas considerações, foram realizados testes utilizando várias funções como a função cosseno, a função gaussiana, a função tangente hiperbólica e a função quadrática. Nestes testes verificou-se algumas restrições quanto à função a ser utilizada:

- a periodicidade é uma característica não desejável para a função, pois provoca uma desordem no espaço das medidas. Por exemplo, a função cosseno (Figura 13a);
- a monotocidade da função no domínio do problema, por outro lado, é uma característica desejável. Por exemplo, a função sigmoide e tangente hiperbólica (Figura 13 b e c).

⁵ Um ponto fixo de um aplicação $f: M \rightarrow M$ é um ponto $x \in M$ tal que $f(x) = x$.

- a função ser uma contração no domínio do problema, também é uma característica desejável. Por exemplo, a função tangente hiperbólica;
- a função deve ser contínua e diferenciável no domínio do problema.

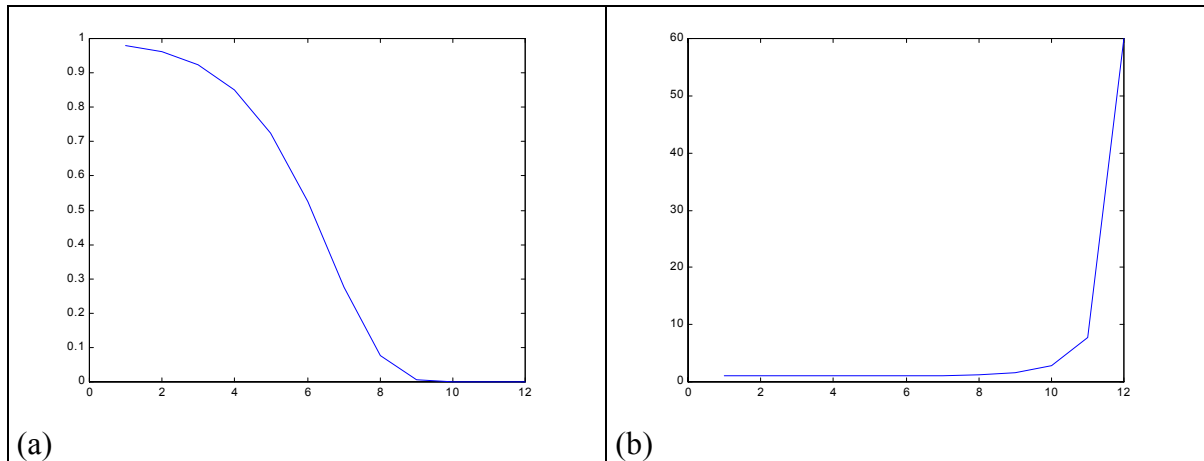


Figura 12: Gráfico das primeiras 12 composições da função $f = x^2$: (a) com $x_0 = 0.99$ a recursão converge para 0; (b) com $x_0 = 1.001$ a recursão diverge

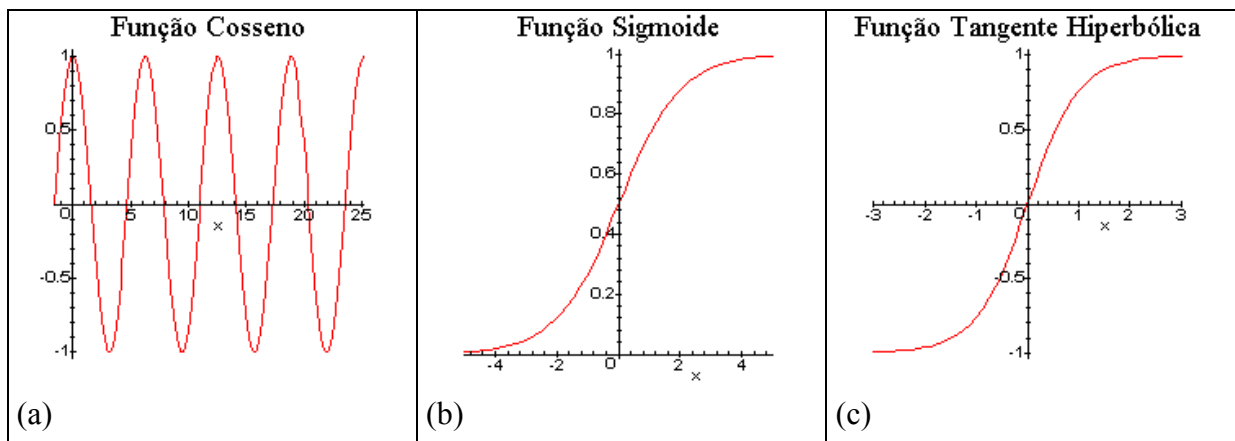


Figura 13 : Gráfico de algumas funções utilizadas para testes

Estas considerações em relação à função f são de caráter intuitivo baseado nos conhecimentos sobre função e seu comportamento gráfico. Deve-se observar que a recursividade realizada para o desenvolvimento do trabalho tem a finalidade de promover uma compensação entre as características do padrão de acordo com a função f utilizada. A composição não é feita de forma simples, existe uma expansão e uma composição em cada passo do processo recursivo.

4.2.2 Originalidade

O objetivo principal do algoritmo proposto é tentar representar as informações de um padrão de uma forma simples. O modelo utiliza um dos conceitos mais primitivos da matemática, número, que além da simplicidade, é de fácil armazenamento e recuperação. Tem-se, então, que captar informação na feitura deste número. A formação de um número através de um processo gradativo e/ou iterativo, não só atrai mas se posta como uma representação de grandezas que não são conhecidas ou não são entendidas.

Considerando-se que é possível representar infinitas informações através dos números de forma única, a questão é como utilizar apenas um número para cada padrão de informação de forma a deixar os que são semelhantes próximos.

Com este objetivo, o modelo proposto realiza uma recursão, ou composição de uma função matemática com o objetivo de encapsular, em um número apenas, a informação contida no padrão para indexação e recuperação. A meta é que, ao indexar as variações dos padrões da mesma classe, obtenham-se índices próximos.

Tem-se a finalidade de apresentar uma forma alternativa de abordar o problema da indexação. Como é possível explorar os números de uma forma mais significativa (potencial do número)? Como utilizar a sua infinidade em grandeza e representação? A meta é construir um número gradativamente, e nesta construção ir colocando características que possam ser identificadas neste número. Sabe-se que os padrões, de uma forma geral, possuem infinitas formas de serem representados, então o espaço ideal para representá-los seria, aparentemente, inviável computacionalmente. Mas, sabe-se também que, dependendo do conhecimento específico sobre o padrão, pode-se identificá-lo com apenas algumas características.

O pensamento aqui está direcionado a entender as características dos padrões como um conjunto independente de valores que contém a informação, não por si só, mas pelas relações umas com as outras. Esta idéia está em consonância com as afirmações de Bezdek e Pal (1992), de que RP é a busca por estrutura dos dados.

Entende-se que uma projeção destes valores para um espaço de dimensão menor seja uma forma forçada de reduzir dimensão. No entanto, acredita-se que escolher um espaço onde estas relações intrínsecas entre características possam ser representadas, seja o grande objetivo de uma técnica de redução e reconhecimento de padrões.

4.2.3 Conclusão

Neste capítulo foi introduzido uma nova forma de redução de dimensionalidade, chamada de INREC, que é baseada na indexação recursiva de uma função matemática para encapsular em um número apenas (índice) a informação contida em um padrão. A função é aplicada de forma recursiva em todas as características do padrão de modo a obter um índice. A meta é, ao indexar as variações de padrões semelhantes ou da mesma classe, obter índices próximos.

5

APLICAÇÕES DO MODELO PROPOSTO (INREC)

5.1 Aplicação do INREC para Indexação de Imagens

A INREC, visando à indexação e recuperação de imagens, gera um índice (tomando a imagem como um vetor de $n \times m$ características) que contenha informação suficiente para recuperar não só a imagem indexada mas pequenas variações da mesma.

Com o objetivo de que variações mudem o índice (índice real), mas não o mude completamente, realiza-se uma varredura na imagem de forma escalada, onde neste caso, o objetivo é provocar uma generalização do índice.

Para elaborar o índice é tomada uma escala de resolução (janela de $q \times q$ pixels, onde $1 \leq q \leq s$, e s representa a escala) na imagem.

Para cada janela ocorrida durante a varredura na imagem, toma-se a média dos tons de cinza, a posição horizontal, a posição vertical e o índice gerado até a janela anterior⁶. Com estes valores atualiza-se o índice parcial para a cor média ocorrida e, com este, o índice geral.

Tanto na formação dos índices parciais como na composição do índice final é utilizada a equação (4.1), onde a função f utilizada é a tangente hiperbólica. A função tangente hiperbólica além de satisfazer as características desejadas (seção 4.2.1.1), foi a função que obteve o melhor desempenho nos testes realizados.

⁶ Na primeira vez se usa zero.

As imagens utilizadas para a elaboração dos testes foram 94 fotos 3x4 cm de faces, obtidas das páginas pessoais dos alunos da Pós-graduação em Engenharia de Produção e Sistemas (PPGEP) da Universidade Federal de Santa Catarina. Estas imagens foram convertidas para 256 tons de cinza, e redimensionalizadas para 50x50 pixels.

Com o intuito de realizar alguns testes para verificar a eficiência deste índice, tomou-se algumas imagens e provocou-se ruídos direcionados (Figura 14). O resultado deste teste é apresentado na Figura 15.

Como mostra a Figura 15, o índice manteve as fotos com ruídos próximas das originais. Esse resultado mostra que o índice, consegue realizar uma boa indexação e, conseqüentemente, uma boa capacidade de recuperação das imagens similares.

Tendo estes resultados com a INREC, antes de realizar-se mais teste com imagens, buscou-se uma forma de validação do modelo proposto, utilizando a mesma idéia do índice para problemas específicos e clássicos de classificação. A próxima subseção apresenta os resultados obtidos.

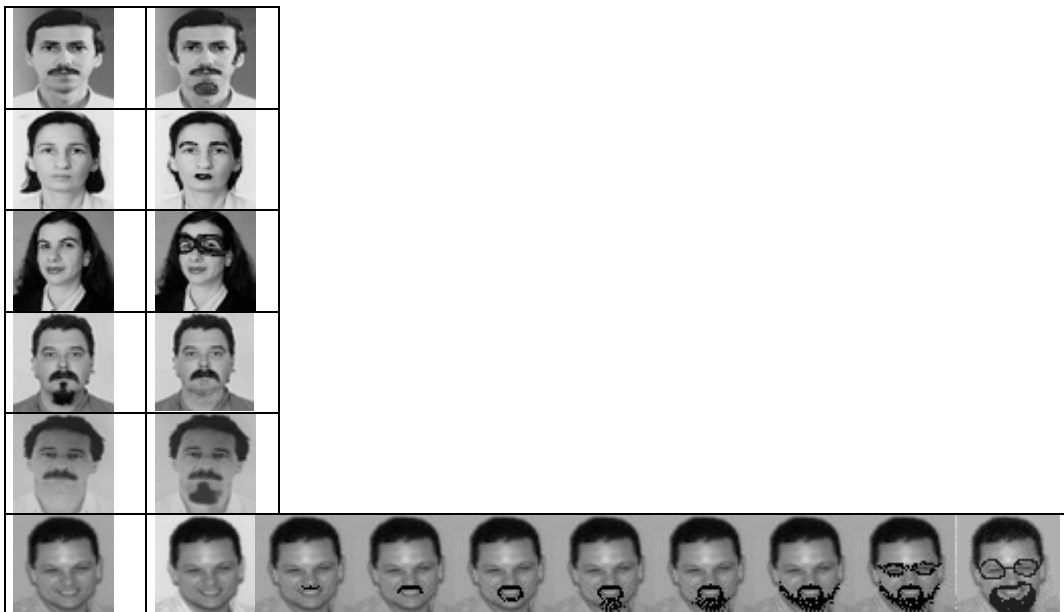


Figura 14: Lado esquerdo são as fotos originais; Lado direito as alteradas



Figura 15: Apresentação das fotos na ordem gerada pelo índice

Alguma restrições com respeito aos dados devem ser comentadas, as fotos utilizadas têm procedências diferentes, por exemplo, umas tem fundo escuro outras fundo claro. O foco, é outro fator preocupante, pois diferencia a dimensão da imagem. Vários fatores, no processamento de imagens, são importantes e deles depende o desempenho dos métodos de indexação e recuperação de imagens. O objetivo nesta aplicação é mostrar que o índice gerado através do modelo proposto captura informações úteis da imagem e pode ser usado num sistema de indexação e recuperação.

5.2 INREC para Classificação de Padrões com Poucas Características

Embora o índice não tenha a finalidade de ser um modelo para a classificação e sim para a redução de dimensão, realizou-se um experimento com o objetivo de verificar o poder discriminatório dos índices gerados a partir das características dos padrões de um problema de

RP. A preocupação é no sentido de verificar o significado da informação no contexto do índice, ou seja, dado o índice de um padrão de teste analisa-se os índices vizinhos no conjunto de índices gerados a partir do conjunto de treinamento, para a tomada de decisão.

A seguir descreve-se o algoritmo utilizado para a classificação dos dados da Íris e dos cromossomos, e faz-se uma análise comparativa dos resultados obtidos com outros métodos de classificação com objetivo de averiguar a quantidade de informação que o índice mantém.

Algoritmo:

1. Realiza-se uma normalização das características.
2. Gera-se o índice usando a função tangente hiperbólica:

$$\rho = \tanh(x_1^{-1} \tanh(x_2^{-1} \dots \tanh(x_p^{-1}) \dots))$$

3. Para testar se um padrão pertence a uma determinada classe, usa-se a frequência de ocorrência para os vizinhos do valor do índice gerado para o padrão em questão.

Observações:

- 1) A característica, quando definida muito precisamente, provoca uma especificação indesejada. Assim, um tratamento no sentido de deixar as características mais livres, ou menos específicas, podem provocar uma generalização melhor. Um cuidado é portanto necessário: quão genéricas devem ser as características para que não percam o poder de discriminação?
- 2) Na fase de testes, a consideração de vários vizinhos ajuda a melhorar a performance quando as classes não são linearmente separáveis.

5.2.1 Aplicação da INREC aos dados da Iris de Fisher

O conjunto de dados da Íris foi popularizado por R. A. Fisher (1936), ao ilustrar os princípios da análise discriminante. Ele contém 6 variáveis com 150 observações. As variáveis representam:

- Número que representa o tipo da espécie da flor;
- Nome da espécie (Íris Setosa, Íris Virgínia e Íris Versicolor);
- Largura da pétala;
- Comprimento da pétala;
- Largura da sépala;
- Comprimento da sépala.

O conjunto de dados contém 50 instâncias de cada classe, onde cada classe refere-se a um tipo de Íris (Setosa, Virgínia e Versicolor). Uma classe é linearmente separável das outras duas; e as outras duas não o são uma da outra. Cada instância no conjunto de dados é descrito por quatro características. O conjunto de dados foi dividido em dois subconjuntos com igual número de instâncias, separando as instâncias de posição ímpar das de posição par (Fu, 1994). Então, realiza-se a validação com este dois conjuntos.

Tabela 1 - Avaliação da performance do INREC para os dados da Íris

MODELO	Taxa de Acerto (%)
INREC (com 5 vizinhos)	94.66
FAN (Free Associative Neurons)	100.00
Extração de regras de redes neurais (KT)	96.70
Árvore de Decisão	93.40

Os resultados obtidos pela aplicação deste método aos dados da Íris são comparáveis aos resultados existentes. O Tabela 1 mostra a taxa de acerto de vários métodos. A INREC apresenta uma taxa de acerto de 94,66%, o FAN 100% de acerto (Raittz, 1997), o KT (extração de regras de redes neurais) 96.7% (Fu, 1994), e a árvore de decisão 93.4% (Fu, 1994).

5.2.2 Aplicação da INREC nos dados dos cromossomas

As características mais utilizadas na classificação automática de cromossomos são o comprimento e posição do centrômero (ou índice centromérico), sendo que com estas duas características é possível separar os cromossomos em 7 grupos, também chamado grupo de Denver (Graham, 1987). A terceira característica utilizada nas comparações é o padrão de

banda do cromossomo ao longo do eixo longitudinal. Estas três características têm sido a base da grande maioria dos estudos desenvolvidos nas duas últimas décadas (Desinov, 1994; Errington, 1993; Graham, 1987; Groen et al, 1989; Keret, 1991; Lundsteen, 1986; van Vliet et al., 1990).

O problema da aplicação nesta fase, consiste em, fornecidas as três características de um cromossoma, identificar a qual dos grupos de Denver, ele pertence.

Nesta aplicação, o INREC obteve um resultado inferior à Radial Basis Functions (Todesco, 1995) e ao Free Associative Neurons (Raittz et al, 1997), e igualou ao resultado obtido por Henrington com o perceptron multicamadas com o algoritmo de aprendizagem Backpropagation (Tabela 2).

Tabela 2 - Avaliação da performance da INREC para os dados dos cromossomos

<i>MODELO</i>	Taxa de Acerto (%)
INREC (com 5 vizinhos)	94.61
FAN (Free Associative Neurons)	95.23
Multi-Layer Perceptron Backpropagation	94.60
Radial Basis Functions	95.96

Os resultados acima comprovam que a INREC apresenta boa performance quando aplicado para classificação de padrões (no caso, Íris e cromossomos). Isso mostra que a simplicidade na elaboração de um método de classificação de padrões ainda pode trazer vantagens não só em termos de eficiência e rapidez, mas também na forma de pesquisa de novos métodos.

Pode-se observar nas tabelas 1 e 2 que a INREC não fornece os melhores resultados para estes dois problemas de classificação, porém eles são comparáveis aos outros citados. Estes resultados mostram que a INREC reduz a dimensão dos dados mantendo poder discriminatório. Desta forma a INREC pode ser considerada uma forma de redução de dimensionalidade.

5.3 Aplicação da INREC para Estagiamento Automático do Sono

5.3.1 Introdução

Os conhecimentos sobre o ciclo vigília-sono desenvolveram-se a partir da década de 30, quando Loomis e colaboradores (Apud Guitot, 1996), utilizando a eletroencefalografia, descreveram o sono como um processo unitário e ativo, composto por estágios, cada um com características próprias.

Com a abertura de um novo campo da fisiologia, e a possibilidade de registro de um número considerável de variáveis durante o sono, o volume de pesquisas sobre o assunto cresceu significativamente. Em 1968 foram definidos os critérios de monitorização e estagiamento do sono por Rechtschaffen e Kales. Apesar de revisado por Carskadon e Rechtschaffen em 1994, o padrão de 68 ainda é adotado.

Seguindo os critérios de estagiamento, grande parte dos médicos e técnicos em polissonografia têm realizado o estagiamento do registro polissonográfico manualmente desde a década de 70. Essa tarefa, quando executada manualmente, é considerada tediosa, pois o registro de uma noite inteira de sono consome, aproximadamente, 3 horas de anotações e cuidadosa observação.

5.3.2 O Sono e Sua Estrutura

5.3.2.1 O Sono

Graças à polissonografia, reconhece-se hoje que o sono não é um estado homogêneo, e que há dois estados distintos de sono. O sono mais surpreendente, e o último a ser descoberto, é o sono em que ocorrem movimentos rápidos dos olhos. Por suas iniciais em inglês, "*rapid eye movements*", este sono é chamado de **REM**. Apesar de ocupar apenas 20% do sono de um adulto, o sono REM é tão importante que o restante é chamado de sono **NREM** (Não REM). O sono NREM pode ser dividido em outras 4 fases, chamadas de estágios, de 1 a 4.

5.3.2.2 A Monitorização

A polissonografia consiste em diversos tipos de registros, como o eletroencefalograma – EEG (ondas elétricas do cérebro), o eletro-oculograma – EOG (movimentos dos olhos), eletromiograma – EMG (tensão muscular), eletrocardiograma – ECG (ondas elétricas do coração), movimentos respiratórios e a oxigenação do sangue (Carskadon e Rechtschaffen, 1994). A polissonografia é realizada em laboratórios do sono por meio de equipamentos especiais, e usados por técnicos capacitados.

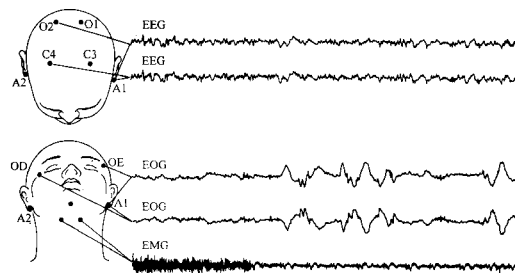


Figura 16 - Diagrama dos principais registros da polissonografia (Guitot, 1996)

Para o estagiamento manual, segundo Carskadon e Rechtschaffen (1994), os canais C3 e C4 do EEG são suficientes para distinguir os estágios do sono. Entretanto, muitos laboratórios também têm utilizado o registro dos canais O1 e O2, entre outros, como canais auxiliares.

5.3.2.3 Considerações Sobre o Estagiamento

O estagiamento⁷ do registro do sono é realizado, normalmente por conveniência, em segmentos chamados de época. O tamanho médio desses segmentos é de 30 ou 20 segundos de amostra, que correspondem, aproximadamente, a uma página de papel de 30 cm registrado a uma velocidade de 10 a 15 milímetros por segundo. Cada época é associada com o estágio que mais apropriadamente caracteriza o padrão predominante durante o intervalo examinado. Assim, o objetivo do estagiamento da época é determinar o simples fator que mais caracteriza a época.

⁷ O padrão de estagiamento manual do sono segue critérios definidos por Rechtschaffen e Kales em 1968, que ainda são utilizados internacionalmente.




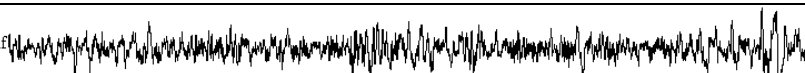
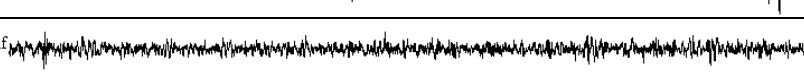
5.3.2.4 O Padrão de Estagiamento

Os padrões do sono descritos por Rechtschaffen e Kales (Carskadon e Rechtschaffen, 1994) são os seguintes :

Quadro 1 : Descrição sucinta das classes do sono

CLASSE	DESCRIÇÃO DO SINAL
Acordado	O EEG exibe ritmo alfa quando o sujeito se encontra relaxado com os olhos fechados. Esse ritmo se altera quando o sujeito está tenso, ou mesmo com os olhos abertos. Também se encontra o ritmo alfa presente, mesmo com os olhos abertos, se o indivíduo está excessivamente sonolento.
Estágio 1	Atividade de relativa baixa voltagem e freqüências mistas. Mista porque é comum a presença de artefatos, além de pequenas faixas de atividade teta.
Estágio 2	Atividade de baixa voltagem e freqüências mistas. Basicamente, a diferenciação do estágio 2 para o estágio 1, se dá em dois padrões específicos conhecidos como fuso e complexo K. Os fusos são compostos por ondas de 12 a 14 Hz com duração de 0,5 a 1,5 seg., enquanto que os complexos K são descritos como “ondas agudas negativas bem delineadas imediatamente seguidas por um componente positivo” (Carskadon e Rechtschaffen, 1994, p. 950).
Estágio 3	O estágio 3 apresenta um mínimo de 20% e um máximo de 50% de ondas de 2 Hz com amplitudes maiores que 75 μ V de pico a pico por época (apud Carskadon e Rechtschaffen, 1994).
Estágio 4	As ondas de atividade lenta devem ser predominantes, estando presentes em pelo menos 50% na época analisada.
REM	O padrão do EEG para o sono REM também é caracterizado por relativa baixa voltagem e freqüências mistas. Embora esse estágio se chame REM o simples movimentos dos olhos não é suficiente para indicar esse estado do sono. Isso acontece uma vez que o fenômeno não é universal, pois muitos indivíduos não apresentam o movimento rápido dos olhos.

Quadro 2 : Aparência dos sinais das classes do sono

CLASSE	APARÊNCIA DO SINAL
Acordado	C3-Rf 
Estágio 1	C3-Rf 
Estágio 2	C3-Rf 
Estágio 3	C3-Rf 
Estágio 4	C3-Rf 
REM	C3-Rf 

5.3.2.5 *Progresso do Sono Durante a Noite*

O primeiro ciclo, em um adulto inicia com o estágio 1 de sono NREM, e em poucos minutos (1 a 7 minutos) avança para o estágio 2. O estágio 2, sinalizado pelo fusos de sono, ou ainda, pelos complexos K detectados no EEG, costuma durar de 10 até 25 minutos. Como progresso natural do estágio 2, em 20 minutos em média, atinge o estágio 3. O estágio 3, de curta duração neste primeiro ciclo, serve de transição para o estágio 4, que costuma durar de 20 a 40 minutos, e em seguida, ocorre um retorno a estágios menos profundos.

5.3.3 *O Estagiamento Automático*

O estagiamento automático proposto por Tafner (1998) é realizado por um sistema de Inteligência Artificial, especificamente, redes neurais artificiais. Essa tarefa é realizada partindo de um sinal digital pré-processado através de diversas técnicas de processamento de sinal digital combinadas.

O sinal biológico utilizado para a realização do estagiamento automático é proveniente do eletroencefalograma, especificamente o canal C3, recomendado inclusive por Carskadon e Rechtschaffen (1994), onde ressaltam que todas as ondas usadas para distinguir os estágios são bem visualizadas utilizando os canais C3 e C4.

Para a aquisição do sinal eletroencefalográfico, utilizou-se o sistema de poligrafia digital POLIWIN da empresa EMSA (1998), que funciona através de um microcomputador ligado a um equipamento de amplificação analógica. O sinal de sono adquirido é realizado a uma taxa de amostragem que pode variar de 150 a 300 Hz. O valor *default* para tomada do sinal é de 200 Hz. A taxa de quantização do sinal é fixa em 12 bits de precisão.

5.3.3.1 *Pré-processamento dos Dados*

Os dados adquiridos pelo sistema POLIWIN são submetidos a uma série de processos com o objetivo de aumentar a representatividade deles em relação ao sistema. Esse aumento da representatividade é necessário pois pode separar o ruído da informação propriamente dita.

O pré-processamento dos dados consiste, então, em 3 fases diferentes, cada uma das fases dá à informação biológica adquirida uma nova interpretação. A atuação de uma fase sobre os resultados obtidos na fase anterior constitui o processamento para este trabalho. As

fases são assim denominadas : Ajuste do Sinal Digital, Geração dos Índices Recursivos e Extração das Características.

5.3.3.1.1 Ajuste do Sinal Digital

O sinal capturado pelo equipamento de polissonografia contém ruídos que podem confundir as fases seguintes no pré-processamento. Assim, é necessário que um processamento anterior procure amenizar a presença de ruídos (sinais de altíssima amplitude em módulo) e ajustar o sinal para as fases que se seguem. A amenização de ruídos é realizada através de um corte sobre os sinais que extrapolam os limites definidos (limite superior e inferior). Após os sinais ajustados aos novos limites, converte-se todos os sinais negativos no seu simétrico.

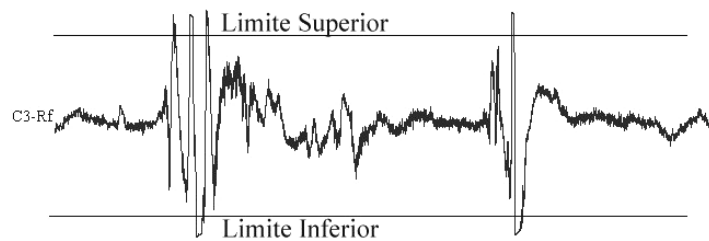


Figura 17 - Trecho do sinal biológico original com os limites superior e inferior

5.3.3.1.2 Geração dos Índices Recursivos

Nesta fase, mostra-se uma utilização concreta do modelo proposto para a redução da dimensão dos dados de um canal do EEG.

A Indexação Recursiva (INREC) é utilizada para reduzir o espaço de características para um espaço de dimensão menor, com o objetivo de reduzir ruídos, aumentar a velocidade de processamento e aumentar a representatividade da informação.

Isso foi realizado aplicando-se a INREC sobre segmentos dos sinais digitais originais para todo o sinal, após a primeira fase. Cada segmento foi substituído pelo seu respectivo índice, construindo assim um novo vetor representando o sinal original, chamado de Vetor de Índices, ou $V_{\text{índice}}$. Para a construção desse vetor de índices utilizou-se também a função *tangente hiperbólica*.

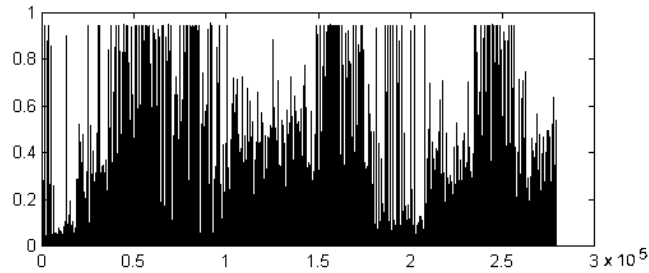


Figura 18 - Vetor de índices

5.3.3.1.3 Extração das Características

A extração de características (técnica desenvolvida especificamente para o problema em questão) é determinada pela contagem da presença de sinais do vetor de índices em determinadas faixas de amplitudes. As faixas de amplitudes são definidas tomando-se a imagem de uma partição do vetor de índices. Esta partição P é tomada em relação ao tamanho do $V_{\text{índice}}$ (domínio) como intervalos de igual tamanho, ou seja, $P = \{[t_i, t_{i+1}] / i = 1, \Lambda, p\}$, onde $\bigcup_{i=1}^p [t_i, t_{i+1}] = V_{\text{índice}}$ e P é a quantidade desejada de elementos da partição.

Uma vez particionado, obtém-se as posições de cada corte (t_i). Em seguida, ordena-se o $V_{\text{índice}}$ e gera-se as faixas de amplitudes tomando-se a imagem de cada corte.

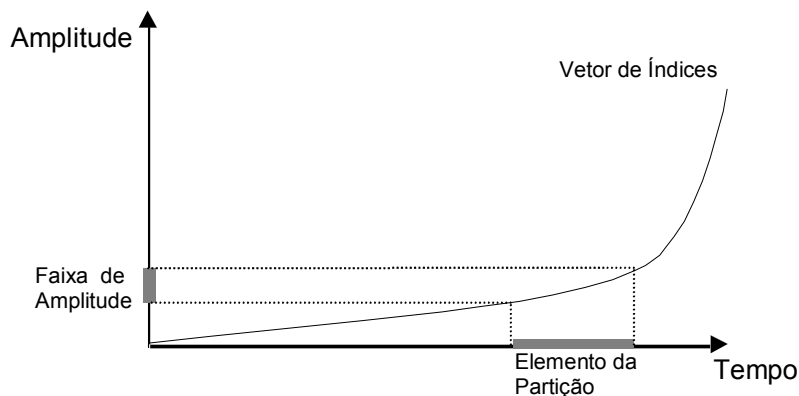


Figura 19 - Faixa de amplitudes e elemento de partição para o vetor de índices

Tendo as faixas de amplitudes delimitadas, a extração de características é realizada tomando-se um determinado tempo do $V_{\text{índice}}$ original (época). As características são geradas

então através da contagem da presença dos sinais nas faixas de amplitudes respectivas. O número de elementos da partição define a quantidade de características.

As faixas de amplitudes operam como segmentadores da informação contida no vetor de índices. Esse processo é realizado devido a uma mudança de contexto da informação presente no vetor. A mudança é, basicamente, a transformação do vetor de índices em segmentos de informação. Assim, o pré-processamento deixa de operar diretamente sobre os valores, para operar sobre as informações extraídas do vetor de índices.

5.3.4 Análise dos Resultados

5.3.4.1 Considerações Iniciais

Para os testes devidos, foi utilizado o canal C3 do EEG dos exames de 4 pacientes adultos do sexo masculino. Os exames fornecidos foram assim denominados :

Tabela 3 - Pacientes usados para o trabalho

Paciente	Idade	Tempo de Registro (horas)
6501	48	03:53:05
7706	25	05:47:58
6740	25	07:35:43
9300	31	07:36:40

Para a realização dos mesmos, modificou-se a quantidade de classes (estágios, rem e acordado) do estagiamento automático. A modificação abrange, basicamente, os estágios 3, 4 e 1. Os estágios 3 e 4 foram agrupados em um único estágio, chamado de estágio 3/4, e o estágio 1 foi somado junto ao estágio 2, denominado de estágio 1/2.

Para a construção dos vetores de treinamento e de testes, na primeira fase do pré-processamento, utilizou-se os limites inferiores e superiores em -1000 e $+1000$. Para gerar o vetor de índices, utilizou-se a recursão a cada 10 pontos em todo o sinal original e, finalmente, o vetor de características (treinamento ou teste) foi gerado em intervalos regulares de 30 segundos, ou 600 pontos do vetor de índices. A quantidade de características extraídas (faixas de amplitudes) para cada época utilizadas por Tafner (1998) foi de 20 faixas.

5.3.4.2 Resultados

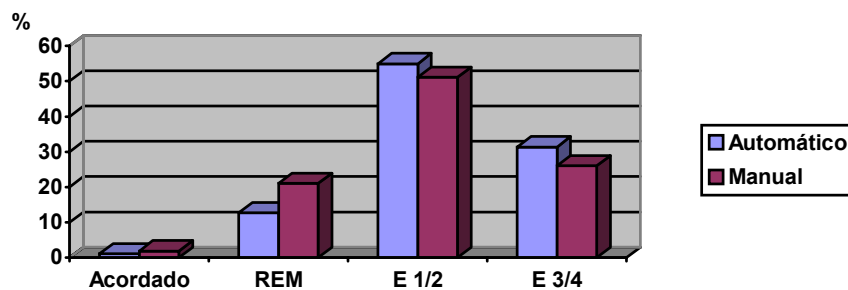
O exame do paciente 7706, submetido à análise da rede neural ARC (Tafner, 1998) treinada com o exame do paciente 6501, obteve uma taxa de acerto de 76,54 % confrontado com o resultado do estagiamento realizado de forma manual. A simulação realizada por Tafner segue a seguinte distribuição:

Tabela 4 - Distribuição do Acerto para o Paciente 7706

Classe	Automático	Manual	Acerto (%)
Acordado	0	13	0
REM	71	146	48,63
Estágio ½	290	355	81,69
Estágio ¾	171	181	94,48
Total	532	695	76,54

Uma representação gráfica, que pode vir a ser bastante útil, é o histograma por classe, que apresenta a quantidade de tempo de cada classe por método (automático e manual).

Gráfico 01 - Distribuição de Ocorrências das Classes



Para os outros dois exames, 6740 e 9300, pode-se observar os seguintes resultados, também obtidos usando as mesmas configurações aplicadas ao exame 7706 :

Tabela 5 - Distribuição do Acerto para os Pacientes 6740 e 9300

Classe	6740			9300		
	Automático	Manual	Acerto (%)	Automático	Manual	Acerto (%)
Acordado	1	95	1,05	4	79	5,06
REM	32	66	48,48	122	232	52,59
Estágio ½	515	659	78,15	321	409	78,48
Estágio ¾	89	91	97,80	111	193	57,51
Total	637	911	69,92	558	913	61,11

5.4 Análise dos Resultados das Aplicações

O índice foi utilizado nas aplicações já descritas ora como parte de um pré-processamento, ora como o próprio reconhecedor de padrões. Essas aplicações têm a finalidade de comprovar, não somente a eficácia do método, mas também a sua versatilidade.

Através dessas aplicações, os resultados demonstraram que :

- 1) A INREC pode auxiliar o processo de indexação e recuperação de imagens;
- 2) A INREC apresenta resultados comparáveis a métodos tradicionais (como redes neurais e estatísticos) quando utilizada para classificação de padrões (dados dos cromossomos e da íris);
- 3) O método proposto pode ser utilizado como parte de um processo de redução de dimensionalidade, como foi o caso da sua participação na formação do conjunto de treinamento e testes do reconhecedor automático do sono;
- 4) O método é matematicamente simples e computacionalmente muito rápido.

Para o caso do reconhecedor automático do sono, pode-se falar da rapidez do estagiamento automático em relação ao manual. O estagiamento automático consome até 3 minutos de tempo máquina, enquanto que o método manual pode consumir até 3 horas. Essa rapidez é devida em parte à redução de dimensão proporcionada pela INREC.

Tafner (1998), em seu trabalho sobre reconhecimento automático das fases do sono, onde utiliza a INREC como uma ferramenta de pré-processamento, aponta como ponto positivo a combinação dessa técnica antecedendo uma outra chamada de “Extração de Características”. Essa combinação, segundo o autor, apresentou uma eficácia satisfatória para o problema do estagiamento automático comparado com a utilização da Transformada Rápida de Fourier (TRF). O sistema utilizando a INREC obteve um desempenho em média 10% superior ao sistema que utilizou a TRF.

6

CONCLUSÕES E RECOMENDAÇÕES

6.1 Conclusões

O modelo proposto de indexação e recuperação de imagens trata-se de um protótipo que busca contribuir para o desenvolvimento na pesquisa relativa à qualidade da informação.

A INREC pode levar em consideração todas as características disponíveis de cada padrão, sem a preocupação da presença ou não de ruído, ou desconsideração de alguma das características, como faz, por exemplo, a técnica das componentes principais.

A natureza intrínseca dos números e das funções matemáticas pode ser utilizada em algoritmos para prover sistemas de indexação e recuperação de imagens de recursos para a associação de padrões que são ainda pouco explorados.

Com apenas um número real construído de forma iterativa, ou seja, com uma formação composta (lapidado), pode-se realizar não só indexação e recuperação de imagens, mas classificar padrões provindos de problemas com um número grande ou não de características.

A INREC apresentou, como performance, resultados comparáveis a métodos tradicionais quando utilizada para classificação de padrões com poucas características. Da mesma forma, também se mostrou eficaz quando utilizada como parte de um processo mais amplo de pré-processamento, como foi verificado no estagiamento automático do sono através do canal de EEG. Esse fato, por si só, revela também versatilidade de aplicação, pois pode ser utilizada em diferentes problemas de RP.

Os resultados obtidos em Tafner (1998) são animadores e indicam uma real possibilidade desta técnica ser usada como parte de processos em processamento de sinais digitais.

Finalmente, cabe ressaltar não somente a simplicidade matemática do método, mas também a sua capacidade de redução de dimensão preservando o significado da informação. Essa redução proporciona, conseqüentemente, maior rapidez no processo computacional.

6.2 Recomendações

A simplificação, através de índices unidimensionais, abre um leque de possibilidades de pesquisas em trabalhos futuros, como a relação do esquema de indexação proposto com técnicas de aprendizado em outras formas de reconhecimento de padrões tais como as redes neurais.

Testes mais elaborados com bases de imagens com aplicações em problemas práticos também podem ser realizados para aumentar a consistência e promover o aperfeiçoamento do método.

Também recomendamos que sejam exploradas outras combinações envolvendo técnicas de processamento de sinal digital e a indexação recursiva, uma vez que os resultados obtidos têm sido animadores.

Para o caso específico do estagiamento do sono, recomendamos a utilização dessa mesma técnica em outros canais, como o eletro-oculograma e o eletromiograma. Se for o caso, a INREC também poderá compor o pré-processamento desses canais junto com outras técnicas como Fourier e Wavelets.

Recomendamos que a INREC seja explorada tanto na utilização de outras funções para geração de índices, como na análise matemática das relações entre o índice gerado e as propriedades das funções utilizadas.

A extração de características, desenvolvida especificamente para o problema de reconhecimento automático do sono, merece maior atenção, uma vez que sua aplicabilidade pode ser estendida a outros problemas de RP como por exemplo, reconhecimento de fala e reconhecimento de imagens.

7

BIBLIOGRAFIA

1. ANDREWS, H. C. Introduction to Mathematical Techniques in Pattern Recognition. Wiley-Interscience, John Wiley & Sons, 1972.
2. ANTONINI, M., BARLAUD, M., MATHIEU, P., DAUBECHIES, I. Image coding using wavelet transform, IEEE Transactions on Image Processing, v. 1, n. 2, Abril, 1992.
3. BACH, J. R., PAUL, S., JAIN, R., A Visual Information Management System for the Interactive Retrieval of Faces, IEEE Trans. Knowledge and Data Engineering, v. 5, n. 4, p. 619-628, Agosto 1993.
4. BACH, J. R., FULLER, C., GUPTA, A., HAMPAPUR, A., HOROWITZ, B., HUMPHREY, R., JAIN, R. C., SHU, C. Virage image search engine: an open framework for image management. Symposium on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases IV, v. 2670, p. 76 - 87, 1996.
5. BELLMAN, R. Adaptive Control Processes: A Guided Tour. Princeton University Press, 1961.
6. BERRY, M. J. A., LINOFF, G. Data Mining Techniques. USA: John Wiley & Sons, Inc, 1997.
7. BESL, P.J., JAIN, R.C. Three-dimensional object recognition, ACM Computing Surveys, v. 17, n. 1, p. 75-145, 1985.
8. BEZDEK, J. C., PAL, S. K. (Editores) Fuzzy Models Pattern Recognition: Methods That Search for Structures in Data, IEEE, 1992.

9. BINAGHI, E., GAGLIARDI, I., SCHETTINI, T. Image retrieval using fuzzy evaluation of color similarity. International Journal of Pattern Recognition and Artificial Intelligence, v. 8, p. 945-968, 1994.
10. BISHOP, Christopher M. Neural Networks for Pattern Recognition. Oxford University Press, 1995.
11. BORDOGNA, G. Pictorial indexing for an integrated pictorial and textual information retrieval environment. Information Services & Use, v. 16, p. 165-173, 1990.
12. BRIGHAM, E. O. The Fast Fourier Transform. New Jersey: Prentice-Hall, Inc, 1974.
13. BROLIO, J., DRAPER, B.A., BEVERIDGE, J.R., HANSON, A.R. SR: a database for symbolic processing in computer vision. Computer, v. 22, n. 12, p. 22-30, 1989.
14. CALIFANO, A., MOHAN, R. Multidimensional Indexing for Recognizing Visual Shapes. IEEE Trans. Pattern Analysis and Machine Intelligence, v. 16, n. 6, p. 373-392, Abril 1994.
15. CARDENAS, A. F., LEONG, I. T., TAIRA, R. K., BREAND, C. N. The Knowledge-Based Object-Oriented PICQUERY Language. IEEE Trans. Knowledge and Data Engineering, v. 5, n. 4, p. 644-657, 1993.
16. CARREIRA-PERPIÑÁN, M. Á. A Review of Dimension Reduction Techniques, Technical Report CS-96-09, 1997. <http://www.dcs.shef.ac.uk/~miguel/papers/cs-96-09.ps.gz> (27 de Maio de 1998, 01:30).
17. CARSKADON, Mary A., RECHTSCHAFFEN, Allan. Monitoring and Staging Human Sleep. In : KRYGER, Meir, ROTH, Thomas e DEMENT, William. Principles and Practice of Sleep Medicine. 2. Ed. USA : W. B. Saunders Company, 1994.
18. CAWKELL, A.E. Imaging systems and picture collection management: a review, Information. Services & Use, v. 12, p. 301-325, 1992.
19. CHAKRAVARTHY, A.S. Toward semantic retrieval of pictures and video. Proc. Riao'94, Intelligent Multimedia Information Retrieval Systems and Management, New York, p. 676-686, 1994.
20. CHANG, S. K., HSU, A. Image Information Systems: Where do we go from here? IEEE Trans. Knowledge and Data Engineering, v. 4 , n. 5, p. 431-442, 1993

21. CHANG, S. K., LIU, S. H. Picture Indexing and Abstraction Techniques for Pictorial Database. IEEE Trans. Pattern Analysis and Machine Intelligence, v. 6, n. 4, p. 475-483, Julho 1984.
22. CHEN, C.Y., CHANG, C.C. An object-oriented similarity retrieval algorithm for iconic image databases. Pattern Recognition Letters, v. 14, p. 465-470, 1993.
23. CHIN, R.T., SMITH, H., FRALICK, S.C. Model-based recognition in robot vision. ACM Computing Surveys, v. 18, n. 1, p. 67-108, 1986.
24. CHOCK, M., CARDENAS, A. F., KLINGER, A. Database Structure and Manipulation Capabilities of a Picture Database Management System (PICDMS). IEEE Trans. Pattern Analysis and Machine Intelligence, v. 6, n. 4, p. 484-492, Julho 1984.
25. CICHOCKI A., UNBENHABEN, R. Neural Networks for Optimization and Signal Processing. John Wiley & Sons inc, 1993.
26. COLOMBO, C., BIMBO, A. D. Color-Induced Image Representation and Retrieval. Pattern Recognition, v. 32, p. 1685-1695, 1999.
27. DENISOV, D.A., DUDKIN, A.K. Model-based chromosome recognition via hypotheses construction/verification. Pattern Recognition Letters, v. 15, n. 3, p. 299-307, 1994.
28. DUDA, O., HART, P. E. Pattern classification and scene analysis. John Wiley & Sons, Inc., 1973.
29. EMSA – Equipamentos Médicos S/A. EMSA. 1998. <http://www.emsamed.com.br> (15 de Outubro de 1998, 23:20).
30. ERRINGTON, P., GRAHAM, J. Classification of Chromosomes using a Combination of Neural Networks. IEEE International Conference on Neural Networks, p. 1236-1241, 1993.
31. FISHER, R., PERKINS, S., WALKER, A., WOLFART, E. Hypermedia Image Processing Reference. England: John Wiley & Sons Ltd, 1996.
32. FISHER, R.A. The Use of Multiple Measurements in Axonomic Problems. 1936. <http://www.comcat.com/~hutch/DASL/dataex.htm> (15 de Outubro de 1997, 23:00).
33. FU, K. S. Syntactic Pattern Recognition and Application. New Jersey: Prentice-Hall, Inc., 1982.

34. FU, L. M. Rule Generation from Neural Networks. IEEE Transactions on Systems: Man and Cybernetics, v. 28, n. 8. p. 1114 - 1124, 1994.
35. _____. Neural Networks in Computer Intelligence. 2^o Ed. New York: McGraw-Hill, Inc., 1994b.
36. GHAHRAMANI, Z., HINTON, G. E. The EM Algorithm for Mixtures of Factor Analyzers. University of Toronto Technical Report CRG-TR-96-1, 1996.
37. GOMES, J., VELHO, L. Computação Gráfica: Imagem. Rio de Janeiro: IMPA/SBM, 424 pp., 1994.
38. GORSKY, W. I., MEHROTRA, R. Index-Based Object Recognition in Pictorial Data Management. Computer Vision, Graphics, and Image Processing, v. 52, p. 416-436, 1990.
39. GRAHAM, J. Automation of Routine Clinical Chromosome Analysis I - Karyotyping by Machine. Analytical and Quantitative Cytology and Histology, v. 9, n. 5, p. 383-390, 1987.
40. GROEN, F.C.A., TEN KATE, T. K., SMEULDERS, A.W.M. , YOUNG, I.T. Human chromosome classification based on local band descriptors. Pattern Recognition Letters, v. 9, n. 3, p. 211-222, 1989.
41. GROTH, R. Data Mining: A Hands on Approach for Business Professionals. USA: Prentice Hall PTR, 1998.
42. GUIOT, Marilene. Polissonografia. In : REIMÃO, Rubens. Sono : Estudo Abrangente. 2. Ed. São Paulo: Atheneu, 1996.
43. HADJIYANNAKIS, K., OGILVIE, R. D., ALLOWAY, C. E. D., SHAPIRO, C. FFT Analysis of EEG During Stage 2-to-REM Transitions in Narcoleptic Patients and Normal Sleepers. Electroencephalography and Clinical Neurophysiology, v. 103, p. 543-553, 1997.
44. HANCOCK, P. J. B., BRADLEY, R. J., SMITH, L. S. The Principal Components of Natural Images. Network: Computation in Neural Systems, v. 3, n. 1, p. 61-70, 1992.

45. HAYKIN, Simon. Neural Networks: A comprehensive Foundation. New York: Macmillan College Publish Company, 1994.
46. HE, Q. Comparison of Content-based Image Retrieval Systems, <http://alexia.lis.uiuc.edu/~hqin/lis429/project.html> (8 de Março de 1998, 23:00), 1996.
47. HERMES, T., KLAUCK, C., KREY, J., ZHANG, J. Image retrieval for information systems. Proc. of SPIE - The Inter. Soc. for Optical Engineering, Storage and Retrieval for Image and Video Databases, 1995.
48. IDRIS, F., PANCHANATHAN, S. Image indexing using wavelet vector quantization, Proc. of SPIE: Digital Image Storage and Archiving Systems, v. 2606, Philadelphia, Oct. 1995.
49. IFEACHOR, E. C., JERVIS, B. W. Digital Signal Processing – A Practical Approach, USA: Addison-Wesley, 1995.
50. JAIN, A. K., VAILAYA, A. Image Retrieval Using Color and Shape. Proc. ACCV, v. 2, p. 529-533, 1995.
51. JAIN, R.C. (Editor) NSF Workshop on Visual Information Management Systems, Workshop Report, Computer Science and Engineering Division, The University of Michigan, Ann Arbor, 1992.
52. JOHNSON, R. A., WICHERN, D. W. Applied Multivariate Statistical Analysis. New Jersey: Prentice Hall inc., 4 ed., 1998.
53. JOSEPH, J. K., RUANAIDH, O., CSURKA, G. Bayesian Approach to Spread Spectrum Watermark Detection and Secure Copyright Protection For Digital Image Libraries. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'99), Fort Collins, Colorado, USA, 1999.
54. JOSEPH, T., CARDENAS, A. F. PICQUERY: A High Level Query Language for Pictorial Database Management. IEEE Trans. Software Engineering, v. 14, n. 5, p. 630-638, Maio 1988.
55. KASABOV, K. N. Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering. USA, Massachusetts Institute of Technology Press, 1996.

56. KERET, D., EZRA, D. An Efficient Algorithm For Automatic Recognition And Classification of Chromosomes By Cluster Seeking. Artificial Intelligence and Computer Vision. Y.A. Dfeldman and A. Bruckstein (Editores). Elsevier Science Publishers B.v., North-Holland, p. 417-430, 1991.
57. KLIR, G. J., YUAN, B. Fuzzy sets and fuzzy logic - theory and applications. Prentice Hall, 1995.
58. KOENDERINK, J. J., VAN DOORN, A. J. Representation of Local Geometry in the Visual System. Biological Cybernetics, v. 55, p. 367-375, 1987.
59. LAMDAN, Y., WOLFSON, H. J. Geometric Hashing: A General and Efficient Model-Based Recognition Scheme. Proc. Second Int'l Conf. Computer Vision, p. 238-249, 1988.
60. LAWRENCE, S., GILES, C. L., TSOI, A. C., BACK, A. Face Recognition: A Covolutional Neural Network Approach. IEEE Trans. On Neural Networks, Special Issue on Neural Network and Pattern Recognition, v. 8, n. 1, p. 98-113, 1997.
61. LEE, J., DICKINSON, B. W. Multiresolution video indexing for subband coded video databases. Proc. Of SPIE: Storage and Retrieval for Image and Video Databases, v. 2185, San Jose, CA, Fevereiro 1994.
62. LEE, T. S. Image Representation Using 2D Gabor Wavelets. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 18, n. 10, Outubro 1996.
63. LEUNG, C.H.C.. Architecture on an image database system. Information Services & Use, v. 10, p. 391-397, 1990.
64. LIMA, Elon Lages. Espaços Métricos. Rio de Janeiro : Impa – CNPQ, 1977.
65. LUNDSTEEN, C., GERDES, T , MAAR, J. Automatic Classification of Chromosomes as Part of a Routine System for Classification Analysis. Cytometry, v. 7, p. 1-7, 1986.
66. MANDAL, M. K., ABOULNASR, T., PANCHANATHAN, S. Image Indexing Using Moments and Wavelets. IEEE Transactions on Consumer Electronics, v 42, n. 3, p. 557-565, Agosto 1996.

67. MILANESE, R., CHERBULIEZ, M., PUM, T. Invariant Content-Based Image Retrieval Using the Fourier-Mellin Transform. International Conference on Advances in Pattern Recognition (ICAPR'98), Plymouth, UK, p. 73-82, 1999.
68. NAGY, G. Image database. Image and Vision Computing, v. 3, n. 3, p. 111-117, 1985.
69. NIBLACK, W., BARBER, R., EQUITZ, W., FLICKNER, M., GLASMAN, E., PETKOVIC, D., YANKER, P., FALOUTSOS, C., TAUBIN, G. The QBIC project: Querying image by content using color, texture, and shape. Proc. SPIE Storage and Retrieval for Image and Video Database, San Jose, Calif., p. 173-187, 1993, .
70. OMIDVAR, O., DAYHOFF, J. (Editores) Neural Networks and Pattern Recognition, USA: Academic Press, 1998.
71. ORENSTEIN, J. A., MANOLA, F. A. PROBE Spatial Data Modeling and Query Processing in an Image Database Application. IEEE Trans. Software Engineering, v. 14, n. 5, p. 611-629, May 1988.
72. PALA, P., SANTINI, S. Image Retrieval by Shape and Texture. Pattern Recognition, v. 32, p. 517-527, 1999.
73. PANDYA, A., MACY, R.B. Pattern Recognition with Neural Networks in C++. CRC Press, 1995.
74. PAO, Y. Adaptive Pattern recognition and neural networks. Addison-Wesley, 1989.
75. PENTLAND, A, PICARD, R.W., SCLAROFF, S. Photobook: Tools for content-based manipulation of image database. Proc. of SPIE on Storage and Retrieval for Image and Video Databases II, San Jose, Calif., p. 34-47, 1994.
76. PETRAKIS, E. G. M., FALOUTSOS C. Similarity searching in large image databases. Technical Report 3388, Department of Computer Science, University of Maryland, 1995.
77. PICARD, R.W., MINKA, T.P. Vision texture for annotation. Multimedia Systems, v. 3, n. 1, p. 3-14, Springer-Verlag, 1995.
78. PIPER, J., GRANUN, E. On Fully Automatic Feature Measurement for Banded Chromosom Classification. Cytometry, v. 10, p. 242 - 255, 1989.

79. POSTON, W. L., MARCHETTE, D. J. Recursive dimensionality reduction using Fisher's linear discriminant. *Pattern Recognition*, v. 31, Issue: 7, p. 881-888, Julho 1998.
80. RAITTZ, R, SOUZA, J.A., DANDOLINI, G. A., PACHECO, R., MARTINS, A., GAUTHIER, F., BARCIA, R. Learning by Means of Free Associative Neurons. NAFIPS 97: Annual Meeting Of The North American Fuzzy Information Processing Society, Syracuse, New York, p. 21-24, Setembro 1997.
81. RAVELA, S, MANMATHA, R. Image retrieval by appearance. Proc. of SIGIR, Philadelphia, Julho 1997.
82. RAVELA, S., MANMATHA, R., Retrieving by Appearance, 1998 <http://hobart.cs.umass.edu/~mmedia/postscript/iccv98.ps.gz> (10 de Abril de 1998, 23:00).
83. RAVELA, S., MANMATHA, R., RISEMAN, E. M. Image Retrieval Using Scale-Space Matching. Computer Vision - ECCV'96, Cambridge, U.K, p. 273-282, Abril 1996.
84. RECHTSCHAFFEN, A., KALES, A. Manual of Standardized Terminology: Techniques and Scoring System for Sleep Stages of Human Subjects. Los Angeles: Brain Research Institute, 1968.
85. RIGOUTSOS, I., HUMMEL, R. Massively Parallel Model Matching. Computer, v. 25, n. 2, p. 33-42, Fevereiro 1992.
86. ROUSSOPOLOUS, N., FALOUTSOS, C., SELLIS, T. An Efficient Pictorial Database System for PSQL. IEEE Trans. Software Engineering, v. 14, n. 5, p. 639-650, Maio 1988.
87. SCHALKOFF, Robert J. Pattern Recognition: Statistical, Structural and Neural Approaches. John Wiley & Sons, Inc., 1992.
88. SCHMID, R. e MOHR, R. Local Greyvalue Invariants for Image Retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 19, no. 5, p.530-535, 1997.

89. SCHOWENGERDT, Robert A. Remote Sensing – Models and Methods for Image Processing. Segunda Edição, Academic Press, pp. 522, 1997.
90. SMITH, J. R., CHANG, S. F. Tools and techniques for color image retrieval. Symposium on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases IV, San Jose, CA, v. 2670, p. 426-437, Fevereiro 1996.
91. SMITH, John. Integred Spatial and Feature Image Systems: Retrieval, Analysis and Compression, Doctor thesis, 1997. <ftp://ftp.ctr.columbia.edu/CTR-Research/advent/public/public/jrsmith/thesis/>, (27 de Maio de 1998, 02:45].
92. SMOLIAR, S. W., ZHANG, H. Content-Based Video Indexing Retrieval. IEEE Multimedia, p. 62-72, 1994.
93. STRICKER, M., ORENGO, M. Similarity of Color Images. Proc. Of SPIE: Storage and Retrieval for Image and Video Databases III, San Jose, v. 2420, p. 381-392, Fevereiro 1995.
94. SWAIN, M.J., BALLARD, D.H. Color Indexing. Int'l J. Computer Vision, v. 7, n. 1, p. 11-32, 1991.
95. SWAIN, M.J., BALLARD, D.H. Indexing via color histograms. DARPA Image Understanding Workshop. Pittsburgh, Pennsylvania, p. 623-630, Setembro 1990.
96. TAFNER, Malcon Anderson. Estagiamento Automático do Sono usando uma Rede Neural Artificial com Alta Representação Cortical. Florianópolis, 1998. Exame de Qualificação - Departamento de Engenharia de Produção, Universidade Federal de Santa Catarina.
97. TANSLEY, J. E., OLDFIELD, M. J., MACKAY, D. J. C. Neural Network Image Deconvolution, 1998. ftp://wol.ra.phy.cam.ac.uk/pub/www/mackay/nn_in_decon.ps.gz (16 de Dezembro de 1998, 02:45).
98. TODESCO, Jose L. Reconhecimento de Padrões Usando Rede Neuronal Artificial com uma Função de Base Radial: uma Aplicação na Classificação de Cromossomas Humanos. Florianópolis, 1995. Tese - Departamento de Engenharia de Produção, Universidade Federal de Santa Catarina, .

99. VALKEALAHTI, K., OJA, E. Reduced Multidimensional Co-Occurrence Histograms in Texture Classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 20, n. 1, p. 90-94, 1998.
100. VAN VLIET, L., YOUNG, I. T. , MAYALL, B. H. The Athena Semi-Automated Karyotyping System. Cytometry, v. 11, p. 51-58, 1990.
101. VIJAYAKUMAR, S., Schaal, S. Local Dimensionality for Locally Weighted Learning, 1997. <http://www.hip.atr.co.jp/~sschaal/pub/vijayakumar-CIRA97.html>, (27 de Maio de 1998, 01:10).
102. WENG, J., SWETS, D. L. Face Recognition. In A. K. Jain, R. Bolle (eds), Biometrics: Personal Identification in Networked Society, Kluwer Academic Press, Hingham, Massachusetts, 1999.
103. WOODS, J. Subband Image Coding. Kluwer Academic Publishers, 1991.
104. WU, J. K., NARASIMHALU, A. D. Identifying Faces Using Multiple Retrievals. IEEE Multimedia, p. 27-38, Summer, 1994.
105. WU, J. K., NARASIMHALU, A. D., MEHTRE, B. M., LAM, P. C., GAO, Y. J. CORE: A Content-based Retrieval Engine for Multimedia Information Systems. Multimedia Systems, v. 3, n. 25, p. 25- 41, 1995.
106. YOSHITAKA, A., KISHIDA, S., HIRAKAWA, M., ICHIKAWA, T. Knowledge-Assisted Content Based Retrieval for Multimedia Databases. IEEE Multimedia, p. 12-21, 1994.