



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

Programa de Pós-Graduação em Computação Aplicada

# **MINERAÇÃO DE DADOS EM BANCOS DE IMAGENS**

Marcelino Pereira dos Santos Silva

Monografia do Exame de Qualificação do Doutorado em Computação Aplicada,  
orientada pelo Dr. Gilberto Câmara Neto.

INPE  
São José dos Campos  
2003



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA

**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

Programa de Pós-Graduação em Computação Aplicada

## **Mineração de Dados em Bancos de Imagens**

**Marcelino Pereira dos Santos Silva**

Monografia do Exame de Qualificação apresentada ao Programa de Pós-Graduação em Computação Aplicada do Instituto Nacional de Pesquisas Espaciais de São José dos Campos, como requisito parcial para obtenção do título de doutor em Computação Aplicada.

Banca Examinadora:

---

Dr. Elbert Einstein Nehrer Macau (Presidente)

---

Dr. Gilberto Câmara Neto (Orientador)

---

Dr. Antônio Miguel Vieira Monteiro

---

Dra. Leila Maria Garcia Fonseca

---

Dra. Sandra Aparecida Sandri

*São José dos Campos, 04 de setembro de 2003.*

## **RESUMO**

Um crescimento explosivo nos acervos de imagens de instituições governamentais, corporativas e científicas tem ocorrido durante os últimos anos. Não obstante, é notória a demanda de diferentes segmentos da sociedade por informações estratégicas oriundas destes acervos. A mineração de dados em bancos de imagens propõe técnicas para extrair conhecimento não explícito (relacionamentos espaciais, padrões relevantes, dentre outros) a partir de grandes acervos de imagens. Este documento aborda temas fundamentais desta área de pesquisa, dentre eles a descoberta de conhecimento em bancos de dados, o sensoriamento remoto, a recuperação de imagens, as características e técnicas de mineração, além do software de mineração ADaM.

## SUMÁRIO

Pág.

### LISTA DE FIGURAS

<b>CAPÍTULO 1 - Introdução.....</b>	<b>6</b>
<b>CAPÍTULO 2 - Contexto de Estudo .....</b>	<b>8</b>
2.1 Descoberta de Conhecimento em Bancos de Dados .....	8
2.2 Sensoriamento Remoto.....	12
2.3 Técnicas Usuais de Extração de Informações em Imagens de Sensoriamento Remoto .....	15
<b>CAPÍTULO 3 - Tecnologias e Desafios na Recuperação de Imagens.....</b>	<b>19</b>
3.1 Busca por Similaridades .....	19
3.2 Linguagem de Consulta para Recuperação por Conteúdo.....	22
3.3 Consultando Imensos Bancos de Imagens .....	26
<b>CAPÍTULO 4 - Mineração de Dados em Imagens .....</b>	<b>30</b>
4.1 Características, Desafios e Potencialidades.....	30
4.2 Arquitetura Dirigida à Informação .....	33
4.3 Técnicas de Mineração de Imagens.....	35
4.4 ADaM – Algorithm Development and Mining .....	37
4.4.1 Arquitetura.....	38
4.4.2 Serviços .....	39
4.4.3 Aplicações .....	41
<b>CAPÍTULO 5 - Conclusões.....</b>	<b>43</b>
5.1 Implicações do Atual Estágio de Mineração de Dados em Imagens.....	43
5.2 Limitações do Processo .....	44
5.3 Perspectivas Tecnológicas .....	44

## LISTA DE FIGURAS

Figura 2.1: Etapas De Kdd [Fay1996].....	11
Figura 2.2: Sistema De Imageamento [Ccr2003].....	12
Figura 2.3: Espectro Eletromagnético [Ccr2003].....	13
Figura 2.4: Segmentação De Imagem .....	16
Figura 2.5: Classificação De Imagem .....	17
Figura 3.1: Imagem Original, Seguida De Segmentação E Rotulação De Objetos.....	22
Figura 3.2: Arg Da Imagem [Pet1997].....	22
Figura 3.3: Exemplo De Resultado De Consulta [Tow2001].....	25
Figura 3.4: Aspecto Da Interface Do Cliente Rasdaman [Bhs2002].....	28
Figura 4.1: O Processo De Mineração De Imagens [Zhl2002] .....	31
Figura 4.2: Arquitetura De Mineração De Imagens Direcionada À Informação [Zhl2001].....	34
Figura 4.3: Duas Imagens Com Mudanças Incomuns Circuladas [Cli2001] .....	37
Figura 4.4: Ambiente Do Adam [Uah2003a].....	39
Figura 4.5: Detecção De Ciclones Tropicais Utilizando Adam E Amsu [Uah2003c] ...	42
Figura 4.6: Detecção De Pontos De Luz Utilizando Adam E Ols [Uah2003c].....	42

## **CAPÍTULO 1**

### **Introdução**

Instituições governamentais, corporativas e científicas têm realizado grandes investimentos na geração, processamento e publicação de imagens de sensoriamento remoto. Tal cenário tem gerado um crescimento explosivo nos acervos e bancos de imagens destas instituições, superando em muito a usual capacidade de interpretar e analisar estes dados.

Não obstante, verifica-se a demanda econômica, política, social e governamental por resultados em tempo hábil, oriundos de informações estratégicas destas imagens. Aplicações relevantes encontram-se em diferentes áreas: uso e cobertura do solo, monitoramento ambiental, planejamento urbano, vigilância territorial, previsão de safra agrícola, dentre outros.

A mineração de dados em bancos de imagens vem ao encontro destas questões através de técnicas, algoritmos e sistemas capazes de extrair conhecimento (informações estratégicas), relacionamentos espaciais e padrões interessantes não explícitos em grandes acervos de imagens. Os investimentos, a disponibilidade e a demanda sobre estes acervos de imagens, aliados à impossibilidade humana de análise manual e individual destes dados, motivam a pesquisa deste tema em direção ao fomento de novas ferramentas e técnicas para análise automática (ou semi-automática) e inteligente de bancos e acervos de imagens de sensoriamento remoto.

Nesta ótica, o presente trabalho busca apresentar e discutir pontos relevantes das tecnologias que suportam a mineração de dados espacial, tencionando contribuir no avanço de fronteiras tecnológicas ligadas a esta área. A organização deste documento é apresentada a seguir.

O capítulo 2 contextualiza a área de estudo, abordando a descoberta de conhecimento em bancos de dados, aspectos relevantes de sensoriamento remoto e técnicas mais utilizadas na extração de informações a partir de imagens de sensoriamento remoto.

No capítulo 3 são tratadas tecnologias e desafios na recuperação de imagens, através da busca por similaridades e de linguagem de consulta para recuperação por conteúdo. São abordadas ainda questões inerentes a consultas em imensos bancos de imagens.

Subseqüentemente, o capítulo 4 focaliza a mineração de dados em imagens, levantando suas características, potencialidades e principais técnicas. Uma arquitetura voltada à informação e um software de mineração espacial são apresentados neste contexto.

O quinto e último capítulo conclui o documento expondo as implicações do atual estágio de mineração de dados em imagens, apontando as limitações do processo e as perspectivas tecnológicas da área.

## CAPÍTULO 2

### Contexto de Estudo

No campo de abrangência deste documento, três elementos apresentam-se de forma relevante, a partir dos quais serão abordadas as questões inerentes à mineração de dados em imagens. O primeiro é o processo maior do qual mineração de dados é uma etapa estratégica, conhecida como Descoberta de Conhecimento em Bancos de Dados (DCBD), ou do termo em inglês Knowledge Discovery in Databases (KDD). O segundo é a área de Sensoriamento Remoto, com seu corpo metodológico e tecnológico de mapeamento e monitoração da superfície da Terra. O terceiro constitui-se do conjunto de técnicas tradicionalmente empregadas na extração de informações de imagens de sensoriamento remoto.

#### 2.1 Descoberta de Conhecimento em Bancos de Dados

Descoberta de conhecimento em bancos de dados é o processo não trivial de identificar em dados padrões que sejam válidos, novos (previamente desconhecidos), potencialmente úteis e compreensivos, visando melhorar o entendimento de um problema ou um procedimento de tomada de decisão [FAY1996]. Examinando estes termos individualmente:

1. Dados: conjunto de fatos  $F$ , como instâncias de um banco de dados. Por exemplo, uma coleção de  $n$  cadastros de pessoas físicas contendo idade, profissão, renda, etc.
2. Padrão: expressão  $E$  em uma linguagem  $L$  descrevendo fatos em um subconjunto  $F_E$  de  $F$ .  $E$  é dito um padrão se é mais simples do que a enumeração de todos os fatos em  $F_E$ . Por exemplo, o padrão “se renda  $< \$r$  então a pessoa não recebe financiamento” seria aplicável para uma escolha apropriada de  $r$ .

3. Processo: geralmente em KDD, processo é uma seqüência de vários passos que envolve preparação de dados, pesquisa de padrões, avaliação de conhecimento, refinamento envolvendo iteração e modificação.
4. Validade: os padrões descobertos devem ser válidos em novos dados com algum grau de certeza. Uma medida de certeza é uma função  $C$  mapeando expressões em  $L$  para um espaço de medidas  $M_C$ . Por exemplo, se um limite de padrão de crédito é ampliado, então a medida de certeza diminuiria, uma vez que mais financiamentos seriam concedidos a um grupo até então restrito a esta operação.
5. Novo: em geral, assume-se que “novidade” pode ser medida por uma função  $N(E,F)$ , que pode ser uma função booleana ou uma medida que expresse grau de “novidade” ou “surpresa”. Exemplo de um fato que não é novidade: sejam  $E = \text{“usa tênis”}$  e  $F = \text{“alunos de colégio”}$  então  $N(E,F) = 0$  ou  $N(E,F) = \textit{false}$ . Por outro lado: sejam  $E = \text{“bom pagador”}$  e  $F = \text{“trabalhador da construção civil”}$  então  $N(E,F) = 0,85$  ou  $N(E,F) = \textit{true}$ .
6. Potencialmente útil: padrões devem potencialmente levar a alguma atitude prática, conforme medido por alguma função de utilidade. Por exemplo, regras obtidas no processo podem ser aplicadas para aumentar o retorno financeiro de uma instituição.
7. Compreensível: um dos objetivos de KDD é tornar padrões compreensíveis para humanos, visando promover uma melhor compreensão dos próprios dados. Embora seja um tanto subjetivo medir compreensibilidade, um dos fatores freqüentes é a medida de simplicidade. O fator de compreensão dos dados está relacionado à intuitividade da representação destes, bem como da granularidade alta o suficiente para que estes sejam compreendidos. Por exemplo: o log de um servidor Web não é uma representação compreensível; já fatos estatísticos extraídos deste log, tais como totais de acesso ou classificação dos acessos realizados, fornecem informação num formato mais intuitivo e de granularidade humanamente compreensível.

O processo de KDD é interativo, iterativo, cognitivo e exploratório, envolvendo vários passos (Figura 2.1) com muitas decisões sendo feitas pelo analista (que é um

especialista do domínio dos dados, ou um especialista de análise dos dados), conforme descrito:

1. Definição do tipo de conhecimento a descobrir, o que pressupõe uma compreensão do domínio da aplicação bem como do tipo de decisão que tal conhecimento pode contribuir para melhorar.
2. Criação de um conjunto de dados alvo (Selection): selecionar um conjunto de dados, ou focar num subconjunto, onde a descoberta deve ser realizada.
3. Limpeza de dados e pré-processamento (Preprocessing): operações básicas tais como remoção de ruídos quando necessário, coleta da informação necessária para modelar ou estimar ruído, escolha de estratégias para manipular campos de dados ausentes, formatação de dados de forma a adequá-los à ferramenta de mineração.
4. Redução de dados e projeção (Transformation): localização de características úteis para representar os dados dependendo do objetivo da tarefa, visando a redução do número de variáveis e/ou instâncias a serem consideradas para o conjunto de dados.
5. Mineração de dados (Data Mining): selecionar os métodos a serem utilizados para localizar padrões nos dados, seguida da efetiva busca por padrões de interesse numa forma particular de representação ou conjunto de representações; busca pelo melhor ajuste dos parâmetros do algoritmo para a tarefa em questão.
6. Interpretação dos padrões minerados (Interpretation/Evaluation), com um possível retorno aos passos 1-6 para posterior iteração.
7. Implantação do conhecimento descoberto (Knowledge): incorporar este conhecimento à performance do sistema, ou simplesmente documentá-lo e reportá-lo às partes interessadas.

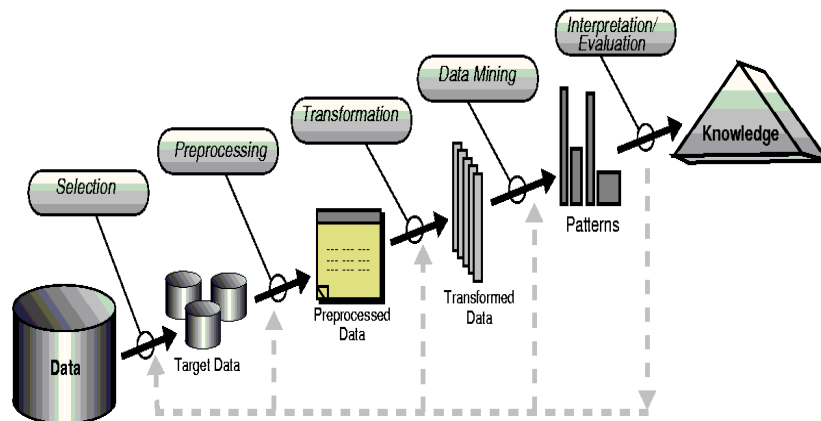


Figura 1: Etapas de KDD [FAY1996]

Visando uma exemplificação da aplicabilidade de KDD, são apresentados a seguir casos onde a descoberta de conhecimento em bancos de dados pode desempenhar tarefas relevantes [WIT2000]:

- Instituições financeiras demandam de seus clientes dados pessoais e financeiros, os quais são utilizados para decidir sobre a consolidação ou não de empréstimos. Num primeiro passo, geralmente métodos estatísticos são utilizados para determinar clientes com perfil bem definido. Os demais casos, que encontram-se no limite entre aceitação e rejeição, precisam de análise humana. A aplicação de KDD sobre um banco de dados histórico, com vários cadastros e diferentes atributos, geraria automaticamente regras que caracterizariam bons e maus clientes. Tais regras poderiam ser aplicadas visando um aumento na taxa de sucesso nas operações de empréstimo.
- A manutenção preventiva de elementos eletromecânicos pode evitar falhas que interrompam processos industriais. A verificação de dispositivos é realizada periodicamente, ocasião na qual técnicos inspecionam componentes, medindo vibrações e outros fenômenos que indicam necessidade de manutenção. Estas aferições são muito ruidosas, devido às limitações dos procedimentos de medição e registro. Estes dados, uma vez estudados por um especialista, conduzem a um diagnóstico. As limitações dos procedimentos técnicos, aliadas à subjetividade

humana, oferecem uma margem de erro considerável. Por outro lado, um universo de 600 falhas, cada uma devidamente registrada representando 20 anos de experiência, pode ser utilizado para determinar tipos de falhas através de procedimentos de KDD, aperfeiçoando assim o processo de busca e correção de problemas.

## 2.2 Sensoriamento Remoto

Sensoriamento remoto é uma área das ciências aplicadas voltada à aquisição de informações da superfície da Terra através de dispositivos que realizam o sensoriamento e a gravação da energia refletida ou emitida, com posterior processamento, análise e aplicação destas informações. Estes dispositivos são denominados sensores remotos, os quais são colocados a bordo de aeronaves ou de satélites de sensoriamento remoto - também chamados de satélites de observação da Terra. Sistemas de imageamento (Figura 2.2) envolvem diferentes elementos e procedimentos, sendo os mais relevantes [CCR2003]:

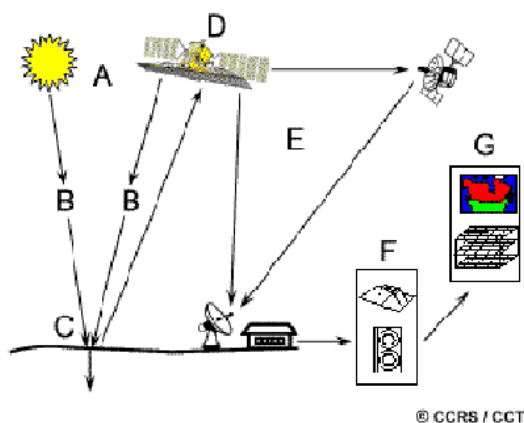


Figura 2: Sistema de imageamento [CCR2003]

- Fonte de energia (A) – provê energia aos alvos de interesse, em forma de radiação eletromagnética. Esta radiação possui propriedades fundamentais e comporta-se de forma previsível, viajando à velocidade da luz. O comprimento de onda é uma característica muito importante da radiação eletromagnética, sendo definida como o comprimento de um ciclo de onda (a medida da distância entre ondas sucessivas).

Outra característica relevante é a frequência, que refere-se ao número de ciclos de uma onda passando por um ponto fixo por unidade de tempo. O espectro eletromagnético (Figura 2.3) varia dos comprimentos de onda mais curtos (raios gama e x) aos mais longos (microondas e ondas de rádio). Diferentes regiões do espectro são úteis ao sensoriamento remoto como, por exemplo, a região ultravioleta, cuja radiação faz com que alguns materiais da superfície terrestre (especialmente rochas e minerais) emitam energia quando iluminados por esta radiação. Já a região do espectro visível pode ser detectada pela visão humana (0.4 a 0.7  $\mu\text{m}$ ). Azul, verde e vermelho são as cores (ou comprimentos de onda) primárias, pois nenhuma delas pode ser criada a partir das outras duas, mas todas as outras cores podem ser criadas combinando-se azul, verde e vermelho. Outras regiões de interesse para o sensoriamento remoto são o infravermelho e o termal.

- Radiação e atmosfera (B) – o percurso da energia, da fonte ao alvo, provoca contato e interação da mesma com elementos atmosféricos (partículas e gases), afetando a radiação através de dois mecanismos. O primeiro, denominado espalhamento, causa o redirecionamento da radiação eletromagnética, segundo o comprimento de onda, a abundância de partículas ou gases e a distância percorrida pela radiação na atmosfera. O segundo mecanismo, a absorção, faz com que moléculas na atmosfera absorvam energia em vários comprimentos de onda.

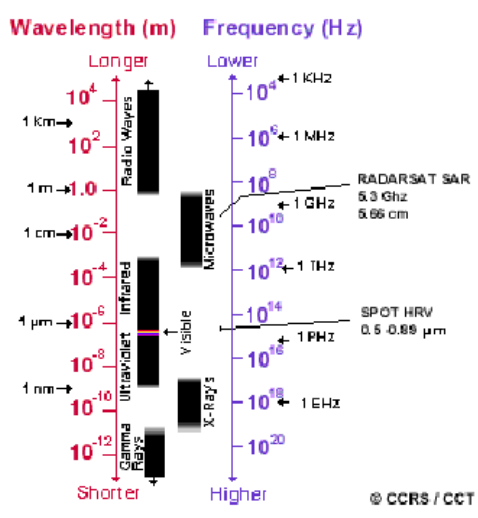


Figura 3: Espectro eletromagnético [CCR2003]

- Interação da radiação com o alvo (C) – após atravessar a atmosfera, a radiação que não foi espalhada nem absorvida interage com alvos terrestres através de absorção (radiação é absorvida pelo alvo), transmissão (radiação passa através do alvo) e/ou reflexão (radiação é refletida e redirecionada). A proporção de cada uma destas formas de interação depende do comprimento de onda e do material e condição do alvo. A reflexão, interação mais relevante para o sensoriamento remoto, pode ser especular (superfície lisa, onde toda ou praticamente toda a energia é redirecionada numa única direção - “espelhada”) ou difusa (quando a superfície é rugosa e a energia é refletida quase uniformemente em todas as direções). A partir de medidas da energia refletida (ou emitida) por alvos terrestres em diferentes comprimentos de onda, pode-se construir respostas espectrais destes elementos, cujos padrões de resposta torna possível distingui-los.
- Gravação da energia pelo sensor (D) – a radiação eletromagnética refletida ou emitida pelos alvos é coletada e gravada pelos sensores. Neste processo, são fundamentais quatro resoluções: espacial, espectral, radiométrica e temporal. A resolução espacial é responsável pelo nível de detalhe que pode ser discernido numa imagem, ou seja, se um sensor possui resolução de 20 metros, então cada pixel representa uma área de 20m x 20m. Já a resolução espectral determina a capacidade do sensor para definir pequenos intervalos de comprimento de onda, pois quanto mais fina a resolução espectral, mais estreita a faixa de comprimento de onda de uma banda em particular. No caso da resolução radiométrica, esta determina a habilidade de discriminação de diferenças muito leves de energia; quanto mais fina a resolução radiométrica de um sensor, mais sensível este é para detectar pequenas variações na radiação emitida ou refletida (níveis de cinza). A resolução temporal determina o tempo necessário para o sensor visitar determinado alvo, ou seja, o tempo necessário que um satélite leva para completar um ciclo de órbita completo; se um sensor pode obter a imagem de determinada área a cada 16 dias, então diz-se que a resolução temporal deste é de 16 dias.
- Transmissão, recepção e processamento (E) – a energia gravada pelo sensor deve ser transmitida para posterior processamento. A transmissão de dados pode ser feita

imediatamente à Terra para Estações de Recepção em Solo, ou então posteriormente transmitidas após gravação dos dados a bordo do satélite. Uma terceira opção seria a transmissão de dados entre uma série de satélites, objetivando atingir a estação de recepção apropriada. Os dados digitais (brutos) recebidos devem ser, quando necessário, processados para a correção de distorções atmosféricas, geométricas e sistemáticas, visando a sua tradução em um formato padrão, gravação em mídia apropriada e posterior disponibilização de imagens.

- Interpretação e análise (F) – a imagem é visualizada e interpretada, visando a identificação, medição e obtenção de informações a respeito dos alvos. A observação destes alvos e seus cenários envolve a avaliação de tons, formas, estruturas, tamanhos, padrões, arranjos espaciais, textura e sombra destes elementos, bem como suas associações e variações. Neste ponto geralmente é realizada a análise de várias bandas com diferentes datas, além da composição e manipulação de imagens visando melhorar sua aparência através de filtragens e transformações, utilizando-se ainda segmentação, classificação e processos de integração de dados.
- Aplicação (G) – o elemento final do processo de sensoriamento remoto é atingido quando a informação extraída é aplicada para a melhor compreensão de alvos e cenários, para revelar novas informações ou auxiliar na solução de um problema em particular nos mais diferentes domínios – levantamento de desflorestamento, vigilância territorial, detecção de manchas de óleo no mar, previsão de safra agrícola, planejamento urbano, dentre outros.

### **2.3 Técnicas Usuais de Extração de Informações em Imagens de Sensoriamento Remoto**

Sendo a imagem a representação de um objeto ou fenômeno, esta produz estímulo visual que caracteriza a cena original. Daí, o fato de a interpretação visual de imagens ser a abordagem primordial para extrair informações a partir destas. Pesquisadores, ambientalistas, cartógrafos e demais profissionais, a partir de técnicas e da sua experiência em determinadas regiões ou “padrões” de paisagem, podem interpretar visualmente uma imagem determinando e caracterizando os elementos e

relacionamentos presentes na mesma. Neste caso, fica evidente o caráter subjetivo da inspeção visual da imagem, influenciada fortemente pela percepção do observador, sua capacidade discriminativa e seu espírito crítico e qualitativo.

No processo de análise e extração de informações a partir de imagens pode-se utilizar algoritmos de segmentação, os quais particionam a imagem em regiões correspondentes às áreas de interesse do ponto de vista da aplicação (Figura 2.4). Uma região é definida como um conjunto de pixels contíguos, com espalhamento bidimensional, os quais apresentam uniformidade em relação a determinado atributo. Área, forma, elementos estatísticos e textura são exemplos de atributos que podem ser obtidos e utilizados na análise da imagem. Os algoritmos de segmentação utilizam basicamente três abordagens: crescimento de regiões, detecção de bordas e combinação de ambas [FON2000].

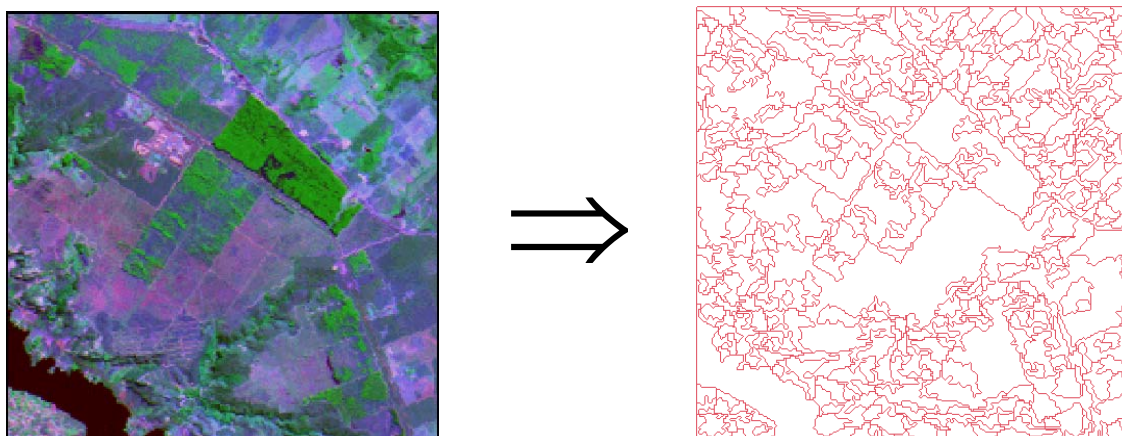


Figura 4: Segmentação de imagem

Na técnica de crescimento de regiões, o algoritmo divide a imagem em um número de regiões homogêneas, cada uma identificada por um rótulo, cujo processo iterativo faz com que regiões espacialmente adjacentes sejam agrupadas segundo algum critério de similaridade. O resultado final é uma imagem rotulada. No método de detecção de bordas, as posições de pixels com variações abruptas de níveis de cinza entre regiões homogêneas (bordas) são extraídas gerando como resultado um mapa de bordas, cujos pontos caracterizam as transições entre objetos distintos. A seleção de uma destas

abordagens ou a combinação delas depende fortemente dos tipos de dados usados na análise e da área de aplicação.

Já a classificação algorítmica de imagens é definida como o processo de extração de informação em imagens para reconhecer padrões e objetos homogêneos com o objetivo de mapear as áreas da superfície terrestre. Como resultado do processo obtém-se uma imagem temática, cujos pixels classificados são representados por padrões gráficos ou cores (Figura 2.5), onde cada cor ou símbolo encontra-se associado a uma classe (tipo de solo, tipo de vegetação, etc.) definida pelo usuário. Os algoritmos de classificação analisam individualmente os atributos numéricos de cada pixel na imagem, caracterizando assim uma abordagem quantitativa, com o intuito de encontrar as fronteiras de decisão entre as classes. A classificação pode ser supervisionada ou não supervisionada [FON2000].

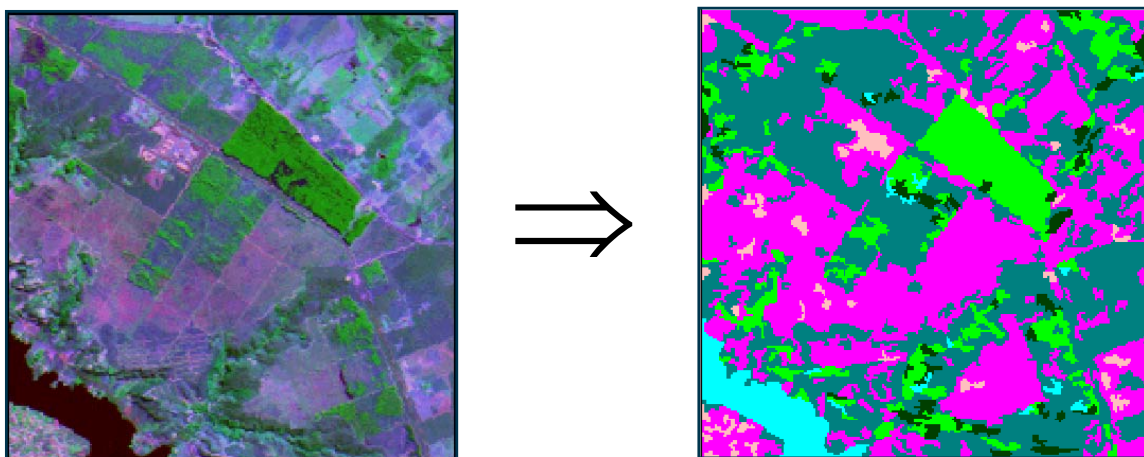


Figura 5: Classificação de imagem

No caso da classificação não supervisionada, cada pixel é associado a uma classe espectral da imagem (o usuário não possui conhecimento prévio do número ou identificação das classes) utilizando para isto algoritmos de agrupamento. Na classificação supervisionada, o usuário indica amostras representativas para cada classe que deve ser identificada na imagem. Através de funções e parâmetros estatísticos, o algoritmo define (a partir das amostras representativas) a que classe cada pixel pertence.

Os classificadores podem ainda ser divididos em classificadores pixel a pixel e classificadores por regiões. Os primeiros utilizam apenas a informação espectral isolada de cada pixel para identificar as regiões homogêneas. Já a classificação por regiões é realizada, basicamente, em duas etapas: primeiro a imagem é particionada em regiões (segmentação) e, posteriormente, cada região é associada a uma classe, baseando assim o processo nas propriedades espectrais e espaciais destas imagens.

É fundamental salientar que a interpretação visual, a segmentação e a classificação são métodos importantes e complementares, que freqüentemente são usados de forma combinada visando obter os melhores resultados na tarefa de extração de informação em imagens. Entretanto, estas técnicas são eficientes para análise de um número limitado de imagens. Daí a necessidade de buscar novas técnicas, algoritmos e metodologias que fomentem a extração automática ou semi-automática de informações estratégicas a partir de grandes acervos de imagens de sensoriamento remoto de forma eficiente.

## CAPÍTULO 3

### Tecnologias e Desafios na Recuperação de Imagens

Novas tecnologias inerentes à extração de conhecimento a partir de acervos de imagens passam, necessariamente, pelo problema da recuperação de imagens, ou seja, acesso eficiente, flexível e funcional a unidades ou conjuntos de imagens em determinado repositório [SDC2002]. Neste contexto, serão apresentadas três linhas de pesquisa que apontam nesta direção: busca por similaridades, linguagem de consulta para recuperação por conteúdo, e consulta a imensos bancos de imagens.

#### 3.1 Busca por Similaridades

A abordagem de busca por similaridades, desenvolvida inicialmente para uso em imagens médicas [PET1997], visa suportar consultas a partir de conteúdos num banco de dados de imagens (BDI). Pontos relevantes neste domínio estão relacionados à extração de características da imagem, representação de conteúdo da imagem, organização da informação armazenada, e estratégias de busca e recuperação.

A implementação do método requer que todas as imagens sejam analisadas antes do armazenamento, pois desta forma descrições apropriadas de seus conteúdos podem ser extraídas e armazenadas no banco juntamente com as imagens originais. Estas descrições são utilizadas então para realizar buscas no BDI, permitindo determinar quais imagens satisfazem o critério de seleção de consultas.

Idealmente, estas consultas deveriam ser especificadas através de uma interface gráfica para o usuário. A “consulta por exemplos” (query by example) permite que o usuário especifique vários objetos com formas complexas e inter-relacionamentos, podendo solicitar todas as imagens contendo objetos similares e relacionamentos semelhantes. Entretanto, imagens recuperadas não precisam ser exatamente semelhantes à consulta, pelo contrário, a pesquisa no banco deve ser aproximada (via graus de similaridade).

Um típico problema deste domínio seria: dado um conjunto de imagens, recupere aquelas que são semelhantes a uma imagem exemplo fornecida. A metodologia proposta ataca este tipo de problema (i) utilizando uma representação eficiente de conteúdo de imagens denominada ARGs (Attributed Relational Graphs), (ii) indexando os ARGs armazenados através de R-Trees, (iii) suportando recuperação por conteúdo aproximado de imagens (baseada em propriedades dos objetos e relacionamentos entre estas) e, (iv) suportando ainda “queries by example” [PET1997].

A partir da coleção de imagens, deve-se derivar representações apropriadas de seus conteúdos e organizar as imagens e suas representações no banco de dados. Desta forma, pode-se buscar eficientemente imagens semelhantes a uma imagem exemplo fornecida. Todas as imagens são segmentadas em contornos fechados correspondentes a regiões ou objetos dominantes da imagem. Assume-se que as imagens possuem um determinado número de objetos rotulados. A rotulação dos objetos é realizada antes de seu armazenamento, e uma classe ou nome é associado a cada um deles (imagens podem conter objetos não rotulados). Neste contexto, considera-se que imagens contenham um número fixo de objetos rotulados e um número variável de objetos não rotulados.

Descrições das imagens são elaboradas em termos de propriedades dos objetos, e em termos de relacionamentos entre os objetos. Estas descrições serão efetuadas através de Grafos Relacionais de Atributos (ARGs) [BAL1982]. Num ARG, os objetos são representados por nós de grafos, e os relacionamentos entre objetos são representados por arcos entre tais nós (Figura 3.2). Tanto nós como arcos são rotulados por atributos que correspondem às propriedades dos objetos e relacionamentos entre estes. As características específicas utilizadas no ARG são derivadas da imagem original e, dependendo da aplicação, podem ser geométricas, estatísticas, texturais ou particulares de algum domínio.

Através desta representação, o problema de recuperação de imagens a partir de um exemplar torna-se um problema de busca num banco de ARGs, pois dada uma consulta seu ARG deve ser computado e buscado dentre os ARGs armazenados. Nestas buscas, a

comparação entre ARGs necessita de uma medida de comparação, a qual foi definida da seguinte forma: Seja  $Q$  uma imagem de consulta contendo  $q$  objetos, e  $S$  uma imagem armazenada que possui  $s$  objetos; seja  $F()$  um mapeamento de objetos em  $Q$  para objetos em  $S$ ; o custo deste mapeamento é definido pelo custo de comparar nós associados mais o custo de comparar os relacionamentos entre estes nós. Este custo é computado como infinito se na imagem armazenada que está sendo comparada algum dos objetos esteja ausente. A avaliação deste custo para recuperação da imagem é um processo que envolve graus de similaridade, pois conteúdos de diferentes imagens não são exatos nem precisos. Assim, o desenvolvimento de funções apropriadas de similaridade é um ponto chave e dependente da aplicação. Outro ponto fundamental é o método de acesso espacial, cuja abordagem escolhida foi o método de indexação R-tree [GUT1984], considerado mais robusto para o endereçamento de espaços com alta dimensionalidade.

Um exemplo de aplicação é ilustrado na Figura 3.1, onde uma imagem médica de tomografia é segmentada e seus objetos rotulados, inclusive os desconhecidos (unlabeled). Os nós correspondem a objetos, e os arcos a relacionamentos entre os objetos. O conjunto de características de objetos utilizado foi o seguinte: tamanho ( $s$ ), computado como a área ocupada; circularidade (roundness) ( $r$ ), definida como a razão do menor pro maior segundo momento; orientação ( $o$ ), computada como o ângulo entre a direção horizontal e o eixo de prolongamento. Já as propriedades utilizadas para descrever os relacionamentos espaciais entre dois objetos foram: distância ( $d$ ), como sendo a distância mínima entre todos os pares de objetos; posição relativa ( $p$ ), definida como o ângulo com a direção horizontal da linha que conecta os centros de massa de dois objetos. Este conjunto de características não é único, pois outros elementos (nível de cinza médio, valores de textura etc.) podem ser utilizados.

Um dos desafios neste cenário é obter busca rápida num banco de ARGs. Para isto, os ARGs foram mapeados em pontos de um espaço multidimensional, visando a utilização de métodos de acesso multidimensional. Estes métodos podem manipular propriedades mais relevantes que outras, através de pesos. Técnicas de redução de dimensionalidade provocam uma melhoria expressiva no tempo de resposta das buscas (embora isto

aumento “falsos alarmes”). Suporte a processamento paralelo e métodos alternativos de indexação de estruturas estão dentre as possíveis melhorias a esta abordagem.

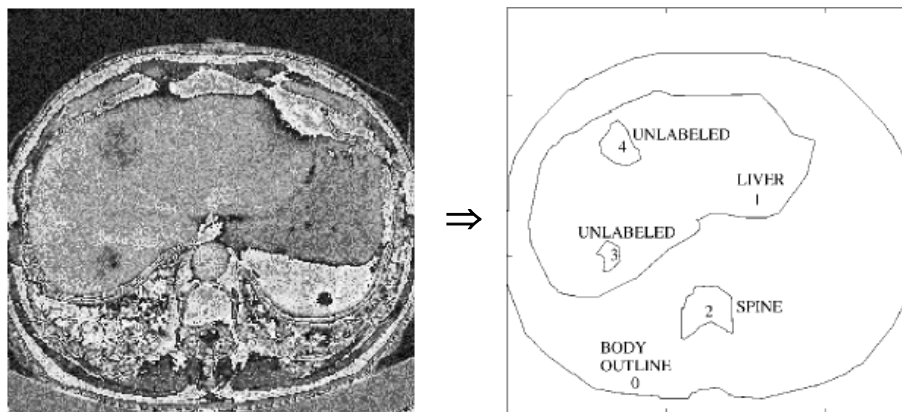


Figura 6: Imagem original, seguida de segmentação e rotulação de objetos

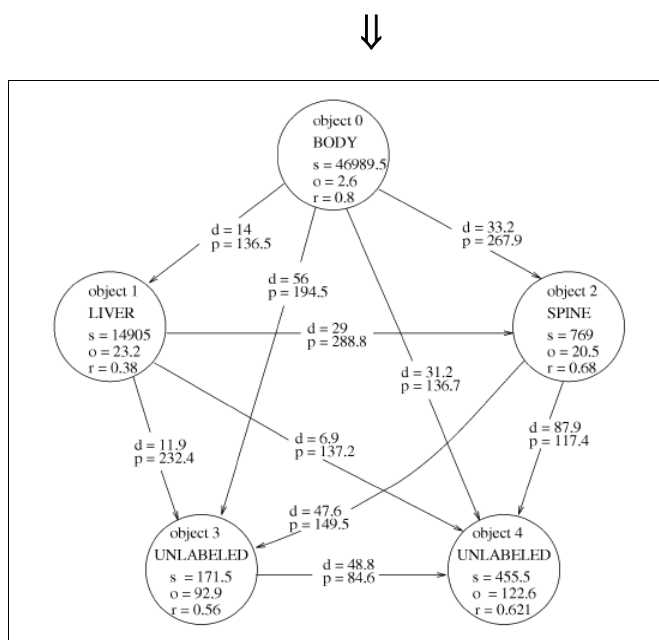


Figura 7: ARG da imagem [PET1997]

### 3.2 Linguagem de Consulta para Recuperação por Conteúdo

Linguagens voltadas à derivação de descrições textuais de conteúdos de imagens são consideradas inapropriadas para recuperação de propósito geral, uma vez que é impraticável gerar representações textuais exaustivas que contenham todas as informações e níveis de detalhe pertinentes quando se trata de acervos de imagens. Por

outro lado, anotações manuais, textos explicativos, documentação de contextos e palavras chaves pré-geradas a partir de imagens conduzem à perda de flexibilidade.

SQL é vista como limitada em seu poder de expressão e extensibilidade, requerendo uma certa experiência e sofisticação por parte do usuário. Mesmo facilitando o processo de interpretação de máquina, e havendo técnicas disponíveis de otimização, escalabilidade e base teórica, SQL não é considerada apropriada e aplicável para o domínio de imagens, pois este possui ambigüidades intrínsecas e incertezas. Já a pré-construção de estruturas de grafos objeto-relacionais é computacionalmente cara e pode requerer consultas complexas para gerar os resultados esperados.

Recuperação de imagens baseada em conteúdo (Content Based Image Retrieval – CBIR) através de linguagem de consulta que permita sentenças curtas tende a ser mais eficiente e intuitivo do que esquemas tradicionais de composição de consultas. A linguagem de descrição de consultas *Oquel* [TOW2001] para recuperação de imagens baseada em conteúdos provê uma arquitetura de linguagem extensível, levando em consideração pontos fundamentais como escopo, complexidade, poder de expressividade, extensibilidade e naturalidade de sintaxe, sendo baseada numa gramática livre de contexto e num vocabulário base. Termos da linguagem representam predicados de características de imagem e direcionam-se a conteúdos em diferentes níveis semânticos, servindo como substantivos, adjetivos e preposições. Sentenças são prescrições de características desejáveis atreladas a imagens relevantes, podendo representar espaço, composição de objetos, e outros relacionamentos entre termos e sub-sentenças.

A linguagem apresenta-se portátil, pois os termos de baixo nível e as funções de avaliação que atuam nesta podem ser modificados ou reimplementados com baixo ou nenhum impacto nos elementos conceituais de mais alto nível da linguagem. Esta portabilidade permite a definição de ontologias específicas de objetos. A linguagem foi implementada no sistema de recuperação de imagem baseado em conteúdo denominado ICON [ATT2002], cujo objetivo é permitir aos usuários visualização, navegação e busca de coleções de imagens digitais de forma prática e eficiente. Utilizando uma

arquitetura cliente-servidor e um projeto orientado a objetos, o software apresenta camadas de abstração e encapsulamento permitindo que o sistema seja parametrizável para requisitos específicos de usuários.

*Oquel* visa facilitar a formulação de consultas (através de características desejáveis da imagem) da forma mais concisa possível, evitando descrições exaustivas e semânticas de imagens. A linguagem busca proporcionar maior naturalidade e flexibilidade através do uso de uma gramática mais complexa. Usuários da linguagem podem elaborar desde simples frases-chave até consultas complexas, utilizando recursos como predicados, relações, conjunções e uma sintaxe de especificação para conteúdos de imagem através de adjetivos de propriedades, relacionamentos, localização de objetos, categorias visuais e termos de mais alto nível para descrição de cenas.

Sua implementação possui um thesaurus de centenas de palavras em linguagem natural, frases e abreviações, provendo facilidades para correção automática de erros em sentenças da consulta. É possível manipular a representação sintática da árvore utilizando uma ferramenta gráfica, garantindo que sentenças sejam construídas de acordo com as regras da gramática sem que o usuário compreenda toda a especificação da linguagem.

Nas cláusulas da gramática é possível especificar quantificadores (“muito”, “nenhum”, etc.), metacategorias (“cidade”, “campo”, etc.), relacionamentos (inclusive espaciais – “mais amplo que”, “perto de”, etc.), categorias visuais (“água”, “pele”, etc.), categorias semânticas (“pessoas”, “veículos”, “animais”, etc.), localização (“abaixo de”, “canto superior direito”, etc.), descrição de forma (“linha reta”, “angular”, etc.), descrição de cores e de tamanhos (“bright red”, “10% da área”, “região mais ampla”, etc.).

Algumas sentenças válidas na linguagem seriam:

- some sky which is close to trees in upper corner size at least 20%
- [indoors] or [outdoors] & [people]

- [some green or vividly coloured vegetation in the centre] which is of similar size as [clouds or blue sky at the top]

O resultado da consulta "[*bright red and stripy*] and [*tarmac in bottom half, size >10%*]" em um banco de dados do ICON é ilustrado na Figura 3.3.

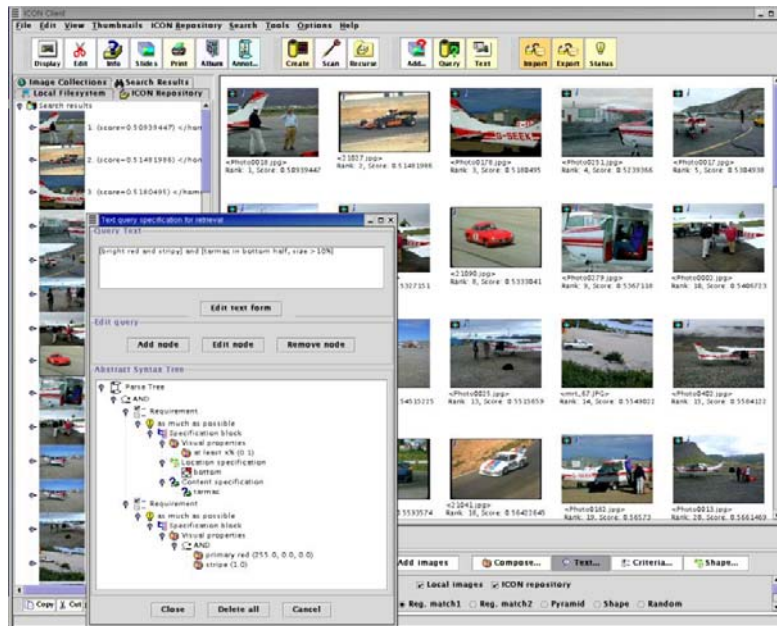


Figura 8: Exemplo de resultado de consulta [TOW2001]

ICON extrai vários tipos de descritores de conteúdo e metadados a partir das imagens através de:

- Segmentação: imagens são segmentadas em regiões disjuntas, sendo gerados conjuntos de propriedades para tamanho, cor, forma, textura de cada região;
- Classificação: descritores de regiões da segmentação alimentam classificadores baseados em redes neurais, os quais são treinados para rotular regiões com probabilidades de membresia em um conjunto de diferentes categorias visuais semânticas (“grama”, “céu”, etc.);
- Grafo de regiões: grafos dos relacionamentos espaciais relativos das regiões (“adjacência”, “distância”, etc.) são gerados;

- Pirâmide Grid: cada categoria visual classificada é representada em diferentes granularidades (espaçamentos - 1x1, quintos de imagem, etc.).

As imagens são recuperadas através da avaliação de uma árvore de sintaxe abstrata derivada da consulta do usuário, visando computar uma probabilidade de relevância para cada imagem (tomando por base metadados, anotações e resultados do processo de classificação). Termos de mais alto nível são expandidos em representações de sentenças ou avaliados através de grafos bayesianos. Em níveis mais baixos, tokens são mapeados diretamente nos descritores de conteúdo. Por exemplo: numa busca por pessoas numa imagem, o sistema analisará a presença e a composição espacial de clusters apropriados de materiais relevantes (roupas, pele, cabelo), relacionando-os com padrões de face e olhos. A evidência é então analisada probabilisticamente para gerar uma estimativa a respeito de presença de pessoas na imagem. Categorias visuais são utilizadas com o intuito de reduzir o espaço de busca, além de prover um recurso intuitivo para a composição de consultas.

Características futuras incluem a aprendizagem autônoma de novos conceitos, técnicas mais sofisticadas de processamento de linguagem natural visando reduzir restrições gramaticais, além da incorporação de conhecimento a priori na estrutura lingüística das consultas.

### **3.3 Consultando Imensos Bancos de Imagens**

Uma vez que a aquisição de dados raster está se tornando mais fácil e menos cara, verifica-se uma crescente demanda por acesso flexível e transparente por parte dos usuários. Recursos eficientes para armazenamento e recuperação de arquivos raster viabilizam a disponibilização em tempo real de acervos de imagens. Geralmente, o armazenamento é realizado de forma que cada imagem seja gravada em um arquivo, orientado mais pelo processo de aquisição do que pelos padrões de acesso do usuário. Bancos de dados são geralmente utilizados apenas para pesquisa em metadados, mas não para a recuperação da imagem em si. A expectativa neste contexto baseia-se na formulação de requisições de imagens raster utilizando consultas que seriam

respondidas por um sistema gerenciador de banco de dados (SGBD) através de imagens apropriadas.

Uma iniciativa neste contexto é o TerraServer [BGS2000], que armazena imagens raster e metadados num SGBD relacional (Microsoft SQL Server), disponibilizando consultas ao acervo via Internet através de navegador web gráfico. Metadados são armazenados em tabelas e as imagens (comprimidas) em campos BLOB. O projeto demonstra escalabilidade na questão do armazenamento e recuperação de grandes acervos de imagens oferecendo, entretanto, um número limitado de funções e apresentando deficiência em pontos estratégicos de manipulação de imagens.

Verifica-se, então, que SGBD's tradicionais (relacionais, orientados a objetos, etc.) não estão prontos para oferecer um conjunto de serviços satisfatórios para dados raster, porque arrays multidimensionais (multidimensional discrete data - MDD) formam uma categoria de informação distinta dos elementos inerentes a estes sistemas (conjuntos, objetos, textos, entre outros). Portanto, o suporte a estes arrays requerem novas técnicas em vários aspectos da arquitetura dos SGBD's, dentre os quais modelagem, suporte a consultas, otimização, gerenciamento de armazenamento.

A proposta do RasDaMan [BHS2002] oferece serviços de um SGBD para dados raster, através da implementação de um middleware (o qual manipula MDD) que roda sobre um sistema de banco de dados relacional (Oracle). Sua principal aplicação é suportar serviços que garantam o fornecimento contínuo de serviços e dados geográficos atualizados por parte do serviço público de cartografia e pesquisa do estado de Bavaria – Alemanha (Bavarian Land Survey Authority – BLVA).

Sistemas anteriores no BLVA possibilitavam o acesso de usuários a imagens apenas através de identificadores. Isto gerava erros de requisição e acesso a dados, obrigava o usuário a adquirir imagens inteiras, quando seu interesse estava voltado a pequenas áreas, além de limitá-lo a imagens com resoluções fixas. Requisitos fundamentais para o RasDaMan incluem: acesso a qualquer área em qualquer nível de zoom dentro de um segundo; manipulação de todos os tipos de dados raster em diferentes níveis de resolução; flexibilidade para realizar (além de zoom e pan) coloração, exportação em

alta resolução, sobreposição, etc.; linguagem de consulta para oferecer flexibilidade e abertura do sistema; armazenamento de imagens em banco de dados, visando obter benefícios clássicos de SGBD's (interface de alto nível, suporte a transações, otimização de consultas, etc.); arquitetura multicamada com interfaces abertas e que respeitem os padrões da área, especialmente o Web Map Server (WMS) e o Web Coverage Server (WCS), ambas especificações do Open GIS Consortium (OGC) [OGC2003].

A interface de cliente do servidor BLVA é baseada no WMS, o qual é implementado em Java utilizando o engenho de servlet Tomcat (Figura 3.4), enquanto componentes do servidor baseiam-se num acoplamento entre o servidor raster multidimensional RasDaMan e o SGBD Oracle. A independência de plataforma da interface OGC WMS possibilita a implementação de uma variedade de front-ends na camada cliente.

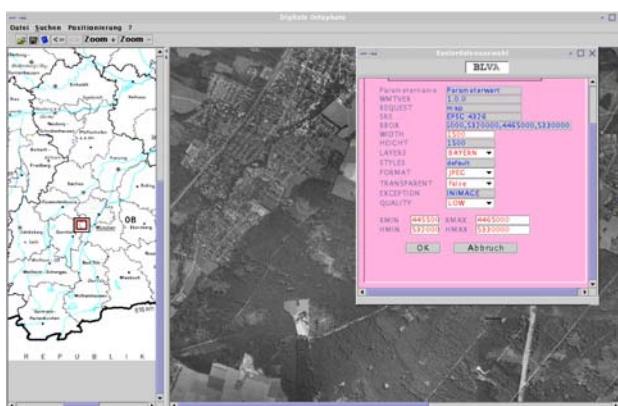


Figura 9: Aspecto da interface do cliente Rasdaman [BHS2002]

O servidor possui uma linguagem de consulta (RasQL) que estende SQL92 com operadores raster declarativos. A avaliação das consultas conta com otimização algébrica e um gerenciador de armazenamento especializado em arrays. O modelo conceitual do servidor está centrado na noção de n-D array (que pode ser de qualquer dimensão, extensão espacial e possuir qualquer tipo de célula), suportando inclusive conjuntos de arrays. As consultas são analisadas (parsing), otimizadas e executadas no servidor RasDaMan. Esta execução ocorre de forma paralela, utilizando paralelismo inter-query.

Um exemplo de requisição ao sistema seria: *“forneça um recorte entre (1000,1000) e (2000, 2000) de todas as orto imagens”*. A respectiva consulta RasQL teria o seguinte aspecto: *“select OrthoColl[1000:2000, 1000:2000] from OrthoColl”*. Recursos da linguagem incluem processamento de sinais, operações sobre imagens, estatísticas, etc.

O armazenamento físico no RasDaMan é implementado através do particionamento de um objeto MDD em sub-arrays (tiles), os quais podem ser opcionalmente comprimidos (especialmente quando são transferidos aos clientes). Um indexador espacial R-tree é utilizado para determinar rapidamente os sub-arrays afetados por uma consulta. Tanto os tiles como os índices são armazenados como BLOBs no SGBD relacional, que também armazena o dicionário de dados utilizado pelo RasDaMan.

O sistema atualmente manipula uma imagem contínua de 950.000 x 1.000.000 de pixels, resultante do mosaico de 13.500 ortofotos digitais georeferenciadas. Para agilizar operações de zoom, uma pirâmide de imagens de nove camadas é mantida pelo servidor. Os dados raster no SGBD relacional ocupam cerca de 700 GB.

Testes demonstraram que o sistema atende aos requisitos iniciais de seu projeto, inclusive a realização de operações sobre dados raster em menos de um segundo em hardware comum. Além do Oracle, RasDaMan pode utilizar os SGBD's relacionais IBM DB2 e IBM Informix, além do SGBD Orientado a Objetos O2. Os próximos passos de projeto incluem técnicas de balanceamento de carga para acesso paralelo pesado, otimização de consultas estatísticas complexas, além de hardware e software redundante e tolerante a falhas com disponibilidade de serviço 7x24.

As diferentes abordagens, técnicas e implementações para a recuperação de imagens apresentadas até o momento evidenciam a relevância estratégica deste tópico, onde observa-se uma busca contínua pelo aumento da eficiência, precisão e poder dos recursos disponibilizados. Alguns autores chegam a considerar as operações de recuperação em bancos de imagem uma técnica de mineração [ZHL2002], haja vista seu peso no resultado final do processo de descoberta de conhecimento, devido à ampla utilização destes serviços nas diferentes etapas de KDD, desde a seleção inicial do conjunto de dados até a aplicação do conhecimento descoberto.

## CAPÍTULO 4

### Mineração de Dados em Imagens

Mineração de dados em imagens não é apenas a aplicação de técnicas de mineração de dados no domínio de imagens. Relevantes diferenças colocam-se entre bancos de imagens e bancos de dados “convencionais” (relacionais, orientados a objeto, etc.). O suporte contextual da semântica de um elemento de imagem, a informação espacial que determina o significado de um dado, e a multiplicidade de interpretações de padrões visuais são alguns dos fatores que influenciam e tornam desafiadora a complexa tarefa de extração de conhecimento a partir de acervos de imagens. Arquiteturas de sistema, técnicas de mineração e softwares especializados tentam avançar nesta tarefa, oferecendo recursos para viabilizar o processo de mineração sobre estes acervos.

#### 4.1 Características, Desafios e Potencialidades

Considerando-se adequadas as metodologias de recuperação de imagens e suas respectivas implementações em determinado contexto de mineração espacial, permanecem ainda questões relevantes em relação ao processo de extração de conhecimento. Banco de dados, processamento de imagens, inteligência artificial e estatística configuram um núcleo do esforço multidisciplinar nesta área, contando ainda com expressivos recursos de outras áreas da ciência, inclusive visão computacional [HAN2001].

O processo de mineração de dados em imagens é apresentado na Figura 4.1. As imagens de um acervo (banco de imagens) são recuperadas segundo critérios inerentes à aplicação. A seguir, uma fase de pré-processamento aumenta a qualidade dos dados, os quais são então submetidos a uma série de transformações e de extração de características que geram importantes informações a respeito das imagens. A partir destas informações, a mineração pode ser realizada através de técnicas específicas, com o intuito de descobrir padrões significativos. Os padrões resultantes são então avaliados e interpretados para a obtenção do conhecimento final, que pode ser aplicado no

entendimento de problemas, na tomada de decisões ou outras aplicações estratégicas [ZHL2002].

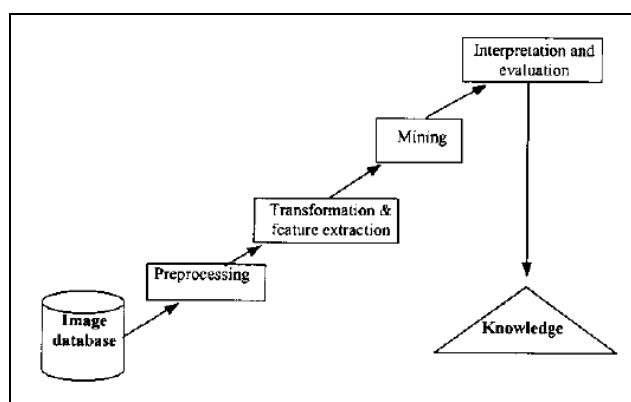


Figura 10: O processo de mineração de imagens [ZHL2002]

Embora o esquema acima assemelhe-se bastante ao da Figura 2.1, deve-se salientar que mineração de dados em imagens não consiste simplesmente na aplicação de técnicas de mineração de dados em bancos “convencionais” ao domínio de imagens. Importantes diferenças entre estes bancos e os bancos de imagens incluem:

- Valores relativos e valores absolutos - em bancos de dados relacionais (por exemplo), os valores de dados são semanticamente significativos. Para ilustrar, tome-se um atributo idade igual a 42; o valor possui um entendimento claro, objetivo. Por outro lado, valores de imagens em si podem não possuir significância sem o suporte de um contexto. No caso de um pixel com valor de escala de cinza igual a 46, torna-se evidente sua dependência dos pixels ao redor para chegar a alguma informação sobre o mesmo (aparência mais clara, mais escura, constitui borda, etc.).
- Informação espacial - outra diferença relevante entre os referidos domínios é que a informação espacial implícita é crítica para a interpretação da imagem, o que não ocorre com bancos de dados relacionais. Com o intuito de minimizar este problema, utiliza-se a extração de características independentes de posição em cada imagem (metadados, anotações, etc.) antes da mineração de padrões espaciais (objetos, seus relacionamentos, etc.).

- Interpretação única e interpretação múltipla - uma característica das imagens é a interpretação múltipla dos mesmos padrões visuais. Neste caso, algoritmos tradicionais de associação de padrões a uma classe (interpretação) não serão aplicáveis. Novas categorias de algoritmos são então necessários, haja vista as necessidades especiais quando realiza-se a mineração e interpretação de padrões a partir de imagens.
- Representação visual dos padrões descobertos - neste tópico, diferentes questões devem ser consideradas: Como representar padrões de imagens de forma que a informação contextual e espacial (além de características importantes da imagem) sejam retidas no esquema de representação? Quais as características relevantes da imagem que devem ser utilizadas no processo de mineração para que os padrões descobertos sejam visualmente significantes? Como representar padrões minerados ao usuário num ambiente visualmente rico?

Um bom exemplo desta problemática é a questão das mudanças de uso e cobertura da terra (Land-Use and Land-Cover Change - LUCC) [LAM2003]. Os temas chave da agenda de pesquisa LUCC incluem: padrões de mudança de cobertura da terra, processos de mudança de uso da terra, respostas humanas à mudança de uso e cobertura da terra, modelos integrados globais e regionais, desenvolvimento de banco de dados de superfícies terrestres, processos biofísicos e seus direcionadores. Estas e outras questões encontram-se certamente retratadas, em maior ou menor grau, nos grandes acervos de imagens das instituições. O problema é exatamente extrair (de forma automática ou semi-automática) estes padrões, tendências e relações de grandes séries temporais de imagens.

Em âmbito nacional, um caso específico refere-se à evolução de padrões de uso da terra em áreas da Amazônia [ESC2003]. Este trabalho, baseado na delimitação de regiões que apresentam padrões espaciais observáveis em séries temporais de imagens (sensor TM dos satélites Landsat), desenvolve uma análise da evolução do uso e cobertura da terra na região Centro-Norte de Rondônia. Dentre os vários elementos abordados na pesquisa, destacam-se:

- Verificação da efetividade da distribuição de terras do INCRA (Instituto Nacional de Colonização e Reforma Agrária);
- Análise da influência das estradas (BR's, secundárias, etc.) na ocupação e na configuração de áreas;
- Preservação de áreas (inclusive indígenas) na atual concepção de assentamentos agrícolas;
- Evolução da vegetação secundária (distribuição espaço-temporal, espécies predominantes, etc.);
- Identificação do tamanho de propriedades e de abandono de áreas;

Estes elementos configuram padrões e tendências espaço-temporais, objetos e relacionamentos de alto nível, que demandam a análise de séries de imagens para extração de informações estratégicas. Neles pode-se identificar os problemas e tópicos supracitados: valores relativos/absolutos, informação espacial implícita, interpretação única/múltipla, representação visual de padrões. Além disso, os métodos geralmente utilizados levam os analistas a selecionarem subconjuntos relevantes destas séries para análise detalhada, incorrendo assim no risco de perda de fenômenos e processos observáveis apenas em conjuntos de pequena granularidade. Neste contexto, metodologias e algoritmos apropriados de mineração de dados espacial certamente aprimorariam o resultado final de muitos trabalhos, podendo até enriquecer ou alterar conclusões.

#### **4.2 Arquitetura Dirigida à Informação**

Um desafio fundamental na mineração de imagens é determinar como a representação de pixel (baixo nível) que está contida numa imagem “bruta”, ou numa seqüência de imagens, pode ser processada para identificar objetos e relacionamentos espaciais em alto nível. A proposta de uma arquitetura direcionada à informação (Information-Driven Framework) [ZHL2001] destaca o papel da informação em vários níveis de representação (Figura 4.2). Quatro níveis de informação são distintos na arquitetura:

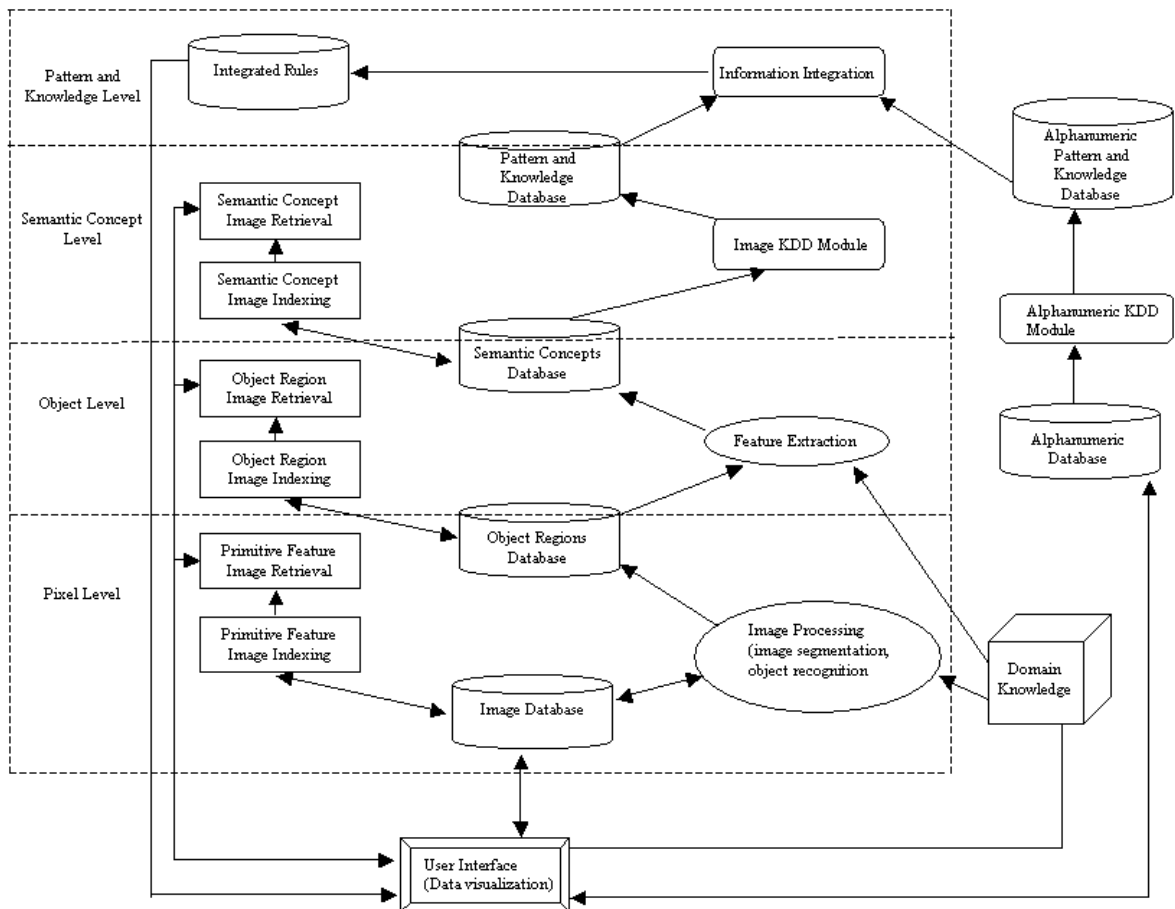


Figura 11: Arquitetura de mineração de imagens direcionada à informação [ZHL2001]

- O nível mais baixo é o Nível de Pixel que consiste das informações da imagem “bruta”, tais como pixels, e de características primitivas da imagem, tais como cor, textura e forma.
- O Nível de Objeto lida com informações de objetos ou regiões homogêneas baseadas nas características primitivas do Nível de Pixel, identificando características específicas ao domínio. Algoritmos de agrupamento (clustering) e segmentação, aliados ao conhecimento do domínio, podem ajudar a particionar as imagens em objetos ou regiões significativas. Modelos de objetos podem ser fornecidos para a identificação de elementos na imagem.

- O Nível de Conceito Semântico coloca os objetos e regiões (identificados no Nível de Objeto) no contexto das imagens, tentando capturar conceitos abstratos no cenário formado. Raciocínio em alto nível e técnicas de descoberta de conhecimento (classificação e agrupamento de imagens, mineração de regras de associação, etc.) são utilizadas para gerar conceitos semânticos de alto nível e para descobrir padrões interessantes.
- O Nível de Padrões e Conhecimento integra dados alfanuméricos relacionados ao domínio com relacionamentos semânticos descobertos nos dados da imagem. Outros passos de mineração são efetuados para descobrir correlações úteis entre dados alfanuméricos e padrões de imagem.

Os quatro níveis de informação podem ser generalizados em duas camadas: os níveis de Pixel e de Objeto formam a camada mais baixa, enquanto os níveis de Conceito Semântico e o de Padrões e Conhecimento formam a camada superior. A primeira contém informações “brutas” e também as extraídas da imagem, caracterizando-se por processos como análise, processamento e reconhecimento da imagem. A camada mais alta lida com operações de imagem em alto nível, como geração de conceito semântico e descoberta de conhecimento a partir de coleções de imagens.

A busca da informação em imagens demanda conhecimento, esforço e habilidade em diferentes domínios, metodologias e ferramentas, as quais são muitas vezes distintas, com pouco ou nenhum grau de integração. Uma arquitetura baseada em níveis da informação avança na questão do fluxo desta informação, permitindo a manipulação de dados em diferentes formatos, estruturas e níveis semânticos, sem perder de vista a integração da informação através destas camadas, permitindo que os diferentes níveis trabalhem em conjunto e de forma acoplada.

### **4.3 Técnicas de Mineração de Imagens**

Componentes de arquiteturas para mineração em imagens dependem de técnicas de extração de informação. A precisão e eficiência destas técnicas são determinantes nos procedimentos de descoberta de conhecimento nos acervos de imagens. Dentre elas

podemos citar: reconhecimento de objetos, classificação e agrupamento de imagens, mineração de regras de associação, e redes neurais [ZHL2002].

O reconhecimento de objetos tem sido amplamente pesquisado no campo do processamento de imagens. Modelos de objetos fornecidos a priori permitem que um sistema de reconhecimento encontre objetos do mundo real em imagens. Algoritmos de aprendizagem de máquina e extratores de informação dependem da identificação e reconhecimento de objetos pelo sistema. Estes sistemas consistem basicamente de banco de modelos (contém os próprios modelos e suas características relevantes), detector de características (avalia e associa características relevantes aos objetos), hipotetizador (atribui probabilidades aos objetos na imagem), verificador de hipóteses (avalia as hipóteses através dos modelos e refina a probabilidade do objeto). Desta forma, o sistema seleciona o objeto com maior probabilidade.

A classificação e o agrupamento de imagens efetuam, respectivamente, a classificação supervisionada e não supervisionada de imagens em grupos. Na classificação supervisionada, é fornecida uma coleção de imagens rotuladas (conjunto de treinamento), e o problema é rotular um novo conjunto de imagens ainda não classificadas. O conjunto de treinamento é utilizado para treinar o algoritmo de aprendizagem de máquina segundo a classe definida para a rotulação das novas imagens. Já a classificação não supervisionada (agrupamento ou clustering) visa agrupar uma coleção de imagens não rotuladas em grupos (clusters) significativos, de acordo com o conteúdo de cada imagem, sem qualquer conhecimento ou rótulo prévio. A classificação inteligente de imagens a partir do seu conteúdo é um meio importante de minerar informações valiosas de grandes coleções de imagens, já que os usuários possuem grande interesse em determinar classes ou agrupamentos de todo um acervo, ora com conhecimento a priori, ora sem ele.

A mineração de regras de associação é freqüentemente utilizada para extrair tendências interessantes, padrões e regras em bancos de dados (quando a classe de uma tarefa de mineração não é determinada como no caso da classificação supervisionada). Regras de associação possuem o seguinte formato:  $D_1 \wedge D_2 \wedge \dots \wedge D_N \Rightarrow R_1 \wedge R_2 \wedge \dots \wedge R_N$ , onde o

próprio algoritmo elege os atributos determinantes (lado esquerdo da regra) e os atributos resultantes (lado direito). Associações entre valores dos atributos são geradas na forma de regras, tendo o algoritmo sua ênfase no compromisso entre precisão e cobertura. Para imagens existem duas abordagens principais: na primeira realiza-se a mineração a partir de acervos de imagens, e na segunda a tarefa é realizada tomando-se uma combinação de imagens e dados alfanuméricos associados.

Redes neurais artificiais tentam alcançar o desempenho humano em vários campos, tais como a fala e a “compreensão” de imagens. Uma rede neural é um processador paralelo massivamente distribuído, composto de unidades básicas de processamento, onde cada uma tende ao armazenamento de conhecimento experimental, tornando tal conhecimento disponível para o uso. Alguns dos pontos fortes destas redes são o reconhecimento de padrões, a previsão de tendências e a construção de modelos de dados. Um projeto relevante, baseado em redes neurais, avalia e detecta mudanças em séries temporais de imagens [CLI2001]. O trabalho apresenta uma técnica para o uso de modelagem preditiva para identificar mudanças incomuns em imagens (sem uma noção pré-definida do que é comum ou incomum). As redes são treinadas para prever valores de regiões numa seqüência de imagens. Diferenças substanciais entre os valores esperados e os reais representam então uma mudança incomum (Figura 4.3).



Figura 12: Duas imagens com mudanças incomuns circuladas [CLI2001]

#### 4.4 ADaM – Algorithm Development and Mining

ADaM, projeto da NASA em conjunto com a Universidade de Alabama em Huntsville, é um conjunto de ferramentas de mineração de dados científicos e de imagens. Suas funcionalidades incluem reconhecimento de padrões, processamento de imagens,

otimização, mineração de regras de associação, dentre outros. O sistema é composto por uma série de componentes individuais que podem ser utilizados em conjunto para realizar tarefas complexas. O software possui módulos implementados em C, C++ e componentes Python. Sua versão atual (4.0.0) possui apenas programas executáveis para linha de comando, sendo que futuras versões com outros componentes e funcionalidades estão em desenvolvimento. Um dos focos do projeto é a implementação eficiente de componentes de desempenho crítico, além do cuidado de manter cada componente do sistema o mais independente possível, visando possibilitar a utilização de subconjuntos de módulos apropriados para determinadas aplicações, inclusive aproveitando componentes de terceiros. O software encontra-se disponível para as plataformas Linux e Windows, possuindo apenas a auto-documentação dos próprios executáveis [UAH2003].

#### **4.4.1 Arquitetura**

O ambiente de mineração de dados do ADaM possui uma arquitetura cliente-servidor (Figura 4.4) que suporta uma grande variedade de funções numa estrutura distribuída, com aplicação voltada à extração de informações de grandes arquivos de dados provenientes das ciências da terra. Ele é capaz de detectar fenômenos ou eventos, armazenando a informação para facilitar posterior processamento ou análise. Processamentos personalizados podem ser desenvolvidos através das aplicações cliente, cujas interfaces podem ser baseadas na Web ou em plataforma desktop.

Aplicações distribuídas, independentes de plataforma, comunicam-se com o engenho de mineração através de uma API (Application Program Interface) comum aos clientes utilizando a rede. Bases de dados locais e distribuídas são utilizadas tanto para a leitura de dados como para o armazenamento de resultados [UAH2003a]. O Sistema de Busca a Eventos/Relacionamentos (Event/Relationship Search System) permite aos usuários conduzir buscas a coincidências e testes de relacionamento entre fenômenos minerados e uma variedade de fatores (tais como regiões geográficas, limites políticos, etc.) para um período específico de tempo.

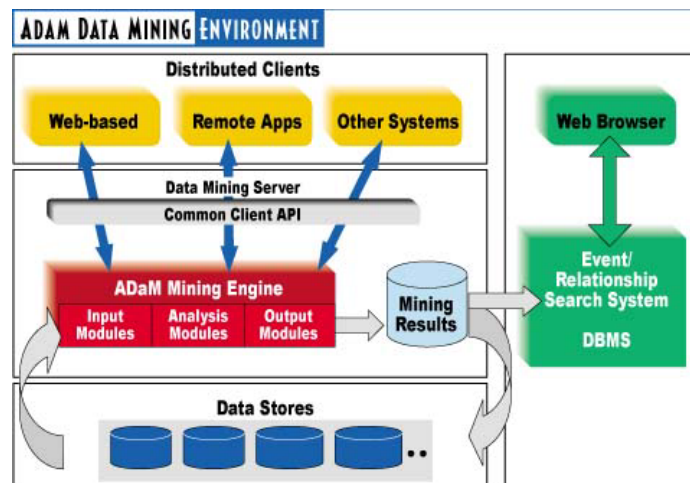


Figura 13: Ambiente do ADaM [UAH2003a]

#### 4.4.2 Serviços

O ADaM possui uma série de serviços relacionados ao reconhecimento de padrões implementados através de técnicas e algoritmos de classificação, agrupamento, seleção de atributos, redução de dados e outros utilitários:

- As técnicas de classificação (supervisionada) geralmente consistem de dois componentes – um módulo de treinamento e outro de aplicação. O primeiro utiliza exemplos de padrões para “aprender” as características da classe de interesse, enquanto o segundo lê as descrições produzidas pelo módulo de treinamento e classifica novos conjuntos de dados. Naive bayes, redes neurais, vizinhos mais próximos e partição recursiva são alguns dos classificadores implementados.
- No que diz respeito às técnicas de agrupamento, suas implementações tomam conjuntos de dados como entrada agrupando-os em classes segundo critérios de similaridade. Estão disponíveis implementações de Isodata, K-Means e Maximin.
- Já algoritmos de seleção de atributos selecionam um subconjunto dos atributos disponíveis através de medidas de relevância destes, enquanto técnicas de redução diminuem o número de instâncias de dados a serem consideradas no processamento através de amostragem e outros procedimentos. Algumas de suas implementações

são eliminação regressiva, seleção progressiva, componentes principais, filtro Relief, remoção específica de atributos.

- Alguns utilitários disponíveis realizam medidas de precisão, remoção de valores fora de faixa, partição/fusão de conjuntos de dados para validação cruzada, cálculo de parâmetros estatísticos, etc.

Um outro conjunto de serviços relevantes presente no software está relacionado a módulos que implementam técnicas de extração de características de imagens. Estes programas de processamento de imagem lêem dados no “formato binário de imagem” (binary image format). ADaM possui utilitários para a conversão entre diferentes padrões de imagem. Operações básicas, segmentação, filtragem e otimização estão dentre os recursos disponíveis:

- Nas operações básicas figuram implementações para mudança de tamanho, de orientação, de escala e de outras propriedades da imagem. Adição e subtração de imagens, equalização de histograma, geração de imagem negativa e estatísticas da imagem também estão presentes no software.
- Para segmentação existem utilitários para localizar bordas, regiões contínuas e polígonos, além de teste de pixels para faixas específicas.
- Operações de filtragem incluem dilatação e erosão da imagem (filtros morfológicos), filtros de média, moda e acoplados com redes neurais (para suavização), dentre outros.
- Características de textura são utilizadas para classificação e segmentação com base na estrutura local da imagem. Existem recursos para computar regras de associação para múltiplas imagens e para encontrar regras de associação que caracterizem uma região da imagem. Além disso, é possível computar filtro de Gabor, co-ocorrência de níveis de cinza, dimensão fractal, dentre outros.

Outros serviços implementados são a mineração de regras de associação (para identificar relacionamentos entre atributos em bases de dados) e as técnicas de

otimização (para encontrar boas soluções para problemas complexos de busca envolvendo espaços de pesquisa muito grandes). Os métodos de otimização no software chamam funções externas específicas visando assim desacoplar técnicas de otimização e funções que a utilizam, permitindo o uso arbitrário e complexo de componentes. Dentre estas implementações de otimização pode-se citar algoritmos genéticos, hill climbing estocástico e “simulated annealing”.

Para acesso aos serviços do ADaM, o usuário pode utilizar os módulos via linha de comando, ou através de uma interface gráfica para o ADaM denominada “Plan Builder”, como também escrever scripts utilizando diferentes linguagens (Perl, Python, etc.).

#### **4.4.3 Aplicações**

Algumas aplicações desenvolvidas pelo Data Mining Center (Universidade do Alabama) [UAH2003b] utilizando o ADaM são voltadas à utilização de técnicas de mineração de dados para problemas das ciências da terra. Dentre os dos projetos estão [UAH2003c]:

- O AMSU (Advanced Microwave Sounding Unit) é um radiômetro de microondas utilizado para detectar temperaturas em diferentes níveis da atmosfera. Com base neste tipo de informação, é possível estimar velocidades de ventos radiais, as quais combinadas com outros fatores podem ser utilizadas para detectar ciclones tropicais. ADaM foi utilizado para implementar o algoritmo desta tarefa, sendo ainda gerada uma imagem a partir dos canais do AMSU pela própria aplicação (Figura 4.5).
- O objetivo dos procedimentos de detecção de luz é identificar pontos luminosos durante períodos noturnos, utilizando-se neste caso o sensor OLS (Operational Line Scan), o qual possui bandas visíveis e termais. Devido à frequência de varredura do sensor, a iluminação aparece como traços horizontais, de um ou dois pixels de comprimento. Entretanto, outros elementos possuem aparência semelhante, como o reflexo da luz da lua nas nuvens. A abordagem utilizada para este caso foi baseada em filtros morfológicos, detecção de gradiente e medidas de textura do ADaM,

através das bandas visíveis e térmicas do OLS. Um conjunto de imagens para treinamento e testes pré-classificadas por especialistas foi utilizado, além de um algoritmo genético do ADaM para ajustar os parâmetros do processo de classificação. Resultados do processo encontram-se na Figura 4.6.

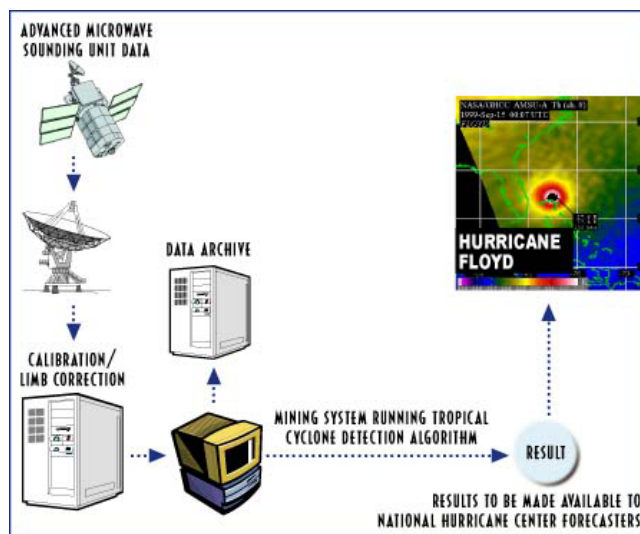


Figura 14: Detecção de ciclones tropicais utilizando ADaM e AMSU [UAH2003c]

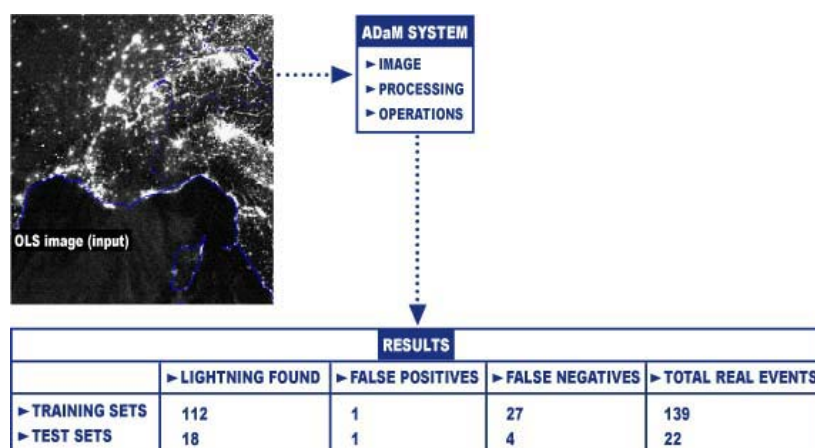


Figura 15: Detecção de pontos de luz utilizando ADaM e OLS [UAH2003c]

## CAPÍTULO 5

### Conclusões

A mineração de dados em bancos de imagens configura-se num processo altamente multidisciplinar e complexo, onde o domínio das metodologias, a competência dos especialistas e o poder das ferramentas muitas vezes não é suficiente para extrair boa parte do conhecimento escondido nas montanhas de acervos de imagens. Parte do problema reside na própria natureza do domínio, uma vez que imagens possuem alto grau de subjetividade, a qual deve ser somada (ou multiplicada) aos desafios da subjetividade da interpretação da informação, a qual visa a obtenção do conhecimento.

#### **5.1 Implicações do Atual Estágio de Mineração de Dados em Imagens**

O processo de descoberta de conhecimento a partir de acervos de imagens demanda um grande compromisso entre eficiência e robustez. Os grandes volumes dos acervos não permitem que infra-estruturas de hardware e software (inclusive aplicativos) desprezem o primeiro (eficiência). Já a relatividade de valores, a complexidade espacial e a multiplicidade de interpretações demandam robustez das implementações, exigindo ainda grande competência na execução de tarefas por parte de especialistas e analistas do domínio considerado.

Alguns elementos nesta área de pesquisa chamam atenção. Um deles é o fato de que a versão atual de um projeto de grande visibilidade (ADaM) utilize arquivos do sistema operacional para persistir imagens e outras informações estratégicas, mesmo havendo todo um direcionamento da comunidade de pesquisa para o armazenamento e recuperação de imagens utilizando SGBD's. Testes preliminares de uso do ADaM apontaram um bom desempenho de seus módulos na execução de diferentes tarefas, embora a compreensão e usabilidade de alguns componentes ainda deixem muito a desejar, especialmente diante da escassez de documentação dos programas (que resume-se a poucas linhas de help on-line, sem código fonte para análise).

## **5.2 Limitações do Processo**

O número de algoritmos e ferramentas disponíveis para mineração de imagens (em comparação com aqueles utilizados na mineração de dados alfanuméricos) ainda é muito limitado, o que impõe severas restrições técnicas às iniciativas nesta área. Outro gargalo tecnológico está relacionado ao investimento em infra-estrutura e material humano: a complexidade e o peso computacional demandam altos investimentos em hardware e software, além do apoio de especialistas das áreas envolvidas no processo multidisciplinar. Organizações como a NASA têm investido consideravelmente no desenvolvimento de técnicas específicas de mineração de imagens, na implementação de software especializado e na aquisição e manutenção de estruturas computacionais grid (poderosos clusters de hardware/software de alto desempenho).

Por outro lado, existem trabalhos e artigos que apresentam-se como pertencentes à área de mineração de imagens, mas na realidade são apenas aplicações de técnicas já conhecidas de processamento e análise de imagens, sem referenciar em momento algum o valor e a demanda dos grandes acervos, nem tampouco apresentando novas abordagens inerentes ao tema.

## **5.3 Perspectivas Tecnológicas**

O ajuste de parâmetros de algoritmos, tarefa considerada crítica no caso de dados convencionais, mostra-se ainda mais exigente no domínio de imagens, haja vista as particularidades e complexidades dos níveis de informação geralmente tratados (pixel, objeto, semântico, conhecimento). Novas abordagens e implementações nesta área certamente contribuirão no avanço tecnológico deste domínio.

Mesmo em se tratando de mineração em imagens, não é possível desprezar atributos alfanuméricos disponíveis nos acervos, pois muitos deles agregam valor à informação extraída das próprias imagens. Daí a demanda por melhores métodos de mineração em diferentes domínios (dados alfanuméricos, imagens, dentre outros).

Sem o acesso eficiente a coleções de imagens, não seria realístico considerar a extração de conhecimento do ponto de vista tecnológico. Sendo assim, o surgimento de novos mecanismos de recuperação de imagens mais eficientes, flexíveis e inteligentes (inclusive visão computacional) beneficiarão a área de mineração de dados espacial.

Verifica-se que ainda há um longo caminho a percorrer, pois até mesmo iniciativas de grandes instituições ainda estão "engatinhando" no grande desafio de descobrir informações estratégicas em grandes acervos de imagens, transformando-as em conhecimento útil e disponível para diferentes aplicações. Certamente o avanço em cada uma das áreas que suportam mineração de dados em bancos de imagens (banco de dados, inteligência artificial, processamento de imagens, estatística, dentre outros) redundará em avanços significativos neste contexto.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [ATT2002] AT&T Laboratories **Image Content Based Organisation and Navigation**. [online]. v. , n. , p. , 2002. <[www.uk.research.att.com/permm/icon.html](http://www.uk.research.att.com/permm/icon.html)>. Visitado em 18 agosto 2003.
- [BAL1982] Ballard, D.; Brown, C. **Computer Vision**. New Jersey: Prentice Hall, 1982.
- [BGS2000] Barclay, T.; Gray, J.; Slutz, D. **Microsoft TerraServer: A Spatial Data Warehouse**. ACM MOD 2002, p. 307-318, 2000.
- [BHS2002] Baumann, P.; Höfner, M.; Schatz, W. **Querying Large Geo Image Databases: A Case Study**. In: IV Brazilian Symposium on GeoInformatics. **Anais**. Caxambu: , 2002.
- [CCR2003] Canada Centre for Remote Sensing **Fundamentals of Remote Sensing**. [online], 2003. <[www.ccrs.nrcan.gc.ca/ccrs/learn/tutorials/fundam/fundam\\_e.html](http://www.ccrs.nrcan.gc.ca/ccrs/learn/tutorials/fundam/fundam_e.html)>. Visitado em 15 agosto 2003.
- [CLI2001] Clifton, C. **Change Detection in Overhead Imagery using Neural Networks**. International Journal of Applied Intelligence, 2003.
- [ESC2003] Escada, M. I. S. **Evolução de Padrões de Uso da Terra na Região Centro-Norte de Rondônia**. São José dos Campos. 155p. ().Tese (Doutorado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, 2003.
- [FAY1996] Fayyad, U. M.; Piatetsky-Shapiro, G.; Smith, P. **Advances in Knowledge Discovery and Data Mining**. Massachusetts: MIT Press, 1996. 560 p.
- [FON2000] Fonseca, L. M. G. **Processamento Digital de Imagens**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, Junho 2000.
- [GUT1984] Guttman, A. **R-Trees: A Dynamic Index Structure for Spatial Searching**. In: ACM SIGMOD. **Anais**. : , 1984. p. 47-57.
- [HAN2001] Han, J.; Kamber, M. **Data Mining: Concepts and Techniques**. Simon Fraser University: Morgan Kaufmann Publishers, 2001. 550 p.
- [LAM2003] Lambin, E. et al **Land-Use and Land-Cover Change Implementation Strategy**. [online]. n. 48, 1999. <<http://www.geo.ucl.ac.be/LUCC/lucc.html>>. Visitado em 07 agosto 2003.
- [OGC2003] Open GIS Consortium, Inc. **Open GIS Reference Model**. [online], 2003. <[www.opengis.org/info/orm/](http://www.opengis.org/info/orm/)>. Visitado em 20 agosto 2003.

- [PET1997] Petrakis, E.; Faloutsos, C. Similarity Searching in Medical Image Databases. **IEEE Transactions on Knowledge and Data Engineering**, v. 9, n. 3, p. 435-447, Maio 1997.
- [SDC2002] Simoff, S.; Djeraba, C.; Zaïane, O. **MDM/KDD2002: Multimedia Data Mining between Promises and Problems**. SIGKDD Explorations, v. 4, n. 2, p. 118-121, 2002.
- [TOW2001] Town, C.; Sinclair, D. **Ontological Query Language for Content Based Image Retrieval**. AT&T Laboratories, p. 75-80, 2001. .
- [UAH2003] University of Alabama in Huntsville **ADaM 4.0.0 Documentation (Draft)**. [online], 2003. <datamining.itsc.uah.edu/adam/docs/ADaM\_4\_Overview.doc>. Visitado em 21 agosto 2003.
- [UAH2003c] University of Alabama in Huntsville **Case Studies**. [online], 2003. <datamining.itsc.uah.edu/case\_studies/index.html>. Visitado em 20 agosto 2003.
- [UAH2003b] University of Alabama in Huntsville **Data Mining Center**. [online], 2003. <datamining.itsc.uah.edu/index.jsp>. Visitado em 22 agosto 2003.
- [UAH2003a] University of Alabama in Huntsville **The ADaM Data Mining Environment**. [online], 2003. <datamining.itsc.uah.edu/environment.html>. Visitado em 21 agosto 2003.
- [WIT1999] Witten, I.; Frank, E. **Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations**. Waikato: Morgan Kaufmann Publishers, 1999. 416 p.
- [ZHL2001] Zhang, J.; Hsu, W.; Lee, M. **An Information-Driven Framework for Image Mining**. In: DEXA. Anais. : National University of Singapore, 2001.p. 232-242.
- [ZHL2002] Zhang, J.; Hsu, W.; Lee, M. **Image Mining: Trends and Developments**. Kluwer Academic, 2002.