

Calendário / CA 05/2003 / Projeto Integrado de Pesquisa - AI

RIBC
Recuperação de Informação com Base no Conteúdo
Visual

Arnaldo de Albuquerque Araújo
Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Av. Antônio Carlos, 6627
31.270-010 Belo Horizonte - MG
Tel.: (31) 3499-5860/5873
Fax: (31) 3499-5858
E-mail: arnaldo@dcc.ufmg.br

Belo Horizonte XX de Fevereiro de 2003.

RESUMO	3
1 EQUIPE	4
2 INTRODUÇÃO	6
3 OBJETIVOS	14
4 NOSSA EXPERIÊNCIA NO ASSUNTO	15
5 RESULTADOS ESPERADOS E BENEFÍCIOS POTENCIAIS	18
6 METODOLOGIA	20
REFERÊNCIAS	27

RESUMO

A facilidade de captura e codificação de imagens digitais tem produzido uma quantidade gigantesca de informação visual on-line. Como consequência, grandes desafios se apresentam na área de sistemas de informação multimídia, envolvendo o armazenamento, a indexação e a recuperação de informação visual em grandes arquivos. O NPDI tem desenvolvido pesquisa envolvendo a recuperação de informação visual em bases de imagens controladas ; recuperação de imagens através da Web e detecção de eventos em seqüências de imagens digitais.

Por outro lado, o NPDI tem apoiado o CECOR/ EBA e o APM no desenvolvimento de sistemas de informação multimídia envolvendo, no caso do CECOR, uma base de informações sobre obras restauradas, seus materiais e técnicas, imagens geradas por métodos de análises, resultados de análises científicas, etc., e a base de informações/catalogação do acervo do pintor Guignard. Já no caso do APM, envolve o grande acervo de fotografias históricas do Estado de Minas Gerais, além de filmes, mapas, plantas, cartas, documentos, etc.

Este Projeto de Pesquisa solicita ao CNPq apoio para realização de pesquisa em sistemas de informação de multimídia e processamento digital de imagens, envolvendo o desenvolvimento de metodologias e soluções na área de recuperação de informação com base no conteúdo visual. Como plataforma, deveremos utilizar o software DB2 da IBM adquirido através do programa mundial IBM Scholars, também disponível para o Brasil. Espera-se, com este auxílio, desenvolver tecnologia para apoiar projetos regionais (CECOR, APM) que visam disponibilizar ao público um material multimídia de importância histórica, artística e cultural cujo acesso é hoje, por razões de conservação e segurança, restrito aos pesquisadores qualificados. Através da Internet, este patrimônio cultural estará acessível não apenas a nível nacional, como internacional, devido a sua importância para a América Latina e Península Ibérica.

Pretendemos, assim, estudar e analisar métodos representativos de extração de características invariantes, tais como: momentos estatísticos, transformadas de Fourier, transformadas de wavelets e redes de pulso acoplado paralelo e seqüencial; além de estudar e analisar estruturas de indexação, que são ferramentas fundamentais nos sistemas de gerenciamento de bancos de imagens envolvendo um grande volume de dados. Com a nossa experiência já adquirida no desenvolvimento de algoritmos e metodologias para a recuperação de informação visual em bases controladas, esperamos poder contribuir com novas metodologias para extração de características e estruturas de indexação visando sua aplicação em imagens fotográficas coletadas na Web e na grande base de fotografias históricas do APM.

Visamos, também, estudar a plataforma DB2 da IBM, para apoiar os projetos do CECOR e assim, contribuir para o desenvolvimento da informatização de dados científico-culturais em Minas Gerais. Além de promover intercâmbio de conhecimento entre os parceiros e formação de competência científico-tecnológica nas áreas de computação, informação multimídia e preservação do patrimônio cultural. A formação humana, a níveis de graduação, mestrado e doutorado, encontra-se entre as metas deste projeto.

1 Equipe

- Arnaldo de Albuquerque Araújo - UFMG/DCC, Professor Adjunto, Doutor em Eng. Elétrica pela UFPB, Pós-doutorado em Informática pela ESIEE Paris, Pós-Graduação em Processamento Digital de Imagens pela RWTH Aachen, Coordenador do Projeto, orientação de doutorandos, mestrandos e bolsistas IC e DTI.
- Luiz Souza – UFMG/EBA/CECOR, Professor Adjunto, Doutor em Química pela UFMG, Pós-Graduação em Restauração de obras de arte pelo Paul Getty Institute, Coordenador do Projeto RECICOR, co-orientação de mestrando, suporte em restauração de obras de arte, fornecimento de bases de imagens.
- Claudina Moresi - UFMG/EBA/CECOR, Professora Adjunto, Doutora em, Coordenadora do Projeto GUIGNARD, suporte em restauração de obras de arte, fornecimento de bases de imagens.
- Sylvie Philipp-Foliguet - ENSEA Cergy-Pontoise, França, Professeur des Universités, Doutora em , Cooperação internacional CAPES/COFECUB, suporte em segmentação de imagens, co-orientação de doutorando.
- Matthieu Cord - ENSEA Cergy-Pontoise, França, Maître de Conference, Doutor em , Cooperação internacional CAPES/COFECUB, suporte em recuperação de informação visual, co-orientação de doutorando.
- Eduardo Alves do Valle Jr. - doutorando em CC/ENSEA Cergy-Pontoise, Mestre em Ciência da Computação pela UFMG, responsável pelo desenvolvimento de técnicas de extração de características e indexação de imagens.
- Camillo Jorge Santos Oliveira - doutorando em CC/UFMG, Mestre em Ciência da Computação pela UFMG, responsável pelo desenvolvimento do protótipo de recuperação de imagens coletadas na *Web*, técnicas de extração de características e de indexação, co-orientação de alunos de graduação.
- Guillermo Cámara Chávez - doutorando em CC/UFMG, Mestre em Ciência da Computação pela USP-SC, responsável pelo desenvolvimento do protótipo de recuperação de imagens com base em atributos invariantes, técnicas de indexação, co-orientação de alunos de graduação.
- Flávio Humberto Cabral Nunes - mestrando em CC/UFMG, Bacharel em Ciência da Computação pela UFMG, responsável pelo desenvolvimento do sistema de informação multimídia do Projeto RECICOR.

- Clarissa Costa e Lima – Bolsista DTI FAPEMIG, Especialista em Engenharia de Software pela UFMG, responsável pelo desenvolvimento do sistema de informação multimídia do Projeto GUIGNARD, estudo da plataforma DB2 da IBM. co-orientação de aluno de graduação.
- Gláucia da Silva Tavares – Bolsista IC FAPEMIG, Graduanda em Ciência da Computação pela UFMG, co-responsável pelo desenvolvimento do sistema de informação multimídia do Projeto GUIGNARD, estudo da plataforma DB2 da IBM.
- 02 Bolsistas IC (a definir)

2 Introdução

Tradicionalmente, a informação visual tem sido armazenada de forma analógica e indexada manualmente usando esquemas proprietários. Atualmente, com o avanço da tecnologia de digitalização e codificação, sistemas de base de dados digitais já são utilizados para armazenar imagens, juntamente com seus metadados e taxonomias associados, muito embora o custo de grandes sistemas seja ainda elevado.

Metadados incluem informação bibliográfica, condições de captura ou geração de imagens, parâmetros de codificação, etc. Taxonomia é uma hierarquia de classes subjetivas (povo, natureza, notícias) usada para organizar assuntos de imagens em vários níveis, incluindo classes semânticas (humor, política) e classes visuais (povo, paisagens). A seleção apropriada de metadados e taxonomias, que devem incorporar características especiais do domínio da aplicação, é, geralmente, o primeiro passo para por em funcionamento uma grande base de imagens. Obviamente, existem sérias limitações no uso destes indexadores, uma vez que requerem anotação manual (dificultando seu uso em grandes arquivos) e que sofrem influência, tanto do domínio de aplicação quanto do conhecimento da pessoa realizadora da tarefa. Enfim, estes indexadores estão sempre limitados na sua viabilidade de capturar todo o conteúdo de uma imagem [Blatter92], [Catell94], [Khoshafian96], [Baeza-Yates99], [Del Bimbo99].

A análise do conteúdo multimídia e indexação baseadas no conteúdo visual sinalizam, juntas, uma direção promissora para complementar a metodologia mencionada anteriormente. Vários sistemas de consulta de imagens/vídeos baseados no conteúdo visual têm sido apresentados na literatura. Tem havido progresso substancial no desenvolvimento de ferramentas que permitem aos usuários especificar consultas de imagens/vídeos através do uso de esboços, seleção de características visuais (cor, textura, forma e movimento), passagem de exemplos, e características espaciais/temporais.

A Figura 1 representa exemplos de imagens que apresentam as características visuais como cor, textura e forma. Nela, o lagarto apresenta forma, textura e cores, enquanto na imagem da onça encontram-se texturas e cores na vegetação, na pele e no solo, além da forma da onça. Quando as seqüências de imagens são consideradas, o problema de indexação torna-se muito mais complicado, pois envolve a identificação e o entendimento de uma cena ou tomada para proporcionar uma indexação destes elementos de forma eficiente. Devido ao formato não estruturado e ao comprimento do vídeo, entenda-se somente seqüência de imagens, esta indexação não é uma tarefa das mais fáceis.



Figura 1: Exemplo de imagens que apresentam as características visuais: cor, textura e forma.

Neste sentido, métodos para indexar e detectar eventos em vídeo, de forma automática, são extremamente interessantes na sociedade moderna, que visa a rapidez e a precisão das respostas às perguntas (consultas) feitas. Um exemplo para o uso destes indexadores está associado à navegação dentro do vídeo, onde é vantajoso ter sistemas capazes de segmentar o vídeo em unidades significativas sem um conhecimento específico *a priori* da natureza do programa ao qual o vídeo está associado [Roth99].

A *World Wide Web* apresenta-se como um enorme manancial de informação e, diariamente, milhares de documentos são adicionados, substituídos ou removidos. Entre estes documentos, encontra-se uma grande coleção de imagens e vídeos digitais, os quais são gerados em grande velocidade a cada dia, através de várias fontes, tais como, sistemas de satélites, experimentos científicos, imagens biomédicas, sistemas de entretenimento domésticos e outros. É evidente a necessidade de indexar, detectar e processar toda esta informação, de tal maneira que se consiga sua consulta e recuperação eficientes. Com relação à consulta baseada em texto, existem sistemas que recuperam a informação de forma eficiente, como o *Lycos*, *Alta Vista* e *Yahoo*. Entretanto, os sistemas de recuperação de imagens na *Web* não acompanharam a velocidade de crescimento das coleções geradas. Não existem (ainda) máquinas de busca de imagens na *Web* que trabalhem em grande escala. Mesmo com a existência de bons trabalhos envolvendo estudos sistemáticos, ainda é necessária uma melhor exploração desta área para que os dispositivos de pesquisa de imagens se tornem compatíveis àqueles baseados em texto [Smith96], [Sclaroff97], [Flickner95], [Frake196].

O problema da recuperação de imagens é um caso particular da área de manipulação e recuperação de informação. De acordo com o paradigma clássico da recuperação automática de informação, o arquivo está organizado na forma de documentos que podem ser recuperados

através de índices sintéticos, por sua vez organizados com uma estrutura de dados que permitam consulta e recuperação rápida. O usuário formula seu problema de recuperação como uma expressão de uma linguagem de consulta. A consulta é traduzida numa linguagem de índices, o índice resultante é “casado” contra aqueles da base de dados, sendo então recuperados os documentos que contém os índices “casados” [Doermann98].

Um documento é a quantidade mínima de informação que pode ser recuperada de um arquivo, sendo identificado por um modelo que define a composição do documento e por uma instância, que é o dado útil a ser arquivado. Partes de instâncias diferentes podem representar aspectos diferentes, que podem ser especificadas e relacionadas umas às outras de acordo com o modelo do documento.

O principal objetivo de um sistema de gerenciamento de informação é poder, facilmente e eficientemente, separar documentos de um arquivo que sejam relevantes para uma determinada necessidade do usuário. Isto requer um esquema de indexação com significado. Índices são representações sintéticas dos documentos: a consulta é formulada e realizada referindo-se aos índices para se obter uma recuperação eficiente.

O problema da indexação pode ser resumido como a atribuição de “palavras-chave”, obtidas de uma “linguagem” descritiva, às entidades dos documentos (palavras, objetos, etc.) para facilitar sua separação. O principal requisito é que a associação linguagem/palavras-chave/documentos tenha a capacidade de discriminação suficiente para eliminar, do espaço de consulta, os dados imprestáveis, sem, no entanto perder informação interessante. Enfim, os índices devem ser significativos, discriminantes e utilizáveis, uma vez que devem estar relacionados, com a maneira na qual o usuário faz sua consulta no ambiente real.

Quando se trata de “documentos” como as imagens ou vídeos que estão associados a diferentes tipos de conteúdo visual, a recuperação de informação visual é ainda mais complexa. Neste caso, os índices podem ser classificados, com respeito à relação que eles têm com a imagem ou vídeo da seguinte maneira [Del Bimbo99]:

- metadados independentes do conteúdo - são dados que não concernem diretamente ao conteúdo da imagem ou vídeo, mas que são, de alguma maneira, relacionados com este, como o formato da imagem, autoria, data, local, condições de iluminação, etc.;
- metadados dependentes do conteúdo - são dados que se referem às características consideradas de nível baixo e médio, como cor, textura, forma, esboço, relação espacial,

movimento e combinações destes. Para alguns tipos de imagens, como as provenientes de satélites, da biomedicina, tomografia computadorizada, etc., é possível descrever o conteúdo destas em termos da geometria intrínseca e de configurações topológicas; metadados descritivos do conteúdo - são dados que se referem ao conteúdo semântico e que concernem às relações das entidades da imagem com entidades do mundo real ou eventos temporais, emoções e significados associados a sinais visuais e cenas.

O conteúdo visual de imagens pode ainda ser classificado em dois tipos principais:

- conteúdo primitivo - que se refere aos elementos básicos que compõem a imagem, assim como às características das imagens que podem ser reconhecidas e extraídas automaticamente pelo computador (análise de imagens, reconhecimento de padrões, visão computacional). Conteúdos primitivos são, em geral, de natureza quantitativa;
- conteúdo complexo - que se refere aos padrões de uma imagem que são percebidos com significado por humanos. Em geral, eles não podem ser identificados automaticamente por computador e são de natureza qualitativa.

A maior vantagem associada com a indexação de conteúdo primitivo é que sua extração pode ser automática. Entretanto, este conteúdo pode não ser suficientemente rico para uma grande variedade de aplicações, uma vez que tipos de objetos e características significativas que podem ser reconhecidos pela máquina, são ainda limitados. Por outro lado, o conteúdo complexo da imagem é semanticamente rico, mas sua extração e indexação são custosas, uma vez que um envolvimento manual considerável é geralmente necessário.

Um problema-chave da indexação de imagens consiste em conseguir estabelecer relações entre os dois tipos de conteúdos de maneira que as desvantagens de cada um possam ser reduzidas ou eliminadas. Assim, um desafio crucial para a indexação e recuperação de imagens pelo conteúdo está no desenvolvimento de mecanismos para a automação da indexação, que começaria com a extração automática de conteúdo primitivo e subseqüentemente faria uso de possíveis regras de conhecimento junto com informação contextual relevante, permitindo uma identificação automática ou inferência de conteúdo complexo. Com tal mecanismo, uma estrutura poderosa de indexação poderia ser construída para possibilitar a recuperação baseada em conceitos semânticos ricos. Trata-se de uma tecnologia que espelha o sistema visual humano.

O problema da indexação está fortemente ligado ao problema da formulação da consulta. Para recuperar imagens estáticas usando metadados dependentes do conteúdo (cor, textura,

forma e esboço), o paradigma básico da recuperação requer que, para cada imagem, seja pré-calculado um conjunto de características distinguíveis.

As consultas são expressas através de exemplos visuais. Para começar a consulta, o usuário seleciona as características (e suas faixas de validade) que são importantes e escolhe uma medida de similaridade. Exemplos tanto podem ser preparados pelo usuário (com ajuda de uma interface homem-máquina) ou extraídos de imagens (amostras). O sistema verifica a “similaridade” entre o conteúdo da imagem usada na consulta e das imagens da base de dados. Como nem sempre os resultados obtidos em resposta a uma consulta são plenamente satisfatórios, em geral, procura-se melhorar este resultado através de uma metodologia onde se tenta manter o número de perdas o mais baixo possível (às custas de um número mais alto de falsas respostas) e permitir uma forma de interação chamada “realimentação por relevância”. Interfaces visuais são usadas para facilitar este tipo de consulta.

A recuperação baseada em similaridade difere da operação de “casamento” (*matching*), que é definida em visão computacional como sendo uma operação para reconhecimento de objetos na qual deve-se decidir se o objeto observado corresponde ou não ao modelo. Assim, a recuperação baseada em similaridade é uma operação de re-ordenamento das imagens da base de dados, de acordo com sua similaridade medida, em relação à imagem-exemplo da consulta. Desta maneira, é uma operação que concerne mais a ordenamento, do que classificação, sem ao menos postular a existência de uma imagem alvo. O re-ordenamento das imagens da base de dados é realizado, mesmo que não haja imagens similares ao exemplo. Neste caso, o usuário faz parte do sistema de recuperação, devendo analisar as respostas apresentadas, e, se for o caso, deve refinar a consulta. Tudo isto realça a importância da existência de interfaces flexíveis e ferramentas de visualização [Del Bimbo99], [Abate99], [Berman99], [Jin96], [Chen99].

Apresentamos a seguir as principais funcionalidades de um sistema de recuperação de imagens com base no conteúdo (RIBC). Na Figura 2, essas funcionalidades estão representadas em módulos, quais sejam, módulo gerenciador de imagens, módulo indexador de características, módulo de recuperação e módulo de interface com o usuário. Observa-se na Figura 2 duas bases de dados que também compõem um sistema típico de recuperação de imagens com base no conteúdo.

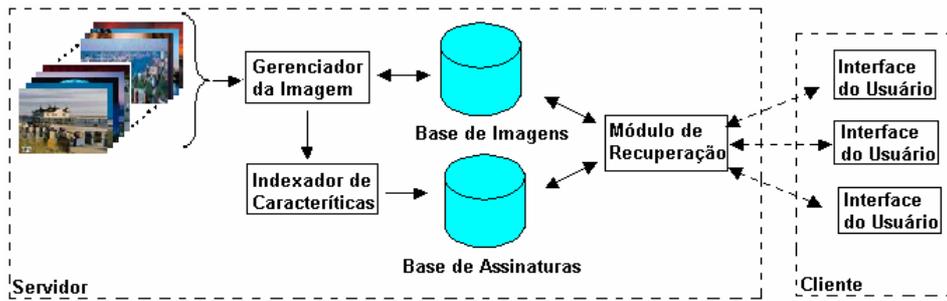


Figura 2. Uma arquitetura típica de um sistema de recuperação de imagens com base no conteúdo.

O módulo gerenciador de imagens é responsável pelo gerenciamento do acesso ao arquivo da imagem, conversão do formato da imagem e armazenamento das informações na base de imagens. O módulo indexador disponibiliza a assinatura da imagem. A assinatura de uma imagem é uma representação da imagem na "visão" de um sistema de recuperação de imagens com base no conteúdo. Os componentes da assinatura são chamados de características. A Figura 3 mostra a relação entre uma imagem, sua assinatura e características.

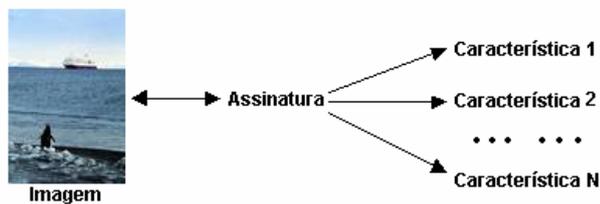


Figura 3. A assinatura de uma imagem é um conjunto de características.

Desta forma, o módulo indexador de características é formado de processos que permitem obter a assinatura da imagem. Processos fazem a extração de características da imagem e disponibilizam estas informações (valores) em um vetor de características que será armazenado na base de assinaturas. O módulo indexador de características é utilizado em dois momentos diferentes no sistema. Inicialmente, o módulo indexador é utilizado como um processo *off-line* que atribui assinaturas para as imagens da base de dados (Figura 4). Em um segundo momento, o módulo indexador será utilizado como um processo *on-line* que atribui assinaturas para a imagem fornecida pelo usuário como base de consulta (Figura 5).

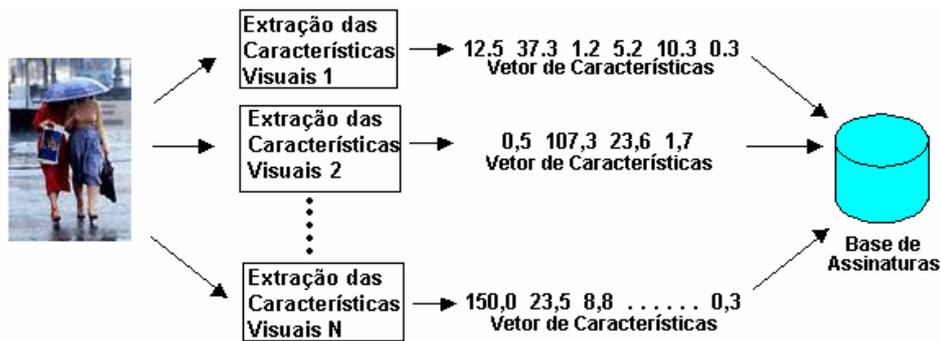


Figura 4. Processo típico de indexação de um sistema de recuperação de imagens com base no conteúdo.

O módulo de recuperação é um servidor de consultas do usuário *on-line* baseado no conteúdo da imagem. O módulo de recuperação recebe do módulo de interface do usuário uma imagem para consulta. Então, o módulo de recuperação faz uso do módulo de indexação para extrair as características dessa imagem fornecida como consulta. A partir da obtenção do vetor de características desta imagem, o módulo de recuperação utiliza seu módulo comparador para procurar na base de assinaturas, imagens que correspondam ao vetor de características da imagem consulta. Depois de obtidas, tais imagens, utiliza-se o módulo de classificação para ordenar os resultados em função do resultado das métricas. Por fim, o módulo de recuperação envia os resultados ao módulo de interface do usuário. A Figura 5 apresenta o processo de recuperação da imagem.

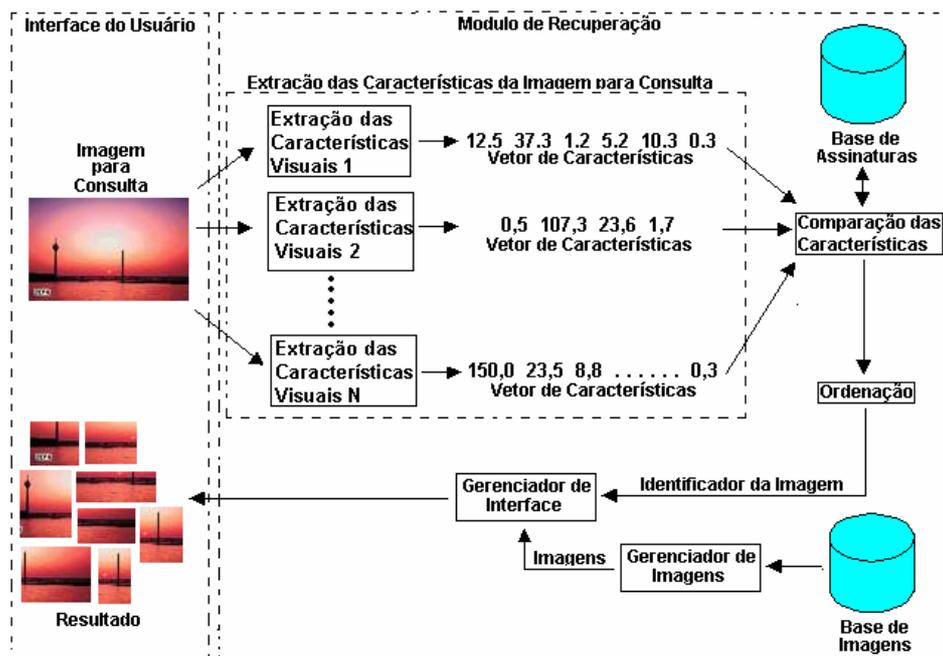


Figura 5. Um típico sistema de recuperação de imagens com base no conteúdo.

O módulo de interface com o usuário é formado de programas clientes que permitem a formulação de consultas baseadas no conteúdo. Aqui, estas consultas podem ser realizadas de duas maneiras diferentes:

- O usuário fornece uma imagem para a consulta.
- O usuário constrói uma imagem através do preenchimento de cores.

O módulo de interface do usuário também permite a visualização dos resultados destas consultas.

3 Objetivos

Estender e consolidar pesquisa em desenvolvimento no NPDI é o objetivo principal deste projeto. Interagindo com outras instituições locais, contando com o apoio de uma Cooperação Internacional CAPES/COFECUB, esperamos desenvolver algoritmos e metodologias para solucionar problemas nas áreas de sistemas de informação multimídia e processamento digital de imagens, envolvendo a recuperação de informação com base no conteúdo visual. Como plataforma deveremos utilizar o software DB2 da IBM adquirido gratuitamente através do programa mundial IBM Scholars, também disponível para o Brasil. Estarão envolvidos no projeto, nas duas linhas de pesquisa, três doutorandos e um mestrando em Ciência da Computação. Alunos de graduação também participarão através de projetos orientados e bolsas IC.

3.1 Objetivos Específicos

- Desenvolver um protótipo de um sistema para a recuperação de imagens com base em características de cor, que torne possível buscar imagens específicas em um banco de dados de imagens coletadas na *Web*.
- Estudar e analisar os métodos representativos de extração de características invariantes, tais como: momentos estatísticos, transformadas de Fourier, transformadas de *wavelets* e redes de pulso acoplado paralelo e seqüencial.
- Estudar e analisar estruturas de indexação, que são ferramentas fundamentais nos sistemas de gerenciamento de bancos de imagens envolvendo um grande volume de dados.
- Estudar a plataforma DB2 da IBM, visando seu uso para apoiar projetos do CECOR: RECICOR, coordenado pelo prof. Luiz Souza e GUIGNARD, coordenado pela profa. Claudina Moresi.
- Contribuir para o desenvolvimento da informatização de dados científico-culturais em Minas Gerais.
- Promover intercâmbio de conhecimento entre os parceiros e formação de competência científico-tecnológica nas áreas de computação, informação multimídia e preservação do patrimônio cultural.

4 Nossa Experiência no Assunto

Em seus doze anos de existência, o NPDI sempre procurou interagir com outras instituições, desenvolvendo pesquisa aplicada e buscando soluções e metodologias para os problemas de seus parceiros, entre eles, podemos citar:

- Centro de Restauração de Obras de Arte - CECOR, Escola de Belas Artes/UFMG
- Departamento de Anatomia Patológica - Escola de Medicina/UFMG
- Departamento de Engenharia Mecânica - Escola de Engenharia/UFMG
- Departamento de Farmacologia - Instituto de Ciências Biológicas/UFMG
- Instituto de Geociências - UFMG
- Centro de Desenvolvimento da Tecnologia Nuclear - CDTN/CNEN
- Fundação Centro de Tecnologia - CETEC/MG
- Fundação Ezequiel Dias - FUNED/MG
- Arquivo Público Mineiro - APM/MG
- Cia. de Processamento de Dados do Estado de Minas Gerais - PRODEMGE
- Instituto de Criminalística - SSP/MG

Através do Acordo de Cooperação Internacional CAPES/COFECUB no. 204/97/99, com duração de jan. 1997 a fev. 2001, conseguimos oficializar interação antiga com a École Supérieure d'Ingénieurs en Électrotechnique et Électronique - ESIEE Paris, França, iniciada em 1990. Este convênio foi renovado, tendo recebido o no. 396/02 e validade entre jan. 2002 a fev. 2004. Temos, desta maneira, seguido a orientação de internacionalização da UFMG. A sistemática destes convênios tem-nos proporcionado quatro missões de trabalho por ano, duas no sentido BR/FR e duas no sentido contrário. Além de nos permitir uma constante atualização tecnológica, esta cooperação com a França tem facilitado nosso contato com as seguintes instituições francesas:

- École Nationale Supérieure d'Électronique et ses Applications - ENSEA, Équipe Traitement d'Images et du Signal - ETIS, Cergy Pontoise. Contatos: Prof. Jean-Pierre Cocquerez e Profa. Sylvie Philipp.
- École Nationale Supérieure des Télécommunications - ENST, Département Images, Paris. Contatos: Prof. Henri Maître e Profa. Isabelle Bloch.

- École Nationale Supérieure des Mines de Paris - ENSMP, Centre de Morphologie Mathématique, Fontainebleau. Contatos: Prof. Jean Serra e Prof. Dominique Jeulin.

Por meio deste intercâmbio internacional, temos trazido convidados a palestras para o Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens - SIBIGRAPI:

- SIBIGRAPI'92 - Prof. Gilles Bertrand - ESIEE
- SIBIGRAPI'96 - Prof. Mohamed Akil - ESIEE
- SIBIGRAPI'97 - Profa. Isabelle Bloch - ENST
- SIBIGRAPI'99 - Prof. Jean-Pierre Cocquerez e Profa. Sylvie Philipp – ENSEA
- SIBIGRAPI'01 - Prof. Michel Couprie – ESIEE

A formação de alunos de Pós-Graduação tem sido uma preocupação constante de nossas atividades. Foi assim que, neste período, conseguimos formar cinco doutores e 19 mestres na área de PDI, além de orientação de cerca de 50 alunos de graduação (Projetos Orientados em Ciência da Computação e Iniciação Científica). Atualmente, nossa equipe envolve três doutorandos e três mestrandos. Nossos resultados foram publicados em cerca de 80 artigos em congressos e periódicos, nacionais e internacionais. Também, fazemos parte da equipe do Projeto SIAM - PRONEX do DCC/UFMG.

Nossas primeiras incursões na área de bases de imagens remontam aos anos 90 e foram desenvolvidas juntamente com o prof. Luiz Souza do CECOR/UFMG. Envolveram a análise de imagens de raios-x de estátuas de madeira [Araújo92a,b], desenvolvimento de protótipo para base de dados para documentos históricos [Araújo93] e a análise de imagens de fotografias, em infravermelho, de pinturas [Araújo96a,b], [Araújo97], [Araújo98]. Atualmente apoiamos dois projetos do CECOR: RECICOR, coordenado pelo prof. Luiz Souza [Nunes02a,b], e GUIGNARD, coordenado pela profa. Claudina Moresi [Moresi02a,b].

Com o Arquivo Público Mineiro (APM), desenvolvemos um protótipo de sistema de informação multimídia para a coleção de documentos do ex-presidente Arthur Bernardes [De Andrade98a,b,c] [De Andrade00]. Este protótipo motivou a direção do APM a levantar recursos externos no sentido de desenvolver novo sistema piloto, visando um programa mais amplo, que tem como objetivo auxiliar, com a utilização de novas tecnologias, as entidades culturais que, no cumprimento de suas funções, guardam ou desenvolvem trabalhos com acervos documentais de valor permanente (...) devendo impulsionar a utilização de novas tecnologias no tratamento de acervos significativos para a memória de Minas Gerais e do País, com ênfase na recuperação do conteúdo documental [Valle02a,b], [Valle03].

Com base na experiência adquirida com o protótipo do sistema de informação da coleção de documentos do ex-presidente Arthur Bernardes, desenvolvemos, também, um protótipo para o Instituto de Criminalística da Secretaria de Segurança Pública de Minas Gerais, visando o acesso de laudos importantes que possam ser considerados como modelos para a elaboração de outros laudos, principalmente, por peritos localizados no interior do estado. Foram realizados testes para verificação da melhor maneira de digitalização dos diversos tipos de dados existentes nos laudos (texto, fotografias, vídeo, resultados de testes laboratoriais, etc.). Optou-se pelo uso de um sistema de gerenciamento de bancos de dados multimídia do tipo relacional/orientado para objetos [Mitre01].

Por fim, em nossa linha de pesquisa em recuperação de informação visual com base no conteúdo, conseguimos realizar um grande levantamento bibliográfico, o que nos capacitou a preparar um curso tutorial [Araújo00] para o SIBGRAPI 2000, assim como tem servido para a apresentação de diversas palestras convidadas. Estes estudos realizados nos levaram a iniciar pesquisa em três aplicações diferentes: recuperação de informação visual em bases de imagens controladas [Rodrigues02], [Rodrigues03a,b]; recuperação de imagens através da Web [Oliveira01], [Oliveira02a,b] e detecção de eventos em seqüências de imagens digitais [Guimarães01], [Guimarães02a,b], [Guimarães03a,b].

5 Resultados Esperados e Benefícios Potenciais

A preservação da memória é uma atividade essencial em um mundo que está a todo o momento buscando no passado material para a análise da história do ser humano e para a construção da identidade cultural dos povos. Entretanto, essa tarefa não se faz sem percalços: além da dificuldade em classificar e armazenar enormes massas documentais de forma sistemática, a fragilidade dos artefatos impõe um severo compromisso entre conservação e acesso. Itens valiosos são guardados em arquivos seguros, disponíveis apenas para uns poucos pesquisadores. Isso, sem dúvida, é frustrante, pois quando os documentos são protegidos, mas não estão ao alcance do público, a tarefa de manter a memória viva não está sendo cumprida adequadamente. A tecnologia digital surge como uma possibilidade de romper esse compromisso, permitindo dar amplo acesso a cópias digitais de alta qualidade de determinados documentos, ao mesmo tempo em que resguarda os originais da manipulação desnecessária. Mais ainda, sistemas computadorizados têm um potencial enorme em facilitar a própria tarefa de organização e descrição dos acervos.

Este projeto visa a capacitação do NPDI/DCC/UFMG na área de sistemas de informação multimídia e recuperação de informação com base no conteúdo visual. Como plataforma deveremos utilizar o software DB2 da IBM adquirido, sem custo, através do programa mundial IBM *Scholars*.

O NPDI interage com o CECOR/UFMG na área de documentação dos processos de restauração de obras de arte lá realizados [Nunes02a,b], [Moresi02a,b]. Apoiamos o APM (Convênio de Cooperação Técnico-científico entre a UFMG e o APM, no. 1450/0/1998, nov. 1998 a out. 2004) nos seus projetos de digitalização, armazenamento e disponibilização da informação de seu rico acervo cultural [Valle02a,b], [Valle03]. Somos membro do Comitê Executivo do Programa: O Uso da Tecnologia Digital no Resgate da Identidade Histórico-Cultural de Minas Gerais, Secretaria de Estado de Ciência e Tecnologia, desde maio 2000 e membro da Comissão Permanente de Acervos da UFMG, Portaria no. 02011, 13/08/01. Fomos membro da Comissão de Avaliação do Edital FAPEMIG 006/2001: O Uso da Tecnologia Digital no Resgate da Identidade Histórico-Cultural de Minas Gerais, Portaria no. 13/2001, 01/06/01 e do Edital 008/2002: Tecnologia Digital, Documento e Memória, Portaria no. 34/2002, 01/11/02.

Acreditamos que a capacitação adquirida com este projeto será de grande utilidade para incrementar nosso apoio aos projetos desenvolvidos no CECOR/UFMG, no APM e que poderá ser utilizada em outros projetos do Estado de Minas Gerais, que apresenta uma posição de

destaque no país nesta área de conservação e disponibilização do patrimônio histórico. Através da divulgação de nossos resultados em congressos nacionais estaremos difundindo este conhecimento a nível nacional.

Além desta capacitação tecnológica, estaremos apoiando a formação de três doutorandos e um mestrando, além dos alunos de graduação que estarão envolvidos em projetos orientados e bolsas de iniciação científica. Em termos de publicações, esperamos publicar pelo menos quatro artigos em congressos (no Brasil e/ou exterior) e ao final da pesquisa, esperamos poder submeter dois artigos em periódicos indexados.

6 Metodologia

Estaremos, neste projeto, atacando as seguintes frentes de pesquisa envolvendo sistemas de informação multimídia e recuperação de informação visual:

- Classificação de imagens coletadas na *Web* com base no conteúdo cor
- Estudo e implementação de métodos usando características invariantes obtidas através de transformadas
- Estudo e implementação de métodos para classificação de cenas e sua aplicação no grande acervo fotográfico histórico do Arquivo Público Mineiro
- Estudo da plataforma DB2 da IBM adquirida, sem custo, através do programa mundial IBM *Scholars*, visando seu uso nas aplicações do CECOR/UFMG
- Desenvolvimento de uma interface para um sistema de recuperação de informação visual

6.1 Classificação de imagens coletadas na *Web* com base no conteúdo cor

Estima-se que existam mais de dez milhões de imagens na *Web* sem falar nas imagens que são trocadas diariamente. A *Web* é grande, distribuída, hipermídia e um sistema de informação não estruturado. O desenvolvimento de ferramentas que tornem possível buscar imagens específicas em um banco de dados tão grande é de uma utilidade inquestionável [Abb99], [Smith95].

Visamos o desenvolvimento do protótipo de um sistema para a recuperação de imagens com base na cor. O sistema analisará as regiões cromáticas de uma imagem, suas adjacências e variações espaciais (modelo estatístico para obter a distribuição da cor no domínio espacial, [Park99]). O processamento da base de imagens para obter estas informações será feito *off-line* e as imagens serão indexadas através de uma estrutura capaz de manipular estruturas multidimensionais, neste caso uma *R*-tree*, [Brown98].

A primeira etapa do trabalho consiste da coleta das imagens na *WWW*. Esta etapa contempla a utilização de um robô (*software*) que busca as imagens na *WWW* armazenando-as no diretório local onde o robô foi executado. A segunda etapa consiste na classificação das imagens, ou seja, as imagens coletadas são separadas em imagens fotográficas e imagens gráficas [Oliveira01]. A próxima etapa consiste da implementação do módulo de extração de características. Como dito anteriormente, este módulo obtém a assinatura das imagens e as

armazena em um banco de assinaturas. A construção do módulo de extração de características prevê a implementação de diversos métodos, entre eles:

- Filtro majoritário, [Gu95]: aplicado na imagem para remover os ruídos que existem próximos das adjacências das regiões cromáticas, tornando as regiões uniformes e as transições entre as regiões mais aguçadas (Figura 6). Isto proporcionará a geração de um grafo mais preciso.



Figura 6. Exemplo da aplicação do filtro majoritário, [Gu95].

- *MCAG (Modified Color Adjacency Graph)*, [Park99]: será utilizado para descrever as regiões cromáticas da imagem e o tamanho das adjacências entre as mesmas. Cada nodo do grafo representa uma cor no modelo RGB da imagem, contendo a quantidade de pixel da sua cor e cada aresta entre os nodos a e b representará a adjacência espacial de a e b , que significará, o número de pixels na cor a que são vizinhas dos pixels na cor b .
- *SVG (Spatial Variance Graph)*, [Park99]: será útil para descrever a distribuição de cor no domínio espacial da imagem. As informações necessária para elaborar este grafo serão: as cores da imagem, a contagem de cor da imagem e coordenadas de cada pixel, para cada cor. Usando equações estatísticas com probabilidade e técnicas de variância, obtém-se a própria variância espacial da cor e a variância espacial relativa de duas cores. Cada nodo contém a própria variância das cores denotadas e cada borda tem como seus atributos a variância espacial entre as cores denotadas pelos nodos ligados.

A implementação do módulo de indexação exige a geração dos índices e a implementação de uma R*-tree [Brown98].

A quarta etapa do trabalho contempla o desenvolvimento do módulo de interface que permite ao usuário realizar consultas ao protótipo. Estão previstos dois tipos de consultas: (a) dada uma imagem para consulta, procurá-la na base de imagens e (b) o usuário poderá compor

as cores da imagem que será submetida para consulta e esta será procurada na base de imagens.

A quinta etapa do trabalho implementa o módulo de recuperação da imagem. Este módulo recebe a consulta do usuário fornecida pelo módulo de interface e submete esta imagem ao módulo de extração de características. Após a obtenção da assinatura da imagem consulta, este módulo implementa duas funcionalidades: comparação das imagens e classificação dos resultados. Para a fase de comparação das imagens é implementada a métrica de similaridade que compara os valores das arestas e os nodos entre os grafos *MCAG* e *SVG* de duas imagens. Pode-se utilizar pesos para cada comparação de grafo, mas são usualmente escolhidos como 1. O valor retornado pela métrica é um número real entre 0 e 1. Pode-se representar grafos *MCAG* e *SVG* como métricas, onde cada elemento ii representa um nodo e cada elemento ij ($i \neq j$) representa uma aresta. Outra matriz S (matriz de seleção), contendo 0's e 1's pode ajudar os grafos *MCAG* e *SVG* das imagens pela eliminação de nodos não importantes (cores) pela consulta. Na fase de classificação, as imagens resultantes do processo de comparação são ordenadas em função dos valores das métricas.

A última etapa do trabalho contempla a finalização do módulo de interface do usuário implementando a fase de apresentação dos resultados obtidos.

6.2 Estudo e implementação de métodos usando características invariantes

Ao gerenciar e manipular informações complexas e volumosas, um dos problemas a serem tratados é o de como armazenar e recuperar tais informações de maneira rápida e precisa. Uma técnica usual é a extração de características (parâmetros) da imagem que idealmente possibilitem a identificação da mesma de forma inequívoca, ou com a menor ambigüidade possível. Uma comparação dos diferentes métodos existentes sob transformações geométricas (rotação, escala e translação), variância de iluminação, tolerância ao ruído mostraram um desempenho com certas limitações [Schmid97]. Muitas vezes, o processo de extração de características é baseado na redução da dimensionalidade dos dados: parâmetros são extraídos das imagens e a seguir utilizados para indexá-las agilizando sua busca e recuperação.

As estruturas de indexação são ferramentas fundamentais nos sistemas de gerenciamento de banco de dados de interesse dentre um grande volume de dados. As primeiras estruturas de indexação desenvolvidas, como as estruturas ISAM (*Indexed Sequential Access Method*), são voltadas à indexação de números e pequenas cadeias de caracteres, sendo que os gerenciadores de dados comerciais atuais as utilizam maciçamente. Essas estruturas não são

úteis para dados complexos, como imagens, sons, vídeo, séries temporais ou estruturas moleculares, que são tipicamente conjuntos de dados de altas dimensões ou mesmo adimensionais. Para o tratamento de dados espaciais com poucas dimensões, já existem diversas estruturas de indexação adequadas, genericamente associadas a Métodos de Acesso Espaciais - SAM (*Spatial Access Methods*), cuja pesquisa iniciou-se com o trabalho pioneiro sobre as *R-Trees* [Guttman84] e prosseguiu com numerosas variações e adaptações (veja [Gaede98] para um apanhado geral sobre estruturas espaciais). No entanto, para dados em espaços métricos adimensionais, nenhuma dessas estruturas são aplicáveis.

Estruturas de dados para espaços métricos (que englobam tanto dados espaciais com dimensão definida quanto dados adimensionais) que endereçam essa necessidade estão começando a ser estudadas, embora existam ainda poucos trabalhos na área, todas apresentando resultados ainda preliminares. Exemplos de dados métricos são imagens como impressões digitais, ou palavras em um dicionário, com funções específicas de dissimilaridade associadas a elas. As principais estruturas métricas existentes são: a *fixed query tree* [Baeza-Yates94], a *mvp-trees (multi-vantage point tree)* [Bozcaya97], *vp-tree (vantage-point tree)* [Chiueh-94], a GNAT [Brin95], a *M-Tree* [Ciaccia97], e a *Slim tree* [Traina99]. Dessas estruturas, apenas a *M-tree* e a *Slim-tree* são dinâmicas, ou seja, permitem a inclusão e remoção de novos objetos depois da árvore já estar criada - as demais estruturas são criadas em uma única operação sobre o conjunto de dados, imutável depois da criação da árvore. Além disso, ambas *M-tree* e *Slim-tree* são escaláveis, e mantêm eficiência quando armazenadas em memória secundária. Um outro trabalho relacionado ao mapeamento de dados multidimensionais é [Faloutsos95].

Características das imagens serão obtidas através da extração de dados que descrevam adequadamente as imagens, bem como características invariantes às transformações geométricas, e através de tais características as imagens serão indexadas e recuperadas utilizando a *Slim-tree*.

O objetivo deste trabalho é desenvolver um conjunto de técnicas para a extração de características de imagens. Através dessas características extraídas, as imagens poderão ser organizadas em um sistema, possibilitando sua recuperação por conteúdo. Estudar e analisar os métodos representativos de extração de características invariantes, tais como: momentos estatísticos, transformadas de Fourier, transformadas de *wavelets* e redes de pulso acoplado paralelo e seqüencial. O módulo de extração de características desempenhará um papel muito importante dentro do sistema geral de recuperação de imagens proposto.

Inicialmente, os dados de entrada são manipulados por uma variedade de métodos que fazem operações, tais como remoção de ruído, segmentação, extração de bordas e melhoramento da qualidade dos mesmos. O problema da indexação pode ser resumido como a atribuição de palavras-chave às entidades dos documentos (palavras, objetos, etc.) para facilitar sua separação. Quando se trata de “documentos” como imagens que estão associados a diferentes tipos de conteúdo visual, a recuperação de informação visual é ainda mais complexa. Os índices devem ser significativos, discriminantes e utilizáveis. O índice é determinado por sua relação com a imagem. Os índices podem representar o conteúdo primitivo e o conteúdo complexo das imagens. Um dos problemas-chave da indexação de imagens consiste em conseguir estabelecer relações entre os dois tipos de conteúdos, de maneira que as desvantagens de cada um possam ser reduzidas. Um desafio crucial para a indexação e recuperação de imagens pelo conteúdo esta no desenvolvimento de mecanismos para a automação da indexação.

Em [Smeulders00], podemos encontrar uma referência geral sobre os esforços realizados para a recuperação de imagens. As técnicas atuais de extração de características estão baseadas em cor, textura e forma, mas estas apresentam certas limitações quando o objeto sofre transformações geométricas. Neste módulo, pretende-se criar uma técnica de extração de características invariantes e tolerantes ao ruído. Existem diferente técnicas tais como os momentos estatísticos [Hu61], [Kan02], as transformadas de Fourier [González93], as transformadas de Wavelets [Wang01], autômatos celulares [Wolfram83] e Redes de Pulso Acoplado [Jonhson94]. Estas duas últimas técnicas visam potenciais métodos para extrair características, tanto de cor, textura e formas, assim como de características invariantes. Estes métodos serão aplicados sobre bases de imagens controladas (bases de imagens internacionais e públicas) já utilizadas em experimentos no NPDI [Rodrigues02], [Rodrigues03a,b].

6.3 Classificação automática de cenas

Classificação automática de cenas (em tipos gerais como “ambientes fechados”, “ambientes naturais”, “cidades”, “animais”, etc.) pode ser útil, tanto porque pode ser usada como um “filtro” antes de buscas mais específicas quanto pode ajudar na identificação de objetos pertencentes unicamente a um único contexto. Um dos primeiros sistemas desse tipo foi o Sistema IRIS [Hermes95], o qual usa uma combinação de cor, textura e relação espacial entre as regiões para derivar uma interpretação da cena, gerando descrições do tipo “montanha”, “floresta”, “lago”, etc. que serve de entrada para um sistema com interface baseada em texto.

Outras abordagens posteriores também seguiram a linha de tentar fazer uma análise de cena. Por exemplo, [Oliveira99] usou filtros de Gabor sobre características de bordas como característica global da imagem para classificar as cenas em duas classes: artificial X natural e ainda ambiente fechado X aberto. [Szummer98] usou uma combinação entre histogramas de cor, textura e Transformada Discreta de Cosseno para treinar um classificador baseado no Algoritmo do Vizinho mais Próximo para distinguir entre cenas ao ar livre e em ambientes fechados. Testes empíricos mostraram que o método obtém até 90% de precisão quando usado para classificar um conjunto de 1300 fotografias coloridas. [Lipson97] propôs uma abordagem diferente, baseada em uma análise semântica qualitativa, usando padrões e uma combinação de distribuição de cor para prototipar cenas como “montanhas” ou “campos”. Foi divulgada uma acurácia de 75% na classificação de fotografias de montanhas, com 12% de falsos positivos. [Vailaya98] desenvolveu um modelo de Classificador Bayesiano para agrupar imagens em um número de categorias, incluindo cidades X paisagens e florestas X montanhas, usando vetores gerados por quantização vetorial a partir de um espaço de momentos de cores e coeficientes de Gabor. A acurácia divulgada foi em torno de 90% para a maioria das imagens classificadas.

[Paek99] desenvolveu um protótipo de um sistema para classificação de fotografias de jornais em cenas de ambientes abertos e fechados, baseado em palavras-chave de títulos de figuras e histogramas de distribuição de cores e bordas. Este sistema conseguiu 86% de acurácia e facilmente superou o método de [Szummer98] o qual conseguiu apenas 74% de acurácia com o mesmo conjunto de testes. Um problema potencial está na escolha das classes que as imagens devem ser assinaladas, uma vez que isso sempre é uma escolha subjetiva e, portanto, limita os resultados experimentais obtidos. Uma exceção para esse problema é o chamado Classificador Hierárquico, desenvolvido por [Vailaya98], o qual usa técnicas automáticas de agrupamento baseadas em classificação subjetiva feita por uma interface independente que gera conjunto de classes.

Estes métodos serão aplicados sobre a grande base digital do acervo de fotografias históricas do Arquivo Público Mineiro (Convênio de Cooperação Técnico-científico entre a UFMG e o APM, no. 1450/0/1998, nov. 1998 a out. 2004) [Valle02a,b], [Valle03], visando sua classificação em cenas do tipo cidades X paisagens, florestas X montanhas, interiores X exteriores, fotos individuais X fotos de grupos, etc. Trata-se de uma base extremamente indexada através de metadados (ano, evento, personalidade, etc.) e que desta maneira servirá perfeitamente para a realização dos testes de classificação previstos neste projeto.

6.4 Plataforma DB2 da IBM e interface para um sistema de RIBC

Atualmente, o NPDI apóia dois projetos do CECOR: RECICOR, coordenado pelo prof. Luiz Souza [Nunes02a,b], e GUIGNARD, coordenado pela profa. Claudina Moresi [Moresi02a,b].

O CECOR reuniu um enorme acervo sobre as obras que foram restauradas em seus 21 anos de funcionamento. Esse acervo é composto por dados bibliográficos, materiais e técnicas utilizados, fotografias e outras informações inéditas no Brasil e na América Latina. Dessa forma, esse acervo se tornou uma rica fonte de pesquisa para os profissionais da área. Mas o acesso a essa informação é lento e trabalhoso, além de necessitar que as pessoas interessadas tenham que se deslocar até o CECOR. Encontra-se em fase de desenvolvimento o Projeto RECICOR, visando a implementação de um sistema de banco de dados que seja acessível pela Internet e torne o cadastro e a consulta ao acervo mais rápido e simples.

Já o Projeto GUIGNARD visa o desenvolvimento de protótipo de um sistema de informação multimídia destinado ao armazenamento de informações textuais, visuais e audiovisuais e acesso às obras de Alberto da Veiga Guignard (1896 - 1962), considerado grande mestre da pintura moderna, formador de escola e de várias gerações de artistas. Teve grande produção artística, desconhecendo-se ainda a totalidade de suas obras, que compreende desenhos, pinturas sobre madeira, pinturas sobre tecido, pinturas murais e forros. Apesar da vasta literatura - livros referentes à sua biografia, análises estéticas da obra, incluindo relatos de depoimentos, livros sobre a criação da escola Guignard, dentre outros, o artista permanece desconhecido do grande público. O projeto pretende promover a compreensão da obra de Guignard, desenvolvendo uma metodologia de análise e caracterização, priorizando o uso de técnicas não-destrutivas. Trata-se de desenvolver, consolidar e difundir nova concepção de pesquisa e conservação, alicerçadas em bases multidisciplinares. Como resultado, o grande público terá acesso a este rico acervo através da Internet.

Pretendemos, neste projeto, estudar a plataforma DB2 da IBM visando seu uso nestas aplicações. Além disto, pretendemos desenvolver uma interface para sistema de RIBC que seja comum a todas as nossas pesquisas nesta área.

Referências

Abate, A.F. et al. IME: an image management environment with content-based access. *Image and Vision Computing*, 17:967-980, 1999.

Abbadeni, N., Ziou, D. & Wang, S. *Image classification & retrieval on the www*. Proceedings of the 4th ACM Conference on Digital Libraries, 1999, Berkeley, CA USA, p. 208 – 209.

Araújo, A. de A., Guimarães, S.J.F. & Cerqueira, G.C. New approach for old movie restoration, IS&T/SPIE Electronic Imaging 2001, San Jose-CA, USA. In: Frank, A.M. (Editor), Conference on High-Speed Imaging and Sequence Analysis III, SPIE, ISBN no. 0 8194 3986 X, vol. 4308, 2001, pp 67-77.

Araújo, A. de A. & Guimarães, S.J.F. Recuperação de informação visual com base no conteúdo em imagens e vídeos digitais, Edição Especial em Computação Gráfica e Processamento de Imagens, Revista de Informática Teórica e Aplicada - RITA, UFRGS, Porto Alegre-RS, Brazil, vol. 7, no. 2, ISSN no. 0103 4308, 2000, pp 43-72.

Araújo, A. de A., Hadad, R.M. & Coelho, B.V. Enhancement of sketch contours on paintings infrared photographs: a comparison, Proceedings of the 9th. European Signal Processing Conference - EUSIPCO, Island of Rhodes, Greece. In: Theodoridis, S., Pitas, I., Stouratis, A. & Kalouptsidis, N. (Editors), Signal Processing IX: Theories and Applications, EURASIP, ISBN no. 960 7620 07 0, vol. II, 1998, pp 853-856.

Araújo, A. de A. & Hadad, R.M. Enhancement of sketch contours on paintings infrared photographs by optimal filtering, Proceedings of the II Workshop on Cybernetic Vision, IEEE Computer Society Press, São Carlos-SP, Brazil, 1997, ISBN no. 0 8186 8058 X, pp 231-236.

Araújo, A. de A. & Bertrand, G. Enhancement of paintings infrared photographs using mathematical morphology and topological processing, Proceedings of the International Symposium on Optical Science, Engineering, and Instrumentation - SPIE, Denver-CO, USA. In Tescher, A.G. (Editor), Applications of Digital Image Processing XIX, SPIE, ISBN no. 0 8194 2235 5, vol. 2847, 1996, pp 677-685.

Araújo, A. de A. & Bertrand, G. Cooperative use of mathematical morphology and topology algorithms to enhance sketch contours on paintings infrared photographs, in Banon, G. (Editor), Proceedings of the II Brazilian Workshop on Mathematical Morphology, São Paulo-SP, Brazil, 1996, <http://www.inpe.br/~banon/URLib2/>.

Araújo, A. de A., Laender, A.F.H., Silva Jr., N.I. & Coutinho, R.N. Image coding for storing historical documents in a database, Proceedings of the 1993 Picture Coding Symposium - PCS, Lausanne, Switzerland, 1993, pp 5.5-6.

Araújo, A. de A., Laender, A.F.H., Silva Jr., N.I., Coutinho, R.N., Veloso, B.R. & Coelho, B.V. An image data base system of brazilian historical documents, Proceedings of the 2nd. International Conference on Automation, Robotics and Computer Vision - ICARCV/IEEE, Singapore, vol. 2, 1992, pp cv-16.6.1-5.

Araújo, A. de A., Davis Jr., C.A., Daker, A.L.V., Souza, L.A.C. & Leal, A.S. Digital processing of x-rays of sculptures: A case study of Aleijadinho's work, Proceedings of the 4th. International Conference on Image Processing and its Applications - IPA/IEE, Maastricht, The Netherlands, ISBN 0 85296 543 5, 1992, pp 617-620.

Athitsos, V., Swain, M.J. & Frankel, C. *Distinguishing photographs and graphics on the www*. Proceedings of IEEE Workshop on Content-Based Access of Image and Video Libraries, Puerto Rico, June 1998.

Baeza-Yatez, R. & Ribeiro-Neto, B. *Modern Information Retrieval*, Wesley, 1999.

Berman, A.P. & Shapiro, L.G. A flexible image database system for content-based retrieval. *CVIU* 75(1/2):175-199, 1999.

Blatner, M.M. & Dannenberg, R.B. *Multimedia Interface Design*, ACM Press, 1992.

Bozcaya, T. & Ozsoyoglu, M. *Distance-Based Indexing for High-Dimensional Metric Spaces*. ACM SIGMOD 1997. pp. 357-368.

Brin, S. *Near Neighbor search in large metric spaces*. VLDB 1995. pp. 574-584.

Brown, L. & Gruenwald, L. *Tree-based indexing for image data*, JVICR, December, 1998, 9(4):300-313.

Brunelli, R., Mich, O. & Modena, C.M. *A survey on the automatic indexing of video data*. Journal of Visual Communication and Image Representation, 10(6):78--112, June 1999.

Catell, R.G.C. *Object-oriented and Extended Relational Database-Systems*, Adison-Wesley, 1994.

Chang, S.F., Huang, Q., Huang, T., Puri, A. & Shahraray, B. *Advances in Multimedia: Systems, Standards and Networks*, chapter Multimedia Search and Retrieval. New York: Marcel Dekker, 1999.

Chiueh, T. *Content-Based Image Indexing*. VLDB 94. pp.582-593.

Ciaccia, P., Patella, M. & Zezula, P. *M-Tree: An efficient access method for similarity search in metric spaces*. VLDB 1997. pp. 426-435.

CONWAY, P. *Preservação no universo digital*. Coordenação de Ingrid Beck, tradução de Olga de Souza Marder. Rio de Janeiro, Arquivo Nacional, 1997.

De Andrade, N.S., & Araújo, A. de A. *Multimídia para acesso a acervos históricos*, Revista iP-Informática Pública, Belo Horizonte-MG, Brazil, vol. 2, no. 1, ISSN no. 1516 697X, 2000, pp 49-66.

De Andrade, N.S., Araújo, A. de A. & De Melo, C.H. *A multimedia information system for governmental historical documents*, Proceedings (CD-ROM) of the Museums and the Web: An International Conference, Toronto, Ontario, Canada, 1998.

De Andrade, N.S., Araújo, A. de A. & De Melo, C.H. *Um sistema de informação multimídia para recuperação de documentos históricos do Arquivo Público Mineiro*, Anais (CD-ROM) do XII Congresso Brasileiro de Arquivologia, João Pessoa-PB, Brazil, junho 1998.

De Andrade, N.S. *Sistemas de Informação Multimídia*. 1998. 73 f. Dissertação (Mestrado em Administração Pública) - Fundação João Pinheiro. Orientador: Araújo, A. de A.

Del Bimbo, A. *Visual Information Retrieval*, Morgan Kaufmann, 1999.

Descrição do Programa *Minas Gerais: Memória, Patrimônio Documental e Informação Digital*. Belo Horizonte, Secretaria de Estado da Cultura, 1999.

Elmasri, R. & Navathe, S. *Fundamentals of Database Systems*. 3^a. Ed. Reading MA, Addison Wesley Longman, Inc. 2000.

Faloutsos, C. & Lin, K. *FastMap: A Fast Algorithm for Indexing, Data-Mining and Visualization of Traditional and Multimedia Datasets*. SIGMOD 1995 pp. 163-174.

Formulário para Apresentação de Projeto Cultural, *Digitalização de Imagens: A preservação da memória fotográfica de Minas*. Belo Horizonte, Associação Cultural do Arquivo Público Mineiro, 1999.

Faulus, D.S. & Ng, R.T. *An expressive language and interface for image querying*, Machine Vision and Applications, 10, 1997, pp 66-73.

Flickner, M. et al. *Query by image and video content: the QBIC system*. IEEE Computer, 1995, v. 28, n. 9, p. 23 - 32, September.

Frakel, C., Swain, M. J. & Athitsos, V. *Webseer: An image search engine for the world wide web*. Technical Report TR-96-14. Computer Science Department, University of Chicago, 1996.

Gaede, V. & Gunther, O. *Multidimensional Access Method*. ACM Computing Surveys. Vol. 30. Nº 2. pp. 170-231, 1998.

Gevers, T. & Smeulders, A.W.M. *Contentbased image retrieval by viewpointinvariant color indexing*. Image and Vision Computing, 17(7):475--488, May 1999.

Gonzales, R.C. & Woods, R.E. *Digital Image Processing*, Addison-Wesley, 1993.

Goodrum, A. & Spink, A. Image searching on the excite web search engine. Information Processing and Management, 2001, v. 37, p. 295 – 311.

Gu, C. & Kunt, M. *Contour simplification and motion compensated coding*, Signal Processing: Image Communications, Elsevier, 1995, 7:279-296.

Gudivada, V. N. & Raghavan, V. V. *Content-based image retrieval systems*. IEEE Computer, 1995, v. 28, n. 9, p. 18 - 22, September.

Guimarães, S.J.F. Identificação de Transições em Vídeo Baseada na Análise de Imagens 2D. 2003. Tese (Doutorado em Ciência da Computação), UFMG – Orientador: Araújo, A. de A.

Guimarães, S.J.F., Couprie, M., Araújo, A. de A. & Leite, N.J. Video segmentation based on 2D image analysis, Pattern Recognition Letters, Elsevier, Amsterdam, Holand, v. 24, n. 07, p. 947-957, 2003.

Guimarães, S.J.F., Couprie, M., Leite, N.J. & Araújo, A. de A. Video fade detection by discrete line identification. In: Kasturi, R., Laurendeau, D. & Suen, C. (Editors), Proceedings of the 16th. International Conference on Pattern Recognition - ICPR, Québec City-QC, Canada, IAPR/IEEE, ISBN no. 0 7695 1699 8, vol. 2, 2002, pp 1013-1016.

Guimarães, S.J.F., Leite, N.J., Couprie, M. & Araújo, A. de A. A multi-scale, directional and parametrized transition detection algorithm, Proceedings of the XV Brazilian Symposium on

Computer Graphics and Image Processing - SIBGRAPI, IEEE Computer Society Press, Fortaleza-CE, Brazil, ISBN no. 0 7695 1846 X, 2002, pp 261-268.

Günsel, B. & Tekalp, A.M. *Shape similarity matching for query-by-example*, Pattern Recognition, 31/7, 1998, pp 931-944.

Gupta, A. & Jain, R. *Visual information retrieval*. Communications of the ACM, 1997, v. 40, n. 5, p. 71 - 79, May.

Guttman, A. *R-Tree: A dynamic Index Structure for Spatial Searching*. ACM SIGMOD 1984. pp. 47-57.

Hermes, T. et al. Image retrieval for information systems. SPIE Storage and Retrieval for Image and Video Databases III, pp 394-405, 1995.

Hu, M.K. *Pattern Recognition by Invariant Moments*, Proc. IRE Transactions on Information Theory, 1961, pp 179-187.

Jin, J.S. & Kurniawati, R. A scheme for intelligent image retrieval in multimedia databases. Journal of VCIR, 7(4):360-377, 1996.

Johnson, J. *Pulse-coupled neural net: translation, rotation, scale, distortion, and intensity signal invariance for images*, Opt. Soc. of America, vol. 33, no 26, 1994, pp 6239-6253.

Kan, C. & Srinath, M.D. *Invariant character recognition with Zernike and orthogonal Fourier-Mellin moments*, Pattern Recognition, vol. 35, 2002, pp 143-154.

Khoshafian, S. & Baker, A. B. *Multimedia and Imaging Databases*, Morgan Kaufmann, San Francisco, 1996.

Lipson, P. et al. Configuration-based scene classification and image indexing. IEEE CVPR, pp 1007-1013, 1997.

Mitre, T.K.. Sistema de Informação Multimídia para Consulta a Laudos Periciais. 2001. Dissertação (Mestrado em Ciência da Computação) - UFMG. Orientador: Araújo, A. de A..

Moresi, C.M.D., Neves, A.R.A., Veloso, B.R., Araújo, A. de A., Valle Jr., E.A., Trindade, S.M.C., Julião, L., Teixeira, I.A.C., Vinhosa, F.L.T. & Hadad, R.M. Estudo e digitalização da obra de Guignard, pintor modernista, Anais (CD-ROM) do III Workshop em Tratamento de Imagens, DCC/UFMG, Belo Horizonte-MG, Brazil, 2002, 6p.

Moresi, C.M.D., Neves, A.R.A., Veloso, B.R., Trindade, S.M.C., Juliao, L., Teixeira, I.A.C., Araújo, A. de A., Hadad, R.M., Vinhosa, F.L.T. & Vieira, I.L. Documentação e estudo da obra de Guignard, Proceedings (CD-ROM) of the Committee International for Documentation Conference - CIDOC/ICOM, Porto Alegre-RS, Brazil, 2002, 2 p.

Nunes, F.H.C., Souza, L.A.C., Araújo, A. de A. & Correa, M.A. Disponibilizando e Preservando o Acervo sobre Conservação e Restauração de Bens Culturais Móveis do CECOR, Revista Eletrônica de Iniciação Científica – REIC, SBC, ano II, vol. II, no. III, ISSN 1519 8219, 2002, <http://www.sbc.org.br/reic>.

Nunes, F.H.C., Souza, L.A.C., Araújo, A. de A. & Correa, M.A. Um sistema de informação multimídia para o CECOR, Proceedings of the VIII Brazilian Symposium on Multimedia and Hypermedia Systems - SBMIDIA, Tools and Applications Workshop, Fortaleza-CE, Brazil, 2002, pp 391-394.

Oliva, A. et al. Global semantic classification of scenes using power spectrum templates. The Challenge of Image Retrieval, UK, 1999.

Oliveira, C.J.S., Araújo, A. de A., Severiano Jr., C.A. & Gomes, D.R. Classificação de Imagens Coletadas na World Wide Web, Proceedings of the VIII Brazilian Symposium on Multimedia and Hypermedia Systems - SBMIDIA, Fortaleza-CE, Brazil, 2002, pp 306-313.

Oliveira, C.J.S., Araújo, A. de A., Gomes, D.R. & Severiano Jr., C.A. Protótipo de um sistema de recuperação de imagens baseado na cor, Proceedings of the VIII Brazilian Symposium on Multimedia and Hypermedia Systems - SBMIDIA, Tools and Applications Workshop, Fortaleza-CE, Brazil, 2002, pp 411-414.

Oliveira, C.J.S., Araújo, A. de A., Severiano Jr., C.A. & Gomes, D.R. Classifying images collected on the World Wide Web, Proceedings of the XV Brazilian Symposium on Computer Graphics and Image Processing - SIBGRAPI, IEEE Computer Society Press, Fortaleza-CE, Brazil, ISBN no. 0 7695 1846 X, 2002, pp 327-334.

Oliveira, C.J.S. Classificador de Imagens Fotográficas Coletadas na World Wide Web. 2001. 73 f. Dissertação (Mestrado em Ciência da Computação) - UFMG. Orientador: Araújo, A. de A.

Paek, S. et al. Integration of visual and text-based approaches for the content labelling and classification of photographs. ACM SIGIR99 – W. on Multimedia Indexing and Retrieval, 1999.

Park, I. K., Yun, I. D. & Lee, S. U. *Color image retrieval using hybrid graph representation*, Image and Vision Computing, Elsevier, 17(7):465-474, 1999.

Rodrigues, P.S.R. Um modelo Bayesiano combinando Análise Semântica Latente, Agrupamento de Imagens e Atributos Espaciais Globais para Recuperação de Informação Visual. 2003. Tese (Doutorado em Ciência da Computação), UFMG – Orientador: Araújo, A. de A.

Rodrigues, P.S.R. & Araújo, A. de A. A Bayesian Network to Improve Content Based Image Retrieval, submetido, IEEE Pattern Analysis and Machine Intelligence, 2003.

Rodrigues, P.S.R. & Araújo, A. de A. A region-based object recognition algorithm, Proceedings of the XV Brazilian Symposium on Computer Graphics and Image Processing - SIBGRAPI, IEEE Computer Society Press, Fortaleza-CE, Brazil, ISBN no. 0 7695 1846 X, 2002, pp 283-289.

Roth, V. Contentbased retrieval from digital video. Image and Vision Computing, 17(7):531--540, May 1999.

Rui, Y., Huang, T. S. & Chang. S.-F. *Image retrieval: current techniques, promising directions, and open issue*, Journal of Visual Communications and Image Representation, 1999, v. 10, p. 39 – 62.

Rui, Y. & Huang, T.S. *Image Retrieval: Current Techniques, Promissing Directions and Open Issues*, The NECI Scientific Literature Digital Library.

Sabharwal, C.L. & Bhatia, S.K. *Image databases and near-perfect Hash table*, Pattern Recognition, 30/11, 1997, pp 1867-1876.

Schmid, C. & Mohr, R. *Image retrieval using local characterization*, Fourth DELOS Workshop, Image Indexing and Retrieval, 1997.

Schweitzer, H. Organizing image databases as visualcontent search trees. *Image and Vision Computing*, 17(7):501—511, May 1999.

Sclaroff, S. & Taycher, L. & La Cascia, M. *ImageRover: A Content-Based Image Browser for the World-Wide Web*. IEEE Workshop on Content-based Acess for Image and Video Libraries. V. 6. 1997.

Sethi, I.K. & Coman, I. *Image retrieval using hierachical self-organizing feature maps*. *Pattern Recognition Letters*, no. 20, 1999, pp 1337-1345.

Shen, D., Wai, W.H. & Horace, H. S. . *Affineinvariant image retrieval by correspondence matching of shapes*. *Image and Vision Computing*, 17(7):489--499, May 1999.

SILVA JR., N.I. Um Sistema de Compressão de Imagens Aplicado a Documentos Históricos. 1993. 132 f. Dissertação (Mestrado em Ciência da Computação) - UFMG. Orientador: Araújo, A. de A.

Smeulders, A.W.M et al. *Content-Based Image Retrieval at the End of the Early Years*. IEEE Trans. on PAMI, vol. 22, no. 12, 2000, pp 1349-1381.

Smith, J.R. & Chang, S.F. Searching for images and videos on the www. TR CU/CTR 459-96-25, Columbia Univ. 1996.

Stricker, M. & Dimai, A. *Spectral covariance and fuzzy regions for image indexing*, *Machine Vision and Applications*, 10, 1997, pp 66-73.

Szummer, M. & Picard, R. Indoor-outdoor image classification. IEEE Int. W. on Content-based access of Image and Video Databases, pp 42-51, 1998.

Traina Jr., C., Traina, A. J. M. & Faloutsos, C. *Distance Exponent: A New Concept for Selectivity Estimation in Metric Trees*. Technical Report. CMU-CS-99-110. 1999.

Vailaya, A., Jain, A. & Zhang, H.J. *On image classification: city vs. Landscape*. Proceeding of IEEE Workshop on Content-Based Access of Image and Video Libraries, Santa Barbara, California, June 1998.

Valle Jr., E.A. Sistemas de informação multimídia para documentos históricos. 2003. Dissertação (Mestrado em Ciência da Computação) – Orientador: Araújo, A. de A.

Valle Jr., E.A. & Araújo, A. de A. [Preserving Historical Collections Using Multimedia Information Systems](#), Proceedings of the VIII Brazilian Symposium on Multimedia and Hypermedia Systems - SBMIDIA, Thesis and Dissertation Workshop, Fortaleza-CE, Brazil, 2002, pp 317-324.

Valle Jr., E.A., Araújo, A. de A., Vieira, F.M. & Costa, C.C.P. [A tool for workflow management in the composition of multimedia databases from preexistent documents](#), Proceedings of the VIII Brazilian Symposium on Multimedia and Hypermedia Systems - SBMIDIA, Tools and Applications Workshop, Fortaleza-CE, Brazil, 2002, pp 379-382.

Wang, J.Z. *Wavelets and Imaging Informatics: A Review of the Literature*, *Jornal pf Biomedical Informatics*, vol. 34, 2001, pp 129-141.

Watanabe, S. *Pattern Recognition: Human and Mechanical*, Wiley, 1985, New York.

Wolfram, S. *Statistical Mechanics of Cellular Automata*, *Rev. Mod. Phys.* vol. 55, 1983, pp 601-644.