



Pós-Graduação em Ciência da Computação

MetaMídia – Um Modelo de Metadados na Indexação e Recuperação de Objeto Multimídia

por

Maria Salete Marcon Gomes Vaz

Tese de Doutorado



Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
<http://www.cin.ufpe.br/~posgraduacao>

RECIFE, DEZEMBRO/2000



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

MARIA SALETE MARCON GOMES VAZ

METAMÍDIA – UM MODELO DE METADADOS NA
INDEXAÇÃO E RECUPERAÇÃO DE OBJETOS MULTIMÍDIA

ESTE TRABALHO FOI APRESENTADO À PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO DO CENTRO DE INFORMÁTICA DA UNIVERSIDADE FEDERAL DE PERNAMBUCO COMO REQUISITO PARCIAL PARA OBTENÇÃO DO GRAU DE DOUTORA EM CIÊNCIA DA COMPUTAÇÃO.

ORIENTADORES: *Fernando da Fonseca de Souza*
Alexandre M. Lins de Vasconcelos
Ana Carolina Salgado

RECIFE, DEZEMBRO/2000

Resumo

Um banco de dados convencional é um repositório de dados, que contém valores numéricos, alfanuméricos, booleanos e datas. Um banco de dados multimídia, adicionalmente, contém imagens gráficas, clipes de vídeo, arquivos de som, textos, entre outros.

O dado armazenado no banco de dados deve ser estruturado de tal modo que a informação desejada possa ser facilmente localizada e recuperada. A modelagem de dados é o termo usado para determinar quais dados devem ser armazenados e refletir os relacionamentos entre eles.

A representação dos dados convencionais é feita essencialmente por atributos que descrevem objetos do mundo real, enquanto que dados multimídia são objetos do mundo real que são mantidos no banco de dados.

A dificuldade em descrever (indexar) objetos multimídia com o objetivo de recuperá-los motivou o desenvolvimento desta tese. Um modelo de metadados é proposto para descrever objetos multimídia e permitir a indexação e recuperação desses objetos.

O uso de metadados na descrição de objetos multimídia justifica-se pela dificuldade em efetuar casamento exato de padrões; dificuldade na pesquisa baseada em conteúdo, devido à análise em grande conjunto de dados; e quando a pesquisa baseada em conteúdo é possível de ser realizada, a mesma não pode ser feita de fato por razões de desempenho.

A importância, relevância e originalidade desta proposta estão relacionadas aos seguintes itens: (i) integração de dados de diferentes bases, com tratamento uniforme, através do uso de uma base de metadados descritivos; (ii) descrição de objetos multimídia, usando um modelo de metadados; (iii) compartilhamento de dados entre objetos, através do mecanismo de herança; e (iv) utilização da composição de objetos para reduzir os resultados de pesquisas.

Abstract

A database is a repository of data, that conventionally contains values of types numerical, alphanumeric, boolean and date. A multimedia database, additionally, contains graphical pictures, clips of video, archives of sound, texts, among others.

The stored data inside a database must be structured in such a mode that the desired information can easily be located and retrieved. The modeling of data is the term used to determine which data must be stored and to reflect the relationships between the data.

The representation of the conventional data is made essentially by attributes that describe objects of the real world, while multimedia data are objects of the real world that are kept in the database. The difficulty in describing (indexing) multimedia data with the objective of retrieval motivated the development of this thesis. A model of metadata is proposed to describe multimedia data and to allow the indexation and retrieval of these data.

The use of metadata in the multimedia object description is justified due to the difficulty in doing pattern matching; difficulty in searching based on content due to the analysis in large; and also because when the search based even on content is possible to be carried through, it cannot be done due to performance limitations.

The importance, relevance and originality of this proposal are related to items: (i) integration of data of different bases, with uniform treatment, through the use of a base of descriptive metadata; (ii) multimedia object description, using a model of metadata; (iii) sharing of data between objects, through the inheritance mechanism; e (iv) use of the object composition to reduce the results of search.

Dedicatória

Dedico esta tese aos meus amigos Anjolina Grissi,
Álvaro Justus, Vanderlei Vilanova, Mariléia Pimentel, Virgínia de Paula, Adeilva
Tenório, que sem eles tudo seria muito mais difícil.

Vocês foram amigos acima de tudo! Obrigada!

Dedico a minha filha Rafaela e ao meu marido Gil por terem esperado a finalização
deste trabalho. Foram muitos momentos sem a minha presença.

Obrigada pelo Amor e Compreensão!

Agradecimentos

Ao professores Fernando Fonseca, Alexandre Vasconcelos e Ana Carolina Salgado, pelo apoio, orientação e estímulo recebido durante o desenvolvimento deste trabalho.

Aos professores, componentes da banca examinadora, José Valdeni de Lima, Guido Lemos, Judith Kelner, Carlos Ferraz e Alexandre Vasconcelos pelas críticas e sugestões importantes para enriquecimento deste trabalho.

Aos professores Alexandre Vasconcelos, Kátia Guimarães e Fábio Silva pela orientação competente recebida durante o desenvolvimento dos exames de qualificação para o doutoramento. Obrigada!!!

Ao amigo Vanderlei, por todos os momentos dispensados para o meu crescimento intelectual e científico. Pelas dicas, pelas críticas, pelo incentivo, pelo papo, pelo ombro, enfim por tudo. Obrigada!

Ao Ricardo Czelusniak meu agradecimento especial pela dedicação e competência na implementação do Protótipo *MetaMídia*. Muito Obrigada!!!!

Ao Gil, por estar sempre ao meu lado; pela alegria, pelo incentivo, pelo apoio nos momentos mais difíceis, pela força e pelo amor. Obrigada!!!

À Luciana, minha fiel “escudeira”, que tratou a minha filha com amor e dedicação, tornando a minha caminhada menos árdua. Obrigada!!!

A minha amiga Mariléia, pelo apoio direto e indireto. Pela dedicação e ajuda durante todo o tempo. Obrigada pelo carinho e amizade!

Aos meus amigos especiais Álvaro, Anjolina, Celso, Virginia, Carla, que estavam sempre dando apoio e carinho. Muito agradecida!

Aos meus pais e meus irmãos, pelo amor, incentivo e compreensão recebidos sempre.

Aos meus colegas e amigos do CIN, obrigada pela força. E também a todos que direta e indiretamente contribuíram com carinho, amizade e incentivo. Obrigada!!!

Ao Centro de Processamento de Dados da Universidade Estadual de Ponta Grossa e a própria Universidade, pela oportunidade e apoio recebidos sempre. Obrigada!!!

Conteúdo

1 INTRODUÇÃO	1
1.1 MOTIVAÇÃO	1
1.2 OBJETIVOS	3
1.3 ESTRUTURA	5
2 MODELO DE METADADOS E BANCO DE DADOS MULTIMÍDIA	7
2.1 INTRODUÇÃO	7
2.2 CONCEITOS BÁSICOS	9
2.2.1 Multimídia.....	9
2.2.2 Dado e Informação.....	13
2.2.3 Objeto Multimídia.....	14
2.3 MODELO DE METADADOS	15
2.3.1 Classificação de Metadados.....	16
2.3.2 Metadados no Contexto Organizacional.....	19
2.3.4 Gestão de Metadados.....	21
2.4 INDEXAÇÃO DE INFORMAÇÃO	23
2.5 RECUPERAÇÃO DE INFORMAÇÃO	26
2.5.1 Métodos de Recuperação.....	27
2.5.2 Algoritmos de Pesquisa.....	30
2.6 ARMAZENAMENTO DE OBJETOS	33

2.7 CONSIDERAÇÕES FINAIS	36
3 METAMÍDIA – UM MODELO DE METADADOS PARA OBJETOS	
MULTIMÍDIA	38
3.1 INTRODUÇÃO	38
3.2 MODELO DE METADADOS	39
3.2.1 Hierarquia de Metadados/Valores.....	41
3.2.2 Meta-Objeto Multimídia.....	44
3.2.3 Criação de Meta-Objeto Multimídia.....	46
3.2.4 Disjunção de Meta-Objetos Multimídia.....	47
3.2.5 Conjunção de Meta-Objetos Multimídia.....	49
3.2.6 Mecanismo de Herança entre Meta-Objetos Multimídia.....	51
3.3 INDEXAÇÃO DE OBJETOS MULTIMÍDIA E ANÁLISE DE DESEMPENHO	55
3.4 RECUPERAÇÃO DE OBJETOS MULTIMÍDIA	61
3.5 CONSIDERAÇÕES FINAIS	63
4 PROTÓTIPO METAMÍDIA.....	66
4.1 INTRODUÇÃO	66
4.2 LINGUAGEM UML PARA DESCRIÇÃO DO MODELO METAMÍDIA	68
4.3 DIAGRAMA DE CASOS DE USO	69
4.4 CENÁRIOS.....	71
4.5 PROJETO DE INTERFACE HOMEM/MÁQUINA	73
4.6 DIAGRAMA DE CLASSES	76

4.7 CODIFICAÇÃO	80
4.8 CONSIDERAÇÕES FINAIS	80
5 TRABALHOS RELACIONADOS	82
5.1 INTRODUÇÃO	82
5.2 MODELOS E SISTEMAS	85
5.2.1 Sistema VIMSYS.....	85
5.2.2 Sistema OVID.....	87
5.2.3 Sistema QBIC.....	90
5.2.4 Sistema CIARS	92
5.2.5 Sistema FIBSSR.....	94
5.2.6 Sistema Piction.....	96
5.2.7 Sistema MARS.....	98
5.3 ANÁLISE COMPARATIVA	101
5.4 DO VIMSYS AO METAMÍDIA.....	104
5.5 CONSIDERAÇÕES FINAIS	107
6 CONCLUSÕES E TRABALHOS FUTUROS	108
6.1 ENFOQUE GERAL	108
6.2 OBJETIVOS ALCANÇADOS	109
6.2.1 Descrição de Objetos Multimídia.....	109
6.2.2 Compartilhamento de Metadados.....	110
6.3 CONTRIBUIÇÃO PRINCIPAL	111

6.4 TRABALHOS FUTUROS	112
6.4.1 Modelo de Indexação e Processo de Ensino/Aprendizagem.....	113
6.4.2 Método de Recuperação Baseado em Similaridade.....	114
6.4.3 Mecanismo de Versões.....	114
REFERÊNCIAS BIBLIOGRÁFICAS	116

Lista de Figuras

Figura 1: Dado e Informação.....	13
Figura 2: De Dados a Conhecimento.....	15
Figura 3: Metadado para Representação de tipos de Mídia.....	17
Figura 4: Transformação de Dados em Informação.....	20
Figura 5: Hierarquia de Indexação Multimídia.....	24
Figura 6: Indexação de Informação Multimídia.....	25
Figura 7: Arquitetura do Modelo de Metadados.....	41
Figura 8: Hierarquia de Metadados/Valores.....	43
Figura 9: Meta-Objeto Multimídia	44
Figura 10: Criação de Meta-Objeto Multimídia.....	46
Figura 11: Disjunção de Meta-Objetos Multimídia.....	47
Figura 12: Conjunção de Meta-Objetos Multimídia.....	50
Figura 13: Herança Simples	52
Figura 14: Herança Múltipla entre Meta_Objetos Multimídia.....	54
Figura 15: Índices de Clusters do Modelo <i>MetaMídia</i>	56
Figura 16: Índices em uma Array Triangular.....	57
Figura 17: Fase de Pré-Processamento	60
Figura 18: Representação de Ator e Caso de Uso, em UML.....	69
Figura 19: Diagrama de Casos de Uso do Modelo.....	70

Figura 20: Tela Principal do Protótipo MetaMídia.....	74
Figura 21: Tela com Criação de metadados/valores para imagens.....	74
Figura 22: Tela de Conjunção ou Disjunção.....	75
Figura 23: Tela de Pesquisa.....	75
Figura 24: Diagrama de Classes do Protótipo.....	77
Figura 25: Herança de Inclusão de Intervalos no OVID.....	89
Figura 26: Do VIMSYS ao MetaMídia.....	105

Lista de Tabelas

Tabela 1: Metadados Descritivos.....	42
Tabela 2: Metadados Estruturais	45
Tabela 3: Análise Comparativa dos Modelos/Sistemas.....	101

Lista de Definições

Definição 1: Meta-Objeto Multimídia	44
Definição 2: Disjunção de Meta-Objetos Multimídia.....	48
Definição 3: Conjunção de Meta-Objetos Multimídia	50
Definição 4: Herança Simples entre Meta-Objetos Multimídia.....	52
Definição 5: Herança Múltipla entre Meta-Objetos Multimídia.....	54
Definição 6: Cálculo do número possível de índices.....	57
Definição 7: Cálculo da Quantidade de <i>Clusters</i> que são visitados.....	59
Definição 8: Cálculo da quantidade de clusters para todos os casos possíveis.....	59
Definição 9: Cálculo da quantidade máxima de descrições a serem pesquisadas....	60
Definição 10: Cálculo da quantidade máxima de elementos a serem pesquisados..	61
Definição 11: Algoritmo Pesquisa e Recuperação de Objetos Multimídia.....	62
Definição 12: Ordem de Complexidade do Algoritmo MetaMídia.....	62

Capítulo 1

Introdução

Tempos difíceis têm um valor científico.

Eles são as oportunidades

Que um bom aprendiz jamais perde.

Ralph Waldo Emerson

Neste capítulo é apresentada a motivação para o desenvolvimento desta tese, onde são enfatizadas a importância, originalidade e relevância do trabalho. Os objetivos da tese são apresentados, bem como um resumo dos objetivos de cada capítulo.

1.1 Motivação

Nos últimos anos, muita ênfase tem sido dada ao desenvolvimento de sistemas de banco de dados multimídia [Nar96, Gro97, HSHA98, GD98, KSC99, GCJ00], os quais fazem uso eficiente de tipos de dados complexos. Esses tipos de dados podem ser imagens, gráficos, seqüências de vídeo, imagens de satélite e dados científicos, entre outros.

O tratamento desses tipos de dados, em Sistemas de Banco de Dados tradicionais [Dat99, EN99, SKS99], tem sido problemático devido à complexidade de suas estruturas, à análise de grande conjunto de dados armazenados e ao baixo desempenho de recuperação quando a pesquisa baseada em conteúdo é efetuada.

Assim, a recuperação de informação baseada na descrição de conteúdo tem sido alvo de inúmeras pesquisas [CC98, WS99, Gor00, HMM00]. A definição das informações que devem ser fornecidas com o objetivo de efetuar futura recuperação, de forma fácil, rápida e correta, não é uma tarefa trivial.

Para tanto, em banco de dados multimídia, o processo de recuperação deve incluir os tipos de consultas conhecidos dos bancos de dados tradicionais, além das funcionalidades de pesquisas efetuadas pelo campo da recuperação da informação [YN99]. Os usuários podem navegar através dos dados e construir consultas estruturadas baseadas na descrição do conteúdo dos objetos multimídia. No contexto desta tese, um objeto multimídia é um objeto composto que referencia vários tipos de objetos, tais como vídeo, texto, som, imagem, entre outros [LZ96, YY97, CSBB97].

Um modelo básico de recuperação pressupõe que o objeto é armazenado para ser recuperado posteriormente. Para que o objeto seja armazenado é necessário indexá-lo (descrevê-lo). Porém, a indexação [GFSC94, AS94, CC98, BSH00] é uma atividade subjetiva que corresponde à identificação e representação do conteúdo ou do assunto de um documento de forma a caracterizá-lo. Diferentes pessoas, embora treinadas de acordo com os mesmos critérios, podem indexar de forma diferente o mesmo objeto.

A indexação de objetos multimídia pode ser feita através de um conjunto de metadados[AS94, CC98], os quais podem ser atribuídos aos objetos multimídia com interferência do usuário (manual), sem essa interferência (automática) ou com alguma interferência (semi-automática).

A recuperação pode depender da disponibilidade desses metadados introduzidos de forma automática ou não pelos usuários. Esses metadados podem ser baseados em conhecimentos adicionais incorporados à semântica dos dados e utilizados em uma aplicação particular. Algumas das principais razões que justificam o uso de metadados, em multimídia, são: dificuldade de efetuar casamento exato de padrões; dificuldade de pesquisa baseada em conteúdo, visto que há necessidade de análise em grande conjunto de dados; e, quando a pesquisa baseada em conteúdo é efetuada, a mesma não pode ser feita freqüentemente devido à degradação no desempenho.

1.2 Objetivos

O objetivo principal desta tese é apresentar um modelo de metadados na indexação e recuperação de objetos multimídia. Enquanto na modelagem de dados, um conjunto de ferramentas conceituais é utilizado para a descrição de dados, relacionamentos entre eles, a semântica de dados e regras de consistência, na modelagem de metadados, no contexto desta tese, é definido e usado um conjunto de ferramentas para indexar e recuperar objetos multimídia.

Para gestão de metadados, na indexação e recuperação multimídia, é necessário decidir quais metadados devem ser coletados e mantidos, e isso não é uma tarefa trivial.

Uma arquitetura de informação deve ser flexível para permitir inclusão ou exclusão de metadados à medida que novas necessidades apareçam.

Assim, existem alguns problemas relacionados à gestão de metadados que devem ser atacados, tais como: (i) metadados tomam uma variedade de formas; (ii) novos conjuntos de metadados são criados; (iii) diferentes usuários podem propor diferentes tipos de metadados; (iv) existem muitos usuários de metadados; (v) adoção de diferentes vocabulários de metadados significa aumento de buscas usando vocabulários que não são familiares.

Assim, estratégias para gestão de metadados devem ser estabelecidas. As tarefas para criar um ambiente de gestão de metadados são: definir requisitos para metadados que devem estar disponíveis para os usuários, desenvolver a arquitetura de gestão de metadados e implementar um sistema de gestão de metadados.

O fato é que representar o conteúdo semântico de objetos multimídia não é uma tarefa fácil, visto que um mesmo objeto pode ser descrito de várias formas, tanto por usuários distintos quanto por um mesmo usuário, em momentos distintos. Assim, um modelo de metadados de objetos multimídia baseado na descrição do conteúdo necessita ter grande expressividade, permitindo flexibilidade de uso.

O modelo de metadados proposto nesta tese permite a indexação e recuperação de objetos multimídia utilizando uma hierarquia de metadados/valores. Através da hierarquia é possível definir os tipos de metadados que são permitidos pelo modelo.

A hierarquia de metadados/valores corresponde a uma especificação de metadados. Essa hierarquia é uma descrição dos objetos multimídia e dos relacionamentos que podem existir entre esses objetos, gerando assim os meta-objetos multimídia.

A definição de meta-objetos multimídia é feita a partir de operações para criação, conjunção e disjunção. Essas operações permitem as relações de agregação, generalização e especialização de metadados/valores entre meta-objetos multimídia.

Na descrição de meta-objetos multimídia o mecanismo de herança é permitido. Assim, os meta-objetos multimídia podem herdar metadados/valores de meta-objetos existentes. Esse mecanismo contribui para melhorar a produtividade no desenvolvimento de sistemas de banco de dados multimídia, visto que muitas descrições não precisam ser dadas, uma vez que podem ser herdadas.

1.3 Estrutura

Além deste capítulo introdutório, esta tese é composta por mais cinco capítulos. No Capítulo 2, **Modelo de Metadados e Banco de Dados Multimídia**, são abordados a conceituação básica e os aspectos relacionados ao desenvolvimento de modelo de metadados e Banco de Dados Multimídia, evidenciando a importância desses conceitos para a tese.

No Capítulo 3, **MetaMídia – Um Modelo de Metadados para Objetos Multimídia**, é apresentada a parte central desta tese. Neste capítulo, é mostrado como representar objeto multimídia associado a metadados. Para tanto, um conjunto de

definições ligado a meta-objetos multimídia, indexação, pesquisa e recuperação de informação multimídia é apresentado.

No Capítulo 4, **Protótipo MetaMídia**, é apresentada a especificação parcial, em UML, do protótipo do modelo de metadados na indexação e recuperação de informação multimídia.

No Capítulo 5, **Trabalhos Relacionados**, são discutidos alguns dos principais trabalhos relacionados e é apresentada uma análise comparativa entre os mesmos. Os pontos destacados na análise são: as mídias abordadas na recuperação, modelo de indexação, os métodos de recuperação e os algoritmos de pesquisa implementados. A ênfase deste capítulo está nos trabalhos que abordam modelo de metadados.

No Capítulo 6, **Conclusões e Trabalhos Futuros**, a tese é analisada como um todo, apresentando as principais dificuldades enfrentadas para o seu desenvolvimento, as principais contribuições e quais as perspectivas de trabalhos futuros.

Capítulo 2

Modelo de Metadados e Banco de Dados

Multimídia

*O verdadeiro perigo não é que
os computadores comecem
a pensar como seres humanos,
mas que os seres humanos comecem
a pensar como computadores.*

Sydney J. Harris

Neste capítulo são abordados os aspectos inerentes a modelo de metadados e banco de dados multimídia. As propriedades e características relacionadas ao tratamento da informação multimídia são discutidas.

2.1 Introdução

Os sistemas de banco de dados tradicionais [Dat99, EN99, SKS99] utilizam tipos de dados simples, tais como números e *strings*, para representar e manipular informações

textuais tais como registros de empregados ou dados científicos. A formalização das consultas é geralmente expressa através da álgebra relacional e os resultados requerem pouco mais que um terminal textual para a apresentação.

Os sistemas de banco de dados multimídia [YY97,CSBB97,Gro97,GM98, Ass98,GD98,CNY98, KSC99,GCJ00] estão voltados para aplicações que interagem com o usuário, fazendo uso simultâneo de diversos tipos de mídia, tais como áudio, imagens estáticas, imagens em movimento, gráficos, textos, entre outras.

Esses sistemas, além de suas características peculiares, têm que manter, também, as propriedades inerentes aos sistemas convencionais: consistência de dados, segurança de dados, restrições de integridade, consulta e recuperação de dados.

Durante o processo de atualização (criação, remoção ou modificação), a visualização consistente dos dados tem que ser mantida, a fim de permitir que todos os usuários tenham a mesma visão desses dados. A inconsistência entre duas entradas representando o mesmo “fato” é um exemplo de perda de integridade. Procedimentos de validação são definidos para serem executados sempre que seja tentada uma operação de atualização no banco de dados.

Os sistemas devem garantir que a grande quantidade de informações diferentes, que dão entrada no banco de dados, possa ser consultada e recuperada de forma segura e consistente.

Para descrever os aspectos inerentes a Modelo de Metadados e Banco de Dados Multimídia, este capítulo é organizado como segue. Na Seção 2.2 são tratados os

conceitos básicos para entendimento desta tese. As características inerentes aos sistemas de banco de dados multimídia são descritas. Na Seção 2.3 são descritos os aspectos inerentes a modelos de metadados. Na Seção 2.4, a indexação de informação multimídia é descrita. Na Seção 2.5, a recuperação de informação é descrita, enfatizando os métodos de recuperação e os algoritmos que os implementam. Na Seção 2.6, os aspectos para armazenamento de dados são abordados. Finalmente, na Seção 2.7, é concluído o capítulo com algumas considerações gerais sobre o mesmo.

2.2 Conceitos Básicos

Nesta seção são tratados os conceitos básicos inerentes a banco de dados multimídia. Esses conceitos são: multimídia, dado, informação e objeto multimídia.

2.2.1 Multimídia

O termo *Multimídia* é aplicado a sistemas de software ou hardware, que processam várias formas de mídia eletrônica, tal como vídeo, imagem, gráfico, animação e textos, entre outras. A multimídia descreve os tipos de dados que apresentam tanto conteúdo estático como dinâmico [SD99]. No contexto deste trabalho é utilizado o termo multimídia tanto para dados dependentes de tempo (vídeo, por exemplo) quanto para os independentes de tempo (imagem estática, por exemplo).

O termo *mídia* pode ser associado a diferentes significados, como mostra a classificação [SL95] apresentada a seguir:

-
- **Mídia de Percepção** – Este tipo caracteriza-se pela maneira com que os usuários tomam conhecimento da informação contida no ambiente de computação. Essas informações podem ser percebidas através das mídias visuais (por exemplo, imagem) e auditivas (por exemplo, discurso).
 - **Mídia de Representação** – Este tipo caracteriza-se pelo formato das mídias quando representadas no ambiente computacional. Por exemplo, as informações do tipo texto podem ser representadas no formato ASCII, EBCDIC, entre outros.
 - **Mídia de Apresentação** – Este tipo refere-se às ferramentas e dispositivos para efetuar a entrada e saída de dados. Por exemplo, os formulários, monitores de vídeo ou caixas de som são mídias usadas para permitir a saída de informação para o usuário. Já o teclado, mouse, câmara de vídeo e microfones são mídias de coleta de informação dos usuários.
 - **Mídia de Armazenamento** – Este tipo caracteriza-se pela forma como as informações multimídia são armazenadas. Por exemplo, algumas informações são armazenadas através de microfilmagem, em disco rígido ou CD-ROM.
 - **Mídia de Transmissão** – Este tipo caracteriza-se pelos meios físicos de transmissão de dados contínuos de um local para outro. Por exemplo, dados podem ser transmitidos através de uma rede.
 - **Mídia de Intercâmbio de Informação** – Este tipo caracteriza-se pelos dispositivos que são utilizados na troca de informação entre diferentes locais.

Por exemplo, a informação pode passar por mídias de armazenamento, através da transmissão direta, via rede de computadores ou através do uso combinado de mídia de transmissão e armazenamento. Esta última corresponde tipicamente ao sistema de correio eletrônico.

Os sistemas que permitem suporte especial para informações multimídia tratam aspectos de armazenamento, transmissão, manipulação e apresentação de mídias. Esses aspectos têm avançado muito nos últimos anos devido às tecnologias que se tornaram disponíveis, tais como: discos óticos para armazenamento de informações que fornecem grande espaço de armazenamento, redes de alta velocidade para compartilhamento de recursos eficientemente, e estações de trabalhos sofisticadas para processamento e apresentação de mídia.

As aplicações da vida real ficaram mais interessantes de serem abordadas através do uso da multimídia. Por exemplo, o uso de áudio e vídeo em aplicações educacionais permite melhor apresentação de informação aos usuários. Além do mais, a multimídia tem maior impacto em aplicações voltadas para usuários sem nenhum conhecimento de processamento de dados. Algumas dessas aplicações estão relacionadas a: educação, automação de escritório, área médica, bibliotecas, museus e turismo, entre outras.

Na educação, as apresentações possuem tipicamente palavras, figuras, vídeos e vozes; livros possuem palavras e figuras, mas não existem razões para que um livro eletrônico não possua também narrações e animações.

Em automação de escritório, as aplicações incluem ferramentas para criação e recuperação de documentos, manutenção de calendários de compromisso, entre outras. Um banco de dados para sistemas de automação de escritórios deve permitir solicitações pertencentes a agendas, documentos e conteúdo de documentos. As informações podem ser representadas de várias formas, além de serem processadas/armazenadas de várias maneiras diferentes. Por exemplo, mensagens de telefone, fax e documentos.

Na área médica, os registros de pacientes podem consistir de históricos de casos, raios-X, anotações de consultas, resultados de testes, etc. E ainda, várias técnicas recentes de processamento de imagens geram dados digitais ou vídeo, diretamente, tais como: tomografia computadorizada, ressonância magnética e ultrasonografia.

Bibliotecas e museus possuem grandes quantidades de dados da literatura e artefatos arqueológicos que podem, convencionalmente, ser recuperados através de índices manualmente gerados. A recuperação pode ser feita, também, através de um *tour* multimídia, onde amostras (imagens estáticas, som de instrumentos exibidos, etc.) do que o visitante encontrará em cada uma das galerias são mostrados. Assim, ao visitante é permitido selecionar o roteiro através de uma tela sensível ao toque, por exemplo.

Em conferências podem existir aplicações para preparação de relatórios de viagens, de criação de documentos multimídia compostos de dados sobre despesas, documentos adicionais digitalizados (por exemplo, recibos de hotéis e passagens), comentários de voz sobre a conferência e uma amostra de uma ou mais palestras, obtidas através da digitalização de um vídeo.

Em turismo podem existir aplicações que permitam navegar virtualmente através de um museu, conhecer as principais obras, ter comentários de voz das obras, vídeos mostrando a vida dos autores e assim por diante

Enfim, as aplicações multimídia podem ser as mais diversas possíveis e trazem benefícios a todas as áreas do conhecimento e em todos os níveis sociais, econômicos e culturais.

2.2.2 Dado e Informação

Os *dados* são fatos relativos a pessoas, lugares, eventos ou outros objetos e conceitos. Esses dados são armazenados em algum dispositivo. Um dado é uma abstração, podendo ter valor e ser usado de diferentes formas.

A um dado, em geral, são associadas as seguintes características identificação, valor e tempo. O valor dos dados depende da habilidade dos usuários para inteligentemente acessá-los, através de pesquisas baseadas em descrição de conteúdo, habilidade em determinar quais dados são úteis, além de estimar o tempo para recuperação mais eficiente.

O *dado* corresponde a matéria-prima originalmente obtida de uma ou mais fontes e *informação* corresponde ao resultado do processamento e/ou refinamento de um dado (Figura 1).



Figura 1: Dado e Informação

A informação deve ser mostrada em um formato conveniente para fins de tomada de decisão ou para ser utilizada em outras atividades organizacionais. Em outras palavras, informação é um dado com associação semântica.

2.2.3 Objeto Multimídia

Um objeto multimídia é um objeto composto que referencia vários tipos de objetos, tais como vídeo, texto, som, imagem, entre outros. Ele pode ser dividido em partes que contêm textos, imagens, e assim por diante [LZ96,YY97,CSBB97].

Um modelo de objetos multimídia deve possibilitar a captura da estrutura de objeto, incluindo seus atributos, conteúdo, comportamento e funções. Os atributos e conteúdo dos mesmos estão diretamente relacionados. Já o comportamento de um objeto é definido como um conjunto de mensagens que ele entende e responde. As funções de um objeto são definições explícitas do mundo real, representadas no Sistema de Gerenciamento de Banco de Dados.

Além de capturar informações a respeito dos objetos, um modelo deve capturar informações a respeito dos relacionamentos entre os objetos. Agregação e herança são dois relacionamentos populares e largamente utilizados em modelos pósrelacionais. O relacionamento de agregação captura a noção de objetos complexos. O relacionamento de herança captura a hierarquia de classes bem como instancia um tipo de objeto.

A definição de informações a respeito de objetos multimídia pode ser feita através de modelos de metadados. Na seção seguinte este ponto é abordado.

2.3 Modelo de Metadados

Metadado, basicamente, corresponde a uma informação que descreve um dado [AS94, CC98]. Metadados são dados que descrevem atributos de um recurso. Eles dão suporte a um número de funções: localização, descoberta, avaliação, documentação, seleção, entre outras. Metadados são dados associados com objetos que ajudam usuários a conhecerem sua existência ou características. Um Modelo de Metadados fornece o contexto para entender os dados através do tempo.

Os metadados fornecem as ferramentas para transformação dos dados em conhecimento (Figura 2). No enfoque desta tese, o modelo de metadados permitirá transformar objetos multimídia armazenados em conhecimento dos mesmos, através da indexação e recuperação.



Figura 2: De Dados a Conhecimento

Os metadados provêm uma descrição concisa a respeito do dado. O dado pode ser um documento, uma coleção de documentos, gráficos, tabelas, etc. Por exemplo, a definição de uma classe, em orientação a objetos, é um metadado. Os modelos são inerentemente metadados, visto que descrevem os tipos sendo modelados. As aplicações do mundo real, tais como catálogos e dicionários, têm metadados. Implementações de linguagens também utilizam metadados.

Em banco de dados, informações a respeito dos dados são tão importantes quanto os dados. Os Sistemas de Gerenciamento de Banco de Dados Relacionais [SKS99] também utilizam metadados. As tabelas no banco de dados são usadas para armazenar informações. De forma similar, um Sistema de Gerenciamento de Banco de Dados Relacional tem várias meta-tabelas que armazenam definições das tabelas. Assim, uma tabela de dados pode armazenar o fato de que a capital do Brasil é Brasília, a capital da Inglaterra é Londres e a Capital da França é Paris. Uma metatabela armazenaria o fato que um país tem uma cidade como capital.

2.3.1 Classificação de Metadados

No contexto multimídia, metadados podem corresponder a referências a dados, as quais são utilizadas para facilitar a recuperação baseada em descrição de conteúdo. A seguir é descrita uma classificação para metadados[BR94].

- **Metadado para Representação de Tipos de Mídia** – São informações adicionais que são importantes para fins de representação de dados multimídia. Essas informações podem ser utilizadas para apresentação desses dados. Exemplos deste tipo de metadado incluem formato de codificação e técnicas aplicadas para compressão de dados. Na Figura 3, para cada tipo de mídia o metadado **Formato** pode assumir um valor. No caso de vídeo, que pode ser uma seqüência combinada de áudio/vídeo, sua codificação pode ser em diferentes formatos, tais como PAL (*Phase Alternating Line*), SECAM (*Sequential Colour Avec Memoire*) e NTSC (*National Television Systems*

Committee), e armazenados no computador usando um formato MPEG (*Motion Picture Expert Group*).

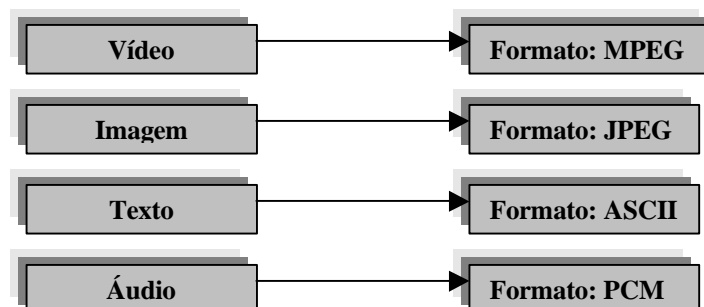


Figura 3: Metadado para Representação de tipos de Mídia

O metadado para formato de imagem pode ser o JPEG (*Joint Photographic Expert Group*). No caso de texto, o formato pode ser o ASCII ou EBCDIC. O metadado para formato de áudio pode ser o PCM (*Pulse Coded Modulation*)

- **Metadado para Descrição de Conteúdo** – Esse tipo de metadado descreve o conteúdo do objeto ou dos componentes do mesmo. A descrição pode ser feita manualmente, de forma automática ou semi-automática, dependendo do tipo de mídia tratada. Um exemplo típico é uma lista de pessoas ou instituições relacionadas ao conteúdo de um objeto multimídia. Essa lista de pessoas é fornecida pelo usuário de forma manual. Vale ressaltar que o objetivo desta tese é criar um modelo de metadados para descrição de conteúdo, com operações que permitam o reuso de descrições de objetos multimídia já efetuadas.
- **Metadado para Classificação de Conteúdo** – A classificação de conteúdo consiste em informações adicionais que podem ser derivadas do conteúdo do documento. Este tipo de metadado pode ser gerado de forma automática ou

semi-automática. Por exemplo, a classificação pode ser feita pelo nível de conhecimento do leitor sobre determinado tema.

- **Metadado para Composição de Documentos** – Componentes lógicos de documentos multimídia têm certa semântica como parte do documento. Metadado para composição de documentos permite conhecimentos de características dos relacionamentos entre componentes de documentos. Este tipo de metadado só faz sentido para documentos multimídia e não para tipos de dados simples. Por exemplo, os modelos de hipermídia podem ter metadados para descrever os componentes e os relacionamentos entre eles.
- **Metadado para História de Documentos** – Em ambientes de publicações multimídia é comum registrar o *status* de dados multimídia. Este tipo de metadado serve tanto para um documento como para componentes individuais, tais como data da última atualização de um determinado documento. Metadado desse tipo pode ser registrado automaticamente ou não. O nome do autor de documentos, data de autoria e os relacionamentos derivados em modelos de versões são exemplos desse tipo de metadado.
- **Metadado para Localização de Documentos** – Os objetos multimídia não são duplicados ou distribuídos, desta forma são acessados de acordo com a necessidade de cada usuário. Metadados são utilizados para localizar documentos. A informação sobre a localização de certos documentos e como esses podem ser acessados são exemplos desse tipo de metadado.

- **Metadado Estatístico** – Este tipo de metadado quantifica a frequência dos documentos ou componentes de documentos com certas características. Os aspectos de desempenho, acessos e frequência de apresentação de documentos em determinado período de tempo também são exemplos de metadado estatístico.

2.3.2 Metadados no Contexto Organizacional

Inicialmente, os metadados eram incorporados aos programas de aplicação. Depois, os Sistemas de Gerenciamento de Banco de dados incluíram metadados em seu dicionário de dados e catálogos. *Data Warehouse* [Inm99,SKS99], por exemplo, pode possuir modelos de metadados que ajudam a mapear dados de diversas fontes e armazenar esses dados sob um esquema único.

Os metadados no contexto da organização estão relacionados em duas categorias: técnicos e de negócio.

- **Metadado Técnico** é a descrição dos dados necessária a várias ferramentas que permitem o armazenamento, manipulação ou movimentação dos dados. Dentro dessa categoria estão todos os tipos de metadados classificados na Seção 2.3.1.
- **Metadado de Negócio** é a descrição de dados necessária aos usuários de negócio para entender o contexto do negócio e o significado dos dados. Por exemplo, a descrição de sistemas de aplicações financeiras. Muitas organizações não fazem o melhor uso de seus dados porque os mesmos não são bem gerenciados. Uma das razões é a falta de conhecimento do significado dos mesmos. A informação gerada não obedece às

mesmas leis econômicas de outros patrimônios. O valor da informação aumenta com o uso e diminui com o tempo. O valor da informação aumenta quando combinada com outra informação. Até hoje a transformação de dados em informação tem sido conduzida por tecnologia.

Os avanços tecnológicos estão sendo integrados pela organização. Esta integração ajuda a gerenciar não só a informação, mas também as formas nas quais as pessoas aprenderão, influenciarão e compartilharão a informação com outros. Com isso, quando o dado é somado com o contexto, o resultado é a informação (Figura 4). Um modelo de metadados vai permitir definir o contexto onde os dados são manipulados. Esse contexto envolve a representação da estrutura e o tempo aplicado a cada tipo de informação.



Dado + Contexto = Informação

Figura 4: Transformação de Dados em Informação

A tecnologia de metadados está surgindo em função das necessidades das organizações conhecerem melhor os dados que elas mantêm e conhecer com mais detalhes os dados de outras organizações. A catalogação dos dados propiciará a maior utilização deles por usuários com múltiplos interesses.

organizações que não documentam seus dados, com o decorrer do tempo, ficam sujeitas à superposição de esforços de coleta e manutenção de seus dados, vulneráveis a problemas de inconsistência e pagam um alto custo pelo não uso ou uso impróprio dessa informação.

2.3.4 Gestão de Metadados

Na estratégia de gerenciamento de metadados é difícil decidir quais metadados devem ser coletados e mantidos. Em geral, uma arquitetura de informação deve ser flexível para permitir aumentar ou diminuir a quantidade de metadados à medida que novas necessidades apareçam.

Alguns dos problemas relacionados à gestão de metadados são:

- **Os metadados tomam uma variedade de formas** – metadados podem ter a forma de atributos, funções, modelos, objetos, entre várias outras.
- **Novos conjuntos de metadados são criados** – cada modelo de metadados pode flexibilizar a criação de conjunto de metadados. Diferentes usuários ou comunidade de usuários podem querer novos conjuntos de metadados.
- **Diferentes usuários podem propor diferentes tipos de metadados** – similar à criação de conjunto de metadados, um gerenciador de metadados deve tornar acessível ao usuário a criação de vários tipos de metadados. Porém, neste contexto deve existir a preocupação de criar metadados com mesma semântica. Por exemplo, migrar dados de diferentes fontes de dados para um *data warehouse*. Os usuários dessas diversas fontes podem criar diferentes tipos de metadados que pode não resultar integração.
- **Existem muitos usuários de metadados** – a grande quantidade de usuários gera grande quantidade de informações que não são conhecidas por todos os usuários. Assim, no gerenciamento de metadados é necessário pensar em

como esses usuários podem compartilhar esses dados. Um modelo de metadados pode permitir o conhecimento dos dados existentes no ambiente.

- **Adoção de diferentes vocabulários de metadados** - significa aumento de buscas usando vocabulários que não são familiares. É necessário permitir que um modelo uniformize a gestão dos metadados e o vocabulário usado pela comunidade de usuários.

As tarefas para criar um ambiente de gestão de metadados são: definir requisitos para metadados que devem estar disponíveis para os usuários; desenvolver a arquitetura de gestão de metadados; selecionar quais as ferramentas devem ser incorporadas na infraestrutura de gestão de metadados; desenvolver os programas que integram as ferramentas selecionadas para atender as necessidades específicas da organização; desenvolver e executar um programa de treinamento para os usuários.

Existem poucas organizações que constroem suas próprias ferramentas de gestão de metadados. A *Web e Data Warehouse*, por exemplo, evidenciam a necessidade da gestão de metadados através da utilização de padrões internacionais e novas ferramentas de gerenciamento. MPEG-7 - *Multimedia Content Description Interface* [Mar01], constitui um padrão para descrição de conteúdo de dados multimídia, o qual dá suporte às interpretações do significado das informações.

No contexto desta tese, um modelo de metadados é proposto para indexação e recuperação de informação multimídia. Para tanto, nas seções seguintes são abordadas a indexação e recuperação de informação.

2.4 Indexação de Informação

A indexação[GFSC94,AS94, Kow97,CC98,BSH00], originalmente chamada catalogação, é a técnica mais antiga para identificação do conteúdo de um dado para sua posterior recuperação. O objetivo da indexação é permitir que uma coleção de dados possa ser recuperada posteriormente pelos usuários da informação.

A indexação permite identificar e representar o conteúdo ou o assunto de um documento de forma a caracterizá-lo. Um documento deve ser indexado tendo-se em mente não apenas o seu armazenamento, mas principalmente sua futura busca e recuperação.

A processo de indexação, em multimídia, pode ser caracterizado como uma atividade subjetiva. Diferentes pessoas, embora treinadas de acordo com os mesmos princípios (critérios), poderão indexar de forma diferente o mesmo objeto.

A Indexação, em sistemas de banco de dados multimídia, pode ser hierárquica e em quatro níveis (Figura 5) [Gro97]: de aplicação, de atributos e valores, de estrutura e de mapeamento entre objetos implementados e o modelo de objetos. No nível de aplicação, existem índices para aplicações multimídia. No segundo nível, os atributos e valores de objetos multimídia são tratados como índices. No nível de estrutura, existem índices para os objetos multimídia propriamente ditos. No nível de mapeamento entre objetos implementados e o modelo de objetos, a indexação é feita entre os objetos codificados e o modelo de objetos multimídia.

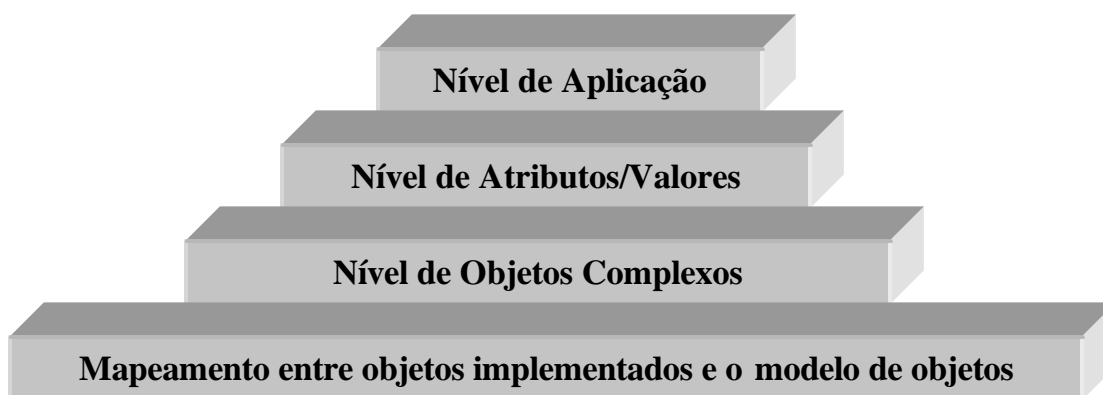


Figura 5: Hierarquia de Indexação Multimídia

O sistema multimídia deverá permitir a construção dos quatro níveis de indexação. Para que recuperações sejam feitas no banco de dados multimídia é necessário que os objetos sejam analisados, de modo que descrições de seu conteúdo possam ser extraídas e armazenadas no banco de dados juntamente com os objetos originais. As descrições são, então, utilizadas para recuperação e permitem determinar quais informações satisfazem os critérios de seleção das consultas ao banco de dados.

Os componentes (Figura 6) que aparecem no processo de indexação são objeto multimídia, indexador, análise, linguagem de indexação, modelo real e objetos armazenados, os quais são descritos a seguir.

A atividade de atribuir metadados a um objeto multimídia é sempre precedida da análise conceitual do mesmo, usando a linguagem de indexação. Tal linguagem possui metadados/valores que podem ser associados aos objetos e funções que os manipulam. Os objetos multimídia podem ser descritos tendo como referência um conjunto de metadados,

os quais podem ser extraídos manualmente ou mecanicamente, sem interferência do usuário.

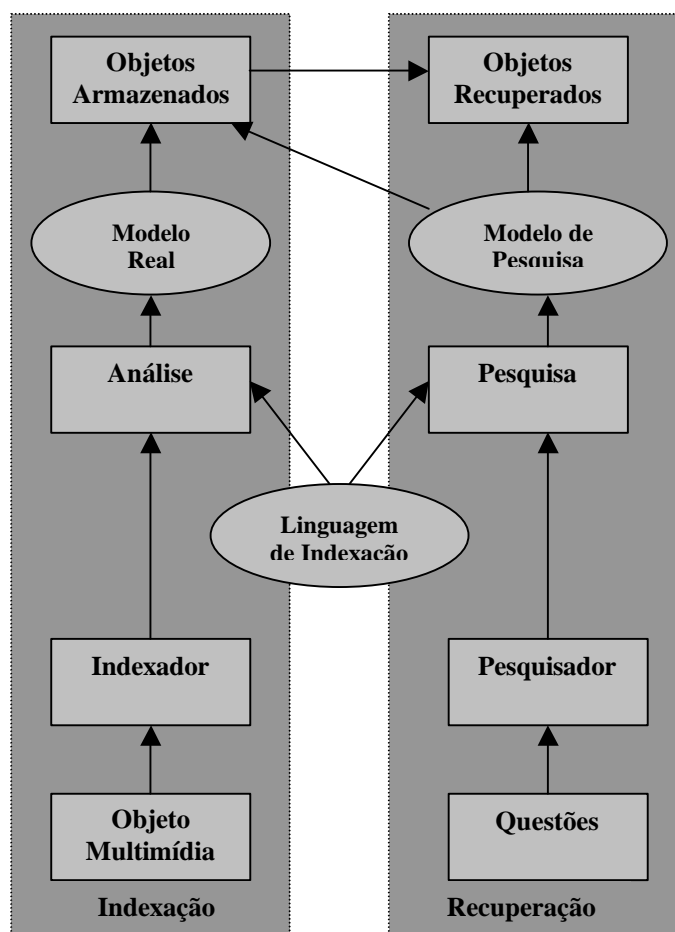


Figura 6: Indexação de Informação Multimídia

A partir da associação de metadados, é construído o modelo de indexação, que será tanto melhor quanto maior for sua fidelidade ao modelo real. Quando a consulta for efetuada, o usuário deve construir um modelo de pesquisa, seguindo o mesmo modelo de indexação.

No processo de recuperação, aparecem os seguintes componentes: questões, pesquisador, pesquisa, modelo de pesquisa e objetos recuperados.

Quando se pretende recuperar um objeto multimídia, o processo é bastante similar. Primeiramente, tem que ser decidido, através da análise conceitual, o que recuperar. Em seguida, tem que ser feita a tradução do modelo obtido para um modelo passível de processamento pelo sistema de recuperação de informação que esteja disponível.

Para esta tradução é indispensável utilizar os metadados da mesma linguagem que foi utilizada durante o processo de indexação dos objetos. É visível, portanto, que a linguagem de indexação é uma ferramenta comum ao indexador e ao pesquisador.

2.5 Recuperação de Informação

Um dado armazenado dentro de uma base deve ser estruturado de tal modo que a informação desejada possa ser recuperada. A modelagem de dados é o termo utilizado para determinar quais informações devem ser armazenadas na base e quais devem refletir os relacionamentos entre os itens de dados. Tanto os usuários como os projetistas de aplicações de banco de dados necessitam conhecer como os dados são modelados.

A recuperação é influenciada por fatores tais como sincronização, linguagem de consulta, indexação, estratégias de pesquisa e recuperação baseada em conteúdo. Os objetos multimídia complexos impõem restrições de sincronização no armazenamento e recuperação de informação multimídia. Estratégias de recuperação de dados devem considerar restrições de sincronização quando executando uma consulta. Para a linguagem deve ser considerado que um usuário, no momento da pesquisa, pode não ter uma estratégia bem definida.

As estratégias para sistemas multimídia não devem ser projetadas somente para casamento exato, mas também para aproximado. As pesquisas devem utilizar medidas de similaridade como construção de blocos fundamentais. Estas medidas de similaridade podem ser computadas em mais de uma dimensão. O resultado da consulta deve ser uma lista classificada em função da medida definida.

A recuperação de informação baseada no conteúdo caracteriza-se pelos dispositivos que são utilizados na troca de informações entre diferentes mídias de armazenamento. Por exemplo, na passagem entre dispositivos óticos e magnéticos, o armazenamento tem organizações distintas e assim, diferentes modelos devem ser construídos. Cada dispositivo, ainda, tem seu próprio método de acesso, cuja combinação introduz mais um fator de complexidade no processamento de consultas.

Nas seções seguintes são descritos métodos de recuperação de informação multimídia e os algoritmos de pesquisa que implementam tais métodos.

2.5.1 Métodos de Recuperação

A maneira pela qual as informações são armazenadas e descritas é tratada de acordo com o modelo de recuperação adotado, porém vale salientar que a eficiência na recuperação depende dos tipos de dados tratados, de como se pretende manter as informações utilizadas para descrever o conteúdo de forma correta, dos modelos de recuperação permitidos às informações e dos algoritmos de pesquisa implementados para tais métodos.

Na recuperação de informação, um objeto multimídia é representado como uma coleção de aspectos. O usuário especifica o que deseja na forma de uma consulta. Dada a

representação da informação do usuário e a coleção de objetos, o sistema estima a semelhança dos dados fornecidos pelo usuário e as representações dos objetos armazenados.

Existem vários métodos de recuperação de informação textual na literatura, tais como [SKS99]: varredura completa de texto, inversões *eclustering*, entre outros.

- **Varredura Completa de texto** – O usuário entra com um padrão a ser pesquisado e o sistema varre o banco de dados até que um texto igual seja recuperado. Apesar da existência de algoritmos de pesquisa de *strings* [CLR90] muito rápidos, varrer um banco de dados pode levar muito tempo.
- **Método de Inversões** – O método de inversões é um dos mais rápidos e adequados para ambientes estáticos, ou seja, onde haja muitas consultas, porém nenhuma ou poucas modificações. Nesse método, cada entrada consiste de um padrão e uma lista de endereços para os objetos onde o padrão ocorre. O padrão pode ser implementado como uma palavra e os endereços podem ser a localização exata da ocorrência ou a identificação do objeto.
- **Clustering** – corresponde ao agrupamento de objetos similares formando *clusters*. Neste método pode ser necessária a reorganização da estrutura de *clusters* quando acontecer inclusão de novos objetos.

Existem vários modelos de recuperação de objetos multimídia, entre os quais podem ser citados: recuperação através de um identificador [BPJ93, OT97], por sentenças

condicionais [BPJ93, OT97], por similaridade [BPJ93, PF95] e recuperação semântica [KKH94, VAB95].

- **Recuperação Através de um Identificador** – Na recuperação através de um identificador, esse é dado como índice e o objeto é recuperado.
- **Recuperação por Sentenças Condicionais** – os objetos são recuperados somente se eles respondem precisamente às restrições expressas na consulta.
- **Recuperação por Similaridade** – No caso da recuperação de informação por similaridade não é necessário que uma informação seja exatamente igual à da consulta. Ao invés disso, a pesquisa ao banco de dados deve ser aproximada, para que todas as informações sejam recuperadas até o grau de similaridade pré-especificado. Este método de recuperação corresponde a um processo de casamento onde a relevância dos objetos é julgada de acordo com a similaridade baseada nas entradas de cada objeto e nas entradas de cada consulta. Petrakis e Faloutsos [PF95] criaram um método para efetuar pesquisa por conteúdo baseada em similaridade, em banco de dados de imagens. O método parte do princípio que existe um número fixo de características comuns que são esperadas para todas as imagens e um número variável de características não definidas antecipadamente (não esperadas). O método pode responder às consultas efetuadas através de exemplos, tais como: *“Obter todas as radiografias que são similares à radiografia de José”*.

- **Recuperação Semântica** – O usuário expressa a informação necessária como um conceito e não como uma coleção de palavras-chave conectadas através de operadores lógicos. Esse método tenta prover mais poder de abstração e mecanismos de estruturação para especificação das consultas.

Os dois primeiros são recuperações simples que qualquer sistema permite, porém os dois últimos são mais complexos visto que necessitam de algoritmos especiais.

2.5.2 Algoritmos de Pesquisa

No processamento de informações multimídia, muitas vezes, os dados não podem ser decompostos em registros independentes, com pequenas partes identificáveis. Assim, os tipos multimídia podem ser descritos através de *strings*, ou seja, seqüência de caracteres.

Strings são, obviamente, centrais em sistemas de processamento de palavras, os quais possibilitam inúmeras funções para a manipulação de texto. Estes sistemas podem manipular objetos extremamente grandes, tais como um livro com milhões de caracteres. Como conseqüência, algoritmos eficientes são de grande importância para manipular tais objetos.

Em multimídia, a recuperação de dados pode recair na área de busca e comparação de seqüências. Um problema clássico nesta área é o casamento (exato) de padrões (*string matching*) [Mel96]. Tal problema consiste em, a partir de um *string* texto de tamanho n e um padrão de tamanho m , obter uma ou todas as ocorrências do padrão no texto.

Os principais algoritmos de busca através de strings [Mel96] são: KMP, Boyer Moore e algoritmo de *string-matching* via autômatos.

2.5.2.1 Algoritmo KMP

O algoritmo KMP (*Knuth-Morris-Prat*) compara um padrão com um texto, caractere a caractere. Porém, quando caracteres de posições correspondentes diferem, ou seja, quando ocorre um *mismatch*, o KMP aproveita a informação que os caracteres do padrão (ou parte dele) casaram com o texto, evitando assim comparações redundantes.

O algoritmo KMP possui uma fase de pré-processamento, onde é guardado o tamanho do maior prefixo do padrão casado, que é igual ao sufixo desse padrão casado, com o texto. Esse tamanho é computado e armazenado numa tabela. O padrão pode então ser deslocado através da diferença entre o tamanho do padrão e o tamanho do prefixo, sem que nenhum casamento seja perdido e sem efetuar comparações desnecessárias.

Esse algoritmo nunca utiliza mais que a soma do tamanho do padrão e o tamanho do texto para comparações de caracteres. O custo de tempo do KMP, para o pior caso, é de $O(n+m)$, incluindo a fase de pré-processamento. Portanto, é uma boa escolha para o problema de *string matching* no caso em que o alfabeto possua uma quantidade grande de elementos.

2.5.2.2 Algoritmo Boyer-Moore

O algoritmo Boyer-Moore compara o padrão com o texto, da direita para a esquerda, ou seja, a primeira comparação é feita entre o último caractere do padrão e o caractere do texto correspondente à última posição do padrão. Se estes são iguais, então os caracteres

da posição precedente do padrão e do texto são comparados, e assim por diante. Ou seja, a busca por ocorrências do padrão no texto é feita normalmente, da esquerda para a direita, apenas a comparação é feita da direita para a esquerda.

Na ocorrência de um *mismatch*, o algoritmo move o padrão para a direita, escolhendo o maior dos deslocamentos propostos por duas heurísticas: mau caractere e bom sufixo. O deslocamento do padrão é efetuado com base nas heurísticas que na ocorrência de um caractere do padrão diferente do respectivo caractere do texto, buscam pela ocorrência mais à direita do caractere do texto que provocou o *mismatch* – na heurística do mau caractere – ou aproveitam as informações de que um sufixo do padrão casou com o texto, no caso da heurística do bom sufixo.

Esse algoritmo utiliza mais que a soma do tamanho do padrão N com o tamanho do texto M de comparações de caracteres. O algoritmo utiliza N/M passos se o alfabeto não for pequeno e o padrão não for longo. Portanto, esse algoritmo é a melhor escolha quando textos de tamanho grande são considerados.

No Boyer-Moore, no pior caso, o custo de tempo é quadrático, porém para casos práticos, ele é sublinear.

2.5.2.3 Algoritmo de String Matching via Autômatos

Os algoritmos anteriores fazem uso de pré-processamento, a fim de reduzir o tempo de execução. Outro modo de efetuar tal pré-processamento é através da construção de autômatos finitos.

Um algoritmo de *string matching* via autômatos examina cada caractere do texto exatamente uma vez. Um conjunto finito de estados é rotulado de forma ordenada. O primeiro e o último elemento desta ordem são os estados inicial e final do autômato.

O alfabeto de entrada é o mesmo sobre o qual estão definidos o padrão e o texto. No autômato, números são associados aos estados. Alcançar um determinado estado indica que houve casamento dos primeiros caracteres, correspondentes ao número desse estado, do padrão com o texto. Em outras palavras, a transição para um estado n é realizada apenas se ocorrer casamento com $n-1$ caracteres do texto.

O custo de tempo de um algoritmo de *string matching* via autômatos, após a construção do autômato, para o pior caso, é de $O(n)$.

2.6 Armazenamento de Objetos

Atualmente, podem existir dados multimídia armazenados fora do banco de dados, em sistemas de arquivos. Quando o número de objetos multimídia é relativamente pequeno, os recursos fornecidos por um banco de dados não são importantes.

A funcionalidade e integridade do banco de dados tornam-se importantes quando é grande o número de objetos multimídia armazenados. Por exemplo, um filme pode ter três horas de duração e pode necessitar de vários *gigabytes*, mesmo depois de compactado.

As informações multimídia podem ser armazenadas em um sistema de banco de dados por diferentes mecanismos. A seguir são apresentados quatro tipos, os quais podem aparecer, juntos ou não, em um sistema de gerenciamento de banco de dados [RNL95].

- **Referências Externas**

Os bancos de dados possuem referências para os objetos multimídia originais. Tais referências podem ser nomes de arquivos ou outros identificadores que permitem localizar o dado no ambiente. Adicionalmente, as referências contêm dados descritivos modelados como atributos ou relacionamentos. Por exemplo, atributos de uma fita de vídeo [JH94] podem ser seu comprimento, o dispositivo de saída aplicado, seu formato de compressão e a descrição textual do conteúdo. Vale ressaltar que o protótipo do modelo MetaMídia proposto nesta tese implementa esse tipo de mecanismo.

- **Campos Longos**

Os campos longos podem armazenar vários *Gbytes* de dados e podem ser usados como domínio para atributos. Essencialmente, podem ser acessados parcialmente. Um sistema de gerenciamento de banco de dados dá suporte para dados armazenados em campos longos. Entretanto, o conteúdo ainda não é interpretado e as operações oferecidas para operar nos dados são somente genéricas.

- **Funções Externas**

Alguns bancos de dados permitem chamar funções externas para processamento de dados armazenados. As limitações em linguagem de manipulação de dados, tal como SQL, são razões para introduzir chamada a funções externas. O suporte multi-usuário e autorização podem ser aplicados

para o acesso a funções externas, mas sua execução não pode ser controlada por um Sistema Gerenciador de Banco de Dados. Por exemplo, se um áudio é acionado em um dispositivo usando uma função externa, o acesso exclusivo não pode ser garantido pelo sistema de banco de dados. Porém, as funções externas são adequadas para reuso de algoritmos existentes e ferramentas no contexto de captura e apresentação multimídia.

- **Sistemas Orientados a Objetos ou Relacionais Estendidos**

Sistemas relacionais estendidos permitem que o programador defina tipos de dados abstratos e os referencie em aplicações. Especialmente, a abordagem orientada a objetos dá suporte à modelagem avançada, pela construção de hierarquia de tipos de dados e definição de relacionamentos entre classes. Em alguns sistemas, mecanismos de indexação específicos podem ser estabelecidos por tipos de dados definidos pelos usuários. Os sistemas orientados a objetos oferecem suporte mais adequado para informação multimídia, mas faltam alguns aspectos tais como suporte a dados dependente de tempo, interação do usuário, consultas baseadas em conteúdo e técnicas de recuperação.

Muitos sistemas de banco de dados dão acesso ao usuário do banco de dados que reside em uma máquina remota. Assim, a distribuição é permitida se a informação multimídia é armazenada em campos longos ou se um sistema orientado a objetos ou relacional estendido é usado.

2.7 Considerações Finais

Os sistemas de banco de dados multimídia gerenciam grandes quantidades de dados, que podem ser usados para dar suporte às aplicações médicas, aos sistemas de informação geográfica, aos projetos de arquitetura, às bibliotecas fotográficas, às galerias de artes, aos museus, às investigações criminais, aos sistemas educacionais, entre uma série de outras aplicações.

Esses tipos de aplicações motivaram, nos últimos anos, pesquisas nos aspectos relacionados aos sistemas de gerenciamento de banco de dados, tais como: técnicas de modelagem, indexação e recuperação de objetos multimídia.

As informações multimídia podem ser representadas na sua forma original, acrescidas de informações descrevendo o seu conteúdo. Essas informações são definidas usando metadados.

No capítulo seguinte é apresentado um modelo de metadados na indexação e recuperação de informação multimídia. Os trabalhos relacionados com esta tese são apresentados no capítulo seguinte ao modelo de metadados, e é feita uma análise comparativa do modelo proposto com os sistemas existentes. A evolução de modelos na última década é apresentada no final, para enfatizar a relevância, originalidade e abrangência do modelo proposto. O protótipo do mesmo é descrito no Capítulo 4.

Capítulo 3

MetaMídia – Um Modelo de Metadados para Objetos Multimídia

*Existe o mundo das idéias
e o mundo da prática.*

Mathew Arnold

Neste capítulo é apresentada a principal contribuição desta tese. A arquitetura e os aspectos inerentes ao modelo de metadados multimídia são definidos. A indexação, o método de recuperação e o algoritmo que implementa tal método são também apresentados.

3.1 Introdução

Um banco de dados multimídia contém imagens gráficas, clipes de vídeo, arquivos de som, textos, entre outros. E assim, o objeto armazenado no banco de dados multimídia deve ser descrito de tal modo que o mesmo possa ser facilmente localizado e recuperado.

A indexação é usada para determinar quais dados devem ser armazenados para descrever os objetos multimídia e permitir a pesquisa e recuperação dos mesmos.

As pesquisas realizadas pelos usuários de banco de dados convencionais tendem a ser altamente estruturadas, enquanto usuários de sistemas de banco de dados multimídia podem desejar realizar consultas baseadas na descrição do conteúdo de um objeto multimídia.

Esta tarefa pode ser difícil se não existirem modelos que permitam a indexação de objetos multimídia no momento de sua criação. No caso de um processo demorado para identificar uma característica, pode ser indexado após a criação. Assim, nesta tese é apresentado um modelo de metadados para auxiliar no processo de indexação e recuperação de objetos multimídia.

Uma das vantagens obtidas com a definição do modelo de metadados é o ganho em termos de tempo, visto que não será necessário acessar inúmeros objetos até encontrar o objeto desejado. Outra vantagem é que no uso de descrições de objetos existentes, as informações desses objetos são herdadas pelo novo objeto, permitindo o compartilhamento de metadados/valores entre meta-objetos multimídia.

Para apresentar tal modelo de metadados, o capítulo é estruturado como segue. Na Seção 3.2, é apresentada a estrutura do modelo de metadados. Na Seção 3.3, é apresentada a indexação de informação multimídia. Na Seção 3.4, são apresentados os aspectos relacionados com a recuperação de informação. Finalmente, na Seção 3.5, são apresentadas as considerações gerais do capítulo. No Capítulo 4 são descritos e analisados alguns trabalhos relacionados. Também são apresentados no Capítulo 4, uma análise

comparativa do *MetaMídia* com esses trabalhos e um estudo evolutivo desses trabalhos, na última década.

3.2 Modelo de Metadados

O processo tradicional de indexação envolve determinar os dados que serão armazenados e os relacionamentos que existem entre eles. As tabelas de banco de dados tradicionais podem ser usadas para descrever os dados formatados.

Atualmente, existem duas abordagens para modelo de dados multimídia. A primeira considera que as técnicas e metodologias usadas em indexação convencional são estendidas para banco de dados multimídia. A segunda reconhece a natureza particular dos dados multimídia, como um objeto complexo. Para tanto, podem existir informações associadas com tais objetos que são classificadas como segue:

- **O próprio objeto**, que é composto de outros objetos. Um usuário pode visualizar todo ou somente parte do objeto. Por exemplo, uma seqüência de vídeo é um agregado de objetos dos tipos vídeo, áudio e texto.
- **As características e propriedades que definem o objeto**, tais como o formato do registro, duração de vídeo, tamanho do frame e técnicas de compressão.
- **Atributos que descrevem o objeto complexo e seus relacionamentos com outros objetos do banco de dados**, muitas vezes descritos como atributos abstratos, tais como título, diretor, data da produção e estúdio de produção.

Adicionalmente, operações/métodos sobre os objetos multimídia são definidas e essas operações diferem daquelas que operam sobre os dados tradicionais mantidos dentro de

um banco de dados. Os objetos multimídia são objetos do mundo real, tais como cliques de vídeo ou imagens gráficas. Em bancos de dados convencionais são conceitos abstratos que descrevem objetos do mundo real. Por exemplo, o banco de dados de empregado mantém as informações a respeito de empregados de uma organização e não o próprio empregado. O banco de dados multimídia mantém as informações do mundo real e não atributos dos mesmos.

Tendo em vista o exposto, nesta tese é proposto um modelo de metadados que permite indexar (descrever) objetos multimídia com o objetivo de recuperá-los. A arquitetura do modelo de metadados (Figura 7) é composta por três módulos: hierarquia de metadados/valores, esquema de metadados e os objetos multimídia propriamente ditos.

O primeiro módulo é a hierarquia de metadados/valores que corresponde à expressão semântica e pragmática de aplicações particulares. Conceitos são definidos como metadados/valores pelos usuários. Neste contexto, o usuário definirá os tipos de metadados e valores que fazem parte da aplicação.

O segundo, esquema de metadados, é criado para os objetos multimídia indexados. Os metadados/valores são criados para os objetos, gerando os metaobjetos multimídia. O terceiro módulo corresponde aos objetos multimídia armazenados e descritos para futura busca e recuperação.

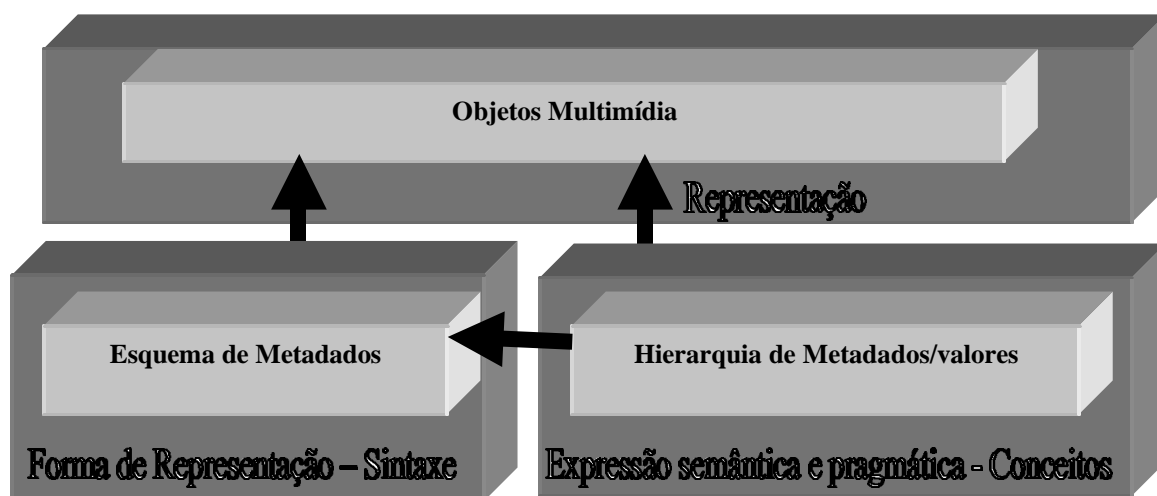


Figura 7: Arquitetura do Modelo de Metadados

Nas subseções seguintes é apresentada a semântica dos elementos que compõem o modelo de metadados: hierarquia de metadados/valores, metaobjeto multimídia, operações básicas, herança em uma hierarquia de generalização/especialização. Esses elementos são definidos através de um formalismo matemático [Oom94], baseado na teoria dos conjuntos.

3.2.1 Hierarquia de Metadados/Valores

A hierarquia de metadados/valores, apresentada nesta tese, é uma especificação de conceitos, isto é, uma descrição dos conceitos e relacionamentos que podem existir para um conjunto de objetos multimídia.

A generalização de uma representação para vídeo, som, imagem e texto, pode ser feita em um domínio de aplicação. Para apresentar o *MetaMídia*, um conjunto de metadados/valores é apresentado (Tabela 1), como exemplo, para uma agência de Publicidade, que trabalha com todos esses tipos de objetos multimídia.

M	Metadados	Valores
M ₁	Localização	Central, Esquerda, Direita
M ₂	Tipo de Imagem	Paisagem, Objetos, Retrato, Nu
M ₃	Classe da Imagem	Realismo, Pictorialismo
M ₄	Coloração	Colorida, Preto e Branco
M ₅	Profissional	Experiente, Amador
M ₆	Formato do Filme	35mm, Slide
M ₇	Perspectiva	Aérea, Terrestre, Close
M ₈	Orientação	Vertical, Horizontal
M ₉	Ângulo	Baixo, Nível do Olho, Alto
M ₁₀	Ambiente	Interior, Exterior, Nublado
M ₁₁	Textura	Randômica, Regular, Orientada
M ₁₂	Cor Predominante	Azul, Vermelho, Verde
M ₁₃	Tempo do Frame	A que ponto pertence a história
M ₁₄	Classe de vídeo	Esporte, documentário, romance
M ₁₅	Tipo de vídeo	Tem objetos, não tem objetos
M ₁₆	Luminosidade do ambiente	Azul, vermelha, ...
M ₁₇	Cor do objeto	Verde, azul, amarelo, ...
M ₁₈	Palavras-chaves de áudio	Início, fim
M ₁₉	Informação de produtor	Nome
M ₂₀	Data de produção	Data
M ₂₁	Mídia original	Filme, vídeo
M ₂₂	Título	Do Filme, da música, do artigo, ...
M ₂₃	Tipo de Peça	Outdoor, spot, vídeo, anúncio, jingle

Tabela 1: Metadados Descritivos

A representação tabular (Tabela 1) é utilizada para melhor evidenciar os metadados/valores sugeridos para a aplicação, os quais são nomeados como descritivos.

Esses metadados são gerados manualmente pelo usuário. As colunas correspondem a: (i) **M** é a designação simbólica dos metadados que são utilizados na construção da seqüência de caracteres (strings) para a descrição de cada objeto; (ii) os metadados que podem ser utilizados na aplicação; e (iii) alguns valores que esses metadados podem assumir (entre outros).

Na Tabela 1, o M_1 representa o metadado *Localização* e possui os seguintes valores: *Central*, *Esquerda* e *Direita*. Assim, a lista $M = [M_1, M_2, \dots, M_n]$, onde o n variará em função do número de metadados definidos para a aplicação, corresponde aos metadados. Os valores dos metadados são listados da seguinte forma: $M_1 = [V_{11}, V_{12}, \dots, V_{1m}]$, $M_2 = [V_{21}, V_{22}, \dots, V_{2m}]$, e assim por diante, onde o m variará em função do número de valores assumidos para cada metadado.

Os metadados/valores, que aparecem na Tabela 1, são colocados em uma hierarquia (Figura 8) representada por um grafo acíclico dirigido [SKS99]. Neste tipo de grafo, um metadado/valor pode ter mais de um metadado de mais alto nível. Por exemplo, o valor **Azul** tem dois metadados de mais alto nível **Luminosidade do ambiente** e **Cor**.

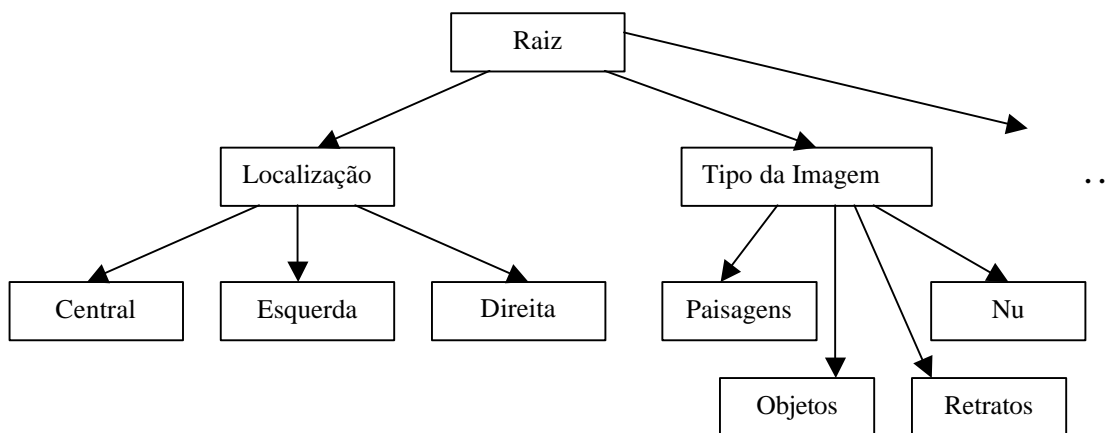


Figura 8: Hierarquia de Metadados/Valores

Assim, por exemplo, o metadado *localização*, pode ser visto como um objeto genérico, dos valores: *central*, *esquerda* e *direita*. O metadado *raiz* representará a aplicação multimídia que pode ser descrita com os metadados/valores definidos

3.2.2 Meta-Objeto Multimídia

Um meta-objeto multimídia (Figura 9) consiste no identificador de objeto multimídia (Id_o) e em uma lista de metadados/valores $d = [m_1:v_1, \dots, m_n:v_n]$, que descreve o objeto multimídia. Assim, os meta-objetos multimídia correspondem às descrições de objetos multimídia. Esses objetos podem descrever objetos do tipo vídeo, imagem, som, texto e documento.



Figura 9: Meta-Objeto Multimídia

Os valores dos metadados tanto podem ser um dado simples como um dado complexo. Esse dado pode ser um conjunto de valores e um meta-objeto multimídia definido recursivamente. Na Definição 1 é apresentado o meta-objeto multimídia formalmente representado.

Definição 1: Meta-Objeto Multimídia

Um meta-objeto multimídia é uma tupla (id_o, d) , onde id_o é o identificador do objeto multimídia, e d é uma lista $[m_1:v_1, \dots, m_n:v_n]$, onde cada m_i ($1 \leq i \leq n$) é um nome de metadado e v_i é o valor do mesmo. O valor v_i pode ser do tipo atômico (números ou *strings*), do tipo conjunto ou do tipo lista definido recursivamente.

Dados dois meta-objetos multimídia $Obj_1 = (id_1, d_1)$ e $Obj_2 = (id_2, d_2)$, com seus metadados/valores $d_1 = [m_{11}:v_{11}, \dots, m_{1n}:v_{1n}]$ e $d_2 = [m_{21}:v_{21}, \dots, m_{2m}:v_{2m}]$, $metadados(d_1)$ denota o conjunto de metadados em d_1 , e $valor(d_1)$ denota os valores de metadados do meta-objeto multimídia d_1 . $Concat(d_1, d_2)$ denota a concatenação da descrição d_1 com a descrição d_2 dos meta-objetos multimídia Obj_1 e Obj_2 . $d_1.m_i$ e $Obj_1.m_i$ denotam o valor v_i .

Como dito na seção anterior, os metadados descritivos apresentados na Tabela 1 são aqueles gerados (criados) manualmente pelo usuário na operação de criação (apresentada na Seção 3.2.3). Além desses metadados existem os metadados gerados automaticamente (dinamicamente) nas operações de disjunção (apresentada na Seção 3.2.4) e conjunção (apresentada na Seção 3.2.5). Esses metadados são aqueles relacionados à criação simples, por conjunção ou por disjunção de metaobjetos multimídia (Tabela 2). Na criação simples, o metadado Criação assume o valor **CR**. Quando o meta-objeto multimídia é gerado a partir de uma conjunção, então o valor do metadado assume o valor **CO**. Quando é gerado a partir da disjunção o metadado Disjunção recebe o valor **DI**. Na Tabela 2 são apresentados esses metadados e seus respectivos valores.

M	Metadados	Valores
M ₁	Criação	CR
M ₂	Conjunção	CO
M ₃	Disjunção	DI

Tabela 2: Metadados Estruturais

Em resumo, os metadados estruturais são gerados a partir de cada operação aplicada aos meta-objetos multimídia. Esses metadados podem ser usados pelas operações de

conjunção e disjunção na pesquisa e recuperação de objetos multimídia. Nas seções seguintes são apresentadas essas operações.

3.2.3 Criação de Meta-Objeto Multimídia

A criação corresponde à descrição (meta-objeto multimídia) de um objeto multimídia dado (*Obj*). Cada objeto multimídia no banco de dados deve possuir um metaobjeto multimídia associado. Existem três formas de criação de meta-objetos multimídia: (i) descrição de objeto multimídia (Figura 10); (ii) disjunção de metaobjetos multimídia existentes (Seção 3.2.4); (iii) conjunção de metaobjetos existentes (Seção 3.2.5).



(Obj₁, [localização: central,
nome: Rafaela,
cor: cinza])

Figura 10: Criação de Meta-Objeto Multimídia

Na Figura 10, o objeto multimídia *Obj₁* é descrito através dos metadados **localização**, **nome** e **cor**, que possuem os valores **central**, **Rafaela** e **cinza**, respectivamente.

A criação de um meta-objeto multimídia a partir de um objeto multimídia ocorre quando um sistema multimídia armazena um objeto. Assim, o usuário vai escolher os metadados/valores para descrever tal objeto e neste momento é criado o metaobjeto multimídia. Nas seções seguintes são descritas as demais operações.

3.2.4 Disjunção de Meta-Objetos Multimídia

A disjunção consiste na criação de um meta-objeto multimídia a partir de meta-objetos existentes (Figura 11). Na edição de objetos multimídia, os valores dos metadados dos meta-objetos existentes são unidos (agregados) em um único meta-objeto multimídia.

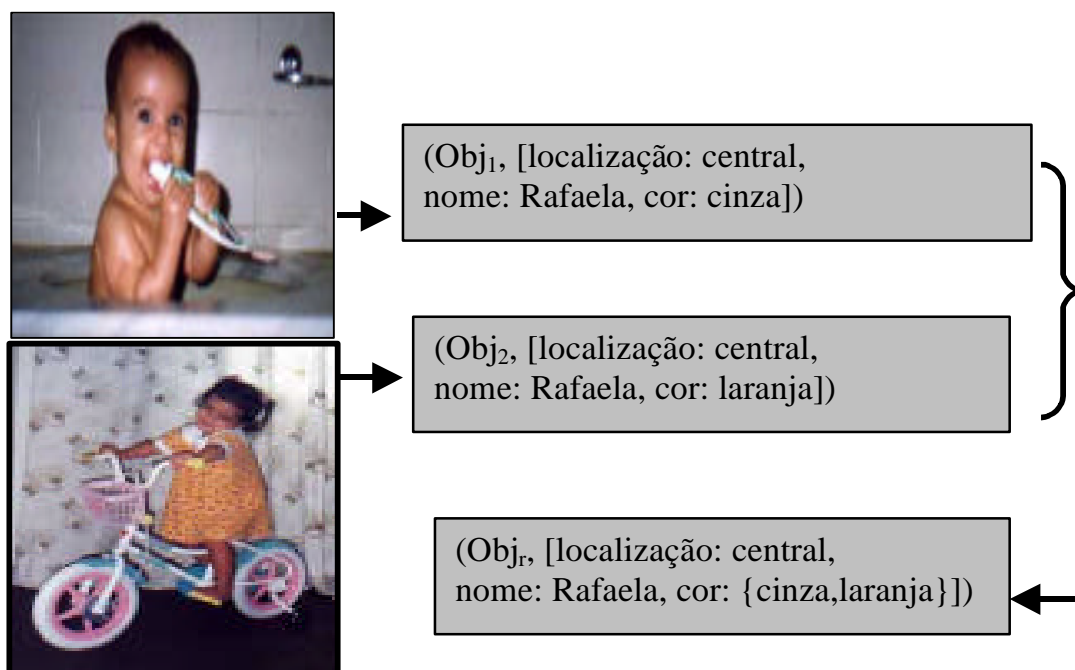


Figura 11: Disjunção de Meta-Objetos Multimídia

Essa operação é usada para derivação automática de descrições. No processo de recuperação de informação, esta descrição será de suma importância visto que não precisará recuperar todos os objetos multimídia, mas sim os que aparecem na composição.

Os metadados dos objetos são agregados e definidos para o novo objeto. O valor associado a cada metadado no novo objeto é dado em função de cada valor de metadado nos objetos existentes.

A disjunção cria a descrição do objeto, a partir dos objetos existentes, através de regras de geração simples. Se um meta-objeto possui metadados que não aparecem nos

outros meta-objetos, então o mesmo será gerado para o objeto resultante. Se existir o mesmo metadado nos dois objetos com valores diferentes, os valores do metadado serão atribuídos para o objeto resultante em um valor do tipo conjunto.

Definição 2: Disjunção de Meta-Objetos Multimídia

A operação disjunção entre meta-objetos multimídia $Obj_1 = (Id_{o1}, d_1)$ e $Obj_2 = (id_{o2}, d_2)$, denotada por $Disj(Obj_1, Obj_2)$, corresponde ao meta-objeto multimídia $Obj_r = (id_r, d_r = Disj(d_1, d_2))$, tal que $d_r = [m_1:v_1, \dots, m_n:v_n]$, onde $m_i \in metadados(d_1 \cup d_2)$ e $(1 \leq i \leq n)$, corresponde a:

1. Para todo $i=1,2,\dots,n$, se $m_i \in metadados(d_1)$ e $m_i \in metadados(d_2)$ fazer:

a) Se $d_1.m_i = d_2.m_i$ então $Concat(d_r, [m_i: d_1.m_i])$

b) Se $d_1.m_i \neq d_2.m_i$ então

i. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo atômico então $Concat(d_r, [m_i: \{d_1.m_i, d_2.m_i\}])$

ii. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo conjunto então $Concat(d_r, [m_i: d_1.m_i \dot{\cup} d_2.m_i])$

iii. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo lista então $Concat(d_r, [m_i: Disj(d_1.m_i, d_2.m_i)])$

iv. Se $d_1.m_i$ é um valor do tipo atômico e $d_2.m_i$ é um valor do tipo conjunto então $Concat(d_r, [m_i: \{d_1.m_i\} \cup d_2.m_i])$

-
- v. Se $d_1.m_i$ é um valor do tipo conjunto e $d_2.m_i$ é um valor do tipo atômico então $Concat(d_n, [m_i: d_1.m_i \cup \{d_2.m_i\}])$
2. Para todo $i=1,2,\dots,n$, $m_i \in metadados(d_1)$ e $m_i \notin metadados(d_2)$ então $Concat(d_n, [m_i: d_1.m_i])$.
3. Para todo $i=1,2,\dots,n$, se $m_i \notin metadados(d_1)$ e $m_i \in metadados(d_2)$ então $Concat(d_n, [m_i: d_2.m_i])$.

A operação de disjunção permite que as descrições dos objetos sejam agregadas e definidas para o novo objeto. Esta operação permite que meta-objetos multimídia sejam concatenados.

Vale salientar que, todos os metadados/valores existentes nos meta-objetos multimídia são agrupados e passam a ser os metadados/valores do meta-objeto multimídia resultante. Na existência de mais de dois meta-objetos multimídia, a operação de disjunção é reaplicada quantas vezes forem necessárias.

3.2.5 Conjunção de Meta-Objetos Multimídia

A conjunção cria um novo meta-objeto a partir de meta-objetos existentes, onde o novo meta-objeto corresponde à interseção dos metadados/valores dos objetos existentes (Figura 12).

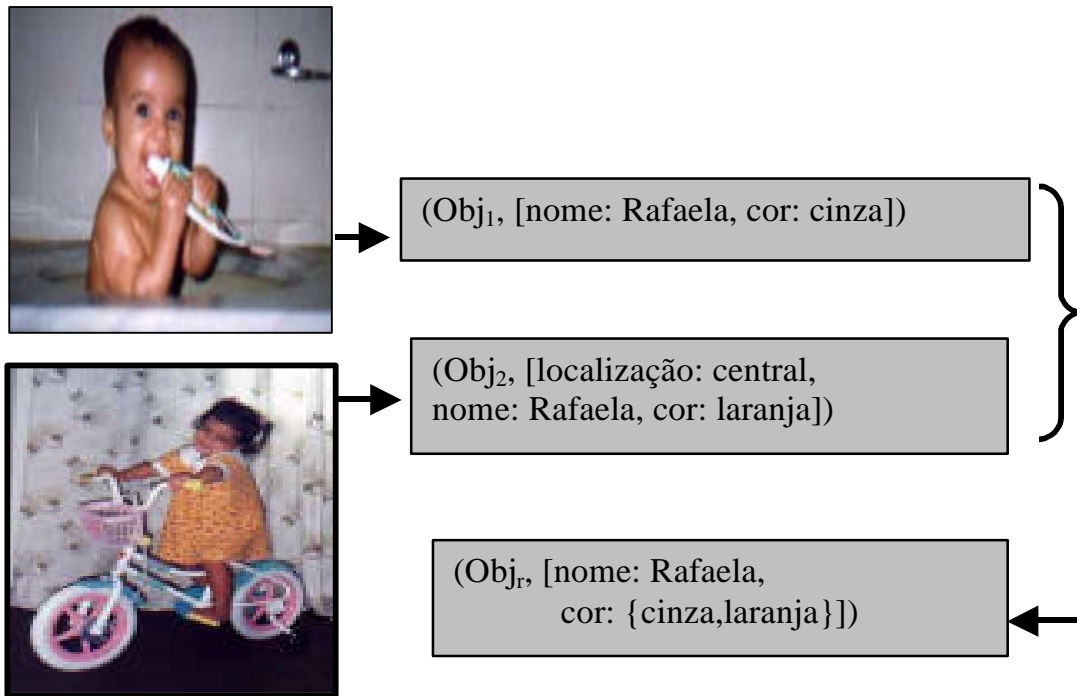


Figura 12: Conjunção de Meta-Objetos Multimídia

Os metadados que são comuns em ambos os objetos são agregados pelo novo objeto. No exemplo da Figura 12, o metadado **Localização** não aparece com seu valor no objeto resultante, visto que ele não é comum a ambos os meta-objetos multimídia.

Definição 3: Conjunção de Meta-Objetos Multimídia

A conjunção entre meta-objetos multimídia $Obj_1 = (Id_{o1}, d_1)$ e $Obj_2 = (id_{o2}, d_2)$, denotada por $Conj(Obj_1, Obj_2)$, é o meta-objeto multimídia $Obj_r = (id_{or}, d_r = Conj(d_1, d_2))$, tal que $d_r = [m_1:v_1, \dots, m_n:v_n]$, onde $m_i \in (metadados(d_1) \cap metadados(d_2))$ e $(1 \leq i \leq n)$, corresponde:

1. Para todo $i=1,2,\dots,n$, $m_i \in metadados(d_1)$ e $m_i \in metadados(d_2)$ faça
 - a. Se $d_1.m_i = d_2.m_i$ então $Concat(d_r, [m_i: d_1.m_i])$
 - b. Se $d_1.m_i \neq d_2.m_i$ então

-
- i. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo atômico então $Concat(d_r, [m_i: \{d_1.m_i, d_2.m_i\}])$
 - ii. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo conjunto então $Concat(d_r, [m_i: d_1.m_i \tilde{\cap} d_2.m_i])$
 - iii. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo lista então $Concat(d_r, [m_i: Conj(d_1.m_i, d_2.m_i)])$
 - iv. Se $d_1.m_i$ é um valor do tipo atômico e $d_2.m_i$ é um valor do tipo conjunto então $Concat(d_r, [m_i: \{d_1.m_i\} \tilde{\cap} d_2.m_i])$
 - v. Se $d_1.m_i$ é um valor do tipo conjunto e $d_2.m_i$ é um valor do tipo atômico então $Concat(d_r, [m_i: d_1.m_i \tilde{\cap} \{d_2.m_i\}])$

A conjunção permite que os metadados/valores do meta-objeto multimídia resultante sejam aqueles que são comuns para os objetos envolvidos. Assim, permitirá que a consulta de objetos multimídia retorne como resultado somente os objetos que casam exatamente com os padrões especificados na consulta.

3.2.6 Mecanismo de Herança entre Meta-Objetos Multimídia

A herança, numa hierarquia de generalização e especialização, é o meio pelo qual valores de metadados de meta-objetos multimídia mais específicos são herdados do objeto mais genérico.

A generalização é uma forma de abstração em que um conjunto de objetos semelhantes é considerado em um objeto de nível mais alto, com os detalhes de nível mais baixo suprimidos. A generalização corresponde a um relacionamento onde, a partir

de propriedades comuns de diferentes meta-objetos, é criado outro meta-objeto. A especialização corresponde ao processo inverso à generalização.

A generalização de objetos permite definir valores de metadados para os meta-objetos multimídia de baixo nível a partir de objetos de mais alto nível. Existem dois tipos de herança no modelo de metadados: simples e múltipla.

A herança simples (Figura 13) é o mecanismo que, numa hierarquia de generalização, as propriedades de somente um objeto de nível mais alto são herdadas.

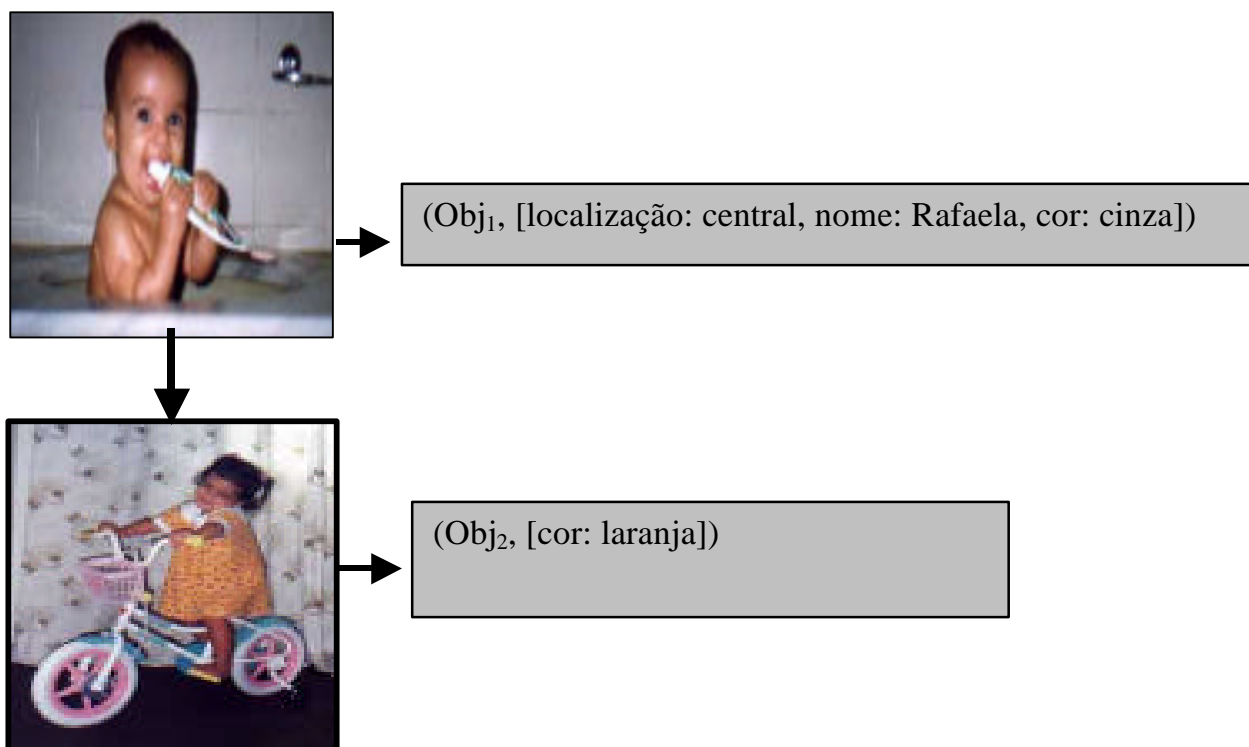


Figura 13: Herança Simples

Definição 4: Herança Simples entre Meta-Objetos Multimídia

A herança simples entre dois meta-objetos multimídia $Obj_1 = (Id_{o1}, d_1)$ e $Obj_2 = (id_{o2}, d_2)$, denotada $\acute{E}_{Um}(Obj_1, Obj_2)$, onde Obj_2 herda os metadados/valores do

meta-objeto Obj_1 . O $Obj_2 = (id_{o2}, d_2 = \hat{E}_{Um}(d_1, d_2))$, tal que $d_1 = [m_1:v_1, \dots, m_n:v_n]$,

onde $(1 \leq i \leq n)$, corresponde a:

1. Para todo $i=1,2,\dots,n$, se $m_i \in metadados(d_1)$ e $m_i \in metadados(d_2)$ então
 - a) Se $d_1.m_i \neq d_2.m_i$ então
 - i. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo atômico então $d_2.m_i = \{d_1.m_i, d_2.m_i\}$
 - ii. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo conjunto então $d_2.m_i = d_1.m_i \hat{E} d_2.m_i$
 - iii. Se $d_1.m_i$ e $d_2.m_i$ são valores do tipo lista então $d_2.m_i = \hat{E}_{um}(d_1.m_i, d_2.m_i)$
 - iv. Se $d_1.m_i$ é um valor do tipo atômico e $d_2.m_i$ é um valor do tipo conjunto então $d_2.m_i = \{d_1.m_i\} \hat{E} d_2.m_i$
 - v. Se $d_1.m_i$ é um valor do tipo conjunto e $d_2.m_i$ é um valor do tipo atômico então $d_2.m_i = d_1.m_i \hat{E} \{d_2.m_i\}$
2. Para todo $m_i \in metadados(d_1)$ e $m_i \notin metadados(d_2)$ então $d_2 = Concat(d_2, [m_i: d_1.m_i])$.

A herança múltipla (Figura 14), por outro lado, permite que as propriedades sejam herdadas de mais de um objeto de nível mais alto.

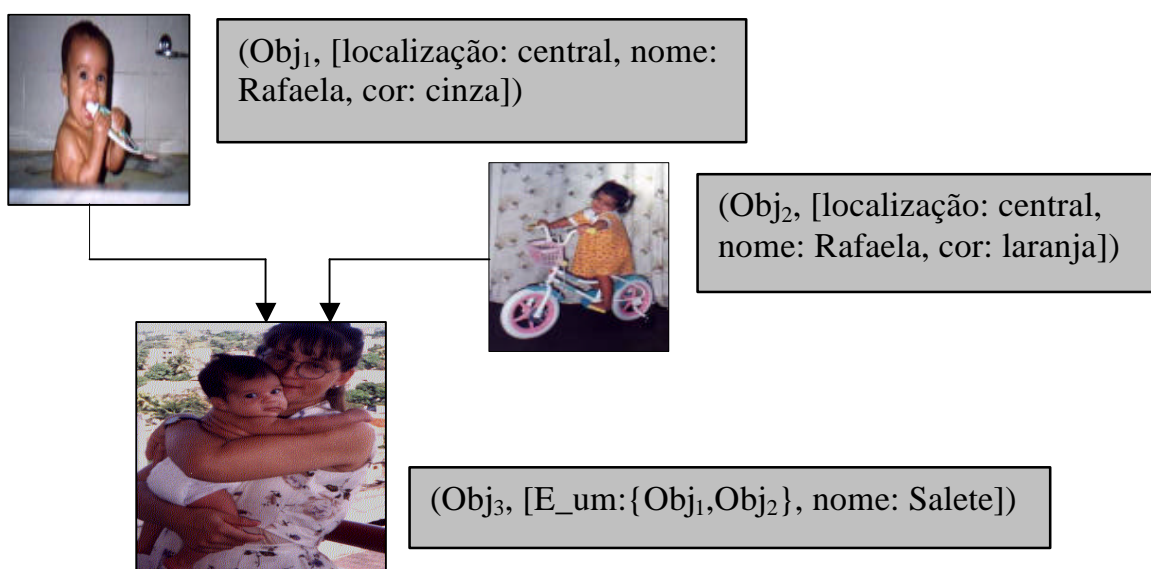


Figura 14: Herança Múltipla entre Meta_Objetos Multimídia

Definição 5: Herança Múltipla entre Meta-Objetos Multimídia

Supondo que existe um meta-objeto multimídia $Obj_I = (id_{oI}, d_I)$. Dado $Obj = \{Obj_{I1}, Obj_{I2}, \dots, Obj_{In}\}$, o conjunto de todos os meta-objetos multimídia no banco de dados, tal que $Obj \acute{E}_{um} Obj_I$ para cada $Obj_{Ii} = (id_{Ii}, d_{Ii})$. O meta-objeto multimídia $Obj_I = (id_I, d_I)$ onde: $d_I = \acute{E}_{Um}(d_{I1}, d_I) \cup \acute{E}_{Um}(d_{I2}, d_I) \cup \dots \cup \acute{E}_{Um}(d_{In}, d_I)$.

Neste tipo de herança pode haver conflitos de metadados entre o objeto de nível mais baixo e o de nível mais alto. No modelo proposto este problema é solucionado através da união de metadados/valores. Assim, o modelo, pela herança de metadados/valores, facilita o compartilhamento de dados.

O mecanismo de herança do Modelo *MetaMídia* não implica em efeito cascata, o qual exigiria muito processamento, visto que esse mecanismo de herança é controlado via hierarquia de metadados/valores.

3.3 Indexação de Objetos Multimídia e Análise de Desempenho

O modelo proposto é baseado na composição de meta-objetos multimídia através de metadados/valores. No método baseado no casamento exato, dado um padrão e a descrição de um objeto multimídia, o resultado tem todos os documentos que satisfazem pelo menos os metadados/valores introduzidos na consulta.

O tempo utilizado numa busca por ocorrências do padrão pode influenciar significativamente o desempenho de um algoritmo. Uma forma de efetuar a busca por tais ocorrências de um modo mais eficiente é através da construção de um índice de texto, ou seja, uma coleção de elementos que apontam para posições estratégicas no texto.

Assim, a representação da informação multimídia é organizada em *Clusters* (Figura 15), que possuem índices apontando para posições estratégicas no banco de dados. *Os Clusters* são agrupamentos de informações que possuem determinados metadados/valores (AB e ABC, por exemplo). Esses metadados são utilizados para a construção de índices.

As organizações em *Clusters* podem levar a diferenças significativas na eficiência de um sistema, visto que agrupam dados que são freqüentemente acessados, e permitem que os mesmos sejam fisicamente armazenados de uma forma que maximiza o acesso

seqüencial. Em suma, o agrupamento de objetos em Clusters tem dois objetivos principais: facilitar o intercâmbio entre estes objetos e diminuir o espaço de busca.

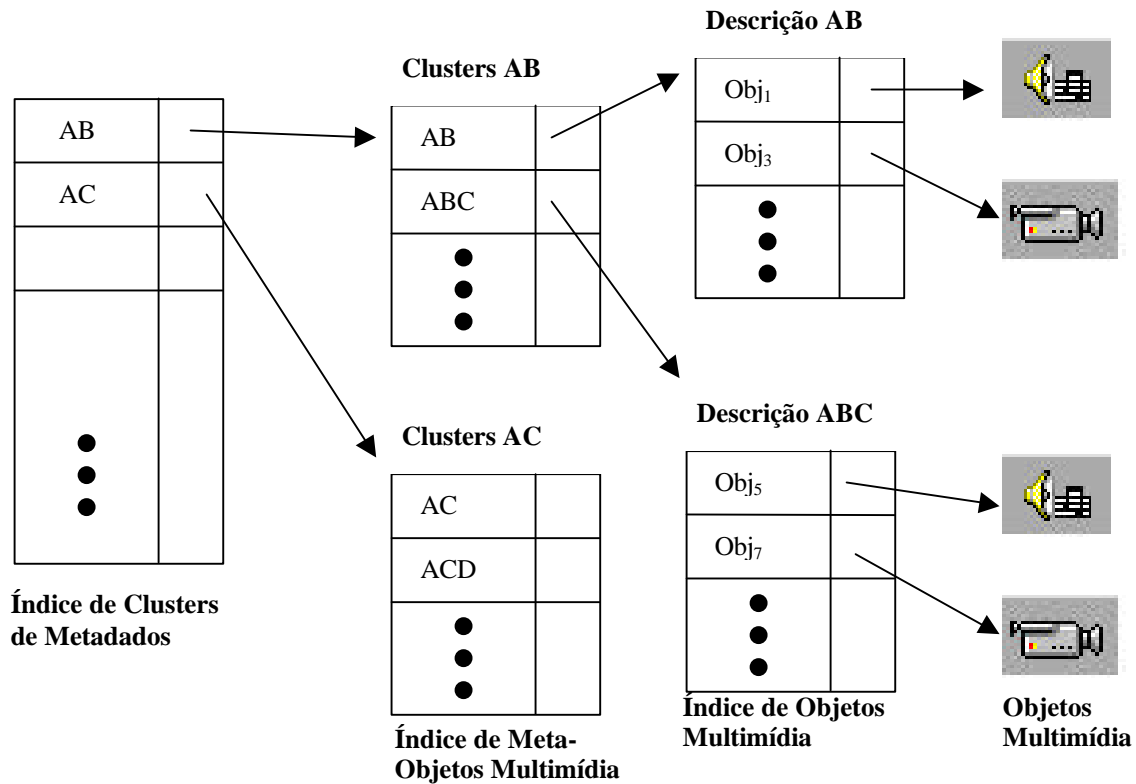


Figura 15: Índices de Clusters do Modelo *MetaMídia*

Na Figura 16 são apresentados os índices organizados de forma seqüencial, em um *array* triangular [CH97] W , que é usado para otimizar o acesso aos *Clusters*. Os índices são construídos através do uso de pelo menos dois valores de metadados, os quais são representados por letras maiúsculas.

Símbolo/Índice	B/1	C/2	D/3	E/4	F/5	G/6	H/7	I/8	J/9	K/10	L/11
A/0	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
B/1		BC	BD	BE	BF	BG	BH	BI	BJ	BK	BL
C/2			CD	CE	CF	CG	CH	CI	CJ	CK	CL
D/3				DE	DF	DG	DH	DI	DJ	DK	DL
E/4					EF	EG	EH	EI	EJ	EK	EL
F/5						FG	FH	FI	FJ	FK	FL
G/6							GH	GI	GJ	GK	GL
H/7								HI	HJ	HK	HL
I/8									IJ	IK	IL
J/9										JK	JL
K/10											KL

Figura 16: Índices em uma Array Triangular

O *array* triangular otimiza o acesso aos dados visto que, dado um índice, podem ser localizados mais rapidamente quais os *Clusters* a serem pesquisados.

A estrutura de índices para *clusters*, dentro de um array triangular, é construída em função da quantidade de metadados/valores definidos para a aplicação. O critério utilizado nesta tese é a combinação do número de metadados, 2 a 2. Por exemplo, se forem utilizados 12 metadados, tem-se 66 índices para *Clusters* possíveis. A preocupação é criar índices para o número de metadados especificados. A Definição 6, generaliza, para o número possível de índices.

Definição 6: Cálculo do número possível de índices

$$i = C_2^m$$

onde m corresponde ao número total de metadados. Portanto, o número máximo de *Clusters* com índices correspondentes é a combinação de m , 2 a 2.

A principal vantagem desta abordagem é o fato de não ser necessário pesquisar uma parte significativa do banco, mas apenas subconjuntos dele. Cada vez que um dado é

armazenado no banco de dados, o mesmo é organizado de acordo com critérios estabelecidos.

A seguir é mostrado como é pesquisado um padrão: AxY , onde $x \in X = [B, C, D, \dots]$ e Y é qualquer palavra construída ordenadamente e de qualquer tamanho com símbolos pertencentes a X . Por exemplo, $ABCDEFGF$ é um padrão para a pesquisa. O x corresponde a B e o Y corresponde a $CDEFG$.

Para $x = B$, o *Cluster* a ser pesquisado será aquele com índice AB . No *cluster* AB podem existir descrições ABC , $ABCDEFGF$, entre outras. Nestas descrições, o y é o complemento além do AB . Para $x = C$, os *clusters* a serem pesquisados serão aqueles com índices AB e AC . Para $x = D$, três *clusters* têm que ser pesquisados, AB , AC e AD . Portanto, à medida que o x aumenta, o número de *clusters* a serem pesquisados também aumenta. Então para $x = L$, os *clusters* que devem ser pesquisados serão todos os Ax , isto é, AB , AC , ..., AL . Assim sendo, no melhor caso, é pesquisado em só um *cluster*, mas no pior caso a pesquisa seria feita em onze *clusters*.

É importante observar que à medida que as informações não são fornecidas, mais *clusters* terão que ser pesquisados para verificar a existência da descrição. A vantagem desta abordagem é que, mesmo que poucas informações sejam fornecidas, a busca não será feita em todo o banco. Isto porque, mesmo no pior caso, no uso de 12 (doze) metadados, a busca será efetuada em 56 *clusters* e esses não correspondem à totalidade do banco, o qual possui 66 *Clusters*. A Definição 7, generaliza, para o pior caso, o cálculo da quantidade de *clusters* que são visitados.

Definição 7: Cálculo da Quantidade de *Clusters* que são visitados

$$c = i - (m - 2)$$

onde i é o número de índices para *clusters* existentes (Definição 7) e m é o número máximo de metadados utilizados.

A Definição 8, fornece, para um dado índice, quantos *Clusters* são visitados em todos os casos possíveis, inclusive o pior caso.

Definição 8: Cálculo da quantidade de *clusters* para todos os casos possíveis

$$q = (y - x) + \sum_{j=0}^x \mathbf{a}^j$$

onde x é o índice do primeiro símbolo e y é o índice do segundo símbolo.

Por exemplo, para o índice **DI**, o índice x será igual a **3** e o índice de y será igual a **8**. Então, aplicados esses valores na Definição 8 a pesquisa é feita em onze *Clusters*.

De modo a diminuir o número de pesquisas desnecessárias em *clusters* que não possuem ocorrências da informação desejada, foi criada uma fase de pré-processamento, onde são armazenadas as informações que são encontradas na *cluster*.

A fase de pré-processamento corresponde ao registro da ocorrência de determinados padrões nas descrições armazenadas nos *clusters*. O número de padrões varia de *cluster* para *cluster*. Esse registro vai ser feito em uma lista que conterà todos os padrões possíveis para um dado *cluster* com suas respectivas quantidades de meta-objetos multimídia.

O x_n , na Figura 17, é a quantidade de descrições, que existe para o padrão, n *cluster* específico. **A, B, C, ..., AB, AC, ..., etc.** são *substrings* que podem estar em uma

dada descrição no *cluster AB*. Se, por exemplo, no cluster não existem descrições com *BC*, então a quantidade de descrições de *BC* é zero. Vale ressaltar que todos os possíveis índices devem estar na lista.

A	B	C	...	Z	AB	AC	...	BC	BD	BE	...
x_1	x_2	x_3	...								x_n

Figura 17: Fase de Pré-Processamento

Na Figura 17, para o padrão AxY , o número de elementos de cada *Cluster* é o próprio tamanho do *Cluster*. Assim, se o *Cluster* com índice *AB* não tem elementos, o mesmo não será pesquisado e assim diminui-se o tempo de pesquisa. Já para o padrão BxY , o número de ocorrências de Bx em *AB* é colocado no *array*. Por exemplo, se x for *C* então é verificado se existem ocorrências de *BC* nos *Clusters* com índices *AB* e *BC*. Se em *AB* não ocorre *BC* então o *Cluster* não é pesquisado. Assim, se não houver ocorrências do padrão no *Cluster*, nenhuma leitura sequencial é feita, e como consequência o tempo de resposta diminui. Caso ocorra, uma pesquisa sequencial é efetuada dentro do *Cluster*.

Depois de agrupados dentro dos *clusters*, todos os meta-objetos multimídia com índices correspondentes, a quantidade máxima de meta-objetos multimídia que o *cluster* possui e, para o pior caso, a serem pesquisadas será calculada pela Definição 9.

Definição 9: Cálculo da quantidade máxima de meta-objetos a serem pesquisados

$$cl = 1 + \sum_{j=1}^{m-1} C_j^{m-2}$$

onde m é o número máximo de metadados permitidos. No pior caso, a quantidade máxima de meta-objetos a serem pesquisados é calculada pela Definição 10.

Definição 10: Cálculo da quantidade máxima de elementos a serem pesquisados

$$v = c.cl$$

onde c é a quantidade de *clusters* a serem pesquisados e cl é o máximo de elementos para um determinado *cluster*, calculado pela Definição 9. O algoritmo verifica quais os *clusters* a serem pesquisados e em seguida busca a informação.

3.4 Recuperação de Objetos Multimídia

Na Seção anterior foi apresentada a indexação de objetos multimídia via o modelo de metadados proposto. As estruturas de indexação foram apresentadas e definições foram feitas. Nesta seção é apresentado o algoritmo que implementa o método de recuperação de objetos a partir da estrutura de indexação apresentada.

O algoritmo que implementa o método *MetaMídia* (Definição 11) é apresentado através do casamento exato de padrão. Dado um conjunto de metadados/valores, o resultado será no mínimo, os objetos que casam com a especificação dada. Objetos com informações adicionais são apresentados, mas possuem as informações que foram introduzidas para a consulta.

```

Algoritmo
início
  Padrão  $P = M_1M_2...M_n$ 
   $k = 0$ 
   $ID = M_1M_2$ 
  para  $i=0, \dots, x-1$  efetuar
    para  $j=0, \dots, (x-(i+1))$  efetuar
       $IC_k = W[i][j]$ 
       $INFO[k] = (IC_k, CONT_k)$ 
       $k = k + 1$ 
    para  $j=0, \dots, (y-(x+1))$  efetuar
       $IC_k = W[x][j]$ 
       $INFO[k] = (IC_k, CONT_k)$ 
       $k = k + 1$ 
    para  $i=0, \dots, (k-1)$  efetuar
      se  $CONT_i > 0$  então
         $IC_i = [DESC_0, \dots, DESC_w]$ 
        para  $j=0, \dots, w$  efetuar
          se  $ID$  ocorre em  $DESC_j$  então
            Mostrar mídias associadas
            Ler próxima descrição
Fim.

```

Definição 11: Algoritmo Pesquisa e Recuperação de Objetos Multimídia

A partir de um padrão, todos os objetos multimídia associados a tal padrão são pesquisados e recuperados. No algoritmo, **ID** é o índice para o padrão **P** e é dado pelo usuário que está pesquisando alguma informação; **x** e **y** são os índices do primeiro e segundo símbolo no **ID**, respectivamente; **W** é o array que contém os índices; **CONT_k** é o número de ocorrências do padrão **ID** no *cluster* **IC_k** (esta informação é adquirida durante a fase de pré-processamento); e finalmente, **DESC_j** são as descrições que ocorrem dentro do *cluster* **IC_k**.

O custo de tempo, para o algoritmo *MetaMídia*, é medido pela Definição 12.

Definição 12: Ordem de Complexidade do Algoritmo MetaMídia

$$O(mv)$$

Onde m é o número máximo de metadados permitidos e v é o número máximo de elementos que são verificados, para o pior caso (Definição 10).

As principais características desejadas na construção e utilização de objetos e meta-objetos multimídia são: rapidez na recuperação dos objetos (aspectos de eficiência), processo de atualização fácil (aspectos de tempo) e economia de armazenamento (aspectos de espaço).

O processo de armazenamento de grandes objetos e seus respectivos metadados/valores não é difícil, mas a recuperação conforme metadados/valores de uma consulta requer cuidados mais elaborados. Assim, ao construir a estrutura de indexação para auxiliar a recuperação de objetos multimídia, foi verificado que os objetos multimídia não necessitam ser alterados, mas sim suas descrições e os índices correspondentes.

3.5 Considerações Finais

Neste capítulo foi apresentado o modelo de metadados na indexação e recuperação de objetos multimídia. Esse modelo de metadados corresponde a um conjunto de ferramentas conceituais usadas para a descrição de objetos multimídia, com o objetivo de indexar e recuperar os mesmos.

O modelo de metadados tem como base um conjunto de metaobjetos multimídia, os quais correspondem a estruturas que descrevem objetos multimídia. Os metaobjetos multimídia possuem uma identificação do objeto multimídia e a sua descrição através de metadados e valores.

Para manipular essa estrutura foram definidas as operações de criação, disjunção e conjunção de meta-objetos multimídia. Para definição dessas operações teve-se como dificuldade inicial o fato que os valores de metadados podem ser de vários tipos (por exemplo, atômicos, conjuntos e listas).

Uma vez criada esta estrutura, os mesmos têm que ser indexados, em um banco de dados, de forma a permitir a pesquisa e recuperação de objetos multimídia. Assim, os metadados são analisados e uma estrutura de índices é construída. Esses índices são atribuídos a *clusters* de meta-objetos multimídia, os quais são agrupamentos de metadados/valores conforme as definições criadas.

As organizações de índices têm suas próprias vantagens e desvantagens. A escolha da estrutura apropriada para um dado conjunto de dados pode ter um impacto significativo no desempenho. Se pesquisas são freqüentes, classificar o arquivo ou construir um índice é importante. Um índice é uma coleção de dados de entrada que dá suporte à recuperação eficiente de dados que tem um dado padrão de pesquisa (metadados/valores).

Os índices podem ser classificados como agrupados (clusters) ou não agrupados. A diferença tem importante implicação no desempenho das pesquisas e recuperações de objetos multimídia. Essas variações são independentes das estruturas de índices usadas para localizar os objetos multimídia.

Em todo o processo de indexação dos metadados/valores no banco de dados, são calculadas as funções para o pior caso. Assim, foi possível comprovar a viabilidade do modelo de metadados na indexação e recuperação de informação multimídia.

A partir do modelo de indexação do banco de dados, um algoritmo de pesquisa e recuperação foi criado. No próximo capítulo é apresentado o protótipo do modelo *MetaMídia*. No Capítulo 5 é apresentada uma análise comparativa de alguns modelos e o modelo proposto neste capítulo. Além disso, a evolução desses modelos na última década é também analisada.

Capítulo 4

Protótipo *MetaMídia*

*O melhor jeito de multiplicar o conhecimento
é dividi-lo.*

David Cohen

Neste capítulo é apresentada a especificação parcial, em UML, do protótipo do modelo de metadados na indexação e recuperação de informação multimídia.

4.1 Introdução

Os sistemas multimídia são ferramentas poderosas que estão mudando a maneira de adquirir o conhecimento e que dão uma nova visão da realidade. Um sistema multimídia caracteriza-se pelo tipo de informação que é especificada, manipulada, editada e armazenada.

O modelo de indexação multimídia aumenta o poder de expressividade semântica para os dados desses sistemas, possibilitando ao usuário extrair o máximo de informação do banco de dados.

O uso de abstrações permite ao usuário modelar e ver os dados em muitos níveis, facilitando a modelagem de situações do mundo real, visto que metaobjetos multimídia em muitos níveis é consistente com o modo que as pessoas vêem o mundo. Por exemplo, um objeto multimídia pode herdar meta-objetos multimídia de objetos já existentes. A generalização e especialização de meta-objetos permitem que descrições sejam feitas uma vez e não inúmeras vezes.

Nesta tese, foi proposto um modelo metadados na indexação e recuperação de objetos multimídia. Essas aplicações impõem um modelo de dados que permite uma definição natural e flexível, e um esquema que permita representar a composição de objetos e os relacionamentos entre seus componentes.

Neste apêndice é descrito o protótipo que implementa o modelo de indexação e recuperação. Essa especificação está elaborada segundo a Metodologia UML [Con99, Lar99]. Os diagramas de casos de uso, diagrama de acompanhamento de eventos, projeto interface homem/máquina – telas do sistema, diagrama de classes e descrição das classes, métodos e atributos são apresentados.

Para validar o modelo apresentado, este capítulo foi estruturado como segue. Na Seção 4.2, é apresentada uma breve descrição da linguagem UML. Na Seção 4.3, são apresentados os diagramas de casos de uso do protótipo. Na Seção 4.4, são apresentados os diagramas de acompanhamento de eventos. Na Seção 4.5, é apresentado o projeto de interface homem/máquina, as telas do sistema. Na Seção 4.6, é apresentado o diagrama de classes do modelo. Nessa mesma seção, são descritas cada uma das classes e suas

propriedades (atributos e operações). Na Seção 4.7 são descritos aspectos de codificação. Na Seção 4.8 é feita uma avaliação da implementação do protótipo e feita as considerações gerais do capítulo.

4.2 Linguagem *UML* para Descrição do Modelo *MetaMídia*

A Linguagem UML (*The Unified Modeling Language*) [Cor99, Lar99] é uma linguagem padrão para especificar, visualizar, documentar e construir artefatos de um sistema e pode ser utilizada em todos os processos do ciclo de desenvolvimento de *software*, e em diferentes tecnologias de implementação.

Os principais objetivos desta linguagem são fornecer aos usuários uma linguagem de modelagem visual expressiva e pronta para uso visando o desenvolvimento de modelos de negócio; fornecer mecanismo de extensibilidade e de especialização para apoiar conceitos essenciais; ser independente de linguagens de programação e processos de desenvolvimento; prover uma base formal para entender a linguagem de modelagem; encorajar o crescimento no número de ferramentas orientadas a objeto no mercado; dar suporte a conceitos de desenvolvimento de nível mais devado tais como colaborações, estrutura de trabalho, padrões e componentes; e integrar as melhores práticas.

Em geral, essa linguagem pode ser usada para mostrar as fronteiras de um sistema e suas funções principais utilizando atores e casos de uso; ilustrar a realização de casos de uso com diagramas de interação; representar uma estrutura estática de um sistema utilizando diagramas de classe; modelar o comportamento de objetos com diagrama de

transição de estado; revelar a arquitetura de implementação física com diagramas de comportamento e de implantação; e estender sua funcionalidade através de estereótipos.

Nas seções seguintes é usada a Linguagem UML para especificar o protótipo do modelo de indexação e recuperação de informação multimídia proposto nesta tese.

4.3 Diagrama de Casos de Uso

O comportamento do sistema é capturado através da análise de casos de uso do sistema. Um diagrama de casos de uso especifica a funcionalidade do sistema. Esses diagramas são usados para descrever o comportamento nas mais variadas situações que ocorrem durante operação. O sistema e seu ambiente são descritos, além da relação entre os dois. O diagrama é constituído de atores e casos de uso. Na Figura 19 está a representação desses componentes.

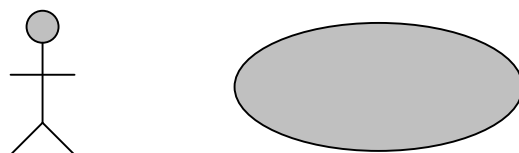


Figura 18: Representação de Ator e Caso de Uso, em UML

Um ator representa o papel que o usuário vai desempenhar em relação ao sistema. O ator não é parte do sistema. O ator pode interagir ativamente, sem receptor, passivo de informação ou representar um ser humano, uma máquina ou outro sistema.

Caso de uso é a descrição de uma transação que especifica uma utilização do sistema. É uma seqüência de ações que o sistema executa e produz um resultado para o

ator. Ele modela o diálogo entre atores e o sistema. O conjunto de todos os casos de uso representa todas as situações possíveis de utilização do sistema.

Em resumo, um caso de uso é um modelo de funções a serem executadas pelo sistema e das interações com suas fronteiras. A sua principal aplicação é confirmar aos usuários e clientes as suas funcionalidades e comportamento.

Na Figura 19, é apresentado o diagrama de casos de uso do protótipo do Modelo de Metadados na Indexação e Recuperação de Objetos Multimídia.

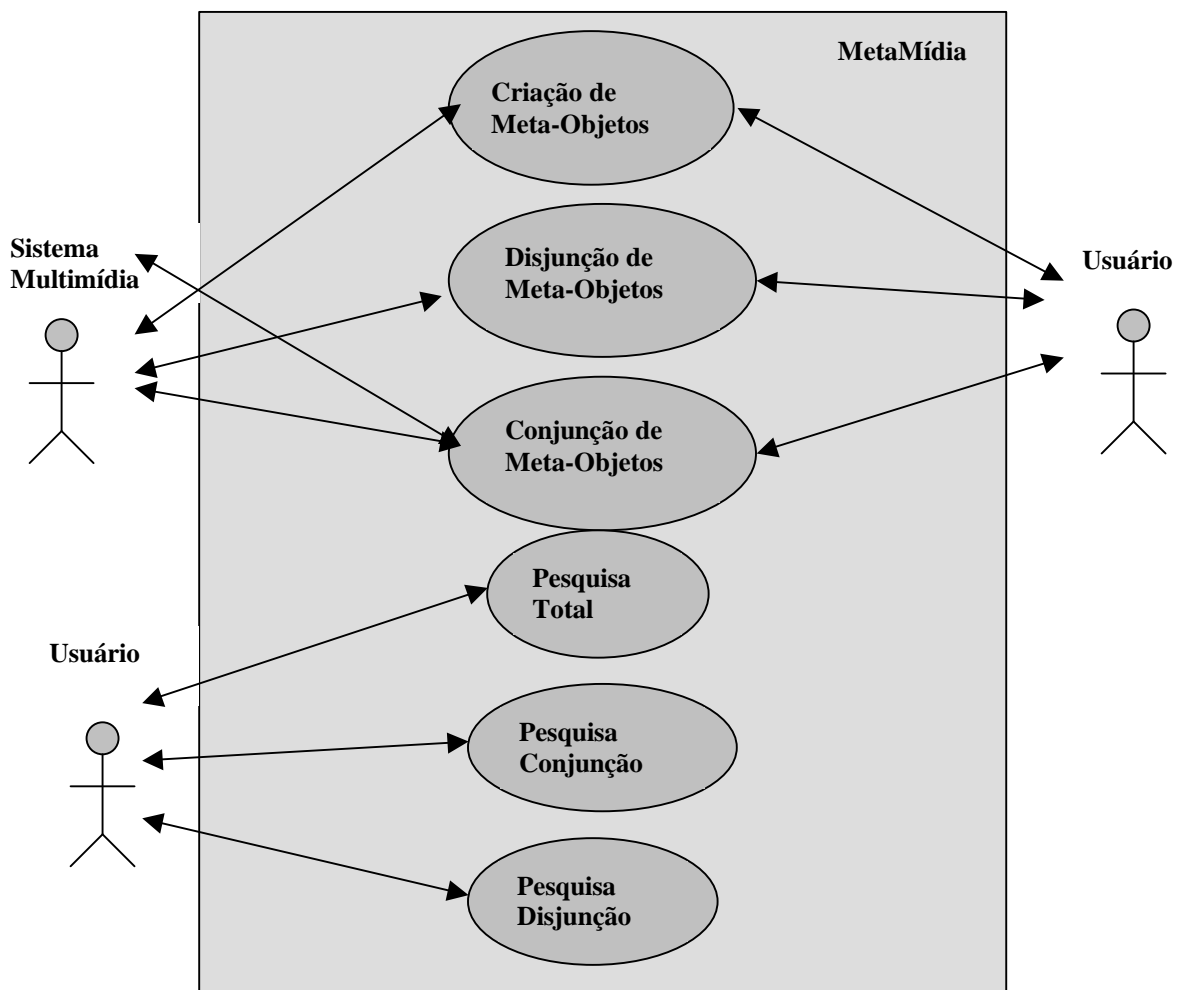


Figura 19: Diagrama de Casos de Uso do Modelo

Uma instância de um caso de uso corresponde a um cenário. O caso de uso deve ser descrito através de vários cenários. Devem ser construídos tantos cenários quantos forem os casos de uso do sistema. Na Seção 4.4 são apresentados os cenários dos principais eventos do sistema.

4.4 Cenários

Os cenários correspondem a uma seqüência ordem-tempo de eventos que podem ocorrer entre objetos. Os diagramas de acompanhamento de eventos representam graficamente os cenários. A seguir são apresentados os cenários do protótipo *MetaMídia*: Criação de Meta-Objeto Multimídia, Disjunção, Conjunção de Meta-Objetos multimídia e Pesquisa de objetos multimídia a partir dos meta-objetos.

- **Criação de Meta-Objeto Multimídia**

O **Sistema Multimídia** armazena o objeto multimídia. O **MetaMídia** exibe os metadados/valores para criação do meta-objeto. Então o **Usuário** informa os metadados e seus valores para indexação. O **MetaMídia** cria o meta-objeto multimídia e confirma a criação do meta-objeto multimídia. O **Sistema Multimídia** confirma a criação do objeto multimídia e finaliza a criação do meta-objeto multimídia.

- **Disjunção de Meta-Objeto Multimídia**

O **Usuário** dá entrada de metadados/valores que deseja para que estejam na disjunção. O **MetaMídia** pesquisa no banco de dados os objetos multimídia que casam com os

metadados/valores digitados. O **Usuário** seleciona os objetos multimídia que deseja para a disjunção. O **MetaMídia** faz a disjunção de meta-objetos multimídia e então confirma disjunção de objetos multimídia. O **Sistema Multimídia** confirma a edição de objetos multimídia.

- **Conjunção de Meta-Objetos Multimídia**

O **Usuário** dá entrada de metadados/valores que deseja para que estejam na conjunção. O **Sistema Multimídia** pesquisa na banco de dados os objetos multimídia que casam com os metadados/valores digitados. O **Usuário** seleciona os objetos multimídia que deseja para a conjunção. O **Sistema Multimídia** dá entrada de objetos multimídia. O **MetaMídia** exibe os metadados/valores dos objetos envolvidos e faz a conjunção dos meta-objetos multimídia. Finalizando, confirma a conjunção dos objetos multimídia.

- **Pesquisa de Objeto Multimídia**

O **Usuário** deseja pesquisa para o **Sistema Multimídia**. O **Sistema Multimídia** exibe a tela principal de **MetaMídia**. O **MetaMídia** exibe a tela de pesquisa. O **Usuário** informa os metadados e seus valores para pesquisa. O **MetaMídia** executa o algoritmo de pesquisa (Definição 11). O **MetaMídia** exibe na tela o resultado para o **Usuário**. Esse seleciona os objetos multimídia resultantes. E então finaliza o sistema.

- **Pesquisa de Objeto Multimídia por Disjunção**

O **Usuário** deseja pesquisa para o **Sistema Multimídia**. O **Sistema Multimídia** exibe a tela principal de **MetaMídia**. O **MetaMídia** exibe a tela de pesquisa. O **Usuário**

informa os metadados e seus valores para pesquisa. O **MetaMídia** pesquisa pelos objetos criados por disjunção O **MetaMídia** exibe na tela o resultado para o **Usuário**. O **Usuário** seleciona os objetos multimídia resultantes e finaliza o sistema.

- **Pesquisa de Objeto Multimídia por Disjunção**

O **Usuário** deseja pesquisa para o **Sistema Multimídia**. O **Sistema Multimídia** exibe a tela principal de **MetaMídia**. O **MetaMídia** exibe a tela de pesquisa. O **Usuário** informa os metadados e seus valores para pesquisa. O **MetaMídia** executa o algoritmo de pesquisa. O **MetaMídia** exibe na tela o resultado para o **Usuário**. O **Usuário** seleciona os objetos multimídia resultantes e finaliza o sistema.

4.5 Projeto de Interface Homem/Máquina

O projeto de Interface Homem/Máquina tem por objetivo apresentar um protótipo do sistema. Nesta seção são apresentadas as principais telas do sistema, com as quais o usuário irá interagir durante a realização das operações que o sistema lhe permite efetuar

A tela principal (Figura 20) permite que o usuário inclua objetos multimídia do tipo vídeo, imagens estáticas (fotos) e som. No momento da inclusão é dada a opção de criação de meta-objetos multimídia, que permitirão a busca baseada em metadados. Se o objeto que está sendo criado já existe na base, a operação de criação não é habilitada.

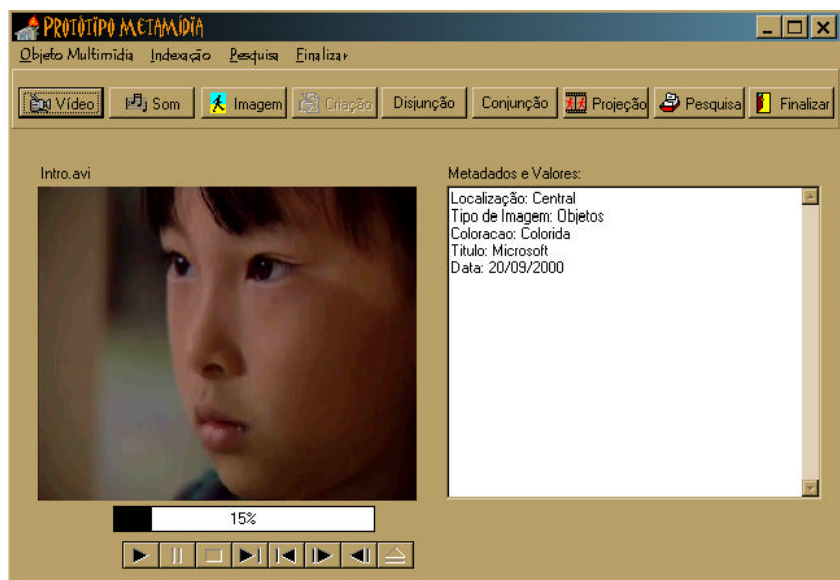


Figura 20: Tela Principal do Protótipo MetaMídia

Na Figura 21, o objeto foi carregado, mas não existia o metaobjeto criado. Assim, a operação de criação é permitida.

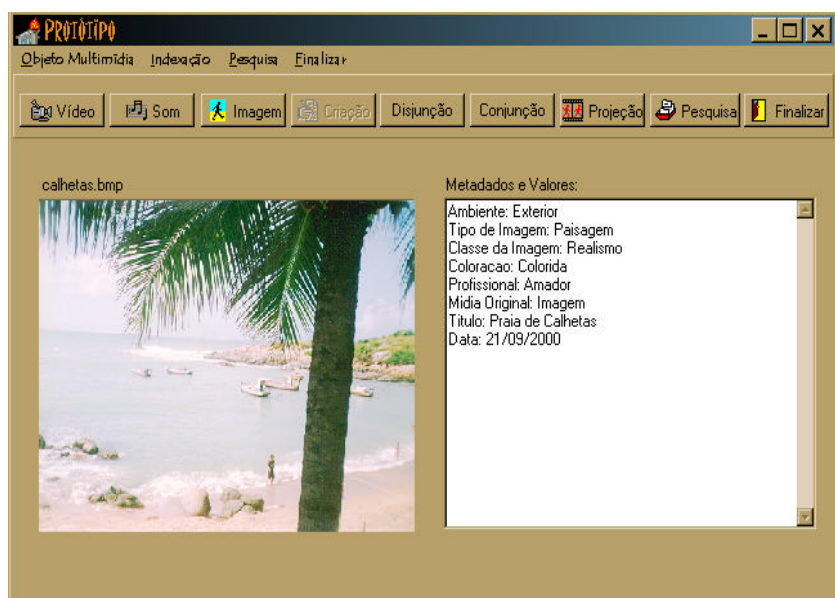


Figura 21: Tela com Criação de metadados/valores para imagens

Após a criação, já irão aparecer no lado direito da tela principal os metadados/valores correspondentes ao objeto. A tela da Figura 22 corresponde à conjunção e disjunção de objetos multimídia. Primeiro, é feita uma pesquisa de objetos que se pretende fazer a

conjunção ou disjunção. Com o resultado, o usuário decide se indexa ou não o objeto com os outros objetos.

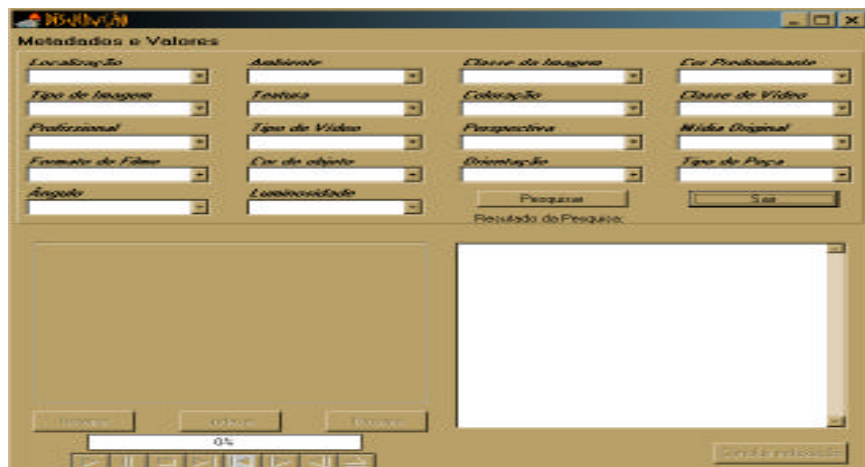


Figura 22: Tela de Conjunção ou Disjunção

A tela (Figura 23) permite a pesquisa e recuperação de objetos multimídia. O usuário dá entrada aos dados para a pesquisa e os resultados são colocados. O retorno pode ser qualquer tipo de mídia que case com os metadados/valores inseridos.



Figura 23: Tela de Pesquisa

A pesquisa pode ser feita total, isto é, todos os objetos existentes na base, inclusive os criados por disjunção e conjunção, são recuperados. As outras duas pesquisas são por conjunção e disjunção, respectivamente.

4.6 Diagrama de Classes

Diagrama de classes é um esquema, um padrão ou um modelo para descrever as instâncias possíveis e as classes de objetos. No diagrama de classes são especificados todos os atributos, operações e relacionamentos de todas as classes do protótipo.

A diagrama de classes provê uma notação gráfica para modelagem das classes e os relacionamentos entre elas. Ele é adequado tanto para modelar abstrações como projetar programas, pois é fácil de entender, conciso e bem prático.

Uma classe de objetos é uma descrição de um grupo de objetos com atributos, comportamentos, relacionamentos com outros objetos e semântica comum. Uma classe é uma abstração que enfatiza características relevantes dos objetos, suprimindo outras. Portanto, um objeto é uma instância de uma classe.

A classe de objetos é representada por um retângulo, subdividido em três áreas. A primeira contém o nome da Classe. A segunda contém os atributos e a terceira contém suas operações.

Para a descrição textual da representação que cada classe do sistema, bem como a descrição de seus atributos e operações, é usado dicionário de dados, em UML. O dicionário de dados permite um detalhamento maior do que está descrito no Diagrama de Classes (Figura 24).

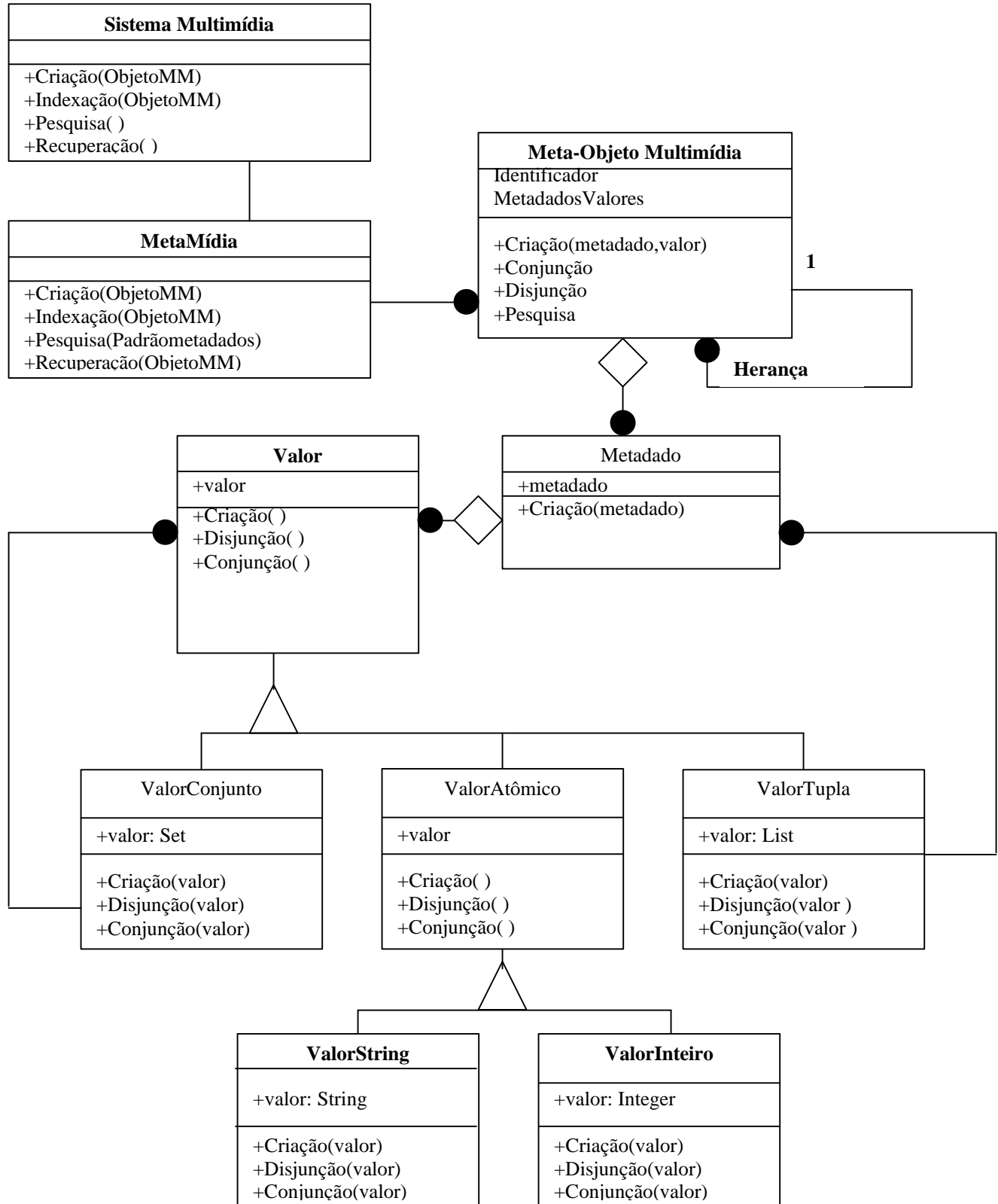


Figura 24: Diagrama de Classes do Protótipo

A seguir é apresentada a descrição das classes, dos atributos e das operações do protótipo.

- **Sistema Multimídia** – representa um sistema multimídia que cria objetos multimídia do tipo vídeo, imagem, som, texto e documento. As operações sobre esta classe são: criação e pesquisa de objetos multimídia. O protótipo *MetaMídia* não edita objetos multimídia. A partir de um sistema multimídia, no momento que o objeto multimídia é armazenado, o mesmo permite criar o meta-objeto multimídia para o mesmo. A operação de criação permite dar entrada do objeto multimídia no banco de dados. A operação de pesquisa permite recuperar objetos multimídia segundo um padrão de entrada.
- **Meta-Objeto Multimídia** – representa todos os meta-objetos multimídia relacionados aos objetos multimídia criados. Esta classe tem as seguintes operações: Criação, conjunção, disjunção e pesquisa. A operação de criação permite criar um meta-objeto multimídia com os seus respectivos metadados e valores (ver descrição dessa operação na Seção 4.2.3). A operação de disjunção permite a união entre meta-objetos multimídia (ver descrição dessa operação na Seção 4.2.4). A operação de conjunção permite a interseção entre meta-objetos multimídia (ver descrição dessa operação na Seção 4.2.5). A operação de pesquisa permite a consulta e recuperação de meta-objetos multimídia.
- **Metadado** – representa os metadados atribuídos a um meta-objeto multimídia. Esta classe permite as seguintes operações: criação, exclusão e alteração de

metadados. Essas operações, como o próprio nome diz, correspondem a criação de metadados atribuídos ao meta-objeto multimídia, respectivamente.

- **Valor** – representa uma classe abstrata. Ela permite a atribuição de valores aos metadados de um meta-objeto multimídia. Esta classe permite a criação de valores. Essa operação, como o próprio nome diz, correspondem a criação de valores de metadados atribuídos ao meta-objeto multimídia, respectivamente.
- **ValorConjunto** – corresponde a uma subclasse da classe Valor. Essa classe representa os valores do tipo conjunto. Esta classe redefine as operações de criação, disjunção e conjunção de valores de metadados do tipo conjunto.
- **ValorAtômico** – corresponde a uma subclasse da classe Valor. Esta classe representa os valores do tipo atômico e é superclasse das classes ValorString e ValorInteiro. Esta classe redefine as operações de criação, disjunção e conjunção de valores de metadados do tipo atômico.
- **ValorLista** – corresponde a uma subclasse da classe Valor. Esta classe representa os valores do tipo Lista. Esta classe redefine as operações de criação, disjunção e conjunção de valores de metadados do tipo Lista.
- **ValorString** – corresponde a uma subclasse da classe ValorAtômico. Esta classe representa os valores do tipo String. Esta classe redefine as operações de criação, disjunção e conjunção de valores de metadados do tipo String.

- **ValorInteiro** – corresponde a uma subclasse da classe ValorAtômico. Esta classe representa os valores do tipo Inteiro. Esta classe redefine as operações de criação, conjunção e disjunção de valores de metadados do tipo Inteiro.

4.7 Codificação

Após a confecção de todos os diagramas, a parte final é a codificação. A codificação corresponde à implementação de todas as classes, com os atributos, operações e relacionamentos entre elas.

A codificação deste protótipo foi feita em Delphi 5.0 e usando todas as técnicas de análise orientada a objetos.

4.8 Considerações Finais

Neste capítulo foi apresentado o Protótipo *MetaMídia*. Existem várias observações a serem consideradas na implementação do modelo de metadados na indexação e recuperação de objetos multimídia, as quais são: implementação das operações de composição e do algoritmo de pesquisa, e o grande conjunto de dados armazenados.

A individualização das informações multimídia, em cada conjunto recuperado, é feita por inspeção visual (característica subjetiva da atividade). Diferentes pessoas, mesmo treinadas de acordo com os mesmos critérios, podem indexar, tanto para armazenamento como para pesquisa e recuperação, de forma diferente uma mesma mídia, da mesma forma que diferentes pintores, mesmo seguindo uma mesma escola, expressam de forma diferente uma mesma paisagem.

Com o Protótipo *MetaMídia*, o problema das diferentes visões subjetivas de uma mídia, por pessoas diferentes é minimizado através do conjunto de metadados/valores criados para uma determinada aplicação. Esse conjunto de metadados/valores está disponibilizado via hierarquia de metadados/valores.

A composição de meta-objetos multimídia é feita através da implementação das operações de conjunção de disjunção de meta-objetos multimídia. A composição, no Protótipo *MetaMídia*, permite que um usuário solicite todos as mídias que contenham os metadados/valores especificados, ou alguns deles.

Neste protótipo, a recuperação é exata. Isto é, a pesquisa é feita para que a recuperação seja executada pelo menos para todos os metadados/valores especificados.

No próximo capítulo são analisados trabalhos relacionados e uma análise comparativa com o *MetaMídia* é feita.

Capítulo 5

Trabalhos Relacionados

*A grande finalidade do conhecimento
não é conhecer, mas agir.*

Thomas Henry Huxley

Neste capítulo são apresentados alguns trabalhos relacionados, dando ênfase aos processos de indexação, aos métodos de recuperação e aos algoritmos de pesquisa. Uma análise comparativa dos trabalhos, com o modelo proposto, é efetuada. A evolução dos modelos na última década é analisada.

5.1 Introdução

Atualmente, o problema não reside em como armazenar grandes quantidades de informação em computadores, mas como recuperar as informações de forma fácil, rápida e correta. A eficiência da recuperação pode depender dos tipos de dados que são tratados, das informações utilizadas para descrição do conteúdo, dos métodos de recuperação permitidos e dos algoritmos de pesquisa implementados.

A recuperação de informação multimídia [AZP96,CL96,ZCA96,GSJ97] não trata somente tipos de dados alfanuméricos, mas também imagem, vídeo, som, entre outros. Portanto, a eficiência na recuperação depende dos tipos de dados tratados. Esses tipos de dados estão, cada vez mais, sendo parte de ambientes de computação com processamento de *strings* e gerenciamento de banco de dados.

As pesquisas em recuperação de informação multimídia são divididas em duas categorias. A primeira é baseada em anotação de dados multimídia com texto. Nessa categoria são utilizados sistemas de recuperação de textos existentes, os quais permitem pesquisar uma informação visual, indiretamente, através de anotações.

A segunda é representar objetos multimídia no banco de dados e usar seus atributos visuais diretamente. Podem ser usadas técnicas de visão computacional para extrair atributos visuais de baixo nível dos objetos multimídia, como cor, textura para imagens e parâmetros de movimento para vídeo, aspectos de som, etc.

Para um dado atributo, a representação e a medida de similaridade são determinadas. A representação é feita através de uma coleção de atributos e a recuperação de objetos é realizada com base na computação das distâncias de similaridade entre esses atributos.

A descrição do conteúdo pode ser realizada de forma manual, automática ou semi-automática. O processo manual geralmente é uma tarefa um tanto difícil, visto que pode depender do ponto de vista de quem estiver descrevendo o objeto. Em geral, a tarefa de descrição manual inclui os seguintes pontos: escolher uma lista de metadados, utilizar

regras para definição de metadados para objetos complexos, e varrer objetos a fim de identificar quais metadados devem ser selecionados. Todos esses pontos resultam do fato de que o usuário que está descrevendo tem que conhecer os objetos, além de ter que lembrar as regras para descrição.

Em objetos textuais, o processo automático é muitas vezes baseado em análises estatísticas das palavras e frases usadas no texto do documento. Porém, no caso de outras mídias outras estratégias são usadas. No caso de figuras, por exemplo, são extraídas suas cores, considerando padrões internacionais de definições de cores.

Em geral, existem quatro métodos de recuperação, os quais foram descritos na Seção 2.4.1: recuperação através de um identificador, por sentenças condicionais, baseada em similaridades e o método de recuperação semântica.

Os algoritmos de pesquisa implementados devem não somente recuperar, mas fazê-lo de forma eficiente, rápida e correta. O algoritmo de recuperação e os processos de descrição são os principais fatores que resultam em uma recuperação eficiente.

Imagem é uma fonte rica e subjetiva de informação. Por exemplo, diferentes pessoas extraem diferentes significados de uma mesma imagem. Como consequência, sistemas de informação multimídia necessitam de recursos para transmissão, armazenamento e processamento. Esses fatores fazem a indexação, recuperação e gerenciamento de informações multimídia um grande desafio.

Neste capítulo são analisados os principais trabalhos relacionados à indexação e recuperação de informação multimídia. Para tanto, o capítulo está organizado como

segue. Na Seção 5.2 é apresentado um estudo dos principais trabalhos, procurando enfatizar os tipos de mídias tratadas, o processo de indexação, os métodos de recuperação permitidos às informações e os algoritmos de pesquisa implementados. Na Seção 5.3 é apresentada uma análise comparativa dos trabalhos. Na Seção 5.4 é apresentado a evolução dos modelos até o *MetaMídia*. Finalmente, a Seção 5.5 é destinada às conclusões do capítulo.

5.2 Modelos e Sistemas

Nesta seção são analisados sete trabalhos relevantes sob o prisma dos tipos de mídia, indexação de informação, métodos de recuperação e algoritmos de pesquisa implementados.

5.2.1 Sistema VIMSYS

Bach et al. projetaram o Sistema VIMSYS (*Visual Information Management System*) [GWJ91,BPJ93,GSJ97], com o objetivo de gerenciar informações visuais para reconhecimento de faces humanas.

A indexação é feita de forma incremental e semi-automática. Isso é possível por meio de um conjunto de atributos básicos, onde a partir desses o usuário interage com o sistema, introduzindo novos atributos.

Quando uma nova imagem é inserida no banco de dados, todas as informações associadas à imagem devem ser armazenadas no banco. Os metadados e seus valores definidos pelo usuário podem ser armazenados de uma maneira padrão. Os metadados

derivados diretamente da imagem, não podem ser facilmente armazenados. Por exemplo, um grupo de *pixels*, formando um objeto em uma imagem, deve ser armazenado no banco de dados, mas seria impraticável armazenar todos os *pixels* individualmente.

A recuperação pode ser realizada através de um identificador, por sentenças condicionais, baseada em similaridades e semântica. A recuperação através de um identificador pode ser efetuada mediante consultas a qualquer um dos atributos das imagens no banco de dados.

A recuperação por sentenças condicionais permite que a pesquisa seja realizada em estágios, permitindo que os critérios sejam direcionados, permitindo que o usuário ajuste os critérios das consultas a cada estágio e progressivamente consiga os resultados desejados. Este tipo de recuperação é adequado quando a especificação do conteúdo da figura (nas consultas) está ambígua. Porém, as pesquisas ao banco de dados são caras, visto que cada vez que uma consulta é efetuada, uma parcela substancial do banco de dados é pesquisada. Isso pode consumir muito tempo para banco de dados com grande quantidade de informações.

A recuperação por similaridade é efetuada através da descrição de uma imagem. O sistema recupera as imagens com maior grau de similaridade àquela descrição. O cálculo de similaridade é baseado nos valores efetuados para descrever a imagem da consulta e que estão contidos nas imagens.

A recuperação semântica é permitida, mas em pequenas proporções. As pesquisas são restritas a poucos atributos e são limitadas, visto que a geração das consultas é

baseada em descrições de imagens recuperadas através de consultas anteriores. No caso de não existirem consultas anteriores, o sistema permite que as consultas sejam geradas utilizando descrições alfanuméricas padrão.

Quando o usuário inicia a pesquisa, características como sexo, idade, etc., são especificadas, embora quaisquer metadados possam ser usados. Para todos os metadados fornecidos pelo usuário são atribuídos valores, sendo esses baseados em estatísticas de metadados, a fim de estabelecer a pesquisa inicial. Depois disso, esta será executada e as imagens casadas com as especificações desta pesquisa são retornadas. A partir daí, o usuário especificará novas consultas baseadas nos resultados obtidos.

O sistema VIMSYS utiliza um algoritmo de segmentação para localizar cada objeto em uma imagem. Para cada objeto do domínio, os atributos e funções necessárias serão mantidos para permitir que esse objeto seja identificado e segmentado em uma imagem.

5.2.2 Sistema OVID

Oomoto e Tanaka projetaram o sistema OVID (*Object-oriented Video Information Database*) [OT93, Oom94, OT97, Vaz97, VVSS98] que trata informação do tipo vídeo. A informação pode ser recuperada através de um identificador e por sentenças condicionais. Essas recuperações são efetuadas a partir de textos, que são descrições (pares de atributos/valores) associadas a cenas significativas de vídeo.

As descrições de cenas de vídeo ou objetos do tipo vídeo são subdivididas em objetos menores, e esses herdam as descrições do objeto original. Um objeto do tipo vídeo

também pode ser valor de um atributo em um outro objeto. Assim, cenas de vídeo podem ser descritas mais naturalmente através de outra cena de vídeo. As descrições são definidas durante o processo de codificação e estruturação. Essas descrições são armazenadas em listas.

A consulta e recuperação dos objetos do tipo vídeo são feitas através de uma linguagem de consulta chamada VideoSQL. O resultado da consulta é uma coleção de objetos que satisfazem uma condição especificada. Antes das consultas serem avaliadas, os objetos são avaliados com base no mecanismo de herança de intervalos de vídeo.

O processo de herança, deste sistema, é diferente dos sistemas orientados a objetos. A herança de inclusão de intervalos ocorre entre instâncias e não entre classes. Os objetos do tipo vídeo pertencem a uma mesma classe.

A herança de inclusão de intervalos permite que informações de um objeto A sejam herdadas por um objeto B, desde que os intervalos de B sejam definidos sobre alguma porção dos intervalos de A e os atributos estejam incluídos em um conjunto de atributos herdáveis. Esse conjunto de atributos é definido pelo usuário.

Na Figura 25, por exemplo, i_1 pode ser uma cena de vídeo a respeito do Presidente Fernando Henrique Cardoso, i_2 a visita do Presidente na comemoração dos 500 anos do descobrimento do Brasil, a qual é uma porção de i_2 . Sendo i_3 uma cena do Presidente fazendo um discurso, que também é uma porção da cena i_1 , então é razoável que os atributos e valores sejam herdados por i_3 . Porém, nem todos os atributos e valores de i_1

devem ser herdados. Neste momento, o usuário deve definir um conjunto de atributos herdáveis.

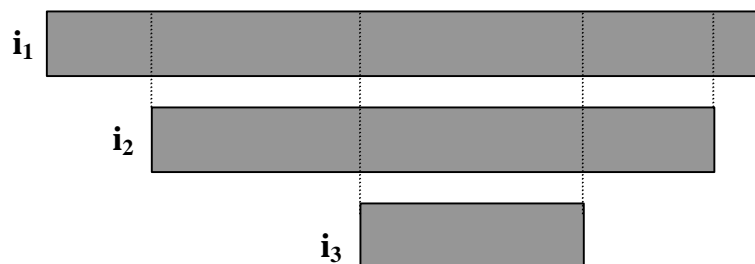


Figura 25: Herança de Inclusão de Intervalos no OVID

O modelo OVID foi especificado formalmente [VVSS98], na linguagem de especificação MooZ (*Modular object-oriented Z*) [MCS93], onde foram expostos de maneira mais clara os conceitos apresentados neste modelo. Na especificação original de OVID [OT97], foi considerado que conjuntos de intervalos também podem ser valores de atributos. No entanto, ao especificar o modelo de dados OVID, chegou-se a conclusão que a especificação original está inconsistente ou então incompleta, pois: (i) nos exemplos apresentados não há menção de intervalo como valor de atributo; (ii) intervalos já são componentes de objetos e assim não há sentido em colocá-los como valores de atributo. Se realmente houver necessidade, um objeto **A** pode ser colocado como valor de um atributo em um objeto **B**, e como consequência, o intervalo associado a **A** será associado ao atributo em **B**; (iii) para um intervalo ser um valor, o mesmo deveria ter a operação **Is_a** e o trabalho não menciona tal operação sobre intervalos.

A indexação efetuada de forma manual, como ocorre em OVID, pode ter diferentes imagens armazenadas, que deveriam possuir a mesma descrição, porém, as descrições podem ser totalmente diferentes. O problema é que não existem definições claras de como

cada imagem deve ser descrita. No sistema OVID não existe um modelo de indexação a ser seguido, tornando o trabalho tedioso e podendo resultar em descrições errôneas.

No sistema OVID, a herança é bastante adequada, pois ajuda o usuário na navegação em objetos, indo de níveis abstratos para detalhes de cenas, de forma gradual.

A consulta em texto, apresentada por OVID, é adequada apenas se as informações a respeito da imagem de vídeo forem necessárias, porém inadequada se for necessário pesquisar um conjunto de imagens diferentes, onde as mesmas, ou similares a estas, devem ter a mesma descrição. Pode ocorrer que essas imagens tenham descrições totalmente diferentes. O problema poderia ser resolvido se definições claras de como cada imagem deve ser descrita fossem estabelecidas. Porém, como usuários elaboram essa descrição, possibilidades de erros existem.

5.2.3 Sistema QBIC

O sistema QBIC (*Query By Image Content*) [NBE+93, FBF+94, LBN+94a, LBN+94b, FSN+95, TOV+95] foi projetado para tratar informações do tipo vídeo e imagens estáticas. Tanto no OVID como nesse sistema, a informação do tipo vídeo consiste de uma seqüência de *frames*, correspondendo a cenas significativas de vídeo. Um frame de cada clipe é extraído ou gerado como um frame representativo do clipe.

O sistema QBIC indexa fotografias. Esse sistema apresenta um conjunto básico de atributos, o que torna mais adequada à descrição dos dados multimídia, visto que tem uma base inicial a ser seguida. OVID, por não ter esse conjunto básico, fica sujeito às descrições e pontos de vista de usuários distintos.

A escolha do frame representativo, de forma estática, pode ser tão simples como selecionar o primeiro frame, o último ou do meio. Porém, na seqüência em movimento, pode não existir um único frame que é representativo para toda a cena. Para esse caso, o QBIC constrói um frame “conhecimento”, a partir da seqüência de frames que quer representar. Esse frame representativo é armazenado no banco de dados como imagem estática.

O usuário pode recuperar imagens e/ou vídeos por padrões de cor. Ele compõe a consulta graficamente com as cores desejadas. O banco de dados é pesquisado e as imagens com um histograma de cor similar, para aquela consulta do usuário, são recuperadas.

Todas as comparações são feitas em imagens estáticas ou frames representativos. Se o frame representativo é recuperado, como resultado de uma consulta, então o usuário pode dar um duplo *click* na imagem para ver o vídeo clipe.

Os atributos são gerados a partir da consulta gráfica, então são usados algoritmos de casamento que recuperam imagens ou vídeos do banco de dados com atributos similares.

Esse sistema apresenta a recuperação semântica e por similaridade. Os métodos de recuperação são baseados em conteúdo de vídeo. As consultas podem ser feitas com base em padrões de cor, textura, formato e posição dos objetos.

O objeto pode também ser identificado pelo seu contorno. O usuário desenha um objeto e a curva desenhada é alinhada com o contorno do objeto. Isso pode ser feito em

tempo real, tanto que o usuário desenha e a curva é automaticamente casada com objeto armazenado.

O algoritmo de segmentação é utilizado para gerar os atributos das imagens. Para facilitar a segmentação automática, uma representação em camada de vídeo é apresentada. As diferentes camadas de vídeo são geradas pelo algoritmo e usadas para identificar objetos significativos na cena, para computações de atributos e processamento de consultas.

Nas recuperações de vídeo são acrescentados os parâmetros de movimentação de câmara, tal como *zoom*. As imagens podem ser recuperadas baseadas em exemplo, isto é, dada uma imagem, recupera-se as imagens similares àquela.

No sistema QBIC, a recuperação é baseada em similaridade e semântica. As consultas são efetuadas ao banco de dados utilizando imagens exemplos, esqueletos e desenhos construídos pelos usuários, ou padrões de textura e cores selecionados.

O algoritmo de pesquisa permite que informações de vídeo sejam segmentadas a fim de permitir a descrição e recuperação. O algoritmo faz uso de exemplos e representação em camadas de vídeo.

5.2.4 Sistema CIARS

O sistema CIARS (*Color Image Archival and Retrieval System*) [BMK95] aborda a recuperação por similaridade. Esse sistema permite indexação por cor. Os atributos de cor são extraídos automaticamente usando um algoritmo de *clustering* de cores. A indexação é realizada num espaço de similaridade, onde distância entre dois pontos para uma medida

de similaridade definida é proporcional à similaridade entre os objetos correspondentes aos pontos.

O sistema CIARS se divide em três fases. A primeira fase corresponde a definir os *clusters* de cores. Esses *clusters* são extraídos através do algoritmo de *clusters*. Esses *clusters* são representados como pontos em um espaço tridimensional de cores. Na segunda fase, esses pontos são separados em grupos disjuntos e retângulos com limite mínimo são construídos. Finalmente, na terceira fase, esses retângulos são tratados como objetos espaciais e uma estrutura de indexação *R-Tree* é construída.

Existem três tipos de consultas que podem ser efetuadas com a *R-Tree*: (i) Obter toda a área na forma de retângulos, os quais são fechados em um ponto particular; (ii) obter todos os retângulos base que têm como interseção um dado retângulo consulta; (iii) Obter todos os retângulos base que são fechados por um retângulo consulta.

Por exemplo, para conseguir pesquisar todas as imagens que têm uma cor particular especificada pelos valores (R,G,B) é preciso executar o primeiro tipo de consulta, isto é, consulta por pontos. Já para recuperar todas as imagens que têm uma cor que é similar a uma cor da imagem da consulta, é preciso executar o segundo tipo de consulta.

O sistema CIARS, da mesma forma que o sistema QBIC, apresenta informação do tipo imagem estática. Existe uma variedade de aplicações que podem ser modeladas no sistema CIARS, tais como a recuperação de objetos de museus e padrões têxteis.

A indexação de atributos é feita de forma automática. Essa indexação é realizada em um espaço de similaridade. Os atributos correspondem a atributos de informações visuais relacionados a cores.

No sistema CIARS, a recuperação baseada em similaridade utiliza um espaço de similaridade, onde os objetos contidos em um mesmo grupo são mais similares comparados com os objetos pertencentes a grupos distintos.

Um algoritmo de *cluster* é utilizado para busca. Este algoritmo é tratado da mesma forma que nos sistemas tradicionais, isto é, agrupa objetos similares formando clusters.

5.2.5 Sistema FIBSSR

O Sistema FIBSSR (*Feature Index-Based Similar-Shape Retrieval*) [MG95] permite a recuperação por similaridade. O mesmo possui uma técnica para recuperação de figuras similares. Esta técnica envolve três aspectos: representação de figuras, estrutura de índices e método de acesso.

Na representação de figuras, uma figura a ser armazenada no banco de dados é processada para obter os limites de contorno, também chamados pontos de interesse. Os limites de contorno são codificados como uma seqüência ordenada de pontos de interesse. Cada atributo é codificado por invariantes de escala, rotação e translação. O FIBSSR define um sistema de coordenadas, para tratar o ponto de interesse base, ao longo da coordenada x. Todos os outros pontos de interesse do atributo são transformados para o sistema de coordenadas. O ponto de interesse base pode ser qualquer ponto de interesse da seqüência.

Os vetores de atributos possuem a representação dos pontos de interesse, que ajudam a formar os índices de atributos para o banco de dados de figuras. A distância euclidiana entre dois vetores define a similaridade entre os dois atributos. Dado esta representação de atributos e a medida de similaridade, o índice pode ser obtido através de qualquer método de acesso a pontos multidimensionais.

O processamento da consulta envolve seleção de metadados, formação de figuras e um conjunto de figuras similares como resposta à consulta. Tal processamento utiliza um algoritmo que implementa um método de acesso a pontos multidimensionais, visando maximizar as chances de obter resultado satisfatório na pesquisa. O índice é então pesquisado para metadados similares na figura. Na consulta é especificada a tolerância para a similaridade. As figuras, onde os metadados são apresentados, são adicionadas a um conjunto de figuras similares. Assim, as figuras que são globalmente similares são adicionadas a um conjunto final.

O FIBSSR trata imagens estáticas que são aplicadas em áreas, tais como meteorologia, medicina, educação e exploração do espaço.

A indexação é efetuada baseada nos limites de contorno das imagens. Esses limites são construídos a partir de um sistema de coordenadas de imagens e essas coordenadas são armazenadas em vetores.

A recuperação é baseada na similaridade entre dois atributos, que é definida como uma distância euclidiana entre dois vetores. Vale lembrar que os vetores possuem as coordenadas dos limites de contorno das imagens.

O processamento de consulta faz uso de um algoritmo que implementa um método de acesso a pontos multidimensionais, visando maximizar as chances de obter resultado satisfatório na pesquisa.

5.2.6 Sistema Piction

O Sistema Piction [Sri92, Sri95] permite o reconhecimento e indexação de imagens de pessoas que aparecem em fotografias de jornais. Este reconhecimento é feito através de informações textuais e fotográficas. Este sistema é baseado na análise das informações que acompanham as fotos dos jornais, identificando os indivíduos que aparecem nas fotografias. Existem dois pontos chaves: interpretação das informações e localização facial dentro da fotografia.

O primeiro estágio de interpretação envolve determinar os substantivos próprios dentro do texto. Por exemplo, “O presidente da república, Fernando Henrique Cardoso”, lugares como “Brasília” e organizações tais como: “Universidade Federal de Pernambuco”. O objetivo principal de *Piction* é referenciar pessoas.

Em seguida, é necessário verificar quais das pessoas referenciadas realmente aparecem na fotografia. Por exemplo, a descrição: “Fernando Henrique Cardoso, esteve em Porto Seguro nas comemorações dos 500 anos do Brasil” implica que Fernando Henrique Cardoso pode ser visto na fotografia. Srihari observou que palavras tais como, antes e depois são significativas para identificar quem aparece na fotografia.

Quando determinado quem aparece na fotografia, então é necessário identificar dentro da fotografia quem é a pessoa. As frases mais comuns, que possibilitam identificar, são: em frente de, da esquerda, da direita, que podem ser usadas.

A recuperação semântica é efetuada tendo como objetivo identificar faces humanas, em fotografias de jornal, baseada nas informações descritas a respeito das fotografias.

As informações são representadas através de um conjunto de restrições. Esse conjunto é dividido em três tipos: espacial, caracterização e contextual. As restrições espaciais são as geométricas ou topológicas, tais como à direita, à esquerda e acima.

As restrições de caracterização descrevem propriedades de objetos e são unárias por natureza, tais como cor do cabelo. As restrições contextuais descrevem as informações relacionadas ao local onde ocorreu o fato, o contexto da cena. Por exemplo, apartamento, aeroporto, etc.

Esse sistema utiliza um algoritmo de processamento de imagem para recuperação, o qual possui duas funções básicas que são localizar e segmentar os objetos, e extrair propriedades visuais dos mesmos. Qualquer propriedade (por exemplo, identificação de pessoas ou objetos) é representada no banco de dados através de coordenadas das imagens. Similarmente, qualquer informação de caracterização da imagem que seja visualmente verificada (por exemplo, cor de cabelo) é também anotada.

As informações baseadas nas imagens determinam a presença de um indivíduo ou objeto. Essas informações podem ser quantificadas de acordo com os seguintes atributos:

(i) identificação da face; (ii) tamanho e orientação da face; e (iii) método usado para identificar a face. Esses três atributos correspondem às medidas de similaridade usadas no algoritmo de processamento de imagem para recuperação.

A indexação é efetuada baseada na extração de informações de legendas de jornais e revistas, do texto explicativo que acompanha uma foto. Esta informação pode ser usada para recuperar a foto e identificar as pessoas que aparecem nas fotos.

O sistema Piction utiliza a recuperação baseada em similaridade e semântica como o sistema QBIC. Porém, Piction identifica fontes de informações tais como similaridade objetiva de texto (casamento exato), baseada em conteúdo de texto (casamento não exato), objetiva de imagem (casamento exato) e conteúdo de imagem (casamento não exato).

O sistema VIMSYS possui uma descrição para cada objeto do domínio. Essa descrição estabelece o processo de avaliação do objeto. Já o sistema *Piction* utiliza algoritmo de processamento de imagem, mas não cita qual algoritmo. Porém, exemplifica que para cada palavra contida na consulta, um contador especifica a quantidade de ocorrência desta palavra na descrição da imagem. Qualquer identificação positiva de pessoa/objeto feita por *Piction* é representada no banco de dados pelas coordenadas nas imagens.

5.2.7 Sistema MARS

O Sistema MARS [ORC+97, RHM98, PC99] (*Multimedia Analysis and Retrieval System*) foi desenvolvido por um grupo de pesquisa na Universidade de Illinois. Esse sistema

explora a recuperação de imagens baseada em conteúdo. Uma imagem é representada como uma coleção de atributos, tais como cor, textura, forma *layout*. Esses atributos são extraídos automaticamente do conteúdo da imagem. Na forma manual são feitas descrições textuais.

Na recuperação baseada em conteúdo, atributos visuais são extraídos de imagens, tais como: cor, textura, forma e *layout*. As técnicas de recuperação de informação foram modificadas para dar suporte aos atributos visuais. Especificamente, o modelo booleano foi estendido baseado em uma interpretação *fuzzy* e probabilística de operadores booleanos, os quais são usados para dar suporte ao resultado classificado da recuperação.

Rui, Huang e Mehrotra [RHM98a] construíram uma estrutura de cenas de vídeo, onde imagens são agrupadas em função da semântica atribuída à mesma. A saída do algoritmo é um vídeo estruturado que facilita o acesso do usuário. A construção não é feita totalmente automática, uma análise de estrutura de vídeo pode ser feita manualmente através de uma ferramenta de análise assistida pelo usuário. O número de frames, imagens, grupos e cenas são resultados da construção.

O usuário constrói graficamente uma consulta pela seleção de imagens, a partir de uma coleção. Atributos específicos das imagens são selecionados. Por exemplo, o usuário pode especificar a recuperação de uma imagem similar a uma outra imagem levando em conta o atributo cor.

A consulta do usuário é interpretada como uma expressão booleana de atributos de imagens, e um modelo de recuperação booleana, adaptado para recuperação de imagens, é

usado para recuperação de um conjunto de imagens classificadas com base no grau de similaridade.

Consultas booleanas oferecem uma interface natural para o usuário formular e refinar as consultas conceituais para o sistema, usando atributos de imagens de mais baixo nível.

A forma mais simples para adaptar o modelo booleano para recuperação de imagens é associar o grau de tolerância para cada atributo, tal que uma imagem caia com uma consulta se a distância do histograma da imagem e o histograma da consulta for menor ou igual ao grau de tolerância associado.

Em MARS, é implementada uma arquitetura de *feedback* relevante, através de experimentos, permitindo melhorar o desempenho consideravelmente. A abordagem é baseada na interação com o usuário, onde este guia o sistema durante a recuperação usando esses experimentos.

A indexação é feita extraindo atributos de imagens, tais como cor e textura, automaticamente ou ainda descrições feitas manualmente.

No Sistema MARS, o método de recuperação é baseado em similaridade e por sentenças condicionais. No caso de sentenças condicionais, a consulta é interpretada através de uma expressão booleana sobre os atributos de imagens. No caso de similaridade, um conjunto de imagens é classificado em função de um grau de similaridade.

5.3 Análise Comparativa

Nesta seção é apresentado um estudo comparativo dos sistemas descritos. A representação tabular abaixo é utilizada para melhor evidenciar o resultado do estudo feito.

Aspectos Sistemas	Mídia	Indexação	Recuperação	Algoritmos
OVID	Vídeo	Manual	Identificador Condicional	(*)
VIMSYS	Vídeo	Manual Automática	Identificador Condicional Similaridade Semântica	Segmentação
CIARS	Imagem	Automática	Similaridade	<i>Clustering</i>
FIBSSR	Imagem	Automática	Similaridade	Método de acesso a pontos multidimensionais
QBIC	Imagem Vídeo	Manual Automática Semi-Automática	Semântica Similaridade	<i>Clustering</i>
Piction	Imagem	Automática	Semântica	Processamento de imagem
MARS	Imagem	Manual Automática	Similaridade Condicional	Baseado em lógica <i>Fuzzy</i>
MetaMídia	Qualquer tipo de mídia	Manual Automática Semi-automática	Identificador Condicional	<i>Clustering</i>

Tabela 3: Análise Comparativa dos Modelos/Sistemas

As linhas da Tabela 3 correspondem aos trabalhos analisados, enquanto que as colunas correspondem aos atributos que foram considerados relevantes no estudo. O (*) que aparece na Tabela 3, corresponde ao fato de não ter sido mencionado na literatura analisada. A recuperação em OVID é feita por uma linguagem de consulta visual. A seguir, são feitas algumas considerações a respeito de cada item citado.

Em geral, os sistemas tratam de recuperação de informação do tipo imagem. Os sistemas OVID e QBIC tratam informações do tipo vídeo. O sistema VIMSYS trata imagem para reconhecimento de faces. Já o *MetaMídia* não faz distinção entre os vários tipos de mídia. A hierarquia de metadados/valores pode ser usada e o tratamento é dado a qualquer tipo de mídia.

Cada modelo indexa (descreve) seus objetos de forma particular. O sistema QBIC, por exemplo, indexa fotografias partindo de um conjunto de atributos e vai adequando ou acrescentando atributos conforme a necessidade, durante a consulta.

A forma mais utilizada de indexação é a manual, onde os usuários constroem o conjunto mínimo de atributos necessários para indexação de objetos multimídia. Porém, alguns sistemas possuem várias formas de indexação, incluindo a automática ou a semi automática. Esse é o caso do VIMSYS, que permite reconhecimento de faces a partir dos atributos inseridos.

O *MetaMídia* permite a indexação manual. Uma vez definido o conjunto de metadados/valores para uma aplicação particular, os mesmos são utilizados. O modelo de metadados não é restrito em sua estrutura, ele permite que outros metadados/valores sejam incorporados e nada impede que estruturas de indexação existentes sejam adicionadas ao sistema.

A estrutura básica do modelo (meta-objeto multimídia) permite que isso seja feito, através da definição de um metadado e de valores para o mesmo. Assim, não impede que indexação automática ou semi-automática seja incorporada ao modelo, não afetando a

estrutura básica do mesmo. Por exemplo, supondo que fosse feita a indexação automática em uma imagem para extração das cores da mesma. O metadado *cor* com todas com todos os seus valores (cores existentes) daquela imagem seriam criados.

O processo de recuperação de informação pode ser feito através de identificador, condicional, semântica e por similaridade. O *MetaMídia* tem implementado a busca por identificador e condicional. O único modelo que aborda os quatro métodos de recuperação é o VIMSYS. Maiores detalhes são explorados na próxima seção, quando são analisados os sistemas do ponto de vista de evolução.

Os algoritmos que implementam os métodos de recuperação de cada um dos sistemas são baseados em *clustering*, processamento de imagens, baseado em lógica *fuzzy* ou segmentação de imagem. O *MetaMídia* propõe um algoritmo baseado em *clustering*, mas possui uma fase de pré-processamento para verificar se existem descrições ou não para o *cluster* a ser pesquisado.

O modelo *MetaMídia* tenta uniformizar as diferentes visões subjetivas de uma mídia por pessoas diferentes, através da hierarquia de metadados/valores. Através desta hierarquia os usuários podem descrever as mídias, não distanciando-se dos metadados/valores utilizados.

5.4 Do VIMSYS ao *MetaMídia*

O modelo *MetaMídia* apresentado, nesta tese, pode tratar qualquer tipo de mídia, seja vídeo, imagens estáticas, som, entre outras, podendo ser utilizado em qualquer aplicação não convencional.

Nesse modelo, a indexação é feita de forma manual, através do uso de metadados/valores que são pré-definidos pelos usuários. Além disso, esses metadados podem ser acrescidos de outros metadados/valores, caso não sejam suficientes para descrever o objeto multimídia. O modelo possui abstrações que permitem descrever informações de forma organizada.

Para ressaltar a relevância e importância do modelo *MetaMídia*, nesta seção é feita uma análise comparativa com os sistemas descritos, procurando enfatizar a evolução durante a última década.

Na Figura 26, é analisada a evolução dos modelos, desde o VIMSYS até o modelo *MetaMídia*. A análise é feita em função da estrutura de representação, indexação e recuperação de informação multimídia. Além de analisar os tipos de objetos multimídia que os modelos tratam e recuperam.

A interseção Sistema/ano representa uma publicação relacionada ao sistema em questão. A linha cheia mostra a evolução do sistema ao longo do tempo. A seta pontilhada representa aspectos relacionados ao modelo *MetaMídia* e que influenciaram o desenvolvimento do mesmo.

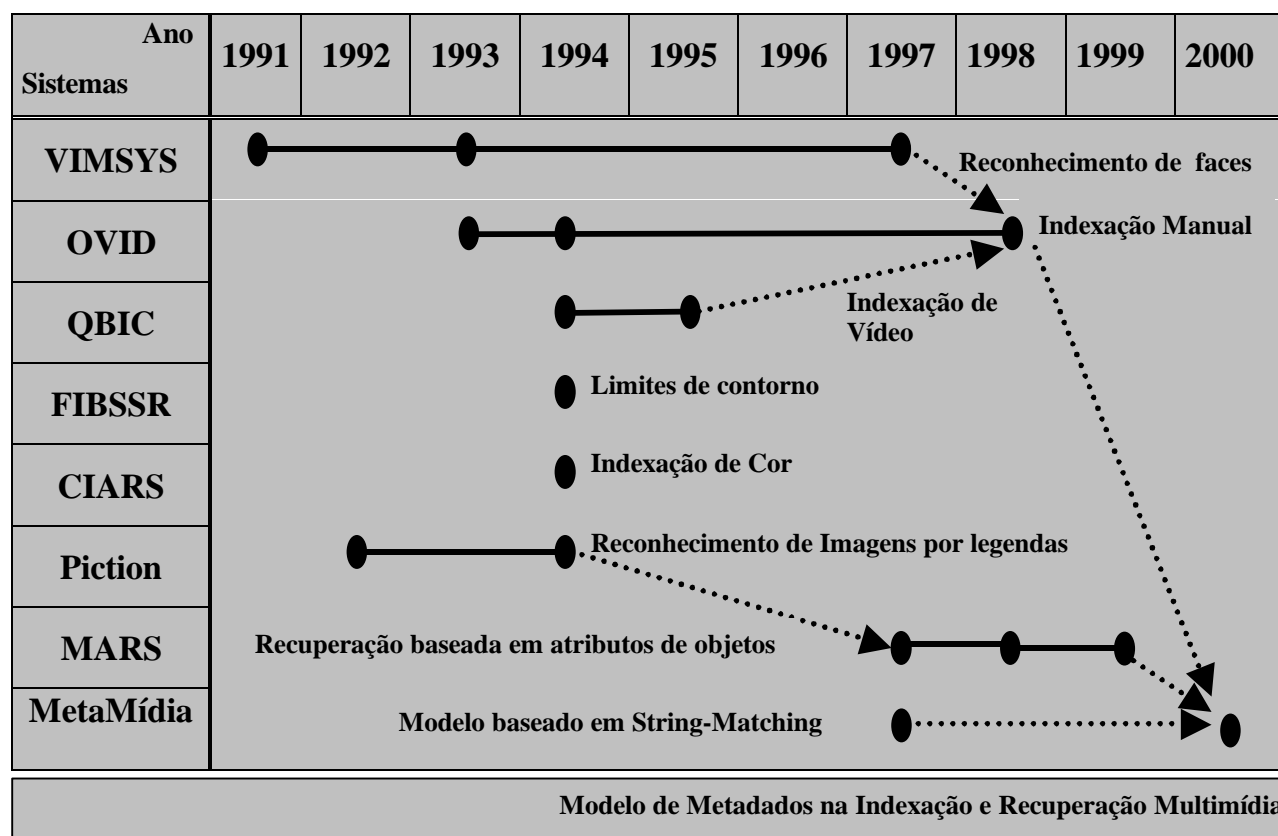


Figura 26: Do VIMSYS ao MetaMídia

O modelo VIMSYS apresentou no início da década um sistema para reconhecimento de faces, que incrementalmente acrescentava atributos para recuperação de imagens. O *MetaMídia* permite que sejam adicionados novos metadados para recuperar objetos multimídia. À medida que novos metadados e valores são acrescentados, um novo conjunto de objetos é retornado. O *MetaMídia* pode facilmente ser utilizado para reconhecimento de faces, desde que o conjunto de metadados seja definido pelo usuário ou extraído automaticamente das imagens, tornando o *MetaMídia* flexível. O VIMSYS parte de um conjunto padrão e a partir dele vai construindo a face.

O sistema OVID possui um mecanismo de herança apenas entre atributos definidos para intervalos de vídeo. O modelo *MetaMídia* permite a herança entre meta-objetos

multimídia para qualquer tipo de mídia, desde que esse contenha o novo objeto. O *MetaMídia* traz como original a herança entre instâncias. Atualmente, os sistemas convencionais permitem identificar um conjunto de objetos através da definição de um atributos comuns que os identificam. No caso desta tese, não é definido um conjunto de atributos para uma classe de objetos, mas sim para cada objeto, que corresponde ao meta objeto multimídia.

O modelo *MetaMídia* permite indexar qualquer tipo de mídia. O modelo é original no sentido de integrar representação, indexação e recuperação de objetos multimídia em um mesmo enfoque. O Sistema *Piction* faz reconhecimento de imagens através de legendas.

O modelo proposto nesta tese não apresenta metadados/valores extraídos automaticamente dos objetos multimídia, como é o caso do MARS. Uma característica importante do MARS e que o *MetaMídia* apresenta é a recuperação baseada em atributos de objetos.

A recuperação é baseada no casamento de padrões, onde o padrão é um subconjunto dos metadados utilizados para descrever o objeto multimídia armazenado. O método de pesquisa e recuperação tem como vantagem não pesquisar uma parte substancial do banco de dados, mas sim apenas parte dele, mesmo para o pior caso. Isso é alcançado devido à indexação de informação ter sido organizada em clusters no modelo *MetaMídia*.

O método de recuperação proposto tem um algoritmo de pesquisa implementado, baseado na estrutura de clusters, e utiliza uma representação de *array* triangular que otimiza o acesso aos clusters.

5.5 Considerações Finais

Neste capítulo foram descritos e analisados os principais trabalhos relacionados com o tratamento de dados multimídia. O enfoque foi dado aos tipos de mídia tratados, o modelo de indexação, método de recuperação e os algoritmos que implementam o método de recuperação. Uma análise comparativa foi feita entre esses trabalhos.

Os metadados são definidos para cada item de dado e a recuperação é feita tendo como base as suas descrições e não o item multimídia. Assim, os algoritmos de *string matching* podem ser adequados devido ao fato de ser efetuada a pesquisa considerando as descrições, as quais podem ser definidas de forma textual. Os metadados são considerados *strings* no processo de recuperação. A representação multimídia é genérica a ponto de permitir a descrição de qualquer tipo de mídia

Finalmente, uma análise comparativa do modelo *MetaMídia* foi feita com os principais modelos, desde o VIMSYS, que surgiu no início dos anos 90. Esta análise permitiu mostrar o que o modelo *MetaMídia* tem de inédito e relevante para o estado da arte.

Capítulo 6

Conclusões e Trabalhos Futuros

*Concluindo, a nova economia não está
na tecnologia, seja ela o microship ou
a rede mundial de telecomunicações.*

Está na mente humana.

Alan Webbe.

Neste capítulo são apresentadas as conclusões obtidas no desenvolvimento desta Tese. As principais contribuições, a originalidade desse trabalho, relevância do mesmo e as perspectivas de trabalhos futuros são evidenciadas.

6.1 Enfoque Geral

Um problema muito importante em Sistemas de Banco de Dados Multimídia, atualmente, é prover mecanismos que permitam que os usuários localizem as informações desejadas.

A recuperação da informação multimídia necessita de um modelo comum capaz de acomodar estrutura de informação semanticamente rica.

Neste trabalho, foi proposto um modelo de metadados na indexação e recuperação de objetos multimídia. Tal modelo de metadados é baseado em uma hierarquia de metadados/valores.

O método de recuperação é implementado com um algoritmo de *clusters* com índices, usando metadados, como descrito no Capítulo 3. No Capítulo 4 vários trabalhos foram analisados e comparados com o modelo *MetaMídia*. Um enfoque de evolução foi apresentado e as vantagens do modelo foram destacadas.

Neste capítulo final são abordados os objetivos alcançados durante o desenvolvimento desta tese, sua contribuição e os aspectos relacionados a trabalhos futuros.

6.2 Objetivos Alcançados

O modelo proposto permite definir os critérios de indexação, pesquisa e recuperação de objetos multimídia, e atinge os seguintes objetivos:

6.2.1 Descrição de Objetos Multimídia

Na modelagem, objetos podem ser especificados sem que o indexador conheça algum(s) valor(es) de metadados do domínio. Além disso, podem existir alguns objetos multimídia que não possuem evidências sugestivas para serem descritos e outros podem gerar

descrições muito resumidas. Isso ocorre, principalmente, na descrição de objetos em fase de preparação, onde algumas partes do mesmo ainda podem ser desconhecidas.

É permitida a criação de um ou mais meta-objetos multimídia em diferentes estágios de criação do objeto multimídia. Um meta-objeto multimídia pode representar um objeto mais significativamente do que um meta-objeto multimídia criado anteriormente.

A descrição de informações estruturadas é permitida, visto que o gerenciamento de grandes quantidades de dados requer instrumentos adequados para a organização de partes desses dados. É disponibilizada uma representação que permite a descrição estruturada de informações. A dificuldade de obter descrições coerentes exige que se padronize, na medida do possível, os metadados e os valores que eles podem assumir.

Os objetos multimídia são descritos através de mecanismos de reuso de metadados/valores, que podem ser importados de outros metaobjetos multimídia, feito através de herança de metadados/valores.

Esta tese tem como consequência direta, permitir descrever com uma semântica mais precisa o modelo e as operações de criação, conjunção, disjunção e herança de meta objetos.

6.2.2 Compartilhamento de Metadados

Compartilhamento de metadados/valores entre múltiplos meta-objetos multimídia, visto que diferentes usuários estão interessados em diferentes componentes ou diferentes

apresentações do mesmo documento. Através das estruturas de abstração é permitido compartilhamento de metadados/valores entre múltiplos objetos.

A representação de objetos multimídia, usando metadados/valores, permite a recuperação baseada na composição de objetos. Além disso, abstrações tais como generalização e especialização de metadados/valores podem ser utilizadas para descrever os objetos multimídia.

6.3 Contribuição Principal

Esta tese tem como principal contribuição, um modelo de metadados na indexação e recuperação de objetos multimídia. Esse modelo possui operações para criação e manipulação de meta-objetos multimídia. O meta-objeto multimídia corresponde a uma descrição de mídia.

A descrição pode ser feita para qualquer tipo de mídia e através dela pode ser recuperado qualquer tipo de objeto multimídia, de forma natural e transparente. Com o modelo de indexação, o usuário efetuará menos consultas desnecessárias, visto que o resultado da consulta será o agrupamento de objetos.

O agrupamento de meta-objetos em *clusters* permite menor tempo de busca e menos informações são analisadas quanto a sua relevância. O desempenho na busca é linear em alguns casos. Assim, através do modelo, facilidades são oferecidas para pesquisa e recuperação de informação multimídia, através da composição de objetos.

O modelo permite que seja feita a herança entre instâncias de meta-objetos multimídia. Em orientação a objetos, existe a herança de estrutura e comportamento. O modelo permite a herança entre instâncias.

Os meta-objetos multimídia representam significativamente um objeto multimídia. Uma vez alterado o objeto multimídia, novos meta-objetos são criados, não substituindo os meta-objetos anteriores. Os novos meta-objetos herdam os metadados/valores dos objetos anteriores. O modelo de metadados não é afetado em função das mudanças ocorridas. Assim, a evolução e edição dos objetos multimídia implicam em mudanças do seu conteúdo, porém não significa que os meta-objetos multimídia ficam não atualizados.

6.4 Trabalhos Futuros

Como perspectivas de trabalho futuro sugere-se a validação das idéias apresentadas nesta tese através do uso de uma base real objetos e meta-objetos multimídia, e usando tipos de mídias disponibilizadas na Internet. Outro aspecto importante a ser tratado é a formalização formal do modelo usando uma linguagem de especificação para descrever os meta-objetos multimídia. Além desses aspectos o modelo de metadados pode ser evoluído a ponto de ser tratado em informática na educação, recuperação baseada em similaridade e mecanismo de versões para meta-objetos multimídia. Abaixo são descritos os aspectos relacionados a cada uma dessas áreas.

6.4.1 Modelo de Indexação e Processo de Ensino/Aprendizagem

Estudos envolvendo o modelo aqui apresentado podem ser aprofundados sob o ponto de vista educacional [GM98], de modo a auxiliar tanto o aluno no exercício do aprendizado, como também auxiliar na avaliação do próprio processo de ensino/aprendizagem.

As novas tecnologias de informação apresentam perspectivas para a educação. Considerando que essas tecnologias estão influenciando todas as áreas da sociedade, a educação precisa atender às expectativas provocadas pelas mudanças sociais, enfatizando a interação criativa, o pensamento crítico e o julgamento de valores, de modo a ajudar os alunos a desenvolverem a capacidade de aprender, condição fundamental para a vida nas décadas futuras.

O processo de aprendizado não é uma atividade linear, com começo, meio e fim bem determinados, mas sim, uma atividade dinâmica onde, a partir de um conhecimento inicial, vão se agregando continuamente modificações sugeridas pela aplicação prática dos conhecimentos.

A utilização do modelo abrange situações variadas e entre essas situações pode-se registrar o ciclo de desenvolvimento de objetos criados por alunos, visto que permite a composição de objetos já existentes através das operações apresentadas no modelo.

Uma ferramenta pode ser criada para permitir a avaliação do processo educacional ao longo da hierarquia de documentos, com o objetivo de analisar a capacidade de aprender, criar, integrar e explorar o conhecimento.

Algumas das vantagens adquiridas serão: melhor análise do processo educacional e melhor aplicação dos conceitos pelos alunos. Como universo inicial de estudo, podem ser tomados professores, alunos e a realidade educacional do ensino médio brasileiro.

6.4.2 Método de Recuperação Baseado em Similaridade

Atualmente, o método de recuperação é baseado no casamento exato de padrões. A definição de medidas de similaridade para o método de recuperação pode ser alvo de estudo.

A pesquisa de objetos multimídia, baseado no *MetaMídia*, é feita em função dos metadados/valores atribuídos à consulta. Todos os objetos multimídia que possuem pelo menos os metadados/valores indicados são recuperados. Vale ressaltar que, os objetos multimídia que não tiverem pelo menos um dos metadados/valores indicados, não são recuperados.

A implementação do método de recuperação baseado em similaridade permitirá que esses objetos sejam recuperados. Mas para tanto, medidas devem ser definidas com critérios de busca e recuperação.

6.4.3 Mecanismo de Versões

O conceito de versão num Sistema de Banco de Dados pode ser definido como: descrição de um objeto num determinado momento; descrição completa de uma entidade, sob um determinado ponto de vista; diferentes representações de um mesmo objeto; ou diferentes estágios de desenvolvimento de um objeto.

A natureza das aplicações multimídia requer flexibilidade dos sistemas, de modo a permitir a existência de várias versões de um mesmo objeto.

O uso de versões apresenta a vantagem de permitir atualizações não destrutivas, gerando novo objeto com novo valor em vez da substituição de valores antigos. Além disso, a seqüência de versões de um objeto pode ser utilizada para fins de consultas.

Um mecanismo de versões para o modelo de indexação de informação apresentado nesta tese pode ser criado. A vantagem da criação de tal mecanismo é que os usuários utilizando na consulta também versões de meta-objetos multimídia, o resultado da pesquisa pode ter ainda mais sucesso.

Referências Bibliográficas

- [AS94] J. T. Anderson and M. Stonebraker. SEQUOIA 2000 metadata schema for satellite images. *SIGMOD RECORD*, 23(4):42-48, December 1994.
- [Ass98] N. A. Assimakopoulos. Systemic approach for a multimedia presentation system with pre-planned knowledge rules. *Proceedings of World Multiconference on Systemics, Cybernetics and Informatic, and 4th International Conference on Information Systems, Analysis and Synthesis*, Orlando, Florida, 3:1-8, July 1998.
- [AZP96] F. Aigrain, H. Zhang, and D. Petkovic. Content-based representation and retrieval of visual media: A state-of-the-art review. *Multimedia Tools and Applications*, 3(3):179-202, November 1996.
- [BMK95] G.P. Babu, B. M. Mehtre, and M. S. Kankanhalli. Color indexing for efficient image retrieval. *Multimedia Tools and Applications*, 1(4):327-348, November 1995.
- [BPJ93] J. R. Bach, S. Paul, and R. Jain. A visual information management system for the interactive retrieval of faces. *IEEE Transactions on Knowledge and data Engineering*, 5(4):619-628, August 1993.
- [BR94] K. Böhms and T. C. Rakow. Metadata for multimedia documents. *SIGMOD RECORD*, 23(4):21-26, December 1994.

-
- [BSH00] S. Bradshaw, A. Scheinkman and K. Hammond. Guiding people to information: providing an interface to a digital library using reference as a basis for indexing. *Proceedings of the 2000 international conference on Intelligent user interfaces*. Pages 37-43. January 2000.
- [CC98] G. Cha and C. Chung. Object-oriented retrieval mechanism for semistructured image collections. *Proceedings of the 6th ACM international conference on Multimedia*. Bristol United Kingdom, pages 323-332, September 1998.
- [CH97] G. Cornell and C. S. Horstman. Core Java. Sunsoft Press, Makron Books, 1997.
- [CL96] J. Cha and S. Lee. COMID: Composite icon browser for multimedia databases. *Multimedia Tools and Applications*, 3(3):203-224, november 1994.
- [CLR90] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. Introduction to algorithms. *MIT Press*, 1990.
- [CNY98] M. T. Chan, T. S. Ng, and N. H. C. Yung. MVM – a multimedia virtual machine for design modeling and performance simulation. *Proceedings of World Multiconference on Systemics, Cybernetics and Informatic, and 4th International Conference on Information Systems, Analysis and Synthesis*, Orlando, Florida, 3:9-15, jully 1998.
- [Con99] Jim Conallen. Modeling WEB Application Architectures with UML. *Communication of the ACM*, 42(10):63-70, October 1999.

-
- [CSBB97] S. Chang, J. R. Smith, M. Beigi, and A Benitez. Visual information retrieval from large distributed online repositories. *Communications of the ACM*, 40(12):63-71, December 1997.
- [Dat99] C. J. Date. An Introduction to Database Systems. Addison-Wesley Book, Seventh Edition, 1999.
- [EN99] R. Elmasri and S. B. Navathe. Fundamentals of Database Systems. Addison-Wesley Book, Third Edition, 1999.
- [FBF+94] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic and W. Equitz. Efficient and Effective querying by image content. *Journal of Intelligent Information Systems 3*, Pages 231-262, 1994.
- [FSN+95] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Styele, and P. Yanker. Query by image and video content: the QBIC system. *IEEE COMPUTER Innovative Technology for computer Professionals – Finding the Right Image – Content-based Image Retrieve Systems*, pages 23-31, September 1995.
- [GCJ00] R. Gonzalez, G. Cranitch and J. Jo. Academic directions of multimedia education. *Communications of the ACM*. Vol. 43, pages 89-95, 2000.
- [GD98] F. Golshani and N. Dimitrova. A Language for ContentBased Video Retrieval. *Multimedia Tools and Applications 6*, pages 289-312, 1998.

-
- [GFSC94] W. I. Grosky, F. Fotouhi, I. K. Sethi, and B. Capatina. Using metadata for the intelligent browsing of structured media objects. *SIGMOD RECORD*, 23(4):49-56, December 1994.
- [GM98] R. Goularte and E. S. Moreira. Helping authoring educational multimedia material through the use of metadata-based browsing tools. *Proceedings of World Multiconference on Systemics, Cybernetics and Informatic, and 4th International Conference on Information Systems, Analysis and Synthesis*, Orlando, Florida, 3:24-29, July 1998.
- [Gor00] A. S. Gordon. Using annotated video as an information retrieval interface *Proceedings of the 2000 international conference on Intelligent user interfaces*. Pages 133-140. January 2000.
- [Gro97] W. I. Grosky. Managing multimedia information in database systems. *Communications of the ACM*, 40(12):72-80, December 1997.
- [GSJ97] A. Gupta, S. Santini, and R. Jain. In search of information in visual media. *Communications of the ACM*, 40(12):34-42, December 1997.
- [GWJ91] A. Gupta, T. Weymouth, and R. Jain. Semantic queries in image databases. *IFIP 2nd Working Conference of Visual DBS, Budapest, Hungary*, September 1991.

-
- [HMM00] W. Hürst, R. Müller and C. Mayer. Multimedia information retrieval from recorded presentations. *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*. Pages 339-341. July 2000.
- [HSHA98] S. Hollfelder, F. Schmidt, M. Hemmje, and J. Aberer. Transparent integration of continuous media support into a multimedia DBMS. *International Workshop on Issues and Applications of Database Technology, Berlin, Germany, July 1998*.
- [HSS96] A. Hoff, S. Shaio, and O. Starbuck. *Ligado em Java*. Makron Books, 1996.
- [Inm99] W. H. Inmon. *Gerenciando Data Warehouse*. 1. ed. São Paulo: Makron Books, 1999.
- [JH94] R. Jain and A. Hampapur. Metadata in video databases. *SIGMOD RECORD*, 23(4):27-33, December 1994.
- [KKH94] Y. Kiyoki, T. Kitagawa, and T. Hayama. A meta-database system for semantic image search by a mathematical model of meaning. *SIGMOD RECORD*, 23(4):34-41, December 1994.
- [Kow97] G. Kowalski. *Information Retrieval Systems: Theory and Implementation*. Kluwer Academic Publishers, 1997.

-
- [KSC99] Y. Kim, C. Sim and J. Chang. Spatial match representation scheme supporting ranking in iconic images databases. *Proceedings of the eighth international conference on Information knowledge management*, Pages 450 – 457, November 1999.
- [Lar99] G. Larsen. Designing component-based frameworks using patterns in the UML. *Communications of the ACM*, 42(10):38-45, October 1999.
- [LBN+94a] D. Lee, R. Barber, W. Niblack, M. Flickner, J. Hafner, and D. Petkovic. Query by image content using multiple objects and multiple features: User interface issue. *In Proceedings of the ICIP*, 1994.
- [LBN+94b] D. Lee, R. Barber, W. Niblack, M. Flickner, J. Hafner, and D. Petkovic. Indexing for complex queries on a query by content image database. *In Proceedings of the ICPR*, 1:142-146, 1994.
- [LZ96] D. Lucarella and A. Zanzi. A visual retrieval environment for hypermedia information systems. *ACM transactions on Information Systems*, 14(1):3-29, January 1996.
- [Mar01] J. M. Martinez. Overview of the MPEG-7 Standard. International Organisation for Standardisation. Singapore, March 2001.
- [MCS93] S. R. L. Meira, A. L. C. Cavalcanti, and C. S. Santos. *The Unix File System: A MooZ Specification*. In K. Lano and H. Haughton, editors, Object-Oriented Specification Case Studies, Object-Oriented Series. Prentice-Hall, 1993.

-
- [Mel96] J. C. B. Melo. Diferentes abordagens em comparação de seqüências. Dissertação de Mestrado, Universidade Federal de Pernambuco, 1996.
- [MG95] R. Mehrotra and J. E. Gary. Similar-shape retrieval in shape data management. *COMPUTER Innovative Technology for Computer Professionals – Finding the Right Image – Content-Based Image Retrieve Systems*, pages 57-62, 1995.
- [Nar96] A. D. Narasimhalu. Multimedia databases. *Multimedia Systems*, 4:226-249, 1996.
- [NBE+93] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic and P. Yanker. The QBIC Project: Querying Images by content using color, texture and shape. *SPIE Conference on Storage and Retrieval for Image and Video Databases*, pages 173-187, 1993.
- [Oom94] E. Oomoto. Study on data models and visual query languages for multimedia and historical databases. *Doctoral Dissertation, Kobe University*, January 1994.
- [ORC+97] M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, and T. S. Huang. Supporting similarity queries in MARS, *Proceedings of ACM Multimedia '97, Seattle, Washington*, pages 403-413, Seattle, Washington, November 1997.
- [OT93] E. Oomoto and K. Tanaka. OVID: design and implementation of a video object database system. *IEEE Transactions on Knowledge and Data Engineering*, 5(4):629-643, August 1993.

-
- [OT97] E. Oomoto and K. Tanaka. Video database systems – recent trends in research and development activities. *The Handbook of Multimedia Information Management, Prentice Hall*, 1997.
- [PC99] K. Porkaew and K. Chakrabarti. Query refinement for multimedia similarity retrieval in MARS. *Proceedings of the conference on ACM multimedia '99*, Pages 235-238, Orlando, Flórida, October 1999.
- [PF95] E. G. M. Petrakis and C. Faloutsos. Similarity searching in large image databases, *Technical Report CS-TR-3388, University of Maryland*, 1995.
- [RHM98] Y. Rui, T. S. Huang, and S. Mehrotra. Exploring video structure beyond the shots. *Proceedings of IEEE International Conference on Multimedia Computing and Systems (ICMCS)*, June 1998.
- [RNL95] T. C. Rakow, E. J. Neuhold, and M. Loehr. Multimedia database systems – the notions and the issues. *Datenbanksysteme in Büro, Technik und Wissenschaft (BTW), GI-Fachtagung, Dresden*, pages 1-29, March 1995.
- [SD99] D. Sitaram and A. Dan. Multimedia Servers: Applications, Environments and Design. *Morgan Kaufmann Publishers*, October 1999.
- [SKS99] A. Silberschatz, H. F. Korth and S. Sudarshan. *Sistema de Banco de Dados*. Terceira Edição. São Paulo: Makron Books, 1999.
- [SL95] R. Steinmetz and B. Lindsay. *Multimedia: Computing, Communications, and Applications*. Prentice Hall, 1995.

-
- [Sri92] R. K. Srihari. Extracting visual information from text: using captions to label faces in newspaper photographs. Doctoral Dissertation. *CEDAR – Center of Excellence for Document Analysis and Recognition*, 1992.
- [Sri95] R. K. Srihari. Automatic indexing and content-based retrieval of captioned images. *COMPUTER Innovative Technology for computer Professionals – Finding the Right Image – Content-Based Image Retrieve Systems*, pages 49-56, 1995.
- [TOV+95] H. Treat, E. Ort, M. Vo, J. Jang, L. Hall, F. Tung and D. Petkovic. Searching images using ultimedia manager. *In SPIE Proceedings Storage and Retrieval for Image and Video Databases III*, pages 204-213, 1995.
- [VAB95] M. Volz, K. Aberer, and K. Boelm. A flexible approach to combine IR semantics and database technology and its application to structured document handling. *GMD Technical Report N. 891, Sankt Augustin*, January 1995.
- [Vaz97] M. S. M. G. Vaz. Recuperação da Informação Multimídia através de algoritmos de string matching. *Exame de Qualificação, Universidade Federal de Pernambuco*, 1997.
- [VVSS98] M. S. M. G. Vaz, A. M. L. Vasconcelos, F. F. Souza, e A. C. Salgado. Descrevendo um modelo de informação de vídeo através de uma linguagem de especificação orientada a objetos. *IDEAS'98 – Workshop Iberoamericano de Engenharia de Requisitos e Ambientes de Software*, 1998.

-
- [WS99] M. Wechsler and P. Schäuble. A new ranking principle for multimedia information retrieval.. *Proceedings of the fourth ACM conference on Digital libraries*. Pages 146-151, August 1999.
- [YN99] R. Baeza-Yates and B. Ribeiro-Neto. Modern Information Retrieval. *Addison-Wesley-Longman*, May 1999.
- [YY97] B. Yeo and M. M. Yeung. Retrieving and visualizing video. *Communications of the ACM*, 40(12):43-52, December 1997.
- [ZCA96] A. Zhang, B. Cheng, and R. Acharya. A fractal-based clustering approach in large visual database systems. *Multimedia Tools and Applications*, 3(3):225-244, November 1996.