

VIII ENANCIB – Encontro Nacional de Pesquisa em Ciência da Informação
28 a 31 de outubro de 2007 • Salvador • Bahia • Brasil

GT 2 – Organização e Representação do Conhecimento
Comunicação Oral

**ONTOLOGIAS: INDEXAÇÃO E RECUPERAÇÃO DE FOTOGRAFIAS
BASEADAS NA TÉCNICA FOTOGRÁFICA E NO CONTEÚDO DA
IMAGEM**

***ONTOLOGIES: PHOTOGRAPHIC INDEXING AND RETRIEVE
BASED ON PHOTOGRAPHIC TECHNIQUE AND IMAGE CONTENT***

Miriam Paula Manini (PGCInf/UnB, mpmanini@uol.com.br)
Mamede Lima-Marques (PGCInf/UnB, mamede@unb.br)
Alex Sandro Santos Miranda (PGCInf/UnB, alexsandromiranda@gmail.com)

Resumo: Este trabalho aborda o registro e a recuperação de imagens no contexto de sistemas informáticos. Levanta os aspectos de transferência da informação contida na imagem para a representação verbal e a maneira como os padrões da *web* semântica podem ser usados para indexar e recuperar fotografias. Ontologias são usadas como um tesouro com a finalidade de auxiliar o indexador no processo de seleção de conceitos para indexar fotografias; além disso, as ontologias são usadas no mecanismo de recuperação como alternativa ao paradigma de recuperação de imagens com base em palavras-chave. Enfoca a indexação e a recuperação de imagens sob o ponto de vista da transformação da informação da imagem, que ocorre desde o momento da transposição da informação imagética para a informação verbal até o momento da busca pelo usuário final através de um sistema de busca de informações de imagens fotográficas. Uma ferramenta de busca de imagens foi construída, com base nas reflexões propostas e desenvolvidas.

Palavras-chave: indexação de fotografias, dimensão expressiva, web semântica, ontologia, ferramenta de busca.

Abstract: *This work approaches the register and the recovery of images in the context of computer science systems. It raises the aspects of transference of the information contained in the image for the verbal representation and the way as the standards of web semantics can be used to make the documentary analysis and to recoup photographs. Ontologies are used like a tesouro with the purpose of assisting the information professional in the process of election of concepts is used as one to make the documentary analysis of photographs; moreover, the ontologies are used in the mechanism of recovery as alternative to the paradigm of recovery of images on the basis of word-key. It focuses the indexation and the recovery of images under the point of view of the transformation of the information of the image, that occurs since the moment of the transposition of the visual information for the verbal information until the moment of the search for the final user through a system of search of information of photographic images.*

Keywords: *documentary analysis of photography, expressive dimension, web semantics, ontology, tool of search.*

Introdução

O foco deste trabalho é o processo de indexação e recuperação de imagens no contexto de sistemas informáticos. Abordamos tanto a indexação quanto a recuperação de imagens fotográficas, visando a melhorar a precisão da recuperação de imagens em sistemas de informação.

Existem atualmente duas abordagens para o problema da indexação de imagens: a primeira é conhecida como indexação com base no conteúdo e a segunda é conhecida como indexação com base em conceitos. Na abordagem que se baseia no conteúdo, a indexação é realizada de modo automático, por algoritmos computacionais, através da extração de características de cor, textura e arranjo de figuras geométricas presentes na imagem. Na indexação com base em conceitos, as imagens são representadas por uma lista de palavras-chave que se referem às informações presentes na imagem. A indexação por conteúdo opera no nível sintático, enquanto a indexação por conceitos opera no nível semântico.

Nosso enfoque é a indexação por conceitos. Neste trabalho apresentamos os aspectos mais importantes deste paradigma; além disso, descrevemos um modelo de indexação que tenta unir os avanços das pesquisas em indexação de imagens desenvolvidas pelas áreas de Processamento de Imagens e Ciência da Informação.

Discutimos os modelos de indexação de imagens propostos pela Ciência da Informação e vemos em que medida a Ciência da Computação pode contribuir para o aperfeiçoamento dos modelos de indexação de imagens através do desenvolvimento de metadados e linguagem de representação do conhecimento como ontologias. O processo de indexação e recuperação de imagens pode ser simplificado no seguinte diagrama:

Imagem → indexação → representação ← recuperação ← expressão de busca ← usuário.

Um dos objetivos deste trabalho é analisar este processo e apresentar uma modelagem conceitual para indexação e recuperação de imagens embasadas nos avanços das pesquisas sobre estes temas. Seguindo o diagrama, vimos que, inicialmente, a imagem passa pela etapa de indexação, ou seja, etapa de seleção de conceitos que representarão a informação contida na imagem.

Abordamos estes problemas: o que indexar, como indexar e como recuperar, sob o ponto de vista tecnológico; e desenvolvemos, ao longo deste trabalho, uma modelagem conceitual para indexação e recuperação de imagens levando em consideração três categorias de informações relacionadas à imagem: conteúdo informacional (o que a imagem mostra), técnica fotográfica e informação não-visual (fotógrafo, título, data, etc.).

Sobre o problema da indexação (o que indexar) existem métodos e técnicas que são empregadas para a extração de descritores ou de palavras-chave com base no conteúdo da imagem. Alguns pesquisadores têm explorado a questão da técnica fotográfica como recurso de indexação. Smit (1997, p. 3) categoriza três parâmetros de análise:

- o que a fotografia mostra (ou seja, seu conteúdo informacional);
- como a fotografia mostra (a forma adotada para mostrar o conteúdo informacional, ou seja, a expressão fotográfica);
- onde a fotografia mostra (dedicado à questão do documento fotográfico enquanto objeto físico).

Os dois primeiros itens – o que a fotografia mostra e como a fotografia mostra – estão estreitamente relacionados à indexação de imagens fotográficas, enquanto o último item não se encaixa neste estudo por não ser de interesse da área de indexação de imagens.

A maior parte dos mecanismos de busca de imagens fotográficas é indexada com base em informações extra-imagéticas, isto é, informações que não estão presentes na imagem, tais como: fotógrafo, título, data, descrição textual, etc. Embora estas informações sejam relevantes, elas não representam a informação presente na imagem. Para melhorar a recuperação de imagens fotográficas há a necessidade de representar o conteúdo informacional (o que a imagem mostra) e a sua expressão fotográfica (como a imagem mostra). A expressão fotográfica é produzida pelo

fotógrafo com o uso da técnica fotográfica. Duas imagens sobre o mesmo objeto podem ter uma recepção diferente por causa da técnica fotográfica empregada.

Objetivos

O objetivo geral deste trabalho é propor um modelo, com base em ontologia, para melhorar a descrição fotográfica com vistas a melhorar a recuperação das imagens através de uma ferramenta de indexação; elaborar uma ontologia de representação de imagens que permita a indexação e a recuperação com base no conteúdo da imagem e na técnica fotográfica.

Para atingir o objetivo principal, tivemos como objetivos específicos:

- realizar um estudo das técnicas de representação da imagem;
- comparar as técnicas de representação da imagem por conceitos e por conteúdo;
- descrever a importância da técnica fotográfica na representação da imagem;
- identificar descritores de imagens fotográficas com base no estudo das técnicas de representação da imagem;
- elaborar uma representação da imagem usando metadados;
- propor um esquema de descrição;
- propor um mecanismo de descrição com base neste esquema;
- definir as classes principais da ontologia para ser empregada na indexação e recuperação de imagens;
- desenvolver um protótipo de indexação e recuperação de imagens fotográficas.

Justificativa

Um dos grandes problemas dos sistemas de recuperação disponíveis na Internet é a baixa precisão da pesquisa, ou seja, muitos registros irrelevantes são recuperados no momento da busca; isto acontece, em parte, devido ao próprio paradigma de consulta com base em palavras-chave. Em se tratando de sistemas de recuperação de imagens, o problema agrava-se, pois a documentação de imagens exige técnicas que, na maioria das vezes, são desconhecidas pelos informatas. O problema é que as imagens se tornam virtualmente irrecuperáveis se não existir uma descrição associada a elas. E a maneira como esta descrição é realizada tem um enorme impacto na recuperação da imagem.

Ainda não existe um modelo de indexação que represente de modo preciso a informação presente numa imagem, mas existem técnicas que recuperam as informações mais relevantes. Uma vez obtida a lista de indexadores, surge uma nova questão: como representar os indexadores? Existem várias alternativas: palavras-chave, resumo textual, tesouro, ontologias, etc. A outra questão é: como o usuário formulará a sua expressão de busca?

O número de banco de imagens na Internet cresce de modo exponencial. Encontrar uma fotografia para uma necessidade particular é tarefa cada vez mais difícil. Ao tentar recuperar uma imagem em um sistema de busca, o usuário depara-se com o problema da quantidade excessiva de registros irrelevantes e com a dificuldade de selecionar termos adequados e mais precisos para formular a consulta.

A causa da quantidade excessiva de registros irrelevantes é, em parte, o próprio paradigma de busca com base em palavras-chave, que não trata adequadamente os termos sinônimos e antônimos; além disso, a indexação é realizada no nível do léxico, o que torna difícil obter resultados semânticos na recuperação da informação.

O processo de transferência da informação imagética para a informação verbal é negligenciado; conseqüentemente, a indexação da imagem é, muitas vezes, realizada levando em conta somente a legenda da fotografia, que, de acordo com as metodologias de documentação de imagem, é apenas uma pequena parte da informação que pode estar associada à imagem. Se a

indexação é realizada com base numa informação que não retrata a informação contida na imagem isso certamente ocasionará problemas na recuperação da informação.

Uma modelagem conceitual de dados que represente o conteúdo informacional, a expressão fotográfica e a informação não-visual somada a uma alternativa de recuperação que facilite a recuperação de imagens usando estas três categorias de informações pode tornar os sistemas de recuperação de imagens fotográficas mais eficientes e capazes de atender as necessidades específicas dos usuários de fotografias.

Ontologias podem ser consideradas uma boa alternativa para minimizar o problema da quantidade excessiva de registros irrelevantes, uma vez que a recuperação se baseia numa estrutura de relacionamentos semânticos e não mais na unidade léxica, como no mecanismo de busca com base em palavras-chave. As hierarquias de conceitos da ontologia podem ser utilizadas no momento da busca para auxiliar o usuário a selecionar e a combinar conceitos para recuperar a informação.

O usuário final interessa-se tanto pelo objeto fotografado quanto pela dimensão expressiva da imagem. Dimensão Expressiva é, por assim dizer, a parte da imagem fotográfica dada pela técnica: é a aparência física através da qual a fotografia expressa seu conteúdo informacional, é a extensão significativa da fotografia manifesta pela forma como a imagem se apresenta (revelada pela técnica). A dimensão expressiva introduz critérios para selecionar imagens com conteúdos idênticos, ou seja, serve para refinar a busca, introduzindo modificadores com relação ao conteúdo informacional.

Esses critérios que a dimensão expressiva introduz servirão de elemento para a recuperação de fotografias. No momento da recuperação, o usuário pode selecionar uma fotografia com base no conteúdo informacional, mas o sistema de recuperação pode, de modo inteligente, sugerir fotografias com base na dimensão expressiva da imagem.

Este trabalho se justifica, então, devido aos problemas de recuperação de imagens mencionados anteriormente e à falta de tratamento adequado na documentação de imagens no contexto de sistemas informáticos; além disso, leva em consideração a questão da dimensão expressiva, que ainda não foi empregada no contexto de sistemas informáticos e pode contribuir muito para facilitar a recuperação de imagens.

Metodologia e Desenvolvimento

A pesquisa consistiu, inicialmente, na sistematização dos descritores de imagens obtidos a partir da revisão de literatura sobre a representação da imagem. Abordamos os aspectos de indexação de imagens com base nos paradigmas de indexação de imagens por conceitos e por conteúdo. Com base neste estudo das técnicas de representação, especificamos uma modelagem conceitual que servirá para auxiliar a implementação de um metadado de descrição da imagem. Adicionalmente, é definida uma ontologia de representação da imagem com base no conteúdo da imagem e na técnica fotográfica. A construção desta ontologia foi embasada nas técnicas de representação da imagem por conceito. Ela servirá para auxiliar o indexador no momento da descrição da imagem por meio de uma ferramenta de indexação (*software*) e para auxiliar o usuário a formular uma expressão de busca num sistema de recuperação de imagens. A ontologia de representação da imagem fornecerá uma estrutura conceitual para representar o conteúdo da imagem.

Com o objetivo de viabilizar a construção do metadado de descrição da imagem, usando a ontologia de representação da imagem, desenvolvemos um protótipo de indexação e recuperação de imagens com base em ontologias. Este *software* – batizado de *Splash* – permite o indexador realizar a descrição da imagem usando a ontologia; além disso, permite ao usuário realizar uma consulta de imagens fotográficas selecionando conceitos dentro da ontologia. O registro da imagem gerada pelo protótipo é uma descrição da imagem em RDF, que agrega diferentes facetas da representação da imagem.

A seguir, apresentamos a parte que consiste no desenvolvimento da modelagem conceitual, na elaboração do metadado de descrição de imagens, no desenvolvimento de uma ontologia de representação do domínio de imagens e na construção de um sistema de indexação e recuperação de imagens.

Modelagem Conceitual

A modelagem conceitual consiste em estabelecer os conceitos, os atributos dos conceitos, e os relacionamentos que existem entre os conceitos da representação da imagem. Com base no que foi exposto sobre os modelos de indexação de imagens por conceito e por conteúdo, partimos para a elaboração de uma modelagem conceitual da imagem. Usamos a linguagem de modelagem de sistemas conhecida como UML (*Unified Modeling Language*), que possui muitos artefatos para representar conceitos, atributos e relacionamentos entre os conceitos, tais como hierarquia de classe, subclasse, composição de objetos, além de instrumentos que permitem visualizar o fluxo de dados dentro de um processo ou funcionalidade de um sistema.

A partir da classificação de metadados de Gilliland-Swetland (2000) em administrativo, descritivo, de uso, técnico e de preservação, realizamos a modelagem conceitual de descrição da imagem. As categorias administrativo, de uso, técnico e de preservação correspondem ao que chamamos de conteúdo não-visual da imagem, ou seja, aquelas informações que não estão presentes no conteúdo da imagem. Para o domínio de descrição de imagem, identificamos os seguintes descritores relacionados com cada categoria:

Administrativo – título, doador, copirraite, autor, dimensão, local, identificador, coleção, data e palavras-chave. A dimensão da imagem é a largura versus a altura, em centímetros. O identificador corresponde a um número ou uma URL que identifique a imagem unicamente. A coleção corresponde ao nome da coleção a que a imagem pertence. Palavras-chave consistem numa lista de palavras que representam o conteúdo da imagem;

De Uso – educativo, exposição, jornal e revista. Educativo corresponde a informações sobre o uso da imagem para fins de ensino. Exposição consiste de informações sobre o uso da imagem em exposição, como local, data e título da exposição;

Técnico – formato e resolução. Exemplos de formato: jpeg, gif, png, etc. Resolução corresponde à resolução da imagem digital em pixels;

De Preservação – legibilidade.

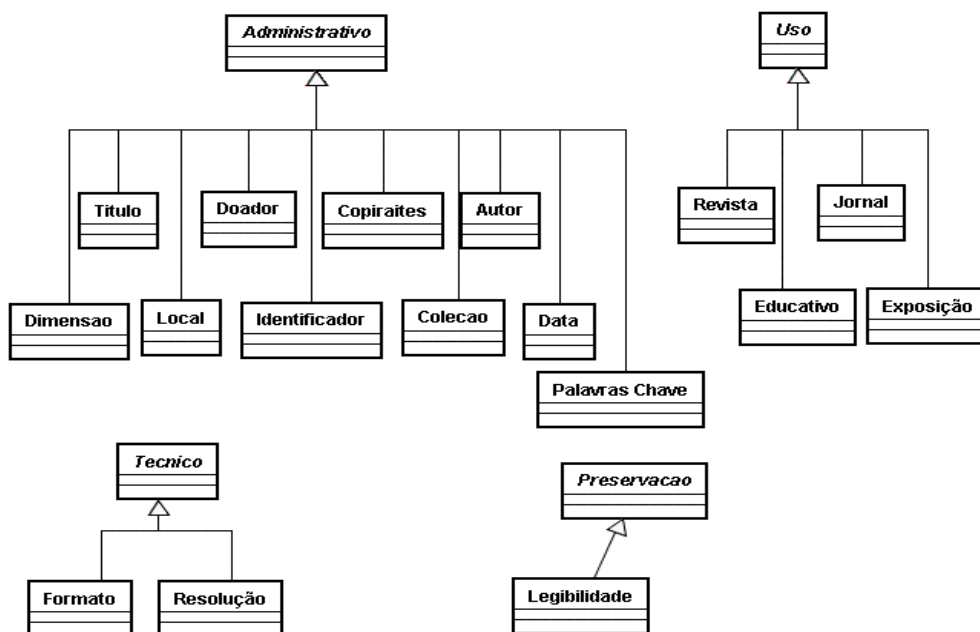


Figura 1: Diagrama de classe do conteúdo não-informacional da imagem.

As contribuições de Shatford (1986) quanto à indexação de imagens e a de James & Chang (2000) quanto à classificação de recursos visuais podem ser categorizadas dentro da classificação de Gilliland-Swetland (2000) como metadados descritivos, uma vez que são utilizados para descrever o conteúdo da imagem. O conteúdo informacional (o que a imagem mostra) denominamos de conteúdo semântico. A figura 2 mostra o diagrama de classe da representação sintática e semântica da imagem. Detalhando um pouco mais o conteúdo semântico por meio do diagrama de classe, podemos observar que os elementos O Que, Quando, Como, Onde e Sobre estendem a classe conteúdo semântico; isto é mostrado no diagrama com o uso de uma seta ligando o elemento ao conteúdo semântico. Este relacionamento é denominado de herança, também conhecido como É UM, ou seja, O QUE É UM conteúdo semântico, ONDE É UM conteúdo semântico.

A figura 2 mostra a representação da imagem por meio do diagrama de classe.

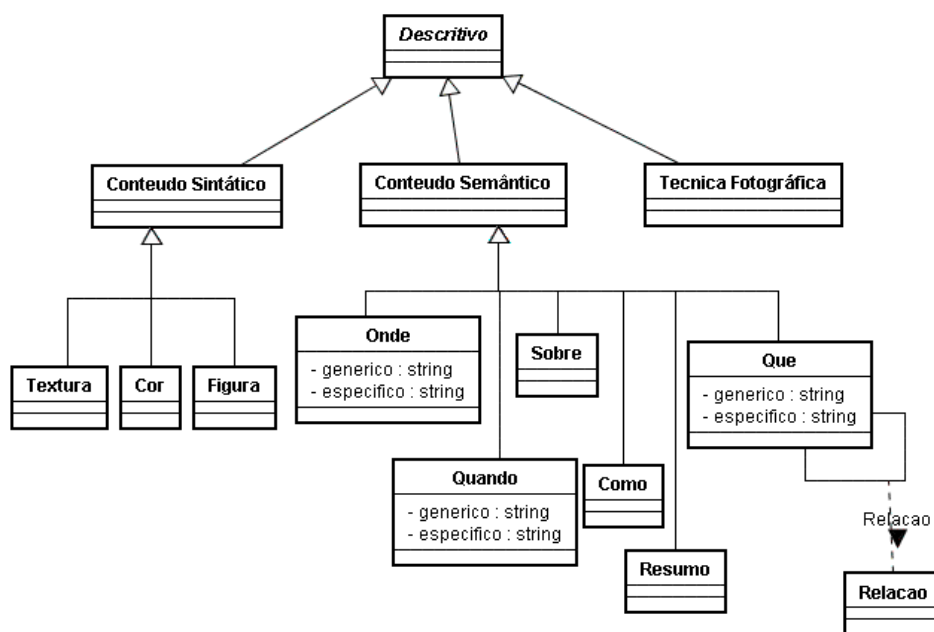


Figura 2: Diagrama de classe da representação da imagem.

O nível sintático é representado no diagrama pela classe conteúdo sintático e suas subclasses cor, textura e figura. A técnica fotográfica está representada pela classe técnica fotográfica.

Construção da ontologia de indexação e recuperação

Empregamos algumas técnicas do processo de construção de *software* para auxiliar no processo de desenvolvimento da ontologia, que é semelhante ao processo de construção de *software*, que apresenta, basicamente, as seguintes fases: análise de requisitos, especificação, delimitação do escopo, implementação e teste.

A análise de requisitos caracteriza-se pelo levantamento e entendimento das necessidades de informação do usuário. É necessário extrair do usuário suas necessidades de informação, ou seja, que tipo de informação o sistema deverá fornecer para o usuário a partir de formulários de pesquisa.

A especificação do sistema usando uma linguagem de representação de conhecimento: em sistemas orientados a objetos usa-se UML para a representação do domínio de aplicação.

O terceiro passo consiste em delimitar o escopo do *software*, isto é, estabelecer um número limite de funcionalidades que serão inicialmente implementadas de modo que seja viável a construção do sistema dentro de um determinado período de tempo.

A fase de implementação consiste na codificação do sistema usando uma linguagem de programação; é neste momento que entra o programador de computadores. A partir daí o sistema é testado e validado. Adicionalmente, vale ressaltar que o processo de desenvolvimento de *software* é iterativo, isto é, depois de atingir a última fase, de testes, o processo reinicia-se, ou seja, ele entra na fase de análise de requisito novamente e todas as fases sucessoras se repetem, até que sejam atendidos os requisitos do sistema.

Similarmente ao processo de construção de *software*, o desenvolvimento de uma ontologia passa pelas mesmas fases, e também é um processo iterativo. No nosso trabalho sobre a representação da imagem levantamos os seguintes requisitos da ontologia:

- fornecer uma estrutura conceitual para responder as questões da grade de análise documentária de imagens (QUEM/O QUE, QUANDO, COMO, ONDE e SOBRE) e para auxiliar na indexação da técnica fotográfica e do conteúdo não visual;
- servir como vocabulário comum entre o indexador e o pesquisador de imagens por meio do mecanismo de busca com base em categorias;
- auxiliar o indexador na seleção de termos que irão representar o conteúdo da imagem;
- auxiliar o usuário na pesquisa de imagens.

Antes de iniciarmos o desenvolvimento da ontologia deparamos com o seguinte problema: qual o escopo da ontologia? Uma imagem pode ser sobre qualquer coisa. A representação do domínio do conteúdo da imagem é ampla e impossível de representá-la completamente. Então nós nos baseamos na grade de análise documentária de imagens de Smit (1996) e Manini (2002) e determinamos que um dos requisitos da ontologia é fornecer respostas para questões apresentadas pela grade e, desta forma, guiar o processo de ontologização. Assim, determinamos as classes raízes da ontologia com base em algumas questões da grade de análise documentária de imagens. As classes raízes da ontologia são as seguintes:

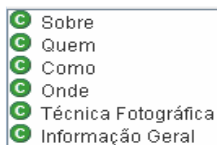


Figura 3: Classes raízes da ontologia.

Estas classes raízes representam os conceitos de mais alto nível de abstração; logo, o escopo da ontologia ficou determinado em representar cada categoria da grade de análise documentária de imagens. Isto conduziu o desenvolvimento da ontologia usando a análise *top down*. Esta técnica consiste em identificar e categorizar os conceitos do nível mais geral para o nível mais específico. A categoria Sobre foi desenvolvida usando categorias de metadados de várias fontes, dentre elas podemos citar o ODP *Open Directory Project* e o IPTC *International Press Communications Council*. O ODP é um mecanismo de busca baseado em categorias. As categorias do *Open Directory Project* são desenvolvidas pela contribuição de muitos editores espalhados pelo mundo. O projeto ODP, assim como é conhecido na comunidade da Internet, disponibiliza suas categorias especificadas em RDF para serem reusadas por outras aplicações.

A Figura 4 mostra o esquema de categorias do conteúdo SOBRE.



Figura 4: Representação da categoria Sobre.

Explorando mais um nível do esquema de categoria da classe principal Sobre temos a Figura 5.



Figura 5 Explorando as subcategorias da categoria Artes.

A categoria Quem, de acordo com a definição de Shatford (1986, p. 49) é usada para descrição de: animado e inanimado, objetos e seres concretos. Smit (1997, p. 4) define a categoria Quem como a identificação do “objeto focado”: seres vivos, artefatos, construções, acidentes naturais, etc. A autora também argumenta que a categoria Quem representa respostas das seguintes questões: de Quem é esta imagem? De que Objetos? De que Seres? A partir destas definições identificamos os seguintes conceitos derivados da categoria Quem: Animal, Construções, Fenômeno da Natureza, Inseto, Objeto, Paisagem, Pessoa e Vegetal.

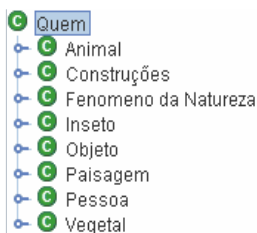


Figura 6: Representação da categoria Quem.

A categoria Como representa, segundo Shatford (1986, p. 49), o que os seres ou objetos na imagem estão fazendo; são ações, eventos e emoções. “Descrição de atitudes ou detalhes relacionados ao ‘objeto focado’ quando este é um ser vivo (p. ex.: cavalo correndo, criança trajando roupa do século XVIII)” (Smit, 1997, p. 4). Com base nas definições anteriores estabelecemos as seguintes categorias:

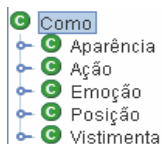


Figura 7: Representação da categoria Como.

A categoria Onde representa, segundo Shatford (1986, p. 53), a localização da imagem no espaço; tipos de lugares geográficos, arquitetônicos e cosmográficos.

A categoria técnica fotográfica foi elaborada a partir da atualização da tabela de recursos técnicos de Manini (2002). A Figura 8 apresenta a representação das categorias da técnica fotográfica.

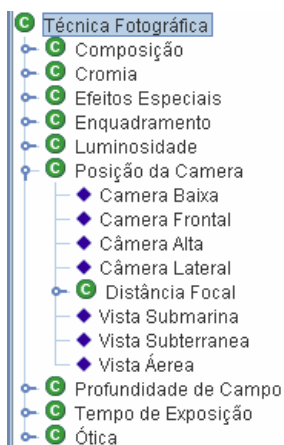


Figura 8: Representação da categoria técnica fotográfica.

Ao indexar uma variável da técnica fotográfica o indexador deve observar a subcategoria Expressão Fotográfica da categoria Sobre para verificar qual conceito abstrato a variável técnica fotográfica expressa. Consideramos que esta parte, referente à expressão, requer ainda reflexão, ponderação e aperfeiçoamento, visto que lida com conceitos abstratos e podem conduzir a uma indexação subjetiva, algo de que os profissionais da informação devem se afastar ao máximo.

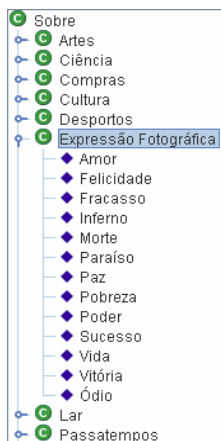


Figura 9 Subcategoria Expressão Fotográfica.

Na próxima seção, descreveremos como implementamos o metadado de descrição que é gerado a partir do emprego da ontologia de representação da imagem, com o uso do protótipo de indexação e recuperação.

Metadado de Descrição de Imagem

O princípio de indexação de imagens com base em ontologias pode ser encontrado em Schreiber (2001); basicamente, o processo consiste em associar conceitos (classes e instâncias) da ontologia do domínio à imagem. Para isto, é necessário o desenvolvimento do esquema de indexação, que está representado na ontologia de indexação e recuperação de imagens pela hierarquia de classes Descrição_Mídia, mais precisamente pela classe Fotografia. O esquema de indexação é uma implementação da representação da imagem que realizamos anteriormente usando UML. A classe fotografia possui os atributos que permitem descrever o conteúdo semântico, o conteúdo sintático e o conteúdo não-visual da fotografia.

A Figura 10 é uma visão geral da ontologia no editor Protege¹ versão 3.2.1. A ontologia que implementamos pode ser dividida em dois tipos: a que é usada para representar o conhecimento do domínio, que, na figura abaixo, corresponde às categorias quem, como, onde, sobre, técnica_fotografica e registro geral; e a de indexação de imagens, que, abaixo, corresponde à hierarquia Descrição_Mídia.

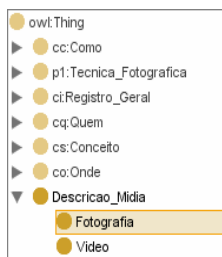


Figure 10: Ontologia de indexação de imagens.

Os atributos da classe fotografia são os que foram especificados na modelagem UML. A descrição de fotografias usando a ontologia de indexação consiste na criação de instância da classe Fotografia. Alguns atributos da classe fotografia, como os atributos do *Dublin Core*, são criados fornecendo valores a partir de um formulário de entrada de dados. Outros atributos, como os que se referem ao conteúdo semântico (quem, onde, como, sobre e técnica fotográfica), são gerados a partir da associação de conceitos da ontologia do domínio ao metadado da classe Fotografia.

Cada registro de uma fotografia gera uma descrição em RDF com base no metadado da classe Fotografia. Isto significa dizer que a descrição de milhares de imagens de um mesmo assunto não aumentará significativamente o tamanho da ontologia do domínio; apenas aumentará o número de instâncias da classe Fotografia.


Protótipo de Indexação de Imagens

Nesta parte mostraremos a geração da descrição de uma fotografia por meio do protótipo de indexação, construído para validar o esquema de indexação. O protótipo foi construído usando a linguagem de programação Java, SPARQL linguagem de consulta para RDF, a representação de ontologias em OWL e a geração descrição de fotografias em RDF.

A descrição consiste, primeiramente, em fornecer valores para os campos, em linguagem natural. Alguns campos ficam em branco, pois não há como os determinar a partir da legenda da fotografia.

A segunda parte da indexação é realizada com o auxílio da ontologia do domínio. Na Figura 11, abaixo, é apresentada a interface de indexação do conteúdo semântico. A indexação consiste em observar a imagem e responder as questões: Quem, Como, Onde, Sobre, Técnica

¹ <http://protege.stanford.edu/>

Fotográfica e Uso por meio da ontologia. Ao clicar no botão  o protótipo abre uma janela mostrando uma visão da ontologia correspondente à questão selecionada. A partir daí a descrição consiste em selecionar conceitos dentro da ontologia para responder a respectiva questão. Os campos da interface de indexação correspondem às questões da grade de análise documentária de imagens (Manini, 2002). Os conceitos genérico e específico (Quem Genérico/Quem Específico) não aparecem na interface de indexação, mas apenas os conceitos quem, onde, como, sobre, etc., devido ao fato da natureza da ontologia ser construída em categorias do nível mais geral para o nível mais específico. Então, quando descrevemos um campo (quem, onde, etc.) com um conceito mais específico da ontologia, a imagem já está implicitamente indexada com o conceito mais geral.

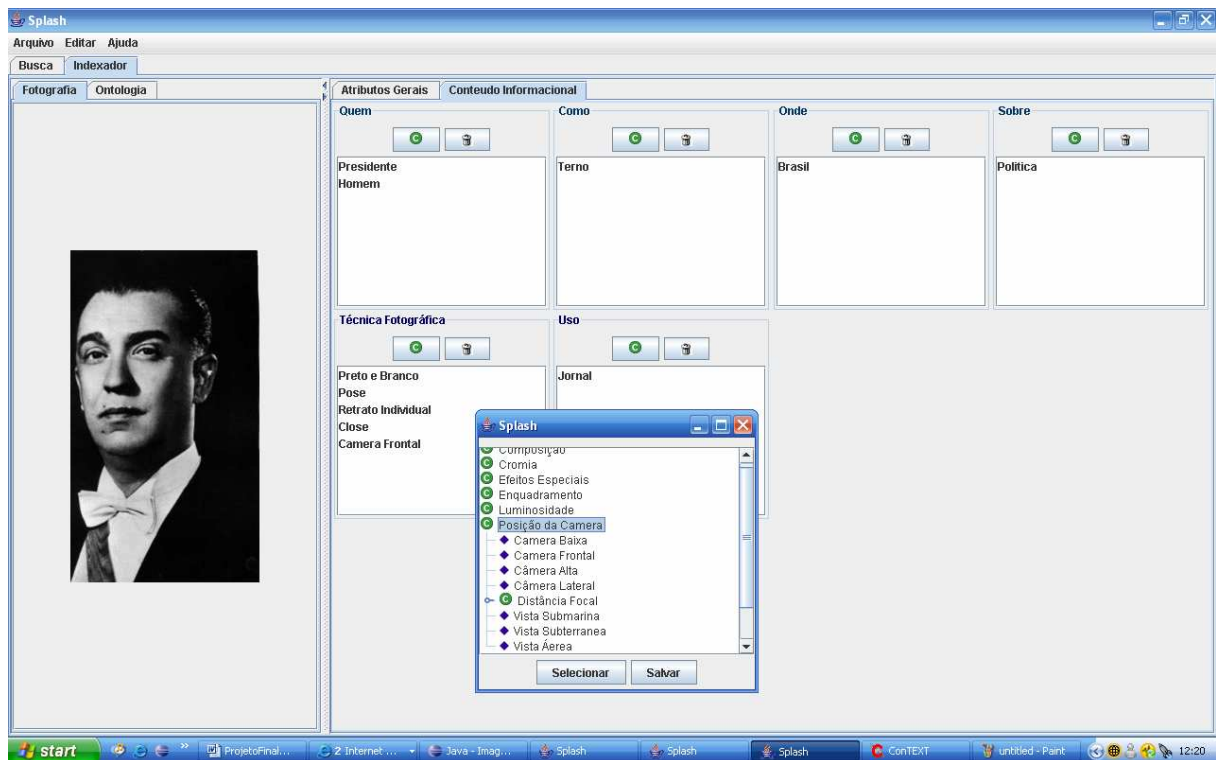


Figura 11: Protótipo de indexação de imagens – interface do conteúdo semântico.

Depois que é realizada a indexação, o protótipo salva a descrição da fotografia em RDF, onde está representado o conteúdo não-visual, o conteúdo semântico. Parte do conteúdo não-visual está representado pelo *Dublin Core* por meio do prefixo dc antes do atributo. A outra parte do conteúdo não-visual foi representada por atributos específicos. Os conteúdos semânticos quem, onde, como, sobre, técnica fotográfica e uso estão representados respectivamente pelos atributos id:quem, id:onde, id:como, id:sobre, id:técnica_fotografica e id:uso. O atributo id:hash usado na descrição e que não aparece na modelagem UML representa o resultado da aplicação do algoritmo de *hashing* à imagem. É um recurso técnico usado para permitir verificar se a imagem já foi anteriormente indexada.

O metadado gerado pelo protótipo pode ser colocado dentro do arquivo da imagem. Os formatos de arquivo de imagens como JPEG, PNG, GIF, etc. permitem a inclusão de informação textual dentro do arquivo. O padrão JPEG, por exemplo, define a seção comentário COM que pode conter até 64K de texto, mais que suficiente para inclusão de metadados, pois os metadados aqui gerados são da ordem de 5K.

Existem várias vantagens em incluir metadados dentre elas citemos:

- o registro da imagem fica acoplado à imagem, evitando que a descrição se perca facilmente;

- permite que outras aplicações que têm acesso à imagem possam categorizá-la e contextualizá-la com base nas informações obtidas do metadado;
- além disso, permite a reindexação da imagem pelos mecanismos de busca de forma mais precisa.

Protótipo de Recuperação de Imagens

O mecanismo de busca é baseado em facetas, que consiste em organizar o conteúdo do repositório em facetas e usá-las extensivamente na interface do usuário para ajudá-lo a formular a consulta e navegar no conteúdo do repositório. O princípio da recuperação é simples: o usuário formula a expressão de consulta por meio da seleção de conceito e o sistema recupera as imagens associadas àquele conceito ou instância no momento da indexação. Se o usuário selecionar conceitos de diferentes categorias o sistema recupera as imagens com base na interseção destas categorias.

Aqui nós usamos este método para organizar o conteúdo da ontologia com base nas facetas do conteúdo semântico (quem, como, onde, sobre, técnica fotográfica e uso). Desta forma, usamos as mesmas facetas da indexação na interface de recuperação. O protótipo de recuperação com base em facetas permite a recuperação de imagens com base no conteúdo semântico e na técnica fotográfica. A recuperação de imagens com base no conteúdo não-visual é realizada por um mecanismo de busca complementar com base em palavras-chave. Observe na figura 13, abaixo, no protótipo de recuperação, que o título das categorias corresponde aos elementos conteúdo semântico da imagem, técnica fotográfica e uso.

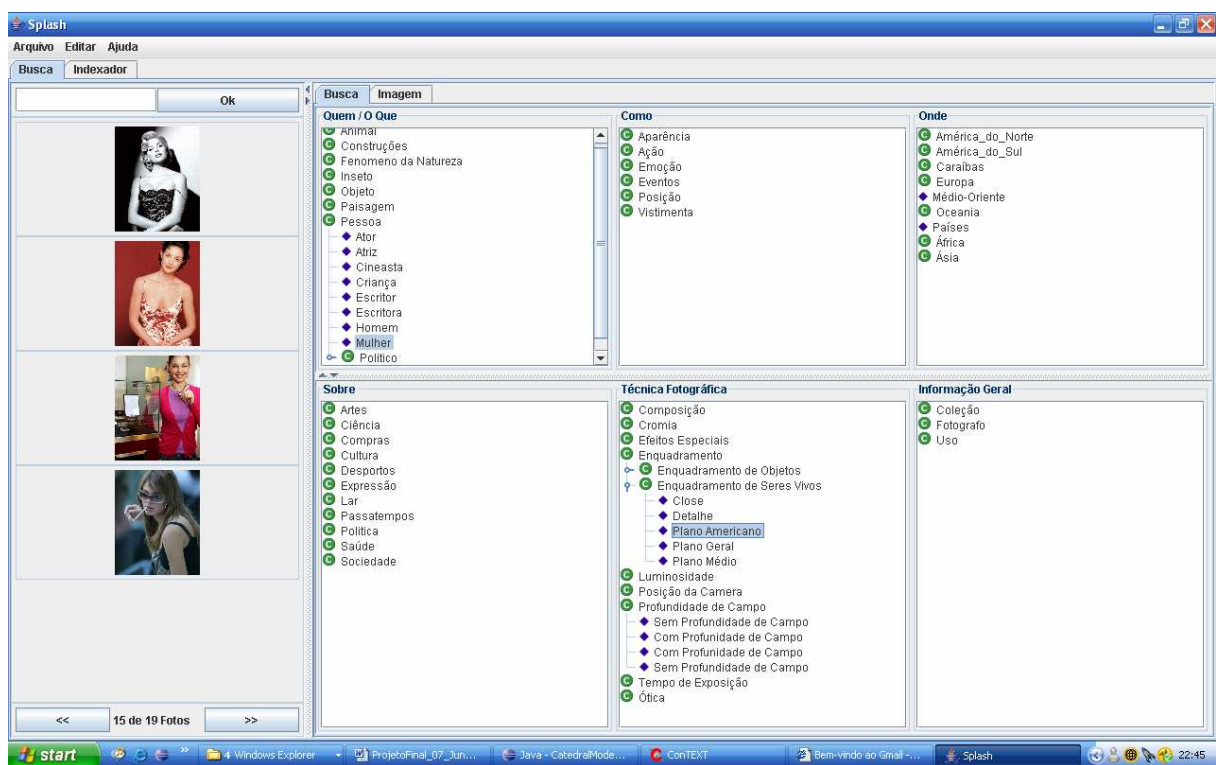


Figura 13: Protótipo de recuperação de imagens com base em facetas.

A figura 13 ilustra também uma expressão de consulta formulada por meio da seleção do conceito mulher da categoria quem e do conceito plano americano da categoria técnica fotográfica. O resultado desta pesquisa é um conjunto de imagens que foram indexadas com o conceito mulher e com o conceito plano americano. A combinação da busca com base em facetas com a recuperação de imagens com base em ontologias traz alguns benefícios: permite que o usuário navegue dentro da ontologia e se familiarize com os seus conceitos; a recuperação baseia-se na

seleção de conceitos dentro da ontologia e isto aumenta a precisão, visto que os usuários selecionam conceitos que foram usados na indexação.

Considerações finais

Este trabalho sistematizou descritores de imagens obtidos do estudo de dois paradigmas importantes da área de indexação de imagens. O resultado deste estudo serviu para a elaboração de uma modelagem conceitual, que, por sua vez, permitiu a elaboração de um metadado para a descrição de imagens. Este metadado agregou vários tipos de atributos para descrição do conteúdo, da técnica fotográfica e de outras classes de metadados como técnico, administrativo, preservação, etc. Além disso, a estrutura do metadado representou a imagem com base em três tipos de método de descrição, a saber: informação estruturada com base em ontologia, dados semi-estruturados com base em pares de atributo/valor e com base em informação não estruturada como texto.

A ontologia de indexação e recuperação de imagens foi desenvolvida para enriquecer a descrição da imagem; ela representou os aspectos do conteúdo semântico da imagem. Delimitamos o escopo de representação com base nas questões do conteúdo semântico. Mesmo assim, o escopo continua muito amplo e foge do controle do ontologista. Diante disso, definimos apenas as classes principais da ontologia com base nas definições de Shatford (1986) sobre a representação da imagem.

Como o processo de desenvolvimento de uma ontologia é iterativo, a representação do conteúdo da imagem pode ser refinada a partir da extensão dos níveis de categorias e da agregação de atributos e na definição de relacionamentos dentro da ontologia. Além disso, é possível melhorar a representação da ontologia e, conseqüentemente, a descrição da imagem a partir do reuso de outras ontologias tal como a *Wordnet*, AAT e ICONCLASS. A *Wordnet* pode ser usada para representar o conteúdo Como. Um tesouro pode ser transformado numa ontologia se ele possui uma hierarquia de termos bem definida. O tesouro AAT pode ser usado para representar parte do conteúdo semântico Sobre. Um trabalho interessante sobre reuso de ontologias para descrição de imagens pode ser encontrado em Hollink et. al. (2006).

O protótipo de indexação e recuperação de imagens com base em ontologias foi inspirado nas idéias de Styrman (2005) e Hyvonen et. al. (2003). Styrman (2005) abordou a indexação e recuperação de imagens com base em ontologias, focando o desenvolvimento da ontologia na representação do domínio, acervos de imagens da Universidade de Helsinki. Aqui nós direcionamos o desenvolvimento da ontologia para a representação da imagem com base no conteúdo semântico, tornado-a independente do domínio. Foi possível verificar a importância da ontologia na descrição de imagens, por meio da associação de conceitos à imagem. Quando uma imagem é descrita com um conceito da ontologia, isto significa que a imagem herda a descrição do conceito associado, permitindo que uma aplicação que tenha posse da imagem navegue dentro da ontologia e extraia mais informação.

O protótipo de busca de imagens empregou dois métodos de recuperação: o método de busca com base em facetas do conteúdo semântico e o método com base em palavras-chave. Embora não tenhamos validado o mecanismo de recuperação por meio da obtenção de dados que possa verificar a precisão e a revocação do sistema, acreditamos que a recuperação de imagens com base em facetas usando ontologias melhora a precisão do mecanismo de recuperação, uma vez que o usuário formula a expressão de busca a partir de conceitos que foram usados pelo indexador no momento da indexação. Além disso, auxilia o usuário a formular a expressão de busca que lida com resultados não vazios por meio da exibição da quantidade de imagens associada a cada conceito.

Este trabalho consistiu no estudo e emprego de técnicas originadas tanto na Ciência da Computação como no emprego de técnicas de representação de imagens originados na Ciência da Informação. O que se observa é que a Ciência da Computação e a Ciência da Informação possuem

objetivos em comum. Ambas estão preocupadas no desenvolvimento de técnicas para melhorar a indexação, a recuperação, a catalogação, organização de informação, etc. Mas o que mais distancia as duas Ciências é o próprio conceito de informação. Na Ciência da Computação, o conceito de informação é confundido com o conceito de dado indivisível e não contextualizado, enquanto que na Ciência da Informação o conceito de informação está associado à semântica, ao significado. Daí o porquê da maioria das técnicas de recuperação da Ciência da Computação serem baseados em unidades léxicas e não em estrutura semântica. Com o advento da Web Semântica começa a se dar os primeiros passos para a recuperação semântica de informação.

Durante a revisão de literatura ficou evidente a inexistência de padrões que possam representar o conteúdo da imagem. Os padrões como o *Dublin Core* foram construídos de modo genérico e não atendem às especificidades da representação da imagem. Enquanto que o padrão VRA foi desenvolvido para representar imagens de um domínio muito específico (Arte e Arquitetura), servindo apenas como um auxílio, mas não como uma solução. Existe uma grande necessidade de metadados que permitam representar os diferentes aspectos da informação imagética.

De um modo mais sistemático, este trabalho contribui com: a definição de uma ontologia de representação da imagem com base no conteúdo semântico e na técnica fotográfica, que permite indexar e recuperar imagens com base no conteúdo semântico e na técnica fotográfica; com uma modelagem conceitual de representação da imagem; com a criação de um metadado capaz de representar as diferentes facetas do conteúdo da imagem; e com a implementação de um sistema de recuperação de imagens usando facetas e ontologias.

A ontologia aqui definida pode ser refinada por meio da criação de vários níveis de categorias, por meio da criação de atributos e por meio da definição de relacionamentos entre os conceitos. Além disso, a ontologia pode ser adequada para domínios específicos e mesmo assim continuar independente do domínio através da especialização de alguns ramos da ontologia.

A descrição da imagem aqui desenvolvida com o uso dos padrões da Web Semântica pode ser anexada dentro do arquivo da imagem e, desta forma, permitir que outras aplicações como mecanismo de busca possa recuperá-la facilmente, extrair o metadado e obter as informações necessárias para contextualizar, reindexar e, desta forma, torná-la acessível globalmente.

Referências Bibliográficas

- Benitez, A. B.; Paek, S. **Object-based multimedia content description schemes and applications for MPEG-7**. 2000. Disponível em: http://www.ctr.columbia.edu/papers_advent/00/ICJ00_ana.pdf. Acesso em: 26 maio 2007.
- GILLILAND-SWETLAND, A. J. **Introduction to metadata: setting the stage**. 2000. Disponível em: <http://www.slis.kent.edu/~mzeng/metadata/Gilland.pdf>. Acesso em: 26 nov. 2006.
- HYVÖNEN, E. et al. **Ontology based image retrieval**. 2003. Disponível em: <http://www2003.org/cdrom/papers/poster/p199/p199-hyvonen.html>. Acesso em: 10 jul. 2004.
- HOLLINK, L., SCHREIBER, G., WIELINGA, B. **Classification of user image descriptions**. Disponível em: <http://www.cs.vu.nl/~guus/papers/Hollink04a.pdf>. Acesso em: 7 out. 2006.
- HOLLINK, L., SCHREIBER, G., WIELEMAKER, J., WIELINGA, B. **Semantic annotation of image collections**. Disponível em: <http://www.cs.vu.nl/~guus/papers/Hollink03b.pdf>. Acesso em: 5 out. 2006.
- JAIMES, A; CHANG, S. **Conceptual framework for indexing visual information at multiple levels**. 2000. Disponível em: http://www.ctr.columbia.edu/~ajaimes/Pubs/spie00_internet.pdf. Acesso em: 2 nov 2006.
- MANINI, M. P. **Análise documentária de fotografias: um referencial de leitura de imagens fotográficas para fins documentários**. 2002. Tese (doutorado) – Universidade de São Paulo, São Paulo, 2002.

- SCHREIBER, A. T. et al. **Ontology-based photo annotation**. 2001. Disponível em: <<http://www.cs.vu.nl/~guus/papers/Schreiber01a.pdf>>. Acesso em: 20 mai. 2004.
- SHATFORD, S. **Analysing the subject of a picture: a theoretical approach**. Cataloging and Classification Quartely, New York, v. 6, n. 3, p. 39-62, 1986.
- SMIT, J. W. **Propostas para a indexação de informação iconográfica**, 1997. (Mimeo).
- SMIT, J. W. **A representação da imagem**. Informare, Rio de Janeiro, v. 2, n. 2 p. 28-36, jul/dez. 1996.
- STYRMAN, A. et. al. **Ontology based image annotation and retrieval**. 2005. Disponível em: <http://www.cs.helsinki.fi/u/astyрман/gradu.pdf>. Acesso em: 6 out. 2006.