

Universidade Federal de Campina Grande  
Centro de Engenharia Elétrica e Informática  
Coordenação de Pós-Graduação em Ciência da  
Computação

PhotoGeo - Uma Biblioteca Digital de Fotografias  
com Suporte de Geoprocessamento

Yuri Almeida Lacerda

Dissertação submetida à Coordenação do Curso de Pós-Graduação em  
Ciência da Computação da Universidade Federal de Campina Grande -  
Campus I como parte dos requisitos necessários para obtenção do grau  
de Mestre em Ciência da Computação.

Área de Concentração: Ciência da Computação

Linha de Pesquisa: Sistemas de Informação e Banco de Dados

Cláudio de Souza Baptista

(Orientador)

Campina Grande, Paraíba, Brasil

Maio - 2009

# PhotoGeo - Uma Biblioteca Digital de Fotografias com Suporte de Geoprocessamento

Yuri Almeida Lacerda

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Campina Grande - Campus I como parte dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Área de Concentração: Ciência da Computação  
Linha de Pesquisa: Sistemas de Informação e Banco de Dados

Cláudio de Souza Baptista  
(Orientador)

Campina Grande, Paraíba, Brasil  
Maio - 2009

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

L131p

2009

Lacerda, Yuri Almeida.

PhotoGeo – Uma biblioteca digital de fotografias com suporte de geoprocessamento / Yuri Almeida Lacerda. – Campina Grande, 2009.

96 f.: il.

Dissertação (Mestrado em Ciências da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

Referências.

Orientador: Profº. Dr. Cláudio de Souza Baptista.

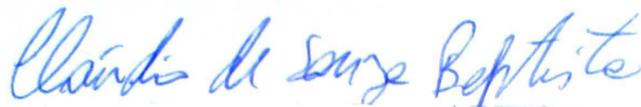
1. Biblioteca Digital. 2. Sistema de Informação Geográfica. 3. Fotografia Digital. I. Título

CDU 004:026.06(043)

**“PhotoGeo – UMA BIBLIOTECA DIGITAL DE FOTOGRAFIAS COM  
SUPORTE DE GEOPROCESSAMENTO”**

**YURI ALMEIDA LACERDA**

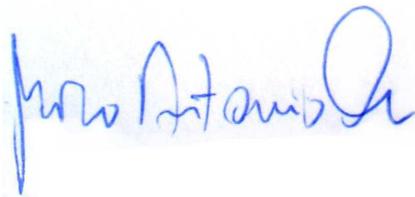
**DISSERTAÇÃO APROVADA EM 12.05.2009**



**PROF. CLÁUDIO DE SOUZA BAPTISTA, Ph.D**  
**Orientador**



**PROF. MARCUS COSTA SAMPAIO, Dr.**  
**Examinador**



**PROF. MARCO ANTÔNIO CASANOVA, Ph.D**  
**Examinador**

**CAMPINA GRANDE - PB**

*Ao meu pai Iari  
e à minha mãe Laice.*

## **Resumo**

A grande facilidade de capturar fotografias com as câmeras digitais fez com que as pessoas estejam produzindo uma quantidade enorme de imagens digitais. Isto acarretou num aumento da complexidade no gerenciamento desse tipo de arquivo. Além disso, os dispositivos móveis com chip GPS e câmera digital embutidos passaram a prover armazenamento nos metadados desses arquivos, além de informações temporais, também informações relativas à posição geográfica (latitude e longitude) do momento da captura da foto, permitindo que seja explorado o aspecto geográfico da fotografia. Os sistemas existentes de organização de fotografias georreferenciadas não exploram de forma eficiente essas informações. Desta maneira, nesta dissertação é proposta a criação de uma ferramenta, chamada PhotoGeo, que permite aos usuários organizarem, navegarem, anotarem e consultarem suas fotos georreferenciadas de maneira simplificada e explorando adequadamente as informações geográficas e temporais. O PhotoGeo propõe um mecanismo novo de organização automática de fotografias, que realiza uma classificação utilizando esses metadados e informações extraídas de fontes externas. Também, é proposto no PhotoGeo uma interface multi-modal para navegar em uma coleção de fotos, dotada também de um mecanismo de consultas avançadas de fotografias que explora os domínios de dados convencional, espacial e temporal. Outra contribuição muito importante do PhotoGeo, é uma nova solução para a sugestão de pessoas que estão presentes nas fotografias, ampliando e utilizando as informações de contextos do usuário.

## **Abstract**

Since digital cameras made photography capturing easy, people are producing and storing an enormous amount of digital images. This caused an increased complexity in managing this kind of media file. Furthermore, the mobile devices with embedded GPS and digital cameras are capable of store temporal (timestamp of the shot) and geographic (latitude and longitude) information in image files metadata, enabling the treatment of the geographical aspects of photography. The existing geographic photo organizers, however, do not efficiently explore spatial-temporal metadata. In this study, we propose a new photo organizer tool – PhotoGeo – which provides the functionalities of organizing, navigating, annotating, and searching geo-referenced photos, which has enhanced usability and adequately explores geographic and temporal information. PhotoGeo proposes a novel automatic photo organizing mechanism which classifies pictures based on metadata and external sources. It also proposes a multi-modal interface to enable navigation within a photo collection, accompanied by a mechanism to perform advanced search in photography collections which explores the conventional, spatial, and temporal domains. Another very important contribution of PhotoGeo, is a new solution to suggest identification of people who are subject in the photos, expanding and utilizing the contextual informations of user.

## Agradecimentos

Primeiramente à Deus, por ter me dado o dom da vida e toda a saúde necessária durante o desenvolvimento deste trabalho.

Ao meu pai, Iari, e à minha mãe, Laice, por todo o apoio, amor, educação e ensinamentos que me deram ao longo da vida, sendo exemplos de verdadeiros pais.

Ao meu orientador, o professor Dr. Cláudio Baptista, pela orientação, empenho, conselhos e confiança que foram fundamentais para que o trabalho acontecesse.

A minha namorada, Thially, por toda amizade, paciência, carinho e amor dedicado durante este período.

Ao meu grande amigo e colega de mestrado Hugo Feitosa, que colaborou bastante com esta pesquisa. Assim como, os professores Dr. Marcus Sampaio (UFMG) e Dr. Anselmo Paiva (UFMA). Ao professor Dr. Marco A. Casanova (PUC-Rio) por ter participado da banca examinadora e contribuído para engradecer esta pesquisa.

Ao amor e amizade dedicados pelos meus familiares, em especial o meu irmão Marco Aurélio e sua esposa Sabrina, e a minha irmã Melina e seu marido Marquinhos.

Aos meus grandes amigos Daniel Fireman e Cláudio Campelo que mesmo longe me apoiaram no desenvolvimento deste trabalho. Também à Nivaldo, Taciano, Fernando, pela amizade e pela ajuda nos experimentos!

Aos meus amigos: Rafael Pietro, Tiago Lamenha, Glauber, Welflen, Marcelo Iury, Arthurzinho (e toda a sua família), Juliana, Marcella, Rossana, Rômulo, Tomás, Flávio (barata), Halley, Danillo, Jemmerson, Welmisson, Pablo Antonino, Felipe, Augusto, Marquinhos (gaúcho), Saulo, Gilson, Guilherme, Aurimar, Taciano Silva e Ricardinho. A todos os amigos, tanto da universidade como dos momentos de lazer, que não foram citados aqui, mas que com certeza foram muito importantes para este trabalho.

Aos meus colegas do LSI pela experiência compartilhada e pela amizade.

A todos os professores do DSC, pela competência e pelo conhecimento transmitido. Aos funcionários da COPIN, em especial a Aninha pela gentileza e simpatia nas diversas vezes que precisei ir na COPIN.

A CAPES, pelo apoio financeiro.

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Contexto e Motivação . . . . .	1
1.2	Definição do Problema . . . . .	2
1.3	Solução Proposta . . . . .	3
1.4	Estrutura do Documento . . . . .	4
<b>2</b>	<b>Referencial Teórico</b>	<b>6</b>
2.1	Dados . . . . .	6
2.1.1	Dados Espaciais . . . . .	6
2.1.2	Dados Temporais . . . . .	7
2.1.3	Dados Multimídia . . . . .	9
2.2	Metadados das Imagens . . . . .	11
2.2.1	Estrutura . . . . .	13
2.2.2	Ferramentas e Bibliotecas para manipulação de Exif . . . . .	20
2.3	Sistemas de Informação Geográfica . . . . .	20
2.4	Bibliotecas Digitais Geográficas . . . . .	22
2.4.1	Organização automática das fotografias . . . . .	24
2.4.2	Navegação multi-modal . . . . .	25
2.4.3	Mecanismo para Auxiliar à Anotação . . . . .	26
2.4.4	Consulta textual simples e avançada . . . . .	27
2.4.5	Consulta temporal . . . . .	27
2.4.6	Consulta espacial simples e avançada . . . . .	28
2.4.7	Repositório Web . . . . .	28

<b>3</b>	<b>Trabalhos Relacionados</b>	<b>30</b>
3.1	Organização Automática de Fotografias . . . . .	30
3.2	Sugestão de Anotação . . . . .	32
3.3	Consultas Avançadas . . . . .	33
3.4	Considerações Finais . . . . .	34
<b>4</b>	<b>Arquitetura do Sistema</b>	<b>37</b>
4.1	Arquitetura do Sistema . . . . .	37
4.2	Módulos da Arquitetura do PhotoGeo . . . . .	40
4.2.1	Navegação multi-modal . . . . .	40
4.2.2	Organização Automática . . . . .	42
4.2.3	Sugestão de anotação de pessoas . . . . .	42
4.2.4	Consultas Avançadas . . . . .	43
4.3	Esquema de Dados . . . . .	44
4.4	Considerações Finais . . . . .	45
<b>5</b>	<b>Organização Automática de Fotografias</b>	<b>48</b>
5.1	Gerenciador de Fotos . . . . .	49
5.2	Coletor de Dados . . . . .	49
5.2.1	Coletor de Dados Espaciais . . . . .	50
5.2.2	Coletor de Dados Temporais . . . . .	51
5.3	Organizador PhotoGeo . . . . .	51
5.3.1	Primeiro Passo: Agrupamento espacial . . . . .	53
5.3.2	Segundo Passo: Segmentação temporal . . . . .	54
5.3.3	Terceiro passo: Aperfeiçoamento . . . . .	55
5.3.4	Quarto passo: Auto-denominação . . . . .	57
5.4	Experimentos . . . . .	57
5.4.1	Análise completa . . . . .	60
5.4.2	Análise parcial . . . . .	61
5.5	Considerações Finais . . . . .	62

---

<b>6</b>	<b>Sugestão de Anotação de Fotografias</b>	<b>63</b>
6.1	Anotações de Fotografias . . . . .	63
6.2	Estimadores para o Ranking . . . . .	64
6.3	Algoritmo de Sugestão . . . . .	67
6.4	Informações de contexto dos usuários . . . . .	69
6.5	Análises e Resultados . . . . .	71
6.6	Considerações Finais . . . . .	74
<b>7</b>	<b>Consultas Avançadas</b>	<b>75</b>
7.1	Consultas Avançadas . . . . .	75
7.1.1	Filtro Convencional . . . . .	76
7.1.2	Filtro Espacial . . . . .	77
7.1.3	Filtro Temporal . . . . .	78
7.2	Considerações Finais . . . . .	79
<b>8</b>	<b>Conclusão</b>	<b>80</b>
8.1	Contribuições . . . . .	80
8.2	Trabalhos Futuros . . . . .	81
	<b>Referências Bibliográficas</b>	<b>89</b>
<b>A</b>	<b>Diagramas do Sistema</b>	<b>90</b>

# Lista de Símbolos

GPS - *Global Positioning System*

PDA - *Personal Digital Assistant*

SIG - *Sistema de Informação Geográfica*

CAD - *Computer Aided Design*

SGBD - *Sistema de Gerenciamento de Banco de Dados*

JPEG - *Joint Photographic Experts Group)*

TIFF - *Tag Image File Format*

JEITA - *Japan Eletronics and Information Technology Industries Association*

Exif - *Exchangeable image file format*

ISO - *International Organization for Standardization*

RGB - *Red Green Blue*

RIFF - *Resource Interchange File Format*

WAVE - *Waveform Audio File Format*

IFD - *Image File Directory*

DCT - *Discrete Cosine Transform*

APP - *Application Market Segment*

SOI - *Start of Image*

OGC - *Open Geospatial Consortium*

SQL - *Structured Query Language*

CBIR - *Content-Based Image Retrieval*

MVC - *Model-View-Controller*

JSP - *Java Server Pages*

HTML - *HyperText Markup Language*

POI - *Point of Interest*

KML - *Keyhole Markup Language*

REST - *Representational State Transfer*

HTTP - *Hypertext Transfer Protocol*

# Lista de Figuras

2.1	Exemplos de dados espaciais . . . . .	7
2.2	Operadores espaciais . . . . .	8
2.3	Operadores para eventos longos. . . . .	9
2.4	Operadores para eventos instantâneos. . . . .	9
2.5	Dados multimídia . . . . .	10
2.6	Exemplo dos metadados de uma imagem . . . . .	11
2.7	Estrutura de uma imagem não compactada. Figura extraída de [10] . . . . .	14
2.8	Estrutura de uma imagem compactada JPEG. Figura extraída de [10] . . . . .	15
2.9	Tipos de dados geográficos definidos pela ISO e OGC. Figura adaptada de [34]	21
4.1	Arquitetura do PhotoGeo . . . . .	38
4.2	Arquitetura detalhada do PhotoGeo . . . . .	40
4.3	Navegação multi-modal . . . . .	41
4.4	Interface da sugestão de anotações do PhotoGeo . . . . .	43
4.5	Diagrama de Entidade-Relacionamento . . . . .	44
5.1	Exemplo de agrupamento espacial de quatro eventos utilizando um mapa extraído do Google Maps . . . . .	54
5.2	Exemplo de segmentação temporal . . . . .	55
5.3	Primeira situação do passo de aperfeiçoamento . . . . .	56
5.4	Segunda situação do passo de aperfeiçoamento . . . . .	56
5.5	Distribuição Temporal das Fotos . . . . .	58
5.6	Resultados do agrupamento em eventos . . . . .	61
5.7	Resultados do Questionário . . . . .	62

---

6.1	Estatísticas das coleções avaliadas . . . . .	71
6.2	Simulador embutido no protótipo do PhotoGeo . . . . .	73
6.3	Análise da influência do segundo passo . . . . .	73
6.4	Análise considerando os contatos mais importantes . . . . .	74
7.1	Interface de Consulta Avançada . . . . .	76
7.2	Resultado de uma consulta . . . . .	77
7.3	Filtros da Consulta Avançada . . . . .	78
A.1	Tabelas do banco de dados do PhotoGeo . . . . .	91
A.2	Diagrama de Classe da Organização Automática . . . . .	92
A.3	Diagrama de Classe do Coletor de Dados . . . . .	93
A.4	Diagrama de Classe da Sugestão de Anotação . . . . .	94
A.5	Diagrama de Classe da Interface Multi-Modal . . . . .	95
A.6	Diagrama de Classe das Consultas Avançadas . . . . .	96

# Lista de Tabelas

2.1	Atributos definidos pelo TIFF. Adaptada de [10] . . . . .	16
2.2	Atributos privados do Exif. Tabela adaptada de [10] . . . . .	17
2.3	Atributos dos ajustes da câmera. Tabela adaptada de [10] . . . . .	19
2.4	Atributos do GPS. Tabela adaptada de [10] . . . . .	19
2.5	Características desejáveis a uma bibliotecal digital especializada em fotografias georreferenciadas . . . . .	24
3.1	Características das pesquisas e sistemas analisados neste trabalho . . . . .	35
4.1	Comparativo entre as características do PhotoGeo e das pesquisas (e dos sistemas) investigadas neste trabalho . . . . .	46

# Capítulo 1

## Introdução

### 1.1 Contexto e Motivação

A substituição das câmeras analógicas pelas digitais fez com que as pessoas estejam gerando cada vez mais uma maior quantidade de dados multimídia, tais como: fotografia, áudio e vídeo. Devido a isso, é muito importante ter disponíveis ferramentas que auxiliem no gerenciamento desses dados.

Grandes empresas e órgãos de pesquisa estão investindo muitos recursos em ferramentas desse tipo. O Google (<http://www.google.com>), com a sua ferramenta Picasa (<http://picasa.google.com>), e o Yahoo! (<http://www.yahoo.com>), com o Flickr (<http://www.flickr.com>), são exemplos dessas empresas.

A fotografia digital é um dos tipos de dados multimídia preferidos pelas pessoas, principalmente devido à popularização da câmera digital e dos dispositivos móveis equipados com câmeras digitais. As pessoas desejam ter mecanismos que auxiliem a armazenar, recuperar, visualizar e compartilhar suas fotografias pessoais. Porém, a grande quantidade de arquivos de fotografias digitais, aliado ao fato de as pessoas possuírem cada vez mais dispositivos para armazená-los (laptops, computadores pessoais, PDAs, celulares, entre outros), dificultam a localização e recuperação destes arquivos. Estes são alguns dos problemas que as bibliotecas digitais multimídia procuram solucionar [3].

Uma forma de auxiliar a recuperação das fotos é através das anotações sobre as fotografias. As anotações podem ser armazenadas dentro dos metadados [10] das próprias fotografias. Essas anotações podem contemplar vários aspectos, desde dados técnicos até in-

formações sobre a semântica da foto. Os dados técnicos são informações que são obtidas automaticamente da câmera digital no momento da captura da fotografia, por exemplo: data e hora, foco da câmera, tempo de exposição e tipo de flash utilizado. Já a informação semântica possuem um maior sentido para o ser humano, por exemplo: descrição da fotografia, pessoas ou objetos que estão na foto, cidade que a foto foi capturada, entre outras. Com o surgimento de câmeras digitais com GPS embutido também tornou-se possível armazenar dados sobre a posição geográfica (latitude e longitude) da câmera no momento da captura da fotografia.

As informações semânticas são as mais importantes para a recuperação em uma biblioteca digital [44]. Estudos mostram que o evento, o lugar e as pessoas que estão presentes nas fotografias são as informações mais marcantes para as pessoas recordarem de uma foto [54]. Um evento é o espaço no tempo onde acontece uma ocasião social ou atividade, por exemplo: casamento, aniversário, congresso, passeio, etc. O lugar pode ser uma cidade, um clube, um estado, uma praia, um país, entre outros. Assim, o exemplo de uma recordação poderia ser: "A foto que capturei do meu amigo João no mês de Julho de 2008 na Praia do Futuro em Fortaleza - Ceará durante uma viagem de férias".

Desta maneira, faz-se necessário ter ferramentas que possibilitem ao usuário gerenciar a sua coleção pessoal de fotografias. Estas devem disponibilizar funcionalidades de navegação nas fotografias, organização, visualização, consulta, anotação, entre outras.

## 1.2 Definição do Problema

Existe a necessidade de se ter ferramentas que auxiliem os usuários a melhor gerenciar suas coleções pessoais de fotografias, contemplando os seus aspectos semânticos.

A posição geográfica (latitude e longitude) de uma fotografia é um dado muito importante para obter informações semânticas das fotografias, como por exemplo, o nome da cidade onde a fotografia foi capturada. Porém, as bibliotecas digitais multimídia existentes exploram de forma muito deficiente essa informação. Desta maneira, faz-se necessário a criação de uma ferramenta que possibilite explorar melhor este tipo de informação, permitindo ao usuário usufruir de melhores recursos.

Uma das principais funcionalidades das bibliotecas digitais multimídia é a consulta. Os

mecanismos de consultas dos organizadores de fotografias na grande maioria são bastante simples, não possibilitando a realização de consultas mais avançadas, principalmente com relação a informação geográfica.

### 1.3 Solução Proposta

Neste sentido propomos neste trabalho a criação de uma biblioteca digital de fotografias com suporte de geoprocessamento para os usuários realizarem uma melhor gerência da sua coleção de fotos pessoais, fazendo um melhor uso da informação geográfica.

Essa solução foi chamada de PhotoGeo. Este sistema possui uma interface multi-modal que permite ao usuário navegar nas fotografias de forma hierárquica ou através de um mapa digital. As principais funcionalidades desse sistema são as seguintes: organização automática de fotografias, sugestão de anotações de pessoas e consultas avançadas.

O módulo de organização automática propõe um novo algoritmo para organizar as fotografias em eventos, utilizando os metadados das fotografias e dados extraídos de uma rede social baseada em mapa e de uma agenda eletrônica web.

Outra funcionalidade muito importante para uma biblioteca digital de fotografias é auxiliar o usuário a anotar informações semântica nas fotografias, como por exemplo: objetos e pessoas presentes nas fotografias. Para esta solução foi desenvolvido um módulo para auxiliar a anotação de pessoas, visto que é uma das informações mais relevantes para os interessados em recuperar uma fotografia. Este módulo possui um algoritmo que classifica as pessoas mais prováveis de estarem presentes nas fotografias.

Como a posição geográfica é um forte fator para os usuários se lembrarem das fotos, então é importante ter mecanismos de consultas espaciais onde se possam utilizar operadores espaciais [17]. Por exemplo, o usuário poderia estar procurando por fotos de praias que não sejam do litoral nordestino, ou de tribos indígenas fora da região amazônica. Não há ainda, na literatura, bibliotecas digitais de fotografias com suporte a mecanismos avançados de consulta espacial.

## 1.4 Estrutura do Documento

Este documento está organizado da seguinte forma:

- Capítulo 2: Este capítulo apresenta o referencial teórico que é necessário para o entendimento do trabalho, estando dividido em cinco seções: a primeira trata sobre os dados espaciais, temporais e multimídia; em seguida, são apresentados os conceitos sobre metadados de fotografias; a seção seguinte apresenta os conceitos fundamentais de sistemas de informação geográfica; logo após, há uma introdução sobre bibliotecas digitais geográficas e são apresentadas as principais características para uma biblioteca digital geográfica com suporte de geoprocessamento.
- Capítulo 3: Apresenta alguns trabalhos que foram desenvolvidos que têm características semelhantes com o trabalho desta dissertação. Este capítulo está dividido nas funcionalidades que são exploradas ao longo do trabalho: organização automática, sugestão de anotações de pessoas e consultas avançadas.
- Capítulo 4: O quarto capítulo desta dissertação apresenta a arquitetura do PhotoGeo e uma descrição geral de cada módulo do sistema. Também, apresenta as tecnologias utilizadas para o desenvolvimento do mesmo.
- Capítulo 5: Apresenta uma solução para a organização automática de fotografias georreferenciadas. Esta solução dá-se através de um algoritmo composto por quatro passos, sendo estes: agrupamento espacial, segmentação temporal, aperfeiçoamento e nomeação. Inicialmente é apresentado o algoritmo e, em seguida, são exibidos os experimentos realizados.
- Capítulo 6: Apresenta uma solução para anotação de pessoas de forma semi-automática a partir de um algoritmo para a sugestão de pessoas. Este algoritmo foi desenvolvido a partir de novos estimadores que são apresentados neste capítulo. Este capítulo está organizado em três partes: a formalização do processo de anotação, a solução proposta e os experimentos realizados.
- Capítulo 7: Apresenta o módulo de consultas que foi implementado pelo PhotoGeo. Dessa maneira, este capítulo mostra o funcionamento do módulo de consultas através

dos filtros convencionais, espaciais e temporais.

- Capítulo 8: Conclui o trabalho com a apresentação dos resultados obtidos a partir do que foi apresentado no trabalho e apresenta um conjunto de trabalhos futuros que podem dar continuidade ao exposto nesta dissertação.

# Capítulo 2

## Referencial Teórico

Este capítulo apresenta um estudo sobre conceitos importantes para a compreensão da dissertação. Inicialmente será apresentado um estudo sobre dados e metadados. Em seguida, são apresentados os conceitos fundamentais de Sistemas de Informação Geográfica (SIG). E por último, a seção 2.5 apresenta uma introdução a Bibliotecas Digitais Geográficas.

### 2.1 Dados

Os dados podem ser classificados em diversos tipos. Para o entendimento dos conceitos apresentados neste capítulo faz-se necessário ter o conhecimento dos seguintes tipos de dados: espaciais, temporais e multimídia. As definições desses tipos são apresentadas a seguir.

#### 2.1.1 Dados Espaciais

Os dados espaciais são aqueles utilizados para representar o espaço, como: pontos, polígonos, linhas e imagens. A Figura 2.1 apresenta alguns exemplos desses dados. O espaço pode ser bidimensional (2D) ou tridimensional (3D). Para este trabalho será considerado o espaço tendo apenas duas dimensões. Desta forma, os pontos são representados através de um par de coordenadas  $(x,y)$ . Por exemplo, as coordenadas  $(15,20)$  representam um ponto que está na posição 15 do eixo-x e na posição 20 no eixo-y de um plano cartesiano  $(x,y)$ . As linhas são representadas através de dois ou mais pares de coordenadas. Já os polígonos, são formados por um conjunto de linhas que se fecham para formar a área de um polígono.

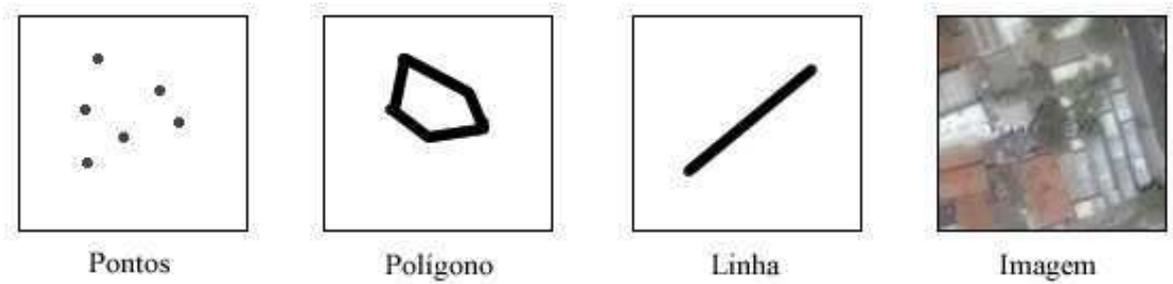


Figura 2.1: Exemplos de dados espaciais

Há um conjunto de operadores espaciais que podem ser utilizados para realizar operações em dados espaciais [40]. Esses operadores foram formalizados por Egenhofer e Franzosa [16]. Alguns desses operadores são apresentados na Figura 2.2 e suas funções estão listadas a seguir:

- **Igual:** Verifica se duas geometrias são iguais.
- **Distância:** Calcula a distância espacial entre duas geometrias.
- **Intercepta:** Verifica se duas geometrias se interceptam.
- **União:** Retorna o resultado da união entre duas geometrias.
- **Área:** Calcula a área de uma determinada geometria.
- **Contém:** Verifica se uma geometria contém completamente outra geometria.

### 2.1.2 Dados Temporais

Os dados temporais são responsáveis por determinar o tempo em que um determinado fato ocorreu. Esse fato poderá ser um evento instantâneo (uma informação pontual) ou um evento longo (um intervalo de tempo). Por exemplo, o momento de captura de uma fotografia é um evento instantâneo. Já um determinado evento, como por exemplo, o “XIV Simpósio Brasileiro de Sistemas Multimídia e Web”, que ocorreu entre os dias 26 e 29 de Outubro de 2008, é um evento longo, pois esse evento aconteceu durante um determinado intervalo de tempo.

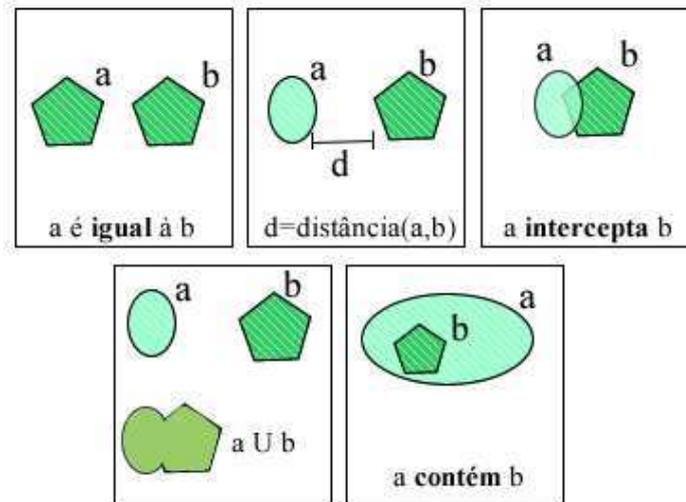


Figura 2.2: Operadores espaciais

Da mesma forma que os dados espaciais, há vários operadores que podem ser utilizados nos dados temporais. Allen [2] propôs um conjunto de operadores para eventos longos, apresentados na Figura 2.3, que analisam o seguinte:

- **Antes:** se um determinado evento ocorre antes de outro.
- **Igual:** um determinado evento ocorre exatamente igual a outro.
- **Encontra:** se dois eventos possuem um determinado período em comum.
- **Sobrepõe:** se o período de duração de um evento sobrepõe outro.
- **Durante:** se um evento ocorre durante outro evento.
- **Inicia:** se um evento inicia igual com outro evento.
- **Termina:** se um determinado evento termina igual com outro.

Neste trabalho serão tratados apenas eventos instantâneos (o momento de captura da fotografia). Desta forma, com base nos operadores citados anteriormente, foi elaborado um conjunto de operadores apenas para tratar de eventos instantâneos. Esses operadores, exibidos na Figura 2.4, são os seguintes:

- **Antes:** se um determinado evento ocorre antes de outro.

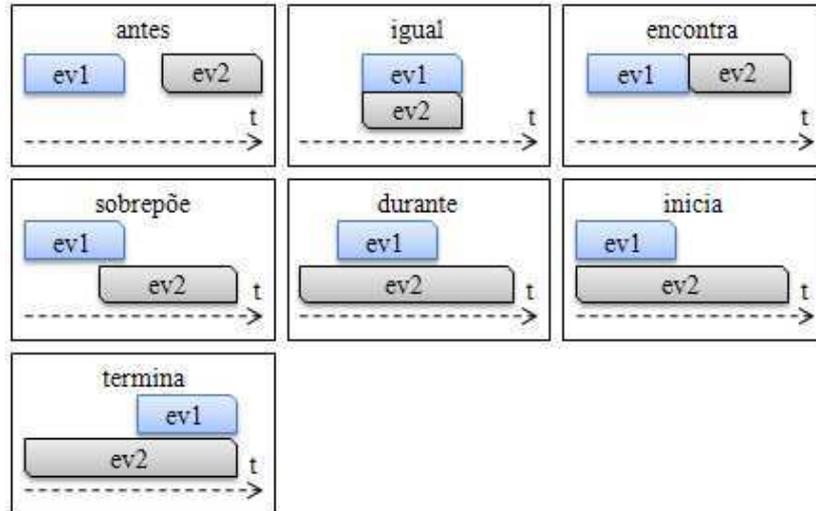


Figura 2.3: Operadores para eventos longos.

- **Igual:** um determinado evento ocorre exatamente igual a outro.
- **Depois:** se um determinado evento ocorre depois de outro.
- **Entre:** se um evento ocorre após um determinado evento e antes de outro.

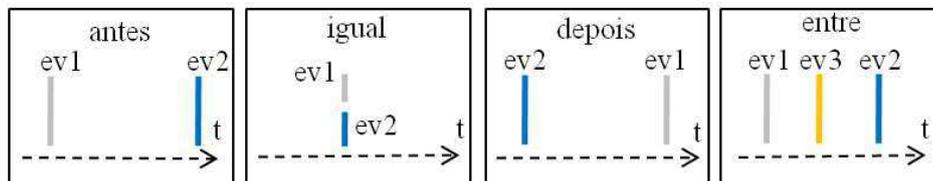


Figura 2.4: Operadores para eventos instantâneos.

### 2.1.3 Dados Multimídia

A evolução nos dispositivos eletrônicos acarretou em um grande aumento na quantidade de arquivos multimídia produzidos. Diversos sistemas de informação necessitam trabalhar com esse tipo de dados [28]: sistemas médicos, aplicações CAD, sistemas na área de entretenimento, SIGs, bibliotecas digitais, entre outros. Esses dados podem ser arquivos de áudio, vídeo, documentos, gráficos, planilhas, textos, entre outros.

Diferentemente dos dados convencionais, que são compostos por caracteres alfanuméricos, os dados multimídia necessitam de estruturas mais complexas para serem armazenados.

Por isso, é necessário que os sistemas multimídia tenham um modelo de dados também complexo devido aos dados suportados por esses sistemas. Os arquivos multimídia possuem tamanhos bem diversificados podendo ir desde um pequeno arquivo de texto até um arquivo de vídeo de muitos gigabytes. Por esse motivo, os sistemas multimídia devem dar o suporte a grandes objetos. A indexação também funciona de forma diferente dos sistemas de informação convencionais, ela deve dar suporte às consultas por similaridade, retornando uma lista ordenada de possíveis resultados. Sendo assim, a interface de um sistema multimídia deve permitir consultas visuais, navegação, visualização de dados complexos, etc. Para alguns tipos de arquivo é possível que seja necessário realizar um pré-processamento para melhorar o desempenho do sistema.

Os dados multimídias podem ser classificados de várias formas. A maneira mais comum é separar de acordo com a dimensão temporal da mídia, ou seja, se ela é ou não contínua no tempo. Utilizando esta classificação, as mídias podem ser separadas em estáticas e dinâmicas [5]. Os dados dinâmicos são aqueles que são baseados em uma linha do tempo, como: áudio, vídeo, animação, fala e música. Já os dados estáticos, são os que não dependem de uma linha de tempo, como: fotografias, gráficos, textos e planilhas. A Figura 2.5 mostra essa classificação dos dados multimídia.

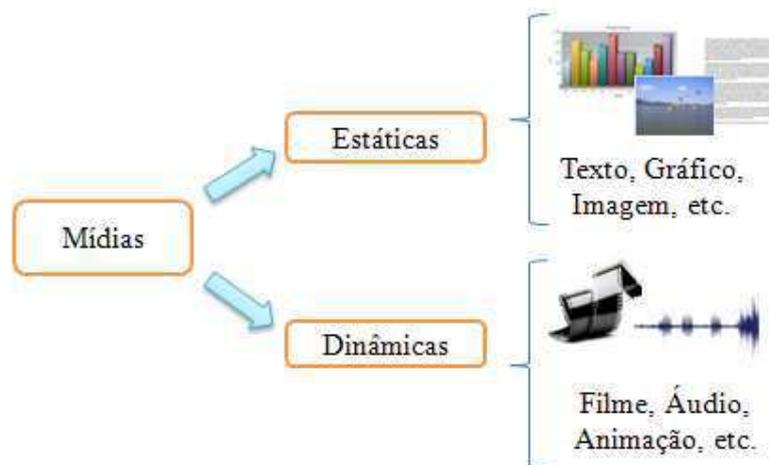


Figura 2.5: Dados multimídia

## 2.2 Metadados das Imagens

Os metadados são dados sobre dados, ou seja, são informações utilizadas para descrever o conteúdo e as características daquele dado [32]. A intenção dos metadados é auxiliar na precisão das consultas, na interpretação semântica e na recuperação dos dados. Em diversos casos há uma dificuldade na separação entre metadado e dado [32].

A integração entre metadados e dados é muito importante para que os metadados estendam as funcionalidades já existentes para o armazenamento de dados, tais como: indexação, busca precisa e recuperação eficiente, estrutura de dados, dentre outras. Uma forma para se realizar essa integração é permitindo que os Sistemas de Gerenciamento de Banco de Dados (SGBDs) manipulem as informações dos metadados.

As informações dos metadados podem ser extraídas a partir de um programa de *parsing* que seja responsável por recuperar informações que sigam um determinado padrão. Geralmente as informações dos metadados estão armazenadas em uma estrutura de dois campos relacionados, onde um é o nome do atributo e o outro é valor do atributo, dentro do próprio arquivo. A Figura 2.6 apresenta um exemplo dos metadados de uma fotografia.

Camera	
Exposure Time	1/158"
F Number	F7
Exposure Program	Manual
ISO Speed Ratings	70
Exif Version	Version 2.2
Date Time Original	2007-03-03 17:18:35
Date Time Digitized	2007-03-03 17:18:35
Components Conf...	YCbcr
Compressed Bits ...	3.8164
GPS	
GPS Version ID	0, 0, 2, 2
GPS Latitude Ref	South latitude
GPS Latitude	7°2'18.73"
GPS Longitude Ref	West longitude
GPS Longitude	34°51'20.81"
GPS Altitude Ref	Sea level
GPS Altitude	0m



Figura 2.6: Exemplo dos metadados de uma imagem

As informações dos metadados variam de acordo com o tipo de dado e do formato do arquivo. Por exemplo, os metadados dos arquivos de imagem são diferentes dos arquivos de áudio, e numa imagem do formato JPEG há metadados diferentes do formato TIFF.

As câmeras digitais mais modernas armazenam nos metadados das fotografias digitais um conjunto de informações relativas ao momento da captura da imagem, como dados do contexto, ajustes da câmera, modelo da câmera, padrão de cores da imagem, áudios relacionados, dentre outras informações [10].

Essas informações armazenadas juntamente com as fotografias têm diversas utilidades. Uma delas, é possibilitar consultar por determinadas fotografias. Por exemplo, tentar localizar a fotografia de um africano em meio das savanas africanas ou então a fotografia de um monumento na cidade de Campina Grande [42]. Outra utilização seria para o uso de técnicas forenses, como por exemplo, a partir de uma determinada fotografia coletada no computador de um criminoso conseguir encontrar a localização geográfica e a data de quando foi capturada a fotografia [51]. As informações de ajustes de câmera podem também servir aos fotógrafos para estudarem como os ajustes influenciam na captura da imagem, para assim aperfeiçoarem suas técnicas. Outra utilidade, é o uso de dados de localização geográfica e informações temporais para organizar automaticamente a coleção de fotos do usuário [43].

O surgimento de diversos equipamentos e softwares para manipular imagens digitais fez com que surgisse a necessidade de haver uma interoperabilidade entre os diversos aparelhos e programas de computadores. Com esse intuito, a Japan Electronics and Information Technology Industries Association (JEITA) criou uma especificação para a anotação em imagens digitais, chamada de Exif [10]. A International Organization for Standardization (ISO) definiu na norma ISO 12234-1 que o Exif era a especificação padrão para anotação realizada pelas câmeras digitais.

A primeira versão do Exif foi lançada em Outubro de 1996 e continha a especificação para as imagens, a estrutura de atributos e as definições dos atributos básicos. Em seguida, foi lançada a segunda especificação, a versão 2.0, que incrementou o espaço de cores RGB, suporte às miniaturas compactadas e arquivos de áudio. Na versão 2.1, lançada em Julho de 1998, foram adicionados os atributos de interoperabilidade. E por último, já em abril de 2002, foi publicada a versão 2.2.

O PRISM Working Group da empresa IDEAlliance criou uma especificação chamada de PRISM Digital Image Metadata Encoding [35], que faz parte do PRISM Specification Package. Essa especificação oferece um novo padrão para auxiliar no gerenciamento de imagens digitais. Diferentemente do Exif, que não possibilita armazenar várias informações

semânticas sobre as fotos, esta especificação fornece um amplo conjunto de metadados para descrever uma determinada fotografia, possibilitando armazenar vários tipos de informações semânticas, como por exemplo: as pessoas, eventos e objetos mostrados nas imagens; informações mais detalhadas sobre o lugar que a fotografia foi capturada; e as pessoas, eventos e objetos que contribuíram para a criação da fotografia. Diferentemente do padrão Exif, a especificação do PRISM utiliza um conjunto de metadados XML, permitindo assim uma estrutura bem mais flexível do que a do Exif. Esse padrão é baseado em vários padrões existentes, tais como: RDF, Dublin Core e algumas especificações do ISO para o armazenamento de informações de localização e de data. O formato proposto pelo PRISM oferece um conjunto simplificado de vocabulário, de forma a permitir uma maior interoperabilidade, intercâmbio e o reuso. Apesar deste padrão aparentemente oferecer mais vantagens do que o padrão Exif, a maioria das câmeras fotográficas ainda utilizam o Exif para armazenar informações nos metadados das fotografias.

A subseção seguinte apresenta como está organizada a estrutura do Exif para as imagens compactadas e descompactadas. Em seguida, serão listadas algumas ferramentas e bibliotecas que podem ser utilizados para visualizar e manipular os atributos do Exif.

### 2.2.1 Estrutura

A especificação Exif é suportada por formatos já existentes, tais como o JPEG, o TIFF Rev. 6.0 e o RIFF WAVE, apenas acrescentando os seus metadados específicos [10]. Estes metadados são muito amplos, ou seja, realizam a cobertura de dados de diversas características, tais como: informações de data e hora da captura, dados técnicos da câmera no momento da captura da imagem, características e configurações das imagens, informações de direitos autorais, entre outros.

O formato Exif pode ser utilizado tanto para realizar anotação em imagens descompactadas como nas imagens compactadas. No caso das imagens descompactadas elas são armazenadas segundo o TIFF Rev. 6.0. A sigla TIFF significa Tag Image File Format (<http://www.fileformat.info/format/tiff/>), que pode ser traduzido como formato de arquivo de imagem com atributos. Este formato é estruturado segundo a Figura 2.7. Ele está dividido em estruturas chamadas IFD (Image File Directory). Cada estrutura IFD contém um conjunto de atributos do Exif. Nesse formato existem quatro IFDs: o primeiro com dados

gerais da imagem; outro com tags privadas do Exif; outro contendo informações de localização geográfica e por último um com informações relativas à miniatura da imagem. No início do arquivo está o cabeçalho TIFF que aponta para o primeiro IFD que contém os dados relativos à imagem e os apontadores para os outros IFDs.

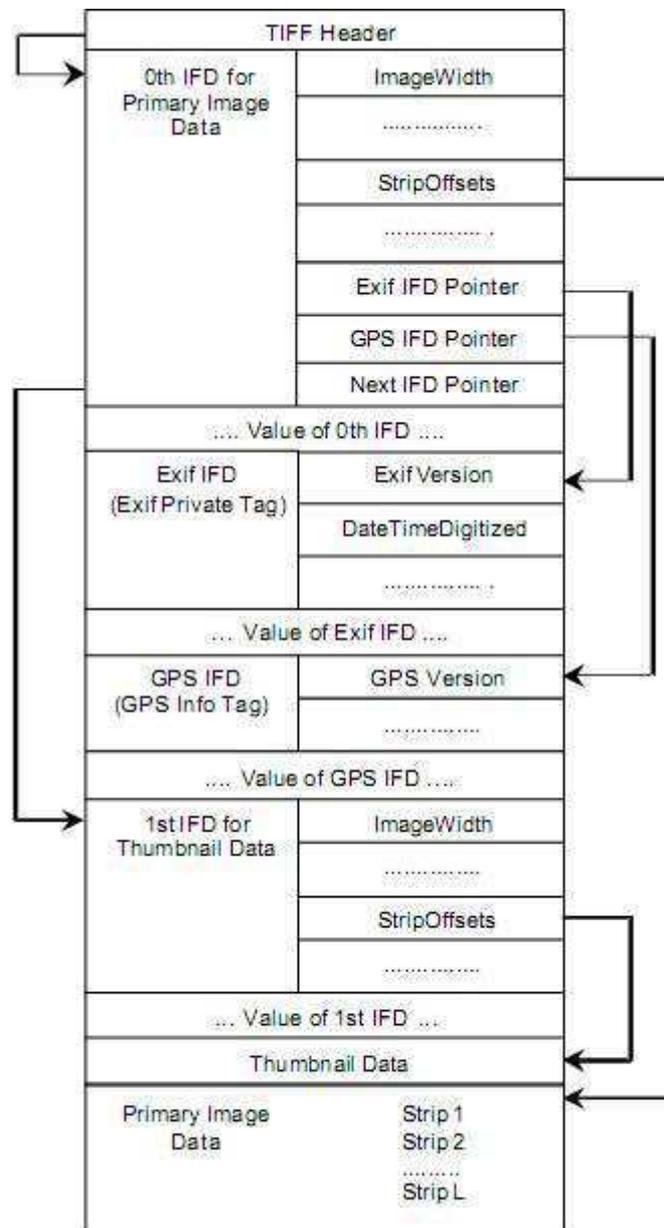


Figura 2.7: Estrutura de uma imagem não compactada. Figura extraída de [10]

Para as imagens compactadas utiliza-se o formato JPEG DCT, especificado pelo ISO/IEC 1091801, com o incremento de um segmento APP1 (Application Market Segment). Este segmento é inserido logo após o início da imagem, o segmento SOI (Start of Image), conforme

mostra a Figura 2.8. A estrutura APP1 é utilizada para armazenar atributos relativos ao Exif utilizando o mesmo padrão do TIFF, com um cabeçalho e os segmentos IFDs.

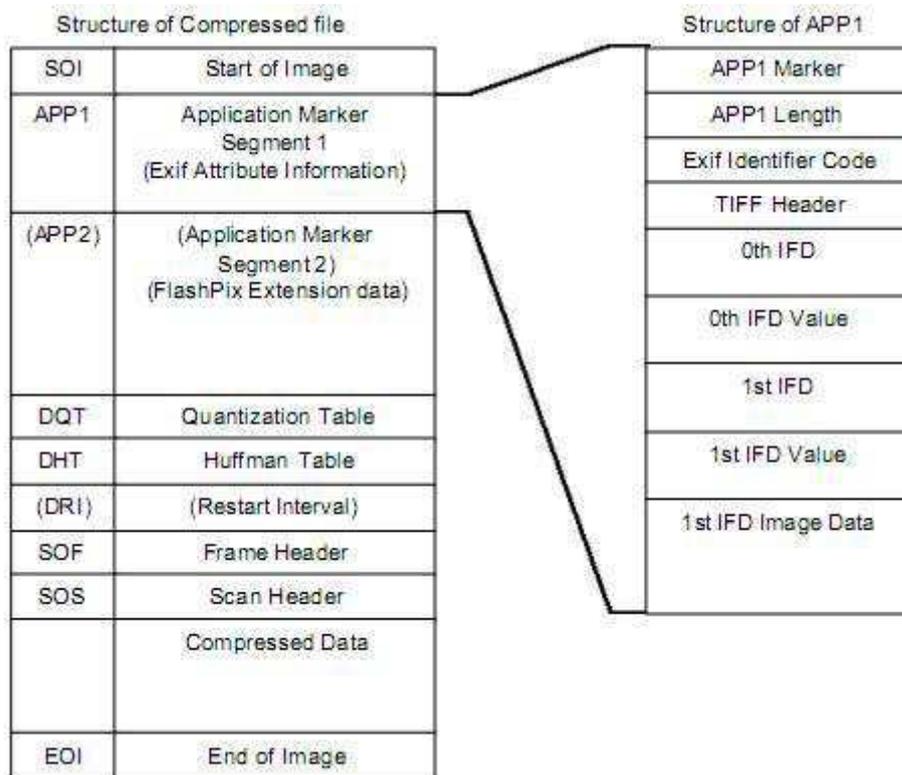


Figura 2.8: Estrutura de uma imagem compactada JPEG. Figura extraída de [10]

### Estrutura de um IFD

A estrutura IFD é utilizada para armazenar as informações relativas aos metadados das imagens. Ela é composta pelos seguintes elementos: atributo, tipo, contador e offset do valor.

O atributo é um campo 2 bytes utilizado para identificar um determinado atributo unicamente. O campo tipo é a definição do tipo do campo, podendo ser: byte, ASCII, short, long, rational, undefined, slong ou srational. O contador é a quantidade de bytes que possui o tipo escolhido. E por último, o offset do valor que é a posição em que o valor daquele atributo está gravado.

### IFDs disponíveis

O segmento reservado para o armazenamento dos metadados está subdividido em IFDs, que são estruturas que contêm os atributos dos metadados. O Exif adota quatro estruturas desse

tipo:

- **Dados Primários:** este IFD contém informações utilizadas pelo Exif que são definidas pelo formato TIFF. Elas são relativas à estrutura da imagem, ou seja, informações como o tamanho da imagem, a resolução, esquema de cores utilizado, data da captura, descrição da imagem, fotógrafo, informações de direitos autorais, entre diversas informações que podem ser vistas detalhadamente na Tabela 2.1.

Tabela 2.1: Atributos definidos pelo TIFF. Adaptada de [10]

Nome da Tag	Nome do Campo
<b>A. Tags relativas a estrutura da imagem</b>	
Comprimento da imagem	ImageWidth
Altura da imagem	ImageLength
Número de bits por componente	BitsPerSample
Esquema de compressão	Compression
Composição dos Pixels	PhotometricInterpretation
Orientação da Imagem	Orientation
...	...
<b>B. Tags relativas ao offset de gravação</b>	
Localização dos dados da Imagem	StripOffsets
Número de colunas por strip	RowsPerStrip
...	...
<b>C. Tags relativas as características da imagem</b>	
Função de transferência	TransferFunction
Cromaticidade do ponto branco	WhitePoint
Espaço de transformação de cor da matriz de coeficientes	YCbCrCoefficients
...	...
<b>D. Outras Tags</b>	
Data e Hora da mudança do arquivo	DateTime
Título da Imagem	ImageDescription
Fabricante do equipamento	Make
Modelo do equipamento	Model
Software utilizado	Software
Pessoa que criou a imagem	Artist
Direitos autorais	Copyright

- **Atributos privados do Exif:** esta estrutura contém um conjunto de atributos definidos pelo Exif relativos a diversas características, vide a Tabela 2.2. Elas estão divididas em oito categorias listadas a seguir:

Tabela 2.2: Atributos privados do Exif. Tabela adaptada de [10]

Nome da Tag	Nome do Campo
<b>A. Tags relativas à versão</b>	
Versão do Exif	ImageWidth
Versão do Flashpix suportada	FlashpixVersion
<b>B. Tags relativas às características da imagem</b>	
Informações do espaço de cores	ColorSpace
<b>C. Tags relativas à configuração da imagem</b>	
Significado de cada componente	ComponentsConfiguration
Modo de compressão de imagem	CompressedBitsPerPixel
Comprimento válido da imagem	PixelXDimension
Altura válida da imagem	PixelYDimension
<b>D. Outras Tags</b>	
Notas do fabricante	MakerNote
Comentários dos usuários	UserComment
<b>E. Tags relativas a informação de arquivos relacionados</b>	
Arquivo de áudio relacionado	RelatedSoundFile
<b>F. Tags relativas a Data e Hora</b>	
Data e hora da geração original de dados	DateTimeOriginal
Data e hora da geração do dado digital	DateTimeDigitized
Subsegundos da data e hora	SubSecTime
Subsegundos da Data e Hora originais	SubSecTimeOriginal
Subsegundos da Data e hora digitalizadas	SubSecTimeDigitized
<b>G. Tags relativas às condições do momento da captura</b>	
Vide Tabela 2.3	
<b>H. Outras Tags</b>	
ID único da imagem	ImageUniqueID

- **Versão do Exif:** contém informações relativas à versão do Exif utilizada e à versão do FlashPix ([http://www.i3a.org/i\\_flashpix.html](http://www.i3a.org/i_flashpix.html)) suportado.
  - **Características da imagem:** possui um campo relativo a informações do espaço de cores utilizado.
  - **Configurações da imagem:** armazena informações relativas à configuração da imagem, tais como modo de compactação, tamanho válido da imagem e configurações de componentes.
  - **Informações do usuário:** possui anotações criadas pelo fabricante e os comentários dos usuários sobre aquela imagem.
  - **Arquivos relatados:** contém a referência para os arquivos de áudio que relatam uma determinada imagem.
  - **Data e hora:** armazena as informações temporais relativas à imagem de forma detalhada.
  - **Condições no momento da captura da imagem:** esta categoria é a maior de todas, possuindo diversas informações de ajustes da câmera no momento da captura da imagem. Informações como tempo de exposição, velocidade do obturador, programa de exposição utilizado, brilho, contraste, modo de utilização do flash, taxa ISO utilizada, entre outras informações são armazenadas nesta categoria. A Tabela 2.3 contém a lista de vários atributos desta categoria.
  - **Outras tags:** Esta categoria contém apenas uma única tag relativa a identificação única da imagem.
- **GPS:** a estrutura GPS é responsável por armazenar todas as informações coletadas de um equipamento de GPS acoplado ou embutido na câmera. Essas informações podem ser latitude, longitude, altitude, velocidade, hora, data, quantidade de satélites, entre outros atributos. A Tabela 2.4 contém a lista de alguns dos atributos disponíveis.
  - **Interoperabilidade:** esta categoria possui o atributo responsável por indicar a regra utilizada para identificar a interoperabilidade, cujo nome do campo é `InteroperabilityIndex`. Esta pode ser relativa às regras recomendadas pelo Exif ou DFC [18] que é o conjunto de regras de design para o arquivo de sistema das câmeras.

Tabela 2.3: Atributos dos ajustes da câmera. Tabela adaptada de [10]

Nome da Tag	Nome do Campo
<b>G. Tags relativas ao momento de captura da foto</b>	
Tempo de exposição	ExposureTime
Número do Foco	Fnumber
Velocidade do ISSO	ISOSpeedRatings
Velocidade do obturador	ShutterSpeedValue
Abertura	ApertureValue
Brilho	BrightnessValue
Fonte de Luz	LightSource
Flash	Flash
Modo de Exposição	ExposureMode
Taxa de Zoom Digital	DigitalZoomRatio
Contraste	Contrast
Saturation	Saturation
Nitidez	Sharpness
...	...

Tabela 2.4: Atributos do GPS. Tabela adaptada de [10]

Nome da Tag	Nome do Campo
<b>G. Tags relativas ao GPS</b>	
Versão da tag do GPS	GPSTimeStamp
Latitude Norte ou Sul	GPSLatitudeRef
Latitude	GPSLatitude
Longitude Leste ou Oeste	GPSLongitudeRef
Longitude	GPSLongitude
Referência de Altitude	GPSAltitudeRef
Altitude	GPSAltitude
Hora do GPS (Relógio Atômico)	GPSTimeStamp
Satélites do GPS utilizados para medir	GPSSatellites
Situação da recepção do GPS	GPSStatus
Unidade de Velocidade	GPSSpeedRef
Velocidade do receptor GPS	GPSSpeed
Correção diferencial do GPS	GPSDifferential
...	...

### 2.2.2 Ferramentas e Bibliotecas para manipulação de Exif

Há uma série de ferramentas e bibliotecas disponíveis para a visualização e manipulação das anotações em Exif das fotografias digitais. Algumas ferramentas levantadas que fornecem ao usuário uma interface para a visualização das anotações das imagens são: Picasa (<http://picasa.google.com>), Exif-Viewer ([http://www.amarra.de/exif\\_uk.htm](http://www.amarra.de/exif_uk.htm)), EXIF Image Viewer ([http://home.pacbell.net/michal\\_k/exif\\_v.html](http://home.pacbell.net/michal_k/exif_v.html)) e Opanda IExif for Mozilla Firefox ([http://www.opanda.com/en/iexif/iexif\\_firefox.htm](http://www.opanda.com/en/iexif/iexif_firefox.htm)).

Também há ferramentas que permitem ao usuário editar as anotações das imagens, tais como: Exifer for Windows (<http://www.friedemann-schmidt.com/software/exifer>), jhead (<http://freshmeat.net/projects/jhead/>), ExifTool (<http://owl.phy.queensu.ca/~phil/exiftool/>) e PhotoMapper (<http://software.copiks.se/photomapper>).

Existem bibliotecas em diversas linguagens de programação disponíveis para os desenvolvedores, tais como: Metadata Extractor para a linguagem Java (<http://www.drewnoakes.com/code/exif>) e para a linguagem .NET (<http://renaud91.free.fr/MetaDataExtractor>); o libexif (<http://libexif.sourceforge.net>) para a linguagem C; o Exiv2 (<http://www.exiv2.org>) para programadores C++; dentre outras.

## 2.3 Sistemas de Informação Geográfica

Pela definição de Bernhardsen [4], “O termo *sistema de informação geográfica* (SIG) é normalmente utilizado de forma genérica para qualquer habilidade baseada em computador para a manipulação de dados geográficos. Um SIG inclui não somente o hardware e software, mas também os dispositivos especiais utilizados para inserir mapas e para criar produtos de mapas, juntamente com o sistema de comunicação necessário para interligar os vários elementos”.

Há diversas áreas em que os SIGs podem ser aplicados, tais como: militar, governo digital, transportes, agricultura, arqueologia, entre outras. O uso desse sistema para analisar o solo de uma determinada região é um exemplo de aplicação na agricultura.

Como já foi visto anteriormente, o dado espacial descreve dados que estejam associados a alguma dimensão espacial. Caso esta dimensão seja a superfície terrestre, então esse dado

pode ser chamado de dado geográfico. Os dados espaciais em um SIG podem representar feições geográficas como ruas, cidades, bairros, países, terrenos, entre outras. Já os dados não-espaciais podem ser no formato textual, numérico, áudio-visual ou de imagem. Esses dados podem ser utilizados para representar nomes de cidades, população de estados, entre outros.

Pela natureza espacial dos objetos de um SIG há a necessidade de se ter um modelo de dados que represente as feições geográficas. Por conta disto, a ISO e a Open Geospatial Consortium (OGC) definiram os tipos de dados geográficos e as funções que devem ser implementadas por um SGBD [34].

A Figura 2.9 apresenta os tipos de dados definidos pela ISO e OGC. O elemento raiz dos tipos de dados é um Geometry que possui os subtipos: Point (ponto), Curve (curva), Surface (superfície) e Geometry Collection (coleção de geometria). Esta coleção de geometria, por sua vez, é composta pelos objetos: MultiSurface, MultiCurve e MultiPoint. Cada um desses objetos pode ser composto por uma coleção de Surface (superfície), Curve (curva) e Point (ponto).

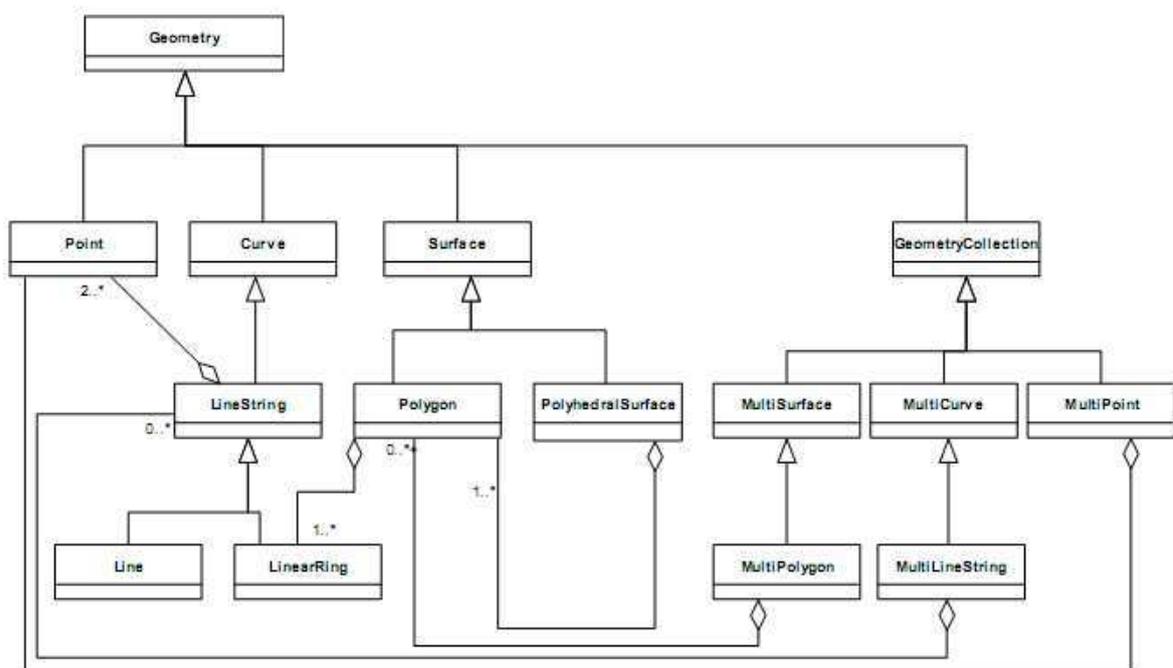


Figura 2.9: Tipos de dados geográficos definidos pela ISO e OGC. Figura adaptada de [34]

A manipulação e análise dos dados espaciais podem ser feitas através das atualizações

e consultas espaciais. Estas podem ser realizadas através da linguagem SQL, utilizando as operações de atualização, seleção, junção e projeção. Por exemplo, para representar as cidades de um país e as ruas de uma cidade, o código em SQL para a criação das tabelas que iriam conter essas informações ficaria semelhante ao Código Fonte 2.1.

Há determinados SIGs que utilizam as informações temporais dos dados como sendo uma dimensão a ser analisada. Por exemplo, em um Sistema de Informação Geográfica que analisa zonas de desmatamento da Floresta Amazônica, o usuário do sistema pode desejar visualizar todas as imagens de satélite de uma determinada região entre os anos de 1980 e 1990.

#### Código Fonte 2.1: Exemplo de Código SQL

```
1 CREATE TABLE cidade (  
2     nome VARCHAR(100) ,  
3     num_habitantes integer ,  
4     geometria Geometry  
5 );  
6  
7 CREATE TABLE ruas (  
8     nome VARCHAR(100) ,  
9     cep VARCHAR(20) ,  
10    geometria Linestring  
11 );
```

## 2.4 Bibliotecas Digitais Geográficas

O advento da Internet juntamente com os avanços da informática impulsionou um enorme aumento na quantidade de arquivos digitais disponíveis. Grande parte desses arquivos são documentos, fotografias, áudio e vídeo clipes que necessitam de mecanismos para auxiliar na recuperação desses documentos.

Com o intuito de auxiliar a recuperação desses arquivos no meio de uma imensidão de informações, foram criadas as bibliotecas digitais [15]. Estas, possuem mecanismos de indexação, consulta, apresentação e análise desses dados. As bibliotecas digitais devem funcionar de modo:

- **Eficiente:** o resultado deve ser o esperado pelo usuário e vir em um tempo satisfatório;
- **Integrado:** os dados devem ser abstraídos em um único modelo de dados, para que a biblioteca possa manipular diversos tipos de dados;
- **Independente de plataforma:** não deve ser dependente de plataforma;
- **Distribuído:** os usuários devem poder acessar o sistema remotamente.

Nos últimos anos, as empresas fornecedoras de mapas têm disponibilizado seus mapas de forma eletrônica, fazendo com que o acesso a essas fontes de dados seja realizado mais facilmente. Aliados a isso, os avanços nos hardware e software disponíveis têm gerado documentos com informações de contexto, tais como localização geográfica e informações temporais.

Por exemplo, um usuário com uma câmera digital com GPS embutido ao capturar uma fotografia do “Monumento dos Tropeiros da Borborema”, na cidade de Campina Grande, a câmera digital automaticamente registrará juntamente com a fotografia as coordenadas geográficas: latitude 7°13’30.73” ao sul e longitude 35°52’39.00” ao oeste; e as informações temporais relativas ao momento da captura.

As bibliotecas geográficas [25] foram criadas com o intuito de suprir a necessidade da recuperação de documentos no domínio espaço-temporal. Elas são bibliotecas digitais com suporte a consultas específicas do domínio espacial e temporal. Dessa forma, elas suportam dados georreferenciados, ou seja, que estão relacionados a uma coordenada geográfica.

Existem dois fatores que dificultam o uso desse tipo de biblioteca. O primeiro é a complexidade que há nos dados, sendo necessário utilizar formas especiais de indexação, consulta e armazenamento. O segundo fator é a dificuldade dos usuários trabalharem no domínio espaço-temporal, onde se exige conhecimentos específicos.

As bibliotecas geográficas podem ser especializadas para determinados tipos de dados, não precisando, necessariamente, suportar todos os tipos de dados. Por exemplo, pode-se ter uma biblioteca geográfica especializada em fotografia digital [43] [52] [53]. As funcionalidades desse tipo de biblioteca devem ser focadas para o tipo de dado a ser tratado.

Então, para o desenvolvimento de uma biblioteca geográfica para fotografias digitais, com base na pesquisa realizada, foi levantado um conjunto de características importantes

para um sistema deste tipo, listadas na Tabela 2.5. Essas características estão descritas a seguir.

Tabela 2.5: Características desejáveis a uma bibliotecal digital especializada em fotografias georreferenciadas

C1	Organização automática das fotos utiliza metadados temporais;
C2	Organização automática das fotos utiliza metadados espaciais;
C3	Organização automática das fotos utiliza o conteúdo;
C4	Uso de informações da agenda de compromissos para auxiliar a organização automática;
C5	Uso de informações de redes baseadas em mapas para auxiliar a organização automática;
C6	Interface de navegação hierárquica temporal;
C7	Interface de navegação hierárquica espacial;
C8	Interface de navegação em mapas;
C9	Anotação automática de lugares;
C10	Ferramenta para auxiliar a anotações de pessoas;
C11	Sugestão de anotação utiliza a recorrência de anotações realizadas anteriormente e metadados espaço-temporal;
C12	Sugestão de anotação utiliza informações da agenda de compromissos;
C13	Sugestão de anotação utiliza o conteúdo das imagens;
C14	Sugestão de anotação utiliza informação dos contatos dos usuários;
C15	Ferramenta para permitir o usuário realizar anotações manuais;
C16	Ferramenta para georreferenciar fotografias;
C17	Expansão de dados de contexto;
C18	Mecanismo de busca textual avançada para consulta em meta dados das fotos;
C19	Mecanismo de busca espacial avançada;
C20	Anotação em ambiente móvel;
C21	Repositório na Web.

### 2.4.1 Organização automática das fotografias

Alguns estudos mostram que o evento e a posição geográfica da fotografia são as características mais marcantes para os usuários recordarem de uma foto [43]. É importante que

o sistema ofereça um mecanismo para automaticamente organizar uma coleção de fotografias em eventos, utilizando informações dos metadados, tais como informações temporais e espaciais, ou informações do conteúdo das fotografias (a partir de técnicas de CBIR).

Existem estudos que levantam uma hipótese da explosão [43]. Esta hipótese afirma que o usuário na maioria das vezes captura fotografias apenas de eventos importantes e nestes eventos ele captura as fotos em intervalos de tempo bem semelhantes.

É desejável que uma biblioteca digital de fotografias possua algoritmos para separar as fotografias dos usuários em eventos de maneira automática, semi-automática e manual. No modo automático o sistema deverá por si só separar as fotografias em eventos sem nenhuma intervenção do usuário. O modo semi-automático implica em o usuário fornecer alguns parâmetros para ajudar na detecção dos eventos. Já no modo manual o usuário separa as fotografias em eventos manualmente, ou seja, sem nenhuma automatização do sistema.

O organizador também poderá fazer uso de informações extraídas a partir de fontes externas para obter maiores informações contextuais das fotografias ou do dono da coleção de fotografias. Por exemplo, consultar Gazetteers ou redes sociais baseadas em mapas para extrair maiores informações da localização da fotografias ou acessar a agenda de compromissos digital do dono das fotos para coletar informações do contexto do mesmo.

### 2.4.2 Navegação multi-modal

Outra funcionalidade essencial para a biblioteca digital é possuir mais de uma forma de os usuários poderem navegar nas fotografias. O usuário deverá escolher àquela que se sentir mais confortável. Uma possibilidade seria o sistema ter uma navegação através de um mapa e outra através de uma hierarquia, esta podendo ser temporal ou espacial.

Uma hierarquia temporal contempla todas as fotos, sejam elas georreferenciadas ou não. O sistema pode prover uma forma de navegação hierárquica em vários níveis, por exemplo, uma hierarquia com dois níveis (ano e mês), quando no mês não tiver muitos eventos. Podendo haver um terceiro nível (dia), no caso de haver uma grande quantidade de eventos em um determinado mês.

Na navegação hierárquica espacial o sistema deverá organizar apenas as fotos georreferenciadas. Para os lugares que não possuam muitas fotos esta hierarquia será de três níveis (país, estado e cidade), caso contrário, essa quantidade de níveis deverá ser aumentada.

Para a navegação a partir de mapas, os eventos deverão ser exibidos no mapa eletrônico que poderá ser navegado utilizando ferramentas de zoom e pan.

### 2.4.3 Mecanismo para Auxiliar à Anotação

A anotação é fundamental para aumentar a cobertura e precisão nas buscas de fotografias. Entretanto, o usuário não deseja realizar muito esforço para anotar, então é importante que o sistema possua mecanismos que possam automatizar ou auxiliar o processo de anotações de fotos. Para isso este sistema deverá prover para o usuário diversos mecanismos de auxílio à anotação.

Há informações que são bastante importantes para as pessoas recordarem das fotografias, tais como, a localização geográfica e as pessoas e os objetos que estavam presentes nas fotografias. Desta maneira, a biblioteca digital deverá fornecer ao usuário uma lista com sugestões do que está presente na fotografia, por exemplo, das pessoas que estão presentes. Esse mecanismo pode utilizar as informações de contexto ou o conteúdo das fotografias para tentar inferir quem está presente na foto. Também, pode-se ampliar as informações de contexto a partir de fontes externas e utilizá-las na sugestão de anotação. Essas fontes podem obter informações da agenda de compromissos do usuário, ou ainda, informações sobre a localização geográfica dos contatos do usuário no instante da fotografia. Com relação à localização geográfica é essencial que a biblioteca digital capture automaticamente as informações do lugar (cidade, estado ou país) que a foto foi capturada.

Uma funcionalidade muito importante é o sistema possibilitar ao usuário editar os metadados das fotos. Uma situação muito comum é de muitas vezes o usuário não ajustar o relógio de sua câmera digital, fazendo com que essa gere fotografias com informações temporais inválidas. Então o sistema deve prover uma forma do usuário poder editar esses metadados.

A maioria das câmeras digitais da atualidade não possui o GPS embutido, o que faz com que capturem imagens sem informações geográficas. A biblioteca deve possuir uma ferramenta que possibilite ao usuário georreferenciar um conjunto de fotos através de um mapa eletrônico.

#### 2.4.4 Consulta textual simples e avançada

É fundamental a biblioteca digital permitir ao usuário a realização de consultas textuais para auxiliar a recuperação das imagens. Existem diversas informações nos metadados das fotografias que podem ser utilizadas como parâmetros para a busca. O usuário deve ter disponível duas interfaces para esta funcionalidade, uma simples e outra avançada.

Na consulta simples, o usuário a partir de um único texto consultará informações nos metadados de descrição da foto. Por exemplo, se o usuário digitar a consulta “monumento”, então o sistema deverá pesquisar todas as fotos que contenham a palavra “monumento” em seus metadados. Se o usuário preferir, poderá filtrar a consulta de forma que apareçam apenas as fotos que estejam dentro da área que está sendo exibida no mapa.

Já no caso da consulta avançada, o usuário poderá detalhar a filtragem da sua consulta. Por exemplo, o sistema deve permitir ao usuário realizar consultas do tipo “recupere as fotografias capturadas em 02 de Dezembro de 2007 na cidade de Campina Grande com a câmera do modelo A500 da marca Samsung utilizando o flash”.

A biblioteca deverá possibilitar a ampliação do contexto das fotografias a partir de seus dados espaciais e temporais visando aumentar a precisão das consultas. Existem *web-services* que fornecem dados de clima, temperatura, estação do ano, altitude, entre outras informações. Estes dados poderão ser utilizados na consulta textual para ampliar a semântica das buscas.

#### 2.4.5 Consulta temporal

A informação temporal é um dado muito importante na semântica da foto. Por este motivo o sistema deverá prover um mecanismo com o qual o usuário possa realizar consultas nas informações temporais das fotografias.

Dessa forma a interface de consulta temporal deverá disponibilizar os operadores topológicos básicos para os usuários: antes, igual, encontra, durante, sobrepõe, inicia e termina. Por exemplo, o sistema deve permitir realizar consultas do tipo “recupere todas as fotografias capturadas entre os anos 2000 e 2001”.

### 2.4.6 Consulta espacial simples e avançada

A biblioteca digital deve possuir um mecanismo de busca espacial, ou seja, que envolvam operações espaciais na filtragem das fotos. Deverão existir duas maneiras de realizar a consulta espacial, uma de forma simples e outra avançada.

Na consulta espacial simples, o usuário através de um único texto poderá localizar as fotos que estão em um determinado local, por exemplo, o usuário digita "Campina Grande" como palavra-chave da sua consulta. A partir daí o sistema deverá processar, a partir das coordenadas geográficas da cidade de Campina Grande, as fotos cujas coordenadas se localizam dentro daquela cidade. A partir daí o usuário poderá visualizar aquelas fotos tanto no mapa, como de forma hierárquica, com as fotos organizadas por eventos e datas.

Já na consulta avançada, o usuário poderá realizar operações espaciais mais avançadas, ou seja, o usuário poderá realizar operações mais específicas, como por exemplo, "recupere as fotos que estão num raio de até 50 km de Campina Grande". O usuário deverá ter disponível operadores espaciais para a realização de consultas. A partir do resultado, o usuário poderá ordená-las de acordo com determinados critérios como "mais recentes", "pela ordem alfabética" e "as mais próximas".

### 2.4.7 Repositório Web

É muito importante que o sistema disponibilize um repositório web para o usuário armazenar suas fotografias. Há dois fatores que impulsionam o sistema a ter esse repositório: a conectividade e a segurança dos dados. O gerenciamento de fotografias que estão armazenadas de maneira distribuída não é um trabalho simples e exige o uso de técnicas para fazer a sincronização de pastas. Uma solução para este problema é apresentado no trabalho de Gargi et al. [22].

A conectividade é motivada pela ampla disseminação da Internet, onde o acesso está disponível nos lugares mais remotos do planeta, e a grande quantidade de dispositivos que permitem realizar a conexão com a Web. Permitindo, com isso, que o usuário possa acessar sua de fotografias onde quer que esteja.

É muito comum os servidores Web possuírem software e hardware menos propícios a falhas. Isto se deve ao fato desses equipamentos possuírem mecanismos de replicação de

dados, rotinas de backup automático, entre outras formas de evitar a perda de dados. Por estes motivos os servidores Web são mais seguros que os computadores pessoais.

# Capítulo 3

## Trabalhos Relacionados

Na literatura existem alguns sistemas propostos que contêm algumas funcionalidades das bibliotecas digitais de fotografias com suporte de geoprocessamento, propostas no capítulo anterior.

Desta maneira, este capítulo é responsável por apresentar um levantamento bibliográfico das pesquisas na área sobre: organização automática de fotografias, sugestão de anotação e sistemas de consultas.

### 3.1 Organização Automática de Fotografias

O problema da detecção automática de eventos tem sido extensivamente estudado recentemente [9] [21] [26] [43] [47] [48]. As principais abordagens propostas são: visual, temporal e espacial. A abordagem visual é baseada no conteúdo da fotografia, por exemplo, características visuais extraídas utilizando técnicas de CBIR [60]. As fotos capturadas em um mesmo evento possuem características visuais similares. Exemplos destas características incluem cor, textura, forma e estrutura. Por exemplo, fotos tiradas durante uma caminhada na praia em um dia de sol, irão conter muita similaridade visual.

Rodden e Wood em [49] analisaram como as pessoas gerenciam as suas fotografias digitais, e constataram que as principais características de um organizador de fotografias são a organização automática e a exibição de miniaturas das fotografias.

Por outro lado, a abordagem baseada nos metadados temporais usa a data e hora extraída dos metadados das fotografias (por exemplo, do Exif, que é utilizado pela maioria das

câmeras digitais). Portanto, esta abordagem é baseada na evidência de que as fotos são tiradas sequencialmente em um dado evento. Por exemplo, durante uma festa de aniversário, pessoas irão tirar fotos em diferentes tempos da festa. Existem alguns trabalhos que usam esta abordagem, como [9] [21] [26] [47] [46]. Entretanto, existem trabalhos que combinam técnicas visuais com temporais, como Ravichandran [48].

A terceira abordagem utiliza tanto metadados espaciais como temporais. Essa abordagem é baseada no fato de que um dado evento normalmente acontece em algum lugar durante algum tempo. Então, um mesmo evento pode conter fotos de localizações próximas. A pesquisa de Chen et al. [8] mostra que os organizadores de fotografias que utilizam essa abordagem possuem resultados melhores do que os que utilizam outro tipo de organização de fotos.

O WWMX [52] utiliza informações espaço-temporais para indexar as fotos, mas não implementa a detecção automática de eventos. O PhotoCompas [43] realiza uma classificação de fotos a partir da localização geográfica e dos eventos. A classificação de fotos é feita utilizando um novo algoritmo de agrupamento (clustering). Além disso, pode-se auto-denominar eventos a partir de nomes de lugares e da informação temporal.

O PhotoMap realiza anotação semi-automática utilizando o contexto espacial, temporal e social de uma foto [53]. Seus metadados incluem amigos do usuário próximos, condições climáticas, endereço de dispositivos e objetos importantes próximos. O PhotoMap provê um mecanismo para organizar, compartilhar e recuperar arquivos de fotos. Entretanto, o PhotoMap não organiza as fotografias automaticamente em eventos.

Dentre os sistemas comerciais, o iPhoto (<http://www.apple.com/br/ilife/iphoto>) detecta eventos automaticamente, mas utilizando somente informações temporais. O Picasa (<http://picasa.google.com.br>) não dá suporte à detecção automática de eventos. Já o Flickr(<http://www.flickr.com>) possibilita ao usuário organizar suas fotos baseado nos metadados espaciais e em tags. Entretanto, este não provê nenhuma organização automática das fotos.

## 3.2 Sugestão de Anotação

As soluções para auxílio a anotações de fotografias podem ser divididas em dois grupos: as que utilizam reconhecimento facial e as que utilizam informações de contexto. Também, existe as técnicas híbridas, que utilizam as duas abordagens.

Um dos trabalhos que utiliza técnicas de reconhecimento facial é o de Girgensohn et al. [23] que permite ao usuário realizar anotações das fotografias e em seguida são exibidas miniaturas de pessoas similares àquelas que foram anotadas, fazendo com que o usuário possa anotar mais rapidamente várias fotografias.

Outra solução utilizando reconhecimento facial é a proposta de Zhang et al. [57]. Segundo os autores dessa pesquisa, esta solução só funciona bem quando estão bem visíveis os dois olhos e o nariz das pessoas nas fotografias. Esta, utiliza o algoritmo do vizinho mais próximo para realizar a similaridade entre as faces. O trabalho de Zhao et. al [58] sugere um mecanismo de anotação automática que utiliza informações extraídas de um algoritmo para reconhecimento facial e corporal aliado a informações contextuais. O trabalho de O'Hare e Smeaton [46] também utiliza uma solução baseada em conteúdo e informações de contexto.

Existem trabalhos que propõem realizar a anotação através de dispositivos móveis equipados com câmeras fotográficas. Há soluções para a anotação utilizando o contexto do usuário para anotar pessoas, objetos ou categorias de atividades [11] [50].

O ACRONYM proposto por Monaghan e O'Sullivan [41] é uma abordagem para anotação semi-automática de fotografias que sugere anotação de lugares e pessoas. Para a sugestão de pessoas o sistema faz uso de informações da rede social do usuário, para analisar os relacionamentos do usuário, e das anotações já realizadas em fotografias da mesma localização geográfica. Já a sugestão de anotação de lugares utiliza informações capturadas de um gazetteer.

O PhotoMap utiliza-se da tecnologia Bluetooth para identificar, através dos respectivos telefones celulares, quais as pessoas que estão próximas a uma determinada fotografia. Este sistema também possui uma interface espaço-temporal para visualização das fotografias. No PhotoMap é proposta a anotação automática das fotografias apenas dos usuários que estão com seus dispositivos móveis ligados no momento da fotografia.

Naaman et al. [44] utiliza a idéia de que há padrões para o aparecimento das pessoas

nas fotografias. Esses padrões são utilizados como estimadores para calcular a chance de uma determinada pessoa estar numa fotografia. São propostos os estimadores: popularidade, co-ocorrência, recorrência espacial e temporal. O sistema sugere uma lista de pessoas com alta probabilidade de estar numa dada fotografia através do cálculo dos estimadores.

O WWMX (World Wide Media eXchange) foi um sistema pioneiro das bibliotecas digitais de fotografias que utilizou as informações espaciais para indexar as fotografias. Este possui interface multi-modal que permite ao usuário navegar nas fotografias tanto através de uma lista de fotografias como em um mapa digital. O WWMX não possui nenhum sistema de auxílio a anotações.

Existem diversos sistemas comerciais que possibilitam a anotação e recuperação das fotografias, porém esses utilizam técnicas ainda muito limitadas. Dentre estes sistemas, destacamos o Photoshop Elements (<http://www.adobe.com/br/products/photoshopelwin>), iPhoto, Picasa e o Flickr.

Há também soluções para anotação de fotografias em ambiente móvel. Uma dessas soluções é o Zone Tag [1]. Este, permite que os usuários enviem fotografias do dispositivo móvel para o sistema e receba a recomendação de tags para anotar estas fotografias. Outro trabalho na área de anotação em dispositivos móveis é o Mobile Media Metadata [55] que investiga a interação dos usuários com o sistema móvel e seus hábitos para anotar, capturar, armazenar, compartilhar e recuperar as fotografias.

As técnicas de análise de imagem para detecção [30] [56] e reconhecimento facial [59] para grandes conjuntos de fotografias ainda são bastante limitadas. As técnicas que utilizam informações de contexto também podem ser aperfeiçoadas com o uso do reconhecimento facial. A solução baseada em estimadores mostra-se promissora e pode ser combinada com outras informações de contexto do usuário e da própria fotografia para melhorar o cálculo das probabilidades de uma pessoa estar presente em uma determinada fotografia.

### 3.3 Consultas Avançadas

O WWMX possui um mecanismo de busca que permite a localização de fotografias a partir de sua localização geográfica. A realização de consultas textuais (a partir da descrição e nome das fotografias) e também através dos usuários do sistema. As fotos são exibidas nos

mapas a partir de media dots, que são agrupamentos de fotografias próximas espacialmente. O mecanismo de busca do WWMX é muito simples, não possibilitando ao usuário realizar consultas mais específicas.

O PhotoCompas permite uma navegação hierárquica nas fotografias de maneira textual. A partir dessa navegação hierárquica se consegue realizar filtragens nas fotografias. Por exemplo, o PhotoCompas possibilita que faça filtro do tipo: “recupere as fotografias capturadas no Rio de Janeiro e que foram capturadas no ano de 2006”. Apenas o operador booleano AND é utilizado nas filtragens. Essa é a única forma que o sistema possibilita ao usuário realizar consultas, não permitindo que se realizem consultas com filtragens mais específicas.

Já o PhotoMap armazena as informações de contexto da fotografia em uma ontologia proposta. O contexto é expandido através da conexão com Web Services (<http://www.w3.org/2002/ws/>) que capturam informações de contexto (clima, temperatura, luminosidade, entre outros). Na literatura não se conseguiu informações se há mecanismo de buscas deste sistema.

O’Hare e Smeaton em [45] propuseram um mecanismo de consulta avançado que permite o usuário realizar filtragens a partir de informações de localização, tempo e das informações extraídas da expansão de contexto da ferramenta. Porém, esta solução não realiza consultas utilizando os operadores espaciais.

Não foram encontradas na literatura informações sobre os mecanismos de consultas das outras pesquisas e sistemas analisados neste trabalho.

## 3.4 Considerações Finais

A Tabela 3.1 apresenta um estudo comparativo, com base nas características levantadas na Tabela 2.5, entre as várias pesquisas e sistemas estudados neste trabalho.

Atualmente, a abordagem que utiliza os metadados das fotografias é mais viável do que a abordagem visual para a organização automática de fotografias. Aliadas a isso, podem-se utilizar fontes de dados externas para estender as informações semânticas das fotografias para auxiliar no processo de classificação das fotografias em eventos. Através de redes sociais baseadas em mapas pode-se capturar informações do lugar (POIs) onde as fotos foram

Tabela 3.1: Características das pesquisas e sistemas analisados neste trabalho

	Ravinchandran et al.	O'Hare et al.	Cooper et al.	Gargi	Graham et al.	PhotoToc	WWMX	PhotoMap	Naaman et al.	Girgerson	Zhang et al.	ACRONYM	Photoshop Elements	Picasa	Flickr	ZoneTag	MMM
C1	X	X	X	X	X	X	-	-	X	-	-	-	X	X	X	-	-
C2	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-
C3	X	-	-	-	-	X	-	-	-	-	-	-	-	X	-	-	-
C4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C6	-	-	X	-	-	-	X	-	X	-	-	-	X	X	X	-	-
C7	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-
C8	-	-	-	-	-	-	X	-	-	-	-	-	-	X	X	-	-
C9	-	X	-	-	-	-	-	X	X	-	-	-	-	-	-	-	X
C10	-	X	-	-	-	-	-	X	X	X	X	X	X	X	-	X	X
C11	-	X	-	-	-	-	-	-	X	-	-	X	-	-	-	X	X
C12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C13	-	X	-	-	-	-	-	-	-	X	X	X	X	X	-	-	-
C14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C15	-	-	-	-	-	-	-	-	-	X	-	-	X	X	X	X	X
C16	-	-	-	-	-	-	-	-	-	-	-	-	-	X	X	-	-
C17	-	X	-	-	-	-	-	X	X	-	-	-	-	-	-	-	-
C18	-	X	-	-	-	-	X	-	-	-	-	-	X	-	X	-	-
C19	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C20	-	-	-	-	-	-	-	X	-	-	-	-	-	X	X	X	X
C21	-	-	-	-	-	-	X	-	-	-	-	-	X	X	X	-	X

capturadas. Outra fonte de dados que pode ser importante para a organização é a agenda de compromissos dos donos das coleções, onde pode-se capturar os eventos que o usuário irá participar.

A proposta de Naaman et al. que utiliza estimadores para a sugestão de anotação é bastante interessante. Uma informação muito importante para a anotação de fotografias é a localização geográfica das pessoas no momento da captura das fotografias. É óbvio que uma pessoa que está há alguns metros da câmera fotográfica no momento da captura tem grandes chances de aparecer na fotografia capturada, e quem está a vários metros da câmera com certeza não irá aparecer na fotografia. Os dispositivos móveis equipados com chip GPS e sistemas baseados em localização permitem que se conheça a localização geográfica das pessoas. Desta maneira, é possível utilizar essa e outras informações de contexto para o algoritmo de sugestão de anotação de pessoas.

É indispensável para uma biblioteca digital de fotografias que se tenha um bom sistema de consultas. A maioria das bibliotecas digitais existentes não exploram bem os mecanismos de consultas, principalmente no que se diz respeito a informação espacial e temporal.

# Capítulo 4

## Arquitetura do Sistema

Este capítulo apresenta a arquitetura do PhotoGeo, que é um protótipo de um organizador de fotografias web que utiliza metadados espaço-temporais, de forma que o usuário pode organizar, anotar e consultar fotografias de maneira simples. Grande parte dos sistemas existentes não exploram bem o metadado geográfico. Há vários problemas detectados nos sistemas existentes: os mecanismos de organização automática existentes são limitados; as interfaces de navegação são simplificadas; os mecanismos de sugestão de anotação não utilizam informações de contexto; e as possibilidades de consultas são rígidas e possibilitam poucas formas de consulta. Além disso, a pesquisas existentes exploram esses aspectos separadamente.

Primeiramente, é apresentada a arquitetura do sistema. Logo em seguida, são apresentados os módulos deste sistema e as suas principais funcionalidades. Na seção 4.3 é apresentado o modelo de dados que foi utilizado para armazenar as informações utilizadas pelo sistema. Por último, são apresentadas as conclusões deste capítulo.

### 4.1 Arquitetura do Sistema

O PhotoGeo é projetado em uma arquitetura distribuída baseada em 4-camadas, em conformidade com o padrão de projeto MVC (Model-View-Controller) [20], onde a apresentação, controle e modelo de negócio estão em camadas distintas, vide Figura 4.1.

O sistema foi implementado utilizando a linguagem de programação Java, devido às suas diversas vantagens, tais como: portabilidade, robustez, orientação a objetos, facilidade de

utilizar a linguagem e por ter diversos frameworks disponíveis na Internet [31].

A camada de apresentação do PhotoGeo utiliza o navegador Web para gerar o conteúdo da apresentação. Essa camada foi implementada através de páginas JSP (Java Server Pages) para os conteúdos dinâmicos e HTML (HyperText Markup Language) para os conteúdos estáticos.

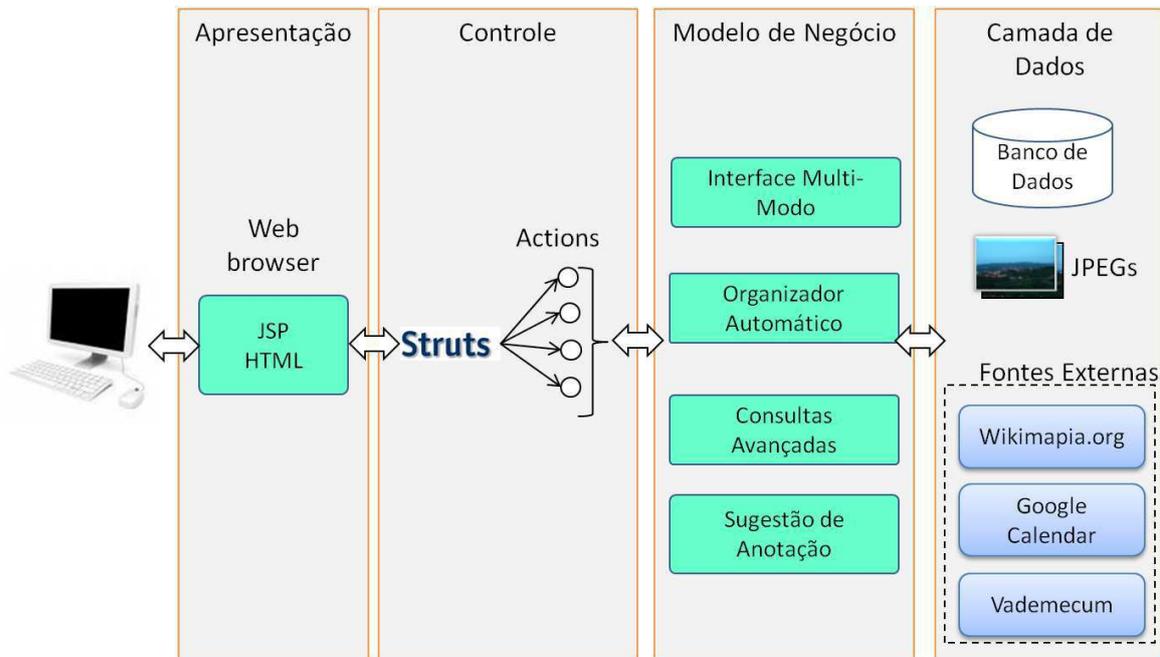


Figura 4.1: Arquitetura do PhotoGeo

A camada de controle é implementada utilizando o framework Struts 2.0 (<http://struts.apache.org/2.x/index.html>). Para essa camada foram construídos *Actions* responsáveis por tratar as requisições vindas da camada de apresentação e acessar o modelo de negócio. O Struts permite a implementação do padrão MVC de maneira simples e com suporte a diversas funcionalidades.

A camada de dados é composta de um servidor de banco de dados, dos arquivos de fotografias dos usuários e das fontes externas. O servidor de banco de dados é responsável por indexar e consultar os metadados textuais, espaciais e temporais.

Foi utilizado o SGBD PostgreSQL (<http://www.postgresql.org>). Este foi escolhido por ser um software livre, ter um certo grau de maturidade e robustez no tratamento de informações geográficas através da sua extensão PostGIS (<http://postgis.refrations.net>). O PostGIS [6] é uma extensão gratuita e de código aberto que obedece ao padrão OpenGIS

(<http://www.opengeospatial.org/standards/is>) da OGC.

As fotografias dos usuários que estão no sistema de arquivos fazem parte da camada de dados do PhotoGeo. Delas são extraídas as informações de seus metadados para serem armazenadas no SGBD. Estas também são requisitadas pelo sistema para serem visualizadas pelos usuários.

A lógica de negócios do PhotoGeo está separada em quatro módulos (Figura 4.2), a saber: navegação multi-modal, organizador automático, sugestão de anotações de pessoas e consultas avançadas. A navegação multi-modal permite que o usuário navegue no sistema através de um mapa eletrônico ou de uma hierarquia. O organizador automático separa um conjunto de fotografias do usuário de maneira automática, utilizando um algoritmo proposto neste trabalho. A sugestão de anotação de pessoas lista, para o usuário, as pessoas mais prováveis de estarem em uma determinada fotografia. Essa sugestão utiliza um algoritmo que se baseia em um conjunto de estimadores. Por último, o módulo de consultas avançadas permite que o usuário do sistema realize consultas mais específicas, utilizando as dimensões espaço-temporal e os dados convencionais. Esses módulos são detalhados na seção seguinte.

O PhotoGeo também permite que se realize anotações manuais nas fotografias, como por exemplo, editar a descrição de uma determinada foto. Caso as fotografias do usuário não sejam georreferenciadas, o PhotoGeo também possui uma ferramenta que permite o usuário georreferenciar as fotografias.

Algumas fontes externas são utilizadas pelo sistema para obter dados. Atualmente, o PhotoGeo utiliza as seguintes fontes: o Wikimapia.Org, o Google Calendar (<http://www.google.com/calendar>) e o Vademecum [14].

O Wikimapia é uma rede social baseada em mapa onde os usuários compartilham pontos de interesses de qualquer lugar do mundo. Com isso, o PhotoGeo extrai informações dos lugares onde as fotos dos usuários foram capturadas.

O Google Calendar é uma agenda de compromissos Web onde os usuários podem armazenar informações relativas a seus compromissos. O PhotoGeo extrai informações dos compromissos que ocorreram durante a captura das fotografias do usuário.

Já o Vademecum é um sistema ciente de contexto para dispositivos móveis. Uma das funções desse sistema é capturar a localização geográfica dos contatos de um determinado usuário. Assim, o PhotoGeo captura, através do Vademecum, a localização dos contatos do

usuário no momento da captura da fotografia.

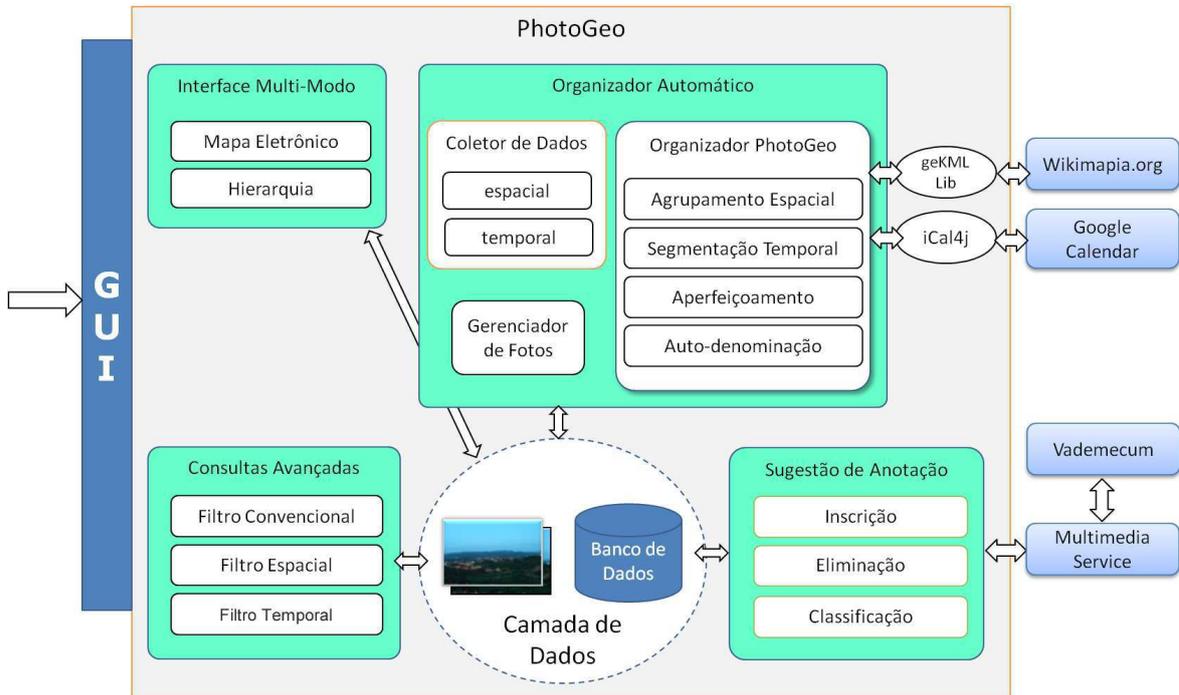


Figura 4.2: Arquitetura detalhada do PhotoGeo

## 4.2 Módulos da Arquitetura do PhotoGeo

### 4.2.1 Navegação multi-modal

Uma funcionalidade essencial para o sistema é possuir várias formas de navegação nas fotografias. Sendo assim, o usuário poderá escolher a interface que lhe for mais atraente. Para o protótipo do PhotoGeo são utilizadas duas formas de navegação, a primeira através de um mapa eletrônico, e a segunda através de uma hierarquia.

Na navegação através de mapas digitais, os eventos são exibidos em um mapa eletrônico para que o usuário possa localizar as fotografias a partir da informação geográfica, podendo utilizar as ferramentas de zoom e pan para a navegação no mapa. Para eventos que estão em pontos geográficos muito próximos devem ser agrupados de acordo com o nível de zoom atual do mapa para facilitar a navegação do usuário.

A Figura 4.3 apresenta a navegação nas fotografias do usuário que está logado no sistema. No lado esquerdo, onde há o retângulo, os eventos estão exibidos na tela através de minia-

turas que representam os eventos. Já no lado direito, dentro de uma elipse, há a navegação hierárquica, tanto na dimensão espacial como temporal.



Figura 4.3: Navegação multi-modal

A navegação hierárquica pode ser separada em dois tipos: espacial e temporal. A hierarquia temporal abrange todas as fotos de uma mesma época, mesmo as não georreferenciadas. Essa hierarquia deve possuir 3-níveis (ano, mês e evento), quando o mês não tiver muitos eventos. No caso de haver muitos eventos, a hierarquia poderá possuir até 4-níveis (ano, mês, dia e evento).

O outro tipo de hierarquia, a espacial, organiza apenas as fotos georreferenciadas. Essa deve ter 4-níveis (país, estado, cidade e evento) para lugares geográficos que não possuam muitas fotos. Caso haja muitas fotografias para um determinado lugar geográfico, essa quantidade de níveis deve ser ampliada.

## 4.2.2 Organização Automática

Este módulo é responsável por organizar de maneira automática um conjunto de fotografias pessoais. Assim, esse módulo classifica um conjunto de fotografias georreferenciadas em eventos de forma automática utilizando os metadados espaço-temporais. Para realizar esta classificação este módulo possui um algoritmo composto de quatro passos: agrupamento espacial, segmentação temporal, aperfeiçoamento e auto-denominação. No final da execução destes quatro passos o sistema terá classificado as suas fotografias em eventos, cada um com um nome específico.

A organização automática é composta de três sub-módulos: o gerenciador de fotos, o gerenciador de dados e o organizador de fotos. O gerenciador de fotos é responsável por monitorar a estrutura de armazenamento das fotos; o gerenciador de dados é responsável por obter os dados espaço-temporal de fontes externas. Por último, o organizador de fotografias automaticamente classifica as fotos em eventos utilizando metadados e dados extraídos do coletor de dados. Também, esse último sub-módulo é responsável por auto-denominar esses eventos.

No Capítulo 5 a solução proposta nesse módulo será apresentada em detalhes.

## 4.2.3 Sugestão de anotação de pessoas

As pessoas que estão presentes nas fotografias são uma das principais informações quando um usuário se recorda de uma fotografia. Sendo assim, essa informação é muito importante para a consulta de fotografias. Com isso, é muito importante que se tenham mecanismos para facilitar a anotação de pessoas que estão presentes em uma determinada fotografia.

Esse módulo é responsável por auxiliar o usuário a anotar as pessoas que estão presentes nas fotografias de sua coleção pessoal. Esse auxílio se dá através de um algoritmo que calcula a probabilidade dos contatos do usuário estarem em uma determinada fotografia. Esse algoritmo utiliza informações extraídas de um sistema ciente de contexto para dispositivos móveis, o Vademecum, e de uma agenda de compromissos do usuário, no caso foi utilizado o Google Calendar.

A interface do PhotoGeo, Figura 4.4, permite que o usuário selecione algumas das sugestões propostas pelo sistema. De imediato o sistema apresenta os resultados mais relevantes,

mas o usuário poderá visualizar os outros resultados. O usuário também poderá inserir uma nova anotação a partir de uma caixa de texto apresentada na mesma interface ou consultar por uma pessoa já existente na base de dados. O funcionamento este módulo é detalhado no Capítulo 6.



Figura 4.4: Interface da sugestão de anotações do PhotoGeo

#### 4.2.4 Consultas Avançadas

Diferente da maioria dos outros sistemas propostos, que permitem o usuário realizar apenas uma consulta simples, o PhotoGeo permite ao usuário realizar consultas avançadas e explorar bem os aspectos espaço-temporais.

O módulo de consultas avançadas do PhotoGeo permite que o usuário crie filtros compreendendo os dados convencionais (campos alfa-numéricos), temporais e espaciais.

O filtro convencional permite o usuário selecionar um campo a ser filtrado, o operador e indicar o valor a ser utilizado. Por exemplo, o usuário poderá criar um filtro para restringir

a consulta apenas das fotografias que possuem a altura da fotografia maior que 800 pixels. O filtro temporal permite que o usuário filtre as fotografias que foram capturadas em um determinado instante de tempo. Para isso, há disponível os operadores para dados temporais (antes, depois, entre e durante) para o usuário realizar a restrição adequada. Por último, o filtro espacial, que permite o usuário utilizar os operadores espaciais para restringir o lugar em que as fotografias foram capturadas. Este módulo é detalhado no Capítulo 7.

### 4.3 Esquema de Dados

Há diversos dados que necessitam ser armazenados no banco de dados do PhotoGeo, desde informações relativas aos usuários do sistema até os metadados das fotografias. O esquema de dados do sistema é apresentado na Figura 4.5 através de um Diagrama Entidade-Relacionamento. Este esquema contém as entidades usuário, álbum, contato, agenda, compromissos, diretórios, evento, foto, país, estado, cidade e ponto de interesse.

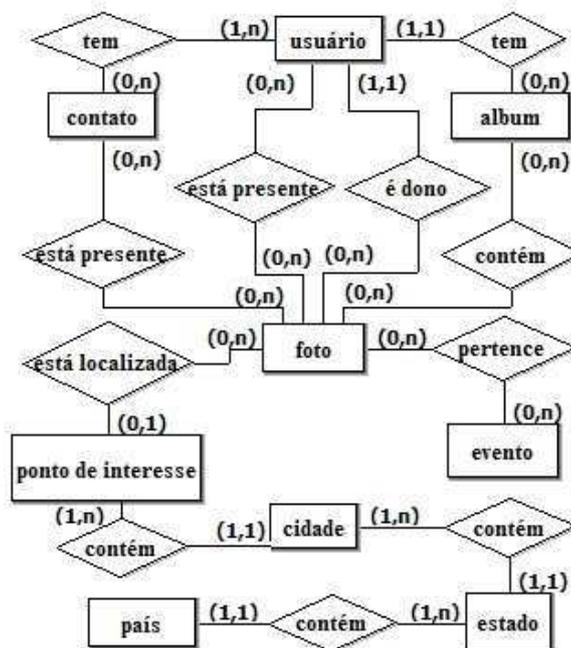


Figura 4.5: Diagrama de Entidade-Relacionamento

A entidade usuário é responsável por armazenar as informações relativas aos usuários do sistema, como por exemplo, o login e a senha de acesso ao sistema. A entidade álbum

é uma coleção de fotografias do usuário que deve monitorar um conjunto de diretórios do sistema de arquivos. Em contato são armazenadas as informações relativas aos contatos do usuário, como por exemplo, o nome e a localização geográfica dos usuários no momento de captura das fotografias. A entidade compromisso representa os compromissos que são extraídos da agenda de compromissos do usuário. Cidade, estado e país são informações espaciais relativas às diversas divisões administrativas do planeta. A entidade foto contém todas as informações que são extraídas dos metadados das fotografias. O evento representa os eventos dos usuários que foram capturadas fotografias. E por último, a entidade ponto de interesse, que é responsável por armazenar todas as informações sobre os lugares extraídos das fontes externas.

No sistema, podem ter vários usuários, onde cada um destes pode ter vários álbuns. Cada álbum irá ter diretórios que serão monitorados. Em cada álbum do usuário, também podem existir várias fotos que estão relacionadas a eventos. As fotos podem estar em uma determinada cidade ou lugar. Um lugar está contido em uma cidade, que está contida em um estado e este em um país. Um determinado usuário pode conter vários contatos. Estes podem estar presentes em uma fotografia e em um compromisso. O usuário também pode estar presente em uma fotografia.

## 4.4 Considerações Finais

O PhotoGeo tem várias vantagens em relação os outros organizadores de fotografias. A Tabela 4.1 apresenta um comparativo, das características levantadas na Tabela 2.5, do PhotoGeo com relação às pesquisas e sistemas já propostos. De forma mais ampla, esta comparação pode ser feita com relação à quatro aspectos separadamente: organização automática, navegação multi-modal, sugestão de anotação e consultas.

A organização automática utilizada pelo PhotoGeo, diferentemente das outras soluções existentes, utiliza informações de fontes externas, tais como o Wikimapia e Google Calendar, no algoritmo de organização automática. A partir destes dados é possível realizar uma melhor separação.

O PhotoGeo tem uma navegação multi-modal, através de uma navegação hierárquica e de um mapa eletrônico. A maioria dos sistemas propostos possuem apenas um tipo de

Tabela 4.1: Comparativo entre as características do PhotoGeo e das pesquisas (e dos sistemas) investigadas neste trabalho

	Ravinchandran et al.	O'Hare et al.	Cooper et al.	Gargi	Graham et al.	PhotoToc	WWMX	PhotoMap	Naaman et al.	Girgerson	Zhang et al.	ACRONYM	Photoshop Elements	Picasa	Flickr	ZoneTag	MMM	PhotoGeo
C1	X	-	X	X	X	X	-	-	X	-	-	-	X	X	X	-	-	X
C2	-	X	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-	X
C3	X	X	-	-	-	X	-	-	-	-	-	-	-	X	-	-	-	-
C4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X
C5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X
C6	-	-	X	-	-	-	X	-	X	-	-	-	X	X	X	-	-	X
C7	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-	X
C8	-	-	-	-	-	-	X	-	-	-	-	-	-	X	X	-	-	X
C9	-	X	-	-	-	-	-	X	X	-	-	-	-	-	-	-	X	X
C10	-	X	-	-	-	-	-	X	X	X	X	X	X	X	-	X	X	X
C11	-	-	-	-	-	-	-	-	X	-	-	X	-	-	-	X	X	X
C12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X
C13	-	X	-	-	-	-	-	-	-	X	X	X	X	X	-	-	-	-
C14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X
C15	-	-	-	-	-	-	-	-	-	X	-	-	X	X	X	X	X	X
C16	-	-	-	-	-	-	-	-	-	-	-	-	-	X	X	-	-	X
C17	-	X	-	-	-	-	-	X	X	-	-	-	-	-	-	-	-	X
C18	-	X	-	-	-	-	X	-	-	-	-	-	X	-	X	-	-	X
C19	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X
C20	-	-	-	-	-	-	-	X	-	-	-	-	-	X	X	X	X	-
C21	-	-	-	-	-	-	X	-	-	-	-	-	X	X	X	-	X	-

navegação. Poucos exploram a navegação através de uma hierarquia espacial.

Outra grande vantagem do PhotoGeo é com relação ao módulo de sugestão de anotações. A maioria das soluções existentes não utilizam informações de contexto dos contatos do usuário. O PhotoGeo propõe utilizar informações de contexto do usuário para a sugestão de anotação de pessoas, por exemplo, a localização geográfica dos contatos no momento de captura da fotografia, ou então, os usuários que estão no compromisso da agenda do usuário do momento da fotografia.

O módulo de consultas é uma funcionalidade que não foi bastante explorado por outros organizadores de fotografia. O PhotoGeo propõe um motor de consultas que permite o usuário realizar buscas mais precisas utilizando tanto as informações convencionais (tamanho da fotografia, descrição, uso do flash, dentre outras), como também informações espaciais (lugar em que a foto foi capturada) e temporais (data da fotografia).

O Apêndice A apresenta as tabelas do banco de dados e os diagramas de classe do PhotoGeo.

## Capítulo 5

# Organização Automática de Fotografias

Para que as imagens sejam localizadas, elas necessitam de anotações em diversos níveis: semântico, sintático e estrutural. Estas anotações são conhecidas como metadados. São exemplos de metadados: autor, data, lugar, descrição, pessoas presentes, fabricante e modelo da câmera, resolução, formato, foco, exposição, abertura, dentre outros. Existem pelo menos três maneiras de anotar fotografias: automaticamente, semi-automaticamente e manualmente. Obviamente que os usuários desejam ter anotações de fotografias com o menor esforço possível, então a anotação automática é a preferida pelos usuários.

Estudos recentes têm mostrado que o evento e a localização geográfica em que foi capturada uma dada fotografia estão entre as características mais importantes para os usuários lembrarem de uma determinada foto [42]. Um evento é caracterizado tanto pelo contexto da foto como do usuário durante a captura da fotografia. Um exemplo de um evento pode ser uma conferência internacional realizada em Campina Grande.

Por outro lado, as redes sociais têm sido bastante estudadas recentemente [7]. Essas redes constituem-se de comunidades online que permitem o compartilhamento de informações. Algumas dessas redes provêem o compartilhamento de informações geográficas. Os usuários dessas redes geralmente anotam e compartilham informações sobre pontos de interesses (POIs) a partir de mapas eletrônicos. Alguns possíveis pontos de interesses são: restaurantes, hotéis, teatros, museus, dentre outros. Essa informação, disponível na Internet, pode ajudar a identificar o contexto de uma localização.

Uma outra importante fonte de informação para obter o contexto do usuário é através das agendas eletrônicas, nas quais os usuários podem anotar seus compromissos e compartilhá-

los com sua rede social.

Este capítulo apresenta o módulo do PhotoGeo que é responsável por organizar automaticamente um conjunto de fotos georreferenciadas, usando metadados espaço-temporais, dados extraídos de uma rede social baseada em mapa e de uma agenda eletrônica. Nenhum dos sistemas propostos explora as informações obtidas de fontes como estas. Para validar este módulo foi utilizado um conjunto de dados contendo 2.412 fotos nos experimentos realizados.

Este capítulo está estruturado da seguinte maneira. A seção seguinte apresenta o módulo que gerencia as fotos dos usuários. A seção 5.3 apresenta o módulo responsável por coletar os dados espaço-temporais. A seção 5.4 apresenta como é executada a organização automática de fotos. A seção 5.5 discute os experimentos e a última seção conclui este capítulo.

## 5.1 Gerenciador de Fotos

O gerenciador de fotos é responsável por percorrer os diretórios do sistema de arquivos que são monitorados pelos álbuns do usuário e por coletar todos os arquivos de fotografias, extraíndo seus metadados. Estes são armazenados no servidor de banco de dados.

Para a extração dos metadados é utilizado o Metadata Extractor (<http://www.drewnoakes.com/code/exif/releases/>), que é uma biblioteca escrita na linguagem de programação Java que realiza o mapeamento dos metadados Exif de uma imagem para objetos Java.

As informações extraídas utilizando o gerenciador de fotos são armazenadas na tabela *photo* no banco de dados (Apêndice A). Esta tabela possui campos para armazenar informações como: nome do arquivo, data e hora de captura, posição geográfica, descrição, flash utilizado, foco, dentre outras.

## 5.2 Coletor de Dados

O coletor de dados é responsável por extrair informações de fontes externas ao sistema. Este está dividido em dois componentes: o coletor de dados espaciais e o coletor de dados temporais. O primeiro obtém informação geográfica de uma rede social baseada em mapa,

como o Wikimapia. O segundo é responsável por obter os compromissos do usuário de uma agenda Web que implementa o padrão iCalendar [12], como por exemplo, o Google Calendar. A seguir serão detalhados esses dois coletores.

### 5.2.1 Coletor de Dados Espaciais

O coletor de dados espaciais obtém informações espaciais de fontes externas. O PhotoGeo utiliza o Wikimapia, que é uma rede social baseada em mapa, onde as pessoas anotam em um mapa global os lugares de interesse em uma determinada localização. O PhotoGeo obtém os metadados através de um Web-Service (<http://wikimapia.org/d/>), que gera um arquivo KML (<http://www.opengeospatial.org/standards/kml/>) com os objetos espaciais que estão dentro de uma determinada área de interesse (também conhecida como bounding box). O KML é uma linguagem baseada em XML para visualizar informações geográficas em mapas Web 2D e 3D. O PhotoGeo extrai dados espaciais de um arquivo KML e os inserem no banco de dados. As tabelas cidade, estado e país do banco de dados do PhotoGeo foram alimentadas manualmente com informações extraídas do IBGE (<http://www.ibge.gov.br>).

As redes sociais permitem o compartilhamento de conhecimento de um dado domínio para uma comunidade particular de interesse, formando uma base de conhecimento. Mais recentemente, essas redes têm utilizado mapas para inserir pontos de interesse. Um exemplo de uma rede deste tipo é o Wikimapia, que visa armazenar informações de pontos de interesse, tais como: shoppings, hotéis, museus, restaurantes, bares, entre outros.

Atualmente, o Wikimapia contém aproximadamente 7.137.841 lugares em torno do mundo que foram inseridos pelos usuários. Este conjunto de dados é constantemente atualizado. Embora os dados obtidos de redes sociais abertas, tais como Wikimapia, não sejam totalmente confiáveis, existem mecanismos de auditoria que visam reduzir as inconsistências destes dados.

A informação coletada de uma rede social pode ser utilizada para organizar e descrever fotos e auxiliar os usuários a anotá-las com localizações geográficas.

## 5.2.2 Coletor de Dados Temporais

O coletor de dados temporais é responsável por adquirir informações dos compromissos do usuário. Atualmente, as agendas eletrônicas têm sido largamente utilizadas. Existem agendas muito complexas, com um rico conjunto de funções, que permitem ser acessadas de qualquer lugar e compartilhadas com outras pessoas.

Então, quando os usuários introduzem seus compromissos em uma agenda, eles estão inserindo informações do contexto do usuário em um determinado tempo do dia. Estes compromissos descrevem informações de “o que”, “quando”, “onde” e “quem” está participando de um determinado compromisso.

Existe um modelo de referência chamado CIDOC/CRM [27] que foi criado com o intuito de possibilitar a troca de informação de herança cultural entre museus. A herança cultural é qualquer registro que pode ser passado entre gerações, como por exemplo: um documento, um objeto, uma fotografia, entre outros. A informação sobre o evento em que foi capturada determinada herança cultural é uma das principais entidades deste modelo de referência. Desta forma, nota-se a importância do registro do evento para as fotografias.

O padrão iCalendar é implementado por muitos calendários, como por exemplo o Google Calendar. O PhotoGeo acessa o Google Calendar através de uma URL que é provida ao usuário para gerar um arquivo utilizando o padrão iCalendar. Este arquivo é lido, utilizando o framework iCal4j [33], e em seguida o dado é inserido no servidor de banco de dados.

## 5.3 Organizador PhotoGeo

Como pode ser visto na arquitetura do PhotoGeo (Figura 4.2), na camada de modelo de negócios, existe um módulo para a organização das fotos. Esta organização é feita automaticamente para o usuário. Esta também pode ser chamada de uma classificação espaço-temporal não-supervisionada como é conhecida no campo da mineração de dados [29].

O Organizador PhotoGeo tem as seguintes hipóteses:

(i) as fotos devem ser de uma coleção pessoal, de forma que não existem eventos em paralelo;

(ii) as fotos devem ter nos seus metadados: latitude, longitude, data e hora; e

(iii) os usuários devem usar sua agenda eletrônica Web, que implementa o padrão iCalendar, para marcar seus compromissos.

O objetivo da organização automática é reduzir o esforço dos usuários durante a organização e anotação de sua coleção de fotografias.

O algoritmo para a organização automática, publicada por Lacerda et al. [38], está dividido em quatro passos: agrupamento espacial, segmentação temporal, aperfeiçoamento e auto-denominação. Depois da execução destes quatro passos, as fotos serão separadas em eventos.

Nos dois primeiros passos do algoritmo - agrupamento espacial e segmentação temporal -, as fotos são agrupadas de acordo com as propriedades similares. No primeiro passo, os grupos de fotos serão formados baseados na proximidade geográfica. Já no segundo passo, cada grupo é dividido em sub-grupos que são próximos no tempo.

Depois da separação das fotos em grupos, algumas inconsistências podem aparecer. Por exemplo, fotos de um mesmo evento que estão em grupos diferentes, ou fotos de eventos diferentes no mesmo grupo. Então, é necessário executar um outro passo para corrigir estes erros. Este passo, chamado de aperfeiçoamento, permite a junção ou separação de grupos e a remoção de elementos de um grupo para adicioná-los em outro grupo. Desta maneira, algumas comparações com os metadados das fotografias são realizadas. Por exemplo, se duas fotos estavam em diferentes grupos devido a uma grande distância espacial entre elas, mas com relação ao tempo estão próximas uma da outra, então provavelmente elas devem estar no mesmo grupo.

O último passo é a auto-denominação dos grupos extraídos. Este passo faz um casamento dos metadados da fotografia com a informação espacial e temporal do banco de dados, atribuindo-se os nomes aos eventos.

Este módulo é composto de um algoritmo de classificação de fotos. É possível ter classificações erradas em determinados eventos ou até a criação de eventos inexistentes. Quando ocorrem essas inconsistências, o sistema permite que o usuário as corrija manualmente. Isto é, o sistema consente a organização manual das fotografias em eventos.

Em seguida, serão discutidos cada um dos quatro passos do algoritmo do PhotoGeo.

### 5.3.1 Primeiro Passo: Agrupamento espacial

O objetivo do primeiro passo é formar grupos de fotos que estão próximas geograficamente. Portanto, é necessário usar os metadados das fotos para proceder neste passo. A localização é representada como um ponto 2-D (latitude, longitude) extraída das fotos georreferenciadas.

Como a localização é expressa em duas dimensões, então não é possível usar uma estratégia baseada em segmentação, que é utilizada no próximo passo. Uma possível abordagem poderia ser utilizar as divisões políticas dos países. Por exemplo, as fotos seriam agrupadas por cidades onde foram capturadas. O problema com essa solução é que algumas fotos podem estar em um mesmo evento, mas terem sido capturadas em cidades diferentes.

Outra possível solução para o problema do agrupamento espacial de fotos seria o uso de um algoritmo não-supervisionado utilizado em mineração de dados. Neste caso, as fotos capturadas em lugares próximos, mesmo se elas são de diferentes cidades, podem ser postas em um mesmo grupo.

Foi adotada a última abordagem no algoritmo apresentado, uma vez que para fazer isto, foram testadas duas abordagens diferentes para compará-las e então escolher a mais adequada para o problema endereçado. Foram escolhidos os algoritmos K-means [39] e DBSCAN [19]. O primeiro realiza uma classificação não-supervisionada baseada no particionamento e o segundo realiza uma classificação não-supervisionada baseada na densidade. Estes algoritmos foram escolhidos por apresentarem parâmetros mais fáceis de estimar, além de possuírem um desempenho melhor para solucionar o problema.

Para testar estes dois algoritmos, foi submetido um conjunto de fotos georreferenciadas para ambos. Os resultados obtidos foram similares usando os dois algoritmos. Foi decidido utilizar DBSCAN por ser mais fácil de ajustar os parâmetros de entrada do que o K-means. Uma análise detalhada desta avaliação encontra-se em Lacerda et al. [36].

No final deste primeiro passo, existirá um conjunto de grupos que irão conter fotos próximas geograficamente. A Figura 5.1 apresenta o resultado do agrupamento espacial para quatro eventos que contêm fotos distribuídas espacialmente (pequenos retângulos representam fotos individuais). É importante dizer que, utilizando somente o agrupamento espacial, não é possível detectar os eventos apropriadamente, pois um dado grupo pode conter mais de um evento. Por exemplo, dois eventos ocorreram na cidade de João Pessoa - Paraíba, Brasil, um durante o Simpósio Brasileiro de Banco de Dados e outro em uma viagem de



A Figura 5.2 apresenta a distribuição temporal de um conjunto de fotos que foram agrupadas espacialmente. As fotografias consecutivas do exemplo que possuem um intervalo maior do que o padrão foram classificadas em eventos distintos.



Figura 5.2: Exemplo de segmentação temporal

O valor de  $t$  é dado pela soma da média, desvio-padrão e uma constante  $k$ . No protótipo desenvolvido foi escolhido  $k = 1$  hora. A constante  $k$  é utilizada para ajustar os casos em que a média e o desvio-padrão têm valores muito pequenos. No final deste passo, cada grupo terá um conjunto de eventos. Este passo pode causar, pelo menos, dois tipos de inconsistência:

- vários eventos que foram unidos em um único evento; e
- um único evento está separado em vários eventos.

### 5.3.3 Terceiro passo: Aperfeiçoamento

Este passo visa reduzir as inconsistências do passo anterior. Nesta etapa, o algoritmo extrai informações temporais externas para aperfeiçoar a qualidade da segmentação.

Para a recuperação dos dados temporais, é requisitado ao Coletor de Dados Temporais que extraia as informações da agenda de compromissos do usuário. Em seguida, há duas possíveis situações para resolver as inconsistências geradas após a segmentação temporal:

- (i) se um dado evento tem fotos que são tiradas durante compromissos diferentes, então este evento será dividido; e
- (ii) se existem dois ou mais eventos consecutivos e eles estão no mesmo compromisso na agenda do usuário, então estes eventos podem ser unificados.

Por exemplo, a Figura 5.3 apresenta um conjunto de fotos que foi classificado como um único evento após o passo de segmentação temporal (Evento A). As datas de captura de algumas dessas fotografias casam com um determinado evento e as datas de captura das outras fotografias casam com outro evento da agenda de compromisso. Após a execução deste passo, o algoritmo irá classificar as fotografias como dois outros eventos (Evento B e Evento C).



Figura 5.3: Primeira situação do passo de aperfeiçoamento

A Figura 5.4 apresenta um exemplo da segunda situação, onde a data de captura das fotografias de dois eventos consecutivos (Evento A e Evento B) casam com a data em que ocorreu um único evento na agenda de compromissos do usuário. Após a execução deste passo, as fotos serão classificadas como sendo de um único evento (Evento C).



Figura 5.4: Segunda situação do passo de aperfeiçoamento

### 5.3.4 Quarto passo: Auto-denominação

O último passo do algoritmo é responsável por atribuir nomes aos eventos obtidos do passo anterior. São utilizadas as informações temporais e espaciais externas para atribuírem nomes aos eventos.

Para obter a informação espacial, o algoritmo a requisita ao módulo Coletor de Dados Espaciais. Este pede por lugares que estão próximos dos grupos gerados no passo 1. O parâmetro de entrada é um *bounding box* de cada grupo e um *buffer* de **N** metros. Nos testes foram utilizados **N = 500 metros**.

A partir das coordenadas geográficas de cada foto, é possível encontrar o nome do lugar, cidade, estado e país em que a foto foi capturada. Esta informação é coletada através do coletor de dados espaciais.

O seguinte formato foi adotado para nomear um evento: “Nome do evento + Nome do Lugar + Cidade + Estado + País + Data Inicial + Duração do Evento”. Da agenda do usuário, obtém-se o nome do evento que casa com a data e hora que a fotografia foi tirada.

Com relação às informações espaciais (lugar, cidade, estado e país), o algoritmo funciona da seguinte maneira:

(i) caso existam fotos de um único lugar, então o evento será anotado com as informações relativas ao lugar, cidade, estado e país; e, caso contrário,

(ii) se existe, em um único evento, fotos de mais de um lugar, o evento será anotado com informações espaciais relativas à cidade, estado e país; e caso contrário

(iii) se existem fotografias de mais de uma cidade, a informação espacial será restrita ao estado e país, e assim sucessivamente.

As informações de data e duração do evento são extraídas do banco de dados. Um exemplo de um evento poderia ser: “IEEE International Symposium on Multimedia (ISM2008) - Double Tree Hotel - Berkeley - Califórnia - EUA - 15/12/2008 - 3 dias”.

## 5.4 Experimentos

A saída do algoritmo de organização automática do PhotoGeo é um conjunto de fotos agrupadas em eventos. Foram realizados experimentos utilizando cinco coleções (A,B,C,D e E) contendo 2.412 fotos georreferenciadas ao todo. Também, foi utilizada a agenda dos donos

das coleções.

A maioria dessas fotos foram capturadas em cidades brasileiras. A coleção A tem 596 fotos que estão distribuídas em aproximadamente um ano e oito meses. Durante esse intervalo de tempo existem meses com zero a três eventos.

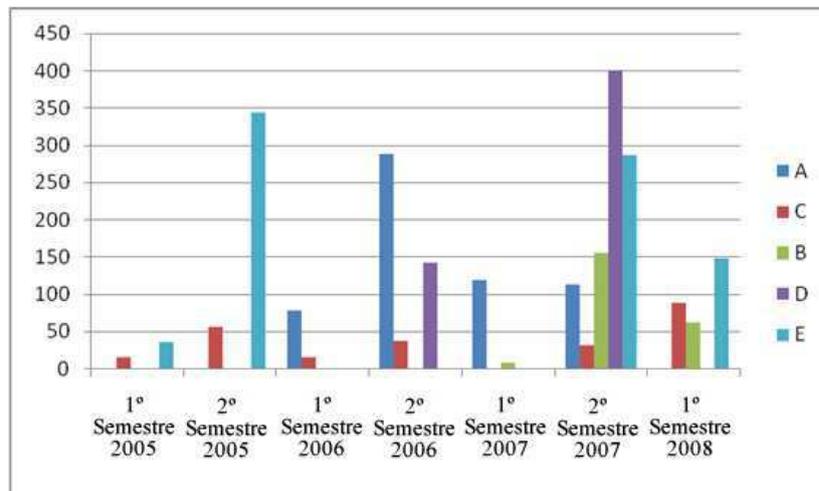


Figura 5.5: Distribuição Temporal das Fotos

As fotografias da coleção A estão distribuídas em um período de aproximadamente um ano e oito meses (veja a Figura 5.5). Dentro deste período há meses que possuem entre zero e três eventos. A maioria das fotos foi capturada em cinco cidades de quatro estados da região Nordeste do Brasil: Campina Grande-PB, João Pessoa-PB, Crato-CE, Farias Brito-CE, Natal-RN, Maceió-AL.

Já a segunda coleção de fotografias está distribuída em um período de aproximadamente onze meses. Ela possui fotografias apenas de quatro eventos. As fotografias foram capturadas em três cidades: Farias Brito-CE, Fortaleza-CE e Campina Grande-PB.

A terceira coleção de fotografias (Coleção C) está distribuída em um período de aproximadamente dois anos e onze meses. Esta possui fotografias relativas às cidades de: Florianópolis-SC, João Pessoa-PB, Campina Grande-PB, Queimadas-PB, Caruaru-PE e Aracaju-SE.

A coleção D está distribuída em um período de aproximadamente um ano. As fotografias desta coleção estão localizadas nos municípios de Ipojuca-PE, Campina Grande-PB, Maceió-AL e Tibau do Sul-RN.

A última coleção (coleção E) possui fotos distribuídas no período de aproximadamente

dois anos e sete meses. Estas fotografias estão distribuídas em Aracaju-SE e nas cidades do Canadá: Toronto, Montreal e Nicarágua.

Nas coleções A e B, quase todas as fotos têm eventos armazenados nas agendas dos usuários. Na coleção D, somente 25% das fotos têm compromissos armazenados na agenda. As coleções C e E não têm eventos descritos nas agendas dos usuários.

Foram utilizadas essas fotos para validar o algoritmo de organização do PhotoGeo. Após a execução dos testes, os usuários avaliaram a efetividade do algoritmo proposto, através da resposta a um questionário com as seguintes questões:

- Questão 1: Quantos eventos foram detectados corretamente?
- Questão 2: Quantos nomes representaram bem os eventos?
- Questão 3: Tente encontrar cinco fotografias distintas de seus eventos. A organização permitiu você encontrá-las em um tempo: ( ) lento ( ) médio ( ) rápido ?
- Questão 4: Qual é a importância que você dá à organização automática das suas fotografias utilizar informações de sua agenda eletrônica de compromissos?  
( ) pequena ( ) média ( ) alta
- Questão 5: No geral, como você definiria a auto-denominação dada a seus eventos?  
( ) péssima ( ) ruim ( ) média ( ) boa ( ) excelente.
- Questão 6: Você teria alguma sugestão para a organização realizada pelo PhotoGeo?

A principal contribuição desta pesquisa é no uso de fontes externas no passo 3 do algoritmo para uma melhor separação das fotos em eventos. Com isto, foram realizados dois tipos de análise: completa e parcial.

A análise completa foi feita através da execução de todos os passos do algoritmo proposto. Foram utilizadas as seguintes métricas:

- a qualidade da classificação - o percentual de eventos que foram separados corretamente;
- a qualidade da auto-denominação dos eventos - o percentual de nomes adequados que foram atribuídos aos eventos; e

- a importância da agenda eletrônica - qual a importância que os usuários atribuíram para o uso das informações da agenda de compromissos.

Por outro lado, a análise parcial foi realizada utilizando apenas os passos 1, 2 e 4. Desta maneira, será analisada a influência do terceiro passo nos resultados do algoritmo. As métricas de **precisão** e **cobertura** são duas métricas clássicas de recuperação da informação que podem ser utilizadas para a avaliação da classificação realizada pelo algoritmo proposto. Em Lacerda et al. [36] foi realizada uma análise utilizando essas métricas. A **precisão** foi definida como o percentual de fotos em um dado evento que foram classificadas corretamente e a **cobertura** foi definida como o percentual entre as fotos que deveriam ter sido classificadas em um determinado evento pelo número de fotos que realmente foram classificadas no evento.

### 5.4.1 Análise completa

De acordo com os resultados obtidos da execução de todo o algoritmo, os donos das coleções A e B consideraram a classificação e a denominação dos eventos excelentes, como podem ser vistos na Figura 5.6. Isto deve-se ao fato de que quase todos os eventos estavam registrados na agenda de compromisso dos donos destas coleções.

Na coleção C, os resultados do PhotoGEO foram avaliados como bom, obtendo 83% de aceitação do dono da coleção. Este mencionou que existia um evento que foi dividido em três outros. Neste evento, em particular, as fotos foram capturadas durante três dias consecutivos. O usuário gostou do uso da agenda de compromissos e mencionou que o sistema PhotoGeo é muito útil para realizar uma primeira classificação (automática) e, em seguida, o usuário poder refiná-la manualmente.

A coleção D teve duas inconsistências durante a separação em eventos. Na primeira, um compromisso havia sido marcado na agenda para um dado dia (ex: uma festa de aniversário), mas este evento passou da meia noite e, além disso, o intervalo de tempo entre a próxima foto após a meia-noite foi maior que a média dos intervalos, assim as fotos foram alocadas em eventos distintos. A segunda inconsistência relatada foi uma viagem do usuário que durou dois dias, esta, que deveria ser um único evento, foi separada em dois eventos distintos. Entretanto, esta viagem não estava anotada na agenda do usuário.

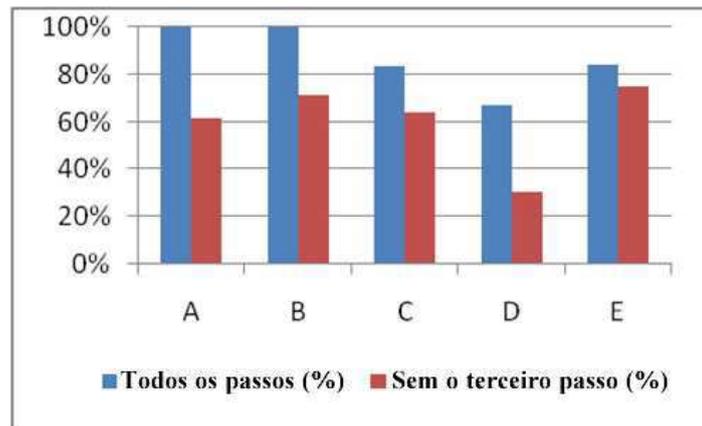


Figura 5.6: Resultados do agrupamento em eventos

Por último, na coleção E, 84% dos eventos foram separados corretamente. Porém, somente 20% dos eventos tiveram uma denominação correta. Isto aconteceu pelo fato de que no banco de dados do PhotoGeo não havia informação sobre as cidades canadenses, com isso, as fotos capturadas naquele país não tiveram a dimensão espacial utilizada propriamente, apenas a temporal. A Figura 5.7 apresenta o resultado do questionário aplicado.

A primeira métrica analisada foi com relação a separação dos eventos. Com base nos resultados da questão 1 (Figura 5.7(a)) houve uma média de 87% de aprovação dos usuários. Outro resultado importante está relatado na Figura 5.7(c), que mostra que a maioria dos usuários teve um acesso rápido às suas fotos.

Com relação à qualidade da auto-denominação dos eventos, foram utilizadas as respostas das questões 2 e 5. A Figura 5.7(b) apresenta o resultado para a questão 2. A coleção E não foi bem avaliada devido ao fato do PhotoGeo não ter dados espaciais do Canadá. Esse problema pode ser resolvido utilizando gazetteers [24] com informações globais. O resultado para a questão 5 (Figura 5.7(d)) variou entre médio e excelente.

### 5.4.2 Análise parcial

O resultado da classificação sem o passo 3 do algoritmo pode ser visualizado na Figura 5.6. Nessa análise foi verificado se a separação está de acordo com os requisitos do usuário.

Note que em todas as coleções o resultado utilizando o algoritmo completo foi melhor. Não utilizando o passo 3, a média de aceitação do usuário caiu de 87% para 60%. Esse resultado pode provar a importância da correção de inconsistências que é feita neste passo.

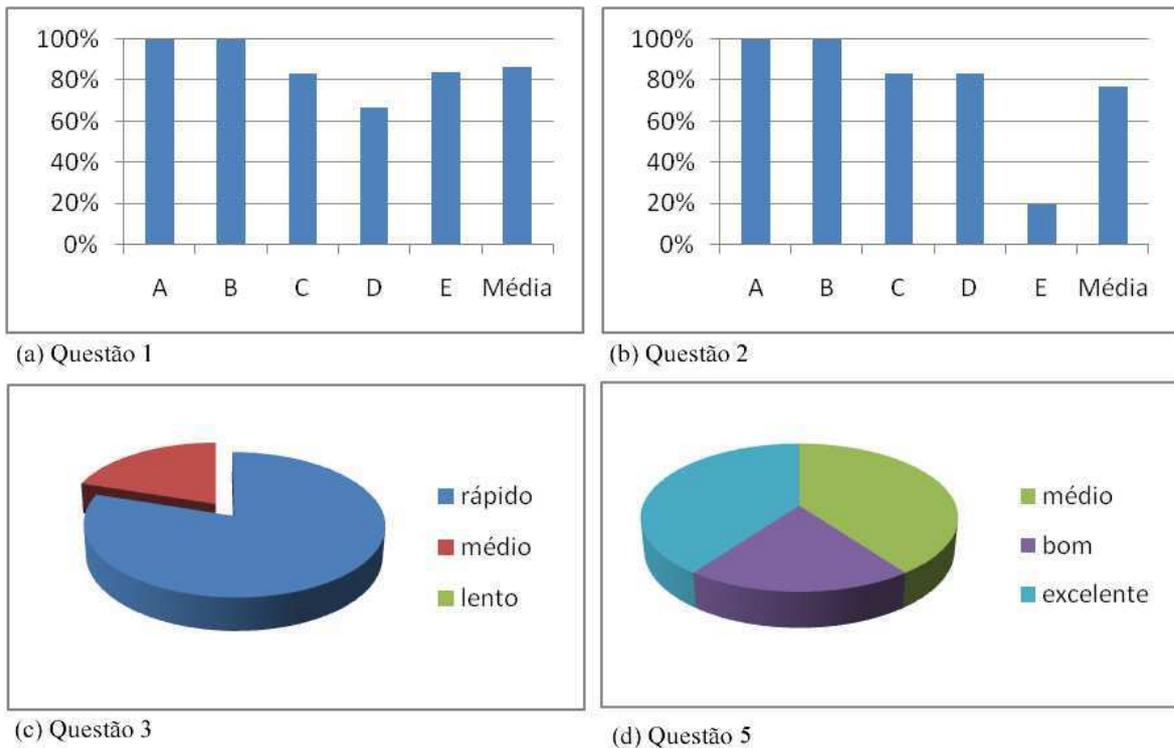


Figura 5.7: Resultados do Questionário

## 5.5 Considerações Finais

Esse capítulo apresentou uma solução para a organização automática de coleções de fotos pessoais georreferenciadas utilizando redes sociais baseadas em mapas e agendas eletrônicas de compromissos. O algoritmo proposto é baseado em quatro passos: agrupamento espacial, segmentação temporal, aperfeiçoamento e auto-denominação. A saída do algoritmo contém um conjunto de fotos separadas em eventos, cada um desses pode ter uma denominação.

Para validar as idéias propostas foi construído um protótipo e utilizado um conjunto de fotos georreferenciadas. Então, foi realizada uma análise com o usuário para verificar a eficácia do protótipo. Os resultados obtidos nessa avaliação foram satisfatórios, e mostram a importância do uso de informações da agenda do usuário e de redes sociais. A auto-denominação teve 77% de aprovação. Por último, o terceiro passo - aperfeiçoamento - comprovou ser bastante útil para melhorar os resultados em aproximadamente 27% na média.

# Capítulo 6

## Sugestão de Anotação de Fotografias

Este capítulo apresenta o mecanismo de anotação semi-automática que foi implementado pelo PhotoGeo. Esta solução, publicada em Lacerda et al. [37], fornece ao usuário um mecanismo que auxilia a anotação das pessoas que estão em uma determinada fotografia, sugerindo nomes de pessoas com maior probabilidade de estarem naquela fotografia. Para isto, utilizam-se informações de contexto obtidas de fontes externas.

A organização deste capítulo está da seguinte forma: a próxima seção apresenta uma formalização sobre o processo de anotação de pessoas em fotografias, em seguida, na seção 6.3 são apresentados os estimadores que são utilizados na solução proposta. Logo após, é apresentado o algoritmo de sugestão de anotações do PhotoGeo. Em seguida, a seção 6.5 apresenta os experimentos realizados para analisar o algoritmo. Por último, são apresentadas as considerações finais do algoritmo de sugestão de anotação.

### 6.1 Anotações de Fotografias

Esta seção apresenta uma formalização do processo de anotação de identidades de pessoas em fotografias, adaptado da proposta de Naaman et al. [44]. Esta adaptação inclui dois novos estimadores. Um estimador desses estimadores é baseado na agenda de compromissos do usuário e o outro é baseado nas informações extraídas de um sistema ciente de contexto. A formalização tem como base dois conjuntos, o conjunto  $\mathbf{F}$  com  $n$  fotografias, que contém vários álbuns de um determinado usuário, e um conjunto  $\mathbf{P}$  com os nomes das pessoas que aparecem nas fotos. Em cada foto  $\mathbf{f}$  aparece um conjunto  $\mathbf{p}$  de pessoas ( $\mathbf{p} \in \mathbf{P}$ ). Assim,

temos ( $\mathbf{p}_f$ ) como o conjunto das pessoas já identificadas na fotografia  $\mathbf{f}$ , e  $\mathbf{K}$  como conjunto de todas as pessoas que já foram identificadas pelo sistema de anotação de fotografias.

Podemos definir formalmente o processo de anotação de pessoas em fotografias como uma seqüência de passos nos quais cada uma das pessoas presentes na fotografia é identificada. Ou seja, em cada passo, a identificação de uma pessoa ( $\mathbf{i} \in \mathbf{p}$ ) é acrescentada ao conjunto  $\mathbf{K}$ .

Em um sistema ideal, dada uma fotografia do conjunto  $\mathbf{F}$  do usuário, um sistema de anotação de pessoas completamente automático geraria um conjunto  $\mathbf{H}_f$  com as pessoas que estão presentes na fotografia. Ou seja,  $\mathbf{H}_f$  seria igual a  $\mathbf{p}_f$ . No entanto, completar esta tarefa de maneira completamente automática é uma atividade muito difícil. Um sistema de sugestão de anotações de fotografias pode gerar um conjunto  $\mathbf{H}_f$  de sugestões de identificações de pessoas presentes na foto. Se  $\mathbf{H}_f \cap \mathbf{p} \neq \emptyset$ , então uma das pessoas em  $\mathbf{H}_f$  está presente na fotografia, portanto, consideramos que há um acerto do sistema de sugestão de anotações. No entanto, se  $\mathbf{H}_f \cap \mathbf{p} = \emptyset$ , consideramos que o sistema errou, pois não apresentou como sugestão nenhuma pessoa presente na fotografia. Prevê-se que a cada interação uma nova pessoa possa ser incluída em  $\mathbf{K}$ , o que pode ser feito do conjunto  $\mathbf{H}_f$  ou diretamente de  $\mathbf{P}$ . Isto é feito para cada uma das fotos  $\mathbf{f}$ , tal que  $\mathbf{f} \in \mathbf{F}$ .

Verifica-se, pois, que o problema central de um sistema de sugestão de anotações de fotografias, é a definição do conjunto  $\mathbf{H}_s$  com as sugestões de identificações de pessoas para anotação.

Para gerar o conjunto  $\mathbf{H}_s$ , o sistema de anotação deve usar vários estimadores para avaliar a probabilidade que uma determinada pessoa em  $\mathbf{K}$  tenha de estar em  $\mathbf{f}$ . Os candidatos com maior probabilidade são incluídos no conjunto  $\mathbf{H}_s$  que é filtrado pela proximidade espacial e apresentado ao usuário.

Em seguida, apresentamos os estimadores que são utilizados neste trabalho e qual o algoritmo proposto para a geração do conjunto  $\mathbf{H}_s$ .

## 6.2 Estimadores para o Ranking

É possível reconhecer padrões que representam o aparecimento de pessoas nas fotografias. Por exemplo, na coleção de fotografias de um determinado usuário, há pessoas que apare-

cerão com mais frequência, tais como: o próprio dono da coleção, a sua família e os seus amigos mais próximos. Este é um exemplo de um padrão de recorrência, ou seja, há pessoas com probabilidade muito maior que outras de aparecer nas fotografias.

Uma forma de sugerir anotações para uma fotografia é com base na probabilidade de uma determinada pessoa estar naquela fotografia. Para estimar esta probabilidade são analisados, neste trabalho, os seguintes estimadores baseados nos propostos por Naaman et al. [44]: recorrência global, recorrência espacial, recorrência temporal, recorrência no evento e co-ocorrência.

**Recorrência Global (r):** é o percentual de vezes que uma determinada pessoa foi anotada nas fotografias. Se uma determinada pessoa já foi anotada várias vezes, ou seja, teve uma recorrência grande, então é grande a probabilidade dela estar em outra foto. Podemos definir a recorrência global de uma pessoa  $i \in \mathbf{K}$  utilizando a equação (1). Para o cálculo da recorrência global,  $\mathbf{X}$  é o conjunto com as  $n_x$  fotografias do conjunto  $\mathbf{F}$ .

$$r_i = \frac{\sum_{f \in \mathbf{X}} g_f(i)}{n_x}, \text{ com } g_f(i) = \begin{cases} 1 & \text{se } i \in p_f \\ 0 & \text{de outra maneira} \end{cases} \quad (1)$$

**Recorrência Espacial (re):** é o percentual de vezes que uma determinada pessoa aparece nas fotografias contidas em uma determinada região. O cálculo deste estimador para uma pessoa  $i \in \mathbf{K}$  pode ser realizado de acordo com a equação (1), onde  $\mathbf{X}$  é o conjunto com as  $n_x$  fotografias do conjunto  $C \subseteq F$ , tal que  $\mathbf{C}$  são as fotografias que foram tiradas em uma determinada cidade.

**Recorrência Temporal (rt):** é o percentual de vezes que uma determinada pessoa aparece nas fotografias que estão contidas num intervalo de tempo  $t$  do instante de captura da fotografia  $f$ . Para calcular o valor para uma pessoa  $i \in \mathbf{K}$  também utiliza-se a equação (1), no qual  $\mathbf{X}$  é o conjunto contendo as  $n_x$  fotos do conjunto  $T \subseteq F$ , tal que  $\mathbf{T}$  são aquelas que estão no intervalo de tempo  $t$  da fotografia  $f$ .

**Recorrência no evento (rev):** é o percentual de vezes que uma determinada pessoa aparece nas fotografias de um determinado evento. Também é utilizada a equação (1) para calcular o valor para uma pessoa  $i \in \mathbf{K}$ , no qual  $\mathbf{X}$  é o conjunto contendo as  $n_x$  fotografias do conjunto  $E \subseteq F$ , tal que  $\mathbf{E}$  são aquelas que estão no mesmo evento da fotografia  $f$ .

**Co-Ocorrência (co):** é o percentual de vezes que duas pessoas apareceram juntas em uma mesma fotografia. Por exemplo: João é namorado de Maria e quase sempre os dois estão juntos nas fotografias, sendo assim, se João aparece em outra fotografia, então a probabilidade de Maria também estar nela é muito grande. Pode-se definir a co-ocorrência da pessoa  $i_1$  com a pessoa  $i_2$  como o número de ocorrências de ambas na mesma foto, dividido pelo número de fotografias em que somente a pessoa  $i_1$  aparece. Caso haja mais de uma pessoa anotada em uma fotografia, será considerado apenas o maior valor de co-ocorrência. Assim, este estimador pode ser calculado por:

$$co_{i_1} = \max_{i_2 \in p_f} \left( \frac{\sum_{f \in F} h_f(i_1, i_2)}{\sum_{f \in F} g_f(i_1)} \right), \text{ com} \quad (2)$$

$$h_f(i_1, i_2) = \begin{cases} 1 & \text{se } i_1, i_2 \in p_f \\ 0 & \text{de outra maneira} \end{cases} \quad \text{e} \quad g_f(i) = \begin{cases} 1 & \text{se } i \in p_f \\ 0 & \text{de outra maneira} \end{cases}$$

Os estimadores podem ser utilizados tanto para descartar determinadas pessoas da sugestão como para ponderar. Por exemplo, geralmente duas pessoas tiram fotografias juntas, mas no momento de uma determinada foto é verificado que elas não estavam próximas, então o sistema deve descartar esta pessoa.

Adicionalmente à proposta de Naaman et. al, este capítulo propõe a utilização do contexto espacial e social do usuário utilizando os estimadores: proximidade espacial e presente no mesmo compromisso.

O estimador **proximidade espacial (pe)** verifica a proximidade espacial entre a câmera fotográfica no momento da captura da fotografia e a posição das pessoas cadastradas no sistema de anotação. Por exemplo, ao registrar uma fotografia; havia dois amigos bem próximos da câmera e outros dois há centenas de metros de distância. Com isto, existe a possibilidade dos dois primeiros amigos estarem na fotografia, entretanto os outros dois certamente não estarão.

A proximidade espacial é utilizada da seguinte maneira. Considere a lista  $\mathbf{H}_s'$  obtida através do uso dos estimadores já apresentados para anotação da fotografia  $\mathbf{f}$  adquirida no ponto de coordenadas  $(\mathbf{lat}_f, \mathbf{long}_f)$ , então a lista  $\mathbf{H}_s$  é formada por todas as pessoas  $i \in \mathbf{H}_s'$ ,

tal que a distância da posição da pessoa  $\mathbf{i}$  para o local da fotografia seja menor que um valor  $\mathbf{R}$ .

Já o estimador **presente no mesmo compromisso (pmc)** verifica se uma ou mais pessoas estão presentes em um determinado compromisso no momento da captura de uma foto. Por exemplo, o usuário pode ter inserido na sua agenda de compromissos um evento no qual estará presente e quais seus amigos que também estarão. Caso o usuário capture fotografias no horário do evento, então o sistema realizará um casamento (*matching*) entre a data das fotografias e o compromisso na agenda do usuário, então descobrirá quais amigos estavam presentes naquele compromisso. Este estimador é computado para pessoa  $\mathbf{i} \in \mathbf{K}$  avaliando a lista de pessoas convidadas para o mesmo evento ( $\mathbf{p}_c$ ), de acordo com a equação (3).

$$pmc_i = \begin{cases} 1 & \text{se } i \in p_c \\ 0 & \text{se } i \notin p_c \end{cases} \quad (3)$$

O cálculo do ranking para cada pessoa  $\mathbf{i} \in \mathbf{K}$  para uma fotografia  $\mathbf{f}$  é realizado através da equação (4).

$$R_{i,f} = w_r \times r_i + w_{re} \times re_i + w_{rt} \times rt_i + w_{rev} \times rev_i + w_{co} \times co_i + w_{pmc} \times pmc_i \quad (4)$$

Onde  $w_r$ ,  $w_{re}$ ,  $w_{rt}$ ,  $w_{rev}$ ,  $w_{co}$  e  $w_{pmc}$  são os pesos de cada estimador que podem ser ajustados para atribuir uma maior ou menor importância a cada um destes.

### 6.3 Algoritmo de Sugestão

Para que o algoritmo funcione de acordo com o desejado para a nossa solução, é importante que sejam levadas em consideração algumas hipóteses:

(i) O sistema ter o conhecimento da posição geográfica de alguns contatos do usuário no momento das fotografias; e

(ii) O usuário utiliza uma agenda de compromissos digital, acessível ao PhotoGeo.

O PhotoGeo deve fornecer para o usuário uma pequena lista das pessoas mais prováveis de estarem em uma determinada fotografia. Esta lista deve estar em ordem decrescente das

peças mais prováveis de estarem na fotografia a ser anotada. O algoritmo para sugestão de anotação está descrito no Código Fonte 6.1 e é composto dos seguintes passos:

**1º Passo (Inscrição):**

No primeiro passo do algoritmo, todos os contatos da rede social do usuário serão adicionados à lista, inclusive o próprio usuário.

**2º Passo (Eliminação):**

O segundo passo do algoritmo é para eliminar as pessoas que possivelmente não estarão na fotografia. Por exemplo, se um contato do usuário, ou ele próprio, no momento da captura da fotografia estiver a uma distância muito grande da câmera, possivelmente ele não estará na fotografia.

Então, nesta fase do algoritmo será utilizado o estimador **pe** para eliminar da lista todos os contatos que no momento da captura estiverem a uma distância maior que um determinado limite **R**. Para o estudo de caso foi utilizado **R = 30 metros**.

Uma dificuldade para o cálculo da distância entre os usuários do sistema é devido à precisão dos atuais equipamentos de GPS. Dependendo do local haverá uma taxa de erro, que varia de acordo com o número de satélites e a qualidade de sinal no lugar desejado. Há alguns equipamentos que utilizam técnicas para melhorar esta precisão, por exemplo, utilizando a triangulação de antenas de celulares.

Note que, se não houver informações sobre a localização de um determinado contato do usuário, este permanecerá na lista de pessoas prováveis de estar na fotografia. Desta forma, mesmo que o usuário não esteja com o celular ligado no momento de captura da fotografia, ou que, por algum motivo, o sistema não conseguiu capturar informações sobre a localização do mesmo, este contato ainda sim poderá ser sugerido pelo algoritmo.

**3º Passo (Classificação):**

O terceiro passo do algoritmo é responsável por classificar os contatos que não foram eliminados no segundo passo. Para isto, serão utilizados os estimadores **r**, **rt**, **re**, **rev**, **co** e **pmc**, descritos na seção anterior. A pontuação de cada contato será calculada de acordo com a equação (4). Cada estimador deverá retornar um valor no intervalo de 0 e 1 que será multiplicado pelo peso que aquele estimador possui.

---

Código Fonte 6.1: Sugestão de Anotações

---

1 Lista\_sugestoes(Foto p, Usuario usuario, Lista contatos)

```
2 inicio
3     // Primeiro passo: inscrição
4     Lista sugestoes = contatos;
5     // Segundo passo: Eliminação
6     while (sugestoes.hasNext()) {
7         Usuario u = sugestoes.next();
8         if (distancia(p.getPosicao(), u.getPosicao(p.getDataHora())) > R) then
9             sugestoes.remove(u)
10        }
11    // Terceiro passo: Classificação
12    While (sugestoes.hasNext()) {
13        User u = sugestoes.next();
14        r = calcularRg(u);
15        re = calcularRe(u, p.getPosicao());
16        rt = calcularRt(u, p.getDataHora());
17        rev = calcularRev(u, p.getEvento());
18        co = calcularCo(u, p.getAnotados());
19        pmc = calcularPmc(u, usuario.getCompromissos());
20        u.rankingGlobal = calcularRanking(r, re, rt, rev, co, pmc);
21    }
22    // Ordenar pelos valores dos estimadores
23    ordenar(sugestoes);
24    return sugestoes;
25 end;
```

---

## 6.4 Informações de contexto dos usuários

O PhotoGeo obtém informações sobre o contexto dos usuários através do acesso ao sistema externo VadeMecum. O modelo de contexto utilizado está descrito por Almeida et al. [13]. A integração do PhotoGeo com o VadeMecum é feita através do Multimedia Service. Este é um serviço responsável por armazenar e fornecer as informações de contexto do usuário e da fotografia e a fotografia propriamente dita.

As informações relativas ao contexto do usuário são aquelas do ambiente em que o usuário está inserido e suas relações com o meio. Por exemplo: quem são seus amigos, quais são os seus compromissos, sua localização geográfica, entre outras.

Já as relativas ao contexto da fotografia são informações acerca do ambiente no momento que a fotografia foi capturada. Por exemplo: localização geográfica, data e hora da captura, fabricante e modelo da câmera, pessoas que estão próximas no momento da captura da fotografia, entre outras.

O ciclo para a anotação da fotografia é composto por três passos: captura, armazenamento e anotação. Estes são descritos a seguir.

O primeiro passo é a captura da fotografia. O usuário através de seu dispositivo móvel equipado com câmera fotográfica e um receptor de GPS realiza a captura da fotografia. O usuário deverá capturar a fotografia utilizando o sistema VadeMecum. Este capturará tanto a fotografia como informações relativas ao contexto do usuário. Estas informações são: os contatos que estão próximos do usuário no momento da captura, a posição geográfica e a data e hora da captura.

No segundo passo, o VadeMecum enviará automaticamente a fotografia para o Multimedia Service. Este oferece uma interface para que possa armazenar a fotografia e as informações de contexto do usuário e da fotografia. Outra responsabilidade neste passo é fornecer essas informações para o PhotoGeo. A comunicação com o Multimedia Service dá-se utilizando o Representational State Transfer (REST). O REST possibilita a comunicação de maneira simples utilizando os métodos padrões do HTTP.

O Multimedia Service é um serviço responsável por gerenciar os arquivos multimídia e suas informações contextuais no sistema VadeMecum. Este serviço foi desenvolvido em Java e pode estar localizado em um servidor Web e acessado remotamente pelo PhotoGeo. O mesmo é responsável por oferecer interfaces para recuperação, consulta e armazenamento de informações multimídia.

O terceiro passo é responsável por realizar a anotação em si. O PhotoGeo requisita as informações ao Multimedia Service e as armazena em seu banco de dados local. Quando o usuário abrir uma fotografia, o sistema automaticamente apresenta a interface com as sugestões de anotações de pessoas.

## 6.5 Análises e Resultados

Para a validação do algoritmo proposto, foi utilizada uma coleção de fotografias pessoais, vide Tabela 6.1, com um total de 1505 fotografias. Todas as fotografias possuem metadados temporais e espaciais. Os metadados espaciais foram obtidos de forma manual, georreferenciando as fotografias a partir de softwares específicos. O dono da coleção escolheu 10 contatos para que fossem simuladas as suas localizações no momento da captura de cada fotografia.

Nesta seção serão analisados os seguintes aspectos: Qual a importância do estimador **pe** (proximidade espacial)? Qual a importância do estimador **pmc** (presente no mesmo compromisso)?

Figura 6.1: Estatísticas das coleções avaliadas

	<b>Coleção</b>
<b>Total de Fotos</b>	1505
<b>Fotos com anotações</b>	1097
<b>Pessoas distintas nas fotos</b>	206
<b>Total de anotações</b>	2416
<b>Média de pessoas por fotos</b>	2.2
<b>Média de fotos por indivíduos</b>	11.73

Para realizar a validação foi desenvolvido um simulador dentro do protótipo do PhotoGeo, vide Figura 6.2. Este é responsável por armazenar qual era a posição geográfica dos contatos no momento da captura da fotografia. O simulador funciona da seguinte forma: ao ser carregada uma fotografia, os contatos que foram escolhidos para serem monitorados serão exibidos em um mapa, no canto direito da figura. Após posicioná-los de acordo com o desejado, essas informações serão armazenadas no banco de dados do PhotoGeo. No exemplo exibido na Figura 6.2, há contatos localizados na cidade de Fortaleza - CE e outros na cidade de Campina Grande - PB.

Para realizar a análise, inicialmente o dono da coleção registrou a posição de seus contatos (que foram selecionados para serem monitorados) e anotou quais as pessoas que estão presentes em cada fotografia. Uma vez que estas informações estão armazenadas no banco de dados, as análises poderão ser realizadas de forma automatizada.

Para todos os valores de  $n$ , com exceção da primeira coluna ( $n = 1$ ), o algoritmo teve um resultado melhor utilizando o estimador **pe**. No geral, houve uma melhora média de 0,68% quando comparado à sugestão que utiliza apenas os estimadores propostos em trabalhos anteriores na literatura. Acredita-se que este valor baixo deva-se ao fato de terem sido registrados a posição geográfica de poucos contatos dos usuários. Para a coleção utilizada, somente em torno de 5% dos contatos possuem as localizações geográficas armazenadas no sistema.

A primeira análise foi para avaliar a influência do estimador **pe**, utilizado no segundo passo do algoritmo. Para isto, este foi executado duas vezes: uma com todos os passos do algoritmo e outra sem o segundo passo. Esta análise foi realizada levando em consideração a taxa de acerto para os  $n$  primeiros resultados do ranking; tal que  $n$  é um número inteiro no intervalo entre 1 e 10, vide Figura 6.3. A taxa de acerto é medida como o número de pessoas presentes na fotografia, dividido pela quantidade de pessoas que foram exibidas entre as  $n$  primeiras posições do ranking. Para as avaliações realizadas nesta seção, todos os estimadores receberam um peso igual com valor igual a 1.

Para que o algoritmo possa alcançar uma melhor precisão na análise anterior seria necessário ter capturado as localizações de uma maior porcentagem de contatos. Desta maneira, foi realizada uma análise utilizando as anotações apenas para os contatos mais recorrentes. Foi realizado um teste utilizando apenas os conjuntos de 50, 40, 30, 20 e 10 contatos mais recorrentes e foram mantidas as mesmas dez pessoas com localizações simuladas. Esta análise foi realizada levando em consideração a taxa de acerto dos cinco primeiros resultados do ranking.

Nesta análise, vide Figura 6.4, houve um ganho médio de 1,86% comparado aos estimadores propostos na pesquisa de Naaman et al. [44]. Analisando separadamente os estimadores, o estimador **pmc** teve um ganho superior ao **pe** por conta que todas as pessoas daquele evento foram informadas na agenda de compromissos. Já no caso do **pe**, a proporção de pessoas que tiveram as localizações pelo total de pessoas foram de: 20%, 25%, 33,33%, 50% e 100%, respectivamente.

**PhotoGEO** 2007 A new concept on digital photo organizer

ALBUM CONTACTS SEARCH

[Voltar](#)

Casamento de Marco - La Maison Dunas Buffet - Fortaleza - 2006-04-29 21:36:13.0 (03:56:46)

Paisagens 132.jpg



**Sugestões:**

- [Yuri Lacerda](#)
- [Thially Goncalves](#)
- [Melina Lacerda](#)
- [Marcos Antonio](#)
- [Aureliano](#)

Name:

Email:

Is Monitored?

**Localização Geográfica dos Contatos:**



Description: OLYMPUS DIGITAL CAMERA  
 Date/Time: 29/04/06 22:32:33.000  
 Width x Height: 320 x 240  
 Camera Manufacturer: OLYMPUS IMAGING CORP.  
 Camera Model: C180,D435

Figura 6.2: Simulador embutido no protótipo do PhotoGeo

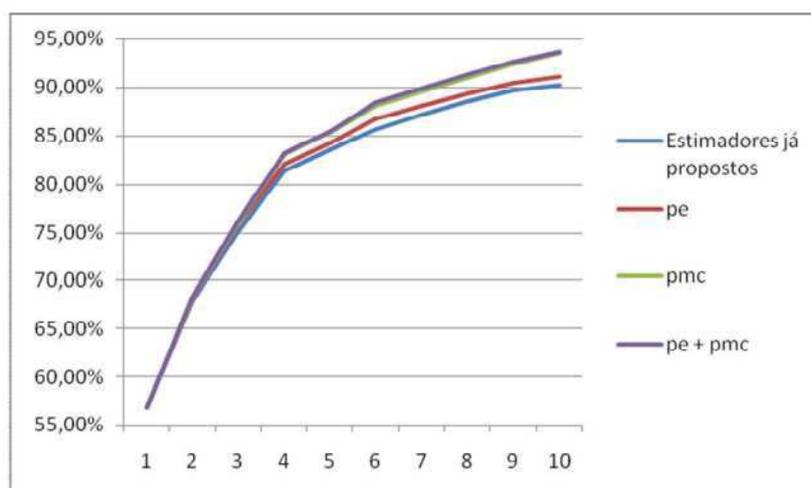


Figura 6.3: Análise da influência do segundo passo

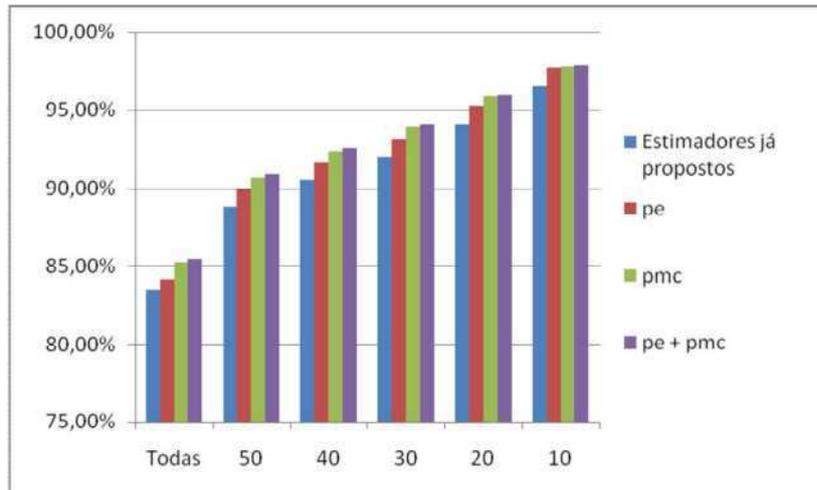


Figura 6.4: Análise considerando os contatos mais importantes

## 6.6 Considerações Finais

Este capítulo apresentou um mecanismo de anotação semi-automático de pessoas implementada pelo PhotoGeo. Esta solução utiliza estimadores baseados em recorrência e no contexto dos usuários. Desta maneira, foram propostos dois novos estimadores para a classificação das: presentes no mesmo compromisso e proximidade espacial.

A partir das análises realizadas observou-se que os estimadores propostos neste capítulo auxiliaram na sugestão de pessoas, melhorando os resultados para a coleção utilizada para os testes. Para que se possam obter resultados mais precisos é ideal que se façam testes utilizando maiores massas de dados.

# Capítulo 7

## Consultas Avançadas

O uso de uma interface moderna e amigável é um requisito fundamental para o sucesso de uma biblioteca digital multimídia. Entretanto, a maioria das bibliotecas digitais de fotografias oferecem um mecanismo de consulta bastante limitado, o que faz com que o usuário não possa realizar buscas mais avançadas. Há estudos que mostram que a localização geográfica é uma das características mais marcantes para os usuários recordarem de uma foto [2]. As consultas espaciais nos organizadores de fotografias têm sido pouco exploradas. Desta maneira, se um usuário desejar realizar consultas mais específicas, como por exemplo a consulta: “Recupere as fotografias que estão em uma distância de até 50km da cidade de Campina Grande nos anos de 2003 e 2004”, a maioria dos sistemas existentes não permite realizar este tipo de consulta. Este capítulo propõe um módulo de consultas avançadas para uma biblioteca digital de fotografias georeferenciadas com três tipos de filtros: espacial, temporal e convencional.

### 7.1 Consultas Avançadas

Como já foi citado anteriormente, o módulo de consultas é uma das principais funcionalidades de um organizador de fotografias. Este deve permitir ao usuário que localize facilmente as fotografias desejadas. Os seguintes requisitos devem ser obedecidos:

(i) O sistema deverá permitir ao usuário montar consultas flexíveis utilizando filtros de acordo com o atributo desejado;

(ii) A interface do sistema deverá ser simples e intuitiva, visto que o sistema deverá ser

utilizado por usuários que não são especialistas em informática.

A interface de consultas do PhotoGeo permite que se realizem filtragens em três domínios: convencional, espacial e temporal. A Figura 7.1 apresenta a interface de consultas proposta. No lado esquerdo há os campos que podem ser exibidos no resultado da consulta, como: descrição da fotografia, data e hora da captura, nome do arquivo, tamanho, entre outros. Já no lado direito há a possibilidade de criar um novo filtro para a consulta e também exibe os filtros que já foram criados para a consulta. Há três tipos de filtros que podem ser criados, abrangendo todos os domínios possíveis de filtragem: convencional, espacial e temporal. Estes são detalhados a seguir.

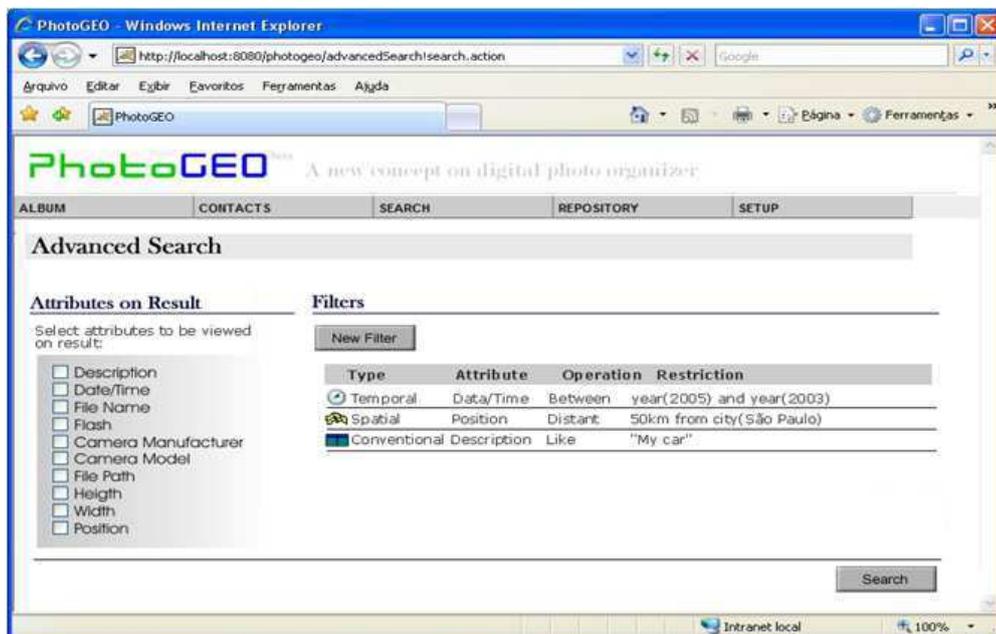


Figura 7.1: Interface de Consulta Avançada

O resultado de uma consulta será exibido como na Figura 7.2. Ao lado esquerdo é exibido um mapa com as fotografias exibidas. No lado direito, há a lista com os resultados. Esta lista é composta dos campos que foram selecionados para serem exibidos antes de executar a consulta.

### 7.1.1 Filtro Convencional

O filtro convencional serve para restringir fotografias a partir de campos de cadeias de caracteres e numéricos. Os seguintes campos podem ser filtrados utilizando este filtro: descrição,

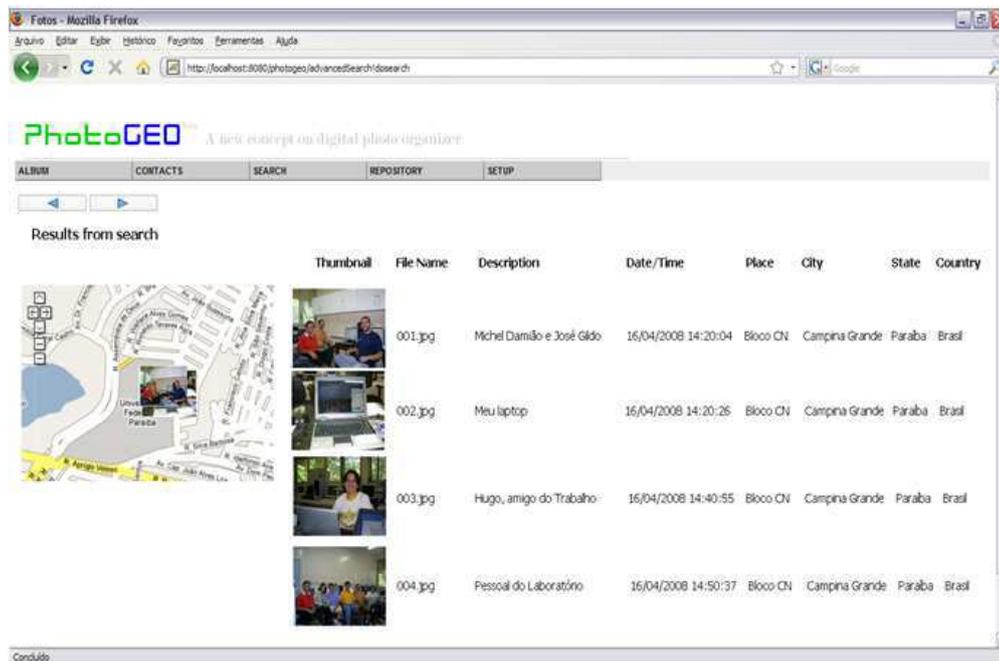


Figura 7.2: Resultado de uma consulta

nome do arquivo, flash, fabricante da câmera, modelo da câmera, diretório, largura e altura.

Para os campos que são cadeias de caracteres (descrição, nome do arquivo, fabricante, modelo e diretório) é permitido utilizar os seguintes operadores: como (like), igual e diferente. Já para campos numéricos (largura e altura) é permitido utilizar os campos ( $=$ ,  $\neq$ ,  $<$ ,  $\leq$  e  $\geq$ ). Na tela do canto inferior direito da Figura 7.3 há o exemplo da criação de um filtro para o campo **largura** utilizando o operador " $<$ ".

### 7.1.2 Filtro Espacial

O filtro espacial serve para ser aplicado à localização geográfica das fotografias, visto que este é o único campo espacial dos metadados das fotografias. Dentre os vários operadores espaciais existentes apenas os operadores **contida** e **distante** fazem maior sentido de serem utilizados. A tela de criação do filtro espacial está exibida no lado esquerdo da Figura 7.3.

Caso seja selecionado o operador **contida** o usuário deverá filtrar a feição que será utilizada na consulta. Por exemplo, se o usuário desejar criar um filtro para as fotografias que estão contidas na cidade do Rio de Janeiro, então o usuário deverá adicionar um filtro da seguinte forma: **município** como **cidade**; **onde** como **nome**; o **operador like** e o **argumento** como "Rio de Janeiro". Em seguida o usuário deverá salvar o filtro.

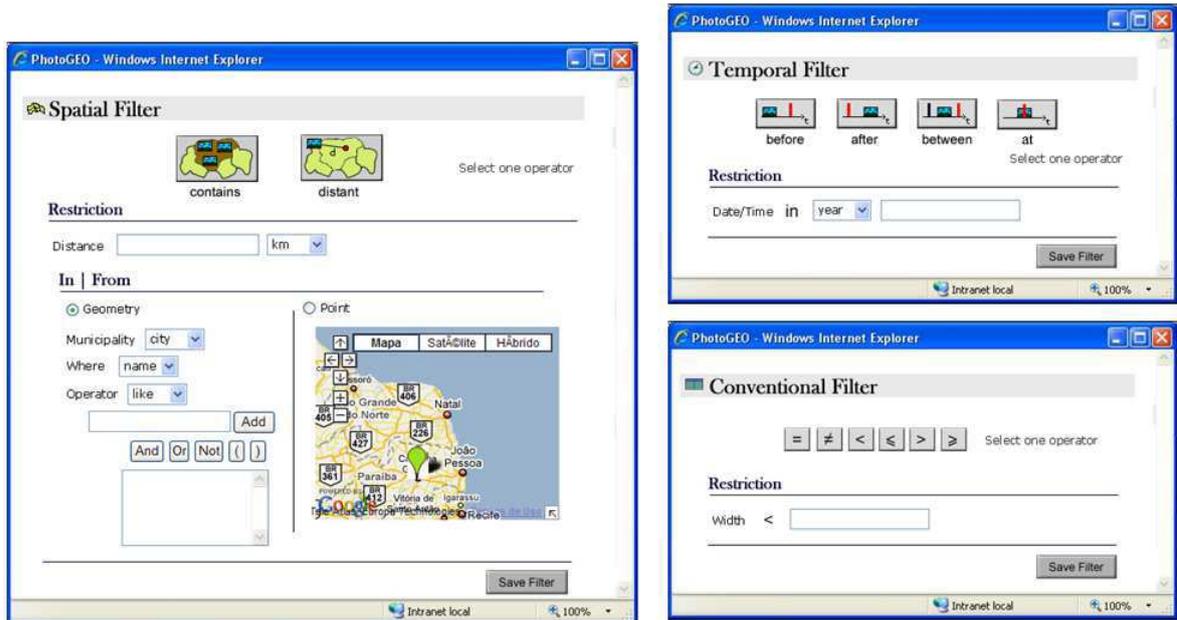


Figura 7.3: Filtros da Consulta Avançada

Para o operador **distante** o usuário deverá selecionar um ponto ou uma feição para ser utilizada no filtro. Por exemplo, se o usuário desejar filtrar as fotografias que estão distantes até 100 km de João Pessoa após selecionar o operador **distante** o usuário tem duas opções: a primeira é selecionar a opção **ponto** em seguida clicar na cidade de João Pessoa através do mapa eletrônico; a segunda é selecionar a opção **sub-filtro** e criar um sub-filtro para a cidade cujo nome é “João Pessoa”.

### 7.1.3 Filtro Temporal

O filtro temporal é aplicado no campo Data/Hora de captura da fotografia. Existem quatro operadores que podem ser aplicados a este campo:

- **antes**: aplicado quando deseja filtrar as fotografias que estão antes de uma determinada data;
- **depois**: aplicado para filtrar as fotografias que estão depois de uma determinada data;
- **entre**: aplicado para filtrar as fotografias que estão entre duas datas;
- **durante**: aplicado para filtrar as fotografias que estão em uma determinada data.

A tela de criação do filtro temporal pode ser vista na Figura 7.3 no canto superior direito.

## **7.2 Considerações Finais**

Os mecanismos de consultas são muito importantes para as bibliotecas digitais de fotografias. Não foi encontrada na literatura nenhuma biblioteca digital de fotografias georeferenciadas que ofereça mecanismos de buscas avançados para os usuários. Neste trabalho, foi proposta a criação de um módulo de consultas avançadas para uma biblioteca digital de fotografias georeferenciadas.

# Capítulo 8

## Conclusão

Neste trabalho, foram expostos os principais conceitos de bibliotecas digitais de fotografias georreferenciadas, descrevendo as principais características dessas ferramentas. Desta forma, foram apresentados os principais conceitos de Sistemas de Informação Geográfica e Bibliotecas Digitais Geográficas. Todos estes temas são necessários para a compreensão do tema. Apresentou-se ainda o estado da arte com relação aos seguintes aspectos abordados neste trabalho: navegação, organização automática, anotação e consultas em coleções de fotografias. No estado da arte foram apresentados os trabalhos desenvolvidos na área e as principais contribuições destes.

Foram apresentadas novas soluções para os problemas dos aspectos explorados neste trabalho. Também foram discutidos aspectos arquiteturais do protótipo apresentado. Este protótipo foi construído para validar as soluções propostas através dos experimentos que foram realizados.

### 8.1 Contribuições

Foi proposta uma nova biblioteca digital de fotografias georreferenciadas na Web para permitir organizar e consultar fotografias buscando simplicidade e eficiência. No PhotoGeo foram exploradas algumas das principais características que deve ter um sistema deste tipo.

Uma destas características é com relação a interface para navegação na coleção de fotografias. É proposto um método de navegação que permite ao usuário navegar através de uma hierarquia espaço-temporal ou por um mapa eletrônico.

Outra contribuição deste trabalho é no âmbito da organização automática de fotografias. O trabalho propôs um novo algoritmo para a organização automática que utiliza informações espaço-temporais de fontes externas para aperfeiçoar a detecção automática de eventos.

Com relação à anotação de fotografias, o sistema propôs um novo mecanismo semi-automático de anotação de pessoas em fotografias que utiliza um algoritmo de sugestão de anotações. Este algoritmo baseia-se em estimadores que calculam a probabilidade de uma pessoa estar em uma dada fotografia. A principal contribuição da solução é o uso de informações de contexto dos usuários coletadas a partir de fontes externas ao sistema.

Outro diferencial proposto neste sistema é o módulo de consultas avançadas de fotografias. Este módulo permite que o usuário realize consultas específicas através de uma interface simples e atraente. O sistema contempla tanto os dados do domínio numérico e textual, como também espaço-temporal.

## 8.2 Trabalhos Futuros

Há diversas possibilidades para dar continuidade a este trabalho. Com relação aos aspectos que foram explorados no PhotoGeo as seguintes sugestões são feitas:

- **Organização Automática:** Há diversas possibilidades de continuidade neste módulo. A primeira delas seria analisar o impacto do uso de outros algoritmos de mineração no primeiro passo do algoritmo. A segmentação temporal é bem limitada, pode-se estudar como melhorar esta etapa do algoritmo. Seria muito importante utilizar gazzeters para coletar mais informações espaciais para melhorar o processo de auto-denominação. Outra sugestão é utilizar uma coleção de dados maior para os experimentos.
- **Sugestão de Anotação:** Muitas pessoas gostam de colecionar fotografias de animais e objetos. Uma proposta de trabalho futuro seria analisar uma solução para a anotação de objetos e animais presentes nas fotografias. Outra sugestão seria utilizar informações de redes sociais para o algoritmo de sugestão de anotação de pessoas. Por exemplo, comparar o aparecimento das pessoas nas fotografias com as informações das redes sociais. Outra possibilidade seria a criação de um estimador baseado em algoritmos de reconhecimento facial. Outra tarefa importante é de ter uma forma de exportar as

anotações do banco de dados para dentro dos metadados das fotografias. Também é importante que se utilizem maiores coleções de dados para os experimentos. Deve-se averiguar também outras informações de contexto que podem ser obtidas de fontes externas para ampliar as informações de contexto da fotografia, como por exemplo, a temperatura no momento da foto.

- **Módulo de Consultas:** Um trabalho futuro possível para este módulo é de realizar testes de usabilidade para averiguar as necessidades e o uso dos usuários para a ferramenta. Outra possibilidade é averiguar maiores possibilidades para as consultas, por exemplo, filtrar as pessoas que estavam presentes nas fotografias.

Como trabalho futuro, também há a possibilidade de agregar novas funcionalidades ao PhotoGeo, tais como: suporte de outras mídias (áudio e vídeos), interface para dispositivos móveis e repositório Web de fotografias.

Alguns desses trabalhos já estão sendo desenvolvidos em paralelo, como é o caso do impacto do uso de outros algoritmos de mineração e de gazettters na organização automática do PhotoGeo. Também está sendo pesquisado sobre o comportamento do algoritmo de organização automática mediante a variação dos valores de entrada do passo de agrupamento espacial e de segmentação temporal.

## Referências Bibliográficas

- [1] S. Ahern, M. Davis, D. Eckles, S. King, M. Naaman, R. Nair, M. Spasojevic, and J.H.I. Yang. Zonetag: Designing context-aware mobile media capture to increase participation. In *Proc. Pervasive Image Capture and Sharing: New Social Practices and Implications for Technology Workshop*, 2006.
- [2] James F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.
- [3] Cláudio de Souza Baptista. StepLib: A Digital Library for Spatio-Temporal and Multimedia Data, 2002. PhD Thesis. The University of Kent.
- [4] Tor Bernhardsen. *Geographic Information Systems: An Introduction*. John Wiley and Sons, 2002.
- [5] Henk M. Blanken, Arjen P. de Vries, Henk Ernst Blok, and Ling Feng. *Multimedia Retrieval (Data-Centric Systems and Applications)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.
- [6] D. Blasby. Building a Spatial Database in PostgreSQL. *Refractions Research*. Retrieved March, 20, 2006.
- [7] D.M. Boyd and N.B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2008.
- [8] C. Chen, M. Oakes, and J. Tait. Browsing Personal Images Using Episodic Memory (Time+ Location). *Lecture Notes in Computer Science*, 3936:362, 2006.

- [9] M. Cooper, J. Foote, A. Girgensohn, and L. Wilcox. Temporal event clustering for digital photo collections. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 1(3):269–288, 2005.
- [10] J. CP. Exchangeable image file format for digital still cameras: Exif Version 2.2. Technical report, 2002. Disponível em: <http://www.exif.org/Exif2-2.PDF> (Acessado em 20 de Março de 2009).
- [11] M. Davis, S. King, N. Good, and R. Sarvas. From context to content: leveraging context to infer media metadata. In *Proceedings of the 12th annual ACM international conference on Multimedia*, pages 188–195. ACM New York, NY, USA, 2004.
- [12] F. Dawson and D. Stenerson. Internet Calendaring and Scheduling Core Object Specification (iCalendar). 1998. Disponível em: <http://www.ietf.org/rfc/rfc2445.txt> (Acessado em 20 de março de 2009).
- [13] Damiao Ribeiro de Almeida, Claudio de Souza Baptista, and Fabio Gomes de Andrade. Using ontologies in context-aware applications. In *DEXA '06: Proceedings of the 17th International Conference on Database and Expert Systems Applications*, pages 349–353, Washington, DC, USA, 2006. IEEE Computer Society.
- [14] Hugo Feitosa de Figueirêdo e Cláudio de Souza Baptista. Um sistema ciente de contexto baseado em uma arquitetura orientada a serviço com suporte a multimídia. In *Workshop de Teses de Dissertações da COPIN 2008*, 2008.
- [15] AP de Vries, B. Eberman, and DE Kovalcin. The design and implementation of an infrastructure for multimedial digital libraries. In *Database Engineering and Applications Symposium, 1998. Proceedings. IDEAS'98. International*, pages 103–110, 1998.
- [16] M.J. Egenhofer and R.D. Franzosa. Point-set topological spatial relations. *International Journal of Geographical Information Science*, 5(2):161–174, 1991.
- [17] A. Eisenberg and J. Melton. SQL: 1999, formerly known as SQL3. *ACM SIGMOD Record*, 28(1):131–138, 1999.
- [18] J. Electronics. Design rule for Camera File system. Technical report, JEITA-49-2, 1998. Disponível em: <http://www.exif.org/dcf.PDF> (Acessado em 20 de março de 2009).

- [19] M. Ester, H.P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining, Portland, OR, AAAI Press*, pages 226–231, 1996.
- [20] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design patterns: elements of reusable object-oriented software*. Addison-Wesley Reading, MA, 1995.
- [21] U. Gargi. Consumer media capture: Time-based analysis and event clustering. Technical report, Technical Report HPL-2003-165, HP Laboratories, August 2003.
- [22] U. Gargi, Y. Deng, and D.R. Tretter. Managing and searching personal photo collections. In *Proceedings of SPIE*, volume 5021, pages 13–21, 2003.
- [23] A. Girgensohn, J. Adcock, and L. Wilcox. Leveraging face recognition technology to find and organize photos. In *Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval*, pages 99–106. ACM New York, NY, USA, 2004.
- [24] M. F. Goodchild and L. L. Hill. Introduction to digital gazetteer research. *International Journal of Geographical Information Science*, 22(10):1039–1044, 2008.
- [25] M.F. Goodchild. The Geolibary. In *Innovations in Gis 5: Selected Papers from the Fifth National Conference on GIS Research UK (GISRUK)*. CRC Press, 1998.
- [26] A. Graham, H. Garcia-Molina, A. Paepcke, and T. Winograd. Time as essence for photo browsing through personal digital libraries. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, pages 326–335. ACM Press New York, NY, USA, 2002.
- [27] Documentation Standards Group. Cidoc crm special interest group definition of the cidoc conceptual reference model: version 4.0, 2004. Technical report, Heraklion, Greece, 2004. Disponível em [http://cidoc.ics.forth.gr/docs/cidoc\\_crm\\_version\\_4.0.pdf](http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.0.pdf).
- [28] V.N. Gudivada. Multimedia systems - an interdisciplinary perspective. *ACM Computing Surveys (CSUR)*, 27(4):545–548, 1995.

- [29] J. Han and M. Kamber. *Data mining: concepts and techniques*. Morgan Kaufmann, 2006.
- [30] E. Hjelmas and B.K. Low. Face detection: A survey. *Computer Vision and Image Understanding*, 83(3):236–274, 2001.
- [31] Cay Horstmann and Gary Cornell. *Core java™, volume I-fundamentals, eighth edition*. Prentice Hall Press, Upper Saddle River, NJ, USA, 2007.
- [32] C. Huc, T. Levoir, and M. Nonon-Latapie. Metadata: Models and conceptual limits. *Second IEEE Metadata Conference*, 1997.
- [33] Framework iCal4j. <http://ical4j.sourceforge.net> (Acessado em 16 de Março de 2008).
- [34] Open Geospatial Consortium Inc. Project document ogc 06-104r3. opengis® implementation specification for geographic information - simple feature access - part 2: Sql option. Technical report, Open Geospatial Consortium Inc., 2006.
- [35] Inc. [IDEAlliance] International Digital Enterprise Alliance. Prism: Publishing requirements for industry standard metadata. Technical report, Março, 2006. <http://www.prismstandard.org/specifications/2.1/>.
- [36] Yuri A. Lacerda, Hugo F. de Figueirêdo, Cláudio de S. Baptista, and Marcus C. Sampaio. A system for self-organizing personal photo collections using spatiotemporal and web-based calendars metadata. (Artigo submetido ao International Journal of Semantic Computing 2009).
- [37] Yuri Almeida Lacerda, Hugo Feitosa de Figueirêdo, Cláudio de Souza Baptista, and Anselmo Cardoso de Paiva. Expandindo e utilizando informações de contexto para a sugestão de anotações de fotografias digitais. In *WebMedia 2008 | Proceedings of the 14th Brazilian Symposium on Multimedia and the Web*, pages 162–169. Brazilian Computer Society (SBC), 2008.
- [38] Yuri Almeida Lacerda, Hugo Feitosa de Figueirêdo, Cláudio de Souza Baptista, and Marcus Costa Sampaio. Photogeo: A self-organizing system for personal photo collections. In *IEEE International Symposium on Multimedia (ISM2008)*, pages 258–265. IEEE Computer Society, 2008.

- [39] J. B. MacQueen. Some methods for classification and analysis of multivariate observations. In L. M. Le Cam and J. Neyman, editors, *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. University of California Press, 1967.
- [40] Y. Manolopoulos, A.N. Papadopoulos, and M.G. Vassilakopoulos. *Spatial databases: technologies, techniques and trends*. Idea Group Pub., 2005.
- [41] Fergal Monaghan and David O’Sullivan. Leveraging Ontologies, Context and Social Networks to Automate Photo Annotation. *Lecture Notes in Computer Science*, 4816:252, 2007.
- [42] M. Naaman, S. Harada, Q.Y. Wang, H. Garcia-Molina, and A. Paepcke. Context data in geo-referenced digital photo collections. In *Proceedings of the 12th annual ACM international conference on Multimedia*, pages 196–203. ACM New York, NY, USA, 2004.
- [43] M. Naaman, YJ Song, A. Paepcke, and H. Garcia-Molina. Automatic organization for digital photographs with geographic coordinates. In *Digital Libraries, 2004. Proceedings of the 2004 Joint ACM/IEEE Conference on*, pages 53–62, 2004.
- [44] M. Naaman, R.B. Yeh, H. Garcia-Molina, and A. Paepcke. Leveraging context to resolve identity in photo albums. In *Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries*, pages 178–187. ACM New York, NY, USA, 2005.
- [45] N. O’Hare, C. Gurrin, H. Lee, N. Murphy, A.F. Smeaton, and G.J.F. Jones. My digital photos: where and when? In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 261–262. ACM New York, NY, USA, 2005.
- [46] Neil O’Hare and Alan F. Smeaton. Context-aware person identification in personal photo collections. *IEEE Transactions on Multimedia, Special Issue on Integration of Context and Content for Multimedia Management (in press)*, 2009.
- [47] JC Platt, M. Czerwinski, and BA Field. PhotoTOC: automatic clustering for browsing personal photographs. In *Information, Communications and Signal Processing, 2003*

*and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on*, volume 1, 2003.

- [48] D. Ravichandran. Clustering Photos to Improve Visualization of Collections. Master's thesis, 2007.
- [49] K. Rodden and K.R. Wood. How do people manage their digital photographs? In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 409–416. ACM Press New York, NY, USA, 2003.
- [50] R. Sarvas, E. Herrarte, A. Wilhelm, and M. Davis. Metadata creation system for mobile images. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 36–48. ACM New York, NY, USA, 2004.
- [51] Richard Sobieraj, Sean e Mislán. *Mobile Phones: Digital Photo Metadata*, 2007.
- [52] K. Toyama, R. Logan, and A. Roseway. Geographic location tags on digital images. In *Proceedings of the eleventh ACM international conference on Multimedia*, pages 156–166. ACM Press New York, NY, USA, 2003.
- [53] W. Viana, JB Filho, J. Gensel, M.V. Oliver, and H. Martin. PhotoMap-Automatic Spatiotemporal Annotation for Mobile Photos. *Lecture Notes in Computer Science*, 4857:187, 2007.
- [54] WA Wagenaar. My memory: a study of autobiographical memory over six years. *Cognitive psychology (Print)*, 18(2):225–252, 1986.
- [55] Anita Wilhelm, Yuri Takhteyev, Risto Sarvas, Nancy Van House, and Marc Davis. Photo annotation on a camera phone. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1403–1406, New York, NY, USA, 2004. ACM.
- [56] M.H. Yang, D.J. Kriegman, and N. Ahuja. Detecting Faces in Images: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 34–58, 2002.

- 
- [57] L. Zhang, L. Chen, M. Li, and H. Zhang. Automated annotation of human faces in family albums. In *Proceedings of the eleventh ACM international conference on Multimedia*, pages 355–358. ACM New York, NY, USA, 2003.
- [58] M. Zhao, Y.W. Teo, S. Liu, T. Chua, and R. Jain. Automatic Person Annotation of Family Photo Album. *Lecture Notes in Computer Science*, 4071:163, 2006.
- [59] W. Zhao, R. Chellappa, PJ Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys (CSUR)*, 35(4):399–458, 2003.
- [60] X.S. Zhou and T.S. Huang. CBIR: From low-level features to high-level semantics. In *Proceedings of SPIE- The International Society for Optical Engineering*, volume 3974, pages 426–431, 2000.

# Apêndice A

## Diagramas do Sistema

Este apêndice apresenta como foi implementado o banco de dados e as classes do PhotoGeo. A Figura A.1 apresenta as tabelas do banco de dados do PhotoGeo. A Figura A.2 e a Figura A.3 apresentam as classes utilizadas pelo módulo de Organização Automática do PhotoGeo. A Figura A.5 apresenta as classes responsáveis por gerar a navegação hierárquica para a interface multi-modal do PhotoGeo. A Figura A.4 apresenta as classes responsáveis pelo módulo de sugestão de anotação. Já na Figura A.6 estão as classes responsáveis pelo módulo de consulta avançada.

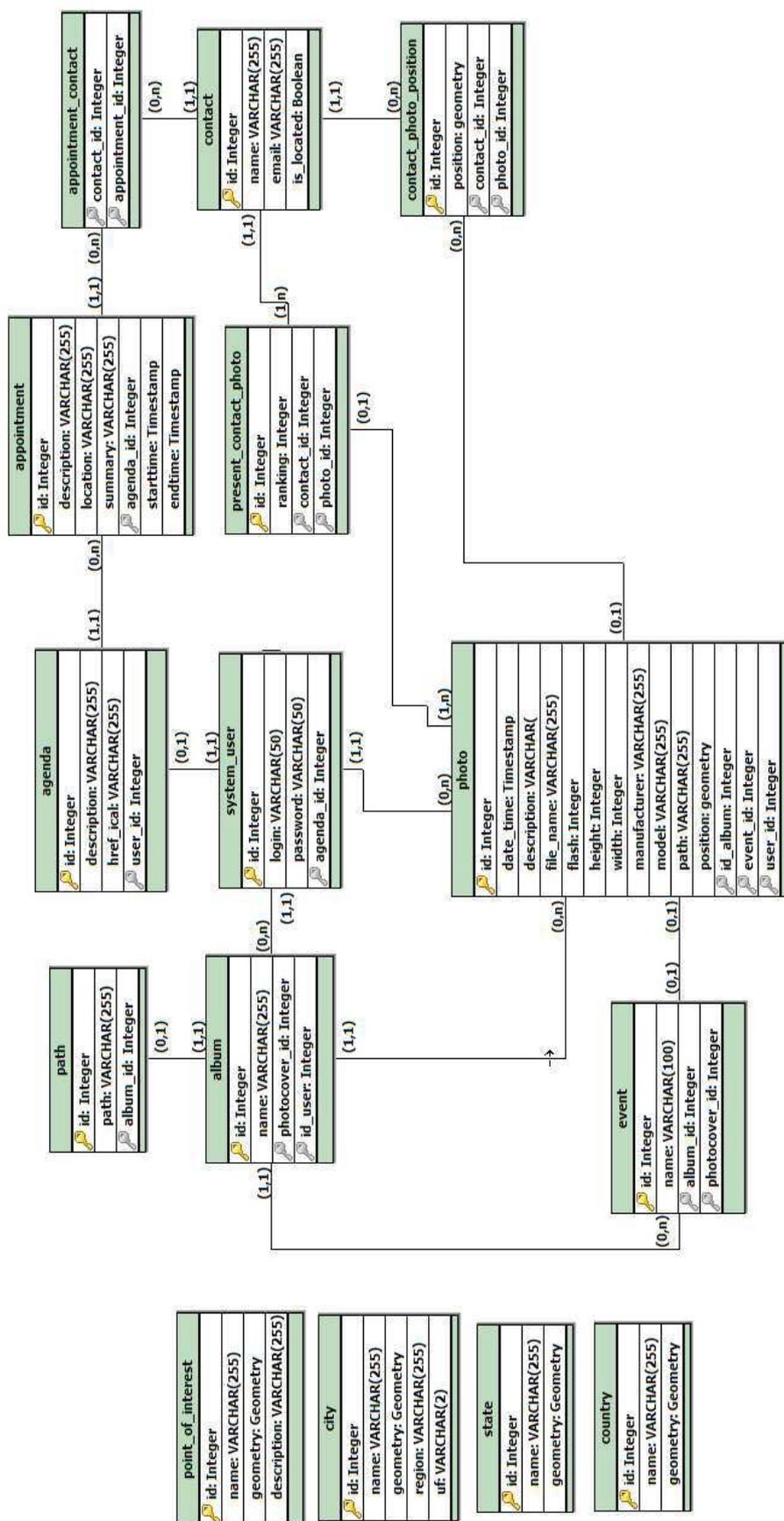


Figura A.1: Tabelas do banco de dados do PhotoGeo

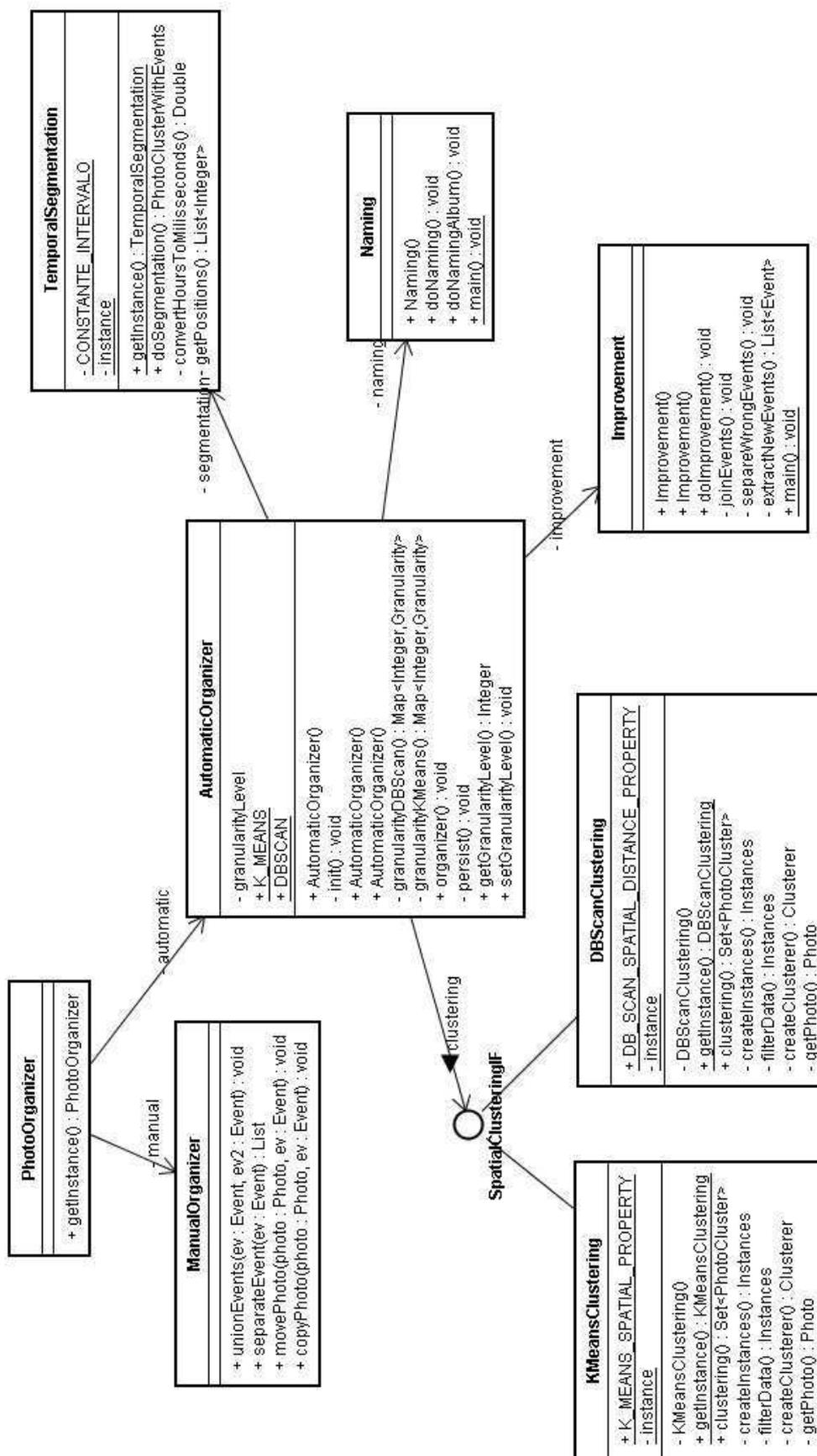


Figura A.2: Diagrama de Classe da Organização Automática

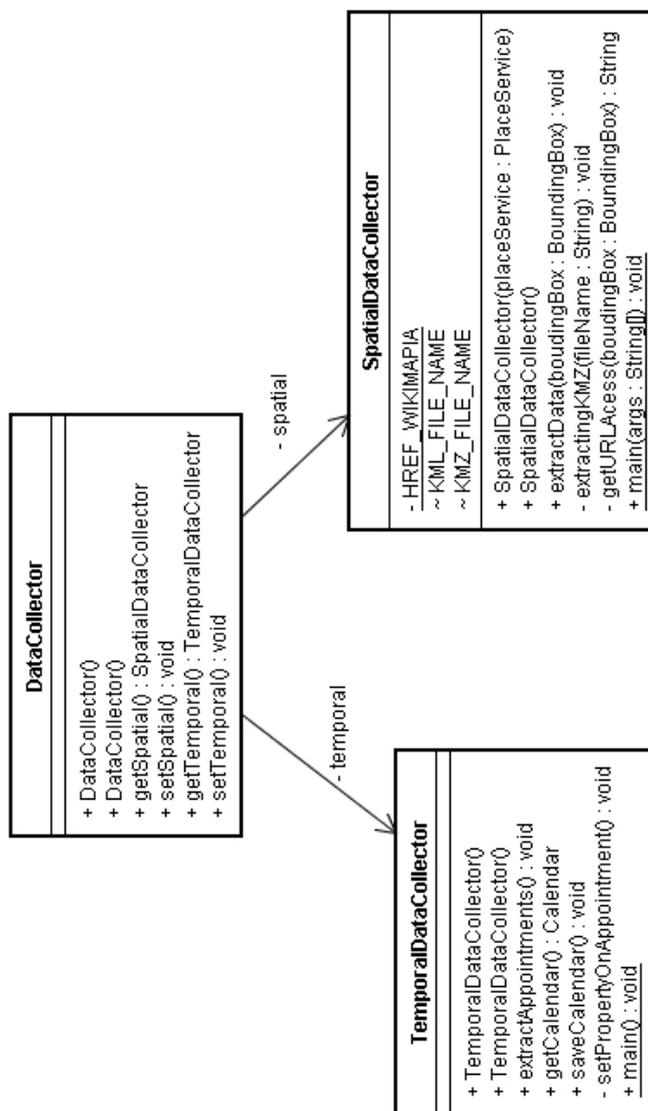


Figura A.3: Diagrama de Classe do Coletor de Dados

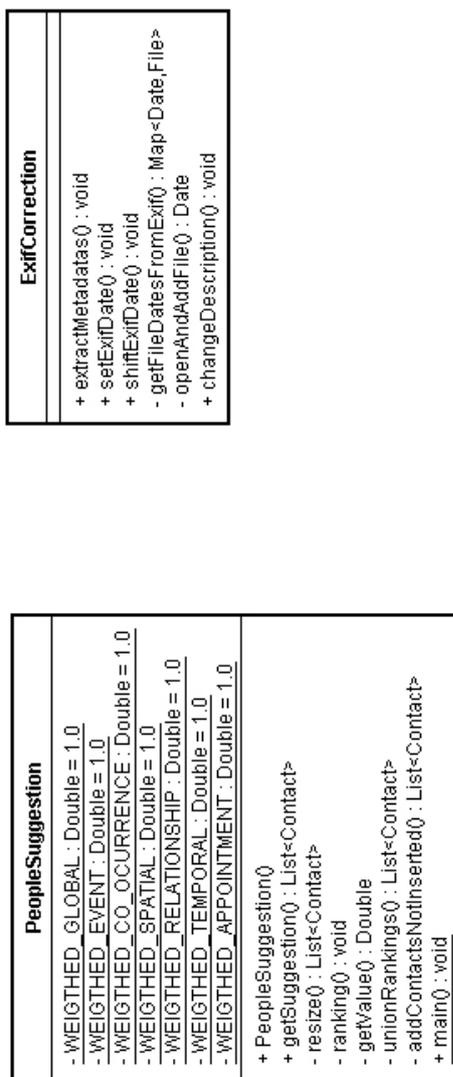


Figura A.4: Diagrama de Classe da Sugestão de Anotação

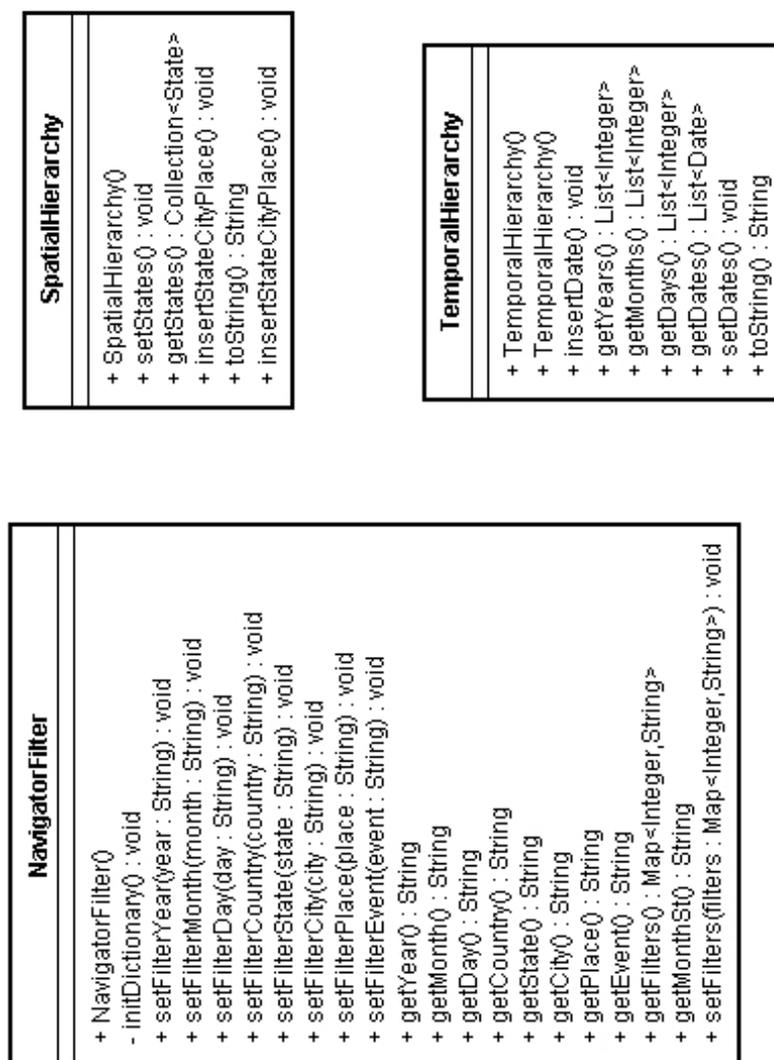


Figura A.5: Diagrama de Classe da Interface Multi-Modal

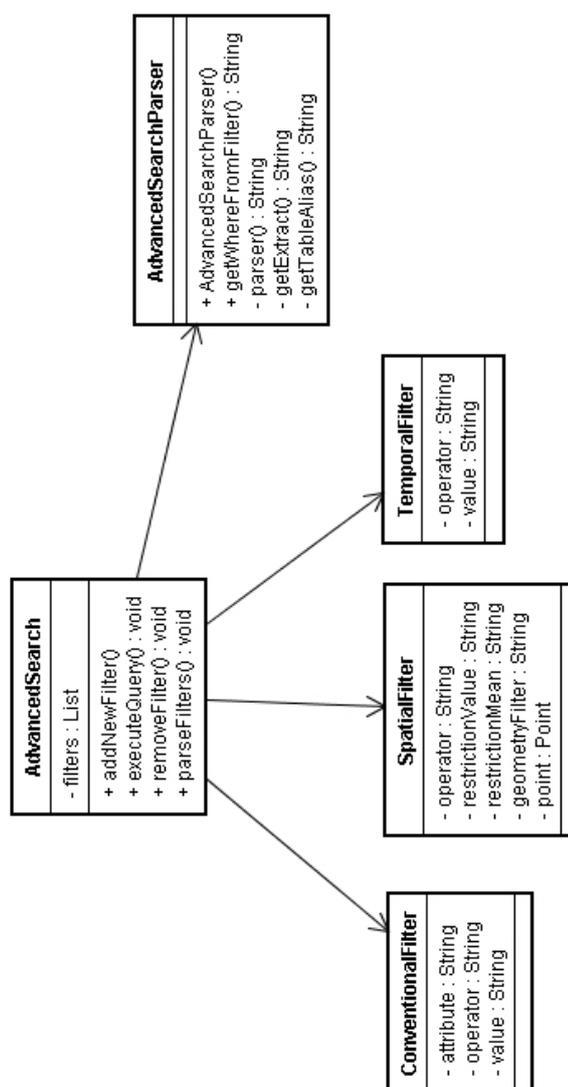


Figura A.6: Diagrama de Classe das Consultas Avançadas