



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA

COORDENAÇÃO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

ANOTAÇÃO DE PESSOAS E EVENTOS EM FOTOGRAFIAS DIGITAIS

Hugo Feitosa de Figueirêdo

Campina Grande, Paraíba, Brasil

©Hugo Feitosa de Figueirêdo, Novembro de 2013



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA

COORDENAÇÃO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

ANOTAÇÃO DE PESSOAS E EVENTOS EM FOTOGRAFIAS DIGITAIS

Hugo Feitosa de Figueirêdo

Tese 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 a obtenção de grau de Doutor em Ciência da Computação.

Área de Concentração: Ciência da Computação

Linha de Pesquisa: Sistemas de Informação e Banco de Dados

Prof. Dr. Cláudio de Souza Baptista

Orientador

Campina Grande, Paraíba, Brasil

©Hugo Feitosa de Figueirêdo, Novembro de 2013

DIGITALIZAÇÃO:
SISTEMOTECA - UFCG

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

F475a Figueirêdo, Hugo Feitosa de.
Anotação de pessoas e eventos em fotografias digitais / Hugo Feitosa de Figueirêdo. – Campina Grande, 2013.
133f. : il. Color.

Tese (Doutorado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

"Orientação: Prof. Dr. Cláudio de Souza Baptista".
Referências.

1. Anotação de Fotografias. 2. Detecção de Eventos. 3. Recuperação de Informação Multimídia. 4. Reconhecimento de Faces. I. Baptista, Cláudio de Souza. II. Título.

CDU 004.4:77(043)

"ANOTAÇÃO DE PESSOAS E EVENTOS EM FOTOGRAFIAS DIGITAIS"

HUGO FEITOSA DE FIGUEIREDO

TESE APROVADA EM 01/11/2013



CLÁUDIO DE SOUZA BAPTISTA, PhD., UFCG
Orientador(a)



WINDSON VIANA DE CARVALHO, Dr., UFC
Examinador(a)



FERNANDO DA FONSECA DE SOUZA, Dr., UFPE
Examinador(a)



LEANDRO BALBY MARINHO, Dr., UFCG
Examinador(a)



JOSÉ EUSTAQUIO RANGEL DE QUEIROZ, D.Sc, UFCG
Examinador(a)



ANSELMO CARDOSO DE PAIVA, DR., UFMA
Examinador(a)

CAMPINA GRANDE - PB

Agradecimentos

Ao meu orientador e amigo, Cláudio Baptista, pelos muitos gritos e broncas, por ter me ensinado muito de como ser um pesquisador e pela paciência que teve comigo.

À minha esposa, Julliana, por ter suportado a falta de atenção que lhe dei durante a redação desta tese e por ter me incentivado nos momentos em que precisei.

Em especial, ao meu filho Igor, que me deu muitas alegrias e muita força nos momentos difíceis.

À minha mãe, por ter me indicado o caminho a ser seguido, pelo apoio financeiro e por ter me ajudado a entender as broncas do meu orientador.

À minha irmã, Marcella, por muitas vezes me colocar para cima quando eu estava cansado do doutorado.

À minha família.

Ao meu grande amigo e parceiro de pesquisas, Yuri Lacerda, com quem realizei e realizarei bons trabalhos de pesquisa.

Ao professor Anselmo, pelos puxões de orelha quando precisei e pelas ótimas recomendações dadas.

Aos colegas de Laboratório, Damião, Fábio Leite, Fábio Gomes, Ana Gabrielle, Júlio Rocha, André, Maxwell, entre outros.

Aos que me auxiliaram na pesquisa do doutorado, Daniel Leite, Davi Serrano, João Paulo Ribeiro, Arthur Sena e Tiago "Baiano".

À CAPES, ao CNPQ e ao IFPB, pelo apoio financeiro.

À UFCG e ao LSI, por toda minha formação acadêmica.

Resumo

A popularidade das câmeras digitais tem gerado um novo problema: como armazenar e recuperar de forma eficiente um grande número de fotografias digitais capturadas e caoticamente armazenadas em múltiplas localizações sem qualquer anotação. Em algumas pesquisas, mostra-se que a anotação de informações relacionadas ao contexto no qual uma determinada fotografia foi capturada auxilia na busca das fotografias, sendo as informações mais relevantes para uma pessoa recordar de uma fotografia: a identificação das pessoas que estão presentes na fotografia (Quem); a localização geográfica (Onde); e informações temporais (Quando). Para anotar essas informações, podem ser utilizados mecanismos manuais, automáticos e semiautomáticos. Os mecanismos manuais não possuem muitos adeptos devido ao processo oneroso e tedioso. Os mecanismos automáticos utilizam os algoritmos de análise de conteúdo, sendo o reconhecimento de faces a principal estratégia, a qual só possui bons resultados com faces frontais e sem oclusões. Já os mecanismos de anotação semiautomática, utilizam algoritmos de recomendação de anotações para auxiliar o usuário. Nesta pesquisa, propõem-se algoritmos para a anotação automática e semiautomática de pessoas e eventos em fotografias. Para a anotação de pessoas, utiliza-se o reconhecimento de faces para a anotação automática e informações de contexto mais conteúdo para a geração de listas de sugestões de pessoas para a anotação semiautomática. Para a anotação de eventos, propõe-se um método para a detecção de eventos em coleções pessoais de fotografias e um método para detecção de eventos compartilhados, nos quais as fotografias são capturadas por usuários distintos em um mesmo evento. A partir do método de detecção de eventos compartilhados proposto, pretende-se melhorar a busca de fotografias de um evento em uma rede social, a realização de anotação cruzada e a detecção de inconsistências nas anotações de fotografias. Nos resultados, comprova-se que a ponderação e a filtragem de estimadores para os algoritmos de anotação semiautomática de pessoas presentes nas fotografias baseados em estimadores melhoram os resultados desses algoritmos. Além disso, é possível fazer a detecção de eventos compartilhados em uma rede social utilizando informações de quem, onde e quando das fotografias.

Abstract

The popularity of digital cameras has created a new problem: how to store and retrieve efficiently a large number of captured and chaotically stored digital photos in multiple locations without annotation. The photo context assists in search of photographs. The most relevant information for a person remember a photograph are: who are present, where and when was captured. To annotate this information, manual, automatic and semi-automatic mechanisms can be used. The manual mechanisms did not have many fans due to the costly and tedious process. Content analysis and face recognition are the main strategy to automatic mechanisms, which has only front faces with good results and without occlusions. In the semiautomatic annotation, recommendation of annotations are used to assist the user. In this research, we propose algorithms for the automatic and semi-automatic annotation of people and events in photos. For the annotation of people, we use face recognition for automatic annotation and content and context information for generate suggestions from people for semi-automatic annotation. For the annotation of events, we propose a method for detecting events in personal photo collections and a method for detection of shared events, in which the photographs are captured by different users in the same event. The method to detect shared events aims to improve the search of photographs of an event in a social network, conducting cross-annotation and detection of inconsistencies in the annotations of photographs. The results prove that the weighting and filtering algorithms estimators for semi-automatic annotation of persons in photographs based estimators improve the results of these algorithms. Furthermore, it is possible to detect shared events in a social network using information of who, where and when of the photos.

Lista de Abreviaturas

CEEI	Centro de Engenharia Elétrica e Informática
UFMG	Universidade Federal de Campina Grande
COPIN	Coordenação de Pós-Graduação em Ciência da Computação
Exif	Exchangeable image file format
XMP	Extensible Metadata Platform
XML	Extensible Markup Language
GPS	Global Positioning System
ARFF	Attribute-Relation File Format

Lista de Símbolos

- F Conjunto de fotografias
- P Conjunto de Pessoas
- f Fotografia $f \in F$
- H_f Conjunto de pessoas para sugestão de pessoas presentes na fotografia f , sendo $H_f \subseteq P$
- P_f Conjunto de pessoas que estão presentes na fotografia f , sendo $P_f \subseteq P$
- i Pessoa $i \in p_f$

Lista de Figuras

Figura 2.1. Detecção e extração de faces de uma fotografia.	12
Figura 3.1. Fotografia e metadados extraídos de seu cabeçalho.....	17
Figura 4.1. Faces extraídas de fotografias com variações de iluminação, pose e expressão. (Fonte: (STONE <i>et al.</i> , 2010)).....	24
Figura 5.1. Gráfico da taxa de acerto (<i>H-Hit Rate</i>) dos algoritmos (servidor e móvel) de sugestão de anotação de pessoas presentes nas fotografias versus número de sugestões (H).	48
Figura 5.2. Gráfico da taxa de acerto dos algoritmos pelo número de pessoas anotadas. ...	49
Figura 5.3. Resultados para a classificação de todos os contatos da coleção 1.	52
Figura 5.4. Resultados para a classificação de todos os contatos da coleção 2.	53
Figura 5.5. Resultados para a classificação de todos os contatos da coleção 3.	53
Figura 5.6. Resultados para a classificação de todos os contatos da coleção 4.	54
Figura 5.7. Resultados para a classificação de todos os contatos de todas as coleções.....	55
Figura 5.8. Resultados para a classificação dos 20 contatos mais populares da coleção 1.	56
Figura 5.9. Resultados para a classificação dos 20 contatos mais populares da coleção 2.	57
Figura 5.10. Resultados para a classificação dos 20 contatos mais populares da coleção 3.	57
Figura 5.11. Resultados para a classificação dos 20 contatos mais populares da coleção 4.	58
Figura 5.12. Resultados para a classificação dos 20 contatos mais populares de cada coleção.	59

Figura 5.13. Gráfico da mediana das taxas de acerto do algoritmo de anotação de pessoas presentes nas fotografias com desvio médio.....	60
Figura 6.1. Precisão do algoritmo completo de detecção de eventos em fotografias utilizando DBSCAN.	77
Figura 6.2. Precisão do algoritmo parcial de detecção de eventos em fotografias utilizando DBSCAN.	77
Figura 6.3. Precisão do algoritmo completo de detecção de eventos em fotografias utilizando K-means.	78
Figura 6.4. Precisão do algoritmo parcial de detecção de eventos em fotografias utilizando K-means.	78
Figura 6.5. Cobertura da anotação de eventos em fotografias com execução completa utilizando DBSCAN.	79
Figura 6.6. Cobertura da anotação de eventos em fotografias com execução parcial utilizando DBSCAN.	79
Figura 6.7. Cobertura da anotação de eventos em fotografias com execução completa utilizando K-Means.....	80
Figura 6.8. Cobertura da anotação de eventos em fotografias com execução parcial utilizando K-Means.....	80
Figura 6.9. Precisão para algoritmo de detecção de eventos utilizando K-Means e DBSCAN variando-se a granularidade.....	81
Figura 6.10. Cobertura para algoritmo de detecção de eventos utilizando K-Means e DBSCAN variando-se a granularidade.....	81
Figura 6.11. Média harmônica para algoritmo de detecção de eventos utilizando K-Means e DBSCAN variando-se a granularidade.....	82
Figura 6.12. Acurácia da função de similaridade versus limiar.....	83

Figura 6.13. Precisão da função de similaridade versus o limiar.	83
Figura 6.14. Sensibilidade da função de similaridade versus o limiar.	84
Figura 7.1. Arquitetura do PhotoGeo (Fonte: (LACERDA, 2009)).	86
Figura 7.2. Arquitetura do protótipo.	87
Figura 7.3. Esquema lógico do Banco de Dados do protótipo.	90
Figura 7.4. Interface que exibe os eventos do usuário.	93
Figura 7.5. Interface que exibe as fotografias de um evento.	94
Figura 7.6. Interface de exibição de eventos compartilhados.	94
Figura 7.7. Visualização da localização geográfica da fotografia no PhotoGeo.	95
Figura 7.8. Visualização de pessoas presentes na fotografia no PhotoGeo.	95
Figura 7.9. Cliente móvel para Windows Mobile.	96
Figura 7.10. Interface do cliente móvel para smartphones com Android.	97
Figura 7.11. Etapas de anotação no cliente móvel Android.	98
Figura A.1. Gráfico das interações dos fatores.	115
Figura A.2. Teste da normal para os resíduos.	115
Figura A.3. Teste de homocedasticidade para os níveis do fator algoritmo.	116
Figura A.4. Teste de homocedasticidade para os níveis do fator coleção de fotografias. .	116
Figura A.5. Resultado da ANOVA.	117
Figura A.6. Teste da normal para os algoritmos.	118
Figura A.7. Intervalos de confiança para os algoritmos.	119
Figura A.8. Teste-T Pareado para comparar os algoritmos EC e RF.	120

Figura A.9. Teste-T Pareado para comparar os algoritmos EC e EC + RF. 120

Lista de Quadros

Quadro 3.1. Comparativo entre Gerenciadores Comerciais de Fotografias.	22
Quadro 4.1. Comparativo entre as abordagens relacionadas.	34
Quadro 4.2. Comparativo entre trabalhos relacionados (Cont.).	35
Quadro 5.1. Fator e Tratamento.	46

Lista de Tabelas

Tabela 5.1. Taxa de acerto dos algoritmos de acordo com o número de sugestões.	48
Tabela 5.2. Taxa de acerto dos algoritmos de acordo com o número de pessoas anotadas.	49
Tabela 5.3. Ponderações dos estimadores calculados utilizando regressão linear.	50
Tabela 5.4. Ponderações dos estimadores calculados utilizando inclinação.	51
Tabela 5.5. Ponderações dos estimadores calculados utilizando correlação.	52
Tabela 5.6. Desvios médios das taxas de acertos dos tamanhos da lista de sugestões para cada estimador e método de fusão.	61
Tabela 5.7. Dados Coletados	61
Tabela 6.1. Níveis de granularidade pré-definidos.	68
Tabela 6.2. Estatísticas da coleção de fotografias de utilizadas na avaliação.	73
Tabela 6.3. Melhores resultados obtidos para cada métrica destacada.	84
Tabela A.1. Cálculo dos Efeitos (α_j e β_i)	112
Tabela A.2. Cálculo dos Resíduos (e_{ij})	113
Tabela C.1. Modelos de requisições REST no sistema.	131
Tabela C.2. Modelos de requisições REST no sistema (continuação).	132
Tabela C.3. Modelos de requisições REST no sistema (continuação).	133

Sumário

Capítulo 1	Introdução.....	1
1.1	Problemática.....	2
1.2	Questões da Pesquisa.....	3
1.3	Hipóteses.....	4
1.4	Objetivos.....	5
1.5	Contribuições.....	5
1.6	Visão Geral da Tese.....	7
Capítulo 2	Anotação em Fotografias.....	9
2.1	Anotação de Pessoas.....	9
2.1.1	Processo de Identificação Automática de Faces em Fotografias.....	11
2.2	Anotação de Eventos.....	12
2.3	Considerações Finais.....	14
Capítulo 3	Gerenciamento de Fotografias.....	15
3.1	Conteúdo e Contexto de Fotografias.....	16
3.2	Metadados de Imagens.....	17
3.3	Gerenciadores de Fotografias Comerciais.....	19
3.3.1	Principais Características dos Gerenciadores.....	21
3.4	Considerações Finais.....	21
Capítulo 4	Pesquisas Relacionadas.....	23

4.1	Anotação de Pessoas.....	23
4.2	Anotação de Eventos	27
4.2.1	Eventos Compartilhados	28
4.3	Anotação de Localização.....	31
4.4	Considerações Finais	33
Capítulo 5	Anotação de Pessoas	36
5.1	Anotação Automática de Pessoas nas Fotografias	37
5.2	Anotação Semiautomática de Pessoas nas Fotografias	38
5.2.1	Estimadores Baseados em Contexto	38
5.2.2	Estimador Baseado em Conteúdo	40
5.2.3	Geração da Lista de Sugestões	40
5.2.4	Combinação de Estimadores	41
5.3	Anotação de Pessoas em Dispositivos Móveis.....	42
5.4	Metodologia de Avaliação.....	43
5.5	Resultados.....	47
5.5.1	Resultados da Anotação de Pessoas em Dispositivos Móveis	47
5.5.2	Resultados da Ponderação na Anotação de Pessoas.....	49
5.5.3	Resultados da Anotação de Pessoas Utilizando Contexto e Conteúdo.....	61
5.6	Considerações Finais	62
Capítulo 6	Detecção de Eventos Pessoais e Compartilhados	63
6.1	Detecção de Eventos Pessoais	65

6.1.1	Granularidade	67
6.2	Detecção de Eventos Compartilhados	68
6.2.1	Similaridade Geográfica (SG).....	69
6.2.2	Similaridade Temporal (ST).....	70
6.2.3	Similaridade de Pessoas (SP).....	71
6.2.4	Observações	71
6.3	Metodologia de Avaliação.....	72
6.3.1	Metodologia de Avaliação do Algoritmo de Detecção de Eventos Pessoais 72	
6.3.2	Metodologia de Avaliação do Algoritmo de Detecção de Eventos Compartilhados.....	75
6.4	Resultados.....	76
6.4.1	Resultados dos Experimentos do Algoritmo de Detecção de Eventos Pessoais	76
6.4.2	Resultados dos Experimentos do Algoritmo de Detecção de Eventos Compartilhados.....	82
6.5	Considerações Finais	85
Capítulo 7	Protótipo de Gerenciador de Fotografias	86
7.1	Arquitetura do Protótipo.....	86
7.2	Banco de Dados	89
7.2.1	Na Anotação de Pessoas Presentes.....	89
7.2.2	Anotação de Pessoas Presentes	90

7.2.3	Detecção de Eventos	92
7.3	Interface Gráfica com Usuário	93
7.4	Avaliação com Usuários	98
Capítulo 8	Conclusões	100
8.1	Proposição para Pesquisas Futuras	101
	Referências Bibliográficas.....	103
	Apêndices	111
	Apêndice A	112
	Apêndice B	121
	Apêndice C	131

Capítulo 1 Introdução

Atualmente, existe uma proliferação de dispositivos móveis com câmeras embutidas, e.g., smartphones, tablets, videogames portáteis. A captura de fotografias é uma das funções mais comuns nos dispositivos móveis atuais. Com a facilidade de sempre estar com uma câmera à mão, os usuários de dispositivos móveis têm capturado uma quantidade cada vez maior de fotografias digitais. Assim, o crescimento acentuado do número de fotografias geradas acarreta uma tarefa onerosa para o usuário no gerenciamento das coleções de fotografias digitais, além de dificultar a realização de buscas em tais coleções.

Em agosto de 2011, o Flickr obteve a marca de 6 bilhões de fotografias¹; enquanto o Facebook possui aproximadamente 100 bilhões de fotografias; por sua vez, o Picasa possui aproximadamente 5 bilhões; e o Photobucket, 8 bilhões². Aproximadamente 250 milhões de fotografias são adicionadas ao Facebook por dia, sendo que, em grandes eventos, esse número aumenta consideravelmente, como por exemplo, no final de semana de 1 de janeiro de 2011, mais de 750 milhões de fotografias foram adicionadas³ por dia. Considerando que somente uma pequena parte das fotografias capturadas são adicionadas às redes sociais, o número de fotografias digitais atinge cifras ainda maiores.

Apesar da grande quantidade de fotografias produzidas, a maioria das pessoas ainda não possui uma organização estrutural no armazenamento de suas fotografias. Alguns o fazem de forma ainda primitiva, tal como a separação por eventos ou diretórios de um nível, não permitindo a manipulação em visualizações diferentes para as coleções de fotografias, devido, principalmente, ao fato de não utilizarem ferramentas para a indexação das fotografias (RODDEN; WOOD, 2003).

O número de soluções acadêmicas (FIGUEIRÊDO *et al.*, 2012a; NAAMAN *et al.*, 2005; O'HARE; SMEATON, 2009; VIANA *et al.*, 2011) e comerciais propostas e postas

¹ <http://blog.flickr.net/en/2011/08/04/6000000000/>, acessado em 01 de outubro de 2011.

² <http://blog.pixable.com/2011/02/14/facebook-photo-trends-infographic/>, acessado em 01 de outubro de 2011.

³ <https://www.facebook.com/press/info.php?statistics>, acessado em 01 de outubro de 2011.

à disposição dos usuários no tratamento de suas fotografias - captura, anotação, armazenamento, indexação, organização e busca - vem aumentando substancialmente. Ainda existem diversos desafios relacionados ao ciclo de vida de uma fotografia, principalmente quando relacionados à anotação de fotografias.

1.1 Problemática

Com o crescimento do número de dispositivos com câmeras embutidas, o barateamento do armazenamento, o crescimento da conectividade com a Internet e o uso das redes sociais, a Web passou a possuir um papel significativo no gerenciamento de fotografias das pessoas. A utilidade de fotografias armazenadas pode ser questionada, visto que os meios disponíveis para sua recuperação e acesso ainda são limitados. Essa limitação ocorre como consequência da falta de acordo no entendimento entre máquinas e humanos: máquinas interpretam fotografias como uma matriz de pixels, enquanto humanos percebem os objetos e o contexto no qual foram capturadas (MONAGHAN, 2008).

Com o intuito de facilitar a recuperação de fotografias em uma coleção, anotam-se tais fotografias com informações sobre o contexto nas quais foram capturadas. Algumas anotações podem ser armazenadas nos metadados das próprias fotografias, e.g., o formato Exif (CP-3451, 2002) permite que sejam armazenadas junto ao arquivo da imagem informações tais como uso do flash, modelo da câmera, localização geográfica e foco, dentre outras.

O processo de anotação de fotografias pode ser realizado de três formas. A primeira é manual, a partir da qual o usuário faz a anotação sem o auxílio de uma ferramenta ou utilizando uma que apenas possua uma interface para a inclusão das anotações. A segunda é semiautomática, a partir da qual o usuário é auxiliado por meio de recomendações para anotar as fotografias. Por fim, a terceira é automática, a partir da qual são obtidas anotações das fotografias utilizando processamento do conteúdo da fotografia e fontes externas (HANBURY, 2008).

A anotação de fotografias é um problema bastante investigado, devido sua importância e complexidade, pois, muitas vezes, é necessário o auxílio do usuário para a realização de anotações (AMES; NAAMAN, 2007). Trata-se, pois, de uma tarefa bastante

onerosa e tediosa. Destarte, busca-se a obtenção de anotações de modo o mais automatizado possível.

Em algumas pesquisas, mostra-se que a anotação de informações relacionadas ao contexto no qual uma determinada fotografia foi capturada auxilia na busca das fotografias. De acordo com pesquisas empíricas (NAAMAN *et al.*, 2004a; RODDEN; WOOD, 2003; WAGENAAR, 1986), as informações mais relevantes para uma pessoa recordar de uma fotografia são: a identificação das pessoas que estão presentes na fotografia (**Quem**); a localização geográfica (**Onde**); e informações temporais (**Quando**).

A maior parte das câmeras digitais adicionam nos metadados das fotografias as informações de data e hora. Com relação à informação de localização geográfica, observa-se que ainda são adicionadas somente em uma pequena parcela de fotografias, apesar de, nos últimos anos, tal parcela ter sido aumentada, devido ao aumento de dispositivos de captura de fotografias que possuem embutido algum método de determinação da localização geográfica. Apesar da informação sobre pessoas presentes nas fotografias ser a mais relevante, a maioria das fotografias não é anotada com essa informação. Esse fato deve-se principalmente à falta de mecanismos eficientes de anotação automática das pessoas presentes nas fotografias, como será abordado na Seção 2.1.

Esta tese trata do desenvolvimento de algoritmos para facilitar o processo de anotação das informações consideradas mais importantes para a recuperação de fotografias: quem, onde e quando.

1.2 Questões da Pesquisa

O desenvolvimento de mecanismos para anotar as fotografias com as informações mais importantes para a sua recuperação geram algumas questões de pesquisa relevantes:

- Q1: É possível realizar anotação automática de informações de quem, onde e quando em fotografias?
- Q2: Quais informações são importantes para melhorar os mecanismos de anotação de pessoas presentes nas fotografias?
- Q3: Como auxiliar usuários na anotação de pessoas presentes nas fotografias?

- Q4: Como detectar erros nas anotações de fotografias?
- Q5: Como propagar anotações de uma coleção de fotografias para outra?
- Q6: Como aumentar a eficácia na geração de anotações e tornar a anotação um processo mais automatizado?
- Q7: Como o processamento nos dispositivos móveis, no momento de captura das fotografias, pode aumentar a eficácia do processo de anotação de fotografias?
- Q8: Como tornar o processo de anotação de quem, onde e quando nas fotografias menos oneroso para o usuário?
- Q9: Quais informações são mais importantes para a anotação semiautomática de pessoas presentes nas fotografias: informações de contexto ou de conteúdo?

1.3 Hipóteses

As seguintes hipóteses foram consideradas e testadas no escopo da pesquisa descrita nesta tese:

- H1: a análise de conteúdo da imagem melhora a taxa de acerto da anotação semiautomática de fotografias baseada em estimadores proposta por Naaman (NAAMAN *et al.*, 2005) e utilizada por Lacerda (LACERDA *et al.*, 2008b);
- H2: mecanismos para ponderar e filtrar estimadores melhoram a taxa de acerto de algoritmos para a anotação semiautomática de pessoas em fotografias;
- H3: é possível detectar eventos a partir de fotografias capturadas por câmeras distintas e armazenadas em diferentes perfis de uma rede social, sendo esses eventos denominados de **compartilhados**;
- H4: as informações das pessoas presentes, carimbo de tempo e localização geográfica são importantes para a detecção de eventos compartilhados em coleções de fotografias de usuários distintos;

1.4 Objetivos

O objetivo geral desta pesquisa é desenvolver técnicas para a anotação das informações de Quem, Onde e Quando em fotografias digitais.

Os seguintes objetivos específicos foram traçados para atingir o objetivo geral:

- O1: desenvolvimento de um algoritmo de anotação semiautomática de pessoas presentes nas fotografias que utilize contexto e conteúdo;
- O2: desenvolvimento de um método para a ponderação e a filtragem de estimadores para o algoritmo de anotação semiautomática de pessoas em fotografias;
- O3: desenvolvimento de um algoritmo de anotação automática de eventos nas fotografias;
- O4: desenvolvimento de um algoritmo de mensuração do grau de similaridade entre eventos de fotografias; e
- O5: desenvolvimento de um algoritmo para a detecção de eventos compartilhados.

1.5 Contribuições

As principais contribuições desta pesquisa são:

- algoritmo de anotação semiautomática de pessoas presentes nas fotografias utilizando contexto mais conteúdo;
- estudo comparativo entre diferentes formas de utilizar conteúdo para auxiliar na anotação semiautomática de pessoas presentes nas fotografias;
- mecanismo de ponderação de estimadores para a anotação semiautomática de pessoas presentes nas fotografias;
- mecanismo para a filtragem de estimadores, que melhora a precisão, a cobertura e o tempo de resposta do algoritmo de sugestões de pessoas presentes nas fotografias;
- algoritmo de detecção de eventos pessoais utilizando níveis de granularidade diferentes;
- proposta de detecção de eventos compartilhados em redes sociais;

- algoritmo de detecção de eventos compartilhados em redes sociais;
- validação da proposta de detecção de eventos compartilhados;
- comprovação de que a detecção de eventos compartilhados auxilia na anotação cruzada de fotografias de diferentes usuários; e
- elaboração de um protótipo para prova de conceitos que abordou a detecção de eventos pessoais, a detecção de eventos compartilhados e anotação de pessoas em fotografias empregando contexto e conteúdo, sendo o protótipo criado para as plataformas Web e móvel.

As publicações geradas diretamente ou indiretamente ao longo desta pesquisa foram:

- On Improving Geotag Quality in Photo Collections. In: Proceedings of The Fifth International Conference on Advanced Geographic Information Systems, Applications, and Services, 2012 (LACERDA *et al.*, 2013);
- Integration Model for Location-Based Services in iDTV Applications. In: Proceedings of The Fifth International Conference on Advanced Geographic Information Systems, Applications, and Services, 2012 (FERNANDES *et al.*, 2013);
- People Photo Annotation Using Context and Content. In: 6th International Workshop on Multimedia and Signal Processing, 2012 (FIGUEIRÊDO *et al.*, 2012c);
- An Approach of Recommendation Based on Non-Negative Matrix Factorization Applied to Photogeo, In: 6th International Workshop on Multimedia and Signal Processing, 2012 (FREIRE *et al.*, 2012);
- Detection of photos from the same event captured by distinct cameras. In: Proceedings of the 18th Brazilian symposium on Multimedia and the Web, 2012 (FIGUEIRÊDO *et al.*, 2012b);
- PhotoGeo: a photo digital library with spatial-temporal support and self-annotation, Multimedia Tools and Applications, Springer, 2012 (FIGUEIRÊDO *et al.*, 2012a);

- Infrastructures for development of context-aware mobile applications, Handbook of Research on Mobility and Computing: Evolving Technologies and Ubiquitous Impacts, IGI Global, 2011 (FIGUEIRÊDO *et al.*, 2011);
- Context and social networks interaction modeling for context-aware alert systems. In: International Conference on Computational Aspects of Social Networks, 2011 (MENEZES *et al.*, 2011);
- Visual Modeling of Workflow with Support to Multimedia and Spatiotemporal Indexing. In: Second International Conference on Electronic Government and the Information Systems Perspective, 2011 (LACERDA *et al.*, 2011); e
- A context-aware solution to annotate people in mobile devices, nos proceedings do "The 7th International Conference on Advances in Mobile Computing & Multimedia (MoMM 2009), ACM, 2009 (FIGUEIRÊDO *et al.*, 2009).

1.6 Visão Geral da Tese

O restante deste documento está organizado como segue.

- Capítulo 2 Anotação em Fotografias: descreve-se a problemática relacionada à anotação de fotografias;
- Capítulo 3 Gerenciamento de Fotografias: são descritos os principais conceitos relacionados ao gerenciamento de fotografias e analisados os principais gerenciadores comerciais;
- Capítulo 4 Pesquisas Relacionadas: mostra-se um levantamento bibliográfico com os principais documentos relacionados à temática desta pesquisa na literatura existentes na área;
- Capítulo 5 Anotação de Pessoas: são apresentados os algoritmos propostos e os resultados obtidos para a anotação de pessoas presentes nas fotografias;
- Capítulo 6 Detecção de Eventos Pessoais e Compartilhados: são apresentados os algoritmos propostos e os resultados obtidos para a detecção de eventos pessoais e compartilhados;
- Capítulo 7 Protótipo de Gerenciador de Fotografias : apresenta-se o protótipo criado, o qual implementa os algoritmos propostos;

- **Capítulo 8 Conclusões:** apresentam-se as conclusões formuladas a partir das investigações e do desenvolvimento atingidos nesta pesquisa. Além disso, são propostos temas para investigações futuras que podem dar continuidade à pesquisa ora documentada;
- **Apêndices:**
 - Apêndices
 - Apêndice A: apresenta-se a análise feita para a validação estatística dos resultados obtidos na avaliação do algoritmo de anotação de pessoas presentes em fotografias utilizando contexto e conteúdo;
 - Apêndices
 - Apêndice B: apresentam-se os procedimentos armazenados no Banco de Dados para a execução dos algoritmos de anotação de pessoas e detecção de eventos; e
 - Apêndices
 - Apêndice C: apresentam-se os formatos de requisições REST do protótipo implementado.

Capítulo 2 Anotação em Fotografias

Neste capítulo, formaliza-se o problema de anotação de pessoas presentes nas fotografias, de eventos e de localização geográfica.

2.1 Anotação de Pessoas

As técnicas destinadas à anotação de pessoas em fotografias podem ser divididas em três grupos. O primeiro grupo utiliza técnicas de reconhecimento de faces. O segundo grupo utiliza técnicas baseadas na utilização de padrões nas informações contextuais relacionadas ao ambiente no qual a fotografia foi capturada. Por fim, o terceiro grupo utiliza técnicas híbridas, melhorando o desempenho do reconhecimento de faces com os padrões das informações de contexto.

A anotação manual de pessoas em fotografias necessita de um grande esforço do usuário, o que ocasiona sua não realização pela maioria dos detentores de grandes coleções de fotografias, apesar de essa ser a principal informação utilizada na busca. Uma solução plausível para este problema seria a anotação automática das pessoas presentes nas fotografias. Contudo, esta solução esbarra no baixo desempenho de técnicas de anotações automáticas, que são baseadas no reconhecimento de faces. Outra solução adotada é a geração de lista de sugestões com as pessoas mais prováveis de estarem presentes na fotografia para o usuário confirmando aquelas que estão realmente presentes, ou seja, uma anotação semiautomática.

Seja F o conjunto de fotografias de um determinado usuário e P o conjunto de pessoas (identificadas pelo nome) presentes nas fotografias, o processo de anotação de pessoas consiste em identificar cada pessoa p presente na fotografia f , onde $p \in P$ e $f \in F$.

O uso de uma técnica para anotar pessoas em fotografias pode diminuir o esforço do usuário. Técnicas que utilizam listas de sugestões geram para cada f uma lista de pessoas $H_f \subseteq P$. A técnica com a solução ótima é aquela que gera H_f , de tal forma que $H_f = P_f$, na qual P_f é a lista de pessoas que realmente aparecem em f e $P_f \subseteq P$.

Dada uma fotografia f , um acerto para uma determinada pessoa p ocorre quando $p \in H_f \cap P_f$ e um erro ocorre quando $p \in P_f - H_f$. A taxa de acerto (*Hit Rate*), razão entre o número de acertos e o número de pessoas presentes na fotografia, pode ser utilizada para mensurar a eficiência das técnicas que sugerem listas de pessoas, sendo H o tamanho de H_f . O valor de H é dependente dos algoritmos e pode variar de acordo com a interface gráfica com o usuário, pois um conjunto H_f com $H = 5$ pode ser adequado para uma interface Web ou Desktop, mas $H \leq 3$ pode ser mais adequado para um dispositivo móvel com dimensões menores de tela.

Assim, o problema da anotação semiautomática de pessoas presentes em fotografias pode ser definido como: encontrar $H_f \subseteq P$ para cada $f \in F$, de forma a se obter a melhor taxa de acerto para um H máximo (*H-Hit Rate*).

Nesta tese, utiliza-se o *H-Hit Rate* como métrica de avaliação dos algoritmos de anotação semiautomática de pessoas presentes em fotografias, sendo essa métrica utilizada na avaliação de várias outras pesquisas relacionadas (O'HARE; SMEATON, 2009) (CHEN *et al.*, 2003).

Algumas soluções para a anotação semiautomática de pessoas nas fotografias (FIGUEIRÊDO *et al.*, 2009; FIGUEIRÊDO *et al.*, 2012a; FIGUEIRÊDO *et al.*, 2012b; LACERDA *et al.*, 2008a; NAAMAN *et al.*, 2005) utilizaram o conceito de estimadores, que são estatísticas baseadas em uma ou mais informações contextuais das fotografias capturadas anteriormente. Alguns exemplos de estimadores são: recorrência geral de um contato na coleção de fotografias, recorrência geográfica de um contato em uma determinada área e recorrência temporal de um contato nas fotografias capturadas em um intervalo de tempo (descritos com detalhes na Seção 5.2.1).

Alguns estimadores apresentam melhores resultados do que outros, em determinadas situações, considerando como métrica *H-Hit Rate*. Porém, na maioria dos casos, são obtidos resultados gerais mais satisfatórios, a partir de uma combinação de vários estimadores, como demonstra-se na Seção 5.5.

Em sua maioria, os algoritmos para a anotação semiautomática de pessoas em fotografias que utilizam estimadores possuem baixa complexidade computacional,

tornando-se apropriados para serem utilizados em plataformas móveis no momento da captura das fotografias e, conseqüentemente, aumentando a probabilidade de a anotação ser realizada.

Porém, um novo fator facilita a incorporação da técnica de reconhecimento de faces em dispositivos móveis, que é o surgimento de dispositivos móveis de alto desempenho, e.g., iPhone 5 da Apple. Estes dispositivos permitem a incorporação de técnicas mais avançadas de visão computacional e reconhecimento de padrões nas aplicações.

Além disso, apesar do reconhecimento de face possuir baixa precisão e cobertura, essa técnica incontestavelmente auxilia na anotação de pessoas, devido a sua automatização. Deste modo, utilizar o reconhecimento de faces, aliado a uma abordagem de anotação semiautomática, que utiliza informações contextuais, torna-se uma solução bastante atrativa para o usuário, pois este não precisará confirmar a anotação de diversas pessoas que irão ser anotadas automaticamente e poderá confirmar a anotação das pessoas que não foram detectadas pelo reconhecimento de faces a partir da lista de sugestões. Outro fator importante é que a taxa de acerto da lista de sugestões pode melhorar, pelo fato de já haver pessoas anotadas e esta informação ser utilizada no algoritmo de elaboração da lista de sugestões.

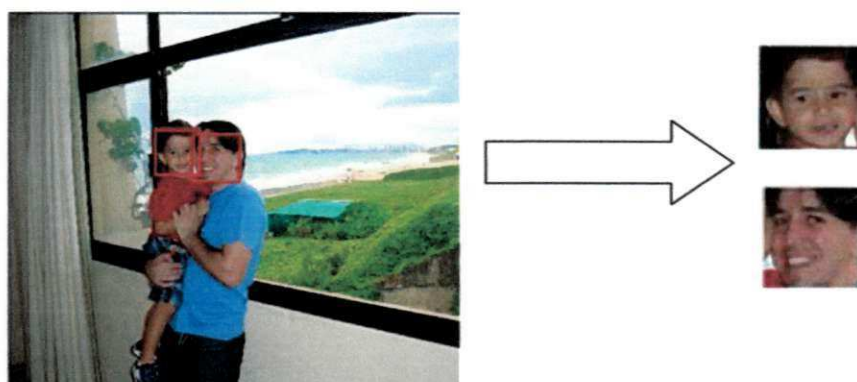
Os problemas mais comuns relacionados ao reconhecimento de faces são: variação da iluminação, oclusão e poses variadas. Um sistema de reconhecimento de faces em dispositivos móveis também pode ser afetado por outros fatores, tais como: baixa qualidade da imagem, poder computacional limitado e memória limitada.

2.1.1 Processo de Identificação Automática de Faces em Fotografias

O processo de identificação automática de faces em uma fotografia é composto basicamente por três etapas: (i) detecção das faces, (ii) extração de características e normalização das faces e (iii) reconhecimento (ZHAO *et al.*, 2003).

Na etapa de detecção de faces, realiza-se uma busca de regiões da imagem que estejam associadas a faces, sendo extraídas essas regiões (YANG *et al.*, 2002). Na Figura 2.1, apresenta-se um exemplo de detecção e extração de faces em uma fotografia.

Figura 2.1. Detecção e extração de faces de uma fotografia.



Dentre os algoritmos propostos para detecção de faces, destaca-se aquele proposto por Viola e Jones (2001), o qual utiliza uma classificação em cascata de características.

Na etapa de extração de características e normalização, realiza-se um processamento nas faces para detecção de pontos de referência, tais como olhos, nariz e boca (YANG *et al.*, 2002). Além disso, uma normalização é feita para que todas as regiões possuam o mesmo formato, sendo alteradas características da imagem, e.g., padrão de cor e tamanho.

Na etapa de reconhecimento, são comparadas as faces detectadas e normalizadas com aquelas contidas na base de dados, retornando um valor de similaridade entre os pares de faces. Os pontos de referência detectados na etapa anterior são comparados com os de outras faces para o reconhecimento de face ser executado. O reconhecimento de faces determina o quanto duas faces são similares. O valor do grau de similaridade entre as faces depende do algoritmo utilizado para o reconhecimento da face. Entre os principais algoritmos de reconhecimento de faces, podem ser citados: *eigenfaces* (TURK; PENTLAND, 1991a) e *fisherfaces* (ZHAO *et al.*, 2003).

2.2 Anotação de Eventos

Um evento é caracterizado por um acontecimento que reúne pessoas com um objetivo comum em um determinado lugar e durante um determinado período de tempo. Podem-se citar como exemplos de eventos: festas de aniversário, concertos de música e

comemoração do Natal. Dessa forma, um evento pode ser descrito por um conjunto de informações contextuais sobre **Quem**, **Quando** e **Onde** de um determinado acontecimento. Assim, a detecção do evento no qual uma fotografia foi capturada é uma informação muito importante para o gerenciamento de fotografias (RODDEN; WOOD, 2003).

A anotação de eventos em coleções de fotografias pode ser definida como a tarefa de geração de subconjuntos de fotografias $Y_{ev} \subseteq F$, que representam a coleção contendo as fotografias que foram capturadas no evento $ev \in Ev$, considerando que Ev representa o conjunto de todos os eventos e F representa o conjunto de todas as fotografias. Esses subconjuntos são gerados de acordo com uma função $\omega: Co \rightarrow B$, que indica se duas fotografias f_i e f_{i+1} consecutivas temporalmente fazem parte de um mesmo evento ev , com $Co = \{(f_i, f_{i+1}) \mid f_i \in F \text{ e } f_{i+1} \in F\}$ e $B = \{0, 1\}$.

Para a detecção de todos os eventos, para cada fotografia $f_i \in F$, é executada a função ω . Se $\omega(f_i, f_{i+1}) = 1$, então $(f_i, f_{i+1}) \in Y_{ev_j}$ e, caso contrário, se $\omega(f_i, f_{i+1}) = 0$, então $f_i \in Y_{ev_j}$, $f_{i+1} \notin Y_{ev_j}$ e $f_{i+1} \in Y_{ev_{j+1}}$.

A detecção automática de eventos pode ser realizada mediante a utilização de informações contextuais das fotografias e fontes externas, como redes sociais (STONE *et al.*, 2008). As principais informações contextuais das fotografias utilizadas para a detecção automática de eventos são: a geográfica e a temporal. O uso da informação temporal é baseada na teoria da explosão de fotografias, na qual fotografias são capturadas em curtos intervalos de tempo durante um evento (COOPER *et al.*, 2005; NAAMAN *et al.*, 2004b).

Um evento possui característica hierárquica, de forma que podem existir subeventos contidos em um evento maior. Por exemplo, uma ida à praia e um churrasco na piscina podem ser considerados subeventos de um evento maior denominado Férias de Janeiro de 2013. Assim, usuários diferentes podem externar diferentes preferências no momento de selecionar em qual evento uma determinada fotografia foi capturada.

Outra característica importante com relação à captura de fotografias durante um evento é a possível existência de várias câmeras diferentes de pessoas distintas. A captura de várias fotografias por cada câmera em um evento é uma situação comum em vários

eventos. Geralmente cada pessoa possuidora de uma câmera fotográfica possui sua organização de fotografias. Como consequência do fato de câmeras distintas possuírem funções distintas, é possível que fotografias capturadas em um mesmo evento possuam tipos de informações de contexto diferentes. Por exemplo, as fotografias capturadas a partir da câmera de um smartphone equipado com GPS possuem a informação de localização geográfica nos seus metadados, enquanto em outra câmera equipada com bússola, as fotografias capturadas possuem apenas as informações de direção e sentido.

No escopo desta pesquisa, os eventos que possuem fotografias capturadas por mais de um usuário (câmeras distintas) serão denominados **eventos compartilhados**. Quando dois eventos forem parte de um mesmo evento compartilhado, significa que possuem um alto grau de similaridade entre suas informações contextuais.

Uma definição complementar àquela de eventos compartilhados é a de **fotografias concorrentes** (JANG; CHO, 2012), que são um conjunto de rolos de fotografias capturadas por um grupo de câmeras em um região específica e em um período de tempo específico. Outras definições complementares dadas pelos autores são: **álbum privado** de fotografias é um conjunto de fotografias obtidas de uma câmera digital fotografadas por um usuário ordinário; e **rolo de fotografias** (*photo roll*) é um conjunto de fotografias capturadas em um período específico que representa um subconjunto de um álbum privado de fotografias com alta similaridade semântica e geográfica.

2.3 Considerações Finais

Neste capítulo, foi apresentado e formalizado o problema relacionado à anotação de pessoas e eventos em fotografias digitais. Além disto, foram apresentadas algumas definições relacionadas à anotação de pessoas e detecção de eventos.

No próximo capítulo, serão apresentados os principais conceitos relacionados ao gerenciamento de fotografias, analisados os principais gerenciadores comerciais e levantadas as características consideradas mais relevantes para a busca de fotografias em gerenciadores de fotografias.

Capítulo 3 Gerenciamento de Fotografias

O gerenciamento de fotografias contendo pessoas possui características peculiares que diferem do gerenciamento geral de imagens. Enquanto em sistemas de busca de imagens baseados em conteúdo são utilizadas características de baixo nível, tais como textura e coloração, para a busca de outras, em sistemas de gerenciamento de fotografias pessoais, as características de baixo nível não são tão importantes, sendo outras informações mais relevantes, tais como as pessoas presentes, o horário e a data que a fotografia foi capturada e a localização geográfica.

Conforme mencionada anteriormente, a falta de organização das fotografias tem dificultado a busca de fotografias pelos usuários. Aliado a este fato, tem-se a falta de informações anotadas nas fotografias, principalmente a de **Quem, Quando e Onde**, consideradas as mais importantes para as pessoas lembrarem-se de uma determinada fotografia (WAGENAAR, 1986).

Considerando um evento como um encontro de pessoas (Quem) realizado em algum lugar (Onde) durante um período de tempo (Quando), verifica-se que o evento possui todas as informações relevantes para serem anotadas nas fotografias, mas, ao capturar as fotografias, as pessoas não têm o hábito, ou ferramentas, para fazer essa anotação.

Neste capítulo, descreve-se como são obtidas as indexações por pessoas presentes nas fotografias e eventos nos quais as fotografias foram capturadas por gerenciadores de fotografias. Em seguida, são apresentados alguns dos principais gerenciadores de fotografias comerciais encontrados e como estes realizam a indexação de fotografias.

O restante deste capítulo está organizado como segue. Inicialmente, descrevem-se o significado de informações de conteúdo e de contexto em fotografias digitais e a importância dessas informações na indexação. Na subseção seguinte, apresenta-se a definição de metadados para fotografias digitais. Em seguida, são apresentados alguns dos principais gerenciadores comerciais de fotografias. Por fim, são apresentadas algumas características consideradas relevantes para gerenciadores de fotografias.

3.1 Conteúdo e Contexto de Fotografias

No escopo deste trabalho, define-se conteúdo de uma imagem como o conjunto de características da imagem geradas a partir dos valores dos pixels. Na recuperação de imagens por conteúdo (*Content-Based Image Retrieval - CBIR*), buscam-se imagens por seu conteúdo, ao invés de utilizar informações textuais (metadados) (LEW *et al.*, 2006). Sistemas de CBIR utilizam estatísticas, reconhecimento de padrões, processamento de sinais e técnicas de visão computacional. Entretanto, ainda existe uma grande lacuna entre informações de conteúdo utilizadas por sistemas de CBIR e os conceitos definidos por usuários com relação às imagens e aos objetos percebidos nas mesmas (semântica dos objetos vistos na imagem) (CAMBRIA; HUSSAIN, 2012).

Os usuários de dispositivos móveis podem se movimentar enquanto executam outras tarefas. Assim, informações podem ser coletadas da situação na qual o usuário está inserido para fornecer serviços e informações personalizadas, realizações automáticas de comandos e armazenamento destas informações para indexação de conteúdo. Esse tipo de informação é chamado de contexto, sendo as aplicações que utilizam este tipo de informação as aplicações cientes de contexto.

Uma das definições mais referenciadas para contexto foi proposta por Dey (DEY, 2001):

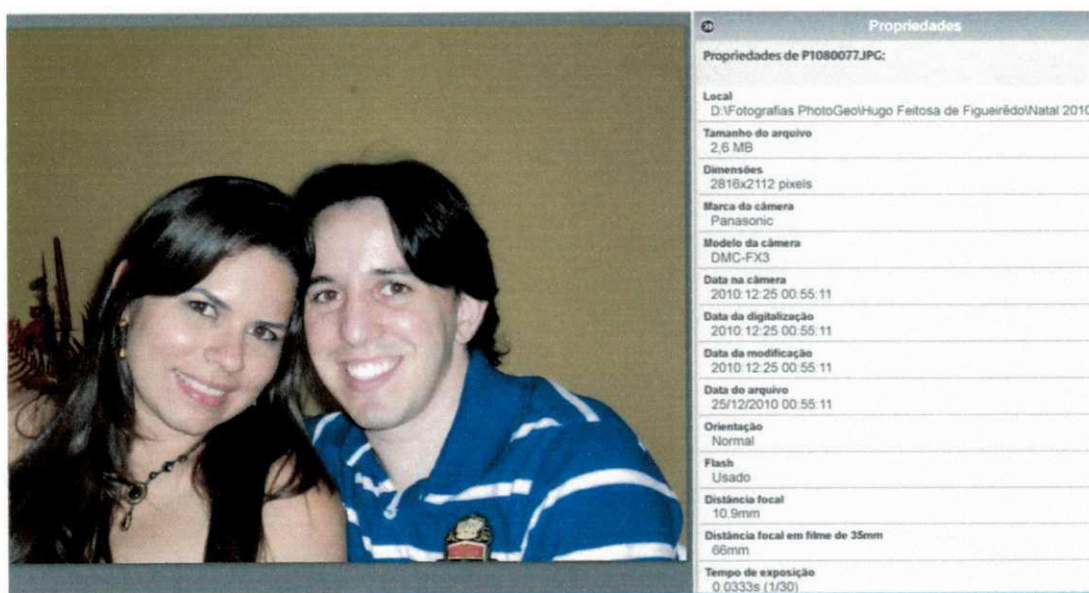
"Contexto é qualquer informação que pode ser utilizada para caracterizar a situação de uma entidade. Uma entidade é uma pessoa, lugar ou objeto que é considerado relevante para a interação entre o usuário e a aplicação, incluindo o próprio usuário e aplicação."

As informações de contexto são importantes para a indexação de fotografias, pois a situação na qual a fotografia foi capturada é utilizada para a busca de fotografias capturadas em um determinado contexto, e.g., um usuário gostaria de recuperar as fotografias que foram capturadas em Junho de 2012, nas quais ele estava presente e que foram capturadas na cidade de Campina Grande.

3.2 Metadados de Imagens

Metadados são dados acerca de dados. Para arquivos do tipo imagem, as informações contidas nos metadados descrevem do que se trata aquela determinada imagem e em quais condições ela foi criada. Nas fotografias, os metadados descrevem as configurações da máquina fotográfica no momento de captura, as condições do ambiente, o dono, dentre outras informações relevantes para expor alguns detalhes da fotografia. Na Figura 3.1, apresenta-se uma fotografia com os seus respectivos elementos armazenados nos metadados.

Figura 3.1. Fotografia e metadados extraídos de seu cabeçalho.



Os metadados de um arquivo digital podem ser armazenados no cabeçalho do próprio arquivo ou em uma base de dados separada. Existem vários formatos diferentes de metadados para os mais variados tipos de arquivos. Cada formato diferente de metadados possui sua forma de armazenamento e elementos obrigatórios e opcionais que devem possuir sobre um arquivo digital, não havendo uma padronização de todos os elementos que devem existir.

Com a intenção de padronizar um conjunto mínimo de elementos que descrevem arquivos digitais, em 1995, um grupo formado por universidades, empresas e organizações não governamentais reuniu-se e elaborou esse conjunto que foi nomeado de

Dublin Core (DC) (DCMI, 2012), sendo mantido pela Dublin Core Metadata Initiative⁴ (DCMI). O conjunto simples de elementos de metadados do DC consiste em quinze elementos: título, criador, assunto, descrição, publicador, contribuinte, data, tipo, formato, identificador, origem, idioma, relação, abrangência e direitos.

O formato mais adotado pelas câmeras digitais para o armazenamento de metadados no cabeçalho do arquivo é o Exif (Exchangeable image file format) (CP-3451, 2002). O formato Exif foi criado pela Japan Electronic Industries Development Association (JEIDA) em 1995, sendo a versão mais recente, a 2.3, elaborada pela Japan Electronics and Information Technology Industries Association⁵ (JEITA), a qual foi criada em 2000, a partir da junção da JEIDA com a Electronic Industries Association of Japan (EIAJ). A International Organization for Standardization (ISO) decretou o Exif como sendo o padrão para o armazenamento de metadados em fotografias pelas câmeras digitais.

A maior parte das fotografias digitais possui um cabeçalho Exif com uma quantidade relevante de informações contextuais, tais como: horário e data em que a fotografia foi capturada, informação de utilização do flash e a localização geográfica (câmeras equipadas com GPS). Sistemas de gerenciamento de fotografias podem utilizar as informações contextuais dos metadados para análise e organização de fotografias.

O Extensible Metadata Platform⁶ (XMP) é um padrão criado pela Adobe para armazenamento de metadados nos cabeçalhos de arquivos. Diversas aplicações já possuem o suporte à edição de informações dos metadados de uma fotografia no formato XMP. Este formato tem a vantagem de armazenar as informações utilizando uma recomendação da W3C⁷, o Resource Description Framework (RDF)⁸.

O International Press Telecommunications Council⁹ (IPTC) é considerado um complemento da especificação Exif, sendo elaborado e mantido por companhias e

⁴ <http://dublincore.org/>. Acessado em outubro de 2011.

⁵ <http://www.jeita.or.jp/english/>. Acessado em Fevereiro de 2013.

⁶ <http://www.adobe.com/products/xmp/>. Acessado em Fevereiro de 2013.

⁷ <http://www.w3.org/>. Acessado em Fevereiro de 2013.

⁸ <http://www.w3.org/RDF/>. Acessado em Fevereiro de 2013.

⁹ <http://www.iptc.org>. Acessado em Fevereiro de 2013.

organizações de notícias. Além disto, os metadados definidos pelo IPTC são baseados no formato XML e adotados pelo XMP.

3.3 Gerenciadores de Fotografias Comerciais

Nesta subseção, são abordados alguns dos gerenciadores de fotografias mais utilizados pelos grandes detentores de fotografias digitais que utilizam algum sistema para organizar suas fotografias.

O Photoscape¹⁰ é um gerenciador de fotografias gratuito que permite visualizar fotografias junto com os metadados no formato Exif e editá-las. Além disto, é possível pesquisar por faces similares na Internet. Porém, este gerenciador não possui mecanismos de indexação de fotografias.

O Picasa¹¹ é o gerenciador de fotografias gratuito – versão desktop – da Google. Neste gerenciador, o usuário pode organizar e visualizar suas fotografias. O Picasa possui anotação manual e automática de pessoas em fotografias, sendo a automática executada por meio de reconhecimento de faces. Porém, o Picasa não utiliza informações contextuais para auxiliar o processo de anotação automática, além de não gerar uma lista de sugestões para auxiliar o usuário em uma anotação semiautomática. Esse gerenciador, além de uma versão desktop, também possui versões Web e para dispositivos móveis – plataforma Android. As fotografias contidas no dispositivo móvel e no desktop podem ser sincronizadas a partir da versão Web. O georreferenciamento das fotografias pelo usuário é realizado mediante a interação com um mapa.

O ACDSec¹² é um gerenciador de fotografias que permite organizar as fotografias por meses e visualizá-las em um mapa. Neste gerenciador, é possível fazer anotação de localização geográfica por meio de um mapa e *tags*. A única forma de obter anotação das pessoas presentes nas fotografias é manual, utilizando *tags*. A versão Web¹³ possui somente algumas funções limitadas, tais como a visualização, a inclusão de *tags* sem sugestões, a avaliação com uma nota das fotografias e o compartilhamento das mesmas.

¹⁰ <http://www.photoscape.org>

¹¹ <http://picasa.google.com/>. Acessado em Outubro de 2011.

¹² <http://www.acdsee.com>. Acessado em Outubro de 2011.

¹³ <http://www.acdseeonline.com>. Acessado em Outubro de 2011.

O Flickr¹⁴, da Yahoo!, é uma rede social de compartilhamento de fotografias digitais e possui as funcionalidades de um gerenciador de fotografias. Esse sistema adota a indexação por *tags*, como forma de gerenciar as fotografias e facilitar busca de fotografias semelhantes. Outra forma de indexar as fotografias é por meio dos metadados contidos no Exif, incluindo a informação de localização geográfica.

O Facebook¹⁵ é uma rede social que permite o gerenciamento de fotografias. Nesta rede social, existe detecção de faces, porém, o reconhecimento de faces não é realizado. Entretanto, o Facebook possui uma grande quantidade de anotações de pessoas presentes nas fotografias, devido a incentivos sociais, tais como o compartilhamento de fotografias e a exibição de alterações no perfil do usuário.

O Facebook é uma rede social largamente usada e como tal, permite que o usuário coloque fotografias organizadas em álbuns. Assim, quando o usuário faz o upload da foto, identificam-se as faces e pede que o usuário nomeie-as.

O Facebook também permite a criação de eventos, no âmbito da qual as pessoas presentes nos eventos podem carregar fotografias. Entretanto, o Facebook não faz a detecção automática de eventos.

O iPhoto¹⁶ desenvolvido pela Apple possui uma tecnologia de detecção e reconhecimento facial que permite identificar quem são as pessoas das fotografias e também combinar imagens semelhantes em uma mesma biblioteca. O iPhoto possui várias possibilidades como localizar e organizar fotografias de pessoas em bibliotecas usando o Faces, que é um dos recursos do programa. Esse recurso ajuda a classificar as bibliotecas a partir das pessoas reconhecidas em cada fotografia. O Faces apresenta um mecanismo de anotação em lote, mediante um agrupamento de todas as faces não anotadas e sugerindo um conjunto de combinações possíveis das faces já nomeadas. Além disso, é possível fazer pesquisas para encontrar amigos que ainda não foram nomeados.

¹⁴ <http://www.flickr.com/>. Acessado em Outubro de 2011.

¹⁵ <https://www.facebook.com/>. Acesso em Novembro de 2013.

¹⁶ <http://www.apple.com/br/creativity-apps/mac/>. Acessado em Novembro de 2013.

3.3.1 Principais Características dos Gerenciadores

Baseado nas hipóteses formuladas na seção 1.3 e após analisar os principais gerenciadores de fotografias comerciais e algumas pesquisas relacionadas (RODDEN; WOOD, 2003) (NAAMAN *et al.*, 2004a) (AMES *et al.*, 2009) (LACERDA, 2009), as seguintes características foram consideradas relevantes para um gerenciador de fotografias digitais:

- C1: Anotação automática de pessoas presentes nas fotografias utilizando reconhecimento de faces;
- C2: Mecanismos para auxiliar na anotação semiautomática de pessoas presentes nas fotografias utilizando contexto e conteúdo;
- C3: Detecção automática de eventos;
- C4: Detecção automática de eventos compartilhados;
- C5: Mecanismos para auxiliar na detecção de inconsistências nas anotações;
- C6: Mecanismos para auxiliar na anotação semiautomática da localização geográfica;

3.4 Considerações Finais

Neste capítulo, foram analisados alguns dos principais gerenciadores comerciais de fotografias digitais e elencadas algumas das características relevantes para um bom gerenciamento de fotografias. Além disso, foram definidos os conceitos de conteúdo e contexto no escopo de fotografias digitais, os quais são muito importantes para indexação de fotografias por gerenciadores de fotografias.

No Quadro 3.1, apresenta-se um comparativo entre os gerenciadores comerciais de fotografias digitais discutidos neste capítulo. As características elencadas na subseção 3.3.1 foram utilizadas para a comparação.

Quadro 3.1. Comparativo entre Gerenciadores Comerciais de Fotografias.

Característica / Sistema	C1	C2	C3	C4	C5	C6
Photoscape	Possui	Não possui	Não possui	Não possui	Não possui	Não possui
Picasa	Possui	Apenas Conteúdo	Possui	Não possui	Não possui	Possui mapa
ACDSee	Não possui	Utiliza tags	Não possui	Não possui	Não possui	Possui mapa
Flickr	Não possui	Apenas contexto para anotação de tags	Utiliza contexto para sugerir tags	Utiliza contexto para sugerir tags	Não possui	Possui mapa
Facebook	Não disponível para os usuários	Detecta faces e utiliza contexto	Não possui	Não possui	Usuários da rede social são alertados e podem corrigir	Utiliza tags
iPhoto	Possui	Apenas conteúdo	Possui	Não possui	Não possui	Possui mapa

Capítulo 4 Pesquisas Relacionadas

Neste capítulo, são apresentadas as pesquisas relacionadas encontradas na literatura na área de anotação e gerenciamento de fotografias.

O restante deste capítulo está estruturado da seguinte forma. Primeiramente, são apresentados os resultados das pesquisas relacionadas à anotação de pessoas presentes nas fotografias, seguido daqueles que tratam sobre a detecção de eventos em coleções de fotografias. Posteriormente, são apresentadas pesquisas encontradas que abordam a anotação da localização geográfica na fotografias. Por fim, são feitas algumas considerações finais sobre as pesquisas relacionadas descritas neste capítulo.

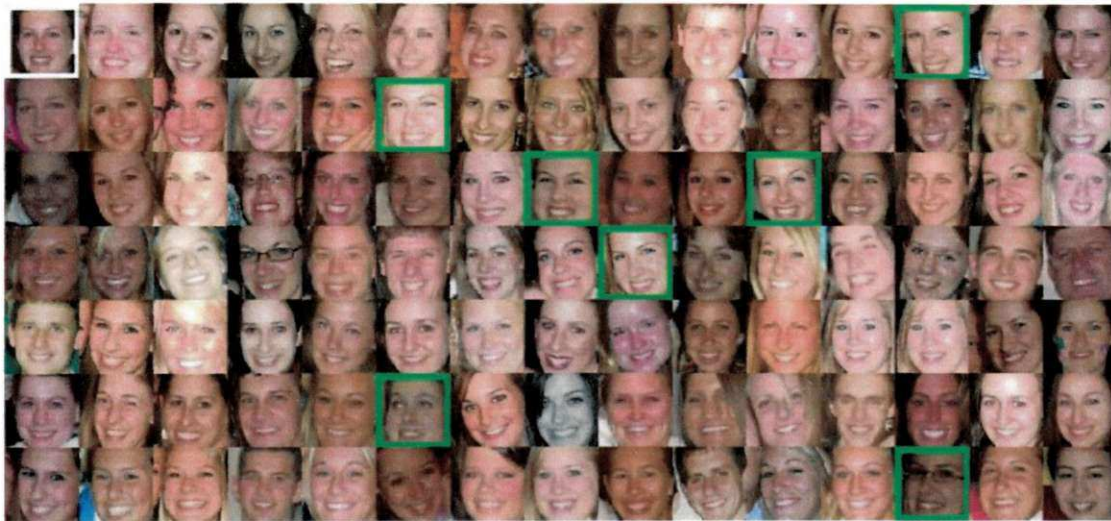
4.1 Anotação de Pessoas

Existem diversas pesquisas documentadas na literatura que abordam a anotação de pessoas presentes em fotografias. As soluções encontradas podem ser classificadas quanto ao grau de automação (automática, semiautomática e manual) ou quanto ao tipo de informação utilizada sobre a fotografia (somente contexto, somente conteúdo ou contexto mais conteúdo). A seguir, os trabalhos relacionados encontrados são descritos.

As soluções que são automáticas e utilizam somente informações extraídas do conteúdo da fotografia empregam técnicas de reconhecimento de faces para a realização das anotações. Essas soluções ainda são imprecisas e possuem baixa cobertura (STONE *et al.*, 2010), onde não existe um controle sobre as variações de pose, iluminação e expressões das faces. Na Figura 4.1, são mostradas algumas faces extraídas de fotografias que possuem essas variações.

Choi *et al.* (CHOI *et al.*, 2009, 2011) criaram um framework para a fusão dos resultados de vários sistemas de reconhecimento de faces existentes nas redes sociais para a execução de uma anotação automática e colaborativa de pessoas presentes nas fotografias. Os mesmos autores, em outro artigo (CHOI *et al.*, 2008, 2010), propuseram um agrupamento das fotografias de uma coleção de acordo com o contexto para, em uma etapa posterior, aplicar o reconhecimento de faces.

Figura 4.1. Faces extraídas de fotografias com variações de iluminação, pose e expressão. (Fonte: (STONE *et al.*, 2010))



Zhang *et al.* (ZHANG *et al.*, 2010) elaboraram uma técnica para agrupamento de pessoas em fotografias pelas roupas para uma posterior utilização em algoritmos de reconhecimento de faces. Em outro trabalho do mesmo grupo (WANG; ZHANG, 2011), foi proposto o aprendizado dos atributos para a busca de fotografias pela roupa. Apesar de nesses trabalhos não ter sido proposta uma solução para anotação automática de pessoas presentes nas fotografias, possuem grande relevância para a utilização de conteúdo para anotação de pessoas.

Gallagher *et al.* (GALLAGHER *et al.*, 2008) utilizaram as informações de agendas pessoais como contexto para anotar as fotografias com as informações do evento. Em outros trabalhos, utilizaram as roupas (GALLAGHER; CHEN, 2008), o tempo (GALLAGHER; CHEN, 2009) e os relacionamentos sociais (WANG *et al.*, 2010) para melhorar o desempenho de reconhecimento de faces. Com relação à localização geográfica (JOSHI *et al.*, 2012), criaram uma máquina de inferência de localização geográfica para fotografias que a partir das *tags* anotadas e uma base de imagens com *geotags* sugere a localização geográfica.

PhotoMap (VIANA *et al.*, 2011) (VIANA, 2010; VIANA *et al.*, 2008) realiza anotação automática de informações temporais, espaciais e contexto social de uma fotografia. Para a anotação do contexto social, utiliza Bluetooth para identificar quais as

peessoas que estão próximas ao local de captura de uma determinada fotografia. Em seguida, o sistema executa a anotação automática das pessoas, incluindo-as como próximas à fotografia. Mas esse sistema não anota as pessoas presentes na fotografia.

O MediAssist (O'HARE; SMEATON, 2009) (COORAY *et al.*, 2006; O'HARE, 2007) é um sistema que oferece navegação, busca e anotação semiautomática de fotografias, utilizando análises de conteúdo de imagem e contexto no qual a fotografia foi capturada. A principal anotação semiautomática realizada pelo MediAssist é o da identidade das pessoas nas fotografias, mediante a análise do conteúdo das imagens e do contexto associado. No MediAssist, para a elaboração da lista de pessoas com maior probabilidade de estarem nas fotografias, foram utilizadas as técnicas de somatório e somatório ponderado para a combinação de estimativas da presença de uma pessoa em uma fotografia, as quais utilizam informações de contexto e conteúdo. Na combinação ponderada, os pesos foram definidos por meio de uma abordagem de força bruta, avaliando todos os possíveis valores para os pesos e adotando aqueles que produziram melhores resultados. Outro problema detectado nesta pesquisa foi a falta de experimentos utilizando informações de contexto e conteúdo aliadas para a solução do problema de anotação de pessoas presentes em fotografias, pois a execução desses experimentos poderia detectar a melhor forma de unir esses dois tipos de informação.

O MMM (DAVIS *et al.*, 2006) (DAVIS *et al.*, 2004, 2005; WILHELM *et al.*, 2004) utiliza uma análise de fatores para combinar informações de contexto com reconhecimento de faces, visando melhorar o desempenho de algoritmos de anotação automática de pessoas presentes nas fotografias. Porém, neste trabalho, não foram utilizadas as informações de contexto e conteúdo para a anotação semiautomática das pessoas presentes nas fotografias.

Anguelov *et al.* (ANGUELOV *et al.*, 2007) propuseram um método probabilístico para aprimorar o reconhecimento de faces utilizando informações da data da captura da fotografias e da aparência das roupas.

Cooray e O'Connor (COORAY; O'CONNOR, 2009) (COORAY, 2008; COORAY *et al.*, 2006; SADLIER *et al.*, 2008) utilizaram a detecção do evento para melhorar o desempenho do reconhecimento de faces. Porém, para a detecção dos eventos não foram

utilizadas informações do contexto da fotografia, mas sim a detecção das roupas que as pessoas estavam utilizando.

PhotoCompas (NAAMAN *et al.*, 2005) utilizaram a ideia de que há padrões para o aparecimento das pessoas nas fotografias. Estes padrões são utilizados como estimadores para calcular a chance de uma determinada pessoa estar numa fotografia. São propostos os estimadores: popularidade, co-ocorrência, recorrência geográfica e temporal. O sistema sugere uma lista de pessoas com alta probabilidade de estarem numa dada fotografia por meio do cálculo destes estimadores. A probabilidade de uma determinada pessoa estar numa fotografia é calculada a partir da soma de todos os estimadores, utilizando o mesmo peso para cada estimador. Dessa forma, não há ponderação dos estimadores de acordo com as coleções de fotografias, comprometendo o *H-Hit Rate*.

O Photogeo (LACERDA, 2009; LACERDA *et al.*, 2008a; LACERDA *et al.*, 2008b) tem como objetivo o auxílio com a anotação de pessoas em fotografias, anotação de eventos em fotografias, armazenamento e busca de fotografias digitais. O PhotoGeo utiliza um somatório dos estimadores para a geração da lista de sugestões para a anotação semiautomática das pessoas presentes nas fotografias, sendo adotado o mesmo peso para todos os estimadores, comprometendo o *H-Hit Rate*.

O ACRONYM (MONAGHAN; O'SULLIVAN, 2007; MONAGHAN, 2008) possui uma solução para a anotação semiautomática de fotografias tanto para sugerir pessoas quanto lugares. Este sistema utiliza informações de GPS, Web semântica, serviços Web e redes sociais de relacionamento. Quando um usuário informa que uma determinada pessoa está na fotografia, então o sistema, mediante os relacionamentos do usuário com esta pessoa, ordena sugestões de outras pessoas, de acordo com sua rede social. Além disto, o ACRONYM utiliza análise de fatores e correlação de Pearson para ponderação de estimadores na anotação semiautomática de estimadores. Entretanto, o ACRONYM não permite o acréscimo de novos estimadores, além de não realizar um filtro dos estimadores mais relevantes para cada usuário.

Stone *et al.* (2010) utilizaram informações de contexto, providas dos metadados das fotografias e da rede social, para melhorar o desempenho do reconhecimento de faces.

Entretanto, esta solução não considera que algumas informações contextuais são mais relevantes para alguns contatos do que para outros.

4.2 Anotação de Eventos

Um dos problemas mais pesquisados na literatura, com relação à anotação de fotografias, é a detecção automática de eventos. A maioria das abordagens propostas utiliza informações visuais, temporais e espaciais, para detecção de eventos.

As abordagens visuais para detecção de eventos são baseadas no conteúdo da fotografia, i.e., as características visuais extraídas utilizando técnicas de CBIR. Fotografias capturadas no mesmo evento, geralmente, contêm características visuais similares, como as cores, textura, formato e estrutura. Por exemplo, fotografias capturadas durante uma caminhada na praia em um dia ensolarado possuem similaridades visuais significantes (VAILAYA *et al.*, 2001; YAVLINSKY *et al.*, 2005).

As abordagens baseadas em metadados temporais utilizam a data e a hora extraídas dos metadados das fotografias, sendo a utilização da informação temporal baseada na teoria da explosão de fotografias (NAAMAN *et al.*, 2004b). Cooper *et al.* (2005) utilizaram essa abordagem. Existem também pesquisas nas quais se combinam técnicas visuais e temporais, como MediAssist (O'HARE; SMEATON, 2009).

A terceira abordagem utiliza os metadados espaciais e temporais, sendo baseadas no fato de que eventos ocorrem em algum lugar, em determinado momento. Dessa forma, um evento contém fotografias de lugares próximos.

PhotoCompas (NAAMAN *et al.*, 2003) (NAAMAN *et al.*, 2004b) detecta eventos utilizando as informações espaço-temporais das fotografias. Além disto, nomeia os eventos automaticamente a partir dos nomes dos lugares e informações temporais. Outra característica desse sistema é a sugestão de anotações de pessoas, que utiliza um algoritmo baseado em estimadores da presença de uma determinada pessoa na fotografia, para listar as pessoas mais prováveis de estarem presentes. Ademais, as anotações do PhotoCompas não são realizadas no momento de captura, sendo necessário que o usuário faça a carga das fotografias no sistema. O PhotoCompas não executa a detecção de eventos compartilhados.

O PhotoGeo (LACERDA *et al.*, 2008b) utiliza um algoritmo para detecção de eventos em coleções de fotografias. O algoritmo utiliza DBScan ou K-Means para realizar um agrupamento geográfico das fotografias e, em seguida, faz uma segmentação temporal por meio da média e desvio padrão entre os intervalos de tempo entre fotografias consecutivas. Apesar deste algoritmo ter obtido bons resultados para eventos de coleções pessoais, não são tratados eventos compartilhados. Também, a localização geográfica nas fotografias é uma informação obrigatória.

Não foram encontrados trabalhos na literatura que consideram as pessoas presentes nas fotografias para a identificação dos eventos. Shimizu et al. (SHIMIZU *et al.*, 2011) utilizaram a co-ocorrência de pessoas para buscar um grupo de fotografias em que uma pessoa tem alta probabilidade de estar presente. Entretanto, não considera a detecção de eventos nas fotografias.

4.2.1 Eventos Compartilhados

As pesquisas que consideram os metadados espaciotemporais para a separação de fotografias pessoais obtiveram bons resultados. Entretanto, para a detecção de eventos compartilhados, estas técnicas esbarram no problema de não considerarem que podem haver diversos eventos em uma área, que ocorreram em um curto período de tempo. Assim, as pessoas presentes nas fotografias podem auxiliar na separação desses eventos distintos e na detecção de eventos compartilhados.

Becker, Naaman e Gravano (2010) utilizaram anotações providas por usuários e metadados das fotografias para agrupar documentos em *sites* com mídias sociais, para tanto, utilizaram uma abordagem de aprendizado de métricas de similaridade para identificação de eventos em mídias sociais. Contudo, a técnica proposta não detecta eventos compartilhados e é aplicada somente ao domínio de mídias sociais.

Jang, Yoon e Cho (2009a) foram os primeiros a utilizarem o fato da existência de várias câmeras em um mesmo evento. Os estudos anteriores em fotografias digitais focavam em bases massivas de fotografias não relacionadas ou coleções privadas de usuários. Jang, Yoon e Cho (2009b) propuseram o agrupamento de fotografias que foram capturadas por um grupo de pessoas durante viagens utilizando informações espaço-temporais, sendo o agrupamento segundo as preferências do dono de cada câmera. O

algoritmo TSA proposto por Cooper et al. (2005) é utilizado para separação pelo *timestamp* das fotografias em grupos. Em outro trabalho (JANG *et al.*, 2010), os autores avaliaram a utilização de outros métodos de classificação supervisionada e não supervisionada.

Nas pesquisas de Jang e Cho (2012) não foi realizada a detecção de eventos compartilhados. O método de agrupamento supervisionado proposto utiliza similaridade espaço-temporal para obter um conjunto de grupos verdadeiros de uma câmera específica de um usuário. Em seguida, são extraídas as características discriminantes dos grupos e aplicadas para agrupar outras fotografias. Métodos não supervisionados utilizam blocos temporais de fotografias para compensar a variação geográfica das fotografias de um usuário. Os autores utilizaram agrupamento hierárquico e baseado em redes neurais. Além disso, os métodos propostos não utilizam as pessoas presentes nas fotografias. Outro problema dos métodos propostos é a necessidade das fotografias estarem georreferenciadas, o que diminui bastante a quantidade de fotografias que podem ser agrupadas em uma rede social. As pesquisas desse grupo são focadas no agrupamento de fotografias concorrentes de acordo com as preferências de um dono de um dos rolos de fotografias presentes nas fotografias concorrentes. Não é abordada a detecção de eventos compartilhados.

CrEve (ZIGKOLIS *et al.*, 2012) é um framework de anotação colaborativa de eventos que utiliza conteúdos. Esse framework extrai informações de sites de mídias sociais com o objetivo de facilitar a anotação de grandes coleções de mídias com informações de eventos. No artigo, comenta-se sobre a grande abundância de mídias compartilhadas em sites, destacando-se que a maioria provém de conteúdo gerado durante grandes eventos, e.g., concertos, jogos de esporte e celebrações. Cita-se também a necessidade de se organizar e indexar esse conteúdo de forma a tornar os mecanismos de busca eficientes. CrEve aborda as seguintes questões: magnitude da coleção de fotografias; inconsistência e propagação de metadados de fotografias; e sugestão de anotação. Nos experimentos, mostra-se que as consultas textuais possuem uma grande relevância para a recuperação de fotografias de um mesmo evento, como também a grande contribuição dos aspectos espaciais e temporais em consultas compostas. Uma das utilidades vislumbradas pelos idealizadores do framework é a utilização desta ferramenta

para a realização das anotações feitas sistematicamente por empresas especializadas em eventos.

No CrEve, a forma de classificação dos estudos com base em eventos foi: (i) representação, (ii) detecção e (iii) anotação. As seguintes dimensões são abordadas sobre eventos: (i) tempo (Quando), (ii) localização (Onde), (iii) participantes (Quem) e (iv) informação (O que). Além disso, considera-se a estrutura hierárquica (eventos consistindo de subeventos).

O CrEve não trata as pessoas presentes no evento ou fotografia como uma anotação diferenciada, podendo esta informação ser uma anotação textual simples. Além disso, não possui a propagação de anotações em distintas coleções de fotografias de um mesmo evento. O CrEve auxilia na criação de eventos de diversas coleções de fotografias, mas não detecta eventos compartilhados de forma automática com o intuito de realizar a detecção de inconsistências e propagação de anotações. Em outro trabalho do grupo de pesquisa (PAPADOPOULOS *et al.*, 2011), os autores concretizaram a detecção de eventos baseada em agrupamento de lugares.

Teleportd¹⁷ é um site para localização de fotografias de um determinado evento em tempo real. Este sistema faz consultas em redes sociais como twitter e busca de fotografias que foram capturadas em um determinado período de tempo em uma determinada localização geográfica. Além disso, o teleportd utiliza as tags anotadas nas fotografias para realizar consultas textuais e fazer uma análise para agrupamento das fotografias.

MediaFinder (RIZZO *et al.*, 2012) procura fotografias e vídeos postados nas últimas 24hs que possuem um termo relacionado, indicando as mídias resultantes como potenciais para serem anexadas a um determinado evento.

Brenner e Izquierdo (2012) propuseram um algoritmo para a detecção de eventos sociais em coleções de fotografias anotadas colaborativamente. O algoritmo foi testado no evento MediaEval 2011 (BRENNER; IZQUIERDO, 2011) com uma coleção de aproximadamente 70.000 fotografias. O algoritmo utiliza informações de fontes externas, descrições textuais e metadados das fotografias para detectar eventos. Dois tipos de eventos foram contemplados nos experimentos: concertos musicais e jogos de futebol.

¹⁷ <http://teleportd.com/>. Acessado em Novembro de 2013.

Presti e Cascia (2012) elaboraram um framework que organiza múltiplas sequências de fotografias concorrentes baseado na proximidade temporal e similaridade do conteúdo visual. Entretanto, o framework não detecta eventos compartilhados, só organiza as fotografias capturadas.

Não foram encontrados trabalhos na literatura que propusessem métodos para a detecção de eventos compartilhados, sendo o trabalho que mais se aproxima da proposta desta tese, é o de Jang, Yoon e Cho (2010). Além disto, não foram encontradas pesquisas que utilizem as informações temporais, espaciais e de pessoas presentes nas fotografias para a detecção de eventos.

4.3 Anotação de Localização

O sistema MMM (DAVIS *et al.*, 2004) constitui-se de uma aplicação cliente num celular com câmera embutida que acessa um servidor responsável por armazenar as fotografias. Neste sistema, o usuário, após capturar uma fotografia, carrega a fotografia para o servidor, incluindo seus metadados. Em seguida, o sistema executa um algoritmo de similaridade com os dados enviados previamente e retorna para o usuário validá-los. Esse sistema sugere apenas nomes de lugares.

WWMX (*World Wide Media eXchange*) (LOGAN; ROSEWAY, 2003) é um gerenciador de fotografias pioneiro na utilização de informações espaço-temporais para indexar fotografias. WWMX permite buscar fotografias baseado na localização geográfica. Entretanto, WWMX não auxilia na anotação de fotografias e não faz inferências para detecção de eventos automaticamente.

No ZoneTag (NAAMAN; NAIR, 2008), as informações de localização podem advir de um GPS ou então das torres de celulares. As recomendações de *tags* são extraídas das anotações já executadas pelo próprio usuário e pelos componentes de sua rede social, RSS feeds, dentre outras. Este sistema possui uma interface simplificada que, além de permitir a anotação, também permite ao usuário compartilhar sua fotografia mediante a carga automática para o sistema Flickr¹⁸. Observa-se que esse sistema não realiza anotações automáticas nas fotografias, dando somente uma assistência aos usuários nas

¹⁸ www.flickr.com. Acessado em outubro de 2011.

anotações de *tags* mais utilizadas naquela determinada área, não sendo detectados os pontos de interesse.

O sistema LOCALE (NAAMAN; PAEPCKE; GARCIA-MOLINA, 2003) propõe a anotação de rótulos automaticamente por meio da associação de fotografias georreferenciadas que não estão rotuladas, com as que estão. Os rótulos não ficam visíveis ao usuário, sendo utilizados para as consultas por fotografias. Para estimar a probabilidade de um termo ser relativo a uma fotografia, são usadas técnicas tais como frequência de aparecimento, ponderação e agrupamento de termos. O LOCALE faz a anotação de forma automática apenas de rótulos, podendo atribuir informações errôneas e não caracterizando o tipo de rótulo que está sendo atribuído à fotografia, ou seja, não identifica se a anotação é de uma pessoa, de um ponto de interesse, de um objeto ou de outro tipo qualquer. Além disto, o LOCALE não utiliza fontes externas, como redes sociais, para a anotação automática.

Conforme mencionado anteriormente, o ACRONYM (MONAGHAN; O'SULLIVAN, 2007) realiza a anotação de lugares, além da anotação de pessoas. Para tanto, as informações sobre os lugares são recuperadas do serviço Web do GeoNames¹⁹, de acordo com a posição geográfica que foi capturada a fotografia.

SPIM (SPatial Image Management) (CHRISTODOULAKIS *et al.*, 2010) realiza anotações de objetos espaciais importantes presentes na fotografia, a partir dos metadados de localização geográfica, data, tempo, direção e sentido. A base de conhecimento do SPIM é formada por ontologias de domínio, e.g., monumentos arqueológicos, igrejas medievais e construções culturais modernas. O SPIM não realiza anotação de pessoas presentes nem de eventos.

Hays e Efros (2008) propuseram uma estratégia chamada de im2gps destinadas à predição de localização geográfica de uma imagem a partir da distribuição probabilística de imagens no globo. Uma base de 6 milhões de fotografias foi utilizada. A imagem é comparada com a base de conhecimento por meio de descritores. As imagens georreferenciadas mais similares são retornadas e utilizadas para estimar a localização

¹⁹ <http://www.geonames.org>. Acessado em Novembro de 2013.

geográfica da imagem da consulta. Para 16% dos casos, o algoritmo acertou a localização geográfica com uma precisão de 200 km de raio.

Serdyukov et al. (2009) realizaram um estudo nas *tags* utilizadas no Flickr e elaboraram um modelo de linguagem para prever a localização geográfica de fotografias. A solução proposta para o georreferenciamento foi dividindo a superfície terrestre em uma matriz e verificando-se a frequência das *tags* em cada célula.

Métodos que utilizaram informações textuais obtiveram melhores resultados do que aqueles que utilizaram informações de conteúdo da imagem no MediaEval 2010 e 2011, que foi um desafio entre grupos de pesquisa para avaliar quem possuía o melhor algoritmo de associar a localização geográfica a vídeos.

4.4 Considerações Finais

Neste capítulo, foram analisadas as principais abordagens relacionadas encontradas na literatura, sendo destacadas aquelas que trataram da anotação de pessoas presentes nas fotografias, da detecção de eventos em fotografias e do georreferenciamento de fotografias.

Analisando as referidas abordagens, observa-se que soluções baseadas em estimadores mostram-se promissoras para a anotação de pessoas em fotografias. Para melhorar o desempenho desse processo, informações de contexto e conteúdo podem ser utilizadas juntas. Porém, não foram encontrados na literatura trabalhos que verifiquem qual a melhor forma de unir esses dois tipos de informações para auxiliar o usuário na anotação de pessoas.

Com relação à anotação de eventos, verificou-se que as redes sociais e as informações de contexto são informações de grande relevância para o agrupamento das fotografias com relação ao evento no qual foram capturadas. Porém, o fato de várias pessoas terem capturado fotografias de diferentes máquinas fotográficas e armazenado em diferentes perfis de redes sociais não foi considerado, apesar de essa informação poder ser utilizada para a anotação automática de fotografias de diferentes usuários.

No Quadro 4.1, apresenta-se um comparativo entre as abordagens apresentadas neste Capítulo, considerando-se as principais funcionalidades destacadas nesta pesquisa.

Quadro 4.1. Comparativo entre as abordagens relacionadas.

	Anotação automática de pessoas	Anotação semiautomática de pessoas	Detecção de eventos pessoais	Detecção de eventos compartilhados	Anotação semiautomática da localização geográfica
Choi et al.	Conteúdo mais contexto	-	Somente tempo.	-	-
Zhang et al.	Conteúdo mais contexto	-	-	-	-
Gallagher	Conteúdo mais contexto	-	Somente tempo.	-	-
PhotoMap	Somente contexto	-	-	-	Mapa
MediAssist	Conteúdo mais contexto	Conteúdo mais contexto. Ponderação estática de estimadores.	Somente tempo.	-	Mapa
MMM	Conteúdo mais contexto	-	-	-	-
Anguelov et al.	Conteúdo mais contexto	-	-	-	-
Cooray e O'Connor	Conteúdo mais contexto	-	Somente tempo.	-	-
PhotoCompas	-	Contexto	Tempo e Localização	-	Mapa
PhotoGeo	-	Contexto	Tempo e Localização	-	Mapa
Acronym	-	Contexto. Ponderação dinâmica de estimadores.	Somente tempo.	-	Sugere geotags
Stone et al.	Conteúdo mais contexto	-	-	-	-

Quadro 4.2. Comparativo entre trabalhos relacionados (Cont.).

	Anotação automática de pessoas	Anotação semiautomática de pessoas	Deteccção de eventos pessoais	Deteccção de eventos compartilhados	Anotação semiautomática da localização geográfica
Jang et al.	-	-	Tempo e Localização	Define fotografias concorrentes. Não detecta eventos compartilhados.	-
CrEve	-	Sugere tags.	-	Busca mídias de um evento em redes sociais utilizando informações textuais, temporais e espaciais.	Sugere geotags
Teleportd	-	-	-	Busca fotografias em tempo real de um evento. Utiliza informações de tempo e localização.	-
MediaFinder	-	-	-	Busca mídias em redes sociais publicadas nas últimas 24hs com um termo.	-
Brenner e Izquierdo	-	-	-	Utiliza fontes externas, descrições textuais, tempo e localização para detectar eventos sociais.	-
Presti et al.	-	-	Tempo e conteúdo.	Utiliza tempo e conteúdo para organizar fotografias concorrentes.	-

Capítulo 5 Anotação de Pessoas

Para solucionar o problema de anotação de pessoas presentes em fotografias digitais pessoais, propõe-se um algoritmo que realize, inicialmente, anotações automáticas e, posteriormente, auxilie o usuário em anotações semiautomáticas.

O algoritmo de anotação de pessoas proposto possui duas versões: uma para arquiteturas cliente-servidor e outra para clientes desconectados. A solução para servidor será útil para a anotação de coleções de fotografias não anotadas, mas já armazenadas em algumas mídias pelo usuário. A solução para clientes desconectados será útil para a implementação em smartphones, tablets e câmeras digitais, sendo possível realizar a anotação mesmo sem conexão com o servidor.

Adota-se como hipótese nesta pesquisa que quanto mais demora-se para anotar uma fotografia, maior a probabilidade dessa fotografia nunca ser anotada. Assim, a versão do algoritmo de anotação de pessoas presentes nas fotografias para clientes desconectados permite que a anotação seja realizada mesmo em situações em que não exista conexão com o servidor, aumentando-se as chances das fotografias serem anotadas.

Verifica-se que, em se tratando de dispositivos móveis, a possibilidade de operações desconectadas é um aspecto importante a ser considerado para garantir a disponibilidade. Por exemplo, em regiões rurais onde não há sinal de operadoras de celulares, o usuário fica impossibilitado de anotar suas fotografias. Este problema também pode ocorrer em lugares em regiões urbanas com muita concentração de pessoas, como por exemplo, estádios de futebol, onde algumas pessoas podem ficar sem sinal de celular, devido à sobrecarga das operadoras telefônicas. Assim, há a necessidade de desenvolver uma solução que atenda tanto dispositivos móveis com conexão quanto sem-conexão com um servidor central.

Para a anotação automática, utilizam-se técnicas de análise de conteúdo de imagem, como detecção e reconhecimento de faces. Já para a anotação semiautomática de pessoas presentes em fotografias, será utilizada uma técnica baseada em combinação linear de estimadores. Informações de contexto e de análise de conteúdo serão utilizadas

peio algoritmo proposto, sendo gerado um conjunto H_f de pessoas mais prováveis de estarem na fotografia f . As pessoas já anotadas de maneira automática ou semiautomática serão retiradas da lista, sendo essas anotações utilizadas para melhorar o *H-Hit Rate* do algoritmo de solução para smartphones. O algoritmo de anotação de pessoas utiliza somente alguns estimadores proposto para o servidor, sendo que outros possuem somente uma implementação simplificada.

5.1 ANOTAÇÃO AUTOMÁTICA DE PESSOAS NAS FOTOGRAFIAS

Uma das técnicas utilizadas para anotação de pessoas, como dito anteriormente, é a baseada na análise do conteúdo da imagem. Mediante essa análise, algumas características importantes da imagem podem ser identificadas para auxiliar na anotação de pessoas.

O primeiro passo para a anotação de pessoas utilizando-se a análise de conteúdo é a detecção das faces. Esta tarefa é realizada por meio da identificação de componentes da face humana na fotografia, e.g., olhos, boca e nariz.

O segundo passo é o reconhecimento das faces detectadas no passo anterior. Para identificar as pessoas presentes na fotografia, o reconhecimento utiliza algumas características detectadas nas faces, como distância entre os olhos. Essas características são comparadas com as contidas na base de treinamento, a qual é incrementada a cada pessoa anotada.

A seguir, serão apresentados os dois passos supracitados para a anotação de pessoas utilizando informações extraídas do conteúdo das fotografias – a detecção e o reconhecimento de faces.

Para a detecção de faces, foi utilizado o algoritmo de classificação que utiliza as características *Haar*, proposto por Viola e Jones (2001) e aperfeiçoado por Lienhart e Maydt (2002).

Após as regiões que representam faces, segundo o classificador, serem detectadas nas fotografias, estas são inseridas no Banco de Dados.

O passo seguinte, após a detecção de faces, é o reconhecimento das regiões armazenadas no Banco de Dados. O reconhecimento utiliza o algoritmo de EigenFaces (TURK; PENTLAND, 1991b).

Uma pontuação para cada contato $c \in C$ é calculada pelo algoritmo de EigenFaces e armazenada no Banco de Dados. Se a pontuação de algum contato for superior a um determinado limiar, a anotação é realizada automaticamente. Entretanto, quando a anotação aparecer para o usuário, ele poderá confirmar a anotação. Somente as anotações confirmadas são utilizadas para a base de treinamento.

Para comparar a anotação automática com a semiautomática, após ser calculada e armazenada a pontuação de cada contato, uma lista H_f é gerada com as H pessoas melhores classificadas para a fotografia f . Caso uma pessoa tenha sido pontuada em mais de uma região para uma mesma fotografia, prevalecerá a maior pontuação.

5.2 Anotação Semiautomática de Pessoas nas Fotografias

Sendo C o conjunto de contatos já identificados pelo sistema, as técnicas propostas por Naaman *et al.* (2005) e Lacerda *et al.* (2008) fazem o uso de estimadores baseados em contexto que avaliam a probabilidade de um contato $c \in C$ estar presente em f . Como descrito na Seção 2.1, os estimadores geram um conjunto H_f , sendo o conjunto dos contatos presentes na fotografia f representado por C_f .

5.2.1 Estimadores Baseados em Contexto

A definição dos três primeiros estimadores utiliza a Equação 1, para cada $c \in C$.

$$Recorrência(c, X) = \sum_{f \in X} \frac{presente(c, f)}{|X|}, presente(c, f) = \begin{cases} 1 & \text{se } c \in C_f \\ 0 & \text{caso contrário} \end{cases} \quad (1)$$

na qual $X \subseteq F$ é um conjunto de fotografias que depende da especificação do estimador.

[RG] Recorrência Global: porcentagem de vezes que um contato foi anotado na coleção de fotografias do usuário. Para o cálculo desse estimador, denota-se X na Equação 1 como o conjunto de todas as fotografias anotadas.

[RGeo] Recorrência Geográfica: porcentagem de vezes que um contato foi anotado em uma fotografia dentro de um raio de d metros da localização geográfica da fotografia f_0 a ser anotada. Para computar esse estimador, considera-se $X = \{f | \text{distância}(f_0, f) \leq d\}$ na Equação 1.

[RT] Recorrência Temporal: porcentagem de vezes que um contato foi anotado num intervalo t do momento em que a fotografia f_0 a ser anotada foi capturada. Para computar esse estimador, considera-se $X = \{f | \text{tempo}(f_0) - t \leq \text{tempo}(f) \leq \text{tempo}(f_0) + t\}$ na Equação 1.

[CO] Co-ocorrência: porcentagem de vezes que um contato está presente na fotografia simultaneamente com outro contato que já foi anotado na fotografia a ser anotada. Considerando dois contatos c_1 e c_2 , sendo $c_2 \in C_f$ (já anotado na fotografia f). A co-ocorrência de c_1 e c_2 é calculada pela Equação 2.

$$CO(i_1, i_2) = \frac{\sum_{f \in F} \text{presente}(c_1, c_2, f)}{\sum_{f \in F} \text{presente}(c_2, f)}, \quad (2)$$

$$\text{presente}(c_1, c_2, f) = \begin{cases} 1 & \text{if } c_1, c_2 \in C_f \\ 0 & \text{caso contrário} \end{cases}$$

na qual C_f e $\text{presente}(c, f)$ são definidos como na Equação 1.

Como deve ser utilizado apenas um valor para o estimador CO para cada $c_1 \in C$, considera-se a Equação 3 para computar o valor de CO para c_1 .

$$CO(c_1) = \text{Max}_{c_2 \in C_f} (CO_{c_1, c_2}) \quad (3)$$

Os valores dos parâmetros d e t interferem no desempenho dos estimadores RGeo e RT, respectivamente. Neste trabalho, consideram-se estimadores diferentes para cada valor adotado nos parâmetros de RGeo e RT. Assim, pode existir um estimador RT para as fotografias capturadas em um intervalo de tempo de 1 hora e outro para 1 semana. Desta forma, os resultados obtidos pelos estimadores RGeo e RT não podem ser considerados para justificar a maior relevância de uma determinada informação de contexto com relação à outra, ou seja, não se pode afirmar que a localização geográfica é mais importante que o tempo para o problema de anotação de pessoas em fotografias.

5.2.2 Estimador Baseado em Conteúdo

O estimador baseado em conteúdo é computado a partir da pontuação obtida pelo contato na etapa de anotação automática das pessoas presentes nas fotografias. O estimador é definido a seguir:

[RF] Reconhecimento de Faces: maior pontuação do contato entre os valores retornados pelas execuções do algoritmo de reconhecimento de faces aplicados as faces detectadas na fotografia. Por exemplo, se duas faces forem detectadas em uma fotografia e o algoritmo de reconhecimento de faces obtiver duas pontuações distintas para o mesmo contato, então a maior pontuação será considerada para este estimador.

Esse estimador é utilizado somente na anotação semiautomática baseada em contexto mais conteúdo.

5.2.3 Geração da Lista de Sugestões

Para gerar a lista de sugestões H_f , para auxiliar na anotação das pessoas presentes na fotografia, podem ser usados vários estimadores. Considerando-se E como sendo o conjunto de estimadores existentes. O valor final da probabilidade do contato estar presente na fotografia é calculado pelo somatório ponderado dos estimadores, conforme a Equação 4.

$$P_{c,f} = \sum_{e \in E} \text{estimador}(e, c, f) * \text{peso}(e, c), \quad (4)$$

na qual, a função *estimador* calcula o valor do estimador e para o contato c na fotografia f e a função *peso* retorna ponderação do estimador e para o contato c .

Após a combinação dos estimadores de cada candidato, uma classificação ordenada de forma decendente a partir do valor de $P_{c,f}$ é gerada. As H pessoas melhores classificadas são adicionadas em H_f . A seguir, serão apresentadas algumas formas de combinar os estimadores e a fórmula utilizada para calcular o peso de um estimador para um contato.

5.2.4 Combinação de Estimadores

A fusão dos estimadores pode ser realizada a partir do preenchimento da lista de sugestões pelo primeiro resultado de cada estimador ou ponderando cada estimador, com o objetivo de obter o melhor resultado.

Na combinação de estimadores por preenchimento, os candidatos são inseridos na lista de sugestões obedecendo à seguinte regra: os primeiros colocados de cada estimador são colocados na lista, sendo os provenientes dos estimadores mais precisos postos primeiro, em seguida, para completar a lista, os candidatos provenientes dos estimadores menos precisos, porém com maior cobertura, são combinados e adicionados.

Ponderação é o ato de aumentar ou diminuir a importância de determinados critérios para se obter um valor ótimo como resposta. O problema de ponderação é conhecido como “otimização”, que implica em achar a combinação que retorna o valor extremo, máximo ou mínimo, de uma função.

Algumas técnicas de ponderação utilizam heurísticas para determinar quais estimadores devem possuir maior peso na combinação do que outros. Porém, essas técnicas não possuem um fundamento matemático para provar que as ponderações adotadas são ideais.

Nesta pesquisa, utiliza-se para a ponderação dos estimadores uma função de inclinação, que é a taxa de variação na linha de regressão entre dois pontos, sendo adotada após a realização de experimentos com alguns métodos de ponderações encontrados na literatura. Para a filtragem dos estimadores, utiliza-se regressão linear de multivariáveis usando o método de mínimos quadrados para a remoção dos estimadores do somatório ponderado, sendo adotado devido à simplicidade da incorporação no algoritmo.

As informações de contexto extraídas de uma fotografia são utilizadas pelos estimadores, os quais analisam estas informações juntamente com as das fotografias já anotadas no PhotoGeo e estimam a probabilidade de cada um dos contatos do usuário estarem presentes na fotografia.

Após cada um dos estimadores calcularem o valor da estimativa de cada contato estar presente na fotografia, é realizada uma combinação dos resultados dos estimadores.

Esta combinação gera um valor único para cada contato, representando a probabilidade do contato estar presente na fotografia. Em seguida, é gerada uma lista com os H contatos mais prováveis de estarem na fotografia, sendo esta lista apresentada ao usuário para auxiliá-lo na anotação da identificação das pessoas presentes na fotografia.

Para o cálculo da ponderação utilizando inclinação, são utilizadas as Equações 5, 6 e 7.

$$peso(e) = \frac{\sum_{f \in F} \left\{ \sum_{c \in C_f} [estimador(e, c, f) - \overline{estimador(e)}] \cdot [presente(c, f) - \overline{presente(c)}] \right\}}{\sum_{f \in F} \left\{ \sum_{c \in C_f} [estimador(e, c, f) - \overline{estimador(e)}]^2 \right\}}, \quad (5)$$

$$\overline{estimador(e)} = \sum_{f \in F} \left[\sum_{c \in C_f} \frac{estimador(e, c, f)}{|C_f|} \right], \quad (6)$$

$$\overline{presente(c)} = \sum_{f \in F} \frac{presente(c, f)}{N_f}, \quad (7)$$

nas quais x é o valor do estimador e y é o valor binário que representa a presença do contato na fotografia.

A inclinação é usada como um peso para majorar a significância dos estimadores que com pequenos valores (valores próximos de 0) indicam a presença da pessoa c na fotografia f .

Atualmente, o cálculo das ponderações dos estimadores para cada contato no PhotoGeo é calculado automaticamente a cada 10 novas anotações do contato. Após serem calculados, os valores das ponderações dos estimadores são armazenados.

5.3 Anotação de Pessoas em Dispositivos Móveis

Para a anotação de pessoas em fotografias nos dispositivos móveis, foi utilizado um algoritmo que pode se conectar ao servidor para requisitar a lista de sugestões de pessoas mais prováveis de estarem presentes nas fotografias ou, caso o dispositivo móvel não esteja conectado a rede, utiliza-se uma versão simplificada do algoritmo de anotação de pessoas proposto para servidor.

Todos os estimadores apresentados foram implementados no sistema PhotoGeo. Entretanto, existe a necessidade de uma versão desses estimadores para dispositivos móveis.

Após uma implementação dos estimadores propostos para o Sistema Operacional Windows Mobile 6.1, verificou-se que os estimadores RG, RT, REv e RGeo são viáveis para serem utilizados no algoritmo final de sugestão de anotação de pessoas em fotografias. A implementação do estimador CO para dispositivo móvel não teve bons resultados, necessitando de um alto custo de processamento ou de armazenamento. Assim, este estimador não foi considerado para ser usado no algoritmo de sugestão.

O cálculo do estimador RG no dispositivo móvel é baseado em dois contadores armazenados na aplicação. O primeiro contador armazena o número total de anotações do sistema. O outro contador armazena o número de anotações de cada pessoa já anotada pelo menos uma vez. O estimador RT foi implementado a partir de uma lista FIFO que armazena as dez últimas pessoas anotadas no sistema. Os estimadores REv e RGeo foram implementados de forma similar ao RG, porém com contadores que abrangem somente as anotações ocorridas no evento corrente ou na área atual, respectivamente.

5.4 Metodologia de Avaliação

Para a avaliação das técnicas de combinação de estimadores foram utilizadas quatro coleções diferentes de fotografias pessoais.

Cada coleção de fotografia foi dividida em duas subcoleções, cada uma contendo 50% da coleção original. A primeira subcoleção foi utilizada para treinamento, ou seja, ponderação dos estimadores. A segunda subcoleção foi utilizada para avaliação das técnicas de combinação dos estimadores. Não foi adotada a técnica de validação cruzada para separação das coleções, pois foi utilizada a ordem cronológica das fotografias. Dessa forma, foi adotado um procedimento mais próximo da realidade do processo de anotação de fotografias.

Foram consideradas duas situações para avaliação. Na primeira, o usuário deseja anotar todas as pessoas que ele conhece que estão presentes na sua coleção de fotografias. Na segunda, o usuário deseja anotar somente os 20 contatos mais populares na sua coleção

de fotografias. Para cada situação foram executados 8 experimentos: 4 experimentos utilizando cada estimador isoladamente, um utilizando o somatório simples de cada estimador, um utilizando o somatório ponderado a partir da regressão linear, um utilizando o somatório ponderado por meio da inclinação e um utilizando o somatório ponderado mediante a correlação para a combinação dos estimadores. Desta forma, para cada coleção foram executados 16 experimentos, totalizando um total de 64.

Para a avaliação do algoritmo de anotação de pessoas utilizando contexto e conteúdo, foi considerada a situação em que o usuário deseja anotar todas as pessoas que ele conhece e que estão presentes na sua coleção de fotografias. A mesma métrica da avaliação da ponderação foi utilizada, o *H-Hit Rate*.

Para avaliar os algoritmos, a seguinte questão de pesquisa foi levantada:

Q10: O algoritmo que analisa informações de contexto e de conteúdo para sugerir pessoas presentes em fotografias possui melhor taxa de acerto em relação aos algoritmos que analisam somente informações de contexto ou somente informações de conteúdo?

Com o intuito de responder a questão de pesquisa, as seguintes hipóteses foram avaliadas:

HA₀: As taxas de acertos resultantes dos algoritmos que sugerem pessoas presentes em fotografias são iguais, ou seja, independem do tipo de algoritmo utilizado.

HA₁: As taxas de acertos resultantes dos algoritmos que sugerem pessoas presentes em fotografias não são iguais, ou seja, dependem do tipo de algoritmo usado.

Se **HA₀** for refutada, as seguintes hipóteses serão avaliadas:

HB₀: O algoritmo que analisa informações de contexto e de conteúdo não apresenta melhor taxa de acerto do que o melhor dos algoritmos que analisam somente informações de contexto ou somente informações de conteúdo

HB₁: O algoritmo que analisa informações de contexto e de conteúdo apresenta melhor taxa de acerto do que o melhor dos algoritmos que analisam somente informações de contexto ou somente informações de conteúdo.

Um estudo empírico foi realizado para comparar o *H-Hit Rate* dos algoritmos. Para a realização de um experimento, foi necessário determinar as variáveis de estudo. As variáveis podem ser classificadas como independentes ou dependentes. As independentes são variáveis que podem ser manipuladas e controladas, as dependentes são variáveis as quais se deseja estudar, ou seja, ver os efeitos causados por mudanças realizadas nas variáveis independentes. Como variáveis independentes do estudo, tem-se: (**I₁**) tipo do algoritmo, (**I₂**) coleção de fotografias, (**I₃**) tamanho da lista de sugestões e (**I₄**) técnica de detecção e reconhecimento de face. Como variável dependente tem-se: (**D₁**) *H-Hit Rate*.

Cada variável independente pode ser chamada de fator. Um fator pode ser secundário ou primário. O fator primário é aquele que se deseja manipular, mudar seus valores e observar o efeito gerado, enquanto o fator secundário é aquele o qual o efeito não é de interesse do estudo, logo esse fator recebe um valor fixo, pré-determinado. Cada fator primário possui um conjunto de valores chamado de tratamento, no qual cada valor representa um nível do tratamento.

Neste estudo, as variáveis **I₁** e **I₂** foram consideradas fatores primários. As variáveis **I₃** e **I₄** foram consideradas fatores secundários. A variável **I₃** recebeu como valor fixo o tamanho de lista com cinco sugestões, valor sugerido pelos autores para a visualização da lista de sugestões em dispositivos com tamanho do visor muito pequeno, por exemplo, smartphones. A variável **I₄** recebeu um valor fixo. Utilizou-se o algoritmo de detecção proposto por Lienhart e Maydt (2002) e o algoritmo de reconhecimento de faces proposto por Turk e Pentland (1991b), ambos implementados pela OpenCV (Open Source Computer Vision). É importante salientar que os parâmetros de entrada utilizados pelos algoritmos de detecção e reconhecimento de faces são os parâmetros padrões da biblioteca OpenCV.

O resumo das variáveis e dos seus respectivos tratamentos é apresentado no Quadro 5.1.

O planejamento de um experimento consiste em determinar as variáveis, os tratamentos, o número de testes (*trials*), e o número de repetições. Os testes consistem da combinação dos níveis de cada fator, que podem ser combinados de várias formas.

Quadro 5.1. Fator e Tratamento.

Fator	Tratamento
Algoritmo (A)	Baseado em Contexto (EC), Baseado em Conteúdo (RF), Baseado em Contexto mais Conteúdo (EC + RF)
Coleção de Fotografias (B)	1, 2, 3, 4

As variáveis e os tratamentos utilizados no experimento foram definidos anteriormente, onde apenas dois fatores primários foram usados. O número de testes foi dado pela combinação de todos os níveis dos fatores definidos anteriormente, ou seja, todas as possibilidades. Não haverá repetição, pois a taxa de acerto permanece inalterada considerando repetições de um mesmo teste.

O design empregado para o experimento foi o fatorial completo com dois fatores e sem repetição. Assumindo que o efeito dos fatores, suas iterações e os erros são aditivos, o seguinte modelo serviu para o cálculo dos efeitos dos fatores:

$$y_{ij} = \mu + \alpha_j + \beta_i + e_{ij}$$

na qual y_{ij} é taxa de acerto para o fator Algoritmo estando no nível j e o fator Coleção de Fotografias no nível i , μ é a taxa média de acerto, α_j é o efeito do fator Algoritmo, β_i é o efeito do fator Coleção de Fotografias e e_{ij} é o erro.

O modelo escolhido foi utilizado para avaliar a importância dos fatores.

Para coletar os dados (taxas de acerto), rotinas que simulam um usuário anotando suas fotografias foram implementadas para cada algoritmo. A taxa de acerto foi calculada utilizando o teste *Gold Standard* (Padrão de Ouro), este teste consiste em comparar os resultados obtidos no experimento com os dados reais, no caso deste estudo, as anotações feitas pelos usuários manualmente.

Três tipos de validações foram considerados: interna, externa e de construção.

A validação interna verifica se existe um relacionamento de causa-efeito entre as variáveis independentes e as dependentes, ou seja, se as variações observadas na variável dependente realmente é resultado da manipulação das variáveis independentes. A externa verifica se as conclusões são verídicas e se podem ser generalizadas para a população de estudo. A validação de construção verifica se o que se deseja medir está sendo realmente medido.

Neste estudo, uma ameaça à construção diz respeito ao fato das fotografias serem manualmente anotadas pelos usuários, proprietários das coleções, para a realização do experimento. Se o usuário deixou de anotar pessoas ou só anotou as pessoas mais populares, a taxa de acerto pode não refletir a realidade. Para contornar esse problema, o experimento poderia ser realizado levando em consideração apenas as pessoas mais populares de uma coleção.

Para mensurar a eficácia dos algoritmos no dispositivo móvel, foram feitos alguns testes e adotada a mesma métrica de *H-Hit Rate*. Além disso, foi realizada uma comparação com os resultados obtidos pelo algoritmo proposto em Lacerda et al. (2008a), que é executado no servidor.

5.5 Resultados

Nesta seção, são apresentados os resultados obtidos com os algoritmos de anotação de pessoas no dispositivo móvel e no servidor, utilizando somente contexto, somente conteúdo e contexto mais conteúdo.

5.5.1 Resultados da Anotação de Pessoas em Dispositivos Móveis

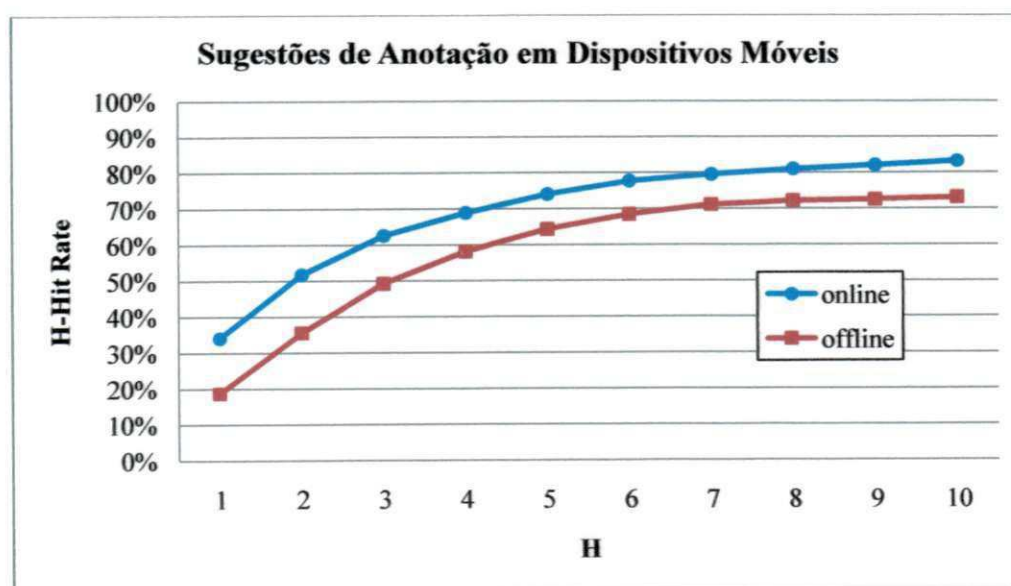
A primeira análise realizada foi com relação ao comportamento da taxa de acerto frente à modificação do número de sugestões. O algoritmo do dispositivo móvel teve uma taxa de acerto de 64,4% e 73,1%, quando o número de sugestões foi igual a 5 e 10, respectivamente. Com o mesmo número de sugestões, o algoritmo executado no servidor obteve uma taxa de acerto de 74,09% e 83,24%. Na Tabela 5.1, são exibidos os resultados das taxas de acerto do algoritmo no dispositivo móvel e no servidor de acordo com o

número de sugestões. Na Figura 5.1, são mostrados os resultados na forma de um gráfico de linha.

Tabela 5.1. Taxa de acerto dos algoritmos de acordo com o número de sugestões.

Sugestões	Servidor	Móvel
1	34,2%	18,8%
2	51,8%	35,6%
3	62,7%	49,2%
4	69,0%	58,2%
6	74,1%	64,4%
7	77,7%	68,5%
8	79,5%	71,2%
9	81,0%	72,1%
10	82,1%	72,6%

Figura 5.1. Gráfico da taxa de acerto (*H-Hit Rate*) dos algoritmos (servidor e móvel) de sugestão de anotação de pessoas presentes nas fotografias versus número de sugestões (H).

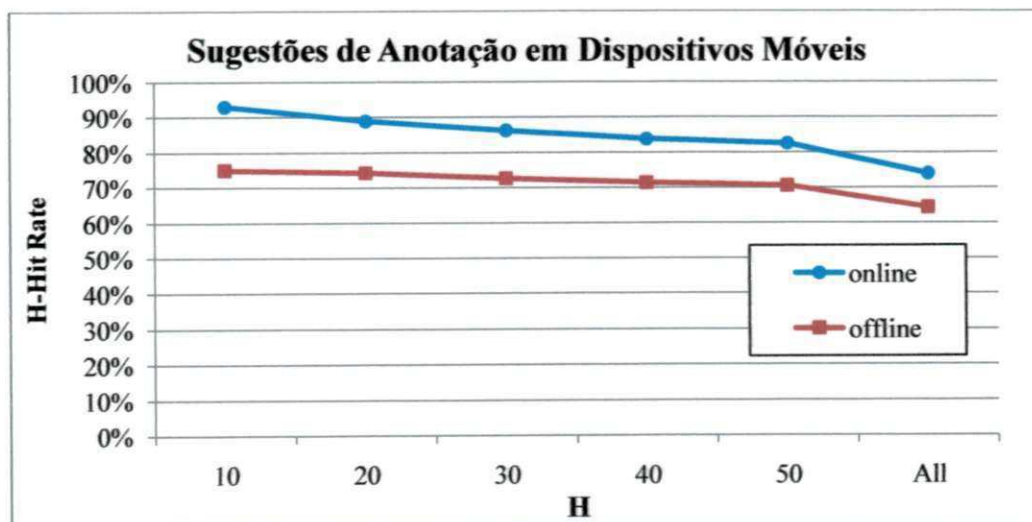


Em outro experimento realizado, utilizou-se uma lista de cinco sugestões para avaliar a taxa de acerto dos algoritmos quando é alterado o número de pessoas que são anotadas pelo usuário. Na Tabela 5.2 e na Figura 5.2, são apresentados os resultados desse experimento.

Tabela 5.2. Taxa de acerto dos algoritmos de acordo com o número de pessoas anotadas.

Número de Pessoas	Taxa de Acerto	
	Servidor	Móvel
10	93.1%	75.1%
20	89.0%	74.3%
30	86.3%	72.8%
40	83.9%	71.5%
60	82.6%	70.6%
Todos	74.1%	64.4%

Figura 5.2. Gráfico da taxa de acerto dos algoritmos pelo número de pessoas anotadas.



5.3.2 Resultados da Ponderação na Anotação de Pessoas

Na Tabela 5.3, na Tabela 5.4 e na Tabela 5.5, são apresentadas as ponderações obtidas para os estimadores para as quatro coleções de fotografias utilizadas no

experimento para as técnicas de ponderação de regressão linear, inclinação e correlação, respectivamente.

Conforme pode ser observado na Tabela 5.3, o estimador de recorrência global (RG) foi descartado para todas as coleções utilizando a técnica de regressão linear para a ponderação dos estimadores. As ponderações das coleções 1 e 2 obtiveram os menores desvios padrões, sendo de 0,1 e a coleção 3 obteve o maior desvio, sendo de 0,23. Estes desvios demonstram a clara diferença existente entre as coleções utilizadas no processo de avaliação, das quais, a coleção 3 destaca-se pela alta razão entre o número de anotações de pessoas e a quantidade de fotografias, sendo o estimador de co-ocorrência o de maior peso. Os desvios das ponderações por estimador não foram tão significativas quanto o por coleção, sendo o de maior valor o de recorrência geográfica, que obteve 0,11 de desvio padrão.

Tabela 5.3. Ponderações dos estimadores calculados utilizando regressão linear.

Coleção	RG	RGeo	RT	CO	Média	Desvio Padrão
1	X	0,31	0,49	0,47	0,34	0,10
2	X	0,37	0,29	0,49	0,31	0,10
3	X	0,17	0,30	0,61	0,30	0,23
4	X	0,14	0,38	0,44	0,25	0,16
Média	X	0,25	0,37	0,5		
Desvio Padrão	x	0,11	0,09	0,08		

Na Tabela 5.4, observa-se que o estimador de recorrência global possui o menor peso para todas as coleções e possui a menor média no valor de 0,15. Da mesma forma, na Tabela 5.5, observa-se que o estimador de recorrência global, apesar de possuir menor

peso em todas as coleções, não foi descartado, pois as técnicas de ponderação utilizando correlação e inclinação não eliminam estimadores que não alteram no resultado.

Ao contrário da técnica de ponderação por regressão linear, os desvios padrões das ponderações por coleção e por estimador não obtiveram grandes valores, sendo mais uma vez o de maior valor, o desvio das ponderações da coleção 3. Os desvios das ponderações nas coleções e nos estimadores para a técnica de ponderação utilizando inclinação foram maiores que os obtidos utilizando correlação. Dessa forma, observa-se que a técnica de inclinação adota ponderações mais variadas que a de correlação.

Tabela 5.4. Ponderações dos estimadores calculados utilizando inclinação.

Coleção	RG	RGeo	RT	CO	Média	Desvio Padrão
1	0,17	0,33	0,41	0,39	0,33	0,11
2	0,17	0,30	0,30	0,32	0,27	0,07
3	0,12	0,27	0,30	0,54	0,31	0,17
4	0,14	0,19	0,22	0,25	0,20	0,05
Média	0,15	0,27	0,31	0,38		
Desvio Padrão	0,02	0,06	0,08	0,12		

Na Figura 5.3, na Figura 5.4, na Figura 5.5 e na Figura 5.6, são mostrados os resultados obtidos no experimento para a anotação de todos os contatos das coleções 1, 2, 3 e 4, respectivamente. Os gráficos apresentados nessa seção possuem no eixo x o valor de H , ou seja, o número de sugestões da lista, e no eixo y a taxa de acerto.

Tabela 5.5. Ponderações dos estimadores calculados utilizando correlação.

Coleção	RG	RGeo	RT	CO	Média	Desvio Padrão
1	0,41	0,54	0,59	0,54	0,52	0,08
2	0,35	0,52	0,52	0,51	0,48	0,08
3	0,33	0,46	0,47	0,62	0,47	0,12
4	0,32	0,35	0,41	0,42	0,38	0,05
Média	0,35	0,47	0,5	0,52		
Desvio Padrão	0,04	0,09	0,08	0,08		

A coleção 3 (Figura 5.5) obteve os piores resultados dentre as coleções para todos os estimadores e valores de H. Analisando os estimadores isoladamente sem combinação, RGeo foi o melhor estimador nas coleções 2 (Figura 5.4) e 3, sendo o de RT, RG e CO os seguintes nessa ordem.

Figura 5.3. Resultados para a classificação de todos os contatos da coleção 1.

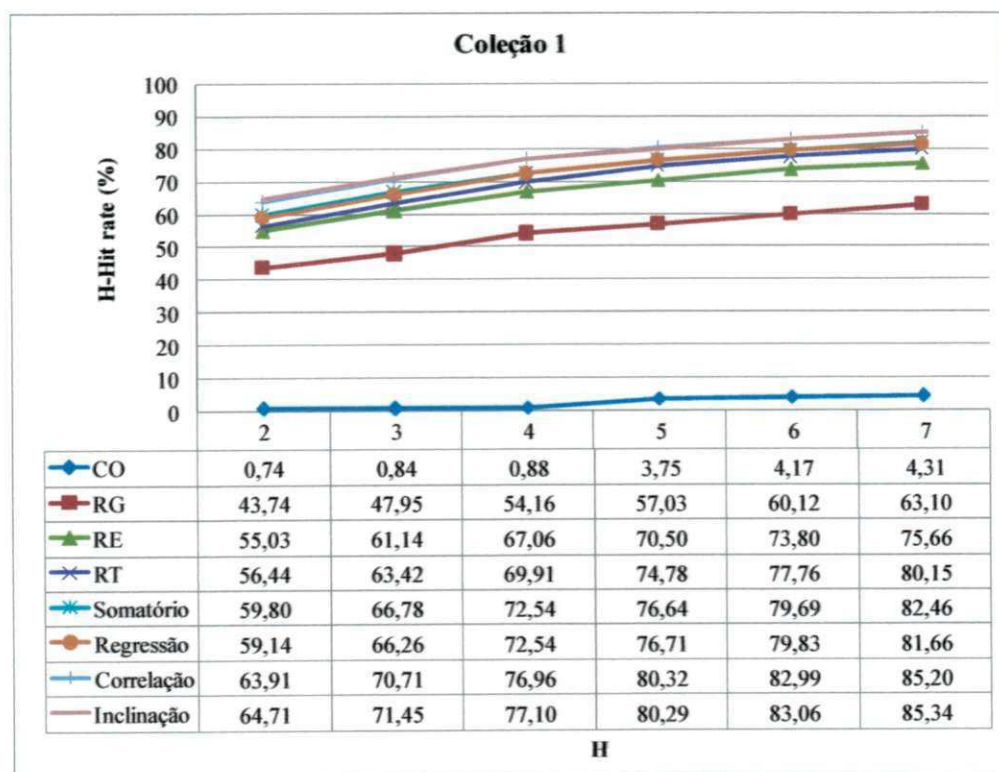


Figura 5.4. Resultados para a classificação de todos os contatos da coleção 2.

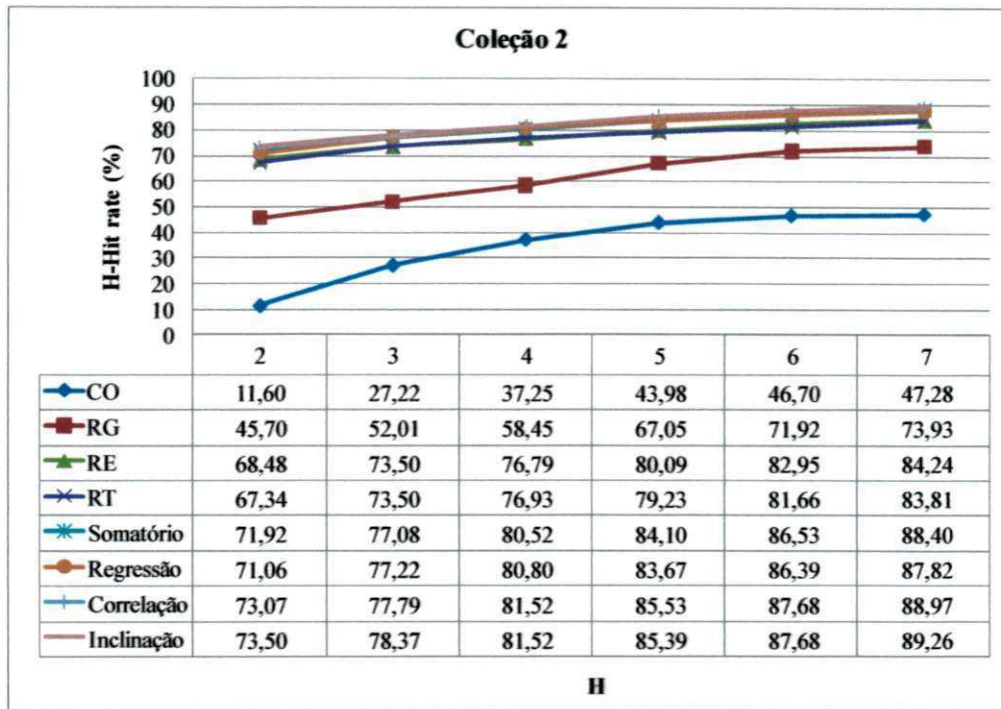
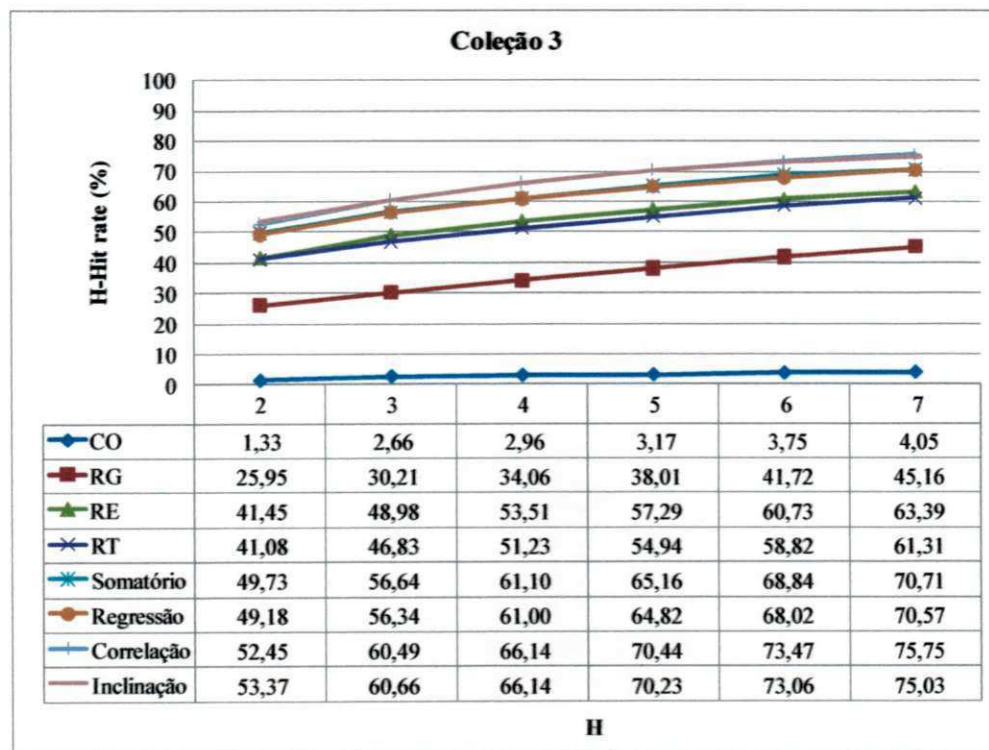
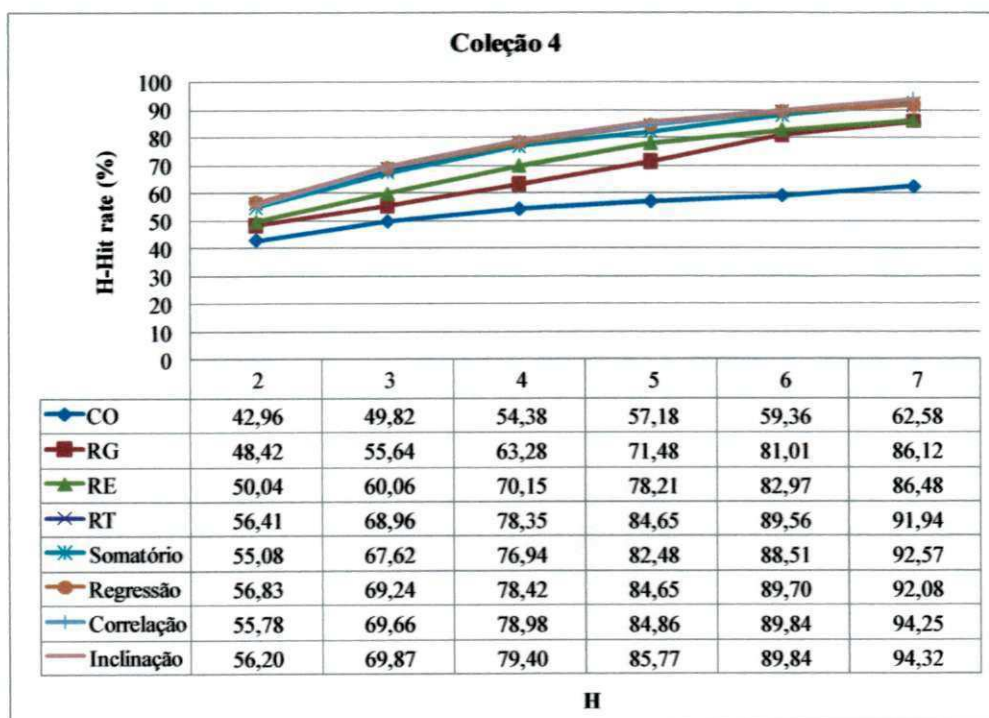


Figura 5.5. Resultados para a classificação de todos os contatos da coleção 3.



Para as coleções 1 (Figura 5.3) e 4, o estimador RT teve os melhores resultados, acompanhado em ordem dos seguintes estimadores RGeo, RG e CO. Dessa forma, pode-se afirmar que cada coleção tem comportamentos distintos e podem possuir diferentes ponderações para cada estimador.

Figura 5.6. Resultados para a classificação de todos os contatos da coleção 4.



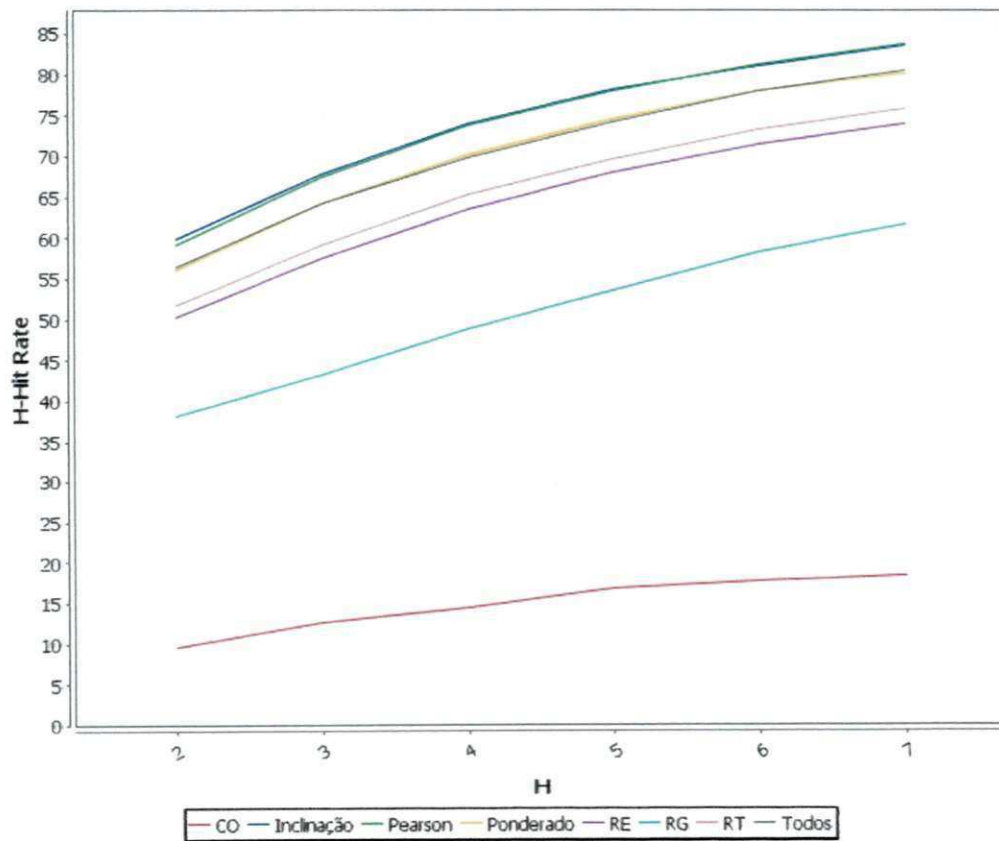
Na Figura 5.7, apresentam-se os resultados para a anotação de todos os contatos de todas as coleções juntas. A variação do tamanho da lista de sugestões (H) afeta a taxa de acerto em aproximadamente 5% entre o tamanho 2 e 3. Porém, essa variação diminui para aproximadamente 1% quando H varia de 6 para 7. Dessa forma, a variação da taxa de acerto diminui consideravelmente a partir de 5 sugestões, que é o tamanho da lista de sugestões adotado no PhotoGeo.

A taxa de acerto para o somatório simples dos estimadores teve uma média de 74% para H igual a 5, enquanto a taxa de acerto para o somatório ponderado utilizando regressão linear obteve uma média de 75%, sendo que para algumas coleções o somatório simples obteve melhores resultados. Deste modo, a ponderação pelo método de regressão

linear obteve um desempenho aproximadamente igual à combinação não ponderada, porém, usando um estimador a menos.

Em contrapartida, o somatório ponderado utilizando o método de inclinação obteve os melhores resultados em todas as coleções para todos os valores de H , com exceção da coleção 3 com o H igual a 6 e 7, que a técnica de correlação obteve melhores resultados. A taxa de acerto usando inclinação aumentou aproximadamente 4% com relação à combinação simples e ponderada empregando regressão linear.

Figura 5.7. Resultados para a classificação de todos os contatos de todas as coleções.



Para H igual a 5, a taxa de acerto utilizando inclinação foi de aproximadamente 79%, chegando a obter uma taxa de 85% para a coleção 2 e 4 para H igual a 5. A pior taxa de acerto para este valor de H usando inclinação foi para a coleção 3 que teve 70%, porém, foi nesta coleção que foi observado o maior crescimento da taxa de acerto com

relação à combinação simples e à ponderada empregando regressão linear, que foi de aproximadamente 5% de aumento.

Com o objetivo de avaliar o efeito no sistema da anotação de uma quantidade menor de contatos diferentes, foram realizados experimentos anotando somente os 20 contatos mais populares de cada coleção.

Nas Figuras 5.8, 5.9, 5.10 e 5.11, são mostrados os resultados obtidos no experimento para a anotação dos 20 contatos mais populares de cada coleção.

A coleção 4 (Figura 5.11) não teve alterações, pois possui apenas 18 contatos diferentes anotados. A coleção 3 (Figura 5.10), que obteve os piores resultados no experimento com todos os contatos, teve um crescimento de 15% em média. A coleção 2 (Figura 5.9) obteve os melhores resultados neste experimento, obtendo um taxa de acerto de aproximadamente 91% para H igual a 5. A melhor taxa de acerto obtido em todos os experimentos foi o da coleção 4 com o H igual a 7, que foi de aproximadamente 95%.

Figura 5.8. Resultados para a classificação dos 20 contatos mais populares da coleção 1.

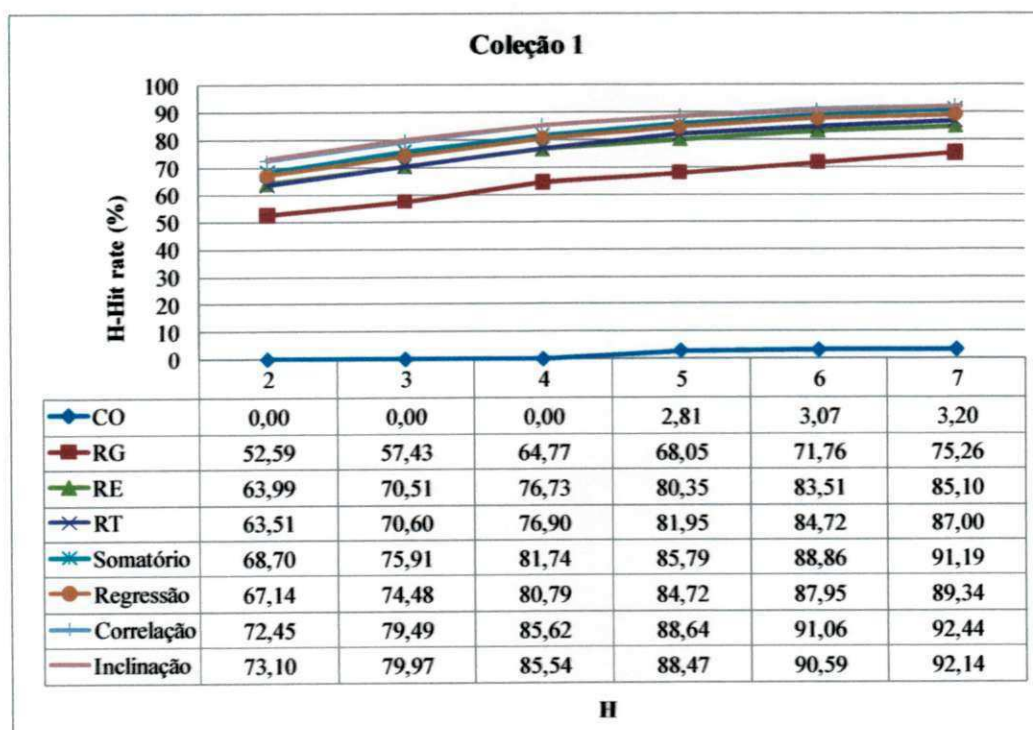


Figura 5.9. Resultados para a classificação dos 20 contatos mais populares da coleção 2.

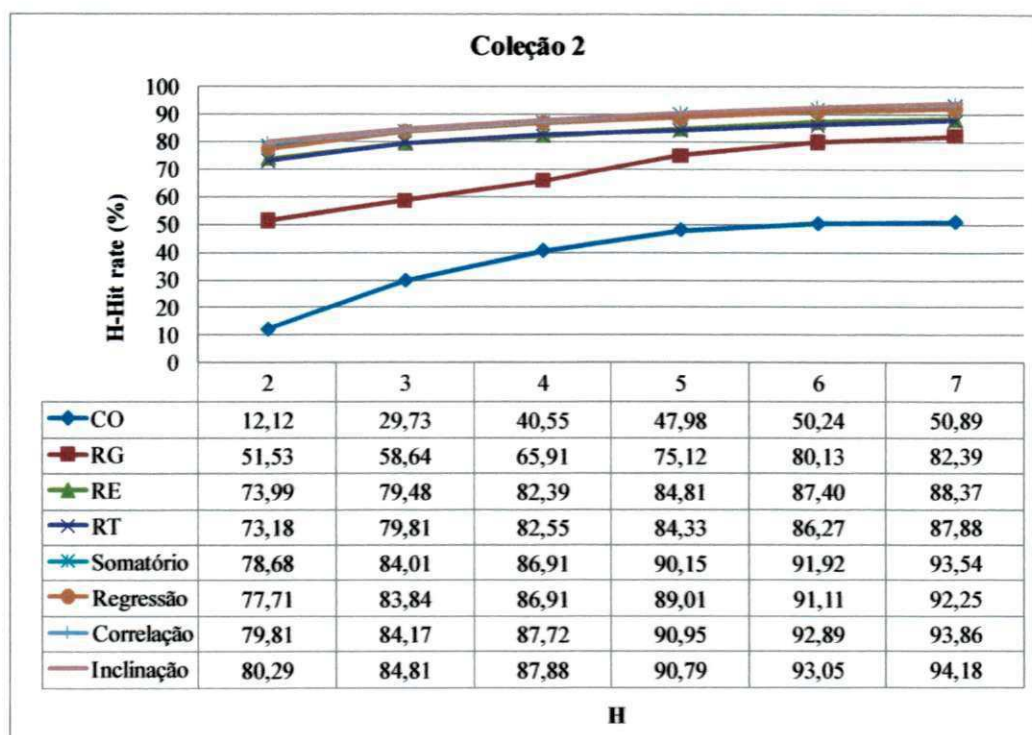


Figura 5.10. Resultados para a classificação dos 20 contatos mais populares da coleção 3.

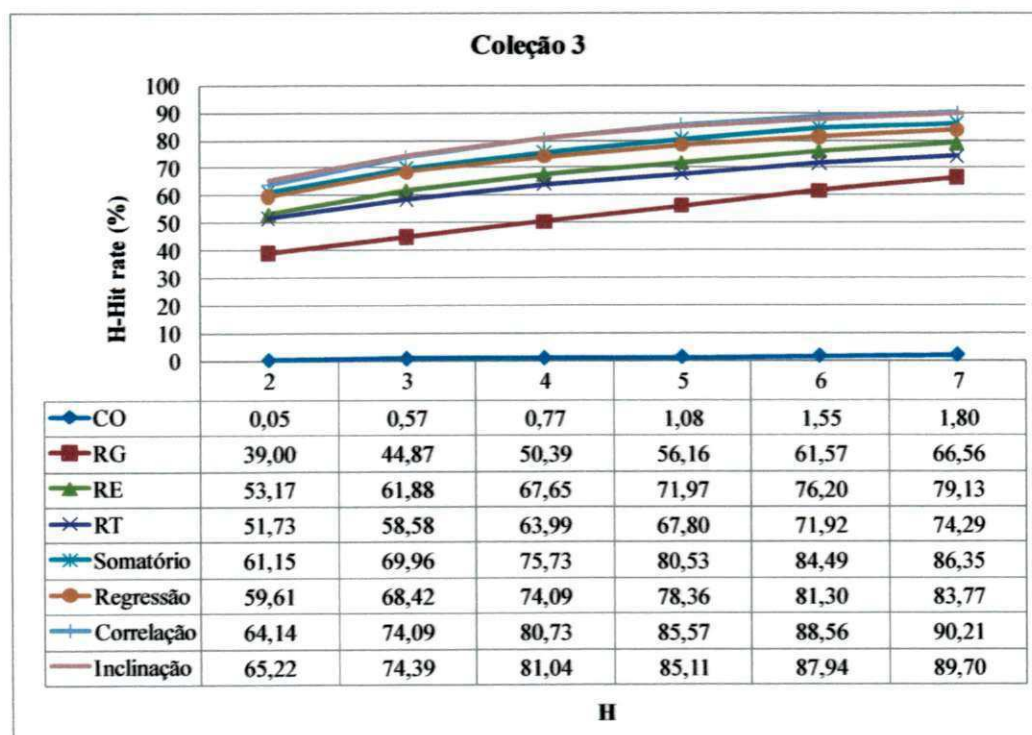
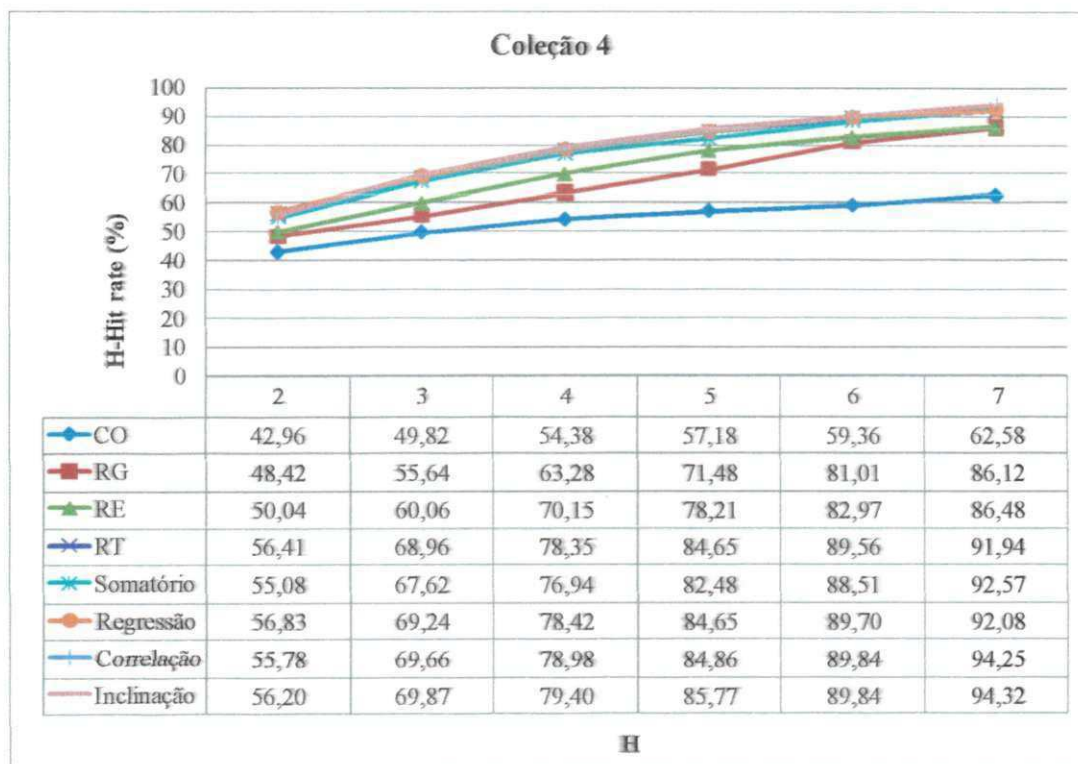


Figura 5.11. Resultados para a classificação dos 20 contatos mais populares da coleção 4.



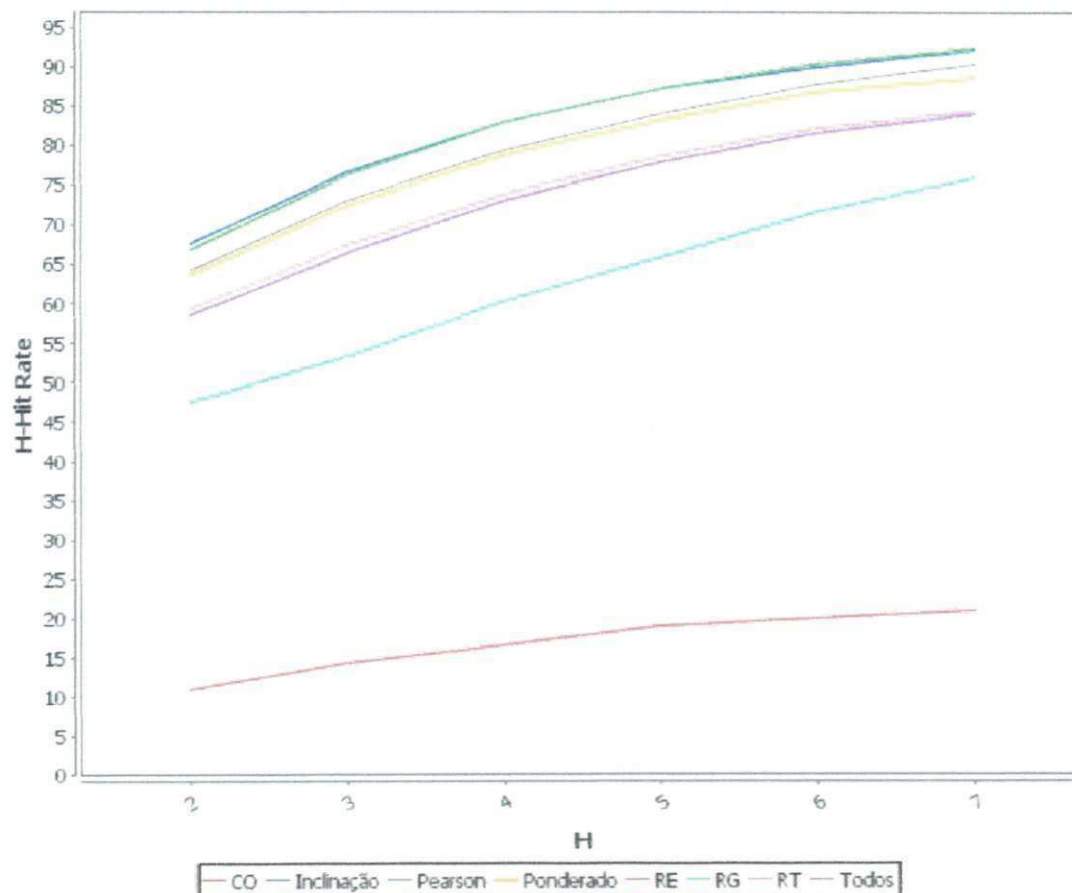
Na Figura 5.12, são mostrados os resultados obtidos no experimento para a anotação dos 20 contatos mais populares de cada coleção para toda a coleção. A taxa de acerto em comparação ao experimento que anotou todos os contatos aumentou em média 9%. A diferença de taxa de acerto entre os estimadores diminuiu devido ao maior aumento dos estimadores isolados sem combinação com relação aos combinados.

Anotando somente os 20 contatos mais populares, a combinação simples obteve resultados aproximados aos da combinação utilizando somatório ponderado pelo método de regressão linear, sendo os do sem ponderação levemente melhores.

Para a combinação usando somatório ponderado pelo método de inclinação, anotando somente os 20 contatos mais populares, obteve mais uma vez os melhores resultados para todos os usuários e todos os valores de H analisados.

O ranking dos contatos para a lista de sugestões pode ser baseado em um estimador. Entretanto, a combinação utilizando o somatório e o somatório ponderado dos estimadores obteve melhores resultados do que os estimadores isoladamente.

Figura 5.12. Resultados para a classificação dos 20 contatos mais populares de cada coleção.



Os resultados obtidos com a combinação de estimadores usando o somatório ponderado a partir do método de regressão linear foram equivalentes ao do somatório simples. Entretanto, o ponderado teve a vantagem de eliminar a utilização de um estimador – recorrência global – diminuindo o tempo de processamento. Dessa forma, esta técnica de ponderação é interessante para avaliar quais estimadores devem fazer parte da composição do valor que será empregado no ranking dos contatos para a lista de sugestões. O uso desta técnica é útil quando existem muitos estimadores a serem considerados.

Em contrapartida, a combinação de estimadores empregando um somatório ponderado utilizando o método de inclinação obteve os melhores resultados utilizando a métrica de *H-Hit Rate*. Porém, este método não realiza uma seleção de estimadores, assim, este método não é recomendado para ser usado com uma grande quantidade de estimadores.

Na Figura 5.13, apresenta-se o gráfico com a mediana dos resultados das taxas de acerto do algoritmo de anotação de pessoas presentes nas fotografias versus o tamanho da lista de sugestões, para cada um dos estimadores e métodos de combinação ora descritos. Na Tabela 5.6, são apresentados os desvios médios das taxas de acertos dos estimadores e métodos de combinação de estimadores para 6 tamanhos de lista de sugestões. Observa-se que o método de combinação de estimadores que utiliza a função inclinação obteve mediana superior para todos os tamanhos de lista de sugestões.

Figura 5.13. Gráfico da mediana das taxas de acerto do algoritmo de anotação de pessoas presentes nas fotografias com desvio médio.

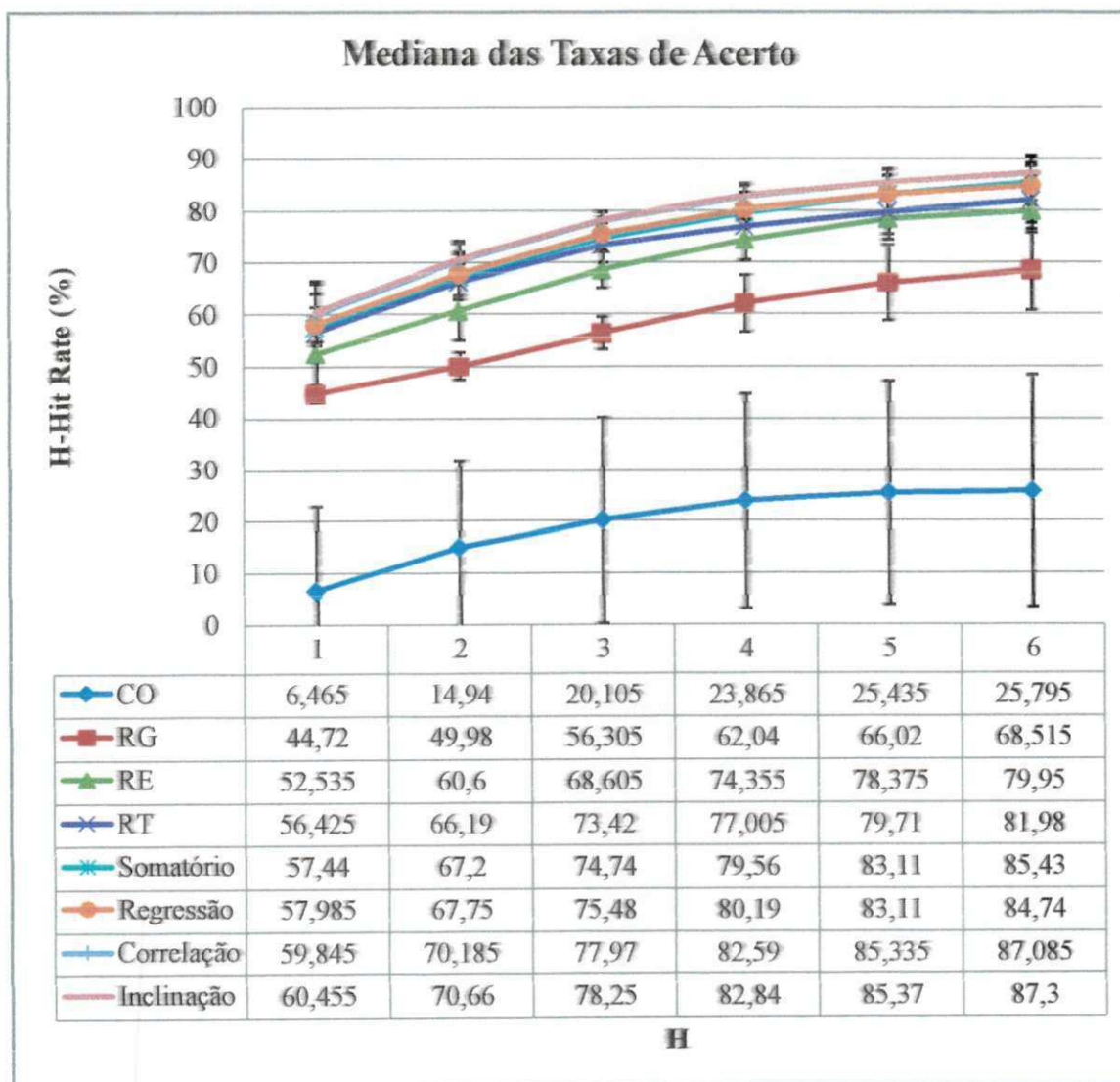


Tabela 5.6. Desvios médios das taxas de acertos dos tamanhos da lista de sugestões para cada estimador e método de fusão.

	1	2	3	4	5	6
CO	16,35111	16,74667	19,97111	20,81333	21,71556	22,49778
RG	1,644444	2,611111	3,1	5,437778	7,264444	7,824444
RE	7,086667	5,733333	3,637778	3,844444	4,071111	4,311111
RT	4,851111	3,471111	3,435556	3,397778	4,377778	4,426667
Somatório	6,435556	4,391111	2,751111	2,955556	3,48	3,566667
Regressão	5,811111	4,208889	3,142222	3,311111	3,651111	3,684444
Correlação	5,877778	3,38	1,577778	2,166667	2,564444	3,184444
Inclinação	5,797778	3,426667	1,493333	2,351111	2,533333	3,12

5.5.3 Resultados da Anotação de Pessoas Utilizando Contexto e Conteúdo

Nesta seção, são apresentados os resultados obtidos com os experimentos realizados com os algoritmos de anotação de pessoas usando somente informações de contexto (EC), apenas análise de conteúdo (RF) e contexto mais conteúdo (RF+EC).

Análises estatísticas foram realizadas para tentar aceitar ou refutar as hipóteses levantadas para responder a questão de pesquisa Q1.

Na Tabela 5.7, os resultados obtidos pelos experimentos são apresentados. Como pode ser observado, o algoritmo que utiliza RF+EC foi o que obteve melhores resultados.

Tabela 5.7. Dados Coletados

Coleção	Algoritmo		
	RF	EC	RF e EC
1	24,51 (%)	74,57	80,20
2	39,02	84,22	84,79
3	71,09	79,31	86,99
4	24,24	64,94	71,32

Após a comparação dos algoritmos, a hipótese HB_0 foi refutada, concluindo-se que o algoritmo que analisa as informações de contexto e de conteúdo para gerar listas de

sugestões para a anotação de pessoa em fotografias produz uma melhor taxa de acerto. Assim, para a questão de pesquisa $Q1$, que questiona se o algoritmo que analisa informações de contexto e de conteúdo para sugerir pessoas presentes em fotografias possui uma melhor taxa de acerto em relação aos algoritmos que analisam somente informações de contexto ou somente informações de conteúdo, a resposta é **sim**. O restante da análise estatística desse experimento encontra-se no Apêndice A.

5.6 Considerações Finais

Neste capítulo, foi descrita a solução para a anotação de pessoas, foi proposto um algoritmo que realiza a anotação semiautomática de pessoas presentes nas fotografias. No referido algoritmo, o usuário é auxiliado na anotação por uma lista de sugestões com as pessoas mais prováveis de estarem presentes nas fotografias. Essa lista é gerada a partir da ordenação decrescente do valor obtido a partir do somatório ponderado de estimadores, que utilizam informações contextuais e de conteúdo para obtenção de suas estimativas.

Capítulo 6 Detecção de Eventos Pessoais e Compartilhados

No escopo desta tese, um **evento** é caracterizado por um acontecimento que reúne pessoas com um objetivo em comum, em um espaço geográfico e durante um determinado período de tempo. Exemplos de eventos incluem festas de aniversário, concertos, casamentos e natal. Dessa forma, um evento pode ser caracterizado por um conjunto de informações contextuais de um determinado acontecimento acerca de **quem, onde e quando**. Essas informações são consideradas as mais importantes para um usuário recordar de suas fotografias (WAGENAAR, 1986), tornando a detecção do evento no qual uma fotografia foi capturada, **uma informação muito importante** para o gerenciamento de fotografias.

Para a anotação de eventos nas fotografias, propõe-se um algoritmo para detecção de eventos nas fotografias de forma automática, mas configurável. O algoritmo proposto para detecção de eventos utilizará informações contextuais adquiridas dos metadados das fotografias do usuário e da sua rede social.

O evento possui características hierárquicas, de forma que podem existir subeventos contidos em um evento maior. Desta forma, propõe-se a utilização de **parâmetros** de entrada para o algoritmo de detecção de eventos que poderão ser **configurados** de acordo com as preferências dos usuários. No escopo desta tese, utiliza-se o termo **granularidade** para definir o conjunto de parâmetros definidos pelo usuário que determinará a magnitude dos eventos detectados automaticamente.

O algoritmo proposto poderá classificar de maneira errônea algumas fotografias pertencentes a um evento ou criar eventos inexistentes. Caso venha a ocorrer **alguma** destas inconsistências, o usuário mediante a utilização de uma ferramenta poderá corrigir manualmente esses problemas, ou seja, o usuário tem a possibilidade de **criação** de eventos de forma manual.

Outro fator importante que será abordado neste capítulo é a detecção de eventos compartilhados. Este termo irá ser utilizado para definir grupos de usuários que capturam fotografias de um mesmo evento para fins de registro. Evento compartilhado é uma situação que ocorre devido à corriqueira existência de várias câmeras de pessoas distintas em um mesmo evento. Em geral, cada pessoa possuidora de uma câmera fotográfica possui sua organização particular de fotografias. Como consequência do fato de câmeras distintas possuírem funções distintas, é possível que fotografias capturadas num mesmo evento possuam tipos distintos de informações nos metadados. Por exemplo, as fotografias capturadas a partir da câmera de um smartphone, equipada com GPS, possuem a informação de localização geográfica nos seus metadados, enquanto em outra câmera, equipada com bússola, as fotografias capturadas possuem apenas as informações de orientação.

A rede de contatos de um usuário é utilizada para a descoberta de conjuntos de fotografias de um mesmo evento, para posterior identificação de que as fotografias fazem parte de um mesmo evento, porém, de usuários diferentes. Para alcançar este objetivo, propõe-se uma função que calcule a similaridade entre eventos para indicar o quanto dois eventos parecem fazer parte de um único evento.

Pretende-se, com a detecção de eventos compartilhados, que outras anotações possam ocorrer de forma automática ou auxiliar os algoritmos de anotação semiautomática. Por exemplo, Maria foi anotada em um evento como presente e em outro evento essa anotação não constava, sendo essa anotação realizada automaticamente para o segundo evento caso os eventos sejam similares. Outra contribuição dos eventos compartilhados está relacionada à anotação semiautomática de pessoas presentes nas fotografias, pois existem estimadores, como o de recorrência no evento, que melhoram o desempenho com um número maior de fotografias.

A solução proposta neste capítulo para a anotação de eventos em fotografias é dividida em dois escopos distintos:

- **Detecção de eventos pessoais:** separação em eventos de coleções de fotografias pessoais, sem a existência de interseção de tempo entre eventos distintos, ou seja, dois eventos não podem ocorrer em um mesmo intervalo de tempo; e

- **Detecção de eventos compartilhados:** descoberta da existência de fotografias de um mesmo evento em coleções de fotografias de donos distintos, ou seja, mais de uma pessoa capturou fotografias do evento e armazenou em sua coleção pessoal de fotografias.

6.1 Detecção de Eventos Pessoais

A detecção de eventos nas fotografias é realizada assumindo que não existem eventos paralelos na coleção de um mesmo usuário. As fotografias devem conter nos seus metadados as informações de data e hora do momento da captura, sendo a informação da localização geográfica utilizada caso exista.

O algoritmo para a anotação de eventos em fotografias proposto nesta tese estende aquele proposto por Lacerda *et al.* (2008b), sendo ambos baseados em quatro etapas. Entretanto, para este último, todas as etapas são obrigatórias, enquanto a extensão realizada pelo algoritmo proposto neste trabalho permite que a primeira e a terceira etapa não sejam realizadas, caso a informação de localização geográfica não exista nas fotografias ou se o usuário não possuir agenda Web. Outro ponto importante é a possibilidade de alterar o nível de granularidade do evento gerado, permitindo que, com a mesma coleção de fotografias, sejam gerados diferentes eventos, sem necessariamente ocorrer em erros.

A entrada do algoritmo de detecção de eventos é uma coleção de fotografias. Na primeira etapa, caso exista a informação de localização geográfica, realiza-se uma classificação geográfica das fotografias. Na segunda etapa, realiza-se uma classificação temporal. Na terceira etapa, são realizadas algumas correções de inconsistências ocorridas após as duas primeiras etapas. Por fim, na quarta etapa, realiza-se a nomeação dos eventos detectados.

Se existirem inconsistências ao final da detecção de eventos, o usuário poderá realizar alterações manualmente, movendo fotografias de um evento para outro, criando e excluindo eventos.

A primeira etapa separa as fotografias em grupos baseado na proximidade geográfica destas, utilizando métodos de agrupamento, sendo o algoritmo DBSCAN

(ESTER *et al.*, 1996) selecionado para esta etapa por ser um método baseado em densidade e possuir os parâmetros de configuração fáceis de serem estimados.

Alguns parâmetros devem ser informados quando uma classificação não-supervisionada é utilizada, tais como: o valor de *épsilon* a distância mínima entre os membros de um mesmo grupo – e o número mínimo de elementos contidos em um grupo no DBSCAN.

Dependendo dos parâmetros de entrada para os algoritmos de agrupamento, podem-se obter diferentes resultados de diferentes granularidades, que resultarão em grupos com diferentes tamanhos. Assim, estes parâmetros podem ser utilizados para personalizar as coleções de fotografias, de acordo com as preferências do usuário.

Como ao final da primeira etapa ainda vão existir muitas inconsistências devido à utilização somente da informação geográfica, na segunda etapa, realiza-se uma classificação temporal com o objetivo de corrigir algumas dessas inconsistências.

A segunda etapa recebe como entrada grupos de fotografias e o parâmetro *t*. O parâmetro *t* é utilizado para modificar o nível de granularidade dos eventos gerados.

Como saída, são retornados grupos de fotografias separados em eventos. Nessa etapa, utiliza-se a teoria do comportamento explosivo no momento de captura de fotografias (COOPER *et al.*, 2005), relatado anteriormente, e as informações de data e hora das fotografias para a realização da classificação.

Na terceira etapa, são extraídas informações externas com o objetivo de corrigir algumas inconsistências remanescentes após as duas primeiras etapas. As primeiras informações a serem utilizadas são extraídas do Wikimapia²⁰, que são os lugares próximos ao evento. Se as fotografias do evento não possuírem a informação de localização geográfica, as primeiras informações não são extraídas. As outras informações utilizadas nesta etapa são adquiridas da agenda do usuário.

Correções são realizadas na terceira etapa nos eventos gerados pelas etapas anteriores para a ocorrência de três tipos de situações. Na primeira situação, detecta-se que

²⁰ <http://wikimapia.org/>. Acessado em Novembro de 2013.

um evento foi realizado em mais de um lugar, porém, as fotografias capturadas em cada um dos lugares são consecutivas no tempo. Para a primeira situação, o evento é dividido em um evento para cada um dos lugares detectados. A segunda situação ocorre quando se detecta, na agenda Web do usuário, que existem eventos diferentes cadastrados durante o intervalo de tempo de um evento detectado pelo algoritmo. Para a segunda situação, o evento é dividido em eventos equivalentes àqueles cadastrados na agenda do usuário. Na terceira situação, detecta-se que durante o intervalo de tempo de um evento na agenda do usuário, foi detectado mais de um evento pelo algoritmo. Para a terceira situação, os eventos detectados são unificados em um único evento.

Na quarta etapa, realiza-se uma nomeação dos eventos detectados nas etapas anteriores, utilizando-se o seguinte formato: “[Nome do evento na agenda do usuário] + [Nome do lugar] + [Nome da cidade] + [nome do estado] + (Nome do país) + (Data inicial) + (Duração do evento)”, sendo os campos entre colchetes opcionais e os entre parênteses obrigatórios. O nome do evento é extraído da agenda do usuário. O nome do lugar é extraído do Wikimapia. Os nomes da cidade, estado e país são obtidos do *gazetteer* GeoNames²¹.

6.1.1 Granularidade

O sistema proposto utiliza parâmetros de entrada para personalizar o algoritmo de detecção de eventos. O algoritmo DBSCAN requer os valores de *épsilon* e o número mínimo de elementos num dado cluster como parâmetros de entrada. Já o algoritmo de classificação temporal requer o valor de *t*. A partir da mudança desses parâmetros de entrada separadamente, os usuários podem personalizar a organização das coleções, de forma que haverá vários eventos em diferentes níveis de granularidade. Para facilitar o uso do sistema, estes parâmetros de entrada têm valores padrões, e o usuário poderá mudá-los a sua escolha para obter uma melhor personalização dos eventos detectados de sua coleção de fotografias, de acordo com sua necessidade.

Na Tabela 6.1, são apresentados estes parâmetros de entrada, usando como exemplo cinco níveis de granularidade pré-definidos. O nível I contém os valores padrões de cada parâmetro. Como pode ser visto, a granularidade temporal – representada pelo

²¹ <http://www.geonames.org/>. Acessado em Outubro de 2011.

parâmetro de entrada t – utiliza a unidade de horas. A granularidade geográfica do parâmetro de entrada ϵ utiliza Km, e o valor de k é percentagem (%). Estes valores são internamente convertidos para preencher os requisitos dos respectivos algoritmos. O parâmetro ϵ no DBSCAN representa a distância máxima que um elemento estará do próximo grupo de elementos.

Tabela 6.1. Níveis de granularidade pré-definidos.

Granularity Level	<i>Temporal</i>	<i>DBSCAN</i>		<i>K-means</i>	
	<i>t</i> (hours)	<i>epsilon</i> (Km)	<i>min number of elements</i>	<i>K</i> (%)	<i>Seed</i>
1	6	1	1	5	7
2	12	50	1	8	7
3	24	100	1	12	7
4	48	250	1	15	7
5	72	500	1	18	7

Os níveis de granularidade de cada parâmetro podem ser seleccionados separadamente, por exemplo, podendo seleccionar DBSCAN no nível 1 e o parâmetro temporal no nível 3.

6.2 Detecção de Eventos Compartilhados

Apesar do algoritmo de detecção de eventos pessoais obter bons resultados para coleções de fotografias pessoais, mostrou-se ineficaz para eventos compartilhados. O principal motivo dos resultados ruins é a existência de eventos que ocorrem ao mesmo tempo.

A detecção de eventos pessoais pode ser o primeiro passo a ser realizado para a detecção de eventos compartilhados em uma rede social, pois a detecção de eventos compartilhados ocorre a partir da comparação entre eventos pessoais.

A detecção de eventos compartilhados pode ser utilizada para as seguintes tarefas: anotação de outras informações por meio de uma anotação cruzada, detecção de erros nas anotações e novas experiências na navegação e busca de fotografias.

As seguintes etapas são realizadas para a detecção de eventos compartilhados:

1. Detecção de eventos pessoais nas fotografias de cada usuário;
2. Cálculo da similaridade de eventos φ (discutida posteriormente) para cada par de eventos $ev_u \in E_u$ e $ev_{U_u} \in E_{U_u}$, sendo E_u o conjunto de eventos do usuário $u \in U$ e E_{U_u} o conjunto de eventos dos usuários que são contatos de u ; e
3. Anotação de evento compartilhado quando a função φ retornar um valor superior a um determinado valor de limiar θ para um par de eventos $\varphi(ev_u, ev_{U_u}) > \theta$.

O cálculo da similaridade entre eventos será realizado por meio da função $\varphi: V \rightarrow S$ que medirá o grau de similaridade entre dois eventos ev_1 e ev_2 , sendo $V = \{(ev_1, ev_2) | ev_1 \in E_1 \text{ e } ev_2 \in E_2\}$, $S = \{s | s \in \mathcal{R} \text{ e } 0 \leq s \leq 1\}$ e E_u o conjunto de eventos do usuário u . A função $\varphi(ev_1, ev_2)$ é utilizada pelo algoritmo de detecção de eventos compartilhados e utiliza as informações de contexto C_{E_u} do conjunto de fotografias $Y_{E_u} \subseteq F$ capturadas no evento E_u . A função $\varphi(ev_1, ev_2)$ é definida pela Equação 8.

$$\varphi(ev_1, ev_2) = SG * PSG + ST * PST + SP * PSP \quad (8)$$

em que, as funções SG , ST e SP medem o grau de similaridade geográfica, temporal e de pessoas, respectivamente, entre os eventos ev_1 e ev_2 . As variáveis PSG , PST e PSP são os pesos dos estimadores de similaridade geográfica, temporal e de pessoas, respectivamente.

6.2.1 Similaridade Geográfica (SG)

Na função de similaridade geográfica de eventos, calcula-se a percentagem da área de intersecção das áreas dos dois eventos, com relação à soma das áreas dos dois eventos. Para isso são realizados os seguintes passos:

- I. Realiza-se uma agregação das geometrias das fotografias de um mesmo evento (função *extent* do OpenGIS (OGC, 1999)), gerando-se uma geometria que representa o raio de abrangência dos eventos a serem comparados;

2. Gera-se um *buffer* de d metros a partir da geometria gerada pela agregação das fotografias dos eventos;
3. Verifica-se a existência de alguma interseção entre as geometrias dos dois eventos;
4. Caso não exista interseção, retorna-se uma similaridade nula;
5. Caso exista uma interseção, calcula-se a área desta interseção;
6. Em seguida, calcula-se a área da geometria de cada evento;
7. Por fim, retorna-se o resultado da aplicação da Equação 9 para o cálculo da similaridade geográfica.

$$SG = \frac{2 \times \text{Área}(\text{Interseção}(ev_1, ev_2))}{\text{Área}(ev_1) + \text{Área}(ev_2)} \quad (9)$$

6.2.2 Similaridade Temporal (ST)

Na função de similaridade temporal de eventos, calcula-se a porcentagem de interseção entre os intervalos de tempo em que ocorrem dois eventos sobre o tempo total dos dois, considerando um intervalo adicional t no início e fim de cada um dos eventos. Para tanto, os seguintes passos são realizados:

1. Identifica-se o início ti_{ev_1} do evento ev_1 , a partir das informações temporais da primeira fotografia capturada entre os dois eventos. A coleção que contiver a fotografia mais antiga será associada ao evento ev_1 ;
2. Identifica-se o final tf_{ev_1} de ev_1 , verificando qual a fotografia mais recente da coleção Y_{ev_1} ;
3. Identifica-se o início ti_{ev_2} de ev_2 , verificando qual a fotografia mais antiga da coleção Y_{ev_2} ;
4. Identifica-se o final tf_{ev_2} de ev_2 , verificando qual a fotografia mais recente da coleção Y_{ev_2} ;
5. Como a primeira fotografia capturada em um evento pode não ter ocorrido no início e a última pode não ter ocorrido no final, consideram-se os seguintes intervalos de tempo $[ti_{ev_1} - \Delta t, tf_{ev_1} + \Delta t]$ e $[ti_{ev_2} - \Delta t, tf_{ev_2} + \Delta t]$,

6. Verifica-se o tempo inicial ti_{ev_c} e final tf_{ev_c} do intervalo em comum entre ev_1 e ev_2 ;
7. Calcula-se ST por meio da Equação 10.

$$ST = \frac{2 \times \text{Intervalo}(ti_{ev_c}, tf_{ev_c})}{\text{Intervalo}(ti_{ev_1}, tf_{ev_1}) + \text{Intervalo}(ti_{ev_2}, tf_{ev_2})} \quad (10)$$

6.2.3 Similaridade de Pessoas (SP)

Na função de similaridade de pessoas em eventos, calcula-se a percentagem de pessoas em comum entre dois eventos sobre o total de pessoas. Para isso, os seguintes passos são realizados:

1. Calcula-se a quantidade de pessoas distintas anotadas $|P_{ev_1}|$ nas fotografias do evento ev_1 ;
2. Calcula-se a quantidade de pessoas distintas anotadas $|P_{ev_2}|$ nas fotografias do evento ev_2 ;
3. Calcula-se a quantidade de pessoas distintas anotadas $|P_{ev_1} \cap P_{ev_2}|$ nas fotografias do evento ev_1 que também foram anotadas nas fotografias do evento ev_2 ;
4. Calcula-se a SP mediante a Equação 11.

$$SP = \frac{2 \times |P_{ev_1} \cap P_{ev_2}|}{|P_{ev_1}| + |P_{ev_2}|} \quad (11)$$

6.2.4 Observações

As funções de similaridade de pessoas, geográfica e temporal indicam o quanto dois eventos são similares. Isoladamente, o resultado da similaridade retornada por essas funções pode ocasionar erros que não ocorreriam na detecção de eventos pessoais, e.g., dois eventos que ocorreram no mesmo horário, mas em localizações diferentes. Como em uma coleção pessoal de fotografias não pode haver dois eventos no mesmo instante de tempo, torna-se mais fácil a detecção de eventos. No entanto, tratando-se de fotografias

armazenadas em uma rede social, podem existir diversos eventos que ocorreram num mesmo período de tempo. Deste modo, as informações de pessoas presentes e a localização geográfica tornam-se essenciais para aumentar a eficácia da função de similaridade. Contudo, mesmo que a função SG ou a SP não possa ser calculada (devido à não existência da localização geográfica ou das pessoas presentes nas fotografias), a função φ pode ser calculada utilizando-se as outras duas funções de similaridade.

A agregação das geometrias que representam a localização geográfica das fotografias de um evento realizada na função de similaridade geográfica é realizada por meio do MBR (*Minimum Bounding Rectangle*), que é uma geometria de menor complexidade computacional (retângulo). A desvantagem na utilização desse método de agregação é que, em alguns casos, grandes áreas que não fazem parte do evento podem ser retornadas. Ameniza-se este problema no método de detecção de eventos compartilhados a partir das funções de similaridade temporal e de pessoas.

Para o cálculo da interseção das geometrias geradas pela agregação supracitada na função de similaridade geográfica, utiliza-se a operação topológica *intersection* do OpenGIS (OGC, 1999), que retorna a geometria representando a área de interseção entre duas ou mais geometrias.

A função de similaridade temporal foi proposta baseada no cálculo da proporção de tempo que os dois eventos possuem em comum sobre o todo. Para tanto, considerou-se um intervalo adicional no início e no fim de cada evento, representando o momento do início e do final do evento que não são capturadas fotografias.

6.3 Metodologia de Avaliação

Nesta seção, apresentam-se as metodologias utilizadas para avaliar os métodos propostos para detecção de eventos pessoais e compartilhados.

6.3.1 Metodologia de Avaliação do Algoritmo de Detecção de Eventos Pessoais

Para a realização de experimentos de avaliação dos algoritmos propostos, foi criada uma base de dados com fotografias e anotações do mundo real. Esta base contém quatro coleções de fotografias de pessoas distintas, com um total de 4.050 fotografias. Há 365 contatos diferentes presentes nas fotografias, com um total de 7.912 anotações de

contatos presentes nas fotografias. Todas as fotografias possuem nos metadados (Exif) a data, hora e a localização geográfica. Estas fotografias foram anotadas com as pessoas presentes, sendo estas informações anotadas pelos proprietários das coleções.

Na Tabela 6.2, são apresentadas as estatísticas relacionadas às coleções de fotografias utilizadas no experimento para avaliação da combinação dos estimadores na anotação semiautomática de pessoas em fotografias.

Como as coleções de fotografias supracitadas não possuem a informação do evento no qual foram capturadas, nos experimentos realizados para a detecção de eventos, foram utilizadas 2.412 fotografias georreferenciadas, divididas em cinco coleções, sendo cada coleção de um usuário diferente. Também foram utilizadas as agendas Web dos donos das fotografias.

Tabela 6.2. Estatísticas da coleção de fotografias de utilizadas na avaliação.

Coleção	Fotografias	Pessoas Diferentes	Anotações de
		Anotadas	Pessoas
A	1.644	110	2.851
B	648	51	698
C	898	186	2.936
D	860	18	1.427
Total	4.050	365	7.912

A coleção A possui 596 fotografias, que estão distribuídas aproximadamente em 20 meses. Durante este intervalo de tempo, existem meses com 0 ou até 3 eventos. A coleção B possui 162 fotografias e possui um intervalo de tempo de 11 meses. A Coleção C possui 238 fotografias e tem um intervalo de tempo de 11 meses. A coleção D possui 541 fotografias e possui um intervalo de tempo de 12 meses. Finalmente, a coleção E possui 812 fotografias capturadas durante 30 meses. Todas as coleções possuem fotografias capturadas no Brasil e Canadá. Nas coleções A e B, quase todas as fotografias possuem eventos registrados na agenda Web dos seus respectivos usuários. Na coleção D, somente 25% das fotografias possuem compromissos registrados na agenda, Nas coleções C e E, não existem eventos cadastrados nas agendas.

Para avaliar a eficácia dos níveis de granularidade, comparações foram realizadas entre os eventos anotados manualmente pelos usuários – considerados ideais – e os eventos automaticamente detectados pelo sistema. Para medir o quanto os eventos são similares (detectados manualmente e automaticamente), foram utilizadas métricas clássicas de recuperação de informação: precisão, cobertura e média harmônica.

Para facilitar a descrição dos experimentos realizados, será utilizada a seguinte nomenclatura: a classificação automática de fotografias será chamada de evento e a classificação manual das fotografias será chamada de diretório.

Nesta avaliação, precisão é a porcentagem de fotografias de um dado evento que foram classificadas corretamente. Cobertura é a porcentagem de fotografias em um diretório que foram classificadas no evento correto. Para calcular a precisão e a cobertura, tornou-se necessário a associação de cada diretório a um evento. Para isso, os seguintes procedimentos foram executados: (i) o evento será relacionado ao diretório que possuir a maior cobertura de fotografias; (ii) se existir mais de um evento com a mesma cobertura relacionada a um diretório, o evento com maior precisão será relacionada ao diretório. Assim, é possível que um evento esteja associado com mais de um diretório; mas um diretório deve estar associado com somente um evento. A precisão, cobertura e média harmônica são calculadas de acordo com as Equações 12, 13 e 14, respectivamente:

$$\text{Precisão}(d, e) = \frac{p(d, e)}{p(d, e) + fd(d, e)} \quad (12)$$

$$\text{Cobertura}(d, e) = \frac{p(d, e)}{p(d, e) + fn(d, e)} \quad (13)$$

$$\text{Média - Harmônica}(d, e) = 2 * \frac{\text{Precisão}(d, e) * \text{Cobertura}(d, e)}{\text{Precisão}(d, e) + \text{Cobertura}(d, e)} \quad (14)$$

Onde:

- $p(d,e)$ é o número de fotografias que estão classificadas corretamente do diretório d no evento e ;

- $fp(d,e)$ é o número de fotografias classificadas no evento e que não pertencem ao diretório d ; e
- $fn(d,e)$ é o número de fotografias do diretório d que não são classificadas no evento e .

6.3.2 Metodologia de Avaliação do Algoritmo de Detecção de Eventos Compartilhados

Para avaliar o método proposto para a detecção de eventos compartilhados, foi utilizada uma coleção de 4.251 fotografias, com as seguintes anotações: informações de hora e data da captura da fotografia, localização geográfica e pessoas presentes nas fotografias, sendo essas informações anotadas pelos proprietários das fotografias. Na avaliação, foi considerada uma situação ideal, na qual todas as fotografias estão anotadas de maneira correta e completa, essa situação foi utilizada para que informações incorretas não afetem a avaliação do método proposto.

As fotografias foram separadas a partir da informação do modelo da câmera fotográfica. Fotografias capturadas por câmeras distintas foram consideradas de diferentes coleções para efeito de avaliação.

Após separar as fotografias de acordo com o modelo da câmera, foi executado o algoritmo de detecção de eventos pessoais para cada coleção de fotografias gerada a partir dos modelos de câmeras. Em seguida, foi executado o algoritmo de detecção de eventos compartilhados para cada par de eventos detectados de usuários distintos, variando-se os diversos parâmetros de entrada para várias combinações. Os parâmetros de entrada do algoritmo alterados foram: os pesos das funções de similaridade geográfica (PSG), de pessoas (PSP) e temporal (PST); e o limiar que separa os eventos considerados similares (θ).

O objetivo do experimento foi detectar os melhores pesos para a função de similaridade de eventos, detectar o melhor limiar e comprovar a hipótese de que as informações de pessoas presentes, localização geográfica, horário e data das fotografias podem auxiliar na detecção de eventos similares em uma rede social.

As métricas utilizadas para avaliar os resultados foram: precisão, sensibilidade, acurácia e coeficiente de correlação de Mathews (Phi). As três primeiras métricas foram utilizadas para avaliar o quanto o método proposto é preciso e abrangente. Já o coeficiente de correlação de Mathews, foi utilizado devido à eficácia no tratamento de quantidades bem distintas de classes. A quantidade de combinações entre eventos que o classificador deve retornar como positivas é 456 e como negativas 18.265.

6.4 Resultados

Nesta seção, são apresentados e discutidos os resultados obtidos na execução dos experimentos para a avaliação dos algoritmos propostos para a detecção de eventos pessoais e para a detecção de eventos compartilhados.

6.4.1 Resultados dos Experimentos do Algoritmo de Detecção de Eventos Pessoais

Nas Figuras 6.1, 6.2, 6.3 e 6.4, são apresentados gráficos representando a precisão versus os níveis de granularidade pré-definidos contidos na Tabela 6.1. Como pode ser percebido, quanto maior a granularidade, menor a precisão para os dois algoritmos de agrupamento geográfico. Isso ocorre devido ao seguinte fato, quando o nível de granularidade aumenta, o número de fotografias de um dado evento aumenta também; como resultado, a precisão diminui. Além disto, o primeiro nível de granularidade possui os eventos mais similares aos diretórios. Para o primeiro nível de granularidade com o DBSCAN, uma precisão média de 92,86% foi obtida; sendo a precisão para o quinto nível de 74,75%. Utilizando-se o K-Means, uma precisão média de 96,96% foi obtida para o primeiro nível de granularidade; sendo a precisão do quinto nível de granularidade de 85,34%.

O uso da terceira etapa do algoritmo de anotação de eventos em fotografias mostrou-se essencial, pois os resultados das métricas de precisão e cobertura melhoraram substancialmente por meio desta etapa. A precisão aumentou em 18.67% quando usado o DBSCAN (vide Figura 6.1 e Figura 6.2), e aumentou 5.9%, quando utilizado o K-means (vide Figura 6.3 e Figura 6.4).

Figura 6.1. Precisão do algoritmo completo de detecção de eventos em fotografias utilizando DBSCAN.

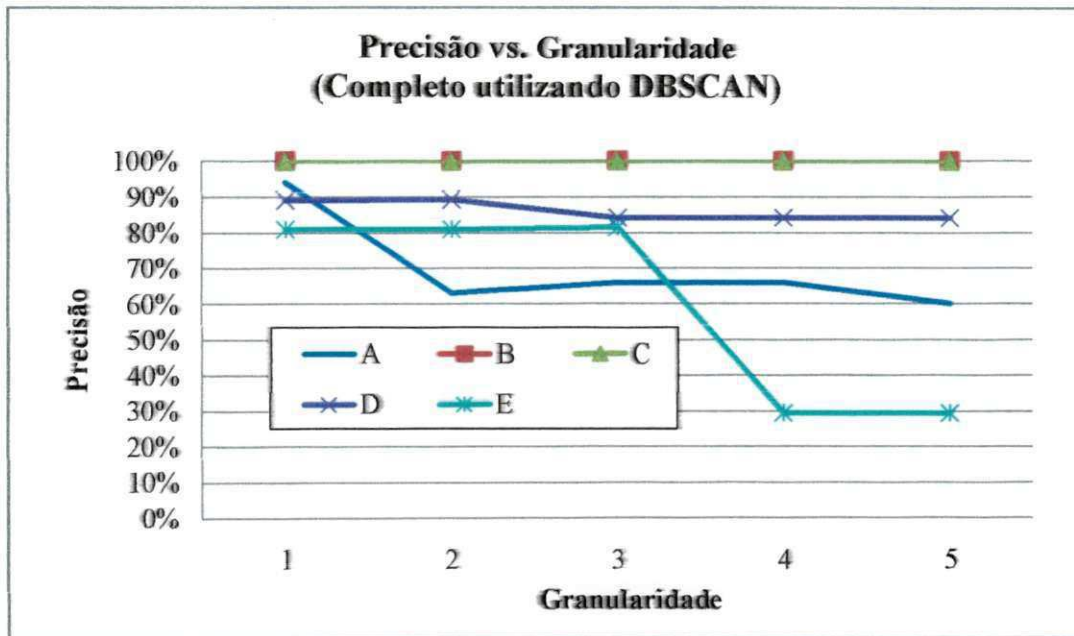


Figura 6.2. Precisão do algoritmo parcial de detecção de eventos em fotografias utilizando DBSCAN.

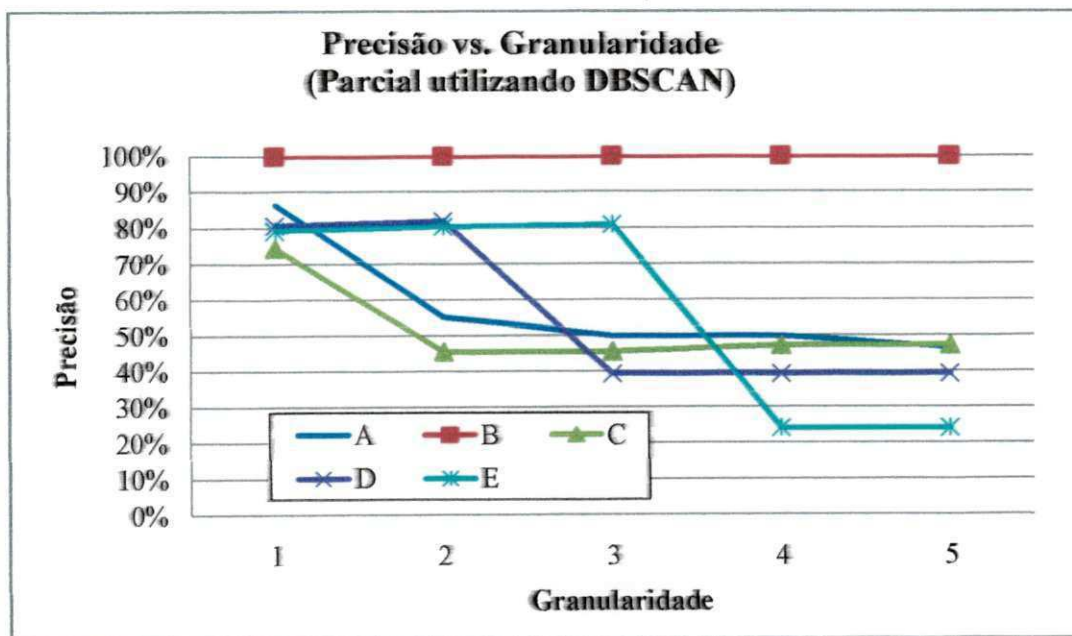


Figura 6.3. Precisão do algoritmo completo de detecção de eventos em fotografias utilizando K-means.

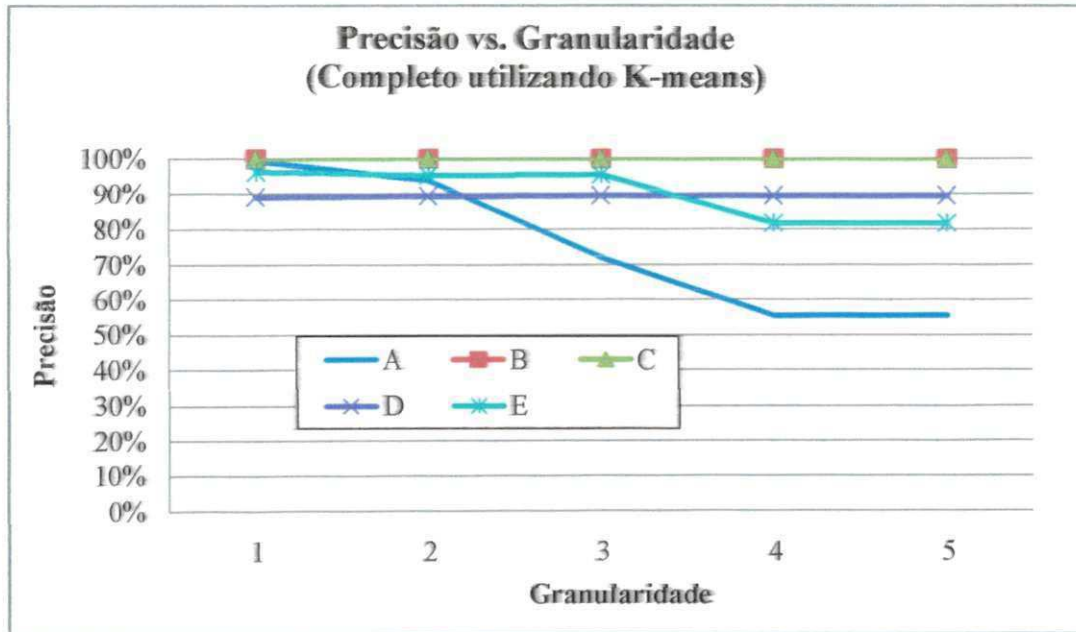
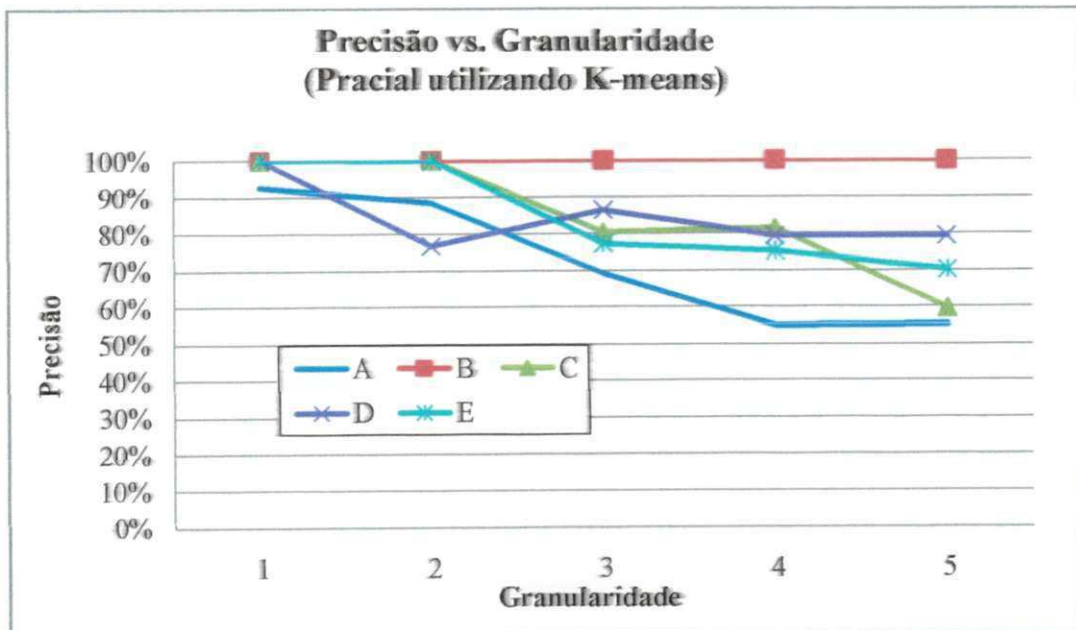


Figura 6.4. Precisão do algoritmo parcial de detecção de eventos em fotografias utilizando K-means.



Nas Figuras 6.5, 6.6, 6.7 e 6.8, são apresentadas as curvas de cobertura de cinco coleções de fotografias utilizadas nos experimentos. Nestes gráficos, a cobertura aumenta à medida que a granularidade aumenta. Isto ocorre devido ao grande número de fotografias no mesmo evento. Observa-se que a cobertura aumentou em 47% com DBSCAN (vide Figura 6.5 e Figura 6.6) e 14.2% com K-means (vide Figura 6.7 e Figura 6.8).

Figura 6.5. Cobertura da anotação de eventos em fotografias com execução completa utilizando DBSCAN.

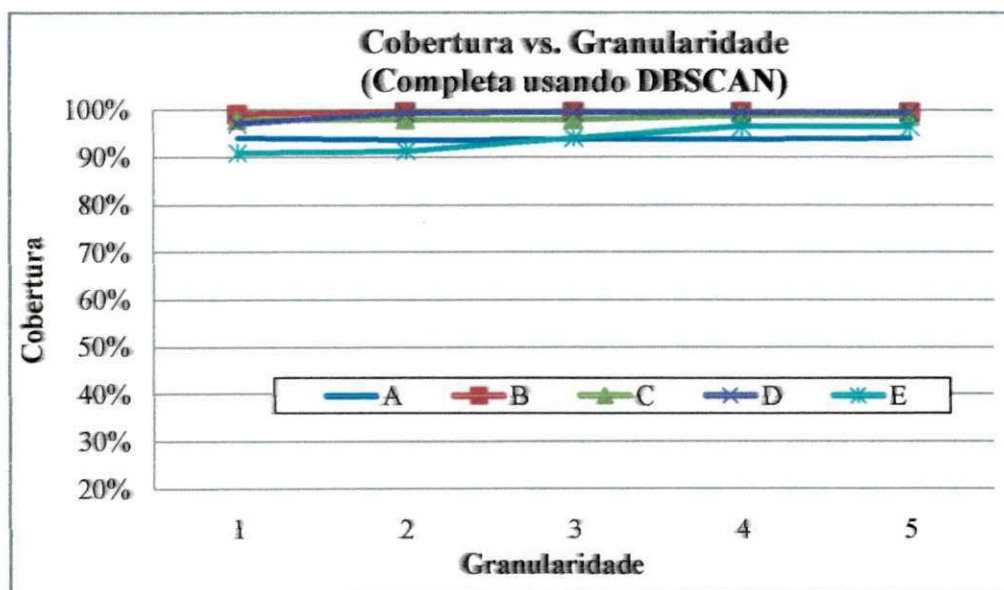


Figura 6.6. Cobertura da anotação de eventos em fotografias com execução parcial utilizando DBSCAN.

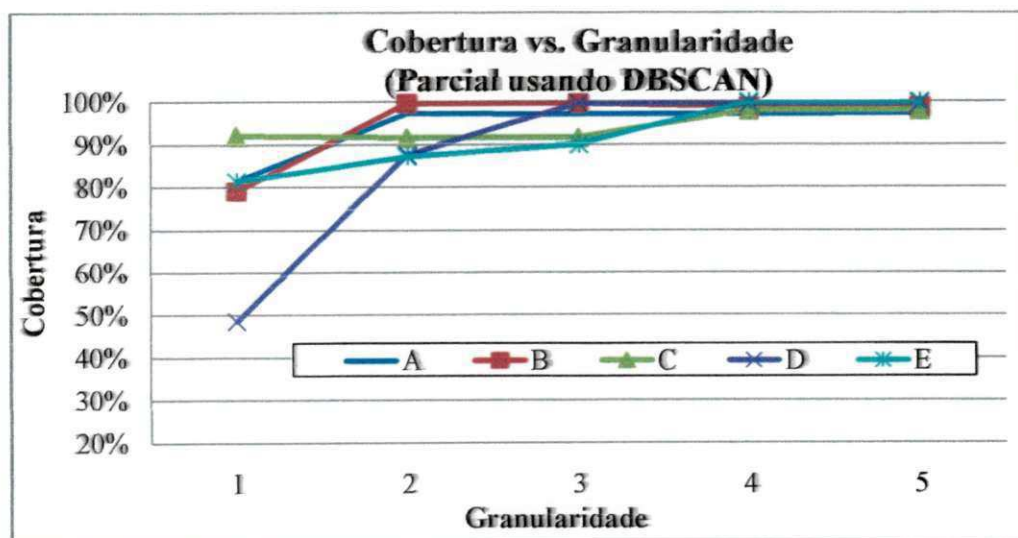


Figura 6.7. Cobertura da anotação de eventos em fotografias com execução completa utilizando K-Means.

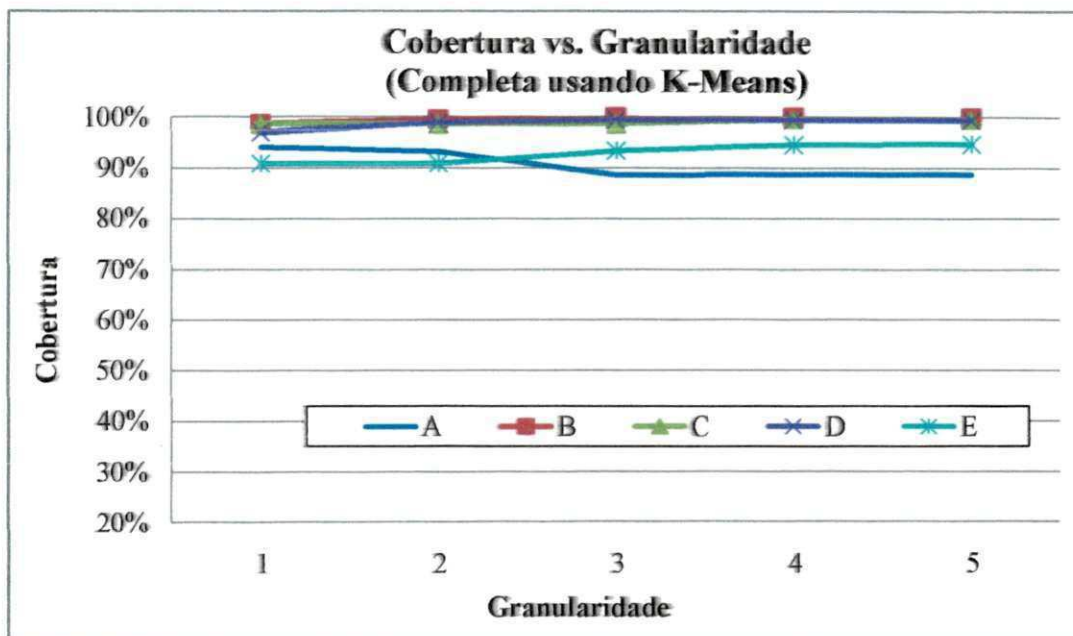
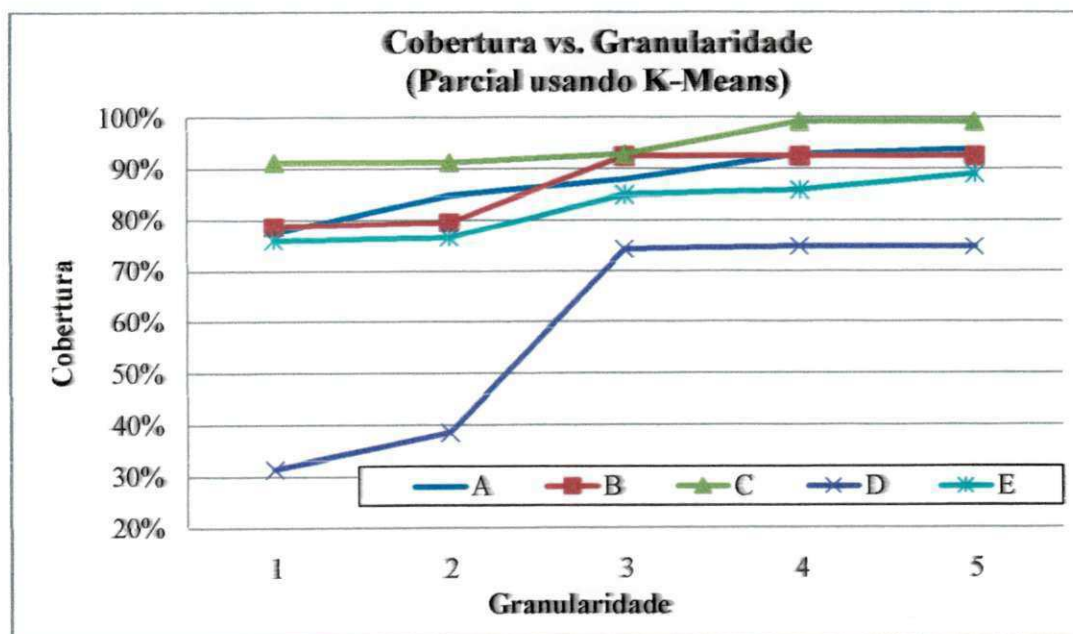


Figura 6.8. Cobertura da anotação de eventos em fotografias com execução parcial utilizando K-Means.



Nas Figuras 6.9, 6.10 e 6.11, mostra-se a comparação entre os algoritmos DBSCAN e K-means. Os resultados indicam que o K-means obteve melhor precisão do que o DBSCAN; entretanto, o último obteve melhor cobertura do que o primeiro. Ainda, como a diferença entre os resultados de cobertura dos dois algoritmos é pequena, a

precisão torna-se mais importante. Conforme os resultados discutidos anteriormente, os parâmetros de entrada têm forte influência nos resultados obtidos.

Figura 6.9. Precisão para algoritmo de detecção de eventos utilizando K-Means e DBSCAN variando-se a granularidade.

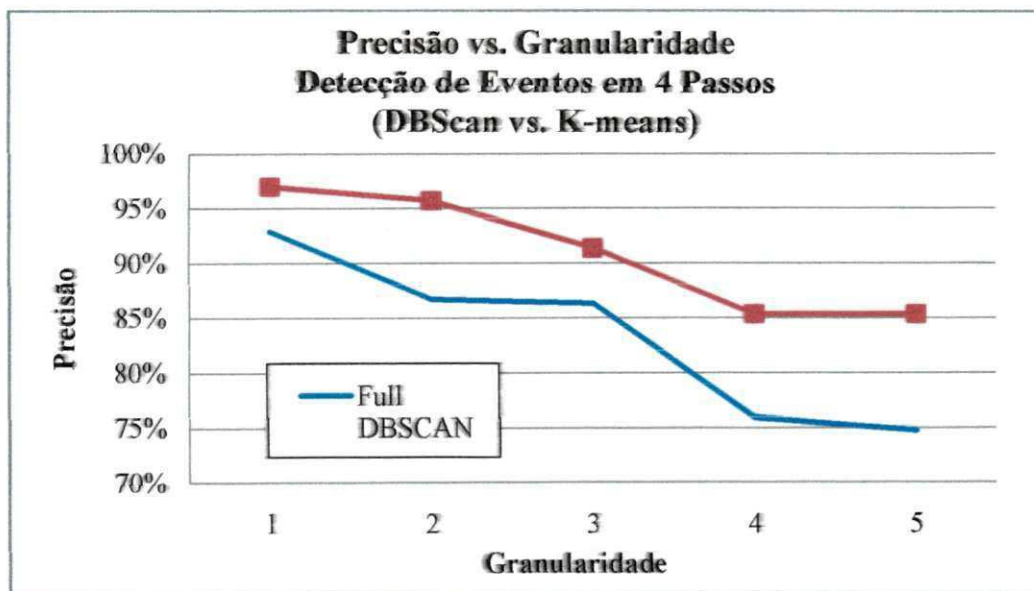


Figura 6.10. Cobertura para algoritmo de detecção de eventos utilizando K-Means e DBSCAN variando-se a granularidade.

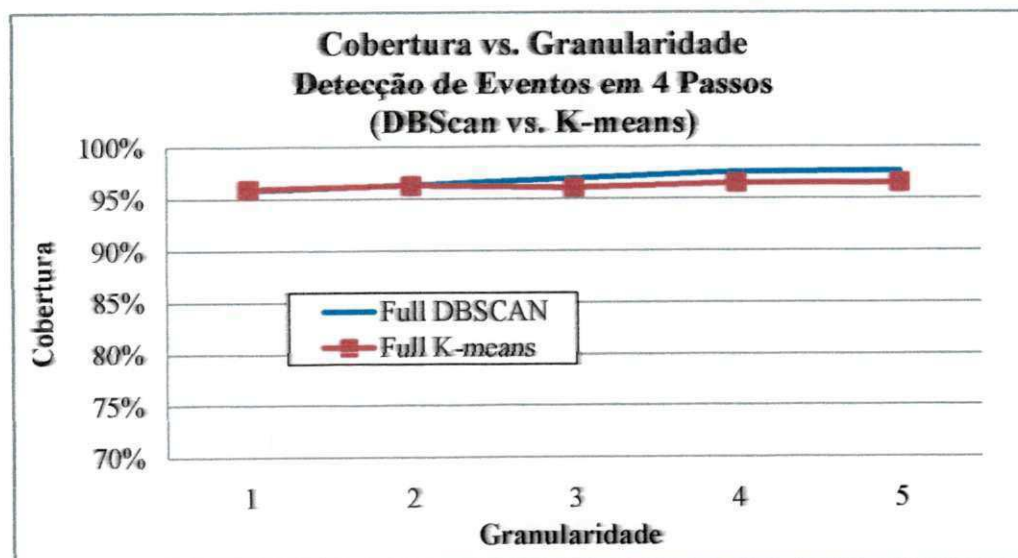
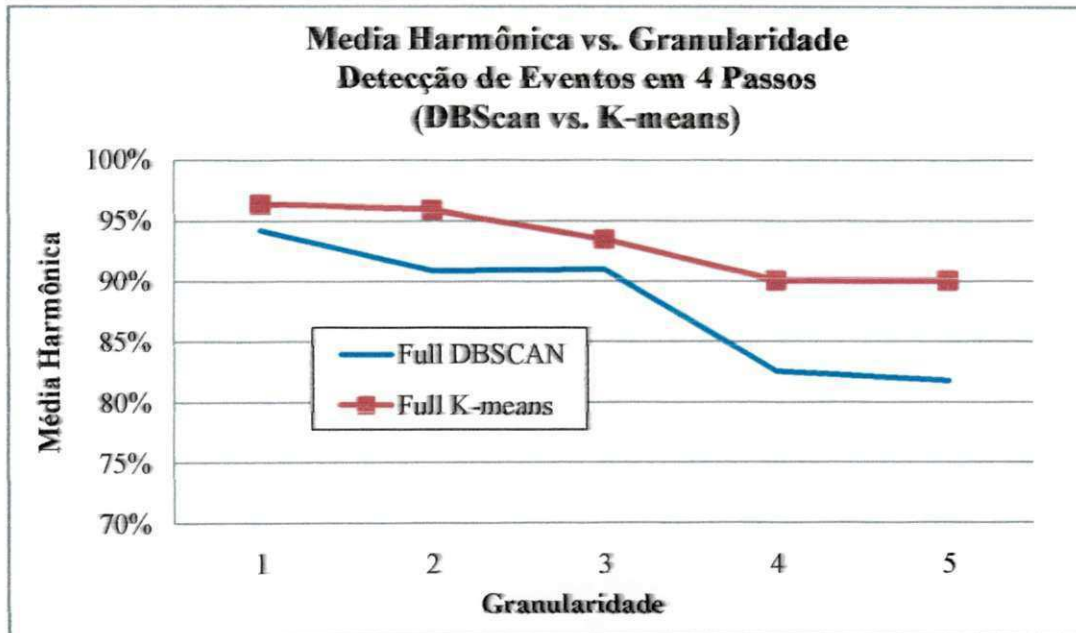


Figura 6.11. Média harmônica para algoritmo de detecção de eventos utilizando K-Means e DBSCAN variando-se a granularidade.



6.4.2 Resultados dos Experimentos do Algoritmo de Detecção de Eventos Compartilhados

Na Figura 6.12, apresentam-se os resultados da métrica acurácia. A partir desta métrica, foram obtidos resultados elevados (o menor foi 79%) devido ao alto número de verdadeiros negativos. Entretanto, foi possível identificar que o detector de eventos similares possui uma melhor acurácia quando utilizadas ponderações para as três funções de similaridade maiores que 0.

Na Figura 6.13, são apresentados os resultados do experimento de variação dos parâmetros utilizando a métrica precisão. A partir desses resultados, verificou-se que o valor do limiar deve ser superior a 0,7 para que seja obtida uma precisão superior a 70%.

Na Figura 6.14, são apresentados os resultados para a métrica sensibilidade, que verifica a quantidade de classificações corretas em relação ao total de eventos similares. Considerando os resultados para os limiares superiores a 0,5, os resultados acima de 0,7 possuem resultados com sensibilidade não satisfatórios. Dessa forma, foram considerados

valores para limiar entre 0,5 e 0,7 para a obtenção de bons resultados para a equação proposta.

Figura 6.12. Acurácia da função de similaridade versus limiar.

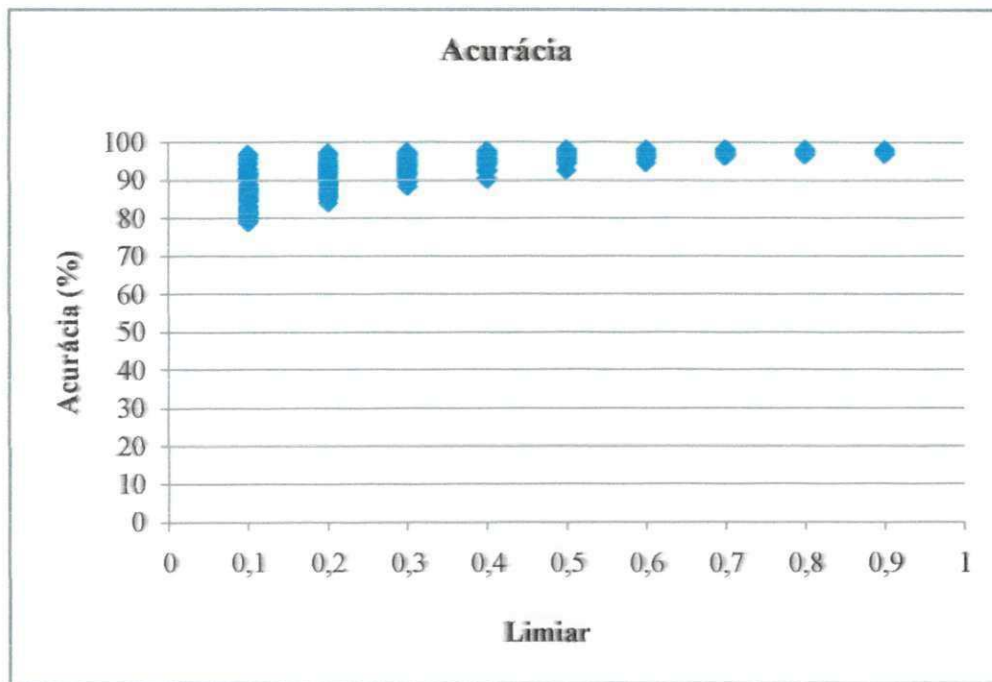


Figura 6.13. Precisão da função de similaridade versus o limiar.

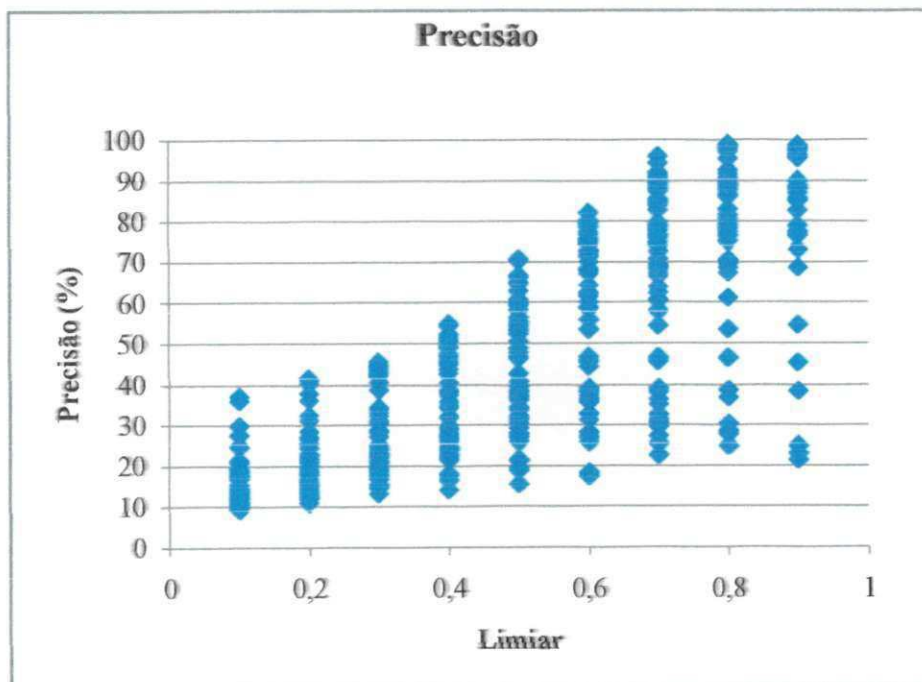
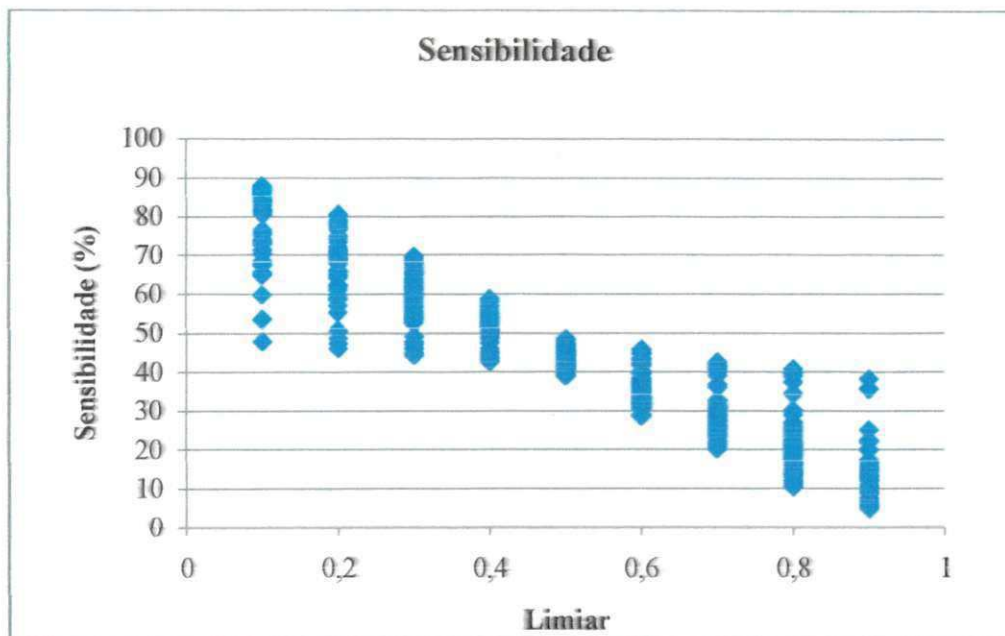


Figura 6.14. Sensibilidade da função de similaridade versus o limiar.



Na Tabela 6.3, são apresentados os melhores resultados obtidos para cada uma das métricas utilizadas na avaliação do método de detecção de eventos compartilhados. Todas as métricas obtiveram seus melhores resultados utilizando pelo menos uma ponderação de 0,2 para as funções de similaridade geográfica, temporal e de pessoas. Para o protótipo implementado, foram selecionados os parâmetros com melhores resultados na métrica Phi, que são: 0,6 para o limiar; 0,2 para a ponderação da função de similaridade de pessoas; 0,3 para a função de similaridade geográfica; e 0,5 para a função de similaridade temporal.

Tabela 6.3. Melhores resultados obtidos para cada métrica destacada.

Limiar	PSP	PSG	PST	Prec	Sens	Acur	Phi
0,6	0,2	0,2	0,6	79,3	34,4	98,2	0,51
0,5	0,2	0,3	0,5	70,4	41,7	98,2	0,53
0,1	0,4	0,4	0,2	10,6	87,9	81,7	0,27
0,8	0,2	0,4	0,4	98,9	19,5	98,0	0,43

6.5 Considerações Finais

Neste capítulo, foi apresentada a solução proposta para a detecção de eventos em fotografias digitais, sendo apresentados algoritmos para detecção de eventos pessoais e compartilhados.

A detecção de eventos pessoais ocorre a partir de uma classificação não supervisionada utilizando informações de contexto e fontes externas, como redes sociais. Também é possível a modificação do nível de granularidade da detecção de eventos, de acordo com as preferências do usuário.

Para a detecção de eventos compartilhados em coleções de fotografias de usuários de uma mesma rede social, foram utilizadas as anotações de pessoas presentes, localização geográfica e hora e data da captura da fotografia. O algoritmo proposto obteve resultados satisfatórios em termos de precisão e acurácia, para uma cobertura aceitável. Para a coleção utilizada nos experimentos, foi considerado como melhor resultado a execução do algoritmo para eventos com similaridade superior a 60%, com peso de 20% para as pessoas, de 20% para a localização geográfica e 60% para a parte temporal. Nessa execução, o algoritmo obteve uma acurácia de 98%, precisão de 79% e cobertura de 34%.

Pretende-se, com a detecção de eventos compartilhados, que outras anotações possam ocorrer de forma automática ou auxiliar os algoritmos de anotação semiautomática. Por exemplo, supondo a existência de um evento compartilhado nas coleções de fotografias de dois usuários, na coleção de um dos usuários, o contato José foi anotado em uma das fotografias, já na coleção do segundo usuário, José não foi anotado nas fotografias, porém, a partir da detecção do evento compartilhado, José é anotado como participante do evento na coleção do segundo usuário. Outra contribuição dos eventos compartilhados é para a anotação semiautomática de pessoas presentes nas fotografias (FIGUEIRÊDO *et al.*, 2012a), pois existem estimadores, tal como aquele de recorrência no evento, que possuem melhor precisão e cobertura com um número maior de fotografias.

Capítulo 7 Protótipo de Gerenciador de Fotografias

Neste capítulo, apresenta-se o protótipo desenvolvido para validação de conceito, o qual implementa os algoritmos propostos neste trabalho. O restante deste capítulo está organizado como segue. Na seção 7.1, descreve-se a arquitetura do protótipo. Na Seção 7.2, são apresentados o esquema lógico do banco de dados e os procedimentos armazenados. Na seção 7.3, apresenta-se a interface gráfica com o usuário do protótipo.

7.1 Arquitetura do Protótipo

A arquitetura do protótipo estende a do PhotoGeo (LACERDA, 2009; LACERDA, 2008a; LACERDA, 2008b), apresentada na Figura 7.1. Na Figura 7.2, mostra-se a arquitetura do protótipo que é dividida em camadas conforme a do PhotoGeo.

Figura 7.1. Arquitetura do PhotoGeo (Fonte: (LACERDA, 2009)).

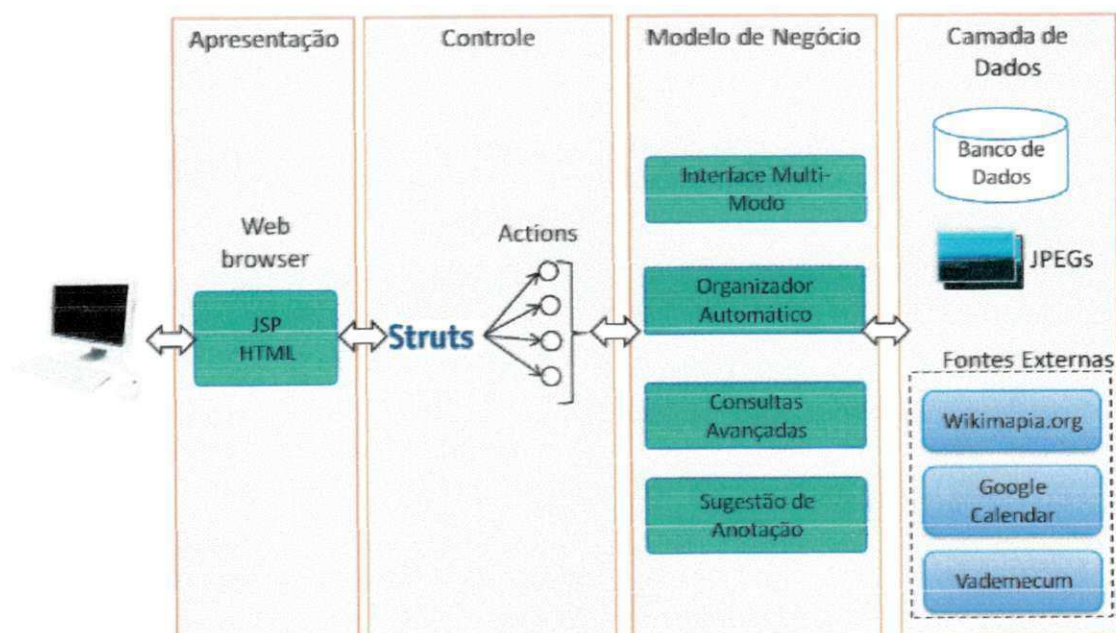
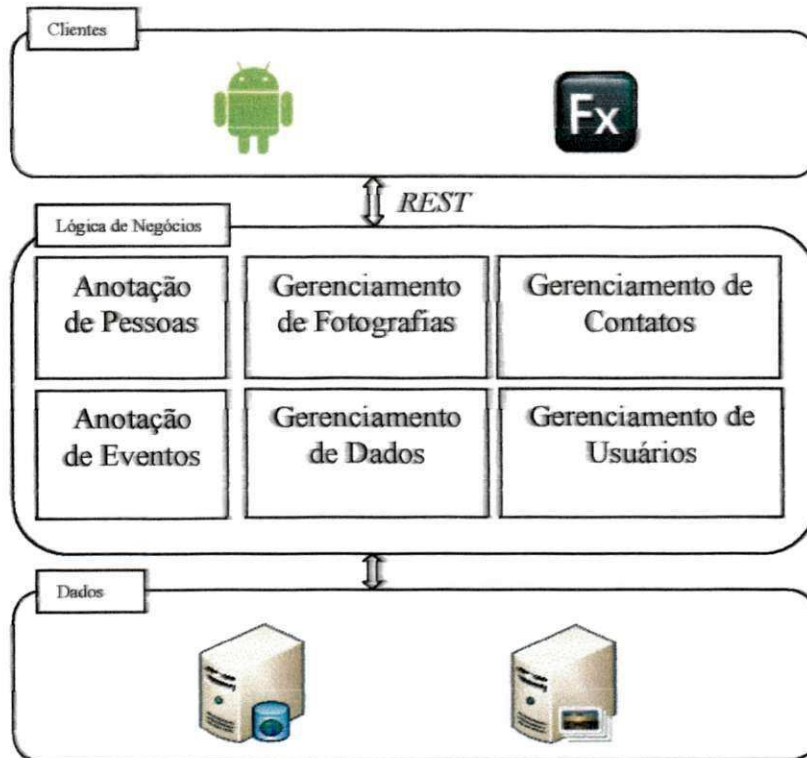


Figura 7.2. Arquitetura do protótipo.



Essa arquitetura utiliza um arranjo tradicional em três camadas: dados, lógica de negócios e clientes. A camada de dados é composta por duas bases de dados: um sistema de banco de dados objeto relacional com suporte espacial e uma coleção de fotografias.

A camada de lógica de negócios é composta por seis módulos principais: gerenciamento de dados, gerenciamento de usuários, gerenciamento de contatos, gerenciamento de fotografias, anotação de pessoas e anotação de eventos.

O módulo de gerenciamento de dados é responsável pelo acesso e mapeamento dos dados em objetos. O módulo de gerenciamento de usuários e o módulo de gerenciamento de contatos são responsáveis pelo gerenciamento dos usuários e dos contatos dos usuários no sistema. Um usuário é uma pessoa que possui permissões de acesso ao sistema, dono de uma coleção de fotografias e possui vários contatos. Os contatos são pessoas que estão presentes nas fotografias dos usuários. Todos os usuários possuem pelo menos um

contato, o qual representa o próprio usuário nas fotografias anotadas com as pessoas presentes. Uma pessoa física só pode ter um objeto do tipo pessoa no sistema, mas podem existir diversos contatos para uma pessoa. Se uma pessoa é contato de mais de um usuário, para cada usuário, existe um contato diferente associado, mas somente uma pessoa. Ao ser adicionado um contato novo no sistema, verifica-se a existência de uma pessoa com o mesmo e-mail no sistema, caso positivo, o contato criado é associado a esta pessoa, caso negativo, é criada uma nova pessoa e o contato é associado a ela.

O módulo de gerenciamento de fotografias é responsável pela adição, remoção, extração de metadados e recuperação de fotografias. Além disto, neste módulo é verificado quem tem permissão de acessar as fotografias. As geometrias das fotografias são adicionadas ao banco de dados de duas formas: extração dos metadados (lat/long) das fotografias (Exif) ou anotadas pelo usuário na interface gráfica com o auxílio de mapa.

O módulo de anotação de pessoas estende o módulo de sugestão de anotação da arquitetura do PhotoGeo. Enquanto o módulo do PhotoGeo apenas sugere uma lista de sugestões para anotação de pessoas, o módulo de anotação de pessoas realiza a anotação de pessoas utilizando reconhecimento de faces além de gerar a lista de sugestões. Outra diferença entre os módulos é que a lista de sugestões do PhotoGeo utiliza somente informações de contexto, enquanto que a do protótipo deste trabalho utiliza também informações de conteúdo da imagem.

O módulo de anotação de eventos estende o módulo organizador automático do PhotoGeo. O módulo organizador automático agrupa as fotografias de um usuário de acordo com similaridades espaço-temporais. Neste trabalho, foi utilizado o conceito de evento ao invés de organizador de fotografias. Além disso, no módulo de anotação de eventos considera-se a hierarquia de eventos e a detecção de eventos compartilhados. A detecção de eventos pessoais apresentada neste trabalho é similar à do PhotoGeo, sendo uma das diferenças a não exigência da localização geográfica nas fotografias. O módulo de anotação de eventos também é responsável por indicar, na visualização de eventos, os eventos similares existentes, detectados pelo algoritmo de detecção de eventos compartilhados.

A comunicação entre o servidor e os clientes Web e Móvel é realizada por meio de REST. No Apêndice C, são descritas as requisições necessárias para acessar e alterar os recursos existentes no sistema. Por exemplo, para autenticar-se no sistema, o cliente realiza uma requisição Http POST para `http://{host}/{systemName}/login/` com os parâmetros `username` e `password`.

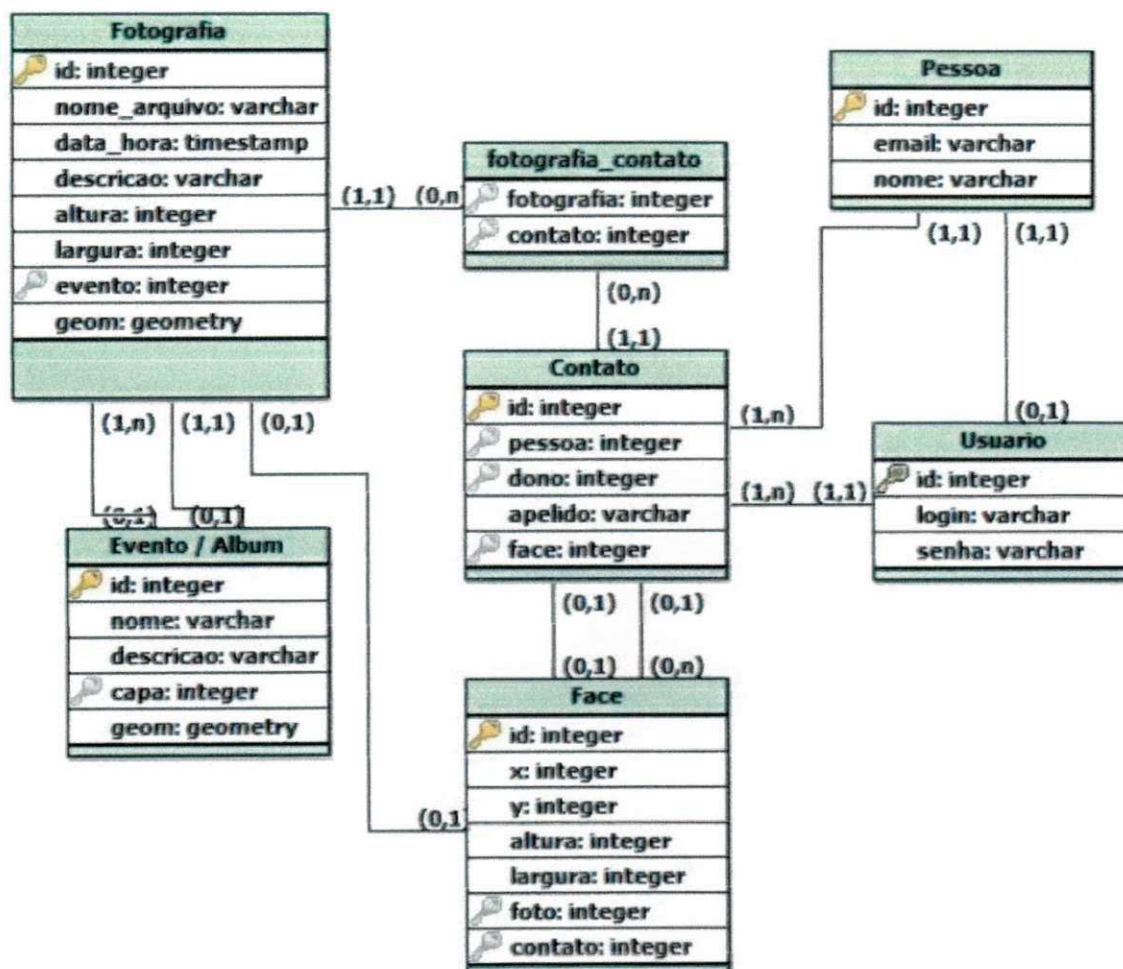
7.2 Banco de Dados

Na Figura 7.3, apresenta-se o esquema lógico do banco de dados do protótipo. A tabela *Usuario* possui os usuários do PhotoGeo. *Fotografia* possui as fotografias pertencentes ao usuário. *Contato* possui os contatos do usuário que estão presentes nas fotografias, sendo o mapeamento dos contatos presentes nas fotografias feito por *fotografia_contato*. Conforme supracitado na descrição do módulo de gerenciamento de usuários e contatos, para representar o próprio usuário em uma fotografia, todos os usuários possuem um contato que o representa. As faces dos contatos extraídas das fotografias ficam na tabela *Face*. Os eventos são representados na tabela *Evento*, existindo outra tabela *Album* responsável por identificar a separação das fotografias realizada pelo usuário.

Algumas tabelas adicionais foram incorporadas ao esquema para a implementação dos algoritmos propostos neste trabalho. Esta incorporação visa à obtenção de um melhor desempenho e escalabilidade na execução dos algoritmos de anotação de pessoas e eventos nas fotografias. Além das tabelas, alguns procedimentos armazenados foram adicionados para executar algumas etapas dos algoritmos.

O Sistema Gerenciador de Banco de Dados (SGBD) PostgreSQL foi utilizado para gerenciar o banco de dados do protótipo. Os procedimentos armazenados foram escritos utilizando a linguagem PL/pgSQL. A escolha do PostgreSQL deveu-se a dois fatores: atende aos requisitos do sistema e um *software* gratuito. Porém, os algoritmos podem ser implementados em qualquer SGBD que dê suporte a dados espaciais e a uma linguagem procedural aliada a SQL.

Figura 7.3. Esquema lógico do Banco de Dados do protótipo.



Nas subseções seguintes, serão descritos detalhes adicionais do esquema lógico adicionados para a execução dos algoritmos propostos neste trabalho.

7.2.1 Anotação de Pessoas Presentes

Para o algoritmo de anotação de pessoas presentes nas fotografias, as seguintes tabelas foram incorporadas ao esquema:

- *Estimador*, que possui os estimadores existentes no sistema com o nome da função que executa o cálculo da estimativa;

- *Pesos Estimadores*, que possui os pesos utilizados em cada estimador para cada contato;
- *Coocorrencia*, que armazena a quantidade de fotografias que dois contatos aparecem anotados juntos; e
- *Pontuacao_Contato_Face*, que possui a pontuação gerada pelo algoritmo de reconhecimento de faces para os contatos sugeridos para as faces detectadas nas fotografias.

Os seguintes procedimentos armazenados foram definidos no banco de dados do protótipo para o algoritmo de anotação de pessoas presentes nas fotografias:

- *inserir_coocorrencia*: função de gatilho que atualiza tabela de *Coocorrencia* quando uma anotação de um contato presente em uma fotografia é realizada;
- *remover_coocorrencia*: função de gatilho que atualiza tabela de *Coocorrencia* quando uma anotação de um contato presente em uma fotografia é removida;
- *recorrencia_global*: função que calcula o valor do estimador RG para um determinado contato em uma determinada fotografia;
- *recorrencia_temporal*: função que calcula o valor do estimador RT para um determinado contato em uma determinada fotografia;
- *recorrencia_geografica*: função que calcula o valor do estimador RG para um determinado contato em uma determinada fotografia;
- *estimador_coocorrencia*: função que calcula o valor do estimador CO para um determinado contato em uma determinada fotografia;
- *estimador_reconhecimento_faces*: função que calcula o valor do estimador RF para um determinado contato em uma determinada fotografia;
- *estimativa_contato*: função que implementa a Equação 4 para o cálculo de $P_{c,f}$, que é a estimativa do contato c estar presente na fotografia f ;
- *sugestoes*: função que retorna os H primeiros contatos com maior valor de $P_{c,f}$;
- *ponderacao*: função que atualiza a tabela *Pesos_Estimadores* ponderação de um estimador e para um contato c mediante o cálculo da Equação 5; e

- *atualizar_ponderacao*: função de gatilho que é disparada quando uma nova anotação de contato presente em uma fotografia é realizada, sendo chamada a função de *ponderacao* para o contato *c* anotado na fotografia *f* quando a quantidade de anotações com *c* é divisível por *D*.

Para que um novo estimador seja adicionado no sistema, é necessária a criação da função no banco de dados que calcula o valor da estimativa e o cadastro do estimador na tabela *Estimador*. A função deve seguir a seguinte assinatura:

nome_do_estimador(id_contato integer, id_fotografia integer) Returns numeric;

Para a seleção dos estimadores que farão parte do modelo gerado, é utilizado um método que remove o estimador com menor ponderação enquanto nenhum melhoramento seja observado na estimativa de erro dada pelo método mínimos quadrados.

7.2.2 Detecção de Eventos

Para a execução do algoritmo de anotação de eventos pessoais, foi adicionado no Banco de Dados o procedimento armazenado *Detecta Eventos Pessoais*, no qual são criados os eventos na tabela *Eventos* e anotadas as fotografias com essa informação.

Em contrapartida, para a detecção de eventos compartilhados, a tabela *Similaridade_Eventos* foi adicionada ao esquema para armazenar o valor do cálculo da similaridade entre cada par de eventos. Os valores de PSG, PSP e PST são armazenados na tabela *Pesos_Similaridade_Eventos*.

O procedimento armazenado *atualiza_similaridade_eventos* foi adicionado ao Banco de Dados para executar o cálculo da Equação 8 para cada par de eventos utilizando os valores de *PSG*, *PSP* e *PST*, sendo o resultado armazenado na tabela *Similaridade_Eventos*. As Equações 9, 10 e 11 são implementadas por meio das funções *Similaridade_Geográfica_Eventos*, *Similaridade_Temporal_Eventos* e *Similaridade_Pessoas_Eventos*, respectivamente.

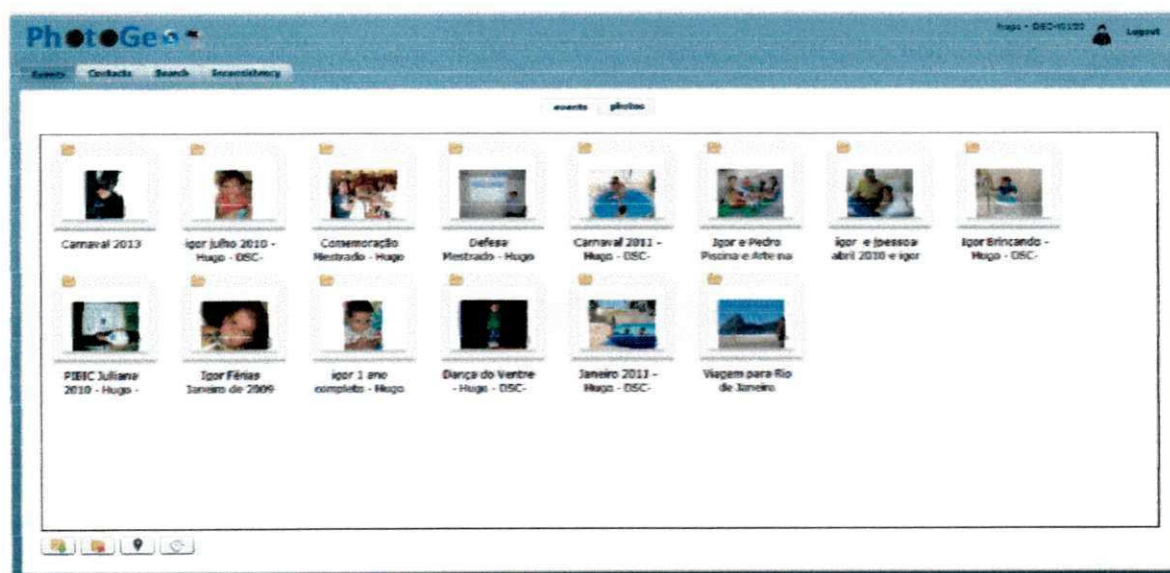
Periodicamente, um gatilho dispara uma ação que procura por eventos compartilhados nas listas de contatos dos usuários no banco de dados. Na implementação atual, o gatilho é disparado uma vez ao dia.

7.3 Interface Gráfica com Usuário

Conforme apresentado na arquitetura do protótipo, existem dois clientes de interface gráfica com o usuário, um Web e outro móvel. O cliente Web foi implementado utilizando a linguagem Flex e possui todas as características descritas na Seção 3.3.1.

Na Figura 7.4, apresenta-se a interface do cliente Web com a lista de eventos existentes do usuário autenticado. Por outro lado na Figura 7.5, mostra-se a lista de fotografias contidas em um determinado evento selecionado, sendo possível adicionar mais fotografias para o evento selecionado.

Figura 7.4. Interface que exibe os eventos do usuário.



Na Figura 7.6, apresenta-se a interface do protótipo, relativa à visualização de fotografias capturadas em um evento, na qual está destacado com um retângulo vermelho a parte que apresenta para o usuário os eventos similares detectados nas coleções de fotografias de seus contatos, que representa o evento compartilhado.

Cada fotografia cadastrada no sistema possui informações específicas do momento de captura da fotografia, podendo estas terem sido adquiridas dos metadados da mesma ou providas de anotações realizadas no próprio sistema. Na Figura 7.7, apresenta-se uma dessas informações de uma determinada fotografia selecionada, que é a de localização

geográfica, sendo possível a alteração por meio do mapa. Na Figura 7.8, são mostradas as pessoas presentes na fotografias que já foram anotadas.

Figura 7.5. Interface que exibe as fotografias de um evento.



Figura 7.6. Interface de exibição de eventos compartilhados.

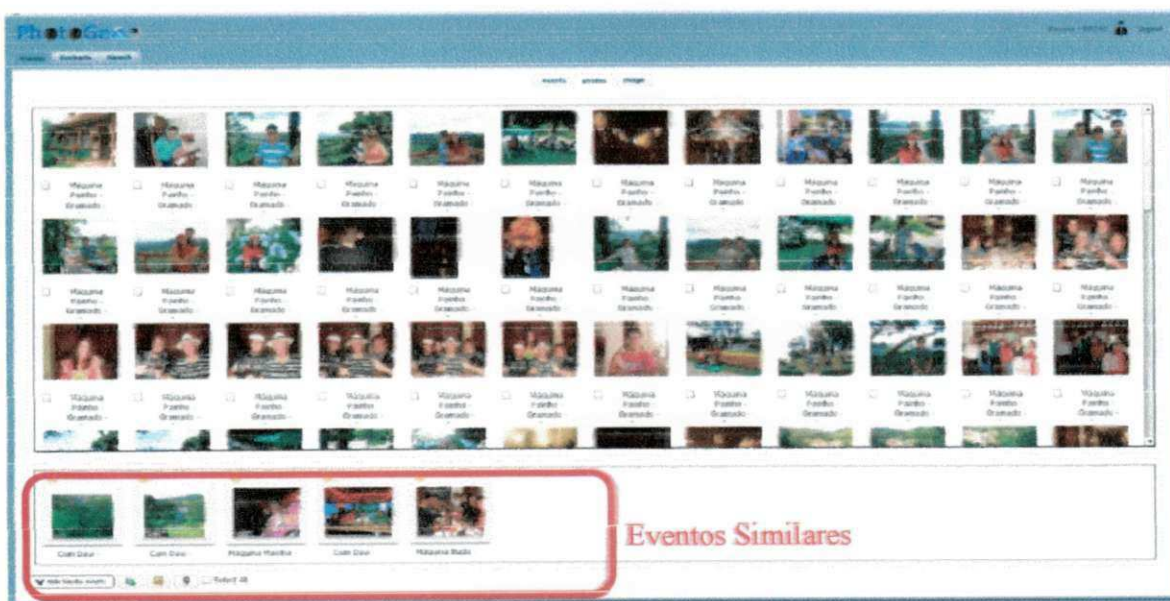


Figura 7.7. Visualização da localização geográfica da fotografia no PhotoGeo.

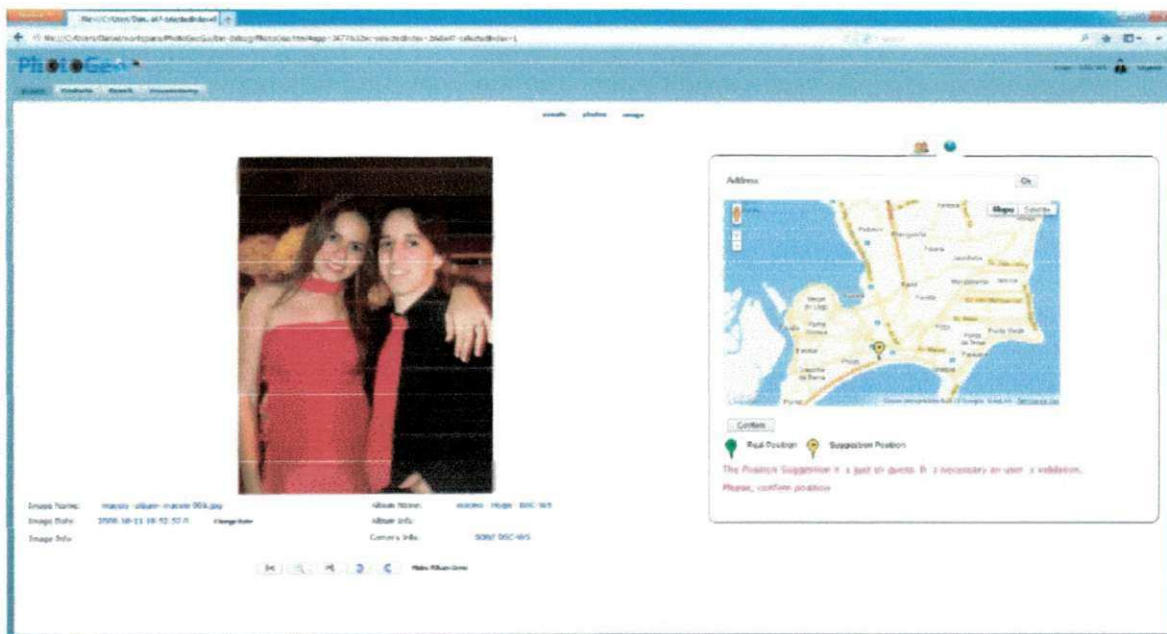


Figura 7.8. Visualização de pessoas presentes na fotografia no PhotoGeo.



Para realizar a anotação de pessoas presentes na fotografia, o usuário possui uma interface que fornece sugestões das pessoas mais prováveis de estarem presentes na fotografia. Na Figura 7.8, apresenta-se a tela com uma lista de sugestões que o usuário pode selecionar para a anotação na fotografia. O cliente móvel possui duas versões, uma para Windows Mobile 6.1 e outra para Android e possui as seguintes funcionalidades:

captura de fotografias, adição de fotografias armazenadas no dispositivo, anotação semiautomática de pessoas presentes nas fotografias, criação de eventos e georreferenciamento automático.

Na Figura 7.9, mostra-se o cliente móvel do protótipo implementado para Windows Mobile 6.1 e testado no HP iPAQ 910c. Esse cliente móvel georreferencia as fotografias capturadas a partir da localização geográfica obtida do GPS embutido no dispositivo. Neste dispositivo, as informações de localização geográfica não são adicionadas automaticamente pelo dispositivo nos metadados da fotografia, sendo necessária a implementação dessa funcionalidade.

Figura 7.9. Cliente móvel para Windows Mobile.

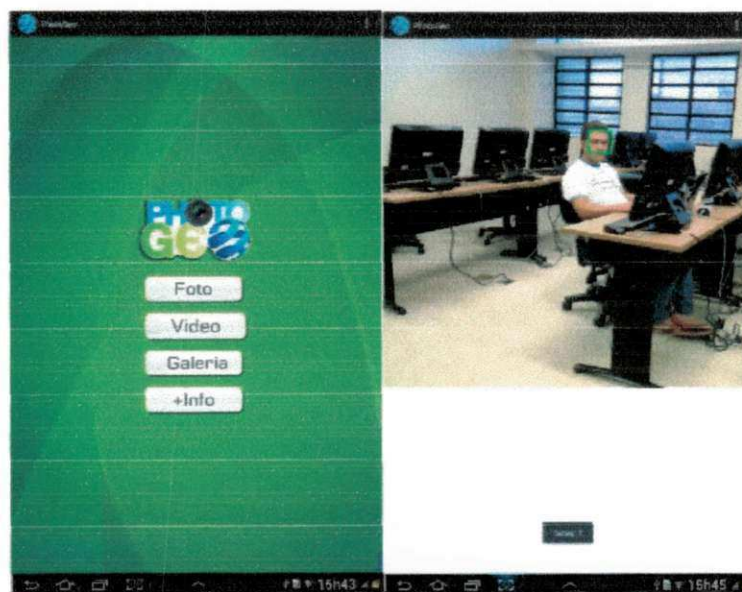


No cliente Windows Mobile, uma lista de sugestões de pessoas é exibida após a captura das fotografias, sendo esta lista gerada pelo algoritmo para dispositivo móvel de

geração de lista de sugestões de pessoas presentes em fotografias, proposto na Seção 5.3 e analisado na Seção 5.5.1. Caso uma determinada pessoa esteja na fotografia, mas não tenha sido sugerida pelo algoritmo, o usuário tem a opção de adicionar o nome desta pessoa na anotação. Antes de capturar uma fotografia, o usuário tem a opção de informar que um evento foi iniciado e, depois de capturar as fotografias, informar que o evento foi encerrado.

Na Figura 7.10, apresenta-se a interface do cliente móvel para smartphones com sistema operacional Android 4.1.2. A captura de fotografias, o georreferenciamento das fotografias e a detecção de faces são funcionalidades providas pela API do Android 4.1.2 que foram utilizadas no cliente.

Figura 7.10. Interface do cliente móvel para smartphones com Android.



Na Figura 7.11, apresentam-se as telas das etapas de anotação de pessoas presentes nas fotografias. Inicialmente, apresenta-se a face destacada ao usuário com a indicação do candidato mais provável com uma interrogação. Quando o usuário toca na interface, apresenta-se a lista de sugestões, na qual ele pode selecionar o contato na lista ou pode adicionar outro nome. Por fim, a fotografia é apresentada com a indicação da anotação da pessoa presente na fotografia. Essa interface foi implementada similar a do Android 2.1.2, modificando-se apenas a lista de sugestões.

Figura 7.11. Etapas de anotação no cliente móvel Android.



7.4 Avaliação com Usuários

O cliente Web foi utilizado por 13 usuários, sendo que todos possuem perfis em redes sociais. Os usuários utilizaram o protótipo Web para carga de fotografias no sistema, anotação dos eventos, anotação das pessoas presentes nas fotografias e georreferenciamento. Algumas melhorias foram implantadas entre os testes dos usuários, sendo as experiências de cada usuário utilizadas para melhorias na interface e detecção de defeitos.

O tempo médio para a anotação de cada fotografia foi de 17 segundos, sendo neste tempo computada a realização das seguintes tarefas: criação dos eventos, carga das fotografias, georreferenciamento, adição dos contatos e anotação das pessoas presentes nas fotografias. A quantidade média de fotografias anotadas por minuto foi de 3,5.

O georreferenciamento das fotografias foi umas das tarefas mais complicadas para os usuários, devido ao não conhecimento das localizações geográficas exatas das fotografias. A partir das dificuldades encontradas no georreferenciamento das fotografias, iniciou-se uma nova pesquisa que visa a elaboração de um algoritmo para propagar a localização geográfica de outras fotografias, utilizando-se inclusive o algoritmo de detecção de eventos compartilhados proposto neste trabalho (LACERDA *et al.*, 2013).

No escopo desta pesquisa, não foram realizados testes de usabilidade, pois os algoritmos propostos para anotação automática não necessitam e o semiautomático para

anotação de pessoas utiliza uma estratégia de lista de sugestões de pessoas, que já foi testada em outra pesquisa (NAAMAN, 2005).

Capítulo 8 Conclusões

Neste capítulo, são apresentadas as conclusões obtidas a partir dos resultados alcançados no desenvolvimento desta tese, sendo feito um paralelo com as hipóteses (Seção 1.3) e os objetivos específicos (Seção 1.4) adotados nesta tese.

Com relação à hipótese H1, foram apresentadas e implementadas três técnicas para gerar listas de sugestões, com o intuito de auxiliar o usuário na anotação de pessoas em fotografias. A primeira técnica utiliza apenas informações de contexto, a segunda utiliza apenas informações de conteúdo e a terceira, proposta nesta tese, combina informações de contexto e de conteúdo. Pôde-se concluir, por meio do experimento realizado, que o algoritmo que combina informações de contexto e de conteúdo produz melhor taxa de acerto do que os demais apresentados. Também foi observado que o algoritmo que utiliza somente informações de contexto possui uma melhor taxa de acerto quando comparado com o que utiliza apenas análise de conteúdo da imagem (detecção e reconhecimento de face).

Com relação à hipótese H2, foram apresentadas duas técnicas de combinação de estimadores utilizando somatório ponderado. A primeira técnica utiliza regressão linear para seleção e obtenção das ponderações dos estimadores, sendo útil quando muitos estimadores estão envolvidos, pois seleciona os melhores para cada usuário. A segunda técnica utiliza inclinação para calcular a ponderação de cada estimador e tem a desvantagem de não verificar quais estimadores não interferem na taxa de acerto quando eliminados na combinação. Os experimentos mostraram que a técnica de ponderação de estimadores pelo método de inclinação obteve melhores resultados quando utilizada a métrica de *H-Hit Rate*, obtendo uma taxa de acerto de aproximadamente 78% para uma lista de 5 sugestões e de 85% para a coleção com melhores resultados. Quando anotados somente os 20 contatos mais importantes de cada coleção a taxa de acerto para uma lista de 5 sugestões aumentou para 87% para todas as coleções juntas e de 91% para a coleção com melhores resultados.

Para a anotação de pessoas presentes nas fotografias em dispositivos móveis, foi proposto um algoritmo com estimadores simplificados com relação ao algoritmo para

servidor. Verificou-se que a versão do algoritmo de anotação de pessoas presentes nas fotografias para servidor possui um taxa de acerto superior ao da versão para dispositivos móveis. Porém, a versão para dispositivo pode ser utilizada quando não for detectada uma conexão com o servidor pelo dispositivo móvel.

Assim, o objetivo específico O1 e O2 foram alcançados com o algoritmo de anotação semiautomática de pessoas em fotografias utilizando contexto e conteúdo e com ponderação e filtragem de estimadores descrito no Capítulo 5.

Com relação à hipótese H3, o algoritmo proposto nesta tese para detecção de eventos compartilhados em coleções de fotografias de usuários de uma mesma rede social foi avaliado e obteve resultados satisfatórios em termos de precisão e acurácia, para uma cobertura aceitável. Nos experimentos do algoritmo de detecção de eventos compartilhados, foi verificada a importância das informações de pessoas nas fotografias, carimbo de tempo e localização geográfica para o cálculo da função de similaridade entre eventos, sendo testada a hipótese H4.

Destarte, os objetivos O3, O4 e O5 foram atingidos pelos algoritmos propostos no Capítulo 6 para detecção de eventos pessoais e compartilhados.

8.1 Proposição para Pesquisas Futuras

Como pesquisas futuras, são sugeridos alguns tópicos pertinentes para dar prosseguimento à pesquisa ora documentada, sendo agrupados por nível de dificuldade:

- Sugestões de pesquisas para o nível de Graduação:

Avaliação dos aspectos relacionados à nova experiência de navegação em fotografias, a partir de eventos pessoais e eventos compartilhados em uma rede social de grande impacto; e

- Avaliação do impacto da granularidade dos eventos nas funções de similaridade entre eventos.

- Sugestões de pesquisas para o nível de Mestrado:

- Investigação da utilização de técnicas de detecção de eventos compartilhados na propagação de anotações em fotografias;

- Utilização das técnicas utilizadas neste trabalho para anotação de outros tipos de arquivo multimídia, e.g., vídeo e áudio;
 - Elaboração de novos algoritmos de detecção de eventos compartilhados, para aumentar a precisão e cobertura;
 - Avaliação de outros algoritmos de reconhecimento na técnica proposta para anotação semiautomática de pessoas presentes nas fotografias;
 - Utilização da detecção de eventos compartilhados para anotação de pessoas presentes nas fotografias;
 - Estudo dos problemas ocasionados por erros de fusos horários e câmeras com relógios com horas erradas; e
 - Testes de usabilidade no gerenciador de fotografias.
- Sugestões de pesquisas para o nível de Doutorado:
 - Utilização de ontologias para descrever os relacionamentos e a utilização de inferência para a propagação de anotações entre eventos compartilhados.

Referências Bibliográficas

AMES, M.; ECKLES, D.; NAAMAN, M.; SPASOJEVIC, M.; HOUSE, N. Requirements for mobile photoware. **Personal and Ubiquitous Computing**, v. 14, n. 2, p. 95–109, 6 jun 2009.

AMES, M.; NAAMAN, M. **Why we tag: motivations for annotation in mobile and online media**. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07. **Anais...** New York, New York, USA: ACM Press. , 2007

ANGUELOV, D.; LEE, K.; GOKTURK, S. B.; SUMENGEN, B. **Contextual Identity Recognition in Personal Photo Albums**. 2007 IEEE Conference on Computer Vision and Pattern Recognition. **Anais...** [S.l.]: IEEE. , jun 2007

BECKER, H.; NAAMAN, M.; GRAVANO, L. **Learning similarity metrics for event identification in social media**. Proceedings of the third ACM international conference on Web search and data mining - WSDM '10. **Anais...** New York, New York, USA: ACM Press. , 2010

BRENNER, M.; IZQUIERDO, E. **MediaEval Benchmark : Social Event Detection in Collaborative Photo Collections**. Working Notes Proceedings of the MediaEval 2011 Workshop. **Anais...** [S.l.]: CEUR-WS.org. , 2011

BRENNER, M.; IZQUIERDO, E. **Social event detection and retrieval in collaborative photo collections**. Proceedings of the 2nd ACM International Conference on Multimedia Retrieval - ICMR '12. **Anais...** Hong Kong, China: ACM Press. , 2012

CAMBRIA, E.; HUSSAIN, A. Sentic Album: Content-, Concept-, and Context-Based Online Personal Photo Management System. p. 477–496, 2012.

CHEN, L.; HU, B.; ZHANG, L.; LI, M.; ZHANG, H.-J. Face Annotation for Family Photo Album Management. **International Journal of Image and Graphics**, v. 03, n. 01, p. 81–94, jan 2003.

CHOI, J. Y.; NEVE, W. DE; PLATANIS, K. N.; RO, Y. M. Collaborative Face Recognition for Improved Face Annotation in Personal Photo Collections Shared on Online Social Networks. **IEEE Transactions on Multimedia**, v. 13, n. 1, p. 14–28, fev 2011.

CHOI, J. Y.; NEVE, W. DE; RO, Y. M.; PLATANIS, K. Automatic Face Annotation in Personal Photo Collections Using Context-Based Unsupervised Clustering and Face Information Fusion. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 20, n. 10, p. 1292–1309, out 2010.

- CHOI, J. Y.; NEVE, W. DE; RO, Y. M.; PLATANIOTIS, K. N. **Face annotation for personal photos using collaborative face recognition in online social networks**. 2009 16th International Conference on Digital Signal Processing. *Anais...* [S.l.]: IEEE. , jul 2009
- CHOI, J. Y.; YANG, S.; RO, Y. M.; PLATANIOTIS, K. N. **Face annotation for personal photos using context-assisted face recognition**. Proceeding of the 1st ACM international conference on Multimedia information retrieval - MIR '08. *Anais...* New York, New York, USA: ACM Press. , 2008
- CHRISTODOULAKIS, S.; FOUKARAKIS, M.; RAGIA, L.; UCHIYAMA, H.; IMAI, T. Picture Context Capturing for Mobile Databases. *IEEE Multimedia*, v. 17, n. 2, p. 34–41, abr 2010.
- COOPER, M.; FOOTE, J.; GIRGENSOHN, A.; WILCOX, L. Temporal event clustering for digital photo collections. *ACM Transactions on Multimedia Computing, Communications, and Applications*, v. 1, n. 3, p. 269–288, 1 ago 2005.
- COORAY, S. H. **Enhancing Person Annotation for Personal Photo Management Using Content and Context**. Dublin, Ireland: DUBLIN CITY UNIVERSITY, 2008.
- COORAY, S. H.; O'CONNOR, N. E. **Enhancing Person Annotation for Personal Photo Management Applications**. 2009 20th International Workshop on Database and Expert Systems Application. *Anais...* [S.l.]: IEEE. , 2009
- COORAY, S.; O'CONNOR, N. E.; GURRIN, C.; *et al.* **Identifying person re-occurrences for personal photo management applications**. IET International Conference on Visual Information Engineering (VIE 2006). *Anais...* Bangalore, India: IEE. , 2006
- CP-3451, J. **Exchangeable image file format for digital still cameras: Exif Version 2.2**. [S.l.]: Japan Electronics and Information Technology Industries Association. , 2002
- DAVIS, M.; KING, S.; GOOD, N.; SARVAS, R. **From context to content: leveraging context to infer media metadata**. Proceedings of the 12th annual ACM international conference on Multimedia - MULTIMEDIA '04. *Anais...* New York, New York, USA: ACM Press. , 2004
- DAVIS, M.; SMITH, M.; CANNY, J.; *et al.* **Towards context-aware face recognition**. Proceedings of the 13th annual ACM international conference on Multimedia - MULTIMEDIA '05. *Anais...* New York, New York, USA: ACM Press. , 2005
- DAVIS, M.; SMITH, M.; STENTIFORD, F.; *et al.* **Using Context and Similarity for Face and Location Identification**. Proceedings of the IS&T/SPIE 18th Annual Symposium on Electronic Imaging Science and Technology Internet Imaging VII. *Anais...* San Jose, California: IS&T/SPIE Press. , 2006
- DCMI. **Dublin Core Metadata Initiative Metadata Terms**. Disponível em: <<http://dublincore.org/documents/dcmi-terms/>>. Acesso em: 4 abr. 2013.

DEY, A. K. Understanding and Using Context. **Personal and Ubiquitous Computing**, v. 5, n. 1, p. 4–7, 28 fev 2001.

ESTER, M.; KRIEGEL, H.; SANDER, J.; XU, X. **A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise**. Proceedings of The 2nd International Conference on Knowledge Discovery and Data Mining. **Anais...** Portland: [s.n.], 1996

FERNANDES, V.; JUNIOR, G. B.; NETO, C. S.; *et al.* **Integration Model for Location-Based Services in iDTV Applications**. The Fifth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2013). **Anais...** Nice, France: IARIA., 2013

FIGUEIRÊDO, H. F. DE; FERNANDES, R. M.; MENEZES, L. C. DE; *et al.* **A context-aware solution to annotate people in mobile devices**. Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia - MoMM '09. **Anais...** New York, New York, USA: ACM Press., 2009

FIGUEIRÊDO, H. F. DE; LACERDA, Y. A.; PAIVA, A. C. DE; CASANOVA, M. A.; BAPTISTA, C. DE S. PhotoGeo: a photo digital library with spatial-temporal support and self-annotation. **Multimedia Tools and Applications**, v. 59, n. 1, p. 279–305, 1 fev 2012.

FIGUEIRÊDO, H. F. DE; SILVA, J. P. R. DA; LEITE, D. F. B.; BAPTISTA, C. DE S. **Detection of photos from the same event captured by distinct cameras**. Proceedings of the 18th Brazilian symposium on Multimedia and the web - WebMedia '12. **Anais...** New York, New York, USA: ACM Press., 2012

FIGUEIRÊDO, H. F. DE; SILVA, J. P. R. DA; PAIVA, A. C. DE; BAPTISTA, C. DE S. **People Photo Annotation Using Context and Content**. 6th International Workshop on Multimedia and Signal Processing. **Anais...** Viena, Austria: [s.n.], 2012

FIGUEIRÊDO, H. F. DE; SILVA, T. E. DA; PAIVA, A. C. DE; QUEIROZ, J. E. R. DE; BAPTISTA, C. DE S. Infrastructures for Development of Context-Aware Mobile Applications. In: CRUZ-CUNHA, M. M.; MOREIRA, F. (Eds.). **Handbook of Research on Mobility and Computing: Evolving Technologies and Ubiquitous Impacts**. [S.l.]: IGI Global, 2011. v. 1p. 1104–1118.

FREIRE, T. P.; JUNIOR, G. B.; FIGUEIRÊDO, H. F. DE; BAPTISTA, C. DE S.; PAIVA, A. C. DE. **An Approach of Recommendation Based on Non-negative Matrix Factorization Applied to PhotoGeo**. 6th International Workshop on Multimedia and Signal Processing. **Anais...** Viena, Austria: [s.n.], 2012

GALLAGHER, A. C.; CHEN, T. **Clothing cosegmentation for recognizing people**. IEEE Conference on Computer Vision and Pattern Recognition. **Anais...** Anchorage, AK: IEEE., jun 2008

GALLAGHER, A. C.; CHEN, T. Using Context to Recognize People in Consumer Images. **IPSJ Transactions on Computer Vision and Applications**, v. 1, p. 115–126, 2009.

- GALLAGHER, A. C.; NEUSTAEDTER, C. G.; CAO, L.; LUO, J.; CHEN, T. **Image annotation using personal calendars as context**. Proceeding of the 16th ACM international conference on Multimedia - MM '08. *Anais...* New York, New York, USA: ACM Press. , 2008
- HANBURY, A. A survey of methods for image annotation. *Journal of Visual Languages & Computing*, v. 19, n. 5, p. 617–627, out 2008.
- HAYS, J.; EFROS, A. A. **IM2GPS: estimating geographic information from a single image**. 2008 IEEE Conference on Computer Vision and Pattern Recognition. *Anais...* [S.l.]: IEEE. , jun 2008
- JANG, C.; CHO, H.-G. A clustering method for concurrent photos obtained from multiple cameras using max-flow network model. *Multimedia Systems*, v. 18, n. 4, p. 295–317, 1 set 2012.
- JANG, C.; YOON, T.; CHO, H.-G. **A smart clustering algorithm for photo set obtained from multiple digital cameras**. Proceedings of the 2009 ACM symposium on Applied Computing - SAC '09. *Anais...* New York, New York, USA: ACM Press. , 2009a
- JANG, C.; YOON, T.; CHO, H.-G. **A New Clustering Methodology for Group Photos Taken by Multiple Travelers**. 2009 Ninth IEEE International Conference on Computer and Information Technology. *Anais...* [S.l.]: IEEE. , 2009b
- JANG, C.; YOON, T.; CHO, H.-G. Digital photo classification methodology for groups of photographers. *Multimedia Tools and Applications*, v. 50, n. 3, p. 441–463, 6 fev 2010.
- JOSHI, D.; GALLAGHER, A.; YU, J.; LUO, J. Inferring photographic location using geotagged web images. *Multimedia Tools and Applications*, v. 56, n. 1, p. 131–153, 16 jul 2012.
- LACERDA, Y. A. **PhotoGeo - Uma Biblioteca Digital de Fotografias com Suporte de Geoprocessamento**. Campina Grande: Universidade Federal de Campina Grande, 2009.
- LACERDA, Y. A.; FIGUEIRÊDO, H. F. DE; BAPTISTA, C. DE S.; PAIVA, A. C. DE. **Expandindo e utilizando informações de contexto para a sugestão de anotações de fotografias digitais**. Proceedings of the 14th Brazilian Symposium on Multimedia and the Web - WebMedia '08. *Anais...* New York, New York, USA: ACM Press. , 2008
- LACERDA, Y. A.; FIGUEIRÊDO, H. F. DE; BAPTISTA, C. DE S.; SAMPAIO, M. C. **PhotoGeo: A Self-Organizing System for Personal Photo Collections**. 2008 Tenth IEEE International Symposium on Multimedia. *Anais...* [S.l.]: IEEE. , dez 2008
- LACERDA, Y. A.; FIGUEIRÊDO, H. F. DE; FEITOSA, R. G. F.; FALCÃO, A. G. R.; BAPTISTA, C. DE S. **Visual Modeling of Workflow with Support to Multimedia and Spatiotemporal Indexing**. Proceedings of the Second International Conference on Electronic Government and the Information Systems Perspective. *Anais...* Toulouse, France: Springer. , 2011

- LACERDA, Y. A.; FIGUEIRÊDO, H. F. DE; SILVA, J. P. R. DA; *et al.* **On Improving Geotag Quality in Photo Collections**. The Fifth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2013). **Anais...** Nice, France: IARIA. , 2013
- LEW, M. S.; SEBE, N.; DJERABA, C.; JAIN, R. Content-based multimedia information retrieval. **ACM Transactions on Multimedia Computing, Communications, and Applications**, v. 2, n. 1, p. 1–19, 1 fev 2006.
- LIENHART, R.; MAYDT, J. **An extended set of Haar-like features for rapid object detection**. Proceedings. International Conference on Image Processing. **Anais...** [S.l.]: IEEE. , 2002
- LOGAN, R.; ROSEWAY, A. Geographic Location Tags on Digital Images. p. 156–166, 2003.
- MENEZES, L. C. DE; RAMOS, A. M.; FALCÃO, A. G. R.; BAPTISTA, C. DE S.; FIGUEIRÊDO, H. F. DE. **Context and social networks interaction modeling for context aware alert systems**. 2011 International Conference on Computational Aspects of Social Networks (CASoN). **Anais...** [S.l.]: IEEE. , out 2011
- MONAGHAN, F.; O’SULLIVAN, D. **Leveraging Ontologies , Context and Social Networks to Automate Photo Annotation**. Proceedings of the semantic and digital media technologies 2nd international conference on Semantic Multimedia. **Anais...** Genoa, Italy: Springer-Verlag. , 2007
- MONAGHAN, F. P. **Context-Aware Photograph Annotation on the Social Semantic Web**. Galway: National University of Ireland, 2008.
- NAAMAN, M. **Leveraging Geo-Referenced Digital Photographs**. Stanford, California: Stanford University, 2005.
- NAAMAN, M.; HARADA, S.; WANG, Q.; GARCIA-MOLINA, H.; PAEPCKE, A. **Context data in geo-referenced digital photo collections**. Proceedings of the 12th annual ACM international conference on Multimedia - MULTIMEDIA '04. **Anais...** New York, New York, USA: ACM Press. , 2004
- NAAMAN, M.; NAIR, R. ZoneTag’s Collaborative Tag Suggestions: **What is This Person Doing in My Phone?** **IEEE Multimedia**, v. 15, n. 3, p. 34–40, jul 2008.
- NAAMAN, M.; PAEPCKE, A.; GARCIA-MOLINA, H. **From Where to What: Metadata Sharing for Digital Photographs with Geographic Coordinates**. Proceedings of The On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE. **Anais...** Catania, Sicily, Italy: Springer. , 2003
- NAAMAN, M.; SONG, Y. J.; PAEPCKE, A.; GARCIA-MOLINA, H. **Automatic organization for digital photographs with geographic coordinates**. Proceedings of the 2004 joint ACM/IEEE conference on Digital libraries - JCDL '04. **Anais...** New York, New York, USA: ACM Press. , 2004

NAAMAN, M.; YEH, R. B.; GARCIA-MOLINA, H.; PAEPCKE, A. **Leveraging context to resolve identity in photo albums**. Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries - JCDL '05. *Anais...* New York, New York, USA: ACM Press. , 2005

O'HARE, N. **Semi-Automatic Person-Annotation in Context-Aware Personal Photo Collections**. Dublin, Ireland: Dublin City University, 2007.

O'HARE, N.; SMEATON, A. F. Context-Aware Person Identification in Personal Photo Collections. *IEEE Transactions on Multimedia*, v. 11, n. 2, p. 220–228, fev 2009.

OGC. **OpenGIS Simple Features Specification For SQL - Revision 1.1**. [S.l: s.n.], 1999

PAPADOPOULOS, S.; ZIGKOLIS, C.; KOMPATSIARIS, Y.; VAKALI, A. Cluster-Based Landmark and Event Detection for Tagged Photo Collections. *IEEE Multimedia*, v. 18, n. 1, p. 52–63, jan 2011.

PRESTI, L. LO; CASCIA, M. LA. Concurrent photo sequence organization. *Multimedia Tools and Applications*, 18 abr 2012.

RIZZO, G.; STEINER, T.; TRONCY, R.; *et al.* **What fresh media are you looking for?** Proceedings of the 2012 international workshop on Socially-aware multimedia - SAM '12. *Anais...* New York, New York, USA: ACM Press. , 2012

RODDEN, K.; WOOD, K. R. **How do people manage their digital photographs?** Proceedings of the conference on Human factors in computing systems - CHI '03. *Anais...* New York, New York, USA: ACM Press. , 2003

SADLIER, D. A.; LEE, H.; GURRIN, C.; SMEATON, A. F.; O'CONNOR, N. E. User-Feedback on a Feature-Rich Photo Organiser. **2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services**, p. 215–218, 2008.

SERDYUKOV, P.; MURDOCK, V.; ZWOL, R. VAN. **Placing flickr photos on a map**. Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval - SIGIR '09. *Anais...* New York, New York, USA: ACM Press. , 2009

SHIMIZU, K.; NITTA, N.; BABAGUCHI, N. **Learning people co-occurrence relations by using relevance feedback for retrieving group photos**. Proceedings of the 1st ACM International Conference on Multimedia Retrieval - ICMR '11. *Anais...* New York, New York, USA: ACM Press. , 2011

STONE, Z.; ZICKLER, T.; DARRELL, T. **Autotagging Facebook: Social network context improves photo annotation**. 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. *Anais...* [S.l.]: IEEE. , jun 2008

STONE, Z.; ZICKLER, T.; DARRELL, T. **Toward Large-Scale Face Recognition Using Social Network Context**. *Proceedings of the IEEE*, v. 98, n. 8, p. 1408–1415, ago 2010.

- TURK, M.; PENTLAND, A. Eigenfaces for Recognition. **Journal of Cognitive Neuroscience**, v. 3, n. 1, 1991a.
- TURK, M.; PENTLAND, A. **Face recognition using eigenfaces**. Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. **Anais...** [S.l.]: IEEE Comput. Soc. Press. , 1991b
- VAILAYA, A.; FIGUEIREDO, M. A. T.; JAIN, A. K.; ZHANG, H.-J. Image classification for content-based indexing. **IEEE transactions on image processing : a publication of the IEEE Signal Processing Society**, v. 10, n. 1, p. 117–30, jan 2001.
- VIANA, W. **Mobilidade e sensibilidade ao contexto para a gestão de documentos multimídia pessoais : CoMMedia**. França: Université de Grenoble, 2010.
- VIANA, W.; BRINGEL FILHO, J.; GENSEL, J.; VILLANOVA-OLIVER, M.; MARTIN, H. PhotoMap: from location and time to context-aware photo annotations. **Journal of Location Based Services**, v. 2, n. 3, p. 211–235, set 2008.
- VIANA, W.; MIRON, A. D.; MOISUC, B.; *et al.* Towards the semantic and context-aware management of mobile multimedia. **Multimedia Tools and Applications**, v. 53, n. 2, p. 391–429, 24 mar 2011.
- VIOLA, P.; JONES, M. **Rapid object detection using a boosted cascade of simple features**. [S.l.]: IEEE Comput. Soc, 2001. p. 1–511–1–518
- WAGENAAR, W. A. My memory: A study of autobiographical memory over six years. **Cognitive Psychology**, v. 18, n. 2, p. 225–252, abr 1986.
- WANG, G.; GALLAGHER, A. C.; LUO, J.; FORSYTH, D. **Seeing People in Social Context : Recognizing People and Social Relationships**. Proceedings of the 11th European Conference on Computer Vision. **Anais...** Crete, Greece: Springer. , 2010
- WANG, X.; ZHANG, T. **Clothes search in consumer photos via color matching and attribute learning**. Proceedings of the 19th ACM international conference on Multimedia - MM '11. **Anais...** New York, New York, USA: ACM Press. , 2011
- WILHELM, A.; TAKHTEYEV, Y.; SARVAS, R.; HOUSE, N. VAN; DAVIS, M. Photo annotation on a camera phone. **Extended abstracts of the 2004 conference on Human factors and computing systems - CHI '04**, p. 1403–1406, 2004.
- YANG, M.; KRIEGMAN, D. J.; MEMBER, S.; AHUJA, N. Detecting Faces in Images : A Survey. **Analysis**, v. 24, n. 1, p. 34–58, 2002.
- YAVLINSKY, A.; SCHOFIELD, E.; STEFAN, R. **Automated Image Annotation Using Global Features and Robust Nonparametric Density Estimation**. Proceedings of the 4th international conference on Image and Video Retrieval. **Anais...** Singapore: Springer-Verlag. , 2005

ZHANG, W.; ZHANG, T.; TRETTER, D. **Clothing-based person clustering in family photos**. 2010 IEEE International Conference on Image Processing. **Anais...** [S.l.]: IEEE. , set 2010

ZHAO, W.; CHELLAPPA, R.; PHILLIPS, P. J.; ROSENFELD, A. Face recognition: A literature survey. **ACM Computing Surveys**, v. 35, n. 4, p. 399–458, 1 dez 2003.

ZIGKOLIS, C.; PAPADOPOULOS, S.; FILIPPOU, G.; KOMPATSIARIS, Y.; VAKALI, A. Collaborative event annotation in tagged photo collections. **Multimedia Tools and Applications**, 21 jun 2012.

Apêndices

Apêndice A

Para avaliar a importância e a significância de cada fator, os seguintes passos foram realizados:

1. Cálculo dos efeitos dos fatores;
2. Cálculo dos resíduos;
3. Cálculo da alocação de variação; e
4. Análise de variância.

Os resultados obtidos dos cálculos dos efeitos e dos resíduos se encontram na Tabela A.1 e na Tabela A.2, respectivamente.

Tabela A.1. Cálculo dos Efeitos (α_f e β_f)

Fator	Nível	Média	Efeito
Algoritmo	RF	39,71	-25,42
Algoritmo	EC	75,76	10,33
Algoritmo	RF + EC	80,83	15,39
Coleção	1	59,76	-5,67
Coleção	2	69,34	3,91
Coleção	3	79,13	13,70
Coleção	4	53,50	-11,94
Média de Todas as Observações		65,43	

Tabela A.2. Cálculo dos Resíduos (e_{ij})

Algoritmo			
Coleção	RF	EC	RF + EC
1	-9,54	4,49	5,05
2	-4,60	4,55	0,05
3	17,6 8	- 10,15	-7,53
4	-3,54	1,11	2,43

A variação da taxa de acerto pode ser devido aos dois fatores e aos erros experimentais. Logo, a variação total foi calculada pela seguinte fórmula:

$$SST = SSA + SSB + SSE,$$

$$SSA = b \sum_j \alpha_j^2,$$

$$SSB = a \sum_i \beta_i^2$$

$$SSE = \sum_i \sum_j e_{ij}^2,$$

nas quais a é o número de níveis do fator A, b é o número de níveis do fator B, SSA é a variação devido ao fator A, SSB é a variação devido ao fator B, SSE é a variação devido aos erros e SST é a variação total.

A porcentagem de variação explicada pelos algoritmos foi

$$100 \times \frac{SSA}{SST} = 100 \times \frac{4020,29}{5823,52} = 69,04\% .$$

A porcentagem de variação devido as coleções foi

$$100 \times \frac{SSB}{SST} = 100 \times \frac{1132,94}{5823,52} = 19,45\% .$$

A variação inexplicada foi

$$100 \times \frac{SSE}{SST} = 100 \times \frac{670,28}{5823,52} = 11,51\% .$$

Analisando essas porcentagens, conclui-se que a escolha do algoritmo é um parâmetro importante na anotação de pessoas em fotografias, porém, ser importante não é o mesmo que ser significativo.

Uma análise de variância é necessária para avaliar a significância de cada fator. Para realizar a análise ANOVA, deve-se assumir que:

- As observações resultantes da combinação dos níveis são independentes, o que pode ser verificado na Figura A.1;
- A distribuição dos resíduos é normal. Para avaliar a normalidade dos resíduos o teste de Anderson-Darling foi aplicado, Figura A.2. Dado que o p-valor do teste de normalidade realizado para os resíduos foi igual a 46,8%, não se pode rejeitar a hipótese nula (H_0) de que os dados são normais a um nível de significância (α) de 5%.
- Os dados nos grupos possuem a mesma variância (homocedasticidade). O teste de Bartlett foi realizado para testar a homocedasticidade dos dados agrupados pelos fatores, Figura A.3e Figura A.4. A hipótese nula do teste é de que as populações possuem a mesma variância. Para um nível de significância (α) de 5%, não se pode rejeitar a homocedasticidade dos dados, pois o p-valor obtido para o fator Algoritmo e para o fator Coleção foi igual a 40,5% e 21,5% respectivamente.

Logo, a análise ANOVA foi realizada e o resultado é apresentado na Figura A.5.

Pode-se concluir que, com um grau de significância (α) de 5%, a escolha do algoritmo pode produzir diferenças significativas na taxa de acerto (p-valor = 0,3%), enquanto a escolha da coleção não produz nenhum impacto significativo (p-valor = 9,5%).

Figura A.1. Gráfico das interações dos fatores

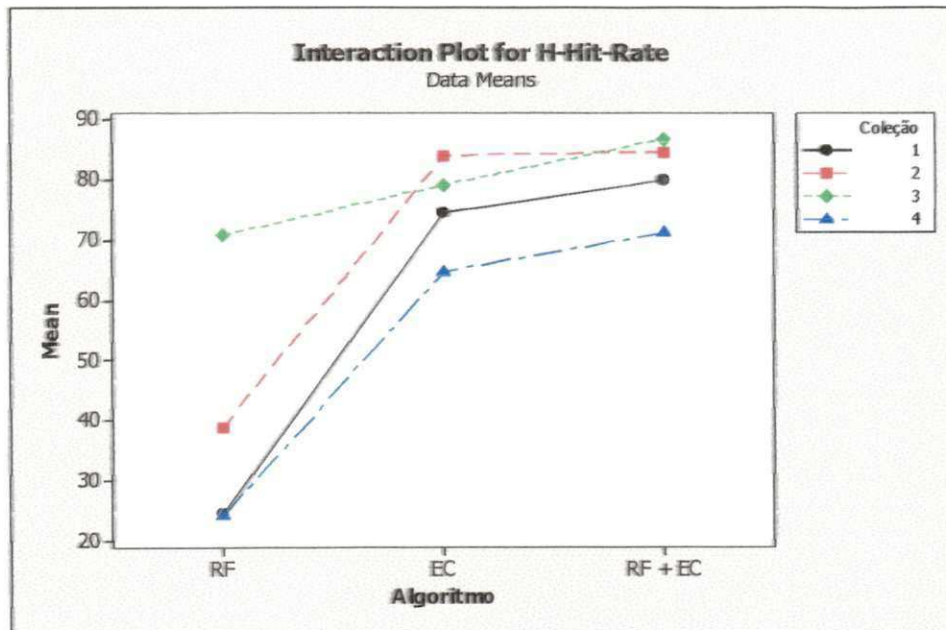


Figura A.2. Teste da normal para os resíduos.

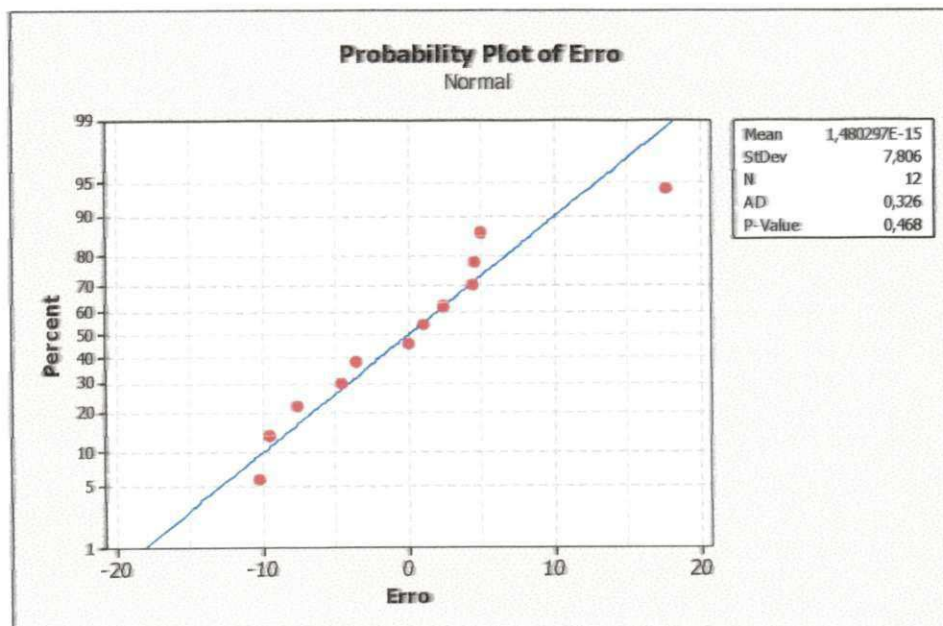


Figura A.3. Teste de homocedasticidade para os níveis do fator algoritmo.

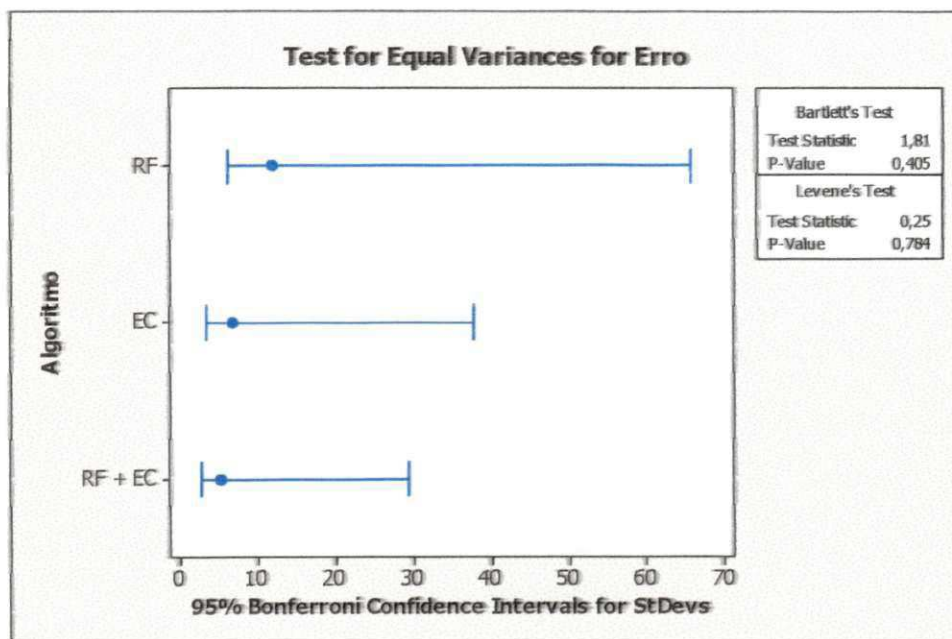


Figura A.4. Teste de homocedasticidade para os níveis do fator coleção de fotografias.

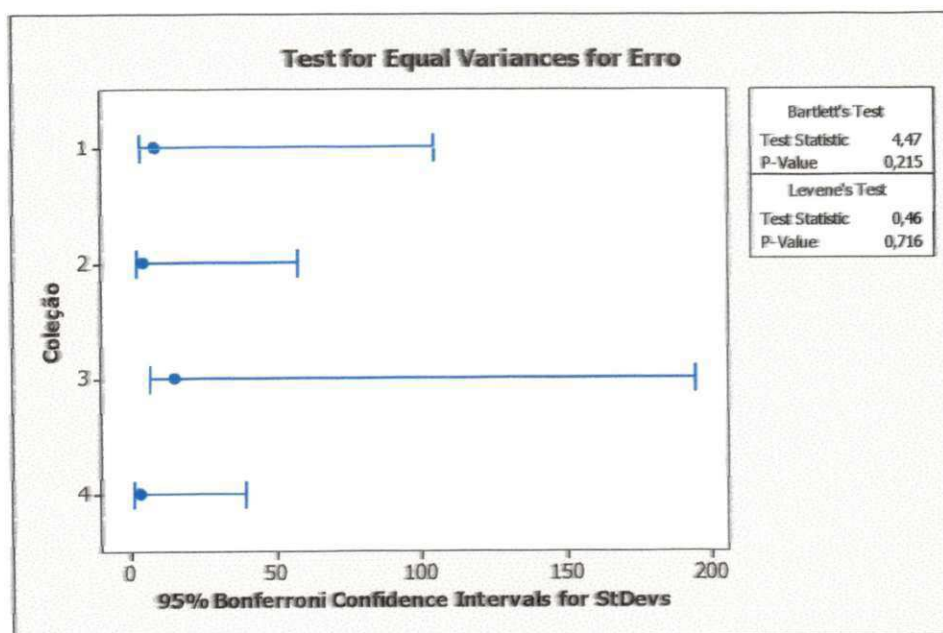
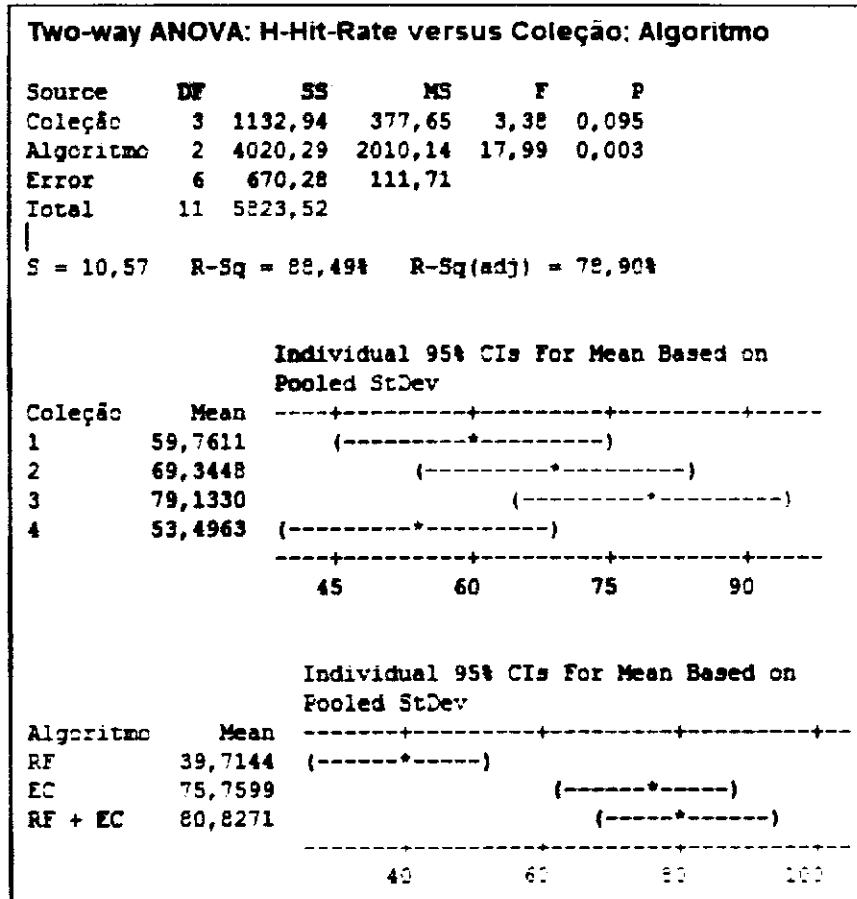


Figura A.5. Resultado da ANOVA.



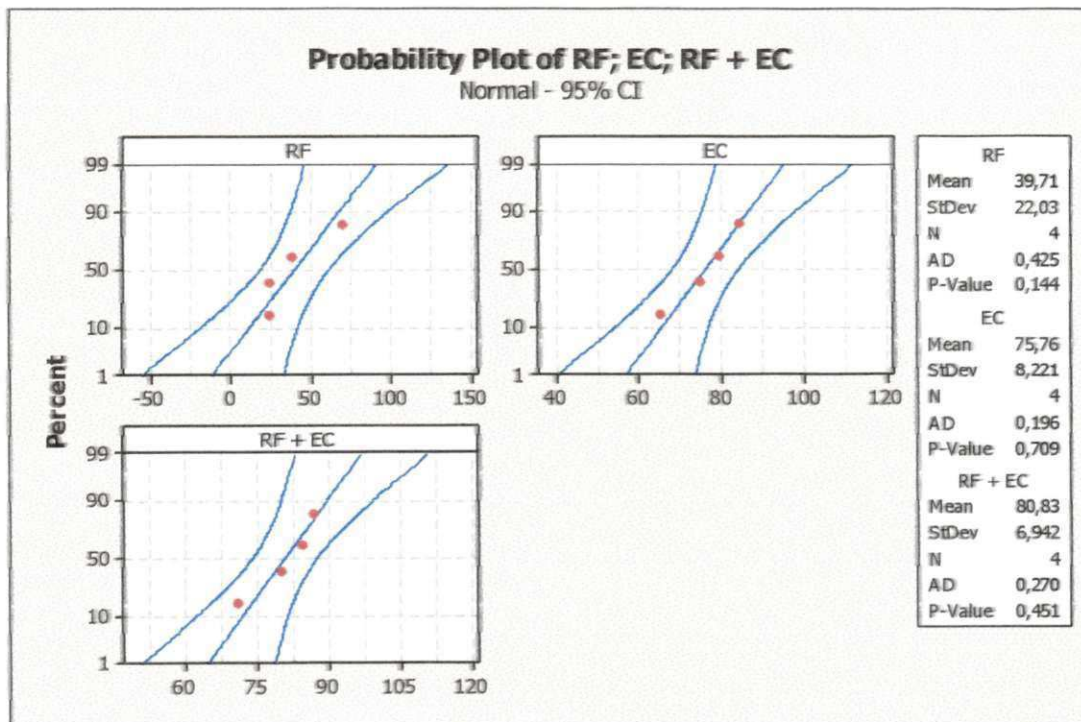
Como consequência da análise de variância, refutamos a hipótese HA_0 de que a taxa de acerto independe do tipo de algoritmo.

Uma vez identificado que o fator Algoritmo produz diferenças **significativas** na taxa de acerto, testes foram realizados para tentar refutar a hipótese HB_0 : o algoritmo que analisa informações de contexto e de conteúdo apresenta uma melhor taxa de acerto do que o melhor dos algoritmos que analisam somente informações de contexto ou somente informações de conteúdo.

O primeiro passo foi verificar a normalidade dos dados para cada nível do fator Algoritmo. O teste de Anderson-Darling foi aplicado e o p-valor para cada nível foi

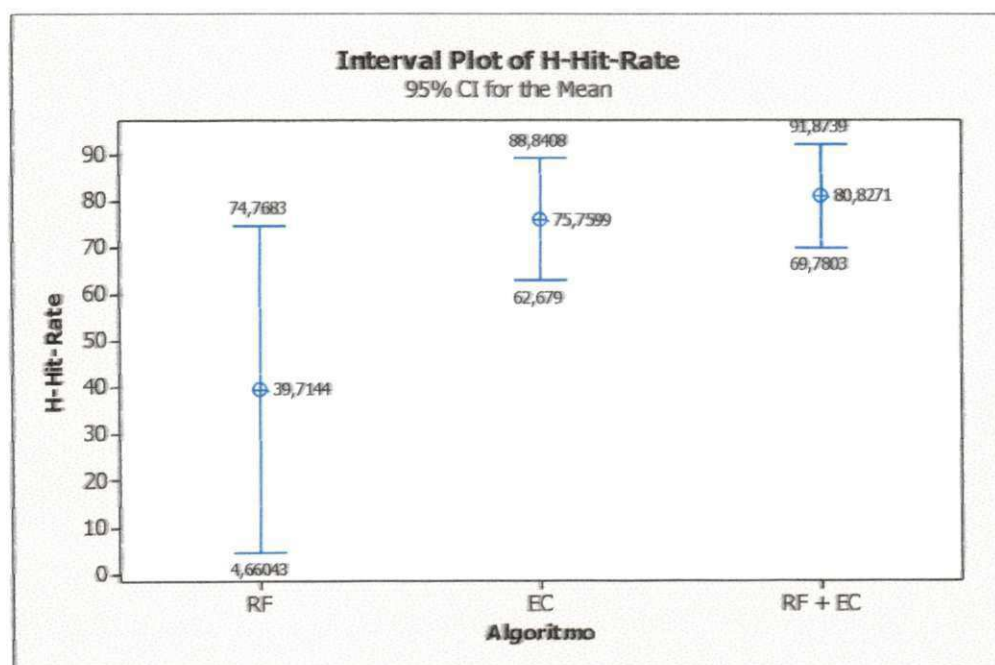
obtido, o que pode ser observado na Figura A.6. Os p-valores obtidos foram 14,4%, 70,9% e 45,1% para os níveis RF, EC e RF + EC, respectivamente.

Figura A.6. Teste da normal para os algoritmos.



Em seguida, assumindo que as populações são normais, pois nenhum p-valor foi menor do que o nível de significância de 5%, os intervalos de confiança foram analisados. Na Figura A.7, pode-se observar que há sobreposição dos intervalos de confiança e um *teste-t de Student* pareado se faz necessário para uma análise mais apurada dos algoritmos.

Figura A.7. Intervalos de confiança para os algoritmos.



Para comparar os algoritmos dois a dois, primeiro foi realizado o *teste-t de Student* pareado para os algoritmos EC e RF. A hipótese alternativa (H_1) utilizada no teste, a taxa de acerto produzida pelo algoritmo EC é maior do que a produzida pelo algoritmo RF, foi aceita, pois o p-valor obtido foi igual a 1,6% e é menor do que $\alpha = 5\%$, Figura A.8.

Por fim, o *teste-t de Student* pareado para comparar os algoritmos EC e RF + EC obteve um p-valor igual a 2,4% (p-valor < 5%), Figura A.9. A hipótese alternativa utilizada no teste, a taxa de acerto produzida pelo algoritmo EC + RF é maior do que a produzida pelo algoritmo EC, foi aceita.

Figura A.8. Teste-T Pareado para comparar os algoritmos EC e RF.

Paired T-Test and CI: EC; RF				
Paired I for EC - RF				
	N	Mean	StDev	SE Mean
EC	4	75,8	8,2	4,1
RF	4	39,7	22,0	11,0
Difference	4	36,05	18,94	9,47
95% lower bound for mean difference: 13,76				
T-Test of mean difference = 0 (vs > 0): T-Value = 3,81 P-Value = 0,016				

Figura A.9. Teste-T Pareado para comparar os algoritmos EC e EC + RF.

Paired T-Test and CI: RF + EC; EC				
Paired I for RF + EC - EC				
	N	Mean	StDev	SE Mean
RF + EC	4	80,83	6,94	3,47
EC	4	75,76	8,22	4,11
Difference	4	5,07	3,11	1,56
95% lower bound for mean difference: 1,40				
T-Test of mean difference = 0 (vs > 0): T-Value = 3,26 P-Value = 0,024				

Apêndice B

Procedimentos armazenados no SGBD PostgreSQL para a execução do algoritmo de geração de lista de sugestões de pessoas presentes nas fotografias e a detecção de eventos compartilhados.

```
CREATE OR REPLACE FUNCTION insert_cooccurrence() RETURNS trigger AS $$
DECLARE
    contact integer;
BEGIN
    FOR contact IN SELECT contactid FROM photo_contact WHERE photoid =
        NEW.photoid AND contactid <> NEW.contactid LOOP
        IF NEW.contactid < contact THEN
            IF EXISTS(SELECT * FROM cooccurrence WHERE contactid1 =
                NEW.contactid AND contactid2 = contact) THEN
                UPDATE cooccurrence SET count = (select count(p1.photoid) from
                    photo_contact p1, photo_contact p2 where p1.contactid =
                    contact AND p2.contactid = NEW.contactid AND p1.photoid =
                    p2.photoid) WHERE contactid1 = NEW.contactid AND contactid2 =
                    contact;
            ELSE
                INSERT INTO cooccurrence (contactid1, contactid2, count) VALUES
                    (NEW.contactid, contact, 1);
            END IF;
        ELSE
            IF EXISTS(SELECT * FROM cooccurrence WHERE contactid1 = contact
                AND contactid2 = NEW.contactid) THEN
                UPDATE cooccurrence SET count = (select count(p1.photoid) from
                    photo_contact p1, photo_contact p2 where p1.contactid =
                    contact AND p2.contactid = NEW.contactid AND p1.photoid =
                    p2.photoid) WHERE contactid1 = contact AND contactid2 =
                    NEW.contactid;
            ELSE
                INSERT INTO cooccurrence (contactid1, contactid2, count) VALUES
                    (contact, NEW.contactid, 1);
            END IF;
        END IF;
    END LOOP;
    RETURN NULL;
END;
$$ LANGUAGE plpgsql;
```

```
CREATE TRIGGER insert_cooccurrence_trigger AFTER INSERT ON photo_contact FOR EACH
ROW EXECUTE PROCEDURE insert_cooccurrence();
```

```
CREATE OR REPLACE FUNCTION delete_cooccurrence() RETURNS trigger AS $$
```

```

DECLARE
    contact integer;
BEGIN
    FOR contact IN SELECT p.contactid FROM photo_contact p WHERE p.photoid =
    OLD.photoid AND OLD.contactid <> p.contactid LOOP
        IF OLD.contactid < contact THEN
            IF EXISTS(SELECT * FROM cooccurrence WHERE contactid1 =
    OLD.contactid AND contactid2 = contact) THEN
                UPDATE cooccurrence SET count = (select
    count(p1.photoid)-1 from photo_contact p1, photo_contact p2 where p1.contactid =
    contact AND p2.contactid = OLD.contactid AND p1.photoid = p2.photoid) WHERE
    contactid1 = OLD.contactid AND contactid2 = contact;
            END IF;
        ELSE
            IF EXISTS(SELECT * FROM cooccurrence WHERE contactid1 = contact
    AND contactid2 = OLD.contactid) THEN
                UPDATE cooccurrence SET count = (select
    count(p1.photoid)-1 from photo_contact p1, photo_contact p2 where p1.contactid =
    contact AND p2.contactid = OLD.contactid AND p1.photoid = p2.photoid) WHERE
    contactid1 = contact AND contactid2 = OLD.contactid;
            END IF;
        END IF;
    END LOOP;
    RETURN OLD;
END;
$$ LANGUAGE plpgsql;

```

```

CREATE TRIGGER delete_cooccurrence_trigger AFTER DELETE ON photo_contact FOR EACH
ROW EXECUTE PROCEDURE delete_cooccurrence();

```

```

CREATE OR REPLACE FUNCTION global_recurrence(cid integer, pid integer)
    RETURNS numeric AS
$BODY$
DECLARE
    rg numeric := 0;
    total numeric;
    com_contato numeric;
    foto RECORD;
    uid integer;
BEGIN
    SELECT * INTO foto FROM photo WHERE id = pid;
    select userid INTO uid FROM album WHERE id = foto.albumid;
    total := (select count(p.*) FROM photo p left outer join album a on a.id =
    p.albumid where a.userid = uid);
    com_contato := (select count(p.*) FROM photo p left outer join photo_contact
    pc ON p.id = pc.photoid left outer join album a on a.id = p.albumid where
    a.userid = uid and pc.contactid = cid);
    IF total > 0 THEN
        rg := (com_contato/total);
    END IF;
    return rg;
END;
$BODY$

```

```

LANGUAGE plpgsql;

CREATE OR REPLACE FUNCTION temporal_recurrence(cid integer, pid integer) RETURNS
numeric AS $$
DECLARE
    rt numeric := 0;
    total numeric;
    com_contato numeric;
    contato RECORD;
    foto RECORD;
    uid integer;
BEGIN
    SELECT * INTO foto FROM photo WHERE id = pid;
    SELECT userid INTO uid FROM album WHERE foto.albumid = id;
    total := (select count(p.*) FROM photo p LEFT OUTER JOIN album a ON
(p.albumid = a.id) where a.userid = uid and age(p.date_time,foto.date_time) > '-1
day');
    com_contato := (select count(p.*) FROM photo p left outer join photo_contact
pc ON p.id = pc.photoid left outer join album a ON p.albumid = a.id where a.userid
= uid and age(p.date_time,foto.date_time) < '1 day' and pc.contactid = cid);
    IF total > 0 THEN
        rt := (com_contato/total);
    END IF;
    return rt;
END;
$$ LANGUAGE plpgsql;

CREATE OR REPLACE FUNCTION spatial_recurrence(contactid_param integer, photoid
integer)
RETURNS numeric AS
$BODY$
DECLARE
    rg numeric := 0;
    total numeric;
    com_contato numeric;
    contato RECORD;
    foto RECORD;
BEGIN
    SELECT * INTO foto FROM photo left outer join album on photo.albumid =
album.id WHERE photo.id = photoid;
    total := (select count(p.*) FROM photo p left outer join album a on a.id =
p.albumid where a.userid = foto.userid and st_distance(p.geom, foto.geom) < 0.01);
    com_contato := (select count(p.*) FROM photo p left outer join photo_contact
pc ON p.id = pc.photoid left outer join album a on a.id = p.albumid where a.userid
= foto.userid and st_distance(p.geom, foto.geom) < 0.01 and pc.contactid =
contactid_param);
    IF total > 0 THEN
        rg := (com_contato/total);
    END IF;
    return rg;
END;
$BODY$
LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION cooccurrence_estimator(cid integer, pid integer)
RETURNS numeric AS $$
DECLARE
    co numeric := 0;
    total numeric;
    com_contato numeric;
    contato integer;
    foto RECORD;
BEGIN
    SELECT * INTO foto FROM photo left outer join album on photo.albumid =
album.id WHERE photo.id = pid;
    FOR contato IN SELECT contactid FROM photo_contact pc WHERE pc.photoid = pid
LOOP
        total := (select count(*) FROM photo_contact c WHERE c.contactid =
contato);
        com_contato := (select c.count FROM cooccurrence c where
(c.contactid1 = cid AND c.contactid2 = contato) OR (c.contactid2 = cid AND
c.contactid1 = contato));
        IF total > 0 AND com_contato/total > co THEN
            co := (com_contato/total);
        END IF;
    END LOOP;
    return co;
END;
$$ LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION face_recognition_estimator(contactid_param integer,
photoid_param integer)
RETURNS numeric AS
$BODY$
DECLARE
    fr numeric := 0;
    score numeric := 0;
BEGIN
    SELECT sf.score INTO score from score_faint sf where sf.contactid =
contactid_param and sf.photoid = photoid_param;
    IF score IS NULL THEN
        score := 0;
    END IF;
    return score;
END;
$BODY$
LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION ranking(pid integer)
RETURNS SETOF contact AS
$BODY$
DECLARE
    contato contact%ROWTYPE;

```

```

    foto photo%ROWTYPE;
    contato_estimador RECORD;
    result Contact;
    count integer;
BEGIN
    SELECT INTO foto * FROM photo p WHERE p.id = pid;
    FOR contato_estimador IN select c.id as contactid, nick,
contact_estimation(c.id, pid) as estimation from contact c where c.userid =
foto.userid AND c.id NOT IN (select pc.contactid from photo_contact pc where
pc.photoid = pid) order by estimation DESC LIMIT 5 LOOP
        SELECT INTO contato * from contact where contato_estimador.contactid
= id;
            return next contato;
        END LOOP;
    return;
END;
$BODY$
LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION contact_estimation(cid integer, photoid integer)
    RETURNS numeric AS
$BODY$
DECLARE
    func varchar;
    es integer;
    gr numeric;
    sql varchar;
    soma numeric := 0;
BEGIN
    FOR es IN select id FROM estimator LOOP
        SELECT sqlfunction INTO func FROM estimator WHERE id = es;
        sql := 'select ' || func || '(' || cid || ', ' || photoid ||
')*w.weight FROM contact_estimator_weight w WHERE w.contactid = ' || cid || ' and
estimatorid = ' || es;
        EXECUTE sql INTO gr;
        IF NOT gr IS NULL THEN
            soma := soma + gr;
        END IF;
    END LOOP;
    RETURN soma;
END;
$BODY$
LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION weighting(cid integer) RETURNS VOID AS
$BODY$
DECLARE
    slope NUMERIC := 0;
    x_ NUMERIC;
    y_ NUMERIC;
    func varchar;
    es integer;
    uid integer;

```



```

sql varchar;

BEGIN
DROP TABLE IF EXISTS arff;
CREATE TEMP TABLE arff(x numeric, y numeric);
select userid into uid from contact where id = cid;
FOR es IN select id FROM estimator LOOP
DELETE FROM arff;
SELECT sqlfunction INTO func FROM estimator WHERE id = es;
--RAISE NOTICE 'Contato: % ; Estimator: %', cid, func;
sql := 'INSERT INTO arff (select '
      || func || '(' || cid || ', p.id), '
      || '(select count(*) from photo_contact pc where
pc.photoid = p.id AND pc.contactid = ' || cid || ') '
      || 'FROM photo p, album a WHERE p.albumid = a.id AND
a.userid = ' || uid || ')';
EXECUTE sql;
DELETE FROM contact_estimator_weight cew WHERE cew.estimatorid = es
AND cew.contactid = cid;
SELECT AVG(x) INTO x_ FROM arff;
SELECT AVG(y) INTO y_ FROM arff;
IF 0 < (select count(*) from arff) THEN
BEGIN
SELECT SUM((x-x_)*(y-y_)) / SUM((x-x_)^2) INTO slope
FROM arff;
EXCEPTION WHEN division_by_zero THEN
--RAISE NOTICE 'caught division_by_zero';
slope := 0.0;
END;
ELSE
slope := 0.0;
END IF;
--RAISE NOTICE 'Peso: %', slope;
INSERT INTO contact_estimator_weight(contactid, estimatorid, weight)
VALUES (cid, es, slope);
END LOOP;
END;
$BODY$
LANGUAGE plpgsql;

CREATE OR REPLACE FUNCTION weighting() RETURNS VOID AS
$BODY$
DECLARE
cid integer;
uid integer;
BEGIN
for uid IN select id FROM photogeouser LOOP
for cid IN select id FROM contact WHERE userid = uid LOOP
PERFORM weighting(cid);
END LOOP;
END LOOP;
END;
$BODY$
LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION update_weighting() RETURNS trigger AS $$
DECLARE
    annotations integer;
BEGIN
    SELECT count(*) INTO annotations FROM photo_contact WHERE id =
NEW.contactid;
    IF (annotation % 10) = 0 THEN
        PERFORM weighting(NEW.contactid);
    END IF;
END;
$$ LANGUAGE plpgsql;

```

```

CREATE TRIGGER update_weighting_trigger AFTER INSERT ON photo_contact FOR EACH ROW
EXECUTE PROCEDURE update_weighting();

```

```

CREATE OR REPLACE FUNCTION photogeo_update_event_similarity(buffertemposegundos
integer, bufferlocalidademetros numeric, peso_pessoas numeric, peso_localidade
numeric, peso_tempo numeric)
RETURNS void AS
$BODY$
DECLARE
    ev1 INTEGER;
    ev2 INTEGER;
    similarity NUMERIC;
BEGIN
    FOR ev1 IN (SELECT DISTINCT id FROM event ORDER BY id ASC) LOOP
        IF ((SELECT COUNT(*) FROM photo WHERE eventid = ev1) > 0) THEN
            FOR ev2 IN (SELECT DISTINCT id FROM event WHERE id > ev1 AND
photogeo_time_event_similarity(ev1, id, buffertemposegundos) > 0) LOOP
                IF ((SELECT COUNT(*) FROM photo WHERE eventid = ev2) > 0) THEN
                    SELECT photogeo_event_similarity(ev1, ev2, buffertemposegundos,
bufferlocalidademetros, peso_pessoas, peso_localidade, peso_tempo) INTO
similarity;
                    IF (similarity > 0.0) THEN
                        DELETE FROM event_similarity WHERE id1 = ev1 and id2 = ev2;
                        INSERT INTO event_similarity VALUES (ev1, ev2, similarity);
                    END IF;
                END IF;
            END LOOP;
        END IF;
    END LOOP;
END;
$BODY$
LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION photogeo_event_similarity(ev1 integer, ev2 integer,
buffertemposegundos integer, bufferlocalidademetros numeric, pp numeric, pg
numeric, pt numeric)

```

```

    RETURNS numeric AS
$BODY$
DECLARE
    sp numeric;
    st numeric;
    sg numeric;
BEGIN
    SELECT photogeo_people_event_similarity(ev1, ev2) INTO sp;
    SELECT photogeo_time_event_similarity(ev1, ev2, bufferTempoSegundos) INTO st;
    SELECT photogeo_geography_event_similarity(ev1, ev2, bufferLocalidadeMetros)
INTO sg;
    return pp * sp + pt * st + pg * sg;
END;
$BODY$
LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION photogeo_geography_event_similarity(ev1 integer, ev2
integer, bufferlocalidademetros numeric)
    RETURNS numeric AS
$BODY$
DECLARE
    box1 geometry;
    areaBox1 numeric;
    box2 geometry;
    areaBox2 numeric;
    areaIntersecao numeric;
    buffer numeric;
BEGIN
    buffer := bufferLocalidadeMetros / 111193;

    SELECT ST_EXPAND(CAST(ST_Extent(p.geom) AS geometry), buffer) INTO box1 FROM
photo p WHERE p.eventid=ev1;
    SELECT ST_EXPAND(CAST(ST_Extent(p.geom) AS geometry), buffer) INTO box2 FROM
photo p WHERE p.eventid=ev2;

    SELECT ST_Area(box1) INTO areaBox1;
    SELECT ST_Area(box2) INTO areaBox2;

    SELECT ST_Area(ST_Intersection(box1, box2)) INTO areaIntersecao;

    IF areaBox1 + areaBox2 > 0 THEN
        return (2 * areaIntersecao) / (areaBox1 + areaBox2);
    ELSE
        return 0;
    END IF;
END;
$BODY$
LANGUAGE plpgsql;

```

```

CREATE OR REPLACE FUNCTION photogeo_people_event_similarity(ev1 integer, ev2
integer)
    RETURNS numeric AS
$BODY$

```

```

DECLARE
    pessoasEvento1 numeric;
    pessoasEvento2 numeric;
    pessoasIntersecao numeric;
BEGIN
    SELECT COUNT(DISTINCT pe.id) INTO pessoasEvento1
    FROM photo p, photo_contact pc, contact c, person pe
    WHERE
        p.eventid = ev1
        AND pc.photoid = p.id
        AND c.id = pc.contactid
        AND pe.id = c.personid;

    SELECT COUNT(DISTINCT pe.id) INTO pessoasEvento2
    FROM photo p, photo_contact pc, contact c, person pe
    WHERE
        p.eventid = ev2
        AND pc.photoid = p.id
        AND c.id = pc.contactid
        AND pe.id = c.personid;

    SELECT COUNT(1) INTO pessoasIntersecao
    FROM (SELECT pe.id
        FROM photo p, photo_contact pc, contact c, person pe
        WHERE p.eventid = ev1
            AND pc.photoid = p.id
            AND c.id = pc.contactid
            AND pe.id = c.personid
        INTERSECT
        SELECT pe.id
        FROM photo p, photo_contact pc, contact c, person pe
        WHERE p.eventid = ev2
            AND pc.photoid = p.id
            AND c.id = pc.contactid
            AND pe.id = c.personid) AS intersecao;

    IF pessoasEvento1 + pessoasEvento2 > 0 THEN
        RETURN (2 * pessoasIntersecao) / (pessoasEvento1 + pessoasEvento2);
    END IF;
    RETURN 0;
END;
$BODY$
LANGUAGE plpgsql;

CREATE OR REPLACE FUNCTION photogeo_time_event_similarity(ev1 integer, ev2
integer, buffertemposegundos integer)
    RETURNS numeric AS
$BODY$
DECLARE
    min1 timestamp;
    min2 timestamp;
    minIntersecao timestamp;
    max1 timestamp;
    max2 timestamp;

```

```
maxIntersecao timestamp;
segundos1 numeric;
segundos2 numeric;
segundosIntersecao numeric;
buffer interval;
BEGIN
buffer := bufferTempoSegundos * INTERVAL '1 second';

SELECT MIN(p.date_time) - buffer, MAX(p.date_time) + buffer
  INTO min1, max1
  FROM photo p
  WHERE p.eventid = ev1;
SELECT MIN(p.date_time) - buffer, MAX(p.date_time) + buffer
  INTO min2, max2
  FROM photo p
  WHERE p.eventid = ev2;

IF min1 < min2 THEN
  minIntersecao := min2;
ELSE
  minIntersecao := min1;
END IF;

IF max1 > max2 THEN
  maxIntersecao := max2;
ELSE
  maxIntersecao := max1;
END IF;

SELECT EXTRACT(epoch FROM max1 - min1) INTO segundos1;
SELECT EXTRACT(epoch FROM max2 - min2) INTO segundos2;
SELECT EXTRACT(epoch FROM maxIntersecao - minIntersecao) INTO
segundosIntersecao;

IF segundosIntersecao < 0 THEN
  RETURN 0;
END IF;

RETURN (2 * segundosIntersecao) / (segundos1 + segundos2);
END;
$BODY$

LANGUAGE plpgsql;
```

Apêndice C

Tabela C.1. Modelos de requisições REST no sistema.

Recurso	Criar	Recuperar	Atualizar	Remover	Listar
User	Método: POST /user/	Método: GET /user/{userId}	Método: PUT /user/{userId} Parâmetros: - username* - password*	Método: DELETE /user/{userId}	Método: GET /users/
	Parâmetros: - username* - password* - name* - mbox* - homepage Resposta: User criado	Resposta: User procurado	Resposta: User atualizado	Resposta: String "Successfully deleted."	Resposta: Set de todos os Users
Event	Método: POST /event/	Método: GET /event/{eventId}	Método: PUT /event/{eventId}	Método: DELETE /event/{eventId}	Método: GET /events/
	Parâmetros: - name* - description - photoId (id para photocover) Resposta: Event criado	Resposta: Event procurado	Parâmetros: - name - description - photoId Resposta: Event atualizado	Resposta: String "Successfully deleted."	Resposta: Set de todos os Events
Contact	Método: POST /user/{userId}/contact/	Método: GET /contact/{contactId}	Método: PUT /contact/{contactId}	Método: DELETE /contact/{contactId}	Método: GET /user/{userId}/contacts/
	Parâmetros: - personId - nickname - depiction Resposta: Contact criado	Resposta: Contact procurado	Parâmetros: - personId - nickname - depiction Resposta: Contact atualizado	Resposta: String "Successfully deleted."	Resposta: Set de todos os Contacts do User com userId

Tabela C.2. Modelos de requisições REST no sistema (continuação).

Recurso	Criar	Recuperar	Atualizar	Remover	Listar
Album	Método: POST	Método: GET	Método: PUT	Método: DELETE	Método: GET
	Parâmetros: - name* - description - photcoverFile (arquivo para foto capa do album)	/album/{albumId}	/album/{albumId}	/album/{albumId}	/user/{userId}/albums/
	Resposta: Album criado	Resposta: Album procurado	Parâmetros: - name - description - photcoverFile Resposta: Album atualizado	Resposta: String "Successfully deleted."	Resposta: Set de todos os Albums do User com userId
Photo	Método: POST	Método: GET	Método: PUT	Método: DELETE	Método: GET
	/album/{albumId}/photo/	/photo/{photoId}	/photo/{photoId}	/photo/{photoId}	/album/{albumId}/photos/
	Parâmetros: - filename*	Resposta: Photo procurada	Parâmetros: - filename* Resposta: Event atualizado	Resposta: String "Successfully deleted."	Resposta: Set de todas as Photos do Album com albumId
	Respostas: Photo criada				
Person	Método: POST	Método: GET	Método: PUT	-	Método: GET
	/person/	/person/{personId}	/person/{personId}		/person/
	Parâmetros: - name* - mbox* - homepage	Resposta: Peron procurada	Parâmetros: - name - mbox - homepage Resposta: Person atualizada		Resposta: Set de todas as Persons
	Resposta: Peron criada				

Tabela C.3. Modelos de requisições REST no sistema (continuação).

Recurso	Criar	Recuperar	Atualizar	Remover	Listar
Contact InPhoto	Método: POST	-	-	Método: DELETE	Método: GET
	/photo/{photoId}/contacts/ / Parâmetros: - contactId* Resposta: String "Successfully created."			/photo/{photoId}/contacts/{contactId} Estados: String "Successfully deleted."	/photo/{photoId}/contacts/ Resposta: Set de todos os Contacts da Photo com photoId