

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

Explorando a localização e orientação de fotografias
pessoais para descoberta de pontos de interesse
baseada em agrupamento

Yuri Almeida Lacerda

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 obtenção do grau de Doutor em Ciência da Computação.

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

Orientadores:

Prof. Dr. Leandro Balby Marinho
Prof. Dr. Cláudio de Souza Baptista

Campina Grande, Paraíba, Brasil
©Yuri Almeida Lacerda, 03/03/2017

Explorando a localização e orientação de fotografias pessoais para descoberta de pontos de interesse baseada em agrupamento

Yuri Almeida Lacerda

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 obtenção do grau de Doutor em Ciência da Computação.

Área de Concentração: Ciência da Computação

Linha de Pesquisa: Sistemas de Informação e Banco de Dados

Orientadores:

Prof. Dr. Leandro Balby Marinho

Prof. Dr. Cláudio de Souza Baptista

Campina Grande, Paraíba, Brasil

Março - 2017

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

L131e Lacerda, Yuri Almeida.
Explorando a localização e orientação de fotografias pessoais para descoberta de pontos de interesse baseada em agrupamento / Yuri Almeida Lacerda. – Campina Grande-PB, 2017.
125 f. : il. color.

Tese (Doutorado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática, 2017.

"Orientação: Prof. Dr. Leandro Balby Marinho, Prof. Dr. Cláudio de Souza Baptista".

Referências.

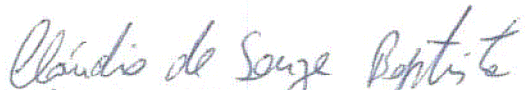
1. Descoberta de Conhecimento. 2. Detecção de Pontos de Interesse. 3. Fotos Orientadas - Aplicação de Descoberta de Conhecimento. 4. Fotos Orientadas - Metadados. 5. Agrupamento de Fotos - Descoberta de Conhecimento. I. Marinho, Leandro Balby. II. Baptista, Cláudio de Souza III. Título.

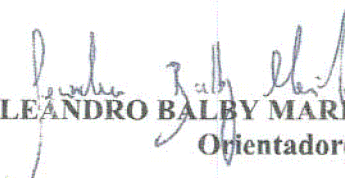
CDU 004.65(043)

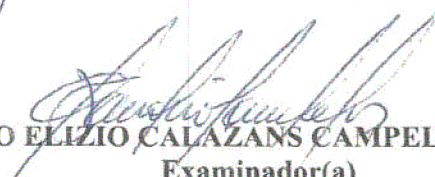
**"EXPLORANDO A LOCALIZAÇÃO E ORIENTAÇÃO DE FOTOGRAFIAS PESSOAIS
PARA DESCOBERTA DE PONTOS DE INTERESSE BASEADA EM AGRUPAMENTO"**


YURI ALMEIDA LACERDA

TESE APROVADA EM 03/03/2017


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


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


CLAUDIO ELIZIO CALAZANS CAMPELO, PhD., UFCG
Examinador(a)


EANES TORRES PEREIRA, D.Sc, UFCG
Examinador(a)

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

MARCELINO PEREIRA DOS SANTOS SILVA, Dr., UERN
Examinador(a)

CAMPINA GRANDE - PB

Resumo

A descoberta de conhecimento a partir de grandes repositórios online de fotografias tem sido uma área de pesquisa bastante ativa nos últimos anos. Isso se deve principalmente a três fatores: incorporação de câmeras digitais e sensores de geolocalização aos dispositivos móveis; avanços na conectividade com a Internet; e evolução das redes sociais. As fotografias armazenadas nesses repositórios possuem metadados contextuais que podem ser utilizados em aplicações de descoberta de conhecimento, tais como: detecção de pontos de interesse (POIs); geração de roteiros de viagens; e organização automática de fotografias. A maioria das abordagens para detecção de POIs parte do princípio que as áreas geográficas onde uma grande quantidade de pessoas capturou fotografias indica a existência de um ponto de interesse. Porém, em muitos casos, os POIs estão localizados a uma certa distância desse local na orientação em que a câmera estava direcionada, e não no ponto exato da captura da fotografia. A maioria das técnicas propostas na literatura não consideram o uso da orientação no processo de detecção de pontos de interesses. Dessa forma, este trabalho propõe novos algoritmos e técnicas para detecção de pontos de interesse em cidades turísticas a partir de coleções de fotografias orientadas e georreferenciadas explorando de diversas formas a orientação geográfica. Esta pesquisa comprovou a importância do uso da orientação nos novos algoritmos voltados para detecção de pontos de interesses. Os experimentos, utilizando uma base de dados real de grandes cidades, demonstraram que os algoritmos considerando a orientação conseguem, em alguns cenários, superar os que não a consideram. Também foram propostas novas métricas de avaliação e uma ferramenta para auxiliar as atividades de descoberta de conhecimento baseada em grandes massas de fotografias.

Palavras-chave: Descoberta de conhecimento. Detecção de pontos de interesse. Fotos orientadas. Metadados. Ponto de interesse. Agrupamento multimídia. Agrupamento de fotos.

Abstract

The knowledge discovery from huge photo repositories has been a very active area of research in the last years. This is due to three facts: the incorporation of digital cameras and geolocation sensors in mobile devices; the advances in Internet connectivity; and the evolution of social networks. The photos stored on those repositories have contextual metadata. Those metadata could be used for many applications of knowledge discovering, such as: Point of Interest (POI) detection; generating of tourist guides; and automatic photo organization. Most approaches for POI detection assume that geographic areas with high density of photos indicate the existence of a point of interest in that area. However, in many cases, the POIs are located in a certain distance of that position in direction where camera was aiming, and not in the exact point of photo shooting. Most of related work do not consider the use of orientation in the process of POI detection. In this way, we propose a set of algorithms and techniques for POI discovery in touristic cities using geotagged and oriented photos collection exploring the geographic orientation in different ways. This research has proven the importance of the usage of orientation in the new algorithms for POI detection. In the experiments with collections related to big cities, the algorithms considering orientation, in several scenarios, have beaten those that do not consider. Also, new metrics of evaluation have been proposed and a new framework to assist all the tasks for knowledge discovery based on huge photo collections.

Keywords: Knowledge discovery. Point-of-interest detection. Oriented photos. Metadata. Point-of-interest. Multimedia clustering. Photo clustering.

Dedico esta tese de doutorado ao meu pai Iari Gomes Lacerda (in memoriam), um grande sábio e estudioso de várias áreas do conhecimento, que dedicou com amor e entusiasmo, em seu tempo livre, o trabalho de pesquisador e historiador. Um homem que nunca buscou os holofortes da fama e do reconhecimento, mas buscou a felicidade na simplicidade, deixando um grande legado de ensinamentos. Dentre eles, a importância da educação na formação do ser humano. Agradeço por ter me mostrado o verdadeiro valor da educação, da família e das coisas simples da vida.

Agradecimentos

Primeiramente, agradeço à Deus pelo dom da vida, por ter me dado saúde e disposição para enfrentar todos os desafios durante o período. Agradeço aos meus pais, por todo o incentivo e investimento em minha educação. Ao meu pai, meu grande mestre e amigo, que muito me ensinou na escolha dos melhores caminhos e tanto me incetivou nesse trabalho de doutorado. A minha mãe, agradeço por todo seu carinho e amor dedicado. Agradeço suas palavras de incentivo e todas as suas orações.

A minha amada esposa Thially, pelo seu apoio e amor em todos os momentos, sempre me apoiando com seus conselhos e carinhos. Agradeço por ter me dado força nas situações mais difíceis e ser essa fortaleza do meu lado.

Aos meus orientadores Dr. Leandro Balby e Dr. Cláudio Baptista, por confiarem em mim e no meu trabalho. Agradeço pelos ensinamentos compartilhados e por me inspirarem como pesquisador.

Aos meus orientadores do doutorado sanduíche Dra. Chiara Renso e Dr. Raffaele Perego, do CNR Pisa, por todo o suporte na Itália durante o período do intercâmbio. Agradeço pelas ótimas contribuições para o meu trabalho de pesquisa e pela excelente receptividade.

Aos examinadores Dr. Cláudio Campelo (UFCG), Dr. Eanes Torres (UFCG), Dr. Marcelino Pereira (UERN) e Dr. Windson Viana (UFC) por todas as contribuições no engrandecimento nesta pesquisa.

Agradeço a Fundação Cearense de Apoio a Pesquisa (Funcap), Coordenação de Aperfeiçoamento Pessoal de Ensino Superior (CAPES) e ao Instituto Federal do Ceará, através do Programa de Formação Doutoral Docente (CAPES) pelo auxílio financeiro.

Ao Instituto Federal do Ceará - campus Crato e todos os colegas servidores pelo apoio e incentivo no desenvolvimento desta pesquisa.

Agradeço aos meus irmãos Marco Aurélio e Melina e familiares por todo o apoio e incentivo no desenvolvimento deste trabalho. Aos meus queridos sobrinhos Eduardo, Mateus e Laís pelos seus beijos, abraços e sorrisos que me incentivaram a seguir em frente.

Aos meus amigos da UFCG: Daniel Leite, Daniel Fireman, Gustavo, Maxwell, José Gildo, Taciano Morais, Ricardo Oliveira, Fabricio Guerra; pelo incentivo, a amizade e os

momentos de descontração. Aos amigos do Bucha's Apartment, Kadu e Hugo por todo o apoio na hospedagem em Campina Grande.

Aos amigos e colegas Hugo Feitosa e Robson Feitosa por colaboração com os trabalhos de pesquisa, incentivo e apoio amigo. Ao meu orientando Johny Moreira por sua colaboração no desenvolvimento da ferramenta PhoCA.

Aos colegas do CNR: Matteo Cattena, Salvatore, Cristina, Franco Nardini e demais; pelos bate-papos, troca de conhecimento e pelas reuniões com vinhos, champagnes, doces, queijos e azeites. Por terem me apresentado um ambiente diferente de trabalho. Ao amigo e *housemate* Vinicius Monteiro pela companhia e amizade.

Aos amigos do Ciência Sem Fronteiras (Pisa, Itália - 2015/2016): André, Mateus, Anderson, Ana, Gustavo e demais; que permitiram matar a saudade do Brasil, pelas viagens e momentos de alegria.

Enfim, agradeço à todos os que me apoiaram direto ou indiretamente em todos os momentos do desenvolvimento desse trabalho. Agradeço à Deus por terem colocado todos vocês no meu caminho.

Conteúdo

1	Introdução	1
1.1	Objetivos	4
1.2	Metodologia de Pesquisa	5
1.3	Contribuições	6
1.4	Visão Geral da Tese	7
2	Fundamentação Teórica	9
2.1	Descoberta de Conhecimento em Banco de Dados	9
2.2	Mineração de Dados	12
2.3	Algoritmos de Agrupamento	14
2.3.1	Introdução às Técnicas de Agrupamento	15
2.3.2	k-means	17
2.3.3	DBSCAN	18
2.4	Mineração de Dados Multimídia	19
2.4.1	Arquitetura de um Sistema para MDM	20
2.4.2	Detecção de Pontos de Interesse	23
2.5	Considerações Finais	25
3	Trabalhos Relacionados	27
3.1	Detecção de Pontos de Interesse	27
3.2	Outras aplicações	32
3.3	Considerações Finais	33
4	Detecção de Pontos de Interesse	35
4.1	O Problema da Detecção de Pontos de Interesse	35

4.2	Compass Clustering	38
4.3	Compass Clustering Lite	40
4.4	Compass Clustering Field of View	41
4.4.1	O campo de visão	41
4.4.2	Descrição do algoritmo	44
4.5	Considerações Finais	45
5	Protótipo - Photo Clustering Analyzer	47
5.1	Introdução	47
5.2	Arquitetura do Sistema	48
5.3	Módulos do Sistema	49
5.4	Considerações Finais	53
6	Experimentos e Resultados	54
6.1	Introdução	54
6.2	Primeira Etapa	55
6.2.1	Base de Dados	55
6.2.2	Metodologia de Avaliação	64
6.2.3	Resultados Experimentais	67
6.3	Segunda Etapa	75
6.3.1	Base de Dados	76
6.3.2	Metodologia de Avaliação	80
6.3.3	Geração das geometricas dos clusters	81
6.3.4	Resultados Experimentais	83
6.4	Considerações Finais	90
7	Conclusões e Trabalhos Futuros	93
7.1	Conclusões	93
7.2	Trabalhos Futuros	95
A	Detalhes de Implementação	104
A.1	Estendendo o PhoCA com um novo algoritmo	104
A.2	Extraindo os metadados	107

B Pontos de Interesses de Referência

Lista de Símbolos

API - Application Programming Interface

CAD - Computer Aided Design

CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CSS - Cascading Style Sheets

EXIF - Exchangeable Image File Format

GPS - Global Positioning Systems

HTML - Hypertext Markup Language

IPTC - International Press Telecommunications Council

KDD - Knowledge Discovery in Databases

LBSN - Location-Based Social Networking Services

MBR - Minimum Bounding Rectangle

MDM - Mineração de Dados Multimídia

OSM - Open Street Map

PhoCA - Photo Clustering Analyzer

PIBIC - Programa Institucional de Bolsas de Iniciação Científica

POI - Point of Interest

SIG - Sistema de Informação Geográfica

XML - Extensible Markup Language

XMP - Extensible Metadata Platform

Lista de Figuras

1.1	Exemplo de agrupamento de fotografias baseadas na localização geográfica.	4
2.1	Passos para um processo de descoberta de conhecimento. A figura foi adaptada de (MAIMON; ROKACH, 2010)	10
2.2	Exemplos de Bases de Dados.	18
2.3	Arquitetura de um sistema de mineração de dados multimídia	20
2.4	Representação dos dados contextuais e visuais	21
2.5	Localização Geográfica das fotografias extraídas na cidade de Paris (pontos pretos) e pontos de interesse (polígonos em verde)	24
4.1	Exemplos de diferentes cenários de captura de fotografia.	36
4.2	Exemplos de localização de fotos capturadas em torno da Torre de Pisa. . .	36
4.3	Fotografia extraída do Flickr com dados de orientação e localização geográfica.	37
4.4	A influência da orientação na descoberta de ponto de interesse baseado em algoritmos de agrupamento.	38
4.5	Exemplo da geração dos pontos de cruzamentos a partir dos segmentos de reta representando a orientação da captura da fotografia.	39
4.6	Representação dos pontos a ser utilizados no agrupamento do Compass Clustering Lite.	41
4.7	Campo de Visão	42
5.1	Arquitetura do PhoCA	48
5.2	O módulo Photo Collector	50
5.3	O módulo Experimentos	51
5.4	O módulo Cluster Analyzer	51

5.5	O módulo Evaluation	52
6.1	Representação gráfica dos passos para a execução dos experimentos.	55
6.2	Fotografias georreferenciadas e orientadas de Paris	56
6.3	Quantidade de fotografias georreferenciadas e orientadas relativas a Paris.	57
6.4	Fotografias orientadas e Pontos de Interesse levantados em Paris.	59
6.5	Coleção Filtrada e Pontos de Interesse sem orientação das fotos	60
6.6	Coleção filtrada e pontos de interesse com a orientação das fotos. Os pontos em vermelho representam os pontos de interesses. Enquanto que os triângulos em verde representam as orientações.	61
6.7	Áreas dos pontos de interesse que são cobertas pelos campos de visão das fotografias. A área coberta está representada em verde, enquanto que a área dos polígonos estão contornadas em rosa.	62
6.8	Coleções de referência (em laranja)	65
6.9	F-measure com o algoritmo DBSCAN	68
6.10	Grupos formados pelo DBSCAN em lilás e grupos de referência em vermelho.	69
6.11	F-measure resultante entre os vários algoritmos.	70
6.12	Grupos formados pelo algoritmo Compass Clustering em lilás e grupos de referência em vermelho.	71
6.13	Grupos formados pelo algoritmo Compass Clustering Lite em lilás e grupos de referência em vermelho.	71
6.14	Exemplo do resultado dos algoritmos no Hôtel des Invalides. O polígono verde no centro da Figura (a) apresenta o ponto de interesse, enquanto que os quadrados em amarelo apresentam as fotografias da base de dados. Na Figura (b) os polígonos em cinza representam os campos de visão das fotografias. Na Figura (c), o polígono laranja representa a coleção de referência, onde os quadrados rosa apresenta os elementos que pertencem a coleção de referência. Na Figura (d) o polígono em azul representa o cluster gerado pelo DBSCAN. Enquanto que na Figura (e) o polígono em vermelho representa o cluster resultante do agrupamento com o Compass Clustering Lite.	73

6.15	Exemplo do resultado dos algoritmos na Basílica Sagrado Coração (Basilique du Sacré-Coeur). O polígono verde no centro da Figura (a) apresenta o ponto de interesse, enquanto que os quadrados em amarelo apresentam as fotografias da base de dados. Na Figura (b), os polígonos em cinza representam os campos de visão de algumas das fotografias que serão consideradas nesse cenário. Na Figura (c), os polígonos em laranja representam as coleções de referência para dois POIs que aparecem na figura: a basílica e o Carrossel do Montmartre (Carrousel du Montmartre), onde os círculos representam os elementos que pertencem aos POIs. Na Figura (d), o polígono em azul representa os clusters gerados pelo DBSCAN. Enquanto que na Figura (e), o polígono em vermelho representa o cluster resultante do agrupamento com o Compass Clustering Lite. Por fim, a Figura (f) apresenta o cluster gerado pelo Compass Clustering.	74
6.16	Número de pontos de cruzamento gerados no Compass Clustering	76
6.17	A distribuição das fotografias (pontos pretos) e os pontos de interesse (polígonos verdes) de Paris, França. A imagem da esquerda representa uma visão global da cidade, enquanto que a imagem da direita representa uma aproximação da área mais turística da cidade.	77
6.18	A distribuição das fotografias (pontos pretos) e os pontos de interesse (polígonos verdes) de Roma, Itália. A imagem da esquerda representa uma visão global da cidade, enquanto que a imagem da direita representa uma aproximação da área mais turística da cidade.	78
6.19	A distribuição das fotografias (pontos pretos) e os pontos de interesse (polígonos verdes) de São Francisco, Estados Unidos. A imagem da esquerda representa uma visão global da cidade, enquanto que a imagem da direita representa uma aproximação da área mais turística da cidade. . .	78
6.20	Geração dos polígonos representando os <i>clusters</i>	82
6.21	Representação gráfica do cálculo da precisão e <i>recall</i>	83
6.22	Resultados - Baseados em Área	85
6.23	Resultados baseado em Área de Paris em formato de mapa. Os POIs da coleção de referência estão representados em laranja.	86

6.24	Resultados baseado em Área de Roma em formato de mapa. Os POIs da coleção de referência estão representados em laranja.	87
6.25	Resultados baseado em Área de São Francisco em formato de mapa. Os POIs da coleção de referência estão representados em laranja.	88
6.26	Estudo de caso em São Francisco. Os pontos de interesse estão representados pelos polígonos em laranja, a localização das fotografias estão representadas como quadradinhos em amarelo e os campos de visão os polígonos em cinza.	91

Lista de Tabelas

2.1	Exemplo de metadados de uma imagem	22
3.1	Aplicações utilizando fotografias georreferenciadas para descoberta de conhecimento	28
3.2	Tabela comparativa entre os trabalhos relacionados sobre detecção de POIs	29
6.1	Área de extração dos dados	56
6.2	POIs extraídos da cidade de Paris no Open Street Map	57
6.3	Relação entre a referência da bússola e a quantidade de fotografias	60
6.4	Relação entre o ponto de interesse e número de fotografias da coleção de referência	62
6.5	Resultado do F-measure do algoritmo DBSCAN	69
6.6	Coleções de fotografias da base de dados.	77
6.7	Categorização dos Pontos de Interesse por cidade	79
6.8	Parâmetros utilizados.	84
6.9	Os melhores parâmetros por cidade e abordagem.	84
B.1	Pontos de Interesse de Paris	110
B.2	Pontos de Interesse de Roma	118
B.3	Pontos de Interesse de São Francisco	123

Lista de Códigos Fonte

A.1	Interfaces principais do PhoCA	104
A.2	Arquivo de configuração dos algoritmos (algorithms.xml)	105
A.3	Implementação de um novo algoritmo de agrupamento	106
A.4	Trecho de retorno de uma consulta no formato JSON.	107

Capítulo 1

Introdução

O crescimento exponencial dos dados digitais e a sua disponibilização na Web tem tornado a descoberta de padrões e conhecimento em grandes massas de dados um processo cada vez mais importante.

A descoberta de conhecimento em bancos de dados, do inglês *Knowledge Discovery in Databases* (KDD), é um tópico multidisciplinar, incluindo Ciência da Computação, que estuda a “análise e modelagem automática e exploratória de grandes repositórios de dados. É um processo iterativo de identificação de padrões válidos, novos, úteis e legíveis de grandes e complexos conjuntos de dados” (MAIMON; ROKACH, 2010).

A grande heterogeneidade e complexidade dos dados digitais produzidos na atualidade e disponibilizados em grandes repositórios torna o processo de KDD mais complexo. Os vídeos, imagens, áudios, planilhas eletrônicas e documentos de texto estão entre alguns desses tipos de dados. Dessa forma, muitos pesquisadores têm se empenhado na área em busca da criação de novos métodos e melhoria dos existentes.

A multimídia tem sido um dos focos de pesquisa na área de descoberta de conhecimento. Isso se deve principalmente à grande produção desse tipo de conteúdo nos últimos anos, além de sua disponibilização na Web, motivados principalmente por: popularização dos dispositivos móveis equipados com câmeras fotográficas; redução dos custos e ampliação da cobertura de acesso à Internet.

O compartilhamento de fotografias é uma das atividades mais importantes nas redes sociais. Segundo os dados apresentados em Ericsson, Qualcomm e Facebook (2013), até o mês de setembro de 2013, a rede social Facebook possuía um compartilhamento médio

diário de aproximadamente 350 milhões de fotografias. Até a publicação desse documento, mais de 250 bilhões de fotografias já haviam sido submetidas nessa rede social.

O dado multimídia é um tipo de dado complexo que necessita de um tratamento próprio para lidar com a descoberta do conhecimento. Por exemplo, quando se trabalha com fotografias, pode-se tentar realizar a descoberta do conhecimento com base em aspectos visuais (SMEULDERS et al., 2000) ou contextuais (HUC; LEVOIR; NONON-LATAPIE, 1997).

Os aspectos visuais são todos aqueles inerentes ao conteúdo da imagem. Os aspectos contextuais são aqueles que podem caracterizar algo sobre o usuário, como por exemplo o modelo da câmera ou o lugar em que a fotografia foi capturada. Segundo Dey e Abowd (1999), "contexto é qualquer informação que pode ser utilizada para caracterizar uma situação de uma entidade. Uma entidade é uma pessoa, lugar ou objeto que é considerado relevante para a interação entre um usuário e uma aplicação, incluindo o usuário e a própria aplicação."

Por exemplo, os algoritmos propostos por Tu e Dong (2010) utilizam os aspectos visuais no processo de KDD em um sistema de recomendação personalizado de moda que leva em consideração estilo, cores favoritas e a cor da pele. Com relação ao contexto, o algoritmo proposto por Yang, Gong e U (2011) usa metadados geo-espaciais (e.g.: coordenadas de latitude e longitude) para detectar pontos de interesse em grandes cidades. Nesse último caso, a informação contextual utilizada seria a localização geográfica armazenada nos metadados das fotografias.

A grande disponibilidade de dados na Web, juntamente com a riqueza dos metadados das fotografias, abre novas oportunidades para aplicações de descoberta de conhecimento e mineração de dados. Alguns outros exemplos desses tipos de aplicação são: organização automática de fotografias (LACERDA et al., 2008), geração automática de roteiros turísticos (POPESCU; GREFENSTETTE; MOËLLIC, 2009) e anotação automática de fotografias (SERGIEH et al., 2012).

Em todas essas aplicações, os algoritmos de agrupamento aparecem como sendo uma das principais técnicas de mineração aplicadas. Por exemplo, a maioria das pesquisas existentes em detecção de pontos de interesse de fotografias digitais georreferenciadas é baseada em versões novas ou modificadas de algoritmos de agrupamento existentes, por

exemplo (KLEINBERG et al., 2009; KISILEVICH; MANSMANN; KEIM, 2010; YANG; GONG; U, 2011).

Na área de sistemas de recomendação, Matyas e Schlieder (2009) e Marinho et al. (2012) propuseram a primeira identificação de áreas de interesse através do agrupamento de fotografias georreferenciadas e, então, recomendaram as áreas de interesses detectadas para os usuários.

O foco deste trabalho está na criação de técnicas voltadas para a detecção de pontos de interesse a partir do uso dos metadados geográficos de grandes bases de dados de fotografias. A execução dessas técnicas exige uma série de atividades, dentre elas: coleta de dados, limpeza, execução de um algoritmo de agrupamento, representação e visualização dos dados. Essas atividades podem ser englobadas em uma ferramenta específica que execute todas as tarefas levando em consideração uma grande massa de dados. Entretanto, durante a execução dessa pesquisa não foi encontrada nenhuma ferramenta *open-source*, extensível e com esse propósito específico.

A maioria das abordagens no estado da arte realiza um agrupamento utilizando principalmente a localização geográfica da câmera no momento da captura da fotografia, através das coordenadas de latitude e longitude, semelhante à Figura 1.1. Entretanto, os locais de interesse das pessoas, na maioria das vezes, estão a uma determinada distância do local de captura da fotografia, e não exatamente no ponto de captura da fotografia, como é considerado por alguns trabalhos. Isto significa que o local de interesse normalmente está localizado, no mínimo, a alguns metros do local da captura e na direção em que a câmera aponta no momento do registro de imagem, semelhante ao apresentado na Figura 1.1(a).

As soluções para detecção de pontos de interesse que utilizam a geo-localização de fotografias em algoritmo de agrupamento, normalmente definem os *clusters* gerados como sendo os POIs detectados. A Figura 1.1(c) mostra como seria uma solução nesse estilo. As áreas desses novos POIs são formadas através do menor polígono concâvo (ou convexo) cobrindo todos os elementos daquele *cluster*.

Outra importante contribuição deste trabalho está na disponibilização de uma ferramenta Web extensível voltada para os pesquisadores da área, afim de facilitar a execução de todo um processo de *clustering* de fotografias.

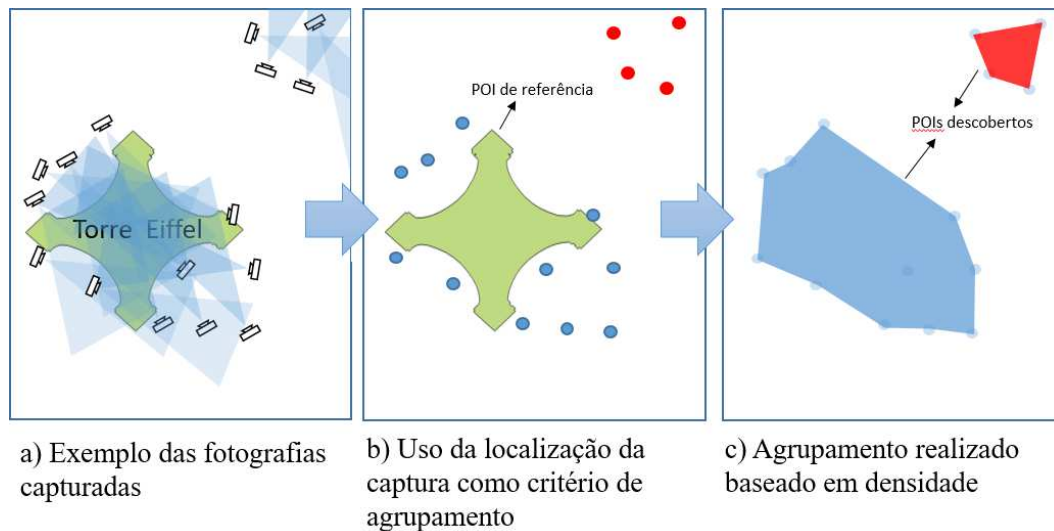


Figura 1.1: Exemplo de agrupamento de fotografias baseadas na localização geográfica.

1.1 Objetivos

Este trabalho tem como hipótese que o processo de detecção automática de pontos de interesse, a partir do uso de algoritmos de agrupamento, pode ser aperfeiçoado a partir do uso da orientação das fotografias digitais.

Portanto, o objetivo geral deste trabalho é propor novas técnicas para detecção de pontos de interesse, utilizando algoritmos de *clustering*, manuseando metadados geográficos de grandes bases de dados de fotografias pessoais georreferenciadas, explorando a orientação geográfica de diversas formas. Além disso, a localização e a orientação devem ser consideradas para a avaliação e representação dos pontos de interesses.

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

- elaboração e implementação de novos algoritmos de agrupamento de fotografias visando a detecção de pontos de interesse, considerando a localização geográfica e orientação das fotografias;
- desenvolver uma nova metodologia de avaliação do processo de detecção de pontos de interesse;
- propor e implementar técnica para a representação geográfica dos pontos de interesses a partir dos *clusters* de fotografias

- propor e implementar um mecanismo automatizado de extração de metadados de fotografias georreferenciadas e orientadas a partir de um repositório de fotografias local ou da Web; e
- desenvolver uma ferramenta Web para auxiliar aos pesquisadores da área na execução de todas as atividades envolvidas no processo de agrupamento (*clustering*) de fotografias.

1.2 Metodologia de Pesquisa

A metodologia de pesquisa adotada nesta tese compreende um conjunto de atividades que foram desenvolvidas ao longo do período da pesquisa. O primeiro passo foi desenvolvido um estudo preliminar dos métodos e técnicas voltadas para a detecção de pontos de interesse a partir do uso de fotografias. Essa atividade foi desenvolvida ao longo de todos os meses de duração do doutorado, pois foi necessária uma frequente atualização acerca do conteúdo que estava sendo estudado. Para isso, as principais publicações da área foram consultadas em busca de novidades na área. Esses trabalhos estão descritos no Capítulo 3.

Em seguida, foram desenvolvidas metodologias para avaliar o processo de detecção de pontos de interesse. Essas metodologias levaram em consideração os métodos existentes na literatura.

A próxima atividade teve o objetivo de elaborar e implementar três novos algoritmos para o agrupamento de fotografias visando a detecção de pontos de interesse. Esses algoritmos levaram em consideração a orientação da fotografia como um critério para o agrupamento.

Logo após, foi proposto e implementado um mecanismo automatizado para extrair e armazenar em banco de dados os metadados de fotografias a partir de um repositório de fotografias local ou da Web.

Destarte, foi proposta e implementada uma nova ferramenta, intitulada PhoCA - Photo Clustering Analyzer, para facilitar a execução de experimentos de agrupamento de fotografias. Isso foi motivado pelo fato de não ter sido encontrada na literatura pesquisada nenhuma ferramenta que auxilie o pesquisador a realizar experimentos de natureza semelhante a proposta desse trabalho.

Algoritmos clássicos da literatura foram implementados e incorporados ao PhoCA

de forma que os parâmetros pudessem ser definidos em tempo de execução. Ademais, a ferramenta permite facilmente incorporar novos algoritmos. Outra funcionalidade importante da ferramenta é a possibilidade de pré-processamentos nos dados para que possam ser utilizados nos experimentos.

Após a execução de um experimento, é importante que haja um mecanismo para facilitar o processo de análise dos dados. Dessa forma, foi implementado um mecanismo de análise estatística dos experimentos e das avaliações.

As atividades de aquisição dos dados, realização dos experimentos, avaliação e análise foram englobadas em um *framework* que permite a extensão dos algoritmos de agrupamento da ferramenta, permitindo que facilmente sejam adicionados ou estendidos novos algoritmos.

Para que fossem realizadas as análises experimentais que comprovem a hipótese levantada neste trabalho, foram definidas as métricas que foram utilizadas nos experimentos. Em seguida, foram realizados os levantamentos de dados utilizando a ferramenta proposta. Destarte, foram executados os experimentos e realizadas as análises dos mesmos.

Durante o desenvolvimento da pesquisa, foram escritos, submetidos e publicados trabalhos científicos em conferências para comprovação do método proposto perante a comunidade científica.

1.3 Contribuições

As principais contribuições dessa pesquisa são:

- novas abordagens para o uso da orientação no processo de detecção de pontos de interesse;
- três novos algoritmos para detecção de pontos de interesse a partir do agrupamento de fotografias;
- novas metodologias de avaliação do processo de detecção de pontos de interesse;
- elaboração de técnica para a geração de áreas geográficas representando os POIs detectados de acordo com os algoritmos;

- estudos comparativos entre algoritmos clássicos da literatura e os algoritmos propostos nesta Tese; e
- proposição, desenvolvimento e disponibilização de uma ferramenta Web extensível para todas as etapas num processo de agrupamento (*clustering*) de fotografias.

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

- “*Exploiting Photo Location and Direction for Clustering-based Points-of-Interest Discovery*” foi aprovado para publicação no ACM SAC 2017, que será realizado em Marraquexe, Marrocos.
- “*Compass clustering: a new clustering method for detection of points of interest using personal collections of georeferenced and oriented photographs*” no WebMedia 2012, realizado em São Paulo.
- “*On Improving Geotag Quality in Photo Collections*” no GEOProcessing 2013, realizado em Nice, França.
- “*PhoCA: An extensible service-oriented tool for Photo Clustering Analysis*” na sessão de Demos e Aplicações no Simpósio Brasileiro de Banco de Dados (SBBD) de 2014 em Curitiba, Paraná.

Uma parte da pesquisa desta Tese foi desenvolvida através do projeto "*The incorporation of compass metadata for the generation of users' trajectory from personal geotagged photos*", durante o Programa de Doutorado Sanduíche no Exterior (PDSE), de Setembro/2015 a Janeiro/2016 no Istituto de Scienza e Technologie A. Faedo, Consiglio Nazionale delle Ricerche (CNR), em Pisa, Itália, sob a supervisão da Dra. Chiara Renso. Esse projeto teve bolsa financiada pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) através do Programa Ciência sem Fronteiras (CsF) sob o processo N° 6393/15-2.

1.4 Visão Geral da Tese

Abaixo estão sumarizados os demais capítulos que compõem esta tese:

- **Capítulo 2 - Fundamentação Teórica:** apresenta uma fundamentação de temas necessários a compreensão do trabalho, tais como: descoberta de conhecimento, mineração de dados, algoritmos de agrupamento e biblioteca digitais multimídia.
- **Capítulo 3 – Trabalhos Relacionados:** apresenta aplicações relacionadas com descoberta de conhecimento a partir de fotografias georreferenciadas, com foco nas que tratam de detecção de POIs.
- **Capítulo 4 - Detecção de Pontos de Interesse:** detalha a problematização do trabalho. Logo após, são apresentadas as soluções propostas neste trabalho.
- **Capítulo 5 – Protótipo - Photo Clustering Analyzer:** apresenta a ferramenta de agrupamento e análise de fotografias digitais georreferenciadas que foi desenvolvida neste trabalho para auxiliar pesquisadores da área.
- **Capítulo 6 – Experimentos e Resultados:** descreve a análise experimental dos algoritmos propostos.
- **Capítulo 7 - Conclusões e Trabalhos Futuros:** apresenta as conclusões com base nos experimentos elaborados nesta pesquisa. Além disso, são propostos novos trabalhos para dar continuidade a pesquisa desenvolvida.
- **Apêndice A - Detalhes de Implementação:** apresenta detalhes de implementação da ferramenta PhoCA.
- **Apêndice B - Pontos de Interesse de Referência:** são apresentados os pontos de interesse utilizados na segunda etapa dos experimentos.

Capítulo 2

Fundamentação Teórica

Este capítulo apresenta os conceitos necessários para o entendimento desta Tese. Dessa forma, como o trabalho lida com áreas multidisciplinares, essas áreas são introduzidas com os seus principais conceitos.

A próxima seção descreve os conhecimentos básicos acerca da área de Descoberta de Conhecimento em Banco de Dados. Logo após, na Seção 2.2 são apresentados os conceitos-chaves de Mineração de Dados. Devido ao trabalho proposto ter foco no uso de algoritmo de agrupamento (*clustering*), na Seção 2.3 são apresentados conceitos básicos, técnicas e algoritmos que são utilizados neste trabalho.

A área de mineração de dados multimídia é detalhada na Seção 2.4, na qual são apresentados conceitos básicos sobre a aplicação de técnicas de mineração de dados multimídia. Ainda nessa seção, são apresentadas algumas aplicações de mineração de dados utilizando fotografias digitais, porém tendo o foco na detecção de pontos de interesse. Por fim, a Seção 2.5 apresenta as considerações finais deste capítulo.

2.1 Descoberta de Conhecimento em Banco de Dados

A grande informatização dos dados ocorrida nos últimos anos aliada aos grandes avanços que aconteceram na Web, principalmente após a criação da Web 2.0, tem produzido uma quantidade massiva de dados nas mais diversas áreas: ciência, indústria, comércio, entretenimento, turismo, entre outras áreas.

Através do uso de técnicas de mineração de dados ou aprendizagem de máquina é

possível descobrir conhecimento através da descoberta de padrões escondidos em dados estruturados ou não estruturados. Por exemplo, supondo a existência de uma imensa base de dados com vídeos de trânsito em uma determinada cidade. Através do uso de técnicas de descoberta de conhecimento será possível descobrir novos padrões em situações de acidentes de trânsito.

Segundo Maimon e Rokach (2010), o termo “Descoberta de Conhecimento em Banco de Dados, do inglês *Knowledge Discovery in Databases* (KDD), é uma modelagem automática e análise exploratória de um grande repositório de dados. O KDD é um processo organizado de identificação de padrões válidos, novos, úteis e compreensíveis de um grande e complexo conjunto de dados”.

A mineração de dados é o núcleo de um processo de KDD, no qual são utilizados algoritmos de inferência para efetuar explorações em um conjunto de dados, com isso são desenvolvidos novos modelos e descobertos novos padrões (MAIMON; ROKACH, 2010).

De acordo com Maimon e Rokach (2010), “o processo de descoberta de dados é iterativo e composto por nove passos”. Sendo este processo dito incremental, ou seja, com exceção do primeiro passo os próximos dependerão do artefato produzido pelo passo anterior. A Figura 2.1 apresenta todos esses nove passos.

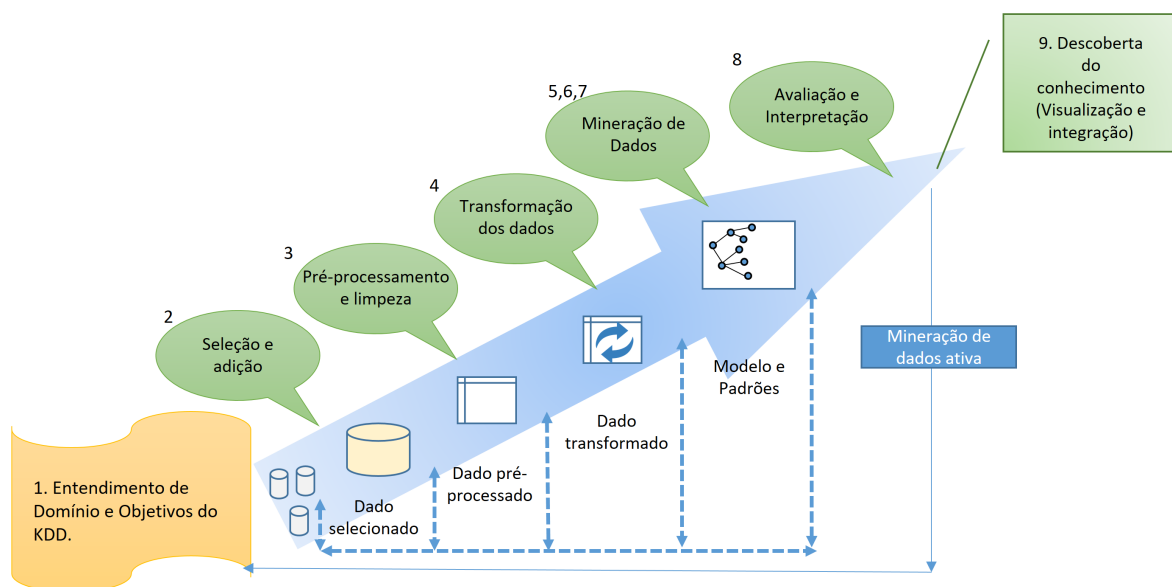


Figura 2.1: Passos para um processo de descoberta de conhecimento. A figura foi adaptada de (MAIMON; ROKACH, 2010)

Os nove passos para o processo são os seguintes:

1. **Entendimento do domínio e objetivos do KDD:** essa fase é responsável pelo entendimento de tudo o que será feito para o projeto de KDD, o entendimento e definição dos objetivos para o usuário final e o ambiente em que o processo de descoberta será utilizado.
2. **Seleção e adição:** O dado que deverá ser utilizado no processo de descoberta de conhecimento deverá ser determinado nesta etapa. Assim, deve-se ter um foco em todos os dados que são necessários para a execução do processo de KDD. É preciso fazer a obtenção de dados adicionais e a definição dos atributos que serão utilizados. Essa fase é crucial em um processo desse tipo, pois a utilização de dados desnecessários, ou até mesmo, a escolha de atributos indevidos, poderá tornar todo o processo um fracasso. O processo de obtenção e armazenamento desses dados é algo que deverá ser pensado cautelosamente, pois a utilização de uma quantidade de dados complexos e massivos poderá tornar o processo de aquisição e armazenamento muito oneroso.
3. **Pré-processamento e limpeza:** A confiança nos dados é algo que é crucial para o KDD. Assim, é importante que se realize o processo de limpeza dos dados de forma a remover ruídos, *outliers* ou dados que são completamente inválidos.
4. **Transformação dos dados:** Os dados devem ser preparados para o processo de mineração de dados. Por exemplo, a redução de dimensões e a transformação de atributos, tal como a discretização de dados numéricos. Essa etapa é fundamental para todo o restante do processo de mineração.
5. **Escolha do tipo de Mineração de Dados:** deve-se escolher um dos tipos de técnicas existentes, por exemplo: classificação, regressão ou agrupamento. Essa técnica deverá ser escolhida de acordo com objetivos do processo de KDD.
6. **Escolha do algoritmo de mineração:** uma vez que a técnica a ser utilizada já tenha sido definida, o próximo passo é realizar a escolha do algoritmo dentro daquela estratégia escolhida. Por exemplo, caso seja escolhida a técnica de agrupamento, deve-se analisar qual dos algoritmos de agrupamento utilizar. Algoritmos de agrupamentos são detalhados na Seção 2.4 deste trabalho.

7. **Execução do algoritmo de mineração:** o próximo passo será empregar aquele algoritmo que foi escolhido e realizar a execução do mesmo. Assim, é possível realizar ajustes no mesmo com base nas primeiras avaliações. Por exemplo, caso um pesquisador ao desenvolver um trabalho de descoberta de conhecimento houvesse optado por uso de um algoritmo como o k-means, detalhado na Subseção 2.4.1, poderia fazer alteração no valor do parâmetro k.
8. **Avaliação e Interpretação:** a avaliação deverá ser feita de acordo com os objetivos traçados no primeiro passo do processo. Assim, deverá ser realizada uma análise se os objetivos traçados no início foram alcançados de acordo com os resultados do algoritmo executado no passo anterior. Assim, poderá, por exemplo, haver alguma mudança nos dados que foram escolhidos na etapa de transformação dos dados.
9. **Descoberta do conhecimento:** o último passo do processo será visualizar a informação que foi descoberta e decidir como fazer uso dessa informação em outro sistema.

É possível observar que as principais etapas do processo de KDD estão relacionadas justamente com a mineração de dados. Dessa forma, a próxima seção trata justamente dos principais conceitos de mineração de dados.

2.2 Mineração de Dados

Segundo Rajaraman e Ullman (2010), "o conceito mais aceito para mineração de dados é a descoberta de modelos para dados". Porém, esse modelo pode ser representado de várias maneiras. Isso vai depender da abordagem utilizada para realizar essa modelagem. Segundo os autores, as mais importantes áreas de direcionamento dessa modelagem são as seguintes:

- **Modelagem Estatística:** a construção de um modelo estatístico que representa a distribuição básica dos dados analisados. Por exemplo, descobrir que um conjunto de dados tem uma distribuição Gaussiana e usar uma fórmula para calcular os parâmetros mais prováveis dessa distribuição.

- **Aprendizagem de Máquina:** alguns tipos de mineração de dados podem utilizar técnicas de aprendizagem de máquina. Nessa área, as técnicas existentes partem do princípio de que uma base de dados deve ser utilizada para fazer o treinamento de um modelo através de algum dos algoritmos de aprendizagem de máquina (RAJARAMAN; ULLMAN, 2010), tais como: redes neurais, máquinas de vetores de suporte, árvores de decisão, etc. Nem todo tipo de situação é possível utilizar esse tipo de abordagem. Por exemplo, o que faz uma pessoa gostar ou não de um filme, um livro ou uma música. Um exemplo típico dessa abordagem é o caso do desafio do Netflix (BENNETT; LANNING, 2007), onde foi oferecida uma boa premiação para quem descobrisse um algoritmo que melhorasse o sistema de recomendação de filmes.
- **Sumarização:** conseguir resumir estruturas complexas em dados simples. Um exemplo tradicional é a forma que o Google utiliza para determinar o “Page Rank” (BRIN; PAGE, 2012) de uma página na Web, mediante uma determinada busca. Esse Page Rank seria um número que representa a importância daquela página para um determinado critério. Os algoritmos de agrupamento (*clustering*) são técnicas bastante conhecida de sumarização, como os algoritmos propostos neste trabalho. Nessa abordagem, os dados são visualizados como pontos em um espaço multidimensional. Esses pontos serão agrupados mediante determinados critérios. Algoritmos como o DBSCAN, k-means e Compass Clustering são exemplos de algoritmos com essa finalidade. Por exemplo, suponha que se deseja analisar dados sobre incidências de epidemias em um país, é possível fazer um mapeamento espacial sobre áreas de maior incidência utilizando algoritmos de agrupamento.
- **Extração de características:** O modelo tradicional baseado em extração de características tenta descobrir os exemplos mais diferentes de um determinado fenômeno e representa os dados por meio desses exemplos. Os mais importantes tipos de extração de características em larga escala são os seguintes:
 - **Conjuntos de itens frequentes:** esse modelo leva em consideração que os dados são formados por “cestas” contendo pequenos conjuntos de itens. Dessa forma, há o cálculo da recorrência de itens que aparecem juntos. Com isso, parte do princípio que se dois itens aparecem sempre juntos é por que há uma relação

entre eles.

- **Itens similares:** os dados são entendidos como uma coleção de conjuntos, na qual existem pares de conjuntos que partilham de elementos em comum. Por exemplo, numa loja on-line os usuários são comparados pelos itens que eles comprarem em comum. Dessa forma, um usuário A que comprou um determinado item é considerado semelhante a um usuário B que fez aquela mesma compra. Então, é recomendado para o usuário A outro produto pelo qual o usuário B também teve interesse. Esse tipo de processo é chamado de “filtragem colaborativa”.

A técnica a ser escolhida de mineração de dados deverá ser adequada ao problema que o usuário esteja desejando solucionar.

2.3 Algoritmos de Agrupamento

Rajaraman e Ullman (RAJARAMAN; ULLMAN, 2010) definem que “clustering é o processo de examinar uma coleção de “pontos” e agrupá-los em *clusters* de acordo com alguma medida de distância (ou similaridade). O objetivo é que os pontos em um mesmo grupo tenham uma distância menor que os pontos que estão em grupos diferentes”, ou seja, o que vai determinar a separação desses grupos seria a técnica utilizada para medir essa distância.

São muitas as aplicações para algoritmos de agrupamento. Por exemplo, no caso deste trabalho, deseja-se aplicar técnicas de agrupamentos para agrupar fotografias de acordo com algumas características a fim de representar lugares que são interessantes para as pessoas. A Figura 1.1 representa fotografias que são agrupadas de acordo com a latitude e longitude utilizando um algoritmo de agrupamento. Dessa forma, uma possibilidade seria utilizar a localização espacial do lugar que a fotografia foi capturada para aplicar uma medida de distância.

2.3.1 Introdução às Técnicas de Agrupamento

Um conjunto de dados adequado para ser realizado o processo de agrupamento deverá ser representado como sendo uma coleção de pontos, os quais estão inseridos em um determinado espaço.

Esse espaço servirá para formular matematicamente as coordenadas que deverão ser representadas de acordo com as dimensões escolhidas para representar cada ponto. Por exemplo, uma possibilidade é representar os dados em um espaço Euclidiano, onde os pontos são vetores de números reais. É importante que, de acordo com o espaço da representação, seja escolhida uma medida adequada para calcular a distância entre os pontos desse espaço.

A distância Euclidiana é muito comum para pontos representados no espaço euclidiano. Ela é representada através do cálculo da raiz quadrada da soma do quadrado da diferença entre cada uma das dimensões das coordenadas. Por exemplo, seja p_1 um ponto com coordenadas (x_1, y_1) e p_2 um segundo ponto com coordenadas (x_2, y_2) . A distância euclidiana entre os pontos p_1 e p_2 seria representada por:

$$d(p_1, p_2) = \sqrt{((x_2 - x_1)^2 + (y_2 - y_1)^2)} \quad (2.1)$$

No caso de coordenadas geográficas, é possível calcular a distância mais precisamente através da Fórmula de Harvesine (SHUMAKER; SINNOTT, 1984). Esta fórmula considera o cálculo da distância numa grande esfera, sendo um cenário mais real para um cálculo de distâncias na superfície terrestre, mesmo que esta não tenha um formato exatamente esférico. Esse cálculo de distância foi utilizado nos experimentos apresentados no Capítulo 6 deste trabalho.

A distância utilizando a Fórmula de Harvesine entre os dois pontos p_1 e p_2 está representada por d_H na Equação 2.2, onde x representa as latitudes dos pontos e y representa as longitudes. A separação de latitude é representada por $\Delta X = x_2 - x_1$ e a separação de longitude é representado por $\Delta Y = y_2 - y_1$. O R representa o raio da esfera, no caso de representar o globo terrestre é possível representar o $R \approx 6.371, 137km$.

$$d_H = 2 \times R \times \arcsen \sqrt{\sen^2\left(\frac{\Delta X}{2}\right) + \cos(x_1) \times \cos(x_2) \times \sen^2\left(\frac{\Delta Y}{2}\right)} \quad (2.2)$$

Há abordagens distintas que vão definir a forma a ser utilizada para agrupar os dados. As

abordagens da literatura são as seguintes:

- **Particionamento:** um algoritmo que utiliza essa técnica organiza os objetos em k *clusters*, onde o desvio total de cada objeto para o centro do seu *cluster* ou para a média da distribuição do *cluster* é minimizado. Cada um dos algoritmos tem sua forma de calcular o desvio com base numa função de similaridade. Alguns exemplos de algoritmos que utilizam esta técnica são: k-means (MACQUEEN, 1967), EM (*Expectation Maximization*) (DEMPSTER et al., 1977) e o algoritmo k-medoid (KAUFMAN; ROUSSEEUW, 2009). O algoritmo k-means é explicado na próxima subseção.
- **Métodos hierárquicos (algoritmos aglomerativos):** os dados são decompostos em uma forma hierárquica (árvore) recursivamente em subconjuntos menores. Existem duas abordagens: “bottom-up” e “top-down”. A primeira inicia com cada elemento em grupos separados e os dados vão sendo aglomerados de acordo com a medida de distância. Na abordagem “top-down”, um *cluster* é quebrado em *clusters* menores de acordo com as medidas até cada objeto estar em um *cluster* ou atingir a condição de parada. Exemplos de algoritmos: AGNES e DIANA (KAUFMAN; ROUSSEEUW, 2009), CURE (GUHA; RASTOGI; SHIM, 1998), CHAMELEON (KARYPIS; HAN; KUMAR, 1999) e BIRCH (ZHANG; RAMAKRISHNAN; LIVNY, 1996).
- **Baseada em densidade:** os algoritmos baseados em densidade vieram com o objetivo de identificar melhor os *clusters* que possuem tamanhos arbitrários. Segundo Han e Kamber (2011), os métodos de particionamento e hierárquico são projetados para encontrar grupos na forma esférica. Assim, alguns algoritmos utilizam a densidade no sentido de identificar *clusters* como sendo áreas mais densas, separando-as pelas regiões menos densas, que são muitas vezes representadas como sendo ruídos. Exemplos: DBSCAN (ESTER et al., 1996), OPTICS (ANKERST et al., 1999) e DENCLUE (HINNEBURG, 2007). Os dois primeiros métodos têm problemas de desempenho quando o número de dimensões é grande.
- **Baseada em grades:** utiliza uma estrutura de dados baseada em grades para ter melhor desempenho. Os dados deverão ser quantizados em um número finito de

células. Alguns exemplos dessa abordagem são: STING (WANG; MUNTZ, 1997), WaveCluster (SHEIKHOLESAMI; CHATTERJEE; ZHANG, 1998) e o CLIQUE (AGRAWAL et al., 1998).

2.3.2 k-means

O k-means (MACQUEEN, 1967) é o algoritmo mais conhecido do tipo particionamento. É assumido que o espaço é euclidiano e é conhecido o número de grupos, o valor de k . O funcionamento do algoritmo é apresentado no Algoritmo 2.3.1. Inicialmente deverá ser definido os k pontos iniciais que vão representar os *centroids* dos possíveis *clusters*. O *centroid* é a representação do elemento médio, ou seja, do elemento central de cada um dos prováveis grupos. Em seguida, o laço deverá associar os outros pontos para os grupos mais próximos, de acordo com a medida de distância.

O *centroid* poderá ser alterado com a adição de novos pontos para um *cluster*. Entretanto, uma vez que foram definidos os *centroids* e só são associados elementos que tem uma certa proximidade não haverá grandes alterações.

Existem duas abordagens para escolha dos *centroids*: na primeira, deve-se pegar pontos que estão a uma maior distância um do outro; na outra agrupa-se uma amostra, uma possibilidade seria hierarquicamente, e escolher pontos mais próximos dos *centroids*.

Algoritmo 2.3.1: Algoritmo k-means

- 1 Escolha k pontos que são prováveis de estar em diferentes clusters;
 - 2 Faça esses pontos como sendo os *centroids* dos seus *clusters*;
 - 3 **repeat**
 - 4 | Atribua (ou re-atribua) cada objeto para o *cluster* que o objeto é mais similar, baseado no *centroid* do cluster;
 - 5 | Atualize o *centroid*, que é calculado baseado na média de objetos do cluster;
 - until** não haja mudança;
-

É possível no final da execução do algoritmo apresentado, executar novamente o algoritmo, porém com os *centroids* sendo os k pontos iniciais dos clusters. Em alguns casos, de acordo com a distribuição do grupo, não haverá mudanças no agrupamento final. Porém, em outras haverá alterações significativas de acordo com a quantidade de novas execuções.

2.3.3 DBSCAN

Suponha um conjunto de pontos distribuídos de acordo com a Figura 2.2. É perceptível a distinção entre os grupos formados, assim como ruídos que não pertencem a nenhum dos grupos. Esses grupos são diferenciados pela densidade dentro do grupo ser maior do que nas áreas externas.

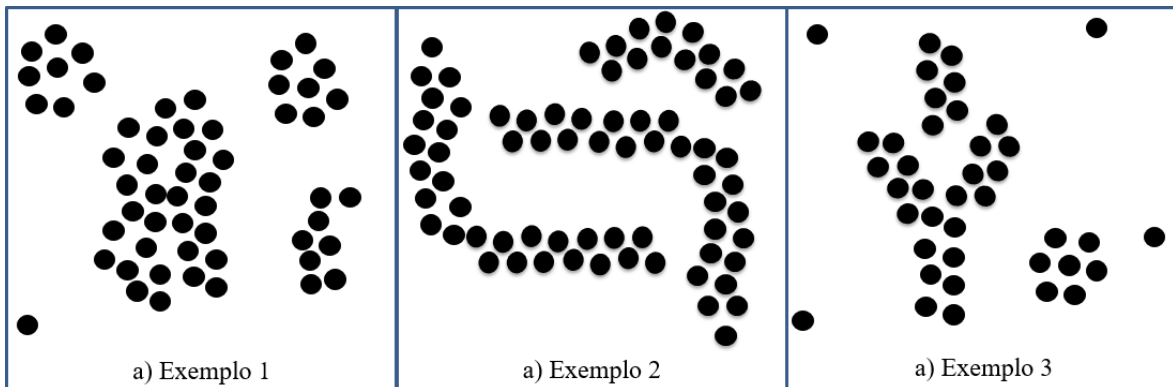


Figura 2.2: Exemplos de Bases de Dados.

Através do algoritmo k-means não seria possível formar grupos como os da Figura 2.2. O DBSCAN (ESTER et al., 1996) é um algoritmo de agrupamento baseado em densidade. Dessa forma, o algoritmo DBSCAN consegue lidar com grupos de formatos arbitrários. Além disso, no DBSCAN não é necessário informar a quantidade de clusters que deverá ser criado, como o parâmetro k , que é exigido no k-means.

A idéia geral do algoritmo DBSCAN, descrito no Algoritmo 2.3.2, é que um *cluster* deverá ser definido como sendo um conjunto máximo de pontos densamente conectados. Dessa forma, é possível encontrar *clusters* com formatos distintos. É possível haver ruídos nesses dados, que não serão classificados em nenhum dos grupos.

Há algumas definições importantes com relação ao DBSCAN: (i) a densidade é justamente o número de pontos que deverá haver dentro de um raio específico (Eps); (ii) um ponto central deverá ter um número mínimo de pontos especificados $MinPts$, dentro do raio (Eps); (iii) um ponto borda é localizado na vizinhança de um ponto central; (iv) um ponto ruído é qualquer ponto que não é classificado como ponto central ou ponto borda.

Algoritmo 2.3.2: Algoritmo DBSCAN

- 1 Selecione aleatoriamente um ponto p ;
 - 2 Realize uma busca com raio Eps ;
 - 3 Identifique todos os pontos densamente conectados a p com relação aos parâmetros Eps e $MinPts$;
 - 4 Se p for um ponto central forma-se um cluster;
 - 5 Se p é um ponto borda e não há pontos densamente conectados a p , é visitado o próximo ponto do conjunto de dados;
 - 6 Continue o processo até todos os pontos do conjunto terem sido analisados;
-

2.4 Mineração de Dados Multimídia

Os avanços computacionais nos últimos anos, principalmente relacionados com a evolução da Web fizeram com que a produção e o compartilhamento de dados multimídia crescessem exponencialmente. Esse fato pode ser constatado pelas próprias redes sociais, nas quais diariamente milhões de fotografias, vídeos, sons, entre outros conteúdos são compartilhados na rede.

O próprio avanço nos dispositivos móveis que possuem a capacidade de capturar fotografias, gravar vídeos, sons e compartilhar diversos tipos de arquivo; tem contribuído bastante para essa grande massa de dados.

A mineração em dados multimídia tem se tornado uma ferramenta bastante utilizada para a descoberta de novos conhecimentos nesse tipo de dados. Dessa forma, houve a criação de uma área chamada Mineração de Dados Multimídia (MDM). Esta é a união entre duas áreas emergentes: multimídia e mineração de dados.

Um dado multimídia é diferente dos dados convencionais pelo fato de necessitar de estruturas mais complexas para ser suportado. No ponto de vista de Banco de Dados, esse tipo de dado é considerado não estruturado e está presente nos mais diversos tipos de aplicação, tais como sistemas médicos, aplicações CAD, sistemas na área de entretenimento, SIGs, bibliotecas digitais, etc. (LACERDA, 2009)

Segundo Wang et al. (2007), mineração de dados multimídia (MDM) remete à descoberta de padrões, extração de regras e aquisição de conhecimento de bancos de dados multimídia.

A mineração de dados multimídia pode ter utilidade em várias áreas, as quais são explicadas a seguir. Na área de bibliotecas digitais (JADHAV; KUMBARGOUDAR, 2007) tem sido utilizadas técnicas para organização de dados multimídia.

Outra aplicação é o uso de vídeos de monitoramento de tráfego de veículos (CHEN et al., 2001) para descobrir conhecimento sobre: áreas de congestionamento, identificação de veículos, detecção de acidentes, etc. Na área de entretenimento, pode-se utilizar mineração de dados para auxiliar na descoberta de preferências dos clientes com intuito de criar recomendações de programação. No campo da saúde, técnicas de mineração de dados podem ser utilizadas na detecção e classificação de imagens médicas para detecção de tumores.

Na área de turismo, alguns pesquisadores têm empenhado esforços em algumas aplicações utilizando fotografias pessoais georreferenciadas, tais como: detecção de pontos de interesse (KLEINBERG et al., 2009; KISILEVICH; MANSMANN; KEIM, 2010; YANG; GONG; U, 2011); geração de roteiros turísticos (POPESCU; GREFFENSTETTE; MOËLLIC, 2009; BRILHANTE et al., 2014); e extração de imagens representativas de pontos de interesse (KENNEDY; NAAMAN, 2008). A detecção de pontos de interesse, tema deste trabalho, será detalhada na Subseção 2.4.2.

2.4.1 Arquitetura de um Sistema para MDM

A Figura 2.3 apresenta a arquitetura de um sistema de mineração de dados multimídia. Os principais estágios de um processo de mineração de dados desse tipo são os seguintes: (1) conhecimento ou entendimento do domínio; (2) a seleção dos dados que se deseja trabalhar, esse estágio também é conhecido como extração das características; (3) pré-processamento dos dados; (4) a descoberta dos padrões; e, por fim, (5) a interpretação dos dados.

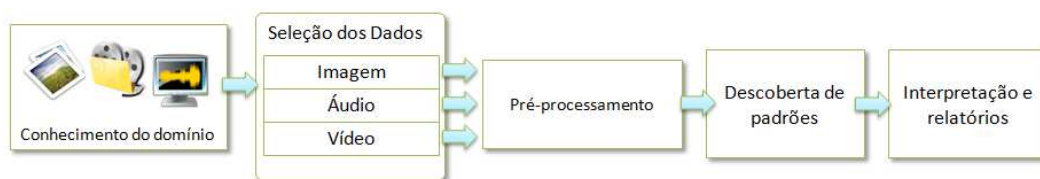


Figura 2.3: Arquitetura de um sistema de mineração de dados multimídia

Como já foi dito anteriormente, um dado multimídia é bem mais complexo do que um dado convencional. Por exemplo, uma imagem possui uma série de dados que podem ser importantes para um processo de mineração. Pode-se fazer uso de aspectos visuais e/ou de conteúdos armazenados nos metadados das mesmas, vide Figura 2.4. Isso vai depender do entendimento do que é preciso para o modelo ou padrão que se deseja descobrir.

Dessa forma, para se trabalhar com o conteúdo visual das imagens, é necessário realizar a aplicação de algoritmos de processamento de imagens em busca de padrões ou objetos desejados. Dependendo da técnica a ser utilizada, esse tipo de algoritmo pode ser bastante oneroso.

Um exemplo de mineração de dados utilizando aspectos visuais seria classificar as fotos que possuem pessoas ou paisagens nas fotografias. Para isso, uma possibilidade seria utilizar algoritmos de reconhecimento facial, por exemplo, o algoritmo proposto em (VIOLA; JONES, 2004), para extrair dados com base no conteúdo visual da imagem e aplicar técnicas de classificação descobrir um novo padrão para classificação de fotografias.



Figura 2.4: Representação dos dados contextuais e visuais

Com relação aos aspectos contextuais, são muitas as características que também podem ser analisadas. Os padrões de armazenamento de metadados, como por exemplo, o Exif (APRIL, 2002) e o IPTC (CORE, 2010), permitem que sejam armazenadas informações relativas ao modelo de câmera utilizado, configurações de lentes, data/hora da captura, localização geográfica, dentre diversas outras, vide Figura 2.4. A Tabela 2.1 apresenta exemplos de metadados armazenados pelo padrão Exif. Na primeira coluna, está identificada a descrição do campo, a segunda coluna é o nome que armazena internamente a informação de acordo com o padrão Exif e a terceira coluna representa um valor de exemplo para aquele campo.

Os dispositivos portáteis modernos capazes de capturar fotografias, tais como

Tabela 2.1: Exemplo de metadados de uma imagem

Campos relativos às configurações da câmera no momento da captura		
Descrição	Nome do Campo	Exemplo
Tempo de exposição	ExposureTime	1/120 sec.
Número do foco	Fnumber	f/2.4
Velocidade do ISSO	ISO Speed Ratings	320
Velocidade do Obturador	ShutterSpeedValue	6.91 EV
Abertura	ApertureValue	2.53 EV
Flash	Flash	No flash function
...
Tags relativas ao GPS		
Referência da Latitude (Norte ou Sul)	GPSLatitudeRef	N
Latitude	GPSLatitude	44.00, 22.12, 0.00
Referência da Longitude (Leste ou Oeste)	GPSLongitudeRef	W
Longitude	GPSLongitude	56.00, 23.98, 0.00
Referência utilizada pelo compasso (norte verdadeiro ou norte magnético)	GPSImgDirectionRef	True North
Angulação do compasso	GPSImgDirection	32.52

smartphones, câmeras digitais, tablets e videogames portáteis cada vez mais estão sendo equipados também com receptores de GPS (Global Positioning Systems) e bússolas digitais. Isso os torna capazes de capturar e armazenar metadados acerca da localização e orientação dos dispositivos no momento da captura das imagens.

Na Tabela 2.1 são apresentados exemplos de valores para os metadados relativos à localização geográfica e da posição da câmera (em relação ao norte geográfico) no momento de captura da fotografia.

Assim, no momento que um usuário captura uma fotografia o mesmo pode estar registrando a localização geográfica daquele espaço geográfico. Uma fotografia poderá ser capturada nos mais diversos tipos de cenários possíveis. Exemplos de algumas situações:

- Um turista visitando uma cidade e captando fotografias dos pontos turísticos. Exemplos: Torre Eiffel em Paris, França; Estátua da Liberdade em Nova Iorque, Estados Unidos; Cristo Redentor no Rio de Janeiro, Brasil; etc.
- Pessoas capturando fotografias de um evento. Exemplos: um show artístico em uma casa de shows; um espetáculo em um teatro, um evento esportivo em uma quadra ou espaço esportivo, um aniversário na casa de algum parente, uma reunião de trabalho em um restaurante, etc.

- Fotografias de objetos de interesse. Exemplos: comidas, quadros, obras de arte, plantas e flores, animais, pessoas, etc.

Através de uma coleção pessoal de fotografias das situações citadas anteriormente é possível inferir uma série de informações sobre o responsável pela captura das fotografias, por exemplo:

- Um usuário que fotografa muitos prédios históricos, certamente tem uma afinidade com arquiteturas antigas.
- Um usuário que costuma capturar fotografias dentro de restaurantes japoneses ou de comidas japonesas, possivelmente tem uma preferência por esse tipo de comida.
- O tempo que o usuário costuma permanecer em determinados tipos de ambientes, por exemplo, se um usuário costuma capturar fotografias num grande intervalo de tempo em parques, pode significar que ele costuma ficar por muito tempo em um lugar desse tipo.
- Os tipos de eventos que costuma ir, por exemplo, se o fotógrafo costuma capturar imagens em shows de um determinado estilo musical é por que certamente gosta daquele estilo de música.
- Quais são as pessoas as quais o fotógrafo mais se relaciona. Por exemplo, os amigos que mais são recorrentes nas suas fotografias, indicam pessoas as quais se tem uma maior afinidade.

Conforme o apresentado anteriormente é possível observar que são muitas as informações que podem ser descobertas a partir de coleções de fotografias pessoais. A subseção 2.4.2 trata de um dos tópicos que tem tido muita atenção dos pesquisadores da área nos últimos anos: a detecção de pontos de interesse.

2.4.2 Detecção de Pontos de Interesse

Conforme o apresentado anteriormente, a partir do uso de coleções de fotografias pessoais é possível descobrir os lugares por onde as pessoas estiveram. Isso se deve às informações contextuais sobre a localização geográfica da câmera no momento da captura das fotografias.

Esse tipo de informação pode ser útil para a recomendação de lugares ou roteiros de viagens com base nas preferências do usuário; descoberta de lugares interessantes nas cidades que muitas vezes não estão em guias turísticos; encontrar lugares legais para fotografar; entre outras aplicações.

Por exemplo, a Figura 2.5 apresenta dois pontos de interesse bem conhecidos, localizados na cidade de Paris: Arco do Triunfo e a Torre Eiffel; e também a distribuição geográfica, por meio da latitude e longitude, dos locais que foram capturadas fotos de uma coleção extraídas do Flickr¹. Observa-se que próximo aos lugares supracitados existe uma maior densidade de pontos. Portanto, essa maior concentração de pontos em determinados lugares seria um indicativo que naquelas proximidades existe um lugar de interesse.

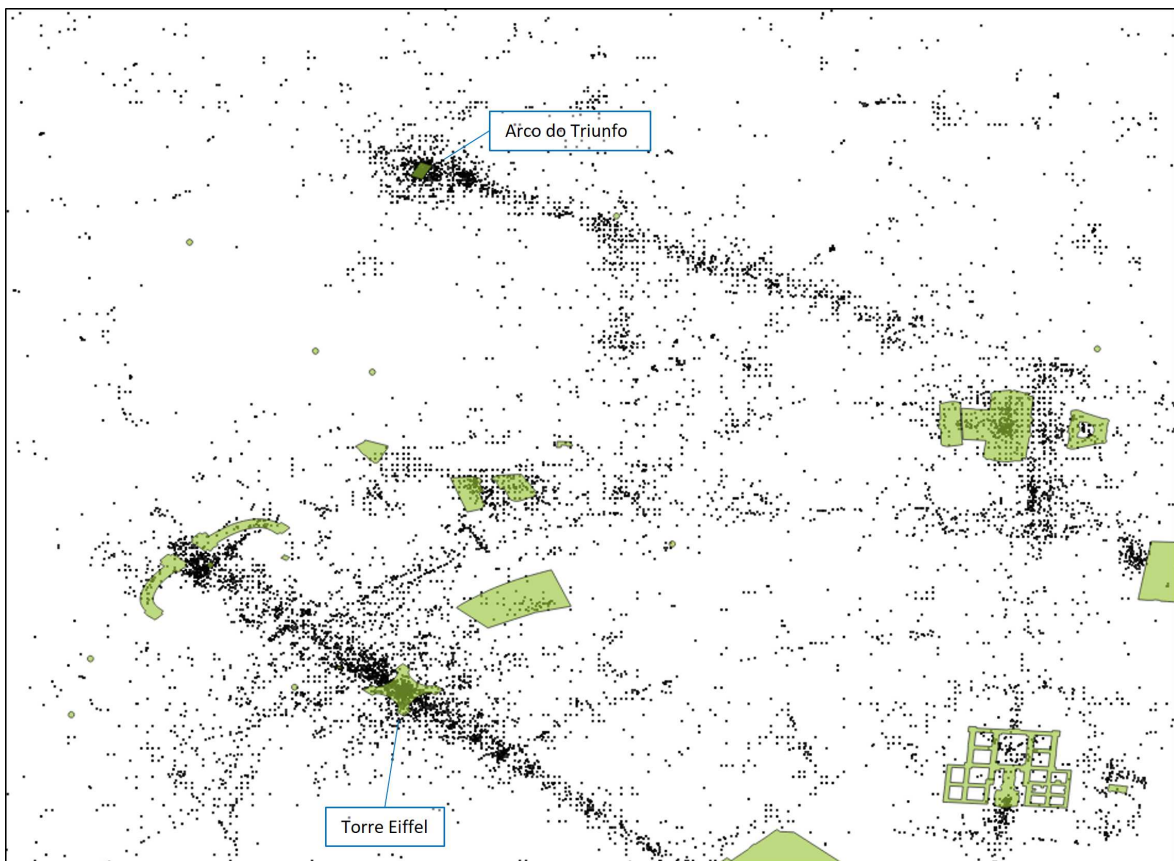


Figura 2.5: Localização Geográfica das fotografias extraídas na cidade de Paris (pontos pretos) e pontos de interesse (polígonos em verde)

Várias pesquisas da literatura já demonstraram que através do uso das fotografias de diversas pessoas é possível descobrir lugares interessantes em grandes cidades.

¹<<http://www.flickr.com>>

A ideia das pesquisas apresentadas basicamente é composta por cinco passos:

- **1º Passo:** extrair todas as fotografias da área geográfica relativa à cidade (ou região) que se deseja descobrir os pontos de interesse;
- **2º Passo:** extrair os metadados da localização geográfica (latitude e longitude) de todas as fotografias georreferenciadas;
- **3º Passo:** realizar algum pré-processamento com os dados com intuito de retirar dados que representem algum ruído;
- **4º Passo:** executar o algoritmo escolhido de agrupamento espacial com parâmetros pré-estabelecidos; e
- **5º Passo:** realizar a avaliação dos *clusters* de acordo com a sua capacidade de representar pontos de interesse previamente conhecidos.

Observa-se que os passos apresentados anteriormente obedecem à arquitetura de um sistema de mineração de dados multimídia, apresentado na Subseção 2.4.1.

2.5 Considerações Finais

Neste capítulo, foi apresentado o embasamento teórico necessário para o entendimento deste trabalho. Inicialmente, tratou-se sobre a descoberta de conhecimento em Banco de Dados, no qual foram apresentados os nove passos do processo de KDD. Em seguida, foram apresentadas as áreas que tratam a mineração de dados, que é o passo central num processo de descoberta de conhecimento.

Logo após, a Seção 2.3 apresentou os conceitos básicos sobre as técnicas de agrupamento, os principais conceitos e dois algoritmos de *clustering* foram apresentados em detalhes: o k-means e o DBSCAN. Por último, tratou-se sobre a mineração de dados multimídia, focando em mineração de fotografias digitais. Foi apresentada a detecção de pontos de interesse utilizando fotografias georreferenciadas.

No próximo capítulo, serão apresentados os principais trabalhos que lidam com descoberta de conhecimento utilizando fotografias georreferenciadas. Além disso, os

trabalhos voltados para detecção de pontos de interesse serão detalhados e comparados com o trabalho proposto nesta Tese.

Capítulo 3

Trabalhos Relacionados

Este capítulo apresenta os trabalhos relacionados ao uso de fotografias georreferenciadas para descoberta de conhecimento. Esse tópico tem sido bastante ativo na comunidade científica nos últimos anos.

Esses trabalhos abordam diferentes aplicações, por exemplo: detecção de pontos de interesse; geração de roteiros de viagens; predição de comportamento do usuário; geração de imagens diversas e representativas de pontos de interesse; extração de informações turísticas; e detecção de eventos. Alguns desses trabalhos estão categorizados na Tabela 3.1.

A próxima seção trata dos trabalhos que estão relacionados à detecção de pontos de interesse. Em seguida, a Seção 3.2 apresenta os outros trabalhos que realizam a descoberta de conhecimento em bases de fotografias georreferenciadas. Por fim, a Seção 3.3 apresenta as considerações finais deste capítulo.

3.1 Detecção de Pontos de Interesse

Vários dos trabalhos investigados voltados para descoberta de pontos de interesse fazem uso de algum tipo de algoritmo de agrupamento espacial (do inglês *spatial clustering*). Isso pode ser constatado em (KLEINBERG et al., 2009; KISILEVICH; MANSMANN; KEIM, 2010; YANG; GONG; U, 2011; CRANDALL et al., 2016). Também há trabalhos que tratam da dedução de outras informações relativas a viagens a partir de conjuntos de fotografias, tais como (ZHENG; CHEN; XIE, 2009).

Dentre os algoritmos utilizados nesses trabalhos estão k-means (RATTENBURY;

Tabela 3.1: Aplicações utilizando fotografias georreferenciadas para descoberta de conhecimento

Aplicação	Trabalhos
Detecção de Pontos de Interesses	(KENNEDY; NAAMAN, 2008) (KISILEVICH; MANSMANN; KEIM, 2010) (LACERDA et al., 2012) (YANG; GONG; U, 2011) (CRANDALL et al., 2016) (THOMEE; ARAPAKIS; SHAMMA, 2016) (LEE; CAI; LEE, 2013)
Geração de Roteiros Turísticos	(POPESCU; GREFFENSTETTE; MOËLLIC, 2009) (De Choudhury et al., 2010) (BRILHANTE et al., 2014)
Organização Automática ou Semi-Automática de Fotografias	(LACERDA et al., 2008) (NAAMAN et al., 2004)
Anotação Automática	(SERGIEH et al., 2012) (SERDYUKOV; MURDOCK; ZWOL, 2009) (LAERE; SCHOCKAERT; DHOEDT, 2010)
Recomendação de Lugares para capturar fotografias	(PHAN et al., 2014)
Detecção de Eventos	(RATTENBURY; GOOD; NAAMAN, 2007) (QUACK; LEIBE; GOOL, 2008) (RATTENBURY; NAAMAN, 2009a) (PAPADOPOULOS et al., 2011)

NAAMAN, 2009b), X-means (POPESCU; SHABOU, 2013), *mean shift* (CRANDALL et al., 2016), DBSCAN (ESTER et al., 1996) e o *Spectral Clustering* (YANG; GONG; U, 2011). Os quatro primeiros são algoritmos de agrupamento espacial. O último é um algoritmo baseado em grafos. A Tabela 3.2 resume algumas das características dos trabalhos que lidam de alguma forma com detecção de pontos de interesse.

Crandall et al. (2016) propuseram uma técnica para o reconhecimento de pontos de interesse em bases de dados de imagens em grande escala. Para essa proposta, eles utilizaram o algoritmo *mean shift* (COMANICIU; MEER, 2002) para descobrir os 500 mais populares POIs no mundo. Eles definiram o parâmetro de raio de 0.0001 graus ($\approx 100m$) e consideraram um máximo de 5 fotos por usuários no mesmo *cluster*. Foi desenvolvida uma técnica que associa uma análise de conteúdo (baseada em dados textuais e na imagem) a uma análise baseada em dados geo-espaciais. Este raio é bastante difícil de estimar, pois

Tabela 3.2: Tabela comparativa entre os trabalhos relacionados sobre detecção de POIs

Característica / Trabalho	Esta Tese	Crandall et al.	Kennedy e Naaman	Kisilevich et al.	Thomee et al.	Yang et al.	Hu et al.	Lee et al.
Algoritmos utilizados	DBSCAN P-DBSCAN Compass Clustering CC LITE CC FOV	Mean shift	k-means	P-DBSCAN	Modelo probabilístico EM DBSCAN X-means Georelevance DBSCAN P-DBSCAN	DBSCAN Spectral Clustering k-means	DBSCAN	DBSCAN
Utiliza lat/long das fotografias na detecção de POIs	SIM	SIM	SIM	SIM	SIM	SIM	SIM	SIM
Utiliza orientação geográfica na detecção de POIs	SIM	NÃO	NÃO	NÃO	SIM	NÃO	NÃO	NÃO
Utiliza conteúdo visual na detecção de POIs	NÃO	SIM	SIM	NÃO	NÃO	NÃO	SIM	NÃO
Utiliza conteúdo textual (tags) na detecção de POIs	NÃO	SIM	SIM	NÃO	NÃO	NÃO	SIM	NÃO
Considera qtd. de fotografias no agrupamento	SIM	SIM	NÃO	SIM	SIM	SIM	NÃO	NÃO
Considera erros de GPS e bússola no agrupamento	NÃO	NÃO	NÃO	NÃO	SIM	NÃO	NÃO	NÃO
Disponibiliza algoritmos e ferramenta para agrupamento	SIM	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO

a área ocupada pelos POIs pode variar bastante. Dessa forma, caso sejam estimados raios maiores que os POIs, serão formados agrupamentos maiores do que determinados POIs, caso contrário, poderá um mesmo POI ser formado por vários pequenos agrupamentos. Outra técnica desenvolvida nesse trabalho foi a descoberta da localização de uma fotografia com base em características visuais e *tags* textuais. O processo de descoberta de POI não foi avaliado. No trabalho proposto nesta Tese foi considerada a orientação da bússola das fotos georreferenciadas no processo de detecção de pontos de interesse. Também, foi proposta a geração de áreas geográficas representando os *clusters* para compará-las com os polígonos dos POIs da coleção de referência.

Kennedy e Naaman (2008) propuseram uma nova técnica baseada em contexto e conteúdo para a geração de resultados diversificados e representativos para consultas de pontos de interesse. Essa técnica leva em consideração as *tags* textuais inseridas pelos usuários nas fotografias, o conteúdo visual das imagens e a sua localização geográfica. Com essa abordagem, consegue-se obter uma associação das *tags* que representam lugares de interesses em uma cidade. O algoritmo de clustering k-means foi utilizado para agrupar as fotografias com base nas características visuais das imagens. Esse trabalho apresentou bons resultados se comparado a resultados derivados somente de *tags* associadas e metadados

sozinhos.

Há trabalhos que fazem o uso do DBSCAN para a detecção de pontos de interesse a partir de fotografias georrefenciadas (KISILEVICH; MANSMANN; KEIM, 2010; THOMEE; ARAPAKIS; SHAMMA, 2016; LEE; CAI; LEE, 2013). Entretanto, este algoritmo não permite que haja grupos com variações de densidade (HAN; KAMBER, 2011). No intuito de solucionar esse problema, (KISILEVICH; MANSMANN; KEIM, 2010) propuseram o algoritmo P-DBSCAN.

O P-DBSCAN é um algoritmo voltado para a análise de lugares e eventos utilizando coleções de fotografias georreferenciadas. Ele foi desenvolvido baseado no DBSCAN, onde a letra P do P-DBSCAN significa "*photo*". Esse algoritmo introduz dois novos conceitos: *limiar de densidade* e *densidade adaptativa*. O primeiro, leva em consideração o número de pessoas na vizinhança, ou seja, ele cria grupos considerando o número mínimo de fotógrafos que capturaram fotografias. Enquanto que o outro permite realizar convergência em grupos nas áreas com uma alta densidade demográfica. Dessa forma, esse algoritmo se propõe a obter *clusters* mais significativos por evitar grupos que apenas uma ou poucas pessoas capturaram fotografias. Além disso, esse algoritmo também se propõe a reduzir os grupos maiores que possuem uma grande diferença na densidade interna. Esse trabalho realizou experimentos utilizando fotografias na cidade de Washington comprovando a eficiência dos novos conceitos.

No trabalho de Yang, Gong e U (2011) foi proposto um algoritmo de agrupamento adaptativo para detecção de pontos de interesse baseado no agrupamento espectral, onde os objetos são representados como grafos não-direcionais e ponderados. Além disso, há uma etapa de reforço que leva em consideração também os aspectos temporais das fotografias. Nessa abordagem, o uso de parâmetros é abolido. Tal solução foi aplicada em coleções de fotografias extraídas do Flickr contendo 216.092 fotografias georrefenciadas das cidades de Paris, Hong Kong e Nova Iorque. Na maioria das análises realizadas no trabalho, esta técnica apresentou melhores resultados em relação aos demais trabalhos analisados na literatura. Os autores propuseram uma avaliação baseada numa área relativa as fotografias capturadas dentro de um percentual da área do MBR (Minimum Bounding Rectangle) dos POIs. Foi considerado que as fotografias que estavam dentro dessa área eram relativas aos pontos. As métricas foram calculadas com base nisto. Nas abordagens consideradas nesta Tese, somente

as fotografias que realmente estavam direcionadas para os POIs são consideradas sendo relativas aos mesmos. Portanto, diferentemente deste trabalho, a orientação foi considerada como critério para melhorar a associação das fotos com os POIs.

Thomee, Arapakis e Shamma (2016) propuseram um método estatístico para identificar a localização geográfica e a geometria de pontos de interesse utilizando os metadados de localização e a orientação da bússola de fotografias georreferenciadas. Diferente da solução proposta neste trabalho, a solução deles considera o erro dos sensores. Os autores analisaram apenas 105 pontos de interesse pelo mundo, enquanto na solução proposta nesta Tese lida com a identificação de todos os pontos de interesse em grandes metrópoles. Nesse trabalho é proposto considerar no máximo 10 fotografias de cada fotógrafo num POI, onde devem existir ao menos 4 fotógrafos diferentes em cada ponto de interesse. No trabalho proposto nesta Tese os *clusters* foram avaliados variando-se o número de fotógrafos por POI. Além disso, o trabalho aqui proposto disponibiliza uma ferramenta voltada para cientistas da área executar todas as atividades de um processo de *clustering* de fotografias.

Hu et al. (2015) propuseram um *framework* para a extração e entendimento das áreas urbanas de interesse (*areas of interests* - AOI) baseadas em fotos georreferenciadas. Eles utilizaram o algoritmo DBSCAN para agrupar através das coordenadas geográficas (lat/long) das fotografias. Os autores extraíram tags textuais e fotos para representar os AOIs. Além disso, um protótipo foi implementado como prova de conceito. Nesse trabalho, os autores investigaram a evolução espaço-temporal das cidades em países desenvolvidos e em desenvolvimento. Diferentemente, em nosso trabalho foi proposta também uma solução que considera um número mínimo de fotógrafos por cluster e que também considera a orientação das fotografias.

Lee, Cai e Lee (2013) utilizaram fotografias georreferenciadas de Queensland, segundo maior estado da Austrália, para detectar pontos de interesse e padrões de captura de imagens pelos turistas. Foram realizadas análises em nível global e local dos pontos de interesse. Para isso, foi utilizado o algoritmo DBSCAN variando os parâmetros de acordo com a análise desejada. Os autores levantaram uma série de avaliações de acordo com a duração da estadia, estações do ano, dia e noite, entre outras.

3.2 Outras aplicações

De Choudhury et al. (2010) propuseram uma abordagem voltada para o usuário para construção de itinerários de viagens automaticamente utilizando os dados geo-temporais deixados por milhões de turistas. Primeiramente, as fotografias foram filtradas para somente serem consideradas fotografias de turistas, com base no casamento com pontos de interesse e em ter um percurso associado. Em seguida, foram gerados grafos com base no percurso de cada “turista” e no tempo de permanência em cada local. A descoberta de um roteiro é realizada levando em consideração a popularidade dos POIs, utilizando um algoritmo guloso de Chekuri e Pal (CHEKURI; PAL, 2005) para solução do problema.

Esses itinerários foram avaliados por 250 pessoas utilizando um sistema de crowdsourcing: Amazon Mechanical Turk (<http://www.mturk.com>). Foram gerados roteiros nas seguintes cidades: Barcelona, Londres, Nova Iorque, Paris e São Francisco. Os usuários avaliaram os roteiros com base em roteiros fornecidos por empresas de turismo profissionais. O resultado dessa avaliação demonstrou que os roteiros gerados automaticamente foram tão bons, ou em alguns casos até melhor, que os roteiros das empresas.

Clements et al. (2010) propuseram uma nova abordagem para recomendação de atrações turísticas em uma cidade com base no histórico das suas fotografias capturadas em outras cidades. Assim, inicialmente é traçado um perfil do usuário com base nos lugares onde ele capturou fotografias. Foram extraídos dados acerca das dez cidades mais visitadas: Londres, Nova Iorque, Paris, São Francisco, Los Angeles, Roma, Chicago, Washington, Barcelona e Berlim. Dessa forma, para recomendar lugares para um determinado usuário é realizada uma ordenação dos lugares mais visitados na cidade, com base em usuários que já visitaram os mesmos lugares que ele em uma dada cidade. Foi utilizado o algoritmo *mean shift* para a descoberta dos lugares mais visitados nas cidades.

O Antourage, proposto em Jain, Seufert e Bedathur (2010), é um sistema para recomendação de roteiros turísticos utilizando dados de geotags extraídos de repositórios web de fotografias. Dessa forma, o sistema requisita ao usuário uma localização inicial e uma distância máxima a ser percorrida. Então, é gerada uma rota turística que não exceda tal limite. Segundo os autores, até o momento desse trabalho, na literatura só se tinha conhecimento de trabalhos que levavam em consideração os pontos de interesse e a duração

da visita. Na abordagem proposta os dados são transformados em um grafo valorado, onde o peso é a distância entre os nós. A solução é baseada nos Sistemas de Formigas Máximo-Mínimo (Max-Min Ant Systems) (STÜTZLE; HOOS, 2000).

Sergieh et al. (2012) apresenta técnicas de como extrair informações válidas acerca de viagens (lugares visitados, tempo gasto e lugares panorâmicos), utilizando dados extraídos dos metadados de coleções de fotografias georreferenciadas. Inicialmente, foi realizado um levantamento de possíveis pontos de interesse da Wikipedia. Para descobrir os lugares visitados, inicialmente é necessário levantar uma lista de possíveis lugares e fazer uma extração de pontos de interesse da Wikipedia. Em seguida, foi desenvolvida uma nova técnica para computar visitas nos pontos de interesse, que leva em consideração o tempo de visitação de cada lugar.

Phan et al. (2014) propõe um novo algoritmo colaborativo para sugestões personalizadas de lugares onde as pessoas podem capturar fotografias. Os pontos de interesse são detectados nesse trabalho a partir da discretização das coordenadas geográficas (latitude e longitude) através de uma função cartográfica de *hashing*. Os autores justificam que essa abordagem tem vantagens em relação aos algoritmos de agrupamento, por poder especificar o tamanho máximo dos grupos.

3.3 Considerações Finais

De acordo com os trabalhos levantados, é possível perceber que descobrir conhecimento a partir de bases de dados de fotografias georreferenciadas tem sido um tópico bastante ativo pelos pesquisadores da área. A detecção de POIs tem sido tema ou parte do problema de vários trabalhos.

A principal contribuição dos algoritmos dessa Tese é avaliar a utilização da direção no processo de detecção de POIs. Diferentemente das abordagens apresentadas, neste trabalho são propostos três diferentes formas de explorar a orientação geográfica das fotografias. Além disso, foram propostas novas metodologias de avaliação para esse tipo de problema. Foram exploradas coleções de fotografias relativas a três grandes cidades turísticas. Além disso, foi proposto uma nova ferramenta Web extensível voltada para a execução de todo um processo de mineração de dados em fotografias, facilmente adaptável a inserção de novos

algoritmos.

O próximo capítulo detalha a problematização da detecção de pontos de interesse utilizando fotografias georreferenciadas e apresenta os algoritmos propostos neste trabalho.

Capítulo 4

Detecção de Pontos de Interesse

Neste capítulo, são apresentadas as novas abordagens para descoberta de pontos de interesse, tema principal desta Tese de Doutorado.

Dessa forma, este capítulo está descrito como segue. A Seção 4.1 apresenta o problema da detecção de pontos de interesse. Em seguida, na Seção 4.2 é apresentado o algoritmo base que é proposto nesta Tese, o Compass Clustering, voltado para detecção de POIs através do agrupamento de fotografias georreferenciadas e orientadas. Em seguida, na Seção 4.3 é descrito o Compass Clustering Lite, uma versão mais simplificada do algoritmo.

Logo após, a Seção 4.4 apresenta o algoritmo Compass Clustering Field of View (CC FOV), que considera o modelo de campo de visão no processo de detecção de pontos de interesse.

Por último, a Seção 4.5 apresenta as considerações finais deste capítulo. As avaliações dos respectivos algoritmos são apresentadas no Capítulo 6.

4.1 O Problema da Detecção de Pontos de Interesse

A maioria dos trabalhos que tratam acerca da detecção de pontos de interesse, conforme apresentado no capítulo anterior, utiliza algoritmos de *clustering* baseado em densidade com as coordenadas de latitude e longitude como entrada de dados.

Dessa forma, utilizar somente a localização geográfica no processo de detecção de pontos de interesse pode ocasionar inferências inexatas, pois esses pontos, em grande parte das situações, não estão localizados no lugar exato em que a fotografia foi capturada,

principalmente quando isso se dá em ambientes ao ar livre.

A Figura 4.1 apresenta duas situações de captura de fotografias em um dos mais famosos pontos turísticos do mundo: a Torre de Pisa. Essa torre está localizada na Praça dos Milagres (*Piazza dei Miracoli*) em Pisa, Itália. Nessa praça estão os monumentos mais importantes de Pisa, apresentados na Figura 4.2, tais como: batistério (I), cemitério monumental (II), catedral (III) e a Torre de Pisa (IV).

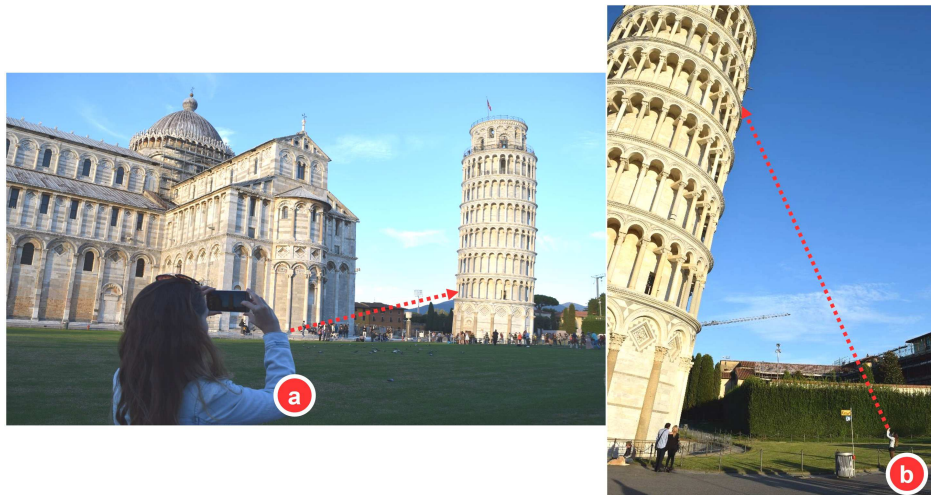


Figura 4.1: Exemplos de diferentes cenários de captura de fotografia.

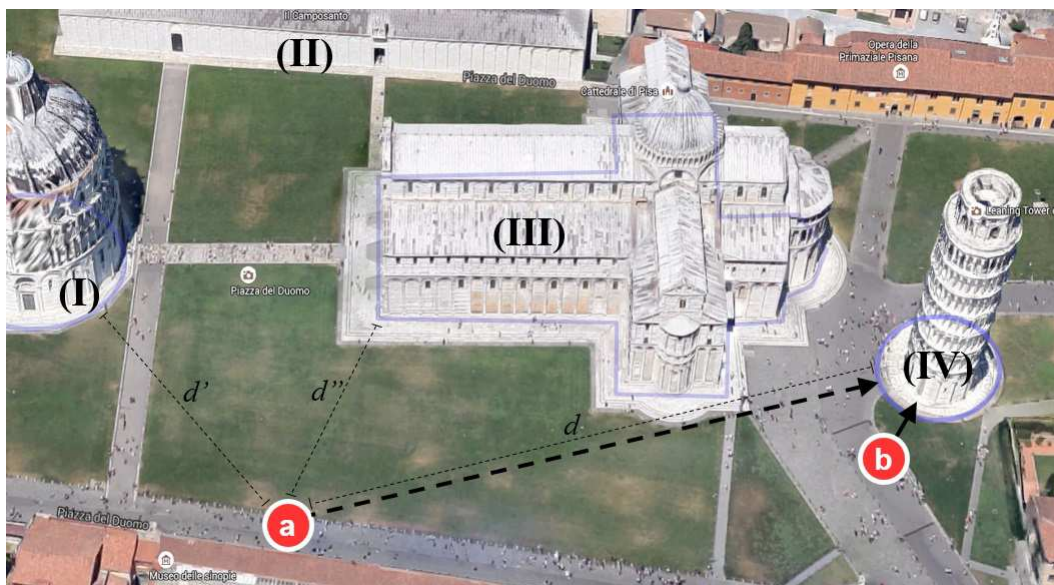


Figura 4.2: Exemplos de localização de fotos capturadas em torno da Torre de Pisa.

O cenário da situação *a*, apresentado na Figura 4.2, é uma das formas mais comuns de capturar fotografias, onde a câmera está localizada d metros do POI e está posicionada

quase que paralela ao horizonte, semelhante ao apresentado na Figura 4.1. As coordenadas geográficas relativas à foto capturada em *a* estão geograficamente mais próxima do batistério (I) e da catedral (III) do que a Torre de Pisa (IV). Uma abordagem para descoberta de pontos de interesse somente baseada em agrupamento das coordenadas geográficas poderia associar essa foto para o outro POI erroneamente.

Em contrapartida, no cenário *b*, apresentado na Figura 4.2, o POI está localizado muito próximo à posição da câmera. Nessa situação, a abordagem baseada somente na coordenada geográfica terá uma boa chance de detectar corretamente o POI relativo à fotografia.

Nos últimos anos, muitos dispositivos inteligentes (*smart devices*), tais como smartphones e câmeras digitais, passaram a ser equipados com um compasso eletrônico (bússola). Essa incorporação tornou possível o armazenamento de informações de orientação da câmera no momento dos registros fotográficos. O mesmo aconteceu anteriormente com relação às coordenadas geográficas, que foram registradas graças aos chips de GPS, também incorporados a esses dispositivos. A Figura 4.3 apresenta o exemplo de uma fotografia capturada no Coliseu, em Roma, extraída do Flickr. Nesse exemplo são apresentados os dados de localização e orientação em formato de mapa (b) e em formato tabular (c).



a) Fotografia



b) Mapa indicando a localização e orientação

Campo	Valor
Latitude	41,89166666
Longitude	12,47133333
Orientação	151,8878049
Referência da Orientação	True North (Norte Verdadeiro)

c) Dados geográficos da Fotografia

Figura 4.3: Fotografia extraída do Flickr com dados de orientação e localização geográfica.

A principal hipótese investigada nesta pesquisa é que a utilização da orientação e localização tornam mais precisa a descoberta de pontos de interesse. A Figura 4.4 apresenta

dois exemplos de algoritmos de agrupamento baseado em densidade para detecção de POIs: (I) considerando apenas as coordenadas geográficas e (II) utilizando as coordenadas geográficas associadas com o compasso da câmera. No lado esquerdo da Figura 4.4 existem dois POIs, p_1 e p_2 , e as fotografias f_1 , f_2 e f_3 entre outras. Apesar de f_1 e f_2 estarem mais próximos a p_2 , eles estão apontando para p_1 . Caso seja utilizado um algoritmo que não considera a orientação da câmera, provavelmente ele irá associar erroneamente f_1 e f_2 ao POI p_2 e f_3 à p_1 , como ilustrado na situação (I). De outra forma, caso a orientação fosse considerada, o resultado do processo de agrupamento provavelmente seria mais similar ao da situação (II).

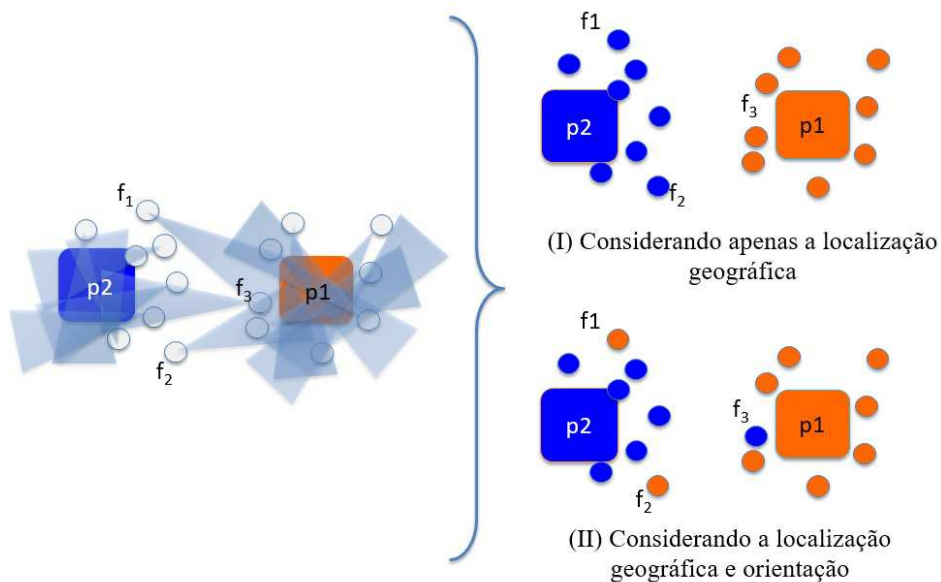


Figura 4.4: A influência da orientação na descoberta de ponto de interesse baseado em algoritmos de agrupamento.

4.2 Compass Clustering

A solução aqui apresentada realiza um agrupamento das fotografias utilizando dados da localização geográfica e também da orientação geográfica no momento de captura da fotografia. Nas fotografias tendem a estar direcionadas para o centro de um determinado ponto de interesse. Por exemplo, a Figura 4.5 apresenta um conjunto de fotografias capturadas próximas a um POI. Assim, quando são traçados segmentos de retas a partir da direção em que as fotografias foram capturadas, estes acabam se cruzando (representado

pelos círculos e quadrados), e o resultado desse cruzamento é a criação de novos pontos que podem se concentrar ainda mais próximos aos POIs (círculos) do que o local exato da captura das fotografias. Outros pontos de cruzamento serão criados de forma mais esparsa, como os representados por quadrados.

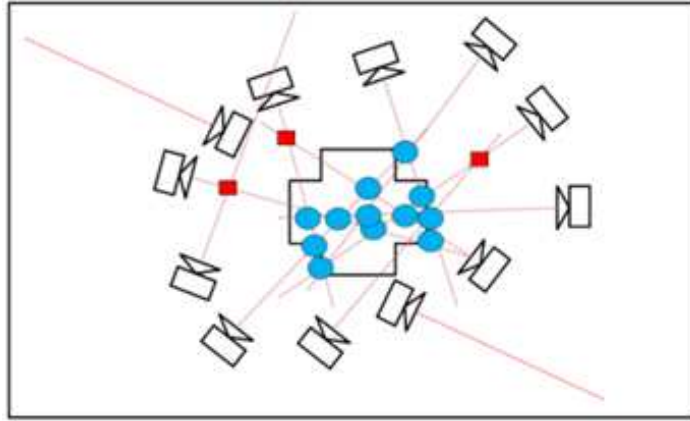


Figura 4.5: Exemplo da geração dos pontos de cruzamentos a partir dos segmentos de reta representando a orientação da captura da fotografia.

Entretanto, podem haver algumas fotografias que não estão direcionadas para o centro do POI. Ademais, pode-se, até mesmo, fazer com que o segmento traçado por sua orientação não gere nenhum cruzamento. Nestes casos, somente as posições geográficas destas fotografias são consideradas para efeito de execução do algoritmo de agrupamento.

O Algoritmo 4.2.1 apresenta um pseudocódigo para o mecanismo de agrupamento aqui descrito. O algoritmo inclui três passos: o primeiro é traçar um segmento de reta com tamanho de d metros de acordo com a orientação da respectiva fotografia (linha 1). Em seguida, devem ser detectados todos os pontos de cruzamentos entre os segmentos traçados no passo anterior; o segundo passo (linha 2) é utilizar um algoritmo de agrupamento (e.g.: DBSCAN), considerando a localização geográfica dos pontos de cruzamento entre os segmentos de reta. Por fim, o terceiro passo (linha 3) é responsável por calcular a quantidade de pontos de cruzamento para cada fotografia, em cada agrupamento. Ao final da execução do algoritmo, as fotografias serão relacionadas ao agrupamento que tiver uma maior quantidade de pontos de cruzamento. Caso haja valores iguais, as fotografias serão classificadas como sendo relativas ao grupo que tiver maior número de pontos de cruzamentos.

Além dos passos citados, o algoritmo realiza uma série de tratamentos específicos para

Algoritmo 4.2.1: Pseudo código ilustrando o funcionamento do algoritmo Compass Clustering.

entrada: Coleção de fotos orientadas; método interno de *clustering*; e d

saída : *Clusters* de fotos orientadas

- 1 pontosDeInterseccao = pegarPontosDeIntersecao(colecaoDeFotos);
 - 2 clustersDeInterseccoes=metodoInternoDeClustering.clustering(pontosDeInterseccao);
 - 3 computarInterseccoesPorCluster(colecaoDeFotos, clustersDeInterseccoes);
 - 4 clusters= criarClustersParaFotos(clustersDeInterseccoes) ;
-

seu correto funcionamento. Um deles é a incorporação dos dados de localização geográfica das fotografias cujo segmento de reta não resultou em nenhum ponto de intersecção. Além disso, outro tratamento realizado é a identificação das fotos que estão localizadas na mesma posição geográfica e com a mesma orientação, da mesma forma que as fotos que não geraram cruzamento. O resultado da intersecção de dois segmentos de retas idênticos gera infinitos pontos de cruzamento. Por esta razão, não há sentido de calcular os pontos de intersecção entre esses dois segmentos de reta.

4.3 Compass Clustering Lite

O primeiro passo do algoritmo Compass Clustering poderá gerar uma grande quantidade de pontos de cruzamento. Esses pontos de cruzamento são utilizados no segundo passo no processo de agrupamento. Dessa forma, uma grande quantidade de pontos de cruzamentos irá impactar no desempenho do algoritmo.

Alguns fatores implicam no aumento do número de pontos de cruzamentos, tais como: a quantidade de dados, a densidade desses dados e o tamanho do segmento de reta. Assim, quanto maior forem esses fatores, maior será o número de pontos de cruzamentos, conseqüentemente reduzindo o desempenho do algoritmo. O algoritmo Compass Clustering Lite (CCLITE) foi desenvolvido com a finalidade de tornar o processo de agrupamento de fotografias mais rápido, não havendo mudanças em seu desempenho pelo aumento do parâmetro d (tamanho do segmento de reta).

O pseudocódigo do Compass Clustering Lite está representado no Algoritmo 4.3.1. A ideia desse algoritmo é bastante simples. Inicialmente, são gerados pontos falsos que substituem a localização da fotografia por um ponto a uma distância d , na mesma direção em

que a fotografia está apontando (linha 1). A Figura 4.6 exibe um exemplo de uma coleção de fotografias e dois pontos de interesse. Os pontos que serão utilizados no agrupamento estão representados pelo “x” em vermelho. Em seguida, é realizado um agrupamento com base nesses pontos (linha 2). Por último, são criados *clusters* com as fotografias em suas localizações originais e os mesmos são retornados.

Algoritmo 4.3.1: Pseudo código ilustrando o funcionamento do algoritmo Compass Clustering Lite.

entrada: Coleção de fotos orientadas; método interno de clustering; e d

saída : Clusters de fotos orientadas

1 pontosFalsos = gerarPosicoesFalsas(colecaoDeFotos, d);

2 clusters=metodoInternoDeClustering.clustering(pontosFalsos);

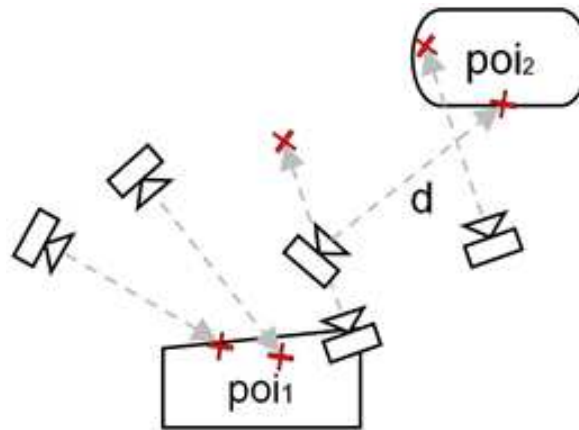


Figura 4.6: Representação dos pontos a ser utilizados no agrupamento do Compass Clustering Lite.

4.4 Compass Clustering Field of View

Nesta seção é formalizada a abstração do campo de visão que é utilizada no algoritmo Compass Clustering Field of View (CCFOV). Em seguida, o algoritmo é descrito e os seus três passos são detalhados.

4.4.1 O campo de visão

O algoritmo Compass Clustering, descrito na Seção 4.2, utiliza um segmento de reta representando a orientação da câmera no momento da captura da fotografia. Essa abstração

serve para descobrir os pontos de intersecção desses segmentos de reta e enfim utilizá-los no processo de *clustering*.

A versão *Lite* do algoritmo, utiliza uma abstração ainda mais simples, na qual são gerados novos pontos no sentido da orientação da câmera no momento da captura da fotografia. Esses pontos ficam d metros de distância do ponto original. Em seguida, esses pontos são utilizados no processo de *clustering*.

Esses algoritmos podem fazer com que seja perdida a precisão quando o objeto do interesse na fotografia não está localizado exatamente no centro da fotografia. Ou ainda, quando mais de um objeto são representados na mesma imagem.

Sendo assim, a ideia é fazer uso do campo de visão no processo de detecção de pontos de interesse. O campo de visão é uma abstração da área visível pela lente da câmera no momento da captura da fotografia, vide o exemplo da Figura 4.7.

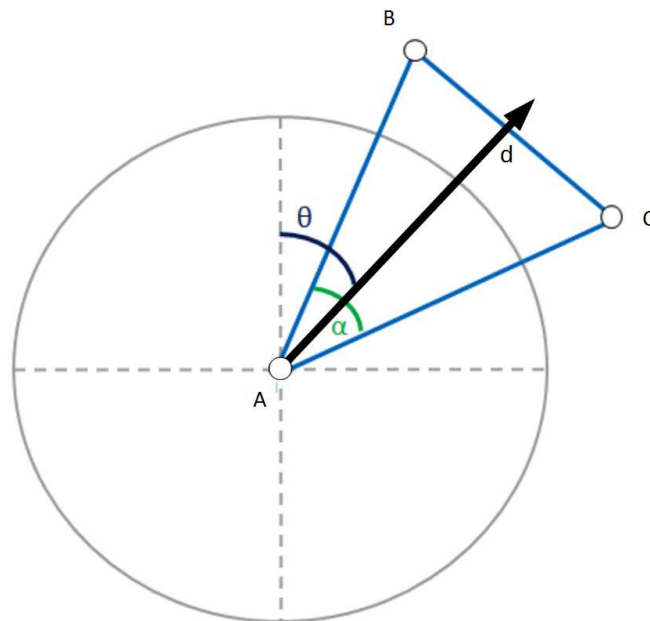


Figura 4.7: Campo de Visão

Esse modelo é mais realístico do que os modelos apresentados nos algoritmos anteriores. Entretanto, a Terra é considerada como sendo um plano geométrico. Além disso, o campo de visão é representado através de um modelo bi-dimensional. Com isso, são desconsideradas as variações de altitude nos terrenos, além de obstáculos que estejam transpondo a imagem, tais como paredes, prédios e montanhas.

O campo de visão será representado na forma de um triângulo T formado pelos segmentos de reta \overline{AB} , \overline{BC} e \overline{AC} , onde os pontos: $A = (x_0, y_0)$, $B = (x_1, y_1)$ e $C = (x_2, y_2)$, vide Figura 4.7.

Esse triângulo é calculado com base em quatro parâmetros: a localização da câmera q , a orientação da câmera θ em relação ao norte geográfico, o ângulo de visão α e o segmento de reta d que representa o alcance da fotografia. No entanto, existe a possibilidade de uma série de erros nos dois primeiros parâmetros, pois os próprios dispositivos estão sujeitos a erro. No modelo adotado, não foram considerados os erros que possam ser provocados pelas imprecisões nos equipamentos.

Os avanços nas tecnologias de geolocalização têm permitido a redução nessas imprecisões. De acordo com a agência americana que trata das informações providas por GPS¹, tipicamente a precisão dos chips de posicionamento de celular conseguem uma precisão dentro de 4.9 metros em áreas abertas. Embora, que dentro ou próximo a prédios, árvores e outros obstáculos reduzem a precisão desses equipamentos.

O cálculo dos segmentos de reta \overline{AB} e \overline{AC} se dá da seguinte maneira: o segmento \overline{AB} é criado inicialmente com as coordenadas $\overline{AB} = (a, b')$, onde $b' = (x_0, y_0 + d)$, ou seja, o ponto b' será um ponto com uma longitude igual à coordenada relativa à localização q da fotografia, porém com uma latitude com d metros de distância. Em seguida, \overline{AB} deverá sofrer uma rotação em sentido anti-horário de $\theta - \frac{\alpha}{2}$ graus. Com relação ao segmento \overline{AC} , o processo será semelhante, porém a rotação no sentido horário será de $\theta + \frac{\alpha}{2}$ graus.

O cálculo do ângulo de visão α depende do conhecimento do tipo de lente e do sensor de captura da imagem da câmera ou dispositivo utilizado para captura da imagem. Dessa forma, é necessário ter uma base com os dados técnicos das câmeras e lentes utilizadas para as fotografias. Essas informações são disponibilizadas em web-sites dos fabricantes. No caso de fotografias onde não sejam encontradas essas informações, é utilizado como valor padrão $\alpha \cong 57,3^\circ$, que é equivalente a 1 radiano. Essa variável foi utilizada nos experimentos, adotando o valor padrão que foi proposto. O aumento desse valor impacta num maior grau de intersecção entre os campos de visão das fotografias (e vice-versa). É possível balancear o aumento (ou redução) desse valor a partir de ajustes dos parâmetros do algoritmo interno

¹<http://www.gps.gov/systems/gps/performance/accuracy/>

que é utilizado no processo de *clustering* (descrito na próxima Seção).

4.4.2 Descrição do algoritmo

O algoritmo Compass Clustering Field of View (CC FOV) é composto por três passos:

- (i) a geração dos polígonos representando os campos de visão de cada fotografia;
- (ii) o cálculo da interseção entre os polígonos; e
- (iii) o processo interno de *clustering*.

O Algoritmo 4.4.1 apresenta o pseudocódigo do algoritmo. O primeiro passo do algoritmo (linha 1) realiza o cálculo do campo de visão de todas as fotografias da base de dados, de acordo com o apresentado na Seção 4.4.1.

Algoritmo 4.4.1: Pseudo código ilustrando o funcionamento do algoritmo Compass Clustering Field of View.

entrada: Coleção de fotos orientadas; método interno de clustering; e d

saída : Clusters de fotos orientadas

- 1 poligonos = gerarPoligonosDosCamposDeVisao(colecaoDeFotos, d);
 - 2 calculaInterseccoes(colecaoDeFotos, poligonos);
 - 3 clusters=metodoInternoDeClustering.clustering(colecaoDeFotos);
-

O segundo passo do algoritmo (linha 2) é realizar o cálculo entre as áreas dos campos de visão. Seja P a base de dados de fotografias, definido como $P = \{p_1, p_2, \dots, p_z\}$, onde z é a quantidade de fotografias georreferenciadas e orientadas. A similaridade entre duas fotos p_x e p_y em termos dos seus campos de visão é definida como sendo o quociente da área de intersecção dos campos de visão correspondentes de p_x e p_y e a soma das suas áreas, vide Equação 4.1.

$$sim(p_x, p_y) = \frac{\text{intersection_area}(p_x, p_y)}{\text{area}(p_x) + \text{area}(p_y)} \quad (4.1)$$

A função $\text{intersection_area}(p_x, p_y)$ retorna um valor (em metros quadrados) representando a área total de intersecção entre os campos de visão de p_x e p_y . Além disso, a função $\text{area}(p_i)$ retorna um valor (em metros quadrados) representando a área total do campo de visão p_i . Uma vez que os algoritmos de agrupamento baseados em distâncias necessitam

de uma função de distância, a função de similaridade $sim(p_x, p_y)$ pode ser convertida em uma função de distância a partir da seguinte equação $dist(p_x, p_y) = 1 - sim(p_x, p_y)$, onde $dist(p_x, p_y)$ representa a distância entre os pontos p_x e p_y .

No terceiro passo (linha 3), é proposto de usar um algoritmo de agrupamento baseado em densidade que utiliza distância no processo de *clustering*, tal como o P-DBSCAN; utilizando como função de distância a intersecção do campo de visão. O P-DBSCAN, conforme descrito no Capítulo 3, tem sido utilizado com algum sucesso para a detecção de POIs. O objetivo desse passo é identificar as áreas onde a intersecção dos campos de visão são mais densos. Os *clusters* resultantes irão representar os POIs descobertos.

4.5 Considerações Finais

Este capítulo apresentou três novos algoritmos de agrupamento de fotografias com o intuito de fazer a descoberta de pontos de interesse. Dessa forma, esses algoritmos propõem uma nova abordagem utilizando os metadados de localização geográfica e orientação das fotografias para uma melhor precisão na detecção de pontos de interesse.

O algoritmo Compass Clustering é um algoritmo dividido em três passos. O primeiro traça segmentos de reta representando a orientação da câmera. Em seguida, são calculados os pontos de cruzamento entre essas retas para serem agrupados internamente utilizando um algoritmo de agrupamento baseado em densidade. Por último, são realizados cálculos e são formados os *clusters* baseado nesse agrupamento dos pontos de cruzamento.

Enquanto isso, o algoritmo Compass Clustering Lite realiza o agrupamento baseado em pontos que são criados a uma certa distância do local que a fotografia foi capturada, indicando uma possível posição do objeto de interesse. Com isso, esses pontos são agrupados para formar os *clusters* de fotografias.

O uso do campo de visão no processo de detecção de pontos de interesse foi apresentando como forma de melhorar o processo de detecção. Em seguida, foi proposto o algoritmo *Compass Clustering Field of View*, um algoritmo de três passos no qual considera o campo de visão no processo de *clustering*.

Apesar da abstração do uso da orientação da fotografia ainda há uma série de aspectos do mundo real que não estão contemplados. A angulação vertical da câmera não está

sendo considerada, ou seja, não há nenhuma consideração se no momento da fotografia a câmera estava apontando para cima, baixo, ou em algum outro ângulo. A mesma é sempre considerada como estando no sentido horizontal.

Outros aspectos também não são considerados: relevo geográfico, construções, diferença de altitude, circunferência da terra e foco da fotografia. Esses fatores poderão num futuro serem considerados para se ter uma maior precisão na detecção de POIs.

No próximo capítulo, é apresentado o protótipo de ferramenta que também é proposto neste trabalho: o PhoCA - Photo Clustering Analyzer. Os algoritmos que foram propostos neste capítulo estão implementados nesta ferramenta.

Capítulo 5

Protótipo - Photo Clustering Analyzer

5.1 Introdução

O processo de descoberta de conhecimento, conforme apresentado no Capítulo 2, envolve um conjunto de nove atividades sequenciais e interativas que fazem com que o processo como um todo seja algo bastante oneroso.

Um processo de descoberta de conhecimento que envolve dados multimídia, como é o caso de fotografias digitais, deverá levar em consideração que se tratam de dados complexos, que possuem um conteúdo visual e uma série de metadados. Sendo assim, o procedimento se torna ainda mais complexo.

O primeiro passo a ser realizado no processo de KDD para fotografias trata da aquisição dos dados. Essa atividade envolve estudar uma série de APIs (*Application Programming Interface*) do repositório de fotos de interesse e escrever códigos específicos para a coleta dessas fotografias e extração dos metadados necessários. Desse modo, seria de grande importância ter aplicações para auxiliar o pesquisador nessa tarefa, tendo em vista que a mesma exigirá uma série de conhecimentos em linguagens de programação e bancos de dados.

Outra questão que deve ser levada em consideração é que para trabalhar com dados multimídia, a exemplo de fotografias, a ferramenta de KDD deverá facilitar a visualização do conteúdo da imagem e de seus metadados. Além disso, quando a localização geográfica é de interesse, é de extrema importância que esse dado possa ser visto em formato de mapa digital, facilitando a leitura dos mesmos.

A fim de atender a essa necessidade, é proposto nesse capítulo o PhoCA ¹ (Photo Clustering Analyzer), uma ferramenta Web extensível para análise de agrupamento de fotografias. O PhoCA visa auxiliar pesquisadores na área disponibilizando as mais importantes tarefas do processo de KDD, tais como: coleta de dados, pré-processamento, mineração de dados, avaliação e interpretação.

Vale ressaltar que não foi encontrado na literatura nenhum trabalho que seja específico para agrupamento de fotografias digitais com características semelhantes à ferramenta proposta neste capítulo.

A próxima seção apresentará a arquitetura do sistema. Em seguida, na Seção 5.3 são apresentados os módulos do sistema. Por último, serão feitas as considerações finais acerca deste capítulo.

5.2 Arquitetura do Sistema

A Figura 5.1 apresenta a arquitetura do PhoCA. O sistema tem uma arquitetura orientada a serviços, significando que suas funcionalidades são disponibilizadas através da invocação de Web Services. Isso permite o acesso as suas funcionalidades a partir de outros *software*, independentemente de linguagem de programação e plataforma do sistema.

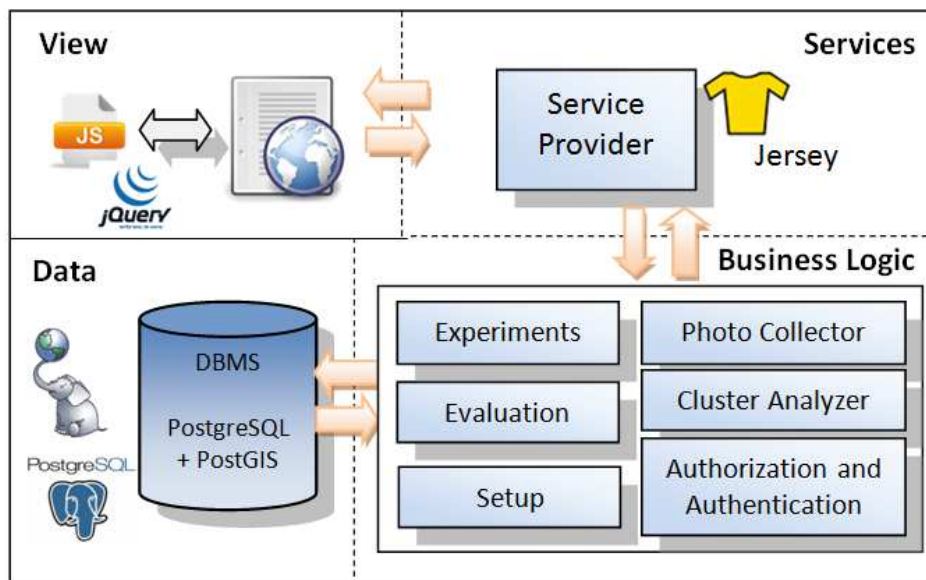


Figura 5.1: Arquitetura do PhoCA

¹<<http://www.yurilacerda.com/phoca>>

O PhoCA é uma aplicação RESTful (RICHARDSON; RUBY, 2008), o que torna mais simples e eficiente a implementação de uma arquitetura de *Web Services*. A interface com o usuário foi desenvolvida utilizando HTML (Hypertext Markup Language), CSS (Cascading Style Sheets), Javascript e JQuery². A biblioteca JQuery foi desenvolvida para permitir o desenvolvimento web rápido, capturando eventos e interações assíncronas, por meio do Ajax, com isso é possível realizar comunicação com os Web Services de forma assíncrona, permitindo melhorar a interação do usuário.

A camada de modelo possui as classes responsáveis pela lógica de negócio. Ela é dividida em seis módulos: Photo Collector; Experiments; Cluster Analyzer; Evaluation; Authorization and Authentication; e Setup.

5.3 Módulos do Sistema

O Photo Collector é responsável pela extração dos metadados de fotografias locais ou do Flickr. O sistema permite o armazenamento e extração de metadados das fotografias enviadas pelos usuários por meio de upload, apresentado à esquerda na Figura 5.2.

No caso de fotografias do Flickr, o usuário deverá informar uma área geográfica para a extração dos metadados das fotos. É possível que essa busca seja limitada através de uma palavra-chave.

A funcionalidade de extração de metadados do Flickr é apresentada à direita na Figura 5.2. De acordo com os parâmetros escolhidos, principalmente de acordo com a área selecionada, esse processo poderá ser bastante dispendioso, podendo levar horas ou até dias para ser completado. No Apêndice A, encontram-se mais detalhes sobre o processo de extração de metadados.

A ferramenta disponibiliza a possibilidade de ajustar alguns parâmetros que permitirão melhorar o desempenho da ferramenta, como por exemplo o número de *threads* utilizado na coleta dos dados. Esses parâmetros poderão ser ajustados a partir do módulo Setup.

O módulo de experimentos permite a execução de um algoritmo de agrupamento para uma coleção de fotografias específica. O PhoCA disponibiliza cinco algoritmos de agrupamento pré-configurados: k-means, DB-Scan, Compass Clustering, Spectral

²<<http://www.jquery.com>>

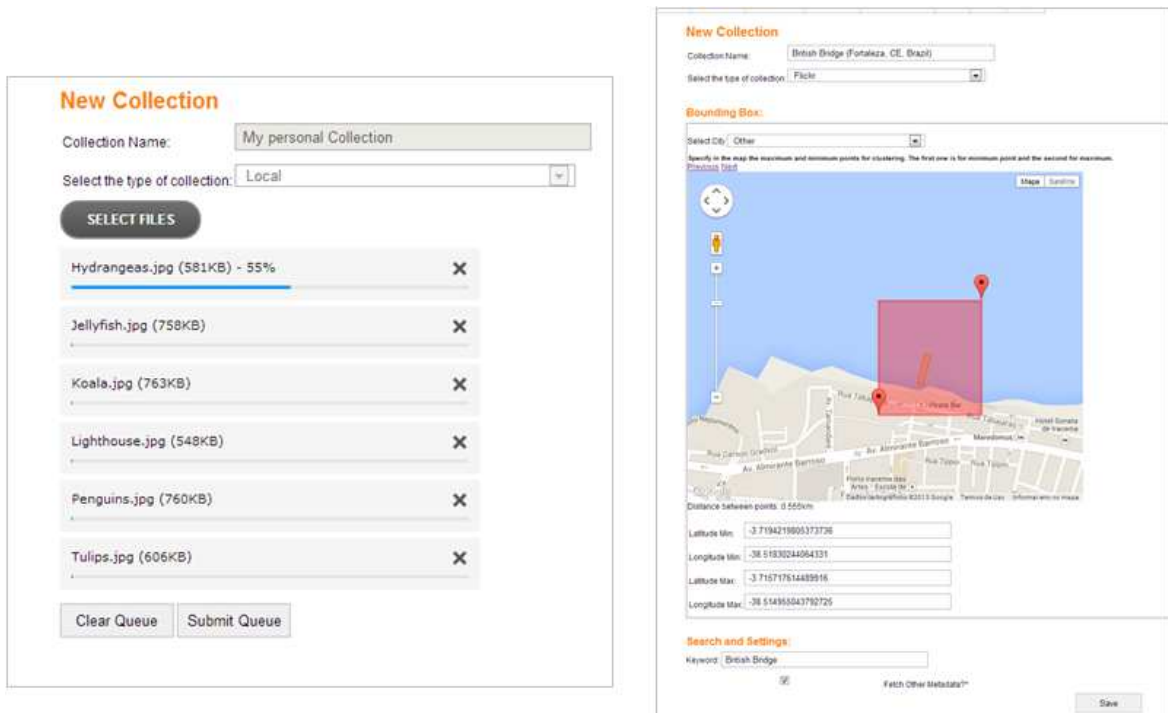


Figura 5.2: O módulo Photo Collector

Clustering e mean shift. A Figura 5.3 apresenta um exemplo da interface gráfica para execução de um experimento. O usuário poderá selecionar o algoritmo, definir seus parâmetros de entrada, selecionar a coleção de fotografias a ser utilizada como entrada e realizar algumas filtragens na coleção. O usuário poderá adicionar novos algoritmos utilizando alguns elementos de programação orientada a objetos e adicionando algumas linhas em um arquivo XML. O Apêndice A apresenta detalhes de como inserir um novo algoritmo na ferramenta.

Figura 5.3: O módulo Experimentos

O próximo módulo, o Cluster Analyzer, é responsável por apresentar um sumário das estatísticas relacionadas aos grupos computados, tais como: algoritmo e parâmetros utilizados; data e horário da execução; tempo de execução; tamanho médio, mínimo, máximo dos grupos formados e o número total de elementos. Além disso, esse módulo é responsável por apresentar o tamanho de cada *cluster*, permitindo a visualização dos mesmos em formato de mapa digital, apresentado na Figura 5.4, ou tabular.

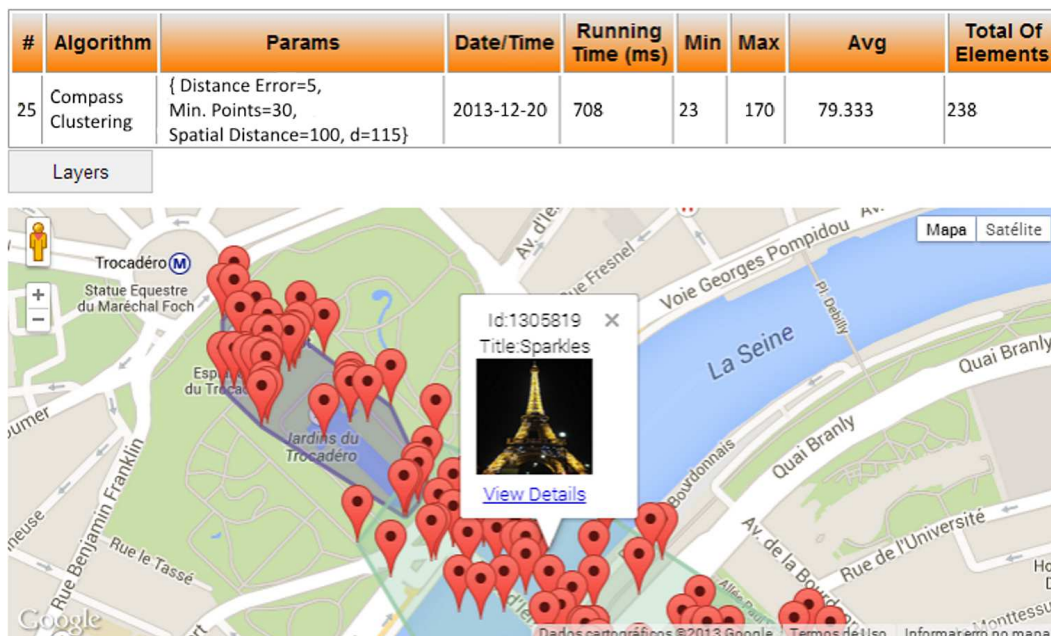


Figura 5.4: O módulo Cluster Analyzer

O módulo Evaluation permite que o usuário possa avaliar se uma determinada fotografia

está corretamente associada com o grupo ao qual ela foi classificada, por meio de uma interface gráfica com o usuário. A Figura 5.5 apresenta a interface gráfica de avaliação. O módulo permite avaliar manualmente cada elemento avaliando se foi agrupado corretamente ou incorretamente. Além disso, esse módulo calcula a métrica de precisão baseado no percentual de fotos agrupadas corretamente. Essa é uma métrica comum no campo de Recuperação da Informação.

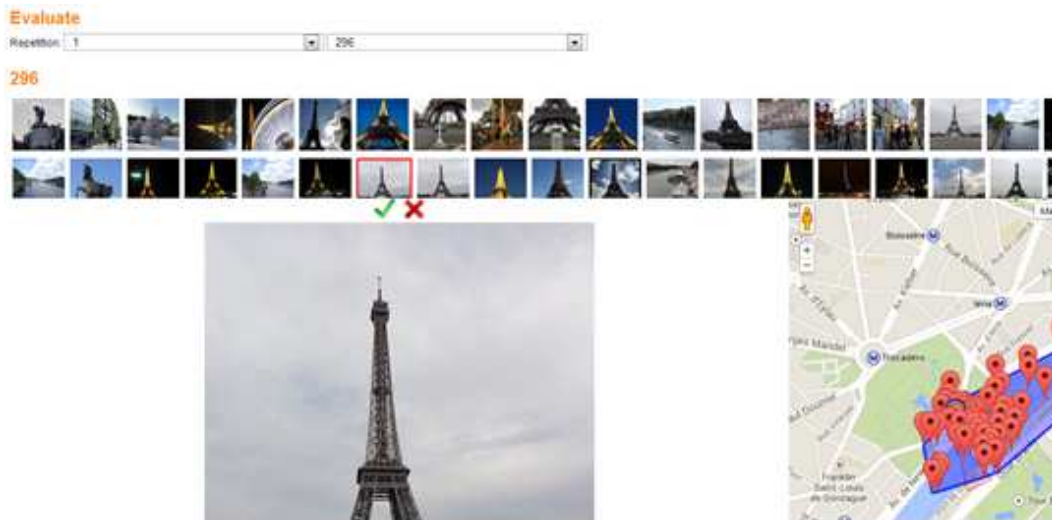


Figura 5.5: O módulo Evaluation

O módulo Authorization and Authentication controla toda a segurança do sistema, ou seja, ele monitora o acesso a recursos e funcionalidades. Além disso, ele permite a criação e manipulação dos usuários do sistema. Existem dois perfis de usuários: administrador e usuário padrão. Somente os usuários com perfil de administrador têm o privilégio para a criação de novos usuários e permite realizar mudanças nas configurações do sistema. O usuário com perfil padrão tem privilégio para utilizar os outros módulos tanto por meio da interface gráfica do sistema, como também invocando os Web Services.

O módulo Setup é responsável pelas configurações do acesso ao banco de dados, chave de acesso a API do Flickr e os ajustes de desempenho. A ferramenta deverá ser instalada em um Web container. O Web container é um componente de um servidor Web que implementa a especificação Java Servlet³. A ferramenta utiliza o Sistema de Gerenciamento de Banco de Dados PostgreSQL⁴ e a extensão espacial Postgis⁵.

³<<http://www.oracle.com/technetwork/java/index-jsp-135475.html>>

⁴<<http://www.postgres.org>>

⁵<<http://www.postgis.org>>

5.4 Considerações Finais

Foi apresentado neste capítulo o PhoCA, uma ferramenta web extensível para análise de agrupamento de fotografias. Essa ferramenta disponibiliza: acesso público as funcionalidades através de Web Services; coleta e extração de metadados de fotografias do Flickr e de fotografias locais; agrupamento de fotografias utilizando algoritmos do estado-da-arte; e a análise de grupos através de uma ferramenta de visualização geo-espacial.

Embora existam muitos trabalhos que utilizam algoritmos de agrupamento para os mais diversos tipos de aplicação, o PhoCA é a primeira ferramenta especializada para apoiar pesquisadores no mais diversos passos de um processo de KDD para esse domínio específico. Na literatura revisada, não foi encontrada nenhuma ferramenta similar ao PhoCA. Ao invés disso, foram encontradas ferramentas para fins genéricos, como o Weka (HALL et al., 2009).

O próximo capítulo apresenta os experimentos realizados para a validação das soluções propostas nesta Tese.

Capítulo 6

Experimentos e Resultados

6.1 Introdução

A avaliação do trabalho aconteceu em duas etapas. A primeira etapa foi realizada de forma a avaliar os dois algoritmos que foram propostos inicialmente: o Compass Clustering e o Compass Clustering Lite. Na segunda etapa, foi avaliado o uso do campo de visão para o processo de detecção de POIs através do algoritmo Compass Clustering Field of View.

As etapas da avaliação realizam todas as fases de um processo de descoberta de conhecimento. Esse processo deverá ser iniciado com a coleta e armazenamento das fotografias que serão utilizadas no experimento e culminará com a execução e avaliação dos experimentos utilizando algoritmos para a detecção de pontos de interesse.

O processo de avaliação apresentado neste capítulo foi realizado apoiado pela ferramenta PhoCA - Photo Clustering Analyzer (LACERDA et al., 2014). A ferramenta foi descrita no Capítulo 5. O PhoCA é responsável por realizar desde a coleta de dados, através do módulo *Photo Collector*, até a execução dos experimentos com o módulo *Experiments*.

A Seção 6.2 apresenta a primeira etapa da avaliação. Em seguida, a Seção 6.3 detalha a segunda etapa de avaliação. Por fim, a Seção 6.4 apresenta as considerações finais do capítulo.

6.2 Primeira Etapa

6.2.1 Base de Dados

A metodologia empregada para os experimentos está apresentada na Figura 6.1. Assim, na primeira fase da metodologia é necessário delimitar o espaço geográfico para que sejam coletadas as respectivas fotografias para a realização dos experimentos. Como cidade-alvo foi selecionada a cidade de Paris, por ser uma das cidades mais visitadas no mundo e por já haver trabalhos na literatura relacionados à mesma. Segundo MasterCard (2014), Paris tem estado entre 2010 até 2014 entre as três cidades mais visitadas no mundo.

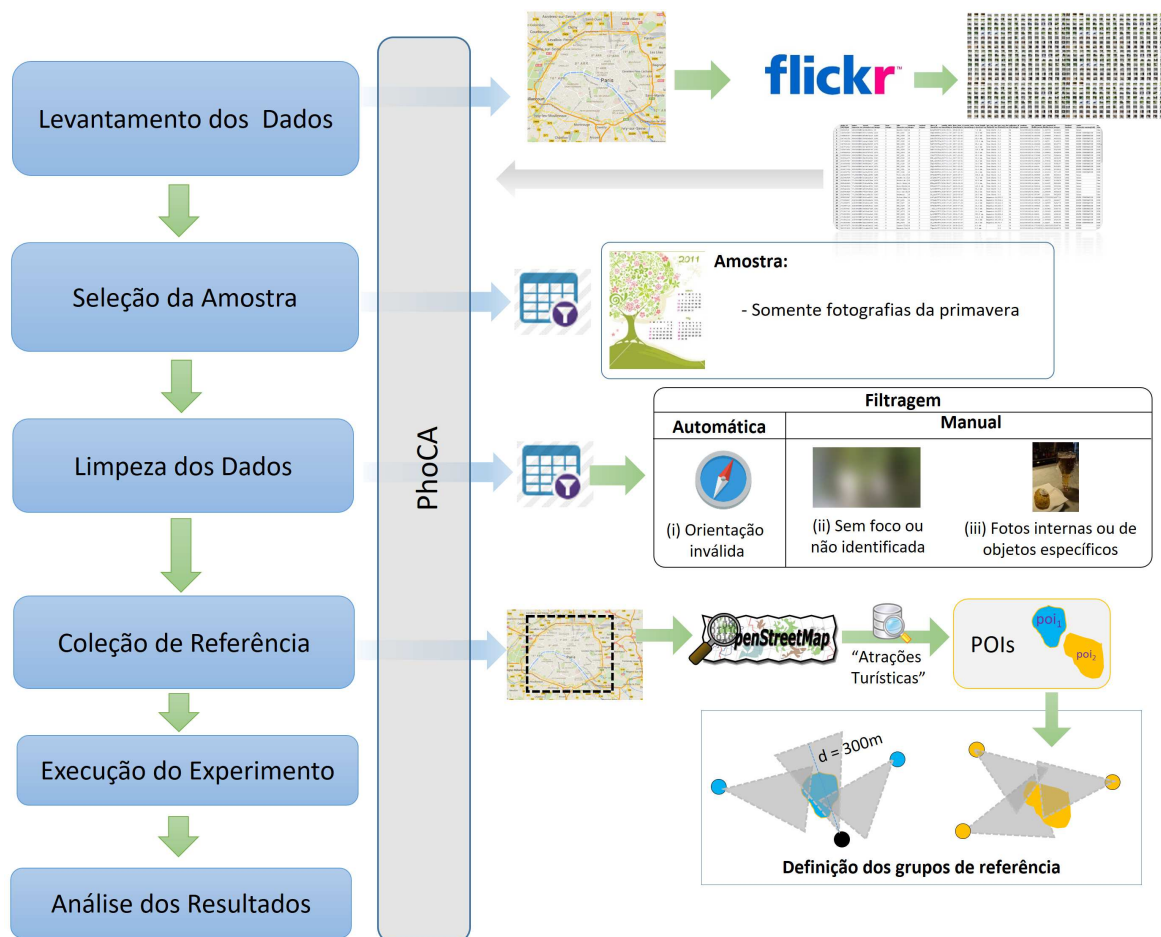


Figura 6.1: Representação gráfica dos passos para a execução dos experimentos.

A Tabela 6.1 apresenta a área delimitada para extração dos dados. Esses dados foram extraídos do Flickr no dia 29/10/2014. O número total de fotografias foi de aproximadamente 972 mil. Desse total, em torno de 7,9% das fotografias são também orientadas. A Figura 6.2

Tabela 6.1: Área de extração dos dados

Latitude e Longitude	[48.8151 48.9030] [2.2232 2.4742]
No. De Fotografias	971.907
Fotografias Orientadas e Georreferenciadas	76.725

apresenta a distribuição espacial dessas fotografias.

Foi necessário escolher uma cidade que tivesse uma grande quantidade de fotografias georreferenciadas, pois o percentual de fotos que possuem o dado acerca da orientação geográfica ainda é muito pequeno comparado ao número total de fotografias sem orientação.

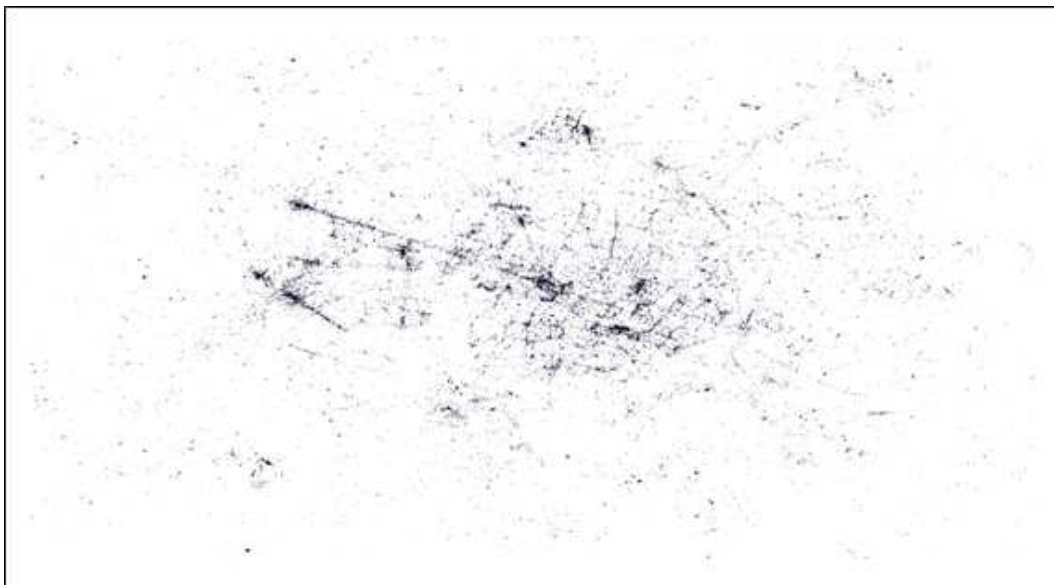


Figura 6.2: Fotografias georreferenciadas e orientadas de Paris

Mesmo com a popularização de dispositivos equipados com bússola digital, o percentual de fotografias sem orientação ainda é muito grande, de acordo com os dados da Tabela 6.1 aproximadamente 92% das fotografias não possuem dado sobre a orientação.

O gráfico apresentado na Figura 6.3 mostra a distribuição das fotografias com orientação geográfica entre os anos de 2004 e 2014. Esses dados se iniciam no ano de 2004, por ser o primeiro ano que aparece esse tipo de dado nas fotografias, e são relativas à data de extração dos dados. Observa-se que houve um crescimento muito grande, principalmente a partir do ano de 2010. Portanto nota-se que abordagens que usam a anotação como fonte adicional de dados estão se tornando cada vez mais viáveis

Foram escolhidos pontos de interesse que estão classificados como “atrações turísticas”

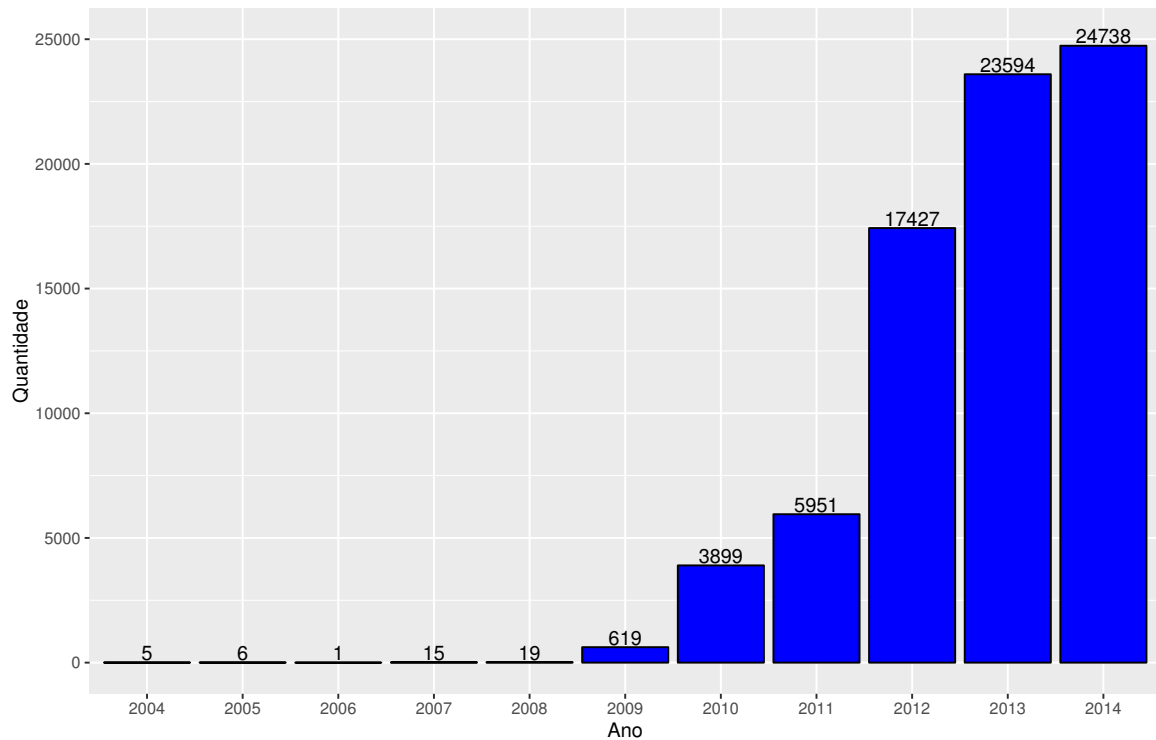


Figura 6.3: Quantidade de fotografias georreferenciadas e orientadas relativas a Paris.

no Open Street Map (OSM), uma vez que esses locais são os preferidos para os turistas capturarem fotografias. Foram extraídos dados do OSM através do MapZen ¹ - uma plataforma livre para acesso a dados de mapas. Assim, 14 categorias relacionadas a turismo foram selecionadas dentre os dados extraídos. Esse processo resultou num total de 55 pontos de interesse. Os dados são providos pelo MapZen no formato de polígono, linha e ponto. Foram utilizados dados representados em formato de polígono, uma vez que as áreas desses polígonos são importantes no momento da criação da coleção de referência, que é utilizada na avaliação. O método empregado na criação da coleção de referência é descrita no final desta seção. Os POIs selecionados estão descritos na Tabela 6.2 e são apresentados em formato de mapa digital na Figura 6.4 pelos polígonos em vermelho. Esses dados foram extraídos e armazenados na base de dados da ferramenta PhoCA.

Tabela 6.2: POIs extraídos da cidade de Paris no Open Street Map

Arc de Triomphe	La Sorbonne
Arc de Triomphe du Carrousel	La Tour aux Figures

¹<https://mapzen.com/>

Argonaute	Laboratoire Aérodynamique EIFFEL
Assemblée nationale	Les Grandes Serres du Jardin des Plantes
Basilique du Sacré-Cœur	Maison de Nicolas Flamel
Bateaux-Mouches	Maison de Tristan Tzara (Adolf LOOS Architecte)
Carrousel de la Tour Eiffel	Moulin
Carrousel de Montmartre	Moulin de la Galette
Cathédrale Notre-Dame de Paris	Opéra Garnier
Cavea des Arènes de Lutèce	Palais de Chaillot
Centre Aquatique	Palais de Justice
Centre Pompidou	Palais de l'Élysée
Centre Wallonie-Bruxelles	Panthéon
Cimetière du Père-Lachaise	Petit Palais
Collège des Bernardins	Place d'Aligre
Colonne de Juillet	Place de la République
Dodo Carousel	Place des Vosges
École Militaire	Plan du quartier Jeanne d'Arc
École Nationale Supérieure des Beaux-Arts	Pont Neuf
Église de la Madeleine	Presbytère
Église Saint-Eustache	Reptiles
Elysées Ceramic	Rotonde
Grande Volière	Rotonde de la Villette
Hôtel de Lauzun	Salle du Livre d'Or
Hôtel de Soubise	Tour Eiffel
Hôtel de Ville	Tour Montparnasse
Hôtel des Invalides	Tour Saint-Jacques
Hôtel Lebrun	

Uma amostra de 9.300 fotografias foi retirada da coleção para que fosse viável realizar o pré-processamento dos dados, tendo em vista que esse processo é manual e oneroso. Considerando que a primavera é um dos melhores períodos para atividades ao ar livre, por conta da amenidade do clima, optou-se por selecionar fotografias deste período para

a amostra. Além disso, os avanços tecnológicos têm tornado mais preciso os sistemas de geolocalização. Portanto, optou-se por utilizar as fotografias do período que envolve a primavera do ano de 2014, que são os dados mais recentes da base de dados extraída. Esse período envolve do dia 21 de março até o dia 20 de junho. Dessa forma, essa amostra ficou com o total de 9.300 fotografias, das quais serão removidas as fotos consideradas inválidas para o experimento. Esse número é viável para uma análise manual.

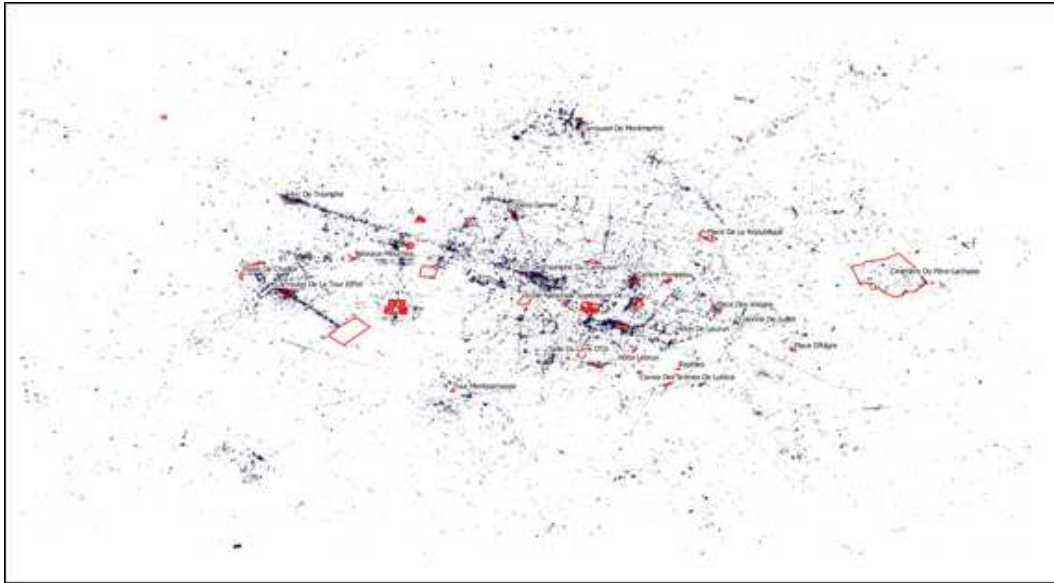


Figura 6.4: Fotografias orientadas e Pontos de Interesse levantados em Paris.

As fotografias consideradas inválidas são aquelas que: (i) não possuem o norte verdadeiro como referência para a bússola; (ii) foram capturadas em ambientes internos; (iii) possuem conteúdo sem foco ou não identificado; (iv) foco somente em pessoas, comidas ou objetos específicos, nas quais não há foco no ambiente ou lugar.

Há dois tipos de referências que as bússolas das câmeras utilizam para o armazenamento de informação relativa à direção da câmera no momento da captura da fotografia: norte verdadeiro e norte magnético. De forma a retirar a incompatibilidade entre esses dois tipos de orientações, optou-se por escolher somente as fotografias do norte verdadeiro, uma vez que o ângulo entre o norte verdadeiro e o norte magnético varia de acordo com a localização e o calendário (THOMEE, 2013). O dado sobre a referência utilizada pela bússola é fornecido pela câmera e armazenado nos metadados. A Tabela 6.3 apresenta os dados relacionados à referência utilizada pelas fotos da amostra.

Assim, o próximo passo após a remoção das fotografias que não possuem o norte

Tabela 6.3: Relação entre a referência da bússola e a quantidade de fotografias

Referência da Bússola	# Fotos
Sem nenhuma referência	53
Norte Magnético	2.258
Norte Verdadeiro	7.008

verdadeiro como referência é a remoção das fotografias consideradas inválidas pelos itens (ii) (iii) e (iv) descritos anteriormente. Esse procedimento foi realizado manualmente em cada uma das 7.008 fotografias, resultando na remoção de 4.944 fotos (70,54%). Isso significa que a coleção resultante ficou com 2.064 fotografias, que corresponde a 29,45% do total, apresentados na Figura 6.5.

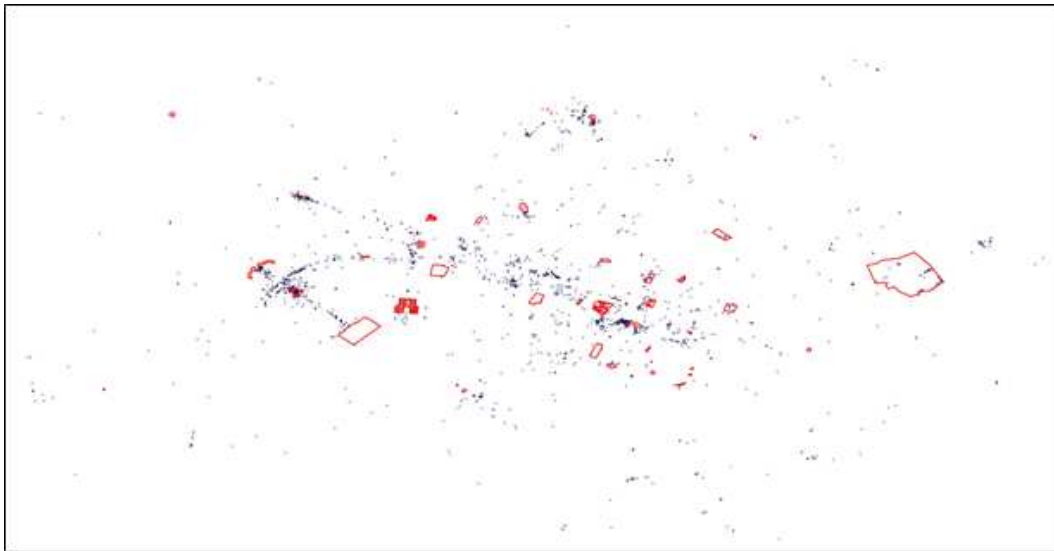


Figura 6.5: Coleção Filtrada e Pontos de Interesse sem orientação das fotos

Logo após, cada uma das fotografias válidas foi classificada manualmente como sendo ou não relativa ao ponto de interesse. Essa avaliação serviu para que fosse formado o conjunto de referência.

A pesquisa desenvolvida por (YANG; GONG; U, 2011) utiliza como sendo o conjunto de referência as fotografias que possuem a localização interna ao MBR relativo ao POI, ou dentro de um incremento da área desse ponto. Dessa forma, esse modelo leva em consideração que, se uma fotografia foi tirada na área próxima ao ponto de interesse, necessariamente ela é relativa a ele. Entretanto, de acordo com o que foi apresentado no Capítulo 4, existem situações em que a fotografia foi capturada próxima a um ponto de interesse, mas ela é relativa a outro.

Dessa forma, foi definido que uma foto é dada como sendo relativa a um POI quando a área relativa ao seu campo de visão faz interseção com a área de algum POI. O campo de visão foi calculado levando em consideração um alcance de 300 metros. Esse valor foi escolhido de forma empírica, baseado no fato que a cidade de Paris possui logradouros grandes, e vários dos pontos de interesses estão em áreas muito amplas. Entretanto, com o avanço da pesquisa será possível estimar melhor este valor. Esse campo de visão é calculado semelhante ao apresentado na Seção 4.4.1.

Caso haja a interseção com mais de um POI, a foto será classificada como sendo relativa ao ponto que houver uma maior área de interseção. A Figura 6.6 apresenta a coleção filtrada com o campo de visão.

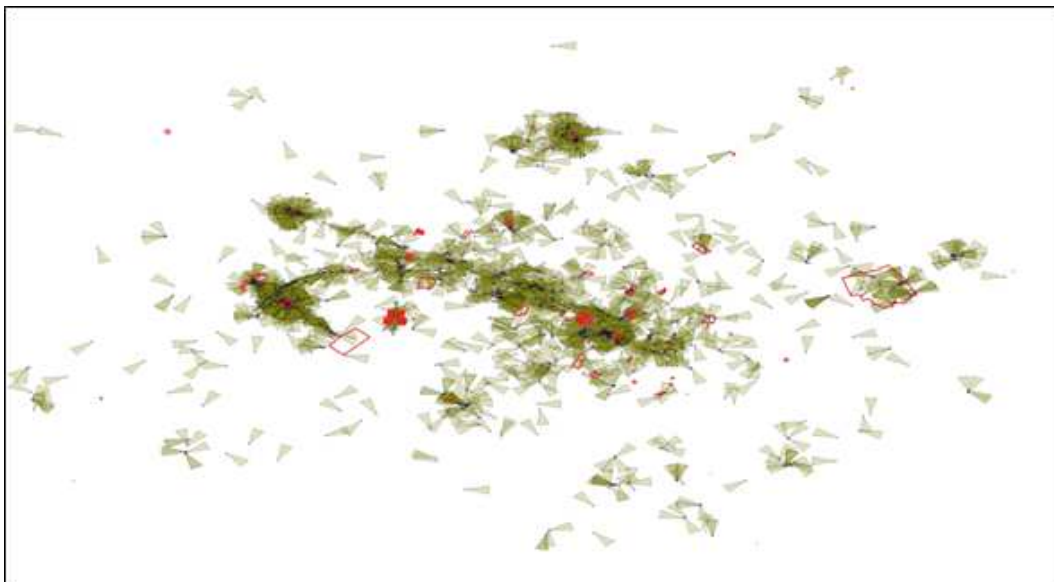


Figura 6.6: Coleção filtrada e pontos de interesse com a orientação das fotos. Os pontos em vermelho representam os pontos de interesses. Enquanto que os triângulos em verde representam as orientações.

A coleção de referência resultou em 686 fotografias sendo classificadas como relativas a algum ponto de interesse. Essas fotografias correspondem a aproximadamente 33,23% das fotografias consideradas válidas para os experimentos. A Figura 6.7 apresenta a área dos pontos de interesse que são cobertas pelos campos de visão das fotografias. Essas áreas estão representadas em verde e os POIs estão contornados em rosa.

Um total de 38 pontos de interesse foram classificados na coleção de referência. Observe na Figura 6.7 que muitos desses pontos tiveram boa parte de sua área coberta pelos campos de visão das fotografias. Enquanto que a Figura 6.8 apresenta grupos de referência, na cor

laranja, que foram formados com as fotografias da coleção de referência. Cada grupo é relativo a um dos 38 pontos de interesse.

A Tabela 6.4 apresenta a quantidade de fotografias para cada ponto de interesse da coleção de referência. Há uma média de aproximadamente 18 fotografias por ponto de interesse, com desvio padrão de aproximadamente 34 fotos. A Torre Eiffel é o ponto que possui maior número de fotografias. Seis pontos de interesse tiveram somente uma fotografia relativa aos mesmos.



Figura 6.7: Áreas dos pontos de interesse que são cobertas pelos campos de visão das fotografias. A área coberta está representada em verde, enquanto que a área dos polígonos estão contornadas em rosa.

Tabela 6.4: Relação entre o ponto de interesse e número de fotografias da coleção de referência

Ponto de Interesse	# Fotos	Percentual (%)
Arc de Triomphe	55	8,017493
Arc de Triomphe du Carrousel	14	2,040816
Assemblée nationale	9	1,311953
Basilique du Sacré-Cœur	26	3,790087
Bateaux-Mouches	3	0,437318
Carrousel de la Tour Eiffel	17	2,478134
Carrousel de Montmartre	8	1,166181

Cathédrale Notre-Dame de Paris	92	13,41108
Centre Pompidou	13	1,895044
Centre Wallonie-Bruxelles	2	0,291545
Cimetière du Père-Lachaise	49	7,142857
Collège des Bernardins	3	0,437318
Colonne de Juillet	1	0,145773
École Militaire	2	0,291545
École Nationale Supérieure des Beaux-Arts	15	2,186589
Église de la Madeleine	1	0,145773
Église Saint-Eustache	1	0,145773
Hôtel de Lauzun	10	1,457726
Hôtel de Soubise	2	0,291545
Hôtel de Ville	17	2,478134
Hôtel des Invalides	20	2,915452
Hôtel Lebrun	1	0,145773
La Sorbonne	6	0,874636
Les Grandes Serres du Jardin des Plantes	2	0,291545
Maison de Tristan Tzara	5	0,728863
Opéra Garnier	14	2,040816
Palais de Chaillot	19	2,769679
Palais de Justice	24	3,498542
Panthéon	6	0,874636
Petit Palais	9	1,311953
Place de la République	10	1,457726
Place des Vosges	9	1,311953
Pont Neuf	5	0,728863
Presbytère	1	0,145773
Rotonde	1	0,145773
Salle du Livre d'Or	2	0,291545

Tour Eiffel	196	28,57143
Tour Montparnasse	16	2,332362
TOTAL	686	100%

6.2.2 Metodologia de Avaliação

Após a inspeção dos dados, a próxima fase é a realização dos experimentos. Esses experimentos deverão levar em consideração a abordagem proposta, que usa o metadado de orientação geográfica na busca de analisar sua influência no uso da orientação da fotografia para uma maior precisão no processo de detecção.

Dessa forma, foi realizado um comparativo entre o algoritmo Compass Clustering em sua versão original; Compass Clustering Lite; e o algoritmo DBSCAN. O algoritmo DBSCAN possui dois parâmetros que necessitam ser definidos previamente: mínimo de pontos (*MinPts*) e a distância espacial mínima (*Eps*). Esse algoritmo já foi utilizado em pesquisas para detecção de ponto de interesse, conforme apresentado no Capítulo 3.

Assim, inicialmente serão realizados experimentos com esse algoritmo na busca de descobrir quais são os melhores parâmetros para serem utilizados com o mesmo para a detecção dos pontos de interesse da coleção levantada. Em seguida, será possível investigar o impacto do uso da orientação geográfica no processo de detecção de POIs. Vale enfatizar que não há conhecimento nenhum sobre a localização dos pontos de interesse durante a execução do algoritmo. Essas áreas são utilizadas apenas para fazer a criação dos grupos de referência que são utilizados para fazer o cálculo das métricas de avaliação.

Para realizar a validação dos experimentos serão utilizadas as métricas de precisão e *recall*² (HAN; KAMBER, 2011) as quais são bastante utilizadas na área de Mineração de Dados e Recuperação de Informação. A precisão (*precision*) relaciona a quantidade de fotos incluídas nos respectivos agrupamentos com as fotos relativas ao POI analisado, classificadas na coleção de referência. Já o *recall* relaciona a quantidade de fotos relativas ao POI analisado com as fotos dos respectivos agrupamentos.

Desta forma, para cada agrupamento, a precisão foi definida como sendo a razão entre o total de elementos do agrupamento que são relativas ao POI e o total de elementos do

²A palavra *recall* não foi traduzida em virtude de não haver uma definição clara da palavra em português que consiga representar bem este conceito.

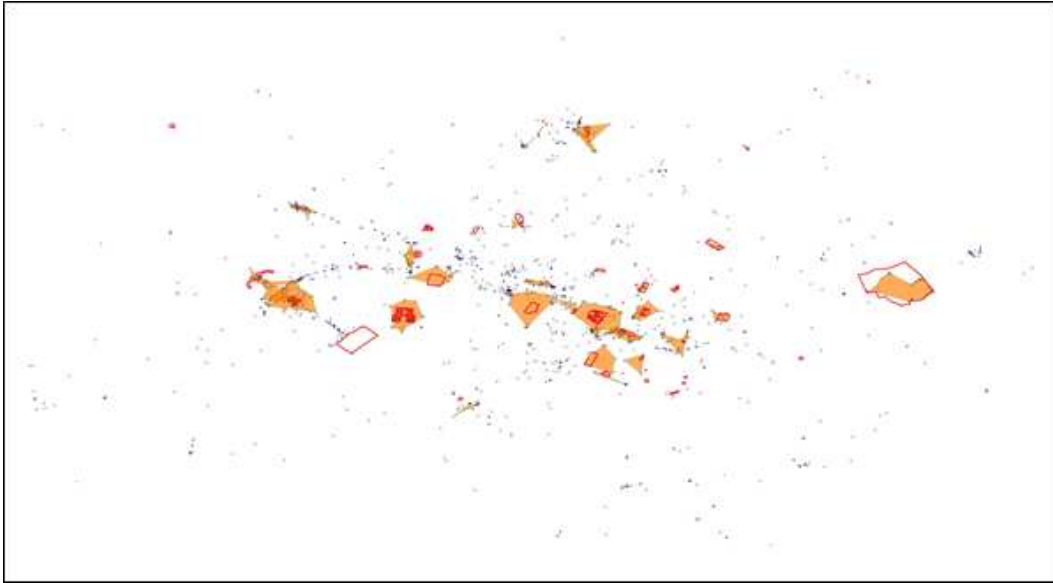


Figura 6.8: Coleções de referência (em laranja)

próprio agrupamento. O cálculo do *recall* de cada agrupamento será a razão entre o número de elementos do agrupamento que são relativos ao POI e o número total de elementos da coleção que interceptam com a coleção de referência.

Essas métricas podem ser dadas formalmente a partir das definições a seguir: Seja C_{ij} o conjunto de fotografias do i -ésimo agrupamento, considerando que estas fotos interceptam algum elemento do grupo de referência relativo ao j -ésimo POI; e P_j as fotos de referência relativas ao mesmo POI. Considerando essa definição, a precisão de um agrupamento em relação ao respectivo POI pode ser dada de acordo com a equação abaixo:

$$p_{ij} = \frac{|P_j \cap C_{ij}|}{|C_{ij}|} \quad (6.1)$$

onde p_{ij} é a precisão do i -ésimo agrupamento relativo ao j -ésimo POI.

O *recall* dado por um agrupamento, pode ser obtido por:

$$r_{ij} = \frac{|P_j \cap C_{ij}|}{|P_j|} \quad (6.2)$$

onde r_{ij} é o *recall* do i -ésimo agrupamento relativo ao j -ésimo POI. Considerando as precisões dos agrupamentos relativas a um POI, dadas pela Equação 6.3, a precisão média é obtida a partir da equação a seguir:

$$\bar{P}_j = \frac{\sum_{i=1}^{n_j} (p_{ij})}{n_j} \quad (6.3)$$

onde (\bar{P}_j) é a precisão média para o j-ésimo POI e n_j é o número total de agrupamentos para o mesmo POI.

O *recall total* de um POI será obtido a partir da média dos *recalls* de seus respectivos agrupamentos, dadas pela Equação 6.4. Assim, formalmente, a equação é definida da seguinte forma:

$$R_j = \frac{\sum_{i=1}^{n_j} r_{ij}}{n_j} \quad (6.4)$$

onde R_j é o *recall total* do j-ésimo POI e n_j é o número total de agrupamentos referentes ao POI.

Sejam X e Y os conjuntos das precisões médias e dos *recalls* totais para todos os POIs, respectivamente, e dados por $X = \{\bar{P}_1, \bar{P}_2, \dots, \bar{P}_j, \dots, \bar{P}_m\}$ e $Y = \{R_1, R_2, \dots, \bar{R}_j, \dots, R_m\}$, tal que $P_j \in X$ e $R_j \in Y$; onde $\{1 \leq j \leq m\}$ e m é o número total de POIs. Os valores mínimos, máximos e médios para as duas métricas podem ser obtidos pelas equações abaixo:

$$P_{min} = \min(X) \quad (6.5)$$

$$P_{max} = \max(X) \quad (6.6)$$

$$P_{mean} = \frac{\sum_{j=1}^m (\bar{P}_j)}{m} \quad (6.7)$$

onde P_{min} , P_{max} e P_{mean} indicam a menor precisão, a maior e a precisão média entre as precisões médias dada a todos os POIs; e

$$R_{min} = \min(Y) \quad (6.8)$$

$$R_{max} = \max(Y) \quad (6.9)$$

$$R_{mean} = \frac{\sum_{j=1}^m (\bar{R}_j)}{m} \quad (6.10)$$

onde R_{min} , R_{max} e R_{mean} indicam a menor, a maior e a média entre todas os *recalls* totais dadas aos POIs.

As métricas de precisão e *recall* podem ser interrelacionadas a partir de medida F-measure. O F-measure é a média harmônica entre a precisão e *recall*. Dessa forma, F representa a média entre a precisão e *recall* para o um dado experimento. Esta fórmula representa o F-measure tradicional ou também conhecida por F-score balanceado. Formalmente, pode ser definido da seguinte maneira:

$$F = \frac{2 \times P_{mean} \times R_{mean}}{P_{mean} + R_{mean}} \quad (6.11)$$

6.2.3 Resultados Experimentais

Alguns trabalhos da literatura já demonstraram que algoritmos de agrupamento baseados em densidades se comportam bem para a detecção de pontos de interesse (KISILEVICH; MANSMANN; KEIM, 2010; THOMEE; ARAPAKIS; SHAMMA, 2016). Dessa forma, resolveu-se adotar um algoritmo desse tipo internamente no Compass Clustering.

Foram executados experimentos variando os parâmetros do DBSCAN com o objetivo de encontrar os melhores resultados. Após a execução desses experimentos utilizando o algoritmo DBSCAN, foram executados experimentos com o algoritmo Compass Clustering. Esse algoritmo, conforme detalhado na Seção 2.3 do Capítulo 2, utiliza internamente um algoritmo de agrupamento de dados e precisa da definição de um parâmetro denominado tamanho do segmento de reta (d).

No caso de realizar uma análise que fosse semelhante para os métodos que utilizam a orientação geográfica para a detecção de pontos de interesse, utilizou-se o mesmo algoritmo (DBSCAN), como sendo o algoritmo interno do Compass Clustering, com os parâmetros que obtiveram um maior valor de F-measure. Com base nessas definições, foi realizada uma variação do parâmetro d para analisar o melhor tamanho do segmento de reta para cada uma das coleções.

DBSCAN

Utilizando a coleção pré-processada de fotografias da cidade de Paris, foram executados experimentos variando o número mínimo de pontos (MinPts) de 5 a 15, com incrementos de 5 unidades. Para cada valor desse parâmetro foi variada a distância espacial mínima de 25 a 200 metros, com incrementos de 25 metros.

Os resultados desse experimento são apresentados também na Figura 6.9 e na Tabela 6.5. É possível observar no gráfico que os melhores resultados para a coleção utilizada foi com os parâmetros: mínimo de pontos = 5 e distância espacial mínima = 100; resultando num F-measure = 0,50.

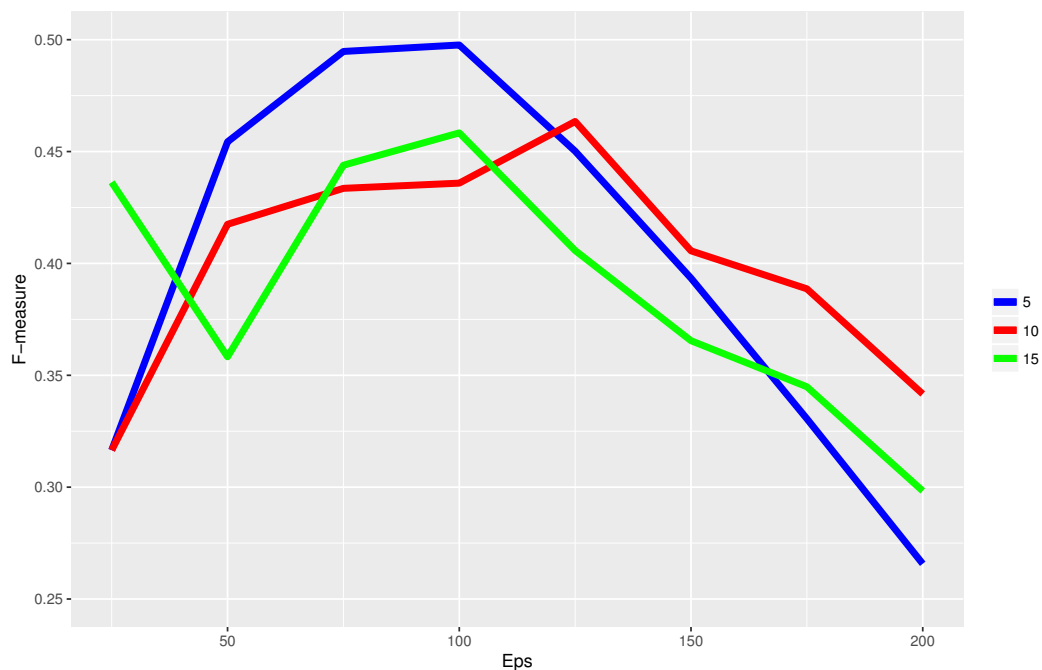


Figura 6.9: F-measure com o algoritmo DBSCAN

O resultado do experimento com o DBSCAN com os melhores parâmetros está disponível na forma de mapa na Figura 6.10. Em vermelho estão os grupos de referência, em lilás os grupos formados pelo DBSCAN.

Compass Clustering

O próximo passo do estudo foi realizar uma série de experimentos com a mesma coleção utilizada nos experimentos anteriores, porém utilizando a orientação geográfica como um

Tabela 6.5: Resultado do F-measure do algoritmo DBSCAN

Dist. Espacial Min.	Min. Points		
	5	10	15
25	0,316511	0,31652	0,436271
50	0,454342	0,417486	0,358262
75	0,49474	0,433555	0,44389
100	0,497622	0,435883	0,458348
125	0,45016	0,463453	0,405773
150	0,393406	0,40562	0,365457
175	0,33059	0,388651	0,344877
200	0,265752	0,341645	0,298319

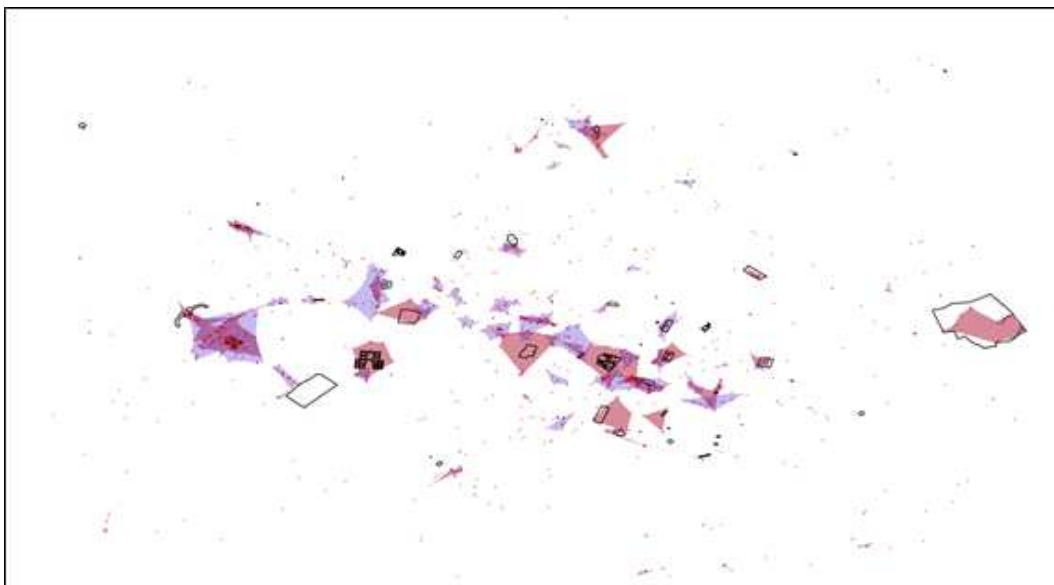


Figura 6.10: Grupos formados pelo DBSCAN em lilás e grupos de referência em vermelho.

dos critérios para o processo de agrupamento. Para isso, utilizou-se dois dos algoritmos propostos no Capítulo 4, o Compass Clustering e o Compass Clustering Lite.

Os melhores parâmetros dos experimentos anteriores foram utilizados para os demais experimentos, tendo em vista que eles foram os melhores valores para os experimentos realizados com o algoritmo DBSCAN, que será utilizado internamente no Compass Clustering.

Para cada um dos algoritmos foram realizados experimentos variando os valores do segmento de reta (d) de 50 a 500 metros, com *steps* de 50 metros. Os valores do F-measure resultantes desses experimentos podem ser conferidos na Figura 6.11.

Com relação ao Compass Clustering original, é possível perceber que houve uma

melhora para valores de d entre 50 a 150 metros. O melhor valor ficou em $d = 100$ metros, obtendo uma melhoria de aproximadamente 2%. Entretanto, ainda não é possível afirmar a real efetividade dessa abordagem, para isso será necessário realizar experimentos com outros conjuntos de dados. Quando houve um aumento muito grande no valor de d , o F-measure reduziu bastante, chegando a uma perda de aproximadamente 9,44% em relação ao DBSCAN, que só considera a posição geográfica.

O resultado do agrupamento utilizando o algoritmo original pode ser conferido na Figura 6.12. Os polígonos em vermelho representam os grupos de referência, que incluem as fotografias que foram tidas como sendo relativas a cada ponto de interesse. Todavia, as áreas formadas pelos polígonos em lilás são os grupos de fotografias agrupadas pelo Compass Clustering. Observa-se que há grupos que deveriam ser menores, como é o caso da área (a), onde há POIs bem próximos uns aos outros. Em outras áreas onde há um ponto de interesse em uma grande área, como em (b), foram formados mais de um grupo, onde deveria existir um único grupo.

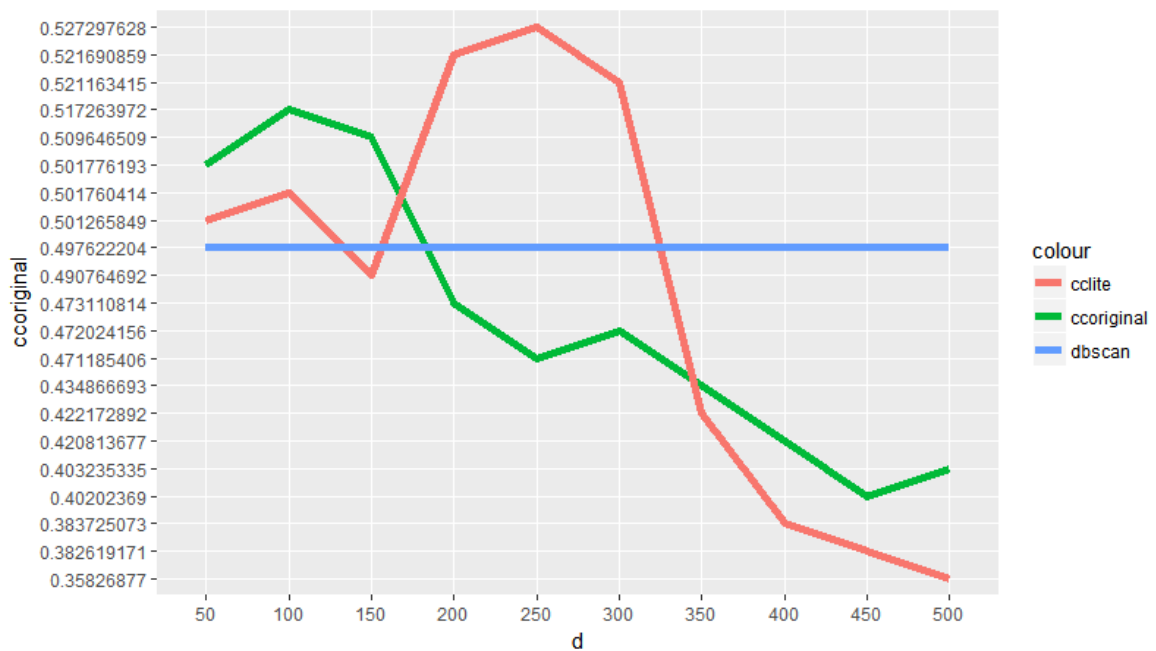


Figura 6.11: F-measure resultante entre os vários algoritmos.

O Compass Clustering Lite foi aquele que obteve um melhor resultado. O mesmo atingiu uma melhoria de quase 3%, com $d = 250$ metros, em relação ao DBSCAN. Observa-se pelo gráfico da Figura 6.11 que, para valores $200 \leq d \leq 300$, o algoritmo obteve os melhores resultados. Igualmente ao Compass Clustering Original, quando o valor de d aumentou

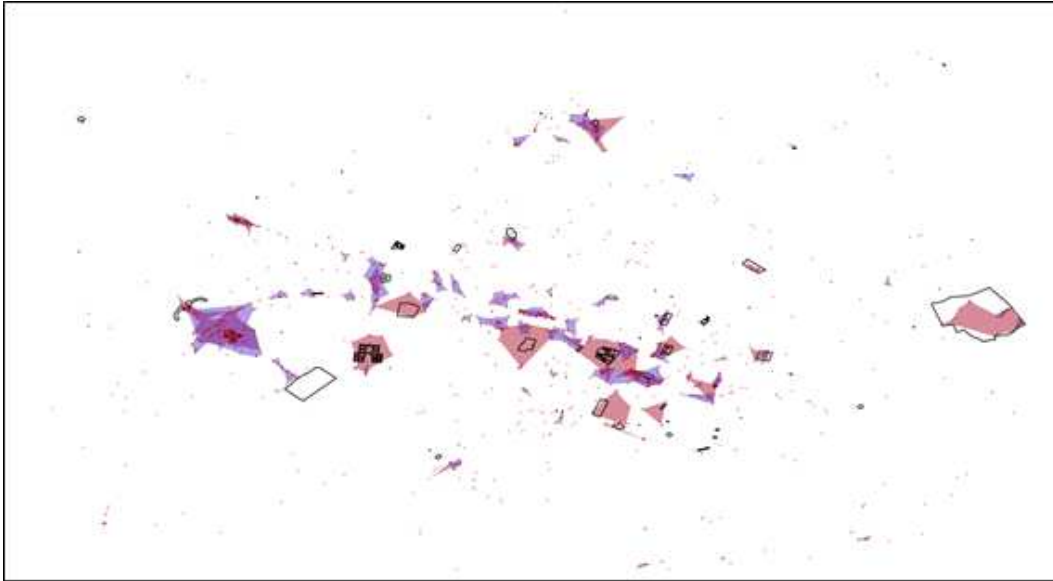


Figura 6.12: Grupos formados pelo algoritmo Compass Clustering em lilás e grupos de referência em vermelho.

bastante, o valor do F-measure caiu bastante, chegando a uma perda de aproximadamente 14% em relação ao DBSCAN. O resultado do agrupamento da versão lite do algoritmo é apresentado na Figura 6.13.

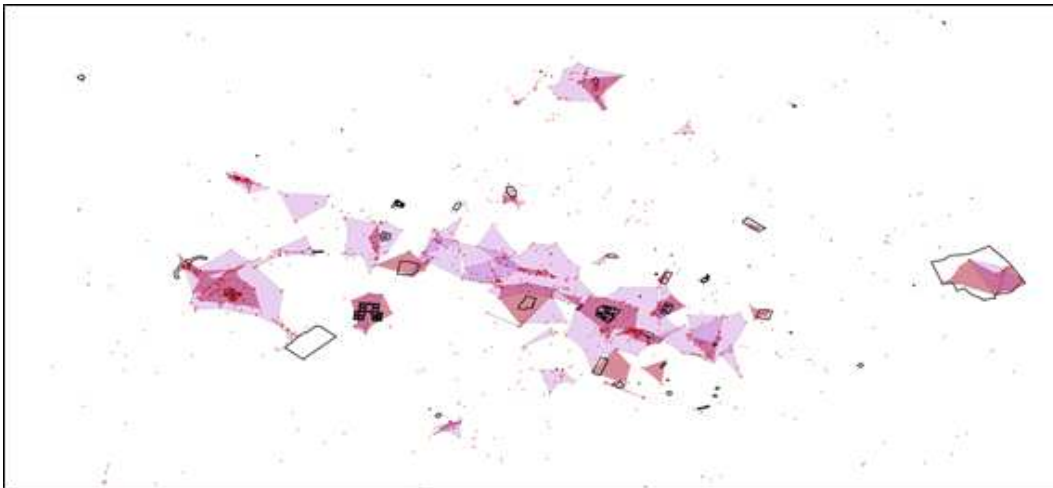


Figura 6.13: Grupos formados pelo algoritmo Compass Clustering Lite em lilás e grupos de referência em vermelho.

A Figura 6.14 apresenta um cenário com as fotografias no entorno do *Hôtel des Invalides*. A Figura 6.14(a) apresenta o ponto de interesse em verde no centro da figura e as localizações que foram capturadas as fotografias, representadas pelos quadradinhos em amarelo. Na Figura 6.14(b), os campos de visão das fotografias são representados por polígonos em cinza.

As fotografias da coleção de referência estão representados na Figura 6.14(c) como sendo os quadradinhos em rosa.

O resultado do algoritmo DBSCAN está apresentado na Figura 6.14(d) pelo polígono em azul, enquanto que o resultado do algoritmo CCLITE é apresentado em vermelho na Figura 6.14(e). Observa-se que neste cenário o algoritmo CCLITE obteve um resultado superior ao DBSCAN, uma vez que conseguiu englobar mais fotografias válidas que o DBSCAN. Inclusive, fotografias que estavam distantes, mas que apontavam na direção do ponto de interesse. O algoritmo DBSCAN conseguiu agrupar somente algumas fotografias que estavam no canto inferior do POI, tendo em vista que nesta área havia uma densidade de fotografias que obedecia aos parâmetros definidos do algoritmo. O algoritmo Compass Clustering Original não obteve nenhum cluster nesta área, indicando que a quantidade de pontos de cruzamentos não foi suficiente para a formação de um cluster. Portanto, o algoritmo CCLITE num cenário semelhante a este se sai melhor do que os outros algoritmos.

A Figura 6.15 (a) apresenta um outro cenário, contemplando as fotografias capturadas no entorno da Basílica Sagrado Coração de Jesus (*Basilique du Sacré-Couer*). A basílica está representada pelo polígono verde ao centro da imagem e as localizações das fotografias representadas por quadrados amarelo. Observa-se uma grande quantidade de fotografias bem próximas a basílica. Na Figura 6.15(a), são apresentados os campos de visão de algumas fotografias que estão distantes do POI. A Figura 6.15(c) apresenta as duas coleções de referências que foram definidas para os dois POIs que estão presentes na imagem, a basílica e o carrossel de Montmartre (Cassousel de Montemartre), representados pelos polígonos em laranja, indicados na figura.

Observa-se através das Figuras 6.15(d), (e) e (f) os resultados dos algoritmos DBSCAN, Compass Clustering Lite e Compass Clustering, respectivamente. O Compass Clustering Lite foi o algoritmo que gerou o maior *cluster*, uma vez que incluiu fotografias que estavam bem distantes, como a exemplo das fotografias (i), (ii), (iii) e (iv), que foram apresentadas na Figura 6.15(b). O DBSCAN e o Compass Clustering nesse caso foram os que saíram melhores, uma vez que criaram grupos mais semelhantes a coleção de referência. O DBSCAN criou um *cluster* adicional em uma área próxima a basílica, mas que não estava identificada como sendo um ponto de interesse.

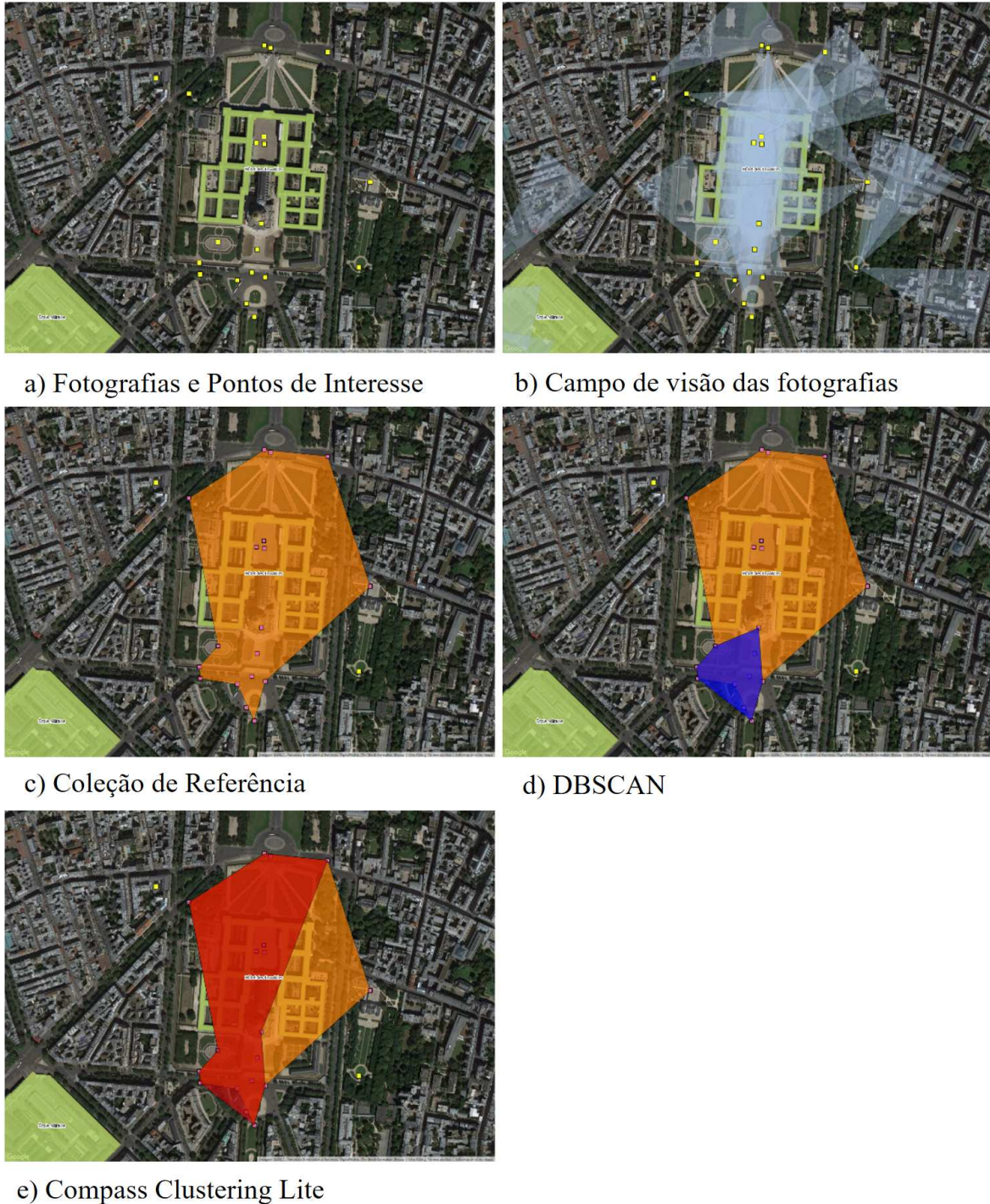


Figura 6.14: Exemplo do resultado dos algoritmos no Hôtel des Invalides. O polígono verde no centro da Figura (a) apresenta o ponto de interesse, enquanto que os quadrados em amarelo apresentam as fotografias da base de dados. Na Figura (b) os polígonos em cinza representam os campos de visão das fotografias. Na Figura (c), o polígono laranja representa a coleção de referência, onde os quadrados rosa apresenta os elementos que pertencem a coleção de referência. Na Figura (d) o polígono em azul representa o cluster gerado pelo DBSCAN. Enquanto que na Figura (e) o polígono em vermelho representa o cluster resultante do agrupamento com o Compass Clustering Lite.

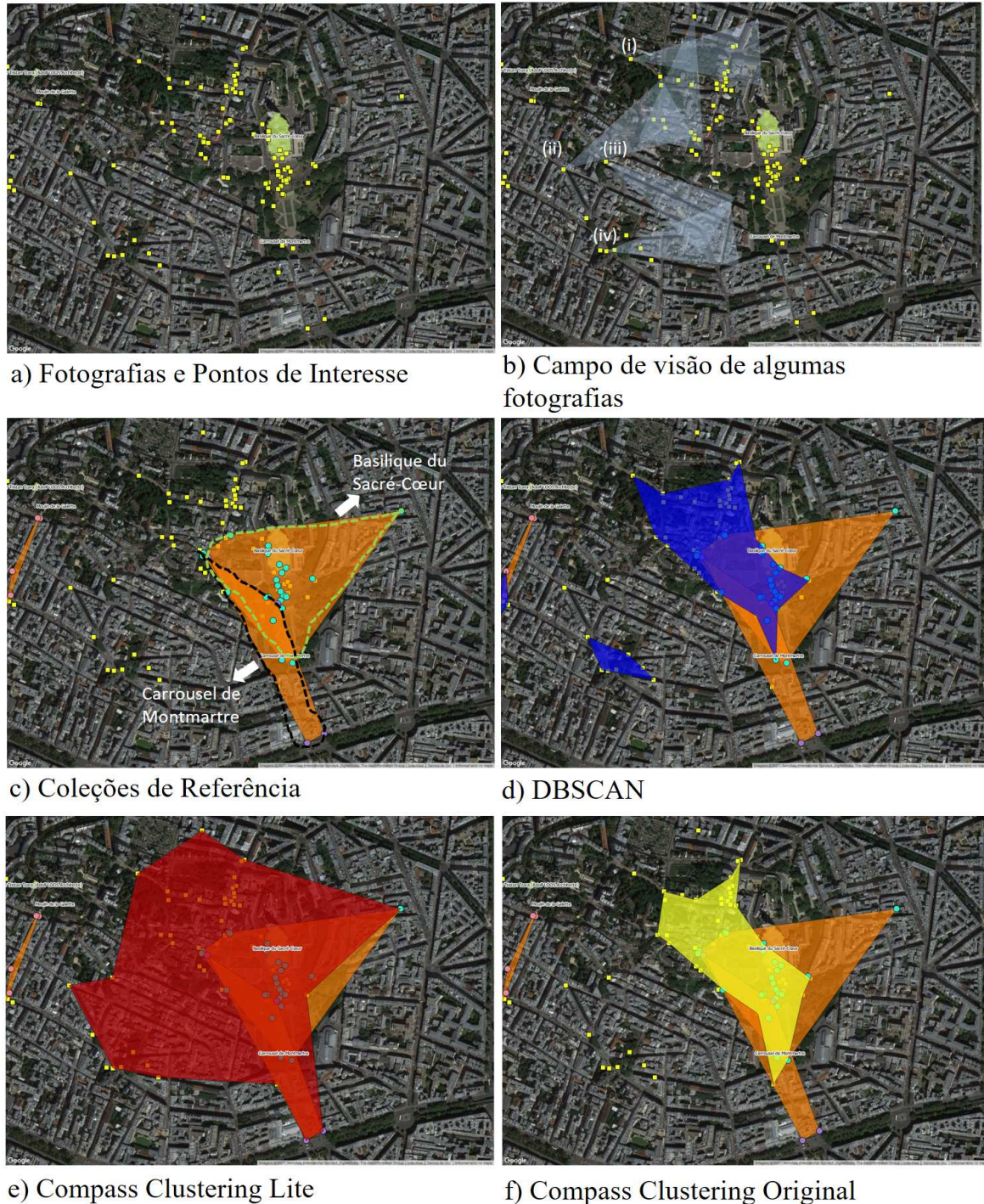


Figura 6.15: Exemplo do resultado dos algoritmos na Basílica Sagrado Coração (Basilique du Sacré-Cœur). O polígono verde no centro da Figura (a) apresenta o ponto de interesse, enquanto que os quadrados em amarelo apresentam as fotografias da base de dados. Na Figura (b), os polígonos em cinza representam os campos de visão de algumas das fotografias que serão consideradas nesse cenário. Na Figura (c), os polígonos em laranja representam as coleções de referência para dois POIs que aparecem na figura: a basílica e o Carrousel do Montmartre (Carrousel du Montmartre), onde os círculos representam os elementos que pertencem aos POIs. Na Figura (d), o polígono em azul representa os clusters gerados pelo DBSCAN. Enquanto que na Figura (e), o polígono em vermelho representa o cluster resultante do agrupamento com o Compass Clustering Lite. Por fim, a Figura (f) apresenta o cluster gerado pelo Compass Clustering.

6.3 Segunda Etapa

A segunda etapa considera todo o conjunto de fotos e não apenas uma amostra do conjunto, uma vez que uma coleção real lida com os mais variados tipos de fotografias. Portanto, não foram realizadas filtragens manuais para eliminar fotografias consideradas inválidas. Nesse caso, espera-se que o algoritmo de detecção de fotografias considere essas fotografias como ruído no processo de agrupamento das fotografias ou ser feita uma limpeza automática nos dados.

Outra razão para considerar todo o conjunto de fotografias é pelo fato de que a metodologia proposta nesta segunda etapa não leva em consideração uma avaliação por elemento do cluster, como era realizada na etapa anterior. A metodologia empregada nesta etapa realiza uma avaliação de acordo com as geometrias formadas pelos clusters em comparação com conjunto de POIs de referência (explicado a seguir).

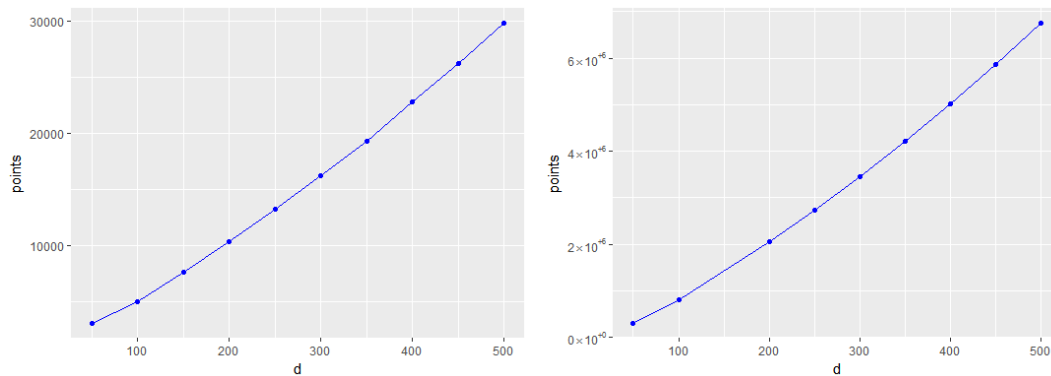
A versão lite do algoritmo Compass Clustering demonstrou superioridade nos resultados da primeira etapa em comparação ao Compass Clustering Original. Ademais, observou-se que o desempenho do Compass Clustering é afetado diretamente pelo número de pontos de cruzamento que são gerados, uma vez que esses pontos são utilizados num algoritmo de agrupamento interno.

O número de pontos de cruzamento cresce de acordo com a quantidade de segmentos de retas. Assim, o número de pontos de cruzamento é diretamente proporcional a quantidade de fotografias da coleção, ao tamanho dos segmentos de retas e a distribuição geográfica dos dados.

Por exemplo, o gráfico na Figura 6.16(a) apresenta a quantidade de pontos de cruzamento que foram gerados pelo algoritmo Compass Clustering para a amostra da coleção que foi utilizada na primeira etapa. Essa coleção contém 2.064 fotografias e gerou até quase 30.000 pontos na primeira etapa para o $d=500$ metros.

O gráfico da Figura 6.16(b) apresenta a quantidade de pontos de cruzamento que são gerados para uma das coleções extraídas para os experimentos da segunda etapa. Essa coleção contém 34.590 fotografias e gerou quase 7 milhões de pontos para o $d = 500$ metros no algoritmo Compass Clustering.

Assim, o algoritmo original do Compass Clustering foi descartado para na segunda etapa.



(a) Utilizando a amostra da coleção de Paris (b) Utilizando a coleção de Roma extraída para a segunda etapa da avaliação.

Figura 6.16: Número de pontos de cruzamento gerados no Compass Clustering

Ao invés dele, foram utilizados os algoritmos Compass Clustering Lite (CCLITE) e Compass Clustering Field of View (CCFOV) para representar os algoritmos com orientação propostos nesta tese.

Nesta seção, é detalhada como foi realizada essa avaliação. Inicialmente, é apresentado como foram levantados os dados para os experimentos realizados nesta seção. Em seguida, será apresentada a metodologia levada em consideração para a avaliação, considerando novas métricas e técnicas. Logo após, os resultados da avaliação são discutidos.

6.3.1 Base de Dados

A coleta de dados foi realizada de forma a extrair metadados de fotografias relativas a três grandes cidades turísticas - Roma, Paris e São Francisco. Para isso, utilizou-se o módulo *Photo Collector* do PhoCA que realizou o processo de extração da base de dados do Flickr.

A Tabela 6.6 apresenta algumas características das bases de dados extraídas. As seguintes colunas são apresentadas na tabela: a cidade da extração; o número total de fotografias extraídas; a quantidade de fotografias orientadas (possuem o metadado da orientação); e, por último, a quantidade de fotografias que foram utilizadas como entrada para os algoritmos.

A maior coleção coletada foi da cidade de Paris, seguida por Roma e São Francisco. Conforme já explicado na Seção 6.2, existem dois tipos de referência de compasso utilizados pelas câmeras para armazenar a informação relativa a orientação da câmeras no momento da captura da fotografia: *norte verdadeiro* e *norte magnético*. Pelos fatores descritos na

Tabela 6.6: Coleções de fotografias da base de dados.

Cidade	# Fotografias	# Orientadas	# Entrada
Paris	1.274.114	123.682 (9,70%)	97.084
Roma	675.618	47.417 (7,02%)	34.590
São Francisco	170.768	46.633 (27,30%)	30.917

Seção 6.2, foram selecionadas somente as fotografias que utilizaram o *norte verdadeiro*. Com isso, resultou um total de 162.591 fotografias a serem utilizadas nos experimentos.

A distribuição dessas fotografias podem ser conferidas visualmente em formato de mapa através das Figuras 6.17, 6.18 e 6.19 relativas respectivamente as cidades de Paris, Roma e São Francisco.

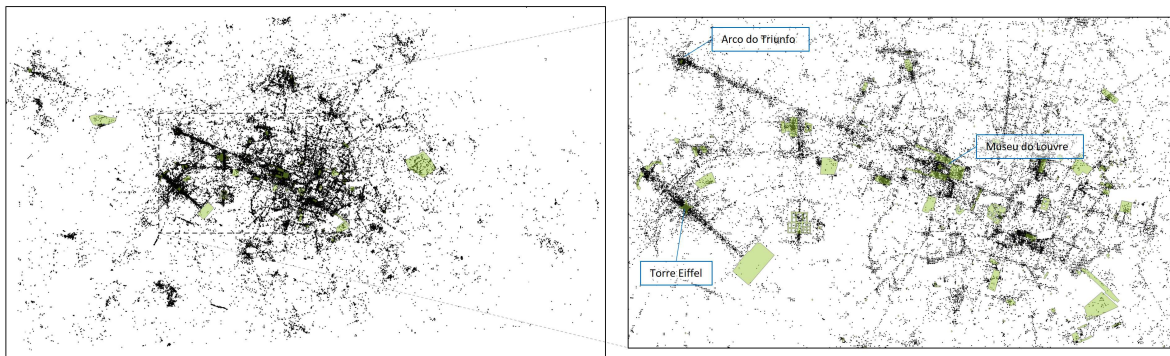


Figura 6.17: A distribuição das fotografias (pontos pretos) e os pontos de interesse (polígonos verdes) de Paris, França. A imagem da esquerda representa uma visão global da cidade, enquanto que a imagem da direita representa uma aproximação da área mais turística da cidade.

Em seguida, os conjuntos de pontos de interesses de referência foram extraídos para analisar e avaliar os *clusters* gerados nos experimentos. Esses POIs foram considerados como sendo o *ground truth*.

Os dados do Open Street Maps (OSM) foram extraídos através do MapZen. Além dos dados em formato de polígono, também foram extraídos dados em formato de pontos. Uma vez que isso permitiu ampliar o número de POIs considerados no conjunto de referência.

Somente os POIs correlacionados a turismo foram considerados, uma vez que este é o principal foco neste trabalho. No caso de pontos, foram gerados *buffers* com raio de 10 metros para representar a geometria desses POIs, uma vez que as métricas utilizadas necessitam da área dos polígonos para serem calculadas.

Foram realizadas filtragens nos POIs, considerando apenas as seguintes categorias:

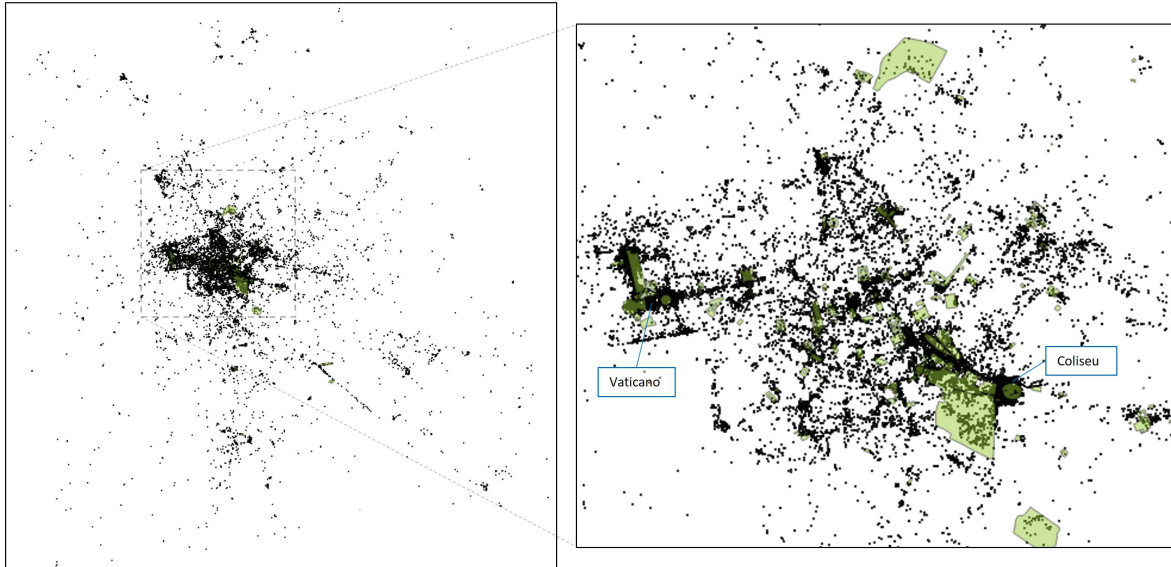


Figura 6.18: A distribuição das fotografias (pontos pretos) e os pontos de interesse (polígonos verdes) de Roma, Itália. A imagem da esquerda representa uma visão global da cidade, enquanto que a imagem da direita representa uma aproximação da área mais turística da cidade.

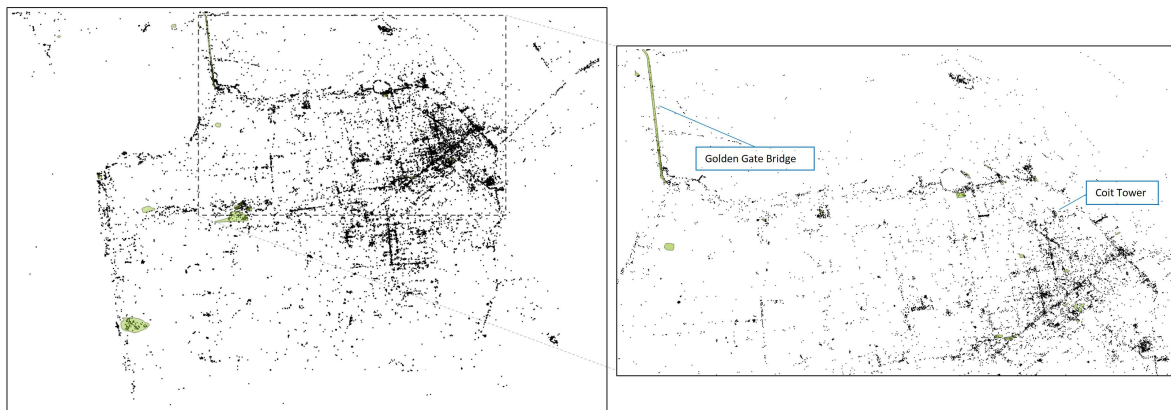


Figura 6.19: A distribuição das fotografias (pontos pretos) e os pontos de interesse (polígonos verdes) de São Francisco, Estados Unidos. A imagem da esquerda representa uma visão global da cidade, enquanto que a imagem da direita representa uma aproximação da área mais turística da cidade.

agriturismo (*agritourism*), aquário (*aquarium*), atração (*attraction*), área de camping (*camp site*), galeria (*gallery*), histórico (*historic*), ponto de referência (*landmark*), monumento (*monument*), museu (*museum*), área de picnic (*picnic site*), ruínas (*ruins*), teatro (*theater*), parque temático (*theme park*), mirante (*viewpoint*) e zoológico (*zoo*).

Essa filtragem resultou num total de 510 POIs para as três cidades consideradas (Paris, Roma e São Francisco) no qual correspondem a 264, 176 e 70 POIs, respectivamente. A

Tabela 6.7: Categorização dos Pontos de Interesse por cidade

Cidade	Categoria	Quantidade
Paris	Aquário	2
Paris	Atração	75
Paris	Galeria	49
Paris	Monumento	1
Paris	Museu	119
Paris	Área de Picnic	2
Paris	Parque Temático	1
Paris	Mirante	13
Paris	Zoológico	2
Rome	Agriturismo	1
Rome	Atração	115
Rome	Galeria	1
Rome	Museu	56
Rome	Mirante	2
Rome	Zoológico	1
San Francisco	Atração	33
San Francisco	Área de Camping	3
San Francisco	Carrosel	1
San Francisco	Ponto de Referência	1
San Francisco	Museu	25
San Francisco	Área de Picnic	2
San Francisco	Ruínas	1
San Francisco	Teatro	1
San Francisco	Mirante	1
San Francisco	Zoológico	2

Tabela 6.7 apresenta a distribuição das categorias dos pontos de interesses por cidade. O Apêndice B lista todos os pontos de interesse utilizados na base de dados de cada uma das cidades e suas categorias.

Os pontos de interesse utilizados na base de dados de referência também são apresentados graficamente (polígonos verdes) nas Figuras 6.17, 6.18 e 6.19 para as cidades de Paris, Roma e São Francisco, respectivamente.

6.3.2 Metodologia de Avaliação

O algoritmo Compass Clustering Field of View (CCFOV) foi avaliado através de uma nova metodologia, utilizando métricas e técnicas que focam na análise da intersecção das áreas dos pontos de interesse detectados. Na avaliação do Compass Clustering na primeira etapa foi utilizada uma abordagem na qual foi analisada cada uma das fotografias dos *clusters* no cálculo das métricas, baseado numa coleção de referência que foi gerada a partir da intersecção dos campos de visão das fotografias com os POIs.

A abordagem utilizada aqui para avaliar os algoritmos utiliza um método que compara a geometria dos *clusters* gerados com a geometria dos POIs de referência (*ground truth*). Foi observado que a avaliação feita desta maneira leva em consideração o aspecto do POI detectado como um todo, e não uma avaliação das fotografias específicas que se referem ao POI. Portanto, as métricas propostas aqui são diferente das métricas propostas anteriormente.

O algoritmo CCFOV foi comparado com o algoritmo Compass Clustering Lite (CCLITE) e com abordagens que só utilizam a localização das fotografias para a detecção de POIs. Como representantes dessas abordagens, foram escolhidos o DBSCAN e o P-DBSCAN, dois algoritmos de agrupamento que têm sido utilizados em trabalhos relacionados para essa tarefa (i.e.: (HU et al., 2015; KISILEVICH; MANSMANN; KEIM, 2010; THOMEE; ARAPAKIS; SHAMMA, 2016)) com algum sucesso.

A versão original do algoritmo Compass Clustering não foi considerada nessa avaliação pelos seguintes motivos: (i) o algoritmo CCLITE obteve melhores resultados na avaliação proposta na primeira etapa; e (ii) problemas de desempenho, em virtude da geração de uma enorme quantidade de pontos de cruzamentos para coleções grandes de fotografias.

O P-DBSCAN foi selecionado como o algoritmo base do CCLITE e do CCFOV. O P-DBSCAN é uma variação do DBSCAN onde o *MinPts* é na prática o número mínimo de

fotógrafos, denominado neste capítulo como *MinOwners*, que acaba evitando *clusters* onde somente uma ou poucas pessoas tiraram fotos. Isso será definido de acordo com o ajuste do parâmetro *MinOwners*. Assim, foram comparados o DBSCAN e P-DBSCAN utilizando somente a localização das fotos como entrada para o CCLITE e CCFOV.

A abordagem proposta de avaliação calcula o percentual de área de intersecção entre os *clusters* e os POIs. As métricas de precisão (*precision*) e *recall* foram utilizadas para realizar esses cálculos. Essas métricas são detalhadas a seguir.

6.3.3 Geração das geométricas dos clusters

A abordagem proposta utiliza as geometrias dos *clusters* gerados pelos algoritmos para avaliar a capacidade na descoberta de POIs. Assim, são necessárias técnicas para definir a geração dessas geometrias que os representem de forma realista, baseado nos critérios que o algoritmo utilizou no seu processo de agrupamento. Essas geometrias são áreas geográficas que devem ser comparadas às áreas geográficas dos POIs.

A Figura 6.20 ilustra o processo de geração dos polígonos representando as áreas dos clusters gerados para cada algoritmo. A Figura 6.20(a) apresenta o exemplo de um conjunto de fotografias georreferenciadas que foram agrupadas formando três *clusters* distintos, representados pelos: círculos, estrelas e quadrados. Os triângulos representam os campos de visão de cada uma das fotografias, enquanto que as geometrias menores representam a localização geográfica de onde as fotografias foram capturadas.

A Figura 6.20(b) ilustra polígonos de *clusters* gerados apenas utilizando a exata localização das fotos. Por outro lado, a Figura 6.20(c) ilustra os polígonos dos *clusters* gerados pelo algoritmo CCLITE, onde a posição de referência está localizada a d metros da localização original da captura da fotografia. No CCLITE, as novas posição foram utilizadas para gerar o polígono representando os *clusters*. Finalmente, a Figura 6.20(d) descreve os polígonos dos *clusters* gerados pelo CCFOV, onde cada foto é representada pelo seu campo de visão correspondente.

Dessa forma, a geometria de cada *cluster* foi gerada a partir de uma combinação de todos os elementos geográficos que representam os componentes do *cluster*. No caso dos algoritmos DBSCAN, P-DBSCAN e CCLITE os elementos são pontos geográficos.

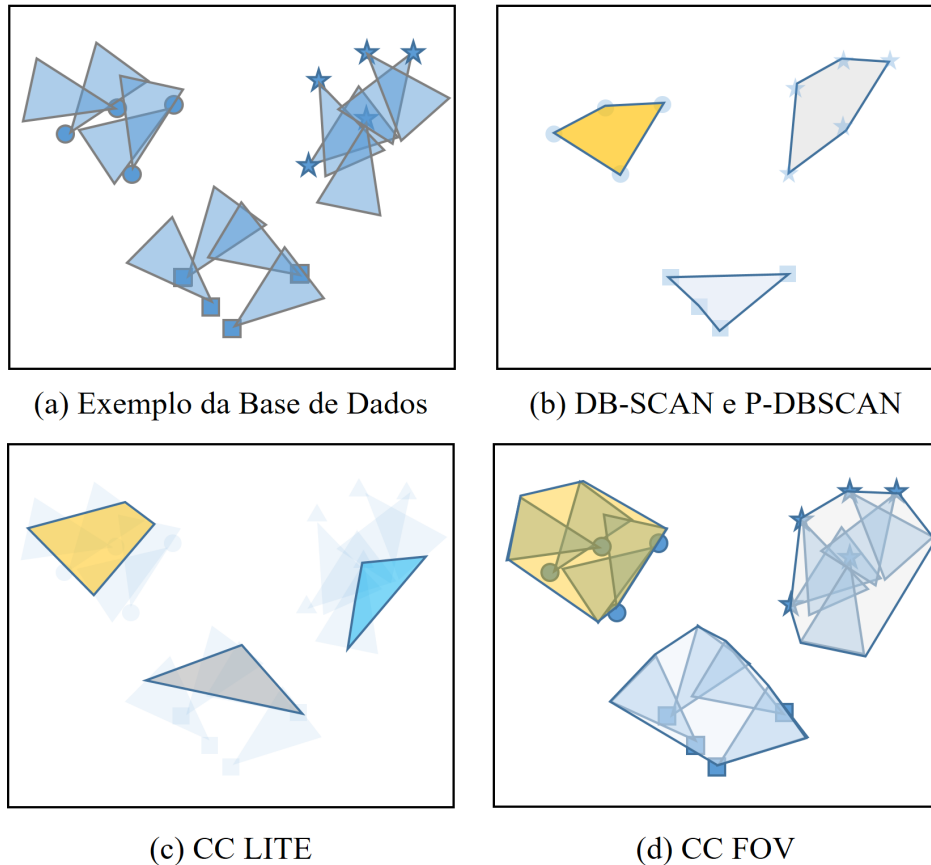


Figura 6.20: Geração dos polígonos representando os *clusters*.

Enquanto que o algoritmo CCFOV os elementos são polígonos. O algoritmo *Concave Hull*³ foi utilizado para gerar polígonos côncavos representando os *clusters*.

Os polígonos que são criados utilizando o algoritmo *Concave Hull* serão utilizados para computar a precisão e *recall* dos algoritmos. O único parâmetro do algoritmo *Concave Hull*, o k , define o nível de suavidade do polígono calculado. Esse parâmetro foi definido como 0.99, que é o máximo de suavidade possível no polígono.

Precisão e *Recall* Baseadas em Área

As métricas baseadas em área propõem uma avaliação dos *clusters*, o percentual de interseção de áreas entre as áreas dos *clusters* e dos POIs. Seja T o conjunto de POIs e C o conjunto de *clusters* gerados para uma dada cidade de interesse. A precisão e *recall* no nível de *cluster* foram definidas em termos de média de intersecção de áreas entre a área de um dado *cluster* $c \in C$ e todos os POIs da cidade, apresentadas nas equações 6.12 e 6.13.

³<http://postgis.net/docs/manual-2.2/ST_ConcaveHull.html>

Observa-se que quando um *cluster* é muito grande, é muito provável dele ter um valor alto de *recall*, mas uma baixa precisão, e vice-versa.

A Figura 6.21 ilustra graficamente como são realizados os cálculos da precisão e do *recall* baseados nas interseções das áreas geométricas dos polígonos que representam os POIs (retângulos) e os *clusters* (círculos). As áreas riscadas representam as áreas que são consideradas no cálculo.

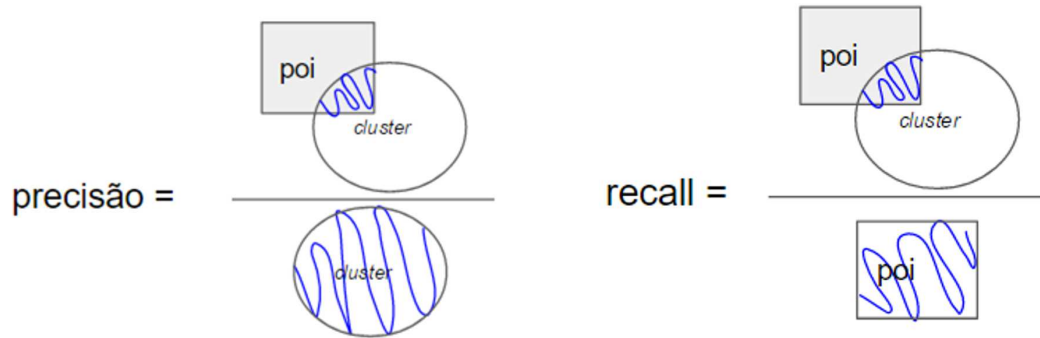


Figura 6.21: Representação gráfica do cálculo da precisão e *recall*

$$\text{cluster-}p_a(c) = \frac{1}{|T|} \sum_{t \in T} \frac{\text{intersect_area}(c, t)}{\text{area}(c)} \quad (6.12)$$

$$\text{cluster-}r_a(c) = \frac{1}{|T|} \sum_{t \in T} \frac{\text{intersect_area}(c, t)}{\text{area}(t)} \quad (6.13)$$

Dessa forma, foram definidas as métricas de precisão p_a e *recall* r_a baseadas em área como a média da precisão e *recall* no nível de cluster sobre todos os *clusters* e POIs respectivamente de uma dada cidade.

$$p_a(C) = \frac{\sum_{c \in C} \text{cluster-}p_a(c)}{|C|}, \quad r_a(C) = \frac{\sum_{c \in C} \text{cluster-}r_a(c)}{|T|} \quad (6.14)$$

6.3.4 Resultados Experimentais

A Tabela 6.8 apresenta os parâmetros que foram utilizados nos experimentos para cada um dos algoritmos. Contudo, a Tabela 6.9 apresenta os parâmetros que encontraram os melhores valores da métrica F-measure.

Tabela 6.8: Parâmetros utilizados.

Algoritmo	Parâmetros
DBSCAN	MinPts = { 10,20,30,40,50} Eps = { 10, 25, 50, 75, 100}
P-DBSCAN	MinOwners = { 10,20,30,40,50} Eps = { 10, 25, 50, 75, 100} Density Drop Threshold = 0.1
CCLITE	d = { 10, 20, 30, 40, 50, 60, 70, 80, 90, 100}
CCFOV	d = {50} Eps = {0.9, 0.95, 0.96, 0.97, 0.98, 0.99}

No caso do algoritmo CCLITE, houve uma interação de 10 a 100 metros com *steps* de 10 metros para definição do melhor valor de *d*. Esse intervalo de valores foi definido baseado na premissa que a maioria dos POIs estão localizados (em média) entre 10 e 100 metros da posição da câmera.

Para o algoritmo CCFOV, o tamanho do campo de visão *d* foi fixado em 50 metros, que representa o valor mediano se considerarmos o valor 100 metros como sendo a distância média máxima entre a câmera e o objeto, e o zero como sendo a distância mínima.

Uma vez que o algoritmo base do CCLITE foi o P-DBSCAN, foram utilizados os melhores parâmetros encontrados para os experimentos realizados com esse algoritmo (vide Tabela 6.9). Para o CCFOV, somente o parâmetro *MinOwners* foi definido usando os melhores valores encontrados para o P-DBSCAN, uma vez que o *Eps* do CCFOV tem uma escala diferente, conforme discutido na Seção 4.4.

Tabela 6.9: Os melhores parâmetros por cidade e abordagem.

	Roma	Paris	São Francisco
DBSCAN	MinPts = 10 Eps = 25	MinPts = 10 Eps = 25	MinPts = 10 Eps = 25
P-DBSCAN	MinOwners = 10 Eps = 25	MinOwners = 10 Eps = 25	MinOwners = 10 Eps = 50
CCLITE	d = 50	d = 10	d = 40
CCFOV	Eps = 0.97	Eps = 0.9	Eps = 0.95

Os resultados de precisão, recall, F-measure e total de *clusters* dos algoritmos considerando os parâmetros da Tabela 6.9 estão resumidos nos gráficos da Figura 6.22. O algoritmo DBSCAN obteve bons resultados gerais com relação a média harmônica (F-measure). Entretanto, observa-se que o mesmo gerou uma quantidade muito grande de

clusters se comparado aos outros algoritmos. Dois fatores motivaram isso: o primeiro, é o fato de que o P-DBSCAN elimina os *clusters* que não possuem uma quantidade mínima de fotografos; o segundo, é que os algoritmos que consideram a orientação ter uma tendência de só agrupar as fotografias que estejam sendo direcionadas para um mesmo local.

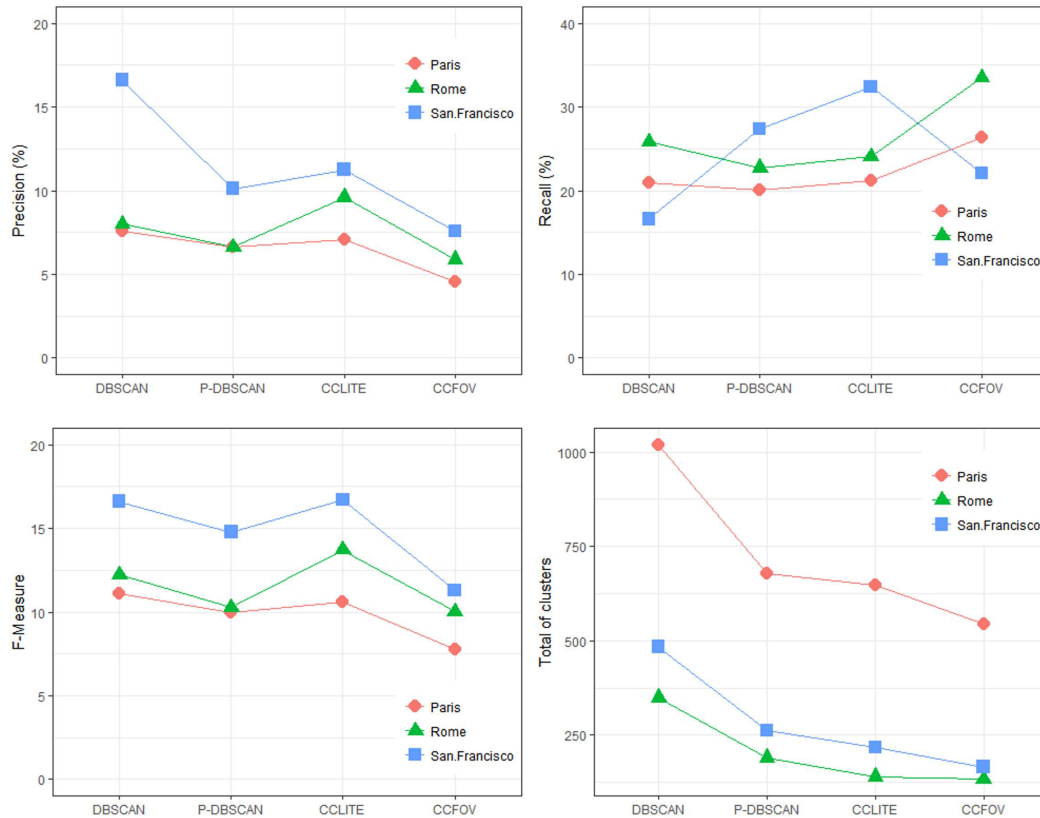


Figura 6.22: Resultados - Baseados em Área

As Figuras 6.23, 6.24 e 6.25 apresentam os resultados em formato de mapa dos experimentos que utilizaram as métricas da Tabela 6.9 para as cidades de Paris, Roma e São Francisco, respectivamente. A grande quantidade de *clusters* gerados pelo DBSCAN pode ser observado nas respectivas figuras. Com isso, observa-se que há uma grande quantidade de *clusters* em áreas que não há nenhum ponto de interesse. Isso é demonstrado com mais evidência na coleção de São Francisco, onde houve uma maior precisão, indicando que houveram vários acertos, mas que de forma geral não conseguiu cobrir as áreas dos pontos de interesse, como demonstrado pelo baixo valor do *recall*.

Outro fator que pode ser observado com relação ao DBSCAN é que os parâmetros de entrada que apresentaram um melhor resultado para o algoritmo possuem valores muito

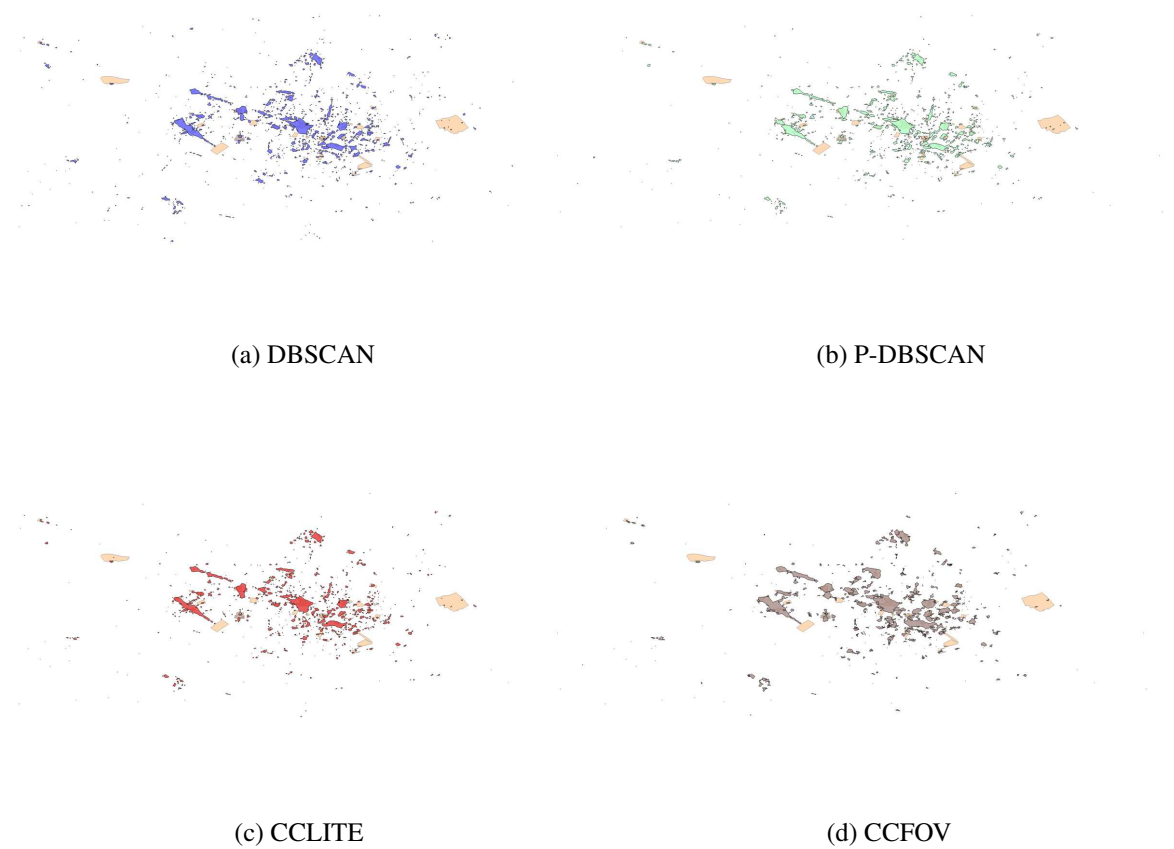


Figura 6.23: Resultados baseado em Área de Paris em formato de mapa. Os POIs da coleção de referência estão representados em laranja.

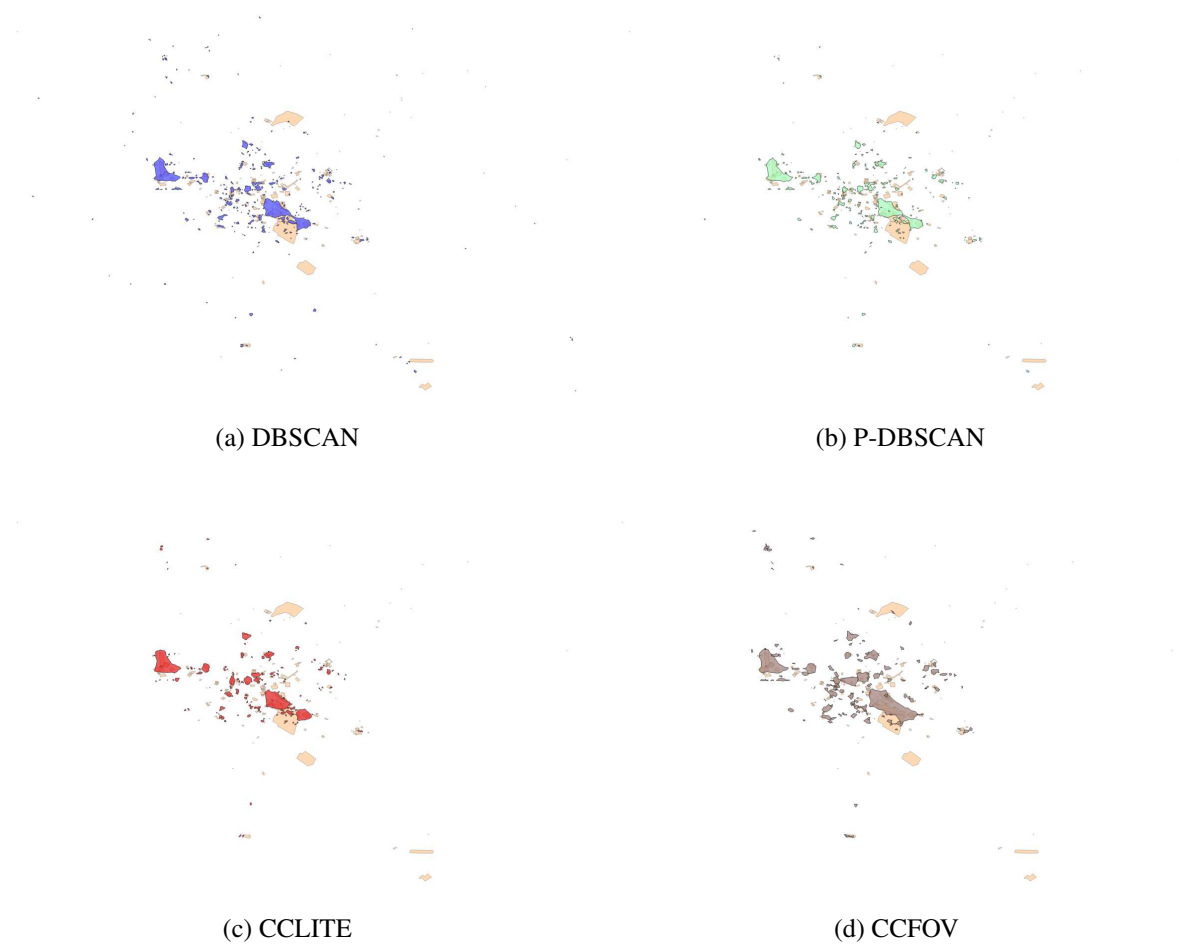


Figura 6.24: Resultados baseado em Área de Roma em formato de mapa. Os POIs da coleção de referência estão representados em laranja.

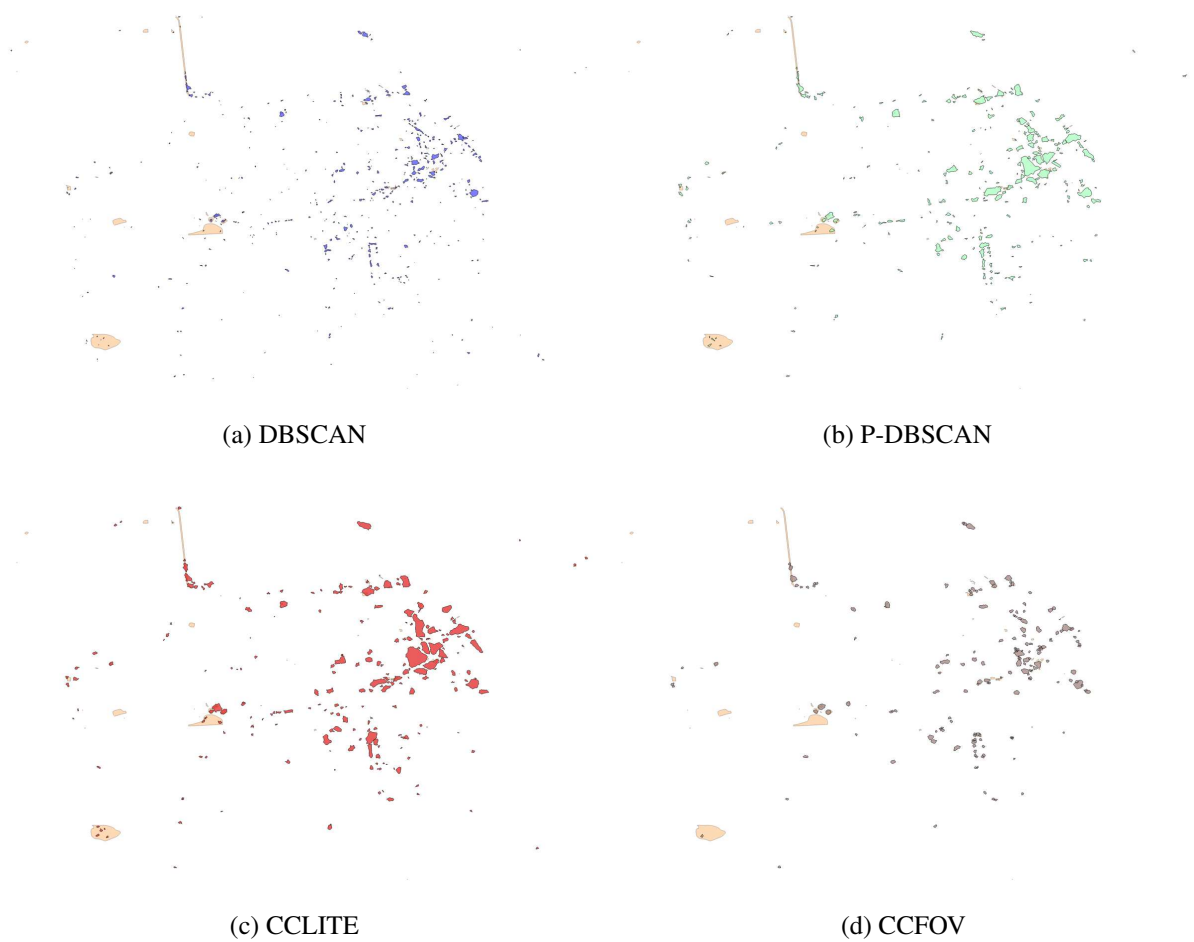


Figura 6.25: Resultados baseado em Área de São Francisco em formato de mapa. Os POIs da coleção de referência estão representados em laranja.

baixos ($MinPts=10;Eps=25$), ocasionando de forma geral a criação de *clusters* menores se comparado aos outros algoritmos.

O algoritmo Compass Clustering Lite obteve um desempenho um pouco melhor que o algoritmo P-DBSCAN, uma vez que a sua média harmônica foi levemente superior em todas as coleções. Na coleção de Roma foi onde o CCLITE conquistou um melhor resultado. Os melhores valores para o parâmetro d do CCLITE foram 10, 40 e 50 para Paris, São Francisco e Roma. Uma possível explicação para isso, pode ser pelo fato da coleção de Paris ter um número bem maior de fotografias e uma quantidade maior de POIs foi exigido um valor menor para o parâmetro d no algoritmo CCLITE. Enquanto que as outras coleções, que possuem bem menos fotografias e menos pontos de interesse, precisem de valores maiores para atingir os pontos de interesse.

O algoritmo CCFOV foi o que proporcionou um maior *recall* para as coleções de Roma e Paris, embora o total de clusters tenha sido o menor entre todos os algoritmos analisados. Entretanto, o algoritmo obteve uma precisão um pouco menor do que os outros algoritmos.

O cenário apresentado na Figura 6.26 apresenta as fotografias ao redor de alguns pontos de interesse em São Francisco, dentre eles o *De Young Museum*, *Japanese Tea Garden* e o *California Academy of Sciences*. A Figura 6.26(a) apresenta as localizações geográficas onde foram capturadas as fotografias no entorno desses POIs. Observa-se que muitas fotos foram capturadas dentro das áreas relativas aos POIs e em seus arredores. A Figura 6.26(b) apresenta os campos de visão das câmeras no momento da captura dessas fotografias.

Observa-se que os quatro algoritmos, representados nas Figuras 6.26(c), (d), (e) e (f) conseguem identificar os POIs. Entretanto, o algoritmo que obtém um resultado mais preciso é o CCFOV, uma vez que os clusters conseguem cobrir com maior aproximação os três *clusters*. Nessa situação, o algoritmo DBSCAN gerou clusters bem menores que não cobrem tão bem os POIs, além de ter gerado três outros clusters que não interceptam nenhum POI e dentro de um dos POIs existem dois clusters gerados.

Os algoritmos P-DBSCAN e CCLITE geraram também apenas três clusters nesta área, mas eles possuem áreas maiores, que englobam mais áreas fora dos pontos. Observa-se que na área tracejada em verde entre o *De Young Museum* e a *California Academy of Sciences* existem fotografias que são direcionadas para ambos os lados, porém mais próximas do museu. Porém, no caso do algoritmo CCFOV, somente as fotografias que

estavam direcionadas para o museu foram agrupadas para o cluster correlacionado. Esses pontos estão apontados pela na Figura 6.26(f). Portanto, o algoritmo CCFOV tende a ter um comportamento melhor que os outros algoritmos em situações semelhantes a esse cenário.

6.4 Considerações Finais

Neste capítulo foram apresentadas duas etapas da avaliação dos algoritmos propostos. A primeira etapa foi responsável por avaliar os algoritmos Compass Clustering e Compass Clustering Lite em uma coleção de fotografias extraída do Flickr relativa a cidade de Paris. Na segunda etapa, foi realizada uma avaliação que considerou o algoritmo Compass Clustering Field of View, utilizando uma quantidade de dados bem maior extraídos de Paris, Roma e São Francisco.

Na primeira etapa, o algoritmo DBSCAN obteve F-measure de 0,497622204 com os parâmetros mínimo de pontos = 5 e distância espacial mínima = 100 metros. Esses valores foram utilizados internamente nos algoritmos Compass Clustering Original e Lite. O algoritmo Compass Clustering Lite apresentou o melhor resultado, obtendo uma melhora de quase 3% em relação ao DBSCAN, para o valor de $d = 250$ metros. Em seguida, o Compass Clustering Original obteve um resultado aproximadamente 2% melhor que o DBSCAN para o $d = 100$ metros.

Enquanto que na segunda etapa, nas métricas *baseadas em área*, observou-se que o CCLITE superou os *baselines* em termos de precisão, recall e F-measure. O CCFOV obteve um resultado melhor em termos de *recall*. Além disso, os algoritmos considerando orientação geraram uma quantidade bem menor de *clusters*, mostrando que os *clusters* detectados possuem uma melhor qualidade. Além disso, os algoritmos com orientação reduziram os *clusters* nas áreas que não existem POIs. Essa é uma qualidade que é desejada na detecção de pontos de interesse.

Além disso, foram apresentadas determinados cenários nas avaliações, avaliando a influência da orientação nos algoritmos. Com isso, foram apresentadas situações onde os algoritmos Compass Clustering tendem a ter melhores resultados.

No próximo capítulo, são apresentados as conclusões desta Tese. As conclusões são descritas e são propostos trabalhos para serem desenvolvidos como continuidade desta

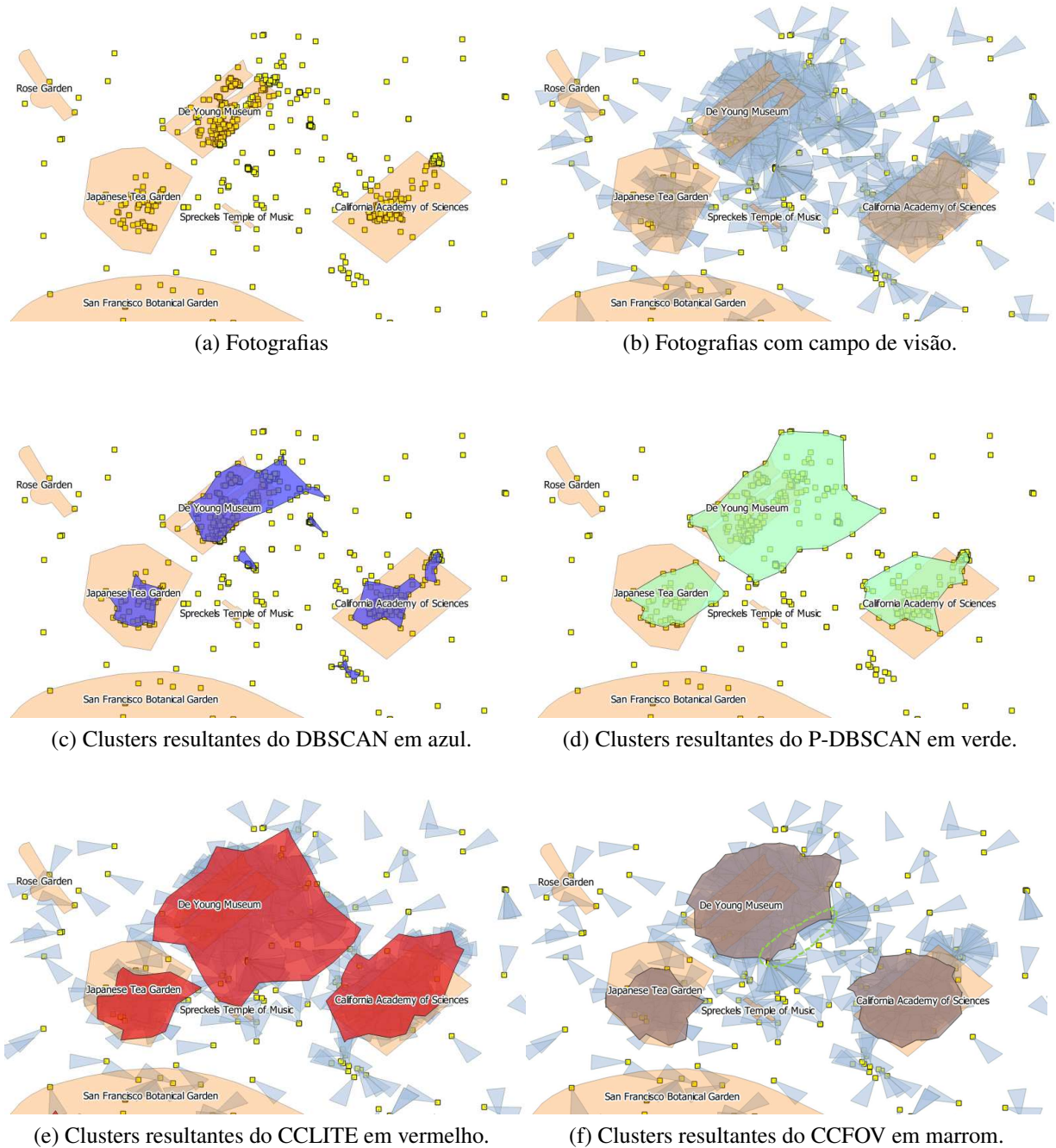


Figura 6.26: Estudo de caso em São Francisco. Os pontos de interesse estão representados pelos polígonos em laranja, a localização das fotografias estão representadas como quadradinhos em amarelo e os campos de visão os polígonos em cinza.

pesquisa.

Capítulo 7

Conclusões e Trabalhos Futuros

Neste capítulo, são apresentadas as conclusões obtidas a partir dos resultados alcançados durante o trabalho de pesquisa de doutorado e as diretrizes para os trabalhos futuros que podem ser desenvolvidos a partir dos achados deste trabalho.

7.1 Conclusões

Esta pesquisa tratou da detecção de pontos de interesses a partir de fotografias georreferenciadas apresentando três diferentes maneiras de explorar os metadados de localização e orientação geográfica das fotografias para auxiliar a esse processo.

Assim, nesta pesquisa foram propostos e implementados três novos algoritmos para a detecção de pontos de interesses a partir do uso de fotografias georreferenciadas e orientadas: o Compass Clustering, Compass Clustering Lite e o Compass Clustering Field of View.

O primeiro algoritmo proposto, o Compass Clustering, possui três passos que levam em consideração os pontos de cruzamento entre segmentos de reta que representam a orientação das fotografias. O segundo algoritmo é uma versão simplificada que utiliza pontos geográficos representando o local para onde as câmeras estavam direcionadas no momento da captura das fotografias. Por fim, o terceiro algoritmo que usa a intersecção entre os campos de visão das fotografias no processo de detecção de POIs.

O algoritmo Compass Clustering foi citado na literatura, em (THOMEE, 2013), como sendo um dos primeiros trabalhos a considerar o aspecto da orientação para detecção de POIs.

A hipótese de que o uso da orientação torna o processo de detecção de POIs mais preciso foi comprovado através dos experimentos apresentados no Capítulo 6, onde os algoritmos considerando a orientação em algumas situações obtiveram vantagens superiores aos algoritmos que não consideravam a orientação.

Além disso, também foram propostas e implementadas novas metodologias de avaliação do processo de detecção de pontos de interesse, tendo em vista que as metodologias da literatura não se aplicavam às abordagens propostas.

A primeira metodologia de avaliação proposta faz uso de uma coleção de referência que é gerada a partir do conhecimento prévio de uma base de POIs. Em seguida, o algoritmo que se deseja avaliar é executado utilizando a coleção de fotografias. Esses algoritmos não conhecem a base de pontos de interesse. Em seguida, os elementos de cada *cluster* gerado são comparados com a coleção de referência em termos de precisão e *recall*.

Essa metodologia foi adotada na avaliação dos algoritmos Compass Clustering e Compass Clustering Lite. Os experimentos demonstraram cenários onde os algoritmos conseguiram sobrepor aos resultados do algoritmo sem considerar a orientação.

A segunda metodologia proposta leva em consideração a análise das intersecções entre as áreas geográficas dos *clusters* e dos POIs de referência. Para isso, foram propostos modelos para criação das áreas representando os *clusters* levando em consideração os critérios utilizados para o agrupamento dos dados. Nesta metodologia foram propostas métricas considerando precisão e *recall* baseados na intersecção de áreas.

Esse método foi aplicado para coleções de fotografias relativas a três grandes cidades: Paris, Roma e São Francisco. O algoritmo Compass Clustering Lite e Compass Clustering Field of View foram avaliados utilizando internamente o algoritmo P-DBSCAN. Esse algoritmo é baseado no DBSCAN, mas considera o número mínimo de fotógrafos em cada grupo. Esses algoritmos foram comparados com duas abordagens sem utilizar a orientação: DBSCAN e o P-DBSCAN.

Nesta avaliação, embora o DBSCAN tenha obtido resultados semelhantes ao CCLITE com relação ao F-measure, ele gerou uma quantidade muito grande de clusters. Além disso, esses clusters gerados de forma geral são bem menores se comparados as outras abordagens, conseguindo cobrir áreas menores dos POIs de referência.

Quando comparado ao P-DBSCAN, o CCLITE obteve resultados um pouco superiores

em termos de precisão e *recall*, apesar da quantidade de *clusters* gerados ter sido menor. Isso indica que houve uma melhor qualidade nos *clusters* que foram gerados.

Com relação ao algoritmo CCFOV, o mesmo proporcionou de modo geral um maior *recall*, mas uma precisão um pouco abaixo dos outros algoritmos. O CCFOV obteve um número bem menor de *clusters*, indicando uma maior qualidade nos *clusters* gerados, uma vez que gerou menos *clusters* em áreas sem POIs.

Outra importante contribuição deste trabalho de doutorado foi a proposta e implementação do PhoCA (Photo Clustering Analyzer) - uma ferramenta Web extensível que permite automatizar todo o processo de *clustering* de fotografias. Essa ferramenta foi utilizada nos experimentos do trabalho, demonstrando sua efetiva aplicação e utilidade no que se propõe.

7.2 Trabalhos Futuros

Esta pesquisa abre a oportunidade para uma série de trabalhos podem ser desenvolvidos para dar continuidade a esta pesquisa. Dessa forma, as propostas de trabalhos futuros foram categorizadas a seguir de acordo com o que elas lidam:

1. **Aperfeiçoamento dos Algoritmos:** algumas questões podem ser analisadas com o intuito de realizar uma melhoria nos algoritmos de agrupamento. Uma possibilidade seria incluir um último passo de aperfeiçoamento para analisar algumas questões relativa a possibilidade de unir ou dividir os *clusters* localizados próximos geograficamente. A detecção de padrões nos dados textuais associados às fotografias (e.g: *tags*) ou a análise do conteúdo visual das imagens através de técnicas de reconhecimento visual podem ser utilizadas nesse aperfeiçoamento de *clusters*. Outro trabalho futuro é a implementação do algoritmo Compass Clustering Dynamic, o qual considera um valor para o tamanho do segmento de reta de forma dinâmico, baseado no valor da distância focal de cada fotografia. Além disso, a geração da geometria relativa aos POIs também poderá ser aperfeiçoada afim de conseguir estimar melhor a área relativa ao POI.
2. **Base de Dados:** uma sugestão de trabalho futuro seria a aplicação das técnicas

apresentadas neste trabalho para outras cidades. Outra questão interessante a analisar seria a possibilidade de utilizar mecanismos automatizados para a limpeza das bases de dados. Por exemplo, pode-se analisar a possibilidade de filtrar automaticamente fotografias que não tem foco em objetos, animais e pessoas. Além disso, também a possibilidade de distinguir fotografias em ambientes fechados (*indoor*) ou ao ar livre (*outdoor*) para tratá-las diferentemente nos algoritmos. Existem técnicas baseada no uso de sensores, por exemplo (ZHOU et al., 2012), e técnicas baseadas em características de baixo nível das imagens, como em (SZUMMER; PICARD, 1998). A ideia seria utilizar alguma técnica neste sentido para auxiliar ao processo de classificação automática.

3. **Análises:** um outro trabalho futuro seria analisar o fator temporal na visitação dos POIs nas cidades, com o intuito de descobrir padrões de visitação ao longo de determinados períodos. Por exemplo, em períodos mais frios há uma tendência em visitar lugares mais fechados ou com programação de inverno, enquanto que nas estações mais quentes há uma tendência de visitar lugares abertos. Outra possibilidade seria analisar a evolução das cidades ao longo dos anos, identificando pontos de interesse que foram criados em determinado período, ou que deixaram de existir.
4. **Detecção de Eventos:** a detecção de eventos está diretamente relacionada a detecção de pontos de interesses, uma vez que um evento acontece em uma determinada localização com uma dada periodicidade. Sendo que muitas vezes a localização que acontece o evento, não é por si só, um ponto de interesse durante o período que não acontece o evento. Dessa forma, um possível trabalho futuro seria analisar as técnicas para detecção de eventos e implantar dentro das abordagens propostas nesta Tese um mecanismo para identificação de eventos. Essa solução, possibilita inclusive aperfeiçoar a detecção de pontos de interesse.
5. **Aplicações:** os resultados da detecção de pontos de interesse devem ser incorporados a sistemas que permitam usuários finais terem acesso aos resultados dos POIs detectados. Para isso, deve-se desenvolver uma ferramenta que possa incorporar os resultados dos algoritmos. Uma possibilidade seria um Sistema de Recomendação de Pontos de Interesse que fizesse a recomendação dos principais pontos a se conhecer

nas cidades de acordo com determinados critérios escolhidos pelos usuários.

Bibliografia

AGRAWAL, R. et al. Automatic subspace clustering of high dimensional data for data mining applications. *Proceedings of the 1998 ACM SIGMOD international conference on Management of data SIGMOD 98*, v. 11, p. 94–105, 1998. ISSN 13845810. Disponível em: <<http://portal.acm.org/citation.cfm?doid=276304.276314>>.

ANKERST, M. et al. Optics: Ordering points to identify the clustering structure. In: *ACM Sigmod Record*. [s.n.], 1999. p. 49–60. ISBN 1581130848. ISSN 01635808. Disponível em: <<http://dl.acm.org/citation.cfm?id=304187>>.

APRIL, E. *Exchangeable image file format for digital still cameras: Exif Version 2.2*. [S.l.], 2002.

BENNETT, J.; LANNING, S. The Netflix Prize. *KDD Cup and Workshop*, p. 3–6, 2007. ISSN 1554351X.

BRILHANTE, I. et al. TripBuilder: A Tool for Recommending Sightseeing Tours. In: *Advances in Information Retrieval*. [S.l.]: Springer, 2014. p. 771–774.

BRIN, S.; PAGE, L. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, v. 56, p. 3825–3833, 2012. ISSN 13891286.

BUCCAFURRI, F. et al. A quad-tree based multiresolution approach for two-dimensional summary data. *Information Systems*, v. 36, p. 1082–1103, 2011. ISSN 03064379.

CHEKURI, C.; PAL, M. A recursive greedy algorithm for walks in directed graphs. In: IEEE. *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*. [S.l.], 2005. p. 245–253.

CHEN, S.-C. et al. Multimedia Data Mining for Traffic Video Sequences. In: CITeseer. *MDM/KDD*. [S.l.], 2001. p. 78–86.

CLEMENTS, M. et al. Using flickr geotags to predict user travel behaviour. *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval - SIGIR '10*, p. 851, 2010. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1835449.1835648>>.

COMANICIU, D.; MEER, P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 24, 2002. ISSN 0162-8828.

CORE, I. *IPTC Photo Metadata: Core 1.1/Extension 1.1*. [S.l.], 2010. 1–55 p.

CRANDALL, D. J. et al. Recognizing landmarks in large-scale social image collections. In: *Large-Scale Visual Geo-Localization*. [S.l.]: Springer, 2016. p. 121–144.

De Choudhury, M. et al. Constructing travel itineraries from tagged geo-temporal breadcrumbs. *Proceedings of the 19th international conference on World wide web - WWW '10*, p. 1083, 2010. Disponível em: <<http://dl.acm.org/citation.cfm?id=1772690.1772815>>.

DEMPSTER, A. P. et al. Maximum likelihood from incomplete data via the EM algorithm. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, v. 39, p. 1–38, 1977. ISSN 0035-9246. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.133.4884>>.

DEY, A. K.; ABOARD, G. D. Towards a Better Understanding of Context and Context-Awareness. *Computing Systems*, v. 40, p. 304–307, 1999. ISSN 00219266.

ERICSSON; QUALCOMM; FACEBOOK. *A Focus on Efficiency*. [S.l.], 2013.

ESTER, M. et al. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In: *Second International Conference on Knowledge Discovery and Data Mining*. [s.n.], 1996. p. 226–231. ISBN 1577350049. ISSN 09758887. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.2930>>.

GUHA, S.; RASTOGI, R.; SHIM, K. CURE: An Efficient Clustering Algorithm for Large Databases. In: *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data - SIGMOD '98*. [s.n.], 1998. p. 73–84. ISBN 0897919955. ISSN 03064379. Disponível em: <<http://portal.acm.org/citation.cfm?doid=276304.276312>>.

HALL, M. et al. The WEKA Data Mining Software : An Update. *SIGKDD Explorations*, v. 11, p. 10–18, 2009. ISSN 19310145.

HAN, J.; KAMBER, M. *Data Mining: Concepts and Techniques*. [S.l.: s.n.], 2011.

HINNEBURG, A. Denclue 2.0: Fast clustering based on kernel density estimation. *Advances in Intelligent Data Analysis VII*, v. 4723, p. 70–80–80, 2007. ISSN 0302-9743. Disponível em: <<http://www.springerlink.com/index/W7236UM043J39L11.pdf>>.

HU, Y. et al. Extracting and understanding urban areas of interest using geotagged photos. *Computers, Environment and Urban Systems*, Elsevier, v. 54, p. 240–254, 2015.

HUC, C.; LEVOIR, T.; NONON-LATAPIE, M. Metadata : Models and conceptual limits. In: *Metadata Conference*. [S.l.: s.n.], 1997.

JADHAV, S. R.; KUMBARGOUDAR, P. Multimedia Data Mining in Digital Libraries: Standards and Features. In: *Proceedings of conference Recent advances in Information Science and Technology READIT*. [S.l.: s.n.], 2007. p. 54–59.

JAIN, S.; SEUFERT, S.; BEDATHUR, S. Antourage: mining distance-constrained trips from flickr. In: *Proceedings of the 19th international conference on World wide web*. [s.n.], 2010. v. 52, p. 1121–1122. ISBN 9781605587998. Disponível em: <<http://dl.acm.org/citation.cfm?id=1772834>>.

- KARYPIS, G.; HAN, E.-H.; KUMAR, V. *Chameleon: hierarchical clustering using dynamic modeling*. 1999. 68–75 p.
- KAUFMAN, L.; ROUSSEEUW, P. J. *Finding groups in data: an introduction to cluster analysis*. [S.l.]: John Wiley & Sons, 2009. v. 344.
- KENNEDY, L. S.; NAAMAN, M. Generating diverse and representative image search results for landmarks. *Proceeding of the 17th international conference on World Wide Web*, p. 297–306, 2008. Disponível em: <<http://portal.acm.org/citation.cfm?id=1367539>>.
- KISILEVICH, S.; MANSMANN, F.; KEIM, D. P-DBSCAN: A Density Based Clustering Algorithm for Exploration and Analysis of Attractive Areas Using Collections of Geo-tagged Photos. In: *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research and Application*. New York, NY, USA: ACM, 2010. (COM.Geo '10), p. 38:1—38:4. ISBN 978-1-4503-0031-5. Disponível em: <<http://doi.acm.org/10.1145/1823854.1823897>>.
- KLEINBERG, J. et al. Mapping the World's Photos. In: *Proceedings of the 18th International Conference on World Wide Web, Madrid*. [s.n.], 2009. p. 761–770. ISBN 9781605584874. ISSN 08963207. Disponível em: <<http://www2009.org/proceedings/pdf/p761.pdf>>.
- LACERDA, Y. A. *PhotoGeo - Uma Biblioteca Digital de Fotografias com Suporte de Geoprocessamento*. Tese (Doutorado) — Universidade Federal de Campina Grande, 2009.
- LACERDA, Y. A. et al. Compass clustering. In: *Proceedings of the 18th Brazilian symposium on Multimedia and the web - WebMedia '12*. New York, New York, USA: ACM Press, 2012. p. 281. ISBN 9781450317061. Disponível em: <<http://dl.acm.org/citation.cfm?id=2382636.2382696>>.
- LACERDA, Y. A. et al. PhotoGeo: A self-organizing system for personal photo collections. In: *Proceedings - 10th IEEE International Symposium on Multimedia, ISM 2008*. [S.l.: s.n.], 2008. p. 258–265. ISBN 9780769534541.
- LACERDA, Y. A. et al. Phoca: An extensible service-oriented tool for photo clustering analysis. In: *29th SBBD - Demos and Applications Session*. [S.l.: s.n.], 2014. p. 227–232.
- LAERE, O. V.; SCHOCKAERT, S.; DHOEDT, B. Towards automated georeferencing of flickr photos. In: ACM. *Proceedings of the 6th Workshop on Geographic information Retrieval*. [S.l.], 2010. p. 5.
- LEE, I.; CAI, G.; LEE, K. Points-of-interest mining from people's photo-taking behavior. In: IEEE. *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. [S.l.], 2013. p. 3129–3136.
- MACQUEEN, J. B. Kmeans Some Methods for classification and Analysis of Multivariate Observations. *5th Berkeley Symposium on Mathematical Statistics and Probability 1967*, v. 1, p. 281–297, 1967. ISSN 00970433. Disponível em: <<http://projecteuclid.org/euclid.bsmsp/1200512992>>.

MAIMON, O.; ROKACH, L. Introduction to Knowledge Discovery and Data Mining. In: *Data Mining and Knowledge Discovery Handbook*. [S.l.: s.n.], 2010. p. 1–17. ISBN 978-0-387-09822-7.

MARINHO, L. B. et al. Extracting geospatial preferences using relational neighbors. *arXiv preprint arXiv:1204.1528*, 2012.

MASTERCARD. *2014 Global Destination Cities Index*. [S.l.], 2014.

MATYAS, C.; SCHLIEDER, C. A spatial user similarity measure for geographic recommender systems. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. [S.l.: s.n.], 2009. v. 5892 LNCS, p. 122–139. ISBN 3642104355. ISSN 03029743.

NAAMAN, M. et al. Automatic organization for digital photographs with geographic coordinates. In: IEEE. *Digital Libraries, 2004. Proceedings of the 2004 Joint ACM/IEEE Conference on*. [S.l.], 2004. p. 53–62.

PAPADOPOULOS, S. et al. Cluster-based landmark and event detection for tagged photo collections. *IEEE MultiMedia*, v. 18, n. 1, p. 52–63, 2011.

PHAN, T. et al. Collaborative Recommendation of Photo-Taking Geolocations. In: ACM. *Proceedings of the 3rd ACM Multimedia Workshop on Geotagging and Its Applications in Multimedia*. [S.l.], 2014. p. 11–16.

POPESCU, A.; GREFENSTETTE, G.; MOËLLIC, P.-A. Mining tourist information from user-supplied collections. In: *Proceeding of the 18th ACM conference on Information and knowledge management - CIKM '09*. [s.n.], 2009. p. 1713. ISBN 9781605585123. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1645953.1646211>>.

POPESCU, A.; SHABOU, A. Towards precise poi localization with social media. In: *Proceedings of the 21st ACM International Conference on Multimedia*. New York, NY, USA: ACM, 2013. (MM '13), p. 573–576. ISBN 978-1-4503-2404-5. Disponível em: <<http://doi.acm.org/10.1145/2502081.2502151>>.

QUACK, T.; LEIBE, B.; GOOL, L. V. World-scale mining of objects and events from community photo collections. In: ACM. *Proceedings of the 2008 international conference on Content-based image and video retrieval*. [S.l.], 2008. p. 47–56.

RAJARAMAN, A.; ULLMAN, J. D. *Mining of Massive Datasets*. 2010.

RATTENBURY, T.; GOOD, N.; NAAMAN, M. Towards automatic extraction of event and place semantics from flickr tags. In: ACM. *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. [S.l.], 2007. p. 103–110.

RATTENBURY, T.; NAAMAN, M. Methods for extracting place semantics from flickr tags. *ACM Transactions on the Web (TWEB)*, ACM, v. 3, n. 1, p. 1, 2009.

RATTENBURY, T.; NAAMAN, M. Methods for extracting place semantics from flickr tags. *ACM Trans. Web*, ACM, New York, NY, USA, v. 3, n. 1, p. 1:1–1:30, jan. 2009. ISSN 1559-1131. Disponível em: <<http://doi.acm.org/10.1145/1462148.1462149>>.

RICHARDSON, L.; RUBY, S. *RESTful Web Services*. [s.n.], 2008. 440 p. ISSN 10897801. ISBN 978-0-596-52926-0. Disponível em: <<http://medcontent.metapress.com/index/A65RM03P4874243N.pdf>>.

SERDYUKOV, P.; MURDOCK, V.; ZWOL, R. V. Placing flickr photos on a map. In: ACM. *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. [S.l.], 2009. p. 484–491.

SERGIEH, H. M. et al. Geo-based Automatic Image Annotation. In: *Proceedings of the 2Nd ACM International Conference on Multimedia Retrieval*. New York, NY, USA: ACM, 2012. (ICMR '12), p. 46:1—46:8. ISBN 978-1-4503-1329-2. Disponível em: <<http://doi.acm.org/10.1145/2324796.2324850>>.

SHEIKHOLESAMI, G.; CHATTERJEE, S.; ZHANG, A. Wavecluster: A multi-resolution clustering approach for very large spatial databases. In: *Proceedings of the International Conference on Very Large Data Bases*. [S.l.: s.n.], 1998. p. 428–439. ISBN 1558605665. ISSN 10477349.

SHUMAKER, B.; SINNOTT, R. Astronomical computing: 1. computing under the open sky. 2. virtues of the haversine. *Sky and telescope*, v. 68, p. 158–159, 1984.

SMEULDERS, A. et al. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, 2000. ISSN 0162-8828.

STÜTZLE, T.; HOOS, H. H. MAX-MIN Ant System. *Future Generation Computer Systems*, v. 16, p. 889–914, 2000. ISSN 0167739X.

SZUMMER, M.; PICARD, R. W. Indoor-outdoor image classification. In: IEEE. *Content-Based Access of Image and Video Database, 1998. Proceedings., 1998 IEEE International Workshop on*. [S.l.], 1998. p. 42–51.

THOMEE, B. Localization of points of interest from georeferenced and oriented photographs. In: ACM. *Proceedings of the 2nd ACM international workshop on Geotagging and its applications in multimedia*. [S.l.], 2013. p. 19–24.

THOMEE, B.; ARAPAKIS, I.; SHAMMA, D. A. Finding social points of interest from georeferenced and oriented online photographs. *ACM Trans. Multimedia Comput. Commun. Appl.*, ACM, New York, NY, USA, v. 12, n. 2, p. 36:1–36:23, jan. 2016. ISSN 1551-6857. Disponível em: <<http://doi.acm.org/10.1145/2854004>>.

TU, Q.; DONG, L. An intelligent personalized fashion recommendation system. In: *2010 International Conference on Communications, Circuits and Systems, ICCAS 2010 - Proceedings*. [S.l.: s.n.], 2010. p. 479–485. ISBN 9781424482238.

VIOLA, P.; JONES, M. Robust real-time face detection. *International journal of computer vision*, v. 57, p. 137–154, 2004. ISSN 0920-5691. Disponível em: <<http://link.springer.com/article/10.1023/B:VISI.0000013087.49260.fb>>.

WANG, D. et al. Learning based neural similarity metrics for multimedia data mining. *Soft Computing*, Springer, v. 11, n. 4, p. 335–340, 2007.

WANG, W.; MUNTZ, R. STING : A Statistical Information Grid Approach to Spatial Data Mining. In: *the 23rd VLDB conference*. [S.l.: s.n.], 1997. p. 1–10. ISBN 1-55860-470-7.

YANG, Y.; GONG, Z.; U, L. H. Identifying points of interest by self-tuning clustering. *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information - SIGIR '11*, p. 883, 2011. Disponível em: <<http://portal.acm.org/citation.cfm?doid=2009916.2010034>>.

ZHANG, T.; RAMAKRISHNAN, R.; LIVNY, M. BIRCH: An Efficient Data Clustering Databases Method for Very Large. In: *ACM SIGMOD International Conference on Management of Data*. [S.l.: s.n.], 1996. v. 1, p. 103–114. ISBN 0897917944. ISSN 01635808.

ZHENG, Y.; CHEN, Y.; XIE, X. GeoLife2. 0: A location-based social networking service. *2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware*, n. 49, p. 4–5, 2009. Disponível em: <<http://www.computer.org/portal/web/csdl/doi/10.1109/MDM.2009.50>>.

ZHOU, P. et al. Iodetector: A generic service for indoor outdoor detection. In: ACM. *Proceedings of the 10th acm conference on embedded network sensor systems*. [S.l.], 2012. p. 113–126.

Apêndice A

Detalhes de Implementação

Neste apêndice serão apresentados detalhes de implementação da ferramenta PhoCA, apresentada no Capítulo 5 desta Tese. A primeira seção apresenta como estender o PhoCA com um novo algoritmo. A segunda seção apresenta em detalhes como se deu o processo de extração de metadados do Flickr.

A.1 Estendendo o PhoCA com um novo algoritmo

A ferramenta foi projetada para incorporar novos algoritmos facilmente. O Quadro A.1 apresenta as interfaces que o usuário deve conhecer para implementar e adaptar um novo algoritmo no PhoCA.

O novo algoritmo deverá implementar a interface *ClusteringIF*. Esta interface tem o método *clustering* que recebe uma coleção de *PhotoIF*, um *Properties* e um *PhotoClusterBuilder*. Este método retorna um *Set* de *PhotoClusterIF*.

Quadro A.1: Interfaces principais do PhoCA

```
public interface ClusteringIF {  
  
    public Set<PhotoClusterIF>  
        clustering (Collection<PhotoIF> photos , Properties p,  
                  PhotoClusterBuilder b);  
}  
  
public interface PhotoIF extends Serializable , Comparable<PhotoIF> {
```

```
public long getId();
double getOrientation();
public String getURL();
public Date getDateTime();
public Double getLongitude();
public Double getLatitude();
public Double getOrientation();
// other methods
}

public interface PhotoClusterIF extends Comparable<PhotoClusterIF>,
    Serializable {

    public void addElement(PhotoIF photo);

    public Collection<PhotoIF> getPhotos();
    // other methods
}

public interface PhotoClusterBuilder {
    public PhotoClusterIF builder(int id);
}
```

O PhotoIF representa cada foto que deverá ser agrupada. Ela é um bean com os métodos gets e sets para acessar os metadados das fotos, tais como: id, descrição, data que a fotografia foi capturada, latitude, longitude, orientação, a URL da fotografia e outros.

O parâmetro *p*, do tipo *Properties*, contém todos os parâmetros exigidos pelo algoritmo. Esses parâmetros devem ser descritos no arquivo de configuração dos algoritmos (*algorithms.xml*).

O Quadro A.2 apresenta as tags que são utilizadas para descrever um novo algoritmo. A tag `<algorithm>` define o nome do algoritmo e o nome da classe. Entretanto, a tag `<params>` define uma lista de parâmetros exigidos pelo algoritmo. Finalmente, a tag `<param>` é utilizada para representar um parâmetro. Essa tag exige que seja informado um nome e uma descrição para cada parâmetro.

Quadro A.2: Arquivo de configuração dos algoritmos (algorithms.xml)

```

<algorithms >
  <algorithm name="MyClustering"
class="com.yurilacerda.phoca.clustering.algorithms.MyClustering">
    <params>
      <param name="n" description=""></param>
    </params>
  </algorithm >
  <!-- other algorithms -->
</algorithms >

```

O Quadro A.3 apresenta o exemplo de um algoritmo de agrupamento chamado *MyClustering*. O algoritmo tem um parâmetro único n , que foi definido no Quadro A.2. As interfaces estão disponíveis em uma biblioteca Java chamada *phoca_base.jar*, e estão disponíveis em <http://www.yurilacerda.com/phoca>.

Quadro A.3: Implementação de um novo algoritmo de agrupamento

```

public class MyClustering implements ClusteringIF {

public Set<PhotoClusterIF>
    clustering (Collection<PhotoIF> photos , Properties p,
                PhotoClusterBuilder b) {
    // getting the parameter of clustering algorithm
    String n = p.getProperty("n");
    // The implementation of algorithm
    Set<PhotoClusterIF> clustersResult = createClusters ();
    // more implementation
    Return clustersResult;
}
// other methods and attributes
}

```

As propriedades de p são de acordo com os nomes dos parâmetros definidos no arquivo de configuração *algorithms.xml*. No exemplo do Quadro A.3, o p teve uma única propriedade com a chave n .

O método *clustering* deve retornar um conjunto de *PhotoClusterIF*. Ele é responsável

por armazenar um conjunto de *PhotoIF* classificadas. Um objeto *PhotoClusterIF* deverá ser instanciado pelo método builder de um *PhotoClusterBuilder* b. Esse método recebe um identificador do grupo (*cluster id*) como um parâmetro.

A.2 Extraindo os metadados

A recuperação de uma coleção pessoal de fotografias georreferenciadas é uma tarefa que demanda uma série de etapas. Assim, o Photo Collector é responsável por coletar um conjunto de fotografias a partir de um grande repositório disponível na Web. O PhoCA extrai dados do repositório de fotos do Flickr, um dos mais conhecidos sistemas de gerenciamento de fotografias digitais.

Além de dispor de uma grande base de dados com fotografias pessoais georreferenciadas, o Flickr também disponibiliza uma API (*Application Programming Interface*) para desenvolvedores de outras plataformas fazerem uso de várias das suas funcionalidades. Entre estas está a consulta de dados relativos às fotografias, implementada sob o método `flickr.photos.search`. Esse método possibilita a realização de consultas de fotografias em toda a sua base de dados pública. Desta forma, existem parâmetros que podem ser utilizados, tais como: a palavra-chave da consulta; o bounding box de uma área que se deseja recuperar as fotos; os dados de retorno; o formato de retorno do método (JSON, XML, etc); entre outros. O Quadro A.4 apresenta um trecho do retorno dessa consulta no formato JSON (*JavaScript Object Notation*).

Quadro A.4: Trecho de retorno de uma consulta no formato JSON.

```
{ "id": "6204695215", "owner": "38937706@N07", "secret": "4ae975089d", "server": "6176", "farm": 7, "title": "les 10 ans du Cube", "ispublic": 1, "isfriend": 0, "isfamily": 0, "latitude": 48.819983, "longitude": 2.25335, "accuracy": "16", "context": 0, "place\_id": "8xjQVDRUV7LYSw\_eVQ", "woeid": "55843833", "geo\_is\_family": 0, "geo\_is\_friend": 0, "geo\_is\_contact": 0, "geo\_is\_public": 1 }
```

Ainda no Quadro A.4 são ilustradas algumas informações relevantes para o processo de obtenção do arquivo com a fotografia. Como se pode perceber, alguns atributos que descrevem as fotografias são retornados em pares chave-valor, tais como: as informações de

latitude e longitude, bem como os atributos id, secret, server e farm, que serão utilizados para formar a consulta de recuperação do arquivo propriamente dito.

Com isso, para recuperar os arquivos com as fotografias, é necessário montar um endereço (URL), construído a partir dos dados que foram apresentados. Uma das formas de montar esse endereço é substituindo os valores entre colchetes, ilustrados na URL a seguir, pelo seu valor correspondente disponível no Quadro A.4. Um dos formatos pelo qual se pode extrair a imagem é a partir do apresentado no endereço (URL): *http://farm[farm-id].staticflickr.com/[server-id]/[id]_[secret].jpg*. Por exemplo, o endereço formado pelos dados apresentados no Quadro A.4 é: `<http://farm7.staticflickr.com/6176/6204695215_4ae975089d.jpg>`.

Também há outros formatos de URL que podem ser montados para recuperar a fotografia conforme descrito na documentação para desenvolvedores . O método `flickr.photos.search` possui a limitação de retornar apenas as 4.000 fotografias iniciais nas consultas. Ou seja, por mais que uma consulta devesse retornar 100.000 resultados, só serão recuperadas as primeiras 4.000. Assim, devem-se submeter consultas bem específicas para retornar apenas o desejado. Contudo, de forma a contornar tal limitação, o módulo Photo Collector possui um mecanismo que identifica quando uma determinada área possui mais de 4 mil fotografias e a separa recursivamente até encontrar áreas que contenham até no máximo 4.000 fotos. Essa abordagem é semelhante à apresentada em (BUCCAFURRI et al., 2011), a qual divide o espaço em espaços menores para coleta de informações.

Para um melhor aproveitamento dos recursos computacionais, esse módulo faz utilização de programação multi-thread com o intuito de tornar o processo de coleta mais rápido. Dessa forma, foi desenvolvido um componente de software responsável por fazer o controle `k threads`, para $k > 1$, que são executadas paralelamente para agilizar o processamento dos dados.

O resultado provido pelo método `flickr.photos.search` não contém todos os dados acerca das informações de georreferenciamento (metadados) das fotografias, tais como: orientação da fotografia, data e hora da captura, dentre outros. Entretanto, os metadados sobre a orientação da fotografia são fundamentais para a técnica proposta neste trabalho. Desta forma, deve-se extrair os metadados das fotografias coletadas.

Assim, é feita uma requisição através do método `flickr.photos.getExif`, disponibilizado

pelo Flickr, para recuperar os metadados de uma determinada fotografia, incluindo: data da captura da fotografia, orientação da foto, tamanho focal, exposição, abertura, e foco. Portanto, este módulo envia requisições dos dados de todas as fotografias que foram coletadas, utilizando o *photo_id* de cada uma delas. Em seguida, é realizada uma conversão dos dados retornados pelo servidor (no formato JSON) para posterior armazenamento na base de dados do PhoCA.

Apêndice B

Pontos de Interesses de Referência

As Tabelas B.1, B.2 e B.3 apresentam os Pontos de Interesse de Referência que foram extraídos na Segunda Etapa da Avaliação do Capítulo 6 relativos as cidades de Paris, Roma e São Francisco, respectivamente.

Tabela B.1: Pontos de Interesse de Paris

Categoria	Ponto de Interesse
Aquário	Aquarium de Paris
Aquário	Palais de la Porte Dorée - Aquarium Tropical
Atração	École Militaire
Atração	École Nationale Supérieure des Beaux-Arts
Atração	Église de la Madeleine
Atração	Église Saint-Eustache
Atração	Alpes
Atração	Ancienne Crèmerie
Atração	Arc de Triomphe
Atração	Arc de Triomphe du Carrousel
Atração	Assemblée nationale
Atração	Basilique du Sacré-Cœur
Atração	Brasserie Bofinger
Atração	Canauxrama
Atração	Carrières des Capucins
Atração	Carrousel de la Tour Eiffel
Atração	Carrousel de Montmartre
Atração	Cathédrale Notre-Dame de Paris

Atração	Cavae des Arènes de Lutèce
Atração	Cévennes
Atração	Centre Georges Pompidou
Atração	Centre Wallonie-Bruxelles
Atração	Chais et entrepôts de Bercy
Atração	Cimetière du Père-Lachaise
Atração	Cinéma - Spectacle Centre Wallonie-Bruxelles
Atração	Collège des Bernardins
Atração	Colonne de Juillet
Atração	Fontaine Cuvier
Atração	Grille du Coq
Atração	Hôtel de Lauzun
Atração	Hôtel de Ville
Atração	Hôtel Lebrun
Atração	Himalaya
Atração	Invalides
Atração	Japon - Chine
Atração	La Sorbonne
Atração	Le Centaure - César
Atração	Le Manoir de Paris
Atração	Le Petit Train
Atração	Les Grandes Serres du Jardin des Plantes
Atração	Lido
Atração	Maison de Madame Violet Trefusis
Atração	Maison de Nicolas Flamel
Atração	Moulin de la Galette
Atração	Moulin Radet
Atração	Musée Grévin
Atração	Opéra Garnier
Atração	Palais de Chaillot
Atração	Palais de Justice
Atração	Palais de Musique
Atração	Panthéon
Atração	Parc des Expositions de Paris - Porte de Versailles
Atração	Pavillon de Musique

Atração	Place d'Aligre
Atração	Place de La Republique
Atração	Place des Vosges
Atração	Place du Carrousel
Atração	Plus Vieil Arbre de Paris
Atração	Point zéro des Routes de France
Atração	Pont Neuf
Atração	Préalpes
Atração	Promenade Canal St. Martin
Atração	Promenade du Canal Saint-Martin
Atração	Proust (dernier lieu de vie et décès)
Atração	Provence
Atração	Salle du Livre d'Or
Atração	Serpent
Atração	Square Georges-Cain
Atração	Statue de Jean Moulin
Atração	Tourbière
Atração	Tour Eiffel
Atração	Tour Montparnasse
Atração	Tour Saint-Jacques
Atração	Trompe-oeil Bach
Atração	USA
Atração	Vedettes de Paris
Atração	Wallaby
Galeria	Addict
Galeria	ALB
Galeria	Anne Barrault
Galeria	Artedomus
Galeria	Art Jingle
Galeria	ArtLigue
Galeria	Artpark
Galeria	Ateliers d'art de France
Galeria	ATS
Galeria	Cloître des Billettes
Galeria	Dix9

Galeria	Espace Angel Orensanz
Galeria	Espace Topographie de l' Art
Galeria	Exit Art Contemporain
Galeria	Galerie 53
Galeria	Galerie Christophe Gaillard
Galeria	Galerie cinéma
Galeria	Galerie de l' angle
Galeria	Galerie de l' Ouest
Galeria	Galerie de multiples
Galeria	Galerie Fatiha Selam
Galeria	Galerie Isaak
Galeria	Galerie J. Kugel
Galeria	Galerie Karsten Greve
Galeria	Galerie les filles du calvaire
Galeria	galerie L.J.
Galeria	Galerie Menouar
Galeria	Galerie OREM
Galeria	Galerie Perrotin
Galeria	Galerie Placido
Galeria	Galerie SEE Studio
Galeria	Galerie Sylvie Lepage
Galeria	Galerie Täiss
Galeria	Galerie Thaddaeus Ropac
Galeria	Galerie thessa herold
Galeria	Histoires de parfums
Galeria	Hors-champs
Galeria	in)(between
Galeria	La galerie de l' instant
Galeria	L' Atelier
Galeria	L' Atelier
Galeria	Laure Roynette
Galeria	LebensonGaleria
Galeria	lecoeur-paris.com
Galeria	Mahatsara
Galeria	Slow

Galeria	Sobering
Galeria	WhiteProjects
Galeria	Zürcher
Monumento	Grande Arche
Museu	Anciens Elèves (Musée)
Museu	Archives Nationales
Museu	Art Ludique - Le Musée
Museu	Atelier Brancusi
Museu	Carrefour Numérique ²
Museu	Catacombes de Paris
Museu	Catacombes de Paris
Museu	Centre Culturel Suédois
Museu	Centre Culturel Suisse
Museu	Choco-Story - Le musée gourmand du chocolat
Museu	Cité des Enfants
Museu	Cité des Sciences et de l'Industrie
Museu	Collection des minéraux - Jussieu
Museu	Crypte Archéologique du Parvis Notre-Dame
Museu	Deyrolle
Museu	Espace Dali
Museu	Espace des Sciences Pierre-Gilles de Gennes
Museu	Espace Reine de Saba
Museu	Fondation Cartier
Museu	Fondation Henri Cartier-Bresson
Museu	Fondation Le Corbusier
Museu	Fondation Louis Vuitton
Museu	Fondation Pierre Bergé Yves Saint-Laurent
Museu	Futur Fondation Galeries Lafayette
Museu	Galerie de Minéralogie et de Géologie
Museu	galerie-musée Baccarat
Museu	Galerie Royale
Museu	Galleries de Paléontologie et d'Anatomie comparée
Museu	Grande Galerie de l'Évolution
Museu	Grand Palais
Museu	Halle Saint-Pierre

Museu	Hôtel des Monnaies
Museu	Hôtel de Sully
Museu	Institut des Lettres et Manuscrits
Museu	Institut du Monde Arabe
Museu	Institut Néerlandais
Museu	Jardin Tino Rossi - Musée de la Sculpture en Plein Air
Museu	Jeu de Paume
Museu	Les Égouts de Paris
Museu	Louvre
Museu	Maison de la Pêche et de la Nature
Museu	Maison de la RATP
Museu	Maison de Victor Hugo
Museu	Maison Européenne de la Photographie
Museu	Mémorial de la Shoah
Museu	Mémorial du maréchal Leclerc de Hauteclocque et de la Libération de Paris - Musée Jean Moulin
Museu	Maxim's Art Nouveau "Collection 1900"
Museu	Musée Adzak - Espace d'Art International
Museu	Musée Arménien de France
Museu	Musée Bible et Terre Sainte
Museu	Musée Boleslas Biegas - Musée Adam Mickiewicz
Museu	Musée Bouilhet-Christofle - Musée d'Orfèvrerie
Museu	Musée Bourdelle
Museu	Musée Clemenceau
Museu	Musée Cognacq-Jay
Museu	Musée Curie
Museu	Musée d'Anatomie Delmas-Orfila-Rouvière
Museu	Musée d'Anatomie Pathologique Dupuytren
Museu	Musée Dapper
Museu	Musée d'Art et d'Histoire du Judaïsme
Museu	Musée d'Art Moderne de la Ville de Paris
Museu	Musée de la Carte à Jouer et Galerie d'Histoire de la Ville
Museu	Musée de la chasse et de la nature

Museu	Musée de la Contrefaàsson
Museu	Musée de la Défense
Museu	Musée de la franc-maàssonnerie
Museu	Musée de la Légion d'Honneur et des Ordres de Chevalerie
Museu	Musée de la magie
Museu	Musée de la Poste
Museu	Musée de la Poupée
Museu	Musée de la Préfecture de Police
Museu	Musée de l'Èrotisme
Museu	Musée de la Serrure
Museu	Musée de l'Assistance Publique Hôpitaux de Paris
Museu	Musée de la Vie Romantique
Museu	Musée de l'Eventail
Museu	Musée de l'Orangerie
Museu	Musée de Minéralogie
Museu	Musée de Montmartre
Museu	Musée des années 30
Museu	Musée des Arts Forains
Museu	Musée des Lettres et Manuscrits
Museu	Musée d'histoire de la Médecine
Museu	Musée d'Orsay
Museu	Musée du Barreau
Museu	Musée du Luxembourg
Museu	Musée du Montparnasse
Museu	Musée du quai Branly
Museu	Musée du Service de Santé des Armées
Museu	Musée du Vin
Museu	Musée Edith Piaf
Museu	Musée en Herbe
Museu	Musée Guimet
Museu	Musée Gustave Moreau
Museu	Musée Henner
Museu	Musée-Jardin Paul Landowski
Museu	Musée-Librairie du Compagnonnage
Museu	Musée Maillol

Museu	Musée Moissan
Museu	Musée national des Arts et Métiers
Museu	Musée National du Moyen Âge
Museu	Musée national Ernest-Hébert
Museu	Musée national Eugène Delacroix
Museu	Musée Pasteur
Museu	Musée Picasso
Museu	Musée Pierre Marly - Lunettes et Lorgnettes
Museu	Musée Rodin
Museu	Musée Valentin Haüy
Museu	Musèè Carnavalet
Museu	Orangerie du Sénat
Museu	Palais de la Découverte
Museu	Palais de Tokyo
Museu	Pavillon de l' Arsenal
Museu	Petit Palais
Museu	Pinacothèque de Paris
Museu	Pinacothèque de Paris 2
Museu	Sainte-Chapelle
Museu	Salle des collections
Museu	Tour de Jean-sans-Peur
Área de Piquenique	Square Commerce
Área de Piquenique	Square de la rue d'Ortolan
Parque Temático	Jardin d' Acclimatation
Mirante	Ballon GENERALI de Paris
Mirante	Belvédère
Mirante	Butte aux Canons
Mirante	Cascade de Buttes-Chaumont
Mirante	Château de Bellevue
Mirante	Espace Observation
Mirante	Esplanade du Trocadéro
Mirante	Labyrinthe - Gloriette de Buffon
Mirante	La Terrasse des Galeries Lafayette
Mirante	Parvis du Sacré-Cœur
Mirante	Temple de la Sybille

Mirante	Temple Grec
Mirante	Vue sur la tour Eiffel
Zoológico	Ménagerie du Jardin des Plantes
Zoológico	Moutons du talus Garibaldi

Tabela B.2: Pontos de Interesse de Roma

Categoria	Ponto de Interesse
Agriturismo	Il Casale Bicocca
Atração	Arco di Costantino
Atração	Arco di Giano
Atração	Area archeologica del Teatro di Marcello
Atração	Area sacra dell' Argentina
Atração	Basilica dei Santi Giovanni e Paolo al Celio
Atração	Basilica dei Santi Trifone e Agostino
Atração	Basilica di San Clemente al Laterano
Atração	Basilica di San Giovanni in Laterano
Atração	Basilica di San Marco
Atração	Basilica di San Paolo fuori le mura
Atração	Basilica di San Pietro in Vincoli
Atração	Basilica di San Sebastiano fuori le mura
Atração	Basilica di Santa Cecilia in Trastevere
Atração	Basilica di Santa Maria degli Angeli e dei Martiri
Atração	Basilica di Santa Maria del Popolo
Atração	Basilica di Santa Maria in Trastevere
Atração	Basilica di Santa Maria Maggiore
Atração	Basilica di Santa Maria sopra Minerva
Atração	Basilica di Sant' Andrea della Valle
Atração	Basilica di Sant' Andrea delle Fratte
Atração	Basilica di Santa Prassede
Atração	Basilica di Santa Pudenziana al Viminale
Atração	Basilica di Santa Sabina all' Aventino
Atração	Basilica Sancti Petri
Atração	Battistero Lateranense
Atração	Bocca della Verità

Atração	bunker "Ferro di Cavallo"
Atração	Campo de' Fiori
Atração	Cannone di Mezzogiorno
Atração	Capo di Bove
Atração	Casa dei Cavalieri di Rodi
Atração	Case medievali e romane a Piazza di Monte Savello
Atração	Castel Sant'Angelo
Atração	Chiesa di San Giovanni dei Fiorentini
Atração	Chiesa di San Luigi dei Francesi
Atração	Chiesa di Sant' Agnese in Agone
Atração	Chiesa di Santa Maria della Pace
Atração	Chiesa di Santa Maria in Via
Atração	Chiesa di Sant' Andrea al Quirinale
Atração	Chiesa di Trinità dei Monti
Atração	Chiesa Nuova
Atração	Circo di Massenzio
Atração	Colosseo
Atração	Crypta Balbi
Atração	Exedra ex Pontificalibus Audientiis Pauli VI
Atração	Fontana dei Libri
Atração	Fontana dell' Acqua Acetosa
Atração	Fontana di Trevi
Atração	Fonte dell' Acqua Paola
Atração	Fonte dell' Acqua Sacra
Atração	Foro di Cesare
Atração	Foro di Traiano
Atração	Foro Olitorio
Atração	Forum Augustum
Atração	Forum Nervae
Atração	Forum Romanum
Atração	Galleria Alberto Sordi
Atração	Gruppo Scout Roma 21
Atração	Gruppo Scout Roma 21
Atração	Idrocronometro
Atração	Insula dell' Ara Coeli

Atração	Ludus Magnus
Atração	Mausoleo di Augusto
Atração	Meta sudans
Atração	Monumentoo al Finanziere
Atração	Ninfeo di Egeria
Atração	Oratorio dei Filippini
Atração	Oratorio del Gonfalone
Atração	Ospedale Santo Spirito in Sassia
Atração	Palatino
Atração	Palazzetto Venezia
Atração	Palazzo Apostolico
Atração	Palazzo Colonna
Atração	Palazzo Corsini alla Lungara
Atração	Palazzo dei Congressi
Atração	Palazzo della Cancelleria
Atração	Palazzo della Civiltà del Lavoro
Atração	Palazzo del Quirinale
Atração	Palazzo di Montecitorio
Atração	Palazzo Doria Pamphilj
Atração	Palazzo Farnese
Atração	Palazzo Lateranense
Atração	Palazzo Madama
Atração	Palazzo Pamphilji
Atração	Palazzo Rospigliosi-Pallavicini
Atração	Palazzo Senatorio
Atração	Palazzo Torlonia
Atração	Pantheon
Atração	Piazza del Campidoglio
Atração	Piazza di San Pietro
Atração	Piazza di Spagna
Atração	Piazza Mincio - Quartiere Coppedé
Atração	Piazza Navona
Atração	Platani Orientali del Seicento
Atração	Ponte Sant' Angelo
Atração	Porta Asinaria

Atração	Reliquia di San Giovanni Battista
Atração	Rovine Mausoleo Santa Costanza
Atração	Sacellum Sixtinum
Atração	San Pietro dal buco della serratura
Atração	Santa Maria Maddalena in Campo Marzio
Atração	Sant' Ignazio di Loyola a Campo Marzio
Atração	Sant' Ivo alla Sapienza
Atração	Santo Stefano Rotondo
Atração	Scala Santa
Atração	Scalinata della Trinità dei Monti
Atração	Sedia del Diavolo (Monumentoo a Elio Callistio)
Atração	Teatro di Marcello
Atração	Tempio della Pace
Atração	Tempio di Adriano
Atração	Templi dell' Area Sacra di Sant' Omobono
Atração	Terme di Caracalla
Atração	Torre delle Milizie
Atração	Villa Farnesina
Atração	Vittoriano
Galeria	F.A.D. room
Museu	Ars Iuncta Porticui Vaticanus
Museu	Aula Ottagona
Museu	Carcere Mamertino
Museu	Casa di Goethe
Museu	Casino Nobile
Museu	Chostro del Bramante
Museu	Esedra di Marco Aurelio
Museu	Explora
Museu	Farnesina ai Baullari
Museu	Galleria Nazionale d' Arte Moderna
Museu	Galleria Spada
Museu	Keats-Shelley Memorial House
Museu	Macro Testaccio
Museu	MAXXI
Museu	Mueso delle Forze Alleate Roma

Museu	Musei Vaticani
Museu	Museo Boncompagni Ludovisi per le arti decorative
Museu	Museo Centrale del Risorgimento
Museu	Museo Centrale del Risorgimento
Museu	Museo dei Frati Cappuccini e Cripta Ossario
Museu	Museo della Civiltà Romana
Museu	Museo della Mente
Museu	Museo dell'Ara Pacis
Museu	Museo dell'Astronomia e Planetario
Museu	Museo delle poste e telecomunicazioni
Museu	Museo di Roma in Trastevere
Museu	Museo Ebraico di Roma
Museu	Museo Etnografico "Luigi Pigorini"
Museu	Museo Hendrik Christian Andersen
Museu	Museo Internazionale del Presepio
Museu	Museo napoleonico di Roma
Museu	Museo Nazionale d'Arte Orientale "G. Tucci"
Museu	Museo nazionale dell'Alto Medioevo
Museu	Museo Nazionale delle Paste Alimentari
Museu	Museo Nazionale del Palazzo di Venezia
Museu	Museo Nazionale Etrusco di Villa Giulia
Museu	Museo Nazionale Romano
Museu	Museo Pietro Canonica
Museu	Museo Storico dei Bersaglieri
Museu	Museo Storico della Liberazione
Museu	Museo Storico dell'Arma dei Carabinieri
Museu	Palazzo Barberini
Museu	Palazzo Braschi
Museu	Palazzo Caffarelli
Museu	Palazzo Clementino
Museu	Palazzo dei Conservatori
Museu	Palazzo delle Esposizioni
Museu	Palazzo Massimo alle Terme
Museu	Palazzo Nuovo
Museu	Parco museo ferroviario Met.Ro Roma Porta San Paolo

Museu	Sacrario delle Bandiere delle Forze Armate
Museu	Sagrestia
Museu	Scuderie del Quirinale
Museu	Technotown
Museu	Terme di Diocleziano
Museu	Vigamus museo del videogioco
Mirante	Balconata del Pincio
Mirante	Belvedere Niccolà ² Scatoli
Zoológico	Bioparco

Tabela B.3: Pontos de Interesse de São Francisco

Categoria	Ponto de Interesse
Atração	Alhambra Theatre
Atração	Alma
Atração	Anglers Lodge
Atração	Aquarium of the Bay
Atração	Balclutha
Atração	Cable Car Turnaround
Atração	Cable Car Turnaround
Atração	Camera Obscura
Atração	C. A. Thayer
Atração	Chinese Pavillion
Atração	City Lights Bookstore
Atração	Coit Tower
Atração	Conservatory of Flowers
Atração	Eppleton Hall
Atração	Eureka
Atração	Ghirardelli Square
Atração	Golden Gate Bridge
Atração	Hercules
Atração	Japanese Tea Garden
Atração	McLaren Lodge
Atração	Murphy Windmill
Atração	North Windmill
Atração	Palace of Fine Arts

Atração	Pioneer Log Cabin
Atração	Point Bonita Lighthouse
Atração	Rose Garden
Atração	San Francisco Botanical Garden
Atração	San Francisco Carousel
Atração	Spreckels Temple of Music
Atração	Statue of Christopher Columbus
Atração	Statue of Mohandas Gandhi
Atração	Sutro Baths
Atração	United Nations Plaza
Área de Camping	Bicentennial
Área de Camping	Kirby Cove
Área de Camping	Rob Hill Campground
Carrocel	Carousel
Ponto de Referência	Bank of America Building
Museu	Asian Art Museu of San Francisco
Museu	Cable Car Barn and Museu
Museu	California Academy of Sciences
Museu	CCA Wattis Institute for Contemporary Arts
Museu	Contemporary Jewish Museu
Museu	De Young Museu
Museu	Farallones Marine Sanctuary Visitor Center
Museu	Haas-Lilienthal House
Museu	John Wickett Museu of Exotica
Museu	National Maritime Museu
Museu	Palace of the Legion of Honor
Museu	Randall Museu
Museu	Ripley's Believe it or Not
Museu	San Francisco Belle
Museu	San Francisco Fire Department Pioneer Museu
Museu	Santa Rosa
Museu	Santa Rosa
Museu	Santa Rosa
Museu	Santa Rosa
Museu	SFMOMA

Museu	SS Jeremiah O'Brien
Museu	USS Pampanito
Museu	Walt Disney Family Museu
Museu	Yerba Buena Center for the ArtsGaleria
Museu	Zeum; Children's Creativity Museu
Área de Piquenique	Picnic Area
Área de Piquenique	Picnic Area
Ruinias	Battery Spencer
Teatro	Yerba Buena Center for the Arts Theater
Mirante	Wilde Overlook Tower
Zoológico	Bison Paddock
Zoológico	San Francisco Zoológico