



**INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA**

**Área Departamental de Engenharia de Electrónica e Telecomunicações e de  
Computadores**

## **Segmentação Automática de Memórias Pessoais**

**DAVID MIGUEL CASACA COELHO**

(Licenciado em Engenharia Informática e de Computadores)

Dissertação para obtenção do Grau de Mestre  
em Engenharia Informática e de Computadores

Orientadores : Doutor João Moura Pires  
Mestre Nuno Miguel Soares Datia

Júri:

Presidente: Doutor Manuel Martins Barata

Vogais: Doutor Artur Jorge Ferreira  
Mestre Nuno Miguel Soares Datia

**Dezembro, 2014**





**INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA**

**Área Departamental de Engenharia de Electrónica e Telecomunicações e de  
Computadores**

## **Segmentação Automática de Memórias Pessoais**

**DAVID MIGUEL CASACA COELHO**

(Licenciado em Engenharia Informática e de Computadores)

Dissertação para obtenção do Grau de Mestre  
em Engenharia Informática e de Computadores

Orientadores : Doutor João Moura Pires  
Mestre Nuno Miguel Soares Datia

Júri:

Presidente: Doutor Manuel Martins Barata

Vogais: Doutor Artur Jorge Ferreira  
Mestre Nuno Miguel Soares Datia

**Dezembro, 2014**



# Resumo

Na era digital há um aumento na produção de fotografias para consumo pessoal. Este acréscimo no número de fotografias produzidas torna mais difícil recuperar e visualizar fotografias relacionadas com um determinado evento social. No entanto a metadata associada a cada fotografia é também mais rica, incluindo frequentemente informação sobre o local onde cada fotografia é tirada.

Para que seja facilitado o acesso às colecções de fotografias, é vantajoso apresentá-las divididas em conjuntos com contexto semelhantes. Dado o grande número de fotografias, é importante que essa divisão seja feita automaticamente.

Neste trabalho pretende-se construir um algoritmo de segmentação de conjuntos de fotografias, que assenta exclusivamente na informação presente na metadata, especificamente a informação temporal e geográfica. O objectivo é conseguir que cada segmento encontrado — um conjunto de fotografias — seja, na medida do possível, representativo de um evento social na vida do utilizador. O algoritmo desenvolvido é uma evolução de um outro, que efectua a segmentação apenas considerando a informação temporal, aqui estendido para incorporar também a informação geográfica. Nesse sentido, com ambas as segmentações é possível avaliar se a informação geográfica é redundante em relação à informação temporal.

Com os testes efectuados, quer de caracterização, quer com utilizadores, verificou-se que a informação geográfica aparenta ter vantagens quando utilizada como complemento à informação temporal. Os testes com utilizadores revelaram essa tendência, sendo necessário uma amostra maior para confirmar estatisticamente esta conclusão.

**Palavras-chave:** fotografia, informação temporal e geográfica, evento social, segmentação de conjuntos de fotografias



# Abstract

In the digital age the production of personal photographic images is increasing. As the number of photographic images increases, it becomes harder to retrieve and to visualize photos from a given social event. However, the metadata associated to each picture is richer, often including information about the place where it was taken.

To ease the access to personal photo collections, one can divide them into sets, each with its own context. Considering the large number of photos in each collection, it is important to automate such division.

In this work, a segmentation algorithm is constructed, which is intended to segment sets of photographic pictures, a process that is based exclusively in the information present in the metadata, specifically the temporal and geographic information. Each segment found — a set of photographic pictures — represents a social event in the life of the user. The algorithm developed in this work is an evolution of another algorithm, which performs the process of segmenting picture sets considering only the temporal information, here extended to include geographic information as well. This way, with both segmentations it becomes possible to evaluate if the geographic information is redundant when we consider that temporal information is always present.

The algorithms were tested using both theoretical tests and user tests. It was verified that the geographical information appears to have advantages when used as a complement to the temporal information. User tests revealed that tendency, despite the fact that a larger sample is needed to increase statistical power of the conclusion.

**Keywords:** photographic picture, temporal information, geographical information, social event, segmentation of personal photographic sets



# Glossário

DAM Desvio Absoluto Mediano ou *Median Absolute Deviation*

DP Desvio Padrão ou *Standard Deviation*

Exif *Exchangeable Image File Format* ou “Formato de Ficheiros de Imagem Permutável”

GPS *Global Positioning System* ou “Sistema de Posicionamento Global”

SIFT *Scale Invariant Feature Transform* ou “Transformada de Características Independente de Escala”

XML *Extensible Markup Language* ou “Linguagem Extensível de Formatação”



# Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objectivos . . . . .	3
1.2	Trabalhos Relacionados . . . . .	3
1.3	Estrutura do Documento . . . . .	6
<b>2</b>	<b>O Problema</b>	<b>7</b>
2.1	Informação Geográfica . . . . .	10
2.1.1	Tecnologias de Aquisição . . . . .	10
2.1.2	Precisão . . . . .	13
2.1.3	Fiabilidade . . . . .	13
2.2	Formalização das Operações . . . . .	14
2.2.1	Adjacência de Segmentos . . . . .	15
2.2.2	Operação de Fraccionamento de Segmento . . . . .	16
2.2.3	Operação de Junção de Segmentos . . . . .	17
2.2.4	Refinamento de Segmentações . . . . .	17
<b>3</b>	<b>Algoritmo de Segmentação</b>	<b>19</b>
3.1	Descrição da Solução . . . . .	19
3.1.1	Tratamento da Informação Geográfica . . . . .	20
3.2	Explicação do Algoritmo . . . . .	22

3.2.1	Parâmetro de Calibração do Algoritmo . . . . .	23
3.2.2	Ordem da Aplicação das Operações . . . . .	24
3.2.3	Descrição da Operação de Fraccionamento . . . . .	24
3.2.4	Descrição da Operação de Junção . . . . .	28
3.2.5	Restrições do Algoritmo de Segmentação no Tempo (AST) .	30
3.3	Detecção de Coordenadas Erradas . . . . .	30
3.3.1	Detecção com Base na Velocidade . . . . .	30
3.3.2	Detecção com Base no <i>Interquartile Range</i> “IQR” . . . . .	32
3.3.3	Detecção com Base no <i>Local Outlier Factor</i> “LOF” . . . . .	32
3.4	Implementação . . . . .	33
3.4.1	Distância entre Coordenadas . . . . .	33
3.4.2	Detecção de Coordenadas Erradas . . . . .	33
3.4.3	Estatísticas de Segmentação . . . . .	33
3.4.4	Operação de Fraccionamento . . . . .	34
3.4.5	Operação de Junção . . . . .	34
3.5	Descrição da Integração com o Trabalho Pré-existente . . . . .	36
<b>4</b>	<b>Avaliação Experimental</b>	<b>39</b>
4.1	Testes de Caracterização . . . . .	40
4.1.1	Algoritmo de Segmentação no Tempo . . . . .	43
4.1.2	Algoritmos de Detecção de Coordenadas Erradas . . . . .	45
4.1.3	Algoritmo de Segmentação no Tempo e Espaço . . . . .	47
4.1.4	Discussão . . . . .	60
4.2	Testes com Utilizadores . . . . .	60
4.2.1	Apresentação dos Testes com Utilizadores . . . . .	62
4.2.2	Caracterização dos Utilizadores . . . . .	68
4.2.3	Resultados dos Testes com Utilizadores . . . . .	74
<b>5</b>	<b>Conclusões</b>	<b>79</b>
5.1	Trabalho Futuro . . . . .	81
<b>A</b>	<b>Descrição Estatística das Coleções de Fotografias</b>	<b>i</b>

# Lista de Figuras

2.1	Exemplo de uma segmentação $S$ com dois segmentos <i>adjacentes</i> . . .	16
2.2	Um segmento é fraccionado em várias partes . . . . .	16
2.3	Junção dos segmentos $s_a$ e $s_b$ . . . . .	17
3.1	Etapas do Processo de Segmentação . . . . .	22
3.2	Um segmento é operado pela operação de fraccionamento . . . . .	26
3.3	Etapas da operação de fraccionamento . . . . .	27
3.4	Dois segmentos operados pela operação de junção . . . . .	29
3.5	Sequência de operações realizadas no ASE . . . . .	35
3.6	Diferenças no processo de activação dos dois algoritmos . . . . .	37
4.1	Exemplo de um refinamento de segmentação . . . . .	42
4.2	Exemplo de duas segmentações compatíveis . . . . .	42
4.3	Exemplo de duas segmentações compatíveis . . . . .	43
4.4	Nº de fotografias por segmento gerado pelo AST . . . . .	44
4.5	Comportamento dos algoritmos de detecção de coordenadas . . . . .	46
4.6	Detalhe da Figura 4.5, restringida a $[0.8; 1]$ . . . . .	46
4.7	Rácio do número de segmentos gerados pelos dois algoritmos . . . . .	48
4.8	Nº de fraccionamentos aceites versus rejeitados . . . . .	50
4.9	Quantidade de junções realizadas com (4.9(a)) e sem (4.9(b)) sucesso	52
4.10	Evolução da métrica $R_{Miss}$ em função do parâmetro $SpaceThreshold$	54

4.11	Evolução de <i>RMiss</i> , para os diversos algoritmos de detecção . . . .	54
4.12	Evolução da métrica <i>Rfa</i> em função do parâmetro <i>SpaceThreshold</i> . .	56
4.13	Evolução de <i>Rfa</i> , para os diversos algoritmos de detecção . . . . .	56
4.14	Evolução da métrica <i>Pk</i> em função do parâmetro <i>SpaceThreshold</i> . .	57
4.15	Evolução de <i>Pk</i> , para os diversos algoritmos de detecção . . . . .	58
4.16	Evolução da métrica <i>WindowDiff</i> em função de <i>SpaceThreshold</i> . . .	59
4.17	Evolução de <i>WindowDiff</i> , para os diversos algoritmos de detecção .	59
4.18	Sequência das páginas <i>web</i> durante os testes com utilizadores . . . .	63
4.19	Detalhes sobre a utilização da aplicação de testes . . . . .	63
4.20	Apresentação das segmentações criadas pelos dois algoritmos . . . .	64
4.21	As segmentações são apresentadas até à última colecção . . . . .	65
4.22	Histograma com Idades dos Participantes . . . . .	68
4.23	Distribuição do tipo de Máquina Fotográfica Tipicamente utilizada	69
4.24	Distribuição da utilização das tecnologias de <i>Geotagging</i> . . . . .	70
4.25	Distribuição da preferência sobre programas de gestão de fotografias	70
4.26	Distribuição da forma de armazenamento de fotografias . . . . .	71
4.27	Distribuição dos <i>web sites</i> utilizados para publicação de fotos <i>online</i>	72
4.28	Distribuição dos métodos de <i>backup</i> de fotografias . . . . .	73
4.29	Distribuição dos factores de organização favoritos . . . . .	73
4.30	Distribuição da Importância da Informação Geográfica . . . . .	74
4.31	Distribuição da Importância da Ordem Cronológica . . . . .	75
4.32	Preferências de algoritmo em função da colecção (ordem normal) .	76
4.33	Preferências de algoritmo em função da colecção (ordem inversa) .	76
4.34	Preferência de algoritmo em função da faixa etária do utilizador . .	77
4.35	Quantidade de participantes que usam georeferenciação . . . . .	77
4.36	Preferências de algoritmo em função da organização preferido . . .	78

# Lista de Tabelas

4.1	Quantidade de testes realizados, em função dos vários tipos de teste	41
4.2	Lista de colecções utilizada, ordenada por ordem directa . . . . .	61
4.3	Parametrização utilizada durante os testes com utilizadores . . . . .	62
A.1	Lista de colecções utilizadas durante os testes de caracterização . . .	ii
A.2	Lista de colecções utilizadas durante os testes de caracterização . . .	iii





# Introdução

O papel da fotografia nas nossas vidas é o de construir uma memória pessoal e social [39, 53]. O uso de redes sociais e a partilha de fotografias através delas demonstra essa necessidade, inerente ao ser humano [53]. Actualmente, a disseminação de dispositivos móveis com capacidades multimédia levou a um aumento do número de fotografias<sup>1</sup> [14, 32]. É esperado que os utilizadores destes dispositivos sejam os responsáveis por uma fatia significativa de todas as fotografias produzidas [4]. Cada uma destas pessoas tem de gerir cada vez mais fotografias, independentemente da forma ou dispositivo de armazenamento. Essa gestão, no entanto, torna-se difícil quer para produzir outros tipos de conteúdos, e.g. álbum, quer para efectuar uma pesquisa [51]. Em 2011, o número de fotografias inseridas diariamente no Facebook era cerca de 200 milhões [1].

Os sistemas de gestão de fotografias ainda não respondem convenientemente aos desejos dos utilizadores, em parte porque a interpretação de uma fotografia é uma tarefa difícil[42]. De entre os vários problemas, os designados “*sensory*” e “*semantic gap*” [44] ainda não estão resolvidos [21]. O “*semantic gap*” é um problema de resolução particularmente difícil se tivermos em conta que a solução ideal terá que compreender o conteúdo de uma colecção de fotografias e reconhecer tanto o contexto — o evento social retratado — como os participantes. Segundo Datta *et al.* [8], uma das formas de diminuir a distância entre o que os utilizadores pretendem e o que o computador lhes é capaz de fornecer, é melhorar

---

<sup>1</sup><http://www.flickr.com/cameras/>, consultado em Setembro de 2014

a descrição do contexto associada a cada fotografia. Esse contexto é geralmente descrito por quatro componentes, conhecidas na literatura por *4Ws*, nomeadamente: Quando, Onde, O Quê e Quem (*When, Where, What e Who*). A automatização do preenchimento destes descritores é muitas vezes inexecutável: muitos dos conceitos que os utilizadores pretendem não têm representação visual. Para além disso, a necessidade de traduzir características de baixo nível (como por exemplo a informação armazenada sob a forma de metadata) em conceitos de alto nível (como é o caso de um evento social do utilizador) revela-se uma tarefa difícil [25]. Desta forma é necessário que o preenchimento destes descritores – o acto de anotar uma fotografia – seja feito à mão, de preferência por alguém com conhecimento sobre o contexto retratado nas fotografias.

Segundo Kindberg *et al.* [26], as razões que motivam a captura de uma fotografia têm origens afectivas e funcionais, cada uma dividida em razões sociais e individuais. Assim, a anotação está intimamente ligada a um contexto pessoal, o qual muitas vezes só pode ser descrito pelo próprio fotógrafo e, potencialmente, pelos intervenientes das fotografias. Porém, o acto de anotar fotografias é um processo moroso que os utilizadores não estão dispostos a efectuar [20]. Uma possibilidade, é diminuir o trabalho de anotação, através de algoritmos que separem grupos de fotografias com um contexto comum.

Consideremos que o propósito de uma fotografia é capturar momentos da vida do utilizador e que este, muito provavelmente, tira várias fotografias em cada um desses momentos. É possível dividir esse conjunto de fotografias em vários segmentos, cada um deles com um contexto partilhado por todas as fotografias que o compõem. Assim, através da anotação em lote, diminui-se o custo associado à criação de anotações algo que aumenta o rácio custo/benefício. Pretende-se desta forma incentivar o utilizador a dar uma descrição textual, i.e uma interpretação das fotografias contidas num segmento, algo que para todos os efeitos é informação que permite lidar com o “*semantic gap*”. Por “*semantic gap*” entenda-se a falta de coincidência entre a informação que se pode extrair dos dados e a interpretação que um qualquer utilizador, num determinado contexto, pode fazer desses dados [45].

No presente trabalho pretende-se desenvolver um método de segmentar conjuntos de fotografias de forma automática em função da informação temporal e geográfica existentes na metadata de cada fotografia. Os diversos segmentos de fotografias – ou sequências de fotografias – são criados única e exclusivamente tendo por base a metadata temporal e espacial. Não se pretende realizar qualquer tipo de processamento de imagem sobre as fotografias. Subjacente a esta decisão está

a assunção de que um conjunto de fotografias que partilham o mesmo contexto – do ponto de vista do utilizador – estarão próximas umas das outras tanto no tempo como no espaço.

## 1.1 Objectivos

Um dos objectivos deste trabalho consiste em otimizar a segmentação resultante do algoritmo de Datia [7], e que recorre apenas a informação temporal para produzir uma segmentação. Essa melhoria passa por refinar a segmentação produzida, utilizando a informação geográfica presente nas fotografias. Neste contexto, o acto de refinar limita-se a alterar uma segmentação em função da informação geográfica também presente nas fotografias que compõem essa segmentação. Pretende-se refinar ao invés de ignorar por completo a segmentação original — segmentando as fotografias somente com base na informação geográfica. O motivo pelo qual se utiliza a informação geográfica para se refinar uma segmentação no tempo já pré-existente, deve-se ao facto do tempo ser a informação dominante para os utilizadores [24, 54, 55].

Outro objectivo do trabalho consiste em determinar se a informação geográfica, potencialmente disponível numa fotografia, é uma mais valia no contexto de um algoritmo de segmentação automática ou se, pelo contrário, é redundante tendo em conta a informação temporal também existente nessa mesma fotografia.

## 1.2 Trabalhos Relacionados

Na literatura, o acto de segmentar um conjunto de fotografias normalmente realiza-se com o intuito de suportar diferentes actividades, por exemplo visualização e anotações. Para criar esta segmentação regra geral utiliza-se a meta-informação presente nas fotografias, nomeadamente a informação temporal.

O trabalho de Plat, et al. [41] agrupa as fotografias em eventos recorrendo ao tempo de criação de fotografias. Quando esta informação aparenta não ser fidedigna<sup>2</sup>, o trabalho recorre à ordem das fotografias e à análise da imagem, determinando o histograma de cores, para efectuar o agrupamento. O algoritmo usa o tempo de criação das fotografias partindo do pressuposto que a existência de

---

<sup>2</sup> Por exemplo, quando a configuração da máquina fotográfica não está ajustada ao local ou, em casos em que as fotografias são digitalizadas.

quebras entre sequências de fotografias indicam o início de um novo segmento [15]. Nos segmentos que contêm mais de 23 fotografias, analisam-se as imagens das mesmas para determinar graus de semelhança entre fotografias consecutivas. Esta análise faz-se com o objectivo de facilitar a visualização das fotografias, evitando a construção de segmentos com um número de fotografias exagerado. De igual modo, Loui, *et al.* [30] recorre ao tempo de criação de fotografias. Depois, para cada evento identificado procuram-se sub-eventos com base na similaridade de imagem, ou seja usa-se um método de correlação entre o histograma de cores entre imagens consecutivas, depois destas serem divididas em blocos. Os autores assumem que a similaridade de imagem tem sempre um contributo a fazer para a segmentação e por isso usa esta técnica de forma sistemática. Por outro lado, Plat, *et al.* acreditam que a informação temporal é suficiente para obter uma boa segmentação e só quando aquela informação não é suficiente para obter aquilo que Plat, *et al.* crêem ser um bom resultado – todos os segmentos com 23 ou menos fotografias – é que se recorre à similaridade de imagem.

Naaman, *et al.* [37] agrupam fotografias e estabelecem uma hierarquia entre grupos com base na informação temporal e geográfica. Este trabalho tem a particularidade de associar nomes geográficos às hierarquias geradas, já que assume que a memória mais forte que um utilizador tem acerca de um evento é o local onde este ocorreu. O algoritmo inicial de segmentação é baseado no tempo. Posteriormente os segmentos são separados em segmentos menores tendo então como critério de separação o tempo e o espaço. Com a informação gerada por estes dois algoritmos é sugerido o nome geográfico a associar às hierarquias criadas. O nome geográfico é gerado com o auxílio de um *dataset* de regiões administrativas disponível comercialmente.

Datia [7] constata que existem separadores naturais (e.g. dia) e culturais (como é o caso da semana), importantes quando se pretende delimitar eventos sociais. Introduz assim o conceito de dia lógico o qual é limitado não pela meia-noite, mas antes pela actividade de cada utilizador. A principal consequência deste conceito é o facto de, para um conjunto de fotografias que esteja distribuído ao longo de vários dias, estas não são particionadas pela meia-noite que separa os vários dias. O conceito de “meia-noite” que separa os dias torna-se desta forma elástico de forma a acomodar eventos que se estendam para além desta hora. Evita-se assim a introdução de um fim de evento e de início de outro quando na realidade não existe. Por outro lado, para manter o conceito de “meia-noite” utiliza-se um parâmetro que regula a elasticidade máxima admitida.

Gozali, et al. [18] explora diversas formas de organizar a apresentação de fotografias aos utilizadores. Pressupõe que as fotografias são uma sequência representativa de um dado evento social na vida do utilizador. O software desenvolvido no âmbito desse trabalho segmenta automaticamente as fotografias em diversos grupos para as organizar. O software também permite que o utilizador corrija os segmentos gerados automaticamente pelo algoritmo descrito em [17], de maneira a que a segmentação corresponda ao seu gosto pessoal. Gozali, et al. conclui que os utilizadores atribuem mais importância à consistência entre segmentos do que à ordem cronológica das fotografias. Entre as características que devem ser tidas em conta para definir um evento, tem-se o momento e o local onde se tirou a fotografia e o objecto fotografado. O critério e a granularidade que rege a operação de segmentação são ambos subjectivos, sendo preferível um algoritmo de segmentação automática que não detecte o início ou fim de um evento, do que detectar erradamente um início ou fim de algo que, para o utilizador não existe.

Em [17], Gozali implementa um algoritmo de segmentação automática no qual usa a informação temporal, três campos da metadata das fotografias: *Focal Length*, *Aperture Diameter* e *LogLight* (mas designando-os conjuntamente como “informação de contexto”) e a própria imagem da fotografia. A análise da imagem é feita em três vertentes: o histograma das cores (à semelhança de Plat em [41]), auto-correlaciona a direcção de gradiente e realiza uma subtracção baseada no algoritmo SIFT<sup>3</sup> descrito em [31] entre imagens consecutivas.

O presente trabalho utiliza, à semelhança de [30, 41] o tempo como informação principal para efectuar a segmentação, utilizando adicionalmente a informação geográfica para refinar os segmentos.

A utilização da informação geográfica e temporal é feita de forma diferente de [37] pois este tinha como objectivo a apresentação de um conjunto num mapa, enquanto que o trabalho proposto pretende facilitar o processo de anotação de um conjunto de fotografias. Baseia-se no trabalho de [7], incorporando a informação geográfica no processo de refinamento dos segmentos.

Vários trabalhos tentam motivar os utilizadores a adicionar informação de contexto, uma anotação que descreva sucintamente o conteúdo de uma fotografia [16, 49] utilizando o conceito de *Crowdsourcing*. No entanto, este tipo de solução não se adequa à tarefa de criar anotações para um conjunto de fotografias pessoais já que este tipo de fotografia, devido à sua natureza pessoal, detém um

---

<sup>3</sup>Um algoritmo de processamento de imagem que transforma os dados de uma dada imagem em coordenadas relativas a características dessa mesma imagem, independentemente da escala da imagem em questão.

contexto que é apenas conhecido pelo utilizador e possivelmente pelos intervenientes na fotografia em questão.

## 1.3 Estrutura do Documento

Neste primeiro capítulo contextualizou-se o leitor com o problema abordado no presente trabalho, enunciando-se os principais objectivos a atingir. Descreveram-se em seguida trabalhos desenvolvidos anteriormente em áreas de investigação relevantes, os quais são brevemente descritos e correlacionados com o presente trabalho.

No capítulo 2 — “O Problema” — discute-se a origem, a aquisição, o tratamento e utilidade dos dados relacionados com informação geográfica. Nesse mesmo capítulo, formalizam-se também as operações que permitem realizar transformações sobre o conjunto de dados para atingir o objectivo do trabalho – fazer reflectir os efeitos que a presença da informação geográfica possibilita.

O capítulo 3 — “Algoritmo de Segmentação” — descreve a solução implementada: levanta questões relevantes acerca da informação geográfica e é explicado o funcionamento dos algoritmos implementados. Descreve-se aqui também a implementação do trabalho e respectiva integração com o trabalho já previamente desenvolvido em [7].

O capítulo 4 — “Avaliação Experimental” — apresenta dados relativos aos testes realizados. Os testes dividem-se em testes realizados no âmbito do desenvolvimento do trabalho, testes que caracterizam o comportamento dos algoritmos e em testes que demonstram as preferências dos utilizadores.

O capítulo final 5 — “Conclusões” — apresenta considerações sobre a utilidade da solução implementada e validade dos resultados obtidos. Expõem-se também os pontos do presente trabalho que poderão ser aprofundados em trabalhos futuros.

# 2

## O Problema

Um artefacto que se perdeu com a revolução digital foi o rolo fotográfico [6]. A alteração foi de tal maneira drástica que provocou a falência de um gigante da indústria fotográfica [35].

Esta revolução digital na área da fotografia trouxe duas inovações extremamente importantes: o custo de se tirar fotografia foi reduzido e o custo associado ao visionamento de uma fotografia foi eliminado.

Ao eliminar o custo associado ao visionamento de uma fotografia – que se reflectia no acto de *revelar* uma fotografia – promoveu-se a distribuição de fotografias. Onde antes era necessário pagar por cada instância revelada, agora simplesmente duplicam-se ficheiros fotográficos – uma tarefa sem custos associados – que são então enviados pela Internet, um acto também ele com um custo associado negligenciável.

O custo associado ao acto de se tirar uma fotografia foi reduzido devido ao facto do rolo fotográfico – um consumível relativamente dispendioso – ter sido substituído por um dispositivo digital de armazenamento de dados, cujo preço tem tendência a cair exponencialmente<sup>1</sup>.

Este dispositivo digital de armazenamento de dados, contém também uma característica inovadora quando comparada com a fotografia analógica tradicional: permite a eliminação de fotografias e conseqüentemente permite que os recursos

---

<sup>1</sup><http://techpageone.dell.com/technology/flash-hdd-synergy-story/>, consultado em Setembro de 2014

“consumidos” por uma fotografia apagada sejam reutilizados por novas fotografias. Estes factores contribuíram para a eliminação da formalidade e cuidados que estavam associados ao acto de se tirar uma fotografia.

A estes factos, acumula-se o facto da qualidade do rolo ser uma característica que foi transferida para o sensor digital – uma característica fundamental no que diz respeito à qualidade da fotografia. Isto implica que um dos factores que mais contribuíam para a custo da fotografia, foi transferido do consumível – pago em função do número de fotografias tiradas – para o hardware, pago uma única vez.

Estes factores, em conjunto, propiciaram uma explosão no número de fotografias tiradas [1], o que exacerbou a necessidade de criar métodos de salvar as fotografias de forma eficiente e, mais importante ainda, encontrar as fotografias desejadas num determinado momento.

Na era digital, uma fotografia pode ser guardada localmente (num dispositivo electrónico do utilizador) ou *online* e.g. em sites de redes sociais. A vantagem da opção local é a garantia de privacidade; só não transferindo a fotografia para um site de redes sociais se pode ter a certeza de que a fotografia nunca será divulgada, apesar das garantias dadas pelos diversos serviços *online* [43]. A questão de garantir a privacidade das pessoas retratadas nas fotografias é cada vez mais premente [33].

Por outro lado, duas das vantagens de se arquivar fotografias num serviço *online* são a disponibilidade imediata dessa fotografia a nível global e, incidentalmente, redundância de dados: se alguma vez se perder cópia local da fotografia, será sempre possível recuperar a mesma realizando o *download* a partir do serviço *online*.

De qualquer uma das duas formas, o acto de se transferir fotografias constitui um arquivamento das mesmas. Uma vez na posse das fotografias, tanto um programa de *software* a ser executado numa máquina local como um serviço *online* têm a possibilidade de processar as fotografias e pedir ao utilizador informações adicionais no sentido de anotar conjuntos de múltiplas fotografias representativos dos eventos sociais ocorridos na vida do utilizador.

Qualquer processamento automatizado que se faça sobre as fotografias, deverá assentar sobre a informação disponível; o que no âmbito deste trabalho implica a informação Exif<sup>2</sup> contida nas mesmas – a metadata de uma fotografia. Esta metadata é gerada pelo próprio dispositivo e dependendo do respectivo fabricante,

---

<sup>2</sup><http://www.digitalpreservation.gov/formats/fdd/fdd000146.shtml>, consultado em Dezembro de 2014

segue um ou mais standards tanto na semântica como no formato dos dados.

As fotografias representativas de eventos sociais ocorridos na vida do utilizador podem ser agrupadas com base em informações presentes nas próprias fotografias, nomeadamente a informação Exif. No âmbito deste trabalho as informações Exif consideradas relevantes são as informações temporal e geográfica, as quais indicam respectivamente quando e onde é que a fotografia foi tirada. É com base na informação Exif que indica o momento em que a fotografia foi tirada, que em [7] se agrupam – segmentam – as fotografias em conjuntos representativos dos eventos sociais do utilizador. Apesar do acto de segmentar as colecções de fotografias não resolver directamente o problema do “*semantic gap*”, é expectável que uma solução que separe as fotografias por contexto propicie a inserção de anotações, e assim atenuar o problema do “*semantic*” e “*sensory gap*”.

O presente trabalho pretende estender essa ideia mas agora para a informação geográfica: utilizar a informação Exif que indica o local onde a fotografia foi tirada para segmentar as fotografias. Especificamente, pretende-se utilizar esta informação para *alterar* a segmentação criada pelo trabalho desenvolvido em [7].

O motivo pelo qual se pretende utilizar a informação geográfica para refinar a segmentação criada com base na informação temporal, prende-se com a fiabilidade destes dois tipos de informação. Devido à forma como cada uma é adquirida, estes dois tipos de informação apresentam fiabilidades diferentes: a aquisição de informação temporal é, regra geral, feita a partir de um relógio interno ao dispositivo móvel; já por outro lado, a informação geográfica é sempre feita recorrendo a componentes externos ao dispositivo móvel os quais nem sempre estão disponíveis e quando estão, apresentam diferentes níveis de precisão. Por este motivo, é preferível recorrer primeiro à informação temporal e só depois, caso esta esteja disponível, recorrer à informação geográfica para refinar o resultado inicialmente obtido a partir da informação temporal. Nos casos em que a informação geográfica não está disponível a segmentação original – criada com base na informação temporal – permanece inalterada.

O foco deste trabalho incidirá sobre fotografias geradas por dispositivos com a capacidade de adicionar estas três *tags* às fotografias. Hoje em dia, de todos os dispositivos móveis com sistema GPS<sup>3</sup> incorporado e com capacidade de tirar fotografias com estas informações, o mais disseminado é o *smartphone*. Esta conclusão deve-se ao facto deste tipo de dispositivo ser o grande responsável pela

---

<sup>3</sup>[www.gps.gov/technical/ps/2008-SPS-performance-standard.pdf](http://www.gps.gov/technical/ps/2008-SPS-performance-standard.pdf), consultado em Dezembro de 2014

maioria das fotografias criadas hoje em dia [38]. Convém no entanto não esquecer que uma fotografia digital pode ser gerada a partir de uma série de fontes: um *smartphone*, uma máquina fotográfica digital, um *tablet* ou um digitalizador de fotografias são apenas exemplos de dispositivos capazes de gerar fotografias digitais. Note-se que no caso do último exemplo, se a metadata não for introduzida manualmente, a informação aí presente não será fiável: estará ausente ou errada. É necessário considerar também que uma fotografia pode ser alterada antes de ser processada pelo algoritmo desenvolvido no presente trabalho, por exemplo com o auxílio de um programa informático. Por este motivo nunca se pode ter a certeza de que a metadata de uma fotografia está correcta ou inalterada desde a sua criação. Um algoritmo que processe metadata errada produzirá uma segmentação incoerente com os eventos retratados na colecção de fotografias, o que será contra-procedente para qualquer tentativa de registar anotações.

## 2.1 Informação Geográfica

O standard Exif 2.2 define 4 *tags* relacionadas com o tempo e mais de 30 *tags* relacionadas com o espaço, em particular com o sistema GPS [11]. Destas, apenas as *tags* que indicam quando e onde (dois dos 4W's — *when* e *where*) é que a fotografia foi tirada são consideradas relevantes no âmbito deste trabalho, e assim a leitura da informação geográfica e temporal relevante consiste na leitura de 3 *tags*: a data e hora de criação, a latitude e a longitude que constam na fotografia.

### 2.1.1 Tecnologias de Aquisição

Hoje em dia no âmbito dos *smartphones*, predominam três tecnologias de georeferênciação: através de sistema GPS, através de pontos de acesso *wireless* 802.11, ou pelas estações base das redes de comunicações móveis [52]. Cada um destes sistemas tem características boas e más e em todos os casos a precisão varia em função do local onde são utilizados.

**Sistema GPS.** Apesar deste sistema de ser o mais preciso, tem um eficácia reduzida em ambientes urbanos. Esta limitação deve-se o facto do sistema necessitar de uma linha de visibilidade directa entre o dispositivo móvel e um conjuntos de satélites de GPS. São necessários um mínimo de 4 satélites GPS todos eles directamente visíveis em simultâneo pelo dispositivo móvel de forma a determinar uma

coordenada com precisão [2, 5]. O sistema pode funcionar com apenas 3 satélites e manter a sua precisão se a altitude do local for de alguma forma conhecida pelo sistema GPS [22]. Como em ambientes urbanos as construções são realizadas em altura, as zonas de sombra – onde não há uma linha de visibilidade directa com satélites GPS – são comuns [36].

Adicionalmente, dependendo do formato da área de céu desimpedida, a precisão do sistema pode piorar significativamente devido ao facto dos satélites visíveis estarem demasiado juntos [28]. Estas interferências no sinal de GPS, poderão ou não impedir o uso do sistema, mas garantidamente reduzem a precisão da coordenada produzida. Mesmo em condições óptimas, a velocidade de transmissão de dados do satélite é de apenas 50bit/s, pelo que o tempo que um dispositivo móvel demora a determinar a sua localização varia entre alguns segundos e 12 minutos e meio [19, 27].

**Sistema aGPS.** Para minimizar a demora a determinar a sua localização, foi desenvolvido o sistema aGPS (*assisted GPS*), o qual delega parte do trabalho de georeferênciação num servidor externo, acessível via rede de comunicações móveis. Este sistema tem vários modos de funcionamento, cada um deles progressivamente mais dependente do servidor aGPS externo [29].

Todos os modos de funcionamento delegam parte do trabalho relacionado com o cálculo da localização geográfica para o servidor aGPS externo. Em função das capacidades do dispositivo móvel, o cliente aGPS pode limitar-se a usar o servidor aGPS para obter mais rapidamente parte dos dados enviados pelo satélite GPS, recebendo-os assim através da rede móvel e todo o trabalho restante é realizado pelo próprio dispositivo móvel. Em dispositivos móveis com menos capacidade de computação, o cliente aGPS pode limitar-se a enviar ao servidor aGPS uma quantidade mínima de informação recebida do satélite GPS e a receber do servidor a localização geográfica do dispositivo móvel, sendo o cálculo realizado pelo servidor aGPS [29].

O sistema de aGPS, apesar de ser mais rápido tem uma precisão mais baixa que o GPS standard e utiliza recursos da operadora móvel – o servidor de aGPS – algo que poderá eventualmente implicar custos financeiros para o utilizador [29].

**Georeferênciação via Redes Wireless 802.11.** Tendo em conta o requisito da visibilidade directa entre o dispositivo móvel e o satélite GPS, torna-se impossível utilizar o sistema GPS em espaços fechados. Neste tipo de ambientes temos um

outro sistema de posicionamento, via redes *wireless* 802.11 o qual tira partido do facto do ponto de acesso *wireless* estar extremamente disseminado em espaços fechados. Este dispositivo transmite de forma omni-direccional sinais que podem ser captados pela maioria dos *smartphones*.

Esta técnica é complementar ao sistema GPS, porque tipicamente é apenas possível em ambiente urbano onde se encontra uma maior densidade de pontos de acesso *wireless*. A maioria das implementações deste sistema de georeferênciação usa os pontos de acesso *wireless* detectados num dado local para criar uma “impressão digital” desse mesmo local. Posteriormente recorre a uma base de dados remota para relacionar essa “impressão digital” com a respectiva coordenada geográfica.

Este sistema quando comparado com o sistema GPS, tem a vantagem de determinar a localização em menos de um segundo com uma precisão apenas um pouco inferior ao sistema GPS [52].

**Georeferênciação via Estações Base de Redes de Comunicações Móveis.** A opção de realizar o posicionamento através das estações base de redes de comunicações móveis existe sempre que o dispositivo móvel esteja equipado com esta tecnologia. Este método assenta sobre a capacidade que um equipamento de comunicações móveis tem de determinar a distância que o separa de uma determinada estação base, medindo a potência do sinal que é emitido dessa torre. Conhecendo esta distância, sabendo que localização da torre é fixa e que o sinal é emitido de forma omni-direccional, é possível determinar com uma precisão razoável que o dispositivo móvel se encontra algures num certo perímetro.

A precisão deste método estará sempre relacionada com a dimensão da área coberta pela torre; quanto maior for a área coberta, maior será o erro. Se o dispositivo móvel estiver dentro do alcance de múltiplas estações base de redes de comunicações móveis, então torna-se possível determinar a localização do dispositivo móvel com maior precisão. Para esse efeito recorre-se a métodos alternativos como é o caso da triangulação. Apesar desta melhoria, o erro da coordenada obtida ronda em média a casa das centenas de metros – entre uma a várias centenas dependendo do número de estações base de redes de comunicações móveis que cobrem o local. Apesar de ser a tecnologia com a pior precisão, é também uma tecnologia que tem uma muito boa cobertura, e em certas situações pode ser a única que está disponível [52].

### 2.1.2 Precisão

Em função do dispositivo, a informação recolhida pode ser proveniente de uma de várias tecnologias; i) GPS, ii) *Wireless* ou iii) redes de comunicações móveis.

Alguns *smartphones* vêm equipados com todas estas tecnologias, utilizando assim as que têm dados disponíveis na altura em que essa informação é necessária. Esta opção, apesar de aumentar a disponibilidade da informação, tem o inconveniente de fazer com que a Informação Geográfica gerada pelo *smartphone* tenha uma precisão variável, em função da tecnologia que foi utilizada num determinado momento.

A tecnologia de georeferênciação mais precisa num dado instante depende das circunstâncias. Factores como o local, hora ou condições atmosféricas condicionam a utilidade das tecnologias disponíveis num dado dispositivo móvel. Assim, a tecnologia seleccionada estará sempre dependente das circunstâncias existentes no instante em que a Informação Geográfica é adquirida. Um problema que resulta desta variedade de tecnologias é o facto de não haver qualquer registo sobre que tipo de fonte de dados produziu uma dada coordenada geográfica. Sem informações sobre a fonte de dados, torna-se impossível determinar qual a precisão ou fiabilidade da Informação Geográfica em causa.

Por norma, o sistema GPS funciona apenas ao ar livre, atingindo a sua máxima precisão apenas em locais com poucos obstáculos onde as interferências de sinal são minimizadas. Os pontos de acesso *Wireless* oferecem um nível de precisão intermédio mas tipicamente existem apenas em meios urbanos e têm um alcance inferior a 1 km. Finalmente o sistema de georeferênciação via redes de comunicações móveis é o sistema mais disseminado, já que as operadoras móveis têm um incentivo comercial para aumentar a cobertura das redes de comunicações móveis, mas infelizmente é também o menos preciso [50].

### 2.1.3 Fiabilidade

Como vimos anteriormente, para um qualquer conjunto de meta-dados, o nível de precisão e fiabilidade variará em função do dispositivo que lhe deu origem, bem como as circunstâncias em que os meta-dados foram gerados. No caso particular da Informação Geográfica, esta pode até nem estar presente, nos casos em que as fotografias são criadas em certos sistemas como por exemplo um digitalizador de fotografias.

Mesmo estando presente, esta informação pode estar errada, dependendo como já vimos da precisão, ou seja da qualidade da informação. A qualidade da Informação Geográfica depende de uma série de factores que interferem com o funcionamento do dispositivo de georeferênciação, tais como o tipo de dispositivo ou mesmo a qualidade ou sensibilidade do mesmo.

Outro factor de relevo são as possíveis interferências ou até mesmo interrupção no sinal de georeferênciação no momento em que se tirou a fotografia [9, 46]. Dependendo das circunstâncias — do tipo de hardware, do momento, do local, etc. — a qualidade da coordenada adquirida pelo dispositivo de georeferênciação pode implicar um erro que varia entre alguns centímetros e alguns quilómetros mesmo quando se considera apenas a tecnologia mais precisa, o GPS [47, 50].

## 2.2 Formalização das Operações

### A Sequência de Fotografias

Considere-se um rolo de fotografias representado pelo conjunto  $P = \{p_1, p_2, \dots, p_N\}$ , ordenadas de forma não decrescente pelo tempo de criação. É possível derivar de  $P$  duas sequências,  $IT$  e  $IS$ , tal que

$$\begin{aligned} IT &= \{It_1, It_2, \dots, It_n\} \\ IS &= \{Is_1, Is_2, \dots, Is_n\} \end{aligned} \quad (2.2)$$

Estas sequências contêm a informação temporal ( $IT$ ) e geográfica ( $IS$ ) de  $P$ , e por isso no âmbito deste trabalho cada  $P_i$  é descrito pelos respectivos  $It_i$  e  $Is_i$ . Garante-se que  $|IT| = |P|$  e  $|IS| = |P|$ . Como a informação geográfica pode não estar presente, é possível que  $IS_i = \emptyset$ .

Assuma-se que cada fotografia do conjunto possui duas etiquetas informacionais – contidas em  $IT$  e  $IS$  – que permitem que duas fotografias sejam comparadas uma com a outra. Note-se que, tendo em conta que é possível que  $IS_i = \emptyset$ , é necessário garantir que cada uma das fotografias que se pretende comparar detém uma etiqueta contida em  $IS$ . Assume-se também que estas duas etiquetas são independentes entre si e que os elementos em  $P$  estão ordenados pela informação contida na etiqueta  $IT$  e não podem ser reordenados.

### A Noção de Segmento

Considere-se a noção de segmento conforme definido em [7], na qual um segmento, denotado por  $s = [It^-, It^+]$ , é uma sub-sequência de  $IT$  contínua e não vazia, delimitada por  $It^-$  e  $It^+$ . Neste caso,  $It^-$  é o limite inferior do segmento  $s$  e  $It^+$  é o limite superior do mesmo segmento, onde  $It^- \leq It^+$ .

### Um Elemento no Âmbito de um Segmento

Um elemento  $It_i \in IT$  fará parte de  $s$  quando:

$$\forall It_i \in IT : It^- \leq It_i \leq It^+ \Leftrightarrow It_i \in s \quad (2.3)$$

Convém notar que um segmento pode ser composto por um único elemento, quando  $It^- = It^+$ , e pode ser igual a  $IT$ , quando  $It^- = It \wedge It^+ = It_{|IT|}$ .

### O Conceito de Segmentação

A um conjunto ordenado e não vazio de segmentos derivados do mesmo  $IT$  dá-se o nome de segmentação. Uma segmentação  $S$  é definida como:

$$S = \{s_a = [It_a^-, It_a^+] : It_a^-, It_a^+ \in IT\} \quad (2.4)$$

onde  $a$  indica a posição do segmento  $s_a$  na segmentação  $S$ . Uma segmentação  $S$  é criada a partir das etiquetas contidas em  $IT$ . Para qualquer segmentação, verificam-se as seguintes propriedades:

- i) o primeiro elemento (limite inferior) do primeiro segmento é sempre o primeiro elemento de  $IT$ :  $It_1^- = It_1$ ;
- ii) o último elemento (limite superior) do último segmento é sempre o último elemento de  $IT$ :  $It_{|S|}^+ = It_{|IT|}$ ;
- iii) não existe qualquer elemento em  $IT$  que separe o último elemento de um dado segmento do primeiro elemento do segmento seguinte:  $(It_a^+)^> = It_{a+1}^-$ , onde  $a \in [1 \dots |S| - 1]$ ;

#### 2.2.1 Adjacência de Segmentos

Considerem-se dois segmentos  $s_a$  e  $s_b$ , sendo  $s_a$  minorado pela etiqueta  $It_a^-$  e majorado pela fotografia  $It_a^+$  e  $s_b$  minorado pela fotografia  $It_b^-$  e majorado pela fotografia  $It_b^+$ :  $s_a = [It_a^-, It_a^+] \wedge s_b = [It_b^-, It_b^+]$  diz-se que  $s_a$  e  $s_b$  são segmentos *adjacentes* quando

**Definição 2.2.1 (Segmentos Adjacentes em S).** Um segmento  $s_a \in S$  é adjacente ao segmento  $s_b \in S$ , quando não há qualquer outro segmento entre  $s_a$  e  $s_b$  tal que:

$$\forall a, b \in \mathbb{N}, \forall s_a, s_b \in S : s_{a+1} = s_b \quad (2.5)$$

Considerem-se dois segmentos adjacentes  $s_a$  e  $s_b$ , contidos na segmentação  $S$ :

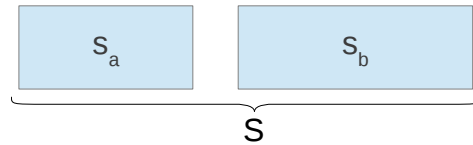


Figura 2.1: Neste exemplo a segmentação  $S$  só contém dois segmentos  $s_a$  e  $s_b$  que consequentemente são adjacentes.

As operações que se podem realizar sobre os mesmos, são a operação de junção – que opera sobre dois segmentos adjacentes e produz um único segmento – e a operação de fraccionamento, que é realizada sobre um único segmento e produz um ou mais segmentos. Ambas estas operações são descritas em seguida:

### 2.2.2 Operação de Fraccionamento de Segmento

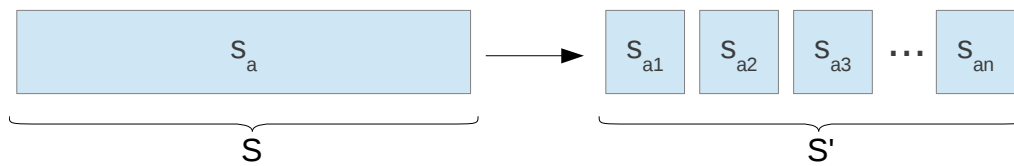


Figura 2.2: Um segmento é fraccionado em várias partes.

Seja  $s_a$  um segmento na segmentação  $S$ . A aplicação da operação de fraccionamento a este segmento produz uma série de segmentos  $\{s_{a1}, s_{a2}, \dots, s_{an}\}$ , tal que:

$$\begin{cases} It_{a1}^- = It_a^- \wedge It_{an}^+ = It_a^+, \\ It_{ai}^+ = It_{ai+1}^-, \end{cases} \quad i \in [0, n - 1]. \quad (2.6)$$

### 2.2.3 Operação de Junção de Segmentos

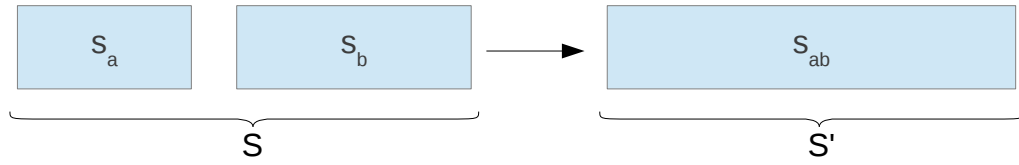


Figura 2.3: Junção dos segmentos  $s_a$  e  $s_b$ .

Sejam  $s_a$  e  $s_b$  dois segmentos adjacentes em  $S$ . A aplicação da operação de junção desses dois segmentos produz um  $s_{ab}$ , tal que:

$$It_{ab}^- = It_a^- \wedge It_{ab}^+ = It_b^+ \quad (2.7)$$

A Figura 2.3 ilustra a aplicação desta operação. A aplicação sucessiva desta operação sobre  $S$ , pode produzir uma segmentação  $S'$  com um único segmento.

### 2.2.4 Refinamento de Segmentações

Refinar a segmentação  $S$  é um processo que resulta numa nova segmentação  $S'$  a qual tem em consideração a informação geográfica contida nas etiquetas de  $IS$ . O processo que cria a nova segmentação  $S'$  poderá fraccionar um segmento, juntar segmentos ou copiar um segmento de  $S$  para  $S'$ . Ao fraccionar um segmento, obtemos um refinamento da segmentação  $S$  original:

$$\forall s' \in S', \exists! s \in S, S' \neq S : s' = s \vee s' \sqsubseteq s \quad (2.8)$$

O processo inverso, juntar segmentos, inverte também a relação entre as segmentações  $S$  e  $S'$ :

$$\forall s' \in S', \exists! s \in S, S' \neq S : s' = s \vee s' \supseteq s \quad (2.9)$$

Tanto o conjunto de fotografias, como a relação de ordem entre essas fotografias, são idênticos entre  $S$  e  $S'$ .



# 3

## Algoritmo de Segmentação

### 3.1 Descrição da Solução

O problema a resolver consiste em realizar uma segmentação efectuada a dois tempos: primeiro é feita com a informação temporal, sendo posteriormente refinada com a informação de localização espacial. A primeira parte do trabalho de segmentação é feita com base na informação temporal e tem em consideração noções importantes para o utilizador, como é o caso do conceito de evento social ou o facto de as regularidades temporais nem sempre coincidirem com os limites desses eventos sociais. Em particular, é tratado o ciclo diário, ajustando os limites superior e inferior de um dia para que este seja compatível com as actividades desenvolvidas.

Pretende-se que a segunda parte do trabalho de segmentação, com base nos dados geográficos, não destrua a segmentação resultante do primeiro processamento. Assim, depois de realizada a segmentação no espaço, efectua-se uma detecção para garantir que as noções consideradas no primeiro processamento – no âmbito do tempo – não foram ignoradas no segundo processamento – no âmbito do espaço.

Como ponto de partida, utiliza-se o algoritmo descrito em [7] para encontrar uma segmentação  $S_t$ . Uma segmentação contém um ou mais segmentos e cada um desses segmentos contém uma ou mais fotografias. A composição de cada um

destes segmentos é determinada apenas em função da informação temporal contida naquelas fotografias, e por este motivo, a segmentação que daqui resulta designa-se uma segmentação no tempo.

Durante o processamento da informação geográfica, para cada segmento presente na segmentação no tempo, avalia-se a informação geográfica contida nas respectivas fotografias. Esta avaliação procura determinar distâncias entre as localizações presentes em cada segmento para, com base da sua proximidade ou afastamento, aplicar as operações de junção e fraccionamento, respectivamente. Este processo que vai resultar na junção e/ou fraccionamento dos segmentos existentes na segmentação no tempo  $S_t$ , produz uma segmentação no tempo e espaço  $S_{te}$ . A decisão de executar uma dada operação sobre um dado segmento é realizada com base em estatísticas relativas ao nível do segmento, e é ponderada por um parâmetro que regula o peso que essas estatísticas têm no algoritmo.

Por fim, garante-se que o processamento da informação geográfica não desfaz por completo o trabalho realizado inicialmente, ou seja, garante que as restrições tidas em conta pelo Algoritmo de Segmentação no Tempo (AST) não são completamente ignoradas pelo Algoritmo de Segmentação no Tempo e Espaço (ASE).

### 3.1.1 Tratamento da Informação Geográfica

Como já foi referido anteriormente, a informação espacial pode conter variações na sua precisão que dificultam a validação deste tipo de informação. Um método de determinar se a coordenada de uma dada fotografia é fidedigna ou não realiza-se calculando a distância e tempo que separa essa fotografia das  $m$  fotografias que a precedem e das  $n$  fotografias que a sucedem. A partir desses dados poderá ser possível obter-se uma velocidade que, caso a coordenada tenha um erro considerável, será obviamente anómala. Este método destina-se a detectar situações em que, segundo os dados, o utilizador deslocou-se uma distância normalmente percorrida a pé, a velocidades que um peão não pode alcançar. Foram testados outros dois métodos que serão posteriormente descritos – detecção com base no *Interquartile Range* “IQR” e com base no *Local Outlier Factor* “LOF” – mas o método baseado no conceito de velocidade foi o que apresentou melhores resultados.

A criação de uma segmentação de fotografias de forma automática pressupõe que se realizem cálculos sobre a informação disponível – neste caso, a informação geográfica presente nas fotografias. Esta informação, ao contrário da informação sobre o tempo, é

- i) potencialmente incorrecta e
- ii) esparsa.

O processamento da informação geográfica tem estes problemas em atenção e tenta detectar e tratar estas anomalias na informação, na medida do possível.

Um método de detecção de coordenadas incorrectas consiste em medir a velocidade. Pode-se usar a informação de tempo e espaço para determinar a velocidade entre as várias fotografias consecutivas de forma a, como proposto anteriormente, detectarem-se situações em que o utilizador se desloca uma distância normalmente percorrida a pé, a velocidades irrealistas para um peão.

Considere-se que num conjunto de fotografias consecutivas, temporalmente próximas entre si, há uma fotografia  $f$  que se encontra geograficamente afastada das restantes. Se este afastamento for suficientemente grande a ponto de indicar uma velocidade irrealista, então justifica-se considerar a informação geográfica dessa fotografia  $f$  como incorrecta. Pode-se até considerar que se a velocidade necessária para tirar a fotografia  $f$  for muito maior que uma velocidade média – ou até mesmo impossível de alcançar – pode ser preferível ignorar a informação geográfica dessa fotografia ou corrigi-la recorrendo à informação geográfica das fotografias adjacentes e a uma série de técnicas estatísticas [48].

Note-se que para este método funcionar não é necessário garantir que a informação temporal de cada uma das fotografias está correcta. Basta apenas garantir que o erro, se existir, é uniforme para todas as fotografias que constituem a segmentação. Na prática, basta garantir que o relógio do dispositivo foi alterado antes da primeira ou depois da última fotografia da segmentação, mas nunca entre estas duas fotografias.

O outro problema que é necessário resolver é a ausência de informação. O que fazer quando a informação que se pretende analisar não está presente? Poderia ser possível recorrer à média ou a outra técnica estatística nas situações em que a omissão é pontual. No entanto, nos conjuntos de dados utilizados durante os testes, constata-se que quando há ausência de informação geográfica, esta ausência nunca é pontual; ocorre sempre ao longo de uma série de várias fotografias.

Para os conjuntos de dados testados no âmbito deste trabalho, na maior parte dos casos a informação geográfica está *quase* sempre presente *ou* ausente para uma dada sequência de fotografias. Isto é, para uma dada sequência de fotografias todas com informação geográfica é pouco provável que haja uma sem informação geográfica e vice-versa. Analisando o conteúdo de vários conjuntos de fotografias

de diferentes fotografos conclui-se que, na maior parte dos casos, as fotografias sem informação geográfica representam na íntegra um evento da vida do utilizador, durante o qual por uma qualquer razão a informação geográfica não estava disponível. Neste contexto torna-se impossível determinar automaticamente a localização geográfica do evento. Assim, o Algoritmo de Segmentação no Tempo e Espaço (ASE) limita-se a não utilizar estas fotografias para alterar o segmento produzido pelo Algoritmo de Segmentação no Tempo (AST). No limite, caso nenhuma fotografia num dado segmento tenha informação geográfica, o Algoritmo de Segmentação no Tempo e Espaço (ASE) não realiza qualquer alteração sobre esse segmento.

## 3.2 Explicação do Algoritmo

Tanto o ASE como o contexto no qual este algoritmo se insere estão ilustrados na Figura 3.1. A parte da figura que está ordenada com números diz respeito ao processo de integração: em que ponto de todo o processo de segmentação é que o ASE é inserido. A sequência de itens ordenados com letras descrevem as várias etapas do próprio ASE.

O processo de segmentação começa por 1) obter a metadata das fotografias que o utilizador pretende processar, 2) realiza os processos de segmentação no tempo e 3) no espaço 4) e finalmente realiza as tarefas dependentes do processamento no tempo e espaço, conforme se indica na Figura 3.1. Partindo da implementação desenvolvida anteriormente [7], que implementa as etapas 1), 2) e 4) da figura, pretende-se adicionar a terceira etapa: refinar a segmentação no tempo com a informação geográfica.

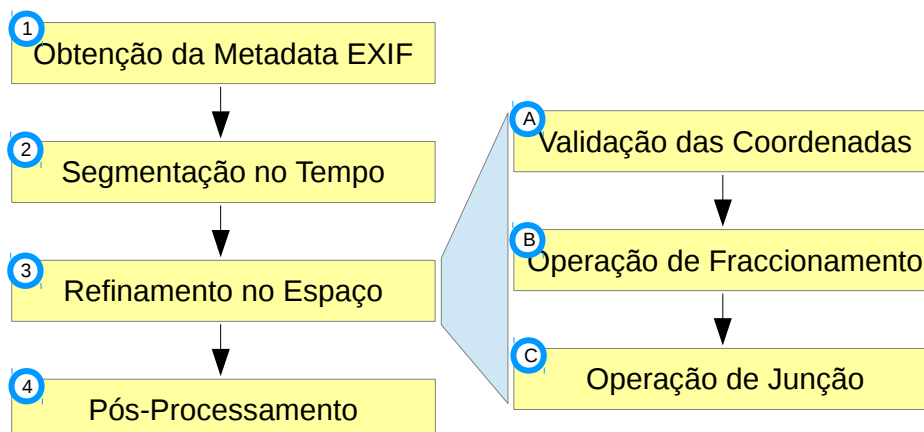


Figura 3.1: Etapas do Processo de Segmentação.

A extensão à implementação existente – representada na figura pela etapa 3) “Refinamento no Espaço” – começa por a) verificar a informação geográfica no sentido de determinar que coordenadas contêm informação que deve ser considerada errada. Os processos utilizados no sentido de tentar identificar esta informação errónea são descritos na secção 3.3 “Detecção de Coordenadas Erradas”. Depois disso, analisa-se a segmentação criada em 2) com o intuito de determinar a forma como a segmentação no tempo distribuiu as fotografias pelos diversos segmentos. Analisando a informação geográfica considerada válida, em B) determina-se se é vantajoso fragmentar cada um dos segmentos contidos na segmentação criada em 2). Depois da operação de fragmentação aplicada a cada um dos segmentos, em C) determina-se se faz sentido unir alguns dos segmentos resultantes da fase anterior, B).

Note-se que as alterações feitas sobre a segmentação são realizadas apenas ao nível dos segmentos: em nenhuma circunstância uma fotografia é eliminada da segmentação ou reordenada em relação às demais.

A análise começa por usar a informação geográfica para calcular a distância que separa as localizações de cada par de fotografias consecutivas. A ideia consiste em determinar a distância percorrida desde a última fotografia tirada; parte-se do princípio que quanto maior for esta distância maior é a probabilidade desta fotografia representar o início de um evento. É por este motivo que a curvatura da terra não é tida em conta: não interessa determinar distâncias com grande precisão, interessa apenas determinar quais foram as fotografias que foram tiradas a grande distância da fotografia precedente. É também por este motivo que as fotografias não podem ser reordenadas: uma reordenação das mesmas destruiria qualquer significado contido nestas distâncias.

### 3.2.1 Parâmetro de Calibração do Algoritmo

O Algoritmo de Segmentação no Tempo e Espaço (ASE) tem a capacidade de variar o seu comportamento, através de um parâmetro designado *SpaceThreshold*, que altera o número de segmentos gerados. O parâmetro tem valores admissíveis no intervalo [0; 100]. Internamente, este parâmetro é utilizado para derivar três patamares: *ReferenceJoinPct*, *ReferenceSplitPct* e *SplitSensitivity*. O primeiro é utilizado na operação de junção, enquanto os dois últimos são utilizados na operação de fraccionamento.

O patamar *ReferenceJoinPct* determina o ganho que o resultado da operação de junção precisa atingir para substituir na segmentação os segmentos operados

pela operação de junção. O patamar *ReferenceSplitPct* tem uma função idêntica ao anterior, mas agora para a operação de fraccionamento: os segmentos gerados pela operação de fraccionamento substituem o segmento que lhes deram origem quando o ganho que apresentam assim o justificar. O patamar *SplitSensitivity* regula a quantidade de segmentos que a operação de fraccionamento gera a partir de um único segmento. Estes patamares serão posteriormente descritos em mais detalhe nas secções 3.2.3 e 3.2.4.

### 3.2.2 Ordem da Aplicação das Operações

Tendo em conta que existe mais do que uma operação, há a necessidade de se estabelecer uma ordem de execução entre as mesmas. Por convenção, é executada em primeiro lugar a operação de Fraccionamento para cada um dos segmentos que existe na segmentação original, gerada pelo Algoritmo de Segmentação no Tempo (AST). Este processo de Fraccionamento gera uma segmentação temporária,  $S'_f$ . Posteriormente, para cada par de segmentos<sup>1</sup> adjacentes na segmentação  $S'_f$ , executa-se a operação de Junção de segmentos.

### 3.2.3 Descrição da Operação de Fraccionamento

Para cada segmento  $s$ , o algoritmo determina se vale a pena separar o segmento  $s$  em vários novos segmentos  $s_1, s_2, \dots, s_n$ . O processo de separação faz-se detectando as fotografias que apresentam uma distância estatisticamente elevada e tornando essas fotografias nas primeiras fotografias dos segmentos  $s_1, s_2, \dots, s_n$ . Para determinar o que constitui uma distância estatisticamente elevada, tem-se em conta as restantes distâncias presentes nesse segmento e o patamar *ReferenceSplitPct* que regula a sensibilidade do algoritmo. Este patamar foi já mencionado na secção 3.2.1.

Assume-se que o primeiro elemento de um dado segmento representa a primeira fotografia de um dado evento na vida do utilizador. O motivo que leva a considerar que a distância indicada nesse primeiro elemento é irrelevante é o facto de, no contexto desse evento, não existirem fotografias anteriores. Desta forma não faz sentido incluir nos cálculos estatísticos qualquer distância que se calcule com base nos dados indicados no primeiro elemento de cada segmento.

Os novos segmentos gerados pela Operação de Fraccionamento têm a particularidade de não serem posteriormente processados na Operação de Junção, para

---

<sup>1</sup>Com excepção dos segmentos indicados na secção 3.2.3.

garantir que a Operação de Junção não desfaz o trabalho feito pela Operação de Fraccionamento. No entanto, estão excluídos desta regra o primeiro e último segmento de cada conjunto de segmentos gerados pela Operação de Fraccionamento. Isto faz-se para que seja possível misturar fotografias localizadas em segmentos gerados por critérios geográficos – obtidos pela Operação de Fraccionamento – com fotografias localizadas em segmentos gerados por critérios temporais, através do Algoritmo de Segmentação no Tempo (AST).

### O patamar *ReferenceSplitPct*

O patamar *ReferenceSplitPct* determina o ganho que os segmentos resultantes da operação de fraccionamento têm que atingir para poderem substituir na segmentação o segmento que lhes deu origem.

$$ReferenceSplitPct = 0 + SpaceThreshold \quad (3.1)$$

A Equação 3.1 determina que é o próprio *SpaceThreshold* que dita o ganho que um dado resultado da operação de fraccionamento deve apresentar, quando comparado com o segmento original; quanto mais alto for *SpaceThreshold* mais resultados da operação de fraccionamento serão aceites na segmentação final.

A Figura 3.2 ilustra a forma como o patamar *ReferenceSplitPct* influencia a segmentação. Considere-se o segmento  $s_a$ . Ao ser processado pela operação de fraccionamento, este segmento resulta no conjunto de segmentos  $\{s_b, s_c, s_d\}$ . Também aqui na Figura 3.2, que ilustra a forma como o segmento  $s_a$  foi fraccionado, as linhas vermelhas e as bandas coloridas representam a “média mais ou menos o Desvio Padrão (DP)” dos vários segmentos e servem apenas para efeitos ilustrativos.

Devido à disparidade entre as distâncias representadas em  $s_b$ ,  $s_c$  e  $s_d$  (ilustradas pelas diferença entre as várias alturas dos pontos que se encontram nestes segmentos) o DP de  $s_a$  é maior que o DP de  $s_b$ ,  $s_c$  e  $s_d$ . Nesta situação os segmentos  $s_b$ ,  $s_c$  e  $s_d$  só serão incluídos na segmentação em substituição de  $s_a$  se

$$100 * \frac{average(DP(S_b), DP(S_c), DP(S_d))}{DP(a)} < ReferenceSplitPct \quad (3.2)$$

Se por exemplo tivermos  $SpaceThreshold = ReferenceSplitPct = 10$ , os segmentos resultantes  $\{s_b, s_c, s_d\}$  só seriam considerados aceitáveis se a média dos DP's destes segmentos for inferior a 10% da média do DP de  $s_a$ . Caso  $SpaceThreshold = ReferenceSplitPct = 70$ ,  $s_b$ ,  $s_c$  e  $s_d$  só serão incluídos na segmentação em substituição

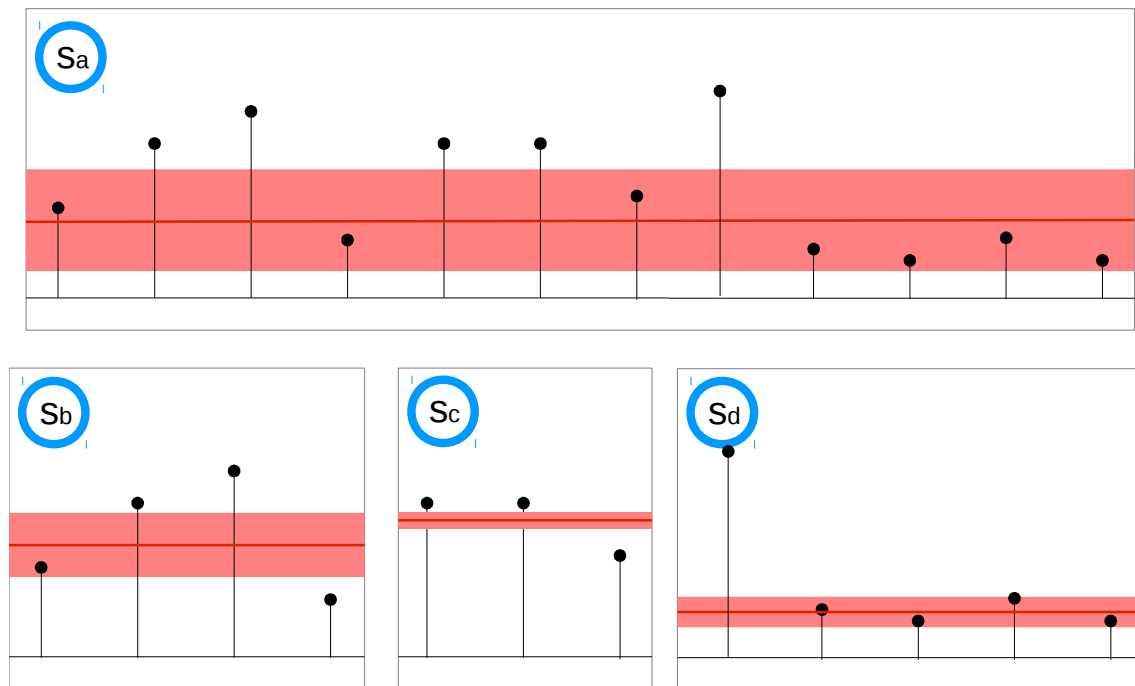


Figura 3.2: O segmento  $s_a$  (em cima) é processado pela operação de fracionamento a qual produz os segmentos  $s_b$ ,  $s_c$  e  $s_d$  (ilustrados em baixo). As linhas vermelhas e as bandas coloridas representam a “média mais ou menos o Desvio Padrão (DP)”.

de  $s_a$  se a média dos DP’s daqueles segmentos for inferior a 70% do DP de  $s_a$ .

### O patamar *SplitSensitivity*

Este patamar diz respeito à operação de fracionamento. Determina quais os pontos onde se fraciona o segmento operado e é também derivado do parâmetro *SpaceThreshold*:

$$SplitSensitivity = 150 - SpaceThreshold \quad (3.3)$$

*SplitSensitivity* regula assim a quantidade de segmentos gerados durante a operação de fracionamento. Quanto mais alto for *SpaceThreshold* maior será o número de segmentos gerados pela operação de fracionamento sobre um dado segmento.

A Figura 3.3 mostra a forma como o patamar *SplitSensitivity* é utilizado na operação de fracionamento. Sobre o segmento a ser fracionado, calcula-se a média das distâncias calculadas para cada coordenada (Figura 3.3(1)).

Por definição a coordenada da primeira fotografia de um segmento não tem distância associada: é a primeira fotografia que retrata um evento social e a distância

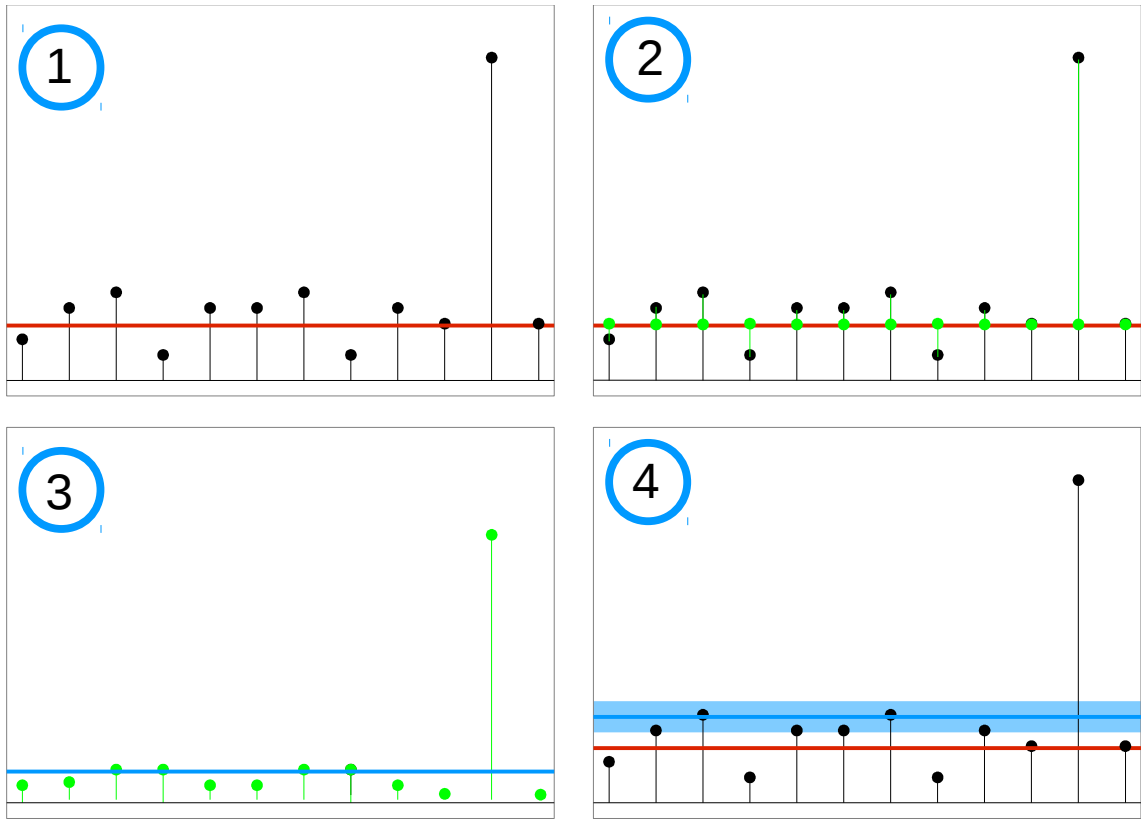


Figura 3.3: A operação de fraccionamento realiza-se ao longo de quatro etapas: 1) cálculo da média das distâncias; 2) cálculo do módulo da diferença de cada distância em relação à média; 3) cálculo do Desvio Padrão do módulo da diferença; 4) determinar os pontos de fragmentação.

que a separa da fotografia precedente não é relevante, uma vez que esta fotografia representa um outro evento social. Assim, a distância associada a uma fotografia de um segmento  $s_a$  com uma etiqueta  $Is_{a_n}$  é definida por:

$$\begin{cases} distancia(Is_{a_n}) = \emptyset, n = 1 \\ distancia(Is_{a_n}) = \|Is_{a_n} - Is_{a_{n-1}}\|, n \geq 2 \wedge n \leq \|IS\| \end{cases} \quad (3.4)$$

e a média das distâncias indicadas pelos elementos de um segmento  $s_a$  é dada por:

$$average(distancias(s_a)) = \frac{\|Is_{a_2} - Is_{a_1}\| + \|Is_{a_3} - Is_{a_2}\| + \dots + \|Is_{a_n} - Is_{a_{n-1}}\|}{n - 1} \quad (3.5)$$

A segunda etapa — Figura 3.3(2) — consiste em calcular o módulo da diferença que cada distância apresenta em relação à média. De seguida, sobre o conjunto

de módulos de diferenças, é calculado o Desvio Padrão(DP) como ilustra a Figura 3.3(3). A etapa final (Figura 3.3(4)) consiste em determinar quais são as fotografias cujas coordenadas apresentam uma diferença à média das distâncias superior ao DP do módulo da diferença (este último calculado na etapa anterior), ponderado pelo patamar *SplitSensitivity*. Estas coordenadas serão o primeiro elemento de um novo segmento. Assim, uma fotografia com uma etiqueta  $I_{s_n}$  será a primeira de um novo segmento se a *distancia*( $I_{s_n}$ ), derivada das coordenadas das fotografias  $I_{s_n}$  e  $I_{s_{n-1}}$  satisfazer a seguinte condição:

$$DP(\text{módulo da diferença}) * \frac{SplitSensitivity}{100} < distancia(I_{s_n}) - media(IS) \quad (3.6)$$

### 3.2.4 Descrição da Operação de Junção

Depois de todos os segmentos na segmentação  $S$  terem sido processados pela operação de Fraccionamento, a segmentação  $S$  é novamente reprocessada desta vez pela operação de Junção. Considerem-se três segmentos consecutivos,  $s_1$ ,  $s_2$  e  $s_3$ . Para cada par de segmentos consecutivos – por exemplo  $s_1$  e  $s_2$  – cria-se um novo segmento –  $s_{1,2}$  – que resulta da junção daqueles dois segmentos. Se o DP de  $s_{1,2}$  for inferior à média dos DP's de  $s_1$  e  $s_2$  (ponderado pelo patamar *ReferenceJoinPct*, já mencionado na secção 3.2.1), removem-se os segmentos originais da segmentação e insere-se em seu lugar o novo segmento  $s_{1,2}$ . A próxima operação de Junção é então realizada sobre  $s_{1,2}$  e  $s_3$  ou sobre  $s_2$  e  $s_3$  em função da operação de Junção anterior ter tido sucesso, ou não.

A operação de Junção opera sobre os segmentos que não foram criados pela operação de Fraccionamento. Sobre os segmentos criados pela operação de Fraccionamento, apenas os que se encontram nas extremidades dos conjuntos de segmentos gerados por esta operação são posteriormente sujeitos à operação de Junção — conforme a exceção expressa na secção 3.2.3. Esta regra pretende evitar que se desfaça completamente o resultado da operação de Fraccionamento. Em simultâneo, esta regra permite que se movam as *primeiras* fotografias de um dado segmento para o segmento *anterior* e vice-versa.

#### O patamar *ReferenceJoinPct*

$$ReferenceJoinPct = 100 - SpaceThreshold \quad (3.7)$$

Este patamar é responsável por determinar qual o ganho que um dado segmento resultante da operação de junção deve apresentar quando comparado com os dois segmentos originais. Quanto mais elevado for *SpaceThreshold* menos segmentos resultantes da operação de junção serão aceites na segmentação final.

A Figura 3.4 ilustra a forma como este patamar influencia a junção. Considerem-se dois segmentos  $s_a$  e  $s_b$  que ao serem processados pela operação de junção resultam no segmento  $s_c$ . Na Figura 3.4 as linhas vermelhas e as bandas coloridas representam a “média mais ou menos o Desvio Padrão (DP)” do segmento em que se encontram. Note-se que tanto as linhas como as bandas são válidas apenas para efeitos ilustrativos.

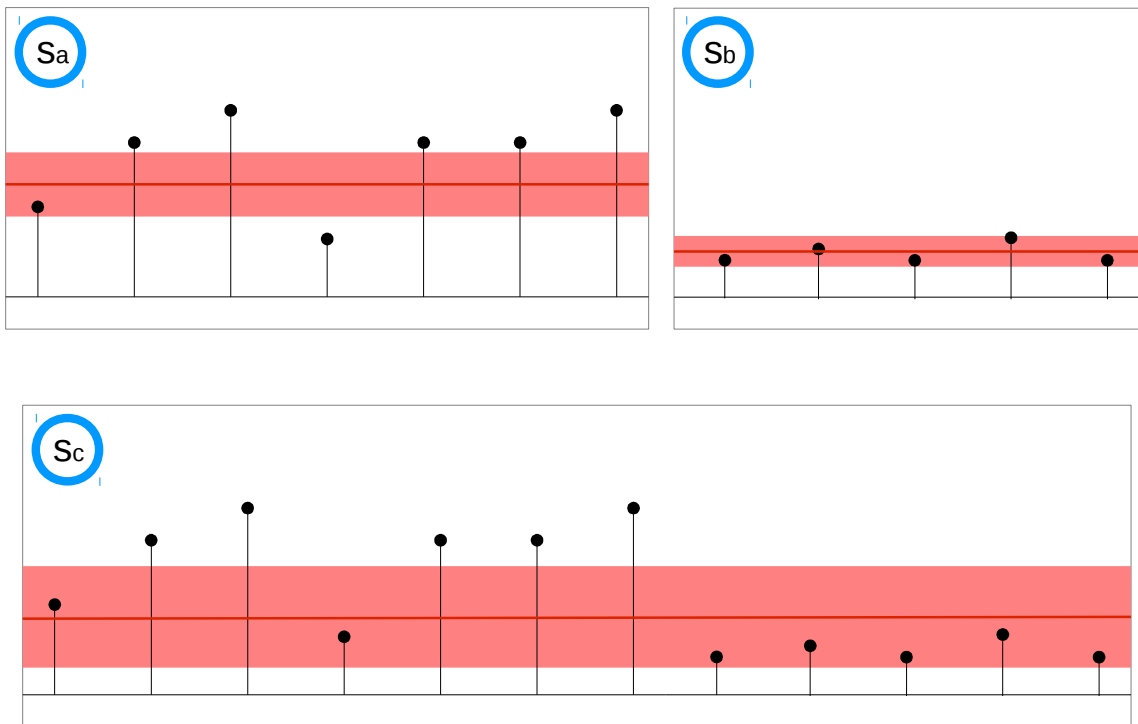


Figura 3.4: Os dois segmentos  $s_a$  e  $s_b$  (em cima) são processados pela operação de junção a qual produz o segmento  $s_c$  (ilustrado em baixo). As linhas vermelhas e as bandas coloridas representam a “média mais ou menos o Desvio Padrão (DP)”.

Devido à disparidade entre as distâncias representadas em  $s_a$  e  $s_b$  (ilustradas pelas diferenças entre as alturas dos pontos que se encontram nestes segmentos) o DP de  $s_c$  é maior que o DP de  $s_a$  ou  $s_b$ . Nesta situação, o segmento  $s_c$  só será incluído na segmentação em substituição de  $s_a$  e  $s_b$  se

$$\frac{100 * DP(S_c)}{average(DP(S_a), DP(S_b))} < referenceJoinPct \quad (3.8)$$

A título de exemplo, para *SpaceThreshold* = 10, temos *ReferenceJoinPct* = 90. Nesse

caso,  $s_c$  só seria considerado aceitável se o DP de  $s_c$  for inferior a 90% da média dos DP's de  $s_a$  e  $s_b$ . Caso  $SpaceThreshold = 70$ , temos  $ReferenceJoinPct = 30$ , e assim,  $s_c$  só será incluído na segmentação em substituição de  $s_a$  e  $s_b$  se o DP de  $s_c$  for inferior a 30% da média dos DP's de  $s_a$  e  $s_b$ .

### 3.2.5 Restrições do Algoritmo de Segmentação no Tempo (AST)

Neste passo, assegura-se que as restrições tidas em conta pelo Algoritmo de Segmentação no Tempo (AST) não foram completamente ignoradas. Em particular, de forma a manter o conceito de dia lógico [7] processa-se cada par de segmentos adjacentes: considerando que a última fotografia do primeiro segmento –  $It_1^+$  – foi tirada no dia  $d_1$ , todas as fotografias do segundo segmento que foram tiradas no dia  $d_2$  e que distam menos de um certo número de horas ao instante de  $It_1^+$  são movidas para o fim do primeiro segmento e inseridas depois de  $It_1^+$ .

Depois deste processo, a fase de “Refinamento no Espaço” ilustrada na Figura 3.1 considera-se terminada e o controlo é devolvido ao código desenvolvido anteriormente em [7].

## 3.3 Detecção de Coordenadas Erradas

### 3.3.1 Detecção com Base na Velocidade

O método de detecção de coordenadas com base na velocidade tenta determinar se uma dada coordenada  $c$  está errada através da velocidade que se pode inferir tanto pela informação temporal como pela informação geográfica presentes na própria fotografia e nas fotografias adjacentes.

O algoritmo implementado permite que se indique tanto a quantidade de fotografias precedentes como a quantidade de fotografias subsequentes que devem ser consideradas no cálculo desta velocidade. A este conjunto de fotografias dá-se o nome de *vizinhança*.

São estas duas quantidades que são parametrizadas neste algoritmo, e como vamos ver posteriormente, foram ensaiados 4 conjuntos de parâmetros: 1:0; 1:1; 2:2; 3:3. Por outras palavras, por cada teste realizado foram ensaiadas quatro iterações para testar cada uma daquelas quatro configurações de forma independente. O significado daquelas parametrizações indica a quantidade de coordenadas anteriores e posteriores à coordenada  $c$  que este algoritmo deve considerar. Assim,

para o primeiro parâmetro deve ser considerada apenas a coordenada anterior a  $c$ ; o segundo parâmetro indica que devem ser apenas consideradas duas coordenadas: a anterior e a posterior; etc.

Para cada coordenada  $c'$  registada nas fotografias incluídas na vizinhança, o algoritmo calcula o espaço  $e$  que separa  $c$  de  $c'$ , bem como a diferença temporal  $t$  entre  $c$  de  $c'$ .

Com esta informação, é calculada a velocidade média  $vm$  – a partir do conjunto de  $v$ 's calculadas com base no:

- espaço  $e$  que foi percorrido entre  $c$  e cada um dos pontos da vizinhança (os vários  $c'$ ) e;
- na diferença temporal  $t$  entre  $c$  e cada um dos pontos  $c'$  da vizinhança.

A partir desta informação determina-se que a coordenada  $c$  está errada se e só se:

- i)  $vm$  é superior a *Mach 1*<sup>2</sup>, porque se considera que uma fotografia que seja tirada a esta velocidade não será processada pelo algoritmo descrito neste trabalho ou;
- ii) a média calculada a partir dos vários  $e$  é inferior a 5km – uma distância normalmente percorrida a pé – e a  $vm$  é superior a 100km/h – uma distância normalmente alcançada através de um meio de transporte.

Pode-se argumentar que uma sequência de fotografias tiradas, por exemplo, durante uma viagem de comboio se enquadra no segundo critério: uma paisagem é fotografada múltiplas vezes enquanto o utilizador está em movimento. Nesta situação, apenas a primeira fotografia irá ser considerada correcta por este algoritmo de detecção de coordenadas, e conseqüentemente apenas esta primeira fotografia irá ser considerada pelo Algoritmo de Segmentação no Tempo e Espaço (ASE). No entanto, o ASE terá a oportunidade de utilizar essa primeira fotografia da sequência para iniciar um novo segmento, no qual também serão inseridas as fotografias subsequentes cujas coordenadas foram incorrectamente consideradas erradas por este algoritmo de detecção de coordenadas.

---

<sup>2</sup>“Mach 1” é o nome normalmente associado à velocidade do som, que no âmbito deste trabalho se considera ser igual a 1225.044 km/h.

### 3.3.2 Detecção com Base no *Interquartile Range* “IQR”

Este método de detecção de coordenadas calcula as distâncias que separam cada par de fotografias adjacentes contidas na segmentação, e considera correctas apenas as coordenadas das fotografias cujas distâncias associadas se encontrem numa gama de valores derivados do primeiro e o terceiro quartil [34].

Depois de calcular a distância percorrida para tirar cada uma das fotografias presentes na segmentação, calcula-se o *Interquartile Range*. Para isso, as distâncias são ordenadas de forma crescente sob a forma de uma lista  $l$  de distâncias e obtêm-se dois elementos nesta lista;  $q_1$  e  $q_3$ :

- i)  $q_1$  é o elemento que se encontra na posição  $\text{dimensao}(l) \div 4$  da lista  $l$ ;
- ii)  $q_3$  encontra-se na posição  $3 \times \text{dimensao}(l) \div 4$  da lista  $l$ ;

Uma dada distância é considerada correcta se e só se estiver compreendida entre:

$$\begin{cases} q_1 - k \times (q_3 - q_1) < \text{distancia} \wedge \\ \text{distancia} < q_3 + k \times (q_3 - q_1) \end{cases} \quad (3.9)$$

onde  $k$  é o factor de IQR, um parâmetro que por omissão é 1.5 na implementação do presente trabalho.

### 3.3.3 Detecção com Base no *Local Outlier Factor* “LOF”

Este último método de detecção de coordenadas utiliza o algoritmo *Local Outlier Factor* descrito em [3] como meio de detecção de coordenadas erradas. No âmbito deste trabalho, tira-se partido do facto do conjunto de coordenadas presentes nas fotografias da segmentação ter uma relação de ordem no contexto do problema, algo que não é considerado na versão original deste algoritmo. Assim sendo, considera-se que o conjunto de coordenadas a avaliar pelo algoritmo LOF está limitado às 10 coordenadas precedentes mais as 10 coordenadas subsequentes à coordenada sob escrutínio. Esta redução faz-se no sentido de se executar em tempo útil este algoritmo o qual é de complexidade  $O(n^2)$ , na sua versão original.

## 3.4 Implementação

Depois da colecção de fotografias ser processada pelo algoritmo descrito em [7], obtém-se uma estrutura de dados representativa da segmentação produzida por aquele algoritmo. Essa segmentação é então processada pelo algoritmo descrito neste trabalho com o objectivo de a modificar.

### 3.4.1 Distância entre Coordenadas

Para todas as fotografias existentes na segmentação, o algoritmo começa por determinar a distância que separa cada par de fotografias adjacentes. É sobre estas distâncias que todos os cálculos estatísticos são realizados. Por definição, a distância associada à primeira fotografia de uma segmentação é desprovida de significado e como tal é excluída dos cálculos descritos em seguida. De forma a simplificar o cálculo destas distâncias, consideram-se que as coordenadas são cartesianas. Este cálculo não é o mais correcto, uma vez que se está a trabalhar com coordenadas geográficas, mas admite-se que o erro que daí possa surgir é desprezável para o problema em questão.

### 3.4.2 Detecção de Coordenadas Erradas

Uma vez determinada a distância que separa cada par de fotografias consecutivas, o algoritmo tenta determinar que coordenadas poderão ser consideradas erradas. Para o efeito, utiliza-se um dos três algoritmos já descritos na secção 3.3 “Detecção de Coordenadas Erradas”: *Local Outlier Factor* (“LOF”) [3], *Interquartile Range* (“IQR”) e um algoritmo que usa o conceito de velocidade. Qualquer coordenada que tenha sido considerada errada não é considerada nos cálculos estatísticos que determinam que alterações devem ser feitas à segmentação existente.

### 3.4.3 Estatísticas de Segmentação

Em seguida calculam-se estatísticas descritivas de toda a segmentação. Este cálculo tem o objectivo de determinar valores estatísticos relativos a todas as fotografias e não apenas a cada um dos segmentos. Estes valores globais são utilizados quando os seus equivalentes relativos a um dado segmento não são úteis para os

cálculos deste algoritmo. A título de exemplo, temos o caso do Desvio Padrão que é igual a 0 para o caso de um segmento com uma única fotografia.

### 3.4.4 Operação de Fraccionamento

Para cada um dos segmentos contidos na segmentação no tempo, ensaia-se a operação de fraccionamento. Esta operação consiste em descobrir quais são as fotografias que apresentam uma distância considerada extrema quando comparada às demais fotografias presentes no segmento escrutinado. Para determinar esta diferença relativa utiliza-se uma de duas métricas estatísticas: o Desvio Padrão (DP) e o Desvio Absoluto Mediano (DAM).

Para o caso do DP, a operação de fraccionamento já foi previamente descrita na secção 3.2.3 “Descrição da Operação de Fraccionamento”. No caso do DAM, o processo é idêntico mas em vez de usar as medidas estatísticas do DP e média são utilizadas em vez disso o DAM e a mediana.

Depois de ensaiada a operação de fraccionamento, compara-se o DP do segmento original contra a média dos DP dos segmentos que resultaram da operação. A ideia subjacente a esta comparação é a consideração de que o DP é uma medida de coerência do conjunto de fotografias que compõem um segmento, ou por outras palavras, o DP determina o grau de afastamento que as fotografias de um segmento apresentam umas em relação às outras.

Desta comparação surge uma de duas possibilidades: ou a coerência do conjunto de segmentos resultantes do fraccionamento está dentro de um limite parametrizável no algoritmo, ou então está fora desse limite.

Se a coerência do conjunto de segmentos resultantes do fraccionamento estiver dentro desse limite, então esses segmentos são inseridos na segmentação original, em substituição do segmento que lhes deu origem, porque afinal, o ganho conseguido em termos de coerência assim o justifica. Caso contrário, a segmentação permanece inalterada e o conjunto de segmentos resultantes do fraccionamento é descartado.

### 3.4.5 Operação de Junção

Depois de executada a operação de Fraccionamento sobre cada um dos segmentos na segmentação original, realiza-se a operação inversa. A operação de Junção é realizada entre cada par de segmentos consecutivos. Descobre-se assim quais

são os segmentos de fotografias que apresentam uma distância considerada próxima entre si.

Depois de dois segmentos consecutivos terem sido juntos num só segmento, a média dos DP's dos segmentos originais é comparada com o valor do DP do segmento resultante da operação de junção. Mais uma vez, considera-se que o DP mede a coerência do conjunto de fotografias que compõem um segmento, ou seja, determina o grau de proximidade que as fotografias de um segmento apresentam umas em relação às outras.

Novamente, desta comparação surge uma de duas possibilidades: ou a coerência do segmento resultante da junção está dentro do limite parametrizado no algoritmo, ou então está fora desse limite.

Se o segmento resultante da junção apresentar uma coerência abrangida por esses limites, então a segmentação original passa a conter esse novo segmento. Os dois segmentos originais são eliminados, porque apresentam uma coerência inferior ao novo segmento. Caso contrário, a segmentação permanece inalterada e o segmento resultante da junção é descartado.

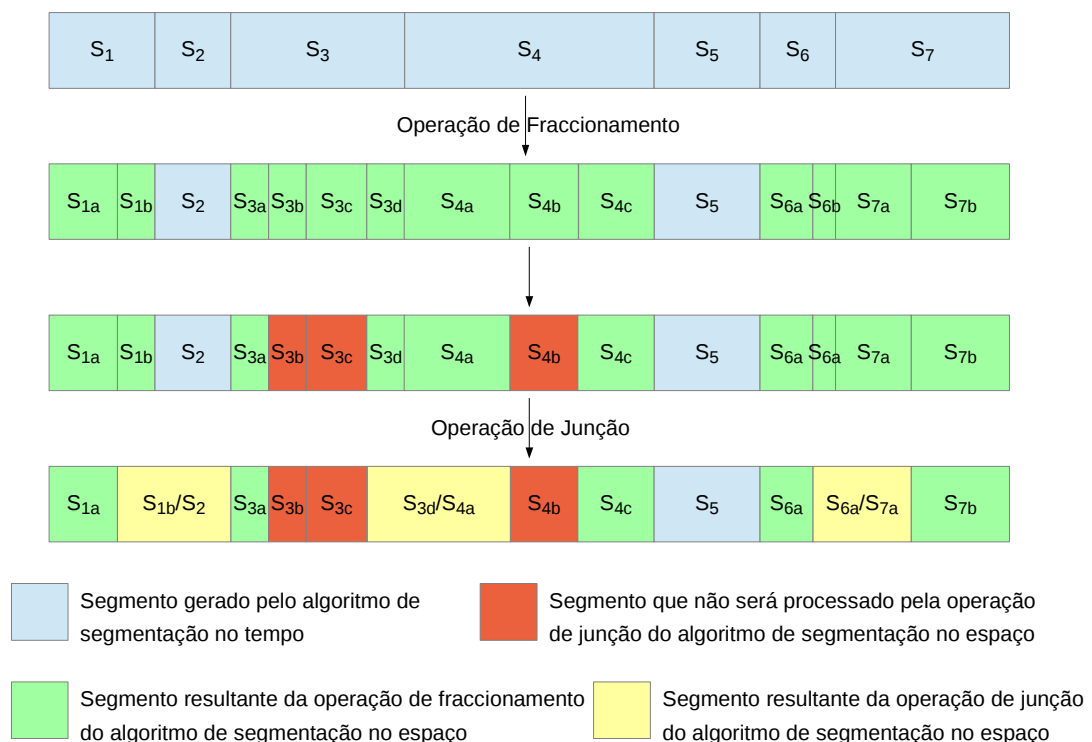


Figura 3.5: Exemplo da sequência de operações realizadas pelo Algoritmo de Segmentação no Tempo e Espaço (ASE) sobre uma segmentação produzida pelo Algoritmo de Segmentação no Tempo (AST).

Garante-se que a operação de junção não desfaz completamente os efeitos da operação anterior, limitando os segmentos sobre os quais a junção é realizada: só são operados os segmentos originais da segmentação no tempo e os segmentos que se encontram nos extremos dos conjuntos de segmentos criados pela operação de fraccionamento. A Figura 3.5 ilustra este processo.

## 3.5 Descrição da Integração com o Trabalho Pré-existente

O algoritmo desenvolvido originalmente no âmbito do trabalho [7] foi estendido no sentido de suportar a segmentação no espaço. A classe *SimpleTimeAlg* (“*Simple Time Algorithm*”) é um ponto fulcral no programa original, uma vez que é a entidade responsável pela execução da segmentação no tempo. Ao substituir esta classe por uma outra controla-se o critério segundo o qual se cria a segmentação. Por este motivo, a classe *SpatialAlgorithm* é estendida a partir da classe *SimpleTimeAlg*.

A classe *SimpleTimeAlg* não foi originalmente pensada para suportar uma extensão, mas no âmbito do desenvolvimento da classe *SpatialAlgorithm* foram efectuadas algumas alterações. Procurou-se minimizar estas alterações, as quais acabaram por ser maioritariamente a alteração da visibilidade de alguns dos atributos – que passaram de *private* para *protected* – e a adição de alguns métodos de conveniência; *getters* para a obtenção de informação relevante tanto para classes derivadas como para entidades cliente que utilizam as instâncias da classe *SimpleTimeAlg*. Uma vez que a nova classe deriva da classe original *SpatialAlgorithm*, as diferenças que se fazem sentir nas entidades cliente são minimizadas. Na Figura 3.6 apresenta-se um diagrama com a sequência de activação das duas classes:

Note-se que a segmentação no espaço invoca internamente o método *Process* da classe *SimpleTimeAlg*, porque se pretende ajustar uma segmentação no tempo utilizando a informação geográfica, também presente nas fotografias. Depois de se garantir que existe uma segmentação no tempo, invocam-se os restantes mecanismos da classe *SpatialAlgorithm* os quais alteram essa segmentação.

Durante a implementação de *SpatialAlgorithm* surgiu a questão do que se consideraria um tempo de execução admissível, pelo que foi introduzido na classe original um método de extracção de estatísticas de execução. À medida que o desenvolvimento da nova classe *SpatialAlgorithm* e os testes de caracterização foram progredindo, as estatísticas que daí resultaram foram tidas em conta para se

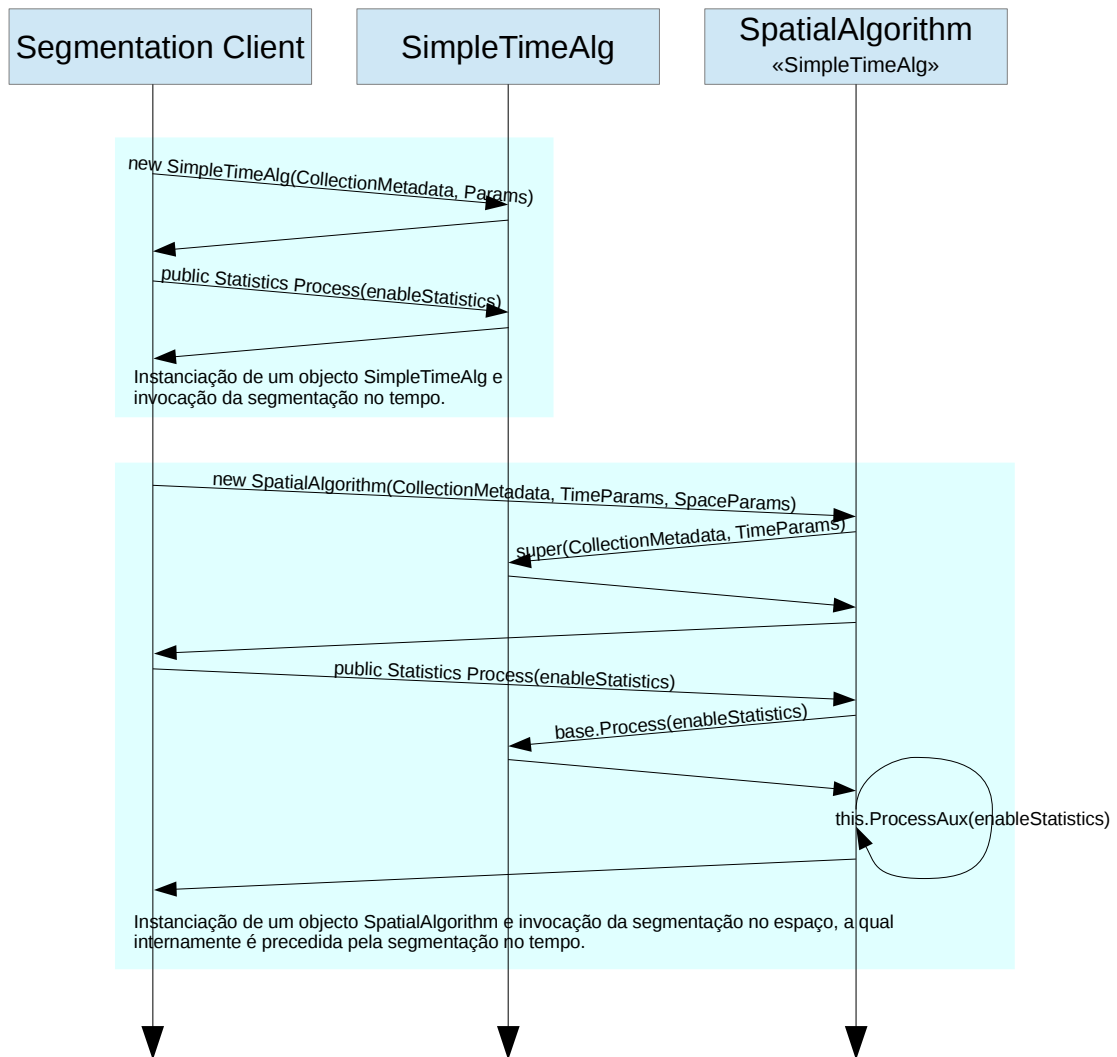


Figura 3.6: Diagrama e pseudo-código ilustrativo das diferenças no processo de activação dos dois algoritmos: a principal diferença é a inclusão do parâmetro de calibração do Algoritmo de Segmentação no Tempo e Espaço; SpaceThreshold.

verificar se o tempo de execução do novo algoritmo era minimamente aceitável (tendo em conta os tempos de execução do algoritmo original) e para determinar se o benefício de uma dada alteração justificava o seu custo. Esta preocupação prende-se com o facto do tempo de execução do algoritmo afectar o tempo de espera por parte dos utilizadores da aplicação *web*.

Ainda no âmbito dos testes de caracterização, foi implementada a funcionalidade de persistir em base de dados os resultados de cada um dos testes aos algoritmos: são persistidas toda uma série de métricas bem como as estatísticas que resultam de cada execução do algoritmo sobre as várias colecções testadas, e para as diversas variações de cada um dos parâmetros disponibilizados pelo algoritmo, conforme se menciona na secção 4.1.

Alguns detalhes da implementação existente foram alterados simplesmente para melhorar a semântica da aplicação. Um exemplo disso é o conceito de coordenada, que passou a ser exprimido sob a forma de uma classe, ao invés da simples utilização de um par de valores decimais sempre que necessário.

Tendo em conta a complexidade de algumas das estruturas de dados do algoritmo e a minúcia com a qual estas são manipuladas à medida que o algoritmo progride, tornou-se necessário a inspecção detalhada destas estruturas de dados. Optou-se por realizar a serialização para XML<sup>3</sup> destas estruturas de dados em diversos pontos do programa de forma a se poder avaliar a forma como o algoritmo progride após a execução deste último. Para o efeito, para além do código que efectivamente realiza a serialização, as classes das estruturas de dados relevantes foram marcadas com a interface *Serializable*.

Foi criada uma interface – *IAlgorithm* – a qual define o que é um algoritmo e a respectiva interface: que questões um algoritmo deve ter a capacidade de responder. Esta interface foi definida depois de se fazer um levantamento ao código já existente e analisar as necessidades dos clientes da classe *SimpleTimeAlg*. Posteriormente e sempre que possível, as referências a *SimpleTimeAlg* foram substituídas por referências a *IAlgorithm*. Esta alteração foi implementada numa altura inicial do desenvolvimento; uma altura em que se planeava desenvolver a classe *SpatialAlgorithm* de forma independente de *SimpleTimeAlg*. No entanto as necessidades de *SpatialAlgorithm* revelaram-se muito semelhantes às de *SimpleTimeAlg* desde cedo. Adicionalmente, uma classe *SpatialAlgorithm* que não emulasse de perto o comportamento e estruturas de *SimpleTimeAlg*, viria a revelar-se difícil de integrar com o código já existente. Por este motivo *SpatialAlgorithm* passou a derivar de *SimpleTimeAlg* e a utilizar as suas estruturas e métodos sempre que possível ou necessário. A interface *IAlgorithm*, apesar de devido a este último desenvolvimento se tornar redundante, foi mantida devido ao peso que o conceito por ela implementado tem no problema.

Na aplicação *web* original, o ponto do código onde o algoritmo original – *SimpleTimeAlg* – é invocado foi alterado de forma a procurar no URL parâmetros que digam respeito ao Algoritmo de Segmentação no Tempo e Espaço (ASE). Quando estes parâmetros são encontrados, *SpatialAlgorithm* é posto em execução e este encarrega-se de internamente invocar o algoritmo de *SimpleTimeAlg*. Só depois deste último concluir o seu trabalho é que *SpatialAlgorithm* invoca os seus próprios algoritmos sobre o resultado de *SimpleTimeAlg*.

---

<sup>3</sup><http://www.w3.org/XML/>, consultado em Dezembro de 2014

# 4

## Avaliação Experimental

Os ensaios práticos realizados consistiram em testes de caracterização dos algoritmos desenvolvidos e testes com utilizadores.

Os testes de caracterização têm como objectivo determinar o comportamento do algoritmo implementado em função dos parâmetros *TimeThreshold* e *SpaceThreshold*. A execução dos testes permitiu verificar as diferenças entre os dois métodos estatísticos (DP e DAM), o número de vezes que as operações de fraccionamento e junção foram efectuadas, o grau de semelhança entre segmentações e o comportamento dos vários algoritmos de detecção de coordenadas.

Os testes de caracterização tiveram também a função de auxiliar o desenvolvimento dos algoritmos: compreender a forma como os algoritmos evoluíam em função das modificações realizadas e assim detectar e remover erros ou problemas introduzidos na aplicação durante o seu desenvolvimento.

Os testes de caracterização incidiram sobre 39 colecções de fotografias disponíveis ao público em geral na *Internet*. Nestas colecções, o número de fotografias contidas varia entre as 17 e as quase 1400 fotografias, como se pode ver nas Tabelas A.1 e A.2 contidas no Anexo A. Estas Tabelas dão-nos também alguns indícios sobre a existência de irregularidades nos dados. A título de exemplo, a Tabela A.1 indica-nos que na colecção *Europe2010* a distância máxima entre duas fotografias consecutivas é de 12 mil e quinhentos quilómetros e na Tabela A.2 é-nos indicado que o tempo máximo entre duas fotografias consecutivas é superior a 5 anos e meio (*Atlantik2008*). No caso do primeiro tipo de irregularidade, a justificação é

simplesmente um erro na aquisição da coordenada geográfica. É necessário que este tipo de problema esteja presente nas colecções de forma a testar os algoritmos de detecção de coordenadas erradas. Quanto às irregularidades no domínio do tempo, após uma análise das fotografias que compõem as colecções afectadas, concluiu-se que esta anomalia se deve à ausência de metadata relativa à data de criação em algumas fotografias. Nesta situação, opta-se por utilizar a data de criação do ficheiro que contem a imagem da fotografia, e tendo em conta que a data de criação de um ficheiro descarregado da *Internet* é a data em que o mesmo foi descarregado, compreende-se o motivo das irregularidades. Uma vez que não há nenhum meio técnico de garantir a presença de qualquer tipo de metadata, optou-se por incluir também estas colecções de forma a observar o impacto nos algoritmos.

Os testes com utilizadores têm o objectivo de avaliar a preferência pelo algoritmo desenvolvido: determinar se os utilizadores preferem as segmentações criadas pelo algoritmo de segmentação no tempo e espaço ou as segmentações resultantes do algoritmo de segmentação no tempo.

A implementação do algoritmo desenvolvido neste trabalho foi integrada com o trabalho desenvolvido em [7]. O foco deste último trabalho consiste precisamente em automatizar o processo de segmentar e apresentar fotografias, tendo em conta apenas a informação temporal presente nessas fotografias.

## 4.1 Testes de Caracterização

Os testes de caracterização consistiram na avaliação dos resultados das várias partes que compõem o Algoritmo de Segmentação no Tempo e Espaço (ASE) e que o caracterizam em função da gama de valores admissíveis para o parâmetro

1. *SpaceThreshold* – mencionado na secção 3.2 – o qual varia entre 0 e 100 e para o parâmetro
2. *TimeThreshold* que varia entre 0.1 e 0.9.

A necessidade deste último parâmetro resulta do facto do Algoritmo de Segmentação no Tempo e Espaço (ASE) utilizar internamente o Algoritmo de Segmentação no Tempo (AST).

A Tabela 4.1 indica a quantidade de iterações que foram realizadas para as diversas combinações de métodos e parâmetros.

	<b>Desvio Padrão (DP)</b>	<b>Desvio Absoluto Mediano (DAM)</b>
<b>Sem detecção</b>	819	819
<b>IQR(1)</b>	819	819
<b>IQR(1.5)</b>	819	819
<b>IQR(2)</b>	819	819
<b>LOF(1)</b>	819	819
<b>LOF(2)</b>	819	819
<b>LOF(4)</b>	819	819
<b>Velocidade(1:0)</b>	819	819
<b>Velocidade(1:1)</b>	819	819
<b>Velocidade(2:2)</b>	819	819
<b>Velocidade(3:3)</b>	819	819

Tabela 4.1: Forma como os 18018 testes foram divididos pelos diversos algoritmos de detecção de coordenadas e métodos estatísticos..

Durante os testes de caracterização executaram-se mais de 18 mil iterações do ASE com o propósito de avaliar a implementação dos algoritmos desenvolvidos; o Algoritmo de Segmentação no Tempo e os Algoritmos de Detecção de Coordenadas Erradas.

Das cerca de 18 mil iterações do ASE realizadas durante os testes de caracterização, metade utilizou o Desvio Padrão (DP) nos cálculos de segmentação. As restantes iterações utilizaram o Desvio Absoluto Mediano (DAM) em vez do DP. Para além desta divisão, os testes também se dividiram em três métodos de detecção de coordenadas (*IQR*, *LOF* e velocidade). Como se pode ver na Tabela 4.1, o método de detecção de coordenadas *IQR* foi testado com 3 parâmetros diferentes (1, 1.5 e 2), o método de detecção de coordenadas *LOF* foi também testado com 3 parâmetros diferentes (1, 2 e 4) e o método de detecção de coordenadas baseado em velocidade foi testado com 4 parâmetros (1:0, 1:1, 2:2 e 3:3).

A prevalência do número 819 na Tabela 4.1 deriva da forma como se variou o parâmetro *SpaceThreshold* – entre [0; 100] com um incremento de 5, o que resulta em 21 valores diferentes – e do número de colecções testadas, 39. Multiplicando estes dois valores obtém-se aquele número.

Cada um dos resultados da segmentação no espaço foi comparado com a respectiva segmentação no tempo que lhe deu origem e verificou-se que:

- em 29.94% dos casos as duas segmentações eram idênticas;
- em 59.37% dos casos uma das duas segmentações era um refinamento da outra;

- em 10.67% dos casos as duas segmentações eram compatíveis e
- em 0.022% casos verificou-se que as duas segmentações eram incompatíveis;

Repare-se que neste contexto as palavras “refinamento”, “compatíveis” e “incompatíveis” têm um significado que é definido quando em [7], se determina a função que mede a distância entre segmentações. Uma segmentação  $S_b$  é um “refinamento” de uma segmentação  $S_a$  quando é possível obter a segmentação  $S_b$  através da aplicação da operação de fraccionamento sobre um ou mais segmentos presentes em  $S_a$ . A Figura 4.1 é um exemplo disso mesmo.

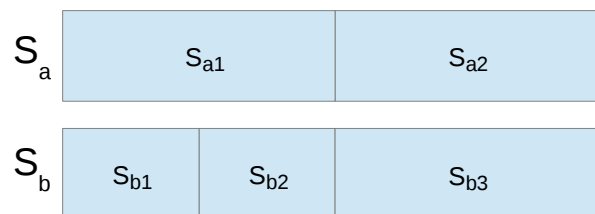


Figura 4.1: Exemplo de um refinamento de segmentação:  $S_b$  é um refinamento de  $S_a$ .

Na Figura 4.2 temos um exemplo de duas segmentações compatíveis uma com a outra. Duas segmentações  $S_a$  e  $S_b$  são “compatíveis” quando, não sendo idênticas nem um “refinamento”, é possível encontrar em cada uma das duas segmentações, um ou mais segmentos que na outra segmentação se encontram fraccionados.

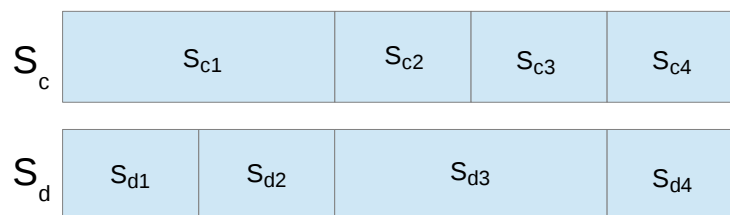


Figura 4.2: Neste caso as segmentações  $S_c$  e  $S_d$  são compatíveis uma com a outra. Os segmentos  $S_{c1}$  e  $S_{d3}$  encontram-se fraccionados nas segmentações  $S_d$  ( $S_{d1}$  e  $S_{d2}$ ) e  $S_c$  ( $S_{c2}$  e  $S_{c3}$ ).

Duas segmentações são incompatíveis uma com a outra, quando não é possível enquadrá-las nos casos anteriores. A Figura 4.3 mostra um desses casos.

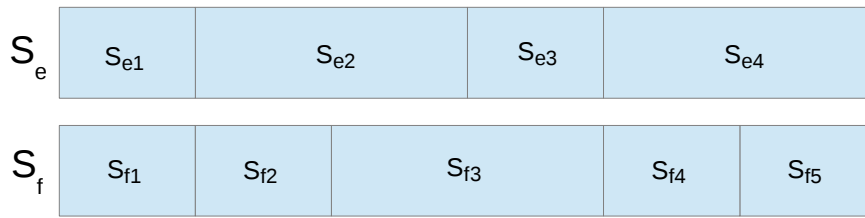


Figura 4.3: Neste caso as segmentações  $S_e$  e  $S_f$  são incompatíveis uma com a outra.

Foram também implementadas quatro medidas para comparar os segmentos: o  $RMiss$  e o  $Rfa$  ambos definidos em [13] e o  $Pk$  e o  $WindowDiff$  (um refinamento do  $Pk$ ) ambos definidos em [40].

Estas quatro métricas comparam a segmentação gerada pelo ASE, contra a segmentação que lhe deu origem — a segmentação gerada pelo AST. Tentam medir as diferenças no que diz respeito à localização das fronteiras que separam os vários segmentos contidos por uma segmentação.

Neste contexto, o  $RMiss$  tenta estabelecer uma relação de equivalência entre os segmentos das duas segmentações e indica quantas fotografias estão em falta nos vários segmentos na segmentação do ASE, quando comparados com os segmentos equivalentes da segmentação do AST. A métrica  $Rfa$  detecta a situação contrária: quantas fotografias estão em excesso nos vários segmentos da segmentação do ASE, ou por outras palavras, quantas fotografias foram incluídas no segmento errado, tendo em conta aquela relação de equivalência.

As métricas  $Pk$  e  $WindowDiff$  utilizam uma janela deslizante de dimensão  $k$  para se focarem nas fronteiras entre segmentos. Enquanto a métrica  $Pk$  determina se os segmentos indicados pelas extremidades da janela deslizante de dimensão  $k$  são o “mesmo” segmento ou segmentos “diferentes”, a métrica  $WindowDiff$  concentra-se no número de fronteiras que se encontram entre as extremidades da janela deslizante de dimensão  $k$ .

#### 4.1.1 Algoritmo de Segmentação no Tempo

Os ensaios começaram por determinar o comportamento do algoritmo em função de  $TimeThreshold$ , para se compreender o seu comportamento quando comparado com o algoritmo ASE. Para isso, o algoritmo foi executado sobre o conjunto de coleções de fotografias utilizadas no âmbito deste trabalho, listadas nas Tabelas A.1 e A.2.

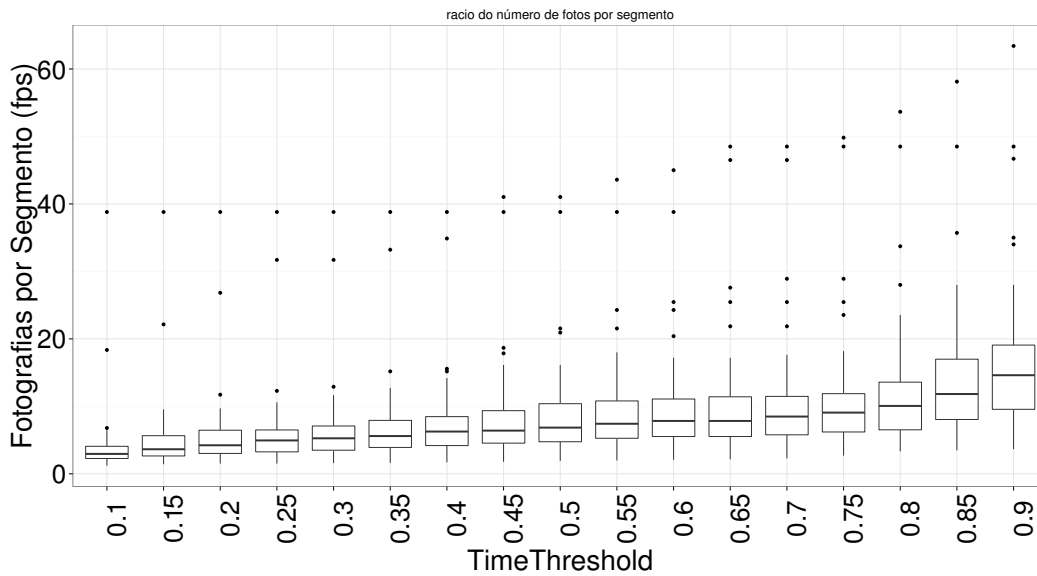


Figura 4.4: O número de fotografias por segmento (fps) gerado pelo algoritmo de segmentação no tempo (AST) varia em função do parâmetro *TimeThreshold*.

Na Figura 4.4, usa-se o número de fps como uma medida normalizada do número de segmentos gerados pelos algoritmos, já que se incluíram nesta análise diversas colecções de fotografias cada qual composta por um diferente número de fotografias.

Ainda na mesma Figura, observando a evolução do número de fotografias por segmento (fps) em função do parâmetro *TimeThreshold*, verifica-se que 0.5 é o melhor valor para este parâmetro.

O motivo pelo qual se considera 0.5 o melhor valor para o parâmetro *TimeThreshold*, prende-se com o facto de i) não ocorrer nenhuma variação drástica no gráfico da Figura 4.4, ii) ser este o valor recomendado em [7] e iii) 0.5 é o valor que permite uma maior latitude de opções para o utilizador, se este pretender observar os efeitos que resultam da variação deste parâmetro.

Assumindo que o número ideal de segmentos gerados depende das preferências do utilizador, pretende-se que o utilizador possa variar ao máximo o valor parametrizado de forma a provocar resultados visivelmente diferentes.

Assim, para as seguintes análises, 0.5 será o valor atribuído ao parâmetro *TimeThreshold* tanto no algoritmo de segmentação no tempo (AST) como no algoritmo de segmentação no tempo e espaço (ASE).

### 4.1.2 Algoritmos de Detecção de Coordenadas Erradas

No âmbito dos testes de algoritmos de detecção de coordenadas erradas, foram apenas considerados os casos que usaram o DP.

O motivo que levou a esta restrição prende-se com o facto de neste contexto, os restantes conjuntos de casos serem redundantes: conforme se descreveu na secção 3.3 “Detecção de Coordenadas Erradas” a etapa de detecção de coordenadas erradas precede o ponto do algoritmo onde a distinção entre DP e DAM é relevante.

Os 8190 casos de teste que daqui resultaram foram divididos em 10 diferentes combinações de algoritmo-parâmetro com 819 iterações cada uma:

- 2457 testes com o algoritmo IQR nos quais usaram-se 3 valores diferentes para o parâmetro K: {1; 1,5; 2};
- 2457 testes com o algoritmo LOF onde se usaram 3 valores diferentes para o respectivo parâmetro: {1; 2; 4};
- 3276 testes com o algoritmo baseado em velocidade nos quais foram parametrizados 4 valores diferentes {1:0; 1:1; 2:2; 3:3};

Através de uma inspecção manual, verificou-se que as coordenadas que se podem considerar decididamente incorrectas constituem um número muito reduzido do total de coordenadas avaliadas. Por este motivo, no âmbito deste trabalho, convencionou-se que quanto mais elevada for a percentagem de coordenadas consideradas correctas, melhor se poderá considerar a qualidade de um dado algoritmo de detecção.

Na Figura 4.5 apresenta-se a distribuição de coordenadas consideradas correctas em função do algoritmo de detecção e respectivo parâmetro.

Conforme se pode constatar, verifica-se que o método baseado no *Inter-Quartile Range* é superior ao método baseado no *Local Outlier Factor*. Esta diferença justifica-se com o facto da maioria das colecções possuir pelo menos 2 fotografias tiradas precisamente no mesmo local, ou seja distam 0 entre si. Esta distância nula introduz anomalias no cálculo do *Local Outlier Factor* o que resulta numa fraca prestação na detecção de coordenadas erradas, por parte deste algoritmo.

Segundo a mesma Figura, o método de detecção baseado na velocidade é o método que avalia a maior quantidade de coordenadas como sendo correctas. Pensa-se que este resultado se deve ao facto deste método avaliar os dados no contexto

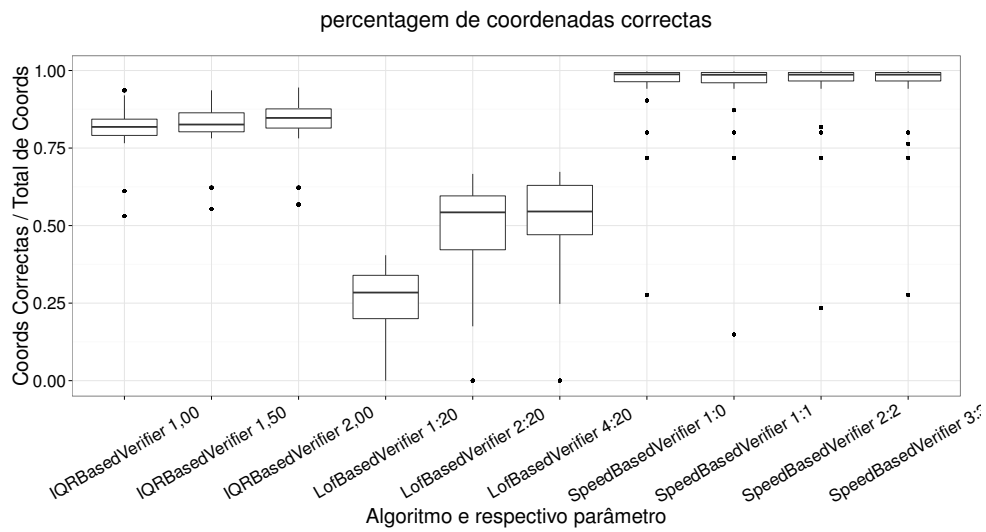


Figura 4.5: Comportamento dos algoritmos de detecção de coordenadas.

do domínio do problema, ao contrário dos outros dois métodos. Com este método, uma coordenada está errada apenas quando viola alguns pressupostos derivadas da semântica do problema, conforme se viu na secção 3.3.1 “Detecção com Base na Velocidade”.

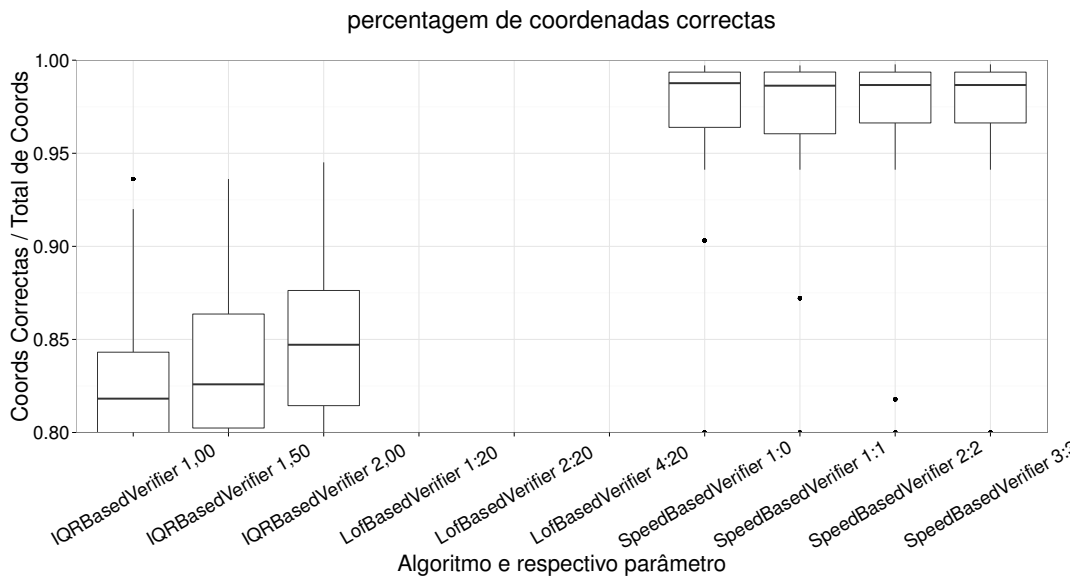


Figura 4.6: Detalhe da Figura 4.5, restringida a  $[0.8; 1]$ .

A Figura 4.6 incide sobre o intervalo  $[0.8; 1]$  da Figura 4.5 e permite-nos observar a evolução do comportamento do método de detecção baseado em *Interquartile Range*, já visto na secção 3.3.2. Constata-se que o número de fotografias com coordenadas consideradas correctas sobe à medida que o factor IQR sobe de 1,

para 1.5 e finalmente para 2. Apesar desta progressão, este algoritmo nunca chega a atingir o nível verificado para o algoritmo baseado na velocidade.

No âmbito do algoritmo baseado na velocidade, o parâmetro 1:1 (que considera a média das distâncias do ponto a ser escrutinado ao ponto anterior e ao ponto posterior) foi a configuração que mais coordenadas erradas encontrou, por uma pequena diferença. Por este motivo, será este o método e parâmetro utilizado nas seguintes análises, sempre que o seu uso se justifique.

### 4.1.3 Algoritmo de Segmentação no Tempo e Espaço

Uma forma de determinar se o parâmetro *SpaceThreshold* está a ter algum efeito na segmentação é verificar o modo como o número de segmentos gerados pelo algoritmo no espaço (ASE) varia em relação ao número de segmentos gerados pelo algoritmo no tempo (AST).

A Figura 4.7 indica a forma como o rácio entre esses dois valores;  $NS_{AST}/NS_{ASE}$  evolui em função de *SpaceThreshold*. Note-se que para *SpaceThreshold* = 0 o  $NS_{AST}$  é superior ao  $NS_{ASE}$ , e à medida que se incrementa *SpaceThreshold* essa proporção desce. Esta descida é razoável, uma vez que à medida que o *SpaceThreshold* sobe, o ASE cria mais segmentos. Em ambos os casos (DP e DAM) usa-se o valor 0.5 para o parâmetro *TimeThreshold*, tal como foi previamente estipulado.

Testaram-se duas métricas estatísticas: o Desvio Padrão (DP, na Figura 4.7(a)) e o Desvio Absoluto Mediano (DAM, na Figura 4.7(b)). Como se pode verificar, o DP apresenta-se mais sensível à variação do parâmetro *SpaceThreshold*. Na Figura 4.7(b)) a mediana atinge o valor 1 quando *SpaceThreshold*=10 e aí permanece até *SpaceThreshold*=95. Na mesma Figura, os *boxplots* permanecem relativamente inalterados em [15; 90]. Por este motivo as análises estatísticas seguintes restringiram-se aos casos de testes que usaram o DP; um universo de pouco mais de 9 mil testes.

A Figura 4.7(a) indica-nos que para *SpaceThresholds* superiores a 30 a mediana do rácio  $NS_{AST}/NS_{ASE}$  estabiliza e deixa de ser influenciada por aquela variável. Verifica-se apenas uma pequena descida na mediana em [45; 55], mas esta alteração é a única excepção que se verifica nos *boxplots* que se mantêm relativamente estáveis. Só se volta a verificar uma alteração quando *SpaceThreshold* atinge 100. Este efeito de estabilização indicia uma má decisão na forma como se calibrou o patamar *ReferenceSplitPct* descrito na secção 3.2.1 “Parâmetro de Calibração do Algoritmo” ou então a existência de um limite inerente aos dados contidos nas

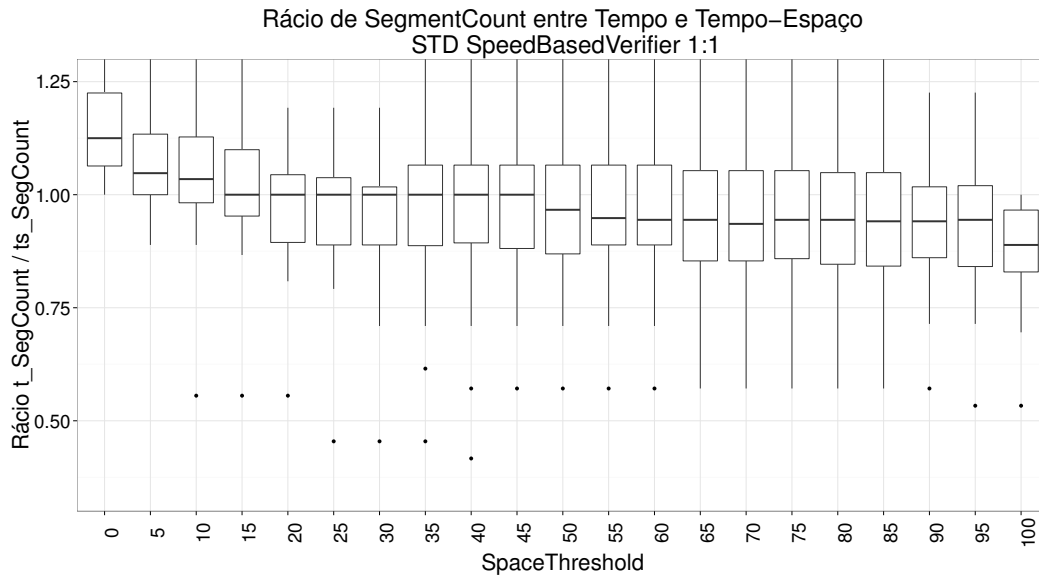
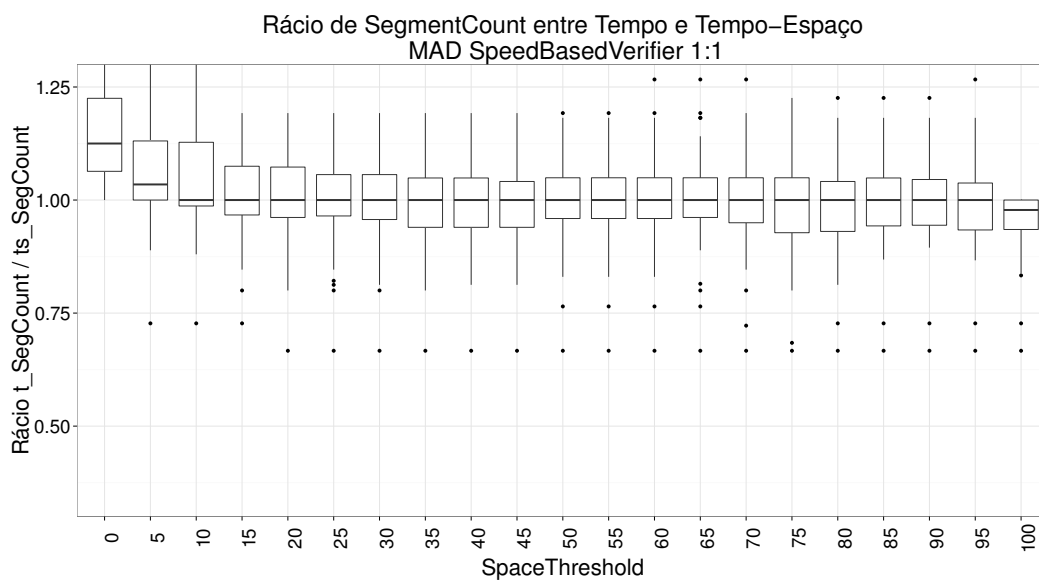
(a) Efeitos de *SpaceThreshold* sobre o rácio  $NS_{AST}/NS_{ASE}$  quando o ASE usa DP(b) Efeitos de *SpaceThreshold* sobre o rácio  $NS_{AST}/NS_{ASE}$  quando o ASE usa DAM

Figura 4.7: Evolução do rácio do número de segmentos (NS) gerados pelos dois algoritmos ( $NS_{AST}/NS_{ASE}$ ), quando o ASE usa o DP (4.7(a)) ou o DAM (4.7(b)). Em ambos os casos a variação ocorre em função de *SpaceThreshold*.

fotografias das colecções usadas.

Posteriormente, verifica-se que este efeito se verifica em vários outros gráficos onde se pretenda ver a forma como uma dada métrica varia em função de *SpaceThreshold*: a métrica evolui até *SpaceThreshold*=30 e depois deste ponto o gráfico estabiliza.

### Quantidade de Operações de Fraccionamento Realizadas com Sucesso

Na Figura 4.8 é apresentada a distribuição do resultado da operação de fraccionamento.

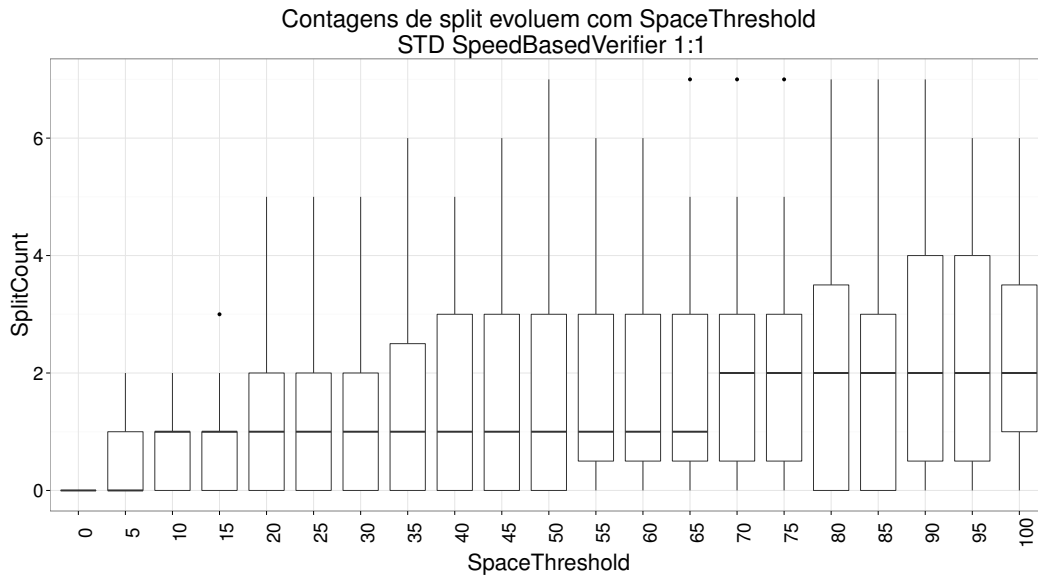
O número de vezes que o resultado da operação de fraccionamento foi considerado admissível variou entre 0 e 7 (4.8(a)), resultados esses que assim sendo acabaram por ser incluídos na segmentação; enquanto o número de resultados descartados variou entre 0 e 46 maus casos (4.8(b)).

Os efeitos do parâmetro *SpaceThreshold* fazem-se sentir no número de fraccionamentos realizados. Tendo sido projectado para promover o aumento do número de segmentos à medida que sobe, o número de fraccionamentos aceitáveis – o *splitcount* – sobe à medida que este parâmetro sobe, como seria de esperar. Por outro lado, o mesmo parâmetro não aparenta ter uma influência perceptível no número de fraccionamentos rejeitados – o *nosplitcount* – sendo a descida mais visível desta métrica observável nos valores extremos máximos.

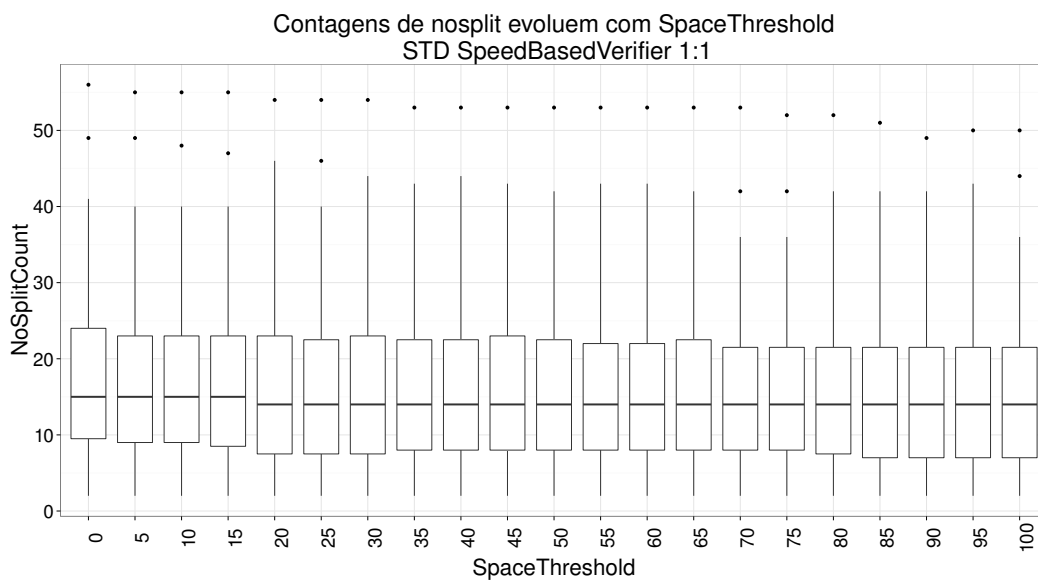
Verifica-se também que a variação no *splitcount* se faz sentir mais acentuadamente quando o *SpaceThreshold* se encontra em [0; 40]. A partir do ponto *SpaceThreshold*=40, a mediana de *splitcount* estabiliza em 1 e só sobe para 2 quando *SpaceThreshold* atinje o valor 70.

Na Figura 4.8(b) temos a situação contrária: o número de vezes que a operação de fraccionamento resultou num conjunto de segmentos considerados pior que o segmento que lhes deu origem – o *nosplitcount*. Neste segundo caso os segmentos resultantes são descartados e inclui-se na segmentação do ASE o segmento originalmente criado pelo AST. O número de resultados descartados aparenta ser insensível ao parâmetro de calibração *SpaceThreshold* uma vez que a mediana ronda sempre um valor ligeiramente abaixo de 15. Esta estabilidade mantém-se praticamente constante durante toda a evolução do *SpaceThreshold*, sendo apenas perceptível uma pequena descida quando o *SpaceThreshold* assume valores em [15; 20].

O facto de ocorrer um aumento do *splitcount* sem que ocorra uma alteração perceptível do *nosplitcount* pode ser explicado tendo em conta que *SpaceThreshold*



(a) O fraccionamento atingiu entre 0 e 7 bons resultados (*splitcount*) por teste.



(b) A quantidade de resultados rejeitados variou entre 0 e 46 (*nosplitcount*) por cada teste efectuado.

Figura 4.8: Quantidade de fraccionamentos realizados com sucesso, em 4.8(a) e o número de fraccionamentos rejeitados, em 4.8(b).

regula o nível de exigência no que diz respeito à aceitação de um dado resultado da Operação de Fraccionamento tal como é indicado na secção 3.2.1 “Parâmetro de Calibração do Algoritmo”. Este parâmetro regula ainda o número de segmentos gerados por esta Operação através do patamar *SplitSensitivity*. À medida que *SpaceThreshold* sobe, a Operação de Fraccionamento realizada sobre um dado segmento pode produzir um ou mais segmentos. Se for produzido apenas um segmento não há nada a decidir. Mas se *SpaceThreshold* for alto a ponto de produzir 2 ou mais segmentos, há a necessidade de determinar qual das duas opções é melhor: o segmento original – incrementando assim o *nosplitcount* – ou os novos segmentos resultantes da Operação – algo que incrementa o *splitcount*. Tendo em conta os dois gráficos, pode dizer-se que os fraccionamentos provocados pelo aumento do *SpaceThreshold* são por norma aceites, o que indica alguma redundância entre os dois patamares *ReferenceSplitPct* e *SplitSensitivity*.

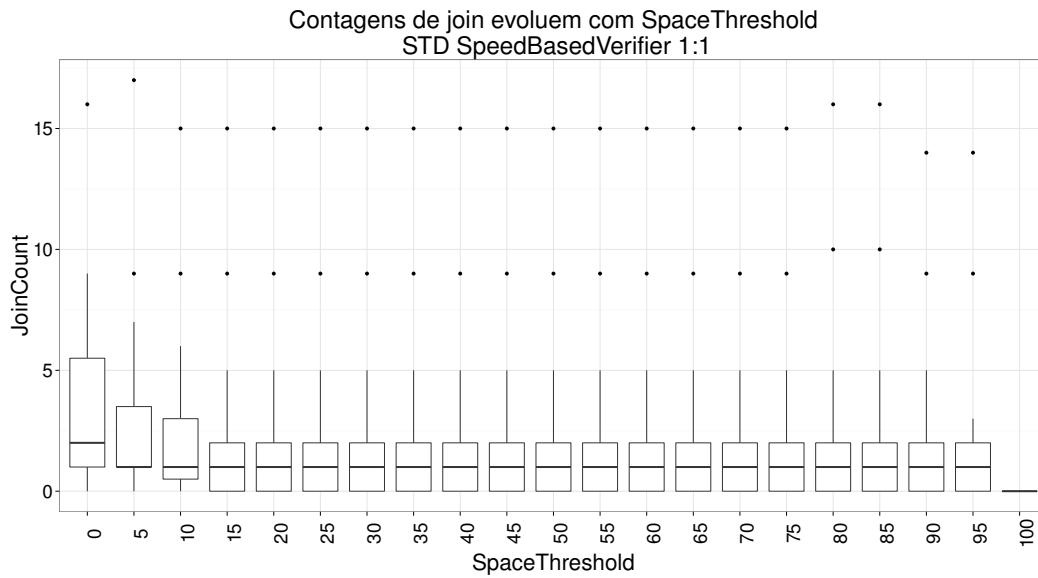
### Quantidade de Operações de Junção Realizadas com Sucesso

Na Figura 4.9(a) é indicado o número de vezes que o segmento resultante da operação de junção foi considerado melhor que os segmentos que lhe deram origem.

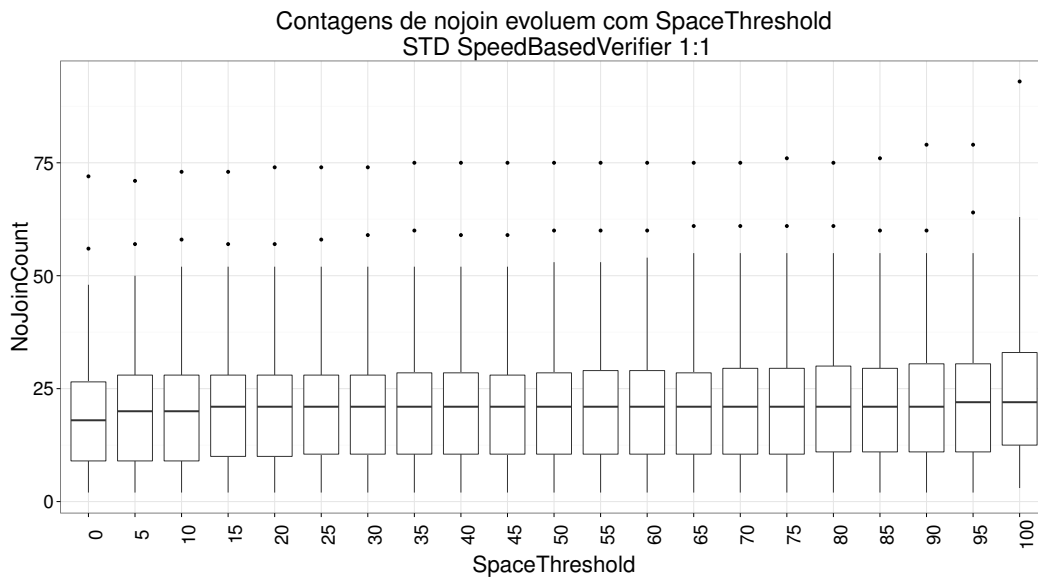
Nesta Figura apresenta-se a distribuição do número de junções realizadas. Os valores indicam a quantidade de resultados de junções que foram considerados aceitáveis e inseridos na segmentação – o *joincount*, ou em alternativa, descartados por não se enquadrarem dentro dos limites parametrizados – o *nojoincount*. Conforme se constata, a quantidade de junções que são aproveitadas (gráfico superior) é baixa. A mediana de *joincount* é 1 quando *SpaceThreshold* é igual ou superior a 5. A sensibilidade de *joincount* a variações de *SpaceThreshold* é muito baixa: a mediana de *joincount* baixa apenas quando *SpaceThreshold* é igual a {5; 100}. À semelhança do que se constatou para *ReferenceSplitPct*, é possível que estes baixos valores se devam a uma má calibração do patamar *ReferenceJoinPct* descrito na secção 3.2.1 “Parâmetro de Calibração do Algoritmo”.

Tanto a Figura 4.8(a) como a Figura 4.9(a) permitem-nos verificar que o número de operações de sucesso é francamente inferior ao número de operações cujo resultado não foi aceite (Figuras 4.8(b) e 4.9(b)). A diferença é quase uma ordem de magnitude: os casos em que as operações tiveram sucesso medem-se na casa das unidades enquanto o número de casos em que as operações não tiveram sucesso medem-se em dezenas.

Pensa-se que o motivo que provoca esta diferença entre os bons e os maus resultados seja a possibilidade de efectivamente haver redundância entre a informação temporal e a informação geográfica. Tendo em conta o tipo de fotografias



(a) Os resultados da junção variaram entre 0 e 2 casos com bons resultados.



(b) O caso oposto variou entre 0 e 40 casos cujos resultados foram rejeitados.

Figura 4.9: Quantidade de junções realizadas com (4.9(a)) e sem (4.9(b)) sucesso.

analisadas neste trabalho – fotografias que reflectem a vida social do utilizador – variações nas diferenças temporais são normalmente acompanhadas de perto por variações nas distâncias percorridas. Por este motivo, o padrão que os dados reflectem no domínio do tempo, provavelmente também se reflectirá no domínio do espaço.

### Métricas de Avaliação de Semelhança entre Segmentos

As quatro métricas utilizadas no âmbito deste trabalho — *RMiss* e *Rfa* descritas em [13] e *Pk* e *WindowDiff* descritas em [40] — medem a diferença entre a segmentação produzida pelo algoritmo de segmentação no tempo e espaço (ASE) e a respectiva segmentação que lhe deu origem, produzida pelo algoritmo de segmentação no tempo (AST).

O algoritmo das métricas *RMiss* e *Rfa* ao comparar as duas segmentações ( $S_t$  e  $S_{te}$ ), tenta estabelecer uma relação de equivalência entre os segmentos das duas segmentações – essencialmente tenta determinar de que forma é que os segmentos de  $S_t$  se “moveram” dentro de  $S_{te}$ . Com base nesta relação, *RMiss* determina quantas fotografias estão em falta nos vários segmentos de  $S_{te}$ , e *Rfa* determina quantas fotografias foram indevidamente incluídas nos vários segmentos de  $S_{te}$ , quando comparadas com a segmentação original  $S_t$ .

Todos os testes apresentados de seguida foram realizados para o conjunto de colecções de fotografias utilizado no âmbito deste trabalho (descritos nas Tabelas A.1 e A.2), com um *timethreshold*=0.5, sendo o Desvio Padrão o método estatístico utilizado. Para as Figuras 4.10, 4.12, 4.14 e 4.16 o método de detecção de coordenadas é o algoritmo baseado em velocidade, parametrizado com “1:1” que considera apenas a coordenada anterior e a coordenada posterior à coordenada avaliada.

#### Métrica *RMiss*

A métrica *RMiss* indica a quantidade de fotografias que, erradamente, não foram incluídas num determinado segmento produzido pelo ASE, quando comparadas com a segmentação original gerada pelo AST. É expressa sob a forma de uma percentagem do número total de fotografias contidas na segmentação.

Como se pode ver na Figura 4.10, é apenas quando o parâmetro *SpaceThreshold* é igual a 0 que a métrica *RMiss* considera que nenhum segmento produzido pelo ASE se “esqueceu” de incluir qualquer fotografia. A partir de *SpaceThreshold*=5, o *RMiss* indica que os segmentos produzidos pelo ASE omitiram várias fotografias que deveriam figurar no segmento avaliado – isto quando comparado com a segmentação original, produzida pelo AST.

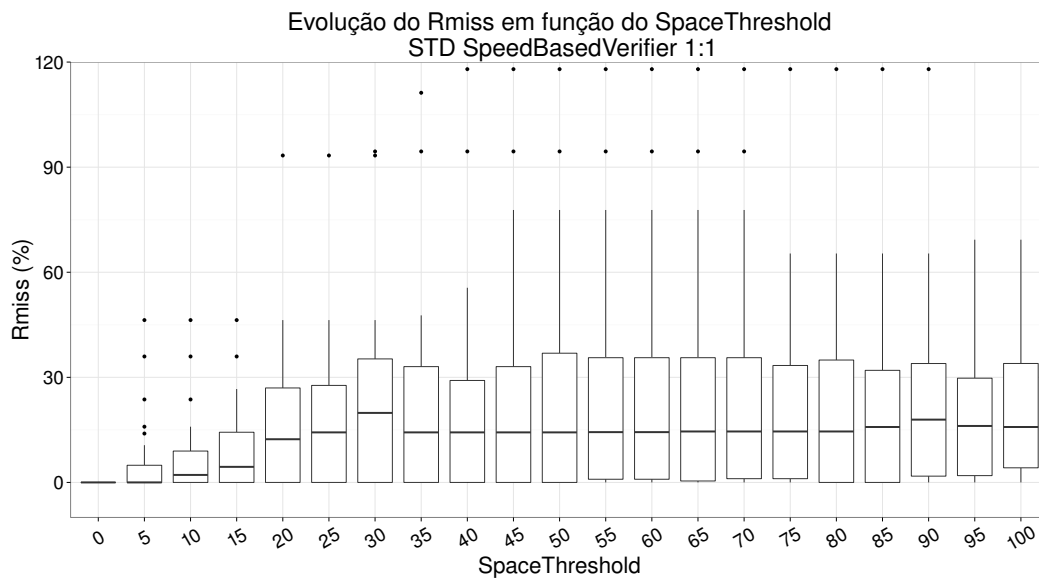


Figura 4.10: Evolução da métrica *RMiss* em função do parâmetro *SpaceThreshold*.

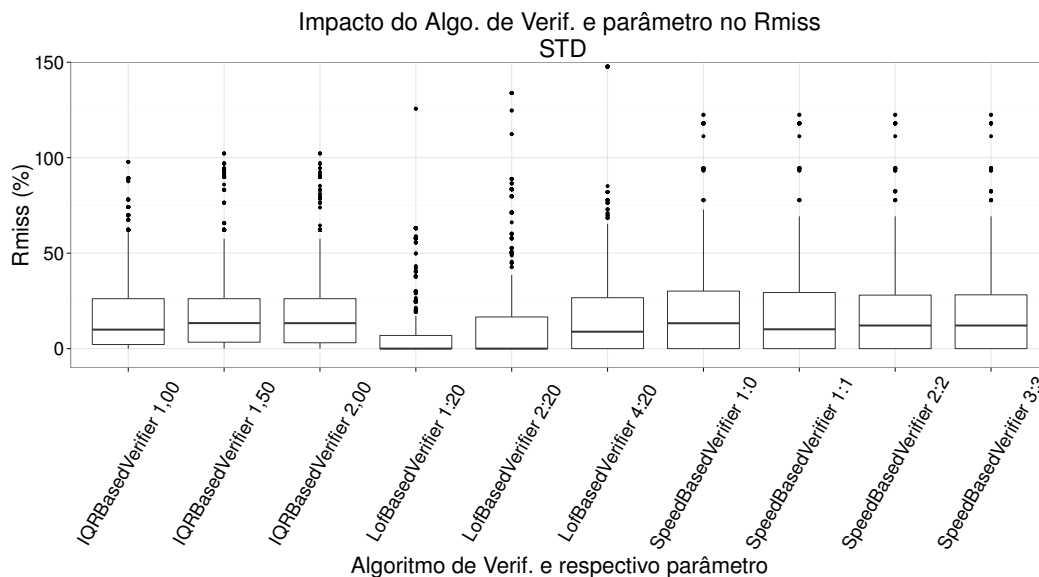


Figura 4.11: Evolução da métrica *RMiss*, em função dos diversos algoritmos de detecção de coordenadas e respectivos parâmetros.

É possível que o facto de  $R_{Miss}$  ser 0 para  $SpaceThreshold=0$  indique uma quantidade mínima ou nula de alterações realizadas à segmentação original. Como se pode ver na Figura 4.8(a), quando o parâmetro  $SpaceThreshold$  é igual a 0 o número de fraccionamentos também é igual a 0. Isto garante que o segmento poderá ter fotografias a mais – caso se tenham realizado com sucesso junções de segmentos – mas nunca terá fotografias em falta.

A partir de  $SpaceThreshold=30$  a mediana desta métrica indica-nos que, cerca de 20% das fotografias presentes nos segmentos da segmentação gerada pelo AST foram movidas para fora do segmento equivalente presente na segmentação gerada pelo ASE.

Também aqui se verifica um acompanhamento da Figura 4.7: tanto o  $R_{Miss}$  como o  $splitcount$  estabilizam para um  $SpaceThreshold$  superior a 30.

As alterações de algoritmo de detecção de coordenadas erradas ou respectivo parâmetro revelaram ter pouco impacto na métrica  $R_{Miss}$ . A excepção é o método de detecção baseado em *Local Outlier Factor*, mas já se verificou pela Figura 4.5 que esse foi precisamente o algoritmo de detecção com pior desempenho.

### **Métrica $R_{fa}$**

A métrica  $R_{fa}$  indica a quantidade de fotografias que foram indevidamente incluídas num determinado segmento produzido pelo ASE, quando comparado com os segmentos existentes na segmentação original gerada pelo AST.

A mediana desta métrica, apresentada na Figura 4.12, atinge os valores mais baixos quando  $SpaceThreshold \in \{10, 15\}$ , o que faz sentido se tivermos em conta a Figura 4.7(a). Esta Figura diz-nos que quando  $SpaceThreshold \in \{10, 15\}$ , verifica-se a menor diferença entre as quantidades de segmentos existente nas segmentações geradas pelo AST e pelo ASE. Assumindo que, quanto menor for esta diferença quantitativa, menores serão as diferenças em termos qualitativos, é razoável admitir que seja este o ponto onde há menor diferença entre as duas segmentações e consequentemente, a menor quantidade de fotografias indevidamente incluídas em segmentos produzidos pelo ASE.

No entanto, seria expectável que na Figura 4.10 a métrica  $R_{Miss}$  atingisse o seu ponto mais baixo quando  $SpaceThreshold=15$ , algo que não se verifica. Entre as possíveis explicações, contam-se um possível erro na implementação das métricas em questão ou então o facto desta ideia pura e simplesmente não poder ser aplicada àquela métrica.

À semelhança da métrica  $R_{Miss}$ , também a métrica  $R_{fa}$  revela-se pouco sensível a

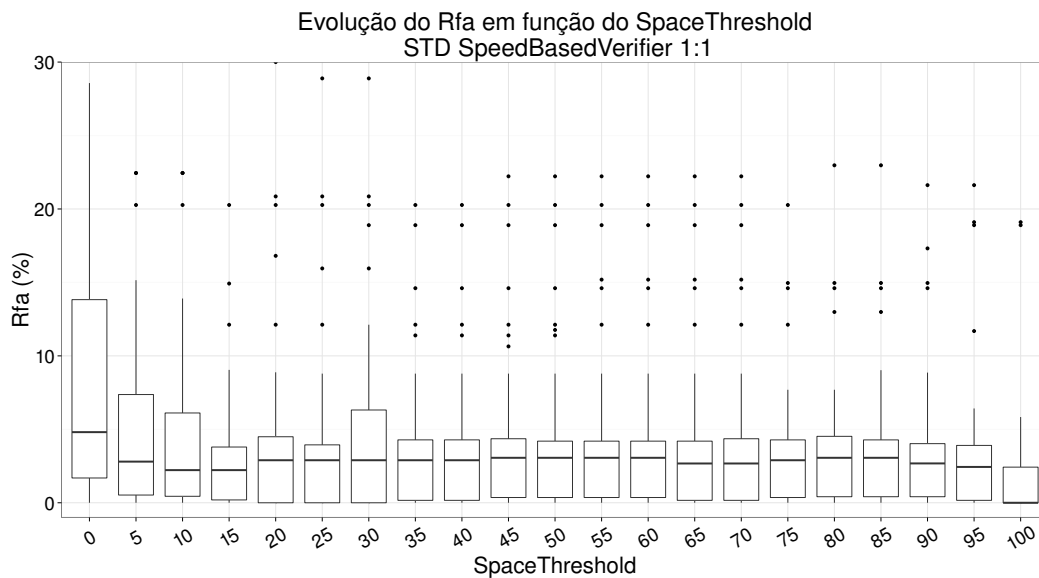


Figura 4.12: Evolução da métrica *Rfa* em função do parâmetro *SpaceThreshold*.

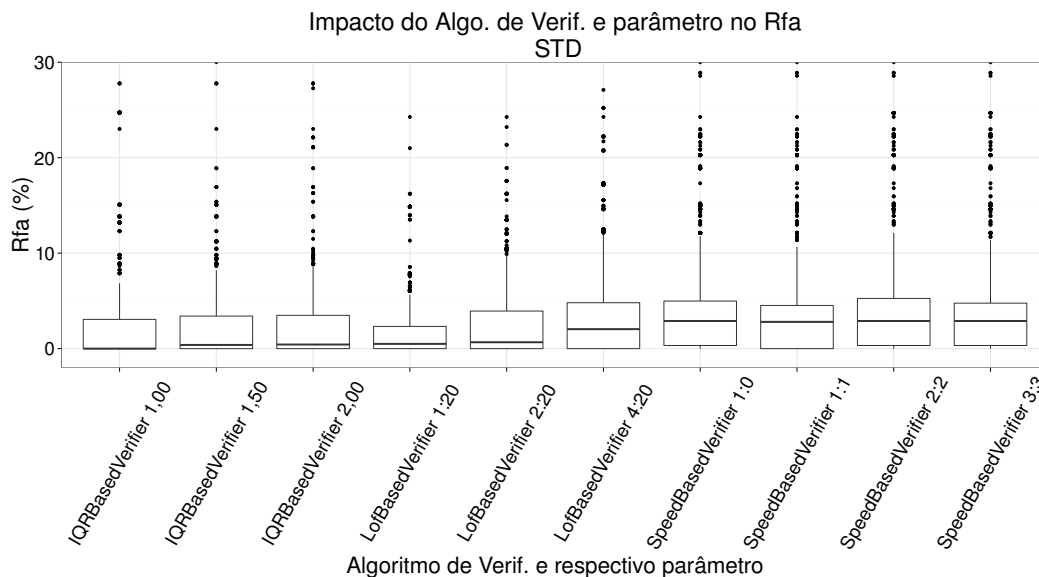


Figura 4.13: Evolução da métrica *Rfa*, em função dos diversos algoritmos de detecção de coordenadas erradas e respectivos parâmetros.

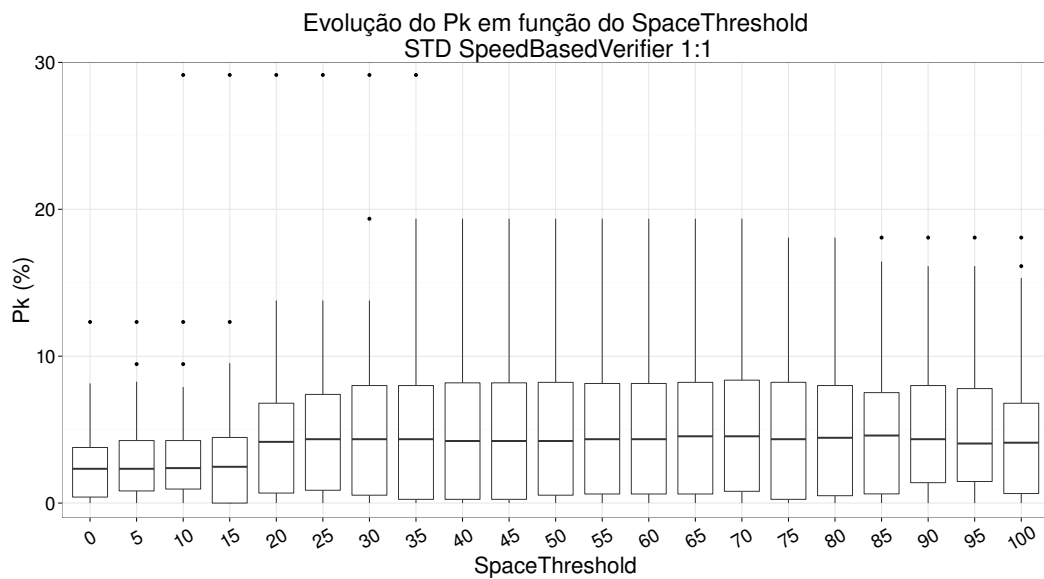


Figura 4.14: Evolução da métrica  $Pk$  em função do parâmetro  $SpaceThreshold$ .

alterações de algoritmo de detecção de coordenadas erradas ou respectivo parâmetro, na Figura 4.13. Mais uma vez, a exceção é o método de detecção baseado em *Local Outlier Factor*.

### Métrica $Pk$

A métrica  $Pk$  foi concebida para, dadas duas segmentações diferentes mas relacionadas, comparar a posição das fronteiras dos segmentos visíveis a partir de uma janela deslizante de dimensão  $k$ . Pretende-se determinar se os segmentos indicados pelas extremidades da janela deslizante de dimensão  $k$  são segmentos “diferentes” ou o “mesmo” segmento. Deslizando esta janela em simultâneo nas duas segmentações, contabiliza-se mais uma diferença sempre que os resultados para as duas segmentações estão em desacordo: sempre que seja indicado que numa das segmentações os segmentos indicados são “diferentes” e na outra são o “mesmo” segmento. Desta forma, pretende-se atribuir diferentes penalizações a fronteiras ligeiramente desviadas ou inexistentes.

Também esta métrica é expressa sob a forma de uma percentagem do número total de fotografias contidas na segmentação.

A Figura 4.14 indica-nos que à medida que  $SpaceThreshold$  sobe – à medida que o número de segmentos criados sobe – aumenta a diferença entre as segmentações geradas por cada um dos dois algoritmos, AST e ASE. Este aumento verifica-se até ao ponto em que  $SpaceThreshold=30$ . A partir daqui a mediana começa a rondar os 5% e aí permanece até  $SpaceThreshold=100$ . Este comportamento de

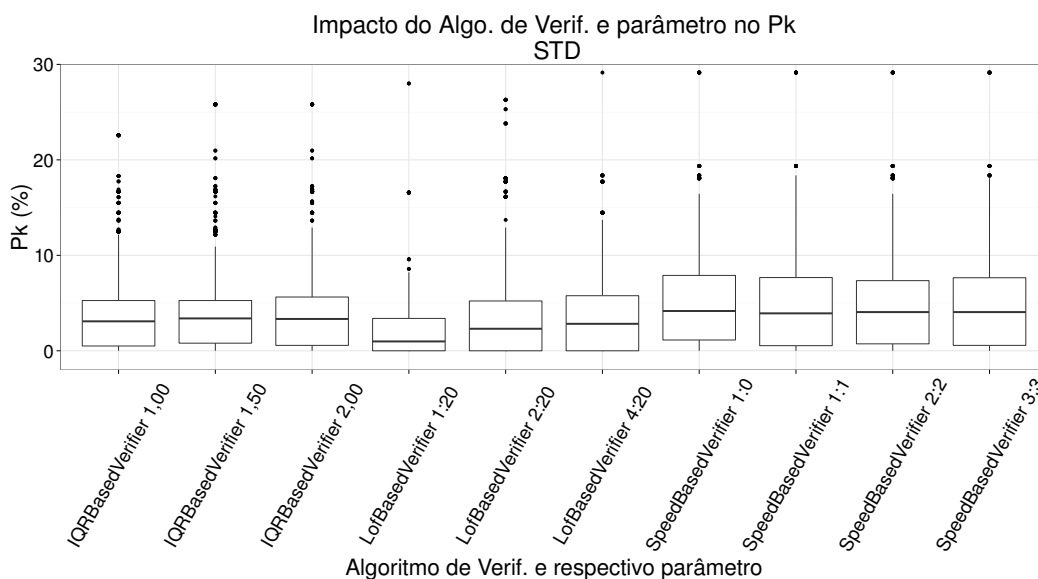


Figura 4.15: Evolução da métrica  $Pk$ , em função dos diversos algoritmos de detecção de coordenadas erradas e respectivos parâmetros.

estabilização a partir de  $SpaceThreshold=30$  tem sido observado em várias figuras desde a Figura 4.7(a), nomeadamente as Figuras 4.10 e 4.12.

Na Figura 4.15, a métrica  $Pk$  demonstra um comportamento semelhante às duas métricas anteriores –  $RMiss$  na Figura 4.11 e  $Rfa$  na Figura 4.13 – no que diz respeito aos diversos algoritmos de detecção de coordenadas.

### Métrica *WindowDiff*

As métricas  $Pk$  e  $WindowDiff$  estão relacionadas entre si e por isso têm um comportamento semelhante. Pode-se constatar a semelhança entre a Figura 4.16, que mostra a evolução do  $WindowDiff$ , e a Figura 4.14.

A principal diferença entre as duas métricas é o facto de, enquanto a métrica  $Pk$  determina se os segmentos indicados pelas extremidades da janela deslizante de dimensão  $k$  são “diferentes” ou o “mesmo” segmento, na métrica  $WindowDiff$  são contabilizadas o número de fronteiras de segmentos que se podem encontrar entre as extremidades da janela deslizante de dimensão  $k$ . É contabilizada uma diferença sempre que o número de fronteiras for diferente para as duas segmentações.

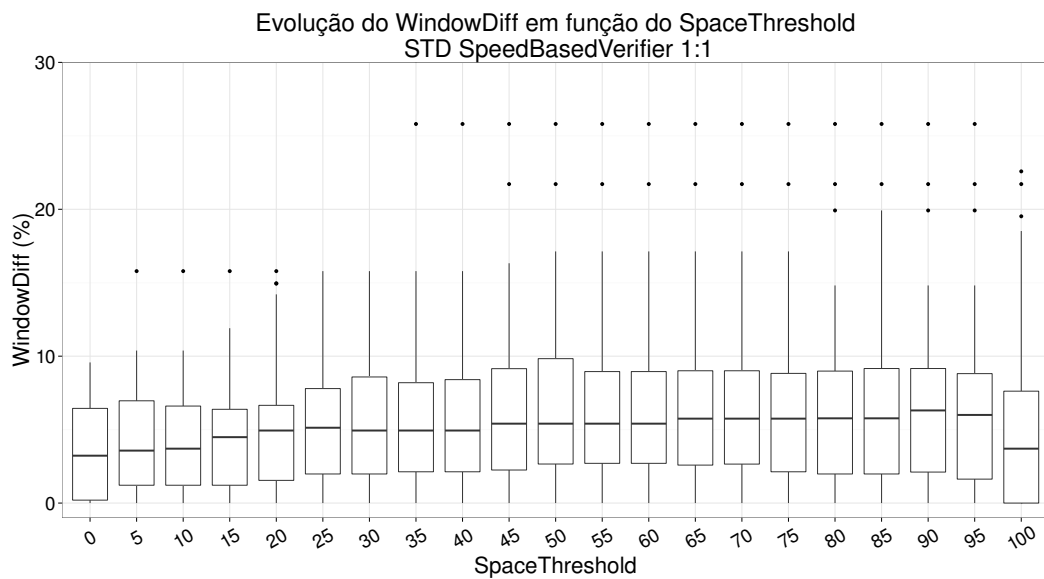


Figura 4.16: Evolução da métrica *WindowDiff* em função do parâmetro *SpaceThreshold*.

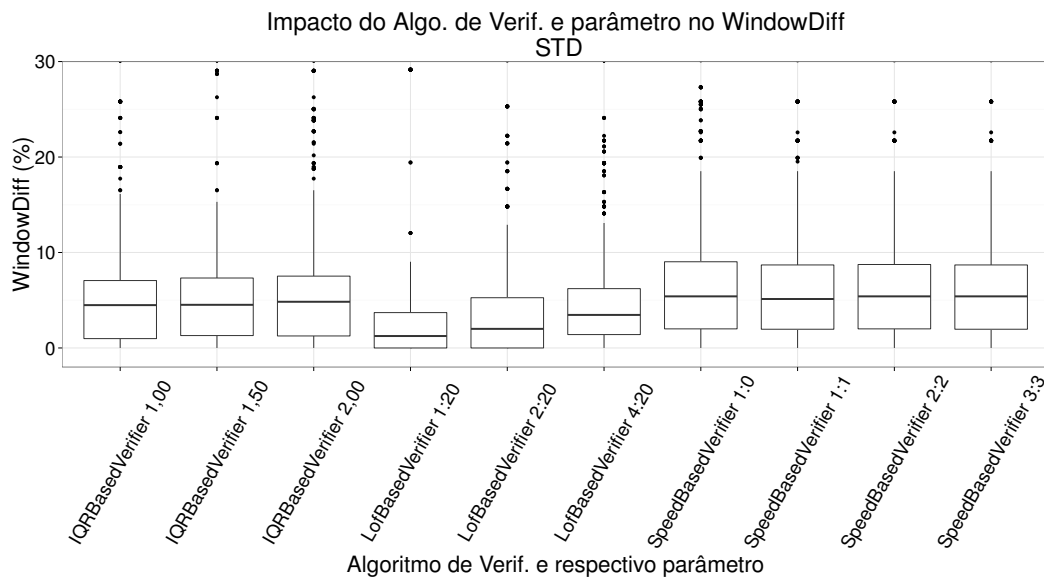


Figura 4.17: Evolução da métrica *WindowDiff*, em função dos diversos algoritmos de detecção de coordenadas erradas e respectivos parâmetros.

Relativamente aos algoritmos de detecção de coordenadas, a Figura 4.17 mostra a forma como a métrica *WindowDiff* apresenta um comportamento semelhante ao apresentado na Figura 4.15, algo que faz sentido uma vez que esta métrica implica apenas a realização de uma pequena modificação à métrica *Pk*.

#### 4.1.4 Discussão

A análise gráfica dos dados permite-nos observar um comportamento de estabilização para valores de *SpaceThreshold* superiores a 30 nas Figuras 4.7(a), 4.10, 4.12, 4.14 e 4.16. Presume-se que este comportamento seja consequência de i) uma má calibração do patamar *ReferenceSplitPct* descrito na secção 3.2.1 “Parâmetro de Calibração do Algoritmo”, ii) do facto do ASE ser incapaz de realizar mais operações, devido às restrições impostas ao algoritmo conforme descritas na secção 3.2.5 “Restrições do Algoritmo de Segmentação no Tempo (AST)”, iii) ou então devido à natureza do próprio conjunto de dados fornecidos pelas colecções. Cada colecção tem uma quantidade fixa de eventos sociais representados nas suas fotografias. Se a partir de certo valor de *SpaceThreshold* o algoritmo criar um segmento por cada um desses eventos sociais, idealmente essa quantidade de segmentos deveria permanecer inalterada, independentemente de subidas adicionais ao valor de *SpaceThreshold*.

Assumindo que aquela estabilização é provocada pela natureza do próprio conjunto de dados é possível que, recalibrando os patamares, apenas se consiga aumentar o número de junções efectuadas. Se os patamares forem recalibrados de forma a provocar um *shift* no eixo do *SpaceThreshold* seria possível confirmar se o algoritmo consegue realizar mais junções. Sabendo que em todas as segmentações criadas pelo AST foram criadas no mínimo 5 segmentos e tendo em conta a Figura 4.7(a) sabe-se que a segmentação produzida pelo ASE nunca cria apenas um único segmento para toda a segmentação, pelo que é legítimo pensar que os patamares podem ser modificados de forma a que os valores da Figura 4.9(a) se aproximem mais dos valores apresentados na Figura 4.8(a).

## 4.2 Testes com Utilizadores

Tendo em conta o objectivo enunciado anteriormente — determinar se a informação geográfica potencialmente disponível numa fotografia é uma mais valia no contexto de um algoritmo de segmentação automática — foram realizados testes com utilizadores.

### Resumo dos Testes com Utilizadores

Foi desenvolvida uma aplicação *web* dedicada à realização de testes. Cada teste é constituído por um questionário ao utilizador e a apresentação de uma série de páginas *web*, cada uma com uma colecção de fotografias segmentada de duas

formas diferentes em simultâneo. As duas segmentações produzidas por cada um dos dois algoritmos (AST e ASE) são colocadas aleatoriamente à esquerda ou à direita da página *web*, com uma probabilidade de 50%. Os utilizadores foram divididos em dois grupos: o primeiro avalia as várias colecções de fotografias por uma certa ordem de apresentação, e para o segundo a ordem de apresentação é a inversa. Esta decisão é realizada uma vez por cada participante, também com uma probabilidade de 50%.

### Parametrização dos Testes com Utilizadores

Os testes realizados contaram com a participação de dezoito (18) utilizadores para decidirem qual a segmentação da sua preferência: a segmentação do Algoritmo de Segmentação no Tempo (AST) – ou a versão feita pelo Algoritmo de Segmentação no Tempo e Espaço (ASE). Foi pedido ao utilizador que, para cada colecção de fotografias, exprimisse a sua preferência sobre as duas segmentações. Cada utilizador teve a oportunidade de apreciar sete (7) colecções de fotografias. A lista de colecções de imagens utilizadas é a indicada na Tabela 4.2.

Colecções	Ordem de Apresentação	Nº de Fotografias
Alpy2011	1º	97
Atlantik2008	2º	90
Gorce2009	3º	88
Marroco2013	4º	75
Chipre2011	5º	72
Macau2008	6º	44
Christmas2008	7º	40

Tabela 4.2: As colecções foram apresentadas a cada um dos utilizadores pela ordem indicada, ou pela ordem inversa. Consistem num subconjunto das colecções apresentadas nas Tabelas A.1 e A.2.

Cada uma das colecções foi segmentada sempre com os mesmos parâmetros, independentemente do utilizador. Os dois algoritmos foram parametrizados conforme se indica na Tabela 4.3.

	AST	ASE
<i>TimeThreshold</i>	0.5	0.5
<i>SpaceThreshold</i>	N/A	0.8
<b>Cálculos Estatísticos</b>	N/A	Desvio Padrão
<b>Detecção de Coordenadas Erradas</b>	N/A	Velocidade (1:1)

Tabela 4.3: Parametrização utilizada em cada um dos dois algoritmos, durante os testes com utilizadores.

Para o algoritmo AST optou-se por seguir as indicações que se encontram em [7], segundo as quais 0,5 é um valor por omissão aceitável para o parâmetro *TimeThreshold*. Quanto ao valor 0,8 para o parâmetro *SpaceThreshold* foi o valor escolhido quando, durante os testes iniciais ao algoritmo de segmentação no tempo e espaço (ASE), se constatou que o algoritmo realizava muito poucas junções e, dependendo da colecção, um baixo número de fraccionamentos. Assim sendo, assumiu-se que havia a possibilidade dos utilizadores não notarem diferenças entre as duas segmentações – produzidas pelo AST e pelo ASE – e assim sendo decidiu-se utilizar um valor que garantidamente produzisse uma segmentação visivelmente diferente.

### 4.2.1 Apresentação dos Testes com Utilizadores

Os Testes com Utilizadores consistem na apresentação de uma série de páginas *web*, a sequência das quais é apresentada na Figura 4.18.

Uma vez iniciado o teste, apresenta-se ao utilizador uma mensagem a explicar que o teste é feito no âmbito de um trabalho de Mestrado e que tem como objectivo determinar a utilidade do algoritmo desenvolvido durante o mesmo. Esta fase inicial corresponde à Figura 4.18(1).

Na etapa seguinte (Figura 4.18(2)), de forma a caracterizar os utilizadores que participam no teste, é realizado um questionário com perguntas que determinam as preferências dos mesmos. São abordadas questões tais como tecnologias e serviços tipicamente utilizados, preferências e opiniões relevantes ao domínio do problema. Tudo sempre relacionado com o acto de se tirar uma fotografia.

Antes de apresentar a primeira página *web* com a primeira colecção de fotografias segmentada pelos dois algoritmos, o utilizador é informado sobre alguns detalhes acerca do uso da aplicação *web* (Figuras 4.19 e 4.18(3)).

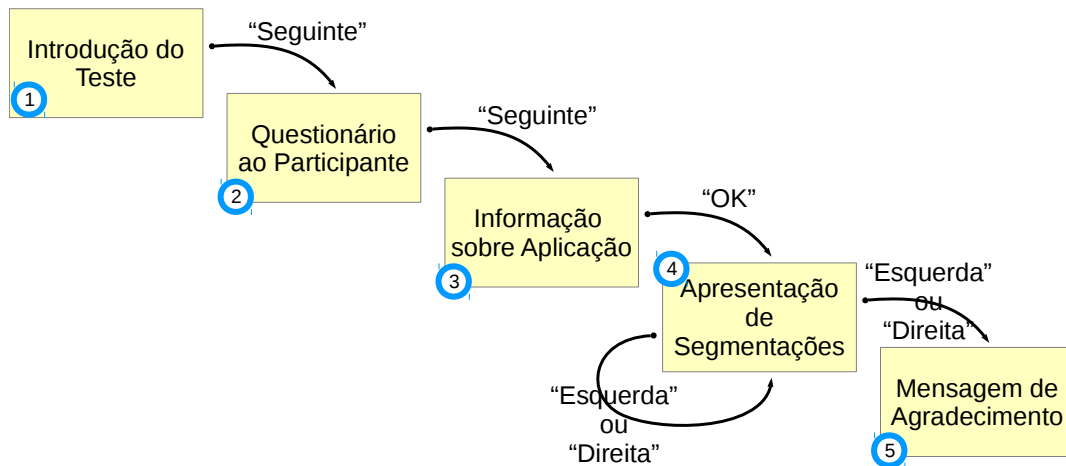


Figura 4.18: Os testes com utilizadores seguem uma sequência de apresentação de páginas *web*.

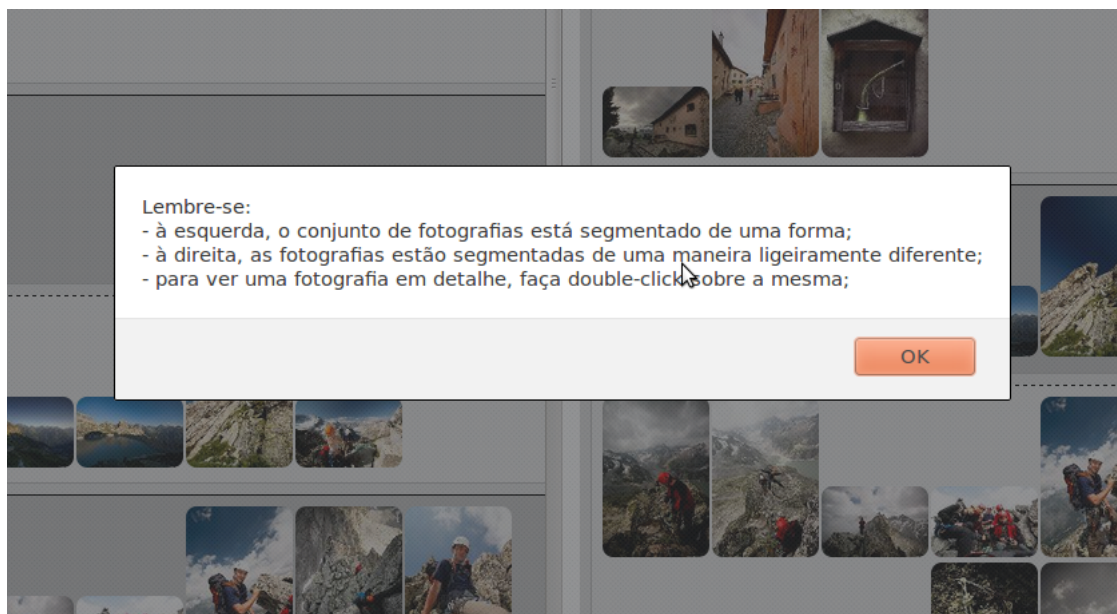


Figura 4.19: Antes de se iniciar a apresentação das segmentações criadas, é apresentada informação sobre a aplicação.

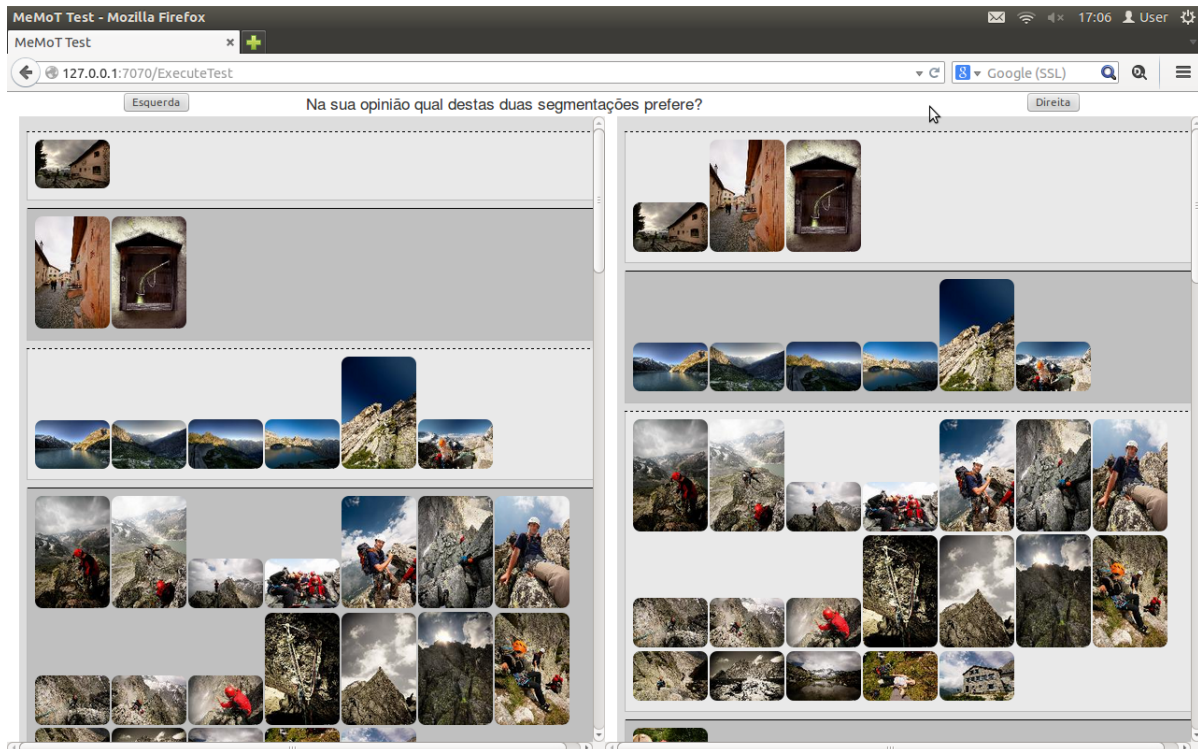


Figura 4.20: Inicia-se a apresentação das segmentações criadas por cada um dos dois algoritmos.

Após a apresentação de uma breve explicação sobre os testes, é apresentada uma sequência de páginas *web*, cada qual com uma colecção de fotografias segmentada de duas formas em simultâneo: de acordo com os Algoritmos de Segmentação no Tempo (AST) e no Tempo e Espaço (ASE) (Figuras 4.20 e 4.18(4)).

Cada uma destas duas segmentações é apresentada à esquerda ou à direita da página *web* sendo essa decisão feita em *run-time* de forma aleatória com uma probabilidade de 50% [10].

Na Figura 4.20 é perceptível uma alteração realizada à segmentação original: o primeiro segmento da segmentação à direita encontra-se fraccionado em dois segmentos na segmentação da esquerda. Também é possível visionar na mesma figura a forma implementada de tornar perceptível a delimitação de cada segmento: com a alternância entre duas tonalidades de cinzento. Isto é feito de forma a facilitar o trabalho do utilizador na detecção de diferenças entre as duas segmentações.

Tanto a gama de colecções de fotografias apresentadas como a sequência de apresentação das ditas colecções é fixa e pré-determinada conforme se pode ver na Tabela 4.2 – as mesmas colecções são apresentadas a cada um dos participantes por uma sequência fixa.

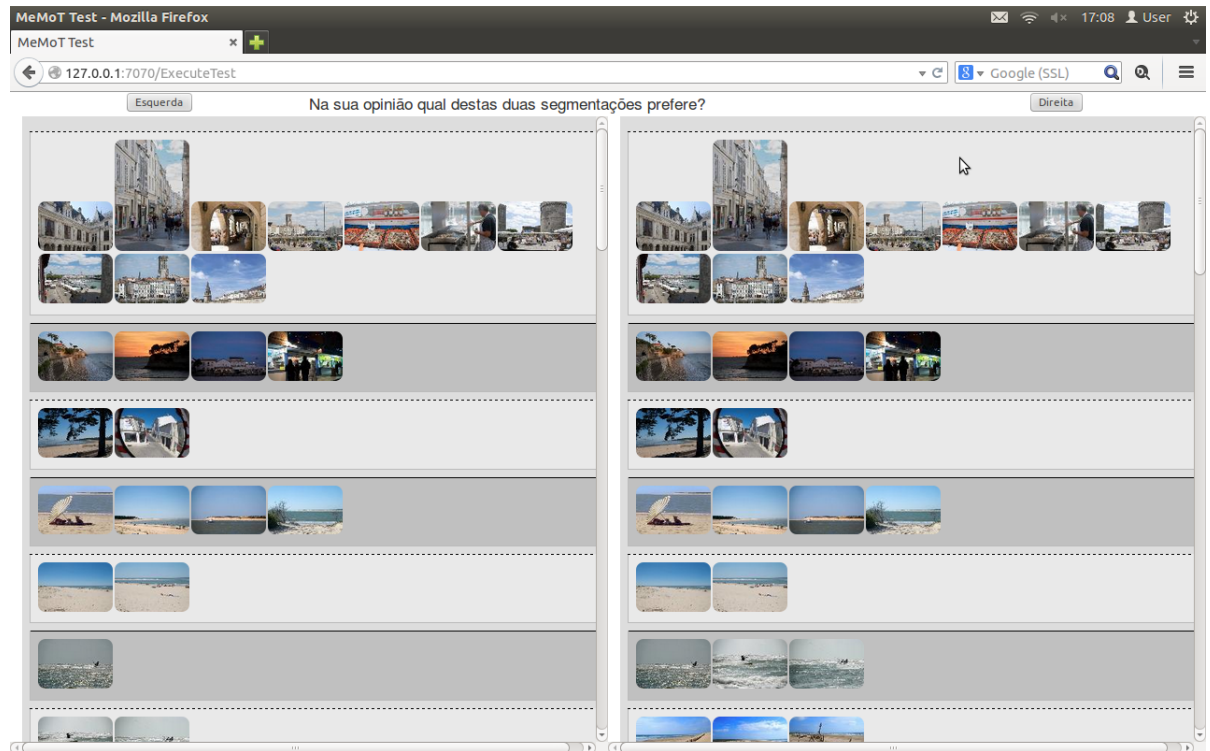


Figura 4.21: É apresentada a segunda colecção de fotografias. As duas segmentações são apresentadas para que o utilizador decida qual é a segmentação que prefere.

A ordem pela qual a sequência de colecções é apresentada pode ser normal ou inversa. Esta decisão é feita de forma aleatória em *run-time*, também com uma probabilidade de 50%. Esta separação dos utilizadores em dois grupos realiza-se de forma a determinar se a ordem de apresentação interfere com os resultados obtidos. Note-se que apenas a ordem de apresentação das colecções poderá eventualmente ser invertida [23]. No âmbito de cada colecção, as fotografias são sempre apresentadas pela mesma ordem e os segmentos criados são também apresentados sempre pela mesma ordem, a cada um dos participantes.

Terminado o carregamento da colecção de fotografias seguinte, os *thumbnails* são enviados ao *browser* e apresentados ao utilizador, como se pode ver na Figura 4.21. Nesta altura a aplicação de testes de utilizador ainda se encontra na fase representada pela Figura 4.18(4).

Este processo é repetido para cada uma das colecções de fotografias listadas na Tabela 4.2. Depois de se registar a preferência relativamente à última colecção, o processo termina com uma mensagem de agradecimento. Na Figura 4.18 esta última etapa é representada pela caixa assinalada com (5).

### Conteúdo do Questionário

De forma a caracterizar os participantes deste teste são pedidas algumas informação aos mesmos. Para além da idade e género, são também pedidas as seguintes informações:

1. Que tipo(s) de máquina fotográfica costuma usar?
  - (a) Máquina Fotográfica do Telemóvel
  - (b) Máquina Fotográfica *Point and Shoot*
  - (c) Máquina Fotográfica SLR
  
2. Quando tira as suas fotografias, regista as coordenadas geográficas do local onde as tira...
  - (a) Não faço ideia do que possa ser uma coordenada geográfica
  - (b) Não registo as coordenadas nas minhas fotografias.
  - (c) A máquina fotográfica faz isso por mim.
  - (d) Recorrendo a um receptor de GPS.
  - (e) Registo as coordenadas à mão, com pouco detalhe.
  - (f) Registo as coordenadas à mão, com o detalhe máximo que conseguir.
  
3. Que programa(s) usa para gerir a sua colecção de fotografias?
  - (a) Nenhum
  - (b) Aperture
  - (c) Digikam
  - (d) F-Spot
  - (e) iPhoto
  - (f) Picasa
  - (g) Lightroom
  - (h) Shotwell
  - (i) Windows Live Photo Gallery
  - (j) Outro. Qual?
  
4. Onde costuma guardar as suas fotos?

- (a) Deixo-as na memória interna ou no cartão de memória da máquina fotográfica.
  - (b) Descarrego-as para o meu computador.
  - (c) Tiro as fotografias da máquina e publico-as num *web site* na Internet.
5. Que *web site(s)* da Internet costuma utilizar para publicar as suas fotos?
- (a) Nenhum
  - (b) Facebook
  - (c) Flickr
  - (d) Instagram
  - (e) PicasaWeb
  - (f) SmugMug
  - (g) Outro. Qual?
6. Que serviços *online* para armazenamento de ficheiros costuma utilizar para salvarguardar (*backup*) as suas fotos?
- (a) Nenhum
  - (b) Dropbox
  - (c) Google Drive
  - (d) iCloud
  - (e) Microsoft Skydrive
  - (f) MyShoebox
  - (g) Outro. Qual?
7. Destes temas, quais considera importantes para organizar as suas fotos?
- (a) Nenhum
  - (b) Nome do evento
  - (c) Espaço (onde se tirou a foto)
  - (d) Tempo (em que momento se tirou a foto)
  - (e) Outro. Qual?
8. Na sua opinião, dos atributos descritos no ponto anterior, o mais importante é o local onde se tirou a fotografia?

- (a) Sim
- (b) Não
- (c) Não sei ou prefiro não responder

9. Na sua opinião a ordem cronológica das fotos (a ordem pela qual as fotos foram tiradas) é importante?

- (a) Sim
- (b) Não
- (c) Não sei ou prefiro não responder

## 4.2.2 Caracterização dos Utilizadores

As respostas obtidas com o questionário indicado anteriormente, permitiu-nos caracterizar os utilizadores que participaram nos testes.

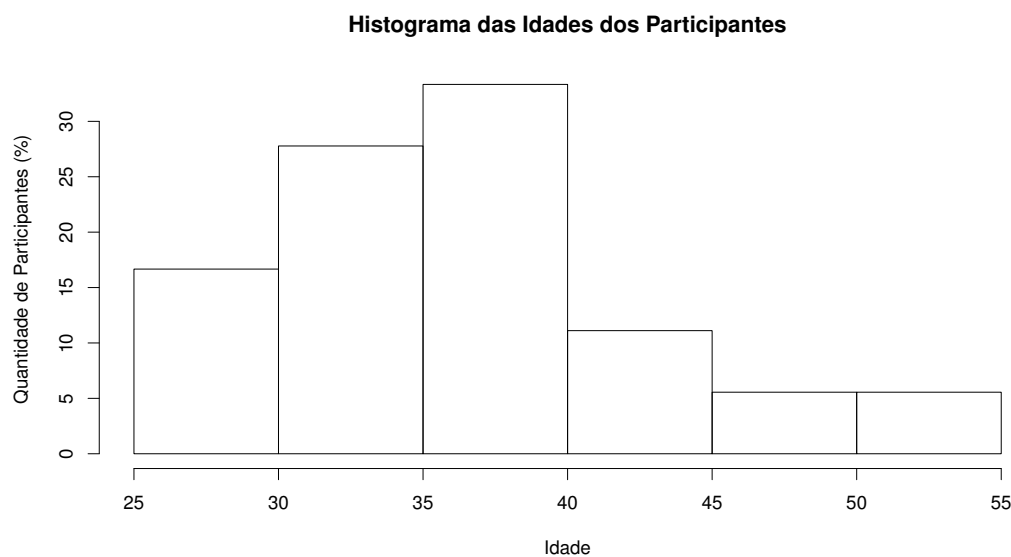


Figura 4.22: Este histograma indica a idade dos vários participantes.

A Figura 4.22 ilustra a faixa etária dos utilizadores que realizam o teste indicado na secção 4.2.1.

A Figura 4.23 indica-nos que a maior parte dos utilizadores — quase 60% — tira as suas fotografias com a máquina fotográfica do seu *Smartphone*. Um número inferior de utilizadores (30%) usa um dispositivo de entrada de gama dedicado

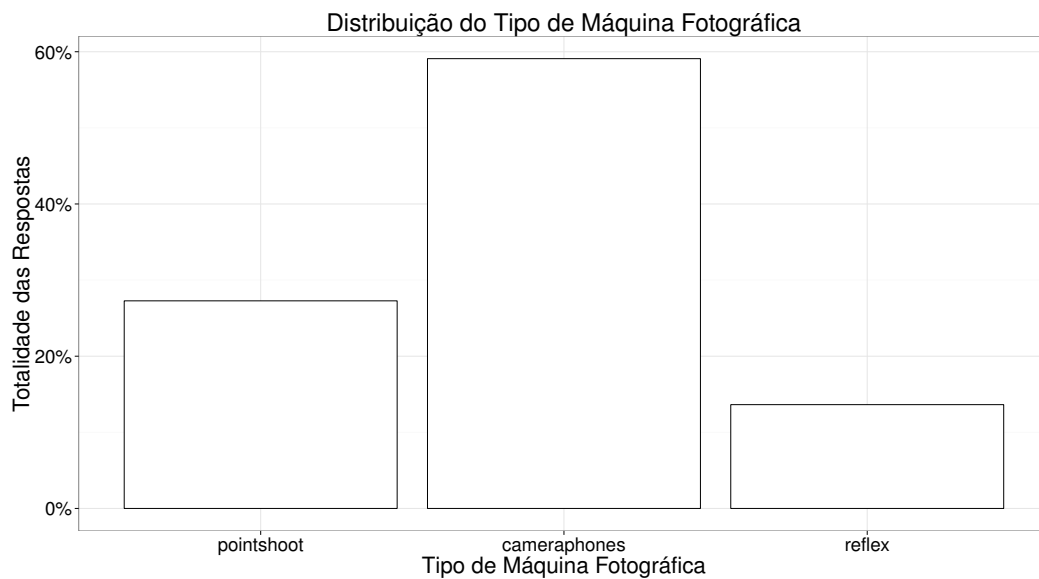


Figura 4.23: Os resultados deste questionário estão em acordo com as conclusões de [38]: tipicamente a máquina fotográfica do *smartphone* é o dispositivo mais utilizado para criar fotografias.

à criação de fotografias — uma máquina fotográfica do tipo *Point & Shoot*. Finalmente, quase 15% dos utilizadores usa uma máquina fotográfica do tipo *Single Lens Reflex*, um dispositivo considerado de gama média ou alta. Note-se que nesta questão os utilizadores podiam assinalar múltiplas respostas — todas as respostas que se aplicassem ao seu caso — mas apenas 3 utilizadores indicaram mais do que um tipo de máquina fotográfica: *Smartphone* e *Point & Shoot*, *Smartphone* e *Single Lens Reflex* e finalmente um último utilizador indicou os três tipos de máquina.

Com a Figura 4.24 ficamos a saber que na grande maioria dos casos (quase 70%), o utilizador não regista as coordenadas nas suas fotografias. No segundo caso mais comum, perto dos 20%, a máquina fotográfica regista automaticamente coordenadas nas fotografias tiradas. Os casos com menos ocorrências, “registo de coordenadas à mão, com o detalhe máximo”, “não faço ideia do que possa ser uma coordenada geográfica” e “registo de coordenadas à mão, com pouco detalhe” foram indicados uma única vez, cada. Mais uma vez, também nesta questão os utilizadores podiam assinalar múltiplas respostas — todas as respostas que se aplicassem ao seu caso. Apenas um utilizador (aproximadamente 5% do conjunto de utilizadores) indicou mais do que uma opção: não regista as coordenadas ou então fá-lo à mão, com grande detalhe.

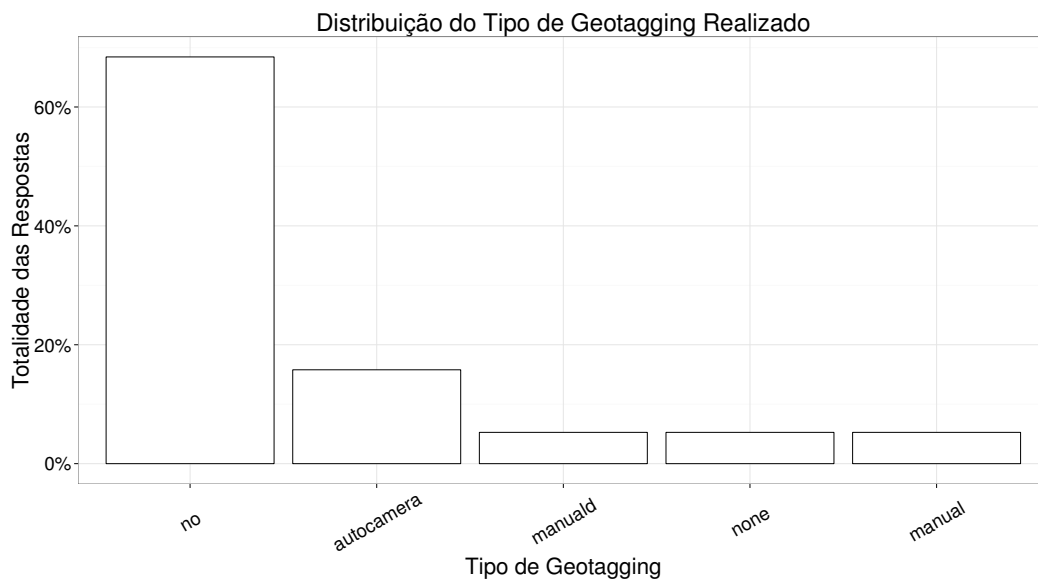


Figura 4.24: A maior parte dos participantes do teste não se preocupa com o uso de tecnologias de *Geotagging*.

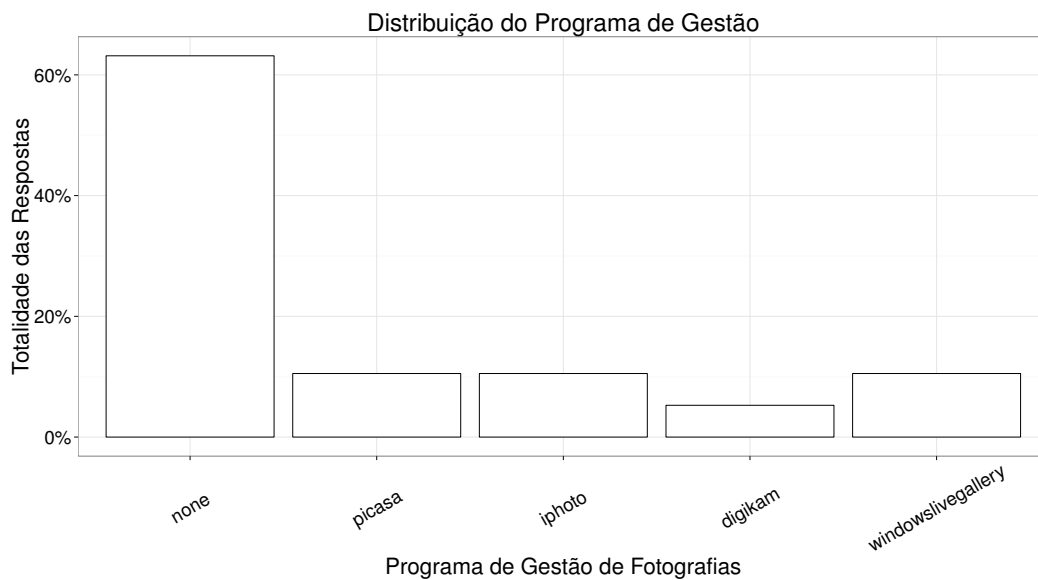


Figura 4.25: A maior parte dos utilizadores não usa qualquer tipo de programa de gestão de fotografias. Poderá ser um indicador que separa os utilizadores que têm por *hobbie* a fotografia, dos utilizadores que tiram fotografias ocasionalmente.

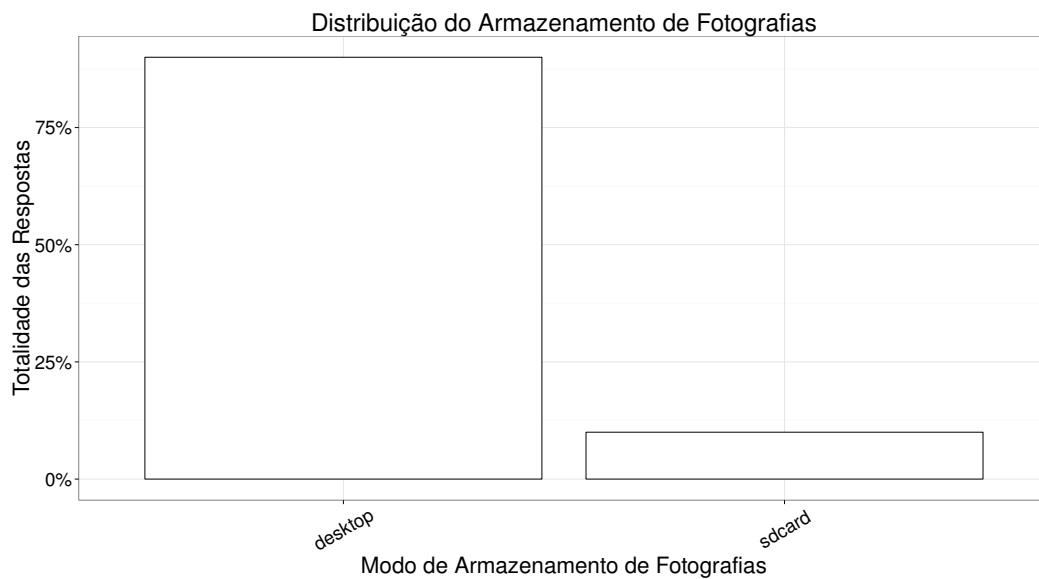


Figura 4.26: Em todos os casos as fotografias são sempre armazenadas num computador pessoal; apenas uma pequena quantidade de participantes indicou armazenar *também* as suas fotografias num cartão de memória.

Na Figura 4.25 são apresentados os programas de gestão de fotografias favoritos dos participantes do teste. A grande maioria (sensivelmente 65%) limita-se a guardar os ficheiros gerados pelas máquinas fotográficas, mas em 35% dos casos é usado um dos programas: “Picasa”, “iPhoto”, “Digikam” ou “Windows Live Gallery”. Os utilizadores podiam assinalar múltiplas respostas, de forma a permitir a indicação de todos os programas normalmente utilizados, mas apenas um utilizador indicou mais do que um programa: “Picasa” e “Digikam”.

O armazenamento das fotografias dos utilizadores continua a ser realizado no computador pessoal, à semelhança de todos os outros documentos digitais. A Figura 4.26 indica-nos também que um número muito restrito de utilizadores armazena as suas fotografias em cartões de memória (“sdcard”). Ninguém indicou utilizar serviços “online” para guardar as suas fotografias. Esta também se tratava de uma questão de escolha múltipla. Todos os utilizadores que indicaram a opção “sdcard” também indicaram a opção “desktop”.

A Figura 4.27 diz-nos que aproximadamente metade das respostas indicam que as fotografias não são publicadas *online*. Quase 30% das respostas indicam a publicação de fotografias no *Facebook*, o que faz sentido tendo em conta a disseminação desse *web site*. De forma a permitir que os utilizadores que usam mais do que um *web site* pudessem exprimir a sua opinião, esta questão permitia a escolha de múltiplas opções. Um dos participantes indicou as opções “Facebook”,

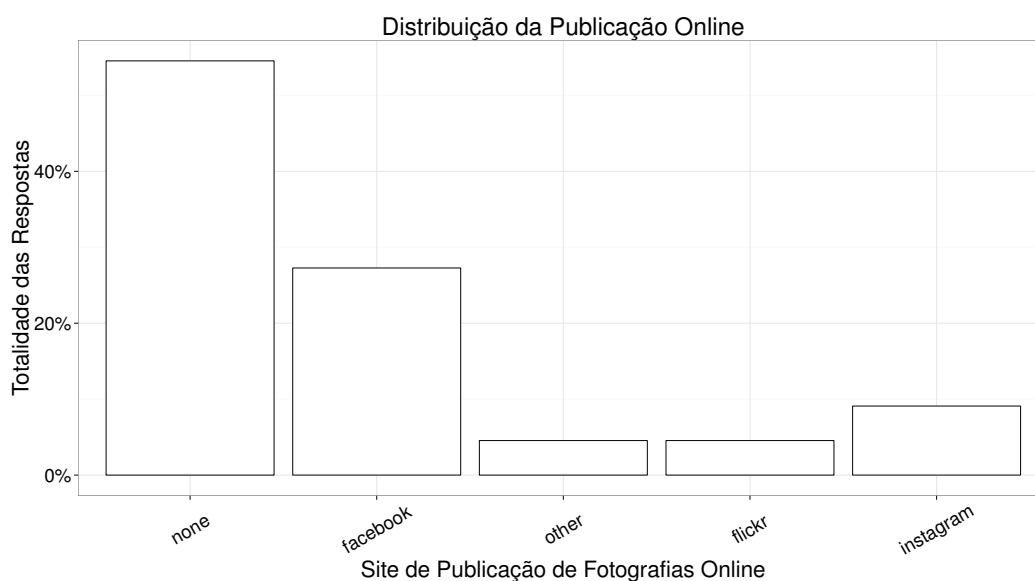


Figura 4.27: A maioria dos utilizadores que publica as suas fotografias na *Internet* fá-lo no *Facebook*. A maior parte dos utilizadores continua a não publicar fotografias *online*.

“Flickr” e “Instagram”. Um outro indicou as opções “Facebook” e “Instagram” e finalmente um terceiro indicou o “Facebook” e “outro” *web site* — o *web site* “Olhares.com”.

De forma a garantir a existência contínua do seu arquivo fotográfico, os utilizadores realizam cópias de salvaguarda das suas colecções de fotografias. Apesar disso, na Figura 4.28 vemos que metade das respostas indicam que não são feitas cópias de segurança do arquivo fotográfico dos participantes. Em dois dos casos o utilizador optou pela hipótese ‘Outro’. Num desses dois casos o utilizador indicou o ‘Disco externo’; no outro foi indicado o ‘MeoCloud’. Esta era uma questão de escolha múltipla de forma a se poder indicar o uso de múltiplos métodos de *backup*, como mandam as boas práticas da Informática.

Mais de 40% das respostas indicaram que o tempo é um factor de organização importante. De acordo com a Figura 4.29, seguem-se o evento social que as fotografias retratam e finalmente a informação geográfica. Ninguém optou pela opção ‘Nenhum’ o que indica que todos os participantes preferem ter um ou mais critérios de organização nas suas colecções de fotografias. Dos inquiridos, 22% indicou que prefere organizar por “espaço” e “tempo”; 27% prefere organizar por “evento”, “espaço” e “tempo” e 33% indicou que prefere organizar por “evento” e “tempo”.

A Figura 4.30 indica que para mais de metade dos participantes a informação

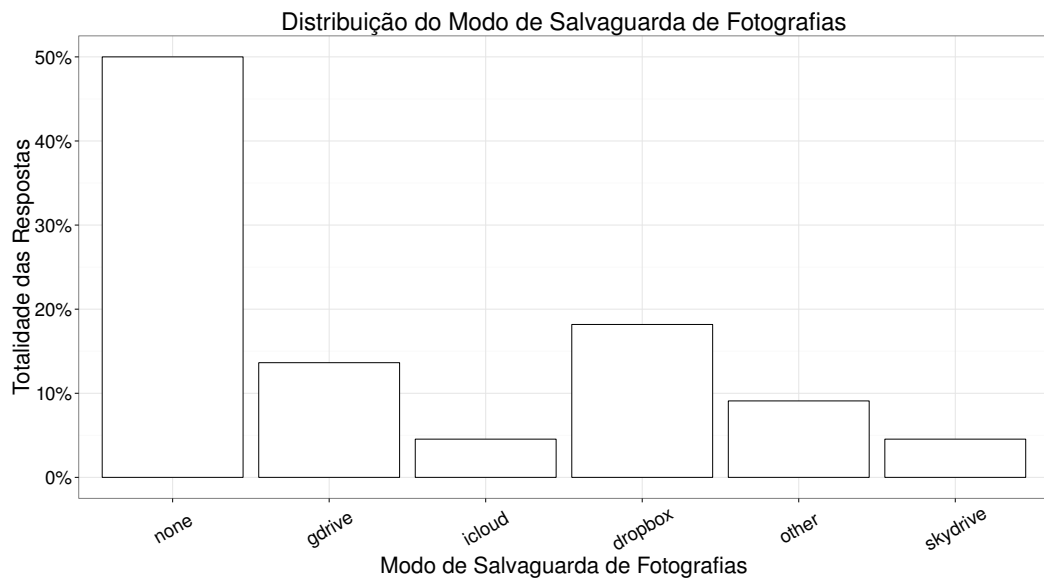


Figura 4.28: Em metade das respostas é indicado a ausência de cópias de salvaguarda das fotografias tiradas.

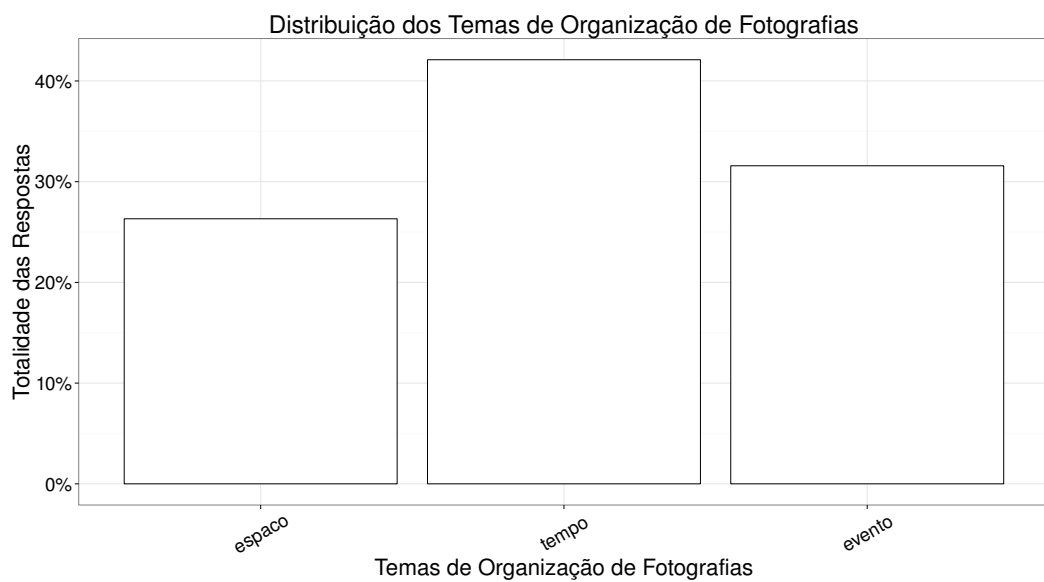


Figura 4.29: O tempo é o factor de organização favorito dos participantes no teste.

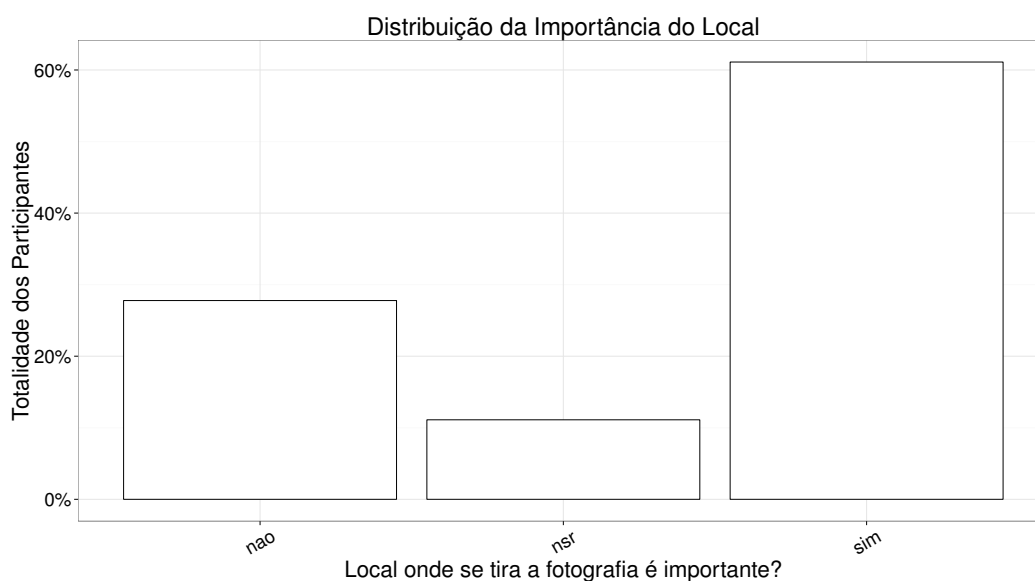


Figura 4.30: Mais de 60% dos participantes indicaram que o mais importante é o local onde se tirou a fotografia.

geográfica é o critério de organização mais importante. Aproximadamente 10% dos inquiridos não sabe ou não responde a esta questão.

A última Figura — 4.31 — confirma que a informação temporal é efectivamente importante para os utilizadores. Indica também que a exigência inicial de garantir a ordem cronológica das colecções de fotografias justifica-se.

### 4.2.3 Resultados dos Testes com Utilizadores

As Figuras 4.32 e 4.33 relatam a preferência dos utilizadores em função da colecção apreciada. Assim, a título de exemplo, parece haver uma preferência pelo ASE para as colecções “Atlantik2008” e “Alpy2011”, quando a sequência de colecções é apresentada pela sua ordem normal.

Quando a ordem é a inversa, esta preferência deixa de se verificar. Por outro lado, no caso da colecção “Christmas2008”, a segmentação favorita é a produzida pelo AST quando a sequência de colecções é apresentada pela sua ordem normal, tendência que efectivamente se inverte quando a sequência de colecções é apresentada por ordem inversa. Não se encontrou uma explicação para este efeito e serão necessários mais ensaios sobre uma amostra maior da população para determinar se esta inversão é algo que resulta dos utilizadores que participaram neste teste ou se é um comportamento característico da população em geral.

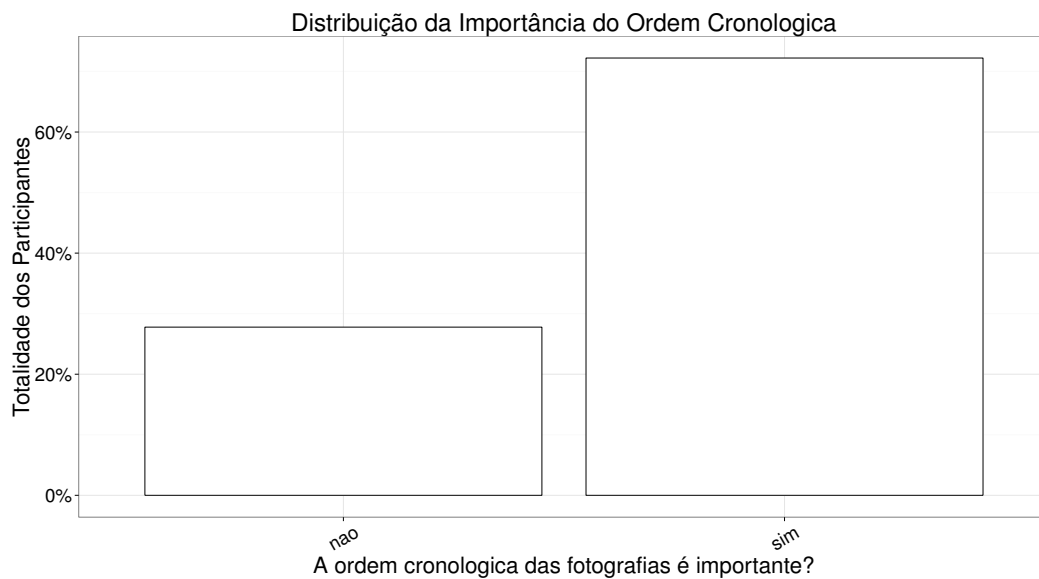


Figura 4.31: Quase três quartos dos inquiridos pensa que a ordem cronológica das fotografias é importante.

Ainda sobre a questão de dividir o conjunto de participantes em dois grupos, dos 18 utilizadores que participaram neste teste, 11 avaliaram as colecções pela ordem normal e os restantes 7 utilizadores avaliaram a colecção pela ordem inversa.

Os utilizadores foram repartidos por três faixas etárias: 1 =  $[0; 32.67[$ , 2 =  $[32.67; 38.33]$  e 3 =  $]38.33; +\infty[$ . Os valores 32.67 e 38.33 representam os quartis  $Q_1$  e  $Q_3$  do conjunto de idades apresentadas pela totalidade dos participantes neste teste.

Apesar de uma análise gráfica da Figura 4.34 indicar uma correlação entre a idade e a preferência do Algoritmo de Segmentação no Tempo e Espaço (ASE), essa correlação não aparenta ser estatisticamente relevante.

Apesar de já ser relativamente fácil encontrar fotografias georeferenciadas nos sites sociais, com a Figura 4.35 ficamos a saber que a grande maioria dos participantes não usa (ou no mínimo acredita que não usa) qualquer tipo de georeferenciação. A mesma Figura diz-nos também que os poucos participantes que fazem uso de tecnologias de georeferenciação são exclusivamente do género masculino, como seria de esperar [12] de uma tecnologia que ainda não está difundida em todo o lado. Infelizmente, devido ao reduzido número de participantes no teste, não se pode fazer qualquer tipo de extrapolação para a população em geral.

A Figura 4.36 contabiliza as vezes que os utilizadores preferiram cada um dos dois algoritmos, discriminando por método de organização de fotografias indicado no questionário. Assim, houve casos pouco expressivos, como foi o caso

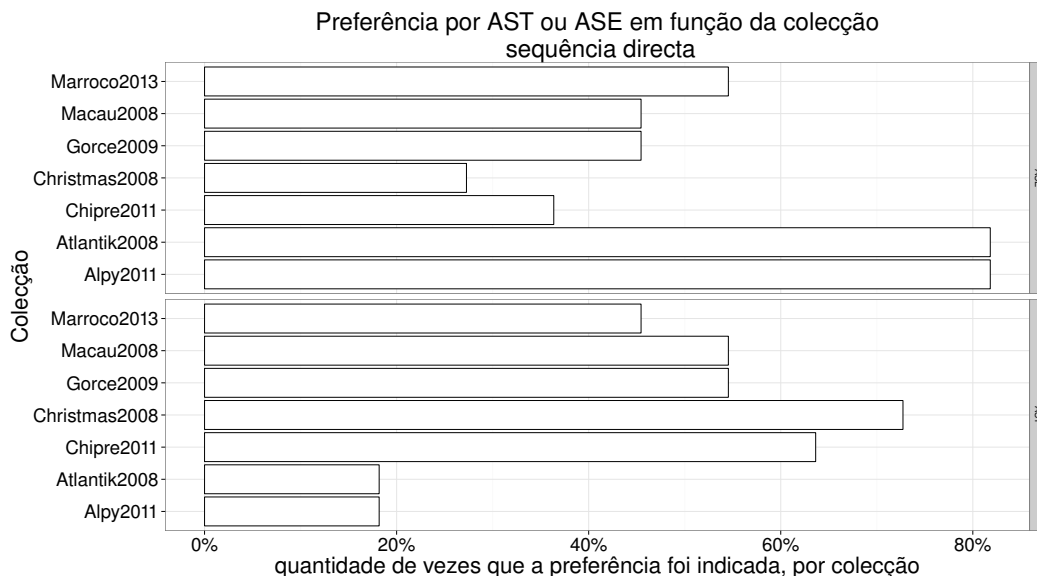


Figura 4.32: Quando as colecções são apresentadas por ordem normal, as contagens de preferências dos participantes indicam uma clara preferência pelo ASE, em algumas colecções.

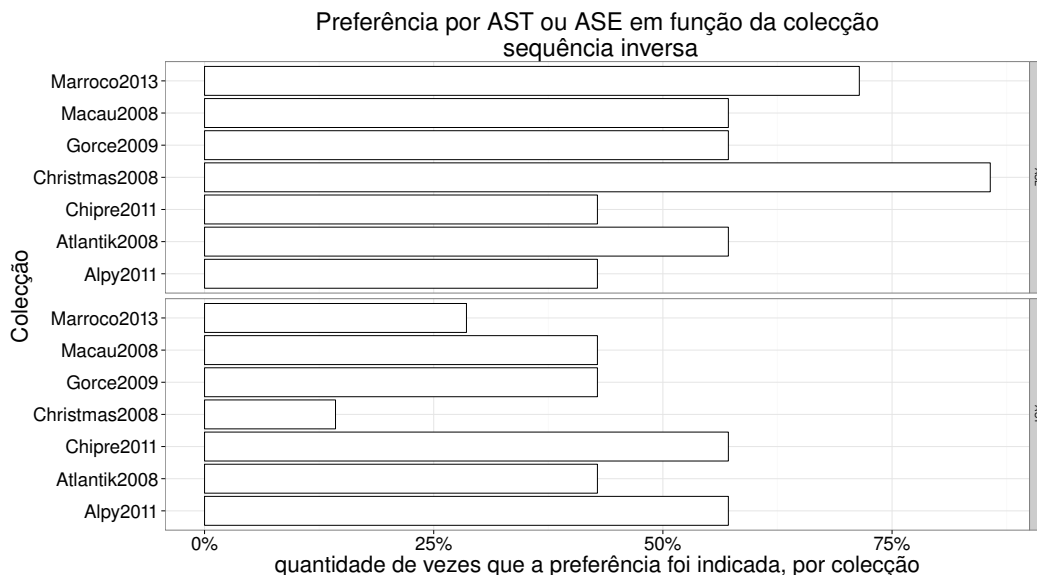


Figura 4.33: Quando as colecções são apresentadas por ordem inversa, os participantes continuam a apresentar preferência pelo ASE se bem que, ao contrário do seria de esperar, em colecções diferentes das verificadas na Figura acima.

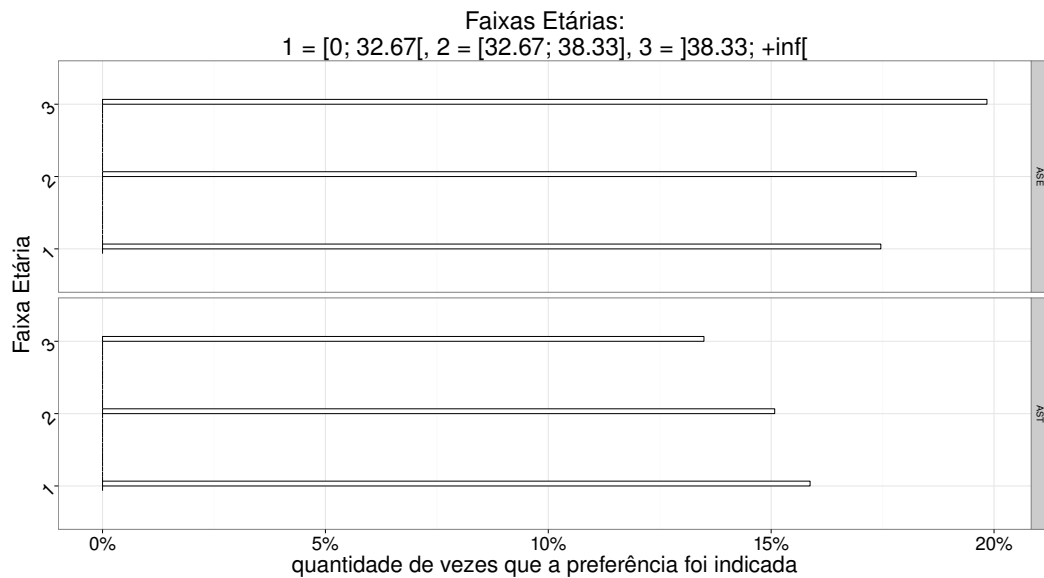


Figura 4.34: As preferências de algoritmo foram divididas em função da faixa etária dos respectivos utilizadores.

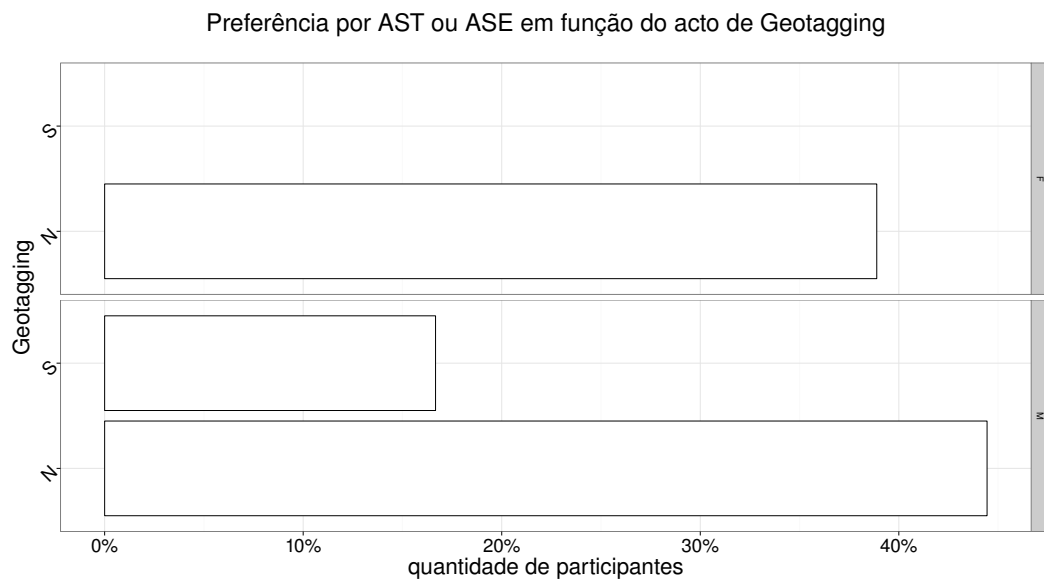


Figura 4.35: Foram pedidas informações aos participantes neste teste que indicassem se incluem informação geográfica nas fotografias que tiram. A figura separa os participantes do género masculino (em baixo) dos do género feminino.

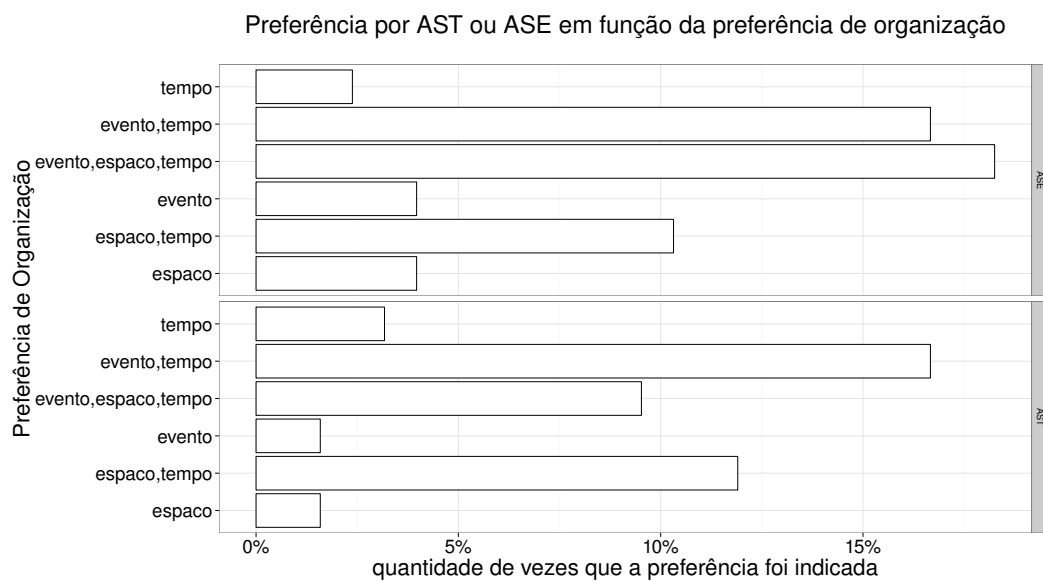


Figura 4.36: As preferências de algoritmo foram contabilizadas consoante o método de organização de fotografias preferido pelos utilizadores.

do “tempo” que efectivamente reflecte o AST como algoritmo favorito mas com pouca diferença, mas também temos casos como “evento,espaco,tempo”, “evento” e “espaco”. Nestes três casos constata-se que os participantes claramente preferiram as segmentações criadas pelo ASE.



## Conclusões

Um dos objectivos deste trabalho consistia em otimizar uma segmentação resultante do algoritmo de Datia [7], gerada com base na informação temporal. Essa optimização utilizou a informação geográfica, também presente nas fotografias que compõem a segmentação, de maneira a modificar a segmentação inicial e transformá-la numa segmentação alternativa.

O outro objectivo do trabalho pretendia determinar se essa informação geográfica, seria uma mais valia no contexto de um algoritmo de segmentação automática. Para este efeito foi construída uma aplicação *web* que, recorrendo aos algoritmos desenvolvidos em [7] e no presente trabalho, recolhe a preferência dos seus utilizadores: qual a melhor forma de segmentar um determinado conjunto de fotografias? Utilizando apenas a informação temporal ou utilizando a informação temporal seguida da informação espacial?

Realizou-se um estudo sobre o problema, em particular a proveniência, correcção e precisão da informação geográfica e ainda a forma de trabalhar com essa informação dentro de certas limitações, nomeadamente a necessidade de manter a ordem cronológica inicial das fotografias.

Foi desenvolvido o Algoritmo de Segmentação no Tempo e Espaço (ASE), que é o assunto principal deste trabalho e ainda diversos algoritmos de detecção de coordenadas erradas, uma necessidade identificada durante o estudo da informação geográfica e dos problemas associados ao uso deste tipo de informação. O processo de desenvolvimento consistiu na realização de várias iterações a projectar,

implementar, testar e avaliar os resultados de cada acréscimo ao projecto. A avaliação dos resultados foi considerada a etapa mais difícil uma vez que determinar a qualidade dos resultados de certos acréscimos é uma questão subjectiva.

Os Testes de Caracterização mostraram que depois de se usar a metadata relativa ao tempo para criar uma segmentação temporal, pode refinar-se essa segmentação usando a metadata que contém as coordenadas geográficas. O número de operações realizadas sobre a segmentação temporal é significativo e analisando manualmente as imagens que compõem os novos segmentos criados pelo ASE confirma-se que as alterações são, na opinião do autor, justificadas.

A variação do parâmetro *SpaceThreshold* em todo o seu domínio mostrou, na maior parte dos casos, uma tendência para aumentar o número de segmentos gerados, quando comparado com a segmentação original. Este efeito está ilustrado nas Figuras 4.8(a) e 4.9(a), segundo as quais ocorreram sempre muitos mais fraccionamentos do que junções.

Durante o desenvolvimento dos Testes de Caracterização, ao apreciar-se o funcionamento da operação de junção, observou-se que em alguns casos os segmentos unidos efectivamente continham fotografias que aparentavam retratar o mesmo evento social e em todos os casos as fotografias dos segmentos unidos tinham sido tiradas em relativa proximidade, tendo em conta a dispersão de todas as fotografias ao longo da segmentação.

Os Testes com Utilizadores apresentaram um conjunto de resultados que indiciam uma inclinação para o ASE, por parte dos participantes. Esses testes revelaram ainda que os utilizadores, à semelhança do já indicado em trabalhos anteriores, utilizam predominantemente o *smartphone*. Mais de metade da amostra da população que participou neste trabalho indicou também que ainda não inclui informação geográfica nas fotografias que criam.

Os participantes do teste foram divididos em dois grupos, com o objectivo de determinar se a ordem de apresentação tem alguma influência no resultado. Esta questão confirmou-se quando, ao avaliar as preferências reflectidas pelos dois grupos, observou-se algumas incoerências entre os dois grupos.

A análise dos resultados destes testes revelou que há indícios de que a informação geográfica, potencialmente disponível numa fotografia, é efectivamente uma mais valia no contexto de um algoritmo de segmentação automática. Intuitivamente, isto faz sentido uma vez que a informação geográfica possui informação com uma semântica diferente da informação temporal; informação essa que responde à questão “onde foi tirada, esta fotografia?”.

Os métodos de detecção de coordenadas erradas tiveram o comportamento esperado, na medida em que os dois métodos de detecção mais genéricos – *Interquartile Range* e *Local Outlier Factor* – tiveram resultados menos satisfatórios que o método direccionado à detecção de características relacionadas com a semântica do problema, o detector com base na velocidade. Como as colecções de fotografias usadas são de domínio público (sobre as quais não existem dados de contexto) qualquer tentativa de definir um valor marginal assume-se como um acto subjectivo, mesmo quando os dados são avaliados manualmente.

Foi, no entanto, possível determinar alguns casos nos quais a coordenada está claramente errada, por exemplo devido ao facto de ser a única a centenas de quilómetros das restantes, ou devido ao facto de localização geográfica não estar de acordo com a imagem retratada – casos óbvios como por exemplo mar versus terra ou cidade versus campo.

O facto das colecções de fotografias não pertencerem aos participantes deste teste significa que o contexto dos vários eventos sociais retratados nas colecções de fotografias é desconhecido dos participantes. Isto implica que qualquer segmentação que os utilizadores prefiram será apenas uma preferência pessoal.

Infelizmente, devido ao reduzido número de participantes nos Testes com Utilizadores, o número de testes efectuados não permite atribuir uma relevância estatística aos resultados obtidos.

Cumulativamente ao baixo número de participantes neste teste, o facto de se ter dividido o conjunto de participantes em dois grupos – o grupo de participantes que avaliou a sequência de colecções de fotografias por ordem directa e o grupo de participantes que avaliou a sequência de colecções de fotografias por ordem inversa – veio realçar os problemas associados ao baixo número de participantes. No entanto, entendeu-se inicialmente que esta divisão era relevante já que se poderiam observar diferentes resultados em função da ordem pela qual se apresentam as colecções de fotografias.

Assim, os resultados permitem identificar um conjunto de tendências, que, para se tornarem estatisticamente relevantes, necessitam de mais testes.

## 5.1 Trabalho Futuro

No sentido de melhorar o algoritmo de detecção de coordenadas baseado em velocidade (3.3.1), deve-se alterar a forma como se determina se uma coordenada é

correcta ou não. Neste trabalho utilizaram-se limites estáticos para determinar o que se considera uma velocidade anómala. É definido um deslocamento máximo e uma velocidade mínima para detectar uma anomalia. Fora deste intervalo, não se conclui que a velocidade é anómala. A título de exemplo, uma velocidade que seja obviamente elevada quando comparada com as restantes velocidade presentes no segmento, poderá não ser considerada errada. Seria preferível ter limites dinâmicos em função dos dados presentes no segmento, de forma a lidar com estas situações. Será no entanto necessário realizar mais testes para chegar a uma medida estatística que se adapte adequadamente a uma quantidade considerável de situações apresentadas por hipotéticos conjuntos de dados.

Todos os algoritmos de detecção de coordenadas erradas realizam um cálculo para determinar distâncias entre coordenadas próximas e só depois determinam que coordenadas devem ser consideradas erradas. Neste trabalho, essas distâncias não são recalculadas depois das coordenadas erradas terem sido detectadas. Do ponto de vista semântico faz sentido recalculer a distância da coordenada subsequente a uma coordenada errada, já que ao considerar que uma dada coordenada está errada implicitamente está-se também a determinar que a distância calculada para a coordenada seguinte também está errada.

O ASE começa por fraccionar os segmentos produzidos pelo algoritmo de segmentação no tempo e só depois de fraccionar, avalia que segmentos deve juntar. Esta ordem pela qual se executam estas duas operações é arbitrária. Não foram realizados estudos para determinar se o mesmo algoritmo, executando as operações pela ordem inversa, apresentaria resultados consistentes com os relatados neste trabalho. Também não se considerou a hipótese de poder existir alguma vantagem em remover a restrição imposta aos conjuntos de segmentos gerados pela operação de fraccionamento segundo a qual, apenas o primeiro e o último segmento destes conjuntos pode ser processado pela operação de junção.

No presente trabalho estudou-se a possibilidade da informação geográfica ser uma mais-valia que possibilita o refinamento da segmentação previamente criada com base na Informação Temporal. Há a possibilidade de um algoritmo que tenha estes dois critérios em atenção *simultaneamente* apresente melhores resultados.

O algoritmo favorece a operação de fraccionamento em detrimento da operação de junção. Um possível desenvolvimento futuro consistiria em aumentar a tolerância do algoritmo que controla a operação de junção de forma a aumentar

a distância mínima admissível para junção de segmentos. Isto provocaria o aumento do número de junções realizadas. Tendo em conta a actual implementação do algoritmo, este processo seria o equivalente a atribuir valores negativos para o parâmetro *SpaceThreshold*, de forma a aumentar o *ReferenceJoinPct* para além dos 100%, tal como é descrito na secção 3.2.4.

Seria desejável determinar se a estabilização observada quando *SpaceThreshold* é superior a 30 (discutida na secção 4.1.4) se deve ao algoritmo ou ao conjunto de dados. Para o efeito, um primeiro passo poderia consistir em substituir todo o conjunto de dados por um outro diferente, mas com uma natureza semelhante e observar os resultados: se nestas circunstâncias os resultados continuarem a estabilizar para um *SpaceThreshold* superior a 30, então será aconselhável investigar o algoritmo para determinar a causa deste comportamento. Nessa situação, um primeiro passo possível será recalibrar o parâmetro à semelhança do que já foi mencionado na secção 4.1.4 e determinar se aquela limitação observada se mantém.

O facto dos participantes não terem passado por uma fase de aprendizagem da aplicação *web* de testes, utilizando para esse propósito um conjunto de dados específico, pode ter deturpado os resultados do teste. Este facto foi em parte deliberado já que se pretendia evitar um possível efeito de saturação junto dos participantes, que não foram remunerados. Assim, decidiu-se reduzir ao mínimo a quantidade de colecções de fotografias que os utilizadores teriam que avaliar no âmbito deste teste. No entanto este é um aspecto cujos efeitos devem ser estudados para determinar se influenciam os resultados do trabalho.



## Referências

- [1] 1000memories. Number of photos ever taken. <http://blog.1000memories.com/94-number-of-photos-ever-taken-digital-and-analog-in-shoebox>, 2012. Online. Accessed: 28/01/2013.
- [2] Geoffrey Blewitt. Basics of the GPS technique: observation equations. *Geodetic Applications of GPS*, pages 2–3, 1997.
- [3] Markus M Breunig, Hans-Peter Kriegel, Raymond T Ng, and Jörg Sander. Lof: identifying density-based local outliers. In *ACM Sigmod Record*, volume 29, pages 93–104. ACM, 2000.
- [4] Christopher Chute. Worldwide digital image print 2011-2014: Forecast the digital photo print bible. Technical report, IDC, 2011.
- [5] James R. Clynych. The global positioning system. [Online; accessed 9-Aug-2013].
- [6] CNN.com. Kodak to sell film business. <http://money.cnn.com/2012/08/24/news/companies/kodak-film/index.html>, 2012. Online. Accessed: 20/08/2013.
- [7] Nuno Datia. Segmenting personal memories. Technical report, CENTRIA, FCT-UNL, 2011.
- [8] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv.*, 40(2):1–60, 2008.
- [9] EarthMeasurement.com. GPS accuracy and limitations. [Online; accessed 06-Aug-2013].

- [10] Michael Eisenberg and Carol Barry. Order effects: A study of the possible influence of presentation order on user judgments of document relevance. *Journal of the American Society for Information Science*, 39(5):293–300, 1988.
- [11] Japan Electronics, Information Technology Industries Association, et al. Exchangeable image file format for digital still cameras: Exif version 2.2. *JEITA CP-3451*, April, 2002.
- [12] Kevin M Elliott and Mark C Hall. Assessing consumer’s propensity to embrace self-service technologies. are there gender differences? *The Marketing Management Journal*, 15(2):98–107, 2005.
- [13] Martin Franz, J Scott McCarley, and Jian-Ming Xu. User-oriented text segmentation evaluation measure. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 701–702. ACM, 2007.
- [14] John F Gantz, Christopher Chute, Alex Manfrediz, Stephen Minton, David Reinsel, Wolfgang Schlichting, and Anna Toncheva. The diverse and exploding digital universe [idc white paper]. *Framingham, MA: International Data Corporation*, 2008.
- [15] Ullas Gargi. Consumer media capture: Time-based analysis and event clustering. *HP-Labs Tech Report*, 2003.
- [16] Google. Google image labeler. [Online; accessed 29-May-2009].
- [17] JESSE PRABAWA GOZALI. *Event Photo Stream Segmentation: Chapter-based Photo Organization For Personal Digital Photo Libraries*. PhD thesis, National University of Singapore, 2013.
- [18] J.P. Gozali, M.Y. Kan, and H. Sundaram. How do people organize their photos in each event and how does it affect storytelling, searching and interpretation tasks? In *Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries*, pages 315–324. ACM, 2012.
- [19] Peter H. Dana. GPS week number rollover issues. [Online; accessed 04-Jul-2014].
- [20] Allan Hanbury. A survey of methods for image annotation. *Journal of Visual Languages & Computing*, 19(5):617 – 627, 2008.

- [21] A. Hanjalic, R. Lienhart, W.-Y. Ma, and J. R. Smith. The holy grail of multimedia information retrieval: So close or yet so far away? *Proceedings of the IEEE*, 96(4):541–547, April 2008.
- [22] Robert R Heselton. *Elevation effects on GPS positional accuracy*. PhD thesis, Virginia Polytechnic Institute and State University, 1998.
- [23] Mu-hsuan Huang and Hui-yu Wang. The influence of document presentation order and number of documents judged on users' judgments of relevance. *Journal of the American Society for Information Science and Technology*, 55(11):970–979, 2004.
- [24] Steve MJ Janssen, Antonio G Chessa, and Jaap MJ Murre. Memory for time: How people date events. *Memory & Cognition*, 34(1):138–147, 2006.
- [25] MS Kankanhalli and Y. Rui. Application Potential of Multimedia Information Retrieval. *Proceedings of the IEEE*, 96(4):712–720, 2008.
- [26] T. Kindberg, M. Spasojevic, R. Fleck, and A. Sellen. The ubiquitous camera: An in-depth study of camera phone use. *IEEE Pervasive Computing*, 4(2):42–50, 2005.
- [27] kowoma.de. Composition of the GPS data signal. [Online; accessed 09-Aug-2013].
- [28] Richard B Langley. Dilution of precision. *GPS world*, 10(5):52–59, 1999.
- [29] Binghao Li, Jiahuang Zhang, Peter Mumford, and Andrew G Dempster. How good is assisted gps? *N/A*, 2011.
- [30] A. Loui and A. Savakis. Automated event clustering and quality screening of consumer pictures for digital albuming. *IEEE Transactions on Multimedia*, 5:390–402, 2003.
- [31] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
- [32] Mathias Lux, Marian Kogler, and Manfred del Fabro. Why did you take this photo: a study on user intentions in digital photo productions. In *Proceedings of the 2010 ACM workshop on Social, adaptive and personalized multimedia interaction and access*, pages 41–44. ACM, 2010.

- [33] Mary Madden, Amanda Lenhart, Sandra Cortesi, and Urs Gasser. Teens and mobile apps privacy. *Book Teens and Mobile Apps Privacy, Pew Research Center, PewInternet.org*, 2013.
- [34] Robert McGill, John W Tukey, and Wayne A Larsen. Variations of box plots. *The American Statistician*, 32(1):12–16, 1978.
- [35] George Mendes. What went wrong at eastman kodak? *N/A*, 2011.
- [36] Marko Modsching, Ronny Kramer, and Klaus ten Hagen. Field trial on gps accuracy in a medium size city: The influence of built-up. In *3rd Workshop on Positioning, Navigation and Communication*, pages 209–218, 2006.
- [37] Mor Naaman, Yee Jiun Song, Andreas Paepcke, and Hector Garcia-Molina. Automatic organization for digital photographs with geographic coordinates. In *JCDL '04: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, pages 53–62, New York, NY, USA, 2004. ACM Press.
- [38] NPD. Consumers now take more than a quarter of all photos and videos on smartphones. [Online; accessed 6-Aug-2013].
- [39] P. Obrador, R. de Oliveira, and N. Oliver. Supporting personal photo storytelling for social albums. In *Proceedings of the international conference on Multimedia*, pages 561–570. ACM, 2010.
- [40] Lev Pevzner and Marti A Hearst. A critique and improvement of an evaluation metric for text segmentation. *Computational Linguistics*, 28(1):19–36, 2002.
- [41] J.C. Platt, M. Czerwinski, and B.A. Field. Phototoc: automatic clustering for browsing personal photographs. *Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on*, 1:6–10 Vol.1, Dec. 2003.
- [42] S Schenk R Troncy, B Huet. *Multimedia Semantics: Metadata, Analysis and Interaction*. Wiley-Blackwell, 2011.
- [43] Bruce Schneier. Terms of service as a security threat. [Online; accessed 28-Aug-2013].
- [44] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-based image retrieval at the end of the early years.

- IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.
- [45] Arnold WM Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1349–1380, 2000.
- [46] New York Suffolk County. Global positioning system GPS data collection guidelines. pages 22–25, 2008.
- [47] New York Suffolk County. Global positioning system (gps) data collection guidelines. pages 11–12, 2008.
- [48] Luis Torgo. Data mining with r. *Learning by case studies. University of Porto, LIACC-FEP*, 7(09):42–48, 2003.
- [49] L. Von Ahn. Games with a purpose. *Computer*, 39(6):92–94, 2006.
- [50] Stephan von Watzdorf and Florian Michahelles. Accuracy of positioning data on smartphones. In *Proceedings of the 3rd International Workshop on Location and the Web*, page 2. ACM, 2010.
- [51] Steve Whittaker, Ofer Bergman, and Paul Clough. Easy on that trigger dad: a study of long term family photo retrieval. *Personal and Ubiquitous Computing*, 14(1):31–43, 2010.
- [52] Paul A Zandbergen. Accuracy of iphone locations: A comparison of assisted GPS, WiFi and cellular positioning. *Transactions in GIS*, 13(s1):5–25, 2009.
- [53] E. Zerubavel. Social memories: Steps to a sociology of the past. *Qualitative Sociology*, 19(3):283–299, 1996.
- [54] Eviatar Zerubavel. *Hidden Rhythms: schedules and calendars in social life*. Univ of California Press, 1985.
- [55] Jiri Zuzanek and Bryan JA Smale. Life-cycle variations in across-the-week allocation of time to selected daily activities. *Loisir et société/Society and Leisure*, 15(2):558–586, 1992.





# Descrição Estatística das Coleções de Fotografias

As tabelas seguintes caracterizam as estatísticas das várias colecções de fotografias utilizadas no âmbito deste trabalho. Em ambas as tabelas é apresentado o DAM (Desvio Absoluto Mediano) e o DP (Desvio Padrão).

<b>Colecção</b>	$n^{\circ}$ <i>fotos</i>	<b>min</b>	<b>max</b>	<b>média</b>	<b>mediana</b>	<b>DAM</b>	<b>DP</b>
Africa2008	179	0.00	888.97	19.14	0.19	0.19	81.11
Alaska	47	0.01	230.35	9.99	2.61	0.86	37.96
Alpy2011	97	0.00	80.68	1.96	0.00	0.00	8.84
Atlantik2008	90	0.00	5752.81	80.01	0.00	0.00	607.05
China2009	171	0.00	1178.80	11.34	0.19	0.15	91.47
Chipre2011	72	0.00	125.22	4.53	0.11	0.11	17.69
Christmas2008	40	0.00	81.52	4.56	0.08	0.08	16.64
Europe2010	513	0.00	12515.76	89.71	0.21	0.19	703.01
Europe2012-2	306	0.00	11807.93	93.44	0.07	0.06	853.82
Europe2012	607	0.00	11636.02	74.26	0.10	0.08	756.90
France2011	164	0.00	12.72	0.43	0.09	0.09	1.56
Gorce2009	88	0.00	39.34	1.50	0.05	0.05	5.86
Hydrobike2009	126	0.00	15.29	2.01	0.63	0.60	3.07
India2010	156	0.00	595.75	11.39	0.08	0.08	62.75
Irlanda2014	258	0.00	10115.74	537.20	0.00	0.00	2009.91
Italy2010	17	0.00	3.54	0.64	0.26	0.26	0.97
Japan2007	81	0.00	3131.34	47.49	0.04	0.04	348.13
Japan2011	448	0.00	1741.02	17.46	0.12	0.09	129.33
Korea2012	280	0.00	1746.19	16.92	0.07	0.06	150.26
Laponia2013	103	0.00	27.51	3.28	0.66	0.66	5.27
Macau2008	44	0.00	21.19	1.43	0.11	0.08	3.87
Marroco2013	75	0.00	586.89	22.98	0.74	0.74	75.69
Mexico2012	263	0.00	12094.57	61.21	0.04	0.04	751.45
Nepal2008	356	0.00	167.18	1.37	0.13	0.12	12.41
Norway2012	238	0.00	970.20	28.53	0.05	0.05	122.16
Oslo2012	53	0.00	7305.04	142.52	0.23	0.23	1003.00
SanFrancisco2008	194	0.00	5.26	0.14	0.00	0.00	0.63
SouthAmerica2010	582	0.00	1053.04	19.79	0.33	0.32	101.33
Swiss2011	186	0.00	80.68	1.65	0.02	0.02	7.10
Swiss2011-1	97	0.00	80.68	1.96	0.00	0.00	8.84
Swiss2011-2	89	0.00	40.09	1.17	0.24	0.24	4.33
Taiwan2008	64	0.00	37.10	3.26	0.17	0.15	7.97
Tajakistan2009	230	0.00	5937.28	38.11	1.13	1.13	392.85
Tajakistan2009-1	157	0.00	351.33	13.19	1.14	1.14	40.94
Tajakistan2009-2	73	0.00	5937.28	92.58	1.07	0.97	694.91
Turkey2010	201	0.00	536.44	7.18	0.03	0.03	53.86
UK2008	111	0.00	378.41	19.18	0.43	0.36	59.33
Vietnam2008	1395	0.00	6837.02	11.43	0.01	0.01	256.05

Tabela A.1: Caracterização das colecções utilizadas durante os testes de caracterização do ASE — as métricas estatísticas avaliam as distâncias que separam cada par de fotografias consecutivas (unidades em km).

A. DESCRIÇÃO ESTATÍSTICA DAS COLECÇÕES DE FOTOGRAFIAS

Colecção	$n^{\circ}$ fotos	min	max	média	mediana	DAM	DP
Africa2008	179	00:00:01	45:19:11	01:43:26	00:08:49	00:07:43	05:08:12
Alaska	47	00:00:02	146:41:27	03:46:55	00:00:11	00:00:09	21:26:27
Alpy2011	97	00:00:07	44:55:15	01:19:36	00:07:05	00:05:51	05:25:52
Atlantik2008	90	00:00:00	48216:19:31	546:21:42	00:15:09	00:15:09	5081:38:34
China2009	171	00:00:07	16:13:00	01:08:30	00:10:36	00:08:21	02:58:40
Chipre2011	72	00:00:10	22:21:27	02:05:50	00:14:29	00:11:53	05:06:32
Christmas2008	40	00:00:07	17:09:04	01:50:48	00:12:29	00:12:09	04:03:19
Europe2010	513	00:00:04	21:59:56	00:55:25	00:08:19	00:06:36	02:48:51
Europe2012-2	306	00:00:02	353:34:12	01:28:16	00:02:19	00:01:44	20:14:55
Europe2012	607	00:00:03	517:37:39	01:06:12	00:03:19	00:02:35	21:01:57
France2011	164	00:00:05	21:49:35	01:09:04	00:21:52	00:19:54	02:55:55
Gorce2009	88	00:00:02	17:15:53	00:59:07	00:05:23	00:05:00	02:52:59
Hydrobike2009	126	00:00:02	16:00:57	00:58:51	00:11:17	00:10:11	02:42:57
India2010	156	00:00:09	44:53:30	02:02:35	00:07:11	00:06:30	06:34:30
Irlanda2014	258	00:00:01	21:20:23	00:34:15	00:01:02	00:00:53	02:53:01
Italy2010	17	00:00:47	14:57:55	01:37:20	00:36:53	00:25:01	03:30:07
Japan2007	81	00:00:05	16:55:29	01:50:22	00:17:40	00:14:50	03:41:05
Japan2011	448	00:00:03	16:45:13	00:36:57	00:06:46	00:05:29	02:10:23
Korea2012	280	00:00:03	13:15:17	00:18:14	00:03:01	00:02:28	01:19:54
Laponia2013	103	00:00:07	24:21:11	02:19:39	00:16:49	00:16:02	05:02:00
Macau2008	44	00:00:16	16:56:13	00:59:01	00:07:59	00:05:04	03:07:23
Marroco2013	75	00:01:10	22:25:09	03:29:33	00:44:09	00:39:43	06:11:43
Mexico2012	263	00:00:04	17402:05:03	67:42:43	00:07:21	00:05:29	1072:58:44
Nepal2008	356	00:00:01	96:26:58	01:41:36	00:07:37	00:07:08	07:01:57
Norway2012	238	00:00:01	19:33:18	00:46:29	00:01:16	00:01:02	03:03:09
Oslo2012	53	00:00:00	24882:11:56	479:04:25	00:08:03	00:05:30	3417:07:18
SanFrancisco2008	194	00:00:00	17:40:49	00:09:11	00:00:10	00:00:09	01:23:53
SouthAmerica2010	582	00:00:04	34:46:05	00:44:18	00:01:35	00:01:18	03:23:45
Swiss2011	186	00:00:01	45:46:17	01:27:48	00:07:19	00:06:56	05:54:03
Swiss2011-1	97	00:00:07	44:55:15	01:19:36	00:07:05	00:05:51	05:25:52
Swiss2011-2	89	00:00:01	37:13:04	01:06:32	00:08:15	00:08:09	04:18:32
Taiwan2008	64	00:00:34	17:04:19	01:10:46	00:11:44	00:09:26	03:05:21
Tajakistan2009	230	00:00:03	67:03:56	04:29:27	01:07:54	01:00:59	08:13:09
Tajakistan2009-1	157	00:00:03	67:03:56	03:42:55	00:50:06	00:44:48	08:07:26
Tajakistan2009-2	73	00:00:11	30:43:16	06:04:10	02:09:59	01:45:00	08:12:17
Turkey2010	201	00:00:00	731:52:36	04:39:29	00:02:25	00:02:10	51:48:20
UK2008	111	00:00:21	23:42:41	01:52:07	00:23:29	00:17:20	04:19:22
Vietnam2008	1395	00:00:00	47635:09:25	34:20:58	00:00:17	00:00:17	1275:22:42

Tabela A.2: Caracterização das colecções utilizadas durante os testes de caracterização do ASE — as métricas estatísticas avaliam os tempos que separam cada par de fotografias consecutivas (unidades em HH:MM:SS).