

メタデータの不均質性を考慮した画像検索結果のクラスタリング手法に関する研究

メタデータ	言語: ja 出版者: 静岡大学 公開日: 2015-04-23 キーワード (Ja): キーワード (En): 作成者: 廣田, 雅春 メールアドレス: 所属:
URL	https://doi.org/10.14945/00008266

静岡大学 博士論文

メタデータの不均質性を考慮した画像検索結果の
クラスタリング手法に関する研究

2014年6月

大学院 自然科学系教育部

情報科学専攻

廣田 雅春

概要

近年、スマートフォンや、デジタルカメラなどのモバイルデバイスが爆発的に普及した結果、一般のユーザが撮影した写真が Flickr などのソーシャルメディアサイト上に大量にアップロードされている。それらの画像を検索する際に、画像検索結果に含まれる画像数の増加に伴い、ユーザが目的の画像を探すための負担が増加している。そのため、ユーザが画像検索結果を効率的に閲覧するための技術の必要性が高まっている。画像クラスタリング・リランキングは、画像検索結果を効率的に閲覧するための技術の1つである。

しかし、画像に付与されているメタデータを用いて画像クラスタリング・リランキングを行う際に、ウェブ上に存在する画像に付与されているメタデータの不均質さへの対処が重要な課題となる。たとえば、ソーシャルタギングの場合、タグが1つも付与されていない場合や、付与されていたとしても、タグと画像の内容に関連性がない場合や、タグの表記ゆれなどがあげられる。そのようなメタデータを用いて画像クラスタリングを行なった場合、クラスタリングの性能が低下し、ユーザの期待するクラスタリング結果が生成されない可能性が高い。メタデータが付与されていない画像は、その画像とメタデータを含む他の画像間の類似度の値を算出することはできない。また、メタデータの一部のみが欠損していた場合、それらの画像間の類似度の算出にメタデータの欠損が強い影響を与えてしまい、本来の意図とは、異なる類似度が算出される場合がある。そのため、画像に付与されているメタデータの修正や、メタデータの欠損を補うことは、ウェブ上の画像をクラスタリングする際の重要な課題の1つである。

本論文では、画像検索結果に含まれる画像に対して、欠損しているメタデータの推定、およびメタデータの修正を行い、推定されたメタデータを用いて画像検索結果のクラスタリングを行う手法を提案した。提案手法では、画像検索の検索クエリと、検索結果の画像に付与されているメタデータのそれぞれの多様性に対処するため、画像検索結果のみを用いて学習することで、メタデータの推定・修正を行った。実験により、メタデータが欠損している場合に、提案手法によって推定されたメタデータを用いて画像クラスタリングを行うことで、クラスタリングの性能の低下を抑えることが可能であることを確認した。

加えて、本論文では、画像検索結果を効率的に閲覧するためのメタデータを用いたクラスタリング手法と、リランキング手法の2つを提案した。

1つは、写真の審美的評価に基づいたリランキング手法である。一般ユーザの写真の撮影技術や、撮影機材の多様さにより、ウェブ上に存在する写真の見た目の美しさも不均質である。そのため、画像検索結果から、審美的評価の高い写真を抽出する技術の必要性が高まっている。提案手法では、画像検索結果において、高価な撮影機材を用いているユーザは、写真の撮影技法に詳しいという仮説を用いて、高価な撮影機材で撮影された写真から得た写真の撮影状況を表すメタデータを用いて、審美的評価に基づいて写真を抽出する手法を提案した。また、人手によるリランキング結果との類似度を評価することで、提案手法の有効性を評価した。結果として、提案手法は、既存手法と比較して、非常に高速に審美的評価に基づいたリランキングが可能であることを示した。

もう1つは、写真の撮影地点の分布に基づいて、地理的なクラスタを抽出し、そのクラスタ間の関連性を抽出する手法である。人々の興味・関心の可視化するため、ソーシャルメディアサイトから取得した撮影地点を表すメタデータである緯度経度情報が付与されている写真を用いて、多くの写真が撮影されている地域をホットスポット(クラスタ)として抽出する。単純にホットスポットを撮影位置情報の密度に基づいて抽出した場合、本来、何らかの関連性があるホットスポットが分離して抽出される場合がある。たとえば、1つのイベントに関連するホットスポットが複数抽出された場合である。これは、単純に撮影位置情報の密度に基づいてホットスポットを抽出した場合、ホットスポットが生成される原因が考慮されていないためである。そのため、本論文では、ホットスポット間の位置関係や、写真に付与されているタグを用いて、独立して抽出されるホットスポット間の関連性を抽出、および可視化する手法を提案した。提案手法では、撮影スポットと被写体の関連性と、タグに基づいた関連性の2種類である。実際に、ソーシャルメディアサイトから取得した画像に提案手法を適用し、地図上に抽出した関連性を可視化することで、その結果を考察した。

目次

第1章 緒言	7
1.1 動機と目的	7
1.1.1 背景	7
1.1.2 画像検索結果	8
1.1.3 画像検索結果のクラスタリング	11
1.1.4 写真の審美的評価	15
1.1.5 地理的なクラスタとその関連性の抽出	16
1.2 本論文の構成	17
第2章 関連研究	18
2.1 撮影状況メタデータの利用	18
2.2 画像クラスタリングを用いた画像検索結果の提示	18
2.3 画像のメタデータ推定	20
2.4 写真の審美的評価	21
2.5 写真の撮影位置情報に基づいたクラスタリング	22
第3章 メタデータの欠損に頑健な 画像検索結果のクラスタリング	23
3.1 概要	23
3.2 類似画像を用いた画像検索結果のメタデータの推定	24
3.3 画像の特徴量の類似度	26
3.4 タグの類似度	27
3.4.1 写真の撮影状況の類似度	27
3.5 実験	29
3.5.1 制約付き凝集法	30
3.5.2 データセット	31
3.5.3 評価指標	31
3.5.4 提案手法によるタグの推定の評価	33

3.5.5	提案手法による Exif の推定の評価	34
3.5.6	推定されたメタデータを用いたクラスタリング手法による 影響の評価	36
3.6	おわりに	37
第 4 章	画素情報を用いない	
	写真の質に基づいた画像検索結果のリランキング	42
4.1	概要	42
4.2	類似画像を用いた写真の質の評価	43
4.2.1	撮影状況類似度を用いたクラスタリング	45
4.2.2	リランキング	45
4.3	実験	46
4.3.1	データセット	46
4.3.2	評価指標	47
4.3.3	提案手法によるリランキング結果の評価	48
4.4	おわりに	49
第 5 章	位置情報付き写真を用いたホットスポットの関連性の抽出	51
5.1	概要	51
5.2	ホットスポットの抽出と関連性の抽出	52
5.2.1	ホットスポットの抽出	52
5.2.2	ホットスポットの分類	53
5.2.3	イベントの分類	54
5.2.4	ホットスポット間の関連性の抽出	55
5.3	実験	56
5.3.1	データセット	56
5.3.2	考察	57
5.4	おわりに	61
第 6 章	結言	62
6.1	成果のまとめ	62
6.2	今後の展望	63

表 目 次

3.1	評価実験に用いたクエリ一覧	28
3.2	Fowlkes-Mallows index における画像のペアのクラス	32
3.3	Exif を推定するためのそれぞれの手法の平均順位	35
4.1	高価な機材と安価な機材によって撮影された写真の質の評価の平均値	43
4.2	Flickr に登録されている撮影機材の種類と数	45
4.3	撮影機材の価格	46
4.4	Unattractive Rejection による提案手法とベースライン手法の比較	49
4.5	提案手法と既存手法の比較 (審美的評価)	50
4.6	提案手法と既存手法の比較 (実行時間)	50
5.1	提案手法によって抽出された関連性の例	57

目次

1.1	2つの異なる種類の画像が含まれている画像検索結果の例	12
1.2	タグが付与されていない画像を含む画像検索結果のクラスタリング の例	13
3.1	欠損したメタデータを推定するアルゴリズム	24
3.2	クラスタリングを用いたノイズタグを削除するアルゴリズム	25
3.3	タグの出現頻度	29
3.4	評価実験の流れ	33
3.5	提案手法の評価 (Tag)	38
3.6	提案手法によって推定されたタグの評価	39
3.7	提案手法の評価 (Exif)	40
3.8	提案手法の評価 (Tag)	41
3.9	提案手法の評価 (Exif)	41
4.1	同じ構図で撮影された2枚の写真とそのExifの値	43
4.2	提案手法の流れ	44
4.3	評価実験で用いる検索キーワードとタグ	46
4.4	提案手法とベースライン手法の比較	49
5.1	ホットスポットの抽出手順	52
5.2	タグに基づいた関連性 (表 5.1 の R0 - R3).	57
5.3	タグに基づいた関連性 (表 5.1 の R4).	58
5.4	撮影スポットと被写体の関連性 (表 5.1 の R7).	58

第1章 緒言

1.1 動機と目的

1.1.1 背景

近年、スマートフォンや、デジタルカメラなどのモバイルデバイスが爆発的に普及している。実際に、デジタルカメラの世帯保有率は、2013年度末に76.5%であり、調査が開始された2001年度末と比較すると約3倍の普及率である [1]。さらに、スマートフォンは更に急速に普及しており、国内のスマートフォンの世帯保有率は、2013年度末時点で、54.7%である [1]。これは、2010年度末時点で、9.7%であったのと比較すると、わずか3年間で約5.6倍の普及率とである。結果として、UGC (User-generated Content) や、CGM (Consumer Generated Media) と呼ばれる、一般のユーザがそれらの機材により生成したコンテンツが増加している。

UGC や、CGM の中でも、ユーザによって撮影された写真¹は、急速に増加している²。人々は、旅行中に訪れる観光スポットにおいて、そこに存在する有名な建築物や、レストランに訪れたときに注文した商品、花火大会などのイベントの様子など、自分が興味を持った様々なものを大量に撮影している。加えて、それらの写真を、Flickr [4] や、Panoramio [5]、Facebook [6] などのソーシャルメディアサイトにアップロードして、共有しているユーザも多い。実際に、Flickr では、2013年3月時点で、8,700万人のユーザがサービスに登録しており、約350万枚の写真が日々アップロードされている [7]。また、2011年4月時点で、Flickr に60億枚以上の写真がアップロードされている [8]。同様に、Facebook では、2012年8月時点で、1日に約3億枚以上の写真がアップロードされている [9]。結果として、これらの膨大な画像を活用するために、ソーシャルメディアサイトなどのウェブ上に存在する膨大な数の画像を検索するための技術の需要が高まっている [10]。

膨大な数の画像からユーザの求める画像を検索するための技術として画像検索

¹本論文において、写真は、デジタルカメラなどで撮影してメディアに保存された画像データのことを表している。

²UGC と CGM として、写真以外にも、動画、ブログ記事、Twitter [2] などのマイクロメディアの記事、口コミサイトなどのレビューなどがあげられる。また、それらに加えて、口コミサイトや、ブログサイトなどの情報を集約共有するウェブサイトそのものも含める場合がある [3]。

が盛んに研究されている。画像検索は、ユーザが入力する検索クエリに基づいて、大きく分けて2種類存在する。1つは、キーワードをクエリとする検索である。ユーザはテキストを入力することで、それを画像と関連付けられたテキスト情報と比較して検索する。このような検索は、TBIR (Text-based image retrieval) と呼ばれる [10, 11, 12]。多くの画像検索システムは、これに該当し、Google [13] や、Yahoo! [14] の画像検索が例としてあげられる。テキスト情報として、ウェブページの場合は画像の周辺に出現するテキスト、ソーシャルメディアサイトに含まれる画像の検索では、ソーシャルタギングによるタグが用いられることが多い。特に、ユーザがクエリとしてタグを入力することで、画像に付与されているタグに基づいて検索を行うのは、TagIR (Tag-based image retrieval) と呼ばれる [15, 16]。しかしながら、欠点として、テキストや、タグを画像に付与する必要があるため、人手が必要なことや、たとえ付与されていたとしても、それらの正確性が問題となる。加えて、画像の特徴を考慮していない点があげられる。また、画像の意味内容ではなく、画像から人々が受ける印象に基づいた検索として、感性語をクエリとした感性検索があげられる。感性検索は、室内ファッションなどのデザイン支援や、美術館や、博物館などでの画像検索が主なニーズである [17]。もう1つは、画像をクエリとする検索である。データベースに含まれる画像から特徴量を抽出して、それとクエリ画像から抽出した特徴量を比較することで、画像の見た目の情報に基づいた検索を実現している。用いられる特徴量としては、オブジェクトの色、場所や、型、画像全体のエッジや、テクスチャなどが用いられることが多い。このような画像の特徴量を用いた検索は、CBIR (Content-based image retrieval) と呼ばれる [18]。CBIR は、画像の見た目に基づいて検索されるため、ユーザが自身の求める画像をクエリとして用いることができた場合は、有効な検索手段である。CBIR で用いられるクエリとして、画像をクエリとする以外にも、QBS (query-by-sketch) と呼ばれる、ユーザがスケッチを描いて検索する手法も提案されている [19, 20]。

1.1.2 画像検索結果

ユーザが入力したクエリに適合する画像が、画像検索結果としてユーザに提示される。大量の画像を含むデータベースから検索した場合は、クエリに適合する画像数の増加に伴い、検索結果に含まれる画像数も増加する。結果として、ユーザが閲覧しなければならない画像の数が増加するため、それらの全ての画像を閲覧することは、ユーザにとって非常に手間である。実際に、ユーザは検索結果の全て

を閲覧することは非常に稀である。例えば、ウェブページの検索結果において、約75%のユーザは、検索結果の3ページ目までしか閲覧しない [21]。これは、画像検索結果においても同様であると考えられるため、ユーザは検索結果に含まれる僅かな画像しか閲覧しないと考えられる。ほとんどの検索システムにおいて、検索結果は、検索クエリとの類似度に基づいたランキング結果をリスト形式でユーザに提示する。リスト形式での表示は、ユーザが必要とする情報が明確で、検索結果の再現率よりも適合率を優先したい場合は有効である [22, 23]。例えば、Windowsのロゴマークの画像を探したい場合は、これに該当する。ユーザが必要な情報が明確で、すべての関連する結果を閲覧せずに上位の数データの閲覧することが想定されるため、検索クエリとの適合率を重視したリスト形式での表示が適している。

しかしながら、検索結果のリスト形式による表示は、画像検索結果の視覚的多様性を保つのが困難であることが指摘されている [12, 23, 24, 25, 26, 27]。検索結果の多様性とは、検索結果に同じような結果しか含まれないような、検索クエリに対する過剰な適合を防ぐべきであると考えに基づいた、検索結果の観点のことである [28]。これを画像検索に置き換えたものが、視覚的多様性である。例えば、特定のブランドの画像を検索した場合、画像検索結果には、そのブランドの商品のほとんど同じような画像が多く含まれる。また、ある観光スポットの画像を検索した場合、同じようなアングルと距離で撮影されたような画像が多く含まれる。このような画像検索結果は視覚的多様性が低いとされ、画像検索結果の閲覧がユーザにとって負担となる [24]。そのため、画像検索結果は、クエリに対する視覚的多様性と高い適合率を両立する必要がある [29, 30]。検索結果の多様性を向上させることは、ユーザの検索結果に対する満足度の向上につながるとされる [31]。

キーワードを用いた検索において、視覚的多様性を低下させる原因として、検索結果の上位にユーザが求める画像を出現させるようなキーワードをクエリとして指定するのが難しいことがあげられる。ユーザによって生成されたキーワードが複数の観点を含んでいる場合や、曖昧な場合がある [24, 32, 29, 33]。例えば、画像検索において、キーワード“apple”で検索した場合、Apple社の製品に関連する画像と、果物のりんごの画像が画像検索結果のほとんどを占める。しかしながら、ユーザが“apple”の画像を検索する際に、これらの2種類の画像のどちらかのみを検索したい場合がほとんどであると考えられる。さらに、ユーザが求める検索結果を得るために適した一般的なキーワードを、ほとんどのユーザは明確にクエリとして指定することはできない [34]。実際に、ユーザの指定するキーワードの多くは、非常に短く、90%以上のクエリに含まれる単語数は、4単語以下である

[35, 36]. これらのことから、画像検索の際に、ユーザは曖昧なキーワードを入力するため、ユーザの求めている画像を多く含む画像検索結果を閲覧することになる。また、曖昧な検索クエリを入力した場合に、類似したデータの位置が固まって表示されないことがある [33]. そのため、ユーザの求める画像と、求めている画像が混同した画像検索結果を閲覧することとなる。結果として、期待とは異なる画像を多く含む検索結果を閲覧することになり、ユーザが必要とする画像を容易に探し出すことが困難となる [37].

そのため、ユーザの検索結果に対する満足度を向上させるため、画像検索結果の閲覧を支援する研究が行われている [38]. 画像検索結果の閲覧を支援するための主なアプローチは、大きく分けて2種類ある。

1つは、ユーザの入力した検索クエリに対する画像検索結果のランキングをリランキングしてユーザに提示するアプローチである。画像検索結果をリランキングすることで、ユーザの求める画像をランキングの上位に出現させるアプローチである [39, 40, 27, 41]. リランキングを実現するために、画像検索結果内のデータのみを用いる手法 [40], 検索結果に対するインタラクション [42], ならびに Wikipedia [43] などの検索結果以外のデータを用いたものなどがあげられる [27, 39].

もう1つは、画像検索結果に対して画像クラスタリング手法を適用することで、類似した画像をグループ化してユーザに提示するアプローチである [24, 44, 45, 46, 47, 48, 25, 47, 33]. このアプローチは、検索結果の画像をすべて提示、または、ランキングの上位の画像のみを提示するのではなく、それぞれのクラスタから抽出された代表画像をユーザに提示する場合が多い。そして、ユーザは、提示された代表画像を選択することで、そのクラスタに属している画像を閲覧することが可能になる。このアプローチでは、曖昧なキーワードによる検索結果から、視覚的多様性に富む画像検索結果を得ることができる。ユーザにとっては、画像検索結果に含まれる多様な画像の代表的な部分のみを閲覧することで、自身の求める画像を見つけることができる。クラスタリングを作成する特徴量としては、画像特徴量や、ソーシャルタギングによるタグが用いられることが多い。

これらの2つのアプローチを比較すると、画像検索結果に対してリランキングを適用する場合は、ユーザの求める画像をランキングの上位に出現させることを目的としているため、ユーザの求める画像が明確な場合は有効である。しかしながら、曖昧なキーワードを用いた検索の場合、画像検索結果の視覚的多様性を保つことは困難である。また、それをランキング手法の改善のみによって解決することは困難である。一方、画像検索結果に対して画像クラスタリングを適用する場合、類似した画像がクラスタにグループ化され、代表画像をユーザに提示するた

め、曖昧な検索クエリや、ユーザ自身が求める画像が明確でない場合などに、画像検索結果に含まれている画像の種類の外観を得ることができると考えられる。そのため、本論文では、画像クラスタリングを用いて、画像検索結果の閲覧を支援する手法の提案を行なった。

1.1.3 画像検索結果のクラスタリング

画像クラスタリングは、前述したように画像検索結果をユーザが効率的に閲覧するための有効な手段の1つであるが、画像の見た目のみを考慮することで、画像の意味を十分に考慮するクラスタリング結果の作成は困難であることが指摘されている [49]。そのため、単にそれらの画像の見た目の違いだけではなく、それらの画像が持つ意味的な違い、すなわち、セマンティックギャップを考慮する必要がある。セマンティックギャップとは、画像の見た目が意味するところと、画像の意味内容が直接的に結びつかないことである。セマンティックギャップの克服は、画像の意味内容の理解や、画像検索における重要な課題として知られている [50, 51, 52, 53]。例えば、図 1.1 は、Flickr においてシロクマについて検索した画像検索結果であるが、ぬいぐるみと実際の動物の2種類のシロクマの画像が含まれていると考えることが可能である。視覚的多様性を考慮する場合、これらの画像がクラスタリングされた際には、異なるクラスタに属するべきであると考えられる。しかしながら、実際には、画像特徴量のみを用いてクラスタリングを行なった場合、それらの画像を区別することは困難なことが多い。そのため、画像の見た目のみを考慮してクラスタリング行なった場合、画像の意味内容と直接結びつかない場合や、画像の意味内容を複数通りに解釈できる場合、本来は異なる解釈としたい画像同士が1つのクラスタに属するように扱われることがある [54]。

画像特徴量のみに基づいた画像クラスタリングでは、セマンティックギャップなどの影響によってユーザの求めるクラスタリング結果の作成が難しいため、ソーシャルタギングによるタグや、Exif (Exchangeable image file format) などのメタデータが用いられることがある [34, 46, 53, 49, 33]。画像特徴量とタグに基づく特徴量などの複数の特徴量を用いる場合に、単純にそれらの特徴量を組み合わせることが、クラスタリングの質を向上させるとは限らない。それどころか、複数の特徴量を考慮したために、クラスタリングアルゴリズムが見当違いなクラスタリング結果を作成する場合がある [55]。

また、画像クラスタリングにソーシャルタギングを用いる際の課題には、画像に付与されているメタデータの不均質さがあげられる。本論文における、メタデー

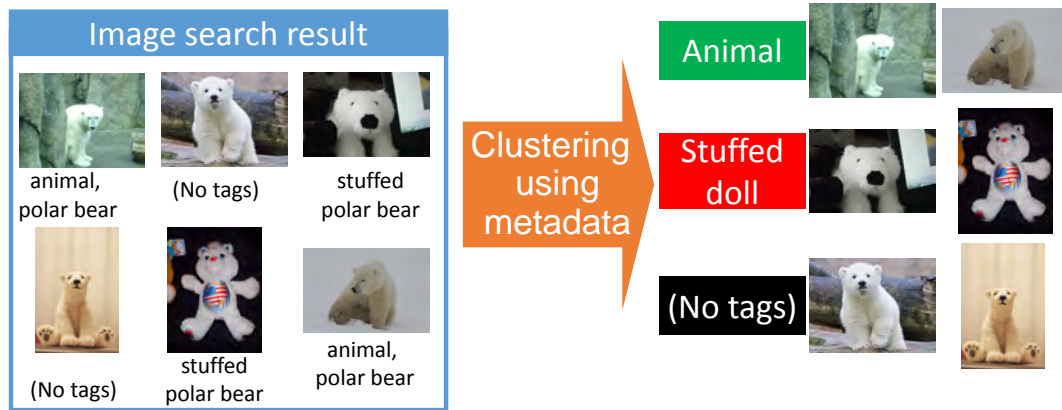


図 1.2: タグが付与されていない画像を含む画像検索結果のクラスタリングの例

が付与されていなかった場合、画像の意味内容を区別することは困難である。言い換えると、図 1.2 において、タグが付与されていない 2 枚の画像 (“No tags”) にたとえ幾つかのタグが付与されていたとしても、他の画像に付与されているような、タグ “animal” や、“stuffed” などが付与されていないければ、それらの画像がどのような意味内容であるかを区別することは困難である。同様に、図 1.2 において、タグ “animal” や、“stuffed” が付与されている画像に、それらのタグが付与されていなければ、実際の動物のシロクマと、ぬいぐるみのシロクマを別々のクラスタにすることは困難である。結果として、画像に十分にタグが付与されていない場合は、クラスタリング結果の質を低下する可能性がある。

もう 1 つの課題は、ソーシャルタギングなどのユーザによって自由に付与されるタグは、必ずしも正確でない場合や、不十分な場合への対処である。そのため、画像間の類似度をソーシャルタギングによるタグを用いて算出するのは、タグの質を考慮する必要がある [34, 56]。実際に、Flickr において、画像とそれに付与されたタグが関連しているのは、約 50% である [57]。また、画像に付与された地理的なタグと、その画像に付与されている緯度経度情報が一致しているかについては、“beach”、“coastline” などの海岸線に関するタグの場合、海岸線から 100m 以下では約 51.9% である、また、500m 以下では約 77.7% である [58]。そのため、画像と付与されているタグが、正確に一致している場合は多くない。最もよく発生すると考えられるのは、単語の表記ゆれである (例: “Louvre”, “Le Louvre”, “Louvre Museum”)。それらのタグが付与された画像を含む画像検索結果をクラスタリングした場合、それらの画像間の類似度を正確に算出することは困難である。結果として、類似度が正確でないため、目的とするクラスタリングの結果を得られない場合がある。例えば、図 1.2 において、2 枚の画像の類似度を単語の共起に基づ

いて算出する場合について考えると、2枚の画像の両方にタグ“polar bear”が付与されている場合と比較して、それらの2枚の画像の1枚に“polar bear”ではなく“white bear”が付与されていた場合、2つの画像の類似度は低くなる。他には、単語の表す意味は同じだが言語の異なるタグ (“Japan”, “日本”), これまでに用いられていなかった新しいタグや、画像の意味内容に関係の無いタグ (場所, 時間, 撮影機材などに関するタグ) があげられる [45].

画像に付与されているメタデータの不均質さ (画像の欠損しているタグ, 不正確なタグ) に対する対処として、画像のタグを修正, および推定するための手法が用いられることがある [59, 60, 61]. そのため、本論文で扱う画像検索結果に対する画像クラスタリングの質を向上させるために、有効な手段の1つであると考えられる。画像に対してタグを推定する場合、多くの研究では、事前に準備したトレーニングデータに対して教師あり学習手法を適用することで、画像のタグを推定する。このトレーニングデータには、NUS-WIDE [62] や、corel などの画像にタグが付与されているデータセットを用いる場合が多い。しかしながら、本論文のようにソーシャルメディアサイトから取得した画像検索結果に含まれる画像のタグの推定修正の場合は、事前にトレーニングデータを準備することには2つの課題がある。1つは、新しく発生したイベントや、トピックに対する対処である。これは、ソーシャルメディアサイトには、日々、ユーザが撮影した画像が投稿されているため、トレーニングデータを作成した後に発生したイベントトピックに関する画像が投稿される。また、それらの画像には、新しいイベントなどに関するタグが付与されている。それらのイベントなどが検索対象とされた場合に、それらにも対応可能なトレーニングデータを事前に作成することは困難である。もう1つは、ユーザの入力する検索クエリの多様性の高さや曖昧性に対する対処である。ソーシャルメディアサイトの画像をユーザが検索する場合、ユーザは検索クエリを自由に入力するため多様性と曖昧性が非常に高い。そのため、どのような検索クエリが入力されたとしても対応可能なトレーニングデータの作成が望ましい。しかしながら、トレーニングデータを用いる場合、トレーニングデータを適用できる範囲に限界があるため、トレーニングデータよりも一般的なケースに対応するのが難しいことが指摘されている [59]. そのため、ソーシャルメディアサイトから取得した画像検索結果に含まれる画像のタグを推定・修正する場合は、トレーニングデータを必要としない手法が必要である。

そのため、本論文では、メタデータの不均質さによって生じる画像クラスタリングの性能の低下を防ぐため、画像検索結果以外の教示用のデータを必要としないメタデータの推定に基づいたクラスタリング手法を提案した。提案するメタデー

タの推定・修正手法は，本論文では，画像検索結果に対して適応し，評価しているが，その他の手法にも適応可能であると考えられる．近年，Web上のメタデータの増加に伴い，メタデータを用いてクラスタリングなどを行うことが増加しているため，その際に解決すべきメタデータの課題の1つであるメタデータの不均質さへの対処が重要となっている．例えば，本論文で後述する写真の審美的評価のためのリランキング手法，および地理的なクラスタとその関連性の抽出手法でもメタデータを特徴量として用いているが，そのような場合にもメタデータの不均質性は課題となる．結果として，本論文で提案するメタデータの推定・修正手法は，それらの課題を解決するための1つの手段として重要であると考えられる．

1.1.4 写真の審美的評価

ユーザに画像検索結果を提示する際に，写真の審美的な評価を考慮することも重要である．ここで，本論文で扱う写真の審美的評価とは，画像の圧縮による生じるアーティファクト(リングングなど)や，ノイズによる劣化ではなく，画像の見た目から受ける審美的な印象のことである．前述したように，ソーシャルメディアサイトに一般のユーザの写真が大量にアップロードされている．しかしながら，それらの写真の多くは，プロの写真家が撮影したものではなく，一般のユーザが撮影した写真である．また，一般のユーザの場合は，写真が趣味であるユーザや，そうでないユーザなど，ユーザの用いる機材や，写真を撮影するための技術と知識に差がある場合がある．そのため，多くのユーザが写真をアップロードしているソーシャルメディアサイトなどでは，それらの写真が混在している．しかしながら，ユーザによって，写真を撮影するための技術，知識，および機材には差があるため，ソーシャルメディアサイトにアップロードされている写真の審美的評価は，不均質である．結果として，ソーシャルメディアサイトにアップロードされている全て写真の審美的評価が高いわけではない．そのため，画像検索結果に審美的評価の低い写真が含まれる場合があり，それを閲覧するのはユーザにとって負担である．実際に，画像を検索する際に，60%以上のユーザは，画像検索結果として，写真の質を重視する [63]．

これらのことより，画像検索や，画像の閲覧のための，写真の審美的な評価手法の研究が行われている．主な手法としては，画像処理技術に基づいた写真の構図の評価があげられる [63, 64, 65, 66, 67, 68]．それらの研究で用いられている構図として，rule of thirds, simplicity, visual weight などがあげられる．例えば，Rule of thirds は，写真の被写体の構図の1つで，等間隔に引いた水平線と垂直線

2本ずつにより画面を9等分した交点に被写体を配置するというものである。そして、機械学習を用いて、トレーニングデータからこれらの構図を学習して、テストデータを評価する手法が一般的である。しかしながら、大量の写真に対して適用する場合、高度な画像処理技術が用いられるため、データの処理量が非常に多くなることが指摘されている [69]。加えて、前述したように、ソーシャルメディアサイトに含まれる写真の多様性によって、十分なトレーニングデータの準備が困難なことがあげられる。そのため、本論文では、画像検索結果において、画像処理技術によらずに、トレーニングデータも用いない写真の質を評価する手法を提案する。

1.1.5 地理的なクラスタとその関連性の抽出

近年、ソーシャルメディアサイト上の画像には、撮影位置情報が付与されているものも多い。人々が多くの写真を撮影した地域には、多くのユーザが興味・関心のある領域が含まれている、または周辺にそのような地域があると考えられる。このような分析は、観光産業や、都市分析、人々の行動を分析するのに有効である。しかし、撮影位置情報が付与された画像は大量に存在するため、そのような分析の際には画像を効率的に閲覧可能にする必要がある。そのため、ソーシャルメディアサイトに投稿された撮影位置情報が付与された写真を用いて、地理的なクラスタ(ホットスポット)を抽出する研究が行われている [70, 71, 72]。ホットスポットとは、多くの写真が撮影されている地域のことである。ホットスポットは、ホットスポットに含まれる写真の撮影方向情報や、ホットスポットの位置に基づいて、関心領域と撮影スポットの2種類に分類することができる。関心領域とは、観光スポットのような、多くの人々が関心を持つ領域のことを指す。撮影スポットとは、遠方の被写体を撮影するため、多くの写真が撮影されている地点のことを指す。しかし、それらのホットスポットには、関連性が存在する場合がある。たとえば、1つのイベントに関連するホットスポットが複数抽出された場合である。このような場合に、ユーザにそれぞれの独立したホットスポットを閲覧させるのは、負担となる。そのため、本論文では、ソーシャルメディアサイトに投稿された撮影位置情報が付与された写真からホットスポットを抽出し、抽出されたホットスポット間の関連性を抽出する手法を提案する。提案手法で抽出する関連性は、撮影スポットと被写体の関連性と、タグに基づいた関連性の2つの種類である。また、本論文では、ソーシャルメディアサイトから取得した画像に提案手法を適用し、地図上に抽出した関連性を可視化することで、その結果を考

察した。

1.2 本論文の構成

2章では、本論文に関連する研究について述べる。画像検索結果に含まれる画像のメタデータの不均質性について、メタデータを推定することで、それを解決することを試みる手法についてまとめる。画像検索結果に対するクラスタリング・リランキング手法についてまとめる。本論文では、視覚的多様性、審美的評価、地理的多様性について扱う。

3章では、画像検索結果のクラスタリングの際に、メタデータが欠損していた場合にも有効にクラスタリングが行えるようなメタデータの推定手法について述べる。提案手法では、推定を k 近傍法を適用することでタグと Exif を推定し、その結果を修正することで、最終的に適切なメタデータの推定を行う。また、画像検索結果に対して、提案手法を適用して推定されたメタデータを用いてクラスタリングを行う。そのクラスタリング結果と、人手によるクラスタリング結果の類似度を評価することで、提案手法の有効性を評価する。

4章では、写真の審美的評価に基づいた画像検索結果のリランキング手法について述べる。提案手法では、画素情報を用いずに、画像に付与されている Exif、および撮影機材の情報を用いてリランキングを行う。人手による写真の評価結果と提案手法のリランキング結果の類似度を評価することで提案手法の評価を行う。また、実行時間とリランキングの性能について、既存手法と比較を行うことで、提案手法の有効性を示す。

5章では、写真の撮影地点の分布に基づいて、地理的なクラスタを抽出し、そのクラスタ間の関連性を抽出する手法について述べる。人々の興味・関心の可視化するために、ソーシャルメディアサイトから取得した撮影地点を表すメタデータである緯度経度情報が付与されている写真を用いて、多くの写真が撮影されている地域をホットスポット(クラスタ)として抽出する。実際に、ソーシャルメディアサイトから取得した画像に提案手法を適用し、その結果を考察した。

6章では、本論文の成果と今後の展望について述べる。

第2章 関連研究

2.1 撮影状況メタデータの利用

画像の撮影状況を考慮するためのメタデータとして Exif がある。これは、デジタルカメラや、スマートフォンなどで撮影された画像に付与されてるメタデータである。記述されるメタデータの例として、写真の撮影日時、機種名、シャッタースピード、焦点距離などがあげられる。加えて、GPS 機器や磁気センサが搭載された機材では、写真の撮影位置を表す緯度経度情報や、撮影方向情報などのジオタグ情報が Exif として自動的に付与される。

撮影状況メタデータを用いた手法がいくつか提案されている。たとえば、個人の写真を整理するための研究である。[73] は、フォトストリームを分割するために、隠れマルコフモデルの特徴量として、[74] の提案した手法、写真が撮影された時の絞り径、色ヒストグラム、およびタイムギャップを利用している。[75] は、絞り、露光時間、焦点距離、撮影時刻などを考慮した EM アルゴリズムを用いて、個人の写真のコレクションからイベントを発見する手法を提案した。また、ソーシャルメディアサイト上に多量に存在する緯度経度情報や、撮影方向情報を用いた研究が行われている。撮影状況メタデータを用いた研究として、[74] は、Exif に含まれる絞り、露光時間、ISO 感度、および焦点距離を利用して、写真が撮影された周辺の明るさを推定する手法を提案している。本論文の提案手法でも、この撮影状況メタデータである Exif 利用した画像のクラスタリングや、リランキングを行う。また、写真に付与されている Exif は不均質である場合や、欠損している場合があるため、画像検索結果のクラスタリングを目的とした Exif の推定を行う。

2.2 画像クラスタリングを用いた画像検索結果の提示

これまでに提案されている、画像クラスタリングを用いたユーザに視覚的多様性に富む画像検索結果を提示するための研究についてまとめる。この研究は、画像クラスタリングに用いる特徴量に基づいて、主に3つのアプローチに分類できる。

1つめは、テキスト情報をのみを用いたアプローチである [48, 76, 77]. これらの手法では、ウェブページに出現する画像の周辺のテキストや、画像に付与されているタグなどを特徴量として用いてクラスタリングを行なっている. [76] は、画像を含むウェブページのテキストや、DOMの構造などを特徴量として、K-meansの欠点の1つであるクラスタ数を決定するパラメータの調整が必要のない Yet another K-means を利用してクラスタリングを行なう手法を提案した. [48] は、はじめに、画像検索ではなくウェブページの検索を行い、その結果に含まれるウェブページの n-grams を特徴量として、[78] で提案されているクラスタリング手法を適用することで、ユーザの入力したキーワードに関連する単語を抽出する. さらに、そのクラスタリング結果から、クラスタを代表する単語を抽出する. それらが類似していた場合は、クラスタを結合することで、得られた単語をキーワードとして画像検索を行い、その結果をユーザに提示する手法を提案した.

2つ目は、画像特徴量のみを用いたアプローチである [24, 79, 80]. これらの手法では、色ヒストグラム、エッジ、テクスチャなどの画像の見た目に基づいてクラスタリングが行われている. [79] は、Color Moments [81] を画像特徴量とした、Affinity Propagation [82] を疎なデータセットにも対応できるように改良した Fast Sparse Affinity Propagation を用いて画像検索結果をクラスタリングしている. [24] は、視覚的多様性に富む画像検索結果をユーザに提示するために、6つの画像特徴量に基づいた画像間類似度を利用した、画像検索結果に対する3つのクラスタリング手法 (Folding, Maxmin, Reciprocal election) を提案している. [80] は、画像検索結果から抽出されるべきクラスタのスケールの違いに対応可能な Dynamic absorbing random walk を提案している. しかしながら、これらの2つのアプローチは、テキスト情報や、画像の見た目のみを考慮して画像クラスタリングを行なっているため、第1章で述べたようなセマンティックギャップを克服することが困難である.

最後に、複数の特徴量を組み合わせたアプローチである [33, 49, 53, 34]. これらの手法の多くは、前述したテキスト情報と、画像特徴量の2つに基づいた画像クラスタリングを適用することで、ユーザに画像検索結果を提示している. [33] は、画像検索結果から、画像特徴量、テキスト情報、およびハイパーリンクの3つの特徴量を抽出することでそれぞれのクラスタリング結果を作成し、それらをユーザに提示した. [49] は、テキスト情報から意味的なクラスタを作成、画像特徴量から視覚的なクラスタを作成することで画像検索結果の提示を行う手法を提案した. はじめに、ユーザの入力したキーワードを用いて得たウェブページからキーフレーズを抽出し、K-lines-based clustering [83] に基づいたクラスタリング手法を適用し

て、意味的なクラスタを作成する。次に、抽出されたキーフレーズで画像検索を行い、それらの画像の画像特徴量に基づいて Bregman Bubble Clustering [84] を適用することで視覚的なクラスタを作成する。これらの結果を階層的にユーザに提示することで、多様性に富む画像検索結果をユーザに提示している。[53] は、画像の周辺のテキストと、画像特徴量に基づいた制約付き多目的最適化を解決することにより画像検索結果のクラスタリングを行なっている。[34] は、density-based なクラスタリング手法である Shared Nearest Neighbor [85] を画像に付与されているタグと、SIFT [86] による Bag of visual words に基づいて適用できるように改良することで、画像検索結果をクラスタリングする手法を提案している。

しかしながら、これらのメタデータを用いたクラスタリング手法では、画像にメタデータが付与されていない場合に、画像間の類似度を正確に算出することができないことが原因で、クラスタリングの性能が低下する場合がある。また、たとえメタデータが付与されたとしても、正確ではないメタデータが付与されている場合がある。そのため、本論文では、画像に不均質なメタデータが付与されている場合や、メタデータが欠損している場合でも、クラスタリングの性能が低下することを防ぐために、メタデータを修正推定する手法を提案する。

本論文で提案する画像クラスタリングを用いた画像検索結果の提示手法では、3つ目のアプローチである複数の特徴量を組み合わせて画像検索結果を提示する。用いる特徴量は、撮影状況を表すメタデータである Exif、ソーシャルタギングによるタグ、および画像特徴量の3つの特徴量である。前述した手法と比較して、提案手法では、Exif を考慮することで、写真の撮影された状況という新たな観点に基づいて画像をクラスタリングする。また、前述した複数の特徴量を考慮することによるクラスタリングの性能の低下を防ぐために、制約付きクラスタリングを用いることで、視覚的多様性に富む画像検索結果を提示するための手法を提案した。

2.3 画像のメタデータ推定

画像に付与されているメタデータを用いたクラスタリングの課題として、メタデータの質や、欠損によってクラスタリングの結果が変わる可能性がある。そのため、メタデータの有無や質に基づく、クラスタリングへの影響を少なくすることはメタデータを用いたクラスタリングの課題の1つである。このような課題を解決するためにこれまでに提案されている、画像のメタデータを推定するための手法についてまとめる。

[56] は、ソーシャルメディアサイトに含まれる画像のタグを修正するために、画

像特徴量を用いて、4つの側面で評価し、凸最適化の手法を適用することで、画像に対して適切なタグを推定している。[59]は、訓練データから準同一クラスタを作成し、そのクラスタリング結果に基づいて重み付け相関ルールを適用することでタグを推定する手法を提案している。抽出された相関ルールに基づいてテストデータの画像にタグを付与する。この方法は、ノイズとなるタグが多く付与されるため、そのタグを最もよく表す画像との類似度を考慮しノイズとなるタグを削除する。[60]は、フォークソノミーに基づいて、ユーザによるタグを objective tags と subjective tags の2つのカテゴリに分け、それに基づいてタグの修正を行う手法を提案している。

これらの手法は、いずれもある程度大規模のデータに対して適用することを前提としており、本論文で扱うような数量の限られたデータをもとにメタデータの推定を行うのには適していない。これらの手法では、トレーニングデータを用いた機械学習により、メタデータの推定修正を行っているが、前述したように、本論文のように、画像検索結果のメタデータを推定する場合には、適さない。これは、検索クエリの多様性によって得られる画像検索結果を想定することが困難であり、結果として、適切なトレーニングデータを用意するのは現実的ではないためである。

そのため、本論文では、メタデータの欠損によって、クラスタリングの性能が低下しないようなメタデータの推定手法を用いた画像検索結果の提示手法を提案した。提案手法では、各画像の正しいメタデータを推定するため、画像検索結果として与えられたデータのみを用い、それ以外に別途教示用の学習データ等を必要としない。

2.4 写真の審美的評価

写真の審美的評価を評価する多くの研究では、三分割法などの写真に含まれる被写体の構図に基づいた評価を行うため、画像処理や機械学習を適用している場合が多い。[87]は、写真のレイアウトについて、画像に含まれる複数の画像特徴量に基づいて空や、被写体の位置を分析することにより写真の審美的評価を行う手法を提案した。[88]は、写真に含まれる被写体の配置や、写真の空と地上の比率などの構図に関する評価を学習することで、写真の審美的評価を行う手法を提案した。また、単一の被写体が写っている写真等について、写真の審美的評価を写真の構成に基づいて評価し、よりその評価が高くなるように写真の被写体の配置を再構成する。[63]は、写真の役割について着目し、トレーニングデータセット

に基づいて写真の役割について学習することで、画像検索結果を写真の審美的評価に基づいてリランキングを行う手法を提案した。[89]は、人手によって写真の審美的な評価を評価したトレーニングデータセットと、写真の被写体と写真の顕著性マップを用いて、写真の質を評価する手法を提案した。[90]は、画像を背景と被写体を分けるのではなく、色空間、テクスチャ、エッジなどの情報から bag-of-aesthetics-preserving features を構築することで写真の審美的な評価を評価する手法を提案した。

これらの手法では、いずれも質の高い写真のトレーニングデータセットを事前に準備する必要があるが、前述したように、画像検索結果に適応する場合は、事前にトレーニングデータセットを用意するのは困難である。また、これらの手法は、審美的に評価したい写真に複数の画像処理手法を適用する必要があるため、処理時間が非常にかかる。そのため、本論文では、事前にトレーニングデータセットが必要なく、画像処理の代わりに画像に付与されている撮影状況メタデータを用いる画像検索結果の写真の審美的な評価手法を提案した。

2.5 写真の撮影位置情報に基づいたクラスタリング

2.2節では、視覚的多様性に富む画像検索結果をユーザに提示することを目的としたクラスタリングについて述べたが、ここでは、画像を写真の撮影位置情報に基づいて地理的な画像のクラスタ(ホットスポット)を抽出する手法についてまとめる。Crandallら[70]は、Web上の大量のジオタグ付き写真と写真の画像特徴を用いてクラスタリングを行うことにより、ホットスポットや関心領域が存在する地域が得られることを示した。Crandallらは、ホットスポットや関心領域の発見には、写真の撮影位置情報が特に重要であると述べている。Kisilevichら[71]は、写真の撮影位置情報および写真の撮影者情報から、写真の密集度に着目したクラスタリング手法を提案し、ホットスポットに該当する地域を発見した。Shiraiら[72]は、写真の撮影位置情報からホットスポットを抽出し、ホットスポットに含まれる写真の撮影方向情報からホットスポットの種類を関心領域や、撮影スポットに分類する手法を提案した。これらの手法では、抽出されたホットスポットは、独立したホットスポットとして抽出される。そのため、本論文では、ホットスポットの間の関連性をホットスポットに含まれるソーシャルタギングによるタグや、撮影方向情報から抽出する手法を提案する。

第3章 メタデータの欠損に頑健な 画像検索結果のクラスタリ ング

3.1 概要

メタデータを用いたクラスタリングは、画像検索結果を効率的に閲覧するために有効な手段のひとつである。しかし、画像に対して十分にメタデータが付与されていない場合も多く、そのことが、クラスタリングの性能を低下させる可能性がある。本論文では、メタデータが十分に付与されていない画像検索結果に対しても有効にクラスタリングが行えるように、画像のメタデータの欠損を推定する手法について述べる。本論文で提案するメタデータ推定手法の目的は、一般的な画像のメタデータを推定する研究とは異なり、その画像に付与されるべきメタデータを正確に推定することではなく、画像クラスタリングが適切に行われるメタデータを推定することである。

画像の最適なメタデータを推定するため、一般的なメタデータの推定手法では、あらかじめメタデータが付与された画像を大量に用意したトレーニングデータを利用する。しかしながら、画像検索結果にそれらの手法を適用する場合は、ユーザの入力する検索キーワードの多様性や、得られた画像検索結果に含まれる画像の多様性から、適切なトレーニングデータを事前に準備することは困難である。そのため、本論文では、画像検索結果として与えられたデータのみを用い、それ以外に別途教示用の学習データを必要としないメタデータの推定手法を提案する。提案手法では、画像検索結果に対して k 近傍法を適用することで、メタデータを推定し、推定されたメタデータの中でクラスタリングに寄与しないメタデータを削除する。提案手法において、推定の対象とするメタデータは、ソーシャルタギングにより付与されたタグと、撮影状況を表すメタデータである Exif の 2 つである。提案手法をソーシャルメディアサイトから取得した画像検索結果に適用することで、本手法が、クラスタリングの性能向上に寄与するようなメタデータを効果的に推定していることを示す。

Algorithm 1 Algorithm for Metadata estimation

INPUT: Image search result P

```
1: for all  $p_i \in P$  do
2:   if  $|t_i| = 0$  then
3:      $t_i \leftarrow kNN(p_i, P)$ 
4:   end if
5:   if  $|e_i| = 0$  then
6:      $e_i \leftarrow kNN(p_i, P)$ 
7:   end if
8: end for
```

図 3.1: 欠損したメタデータを推定するアルゴリズム

3.2 類似画像を用いた画像検索結果のメタデータの推定

提案手法では、画像検索結果 $P = \{p_0, p_1, \dots, p_n\}$ のメタデータの付与されていない画像 p_i に対して、タグ、Exif の2種類のメタデータを推定する。タグの推定の場合は、 k 近傍法を適用することで、類似する画像のタグを用いて画像のタグを推定したのちに、クラスタリング手法を適用することで、適切でないタグを削除する。また、Exifの推定の場合は、タグの推定と同様に、 k 近傍法を適用することで、Exifの値を推定したのちに、クラスタリング手法を適用することで、Exifの値を修正する。今回、メタデータの推定手法として k 近傍法を用いるが、学習に用いるデータには画像検索結果のみを用い、別途トレーニングデータ等を用いない。

提案手法では、ある画像 p_i にタグが付与されていない ($|t_i| = 0$) 場合、類似する画像を用いて、適切なタグ t_i を推定する。しかし、今回は、 k 近傍法で他の画像に付与されている全てのタグが、画像に付与される可能性があるため、画像の内容を表さないタグを修正する必要がある。これは、画像検索結果の画像に初めから付与されているタグや、推定されたタグの中には、クラスタリングの性能の低下に繋がるタグが含まれている可能性があるためである。例えば、表記のゆれや、内容に関連の無いタグである。そのため、クラスタリングの性能を低下させる可能性が高いタグを削除する。

本論文で提案するメタデータの推定のためのアルゴリズムを図 3.1 に示す。メタデータが欠損している画像 p_x を含む画像検索結果 P に対して、他の画像から適切なメタデータを選択し、 p_x に付与するため、 k 近傍法を適用する。そのため、画像 p_x をテストデータとし、画像 p_x 以外の画像 R を用いて k 近傍法を適用す

Algorithm 2 Algorithm for removing noisy metadata
INPUT: Image search result P , thresholds α, β

```
1: for all Image sub set  $P'$  has  $w_i \in T$  do
2:   Clustering result  $C \leftarrow Clustering(P')$ 
3:   for all Cluster  $c_i \in C$  do
4:     if  $|c_i| < \alpha * |P'|$  then
5:       for all  $p_i \in c_i$  do
6:         delete  $w_i$  in  $p_i$ 
7:       end for
8:     end if
9:     if  $|c_i| < \beta * |P'|$  then
10:      for all  $p_i \in c_i$  do
11:        if  $e_i$  is null then
12:           $e_i \leftarrow ave(e_i \in c)$ 
13:        end if
14:      end for
15:    end if
16:  end for
17: end for
```

図 3.2: クラスタリングを用いたノイズタグを削除するアルゴリズム

る。メタデータが欠損した画像に対してメタデータが推定された画像検索結果を P' とする。

図 3.2 に、ノイズとなるメタデータを削除するためのアルゴリズムを示す。ここでは、タグの削除を例にして説明する。アルゴリズムでは、タグが推定された画像検索結果 P' に含まれるそれぞれのタグを $T = \{w_0, w_1, \dots, w_l\}$ とする。それぞれのタグの使われ方を考慮するため、画像検索結果 P' において、タグ w_a を含む画像に対してクラスタリング手法を適用し、そのクラスタリング結果 $\{C_0, C_1, \dots, C_{|T|}\}$ を作成する。ここで、 $|T|$ を P' に含まれるタグの種類数とする。タグ w_a のクラスタリング結果 C_y のクラスタ c_i に含まれる画像数 $|c_i|$ が α より小さい場合、そのクラスタに含まれている全ての画像からタグ w_k を削除する。これにより、 k 近傍法で余分に付与されてしまったタグや、画像にあまり関係のないタグを削除することができる。

今回は、クラスタリングには画像検索結果をクラスタリングする手法である Maxmin [24] を用いる。Maxmin は、クラスタ数などのパラメータを指定が必要ない画像検索結果をクラスタリング手法である。

同様に、Exif の修正について、クラスタリング結果 C_y のクラスタ c_i に含まれる画像数 $|c_i|$ が β より小さい場合、クラスタ c_i に含まれる Exif を推定された画像の Exif の値をそのクラスタ c_i の平均値とする。画像が特徴的かつ視覚的に類似している画像を考慮した Exif を付与することができると思われる。本論文では、

$\alpha = 0.25, \beta = 0.5$ としている.

3.3 画像の特微量の類似度

2つの画像間の視覚的な類似度を求めるために、画像の低次特微量に基づいた類似度を算出する. 類似度の算出には、[24]の論文で用いられている算出方法を用いる. [24]では、6つの低次特微量を用いる (Color histogram, Color layout [91], Scalable color [91], CEDD [92], Edge histogram [91], Tamura [93]).

Color histogram は、RGB色空間の各チャンネル毎の色の分布を表し、それぞれの色が画像中に現れる頻度を表す. 2つの画像の色ヒストグラムの類似度の算出には、Bhattacharyya Distance [94]を用いる. 2つの画像 a, b の Bhattacharyya Distance $d_0(a, b)$ は、以下の式で計算される.

$$d_0(a, b) = -\log \sum_{i=1}^n \sqrt{ch_a(i) * ch_b(i)} \quad (3.1)$$

ここで、 ch_a, ch_b は、画像 a, b の Color histogram である.

Color layout は、画像の代表的な色の空間的な分布を表す特微量である. 画像を64のブロックに分割し、それぞれのブロックの平均の色を取得し、離散コサイン変換を適用した結果にジグザグスキャンを適用して取得する.

Scalable color は、HSV色空間のヒストグラムに Haar 変換を適用したものである. はじめに、画像を HSV 色空間に変換し、Hを16レベル、SVを4レベルにそれぞれ変換することによって、256ビンのHSVヒストグラムを作成する. そして、Haar変換を適用して、128, 64, 32, 16のヒストグラムを求める. 2つの画像の Scalable Color の類似度の算出には、 $L_1 - norm$ を用いる.

CEDD (color and edge directivity descriptor) は、色とテクスチャの両方を考慮したヒストグラムを作成する特微量である. はじめに、HSV色空間で色ヒストグラムを作成する. そして、幾つかのルールに基づいて24ビンのヒストグラムに分け、それに5つのエッジのフィルタを適用することで特微量を得る. 用いるエッジのフィルタは、水平、垂直、45度、135度、向き無しの5つである. 2つの画像の CEDD の類似度は、Tanimoto 係数で算出する.

Edge histogram は、画像の局所的なエッジの分布のヒストグラムである. 画像を 4×4 のグリッドに分割して、5つのフィルタ (水平、垂直、45度、135度、向き無し) を適用する. 結果として、80ビンのヒストグラムが得られる. 2つの画像のエッジヒストグラムの類似度の算出には、 $L_1 - norm$ を用いる.

Tamura は、複数のテクスチャの特徴量によるヒストグラムである。本論文で用いる特徴量は、coarseness, contrast, および directionality の3つである。2つの画像の Tamura の類似度の算出には、標準 $L_2 - norm$ を用いる。

これらの6つの低次特徴量に基づいて2つの画像 a, b の類似度 $I(a, b)$ を以下の式で算出する。

$$I(a, b) = \frac{1}{f} \sum_{i=0}^f \frac{1}{\sigma_i^2} d_i(a, b) \quad (3.2)$$

ここで、 f は用いられる低次特徴量の種類数、 $d_i(a, b)$ は2つの画像 a, b 間の i 番目の低次特徴量の値、 σ_i^2 は i 番目の低次特徴量の分散である。

3.4 タグの類似度

それぞれの画像に付与されたタグに基づいて2つの画像間の類似度を求める。そのため、本論文では、あるタグがどのくらいの写真に付与されているかを評価する IDF を利用する。すべての画像に付与されているタグを $tag = \{tag_1, tag_2, \dots\}$ と表した場合の w 番目のタグの IDF idf_w を以下の式で算出する。

$$idf_w = \log \frac{N}{n_w} \quad (3.3)$$

ここで、 N は全画像数、 n_w は全画像のタグの中で w 番目のタグが付けられている画像数である。

すべてのタグの IDF を算出した後に、それぞれの画像のタグベクトルを作成する。画像 a のタグベクトルを $\vec{a} = \{idf_{1a}, idf_{2a}, \dots\}$ を以下のように定義する。

$$idf_{ia} = \begin{cases} idf_i & tag_i \in tag_a \\ 0 & otherwise \end{cases} \quad (3.4)$$

ここで、 tag_a は、画像 a に付与されているタグの集合である。

2つの画像 a, b のタグベクトルを用いてコサイン類似度を算出し、2つの画像 a, b 間の画像類似度とする。

$$T(a, b) = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|} \quad (3.5)$$

3.4.1 写真の撮影状況の類似度

撮影状況の特徴量とするため、考慮するべき要素として、写真を撮影する際の周辺の明るさ、撮影者の位置や、被写体と撮影者の距離などがあげられる。例え

表 3.1: 評価実験に用いたクエリ一覧

apple	arashi	bicycle	building	cardgame	christmas
earthquake	eel	final fantasy	gundam	hamamatsu	hamster
kyoto	landscape	mercedes	mount fuji	mozart	muffler
computer	disaster	jellyfish	hydrangea	sagrada familia	nuclear plant
sandwich	stationery	tea	tiger	ultraman	windows

ば, 写真に付与されている Exif に含まれる Subject Distance を用いることで, カメラの位置と被写体との距離がわかる. しかしながら, Exif が付与されている写真のなかでも, Focal Length や, Exposure Time などは, 多くの写真に付与されているが, Subject Distance や, 位置情報を表す値は, 多くの場合に付与されていないため, 現段階では, それらを特徴量として用いるのは困難である. 今後, それらのメタデータが付与された写真が増加した場合は, それらを考慮した撮影状況類似度の算出を行うべきである. そのため, 提案手法では, 写真の撮影状況として, 写真の明るさのみを考慮する.

2つの画像間の撮影状況の類似度を求めるために, Exif に基づいた類似度を算出する. そのために, 写真の明るさに関する Exif の値を利用した *LogLightMetric* に基づいて画像の撮影状況特徴量を算出する [74]. ISO Speed は画像素子の感度を示す値である. この値が大きいほど, 画像素子の感度が増す. Exposure Time はカメラのシャッターが開いている時間であり, 値が大きいほどシャッタースピードが遅い. Aperture はレンズを通過する光量を示す値である. この値が大きいほどレンズを通過する光量は少ない. Focal Length は焦点距離を示す値である. この値が大きいほど焦点が遠い. 画像の撮影状況 *LogLightMetric* を以下の式に基づいて求める.

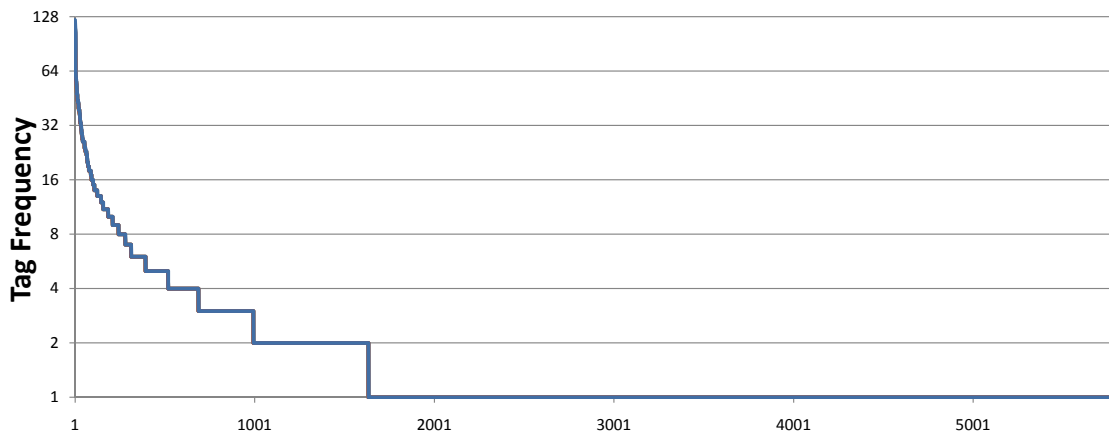
$$\text{LogLightMetric} = \lg(K * ET * A * ISO / FL^2) \quad (3.6)$$

ここで, K は比例定数, ET は Exposure Time, A は Aperture, ISO は ISO Speed, および, FL は Focal Length の値である. LogLight Metric は, 値が小さいほど写真を撮影した際の環境が明るいことを表している.

2つの画像 a, b 間の類似度を以下の式に基づいて $L_1 - norm$ で求める.

$$\text{exif}(a, b) = |LLM_a - LLM_b| \quad (3.7)$$

ただし, 画像 a, b の LogLightMetric をそれぞれ LLM_a, LLM_b とする.



の多くは、カメラメーカーや、英語以外で書かれたタグや、時刻などである。そのため、これらのタグを推定出来たととしてもクラスタリング精度の向上に繋がるとは限らない。

図 3.4 に、評価実験の流れを示す。本実験において、提案手法で用いる k 近傍法のパラメータは、 $k = 5$ とする。欠損したメタデータへの頑健性を評価するために、画像検索結果に含まれる r 枚の画像に付与されているメタデータを全て削除し、削除されたメタデータを提案手法によって推定する。また、メタデータを削除する画像数を $r = 0, 1, 2, \dots, 50$ とする。 $r = 0$ の場合は、メタデータの推定は行わずに、メタデータの削除のみを行なっている。本実験において、 k 近傍法を用いたメタデータの推定を行い、メタデータの削除を行わない手法を比較手法とする。

また、正解データのクラスタリング結果の実験協力者による差を考慮するために、提案手法によるクラスタリング結果と比較手法を 2 人の正解データとそれぞれ比較して、3 つの場合に分けて評価結果の比較を行う。1 つ目は、2 人の正解データに対する評価結果の平均である。2 つ目は、2 人の正解データに対する評価結果に対して、評価の低い方を選出した場合である。最後に、2 人の正解データに対する評価結果に対して、評価の高い方を選出した場合である。

3.5.1 制約付き凝集法

ここで、本実験で用いる制約付き凝集法 [95] について述べる。制約付き凝集法は、制約の作成に用いる類似度を選択し、選択された類似度に基づいて画像検索結果に対して制約を付与し、他の類似度に基づいて制約の付与された画像検索結果に対してクラスタリングを行う。

本実験では、制約として低次特徴量、撮影状況メタデータ、および画像タグに基づいた must-link 制約を用いる [96]。must-link 制約の付与は、撮影状況類似度、画像類似度、およびタグ類似度の中から 2 種類を用いる。例えば、タグと低次特徴量を must-link 制約として用いる場合には、以下の式を用いる。

$$M(a, b) = T(a, b) + I(a, b) \quad (3.8)$$

これは、2 つの画像 a, b 間のタグ類似度 $T(a, b)$ と画像類似度 $I(a, b)$ の和である。この値が小さい順に 2 つの画像 a, b 間の must-link 制約 M_1, M_2, \dots, M_r とする。これらの制約を用いて制約付き凝集法を行う。制約付き凝集法において must-link 制約を用いることによって生じる凝集法との違いは、must-link 制約が結ばれた 2 つ

の画像 a, b が初期状態から同一のクラスタに属していることである。本手法では、制約付き凝集法においてそれぞれのクラスタ間の類似度は、この場合、撮影状況類似度に基づいた群平均法を用いて算出する。

クラスタリングの終了条件は2つある。1つ目は、群平均法によって算出された最も類似したクラスタ間の類似度が閾値を下回った時である。ここでの閾値は、凝集法に用いた類似度の平均とした。2つ目は、制約付き凝集法を用いたクラスタリングの過程においてクラスタ数が閾値以下になったときである。これは、ユーザにクラスタリング結果を提示するのに際して、クラスタ数が少なくなりすぎることを防ぐためである。制約付き凝集法に用いるパラメータは、[96]で最も評価の高い値を用いた。

3.5.2 データセット

表3.1に、本実験で用いた30の検索クエリを示す。これらの検索クエリによって、Flickrより得られた画像検索結果を実験に用いる。用いる画像検索結果は、Flickrに対して指定した検索クエリにより検索結果として返されたリスト型の画像検索結果の中から、ExifのメタデータとしてISO Speed, Exposure Time, Aperture, および Focal Length の4つ全てが付与されており、1つ以上のタグが付与されている画像のランキングの上位50件分を用いる。そして、30人の実験協力者がクラスタリング結果を作成し、それを本実験の正解データとした。ただし、それぞれのクエリについて、2人の実験協力者が正解データを作成したため、得られた正解データは60個である。

3.5.3 評価指標

本実験では、2つの評価指標 (Fowlkes-Mallows index [97] と Variation of information [98]) を用いる。

Fowlkes-Mallows index は再現率, 適合率を用いた評価を行う。画像検索結果 I の2つのクラスタリング結果を C, C' とし, I のすべての画像のペアを表3.2のクラスに分ける。以下のように再現率, 適合率を求める。

$$W_I(C, C') = \frac{N_{11}}{N_{11} + N_{01}} \quad (3.9)$$

$$W_{II}(C, C') = \frac{N_{11}}{N_{11} + N_{10}} \quad (3.10)$$

表 3.2: Fowlkes-Mallows index における画像のペアのクラス

N_{11}	画像のペアが C , C' の両方で同じクラスタに属している
N_{10}	画像のペアが C では同じ, C' は異なるクラスタに属している
N_{01}	画像のペアが C では異なる, C' は同じクラスタに属している
N_{00}	画像のペアが C , C' の両方で異なるクラスタに属している

これらの式を用いて Fowlkes-Mallows index の値を求める. この値が大きいほど 2 つのクラスタリング結果は類似している.

$$FM(C, C') = \sqrt{W_I(C, C')W_{II}(C, C')} \quad (3.11)$$

variation of information は相互情報量とエントロピーを用いた評価を行う. エントロピーを求めるためにランダムに選択したある画像がクラスタ k に属する確率を求める.

$$P(k) = \frac{n_k}{n} \quad (3.12)$$

ここで n_k はクラスタリング結果 C のクラスタ k に含まれる画像数, n は全画像数である. 次に, 以下の式でクラスタリング結果 C に関するエントロピーを求める.

$$H(C) = - \sum_{k=1}^K P(k) \log P(k) \quad (3.13)$$

次に, 2 つのクラスタリング結果 C, C' 間の相互情報量を求める. そのために, ランダムに選択したある画像がクラスタリング結果 C において k に属し, クラスタリング結果 C' において k' に属している確率を求める.

$$P(k, k') = \frac{|C_k \cap C'_k|}{n} \quad (3.14)$$

クラスタリング結果 C, C' 間の相互情報量を $I(C, C')$ とし, 以下の式で表す.

$$I(C, C') = \sum_{k=1}^K \sum_{k'=1}^{K'} P(k, k') \log \frac{P(k, k')}{P(k)P'(k')} \quad (3.15)$$

これらを用いて variation of information を $VI(C, C')$ を以下の式で求める.

$$VI(C, C') = [H(C) - I(C, C')] + [H(C') - I(C, C')] \quad (3.16)$$

Variation of information はある画像のそのクラスタ間の関連に着目し, 2 つのクラスタリング結果のその関連性の違いを評価している. そのため, variation of information は値が小さいほど 2 つのクラスタリング結果は類似している.

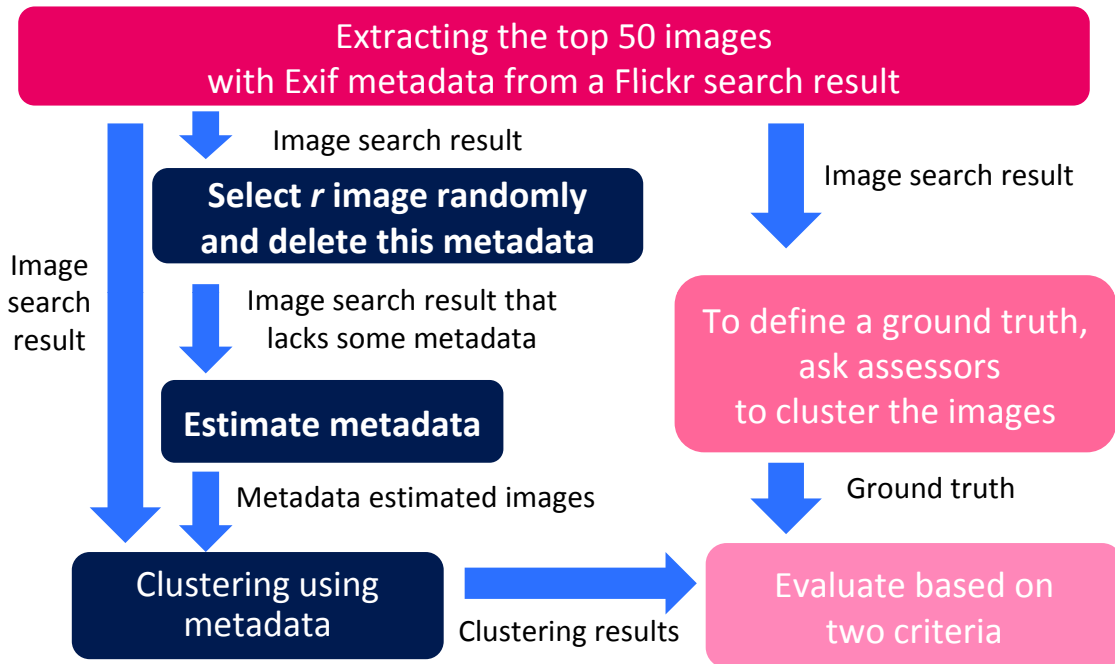


図 3.4: 評価実験の流れ

3.5.4 提案手法によるタグの推定の評価

図 3.5 に、Fowlkes-Mallows index と Variation of information の 2 つの評価指標を用いた、タグの推定によるクラスタリング結果の変化を示す。ここで、Fowlkes-Mallows index は、大きい値の方が評価結果が良く、Variation of information は、逆である。Average(FM), Average(VI) は、Fowlkes-Mallows index と Variation of information それぞれの評価指標において、2 つの正解データに対する提案手法の評価結果の平均値を 30 の実験データに対して算出し、その平均値である。Worst case average(FM), Worst case average(VI) は、2 つの評価結果に対して評価の低い方を選出した場合である。Best case average(FM), Best case average(VI) は、2 つの評価結果に対して評価の高い方を選出した場合である。 r は、メタデータを削除した画像数である。 $r = 0$ の場合は、画像検索結果に付与されているタグを提案手法によって修正したタグを用いたクラスタリング結果の評価結果である。図 3.5 では、5 つの手法を比較している。1 つは、撮影状況類似度に基づいた k 近傍法である。2 つ目は、画像類似度に基づいた k 近傍法である。3 つ目は、画像類似度に基づいた提案手法 ($\alpha = 1.0$) である。これは、 $\alpha = 0, 0.1, \dots, 1.0$ と調整した時に、Fowlkes-Mallows index と Variation of information による評価が最も高かった α である。4 つ目は、画像類似度に基づいた提案手法 ($\alpha = 0.25$) である。最後に、他の画像からランダムに選択して、欠損した画像に付与したものである。

図 3.5 において、多くの場合において、画像類似度を用いた提案手法が最もよい評価結果となっている。しかしながら、幾つかの場合において、提案手法が有効に機能していない。これは、提案手法におけるノイズタグの削除の工程で、ノイズタグだけではなくクラスタリングに有効なタグについても削除しているためだと考えられる。これらのことより、多くの場合において、提案手法によって、適切なタグが推定されており、メタデータが欠損している場合に提案手法を適用することによって、そのような画像検索結果においてもクラスタリングの性能を保つことができていることがわかる。

ここで、提案手法によって推定されたタグが画像の意味内容を表しているか評価するために、2つの評価実験を行った。1つは、メタデータを削除した画像に対して、提案手法によって本来付与されていたタグが復元されたかを評価する。以下の式に基づいて、この評価を行なった。

$$correct = \frac{|estimatedtags \cap originaltags|}{originaltags} \quad (3.17)$$

もう1つは、提案手法によって推定されたタグが、少なくとも間違っていないかを評価する。以下の式に基づいて、この評価を行なった。

$$wrong = \frac{|estimatedtags \cap wrongtags|}{wrongtags} \quad (3.18)$$

これらの式において、*estimatedtags* は提案手法によって推定されたタグ、*originaltags* は本来の画像に付与されていたタグ、*wrongtags* は画像に相応しくないタグである。

図 3.6 に、提案手法 ($\alpha = 1.0$) でタグを推定した結果についての評価を示す。図 3.6a において、 $r = 0$ の場合は、画像に付与されている約 60%のタグが画像の意味内容と一致していることがわかる。そして、 r の増加に伴って、提案手法による復元された割合が徐々に低下していることがわかる。また、図 3.6b において、提案手法によって推定削除が行われた結果の画像に付与されているタグが誤っているのは、約 2%以下であることがわかる。これらのことより、提案手法によって、画像の意味内容を表すタグが十分に推定されているとは言えないが、少なくとも誤っているタグが削除されていることがわかる。

3.5.5 提案手法による Exif の推定の評価

図 3.7 に、Fowlkes-Mallows index と Variation of information の2つの評価指標を用いた、Exif の推定によるクラスタリング結果の変化を示す。 $r = 0$ の場合は、画像検索結果の画像に付与されている Exif の値を提案手法によって修正して

表 3.3: Exif を推定するためのそれぞれの手法の平均順位

(a) Fowlkes-Mallows index		(b) Variation of information	
Approach	Average	Approach	Average
OurApproach: $\beta = 0.4$	1.922	OurApproach: $\beta = 0.4$	1.961
OurApproach: $\beta = 0.5$	2.039	OurApproach: $\beta = 0.5$	2.176
k-NN:TagSimilarity	4.863	k-NN:TagSimilarity	4.922
k-NN:ImageSimilarity	2.902	k-NN:ImageSimilarity	2.490
Random	2.804	Random	2.980

クラスタリングを行なった結果である。図 3.7 では、5つの手法を比較している。1つ目は、タグ類似度に基づいた k 近傍法である。2つ目は、画像類似度に基づいた k 近傍法である。3つ目は、画像類似度に基づいた提案手法 ($\beta = 0.4$) である。これは、 $\beta = 0, 0.1, \dots, 1.0$ と調整した時に、Fowlkes-Mallows index と Variation of information による評価が最も高かった β である。4つ目は、画像類似度に基づいた提案手法 ($\beta = 0.5$) である。最後に、他の画像からランダムに選択して、欠損した画像に付与したものである。

図 3.7 において、他の手法と提案手法の差が小さい。“Random” と提案手法の評価結果にあまり差がない原因は、他の写真の Exif の値をそのまま付与したことが原因である。例えば、Focal Length は連続値であるが ISO Speed は離散値であるが値の種類が固定されている。そのため、他の写真の Exif の値をランダムに取得しても、タグをランダムで付与した場合と比較すると、極端に誤った値が付与されることが少なかったことがあげられる。結果として、誤りの少ない Exif を用いてクラスタリングが行われたため、提案手法によるクラスタリング結果の評価結果との差があまり発生しなかったと考えられる。

ここで、表 3.3a, および表 3.3b に、Fowlkes-Mallows index と Variation of information による評価結果の順位の平均を示す。順位は、 r を $0 \dots 49$ と変化させて、それぞれの r の際に手法が何番目に評価が高かったのかで評価する。例えば、最も評価の高い手法は、1位で、最も悪い手法は5位である。表 3.3a, および表 3.3b において、提案手法が最も高い評価となっていることがわかる。そのため、平均すると、他の手法と比較して、提案手法を適用することでクラスタリング精度が改善される。

3.5.6 推定されたメタデータを用いたクラスタリング手法による影響の評価

本実験では、提案手法によるメタデータの修正に対して、3.5.1節で述べた制約付き凝集法を適用している。しかしながら、メタデータの推定後に適用するクラスタリング手法による性能の差については、評価していない。言い換えると、提案手法によって推定されたメタデータを、制約付き凝集法以外のクラスタリング手法で用いた場合にも有効であるかを評価する必要がある。そのため、本節では、提案手法によるメタデータ推定の性能が、クラスタリングアルゴリズムに対してどの程度影響を与えるかを評価する。

図 3.8, および図 3.9において、Fowlkes-Mallows index と Variation of information の2つの評価指標を用いた、評価結果を示す。ここで、それぞれの結果はメタデータの推定を行い、その結果を用いてクラスタリングを適用した結果である。これらの図において、“Random”は、メタデータが欠損している画像に対して、提案手法によるメタデータの推定を行わずに、ランダムに選択した他の画像のメタデータを付与した結果である。また、“Our Approach”は、メタデータが欠損している画像に対して、提案手法によるメタデータの推定を行った結果である。この時、 $\alpha = 1.0$, $\beta = 0.4$ である。クラスタリング手法については、“AGG”は、制約を用いない一般的な凝集法を表している。また、“CAGG”は、制約付き凝集法を表している。

図 3.8に、提案手法によるタグの推定を行った場合は、凝集法でも制約付き凝集法でも評価が高くなっている。そのため、提案手法によるメタデータ推定は、クラスタリングアルゴリズムに有効であることが言える。また、図 3.9に、提案手法による Exif の推定を行った場合は、提案手法によるメタデータ推定かランダムであるかによらずに、制約付き凝集法を用いた場合の方がクラスタリング結果の評価が高くなっている。これは、前述したように、ランダムに Exif の値を推定した場合でも、ある程度の性能が保たれているためである。結果として、図 3.9は、クラスタリング結果の評価値の差が制約付き凝集法と、制約とを用いていない凝集法のクラスタリングの性能の差を表す結果となっている。これらのことより、提案手法によって、タグのメタデータを推定した場合には、クラスタリングの性能が改善することがわかった。

3.6 おわりに

本章では、ウェブ上に存在する画像検索結果を効果的にクラスタリングするための、画像に付与されるべきメタデータの欠損を推定により補う手法を提案した。本論文で推定の対象とするメタデータは、ソーシャルタギングにより付与されたタグと、撮影状況を表すメタデータである Exif の2つである。それぞれの画像の正しいメタデータを推定するため、画像検索結果として与えられたデータのみを用い、それ以外に別途教示用の学習データ等を必要としない推定手法を提案する。本手法をソーシャルメディアサイトから取得した画像検索結果に適用し、本手法がクラスタリングの性能向上に寄与するようなメタデータを効果的に推定していることを示すために、Fowlkes-Mallows index と Variation of information の2つの評価指標を用いた評価を行なった。結果として、タグの推定については、多くの場合は、提案手法による推定を適用するとクラスタリングの性能が改善することが評価実験より示された。一方、Exifの推定については、提案手法による推定を適用してもクラスタリングの性能は非常にわずかに改善したのみであった。そのため、Exifの推定精度を改善することは、今後の課題である。加えて、今回の実験では、提案手法を k 近傍法と比較をしているが、検索結果以外のデータを参照するような、その他の手法との比較についても今後の検討課題である。

また、提案手法では、画像の低次特徴量、撮影状況類似度と、タグ類似度を用いている。しかしながら、画像に付与されている他のメタデータとして、GPSによる緯度経度情報、撮影方向情報などがある。これらのメタデータは、写真がどの地点で、どの方向を撮影したのかを表しているため重要である。また、写真の緯度経度情報は、タグの地理的な曖昧性の解消に非常に有用である。例えば、写真に地名を表すタグが付与されていた場合に、その地名と実際の場所を関連付けることで区別することが可能になる。そのため、これらのメタデータを推定することは、クラスタリングの性能を改善するために有用であると考えられる。しかしながら、これらのメタデータが付与されている写真は、非常に少ない。また、撮影方向情報が付与されている写真は、Flickr から収集した約 190 万枚の写真の中で、約 10% の写真にしか付与されていない。撮影位置情報を推定するための手法としては、類似する画像を検索することで位置情報を推定する手法などが提案されている [99]。そのような手法を参考にして、本論文で提案している手法を、これらのメタデータも推定可能に拡張することも今後の課題の1つである。

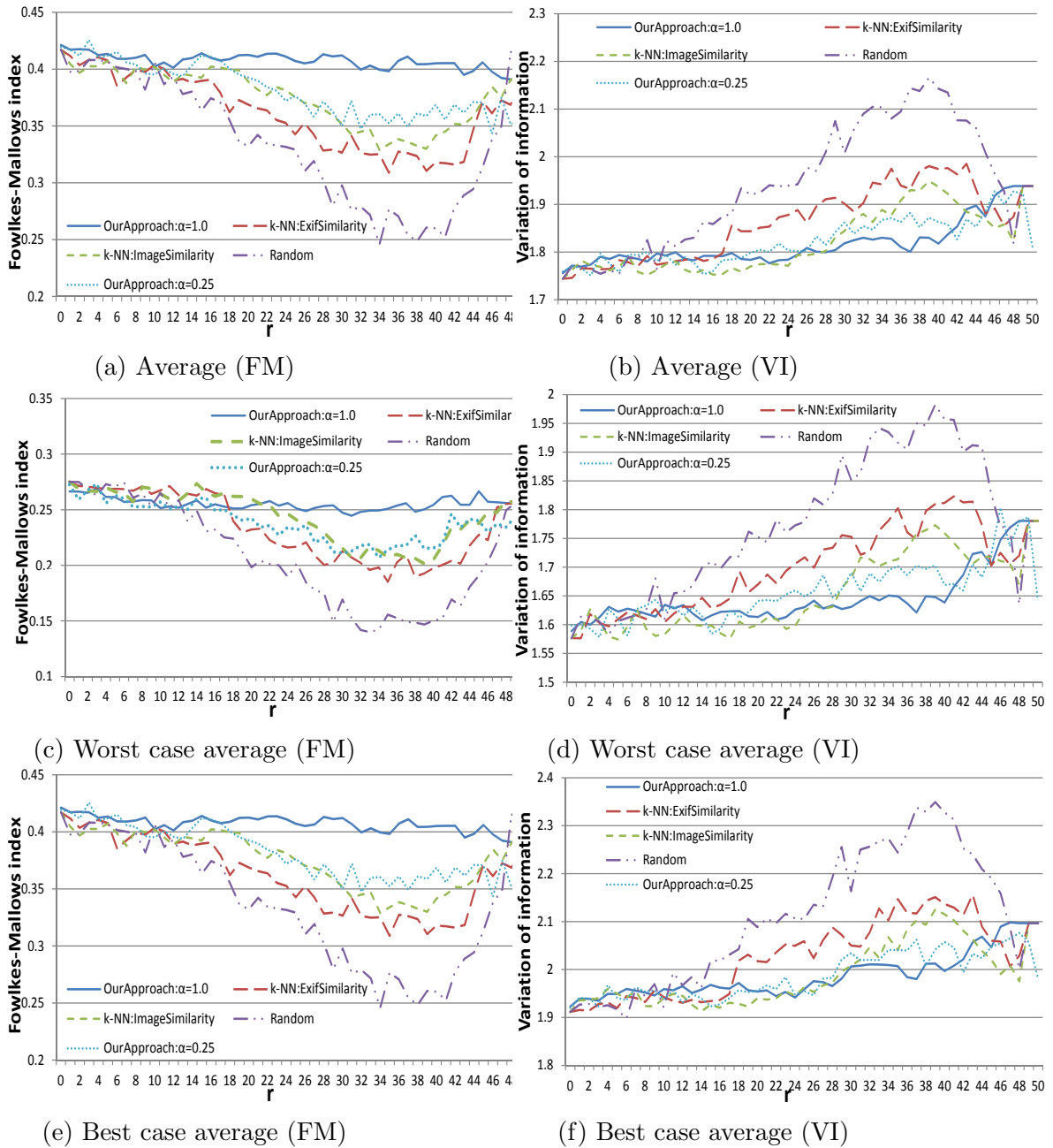
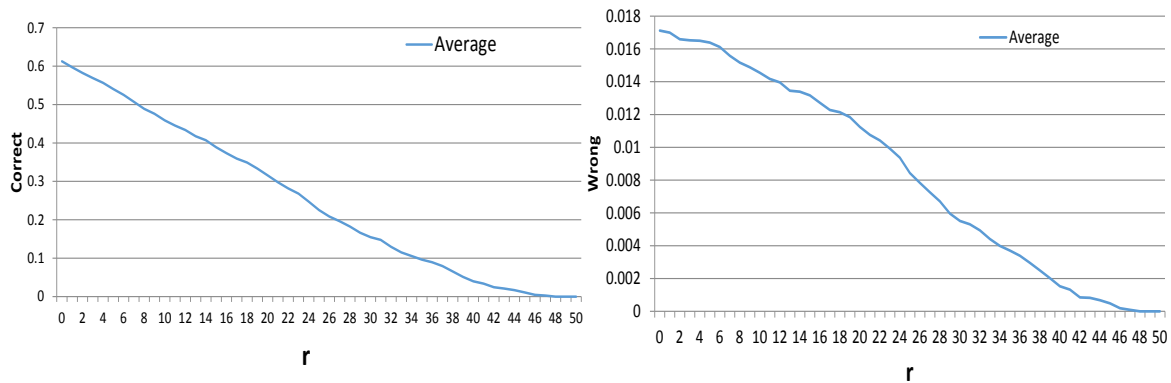


図 3.5: 提案手法の評価 (Tag)



(a) Correct

(b) Wrong

図 3.6: 提案手法によって推定されたタグの評価

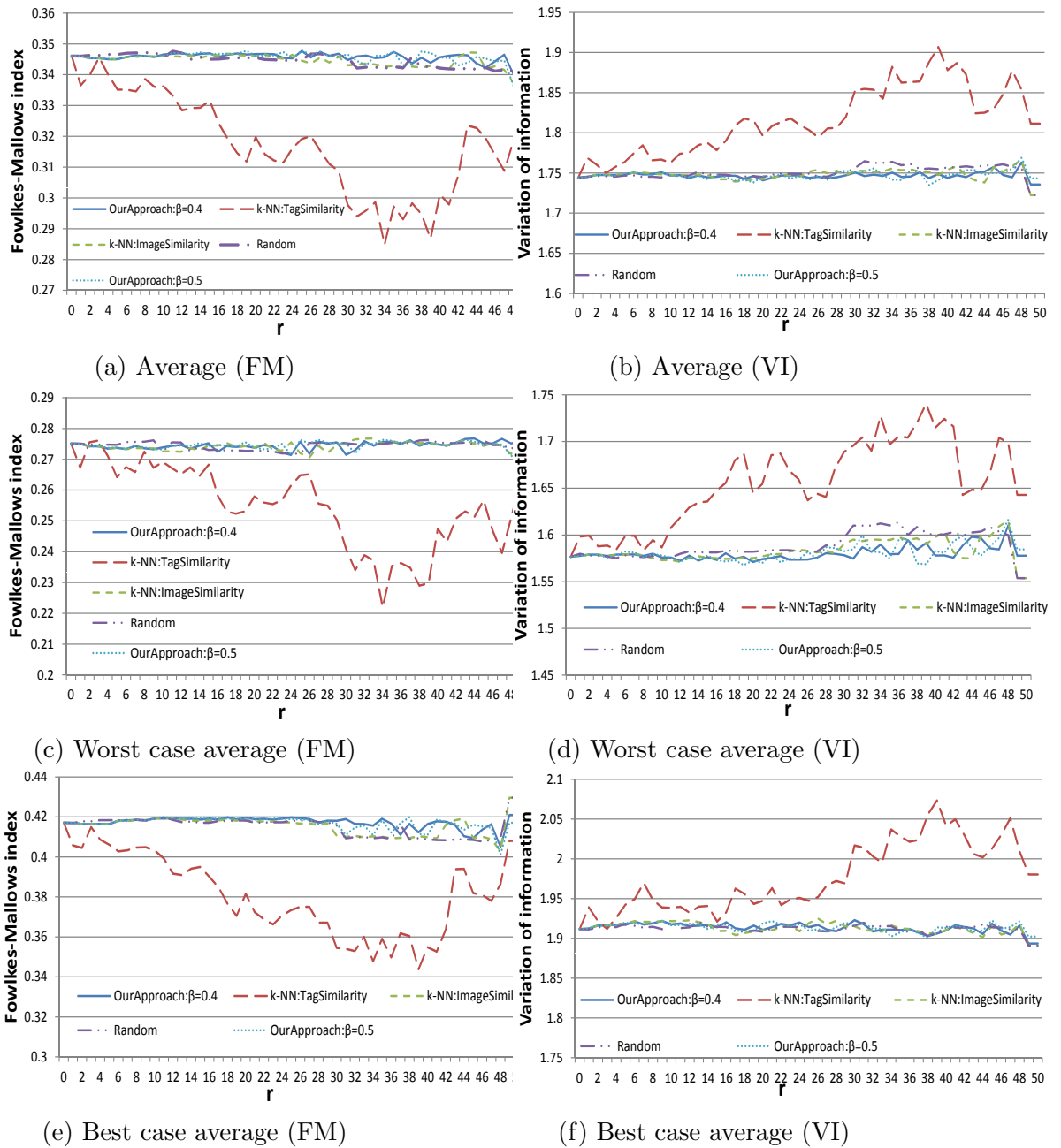


図 3.7: 提案手法の評価 (Exif)

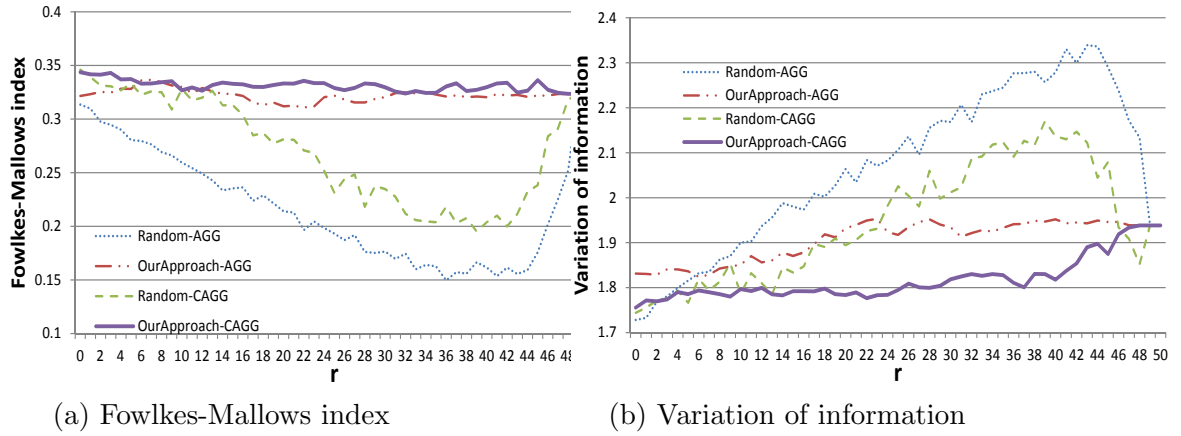


図 3.8: 提案手法の評価 (Tag)

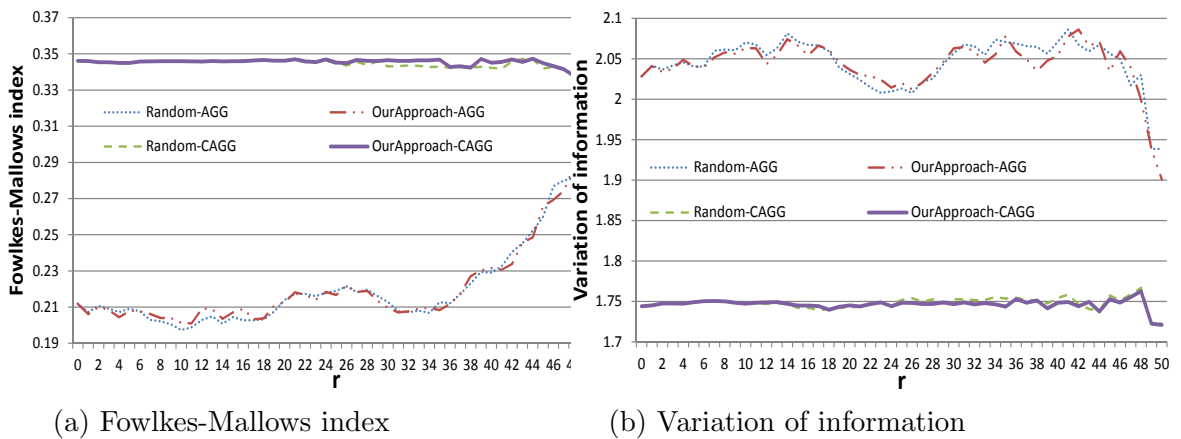


図 3.9: 提案手法の評価 (Exif)

第4章 画素情報を用いない 写真の質に基づいた画像検索 結果のリランキング

4.1 概要

本章では、画素情報を用いない写真の審美的評価に基づいた画像検索結果のリランキング手法について述べる。近年、デジタルカメラや、スマートフォンなどの普及により、それらの機器で撮影された写真がウェブ上に大量に存在する。それらの写真の多くはプロの写真家によらない写真であり、質が低い写真も含まれている。結果として、ソーシャルメディアサイトなどにおいて画像を検索する際などに、画像検索結果に質の低い写真が含まれる場合がある。そのような画像検索結果は、閲覧に対するユーザの満足度を低下させる。そのため、大量の写真の中から、質の高い写真を選択する技術の必要性が高まっている。しかしながら、写真の質を評価する際に、画像の画素情報を用いる場合にはデータ処理量が大きくなるという課題がある。

そこで、本論文では、これらの課題を解決するために、画像の画素情報を用いずに、写真に付与されている Exif や、カメラの情報を用いて写真の質を評価する手法を提案する。ここで、ほとんど同じ構図で撮影された2枚のひまわりを撮影した写真と撮影した際にそれらの画像に付与されたの Exif の値を図 4.1a、および図 4.1b に示す。これらの写真は、ほとんど同じ構図で撮影された写真であるが、図 4.1a の写真は、図 4.1b よりも質の高い写真であると考えられる。これは、図 4.1b の写真は、明らかに光量が不足しているためである。実際、図 4.1b の写真の絞り値は、図 4.1a よりも低い値となっている。これらのことより、Exif は、写真の質に影響する要因であると考えられる。本論文では、Exif に含まれるデータの中でもカメラの情報に着目して写真の質を評価し、画像検索結果に対するリランキングを行い、ユーザに質の高い写真を提示する事を目指す。提案手法の評価として、ソーシャルメディアサイトから取得した画像検索結果に対して、写真の質に基づいた効果的なリランキングが可能であることを示し、写真の質を評価するた



ISO Speed 100
 Exposure Time 0.001 sec
 Aperture f/8.0
 Focal Length 12.8 mm

(a) 質の高い写真



ISO Speed 80
 Exposure Time 0.0001 sec
 Aperture f/4.0
 Focal Length 6.1 mm

(b) 質の低い写真

図 4.1: 同じ構図で撮影された2枚の写真とその Exif の値

表 4.1: 高価な機材と安価な機材によって撮影された写真の質の評価の平均値

	Average
“Alexanderplatz” Low-price	1.350
“Alexanderplatz” High-price	1.980
“Brandenburg Gate” Low-price	1.410
“Brandenburg Gate” High-price	2.240

めに Exif が有効な手段の1つであることを示す。

4.2 類似画像を用いた写真の質の評価

写真の質を評価する際に、画像処理の技術を利用する手法がほとんどであるが、計算コストが非常に高いため、提案手法では、画素情報を用いずに、カメラの情報を用いて写真の質を評価する手法を提案する。提案手法では、画素情報を用いずに写真の評価するために、質の高い写真に Exif のパラメータが類似している写真もまた質の高い写真とする。そのために、提案手法では、最初に用いる質の高い写真を画像検索結果に含まれる写真の中で、高価な撮影機材で撮影されたものとする。これは、高価な撮影機材を用いるユーザは、写真の撮影が趣味で、写真の撮影技法について詳しい可能性が高いためである。

また、適用可能なドメインを限定しないために、トレーニングデータが事前に必要となる教師あり学習を用いた手法ではなく、画像検索結果に含まれる写真に含まれるデータを利用して写真の質の評価を行う。

ここで、高価な撮影機材を用いているユーザが撮影した写真と、そうではない

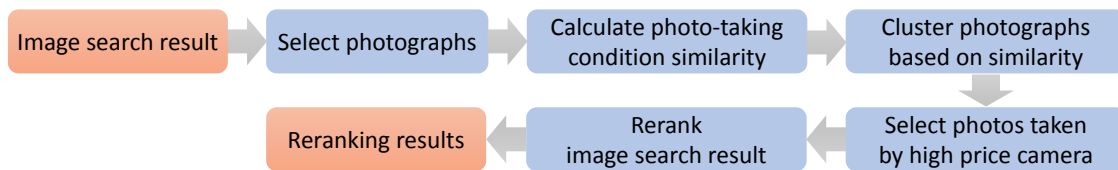


図 4.2: 提案手法の流れ

ユーザが撮影した写真の質を比較する実験を行う。はじめに、Flickr から、2つの検索キーワードを用いて画像を取得し、2つのデータセットを作成した。検索キーワードは、“Brandenburg Gate”と、“Alexanderplatz”である。そして、それらの画像検索結果に含まれる写真を300ドル以上の撮影機材によって撮影された写真と、100ドル以下の撮影機材で撮影された写真の2つに分けた。撮影機材の価格は、Exifに付与されている撮影機材のメーカー名とモデルに名したがって、Amazon [100] を利用して価格を取得した。それぞれ100枚ずつをランダムに選択し、これらの写真を実験参加者に写真の質を評価してもらった。評価は、1=“Bad”, 2=“Normal”, and 3=“Good”の3段階で行なった。高価な撮影機材によって撮影された100枚の写真の評価と、安価な撮影機材によって撮影された100枚の写真の評価の平均値を、表 4.1 に示す。Low-price は、100ドル以下の撮影機材で撮影された写真を表す。また、High-price は、300ドル以上の撮影機材で撮影された写真を表す。表 4.1 より、2つのデータセットともに、高価な機材で撮影された写真の方が質の評価が高いことがわかる。加えて、実際に、その差が有意であるかを検証するために、t-検定を適用した。“Brandenburg Gate”によるデータセットでは、0.1%で有意差が認められた。また、“Alexanderplatz”によるデータセットでも、0.1%で有意差が認められた。これらのことより、高価な機材で撮影された写真は、平均すると質の高い写真が含まれている可能性が高いと言える。

図 4.2 に、本論文で提案する手法の処理の手順を示す。はじめに、Flickr などから取得した画像検索結果から、特定のタグが付与されている写真を選出する。これは、前述したように、ユーザがキーワードをクエリとして検索した際に得られる画像検索結果には、複数の異なる意味内容の写真が含まれている場合があるためである。画像検索結果に含まれる多様性への対処については、3章で提案しているクラスタリングなどで行うとし、ここで提案する手法においては、対象としない。次に、選出した写真に付与されている Exif に基づいて撮影状況類似度を算出し、クラスタリングを適用する。写真の撮影された状況(時刻や、屋外屋内など)の多様さを含む画像検索結果に対して、それらを1つの評価基準で評価することの困難さに対処するためである [40]。最後に、それぞれのクラスタにおいて、高

表 4.2: Flickr に登録されている撮影機材の種類と数

Type	Point and shot	Camera phone	Comcorder	DSLR	Digital Back	Digital Rangefinder	Unknown	Film camera
Number	1545	509	232	136	29	4	4	1
Percentage	65.74 %	17.33%	9.83%	5.76%	1.23%	0.17%	0.17%	0.04%

価な撮影機材で撮影された写真を選出し、クラスタに含まれる画像とそれらを比較することでリランキングする。

4.2.1 撮影状況類似度を用いたクラスタリング

本節では、撮影状況類似度に基づいて画像検索結果をクラスタリングすることで、クラスタを得る。2つの画像間の撮影状況類似度の算出には、3.4.1章と同様である。写真の明るさを考慮するために、写真の LogLight Metric を特徴量として算出し、 L_1 -norm で類似度とする。また、クラスタリングには、パラメータの指定が必要ない画像検索結果をクラスタリング手法である Maxmin [24] を用いる。これにより、写真の明るさに基づいたクラスタが作成される。

4.2.2 リランキング

撮影状況類似度に基づいたそれぞれのクラスタにおいて、高価な撮影機材を用いて撮影された写真の Exif のパラメータに基づいて写真をリランキングする。はじめに、クラスタに含まれる画像の撮影に用いられた撮影機材が高価な順に、上位 $\alpha\%$ を取得する。本論文では、 $\alpha = 20$ である。取得した写真の Exif から ISO Speed, Exposure Time, Aperture, および Focal Length を抽出し、それぞれの平均値を求め、これをそのクラスタにおいて質の高い写真を撮影するために用いられた Exif のパラメータとする。それ以外の写真 p について、以下の式に基づいて、高価な機材を用いて撮影された写真の Exif パラメータの類似度 $rerank_p$ を算出する。

$$rerank_p = \sum_{j=1}^{|\mathcal{I}|} exif_{jp} \quad (4.1)$$

ここで、 $exif_{jp}$ は、高価な機材を用いて撮影された写真の4つの Exif のパラメータである。求められた類似度をリランキングのスコアとし、昇順でユーザに提示する。

pantheon	alexanderplatz	brandenburg gate	eiffel tower	sagrada familia
temple	tower	brandenburg	architecture	architecture
dome	fernsehturm	platz	night	temple
piazza della rotonda	allemagne	kreutzberg	seine	catedral
agrippa	schweizer	sky	torre	modernism
fountain	tv	pariser	monument	obra

図 4.3: 評価実験で用いる検索キーワードとタグ

表 4.3: 撮影機材の価格

	No price	$0 < price$	$100 < price$	$500 < price$	$1000 < price$	$5000 < price$
Number	1131	146	807	193	70	13
Percentage	47.92%	6.19%	34.19%	8.18%	2.97%	0.55%

4.3 実験

本節では、提案手法による画像検索結果のリランキング結果に、人手により評価された質の高い写真がどれぐらい含まれているかを評価する。また、ベースライン手法と提案手法を比較することで、Exifを用いることで写真の質を評価することが可能であることを示す。

4.3.1 データセット

ここでは、評価実験を行うために、撮影機材の価格のデータセットと、画像検索結果のデータセットの作成について述べる。

提案手法で用いる、写真の撮影機材の価格のデータセット作成する。はじめに、Flickrに登録されている2,360の撮影機材のモデル名を取得し、そのモデルのAmazonでの販売価格を取得した。撮影機材の種類と、価格について表4.2と表4.3に示す。表4.3において、No priceとは、Amazonから撮影機材の価格が取得できなかったものである。実験では、価格が取得できた撮影機材で撮影された写真のみを用いて行う。

図4.3に示す検索キーワードでFlickrを検索し、その中で特定のタグを持つ画像を選出した。実験では、この手順によって得た、25の画像検索結果をデータセットとして用いる。評価のために、4人の実験協力者に、リランキング結果を提示して、それぞれの画像に(1=“Good”, 0=“Bad”)を付与してもらった。

2つの方法による結果を提案手法との比較に用いる。1つは、Flickrより取得した画像検索結果にランキングをそのまま用いる。もう1つは、1つ目の結果に含まれる画像をランダムに並び替えたものである。

4.3.2 評価指標

評価指標として、Precision@ k 、NDCG@ k (normalized discounted cumulative gain measure) [101]、および Unattractive Rejection [63] を用いる。

Precision@ k は、検索結果の上位 k 件において、評価の高いものが含まれているほど、ランキングの評価は高くなる。本論文の場合は、上位 k 件に、質が高いと評価された写真 (1=“Good”) の数であり、以下の式で算出する。

$$Precision@k = \frac{1}{k} \sum_{i=1}^k rel_i \quad (4.2)$$

ここで、 rel_i は k 番目の順位の評価点数を表す。

NDCG@ k は、検索結果上位 k 件において、理想的なランキングへの近さを表す評価指標である。Precision@ k との違いとして、NDCG@ k は、ランキングの順位による重みが付与されている。ランキングの上位 k 件において、同じ件数だけ良い評価のものが存在した場合は、上位に出現するほどランキングは高く評価される。NDCGを計算するためのDCGは以下の式で算出される。

$$DCG@k = rel_1 + \sum_{i=2}^k \frac{rel_i}{\log_2 i} \quad (4.3)$$

DCG@ k は順位の下降とともに重みを付け加えている。NDCGは、以下の式で算出される。

$$NDCG@k = \frac{DCG@k}{IDCG@k} \quad (4.4)$$

IDCG@ k は、DCGが最も理想的だった場合の値を表す。本実験では、 $k = \{1, 2, \dots, 10\}$ とする。

また、Unattractive Rejection は、ランキング結果の中に最初に現れる不適切な画像 (本論文の場合は、0=“Bad”) の順位を評価することで、作成したランキングがどれだけ不適切な画像の出現をランキングの上位に出現させることを防いでいるかを評価している。Unattractive Rejection (UR) は、以下の式で算出する。

$$UR = \frac{1}{|Q|} \sum_{i=1}^Q \frac{1}{rank_1} \quad (4.5)$$

ただし、 k 件のランキングに、不適切な画像が含まれていなかった場合は、 $UR = 0$ とする。Unattractive Rejection は、値が低いほど、リランキングによって不適切な画像の出現を防ぐことができていると評価する。

4.3.3 提案手法によるリランキング結果の評価

図 4.4a, 図 4.4b, および表 4.4 に、それぞれ Precision@10, NDCG@10, および Unattractive Rejection の提案手法の評価結果を示す。“Reranking”は、提案手法によるリランキング結果を表す。“Default”は、Flickr より取得した画像検索結果にランキングをそのまま用いたものである。“Random”は、ランダムに画像検索結果を並び替えたものである。これらは、25 のデータセットに対する評価の平均値である。図 4.4a と、図 4.4b において、提案手法は、すべての順位において最も高い評価となった。また、表 4.4 においても、提案手法が最も高い評価となっている。結果として、提案手法は、質の高い写真を画像検索結果の上位に出現させることができた。

次に、提案手法と既存手法の審美的評価の性能と、実行速度について比較する。本実験では、Loらにより提案された手法 [102] を比較手法とする。この手法は、トレーニングデータとテストデータに含まれる写真から複数の画像特徴量を生成し、SVM によって審美的評価の高い写真か、低い写真であるかを分類している。本論文の提案手法では、画像検索結果に対して審美的評価に基づいたランキングを作成するため、SVM ではなく SVR によって学習した結果を用いる。また、この手法は、matlab を用いて実装されているが、octave を利用して実験を行った。さらに、この手法では、トレーニングデータが必要となるため、dpchallenge¹ から取得した写真とその評価値をトレーニングデータとした。実験条件として、3つのクエリに対して、前述した手順で評価を行なった。審美的評価の評価指標も前述した3つの評価指標を用いる。また、実行速度を比較するために、50, 100, 150, 200 枚の画像に対してそれぞれの手法を適用した実行時間を示す。

表 4.5 に、提案手法と既存手法を3つの評価指標に基づいて評価した結果を示す。Unattractive Rejection による評価は、提案手法が優れていることがわかる。しかしながら、他の2つの評価指標では、提案手法の方が優れているが、ほとんど2つの手法に差がない。そのため、提案手法は、既存手法と比較して、少なくとも同等程度の写真の審美的評価の能力があると考えられる。次に、表 4.6 に、提案手法と既存手法の実行時間を計測したものを示す。この表より提案手法の方が

¹<http://www.dpchallenge.com/>

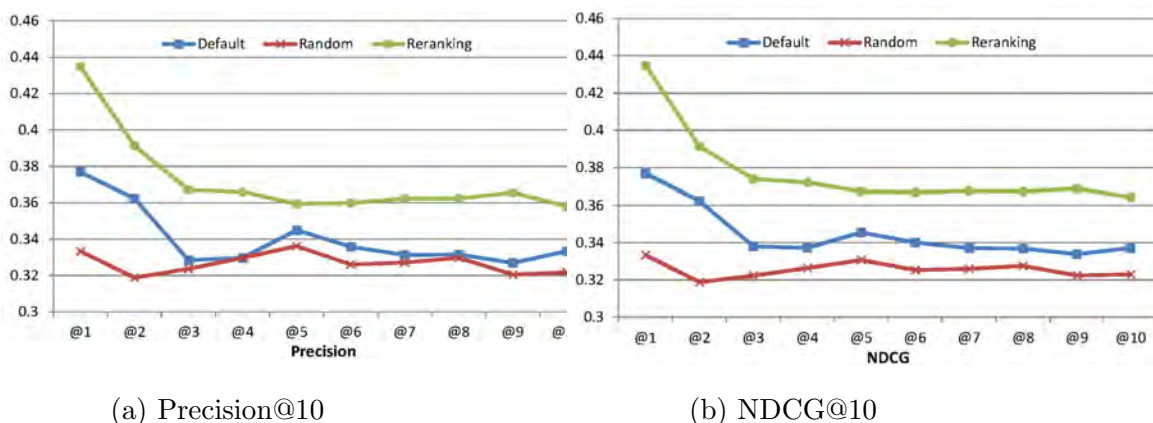


図 4.4: 提案手法とベースライン手法の比較

表 4.4: Unattractive Rejection による提案手法とベースライン手法の比較

	Default	Random	Reranking
UR	0.781	0.806	0.742

1, 000 倍以上高速であることがわかる。これは、提案手法が画像の画素情報、および画像処理技術を全く用いていないためであると考えられる。多くの手法では、高度な画像処理技術を用いるため、画像を複数回にわたって走査する場合や、畳込みなどを行うため、計算コストが非常に高い。一方、提案手法では、写真の撮影状況を表すメタデータである Exif のみを用いており、画像の走査などが必要ないため、計算コストが非常に低い。これらの2つの結果より、既存手法と比較して、提案手法は、写真の審美的評価を評価するための性能を大きく改善するには至っていないが、劣っておらず、実行時間は、非常に高速であるといえる。そのため、提案手法は、高速に写真を審美的評価に基づいてリランキングするために有効な手法であると考えられる。

4.4 おわりに

本章では、画素情報を用いない写真の質に基づいた画像検索結果のリランキング手法を提案した。写真の質を評価するために、画素情報を用いずに、写真のメタデータや、カメラの情報を用いた。提案手法の有効性を示すために、Flickr から取得した写真に対して提案手法を作成し、人手による評価を行った。結果として、提案手法によってリランキングを行うことは有効であることを示した。

表 4.5: 提案手法と既存手法の比較 (審美的評価)

	NDCG@10	Precision@10	Unattractive Rejection
比較手法	0.77	0.78	0.50
提案手法	0.80	0.78	0.32

表 4.6: 提案手法と既存手法の比較 (実行時間)

	50	100	150	200
比較手法	142.497	0289.21	439.26	578.37
提案手法	0.03	0.04	0.07	0.09

今後の課題として、提案手法では、写真の明るさを考慮しているが、それ以外のメタデータを用いた評価手法の提案があげられる。

第5章 位置情報付き写真を用いた ホットスポットの関連性の 抽出

5.1 概要

本章では、画像を閲覧するために、写真の地理的なクラスと、その関連性を抽出し、ユーザに提示するための手法について述べる。近年、GPS機能が付随した撮影機材の普及に伴い、緯度経度情報が付与された写真がウェブ上に存在する。ユーザが写真を撮影する際に、自身の興味のある被写体を撮影することがほとんどである。そのため、写真の撮影地点の集合を可視化することで、ユーザの興味を地理的な観点で可視化することが可能になると考えられる。言い換えると、ウェブ上の位置情報付き写真を利用することで、地理的な観点でユーザの興味を表すホットスポット(クラス)を抽出することが可能である。しかしながら、それらのホットスポットは、独立して抽出されるが、実際には、複数の観点での関連性が存在する。たとえば、特定の被写体を撮影するための撮影スポットを表すホットスポットと、その被写体を含むホットスポットという関連性があげられる。また、特定のイベントに関する写真が多く撮影されたホットスポット同士の関連性があげられる。しかしながら、単純にホットスポットを抽出すると、それらは独立して抽出されるため、ユーザがホットスポットを閲覧する際に、閲覧の負担がかかる。そのため、本論文では、ソーシャルメディアサイトから取得した位置情報付き画像からホットスポットを抽出し、2つの観点における関連性をそれぞれ抽出する手法を提案する。また、提案手法によって抽出したホットスポットの関連性を可視化することによって、提案手法によって抽出された関連性が適切であるかを検討する。

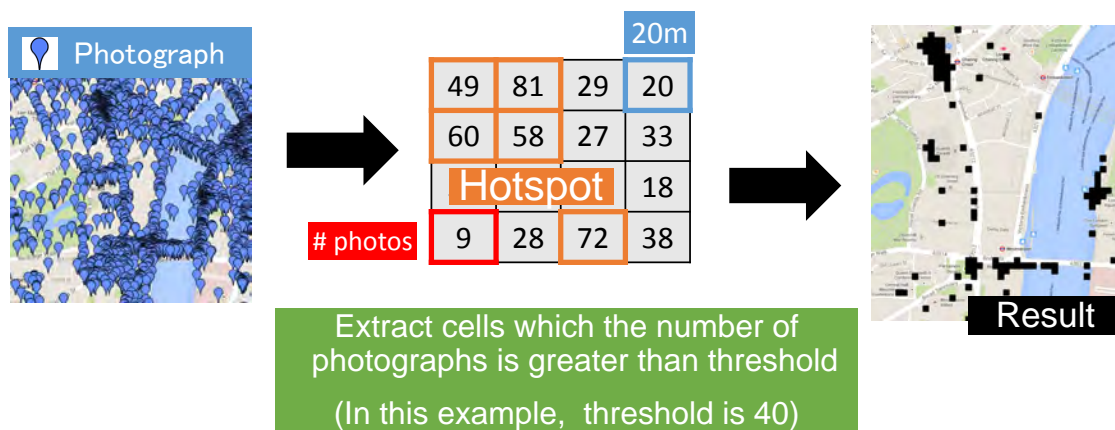


図 5.1: ホットスポットの抽出手順

5.2 ホットスポットの抽出と関連性の抽出

提案手法では、メタデータが付与された画像を利用して、ホットスポットの抽出とその関連性の抽出を行う。用いるメタデータは、緯度経度情報、撮影方向情報、焦点距離、およびソーシャルタギングによるタグである。本論文では、タグに基づいたホットスポットの関連性と、撮影スポットと被写体の関連性の2種類の関連性を抽出する。提案手法では、はじめに、ホットスポットを抽出したい領域で撮影された画像を取得し、それをグリッドに配置して、それぞれのセルに含まれる写真数に基づいて、ホットスポットを抽出する。次に、撮影方向情報と、緯度経度情報に基づいて、ホットスポットを分類する。本論文では、特定の被写体を撮影するための写真の撮影スポットと、ホットスポット内にランドマークのような多くのユーザの興味を集める関心領域と、それ以外の3つに分類する。次に、それぞれのホットスポットでイベントが発生していたかを分類する。最後に、タグに基づいたホットスポットの関連性と、撮影スポットと被写体の関連性の2種類の関連性を抽出する。その際に、タグに基づいたホットスポットの関連性は、ソーシャルタギングによるタグを利用する。また、撮影スポットと被写体の関連性の抽出には、撮影方向情報と、緯度経度情報を利用する。

5.2.1 ホットスポットの抽出

本論文では、写真に付与されている Exif に含まれる GPSLatitude(緯度) と GPSLongitude(経度) を写真の撮影位置情報とする。

取得した写真に付与された緯度経度情報に基づいて、 y 座標と x 座標で表され

るグリッドを作成する。以下の式を適用することで、写真に付与された緯度経度情報を2次元グリッドのセルの座標 (x, y) に変換する。

$$y = M_{height} - \frac{(Lat - Lat_{min}) * M_{height}}{Lat_{max} - Lat_{min}} \quad (5.1)$$

$$x = M_{width} - \frac{(Lng - Lng_{min}) * M_{width}}{Lng_{max} - Lng_{min}} \quad (5.2)$$

ここで、 Lat は、写真に付与された緯度 (Exif の GPSLatitude), Lng は、写真に付与された経度 (Exif の GPSLongitude の値) を表す。 Lat_{max} , Lat_{min} , Lng_{max} , および Lng_{min} は、 Lat と Lng の最小値と最大値を表す。 M_{height} , M_{width} は、生成するグリッドの高さと幅をそれぞれを表す。本論文では、 M_{height} , M_{width} は、生成するグリッドの1つのセルが約20mになるように設定した。結果として、生成されたグリッドの1つのセルには、約20mの範囲で撮影された写真が含まれている。

5.2.2 ホットスポットの分類

撮影方向情報と、撮影位置情報に基づいて、抽出されたホットスポットを関心領域、撮影スポット、その他のホットスポットの3つに分類する。ここでは、[72]において提案されている手法を利用する。

関心領域とは、ホットスポットの中心に関心のある領域 (ランドマークなど) が存在し、その領域の周辺からホットスポットの中心を撮影しているホットスポットである。そのため、写真の画角と、写真とホットスポットの中心の位置関係を利用することで、写真がホットスポットの中心を撮影したかを評価する。写真の画角は、以下の式で算出される。

$$\alpha = 2 \tan^{-1} \left(\frac{l}{2f} \right) \quad (5.3)$$

ここで、 $|x|$ は写真の横幅、 f は写真の焦点距離、 l は映像素子のサイズを表す。焦点距離は、写真に付与された Focal Length を用いる。次に、写真の撮影位置からホットスポットの中心への角度を算出する。

$$\theta = \tan^{-1} \frac{\cos y_2 * \sin(x_2 - x_1)}{\cos x_1 * \sin y_2 - \sin y_1 * \cos y_2 * \cos(x_2 - x_1)} \quad (5.4)$$

この θ は、写真の撮影位置 (x_1, y_1) から、北を0度にした場合のホットスポットの中心 (x_2, y_2) が存在する方向である。画角 α , 撮影方向 β から写真を内向き写真

と外向き写真に分類する.

$$\begin{cases} \text{内向き写真} & (\beta - \frac{\alpha}{2} < \theta < \beta + \frac{\alpha}{2}) \\ \text{外向き写真} & (\textit{otherwise}) \end{cases} \quad (5.5)$$

写真を分類した後に, ホットスポットの中心からある一定の範囲内に存在する内向き写真と外向き写真のそれぞれの枚数を数える. このとき, 内向き写真の枚数が外向き写真の枚数以上になる範囲が存在するときに, そのホットスポット内に関心領域が存在すると定義する.

次に, 撮影スポットとは, ホットスポットに含まれる写真の多くが同じ方向を撮影しているようなホットスポットである. これは, 近くに関心領域が存在する場合に, 多くのユーザがそれを特定の場所(ランドマークがよく見える場所)から撮影しているためである. そのため, ホットスポットに含まれる写真の撮影方向情報に基づいて撮影スポットかを判断する. ホットスポットに含まれる写真の撮影方向を10度ずつのヒストグラムにし, 1つのビンが全体の15%よりも高ければ, そのホットスポットは撮影スポットであるとする.

最後に, 関心領域でも, 撮影スポットでもないホットスポットをその他のホットスポットとする.

5.2.3 イベントの分類

イベントが発生した地点では, 多くのユーザがそのイベントに関する写真を撮影するため, その地点がホットスポットとして抽出される. そのため, 前述した手法で抽出したホットスポットにおいて, イベントが発生したかどうかをホットスポットに含まれる写真に付与されている撮影時間情報に基づいて分類する.

本論文において, ホットスポットにおけるイベントとは, イベントが発生した日以外の日のホットスポット内での写真の撮影枚数と比較して, 特定の日に写真の枚数が極端に多い場合に, そのホットスポットでイベントが発生したとする. そのため, ホットスポットのそれぞれの日付における写真の撮影枚数の増加を検出することで, イベントを検出する. はじめに, Simple Moving Average(SMA)をそれぞれの日付に対して求める. 特定の日付 M の n 日間の SMA は以下の式で算出される.

$$SMA_M = \frac{p_M + p_{M-1} + \dots + p_{M-n}}{n} \quad (5.6)$$

ここで, p_M は, ホットスポットに含まれる特定の日付 M 日に撮影された写真の枚数である. 本論文では, $n = 7$ とした. 最後に, 特定の日付 M における写真の

増加率 RI_M を，以下の式で算出する．

$$RI_M = \frac{SMA_M - SMA_{M-1}}{SMA_{M-1}} \quad (5.7)$$

$RI_M > 8$ の場合，そのホットスポットではイベントが発生したとする．

5.2.4 ホットスポット間の関連性の抽出

ここでは，これまでに抽出したホットスポットについて，ホットスポット間の関連性を抽出する．本論文では，撮影スポットと被写体の関連性と，タグに基づいた関連性の2つの種類のホットスポットの関連性を抽出する．

撮影スポットと被写体の関連性

撮影スポットと被写体を含むホットスポットの位置関係に基づいて，ホットスポット間の関連性を抽出する．撮影スポットに含まれる写真の中で，多くの写真がほぼ同じ方向を撮影しており，その方向にホットスポットが存在している場合，その撮影スポットは，そのホットスポットを撮影するための撮影スポットであると考えられる．本論文では，これをホットスポット間の関連性として抽出する．

撮影スポットと被写体の関連性を抽出するために，撮影スポットに含まれる写真に付与されている撮影方向情報の平均値と，擬似的な画角を用いる．この撮影スポットの擬似的な画角は，その撮影スポットの平均の撮影方向にホットスポットが存在するかの決定に用いる．撮影スポットの擬似的な画角 aov_p は以下の式で算出される．

$$aov_p = \frac{90}{\log_{10}(distance)} \quad (5.8)$$

ここで， $distance$ は，ホットスポットと撮影スポットのそれぞれの中心の Hubeny 距離である．Hubeny 距離は，以下の式で算出される．

$$distance = \sqrt{(M * dP)^2 + (N * \cos(P) * dR)^2} \quad (5.9)$$

P は2点間の平均緯度， dP は2点間の緯度差， dR は2点間の経度差， M は午線曲率半径， N は卯酉線曲率半径を表す．この擬似的な画角は，ホットスポットと撮影スポットの距離が近いほど，大きくなる．最後に，以下の式に基づいて，撮影スポットの平均撮影方向から擬似的な画角の角度以内にホットスポットが存在するかを判断する．

$$dir_s - \frac{aov_p}{2} < angle_{(s,h)} < dir_s + \frac{aov_p}{2} \quad (5.10)$$

ここで, $angle_{(s,h)}$ は, 撮影スポットとホットスポットの角度の差である. この式を満たす場合, 撮影スポットとホットスポットに関連性があると判断する.

タグに基づいた関連性

ホットスポットに含まれている写真に付与されているタグに基づいた関連性を求める. ホットスポットに含まれる写真に付与されているタグは, そのホットスポットに関連するタグが付与されていると考えられる. たとえば, 写真の被写体に関するタグや, イベントに関するタグがあげられる. そのため, 2つのホットスポットに含まれている写真に付与されているタグが類似していた場合, それらのホットスポットでは, 何らかの関連性があると考えられる. そのため, 2つのホットスポットの関連性を求めるために, ホットスポットに含まれるタグの類似度を算出する. それぞれのホットスポットの特徴量を TFIHF (Term Frequency Inverse Hotspot Frequency) を用いて算出する. TF_{ai} は, ホットスポット a に含まれる写真に付与されているタグ i の出現回数である. IHF_i は, タグ i が含まれているホットスポットの割合の逆数であり, 3.4 の IDF の算出と同じ手順で算出する. そして, 得られた TFIHF に基づいて, ホットスポット間のコサイン類似度を算出する. 一定の距離に以内に存在する2つのホットスポットのコサイン類似度が0.5よりも大きければ, それらのホットスポットはタグに基づいた関連性が存在するとする.

5.3 実験

本節では, 提案手法によって抽出されたホットスポットとその関連性を可視化した結果について述べる. 提案手法によって抽出された関連性は, タグに基づいた関連性と, 撮影スポットと被写体の関連性の2つである. 抽出したそれぞれの結果は Google Maps [103] に表示した.

5.3.1 データセット

ここでは, 本実験に用いる写真について述べる. 本実験では, Flickr より取得した写真 1,123,550 枚を用いる. この時, これらの写真の全ては, Exif に含まれる緯度 (GPSLatitude, GPSLatitudeRef), 経度 (GPSLongitude, GPSLongitudeRef), 撮影時間のメタデータが付与されている. これらの写真が撮影された期間は, 2010

表 5.1: 提案手法によって抽出された関連性の例

	Spot	# photos	Relation	# hotspots
R0	Soho Square	367	Tags	2
R1	Oxford Street and Regent Street	4070	Tags	5
R2	Leicester Square	1906	Tags	3
R3	Convent Garden	5735	Tags	3
R4	Tower Bridge and Tower of London	4527	Tags	7
R5	Olympic Park	4440	Tags	6
R6	London Eye	9819	Tags	2
R7	Big Ben	5510	SS	2

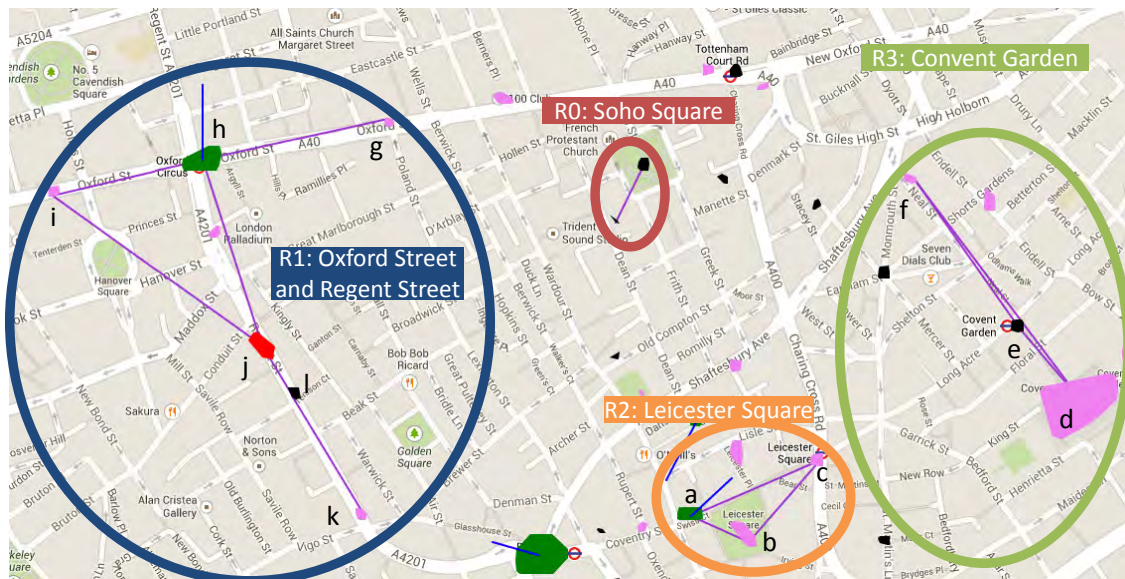


図 5.2: タグに基づいた関連性 (表 5.1 の R0 - R3).

年1月1日から, 2014年6月31日までである. また, 写真を取得した範囲は, ロンドン周辺 (緯度:-0.450439 0.148315, 経度:51.301643 51.669361) である.

5.3.2 考察

表 5.1 に提案手法によって抽出された関連性の例を示す. "Spot" は, 提案手法によってホットスポットから関連性が抽出された地域を表す. "# photos" は, そのホットスポットに含まれる写真数の合計を表す. "Relation" は, 提案手法によって抽出された関連性が2つの関連性のどちらであるかを表す. "Tags" は, タグに基づいた関連性を表す. "SS" は, 撮影スポットと被写体の関連性を表す. "# hotspots" は, その地域において関連性が抽出されたホットスポットの数を表す.

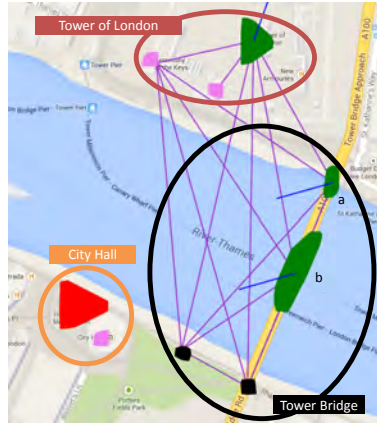


図 5.3: タグに基づいた関連性 (表 5.1 の R4).



図 5.4: 撮影スポットと被写体の関連性 (表 5.1 の R7).

図 5.2 と、5.3 に、提案手法によって抽出されたタグに関する関連性の例を、図 5.4 に提案手法によって抽出された撮影スポットと被写体の関連性の例を示す。これらの図において、緑色のポリゴンは、イベントの発生していない撮影スポット、黄色のポリゴンはイベントの発生した撮影スポット、青色の線は撮影スポットに含まれる写真の平均の撮影方向、赤色のポリゴンはイベントの発生していない関心領域、紫色のポリゴンはイベントの発生してたその他の領域、黒色のポリゴンはイベントの発生していなその他の領域を表す。また、紫色の線はタグに関する関連性を、黒色の線は撮影スポットと被写体の関連性を表す。

図 5.2 に Soho Square, Oxford Street, Covent Garden 周辺で抽出されたタグに基づいた関連性について示す (表 5.1 の R0 - R3)。提案手法を適用することで、この地域では、4つの地域に関するタグに基づいた関連性が抽出された。例えば、R2 (図 5.2 の橙色の丸) では、Leicester Square で3つのホットスポットが抽出され

た。R2の関連性において、最も左のホットスポット(図5.2a)は、Leicester Squareの近くにあるM&M's worldの店の周辺で抽出された撮影スポットである。この撮影スポットでは、多くのユーザが店の前から、店の入り口を撮影していたため、撮影スポットとして抽出された。中央のホットスポット(図5.2b)は、Leicester Squareの中で発生したイベントに関する写真が多く撮影されているホットスポットである。また、右のホットスポット(図5.2c)は、Leicester Squareの地下鉄の駅の周辺で撮影された写真を多く含むホットスポットである。これらのホットスポットは、単にホットスポットを抽出した場合には独立して抽出される。しかし、実際には、これらのホットスポットは1つの地域の中で抽出されたホットスポットであるため、地域に関する関連性があると考えられる。実際に、これらのホットスポットに含まれる写真の多くには、“Leicester”や、“Leicester Square”などの地域を表しているタグが付与されている。そのため、提案手法を適用することで、Leicesterの地域に関する関連性を抽出することができた。加えて、R3(図5.2の緑色の丸)では、3つのホットスポットについての関連性が抽出された。一番下のホットスポットは、Covent Garden Marketと、Big Egg Huntというチャリティイベントに関する写真を多く含むホットスポットである。中央のホットスポットは、Covent Gardenの地下鉄周辺と、その周辺のお店に関する写真を多く含むホットスポットである。一番上のホットスポットは、レストランで行われたパーティの写真と、そのレストランに関する写真を多く含むホットスポットである。提案手法により、これらのホットスポットについても地域に関連するホットスポットを抽出することが出来た。これらのことより、提案手法を適用することで、ホットスポットに含まれる写真に付与されているタグに基づいてホットスポット間の地理的な関連性を抽出できていることがわかる。

次に、R3(図5.2の青色の丸)では、Oxford Streetと、Regent Street周辺で抽出されたホットスポットから抽出された関連性について述べる。図5.2h-kのホットスポットについては、2010年に行われたデモに関するイベントの写真を多く含むホットスポットである。提案手法では、これらのホットスポットの関連性が抽出されている。実際に、これらのホットスポットでは、そのデモに関連するタグが多く含まれている。また、図5.2h-gの関連性については、タグ“Oxford Street”などが付与されている写真が多く含まれているホットスポット間の関連性である。そのため、これらのホットスポットでは、地理的な関連性が抽出された。これらのことより、提案手法を適用することで、ホットスポットに含まれる写真に付与されているタグに基づいてホットスポット間のイベントに関する関連性を抽出できていることがわかる。

図 5.3 に、Tower Bridge, Tower of London および City Hall 周辺で抽出されたホットスポットとその関連性について示す。Tower of London 周辺では、4つのホットスポットが抽出された(図 5.3 の黒色の丸)。Tower Bridge 周辺では、3つのホットスポットが抽出された(図 5.3 の赤色の丸)。City Hall 周辺では、2つのホットスポットが抽出された(図 5.3 の橙色の丸)。また、提案手法によって、Tower of London と、Tower Bridge の関連性は抽出された。実際に、それらのホットスポットでは、お互いの建物を撮影した写真が多く含まれており、それらに関連するタグも多く付与されている。しかしながら、City Hall と Tower Bridge の関連性は抽出されていない。これは、Tower Bridge 周辺で抽出されたホットスポットに含まれる写真の中で、City Hall を撮影している写真が少ないことが原因であると考えられる。提案手法では、ホットスポットに含まれる写真に付与されているタグに基づいて、ホットスポット間の類似度を TFIHF を用いて関連性を抽出している。そのため、ホットスポットに含まれている写真に、本来抽出されるべき関連性に関するタグが十分に含まれていない場合は、提案手法では、関連性を抽出することができない。このような場合の対処は今後の課題とする。

最後に、表 5.1 の R7 に Big Ben 周辺で抽出されたホットスポットと、抽出された関連性を図 5.4 示す。黒線は、提案手法によって抽出された撮影スポットと被写体の関連性を表している。左のホットスポット(図 5.4a)は、Big Ben を撮影した写真を多く含む撮影スポットである。中央のホットスポット(図 5.4b)は、実際の Big Ben の場所を含むホットスポットである。提案手法により、これらのホットスポットについて、撮影スポットと被写体の関連性が抽出された。左の撮影スポットは、Big Ben を撮影するスポットであると考えられるため、提案手法によって適切に関連性が抽出されたと考えられる。しかしながら、図 5.4 において、他の抽出されたホットスポットについても、Big Ben を撮影した写真が多く含まれているが、提案手法によって関連性を抽出することが出来なかった。これについて、これらのホットスポットにおいて、撮影方向情報が付与されている写真が少なかったため、提案手法が有効に機能しなかったと考えられる。

提案手法によって、ホットスポット間の撮影スポットと被写体の関連性と、タグに基づいた関連性を多くの場合に抽出することができた。これは、ホットスポットとして抽出されるような場所では、多くの写真が撮影されているため、関連性の算出に用いるメタデータが付与されている写真が比較的多かったためである。しかしながら、一部のホットスポットにおいて、関連性を抽出できなかった場合もある。これは、ホットスポットに含まれる写真に十分にメタデータが付与されていなかったことが原因の1つである。そのため、4章で述べたようなメタデータの推

定手法と組み合わせることで、このような課題に対処することが今後の課題の1つとしてあげられる。

5.4 おわりに

本章では、ユーザが画像検索結果を閲覧する際に、地理的な観点に基づいて効率的に閲覧可能にするために、地理的なクラスタ(ホットスポット)を抽出し、その関連性を算出する手法を提案した。提案手法では、ホットスポット間の撮影スポットと被写体の関連性と、タグに基づいた関連性の2つの種類の関連性を抽出した。関連性を抽出する際に、画像に付与されたメタデータである、写真の撮影方向情報、撮影位置情報、ユーザによるタグ、および撮影時間情報を用いた。また、提案手法をソーシャルメディアサイトから取得した画像に対して適用し、地図上に可視化することで、提案手法によって抽出された関連性についての検討を行なった。タグに基づいた関連性については、イベント、地域性、または被写体の3つの種類の関連性について抽出することができた。また、撮影スポットと被写体の関連性については、被写体の周辺の撮影スポットから関連性を抽出することができた。

今後の課題として、1つの地点が複数の被写体を撮影可能なホットスポットである場合に、それを複数のホットスポットに分割して、関連性を抽出することがあげられる。これは、1つの地点の周辺に複数のランドマークが存在する場合、提案手法では、複数のランドマークに影響されて1つの関連性を抽出するためである。

第6章 結言

6.1 成果のまとめ

本論文では、画像に付与されているメタデータの不均質さによってクラスタリングの性能が低下することを防ぐためのメタデータの推定・修正手法を提案した。画像に付与されているメタデータは、ユーザによって自由に付与されることが多いため、正確ではない場合や、メタデータが付与されていない場合がある。そのような画像を含む画像検索結果などの画像データセットに対してクラスタリングを適用する場合に、画像間の類似度を正確に算出することができないため、クラスタリングの性能が低下することがある。そのため、提案手法によって、メタデータの不均質さに対処することは、重要な課題である。また、ユーザに画像検索結果を効果的に提示することを目的として、写真の審美的評価のリランキング手法と、地理的なクラスタの抽出とその関連性の抽出手法の2つを提案した。それぞれについて以下にまとめる。

3章では、画像検索結果に対するクラスタリング手法のための、メタデータの推定手法を提案した。提案手法では、画像検索結果に含まれる画像の中でタグや、Exifが付与されていない画像に対して、事前にトレーニングデータを準備する代わりに、画像検索結果内の類似画像を用いてメタデータを推定した。また、ソーシャルメディアサイトより取得した画像検索結果に対して、提案手法を適用した結果を2つの評価指標で評価し、ベースライン手法と比較して、提案手法の有効性を確認した。

4章では、写真の質に基づいた画像検索結果のリランキング手法を提案した。提案手法では、写真の質を評価するために、画素情報に基づいた画像処理を用いずに、写真のメタデータや、撮影機材の情報を用いた。提案手法の有効性を示すために、ソーシャルメディアサイトより取得した画像検索結果に対して、提案手法を適用した結果を3つの評価指標で評価し、既存手法と比較して、非常に高速であることを示した。

5章では、ユーザが画像検索結果を閲覧する際に、地理的な観点に基づいて効率的に閲覧可能にするために、地理的なクラスタ(ホットスポット)を抽出し、そ

の関連性を算出する手法を提案した。提案手法では、ホットスポット間の撮影スポットと被写体の関連性と、タグに基づいた関連性の2つの種類の関連性を抽出した。また、提案手法をソーシャルメディアサイトから取得した画像に対して適用し、地図上に可視化することで、提案手法によって抽出された関連性についての検討を行なった。

6.2 今後の展望

今後の展望として、画像検索結果のメタデータの推定手法に用いるメタデータの種類数の増加があげられる。近年、GPS付きのモバイルデバイスで撮影された写真には、位置情報や、撮影方向情報、高度情報が付与されている場合がある。本論文では、画像に付与されているメタデータとして、ソーシャルタギングによるタグや、Exifに含まれるApertureや、ISO Speedなどを使用しているが、これらのメタデータと組み合わせることで、より効果的な画像クラスタリング結果を作成できると考えている。具体的には、画像の見た目、意味内容、撮影状況の3つの特徴量に加えて、クラスタリングに用いる特徴量の観点に、地理情報が加わる。例えば、画像の見た目がほとんど同じでも、撮影された場所が異なる写真を別のクラスターに分けることができる。そのため、地理情報に関するメタデータを提案手法に取り入れることは、本研究において重要な課題であると考えている。

謝辞

本論文の執筆では、多くの方々からのご支援とご協力をいただきました。ここでは、それらの方々への感謝の気持ちを述べます。本研究を進めるにあたり、指導教員である静岡大学情報学部横山昌平講師には、貴重な時間を割いて懇切丁寧なるご指導・ご鞭撻を頂きました。また、様々な面で議論し支えていただいた首都大学東京システムデザイン学部石川博教授，静岡大学情報学部福田直樹講師に心から感謝いたします。また，論文の内容に関して有益なご助言を頂いた静岡大学情報学部小西達裕教授，佐治斉教授，青木徹教授に厚くお礼申し上げますゼミ等を通じて有益な議論やご指摘を多数頂きました静岡大学情報学部情報科学科横山研究室の皆様，及び友人・知人の皆様に心から感謝いたします。

参考文献

- [1] 消費動向調査. http://www.esri.cao.go.jp/jp/stat/shouhi/menu_shouhi.html.
- [2] Twitter. <https://twitter.com/>.
- [3] 喜連川優, 豊田正史, 田村孝之, 鍛冶伸裕, 今村誠, 高山泰博, 藤原聡子. 10.socio sense : 過去9年に及ぶ web アーカイブから社会の動きを読む (第2部:情報の高信頼蓄積・検索技術等の開発,<特集>学と産の連携による基盤ソフトウェアの先進的開発). 情報処理, Vol. 49, No. 11, pp. 1290–1296, nov 2008.
- [4] Flickr. <http://www.flickr.com/>.
- [5] Panoramio. <http://www.panoramio.com/>.
- [6] Facebook. <https://www.facebook.com/>.
- [7] The man behind flickr on making the service 'awesome again'. <http://www.theverge.com/2013/3/20/4121574/flickr-chief-markus-spiering-talks-photos-and-marissa-mayer>.
- [8] 6,000,000,000. <http://blog.flickr.net/en/2011/08/04/6000000000/>.
- [9] The facebook photo machine. <http://royal.pingdom.com/2012/08/24/the-facebook-photo-machine/>.
- [10] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv.*, Vol. 40, No. 2, pp. 5:1–5:60, May 2008.
- [11] Najlae Idrissi, José Martinez, and Driss Aboutajdine. Bridging the semantic gap for texture-based image retrieval and navigation. *Journal of Multimedia*, Vol. 4, No. 5, pp. 277–283, 2009.

- [12] Dimitrios Skoutas, Enrico Minack, and Wolfgang Nejdl. Increasing diversity in web search results. In *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*, WebSci '10, Raleigh, NC, USA, 2010.
- [13] Google images. <http://www.google.com/imghp>.
- [14] Yahoo search - images. <http://image.search.yahoo.com/>.
- [15] Lin Chen, Dong Xu, I.W. Tsang, and Jiebo Luo. Tag-based web photo retrieval improved by batch mode re-tagging. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 3440–3446, June 2010.
- [16] Aixin Sun, Sourav S Bhowmick, Khanh Tran Nam Nguyen, and Ge Bai. Tag-based social image retrieval: An empirical evaluation. *Journal of the American Society for Information Science and Technology*, Vol. 62, No. 12, pp. 2364–2381, 2011.
- [17] 木本晴夫. 感性語による画像検索とその精度評価 (<特集> 人文科学とコンピュータ). *情報処理学会論文誌*, Vol. 40, No. 3, pp. 886–898, mar 1999.
- [18] 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 Trans. Pattern Anal. Mach. Intell.*, Vol. 22, No. 12, pp. 1349–1380, December 2000.
- [19] Max J Egenhofer. Query processing in spatial-query-by-sketch. *Journal of Visual Languages & Computing*, Vol. 8, No. 4, pp. 403–424, 1997.
- [20] Eugenio Di Sciascio, G. Mingolla, and Marina Mongiello. Content-based image retrieval over the web using query by sketch and relevance feedback. In *Proceedings of the Third International Conference on Visual Information and Information Systems*, VISUAL '99, pp. 123–130, London, UK, UK, 1999. Springer-Verlag.
- [21] 吉川日出行. サーチャーキテクチャ : 「さがす」の情報科学. ソフトバンククリエイティブ, 2007.

- [22] Benyu Zhang, Hua Li, Yi Liu, Lei Ji, Wensi Xi, Weiguo Fan, Zheng Chen, and Wei-Ying Ma. Improving web search results using affinity graph. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '05, pp. 504–511, New York, NY, USA, 2005. ACM.
- [23] T. Leelanupab, M. Halvey, and J.M. Jose. Application and evaluation of multi-dimensional diversity. In *Theseus/ImageCLEF workshop on visual information retrieval evaluation*, ImageCLEF '09, October 2009.
- [24] Reinier H. van Leuken, Lluís Garcia, Ximena Olivares, and Roelof van Zwol. Visual diversification of image search results. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, pp. 341–350, New York, NY, USA, 2009. ACM.
- [25] Roelof van Zwol, Börkur Sigurbjornsson, Ramu Adapala, Lluís Garcia Pueyo, Abhinav Katiyar, Kaushal Kurapati, Mridul Muralidharan, Sudar Muthu, Vanessa Murdock, Polly Ng, Anand Ramani, Anuj Sahai, Sri-ram Thiru Sathish, Hari Vasudev, and Upendra Vuyyuru. Faceted exploration of image search results. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pp. 961–970, New York, NY, USA, 2010. ACM.
- [26] Carl L. Sable and Vasileios Hatzivassiloglou. Text-based approaches for the categorization of images. In *Proceedings of the Third European Conference on Research and Advanced Technology for Digital Libraries*, ECDL '99, pp. 19–38, London, UK, UK, 1999. Springer-Verlag.
- [27] Tao Mei, Yong Rui, Shipeng Li, and Qi Tian. Multimedia search reranking: A literature survey. *ACM Comput. Surv.*, Vol. 46, No. 3, pp. 38:1–38:38, January 2014.
- [28] Marina Drosou and Evaggelia Pitoura. Search result diversification. *SIGMOD Rec.*, Vol. 39, No. 1, pp. 41–47, September 2010.
- [29] Kai Song, Yonghong Tian, Wen Gao, and Tiejun Huang. Diversifying the image retrieval results. In *Proceedings of the 14th Annual ACM International*

- Conference on Multimedia*, MULTIMEDIA '06, pp. 707–710, New York, NY, USA, 2006. ACM.
- [30] Charles L.A. Clarke, Maheedhar Kolla, Gordon V. Cormack, Olga Vechtomova, Azin Ashkan, Stefan Büttcher, and Ian MacKinnon. Novelty and diversity in information retrieval evaluation. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '08, pp. 659–666, New York, NY, USA, 2008. ACM.
- [31] Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, and Georg Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th International Conference on World Wide Web*, WWW '05, pp. 22–32, New York, NY, USA, 2005. ACM.
- [32] W Bruce Croft, Stephen Cronen-Townsend, and Victor Lavrenko. Relevance feedback and personalization: A language modeling perspective. In *DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries*, Vol. 3, 2001.
- [33] Deng Cai, Xiaofei He, Zhiwei Li, Wei-Ying Ma, and Ji-Rong Wen. Hierarchical clustering of www image search results using visual, textual and link information. In *Proceedings of the 12th Annual ACM International Conference on Multimedia*, MULTIMEDIA '04, pp. 952–959, New York, NY, USA, 2004. ACM.
- [34] Pierre-Alain Moëllic, Jean-Emmanuel Haugeard, and Guillaume Pitel. Image clustering based on a shared nearest neighbors approach for tagged collections. In *Proceedings of the 2008 International Conference on Content-based Image and Video Retrieval*, CIVR '08, pp. 269–278, New York, NY, USA, 2008. ACM.
- [35] Michael Bendersky and W. Bruce Croft. Analysis of long queries in a large scale search log. In *Proceedings of the 2009 Workshop on Web Search Click Data*, WSCD '09, pp. 8–14, New York, NY, USA, 2009. ACM.
- [36] Dian W. Tjondronegoro, Amanda H. Spink, and Bernard Jansen. Multimedia web searching on a meta-search engine. In *Twelfth Australasian*

- Document Computing Symposium*, ADCS '07, pp. 80–83, Melbourne Zoo, Australia, 2007. RMIT University.
- [37] Zhicheng Dou, Sha Hu, Kun Chen, Ruihua Song, and Ji-Rong Wen. Multi-dimensional search result diversification. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11, pp. 475–484, New York, NY, USA, 2011. ACM.
- [38] Ritendra Datta, Jia Li, and James Z. Wang. Content-based image retrieval: Approaches and trends of the new age. In *Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval*, MIR '05, pp. 253–262, New York, NY, USA, 2005. ACM.
- [39] Linjun Yang and Alan Hanjalic. Supervised reranking for web image search. In *Proceedings of the International Conference on Multimedia*, MM '10, pp. 183–192, New York, NY, USA, 2010. ACM.
- [40] Adrian Popescu, Pierre-Alain Moëllic, Ioannis Kanellos, and Rémi Landais. Lightweight web image reranking. In *Proceedings of the 17th ACM International Conference on Multimedia*, MM '09, pp. 657–660, New York, NY, USA, 2009. ACM.
- [41] Rong Yan, Alexander Hauptmann, and Rong Jin. Multimedia search with pseudo-relevance feedback. In *Proceedings of the 2Nd International Conference on Image and Video Retrieval*, CIVR '03, pp. 238–247, Berlin, Heidelberg, 2003. Springer-Verlag.
- [42] Zheng-Jun Zha, Linjun Yang, Tao Mei, Meng Wang, and Zengfu Wang. Visual query suggestion. In *Proceedings of the 17th ACM International Conference on Multimedia*, MM '09, pp. 15–24, New York, NY, USA, 2009. ACM.
- [43] Wikipedia. <http://en.wikipedia.org/>.
- [44] Masaharu Hirota, Naoki Fukuta, Shohei Yokoyama, and Hiroshi Ishikawa. A robust clustering method for missing metadata in image search results. *Journal of Information Processing*, Vol. 20, No. 3, pp. 537–547, 2012.

- [45] M. Hirota, N. Fukuta, S. Yokoyama, and H. Ishikawa. Implementing constraint-based clustering for a photo search system using estimated metadata. In *2nd International Symposium on Applied Informatics, ISAI '11*, 2011.
- [46] Masaharu Hirota, Shohei Yokoyama, Naoki Fukuta, and Hiroshi Ishikawa. Constraint-based clustering of image search results using photo metadata and low-level image features. *Computer and Information Science 2010*, Vol. 317, p. 165, 2010.
- [47] Fatemeh Alamdar and MohammadReza Keyvanpour. Effective browsing of image search results via diversified visual summarization by clustering and refining clusters. *Signal, Image and Video Processing*, Vol. 8, No. 4, pp. 699–721, 2014.
- [48] Feng Jing, Changhu Wang, Yuhuan Yao, Kefeng Deng, Lei Zhang, and Wei-Ying Ma. Igroup: A web image search engine with semantic clustering of search results. In *Proceedings of the 14th Annual ACM International Conference on Multimedia, MULTIMEDIA '06*, pp. 497–498, New York, NY, USA, 2006. ACM.
- [49] Haoyang Ding, Jing Liu, and Hanqing Lu. Hierarchical clustering-based navigation of image search results. In *Proceedings of the 16th ACM International Conference on Multimedia, MM '08*, pp. 741–744, New York, NY, USA, 2008. ACM.
- [50] 井手一郎, 柳井啓司. セマンティックギャップを越えて : 画像・映像の内容理解に向けて. *人工知能学会誌*, Vol. 24, No. 5, pp. i–ii, 691–699, sep 2009.
- [51] Ying Liu, Dengsheng Zhang, Guojun Lu, and Wei-Ying Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, Vol. 40, No. 1, pp. 262–282, January 2007.
- [52] Phei-Chin Lim, Narayanan Kulathuramaiyer, and Dayang NurFatimah Awg Iskandar. Towards semantic clustering—a brief overview. *International Journal of Image Processing (IJIP)*, Vol. 4, No. 6, p. 557, 2011.

- [53] Bin Gao, Tie-Yan Liu, Tao Qin, Xin Zheng, Qian-Sheng Cheng, and Wei-Ying Ma. Web image clustering by consistent utilization of visual features and surrounding texts. In *Proceedings of the 13th Annual ACM International Conference on Multimedia*, MULTIMEDIA '05, pp. 112–121, New York, NY, USA, 2005. ACM.
- [54] Meng Wang, Dinghuang Ji, Qi Tian, and Xian-Sheng Hua. Intelligent photo clustering with user interaction and distance metric learning. *Pattern Recogn.*, Vol. 33, No. 4, pp. 462–470, March 2012.
- [55] R. Bellman and Rand Corporation. Dynamic programming. In *Rand Corporation research study*. Princeton University Press, 1957.
- [56] Guangyu Zhu, Shuicheng Yan, and Yi Ma. Image tag refinement towards low-rank, content-tag prior and error sparsity. In *Proceedings of the International Conference on Multimedia*, MM '10, pp. 461–470, New York, NY, USA, 2010. ACM.
- [57] Lyndon S. Kennedy, Shih-Fu Chang, and Igor V. Kozintsev. To search or to label?: Predicting the performance of search-based automatic image classifiers. In *Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval*, MIR '06, pp. 249–258, New York, NY, USA, 2006. ACM.
- [58] 大森雅己, 廣田雅春, 石川博, 横山昌平. Flickr は海岸線を描けるか? 第6回データ工学と情報マネジメントに関するフォーラム, DEIM '14, 淡路夢舞台&ウェスティン淡路, 兵庫県, 2014.
- [59] Yang Yang, Zi Huang, Heng Tao Shen, and Xiaofang Zhou. Mining multi-tag association for image tagging. *World Wide Web*, Vol. 14, No. 2, pp. 133–156, March 2011.
- [60] Sihyoung Lee, Wesley De Neve, and Yong Man Ro. Tag refinement in an image folksonomy using visual similarity and tag co-occurrence statistics. *Image Commun.*, Vol. 25, No. 10, pp. 761–773, November 2010.
- [61] Changhu Wang, Feng Jing, Lei Zhang, and Hong-Jiang Zhang. Scalable search-based image annotation of personal images. In *Proceedings of the 8th*

- ACM International Workshop on Multimedia Information Retrieval*, MIR '06, pp. 269–278, New York, NY, USA, 2006. ACM.
- [62] Tat-Seng Chua, Jinhui Tang, Richang Hong, Haojie Li, Zhiping Luo, and Yantao Zheng. Nus-wide: A real-world web image database from national university of singapore. In *Proceedings of the ACM International Conference on Image and Video Retrieval*, CIVR '09, pp. 48:1–48:9, New York, NY, USA, 2009. ACM.
- [63] Bo Geng, Linjun Yang, Chao Xu, Xian-Sheng Hua, and Shipeng Li. The role of attractiveness in web image search. In *Proceedings of the 19th ACM International Conference on Multimedia*, MM '11, pp. 63–72, New York, NY, USA, 2011. ACM.
- [64] Congcong Li, Alexander C. Loui, and Tsuhan Chen. Towards aesthetics: A photo quality assessment and photo selection system. In *Proceedings of the International Conference on Multimedia*, MM '10, pp. 827–830, New York, NY, USA, 2010. ACM.
- [65] Xiaoshuai Sun, Hongxun Yao, Rongrong Ji, and Shaohui Liu. Photo assessment based on computational visual attention model. In *Proceedings of the 17th ACM international conference on Multimedia*, MM '09, pp. 541–544, New York, NY, USA, 2009. ACM.
- [66] Yiwen Luo and Xiaoou Tang. Photo and video quality evaluation: Focusing on the subject. In *Proceedings of the 10th European Conference on Computer Vision: Part III*, ECCV '08, pp. 386–399, Berlin, Heidelberg, 2008. Springer-Verlag.
- [67] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang. Studying aesthetics in photographic images using a computational approach. In *Proceedings of the 9th European conference on Computer Vision - Volume Part III*, ECCV '06, pp. 288–301, Berlin, Heidelberg, 2006. Springer-Verlag.
- [68] Yan Ke, Xiaoou Tang, and Feng Jing. The design of high-level features for photo quality assessment. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume*

- 1, CVPR '06, pp. 419–426, Washington, DC, USA, 2006. IEEE Computer Society.
- [69] Masaharu Hirota, Shohei Yokoyama, Naoki Fukuta, and Hiroshi Ishikawa. Reranking image search result based on photographic quality assessment without image features. In *IADIS International Conference WWW/Internet 2012*, ICWI '12, Madrid Spain, 2012.
- [70] David J. Crandall, Lars Backstrom, Daniel Huttenlocher, and Jon Kleinberg. Mapping the world's photos. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, pp. 761–770, New York, NY, USA, 2009. ACM.
- [71] Slava Kisilevich, Florian Mansmann, and Daniel Keim. P-dbscan: A density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos. In *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application*, COM.Geo '10, pp. 38:1–38:4, New York, NY, USA, 2010. ACM.
- [72] Motohiro Shirai, Masaharu Hirota, Shohei Yokoyama, Naoki Fukuta, and Hiroshi Ishikawa. Discovering multiple hotspots using geo-tagged photographs. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, pp. 490–493, New York, NY, USA, 2012. ACM.
- [73] Jesse Prabawa Gozali, Min-Yen Kan, and Hari Sundaram. Hidden markov model for event photo stream segmentation. In *Proceedings of the 2012 IEEE International Conference on Multimedia and Expo Workshops*, ICMEW '12, pp. 25–30, Washington, DC, USA, 2012. IEEE Computer Society.
- [74] Pinaki Sinha and Ramesh Jain. Classification and annotation of digital photos using optical context data. In *Proceedings of the 2008 International Conference on Content-based Image and Video Retrieval*, CIVR '08, pp. 309–318, New York, NY, USA, 2008. ACM.
- [75] Tao Mei, Bin Wang, Xian-Sheng Hua, He-Qin Zhou, and Shipeng Li. Probabilistic multimodality fusion for event based home photo clustering. In

- Multimedia and Expo, 2006 IEEE International Conference on*, pp. 1757–1760. IEEE, 2006.
- [76] Hongqi Wang, Olana Missura, Thomas Gärtner, and Stefan Wrobel. Context-based clustering of image search results. In *Proceedings of the 32Nd Annual German Conference on Advances in Artificial Intelligence*, KI '09, pp. 153–160, Berlin, Heidelberg, 2009. Springer-Verlag.
- [77] HengTao Shen, Kian-Lee Tan, Xiaofang Zhou, and Bin Cui. Icicle: A semantic-based retrieval system for www images. *Multimedia Systems*, Vol. 11, No. 5, pp. 438–454, 2006.
- [78] Hua-Jun Zeng, Qi-Cai He, Zheng Chen, Wei-Ying Ma, and Jinwen Ma. Learning to cluster web search results. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '04, pp. 210–217, New York, NY, USA, 2004. ACM.
- [79] Yangqing Jia, Jingdong Wang, Changshui Zhang, and Xian-Sheng Hua. Finding image exemplars using fast sparse affinity propagation. In *Proceedings of the 16th ACM International Conference on Multimedia*, MM '08, pp. 639–642, New York, NY, USA, 2008. ACM.
- [80] Jingdong Wang, Liyan Jia, and Xian-Sheng Hua. Interactive browsing via diversified visual summarization for image search results. *Multimedia Systems*, Vol. 17, No. 5, pp. 379–391, 2011.
- [81] M. A. Stricker and M. Orengo. Similarity of color images. In *Storage and Retrieval for Image and Video Databases III*, Vol. 2420 of *SPIE '95*, pp. 381–392, March 1995.
- [82] Brendan J Frey and Delbert Dueck. Clustering by passing messages between data points. *science*, Vol. 315, No. 5814, pp. 972–976, 2007.
- [83] Igor Fischer. New methods for spectral clustering. *Technical Report No. IDSIA-12-04*, 2004.

- [84] Arindam Banerjee, Srujana Merugu, Inderjit S. Dhillon, and Joydeep Ghosh. Clustering with bregman divergences. *J. Mach. Learn. Res.*, Vol. 6, pp. 1705–1749, December 2005.
- [85] R. A. Jarvis and E. A. Patrick. Clustering using a similarity measure based on shared near neighbors. *IEEE Trans. Comput.*, Vol. 22, No. 11, pp. 1025–1034, November 1973.
- [86] P. C. Ng and S. Henikoff. Sift: predicting amino acid changes that affect protein function. *Nucleic Acids Research*, Vol. 31, No. 13, pp. 3812–3814, 2003.
- [87] S. Dhar, V. Ordonez, and T. L. Berg. High level describable attributes for predicting aesthetics and interestingness. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '11*, pp. 1657–1664, Washington, DC, USA, 2011. IEEE Computer Society.
- [88] Subhabrata Bhattacharya, Rahul Sukthankar, and Mubarak Shah. A framework for photo-quality assessment and enhancement based on visual aesthetics. In *Proceedings of the International Conference on Multimedia, MM '10*, pp. 271–280, New York, NY, USA, 2010. ACM.
- [89] Xiaoshuai Sun, Hongxun Yao, Rongrong Ji, and Shaohui Liu. Photo assessment based on computational visual attention model. In *Proceedings of the 17th ACM International Conference on Multimedia, MM '09*, pp. 541–544, New York, NY, USA, 2009. ACM.
- [90] Hsiao-Hang Su, Tse-Wei Chen, Chieh-Chi Kao, Winston H. Hsu, and Shao-Yi Chien. Scenic photo quality assessment with bag of aesthetics-preserving features. In *Proceedings of the 19th ACM International Conference on Multimedia, MM '11*, pp. 1213–1216, New York, NY, USA, 2011. ACM.
- [91] Phillipe Salembier and Thomas Sikora. *Introduction to MPEG-7: Multimedia Content Description Interface*. John Wiley & Sons, Inc., New York, NY, USA, 2002.
- [92] Savvas A. Chatzichristofis and Yiannis S. Boutalis. Cedd: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval.

- In *Proceedings of the 6th International Conference on Computer Vision Systems*, ICVS'08, pp. 312–322, Berlin, Heidelberg, 2008. Springer-Verlag.
- [93] H. TAMURA. Textural features corresponding to visual perception. *IEEE Trans. Syst. Man. Cybernet.*, Vol. 8, pp. 460–473, 1978.
- [94] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of Cal. Math. Soc.*, Vol. 35, No. 1, pp. 99–109, 1943.
- [95] Ian Davidson and S. S. Ravi. Agglomerative hierarchical clustering with constraints: Theoretical and empirical results. In *Proceedings of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases*, PKDD'05, pp. 59–70, Berlin, Heidelberg, 2005. Springer-Verlag.
- [96] 廣田雅春. A study on clustering method for image search results using constrained clustering and metadata estimation techniques. 静岡大学大学院情報学研究科情報学専攻修士論文, 2011, (非公開).
- [97] EB Fowlkes and CL Mallows. A method for comparing two hierarchical clusterings. *Journal of the American Statistical Association*, Vol. 78, No. 383, pp. 553–569, 1983.
- [98] Marina Meilă. Comparing clusterings—an information based distance. *J. Multivar. Anal.*, Vol. 98, No. 5, pp. 873–895, May 2007.
- [99] James Hays and Alexei A. Efros. Im2gps: estimating geographic information from a single image. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 0, pp. 1–8, 2008.
- [100] Amazon.com. <http://www.amazon.com/>.
- [101] Kalervo Järvelin and Jaana Kekäläinen. Ir evaluation methods for retrieving highly relevant documents. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '00, pp. 41–48, New York, NY, USA, 2000. ACM.
- [102] Kuo-Yen Lo, Keng-Hao Liu, and Chu-Song Chen. Assessment of photo aesthetics with efficiency. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pp. 2186–2189, Nov 2012.

[103] Google maps. <https://www.google.com/maps/>.

論文リスト

学位論文審査資格に係る論文

1. Masaharu Hirota, Naoki Fukuta, Shohei Yokoyama and Hiroshi Ishikawa, A Robust Clustering Method for Missing Metadata in Image Search Results, Journal of Information Processing, Vol.53, No.3, pp.537-547, 7.2012

学位論文内容に係る論文

1. Masaharu Hirota, Naoki Fukuta, Shohei Yokoyama and Hiroshi Ishikawa, Constraint-based Clustering of Image Search Results Using Multiple Features, International Journal of Computer & Information Science, ACIS, Vol.11, No.4, pp. 26-36, 2010

その他の論文

1. 白井 元浩, 廣田 雅春, 石川 博, 横山 昌平, ジオタグ付き写真を用いた関心領域と撮影スポットの発見, 電子情報通信学会論文誌 D, Vol.J97-D, No.4, pp.835-844, 2014.04.01

国際会議

1. Masaharu Hirota, Motohiro Shirai, Hiroshi Ishikawa and Shohei Yokoyama, Detecting Relations of Hotspots using Geo-tagged Photographs in Social Media Sites, first international acm workshop on managing and mining enriched geo-spatial data (GeoRich2014), 6.2014
2. Motohiro Shirai, Masaharu Hirota, Hiroshi Ishikawa and Shohei Yokoyama, A method of area of interest and shooting spot detection using geo-tagged

photographs, ACM SIGSPATIAL Workshop on Computational Models of Place 2013 at ACM SIGSPATIAL GIS 2013 (CoMP2013), 11.2013

3. Motohiro Shirai, Masaharu Hirota, Shohei Yokoyama, Naoki Fukuta and Hiroshi Ishikawa, Discovering Multiple HotSpots using Geo-tagged Photographs, 20th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS 2012), 11.2012
4. Masaharu Hirota, Shohei Yokoyama, Naoki Fukuta and Hiroshi Ishikawa, RERANKING IMAGE SEARCH RESULT BASED ON PHOTOGRAPHIC QUALITY ASSESSMENT WITHOUT IMAGE FEATURES, IADIS International Conference WWW/Internet (ICWI2012), 10.2012
5. Masaharu Hirota, Naoki Fukuta, Shohei Yokoyama and Hiroshi Ishikawa, Implementing Constraint-based Clustering for a Photo Search System Using Estimated Metadata, 2nd International Symposium on Applied Informatics (ISAI2011), 8.2011
6. Masaharu Hirota, Shohei Yokoyama, Naoki Fukuta, Hiroshi Ishikawa, Constraint-based Clustering of Image Search Results Using Photo Metadata and Low-level Image Features, 9th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2010), pp.165-178, 8.2010

国内会議

1. 三浦 惇貴, 廣田 雅春, 野澤 浩樹, 横山 昌平, 地域ブログサイト全量解析による Local Expert ユーザのランキング手法, ARG 第4回 Web インテリジェンスとインタラクション研究会 (ARG WI2 2014), 5.2014
2. 廣田 雅春, 石川 博, 横山 昌平, ジオタグ付き写真を用いた写真の撮影方向の推定手法, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 3.2014
3. 青山 賢, 廣田 雅春, 石川 博, 横山 昌平, 写真に付与されたジオタグに基づいた道草発見, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 3.2014

4. 井上 貴弘, 廣田 雅春, 石川 博, 横山 昌平, コンテキストに基づいたフォトストリームの分割を支援するジオタグ軌跡の集約的分析, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 3.2014
5. 三浦 惇貴, 廣田 雅春, 野澤 浩樹, 横山 昌平, ゲーミフィケーションをユーザーに無負担で導入するための Local Expert 抽出手法～浜松を対象とした地域ブログサイト全量解析による Venue の発見～, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 3.2014
6. 大羽 洋隆, 廣田 雅春, 石川 博, 横山 昌平, ジオタグ付き写真を用いた土地被覆の分析と可視化, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 3.2014
7. 大森 雅己, 廣田 雅春, 石川 博, 横山 昌平, Flickr は海岸線を描けるか?, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 3.2014
8. 野呂 勇太, 三浦 惇貴, 廣田 雅春, 野澤 浩樹, 横山 昌平, イベントの時空間推定に基づくブログとマイクロブログの統合閲覧システムの提案, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 3.2014
9. 白井 元浩, 廣田 雅春, 石川 博, 横山 昌平, ジオタグ付き写真を用いたホットスポットの分類と関連の抽出, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM2014), 3.2014
10. 大羽 洋隆, 廣田 雅春, 横山 昌平, 石川 博, ジオタグ付き写真を用いた地図の彩色システムの構築, 第5回データ工学と情報マネジメントに関するフォーラム (DEIM2013), 3.2013
11. 大森 雅己, 廣田 雅春, 横山 昌平, 石川 博, タグの時空間情報を考慮したジオタグ付き写真の可視化システムの開発, 第5回データ工学と情報マネジメントに関するフォーラム (DEIM2013), 3.2013
12. 野呂 勇太, 廣田 雅春, 横山 昌平, 石川 博, マイクロブログにおけるハッシュタグ自動付与システムの提案, 第5回データ工学と情報マネジメントに関するフォーラム (DEIM2013), 3.2013
13. 白井 元浩, 廣田 雅春, 横山 昌平, 石川 博, ジオタグ付き写真を用いたホットスポットの分類とランドマークの形状抽出手法, 第5回データ工学と情報マネジメントに関するフォーラム (DEIM2013), 3.2013

14. 粟立軍, 廣田 雅春, 横山 昌平, 石川 博, マイクロブログにおけるクラスタリング技術の比較, 第5回データ工学と情報マネジメントに関するフォーラム (DEIM2013), 3.2013
15. 坂本 翼, 廣田 雅春, 横山 昌平, 石川 博, Twitter の多様なハッシュタグ間の同義・階層関係の推定, 第5回データ工学と情報マネジメントに関するフォーラム (DEIM2013), 3.2013
16. 武藤 教宏, 廣田 雅春, 横山 昌平, 石川 博, GPU を用いたグラフデータ分割手法, 第5回データ工学と情報マネジメントに関するフォーラム (DEIM2013), 3.2013
17. 原木 司, 廣田 雅春, 横山 昌平, 石川 博, 道路ネットワークを用いた GPS 軌跡の特徴点を保持する圧縮アルゴリズムの提案, ARG 第1回 Web インテリジェンスとインタラクション研究会 (ARG WI2 2012), 12.2012
18. 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, 画像情報を用いない写真の質判定手法の提案, 第4回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 3.2012
19. 後藤 康路, 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, MapReduce を用いた並列 SOM の高速化手法の提案, 第4回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 3.2012
20. 白井 元浩, 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, ジオタグ付き写真を用いたランドマークの多視点再現システム, 第4回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 3.2012
21. 玉井 孝典, 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, 研究者間コラボレーションを促進する超高解像度 WebGIS, 第4回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 3.2012
22. 坂本 翼, 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, Twitter ストリームのバーストの断続性に着目したキーワード抽出, 第4回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 3.2012
23. 原木 司, 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, 移動情報タグ生成システムにおける加速度センサを用いた消費電力削減手法の提案, 第4回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 3.2012

24. 武藤 教宏, 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, GPU と CPU の並列処理に適した XML インデクス生成と検索手法, 第 4 回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 3.2012
25. 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, 画像共有サイトに対する画像検索結果提示システムの開発, 第 4 回楽天研究開発シンポジウム, 11.2011
26. 廣田 雅春, 福田 直樹, 横山 昌平, 石川 博, 画像検索結果のクラスタリングを目的とした類似画像からのメタデータ推定手法の検討, 信学技報, Vol.111, No.70, AI2011-5, pp.21-26, 5.2011
27. 廣田 雅春, 福田 直樹, 横山 昌平, 石川 博, ウェブ上の類似画像を用いたメタデータの推定と画像検索結果提示への適用, 電子情報通信学会 2011 年総合大会, 3.2011
28. 廣田 雅春, 青島 傳隼, 横山 昌平, 福田 直樹, 石川 博, 知の構造化と発信を助ける可視化技術, ポスター・デモ (DBS 推薦), WebDB Forum 2010, 11.2010
29. 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, 撮影状況メタデータと画像低次特徴量を用いた画像検索結果提示システムの実装, 情報処理学会創立 50 周年記念全国大会, pp.883-884, 3.2010
30. 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, 撮影状況メタデータを用いた制約付きクラスタリングによる画像検索結果提示手法の検討と評価, 第 2 回データ工学と情報マネジメントに関するフォーラム (DEIM2010), 2.2010
31. 廣田 雅春, 横山 昌平, 福田 直樹, 石川 博, 撮影状況メタデータと画像低次特徴量を用いた画像検索結果提示手法の検討, 信学技報, Vol.109, No.386, AI2009-29, pp.61-66, 1.2010