

画像に関する時間情報の推定と画像検索への適用

山下 浩平[†] 山本 祐輔^{††,†††} アダムヤトフト^{††} ニミットパッタナスリ^{††}

田中 克己^{††}

[†] 京都大学工学部情報学科 〒606-8501 京都府京都市左京区吉田本町

^{††} 京都大学大学院情報学研究科社会情報学専攻 〒606-8501 京都府京都市左京区吉田本町

^{†††} 日本学術振興会特別研究員

E-mail: †{yamashita,yamamoto,adam,nimit,tanaka}@dl.kuis.kyoto-u.ac.jp

あらまし 本研究ではウェブ上の画像の撮影時間を推定する手法を提案する．提案手法は次の2ステップに分けられる．第1ステップでは、「ある画像*i*とある時刻*t*との関連が強い場合，画像*i*と類似する画像も時刻*t*および*t*に近い時間と関連が強い」という仮定の下，画像・時間対間の関係分析を行い，対象画像に関連の強い時間を求める．第2ステップでは，求められた関連時間の中からウェブページの作成時間を用いて撮影時間を推定する．評価実験の結果，提案手法はページのURLからページの作成時間を抽出できた場合，平均10ヶ月の誤差で撮影時間を推定できることが示された．

キーワード 画像と時間の関連度，画像の撮影時間，画像検索

1. はじめに

近年，インターネットの普及により誰もがウェブ上に簡単に情報を発信できるようになり，ウェブ上に存在する情報量が爆発的に増え続けている．中でも最近ではテキストだけでなく，画像・音声・動画といったマルチメディアコンテンツが急激に増加しており，インターネットを利用している中でこのようなコンテンツをよく見かけるようになった．しかし，このようなコンテンツがいつ作成されたものであるかわからないような場面が多々ある．例えば10年前に撮影された，現在はすでに存在しない建物の写真にも関わらず，周辺テキストの内容が現在の事を中心に書かれているため，ユーザが写真の内容を現在のものであると誤解してしまう可能性が十分考えられる．また，ある人物の写真の撮影日時が最近であると記述されているが本当は10年前の写真であり，それを見たユーザが誤った記述を信じてしまい，その写真の人物が現在の姿であると誤解してしまう可能性も考えられる．

そこで本研究では，そのようなコンテンツの中でも特に画像に注目して，ウェブページに掲載された画像の撮影時間を推定する手法を提案する．画像に関連する時間の概念は様々なものが存在するが，その中で本研究で画像の撮影時間に限定したのは，他の様々な時間と比べて画像の信憑性と大きく関連していると考えているからである．画像と周辺テキストとの時間的な整合性や，画像と画像のメタデータに含まれる時間情報との整合性は信憑性判断に大きく影響される．例えば先ほど述べた，10年以上前の建物の写真の問題がまさに画像の信憑性の問題の一種である．したがって画像の撮影時間を推定することは画像の信憑性の分析に役立つと期待される．

現実的な問題としてウェブページの画像に関して，画像のメ

タデータに撮影時間情報・編集情報が含まれていなかったりその画像に関連する時間情報が不足していることが多い．また画像の撮影時間・編集時間に関する情報が含まれていても，時間情報に誤りがあったり改ざんされているケースもある．このような問題に対して画像に関連する時間情報の不足を補うために，本研究では類似画像とそれらを含むウェブページを用いて画像に関連する時間情報を増やし，画像の撮影時間の推定を行う．具体的には各々の画像に対して自身を含むウェブページからその画像と関連のある時間を抽出し，画像とその時間の対を生成する．他の類似画像も同様にしてその画像と関連のある時間との対を生成し，得られた全ての対から2部グラフを作成する．そしてある対に対して画像的にも時間的にも似た対が多く存在すればその対の関連は強くなるという仮定のもとで各々の対の関連の強さを計算し，画像と関連の強い時間の集合を抽出してそこから撮影時間の推定を行う．

このように，本研究の提案システムでウェブ上の画像の撮影時間の推定ができれば，先ほど述べた時間情報の不足や誤りによって生じる問題が解決されるだけでなく，画像検索にも適用が可能となる．例えばあるオブジェクトの最も撮影されたのが新しい画像が欲しかったり，ある期間内に撮影された画像だけが欲しかったり，さらには検索結果を時間順に並べ替えるといったことも可能となり，新たな画像検索サービスの実現が期待される．

本研究の構成は次のとおりである．2章では関連研究について述べ，既存の研究と本研究の違いを述べる．3章では提案システムのアルゴリズムの詳細を述べ，4章では現実の状況下での提案システムの実装について述べる．5章では提案システムが画像検索に適用できることを述べる．6章では提案システムがどの程度有効であるかを調べる実験について述べ，7章では

実験の考察を述べる．そして 8 章では本研究のまとめを述べる．

2. 関連研究

本章では本研究の関連研究について 3 つに分けて述べる．

2.1 時間を考慮したウェブ検索

近年、時間を考慮したウェブ検索の研究が盛んに行われている．特に対象となっているのが、ブログや、ニュースなど、書かれた時間が重要な意味を持つことが多いウェブ上の文書である．多くの研究では、時間を考慮しなくてはならないクエリ (time sensitive query) を対象として、ランキングにおいて新しい文書を上位に順位付けするといったことが行われる．

Dakka らは、言語モデルにおける情報検索のランキング手法の中において、時間を考慮したランキングを行うための一般的な手法を提案している [6]．実際に、ニュースアーカイブ検索において、新しいものを求めるようなクエリが与えられた際に、時間を考慮したランキングを行うことで、手法の評価を行っている．Zhang らは、クエリが潜在的に時間指定を含むような場合に、検索エンジンから返された検索結果を再ランキングすることで、適合率を上げる手法について提案している [7]．クエリログを利用して、どのようなクエリが、時間について考慮しなくてはならないクエリであるかということの判定手法の提案を行い、また、ある検索結果の文書がどの時間に適合するかということの推定手法の提案を行っている．Dong らも、時間を考慮しなくてはならないクエリであるかどうかの判定を行い、与えられたクエリが時間を考慮しなくてはならないものであった場合には、より新鮮度の高いものを上位に順位付けするようなランキングのためのモデルを提案している [1]．Arikan らは、時間指定されたクエリが現在の Web 検索エンジンでうまく機能しないことを改善するため、言語モデルベースのランキングアルゴリズムにおいて、時間を考慮する手法について提案している [2]．

これらの研究では文書についての時間を考慮している．特にニュース記事は、それが書かれた時間を推定することが容易であり、これらの研究における対象となっていることが多い．本研究が対象としているのは、文書ではなく、画像である．例えば、ニュースで利用される写真は、必ずしもそのニュースのために撮影された写真ではない場合も存在しており、画像という単位での時間推定についても検討すべきであると考えられる．

2.2 ウェブ画像のメタデータ取得手法

山本らは Web 上に存在する画像・テキスト対の信憑性を評価するモデルを提案している [8][9]．この手法の対象は画像・テキスト対であるが、これを画像・時間対としても有効であると考えられる．したがって本研究では山本らの手法を利用することを考える．しかし、山本らの手法をただそのまま使うだけでは本研究の目的を達成することができないため、本研究ではさらなる改良を加えている．

Wang らはウェブ上のある画像を、それと似た画像を含むページをウェブ上から抽出してそのページに含まれるテキストから画像と関係のあるタグを抽出する手法を提案している [3]．類似画像検索の技術を用いて関係のあるページを抽出して、そ

のページから必要な情報を抽出するという点で本研究と関連があるが、本研究が抽出したいのは時間情報であるという点で異なる．

山根らは指定されたクエリを含むブログ記事の時系列から画像を抽出し、テキスト情報と画像の特徴量を複合的に用いて時期ごとの画像のランキングを行い、提示する手法を提案している [4]．画像を時間的観点から捉え、時系列ごとに画像をランキングして表示するという点で本研究と大いに関連がある．しかし、画像の対象がタイムスタンプが明記されているブログ記事内の画像だけであるのに対し、本研究では時間情報が含まれているとは限らないウェブページに掲載された画像を対象としている点で異なる．

2.3 画像ファイルに対する時間情報

画像ファイルに対して時間情報のメタデータを付与する形式がいくつか存在している．まず、一般的なファイルシステムでは、あらゆるファイルにファイルの作成日時、ファイルの最終更新日時、ファイルの最終アクセス日時などが保持されていることが多く、画像の時間情報として、これらを利用することも可能である．

JPEG、TIFF などの形式のファイルに対しては、Exchangeable image file format (Exif) [11] と呼ばれる、主に写真画像のためのメタデータの形式が存在している．Exif 形式で付与できる情報は、以下のような情報である．

- 撮影日時
- 撮影に利用したカメラのメーカー名、モデル名
- 画像全体の解像度
- 水平・垂直方向の単位あたり解像度

これら以外にも、撮影時のシャッタースピードや、ISO 感度などが記録可能である．最近の多くのデジタルカメラでは、画像ファイルにこれらの情報が自動的に付与されることが多い．Exif 情報は編集可能であるが、あまり編集されることはないため、カメラの設定日時が正しければ、ある程度は信頼できる時間情報を取得することができる．画像編集ソフトでは、保存時に Exif 情報を保存しないものも存在するが、Adobe 社の Photoshop を含む一部のソフトでは、元の Exif 情報は保持した上で、編集日時を記録するというを行うことが可能である．

しかし、実際にはウェブ上の画像は Exif 情報が含まれていないものも多い．したがって画像の時間情報の欠落や誤りによる問題を防ぐために、撮影時間を推定する技術が必要となる

3. 画像撮影時間の推定アルゴリズム

本章では、画像を含むウェブページが与えられたときに、その画像の撮影時間だと思われる時間を推定するアルゴリズムを説明する．図 1 はその大まかな流れ図を表す．

画像の撮影時間を推定するために、まず画像と関連が強い時間の集合を求めて、その関連が強い時間の集合から撮影時間を推定する．画像と関連が強い時間とは、例えばその画像の撮影時間やその画像に写っている物体について言及された時間など様々である．

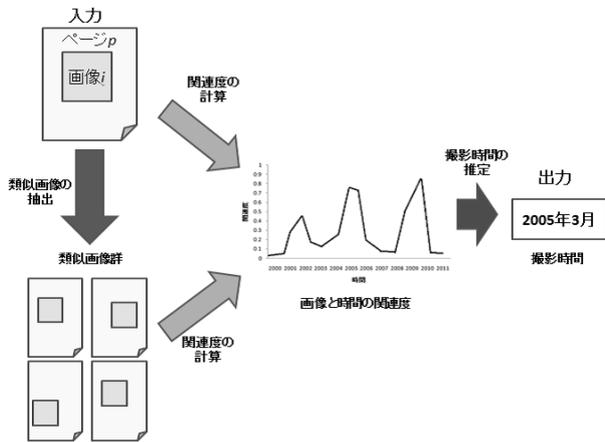


図 1 本手法の流れ図

画像と関連のある時間として、我々は画像を含むウェブページのテキスト中にある時間を利用することを考える。しかし、その 1 ページのテキストだけでは情報が少なすぎるため、画像の撮影時間を推定するのは困難である。そこで本研究では、元画像の類似画像を含むウェブページの集合を利用して、撮影時間を推定する手法を提案する。直観的には類似画像を含むウェブページには類似画像に関連する時間が含まれており、それは類似画像が元画像に似ていれば似ているほど元画像に関連する時間でもあることが多いと考えられる。したがって得られた各々の画像に対して関連する時間を各々のウェブページのテキストから抽出して、関連のある画像と時間の対を作成する。そしてそれらの対の関連の強さを、他の対の関連の強さを考慮しながら計算する。

その結果得られた元画像と関連が強い時間の集合から撮影時間を推定する。画像の撮影時間は当然その画像がウェブ上に現れた時間より前であるが、直感的にはその画像の撮影時間とウェブ上に現れた時間は近いことが多いと考えられる。したがって本アルゴリズムでは、元画像と類似画像を含むウェブページの作成日時をその画像がウェブ上に現れた時間だと考え、それらの作成日時と先ほどの対の関連の強さをを用いて撮影時間を推定する。

本章では、元画像の類似画像とその類似画像を含むウェブページを何らかの方法で集めてくることができ、またウェブページのテキストから画像に関連のある時間を正確に抽出することができ、さらに各々のウェブページの作成時間がわかっているという理想的な状況で撮影時間を推定する方法について述べる。

3.1 画像と時間の関連度の計算

本節では画像と関連が強い時間の集合を求めるために、まず画像 i と時間 t の画像・時間対 $\langle i, t \rangle$ を定義する。画像・時間対 $\langle i, t \rangle$ は画像 i と時間 t に何らかの関連があることを表している。その対の画像と時間の関連の強さを関連度と呼ぶことにし、画像 i と時間 t の関連度を $Score(i, t)$ と表す。画像 i と時間 t の関連が強いほど関連度 $Score(i, t)$ が高くなるものとする。

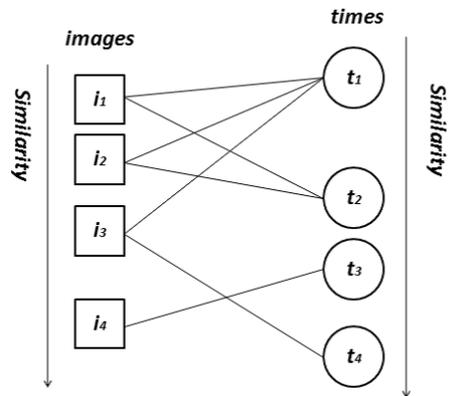


図 2 ウェブページに掲載された画像とその周辺テキストに出現する時間表現との対応を表す 2 部グラフ。

画像 i と関連のある時間は先ほど述べたように、画像 i を含むウェブページのテキスト中の時間のうち画像 i と関連のある時間全てであるとし、その各々の時間 t を抽出して対 $\langle i, t \rangle$ を生成する。これらの操作を元画像と類似画像の各々に対して実行し、得られた対全ての集合を $Pairs$ とする。ここで得られた画像と時間の対の集合 $Pairs$ は、図 2 のような 2 部グラフを形成すると考えられる。

関連度の強さの計算には、山本ら [8] [9] の考え方をを用いる。山本らは「画像 i とその説明テキスト t との整合性が高いケースでは、画像 i に類似する画像の説明テキストもテキスト t と類似した内容であることが多い」と述べている。これは画像・テキスト対だけでなく画像・時間対にも当てはまると考えられるので、本研究に当てはめると画像的にも時間的にも似た対が多く存在すると、その対は関連が強いと考えられる。つまり、ある対 $\langle i, t \rangle$ について、画像 i に似た画像と時間 t に近い時間の対が多く存在し、かつその各々の対の関連度が高い時、関連度 $Score(i, t)$ の値が高くなると考えられる。この仮定から、ある画像 i と時間 t の対の関連度 $Score(i, t)$ は以下のように定義することができる。

$$Score(i, t) = \sum_{\langle i', t' \rangle \in Pairs} Score(i, t) Sim_i(i, i') Sim_t(t, t') \quad (1)$$

ただし、 $Sim_i(i, i')$ は画像 i と i' の画像的な類似度を表し、画像 i と i' が似ているほど値が大きくなる。 $Sim_t(t, t')$ は時間 t と t' の時間的な類似度を表し、時間 t と t' が近いほど値が大きくなる。ここで定義された $Score(i, t)$ は再帰的であり、値が収束するまで繰り返す。そして収束した値を最終的な関連度 $Score(i, t)$ として用いる。

しかし、元画像 i_{ori} について、 i_{ori} を含むウェブページのテキスト中に時間表現 t が含まれていなければ、 i_{ori} と t の関連度 $Score(i_{ori}, t)$ が計算されることはない。このような場合、最終的にアルゴリズムの出力として撮影時間 t を出力することができない。したがって元画像 i_{ori} については、画像・時間対の集合 $Pairs$ に含まれる対の関連度の計算後に、 $Pairs$ に含まれていない元画像 i_{ori} と時間 t の対の関連度 $Score(i, t)$ を近似的に計算する必要がある。そこで t を含む他の対 $\langle i, t \rangle$ の集合

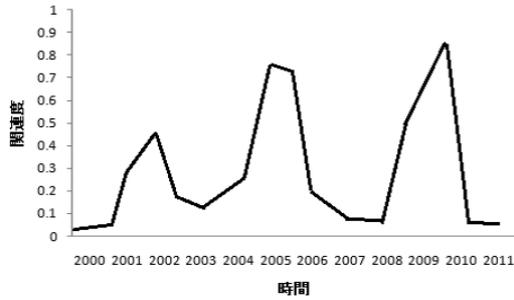


図3 画像と時間との関連度を示すグラフ．横軸は時間を表し，縦軸は画像と特定の時間との関連度の大きさを表している．

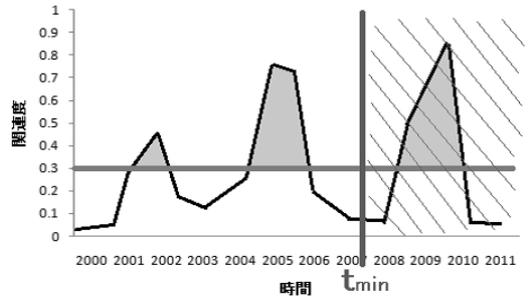


図5 ウェブページの作成時間を考慮した画像と関連の強い時間の絞り込み．時間 t_{min} 以降の関連度は無視される．

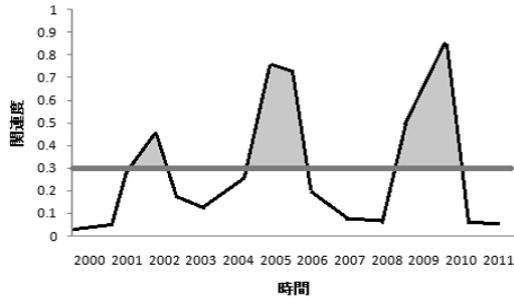


図4 画像と時間の関連度の大きさに注目した関連の強い時間のフィルタリング．関連度が極端に大きくなっている周辺の時間は，画像と特に関連の深い時間であると考えられる．

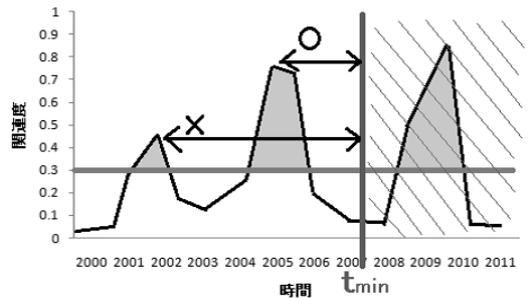


図6 関連度グラフを用いた撮影時間の推定．先に述べた方法で絞り込まれた関連の強い時間のうち時間 t_{min} に最も近い時間を，画像が撮影された時間として抽出する．

の関連度に重みとして i_{ori} と i の画像類似度をかけたものの平均値を $Score(i_{ori}, t)$ の近似値だと考える．それを式に表すと以下ようになる．

$$Score(i_{ori}, t) = \frac{\sum_{i \text{ s.t. } \langle i, t \rangle \in Pairs} Score(i, t) Sim_i(i_{ori}, i)}{|\{i | \langle i, t \rangle \in Pairs\}|} \quad (2)$$

ただし， $|x|$ は集合 x の要素数を表し，分母は集合 $Pairs$ 中の時間 t を含む対の個数を表す．もしその個数が 0 の場合，どの類似画像も時間 t に関連がないので，元画像 i_{ori} も時間 t と関連がある可能性は小さい．したがってこのような場合は

$$Score(i_{ori}, t) = 0 \quad (3)$$

と定義する．

以上の式 (1), (2), (3) から，元画像 i_{ori} に対する全ての時間 t の関連度 $Score(i_{ori}, t)$ が得られる．

3.2 撮影時間の推定

本節では前節で求められた元画像 i_{ori} に対する全ての時間 t の関連度 $Score(i_{ori}, t)$ をもとに元画像 i_{ori} の撮影時間を推定する．

前節で得られた関連度 $Score(i_{ori}, t)$ の値はグラフ化すると例えば図3のようになる．図3のように関連度が高くなっている山の部分がいくつか見られるが，この関連度が高い部分の時間が実際に画像と関連のある時間であり，関連度の低い部分が関連度の計算において発生したノイズであると考えられる．そしてこれらの山のうちの一つが画像の撮影時間を表していると考えられる．したがってまずは図4のように関連度が高くて山をなしている部分の時間を抽出する．これには図4のようにある一

定の閾値を与えて関連度がそれを上回るときの時間を抽出すればよい．

ところで，元画像の撮影時間は当然その元画像を含むウェブページの作成日時より以前である．他の類似画像についても同様のことが言え，結局元画像の撮影時間は元画像と類似画像を含むどのウェブページの作成日時よりも以前である．ただしこれは類似画像が元画像と同一で同じ物体に関する画像である場合であり，類似画像が元画像と似ているが実際には別の物体に関する画像である場合は普通成り立たない．したがって元画像と同一な画像を含むウェブページの作成日時のみを考えると，元画像の撮影時間は類似画像を含むどのウェブページの作成日時よりも以前であると考えられる．このことから撮影時間の推定の際，元画像およびほぼ同一な画像を含むウェブページの作成日時のうち最も古い時間を t_{min} とし，その t_{min} 以降の山は撮影時間にはなりえないので除去する．その様子を図5に示す．

最後に残った山から撮影時間を決定する．先ほど述べた t_{min} は元画像と同一画像のウェブへのアップロード時間のうち，最も古いものであると考えられる．そして最も昔にアップロードされた画像が，ウェブ全体に存在するその同一画像の中で一番最初にウェブに現れたと考えられる．そのような画像は実際に撮影されてからアップロードされるまでの期間が短いと仮定する．その仮定によりこの時点でいくつかまだ山が残っている可能性があるが，図6のようにその時点で最も t_{min} に近い山が撮影時間を含むとし，その山の中で最も関連度の高い時間を最終的に撮影時間とする．

以上より，与えられた画像の撮影時間を求めることができる．

4. システムの実装

前章で述べたアルゴリズムは、元画像の類似画像とその類似画像を含むウェブページを何らかの方法で集めてくることができ、さらにウェブページのテキストから画像に関連のある時間を正確に抽出することができるという理想的な状況を考えていた。しかし、現実には元画像の類似画像全てをウェブ上から収集することやウェブページのテキストから画像に関連のある時間だけを抽出するのは極めて困難である。そこで本章では、類似画像の収集には類似画像検索エンジン TinEye を用い、画像に関連のある時間はウェブページの対象画像付近の時間を抽出するという手法で本アルゴリズムを実装する。

本システムの動作は以下のステップに分けられる。

- (1) 撮影時間を推定したい画像 i とそれを含むウェブページ p を入力する
- (2) 画像 i の類似画像とそれを含むウェブページを収集する
- (3) ウェブページ p から画像 i に関連のある時間を抽出し、画像・時間対を生成する。
- (4) ステップ2で集めたウェブページからも時間を抽出し、その各々のウェブページに対し画像・時間対を生成する。
- (5) 画像と時間の関連度の計算をする
- (6) 画像 i の撮影時間を推定する
- (7) 推定された時間を返す

以下で各々のステップの詳細を述べる。

4.1 類似画像の収集

本節では類似画像検索エンジン TinEye^(注1)を用いた類似画像の収集と、画像と関連のある時間をウェブページのテキスト中から抽出して画像と時間の対を生成する手法を述べる。

類似画像の収集は、類似画像検索エンジン TinEye にクエリとして元画像を与えることで実現する。TinEye はクエリを与えた結果として類似画像とその類似画像を含むウェブページを返すので、その上位 N 件の類似画像を利用する。

4.2 画像と関連のある時間の抽出

次に、元画像や類似画像の含まれる各々のウェブページのテキストから画像に関連のある時間を抽出する。まずウェブページのテキストから時間を抽出する手法だが、これには正規表現を用いる。具体的には以下のような様々な時間をテキストから抽出する。

”August 29, 1988”
”Aug 29th 1988”
”08.29.88”
”in August 1988”
”in 1988”

そして抽出された時間を共通の形式 $t = YYYYMM$ に変換する。YYYY と MM はそれぞれ年と月を表しており、例えばテキストから”August 29, 1988”という時間が得られた場合は $t = 198808$ と変換される。ただし、”in 1988”のように

月に関する表現が含まれていない場合は $t_1 = 198801, t_2 = 198802, \dots, t_{12} = 198812$ のようにその年の1月から12月まで全てに関連があると考え、そしてこの時間 t を後で述べる画像と時間の対を扱う際に利用する。ここで共通の形式として年と月のみを扱い日を扱わなかったのは、推定された撮影時間と本当の撮影時間の誤差が日単位であったとしても現実的にはそれほど問題がない場合が多いと考えられるからである。

このようにしてウェブページのテキスト中から時間を抽出することが可能となる。しかし、ウェブページのテキスト中の様々な時間のうち、どの時間が画像と関連があるかを機械的に判断するのは極めて難しい。そこで本節では、ウェブページ中の対象画像との距離が小さい時間は画像と関連のある可能性が高いという仮定を置いて、その仮定から画像との距離が一定以下の時間だけを利用することを考える。画像と時間の距離は視覚的なものではなく画像と時間の間の文字数とする。Position(x) をテキスト中の x の位置を表す関数とすると、画像 i と時間 t の距離 Distance(i, t) は以下のように定義できる。

$$Distance(i, t) = |Position(i) - Position(t)| \quad (4)$$

そして画像 i との距離が一定以下の時間 t とは以下の条件を満たすものと定義できる。

$$\frac{Distance(i, t)}{TextLength} < \alpha \quad (5)$$

ただし、TextLength はウェブページのテキスト全体の長さを表し、 α はパラメータである。後で述べる実験では α の値は 0.15 としている。

以上の操作により、画像 i に対して i を含むウェブページのテキストから i と関連のある時間 t を抽出し、その組を対 $\langle i, t \rangle$ とする。そして得られた対全体の集合を Pairs とし、画像と時間の関連度の計算に用いる。

4.3 画像と時間の関連度の計算

本節では前節で得られた対の集合 Pairs から画像類似度 $Sim_i(i_1, i_2)$ と時間類似度 $Sim_t(t_1, t_2)$ に基づいて画像に対する時間の関連度 Score(i, t) を計算する。

画像から画像特徴量を抽出する手法は多数提案されている。近年では、色特徴量やテクスチャ特徴量のような大域的特徴量よりも画像のスケールや回転に無関係な特徴をとらえることができる局所特徴量が注目されている。今回の提案手法では、近年最も注目されている SIFT(scale invariant feature transform) 特徴量 [10] を用いて画像特徴量を表現し、画像間の類似度を計算する。画像に SIFT アルゴリズムを適用すると、キーポイントと呼ばれる画像の局所特徴量をよく表現する特徴点が複数抽出される。キーポイントは 128 次元のベクトルで記述されている。提案手法では、画像 i_1 と i_2 が共有するキーポイントの数を Common(i_1, i_2)、画像 i_1 と i_2 が持つキーポイントの総数 n_1 と n_2 の平均と定義して、割った値を画像 i_1 と i_2 の間の類似度として定義する。

$$Sim_i(i_1, i_2) = \frac{Common(i_1, i_2)}{\frac{n_1 + n_2}{2}} \quad (6)$$

時間類似度について、本研究における時間の類似度の役割は

(注1): <http://www.tineye.com/>

2つの時間 t_1, t_2 が互いに画像との関連度どの程度影響を与えるかを定めることであり、例えば t_1, t_2 の差が1, 2ヶ月のように近い場合は互いに影響を与え得るが、その差が1年以上の場合はほぼ影響を与えないと考えてよい。したがって時間類似度 $Sim_t(t_1, t_2)$ は t_1 と t_2 の差が小さい時だけ値が大きくなるような関数がよいので以下のように定義する。

$$Sim_t(t_1, t_2) = e^{-\lambda|t_1 - t_2|} \quad (7)$$

ただし、 λ はパラメータであり、後で述べる実験では λ の値は0.5としている。

以上により式 (1), (6), (7) から関連度 $Score(i, t)$ が計算される。

また、元画像 i_{ori} に対して、 $Pairs$ に含まれていない $\langle i_{ori}, t \rangle$ の関連度についても前章で述べたとおり式 (2), (3) で計算する。これらの操作で元画像 i_{ori} と全ての時間 t の対に関して関連度 $Score(i_{ori}, t)$ が得られるので、この関連度を用いて元画像 i_{ori} の撮影時間を推定する。

4.4 撮影時間の推定

本節では前節で得られた関連度 $Score(i_{ori}, t)$ を用いて元画像 i_{ori} の撮影時間の推定を行う。

前章では t_{min} を求めるに当たり、画像や類似画像が含まれるすべてのウェブページの作成日時が分かっているものとした。しかし実装にあたり、現実には与えられたウェブページの作成日時がわかることは多くない。よって実装では与えられたウェブページのうち、実際にわかっている作成日時だけを利用し、得られた作成日時のうち最小のものを t_{min} とすることにする。ウェブページの作成日時の抽出には、本研究ではウェブページのメタデータ用いるのではなく、ウェブページの URL に含まれる時間を正規表現を用いて抽出する。例えば

- <http://www.xxxxxxxx.com/2003/01/yyyyyyyy.html>
- <http://www.zzzzz.com/2003-01-02abcde.html>
- <http://www.ddddddd.com/aaa/20030102abcde.html>

のような URL からは時間 $t = 200301$ が抽出される。こうして対象のウェブページ全てから得られた時間を用いて t_{min} を決定する。そして今回は $t > t_{min}$ なる時間 t については3章で述べたように t が撮影時間にはなりえないとして撮影時間の推定の際には無視する。

5. 画像検索への適用

本節では4章で実装したシステムを利用して時間を考慮した画像検索が実現できることを簡単に述べる。

近年、技術の発達に伴い様々な画像検索エンジンが登場している。従来の画像検索エンジンは主に入力したクエリと画像を含むページ上のテキストとを比較して画像を検索するテキストベースのものが主流だった。最近ではクエリが画像であり、その画像と特徴の似た画像を検索するといったようなコンテンツベースの画像検索エンジンも登場しつつある。しかし、現在ではまだ、画像に関する時間を考慮した画像検索エンジンは登場していない。画像に関する時間を考慮した画像検索とは例えば次のような2つの検索が考えられる。

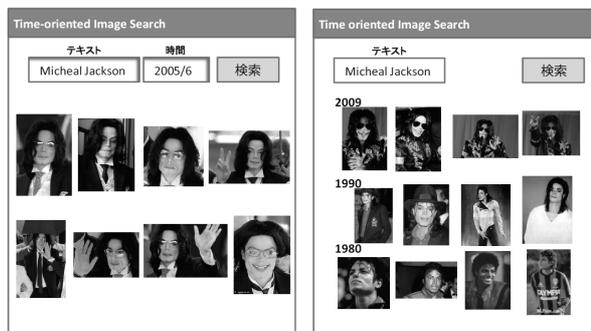


図7 画像に関する時間を考慮した画像検索

1つはあるクエリについて欲しい時期を指定ができる画像検索である。図7の左図のように一般的な画像検索のようなクエリフォームの他に、時間を入力できるクエリフォームを備えたインターフェースが考えられる。図7の左図では実際にクエリ”Michael Jackson”に対し、時間のクエリ”2005/6”も入力し、2005年6月のMichael Jackson画像が得られている。

もう1つは検索結果を時間ごとにソートやクラスタリングできる画像検索である。図7の右図のように一般的な画像検索のようなインターフェースだが、画像を検索すると検索結果を時間ごとにクラスタリングし、代表的な画像を出力する。図7の右図では実際にクエリ”Michael Jackson”に対し、2009年、1990年、1980年の画像が得られている。

そこで画像の撮影時間を推定する本システムが実現すれば、画像の撮影時間を考慮した画像検索も実現することができるようになると思われる。簡単に説明するとあるクエリで得られた画像1つ1つに対し、本システムで撮影時間を推定し、その時間を用いて時間を考慮した画像検索が実現できる。しかし、現実的には画像検索で返ってくる画像の数は膨大で、中にはクエリと関係のない画像がたくさん含まれており、結果としてあまり精度が悪くなってしまおうと考えられる。この問題を解決するには、クエリに対する代表的な画像、つまり得られた画像のうちクエリと関係があり、なおかつ検索結果に似た画像がたくさん含まれるような画像だけを抽出すればよい。これを実現するための方法としては VisualRank [5] が考えられる。VisualRank を用いてクエリに対して代表的な画像だけを集め、その画像の集合に対して本システムを適用することでより精度の高い画像の撮影時間を考慮した画像検索が実現されることが期待される。

6. 実験

本章では提案手法の有効性を確かめるために提案手法だけでなく他の3つのベースライン手法も用いて比較実験を行った。

6.1 実験設定

今回の実験では4章で述べたアプローチにしたがって撮影時間の推定を行った。本実験ではあらかじめ撮影時間のわかっている8つの画像をテストセットとして用意した。アルゴリズム実行に必要な画像収集には TinEye を使い、検索結果の上位100件を撮影時間推定に用いた。以下に3つのテストセット画

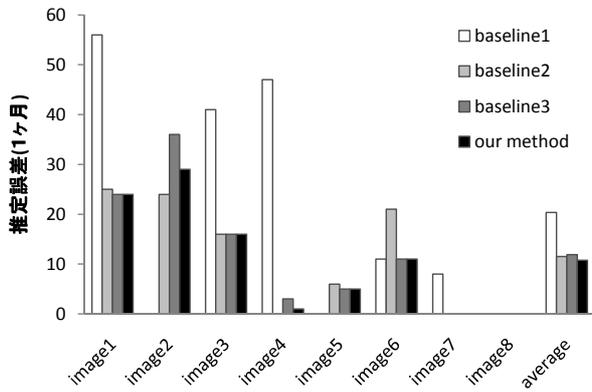


図 8 撮影時間の推定精度の比較 . 0 に近いほど撮影時間の推定誤差が小さいことを意味する .

像についてその画像の説明と URL を示す .

- 画像 1 「自由の女神」
 - ページ URL: <http://www.latestworldbreakingnews.com/statue-liberty.html>
 - 画像 URL: <http://www.latestworldbreakingnews.com/wp-content/uploads/2010/12/94-Statue-of-Liberty.jpg>
- 画像 4 「マイケルジャクソン」
 - ページ URL: <http://english.cctv.com/20090626/101063.shtml>
 - 画像 URL: http://english.cctv.com/20090626/images/1245975228136_1245975228136_r.jpg
- 画像 8 「ニューオーリンズ」
 - ページ URL: <http://www.pre-office.com/amphibious-adaptation/>
 - 画像 URL: http://www.pre-office.com/storage/456px-KatrinaNewOrleansFlooded.jpg?_SQUARESPACE_CACHEVERSION=1243481556940

3章の式(5)で述べたとおり、画像からの距離が一定以下の時間を画像と関連する時間として抽出するためのパラメータ α は 0.15 とし、式(7)で述べたとおり時間類似度の計算において使用されるパラメータ λ は 0.5 とした。また、本実験では画像と関連の強い時間の集合を抽出するために、対象区間を 2000 年 1 月から 2011 年 12 月までとし、その区間での関連度の偏差値を計算し、偏差値が 50 以上の時間を画像と関連の強い時間として利用した。

6.2 撮影時間の推定の精度の判定

本節では提案手法および 3 つのベースライン手法を用いて撮影時間がわかっている 8 つの画像に対して撮影時間の推定を行い、その精度を比較した。

まず本実験で用いた 4 つの手法について説明する。提案手法は今まで述べてきた手法である。ベースライン手法 1 は提案手法と同様に関連度を計算し、関連度の最も高い時間を画像の撮影時間とする手法である。この手法を選んだのは、関連度が一番高い時間がそのまま撮影時間となるかどうかを調べるためである。ベースライン手法 2 はウェブページのテキストから時間

	推定結果	誤差
ベースライン 1	2009 年 6 月	48ヶ月
ベースライン 2	2005 年 6 月	0
ベースライン 3	2005 年 3 月	3ヶ月
提案手法	2005 年 5 月	1ヶ月

表 1 画像 4 の撮影時間の推定結果 . 画像 4 の実際の撮影時刻は 2005 年 6 月である .

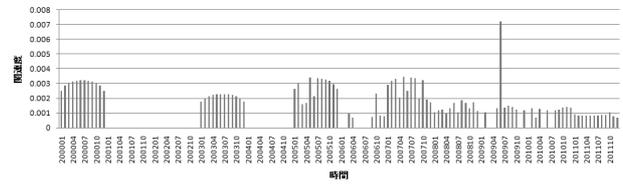


図 9 提案手法を画像 4 に適応した結果

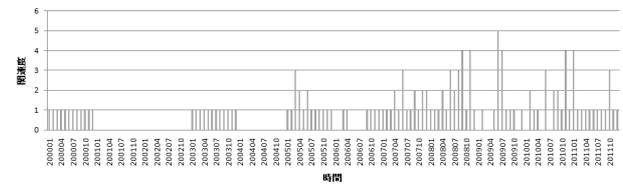


図 10 ベースライン手法 3 を画像 4 に適応した結果

を抽出することは行わず、ただページの URL から得られた時間により決定する t_{min} をそのまま画像の撮影時間とする手法である。この手法を選んだのは、3章で述べたように、最終的に撮影時間が t_{min} に近くなるという仮定が正しければ、 t_{min} 自身も撮影時間に近くて良い結果になりうる可能性があるためである。ベースライン手法 3 は提案手法のような 2 部グラフを用いた再帰的な計算を行わず、 $Score(i_{ori}, t)$ は単純に時間 t が抽出されたページ数とする手法である。この手法を選んだのは、提案手法のように 2 部グラフを用いた再帰的な計算を行わなくても、画像から一定の距離にある時間を使うだけで効果があるかどうかを確かめるためである。以上 4 つの手法に対して 8 つの画像の撮影時間の推定を行った。

推定された撮影時間と実際の撮影時間の誤差について提案手法とベースライン手法を比較した。その結果を図 8 に示す。図 8 によると、撮影時間の推定誤差は画像によって大きく異なるが、平均的には提案手法が最も推定誤差が小さかった。すなわち、3 つのベースライン手法よりも提案手法の方が精度良く撮影時間を推定することができたと言える。

次に 8 つの画像のうち、マイケルジャクソンの画像である画像 4 の実験結果の詳細を述べる。画像 4 の実際の撮影時間は 2005 年 6 月である。提案手法とベースライン手法 3 を適用した結果の関連度のグラフをそれぞれ図 9、図 10 に示す。画像 4 に関しては、ベースライン手法 1 は推定誤差が大きかったが、それ以外は推定誤差が小さかった。ベースライン手法 1 の誤差が大きかった理由は、図 9 からわかるように実際の撮影時間とは異なる時間が最も関連度が高かったからである。この最も関連度の高かった 2009 年 6 月というのはマイケルジャクソンが死亡した月であり、この死亡のニュースに画像 4 の類似画像

が多く使われていたことにより関連度が高くなっていた。他の手法が推定誤差が少なかったのは、今回は t_{min} の値が実際の撮影時間と一致していたからであると考えられる。

7. 考 察

前章で8つの画像について、それぞれ提案手法と3つのベースライン手法で撮影時間の推定を行った。実験結果を見る限りでは、誤差の平均だけ考えるとベースライン手法1が誤差が大きく、その他の手法がおおよそ誤差が近かったがその中でも提案手法が最も誤差が少なかった。しかし、その実験結果と画像が含まれていたウェブページの中身を比較してみると、例えば画像の撮影時間の記述がウェブページ上にあるにも関わらず距離が遠いため抽出されなかったり、ブログなどのタイムスタンプがあるにも関わらず、URLにその時間が書かれていないため、そのページの作成時間が抽出されないといった問題などが多々見られた。

本研究での提案アルゴリズムは、ある画像に関する類似画像が何らかの方法で入手できたり、画像に関連する時間だけをウェブページから抽出できたり、全ての画像を含むページに対してページの作成日時が分かっているという理想的な状況を考えている。しかし、現実的にはこのような状況を実現することは困難であり、何らかの代替案で理想的な状況に近づけようとしているのが現状である。本研究では画像の収集には類似画像検索エンジン TinEye を用い、画像と関連のある時間は画像との距離が近いという仮定に基づいて画像と距離の近い時間を抽出したり、ウェブページの作成日時がわからないのでウェブページの URL に含まれる時間表現で代替していた。その結果、ウェブ上にある全ての類似画像を手に入れられなかったり、画像とは関連のない時間が抽出されたり逆に関連のある時間がなかなか抽出されなかったり、ウェブページの作成時間がうまく抽出できないといったことが起こりうる。したがってこのような問題を解決するのが今後の課題だと考えられる。

8. おわりに

本研究では、ウェブページに掲載された画像の撮影時間を推定する手法を提案した。提案したアルゴリズムでは、まず与えられた画像を含むウェブページ中に含まれる時間のうち、画像と関連のあるものを抽出して画像と時間の対を作成した。そしてその画像の類似画像を含むウェブページに対しても同様の操作を行い、画像と時間の対を作成した。ここで得られた画像と時間の対に対し、画像と時間の関連の強さを関連度として与えた。ある画像と時間の関連の強さは、似た画像と時間の対が多く存在すれば関連度はあがるという仮定のもと、関連度を再帰的な式で定義し計算した。そうして得られた画像との関連度が高い時間から撮影時間を抽出するという手法を取った。

8つの画像を用いた実験の結果、提案手法は与えられた8つの画像に対する推定誤差の平均がベースライン手法と比較して精度が良かった。しかし、 t_{min} の値がうまく取れなかったり、関連度のピークが分散した場合はうまくいかないこともわかった。

今後の課題としてはまず画像と関連のある時間の抽出方法を改善することが挙げられる。そして類似画像の収集方法も改善することが挙げられる。さらにウェブページの作成時間を正確に抽出する方法の改善も挙げられる。これらの手法が改善されれば、画像の撮影時間がより正確に推定できるようになると期待される。

謝辞 本研究の一部は、京都大学 GCOE プログラム「知識循環社会のための情報学教育研究拠点」、および、NICT 委託研究「電気通信サービスにおける情報信憑性検証技術に関する研究開発」(研究代表者: 田中克己)、マイクロソフト産学連携研究機構 CORE 連携研究プロジェクト「Mining and Searching Web for Future-related Information」、文部科学省科学研究費補助金若手研究(B)「Towards time-focused Web Search and Mining」(研究代表者: Adam Jatowt, 番号 22700096) によるものです。ここに記して謝意を表します。

文 献

- [1] Dong A, Chang Y, Zheng Z, Mishne G, Bai J, Zhang R, Buchner K, Liao C, Diaz F, “Towards recency ranking in web search”, Proceedings of the third ACM international conference on Web search and data mining (WSDM 2010), pp.11–20, 2010
- [2] Arikani I, Bedathur S, Berberich K, “Time Will Tell: Leveraging Temporal Expressions in IR”, Proceedings of the Second International Conference on Web Search and Web Data Mining (WSDM 2009), 2009.
- [3] Wang XJ, Zhang L, Li X, Ma WY, “Annotating Images by Mining Image Search Results”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.30, no.11, pp.1919–1932, 2008.
- [4] 山根 遥香, 豊田 正史 “画像類似度の有向グラフを用いた CGM 上の画像時系列ランキング手法”, 情報処理学会創立 50 年 (第 72 回) 記念全国大会, 2010.
- [5] Jing Y, Baluja S, “VisualRank: Applying PageRank to Large-Scale Image Search”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.30, pp.1877–1890, 2009.
- [6] Dakka W, Gravano L, Ipeirotis P, “Answering general time sensitive queries”, Proceeding of the 17th ACM conference on Information and knowledge management (CIKM 2008), pp.1437–1438, 2008.
- [7] Zhang R, Chang Y, Zheng Z, Metzler D, Nie J, “Search result re-ranking by feedback control adjustment for time-sensitive query”, Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp.165–168, 2009.
- [8] 山本 祐輔, 田中 克己 “Web 上の画像・テキスト対の信憑性分析”, DEIM Forum 2009, 2009.
- [9] 山本 祐輔, 田中 克己 “データ対間のサポート関係分析に基づく Web 情報の信憑性評価”, 情報処理学会論文誌: データベース, Vol.3, No.2(TOD46), pp.61–79, 2010.
- [10] Lowe D, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, Vol.60, No.2, pp.91–110, 2004.
- [11] <http://www.jeita.or.jp/cgi-bin/standard/list.cgi?cateid=1&ubcateid=4>