

NAIST-IS-MT0351017

## 修士論文

# 撮影位置・姿勢情報に基づく 写真への索引付加システムの構築

岩崎 季世子

2005年 3月 10日

奈良先端科学技術大学院大学  
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に  
修士(工学) 授与の要件として提出した修士論文である。

岩崎 季世子

審査委員：

横矢 直和 教授

千原 國宏 教授

山澤 一誠 助教授

# 撮影位置・姿勢情報に基づく 写真への索引付加システムの構築\*

岩崎 季世子

## 内容梗概

デジタルカメラをはじめとした撮影機器の普及により、個人が写真を撮影する機会は日常化してきている。しかし、写真を簡便に管理する方法は少なく、膨大な量のデータが未整理のままであることが多い。写真を管理する方法の1つとして、写真の内容を説明する語を付加しておくことが考えられるが、これを人手で行うことは非常に手間がかかる。一方で、完全に自動化されたシステムによってユーザの意図した語を写真に付加することもまた、困難である。そこで本研究では、写真撮影時のカメラの位置・姿勢情報から被写体の位置を推定し、推定位置の地名や施設名の候補をユーザに提示することで、半自動的な索引付加作業を実現するシステムを提案する。索引候補語は、あらかじめ用意された地図データベースから対応する位置の地名や施設名を取得し、データベース内に適当な索引語が含まれていない場合には、web 検索を用いた関連語抽出処理によって新たな候補語を取得しユーザに提示する。ユーザに選択された索引語は、その位置に適当な語であると見なし、これを地図データベースへのフィードバックとして使い、データベースの更新を行う。これにより、提示される候補語が変化し、ユーザの選択作業は効率化される。索引付加作業は撮影と並行して行うことができ、撮影終了後すぐに効率的な閲覧が可能となる。提案手法に基づいたプロトタイプシステムを構築し、実際に撮影した写真に索引付加を行った実験により有用性の確認を行った。

---

\* 奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT0351017, 2005年3月10日.

キーワード

撮影位置・姿勢，写真への索引付加，地図データベース，web 検索，関連語  
抽出

# An Indexing System for Photos Based on Shooting Position and Orientation\*

Kiyoko Iwasaki

## Abstract

With the spread of digital cameras, shooting photos has been becoming an everyday affair. However, there are few methods or systems to manage photos simply, and a huge amount of photo data remains unorganized. Although it is possible to add appropriate words explaining the contents of the photo as one of the methods to manage photos, it requires much time and effort to input such indexes manually. It is also difficult to add indexes intended by a user automatically.

In this paper, we propose a semi-automatic photo indexing system that enables users to generate indexes simply and browse a photo library efficiently. In the proposed framework, place or facility names as index candidates corresponding to subject positions estimated from shooting position and orientation are acquired from a map database prepared in advance. If there are no appropriate candidates in the database, new candidates are obtained by relevant word extraction using web retrieval. The map database can present more appropriate candidates by updating the database using the index selected by users as feedback. The indexing can be done in parallel with shooting, and efficient browsing is attained immediately after a shooting. We have implemented an indexing system prototype based

---

\* Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0351017, March 10, 2005.

on the proposed method, and have confirmed the feasibility of the system with some experiments.

**Keywords:**

shooting position and orientation, photo indexing, map database, web retrieval, relevant word extraction

# 目次

1. はじめに	1
2. デジタル写真の管理に関する従来研究と本研究の位置付け	4
2.1 画像検索における従来手法	4
2.1.1 画像特徴量による検索	4
2.1.2 画像のメタデータによる検索	5
2.2 写真の効率的管理を目的とした従来手法およびシステム	8
2.2.1 メタデータ入力効率化	8
2.2.2 位置情報に基づいた管理	9
2.3 本研究の位置付けと方針	12
3. 撮影位置・姿勢情報に基づく写真への索引付加	14
3.1 索引付加の概要	14
3.2 撮影位置・姿勢情報付き写真の取得	16
3.3 被写体位置の推定	16
3.4 地図データベースからの索引候補語の取得	17
3.4.1 地図データベースの構成	17
3.4.2 位置情報と選択頻度に基づく取得	18
3.5 web 検索を用いた関連語の抽出	18
3.6 ユーザ選択をフィードバックとした地図データベースの更新	22
3.7 索引付加写真閲覧のための写真データベースの構築	24
4. 実験	26
4.1 プロトタイプシステム	26
4.2 撮影した写真への索引付加実験	29
4.2.1 予備実験	29
4.2.2 実証実験	34
5. まとめ	40

謝辭	42
參考文獻	43



## 表 目 次

1	地図データベースに含まれるデータの例 . . . . .	18
2	プロトタイプシステムに用いた機器類 . . . . .	27
3	地図データベースの更新 . . . . .	32
4	抽出された「薬師寺」の関連語(上位20語) . . . . .	32
5	索引付加作業における索引語の提示順位：薬師寺 . . . . .	36
6	索引付加作業における索引語の提示順位：法隆寺 . . . . .	37

## 1. はじめに

デジタルカメラをはじめとした撮影機器の普及により，個人が写真を撮影する機会は日常化してきている．しかし，写真を簡便に管理する方法は少なく，膨大な量のデータが未整理のままであることが多い．

画像検索については，長年に亘って研究が行われており，さまざまな手法やシステムが提案されている [1,2]．また，web 上で画像の検索を行う検索エンジン [3,4] や，図 1 に示すような画像（写真）閲覧用のソフトウェア [5-7] も一般に普及している．画像の検索手法は，画像に含まれる色や形，テクスチャといった画像特徴量を利用するものと，画像に付加された注釈や撮影場所，撮影時刻などのテキスト情報によるメタデータを利用するものに大別できる．前者には，ユーザが描いたスケッチや与えた画像に類似した特徴量をもつ画像を検索するものなどがある [8-11]．後者は，ユーザが与えたキーワードで，画像に付加されたテキスト情報を検索し，該当する画像を取得するものである [12,13]．

写真についても，これらの手法を適用した検索や閲覧を行うことができるが，本研究では個人が撮影した写真を対象としており，これに適切な手法を検討する．個人が撮影する場合の被写体としては，家族や友人などの人物，旅行先などの風景や施設等が一般に想定される．また，撮影した写真を閲覧・検索する状況としては，撮影した本人やその身近な人物が，特定の日時・人物・出来事・場所などに基づいて行うことが想定される．

撮影した日時については，写真のメタデータとして付加されている場合がほとんどであり，日時に基づいた閲覧・検索は一般に行われている．人物については，個人の写真であれば，家族や友人などある程度限られた人物のみが撮影されているものと考えられる．そこで，顔認識手法を用いたシステム [14,15] や写真の人物にラベル付けを行うインタフェース [16] が提案され，人物に基づいた閲覧・検索を実現している．出来事については，個人のスケジュール管理を行うソフトウェアを参照して，撮影日時に対応する予定から写真の場面を類推する手法 [17,18] が提案されているが，適用できる状況は限定されている．場所については，地図を用いた GUI から写真の撮影位置をユーザが手入力するシステム [19] や GPS を利用して位置情報を取得するシステム [17-19] などが提案されている．このよう

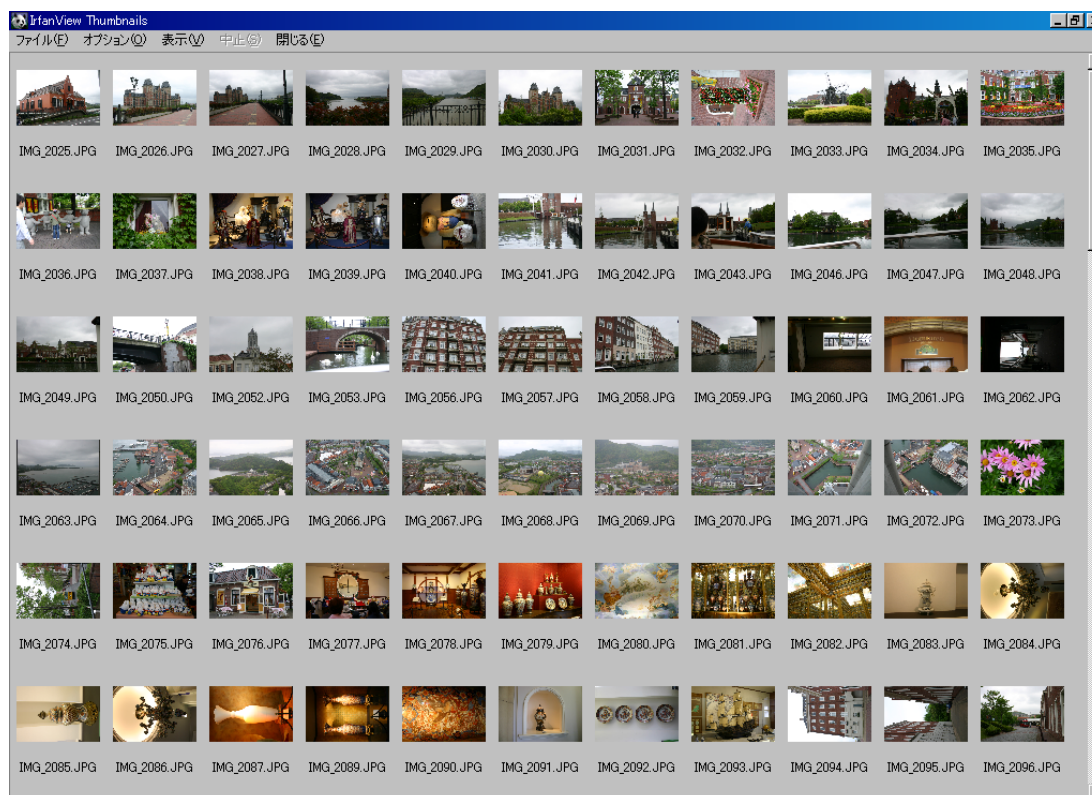


図 1 画像閲覧ソフトウェア IrfanView [7]

なシステムでは位置情報が数値として得られるが、閲覧や検索におけるユーザの入力は、一般に地名や施設名であると考えられる。したがって、得られた位置情報を閲覧や検索に利用するためには、位置情報を数値から地名や施設名といったテキスト情報に変換する必要がある。この変換には地図データを利用すると考えられるが、対応するデータがない場合やユーザの意図とは異なるデータに対応付けられる場合など、適切な変換が行われないことがあり、場所に基づいた閲覧や検索を行うには課題がある。

また、日時・人物・出来事・場所のいずれについても、ユーザが与えた写真と画像特徴量の類似した写真を取得する手法を適用し、閲覧・検索を行った場合、画像間のマッチングに用いる特徴量によって結果として得られる写真が変化する。ユー

ザが与えた写真の意図を反映させる特徴量を決定することは容易ではなく，結果的に，ユーザが意図した写真を得られるまでに時間や労力を要する場合がある．

近年，写真に対する撮影位置情報の付加が一般的になりつつある．例えば，市販されているデジタルカメラにおけるメタデータの標準規格である Exif [20] は，カメラパラメータや GPS で取得した位置情報，写真の内容に関する記述などを JPEG 形式の画像ファイル自体に含めることができる．また，GPS とカメラの付加された携帯電話により位置情報付きの写真を取得することができる．そこで本研究では，写真撮影時のカメラの位置・姿勢情報に着目し，この情報を用いて被写体位置を推定し，推定位置の地名や施設名の候補をユーザに提示することで，半自動的な索引付加作業を実現するシステムを提案する．索引候補語は，あらかじめ用意された地図データベースから対応する位置の地名や施設名を取得し，データベース内に適当な索引語が含まれていない場合には，web 検索を用いた関連語抽出処理によって新たな候補語を取得しユーザに提示する．ユーザに選択された索引語は，その位置に適当な語であると見なし，これを地図データベースへのフィードバックとして使い，データベースの更新を行う．これにより，提示される候補語が変化し，ユーザの選択作業は効率化される．索引付加作業は撮影と並行して行うことができ，撮影終了後すぐに効率的な閲覧が可能となる．

以降，2 章ではデジタル写真の管理に関する従来研究と本研究の位置付けについて，3 章では撮影位置・姿勢情報に基づく写真への索引付加手法について述べる．4 章では提案するシステムのプロトタイプを用いた実験について述べ，最後に，5 章で本論文をまとめ，今後の展望について述べる．

## 2. デジタル写真の管理に関する従来研究と本研究の位置付け

本章では、本研究に関連する従来研究および本研究の位置付けについて述べる。2.1節では、画像検索の分野における従来研究について、2.2節では、デジタル写真の効率的な管理に関する従来研究とその課題について述べる。最後に2.3節で、本研究の位置付けと方針を述べる。

### 2.1 画像検索における従来手法

画像検索の分野では、長年に亘って研究 [1, 2] が行われており、ユーザが複数の画像の中から欲しい画像を取得する方法として、さまざまな手法やシステムが提案されている。画像検索システムの構築においては、ユーザによる検索クエリの与え方、クエリから得られる何らかの特徴量と画像の特徴量のマッチング方法、ユーザに対する検索結果の提示方法、画像を保持するデータベースの構造などが検討されている。

また、画像検索システムは、画像特徴量を利用するもの [8–11] と画像に付加されたテキスト情報などによるメタデータを利用するもの [12, 13] に大別できる。前者には、ユーザが与えた画像、スケッチなどに含まれる画像特徴量やユーザが直接指定した画像特徴量（カラーヒストグラムなど）を用いて類似画像を取得する方法 [8, 10] などがある。また、後者は、ユーザがキーワードを与えて、このキーワードを付加された画像を取得する方法であり、web上の画像検索エンジンやいくつかのシステムが実現されている [12]。以下、2.1.1では、画像特徴量を利用した検索手法について、2.1.2では、画像に付加されたメタデータを利用した検索手法について、それぞれの特徴を述べる。

#### 2.1.1 画像特徴量による検索

画像特徴量を用いた代表的な検索手法として、ユーザが選択した画像や描いたスケッチから色やテクスチャ、レイアウトなどに関する画像特徴量を抽出し、類

似した特徴量をもつ画像を取得する方法が挙げられる [8–11]。図2は、Brunelliらにより開発された画像検索システム Compass [8] であり、ユーザが選択した画像に基づき、類似した画像特徴量をもつ画像をユーザへ提示する。また、図3に示す Dasらによる Focus [9] は、ユーザが画像の一部を選択し、これをテンプレートとして類似画像を検索するシステムである。ユーザが描いたスケッチを利用したシステムとしては、図4に示す Sciascioらの DrawSearch [10] が挙げられる。このような手法では、画像のどの特徴量を用いて画像間のマッチングを行うかで結果が変わってくるため、ユーザの意図した画像が得られない場合がある。このため、ユーザと対話的に検索作業を行い、その過程における入力をフィードバックとして、ユーザの意図した画像を取得する手法などが提案されている [21]。この手法を用いて、図5に示すような、ユーザがキーワードを入力し、キーワードに合った内容の画像を画像特徴量から取得する手法も提案されている [12]。このような手法では、ユーザが入力として与えた画像やスケッチにおいて意図した特徴と、システムが使用する画像特徴量を対応付ける方法が課題となっているが、ユーザが与えた画像の意図を反映させる特徴量を決定することは容易ではなく、結果的に、ユーザが意図した写真を得られるまでに時間や労力を要する場合がある。

### 2.1.2 画像のメタデータによる検索

画像特徴量を用いた検索では、画像の意味的な内容について考慮されることは少ない。一方、画像のメタデータを用いた検索では、画像に対してその意味的な内容をキーワード等として付加しておくことで、内容についての検索を行うことができる。

一般的な web 検索では、ユーザがキーワードを与えることでテキスト検索が行われ、関連する web サイトが提示される。画像についても、あらかじめテキストのメタデータを付加しておくことや、何らかの方法で画像に関連するテキスト情報を取得することで同様のテキスト検索を行い、画像を取得するシステムが運用されている。テキスト検索については、長年に亘って研究の行われている分野であるため、既存の手法を用いることができる。しかし、画像のメタデータに対してテキスト検索を行うためには、あらかじめテキストによる情報を付加しておく



图 2 Compass [8]

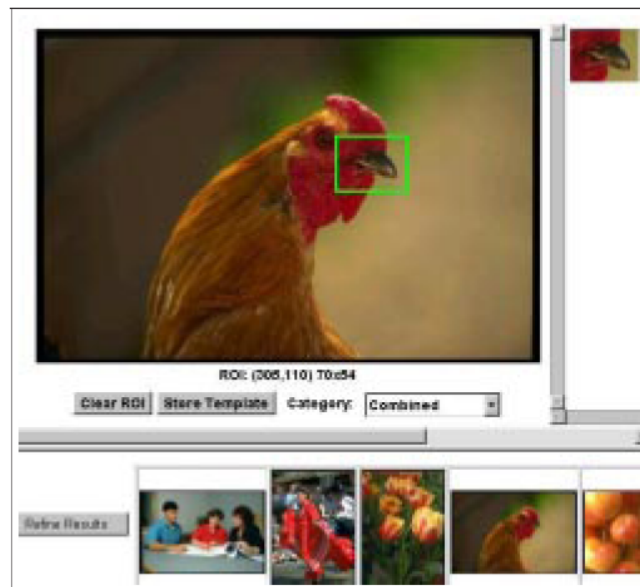


图 3 Focus [9]

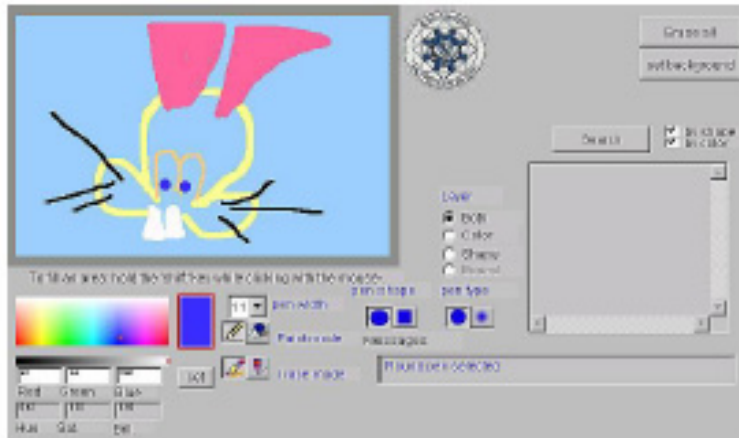


图 4 DrawSearch [10]

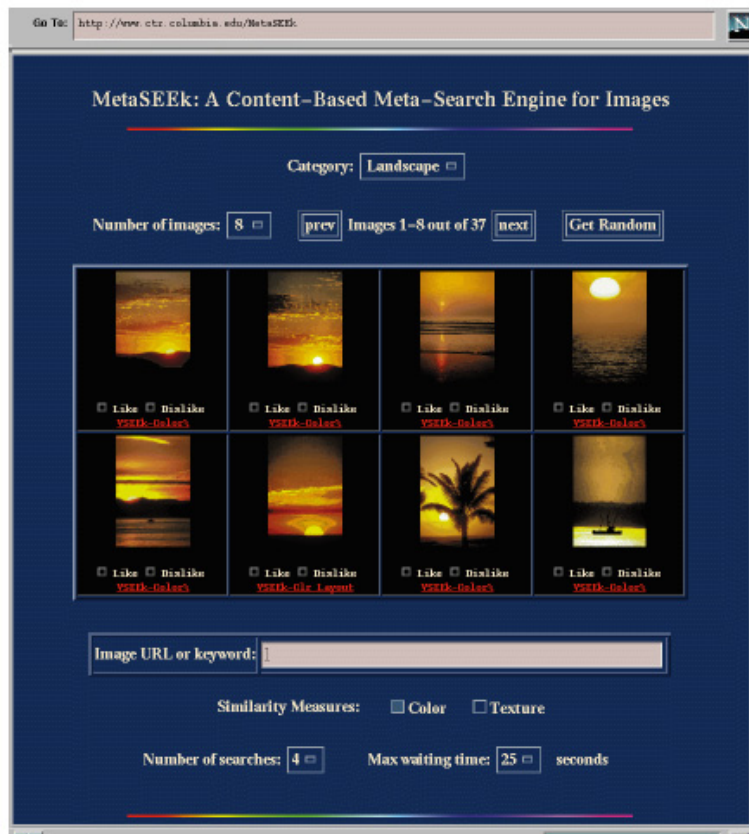


图 5 MetaSEEK [12]



必要がある。画像の作成者や作成日時等の情報については、画像を作成または取得する際に自動的に付加されていることが多いものの、画像の内容についての記述はユーザによる手入力を要する。このため、メタデータ入力のためのユーザインタフェース [16] や、図鑑や web ページ内の画像についてその周辺の文からキーワードを抽出し、画像のメタデータとする手法 [13] などが提案されている。いずれの手法もメタデータ入力にかかるコストの低減を目指しているが、前者については、入力するメタデータの内容が想定するシチュエーションに特化されている場合が多いことや大量の画像に対して入力を行う場合について検討が必要であること、また、後者については、適用できる状況が限定されていることなどの問題が挙げられる。

## 2.2 写真の効率的な管理を目的とした従来手法およびシステム

2.1 節で述べた、画像特徴量を利用する手法と画像に付加されたテキスト情報などによるメタデータを利用する手法は、写真についても適用され、写真を効率的に管理するためのさまざまな手法およびシステムが提案されている [22-31]。Mills ら [24] は上記の 2 つの手法を評価し、個人の写真に関しては、画像特徴量を使った検索よりも、画像にあらかじめ付加した注釈を利用したテキスト検索の方が、ユーザの意図した画像を取得しやすいとしている。このような見地から、ユーザにとって負担の少ない形で写真に何らかのテキスト情報を付加し、効率的な管理を実現することを目的とした手法について研究が行われている [16, 23, 25, 32, 33]。

以下、2.2.1 でメタデータ入力の効率化を図った従来研究について、2.2.2 でメタデータの中でも特に撮影時の位置情報を利用して写真の管理を行った従来研究について述べ、その研究課題について検討する。

### 2.2.1 メタデータ入力の効率化

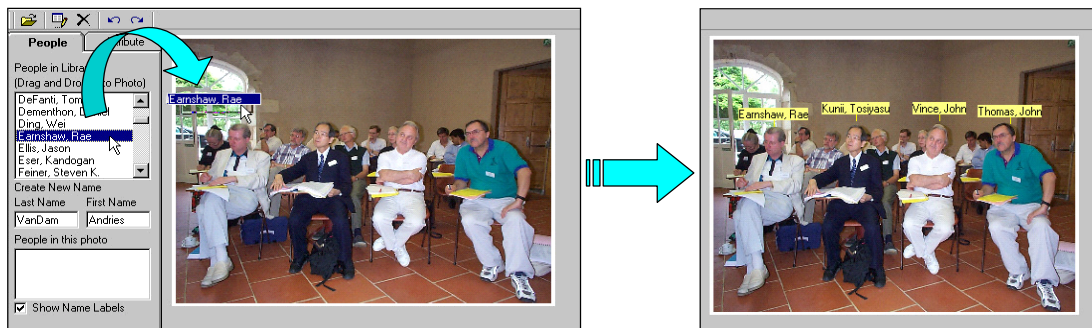
前述したように、写真に対してその内容や撮影日時・場所などのテキストによる情報をあらかじめ付加しておくことで、既存のテキスト検索手法を利用して画像を取得することができる。JEITA 標準の Exif format [20] は、デジタル写真

へのメタデータの記述について定めたもので、内容に関する記述や撮影時の焦点距離等のカメラパラメータ、GPSによって取得される位置情報等を画像ファイル自体に含めることができる。Exifは一般的な市販のデジタルカメラでサポートされており、これらのメタデータを利用した検索を行うこともできる。しかし、撮影された被写体など内容に関する記述については人手で行わなければならない、これは、多くの時間や労力を要するものである。このような情報を効率的に入力するため、いくつかのシステムが提案されている。

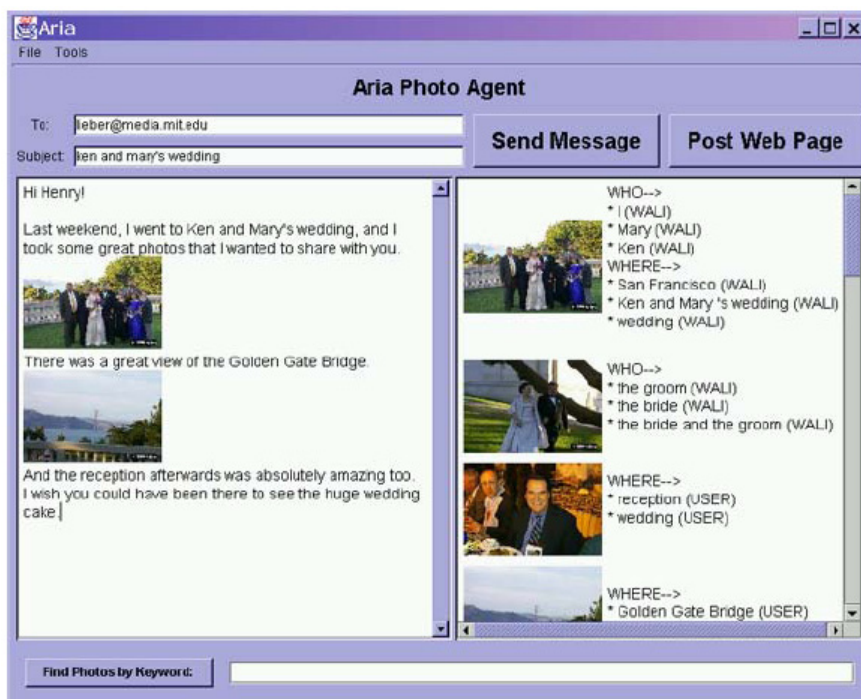
Shneidermanら [16, 34] は、図6に示すように、あらかじめデータベースに登録した人名等のラベルの一覧と、注釈付けを行う写真のサムネイル画像をユーザに提示し、注釈に適したラベルを画像上にマウスで drag-and-drop するシステムを提案している。また、自動的に注釈付けを行う方法として、Liebermanら [33] は、図7に示すように、撮影した写真をメールに添付して送るという行為を利用する手法を提案している。これは、写真を添付したメールの本文が写真を説明しているものとして、出現する単語を写真の注釈とするものである。写真は、メールに添付して送られるごとにその内容を説明する文章が増えることとなり、より適切な検索が可能となる。しかし、Shneidermanらが提案するシステムは、注釈として使用するラベルをあらかじめデータベースに登録しておく必要があること、Liebermanらの手法は、メールに写真を添付するという限定された状況を想定しており、これ以外の状況では利用できないこと、全自動で処理されることによる注釈語の誤りなどの問題が挙げられる。

### 2.2.2 位置情報に基づいた管理

近年、メタデータのうち、特に撮影時の位置情報を利用して写真を管理する手法について研究が行われている [17–19, 35, 36]。また、デジタルカメラのメタデータ標準規格である Exif での位置情報のサポートや、カメラおよび GPS 付きの携帯電話により位置情報付きの写真を取得できることなどから、今後、撮影位置情報をもつ写真の一般化が予想される。写真の撮影位置情報は、その写真の内容に大きく関わるものである。ユーザは、写真に付加された位置情報により、撮影された対象や状況を想起することができ、閲覧や検索の効率化が期待される。



☒ 6 Direct Annotation [16,34]



☒ 7 Aria [33]

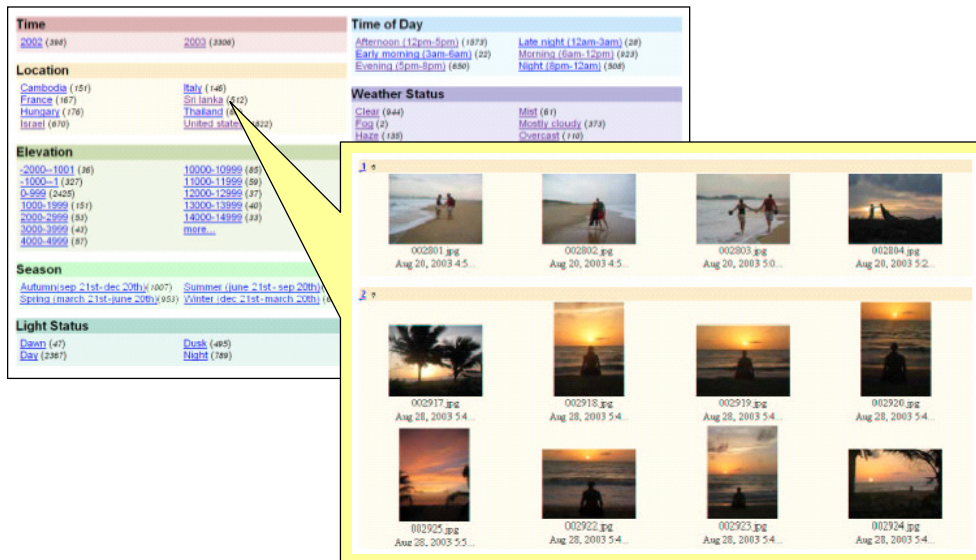


図 8 PhotoCompas [17,18]

また、単に撮影位置の緯度・経度といった数値的な情報を与えるのではなく、ユーザに対して対応地点の地名や施設名を提示したり、写真を地図上に配置して見せたりすることで、より効果が高まるものと考えられる。

Toyama ら [19] は、写真に対して位置情報タグを付加することを提案し、位置情報の取得方法やタグのデータ構造等を検討している。また、タグを利用し、地図上の撮影地点に写真を配置してユーザへ提示するシステムなどの応用例を示している。Naaman ら [17, 18, 36] は、撮影した写真を撮影時刻や位置等に基づいて階層構造のグループに分類する手法を提案している。また、ユーザがこの階層構造を辿って写真を閲覧するシステムとして、図 8 に示す PhotoCompas を構築している。

写真に付加された位置情報を用いて地図データベースを参照すると、その地点（もしくはその付近）の住所や施設名を得ることができる。これによりユーザは、撮影位置に基づいたキーワード検索や閲覧を行うことができる。しかし従来研究では、あらかじめ用意された地図データベースと写真の撮影位置をマッチングしているため、ユーザの意図した索引語が付加されない場合がある。この原因とし

ては、まず、地図データベースに撮影位置に対応するデータが収録されていないことが挙げられる。また、ユーザが閲覧や検索を行う場合には、住所ではなく施設や建物の名称の方が適していると考えられるが、例えば、観光施設などの地図データでは、その施設を代表する名称が対応する敷地内のある一点に与えられており、施設内にあるそれぞれの建物などのデータは収録されていない場合が多い。このため、地図データベースに該当する地名や施設名がない場合には人手による入力が必要となる。また、被写体の位置が撮影位置から離れている場合に、撮影位置に対応するデータが写真の内容に合わないことも原因として挙げられ、写真の被写体位置を考慮した手法の検討が必要であると考えられる。

### 2.3 本研究の位置付けと方針

本研究では、個人が撮影した写真を管理し、ユーザが効率的な閲覧・検索を行うことができることを目指す。そこで、従来から研究の行われている画像検索手法のうち、個人の写真について適しているとされる、メタデータを利用して検索を行う手法を用い、また、メタデータのうち、近年一般的になりつつある位置情報に着目する。これは、位置情報がその写真の内容に大きく関わるものであるため、ユーザが位置情報から撮影された対象や状況を想起し、閲覧や検索の効率化が図られるためである。

位置情報に基づいた閲覧・検索を行う手法においては、緯度・経度等の数値によって取得される情報から、ユーザが閲覧や検索を行うのに適した施設名等のテキスト情報へ適切な変換を行うことが必要である。また、撮影時に位置情報として取得できるのは撮影位置についての情報であるため、被写体位置を考慮し、より写真の内容に適した情報の付加について検討する必要がある。そこで本研究では、写真撮影時のカメラの位置・姿勢情報から被写体の位置を推定し、推定位置の地名や施設名の候補をユーザに提示することで、半自動的な索引付加作業を実現するシステムを提案する。図9に、システムの概要を示す。システムは、ユーザが操作するクライアント、そして写真と撮影時の位置・姿勢情報を処理して索引候補語をユーザへ提示するサーバからなる。サーバには、地図データベースが置かれ、これは、複数のユーザで共有するものである。まず、ユーザは、クライ

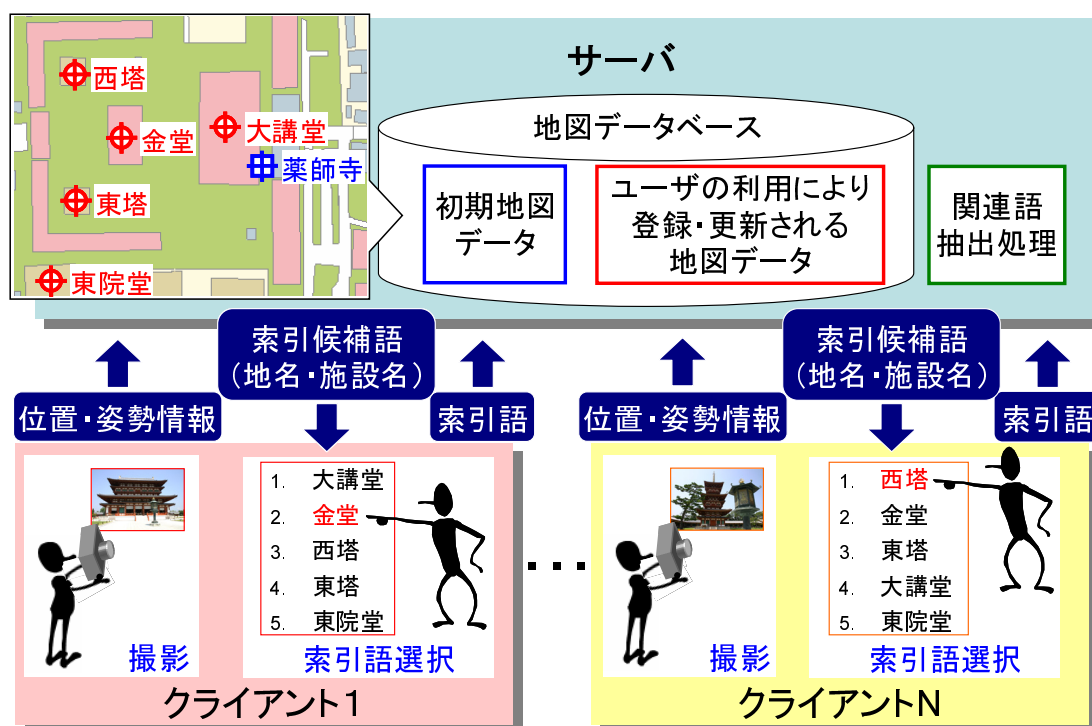


図 9 撮影位置・姿勢情報に基づく写真への索引付加の概要

アントから、写真の撮影位置・姿勢情報を送信する。サーバは、これを処理して写真の撮影位置・姿勢情報に基づく索引候補語を生成する。索引候補語としては、あらかじめ用意された地図データベースから対応する位置の地名や施設名を取得するが、データベース内に適当な索引語が含まれていない場合には、web 検索を用いた関連語抽出処理によって新たな候補語を取得する。生成された索引候補語はクライアントへ返され、ユーザに提示される。ユーザは、提示された索引候補語の中から写真の内容に合った語を選択し、これが写真の索引語となる。また、ユーザに選択された索引語は、その位置に適当な語であると見なし、これを地図データベースへのフィードバックとして、データベースの更新を行う。つまり、ユーザがシステムを利用することで地図データベースが更新され、次に同一の被写体を撮影した写真について索引付加作業を行う場合、ユーザへ提示される候補語は変化する。これにより、ユーザの索引付加作業の効率化を図る。

### 3. 撮影位置・姿勢情報に基づく写真への索引付加

本章では、写真の撮影位置・姿勢情報に基づく索引候補語をユーザに提示することで、半自動的な索引付加作業を実現する手法について述べる。まず、3.1節で、索引付加手法の概要について述べ、次にシステムの処理の流れに沿って、3.2節で撮影時のカメラ位置・姿勢情報の取得方法、3.3節でカメラ位置・姿勢情報からの被写体位置の推定方法、3.4、3.5節で被写体位置に対応する索引候補語の取得方法、3.6節でユーザが選択した索引語を用いて地図データベースを更新する方法について順に述べる。また、3.7節で、索引付加写真閲覧のための写真データベースの構築について述べる。

#### 3.1 索引付加の概要

図 10 に撮影位置・姿勢情報をもつ写真に索引付けを行う手法の概要を示す。

はじめに、ユーザは、GPS やジャイロセンサ、コンパス等のセンサとカメラを用いて、位置・姿勢情報付きの写真を取得する。撮影時に取得される撮影時刻、撮影地点の位置情報を写真の索引とする一方で、この情報を利用して索引候補語を取得する。写真の被写体について索引候補語を取得するため、撮影地点の位置・姿勢情報とカメラパラメータから被写体位置を推定する。次に、推定した被写体位置を用いて地図データベースを参照し、推定位置付近の地名や施設名を取得する。取得した地名や施設名は、写真の索引候補語としてユーザに提示され、ユーザは提示された候補語の中から被写体に適切な語を選択する。しかしこのとき、提示された候補語の中に写真に適した索引語が含まれていないということが想定される。これは、地図データベースがその地点を代表する名称である地名や施設名を含むもので、施設内の建物といったより詳細なレベルの名称を含んでいないためである。例えば、図 11 に示すように、地図データベースは、「薬師寺」という施設名のデータを保持しているものの、「薬師寺」内の建物である「金堂」、「東塔」、「西塔」といった施設名を保持していない。そこで、より詳細なレベルの名称をはじめとした、関連する名称を新たに索引候補語として取得するため、web 検索を用いた関連語抽出を行う。ユーザは、提示された索引候補語の中から付加し

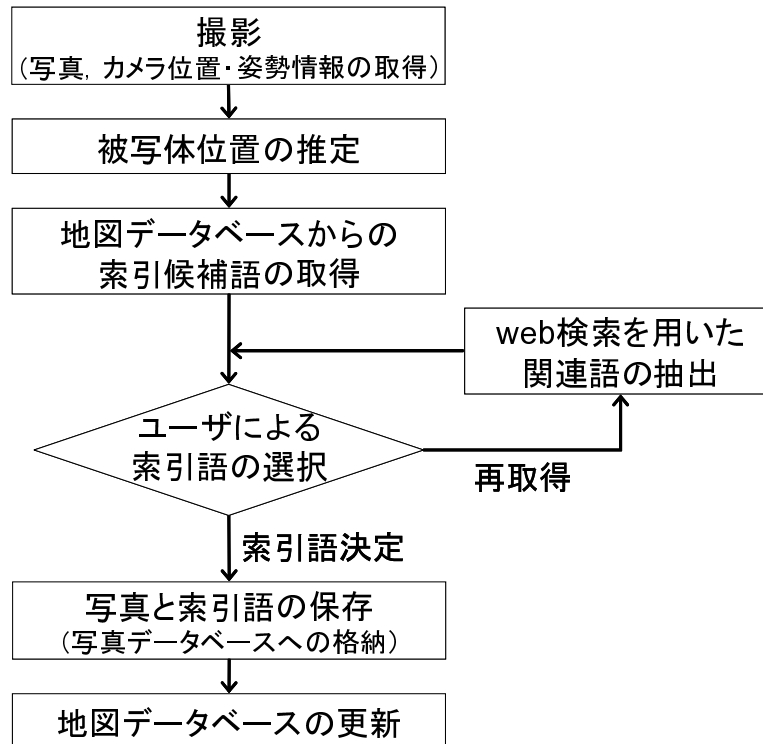


図 10 撮影位置情報付き写真の索引付加

たいと考える索引語に最も関連すると考えられる語を選択する。システムは、選択された語をキーワードとして web 検索を行い、関連語を抽出し、これを新たな索引候補語としてユーザに提示する。ユーザにより選択された索引語は、写真や撮影時刻・位置情報といったメタデータと共に写真データベースへ格納し、さらに、地図データベースへのフィードバックとしても利用する。選択された索引語を用いてデータベースの更新を行うことで、ユーザに提示される索引候補語やその提示順序を変化させ、ユーザによる作業の効率化を図る。また、写真とその索引語およびその他のメタデータの格納された写真データベースが構築されることにより、ユーザは、索引語やメタデータに基づく閲覧や検索を行うことができる。



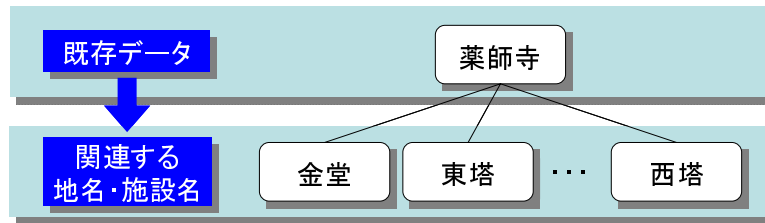


図 11 関連する地名・施設名の取得

### 3.2 撮影位置・姿勢情報付き写真の取得

提案システムを利用するユーザは、撮影位置・姿勢情報を取得することのできるセンサ付のカメラを用いて撮影を行う。センサ付カメラには、カメラに対して GPS やジャイロセンサ、コンパス等のセンサが取り付けられている。位置情報に関しては GPS を用いて、撮影位置の緯度・経度、標高を取得する。姿勢情報に関してはジャイロセンサおよびコンパスを用い、仰角、方位角を取得する。また、写真を記録した JPEG ファイルに含まれる Exif 情報から、カメラパラメータである焦点距離、レンズ F 値等を取得する。

### 3.3 被写体位置の推定

撮影時に取得した位置・姿勢情報から、撮影された被写体の位置を推定する。図 12 は、撮影位置と被写体位置の関係を示したものである。まず、WGS84 (GPS の基準座標系) 上にある撮影位置の緯度・経度・標高を原点とした地平座標系を考える。地平座標系は、地表面付近のある点を原点として、天頂方向、東方向、北方向にそれぞれ軸をとる。この座標系上で、コンパス・ジャイロ・Exif 情報から得られる方位角・仰角・被写体までの距離を用いて推定被写体位置を算出する。これを WGS84 上に変換した緯度・経度を用いて、3.4 で述べる地図データベースからの索引候補語の取得、および 3.6 で述べる地図データベースの更新を行う。

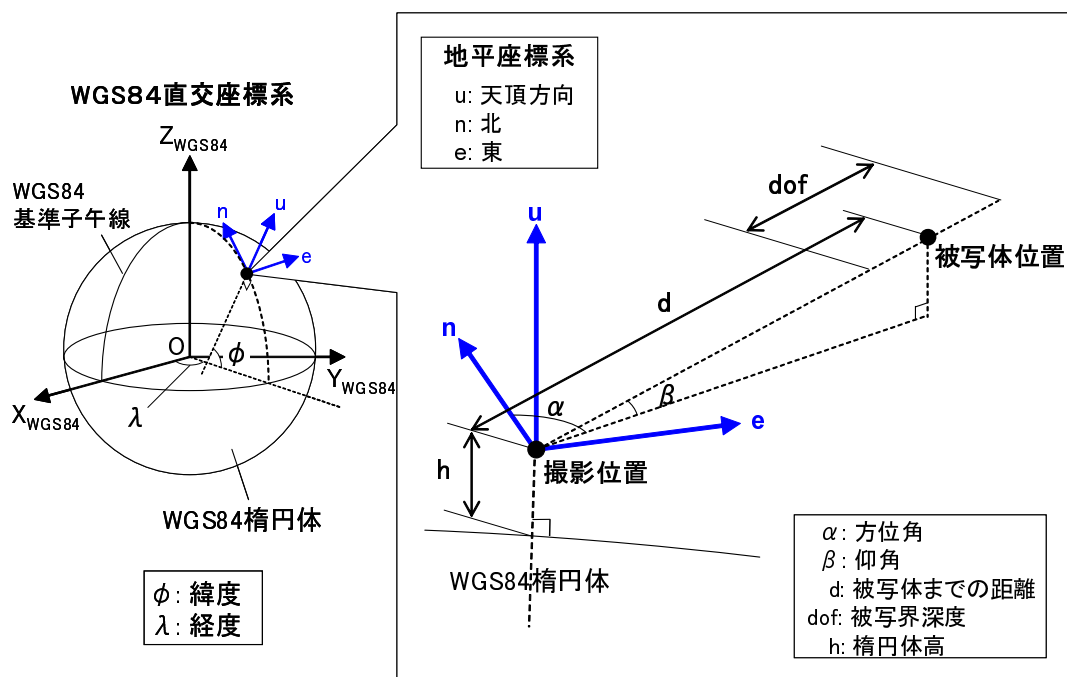


図 12 被写体位置の推定

### 3.4 地図データベースからの索引候補語の取得

推定した被写体位置を用いて地図データベースへの問い合わせを行い、写真に付加する索引の候補語を取得する。以下、3.4.1 では地図データベースの構成要素について、3.4.2 では地図データベースから候補語を取得する際の処理について述べる。

#### 3.4.1 地図データベースの構成

地図データベースの各レコードは、索引候補語となる地名や施設名といった名称、対応する位置の緯度・経度、ユーザに選択された回数である頻度から構成される。表 1 に例を示す。地図データベースは、初期状態では市販の地図ソフトに収録されたデータなどの既存のデータを格納したものであるが、索引付加システムをユーザが利用することで、新たなデータの登録や既存のデータの更新が行わ

れる．データの登録や更新は，ユーザが索引付加作業において選択した索引語に基づいて行われる．詳細については3.6節で述べる．

表 1 地図データベースに含まれるデータの例

name	N/S	latitude[deg]	E/W	longitude[deg]	frequency
薬師寺	N	34.668878	E	135.784313	0
東塔	N	34.668073	E	135.784335	5
唐招提寺	N	34.675775	E	135.784786	4
五重塔	N	34.614133	E	135.733928	13

### 3.4.2 位置情報と選択頻度に基づく取得

推定した被写体位置を用いて地図データベースへの問い合わせを行い，写真に付加する索引の候補語を取得する．図 13 に，地図データベースに登録されたデータから索引候補語を取得する処理のイメージを示す．まず，推定した被写体位置からの距離が一定範囲内の地名や施設名を近い順に取得する．取得した候補をユーザによる選択の有無で分類し，これまでにユーザに選択された履歴のある候補を上位に，選択された履歴のない候補を下位にし，索引候補語としてユーザに提示する．

## 3.5 web 検索を用いた関連語の抽出

提案するシステムでは，写真に対して撮影位置・姿勢情報に基づく索引語を付加することを目的とするが，地図データベースに該当する位置のデータがない場合や取得した索引候補語が写真に適切でない場合には，新たな候補語を何らかの手段で取得する必要がある．本研究ではその手段として，web 検索を用いた関連語の抽出を検討する．

web を利用して関連語を自動収集することを目的として，佐藤ら [37] は，与え

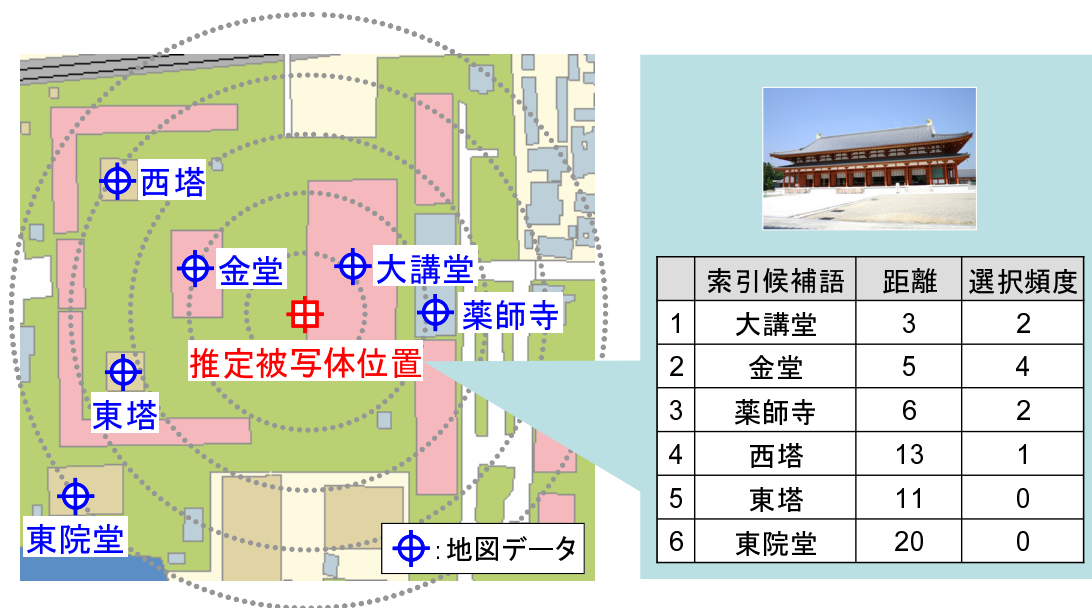


図 13 地図データベースからの索引候補語の取得

られた専門用語についてその用語と関連する用語を web テキストを利用して収集する手法を提案している。佐藤らの手法では、収集した用語を辞典に利用することを目指し、正確な抽出を主眼としている。このため、処理にかかる時間は問題とされていない。一方、本研究で目的とするシステムでは、ユーザがインタラクティブに索引付加作業を行うため、処理にかかる時間を考慮する必要がある。このため本システムでは、web 検索の結果として取得される web ページのうち、抽出に使用するページ数を少なくし、また、web ページの解析による関連語の抽出は以下に述べる方法で行う。

図 14 に関連語を抽出するための処理の概要を示す。抽出処理の入力は、地図データベースを参照して取得した被写体位置付近の地名や施設名の中からユーザにより選択された単語である。この選択された単語を以下、「キーワード」と呼ぶ。まず、このキーワードを用いて web 検索を行い、得られた URL の上位 M 件の web ページを取得する。次に、ページ内の HTML タグ等を除いた、タグ間のテキスト部分に対して形態素解析を行って、文を単語に分割し各単語の品詞情報

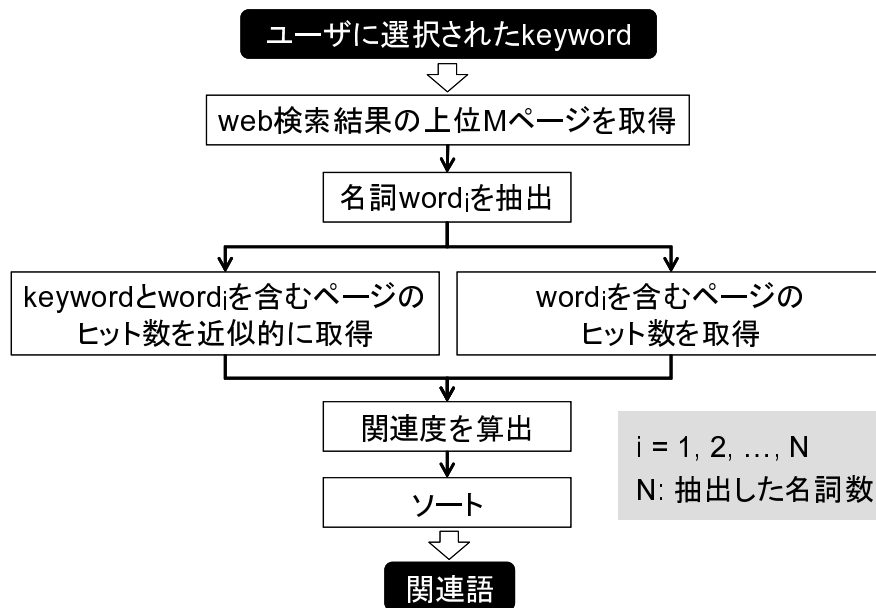


図 14 web 検索を用いた関連語抽出処理

を得る．これに基づき，索引語に適している名詞に分類された単語のみを抽出する．ここで，キーワードと抽出した各単語の関連を示す指標として，式 (1) で定義される関連度を考える．これは，web 内でキーワードの出現するページ集合と抽出した名詞  $word_i$  の出現するページ集合の積集合と和集合の比で表されており，この値が大きいほどキーワードと  $word_i$  の関連は深いといえる．

$$\begin{aligned}
 relevance_i &= hit_{key \cap word_i} / hit_{key \cup word_i} \\
 &= hit_{key \cap word_i} / (hit_{key} + hit_{word_i} - hit_{key \cap word_i}) \quad (i = 1, 2, \dots, N), \quad (1)
 \end{aligned}$$

但し,

$relevance_i$  :  $word_i$  の関連度,

$hit_{key}$  : キーワードを含むページの検索ヒット数,

$hit_{word_i}$  :  $word_i$  を含むページの検索ヒット数,

$hit_{key \cap word_i}$  : キーワードと  $word_i$  を含むページの検索ヒット数,

$hit_{key \cup word_i}$  : キーワードか  $word_i$  のどちらか含むページの検索ヒット数,

$N$  : 抽出した名詞の数.

関連度  $relevance_i$  の算出には, 新たに  $hit_{word_i}$ ,  $hit_{key \cap word_i}$  を取得する必要があるが, この取得には非常に時間がかかる.  $hit_{word_i}$ ,  $hit_{key \cap word_i}$  を取得するためには, それぞれ抽出した名詞の数である  $N$  回, 合わせて  $2N$  回の web 検索が必要である. 例えば, キーワードによる検索で上位 10 件の web ページを取得し, それぞれのページで 50 語の名詞が抽出されたとすると, 全体で 500 語の名詞が抽出される. したがって, (1) 式により関連度を算出するには 1000 回の web 検索が必要となる. これを短縮するため,  $hit_{key \cap word_i}$  は式 (2) により近似的に取得する. これは,  $hit_{key \cap word_i} / hit_{key}$  が, 先に取得したキーワードを含む上位  $M$  件の web ページのうち  $word_i$  が出現する確率とほぼ同じと仮定したものである. この近似により, 関連度の算出に要する検索回数を  $1/2$  にすることができる. いくつかのキーワードについて近似を用いた方法と用いない方法で関連度を算出したところ, 大きな違いは見られなかった.

$$hit_{key \cap word_i} \simeq hit_{key} \times page_{word_i | key} / page_{key}, \quad (2)$$

但し,

$page_{key}$  : キーワードで取得した web ページ数,

$page_{word_i | key}$  : キーワードで取得した web ページのうち  
 $word_i$  が出現した web ページ数.

以上により,  $page_{word_i}$  と  $hit_{word_i}$  を取得することで, 関連度  $relevance_i$  を算出する. この関連度  $relevance_i$  で, 抽出された名詞  $word_i$  をソートし, 索引候補語としてユーザに提示する.

### 3.6 ユーザ選択をフィードバックとした地図データベースの更新

ユーザは、地図データベース参照もしくは web 検索を用いた関連語抽出処理によって取得された索引候補語から、写真の内容に合った地名や施設名を選択する。選択された名称は、索引語として写真に対応付けられると共に、地図データベースへのフィードバックとして用いて、データベースの更新を行う。

図 15 は、ユーザが選択した索引語をフィードバックとした地図データベースの更新処理について示したものである。ユーザによって決定された索引語は、まず、その索引語の取得元によって処理を分岐する。索引語が市販の地図データや国土交通省提供の地図データなど既存の地図データから取得されたものである場合、その位置情報は信頼できると考えられる。そこで、索引語がユーザによって選択された回数を示す頻度を増加させる操作のみを行う。一方、索引語が web 検索を用いた関連語抽出処理により取得されたものである場合、その索引語が対応付けられる位置は写真の推定被写体位置である。このため、位置・姿勢情報の取得におけるセンサの精度によって、推定した被写体位置には誤差が生じている。これを考慮し、以下のような処理を行う。まず、その索引語が既に地図データベースに登録されているか否かで処理を分岐する。索引語が地図データベースに登録されていない単語である場合、対応付けられた写真の推定被写体位置をその単語の位置情報としてデータベースへの登録を行う。索引語が既に地図データベースに登録されている単語である場合、地図データベースにおいて対応付けられた単語の位置情報と、ユーザによる索引付加作業によって新たに対応付けられた写真のもつ位置情報を用いて、式 (3) により算出された値を新たな位置情報としてデータベースを更新する。また、ユーザによる選択頻度についても更新を行う。

$$\begin{cases} latitude_{new} &= (latitude_{prev} \times frequency_{prev} + latitude_{sbj}) / frequency_{new}, \\ longitude_{new} &= (longitude_{prev} \times frequency_{prev} + longitude_{sbj}) / frequency_{new}, \\ frequency_{new} &= frequency_{prev} + 1, \end{cases} \quad (3)$$

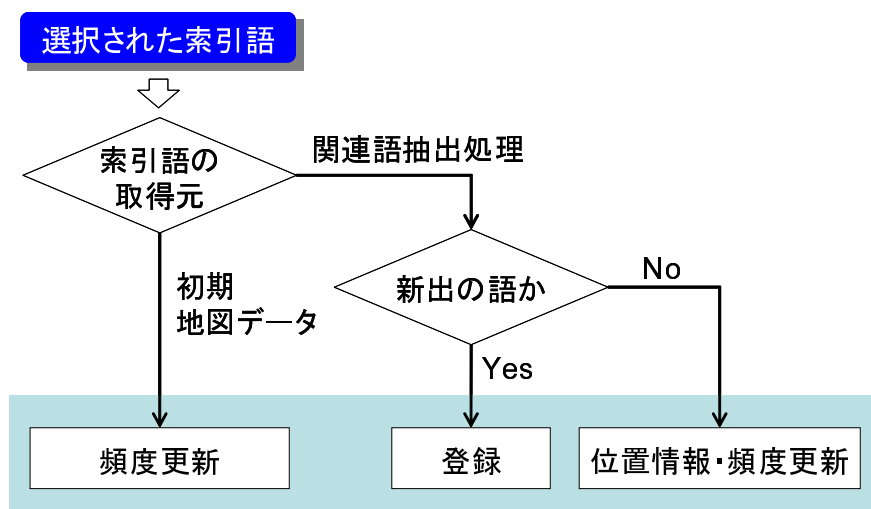


図 15 ユーザ選択の地図データベースへのフィードバック

但し，

$latitude_{new}$ ,  $longitude_{new}$  : 索引語に対応付けられる新たな緯度・経度,

$latitude_{prev}$ ,  $longitude_{prev}$  : データベースに登録されている緯度・経度,

$latitude_{sbj}$ ,  $longitude_{sbj}$  : 被写体推定位置の緯度・経度,

$frequency_{prev}$ ,  $frequency_{new}$  : ユーザによる選択頻度.

図 16 は，地図データベースにおける索引語と対応する位置，また，その更新のイメージを地図上に示したものである．索引語とこれを付加された写真の位置情報の対応関係を，地図データベースに保持しておくことで，次に同一の被写体が撮影された際の索引付加作業は効率化される．また，新たに得られる写真の位置情報を利用して，索引語に対応付けられた位置を更新する．つまり，同一被写体の写真が撮影され，ユーザによる索引付加作業のフィードバックが得られるごとに，地図データベース内の対応する索引語と位置情報からなるデータは，センサによる誤差の影響を低減するなどの目的から位置情報の平均化が行われる．また，よく使用される索引語の選択頻度は増加し，この 2 つの値に基づいた索引候補語の提示を行うことで，ユーザ入力の効率化が図られるものと考えられる．



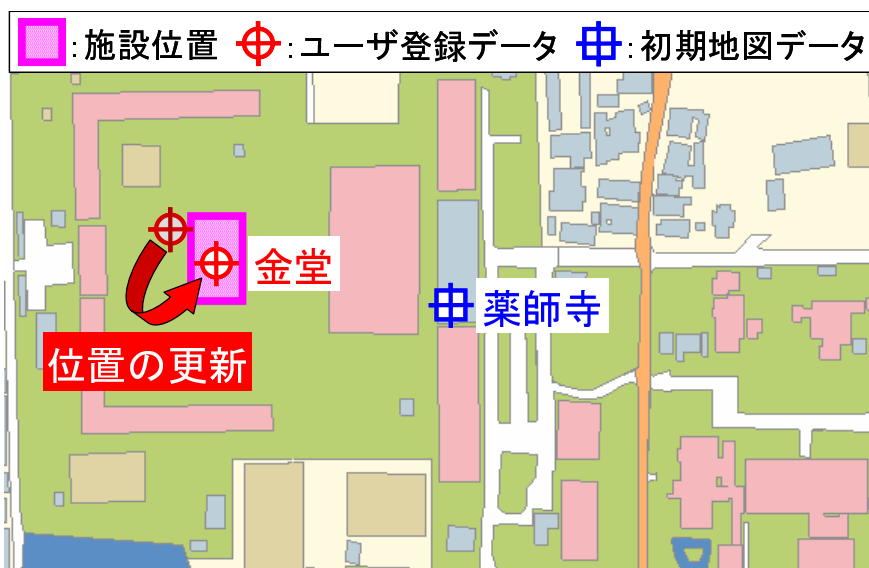


図 16 地図データベース更新のイメージ

### 3.7 索引付加写真閲覧のための写真データベースの構築

ユーザによる索引付加作業が完了すると、写真と付加された索引語は写真データベースに格納される。写真データベース内の各レコードは、写真（画像ファイル名）、索引語、撮影日時、撮影位置・姿勢情報といったメタデータから構成される。ユーザは付加した索引語を見出しとして、写真データベース内の写真をアルバムのように閲覧することができる（図 17）。また、写真に付加された索引語やメタデータはテキストデータであるため、写真に対するキーワード検索を行うことが可能である。

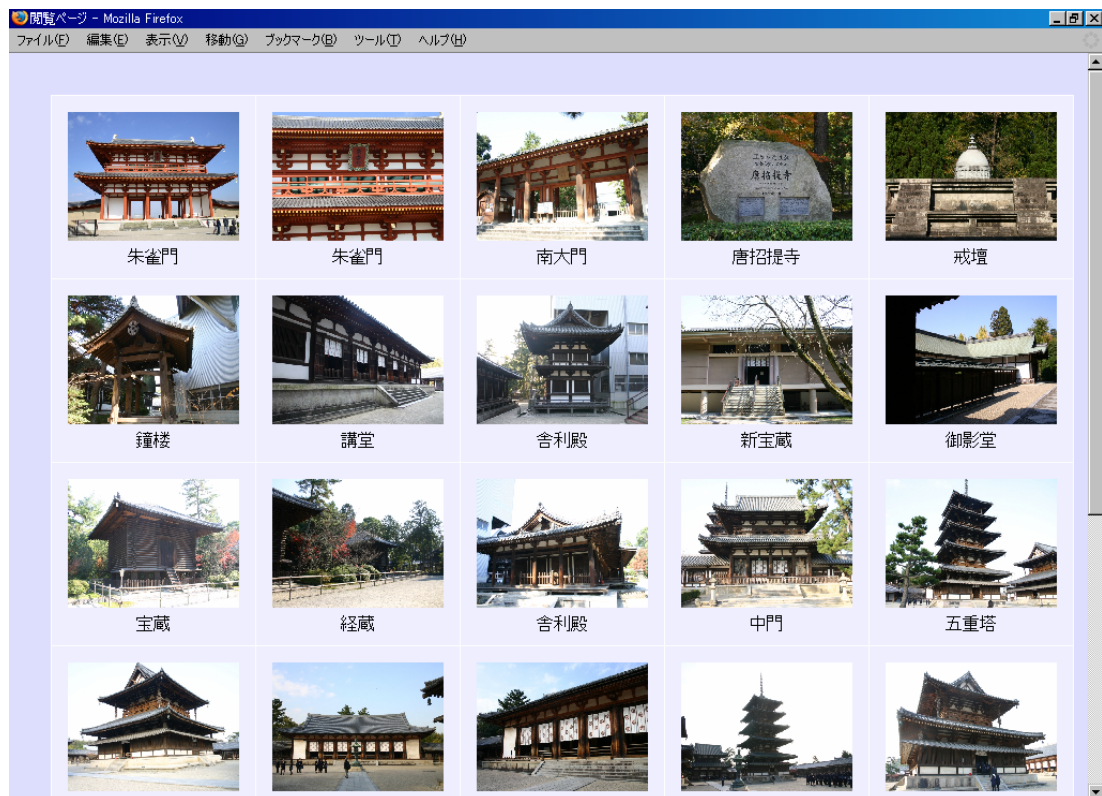


図 17 索引を付加した写真の閲覧

## 4. 実験

提案する写真への索引付加手法の有効性を確認するために、プロトタイプシステムを作成し、実際に観光地等で撮影した写真に対して索引を付加する実験を行った。4.1 節では、提案手法に基づいて作成したプロトタイプシステムの機器構成と各部について説明する。4.2 節では、作成したプロトタイプシステムを使用して、撮影した写真に対して索引付加を行った実験について述べる。

### 4.1 プロトタイプシステム

提案したシステムのプロトタイプを作成し、撮影した写真への索引付加を行った。プロトタイプシステムは、図 18 に示すようにクライアントとサーバからなる。クライアントは、写真と撮影時の位置・姿勢情報を取得するカメラ、GPS、ジャイロからなるセンサ付カメラと、これらの情報を記録し、サーバとの情報のやり取りを行うクライアント PC から構成される。システムに使用した機器類については表 2 に示す。また、サーバは、クライアントから送られる撮影位置・姿勢情報に基づき、格納されている地図データベースの参照および web 検索を用いた関連語抽出処理を用いて、ユーザに写真の索引候補語を提示する。ユーザが図 19 に示すブラウザ上で動作する選択画面から、索引語を選択すると、写真と索引語は写真データベースに格納される。また、索引語をフィードバックとして用いた地図データベースの更新が行われる。

プロトタイプシステムを用いた索引付加作業は、以下のように行われる。まず、ユーザがセンサ付カメラにより撮影を行うと、写真と撮影時の位置・姿勢情報がクライアント PC に記録される。記録した情報をクライアント PC から、ネットワークを介してサーバに送信し、索引付加作業を開始する。ユーザは、web ブラウザを用いた入力画面で索引を付加する写真、位置・姿勢情報を記録したファイルを指定し、サーバに送信する。サーバは、受信した情報から写真の被写体位置を推定し、サーバ内に設置された地図データベースを参照することで対応する位置の地名や施設名を索引候補語として取得する。これをクライアント側に返信すると、図 19 に示す索引語選択画面がユーザに提示される。ユーザは、表示された

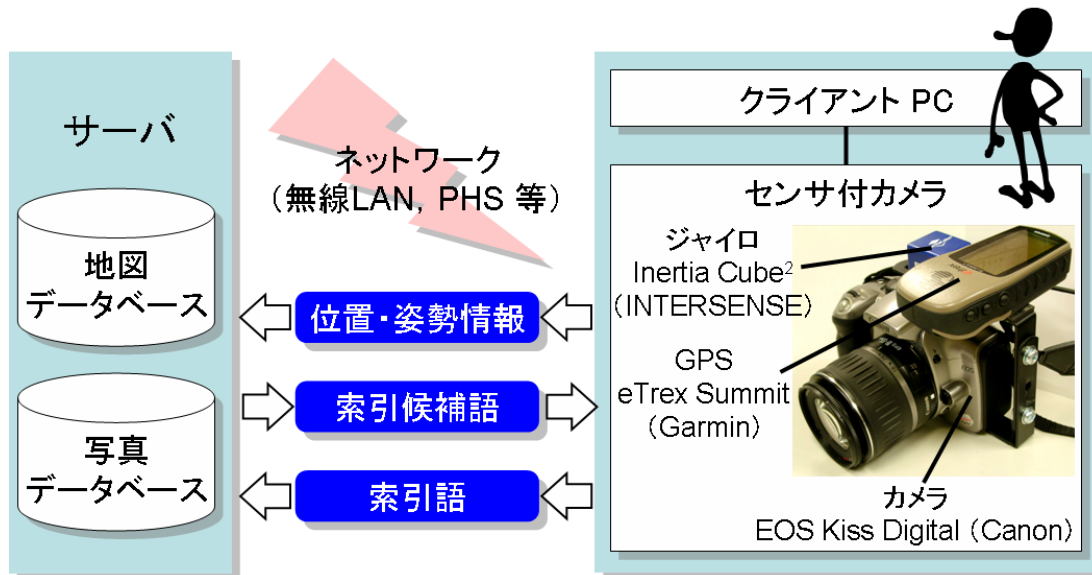


図 18 プロトタイプシステムの構成

表 2 プロトタイプシステムに用いた機器類

ジャイロセンサ	Inertia Cube <sup>2</sup> (InterSense)	
GPS	eTrex Summit (Garmin)	
カメラ	EOS Kiss Digital (Canon)	
クライアント PC	LavieG (NEC)	
	web ブラウザ	Mozilla FireFox 1.0
	PHS	b-mobile (日本通信株式会社)
サーバ PC	web サーバ	Apache 1.3.27
	データベース	PostgreSQL 7.3.2
	サーブレット	Tomcat 5.5.3, JDK5.0

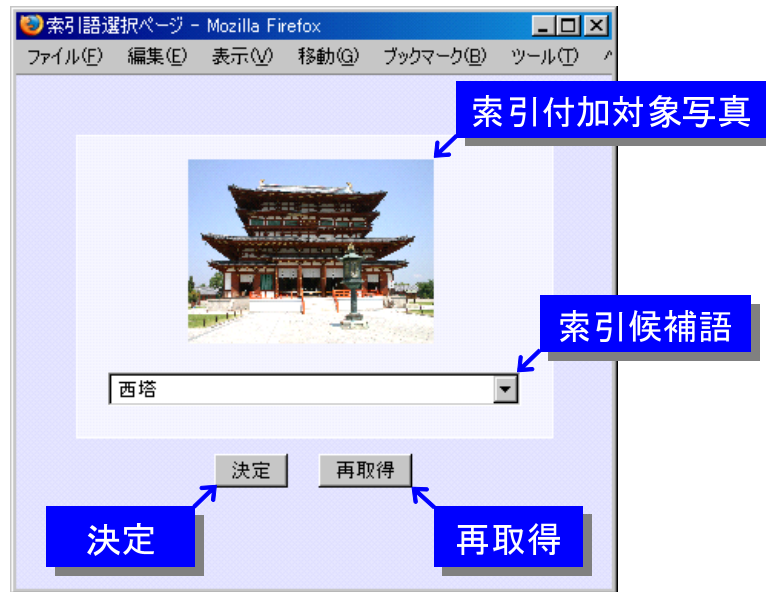


図 19 索引語選択画面

写真に対して付加したい索引語を、プルダウン部分に提示される候補語から選択する。候補の中に付加したい索引語が含まれている場合には、それを選択して送信ボタンを押すことで、写真データベースに写真と索引語、撮影時刻、撮影・推定被写体位置等の関連情報が格納される。候補の中に索引語が含まれていないと判断した場合には、ユーザが写真に付加したい索引語に最も関連すると思う単語を選択して再取得のボタンを押すことで、新たな索引候補語の取得がサーバに要求される。サーバは、クライアントから送られる単語から、web 検索を用いた関連語抽出処理を行う。この結果得られた関連語を再びクライアントに送信し、同様の選択作業がユーザによって行われる。ユーザによる選択作業が行われると、選択された索引語は写真に対応付けられると共に、サーバへと送信され、これをフィードバックとした地図データベースの更新処理が行われる。

初期地図データベースの構築には、国土交通省の国土数値情報・施設データと市販の地図ソフト（アルプス社製「プロアトラス W2」）の施設データを使用し、ネットワーク上のサーバに置いた。web 検索を用いた関連語抽出処理には Google

の提供する Google API [38] を検索エンジンとして使用し，検索結果の上位 10 件のページを取得した（図 14 において  $M=10$ ）．また，形態素解析には日本語形態素解析システム「茶筌」[39] を使用した．

## 4.2 撮影した写真への索引付加実験

### 4.2.1 予備実験

予備実験では，システムが提案手法に基づいた動作をするか確認し，写真に対して適切な索引を付加できる場合とそうでない場合について検証した．図 20(a) ~ 23(a) は，奈良の薬師寺において撮影したものである．これらの写真に対してプロトタイプシステムを用いた索引付加作業を行った．

実験で用いた地図データベースには，索引候補語として「薬師寺」が含まれているが，薬師寺内にある個々の建物の名称は含まれていない．以下では，図 20(a) ~ 23(a) の順に索引付加作業を行った様子について述べる．

図 20(a) は「金堂」を撮影した写真であり，ここでは，この「金堂」という索引語を付加することを想定する．サーバは初期状態の地図データベースを保持しており，この写真とその撮影位置・姿勢情報を送信すると図 20(b) に示すように，「薬師寺」，「西の京派出所」などの索引候補語がユーザに提示される．ここでユーザが「金堂」の上位層の単語である「薬師寺」を選択し，再取得のボタンを押すと，サーバ側で「薬師寺」の関連語抽出処理が行われ，図 20(c) の画面がユーザに提示される．索引候補語の中には「金堂」が含まれており，ユーザがこれを選択し，送信することで写真データベースへの登録と地図データベースの更新が行われ，地図データベースには，表 3 の上段のレコードが追加される．また，このとき「薬師寺」の関連語抽出処理により得られた候補語は，表 4 に示す単語を含む 402 語であり，その上位 100 語程度の中に薬師寺内のほとんどの建築物の名称が得られている．

次に図 21(a) は，図 20(a) と同様に「金堂」を撮影した写真で，図 20(a) とは別の方向から撮影されたものである．これをサーバに送信すると，図 21(b) の画面がユーザに提示される．地図データベースが図 20(a) 登録時に更新されたことが

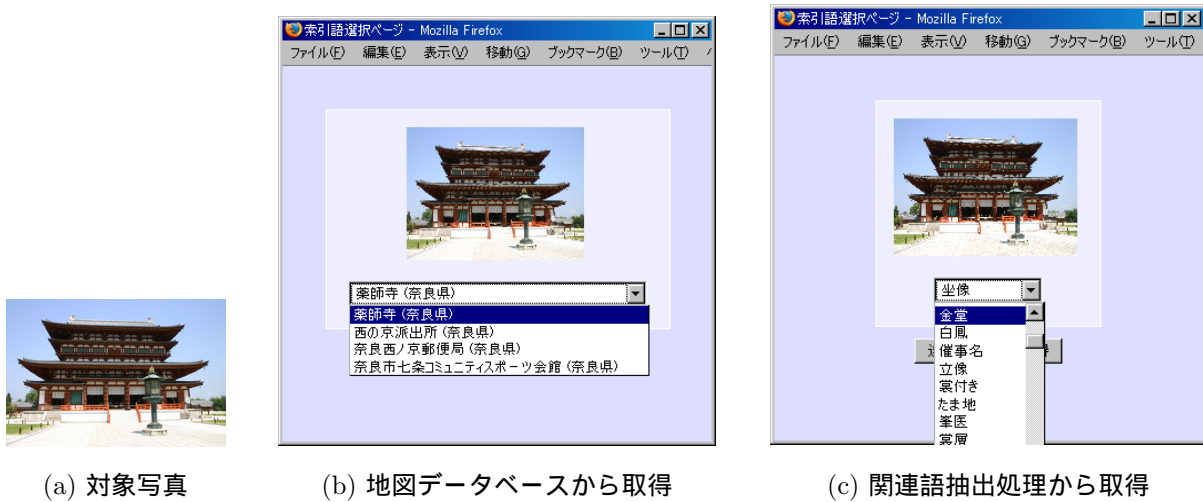


図 20 金堂 1



図 21 金堂 2



(a) 対象写真



(b) 関連語抽出処理から取得

図 22 玄奘三蔵院



(a) 対象写真



(b) 取得不能

図 23 興楽門



表 3 地図データベースの更新

input image	name	N/S	latitude	E/W	longitude	frequency
図 20(a)	金堂	N	34 °40' 4.7"	E	135 °47' 3.5"	1
図 21(a)	金堂	N	34 °40' 6.0"	E	135 °47' 4.1"	2

表 4 抽出された「薬師寺」の関連語 (上位 20 語)

1	薬師尊像	11	東塔
2	薬師寺ホームページ	12	門扉
3	東院堂	13	西塔
4	老春手帳	14	白鳳時代
5	玄奘三蔵院	15	統天皇
6	フェノロサ	16	天武天皇
7	平山郁夫画伯	17	修二会
8	花会式	18	奈良駅
9	三重塔	19	金堂
10	光明皇后	20	催事名

太字は「薬師寺」内の施設名

ら、1度目の問い合わせから候補語として「金堂」が得られており、これを選択、送信するのみで写真データベースへの登録が完了する。また、ここでも地図データベースの更新が行われ、「金堂」に対応付けられる位置情報が、図 21(a) のもつ位置情報を利用して更新される。結果、表 3 上段に示した「金堂」のレコードは表 3 下段のように更新された。この更新前後の 2 点間の距離は約 43m である。なお、撮影時の誤差は GPS から出力された値によると約 10m であった。使用した 2 枚の写真の推定被写体位置間の距離は約 85m 離れており、推定精度が良いとはいえない。これは、算出に使用している Exif 情報から得られた被写体までの距離が、精度の高いものでないことによる。このため今後は、同じ対象を撮影した複数の写真の撮影位置・方向から、その交点を算出し、索引候補語の位置とする方法などを検討する必要がある。

同様の処理によって、図 22(a) については図 22(b) に示すように索引候補語「玄奘三蔵院」が選択され、索引が付加される。なお、「玄奘三蔵院」は、索引候補語中の 5 番目に提示されている。一方、図 23(a) については図 23(b) に示すように、地図データベースへの問い合わせ、関連語抽出処理を行ったものの、適切な索引語である「興楽門」が得られなかった。これは、関連語抽出処理において取得した上位 10 件の web ページに「興楽門」という単語が含まれていなかったことによるものである。

予備実験における索引付加作業により、プロトタイプシステムが提案手法に基づいた動作をすることを確認した。具体的には、地図データベースから写真撮影時の位置・姿勢情報に基づく地名や施設名が取得されること、web 検索を用いた関連語抽出処理により、入力として与えた語に関連する語が取得されることを確認した。これらの処理により取得された単語を索引候補語としてユーザに提示することで、写真に対する半自動的な索引付加作業を実現することができた。また、ユーザが索引付加作業において選択した索引語をフィードバックとして、地図データベースの更新を行うことにより、同一の被写体に対する作業が効率化されることを確認した。一方、提案手法では索引語を付加できない事例も確認されたため、これについては、索引候補語を取得する手法の改良や新たな手段の検討が必要であると考えられる。

#### 4.2.2 実証実験

観光施設を対象として、予備実験よりも多くの写真を撮影し、プロトタイプシステムを用いた索引付加作業によって写真に対して適切な索引を付加することができるか、また、地図データベースの登録・更新について検証するための実験を行った。実験において索引付加の対象とした写真は、薬師寺及び法隆寺において撮影したものであり、その一部を図 24, 25 に示す。薬師寺について 9 施設, 72 枚, 法隆寺について 12 施設, 49 枚の写真に対して索引付加を行った。索引付加作業では、ユーザが写真の被写体であるそれぞれの施設に対して、その施設名を索引として付加することを想定する。地図データベースと web 検索を用いた関連語抽出処理についての条件は、予備実験と同様である。

予備実験と同様の索引付加作業を、写真を撮影した順序で行った。表 5, 6 は、被写体とした各施設について、その名称と撮影枚数、web 検索を用いた関連語抽出処理で取得した際のユーザへの提示順位、地図データベースから取得した際の平均提示順位を示している。薬師寺では、付加する索引語は索引候補語中、関連語抽出処理において 6 枚の写真を対象に平均 23.8 位、地図データベースからの取得において 56 枚を対象に平均 1.4 位に提示された。ほとんどの場合で、選択する索引語は候補語リスト中の上位にあり、ユーザによる選択作業は簡単なものであった。法隆寺では、付加する索引語は索引候補語中、関連語抽出処理において 8 枚を対象に平均 117.6 位、地図データベースからの取得において 37 枚を対象に平均 2.8 位に提示された。関連語抽出処理では、選択する索引語がリストの下位にあったため選択作業の負担は大きいものであった。しかし、索引語が、一度地図データベースへ登録された後の作業ではリストの上位に提示されており、選択作業の負担は軽減された。実験を通して、地図データベースに一度登録された索引語については、ユーザにそれほど負担をかけることなく索引付加作業を行うことができた。但し、施設によって提示順位にはばらつきがあった。この原因としては、施設同士が近接している場合や施設が広がりをもつ場合に、対象施設の周囲の施設の方が上位に提示されることが挙げられる。写真撮影時にセンサから取得した位置・姿勢情報に含まれる誤差も原因の一つと考えられる。

また、薬師寺では 9 施設中 3 施設 (72 枚中 10 枚)、法隆寺では 12 施設中 2 施設



図 24 索引付加対象写真 1：薬師寺（一部）



図 25 索引付加対象写真 2：法隆寺（一部）

(49 枚中 4 枚) で、適切な施設名を索引語として付加することができなかった。これは、web 検索による関連語抽出処理において取得した web ページ中に、それらの単語が含まれていなかったことが原因である。索引の付加できなかった施設は、観光施設としてはあまり注目されない施設であり、そのため、取得した web ページにも含まれていなかったと考えられる。同時に、一般の観光客が本システムを使用する際にも、撮影される可能性は低いものと考えられるため、実用上それほど問題にはならないと考えられる。

索引付加作業による地図データベースへの登録・更新の様子として、薬師寺内で撮影した写真のうち、大講堂の写真 11 枚に索引付加を行った際に、地図データベースにおいて索引語「大講堂」の対応付けられる位置が変化していった様子を図 26 に示す。図 26(a) に示す、1 枚目の写真に索引を付加した際の登録位置から、索引付加枚数の増加と共に、索引語「大講堂」の対応付けられる位置が、実際の建物位置を示した領域の中央付近に近づいていることがわかる。したがって、薬師寺の大講堂について適切な地図データが作成され、今回の索引付加作業後に別のユーザが大講堂の写真を撮影し、索引を付加する際には、この地図データを

表 5 索引付加作業における索引語の提示順位：薬師寺

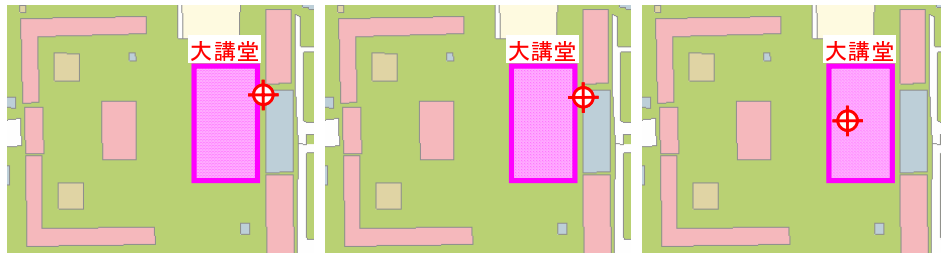
索引語	撮影枚数	関連語抽出処理における提示順位	地図 DB による平均提示順位
西塔	7	11	1.5
東塔	7	10	1.3
大講堂	11	100	1.7
玄奘三蔵院	19	2	1.0
金堂	12	15	1.7
東院堂	6	5	1.2
鐘楼	4	取得なし	-
與樂門	4	取得なし	-
南門	2	取得なし	-
平均	8	23.8	1.4

利用した効率的な索引付加作業が行われるものと考えられる。

また、索引付加作業による地図データベースへの登録・更新の結果を図 27、28 に示す。初期データに加えて、索引語として得られた施設名が地図データベースに登録されている。索引語（施設名）に対応付けられた位置は、ほとんどが施設位置上またはその付近にあり、索引付加作業によって適切な地図データが作成されたことがわかる。しかし、図 27 の「東院堂」や図 28 の「中門」、「大講堂」のように実際の施設位置から外れた位置に対応付けられた索引語（施設名）もあり、これは主に、撮影位置・姿勢情報を取得した際のセンサ類の誤差や Exif 情報から取得する被写体までの距離の精度の低さのために、正確な被写体位置が推定できていない場合があるためと考えられる。したがって、今後、被写体位置の推定や地図データベースにおける位置情報の更新方法等について検討が必要であると考えられる。

表 6 索引付加作業における索引語の提示順位：法隆寺

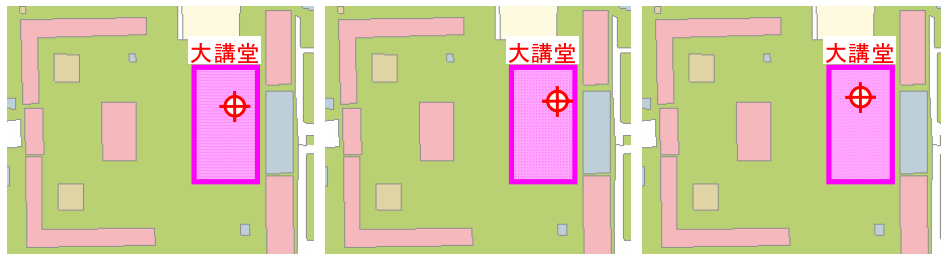
索引語	撮影枚数	関連語抽出処理における提示順位	地図 DB による平均提示順位
大講堂	3	163	4.5
中門	4	146	2.7
金堂	7	6	3.3
綱封蔵	1	42	-
大宝蔵院	4	100	1.3
礼堂	2	98	1.0
絵殿	2	265	-
聖霊院	2	121	1.0
夢殿	8	処理なし（初期データ）	2.6
五重塔	12	処理なし（初期データ）	3.3
東大門	2	取得なし	-
四脚門	2	取得なし	-
平均	4.3	117.6	2.8



(a) 1 枚目：登録

(b) 2 枚目：更新

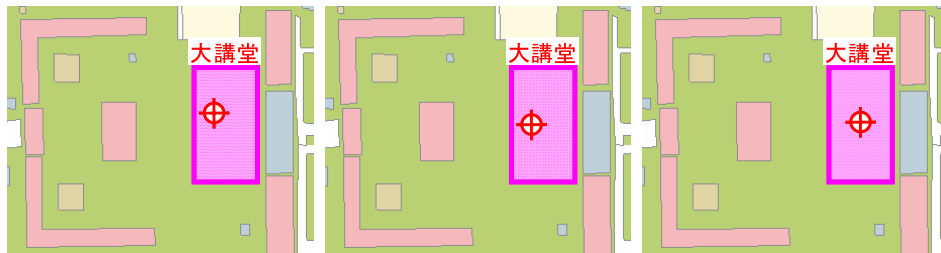
(c) 3 枚目：更新



(d) 4 枚目：更新

(e) 5 枚目：更新

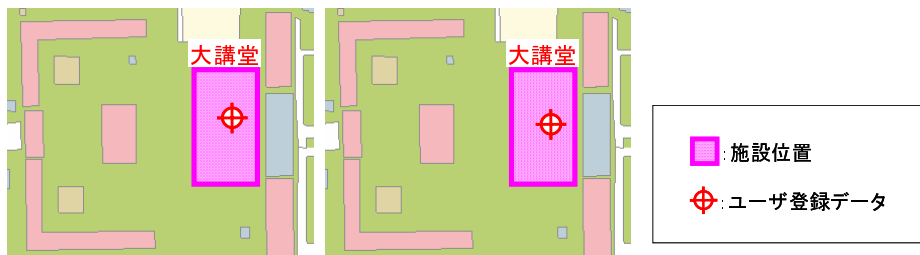
(f) 6 枚目：更新



(g) 7 枚目：更新

(h) 8 枚目：更新

(i) 9 枚目：更新



(j) 10 枚目：更新

(k) 11 枚目：更新

図 26 索引付加作業による写真毎の地図データの登録・更新：大講堂

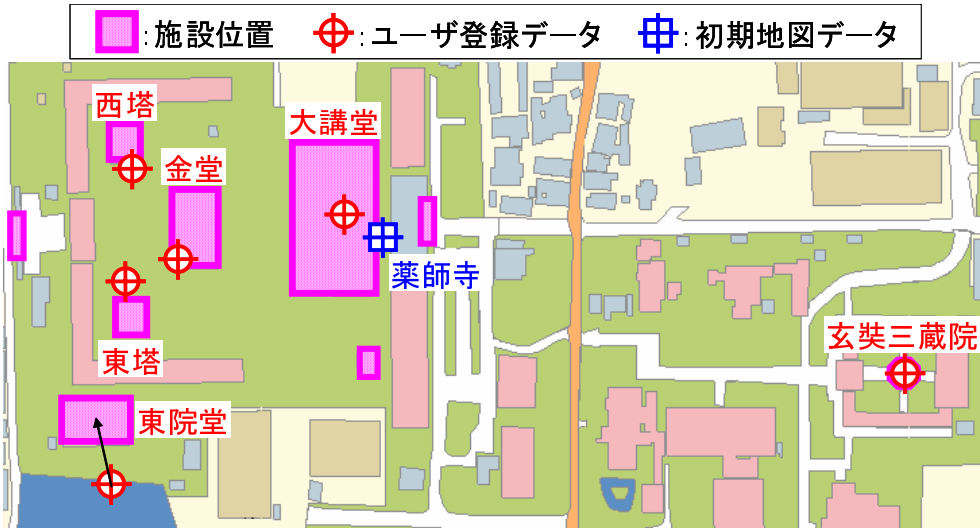


図 27 地図データベースへの登録・更新：薬師寺

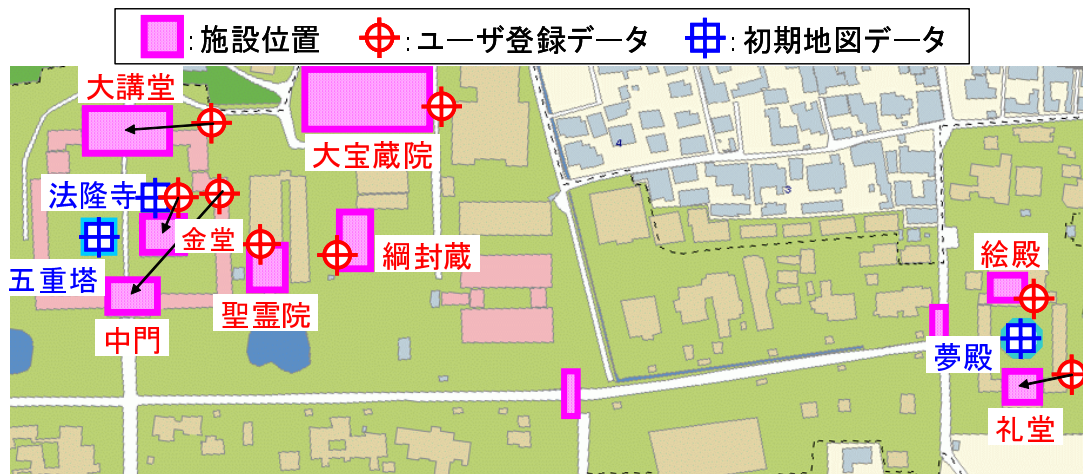


図 28 地図データベースへの登録・更新：法隆寺



## 5. まとめ

本研究では、個人が撮影した写真を効率的に管理することを目的とし、従来から研究の行われている画像検索手法のうち、個人の写真について適しているとされる、メタデータを利用して検索を行う手法について検討した。また、メタデータのうち特に、写真の撮影位置情報に着目し、これをユーザが閲覧や検索を行うのに適した被写体の地名や施設名等のテキスト情報として、写真に付加しておくための手法を検討した。

以上の検討に基づき、写真の撮影位置・姿勢情報を利用して被写体の位置を推定し、推定位置の地名や施設名の候補をユーザに提示することで、半自動的な索引付加作業を実現するシステムを提案した。索引候補語は、あらかじめ用意された地図データベースから対応する位置の地名や施設名を取得し、写真に適切な索引語が含まれていない場合には、web検索を用いた関連語抽出処理によって取得した。ここでユーザにより選択された索引語は、その位置に適切な語であると見なし、これを地図データベースへフィードバックさせることで提示される候補語を変化させ、ユーザの選択作業の効率化を図った。また、地図データベースは複数のユーザ間で共有するため、ユーザのうち一人が、ある被写体について索引を付加すると、その後他のユーザが同じ被写体を撮影し、索引を付加する際の選択作業は効率化される。

提案したシステムのプロトタイプを用いて、撮影した写真に対して位置・姿勢情報に基づいた索引語の付加実験を行い、適切な索引語の付加が可能であることを確認した。また、システムの利用により、地図データベースへの新たな地図データの追加や更新が行われることで、索引候補語をより適切に提示することができたと考えられる。

今後の課題としては、まず、提案手法で取得できていない地名や施設名の取得方法の検討が挙げられる。提案した取得方法では、地図データベースにない単語を関連語抽出処理によって取得しているが、この処理を改良することや簡単な入力インタフェースの導入などを検討する。次に、地図データベースに登録される候補語の位置情報について、実際の位置とずれが生じる点について検討する必要がある。この主な原因は、取得した位置・姿勢情報等に含まれる誤差であるため、

今後，撮影方向や撮影時のセンサの受信状況による誤差や被写界深度，被写体の大きさ等を考慮した方法を検討する．また，撮影場所および利用者に関してより広範な評価実験を行うこと，構築した写真データベースにユーザがアクセスするためのインタフェースを検討することなどが必要である．

## 謝辞

本研究の全過程を通して，懇切なる御指導，御鞭撻を頂いた視覚情報メディア講座 横矢 直和教授に心より感謝申し上げます．

また，本研究の遂行にあたり，有益な御助言，御鞭撻を頂いた 像情報処理学講座 千原 國宏教授，ならびに本研究を進めるにあたり多くの御指導，御助言を頂いた 視覚情報メディア講座 山澤 一誠助教授に厚く御礼申し上げます．

さらに，物心両面において温かい御支援を頂いた 視覚情報メディア講座 神原 誠之助手，的確な御助言を頂いた 視覚情報メディア講座 佐藤 智和助手に深く感謝致します．また，研究室での生活を支えて頂いた 視覚情報メディア講座事務補佐員 守屋 智代女史に心より感謝申し上げます．

最後に，研究活動において常に温かい御助言を頂いた視覚情報メディア講座の諸氏に深く感謝致します．

## 参考文献

- [1] 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. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 12, pp. 1349–1380, 2000.
- [2] Remco C. Veltkamp and Mirela Tanase. Content-based image retrieval systems: A survey. Technical Report TR UU-CS-2000-34 (revised version), Department of Computing Science, Utrecht University, 2002.
- [3] AltaVista Photofinder. <http://www.altavista.com/image/>.
- [4] Google Image Search. <http://images.google.com/>.
- [5] Adobe Photoshop Album. <http://www.adobe.com/>.
- [6] Apple iPhoto. <http://www.apple.com/>.
- [7] IrfanView. <http://www.irfanview.com/>.
- [8] Roberto Brunelli and Ornella Mich. Efficient image retrieval by examples. In *Proc. 5th Working Conf. on Visual Database Systems*, pp. 145–162, 2000.
- [9] Madirakshi Das, Edward M. Riseman, and Bruce A. Draper. Focus: Searching for multi-colored objects in a diverse image database. In *Proc. 1997 Conf. on Computer Vision and Pattern Recognition (CVPR '97)*, p. 756, 1997.
- [10] Eugenio Di Sciascio, G. Mingolla, and Marina Mongiello. Content-based image retrieval over the web using query by sketch and relevance feedback. In *Proc. 3rd Int. Conf. on Visual Information and Information Systems*, pp. 123–130, 1999.

- [11] John R. Smith and Shih fu Chang. Querying by color regions using VisualSEEk content-based visual query system. In *Intelligent Multimedia Information Retrieval*, pp. 23–41. MIT Press, 1997.
- [12] Mandis Beigi, Ana B. Benitez, and Shih-Fu Chang. MetaSEEk: A content-based meta-search engine for images. In *Proc. SPIE 1998 Conf. on Storage and Retrieval for Image and Video Databases VI IST/SPIE 1998*, Vol. 3312, pp. 28–30, 1998.
- [13] 相良直樹, 砂山渡, 谷内田正彦. HTML テキストの重要文を用いた画像ラベリング手法. 電子情報通信学会論文誌 (D-I), Vol. J87-D-I, No. 2, pp. 145–153, 2004.
- [14] Yuksel Alp Aslandogan and Clement T. Yu. Multiple evidence combination in image retrieval: Diogenes searches for people on the web. In *Proc. 23rd Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*, pp. 88–95, 2000.
- [15] Yuksel Alp Aslandogan and Clement T. Yu. Diogenes: a web search agent for person images. In *Proc. 8th ACM Int. Conf. on Multimedia*, pp. 481–482, 2000.
- [16] Ben Shneiderman and Hyunmo Kang. Direct Annotation: A drag-and-drop strategy for labeling photos. In *Proc. Int. Conf. on Information Visualization*, pp. 88–95, 2000.
- [17] Mor Naaman, Yee Jiun Song, Andreas Paepcke, and Hector Garcia-Molina. Automatic organization for digital photographs with geographic coordinates. In *Proc. 2004 Joint ACM/IEEE Conf. on Digital Libraries*, pp. 53–62, 2004.
- [18] Mor Naaman, Yee Jiun Song, Andreas Paepcke, and Hector Garcia-Molina. Automatically generating metadata for digital photographs with geographic coordinates. In *Proc. 13th Int. World Wide Web Conf. on Alternate track papers & posters*, pp. 244–245, 2004.

- [19] Kentaro Toyama, Ron Logan, Asta Roseway, and P. Anandan. Geographic location tags on digital images. In *Proc. 11th ACM Int. Conf. on Multimedia*, pp. 156–166, 2003.
- [20] Japan Electronics and Information Technology Industries Association(JEITA). Exchangeable image file format for digital still cameras: Exif version 2.2, 2002.
- [21] Ana B. Benitez, Mandis Beigi, and Shih-Fu Chang. Using relevance feedback in content-based image metasearch. *IEEE Internet Computing*, Vol. 2, No. 4, pp. 59–69, 1998.
- [22] John Adcock, Matthew Cooper, John Doherty, Jonathan Foote, Andreas Girgensohn, and Lynn Wilcox. Managing digital memories with the fxpal photo application. In *Proc. 11th ACM Int. Conf. on Multimedia*, pp. 598–599, 2003.
- [23] David Frohlich, Allan Kuchinsky, Celine Pering, Abbe Don, and Steven Ariss. Requirements for photware. In *Proc. 2002 ACM Conf. on Computer Supported Cooperative Work*, pp. 166–175, 2002.
- [24] Timothy J. Mills, David Pye, David Sinclair, and Kenneth R. Wood. Shoebox: A digital photo management system. Technical Report 2000.10, AT&T Laboratories Cambridge, 2000.
- [25] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. *Int. Journal of Computer Vision*, Vol. 18, No. 3, pp. 233–254, 1996.
- [26] Jiebo Luo, Amit Singhal, Stephen P. Etz, and Robert T. Gray. A computational approach to determination of main subject regions in photographic images. *Image and Vision Computing*, Vol. 22, pp. 227–241, 2004.

- [27] Andreas Girgensohn, John Adcock, and Lynn Wilcox. Leveraging face recognition technology to find and organize photos. In *Proc. 6th ACM SIGMM Int. Workshop on Multimedia Information Retrieval*, pp. 99–106, 2004.
- [28] Kerry Rodden and Kenneth R. Wood. How do people manage their digital photographs? In *Proc. Conf. on Human Factors in Computing Systems*, pp. 409–416, 2003.
- [29] Philippe Mulhem and Joo-Hwee Lim. Home photo retrieval: Time matters. In *Int. Conf. on Image and Video Retrieval, Lecture Notes in Computer Science*, Vol. 2728, pp. 308–317, 2003.
- [30] Adrian Graham, Hector Garcia-Molina, Andreas Paepcke, and Terry Winograd. Time as essence for photo browsing through personal digital libraries. In *Proc. 2nd ACM/IEEE-CS Joint Conf. on Digital Libraries*, pp. 326–335, 2002.
- [31] Risto Sarvas, Mikko Viikari, Juha Pesonen, and Hanno Nevanlinna. Mob-share: controlled and immediate sharing of mobile images. In *Proc. 12th annual ACM Int. Conf. on Multimedia*, pp. 724–731, 2004.
- [32] Liu Wenyin, Susan Dumais, Yanfeng Sun, HougJiang Zhang, Mary Czerwinski, and Brent Field. Semi-automatic image annotation. In *Proc. Human-Computer Interaction - Interact'01*, pp. 326–333, 2001.
- [33] Henry Lieberman, Elizabeth Rosenzweig, and Push Singh. Aria: An agent for annotating and retrieving images. *IEEE Computer*, Vol. 34, No. 7, pp. 57–61, 2001.
- [34] Hyunmo Kang and Ben Shneiderman. Visualization methods for personal photo collections: Browsing and searching in the PhotoFinder. In *Proc. IEEE Int. Conf. on Multimedia and Expo (ICME2000)*, pp. 1539–1542, 2000.

- [35] Diomidis D. Spinellis. Position-annotated photographs: A geotemporal web. *IEEE Pervasive Computing*, Vol. 2, No. 2, pp. 72–79, 2003.
- [36] Mor Naaman, Andreas Paepcke, and Hector Garcia-Molina. From where to what: Metadata sharing for digital photographs with geographic coordinates. In *Proc. 10th Int. Conf. on Cooperative Information Systems*, pp. 196–217, 2003.
- [37] 佐藤理史, 佐々木靖弘. ウェブを利用した関連用語の自動収集. 自然言語処理, Vol. 153, No. 8, pp. 57–64, 2003.
- [38] Google Web API. <http://api.google.com/>.
- [39] 松本裕治. 形態素解析システム「茶筌」. 情報処理, Vol. 41, No. 11, pp. 1208–1214, 2000.