

## モバイル環境における写真情報の構造化手法の提案

後藤 孝行<sup>†1,†4</sup> 塚田 浩二<sup>†1,†2</sup>  
濱崎 雅弘<sup>†3</sup> 武田 英明<sup>†4</sup>

写真内容に基づいた閲覧や探索を実現させるためにメタデータ付与を継続的に行うことはユーザにとって大きな負担である。本研究では写真間にリンクという関係性を示す構造的なメタデータを記述することで写真情報を構造化する「リンクに基づく情報構造化」を提案して、このリンク付けを撮影しながら行うことができる「RefleCam」という新しいカメラアプリケーションを試作した。リンクに基づく情報構造化は、写真同士をユーザがリンク付けしていくことで、リンクを利用した関連写真の閲覧や探索を行うことができる。そして、RefleCamは撮影した写真に基づき過去に撮影した写真を提示することで、撮影しながら思い出を振り返ることができ、さらに提示された写真の中からユーザが関係すると思った写真を選択することでリンク付けを行うことができる。評価実験を行った結果、RefleCamは、ユーザにほとんど負担を与えずに継続的なリンク付けを実現していることを確認した。

### Link-based Picture Information Organizing Method in the Mobile Environment

TAKAYUKI GOTO,<sup>†1,†4</sup> KOJI TSUKADA,<sup>†1,†2</sup> MASAHIRO HAMASAKI<sup>†3</sup>  
and HIDEAKI TAKEDA<sup>†4</sup>

A lot of people find it difficult to add metadata to pictures for browsing and searching them based on the content. To solve this problem, we propose the link-based information organizing method. Using this method, users can browse and search related pictures just by linking pictures. Addition to this proposal, we developed a novel camera application called the RefleCam that helps users easily link pictures while taking pictures. In this paper, we describe the concept, development and the experiment of the RefleCam, and introduce features and problems of link-based organizing method based on the results of the experiment.

#### 1. はじめに

近年、カメラ付き携帯電話や小型デジタルカメラなどが普及し、日常的に写真撮影を行うようになったことで、一般ユーザが多く写真の蓄積・共有するようになった。それに伴い、膨大な写真を多様な視点から手軽に探索・閲覧したいという要求が高まっている。例えば、以前撮影したあの猫の写真をさがしたい、いままで撮った食事だけを閲覧したい、といった要求に応えるためには、時間軸だけでなく写真の内容を対象

とした探索が必要であり、これまで多くの画像内容に基づく検索手法が研究されてきた<sup>7)</sup>。しかし、写真の意味解釈の実現は、顔認識に限れば実用化され始めているが(例:iPhoto<sup>☆</sup>)まだまだ研究途上であり、現実的にはタグのようなメタデータを付与するケースが多い。よって、ユーザにできるだけ負担無くメタデータを付与させるための支援が、ヒューマンコンピュータインタラクションの領域で数多く行われてきた<sup>3),4)</sup>。ただし、それらメタデータ付与支援の研究は主にタグのような写真の注釈が主体であり、基本的に文字入力が必要で、特にモバイル環境での継続的な利用は困難であった。一方、写真に継続的にメタデータ付与を行うには撮影直後が効果的という報告があり<sup>1)</sup>、モバイル環境でのメタデータ付与行為は重要である。

そこで、我々はモバイル環境における写真情報を構造化する仕組みとして「リンクに基づく情報構造化手

†1 お茶の水女子大学 お茶大アカデミックプロダクション

Ochanomizu University

†2 科学技術振興機構 さきがけ

JST PRESTO

†3 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology

†4 国立情報学研究所

National Institute of Informatics

☆ <http://www.apple.com/ilife/iphoto/>

法」を提案する。リンクに基づく情報構造化は、ユーザが写真間にリンク付けしていくことで写真情報を構造化していく。リンクは、タグのように写真の内容を直接記述することはできないが、関連写真の閲覧や探索に利用可能である。我々は、このリンク付けをユーザに負担なく継続的に付与させるため、写真撮影行為にリンク付けを統合した「RefleCam(Reflect+Camera)」という新しいカメラを試作した。RefleCamは、撮影した写真情報に基づき過去に撮影した写真を提示して、ユーザがそれらの写真の中から撮影写真と関係すると感じた写真を選択することでリンク付けを行う。これによって、ユーザはほとんど負担を感じずにメタデータ付与行為を継続的に行うことができる。

## 2. RefleCam

RefleCamは、撮影写真の情報に基づいて過去に撮影した写真を検索・提示して、ユーザが任意に関連写真を選ぶことでリンク付けを行えるカメラである(図1)。ここで、通常の撮影行為をほとんど妨げない候補写真の数を限定し、「選択しない」行為を許容することでほとんど負担のないメタデータ付与を実現する。一方、メタデータの付与には構造化という長期的なインセンティブだけでなく、すぐに実感できるようなインセンティブも重要である<sup>6)</sup>。そこで、RefleCamでは、撮影直後に写真提示を行うことで、過去の記憶の想起を促し新しい撮影体験の創出も目指している。例えば、過去に撮影した付近のお店が提示されることで再び同じお店に行ってみるなどのように、撮影行為がユーザの行動を促す可能性がある。さらに、見覚えのあるものを撮影すると過去にも撮影したことに気が付くというように記憶を探索するといった効果を期待している。このように、RefleCamは、撮影直後の思い出の振り返りと超シンプルなメタデータ付与行為によって継続的な構造化体験を目指す。

### 2.1 過去に撮影した写真の提示と選択

RefleCamはスマートフォン<sup>\*</sup>上で写真撮影を行うと、撮影写真の確認画面においてその写真との類似性に基づいて過去に撮影した写真のサムネイルを提示する。提示されたサムネイルを指でなぞっていくと大きなサイズで写真閲覧を次々に行うことができる(図2)。このとき、表示してある写真の撮影時間、場所も提示することで、その写真がどのような状況で撮られたかを思い出す支援も行う。そして、撮影した写真と関連するものがあれば、サムネイルをクリックするこ

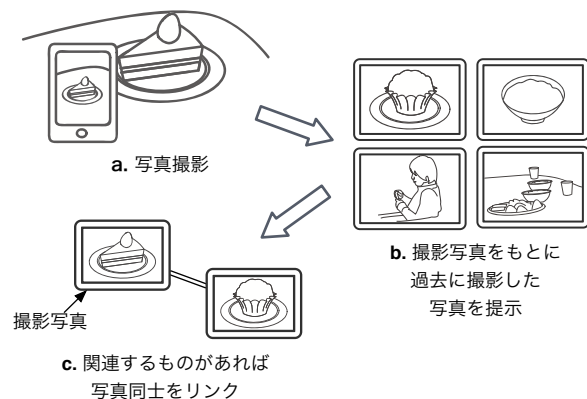


図1 RefleCam のコンセプト

Fig.1 the Concept of the RefleCam

とでリンク付けを行うことができる(図3)。リンクした写真はサムネイルの縁が光ることでリンクしていることを表す。また、写真確認行為を妨げないように提示する関連写真は一目で一覧できる(スクロールしなくてもよい)枚数(6枚)に制限している。

撮影写真の類似性に基づく写真提示のアルゴリズムは、撮影場所、撮影時間、デバイスの姿勢、明るさ、画像の色合い、画像の形状の六つの写真特徴を利用する。そして、各特徴に対して撮影写真と検索対象の写真との差を1から0の値に正規化し、その合計値の高い上位6枚を提示写真にしている。位置情報は、徒歩圏を一つの閾値として考え、距離差0km~3kmを1から0の値に正規化し、3km以上のものは0にした。なお、実験前の試験的な運用において、撮影場所、撮影時間、明るさは値が高くなる傾向があったため、正規化した値に0.5の重みを掛けた。画像特徴を利用して類似画像を検索するにあたって、写真に写っている対象そのものを探す特定物体認識と写っている対象のカテゴリに属するものを探す一般物体認識の二つの方法がある。本研究では、同一カテゴリの写真にリンクが付与できるようにするため、形状特徴の表現に一般物体認識でよく使われているBag-of-keypoints<sup>2)</sup>を用いた。keypointsを構築するのに必要な学習画像には、Caltech 101<sup>\*\*</sup>を利用した。

### 2.2 リンクブラウジング

RefleCamで撮った写真は、一般的な写真閲覧と同じように写真を左へフリックすることで一つ前に撮った写真へ移動し、右へフリックすることで次に撮った写真へ移動するといった、時間軸にそった閲覧が行える。この時、RefleCamは撮影写真とともにリンクし

<sup>\*</sup> iOS アプリケーションとして実装

<sup>\*\*</sup> [http://www.vision.caltech.edu/Image\\_Datasets/Caltech101/](http://www.vision.caltech.edu/Image_Datasets/Caltech101/)

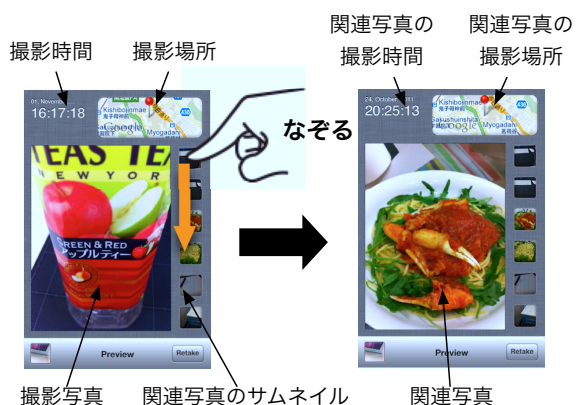


図 2 撮影写真に基づく写真提示

Fig. 2 The RefleCam shows pictures related to the target picture.

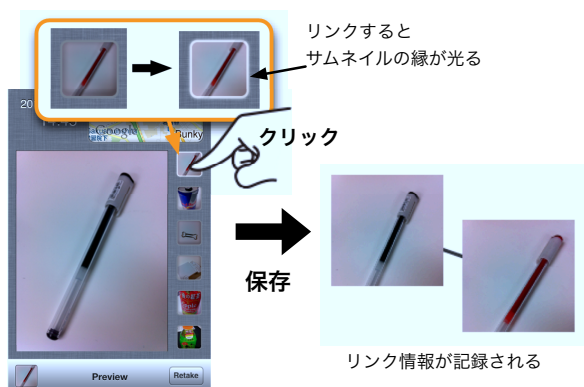


図 3 クリックによるリンク付け

Fig. 3 Users easily link pictures just by clicking.

た写真のサムネイルも表示する。このサムネイルをクリックすることでリンクした写真を見ることができる(図4)。これによって、ユーザの作成した関係性に基づいた写真閲覧が行えるようになる。リンクを通して表示した写真は新しい時間軸に移動しているため、リンクと時間軸を行き来しながら興味ある写真を可逆的に探索することができる。このように、写真間にリンクを付けることで時間軸を飛び越えた新しい閲覧手法を実現できる可能性がある。

### 3. ユーザ評価実験

RefleCam で負担なくリンクが付けられるのか、またどのようなリンクが付けられるのかを調査するため実環境での継続的な評価実験を行っている。

RefleCam はユーザの負担を考慮して関連写真の提示枚数を限定しているため、写真の選択アルゴリズムがリンク付けに大きく影響することが考えられる。そこで、アルゴリズムを(1)ランダム提示、(2)撮影写真の類似性に基づく提示を比較して、提示手法の効果



図 4 リンクブラウジング

Fig. 4 Users can brows related pictures based on the links

を確認する。また、撮影写真の類似性に基づく提示においてどのような写真の特徴がリンク付けに影響を与えているのかを調べることで、ユーザが付けるリンクの性質を明らかにして、今後の設計指針を考察する。

#### 3.1 実験方法

被験者に RefleCam を使って写真撮影とリンク付けを行ってもらい、そのリンク付けの事例をもとにユーザ自身にその意図や、操作の負担をインタビューで説明してもらうことで、構造化の効果を検証する。

ランダム提示と撮影写真の類似性に基づく写真提示の各手法におけるリンク付与数を調べることで、それぞれの効果を確認する。また、この際に、イベント(同じ日に撮影された写真集合)間を繋げるリンクが存在しているかも確認する。現在多くの写真管理アプリケーションは同じ日に撮られた写真を同一イベント(グループ)として扱って写真を管理しているため、この別イベントをつなぐリンクがあることで、従来とは異なった視点による管理が実現できる可能性がある。

撮影写真の類似性に基づく提示を行ったとき、利用している写真特徴のどれがリンク付けに影響を与えているのかを調査する。調査に利用する提示写真において、写真撮影後に提示される6枚の写真集合のどれかが選択されなければその写真集合(6枚)は除外する。これは、提示された6枚のうちどれにも選択されていない(リンクが付けられていない)と、関係性にかかわらずそもそもリンク付けが行える状況でなかった可能性があるためである。

実験の被験者は3名(男性1名、女性2名)、実験期間は約3週間である。なお、提示アルゴリズムは、

ランダム提示を2週間、次に写真の類似性に基づく提示を1週間利用させた。撮影枚数などには、特にノルマを与えず普段通りの撮影を依頼した。

### 3.2 結 果

撮影枚数とリンク数を表1に示す。また、提示写真においてユーザが選択した写真（リンク付けした写真）と、ユーザが選択しなかった写真において、撮影写真との各特徴における類似度の平均を表2に示す。そして、各ユーザの主だったリンク事例を図5, 6, 7に示す。図中に書かれてある写真のつながりの意味は、インタビュー中に被験者自身が回答したものである。

まず、ランダム提示と類似性に基づく提示を比較すると、すべてのユーザにおいて類似性に基づく提示のほうが一枚あたりのリンク数が多かった（表1）。このことから、提示アルゴリズムの効果はあったと考える。インタビューにおいても、アップデート後（類似性に基づく提示）のほうが関係する写真がでてきたという意見があった。リンク付けの負担は、通常のカメラ撮影と比べてとくに負担ではないという意見だった。

次に、提示写真の類似度をユーザが選択した写真と、選択しなかった写真とで比較すると、ユーザAでは、すべての特徴において、リンク付けした写真の方がリンクしてない写真よりも高い値になっている（表2）。ユーザBは、位置情報の類似度は顕著に違うが、他の特徴が目立ってリンク付けに貢献していない。ユーザCも、ユーザBと同じく、特定の特徴が際立ってリンクに寄与しているわけではない。ユーザAのリンク事例を見ると、同じカメラアングルの写真を中心にリンク付けを行っていることが確認できる（図5）。ユーザAは同じ日に何枚も撮影を行っており、同一イベント内で類似する写真を中心にリンク付けを行ったため、すべての特徴において高い類似性を示していたと考えられる。ユーザBのリンク事例をみると、写真をみただけでは関係がわからないリンクのつながりが多く存在した（図6）。また、リンクのつながりによるグループが存在しているが、同一グループでも、リンク付けの視点が異なったものが混在していた。このため、各特徴の類似度が選択した写真と選択しなかった写真で差があまりなかったと考えられる。ユーザCもリンク事例を見ると、写真の特徴に関係なく同じイベントにおいて撮影した写真を中心にリンクを繋げていた（図7）。

### 3.3 考 察

ユーザの写真選択負担を抑えるために関連写真の提示枚数を抑えた設計でも、撮影写真の類似性に基づく提示を行うことで高い頻度でリンク付与が行える。

撮影写真の各特徴の類似度を単純に集計するアルゴリズムでは、一日の撮影枚数が多いと、リンク付けの対象が同一イベント内の写真になる傾向があった。同一イベント内でもリンクが付くことで、タグの共有などにリンクが役立ちそうだが、過去の思い出を振り返るというコンセプトにおいては不適切であった。特定の特徴が選択に寄与するわけでもないのに、利用する写真の特徴を切り替える提示アルゴリズムを利用するのが良いかもしれない。

改善点として、近い時間に撮影した写真は提示させないようにするとか、また、関連写真がない場合は提示写真を入れ替えるといったことが考えられる。

リンク付けの目的によるが提示写真の類似度において、選択した写真、選択しなかった写真において顕著な違いが存在しない場合があり、有益な関係付けを行うには、ユーザ自身の判断が重要である。そして、ユーザがつけた関係は、イベント間をつなげるものも多く閲覧体験を向上させるようなつながりである。

## 4. 関連研究

RefleCamが持つ、ある写真に対して関連する写真を提示する機能は、情報検索の分野ではquery-by-example<sup>9)</sup>と呼ばれ、様々な研究がなされており、近年のスマートフォンの普及によってモバイル環境においても多くの試みがある（例えば、GoogleGoggles<sup>\*</sup>）。本研究では、リンク付けの候補の提示と、撮影写真に基づく写真閲覧を実現するためにこの技術を用いた。

これまで、メタデータ付与を軽減するため多くの研究が行われてきた。このアプローチを大きく分けるとメタデータ付与を支援する方法、コミュニケーションやコンテンツ作成の中でメタデータが自然と付けられる方法がある。例えば、メタデータ付与の支援システムとして、Shneidermanら<sup>4)</sup>はあらかじめデータベースに登録した人名等のラベルを、写真上へDrag&Dropしメタデータを付与するシステムを提案している。また、Sigurbjörnssonら<sup>5)</sup>のようなタグ推薦システムなどもある。コミュニケーションやコンテンツ作成の中で自然にメタデータが付けられる方法として、Liebermanら<sup>3)</sup>は、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し、メール本文の写真の説明を写真のメタデータとして付与を行っている。Watanabeら<sup>8)</sup>は撮影者の注目部分を矢印で指し示した写真を作成することができるカメラを提案している。

メタデータを継続的に付与するには、ユーザの負荷

<sup>\*</sup> <http://www.google.com/mobile/goggles/>

表 1 撮影枚数とリンク数

Table 1 The number of pictures and number of links.

ユーザ	撮影枚数	リンク数	ランダム提示			類似性に基づく提示		
			撮影枚数	リンク数 (別イベント)	リンク数/枚	撮影枚数	リンク数 (別イベント)	リンク数/枚
ユーザ A	113	231	54	10(5)	0.19	59	221(15)	3.75
ユーザ B	35	35	28	25(17)	0.89	7	10(7)	1.43
ユーザ C	37	40	26	18(1)	0.69	11	22(1)	2.00

表 2 提示写真の特徴別類似度 (平均 ± 標準偏差)

Table 2 The similarities of picture (average±SD)

		位置	時間	明るさ	姿勢	色	形状
ユーザ A	選択写真	0.49±0.01	0.48±0.02	0.49±0.01	0.95±0.09	0.72±0.14	0.61±0.13
	非選択写真	0.25±0.21	0.42±0.05	0.46±0.05	0.82±0.13	0.45±0.22	0.39±0.25
ユーザ B	選択写真	0.48±0.01	0.48±0.02	0.45±0.05	0.86±0.12	0.54±0.20	0.50±0.06
	非選択写真	0.23±0.23	0.46±0.03	0.43±0.05	0.82±0.13	0.65±0.24	0.47±0.11
ユーザ C	選択写真	0.5±0.00	0.49±0.02	0.49±0.01	0.90±0.08	0.49±0.11	0.43±0.08
	非選択写真	0.47±0.10	0.38±0.06	0.45±0.03	0.88±0.07	0.49±0.09	0.49±0.07

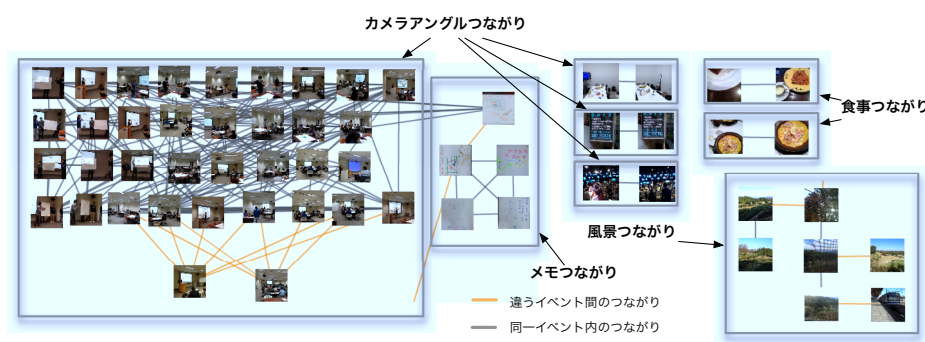


図 5 ユーザ A のリンク事例

Fig. 5 The linking case of user A.

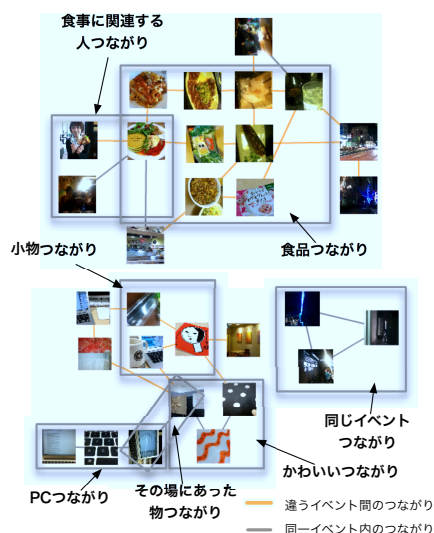


図 6 ユーザ B のリンク事例

Fig. 6 The linking case of user B.

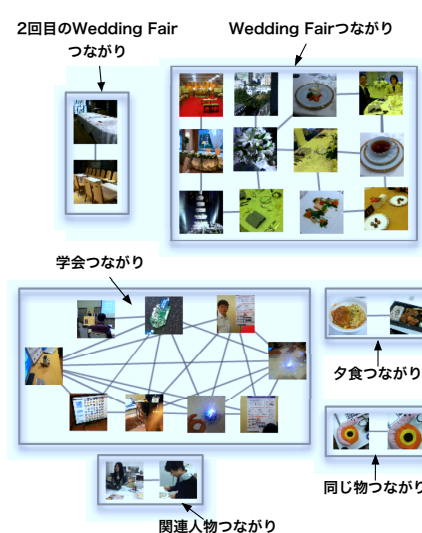


図 7 ユーザ C のリンク事例

Fig. 7 The linking case of user C.

を下げるだけでなく、短期的な効果が得られるような設計が重要である。RefleCam では、過去の写真を撮影直後に提示することで、過去の記憶を振り替わりながら写真の構造化を行う機会を提供する。

## 5. 議 論

リンク構造は Web の特徴の一つでありリンク付けは広く利用されているが、ユーザが日常的に付けるメ

タデータとしてはあまり利用されていない。リンク付けにはリンク先のリソースを選択する必要があり、これを探し出すのは必ずしも容易ではないからだと考える。つまり、リンク付けは、文字入力が必要ない反面、リンク先の選択という点でタグより複雑である。タグ付けは、適切な内容がすぐ思いつかなかつたり、モバイル環境では文字入力がしにくかったりと継続性に難がある。リンク付けはタグにくらべ構造化が限定的ではあるがクリックだけでおこなえることができ、また相対的な判断がおこなえるため継続しやすいと考える。そして、限定的であっても継続的にリンクが付与できると、より表現豊かなスライドショーの実現など、写真閲覧体験を向上させることができると考える。また、写真間にリンクが存在するとタグ付きの写真からタグを取り込むことができるため、限定的な構造化というリンクの欠点がある程度補えるのではと考える。

写真への継続的なメタデータ付与には、撮影直後というタイミングにおいていかにユーザに付与させるかが重要であるが、様々な状況が想定されるモバイル環境では付与手段が一つではなく場合に依じた多様な手段が存在していることが望ましい。それら多様な手段を併用していくことによって継続的な写真情報の構造化が実現すると考える。

## 6. 結 論

本研究では写真間にリンクという関係性を示す構造的なメタデータを記述することで写真情報を構造化する「リンクに基づく情報構造化」を提案した。リンクに基づく情報構造化は、写真同士をユーザがリンク付けしていくことで、リンクを利用した関連写真の閲覧や探索を行うことができる。このリンク付けを撮影しながら行うことができる「RefleCam」というカメラアプリケーションを試作して評価実験を行った結果、ユーザ負担をより軽減するため提示枚数を抑えた提示でも、写真情報を利用することによって多くのリンク付け行為が行われており、ユーザの判断を反映した多様なリンク構造が作られていることを確認した。これにより、リンクを利用した閲覧、探索が実現できると考える。今後は、評価実験で得た課題を反映しつつ、実環境での長期的な運用を通してシステムの効果を検証・改善していきたい。

**謝辞** 本研究の一部は、科学技術振興機構さきがけプログラムの支援を受けた。

## 参 考 文 献

- 1) Ames, M. and Naaman, M.: Why we tag: motivations for annotation in mobile and online media, *Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, USA, ACM, pp.971–980 (2007).
- 2) Csurka, G., Dance, C.R., Fan, L., Willamowski, J. and Bray, C.: Visual categorization with bags of keypoints, *In Workshop on Statistical Learning in Computer Vision, ECCV*, pp.1–22 (2004).
- 3) Lieberman, H., Rosenzweig, E. and Singh, P.: Aria: An Agent For Annotating And Retrieving Images, *IEEE Computer*, Vol.34, No.7, pp. 57–61 (2001).
- 4) Shneiderman, B. and Kang, H.: Direct Annotation: A Drag-and-Drop Strategy for Labeling Photos, *Proceedings of the International Conference on Information Visualisation*, pp.88–95 (2000).
- 5) Sigurbjörnsson, B. and van Zwol, R.: Flickr tag recommendation based on collective knowledge, *WWW '08: Proceeding of the 17th international conference on World Wide Web*, ACM, pp.327–336 (2008).
- 6) Takeda, H. and Ohmukai, I.: Building semantic web applications as information/knowledge sharing systems, *in End User Aspects of the Semantic Web, Colocated with ESWC 2005* (2005).
- 7) Veltkamp, R. C. and Tanase, M.: Content-Based Image Retrieval Systems: A Survey, Technical Report Technical Report UU-CS-2000-34, Department of Computing Science, Utrecht University (2000).
- 8) Watanabe, K., Tsukada, K. and Yasumura, M.: WillCam: a digital camera visualizing users. interest, *CHI '07 extended abstracts on Human factors in computing systems*, ACM, pp.2747–2752 (2007).
- 9) Zloof, M.M.: Query-by-example: the invocation and definition of tables and forms, *Proceedings of the 1st International Conference on Very Large Data Bases*, New York, NY, USA, ACM, pp.1–24 (1975).