

## 色名抽出と色特徴量変換に基づく 典型的画像の Web 検索

Search the Web for Typical Images based  
on Extracting Color-names from the Web  
and Converting them to Color-Features

服部 峻<sup>▼</sup>      田中 克己<sup>◆</sup>

Shun HATTORI      Katsumi TANAKA

近年, Web 上には文書データだけでなく画像データも大量に存在するようになって来ており, これらの情報を有効に活用するため, Web 文書検索だけでなく Web 画像検索に対しても多種多様な要求が生まれて来ている. 対象オブジェクトが画像中に含まれているという条件を満たせば良い単なる適合画像ではなく, 対象オブジェクトの外観的な特徴を把握し易く, また, できるだけ多くの人が画像を見るだけで名称まで想起できるような「典型的画像」を検索したい場合も多い. Google イメージ検索など従来の一般的な Web 画像検索エンジンに対して, 対象オブジェクトの名称をキーワード条件として入力するだけで, そのオブジェクトを画像中に確かに含むような適合画像を検索できる精度は高くなって来てはいるが, 「典型的画像」を精度良く検索することは必ずしもできていない. そこで, 我々は, Web 画像の周辺テキストと対象オブジェクトの名称とのキーワードマッチングだけでなく, テキストマイニング技術により Web から抽出して来た対象オブジェクトの典型的な色名(外観記述)を色特徴量に射影した上で, 検索対象の各 Web 画像に対して, 対象オブジェクトの典型的な色特徴量を含む度合いを評価することによって, 対象オブジェクトの「典型的画像」を Web から検索する手法を提案する.

As the number of Web images as well as Web documents available on the Web has grown exponentially in recent years, Web image searches as well as Web document searches are being pressed more various demands for by their users. We often want to search the Web for not only acceptable images of a target object that simply contain it but also its “typical images” that assuredly represent its typical appearance features and that most people just have to see to recall its name. If we submit its name as a keyword-based query to such a conventional Web image search engine as Google Image Search, we could acquire its acceptable images enough precisely to some extent. However, we can not always acquire its “typical images”. In this paper, we propose a method to search the Web for “typical images” of a target object based on not only matching surrounding text of Web images with its name, but also converting from its Web-extracted appearance descriptions, especially color-names, to image features.

<sup>▼</sup> 学生会員 京都大学大学院 情報学研究科 博士後期課程  
日本学術振興会特別研究員 DC

[hattori@dl.kuis.kyoto-u.ac.jp](mailto:hattori@dl.kuis.kyoto-u.ac.jp)

<sup>◆</sup> 正会員 京都大学大学院 情報学研究科

[tanaka@dl.kuis.kyoto-u.ac.jp](mailto:tanaka@dl.kuis.kyoto-u.ac.jp)

### 1. はじめに

近年, Web 上には文書データだけでなく画像データも大量に存在するようになって来ており, これらの情報を有効に活用するため, Web 文書検索だけでなく Web 画像検索に対しても多種多様な要求が生まれて来ている. 従来の一般的な Web 画像検索エンジンは, 対象オブジェクトの名称が指定されると, そのオブジェクトが画像中に含まれているという条件を満たす単なる適合画像の検索精度は良くなって来てはいる. しかしながら, 対象オブジェクトの外観的な特徴を把握し易く, また, できるだけ多くの人が画像を見るだけで名称まで想起できるような「典型的画像」や, 印象語も指定して, その印象に合った対象オブジェクトの画像を検索したい場合もあるが, このような特別な画像を精度良く検索することは必ずしもできない. 例えば, 「東京駅」の「典型的な」画像を検索しようと, [“典型的な東京駅”]という長いフレーズを検索クエリとして Google イメージ検索[1]に与えると検索結果は 0 件であった. やや緩和して, [“東京駅”AND “典型的な”]という複数のキーワードを連言で検索クエリとして与えると検索結果の件数は十分に多かったが, 赤レンガ造りが顕著な外観である「東京駅」の典型的画像は容易には見当たらなかった. また, 同様に「東京駅」の「ロマンチックな」画像を検索しようとしても上手く行かなかった. これら画像検索への多種多様な要求のうち, 印象語に基づく画像検索に関しては, 予め学習された印象語と画像特徴量との対応関係を利用して画像フィルタリングを行う K-DIME[3]など, 既に多くの研究が行われている[4,5]. 一方で, 「典型的画像」の検索に関しては, あまり言及されていない.

本論文で我々は, 対象オブジェクトと言えどどのような外観をしているのかを把握することが容易であり, また, できるだけ多くの人が画像を見るだけで名称まで想起できるような画像を「典型的画像」と呼び, 単に対象オブジェクトを画像中に含むだけでなく, 対象オブジェクトの典型的な外観的特徴を確かに備えていることをその要件とする. 検索対象の各 Web 画像に対して, この要件を満たすか否か, その度合いを評価するためには, 対象オブジェクトの典型的な外観的特徴を求める必要がある. 対象オブジェクトの名称をキーワード条件として検索し, その適合画像を高精度に含むことが期待される画像集合を解析することによって, 典型的な画像特徴量を求める方法[6]が考えられるが, 対象オブジェクトの適合画像の多様性の問題から, 各 Web 画像において, 対象オブジェクトと多種多様な背景や周辺オブジェクトとを精確に切り分ける必要があり容易ではない. そこで, 我々は, 対象オブジェクトの適合画像が非常に多様であることに比べて, 対象オブジェクトの外観に関する言語的記述がやや非多様であることに注目する. まず, テキストマイニング技術により, 対象オブジェクトに関する大量の Web 文書を解析し, 対象オブジェクトの典型的な色名(外観記述)を抽出する. 次に, 色名を HSV 色空間上の色特徴量に射影することで, 対象オブジェクトの典型的な外観的特徴を得る. 以上により, ユーザが指定した対象オブジェクトの名称, 及び, Web から抽出して来た対象オブジェクトの典型的な色名, それを射影した対象オブジェクトの典型的な色特徴量に基づいて, テキストと内容の両方に基づく画像クエリを生成する. 最後に, 統合クエリのうちテキストに基づく条件に合致した各 Web 画像に対して, 対象オブジェクトの典型的な色特徴量を含む度合いを評価することによってランキングし, 対象オブジェクトの「典型的画像」を Web から検索する手法を提案する.

## 2. 提案手法

本章では、対象オブジェクトの典型的な外観の特徴を備えている「典型的画像」を Web から検索するための手法について提案する。まず、提案手法の基本的な概要について述べた後、提案手法を構成する各機構について述べて行く。

### 2.1 概要

対象オブジェクトの名称が言語的に与えられると、そのオブジェクトの外観的な特徴を把握し易く、また、多くの人が画像を見るだけで名称を想起できるような「典型的画像」を検索結果として返す。提案手法の内部の基本的な概要を示すと図1のようになる。まず、ユーザから与えられた対象オブジェクトの名称に基づいて、テキストマイニング技術により、対象オブジェクトに関する大量の Web 文書を解析し、対象オブジェクトの典型的な外観記述（本論文では色名のみ）を抽出する。次に、言語的な外観記述と画像特徴量との対応関係に基づいて射影することで、対象オブジェクトの典型的な外観特徴量（本論文では HSV 色空間上の色特徴量のみ）を得る。以上により、ユーザによって与えられた対象オブジェクトの名称、及び、Web から抽出して来た対象オブジェクトの典型的な外観記述、それを射影して得られた対象オブジェクトの典型的な外観特徴量を手掛かりに、テキストと内容の両方に基づく画像クエリを統合的に生成する。最後に、統合クエリのうちテキストに基づく条件に合致した各 Web 画像に対して、対象オブジェクトの典型的な外観特徴量を含む度合いを評価することによってランキングし、対象オブジェクトの「典型的画像」を検索結果として返す。

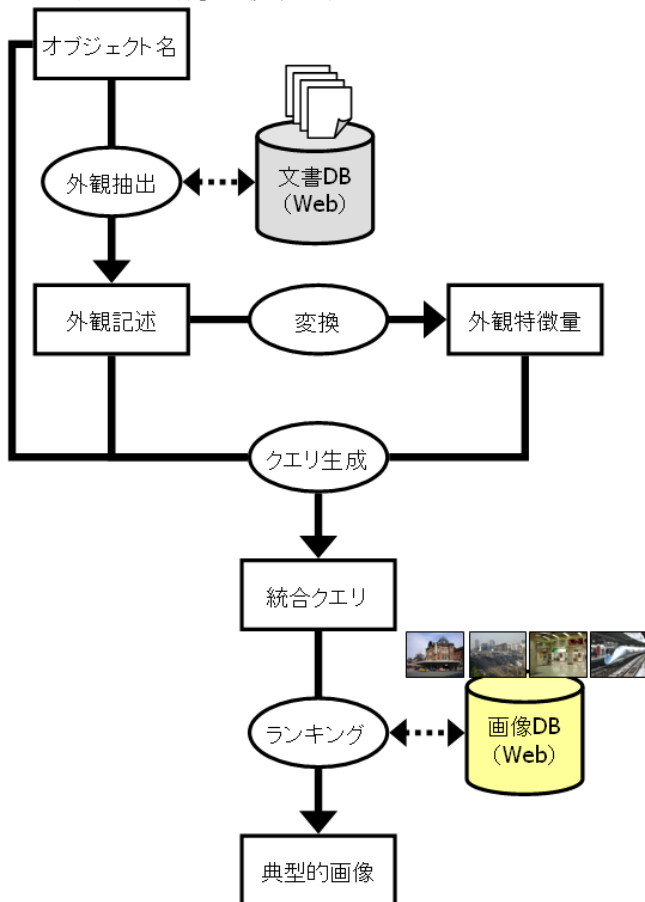


図1 「典型的画像」の検索手法の概要

Fig.1 Overview of our search for “typical images”

### 2.2 オブジェクトの典型的な色名の抽出

我々はこれまで、色名だけでなく、形状や質感など、対象オブジェクトの様々な外観記述を Web から精度良く抽出する手法について研究して来た[7-9]。これらの知見を参考にし、本論文では、ユーザによって与えられた対象オブジェクトの名称に対して、そのオブジェクトの典型的な色名を構文パターンに基づいて抽出する手法を採用する。構文パターンを用いて大量の Web 文書から抽出された各候補  $c_t$  に対して、次式を用いることで、対象オブジェクト  $o_t$  の典型的な色名としての相応しさの度合いを評価する。

$$\text{weight}(c_t, o_t) := (\text{df}("c_t \text{色} の / i o_t") + 1) \cdot (\text{df}("o_t \text{色} / i c_t"))$$

但し、 $\text{df}(q)$ は、検索質問  $q$  を Web 文書検索エンジンで検索した結果の件数を表す。本論文の実験では、Web 文書検索エンジンとして、Google ウェブ検索[2]を用いている。そして、この評価値が他に比べて十分に大きな候補一つを、対象オブジェクトの典型的な色名として採択する。

### 2.3 色名から色特徴量への変換

基本的に必要な機能は、言語的な外観記述から、画像的な外観特徴量へ変換することである。外観記述には、色名だけでなく、形状や表面の材質・質感に関する記述なども含まれる。色名から色特徴量への変換は、JIS 慣用色名とそのマンセル値との対応関係が日本工業規格[9]で規定されており、他種の外観記述から画像特徴量への変換と比べて利用し易い。また、色特徴量は対象とする画像のドメインを限定せず、単純な処理で画像の特徴を表現できる汎用的な手法であるため、外観特徴量として第一に採用することにした。本論文では、日本工業規格で規定されている各色名のマンセル値に基づいて、機械的に HSV 色空間上に射影することによって、各色名を表現する画像的な色特徴量として用いる。

### 2.4 典型的画像の検索クエリ生成

色名抽出と色特徴量変換により、対象オブジェクトの名称、及び、その典型的な色名が言語的な手掛かりとして、また、典型的な色特徴量が画像的な手掛かりとして得られている。対象オブジェクトの「典型的画像」を検索するための検索クエリとして、以下のような様々な候補が考えられる。

1. 対象オブジェクトの名称だけをテキストに基づく条件として、対象オブジェクトの典型的な色特徴量を内容に基づく条件として統合した画像クエリ
2. 対象オブジェクトの名称と典型的な色名とをテキストに基づく条件として、対象オブジェクトの典型的な色特徴量を内容に基づく条件として統合した画像クエリ
3. 対象オブジェクト名と「典型的な」という語をテキストに基づく条件として、対象オブジェクトの典型的な色特徴量を内容に基づく条件として統合した画像クエリ

しかしながら、いずれの検索クエリ候補においても、内容に基づく条件として用いているのは対象オブジェクトの典型的な色特徴量である。従って、各クエリ候補のうちのテキストに基づく条件だけを用いて検索した場合の精度が最も良くなるような検索クエリ候補を採用すべきである。テキストに基づく画像検索である一般的な Web 画像検索エンジンでは、キーワード条件が複数の語から成ると、特に、外観語や印象語を含むと精度が著しく悪化することが観察されるため、本論文では、対象オブジェクトの「典型的画像」を検索するための統合クエリとして1番目の検索クエリ候補を採用する。但し、3章の実験において、2番目の検索クエリ候補との精度比較も行っている。

### 2.5 統合クエリに基づく重み付け

「典型的画像」を検索するための統合クエリとして、テキストに基づく条件  $o_t$  と内容に基づく条件  $c_c$  が与えられた場合に、検索対象の各 Web 画像  $i$  に対して、対象オブジェクトの「典型的画像」としての相応しさの度合いを評価する必要がある。基本的には、テキストに基づく重み付けと内容に基づく重み付けの評価値とを何らかの合成関数で結合する。本論文では、最終的に検索結果として返される全ての画像はテキストに基づく条件に合致し、かつ、重みに関しては差を付けず、その上で、内容に基づく重み付けだけで対象オブジェクトの「典型的画像」としての相応しさの度合いを評価する。つまり、以下の式で定義する。

$$\text{weight}(i, o_t, c_c) := \text{weight}(i, o_t) \cdot \text{weight}(i, c_c)$$

$$\text{weight}(i, o_t) := 1 \quad \text{iff Web 画像 } i \text{ が条件 } o_t \text{ を満たす}$$

但し、Web 画像  $i$  が条件  $o_t$  を満たすとは、Web 画像のファイル名や ALT 属性、周辺テキスト中にキーワード条件  $o_t$  が含まれている場合を表す。本論文の実験では、Web 画像検索エンジンとして Google イメージ検索[1]を用いており、キーワード条件  $o_t$  による検索結果上位  $n$  件の Web 画像に対して、テキストに基づく重み付けの評価値  $\text{weight}(i, o_t)$  として 1 を等しく与え、それ以外の場合には 0 を与える。

最後に、検索対象の各 Web 画像  $i$  に対して、内容に基づく条件である対象オブジェクトの典型的な色特徴量  $c_c$  による重み付けの評価値を次式で定義する。

$$\text{weight}(i, c_c) := \sum_{c \in c_c} \text{sim}(c, c_c) \cdot \text{prop}(c, i)$$

但し、 $\text{sim}(c, c_c)$  は、何らかの色空間における色特徴量  $c$  と  $c_c$  との間の類似度を表す。本論文の実験では、HSV 色空間における色の類似度[11]を用いて算出した上で、その値が 0.8 以下である場合には無視 (0.0 に) している。また、 $\text{prop}(c, i)$  とは、検索対象の各 Web 画像  $i$  において、色特徴量  $c$  が占有する画素面積の割合を表す。

### 3. 実験結果

本章では、Web からの色名抽出とその色特徴量への変換に基づく「典型的画像」の Web 検索手法の精度を検証する。まず、対象オブジェクトの典型的な色名の抽出精度を確認する。次に、抽出された色名から変換した色特徴量に基づいてランキングすることによって、キーワード条件として対象オブジェクトの名称を用いて検索した場合に比べて、典型的画像を検索する精度が改善されるかを確認する。

対象オブジェクトの名称が与えられた場合に、その典型的な色名を構文パターンに基づいて Web から抽出する手法を、10 種類の地物名に適用した結果が表 1 である。7 種では期待通りに抽出できているが、「京都駅」と「京都大学」では抽出できておらず、これら 2 種に関しては提案手法によって典型的画像の検索精度を改善できない。さらに、「東京タワー」では典型的な色名である「赤い」ではなく世界糖尿病デーなど特殊な日だけの色名である「青い」が採択されてしまっており、典型的画像ではなく特殊画像が検索され得る。

次に、対象オブジェクトの名称をテキスト条件として画像検索した結果をベースラインとして、対象オブジェクトの典型的な色特徴量を内容条件として追加した画像クエリ (2 章 4 節の 1 番目) によってランキングした結果の上位  $k$  件の平均適合率を比較すると図 2 のグラフになる。典型的画像よりもノイズ画像を上位数件に引き上げてしまう傾向が見られ、ベースラインよりも典型的画像の検索精度が悪化してし

表 1 オブジェクトの典型的な色名の抽出例

Table 1 Examples of Web-extracted color-names

名称 $o_t$	色名 $c_t$	weight( $c_t, o_t$ )
東京駅	赤い レンガ色	63 (= 7 · 9) 39 (= 3 · 13)
京都駅	抽出できず	
東京タワー	青い 赤い	7843 (=341 · 23) 3045 (=105 · 29)
京都タワー	白い 青い	192 (=32 · 6) 142 (=71 · 2)
東京大学	赤い 青い	15 (=5 · 3) 3 (=1 · 3)
京都大学	抽出できず	
金閣寺	黄金色 黄色	36 (=6 · 6) 6 (=3 · 2)
銀閣寺	茶色 白い	6 (=3 · 2) 3 (=1 · 3)
厳島神社	赤い オレンジ色	125 (=5 · 25) 2 (=1 · 2)
国会議事堂	白い 赤い	40 (=10 · 4) 36 (=2 · 18)

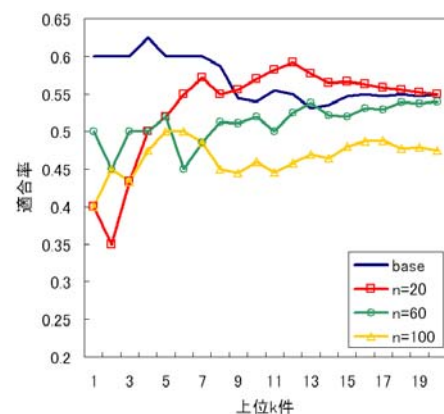


図 2 典型的画像の検索結果の上位  $k$  件の平均適合率 (条件  $o_t$ : オブジェクト名, 条件  $c_c$ : 色特徴量)  
Fig.2 Top  $k$  avg. precision of “typical images” search ( $o_t$ : object-name,  $c_c$ : color-feature)

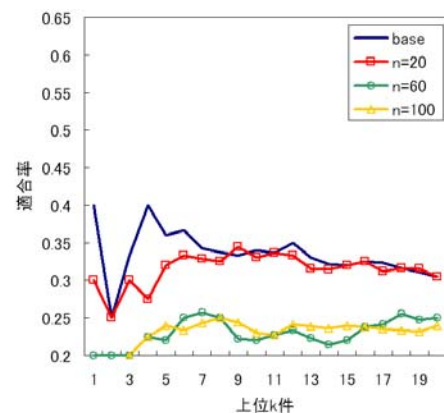


図 3 典型的画像の検索結果の上位  $k$  件の平均適合率 (条件  $o_t$ : オブジェクト名&色名, 条件  $c_c$ : 色特徴量)  
Fig.3 Top  $k$  avg. precision of “typical images” search ( $o_t$ : object-name & color-name,  $c_c$ : color-feature)

表 2 典型的画像の検索の上位 20 件の個々の適合率  
Table 2 Top 20 each precision of "typical images" search

名称 $o_t$	base	n=40	n=60	n=80	n=100
東京駅	0.60	<b>0.75</b>	0.65	0.45	0.30
京都駅	0.70	0.70	0.70	0.70	0.70
東京タワー	<b>0.85</b>	0.70	0.60	0.65	0.75
京都タワー	<b>0.85</b>	0.75	0.70	0.70	0.55
東京大学	0.00	0.00	<b>0.15</b>	<b>0.15</b>	0.10
京都大学	0.05	0.05	0.05	0.05	0.05
金閣寺	<b>0.85</b>	0.75	0.80	0.70	0.70
銀閣寺	<b>0.80</b>	0.65	0.45	0.50	0.45
厳島神社	0.55	0.70	<b>0.80</b>	0.65	0.70
国会議事堂	0.25	0.40	0.50	<b>0.60</b>	0.45

まっている。しかしながら、上位 20 件での個々の適合率を比較すると表 2 のようになり、「東京タワー」や「金閣寺」など適合率が元々高い地物名に対しては提案手法で改善できていないが、「東京大学」や「国会議事堂」など適合率が元々低い地物名に対しては提案手法で改善できている。対象オブジェクト毎に最適な  $n$  を求めることができれば、平均適合率においてもベースラインより改善できる可能性がある。

また、対象オブジェクトの名称と典型的な色名とをテキスト条件として画像検索した結果をベースラインとして、対象オブジェクトの典型的な色特徴量を内容条件として追加した画像クエリ (2 章 4 節の 2 番目) によってリランキングした結果の上位  $k$  件の平均適合率を比較すると図 3 のグラフになる。テキスト条件として対象オブジェクトの名称だけでなく典型的な色名も追加してしまうと典型的画像の検索精度が悪化することが確認できる。

#### 4. まとめと今後の課題

本論文で我々は、対象オブジェクトの典型的な外観的特徴を確かに備えていることを「典型的画像」の要件とし、対象オブジェクトの名称がユーザから与えられた場合に、テキストマイニング技術により Web から抽出して来た対象オブジェクトの典型的な色名を HSV 色空間上の色特徴量に変換した上で、検索対象の各 Web 画像に対して、対象オブジェクトの典型的な色特徴量を含む度合いを評価することによって、対象オブジェクトが画像中に含まれているだけの単なる適合画像ではなく、対象オブジェクトの外観的特徴を把握し易い「典型的画像」を Web から検索する手法を提案した。

10 種類の地物名に対して提案手法を適用した実験の結果、平均適合率においてはベースラインを上回ることはできなかった。しかし、個々の適合率を見ると、適合率が元々高い地物名に対してはやや下がるものの高いレベルを維持し、適合率が元々低い地物名に対しては大きく改善する場合もあり、より多くの地物名に対して高精度を実現できたと言える。

人物名などのように、対象オブジェクトの名称に対して、典型的な色名が Web から抽出できない(典型的な色名が元々存在しない) 場合には、典型的な色特徴量による重み付けを行うことはできないが、色名以外の外観記述や印象語も Web から抽出することによって、「典型的画像」を精度良く検索できる対象ドメインを広げることができると考えている。

また、典型的な色名が抽出できなかった「京都大学」における「時計台」、典型的な色名は抽出できたが色特徴量による重み付けだけでは精度が悪くなかった「東京大学」における「赤門」「安田講堂」など、対象オブジェクトの外観記述だけでなく、構成要素名などの併用も検討して行く。

#### 【謝辞】

本研究の一部は、科学研究費補助金特別研究員奨励費「モバイル・ユビキタス環境における空間情報アクセスに関する研究」(研究代表者：服部峻，課題番号：1955301)，及び、京都大学グローバル COE プログラム「知識循環社会のための情報学教育研究拠点」(研究代表者：田中克己) の助成を受けたものである。ここに記して謝意を表す。

#### 【文献】

- [1] Google イメージ検索: <http://images.google.co.jp/> (2007).
- [2] Google ウェブ検索: <http://www.google.co.jp/> (2007).
- [3] Inder, R., Bianchi-Berthouze, N., and Kato, T.: "K-DIME: A Software Framework for Kansei Filtering of Internet Material," Proceedings of IEEE International Conference on Systems, Man and Cybernetics (SMC'99), Vol.6, pp.241-246 (1999).
- [4] 栗田多喜夫, 加藤俊一, 福田郁美, 坂倉あゆみ: "印象語による画像データベースの検索," 情報処理学会論文誌, Vol.33, No.11, pp.1373-1383 (1992).
- [5] 木本晴夫: "感性語による画像検索とその精度評価," 情報処理学会論文誌, Vol.40, No.3, pp.886-898 (1999).
- [6] 柳井啓司: "キーワードと画像特徴を利用した WWW からの画像収集システム," 情報処理学会論文誌 (トランザクション) データベース, Vol.42, No.SIG10(TOD11), pp.79-91 (2001).
- [7] 服部峻, 手塚太郎, 田中克己: "オブジェクトの外観情報の Web マイニング," 第 18 回データ工学ワークショップ (DEWS'07) 論文集, L4-6 (2007).
- [8] 服部峻, 手塚太郎, 田中克己: "文書中の地物画像を言語的記述で代替するための地物の外観情報の Web からの抽出," 情報処理学会論文誌 (トランザクション) データベース, Vol.48, No.SIG11(TOD34), pp.69-82 (2007).
- [9] Hattori, S., Tezuka, T, and Tanaka, K.: "Mining the Web for Appearance Description," Proceedings of the 18th International Conference on Database and Expert Systems Applications (DEXA'07), LNCS Vol.4653, pp.790-800 (2007).
- [10] 日本工業規格: 物体色の色名, JIS Z 8102:2001 (2001).
- [11] Smith, J. R. and Chang, S.-F.: "VisualSEEK: A Fully Automated Content-Based Image Query System," Proceedings of the 4th ACM International Conference on Multimedia (ACM Multimedia'96), pp.87-98 (1996).

#### 服部 峻 Shun HATTORI

京都大学大学院情報学研究科社会情報学専攻博士後期課程在学中、及び、日本学術振興会特別研究員 DC。2006 年京都大学大学院情報学研究科社会情報学専攻修士課程修了。主にモバイル・ユビキタス環境における社会基盤技術の研究に従事。情報処理学会、電子情報通信学会、日本データベース学会各学生会員。

#### 田中 克己 Katsumi TANAKA

京都大学大学院情報学研究科社会情報学専攻教授。1976 年京都大学大学院修士課程修了。博士 (工学)。主にデータベース、マルチメディアコンテンツ処理の研究に従事。IEEE Computer Society, ACM, 人工知能学会, 日本ソフトウェア科学会, 情報処理学会, 日本データベース学会等各会員。