

# 画像内オブジェクトの相対的特徴に基づく類似検索手法

Similar Image Retrieval Method based on Relative Features of Image Objects

辻 健太<sup>▼</sup>

Kenta TSUJI

川越 恭二<sup>◆</sup>

Kyoji KAWAGOE

本稿では、画像内オブジェクトの相対的特徴に基づく類似画像検索手法の提案を行う。現在の画像内容型検索では、問合せ画像から抽出した特徴を用いて類似画像検索を行っている。このとき、利用者は他の閲覧画像と比較して特定の問合せ画像を選択する。しかし既存の類似画像検索では、問合せ画像と閲覧画像における画像内オブジェクトの相対的特徴を考慮していないため、問合せ画像に特徴量の小さい画像内オブジェクトが含まれていた場合、その画像内オブジェクトの重要度が低くなるという問題がある。そこで本研究では、画像にあらかじめ付加されているアノテーション情報から抽出した画像内オブジェクトに対し、文書検索の際に用いられるTF-IDF法を適用することで、問合せ画像と閲覧画像における画像内オブジェクトの相対的特徴を抽出する。このことにより、問合せ画像と閲覧画像における画像内オブジェクトの相対的特徴を考慮した類似画像検索が可能となる。

In this paper, we propose a similar image retrieval method based on image object features. Although many content-based image retrieval methods have been proposed so far, they don't consider the relative features contained in a query image and one of them is small, the importance of such small sub-image object becomes low with these existing methods. In order to solve the problem, we introduce new features for sub-image objects, called relative features, which are used to calculate similarity between a query and browsing images with use of TF-IDF. The relative features of an image are defined as a relative area of an individual sub-image inside the image, extracted from annotation information attached to the image beforehand. From our some experiments, we can confirm that our proposed method has better accuracy than the baseline method.

## 1. はじめに

現在、Web 上に存在する膨大な画像から適合画像を検索するための研究が数多く行われている[1][2][3]。画像検索手法には、画像と共に起る文書情報やタグ情報などのキーワードによる画像検索 TBIR(Text-based Image Retrieval)と、画像の内容に基づいて画像検索を行う画像内容型検索

CBIR(Content-based Image Retrieval)がある。この2種類の画像検索手法で、後者の画像内容型検索 CBIR には、画像内オブジェクトに着目した研究がある[4][5]。しかし、この類似画像検索手法では、問合せ画像と閲覧画像における画像内オブジェクトの相対的特徴を考慮しておらず、問合せ画像の特徴のみを考慮しているにすぎない。そのため、例えば問合せ画像に特徴量の小さい画像内オブジェクトが含まれていた場合、その画像内オブジェクトの重要度が低くなるという問題がある。仮に利用者が CBIR を用いた類似画像検索システムに、特徴量の大きい画像内オブジェクトAと特徴量の小さい画像内オブジェクトBが含まれている画像を問合せ画像として入力したとする。このとき、システムは入力された問合せ画像において画像内オブジェクトBの特徴量が画像内オブジェクトAの特徴量よりも小さいため、画像内オブジェクトBよりも画像内オブジェクトAの重要度が高いと処理することが多い。そのため、利用者は、画像内オブジェクトAの特徴をより考慮した類似画像をシステムから検索結果として得ることとなる。しかし、利用者が複数の画像を閲覧し、これらの閲覧画像の中から問合せ画像を選択している場合には、利用者にとって必ずしも画像内オブジェクトAの重要度が高いとは限らない。例えば、閲覧画像集合内の問合せ画像以外の画像すべてに画像内オブジェクトAが含まれているが、画像内オブジェクトBが含まれていない場合には、利用者の入力した問合せ画像において画像内オブジェクトBの重要度の方が高いことも多い。つまり、画像内オブジェクトに着目した類似画像検索において、問合せ画像と閲覧画像に含まれる画像内オブジェクトの相対的特徴を考慮する必要があるといえる。

そこで本研究では、画像内オブジェクトの相対的特徴を用いた類似画像検索手法を提案する。本提案手法では、まず画像にあらかじめ付加されているアノテーション情報から画像内オブジェクトを抽出する。抽出した画像内オブジェクトに対して、文書検索の際に用いられるTF-IDF法を適用することで、問合せ画像と閲覧画像における画像内オブジェクトの相対的特徴を抽出する。具体的には、問合せ画像と閲覧画像における画像内オブジェクトから二つの相対的特徴を抽出する。一つ目の相対的特徴は、問合せ画像や閲覧画像における画像内オブジェクトの特徴量の差である。問合せ画像における画像内オブジェクトの特徴量が閲覧画像集合におけるその平均値と差が大きい場合、その画像内オブジェクトは問合せ画像において特徴的であると考えられる。そこで、問合せ画像における画像内オブジェクトの面積と閲覧画像集合におけるその平均値との差に、文書検索の際のTF値を適用する。もう一つの相対的特徴は、問合せ画像や閲覧画像における画像内オブジェクトの出現頻度である。閲覧画像集合において出現頻度の低い画像内オブジェクトが問合せ画像に含まれていた場合、その画像内オブジェクトは問合せ画像において特徴的であると考えられる。また閲覧画像集合において出現頻度の高い画像内オブジェクトが問合せ画像に含まれていなかった場合、その画像内オブジェクトは問合せ画像において特徴的であると考えられる。そこで、問合せ画像と閲覧画像における画像内オブジェクトの出現頻度の逆数に、文書検索の際のIDF値を適用する。このように算出したTF値とIDF値を用いて画像の特徴ベクトルを生成することで、問合せ画像と閲覧画像における画像内オブジェクトの相対的特徴を考慮した類似画像検索が可能となる。

<sup>▼</sup> 学生会員 立命館大学大学院理工学研究科博士前期課程 [ktsuji@coms.ics.ritsumei.ac.jp](mailto:ktsuji@coms.ics.ritsumei.ac.jp)

<sup>◆</sup> 正会員 立命館大学情報理工学部 [kawagoe@is.ritsumei.ac.jp](mailto:kawagoe@is.ritsumei.ac.jp)

## 2. 相対的特徴を用いた類似画像検索

### 2.1 基本的な考え方

従来の画像内容型検索CBIRは、問合せ画像に特徴量の小さい画像内オブジェクトが含まれていた場合、その画像内オブジェクトの重要度が低くなる。図1と図2に示す画像を用いて具体的に説明する。例えば、図1に示す画像を問合せ画像として、利用者が類似画像検索を行ったとする。このとき、問合せ画像内に特徴的なオブジェクトとして「金閣寺」や「紅葉」が含まれている。従来の類似画像検索システムの場合、「紅葉」の画像内オブジェクトの特徴量よりも「金閣寺」の画像内オブジェクトの特徴量が大きいため、システムは「紅葉」の重要度よりも「金閣寺」の重要度が高いものとして処理される。しかし、利用者が図2に示す4枚の画像を閲覧画像集合として、それらの閲覧画像と比較して問合せ画像を選択した場合、「金閣寺」の重要度よりも「紅葉」の重要度の方が高いと考えることができる。つまり、画像内オブジェクトに着目した類似画像検索において、問合せ画像と閲覧画像における画像内オブジェクトの相対的特徴を考慮する必要がある。

先に述べたように、本研究では、画像内オブジェクトの相対的特徴を二つ定義する。一つ目の相対的特徴は、問合せ画像や閲覧画像における画像内オブジェクトの特徴量の差である。問合せ画像における画像内オブジェクトの特徴量が閲覧画像集合におけるその平均値と差が大きい場合、その画像内オブジェクトは問合せ画像において特徴的であると考えられる。二つ目の相対的特徴は、問合せ画像や閲覧画像における画像内オブジェクトの出現頻度である。閲覧画像集合において出現頻度の低い画像内オブジェクトが問合せ画像に含まれていた場合、その画像内オブジェクトは問合せ画像において特徴的であると考えられる。また、閲覧画像集合において出現頻度の高い画像内オブジェクトが問合せ画像に含まれていない場合、その画像内オブジェクトは問合せ画像において特徴的であると考えられる。そこで、一つ目の相対的特徴である問合せ画像や閲覧画像における画像内オブジェクトの特徴量の差をTF値、二つ目の相対的特徴である問合せ画像や閲覧画像における画像内オブジェクトの出現頻度の逆数をIDF値とし、文書検索の際に用いられるTF-IDF法を適用する。ここで、TF-IDF法とは、単語の出現頻度をTF値とし、総文書に含まれる単語の出現頻度の逆数をIDF値とすることで、文書中の特徴的な単語を抽出するためのアルゴリズムである。つまり、ある文書中で出現回数の多い単語や特定の文書にしか出現しない単語の重要度を高くする。本研究にTF-IDF法を適用することによって、閲覧画像集合から利用者が選択した問合せ画像における画像内オブジェクトの相対的特徴を抽出できると考える。

以上述べたように本提案手法は、文書が単語から構成されていることに対し、画像が画像内オブジェクトから構成されているといった対応関係に着目したことに基づいている。文書検索の際のTF-IDF法では、各単語を次元としたベクトル空間上に文書を表現するのが一般的であるが、本提案手法では、画像内オブジェクトのタグ情報を次元としたベクトル空間上に画像を表現する。

### 2.2 画像内オブジェクトの相対的特徴抽出

画像  $p$  にタグ  $t$  で示される画像内オブジェクトが複数存在するとする。このとき、画像  $p$  におけるタグ  $t$  で示される画像内オブジェクトの特徴量を  $f(t, p)$  とする。また、閲覧画



図1 問合せ画像

Fig.1 Query image



図2 閲覧画像集合

Fig.2 Browsing images

像集合におけるタグ  $t$  で示される画像内オブジェクトの特徴量の平均値を  $ave(t)$  とする。  $f(t, p)$  と  $ave(t)$  の差が大きいほど、画像におけるタグ  $t$  の画像内オブジェクトが特徴的であるとする。そこでTF-IDF法におけるTF値の考え方を適用し、  $f(t, p)$  と  $ave(t)$  の差を画像内オブジェクトの局所的な重要度とする。画像  $p$  におけるタグ  $t$  で示される画像内オブジェクトの局所的な重要度  $tf(t, p)$  を(1)式で算出する。

$$tf(t, p) = f(t, p) - ave(t) \quad (1)$$

閲覧画像集合において出現頻度の低い画像内オブジェクトが問合せ画像に含まれている場合、その画像内オブジェクトは問合せ画像において特徴的であるとする。また、閲覧画像集合において出現頻度の高い画像内オブジェクトが問合せ画像に含まれていない場合も、同様のことがいえる。そこで、TF-IDF法におけるIDF値の考え方を適用し、閲覧画像集合における画像内オブジェクトの出現頻度の逆数をその画



図3 システム概要

Fig.3 Outview of our image retrieval system



図5 アノテーション情報

Fig.5 Image annotation

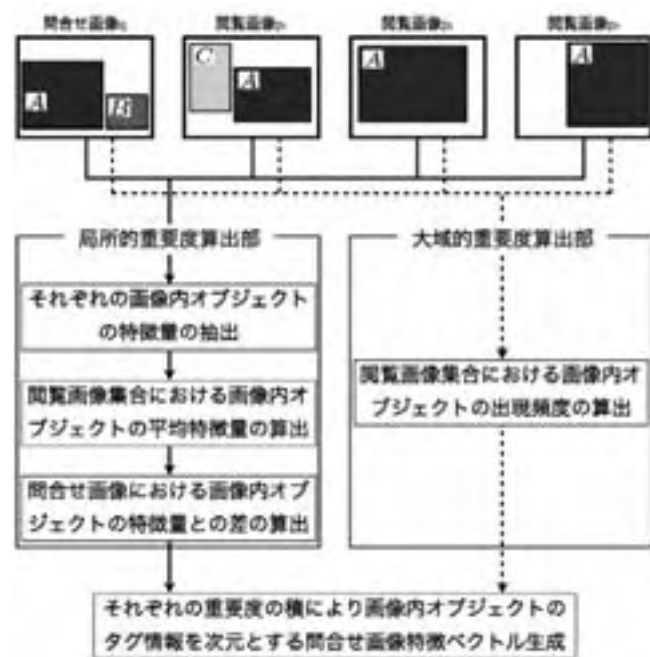


図4 システム詳細

Fig.4 Detail of our image retrieval system

像内オブジェクトの大域的な重要度とする。閲覧画像中にタグ  $t$  で示される画像内オブジェクトが1回以上生起する画像数とする。閲覧画像数を  $N$  とし、タグ  $t$  で示される画像内オブジェクトの大域的な重要度  $idf(t)$  を(2)式で算出する。ここで、問合せ画像中にタグ  $t$  で示される画像内オブジェクトが含まれている場合は、閲覧画像集合における画像内オブジェクトの出現頻度の逆数を用いる。このことにより、閲覧画像集合において出現頻度の低い画像内オブジェクトが問合せ画像に含まれていた場合、その画像内オブジェクトの重要度を高くすることが可能となる。また、問合せ画像中にタグ  $t$  で示される画像内オブジェクトが含まれていない場合は、閲覧画像集合における画像内オブジェクトの非出現頻度の逆数を用いる。このことにより、閲覧画像集合において出現頻度の高い画像内オブジェクトが問合せ画像に含まれていない場合、その画像内オブジェクトの重要度を高くすることが可能となる。

$$idf(t) = \begin{cases} \log(N/f_t) + 1 & \langle S(t, p) \rangle > 0 \\ \log(N/N - f_t) + 1 & \langle S(t, p) \rangle = 0 \end{cases} \quad (2)$$

(1)式、(2)式で算出されるそれぞれの重要度から、画像  $p$  におけるタグ  $t$  で示される画像内オブジェクトの重要度  $w(t, p)$  を(3)式で算出する。

$$w(t, p) = tf(t, p) \cdot idf(t) \quad (3)$$

(3)式で算出した値を要素とし、各画像内オブジェクトのタグ情報を次元とした画像の特徴ベクトルを生成する。本研究では、ベクトル間のコサイン尺度を算出することによって、画像間の類似度を算出する。

### 3. 画像内オブジェクトの相対的特徴に基づく類似画像検索システム

#### 3.1 システム概要

図3に本稿で提案した画像内オブジェクトの相対的特徴を用いた類似画像検索システムの概要図を示す。

利用者はまず、システムに問合せ画像と閲覧画像を入力する。システムはその画像集合に含まれる画像内オブジェクトを抽出する(STEP1)。システムは利用者の問合せ画像と閲覧画像から、問合せ画像の相対的特徴を抽出し、特徴ベクトルを算出する(STEP2)。システムは問合せ画像の特徴ベクトルを用いて画像データベース内の画像との類似度を算出し、問合せ画像と類似する画像を利用者に出力する(STEP3)。

STEP2のより詳しい処理の流れを具体的に図4に示す。局所的な重要度算出部でまず、STEP1で抽出した画像内オブジェクトの特徴量を抽出する。そして閲覧画像集合における画像内オブジェクトの特徴量の平均値を算出し、問合せ画像における画像内オブジェクトの特徴量との差を、局所的な重要度とする。次に大域的な重要度算出部で閲覧画像集合における画像内オブジェクトの出現頻度を算出し、大域的な重要度とする。最後にそれぞれの重要度の積を算出することによって、画像



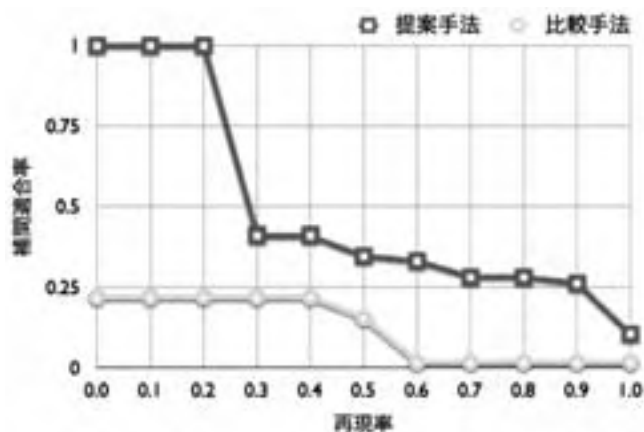


図6 予備実験における再現率適合率曲線

Fig.6 Precision recall curve(1)

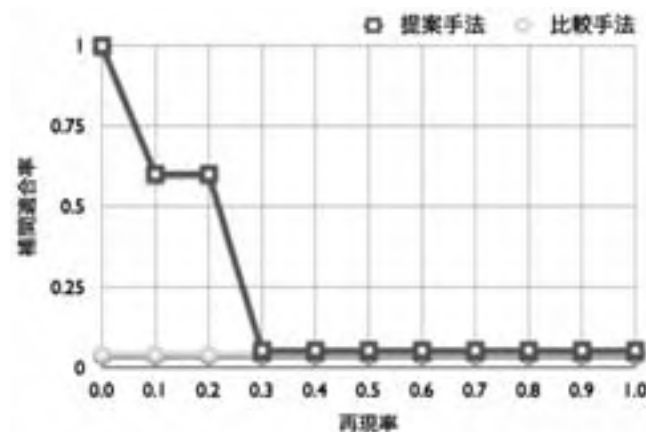


図7 評価実験における再現率適合率曲線

Fig.7 Precision recall curve(2)

内オブジェクトのタグ情報を次元とする問合せ画像の特徴ベクトルを生成する。

### 3.2 アノテーション情報を用いた画像内オブジェクト抽出

画像内オブジェクトに着目した類似画像検索では、画像内オブジェクトの抽出精度が類似画像検索の検索結果に大きな影響を与える。つまり、画像内オブジェクトに着目した類似画像検索において、人が画像を閲覧したときに特徴的であると判断した画像内オブジェクトを抽出することが必要である。

LabelMe[6]では、画像内オブジェクトに対し人手で多角形領域とタグ情報を入力することによって、画像内の特徴的なオブジェクトを抽出することを可能にしている。そこで本研究では、それぞれの画像には図5に示すようなアノテーション情報が登録されていると仮定し、このアノテーション情報を用いて画像内オブジェクトを抽出する。具体的には、域とその画像内オブジェクトのタグ情報で構成されている情報を含んでいるとする。利用者があらかじめ画像内の特徴的なオブジェクトに対して付加したアノテーション情報を用いることによって、画像内の特徴的なオブジェクトを抽出することが可能となる。

## 4. 予備実験

本稿で提案した画像内オブジェクトの相対的特徴を類似画像検索に適用することによる検索結果の変化を確認するため、まず予備実験を行った。予備実験における画像内オブジェクトの特徴量抽出には、画像内オブジェクトの面積値を用いた。つまり、面積の大きい画像内オブジェクトの特徴量が大きく、面積の小さい画像内オブジェクトの特徴量が小さいとする。予備実験では、問合せ画像は特徴量の小さい画像内オブジェクトを含んでおり、閲覧画像集合においてその出現頻度が低い場合を想定する。閲覧画像は、5種類の画像内オブジェクトから2もしくは2種類含む画像6件とする。データベース内の検索対象となる画像は、5種類の画像内オブジェクトを1から3種類ランダムに生起させた画像100,000件とする。問合せ画像に含まれる画像内オブジェクトの特徴量から生成したベクトルを用いてコサイン尺度によりラン

キングした検索結果上位100件と、その検索結果に対して提案手法を用いてリランキングした結果を比較する。それぞれの手法における検索結果に対して、補間適合率を用いた再現率適合率曲線によって、提案手法の有効性を確認する。図6に提案手法と比較手法における再現率適合率曲線を示す。図6に示すように、提案手法を用いてランキングされた結果の方が上位に正解画像を多く含んでいることがわかる。これは提案手法における画像内オブジェクトの相対的特徴を用いることにより、特徴量の小さい画像内オブジェクトの重要度が高くなったためであると考えられる。このことより、提案手法を用いて類似画像検索を行うことによって、正解画像がより上位に出力されることを確認できた。

## 5. 評価実験

### 5.1 実験方法

本稿で提案した画像内オブジェクトの相対的特徴を類似画像検索に適用することによる有効性を確認するため、まず評価実験を行った。評価実験における画像内オブジェクトの特徴量抽出には、予備実験と同様に、画像内オブジェクトの面積を用いた。問合せ画像は図1に示す画像を用いた。閲覧画像は図2に示す画像4件とする。データベース内の検索対象となる画像は、画像内オブジェクトとして金閣寺を含む画像291件、紅葉を含む画像32件(内、どちらも含む画像13件)から構成される画像311件とする。あらかじめ全ての画像内の特徴的なオブジェクトに対してアノテーション情報を付加している。

### 5.2 実験結果

それぞれの手法における検索結果に対して、補間適合率を用いた再現率適合率曲線によって、提案手法の有効性を確認する。図7に提案手法と比較手法における再現率適合率曲線を示す。図7に示すように、提案手法を用いてランキングされた結果の方が上位に正解画像を多く含んでいることがわかる。ここで、それぞれの手法における検索結果をそれぞれ図8と図9、図10と図11に示す。図8は、提案手法を用いてランキングされた検索結果1件目に出力された画像で、図9は、検索結果2件目に出力された画像である。図10は、比較手法を用いてランキングされた検索結果1件目に出力された画像で、図11は、検索結果2件目に出力された画像





図8 提案手法検索結果1件目

Fig.8 Top placement in search results with our proposed method



図9 提案手法検索結果2件目

Fig.8 Second placement in search results with our proposed method

である。提案手法を用いてランキングされたそれぞれの検索結果の画像には、「金閣寺」の画像内オブジェクトだけでなく「紅葉」の画像内オブジェクトも含んでいるが、比較手法を用いてランキングされたそれぞれの検索結果の画像には、「金閣寺」の画像内オブジェクトだけしか含んでいない。これは提案手法における画像内オブジェクトの相対的特徴を用いることにより、問合せ画像に含まれていた「紅葉」の画像内オブジェクトの重要度が高くなったためであると考えられる。つまり、問合せ画像の特徴量のみを考慮して類似画像検索を行った場合、特徴量の小さい「紅葉」の画像内オブジェクトは重要度が低いと処理されていたが、問合せ画像と閲覧画像における「紅葉」の画像内オブジェクトの相対的特徴を考慮することによって、「紅葉」の画像内オブジェクトの重要度を高くすることが可能となった。このことより、類似画像検索に本研究で提案した画像内オブジェクトの相対的特徴を適用することによる有効性を確認することができた。

しかし、提案手法を用いてランキングされた検索結果の下



図10 比較手法検索結果1件目

Fig.10 Top placement in search results with baseline method



図11 比較手法検索結果2件目

Fig.11 Second placement in search results with baseline method

位にも正解画像が含まれていた。これは、矩形領域によってアノテーション情報を付加することが困難な画像内オブジェクトが画像に含まれていた場合、画像内オブジェクトの抽出精度が悪いたと考えられる。そのため、今後は、画像内オブジェクトの抽出精度を上げる必要がある。

## 6. 関連研究

本研究に関連する研究として、画像の類似検索に情報検索における相対的問合せ表現[7]を適用した研究がある。中島ら[8]は、“利用者が正事例として選択した画像”と“利用者が閲覧した画像の中で正事例と類似する画像”との差を利用することによって、利用者の検索意図をよりの確に反映することを可能にする差異増幅型適合フィードバックを提案している。問合せ画像と閲覧画像の差を類似画像検索に利用する点で、本研究と類似しているが、画像内オブジェクトの出現頻度を考慮している点で異なる。

## 7. おわりに

本稿では、画像から抽出した画像内オブジェクトに対して、文書検索の際に用いられる TF-IDF 法を適用することによって、問合せ画像と閲覧画像の相対的特徴に基づく類似検索手法を提案した。評価実験結果より、画像内オブジェクトの相対的特徴を類似画像検索に適用することの有効性を示した。本提案手法により、画像内オブジェクトに着目して類似画像検索を行う利用者が、他の閲覧画像と比較して問合せ画像を選択した場合に、従来手法より適合画像を得やすくなると考えられる。

今後は、人手でアノテーション情報を付加するのではなく、自動アノテーション手法[9]や Visual Phrase[10]を本研究に適用することを検討する。また、画像内オブジェクトの抽出精度を上げるために、画像の特徴量抽出の際に、SIFT(Scale-Invariant Feature Transform)特徴量[11][12]を用いることも検討する。さらに、画像適合フィードバックとの比較実験を行う必要がある。

## [文献]

- [1] Y. Jing and S. Baluja : “PageRank for Product Image Search”, WWW2008/Refereed Track: Rich Media. pp. 307-315 (2008).
- [2] Y. Jing and S. Baluja : “VisualRank Applying PageRank to Large-Scale Image Search”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.30, No.11, pp. 187-1890 (2008).
- [3] M. Ambai and Y. Yoshida : “Multiclass VisualRank: VisualRanking Method in Clustered Subsets Based on Visual Features”, Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp. 732-733 (2009).
- [4] Q. F. Zheng, W. Q. Wang and W. Gao : “Effective and Efficient Object-based Image Retrieval Using Visual Phrases”, Proceedings of the 14th annual ACM international conference on Multimedia. pp. 77-80 (2006).
- [5] 森拓矢, 松本哲也, 大西昇 : “画像内のオブジェクトに注目した画像内容検索”, 情報処理学会研究報告. CVIM, Vol.18, pp. 167-174 (2005).
- [6] B. Russell, A. Torralba, K. Murphy and W. Freeman : “LabelMe: a database and web-based tool for image annotation”, International Journal of Computer Vision, Vol.77, No.1, pp. 157-173 (2008).
- [7] 中島伸介, 田中克己 : “差異増幅型適合フィードバックに基づく画像データベース検索”, 電子情報通信学会論文誌 D-1, 情報システム, I-情報処理, Vol.87, No.2, pp. 164-174 (2004).
- [8] 中島伸介, 木下真一, 田中克己 : “相対的マッピング処理に基づく相対的情報検索手法”, 情報処理学会論文誌, データベース, Vol.45, No.4, pp. 63-75 (2004).
- [9] 吉田悠一, 安倍満 : “領域間で共通する特徴点の数を利用した代表画像への自動アノテーション手法”, 電子情報通信学会技術研究報告. PRMU, Vol.108, No.363, pp. 171-176 (2008).
- [10] S. Zhang , Q. Tian, G. Hua, Q. Huang and S. Li : “Descriptive Visual Words and Visual Phrases for Image Applications”, ACM Multimedia (2009).
- [11] D. G. Lowe : “Object recognition from local

scale-invariant features”, Proc. Of the 7th international Conference on Computer Vision. pp. 1150-1157 (1999).

- [12] 永橋知行, 藤吉弘亘, 金出武雄 : “領域分割に基づく SIFT 特徴量を用いた物体識別”, 電気学会研究会資料. SC, システム制御研究会, Vol.2007, No.1, pp. 39-44 (2007).

## 辻 健太 Kenta TSUJI

立命館大学大学院理工学研究科博士前期課程在学中。立命館大学情報理工学部情報コミュニケーション学科卒業。情報検索に関する研究に従事。ACM SIGMOD, 情報処理学会, 日本データベース学会各学生会員。

## 川越 恭二 Kyoji KAWAGOE

立命館大学情報理工学部情報コミュニケーション学科教授。情報システム, 情報検索, データ工学に関する研究に従事。IEEE, ACM SIGMOD, 情報処理学会, 電子情報通信学会, 情報システム学会, 日本データベース学会各会員。