

# キーワード検索との組み合わせによる オブジェクト画像検索方式の評価検討

## Evaluation of an object based image retrieval method with combination of keywords

高橋 祐介  
Yusuke Takahashi

平田恭二  
Kyoji Hirata

### 1. まえがき

インターネットのような大量の画像データの中から、特定の画像を探し出す方法として、キーワードによる画像検索法が広く用いられている。画像が含まれる HTML ページ内の周辺テキストをその画像のメタデータとして抽出し、キーワード・マッチングにより画像を抽出する。この手法では、画像解析が行われないので、抽出されたメタデータと画像内容とが無関係なものが数多く抽出される問題があり、ユーザは所望の画像を探すために大量の結果画像を閲覧しなければならない。

AMORE[1]は、HTML 解析によるキーワード抽出および画像内の領域の色・位置関係などの構成情報を抽出し、類似画像検索を行っている。このビジュアル・意味的な類似度による画像検索が美術館などで利用され、キーワードと画像による検索の併用の有効性を示している。しかし、インターネットのような無秩序な画像集合では、画像の構成が雑多であるため、所望の画像を抽出することは困難である。

異なる構成の画像間における検索では、画像中に含まれる対象物(オブジェクト)の利用が有効である。ユーザは構図に関わらず該当オブジェクトの含まれる画像を検索することができる。我々はこれまで、画像を視覚的特徴に基づき分割し、複数の分割領域の組み合わせをクエリーとすることでオブジェクトベース検索を実現する画像検索方式を提案してきた[2]。この方式では、登録画像 1000~5000 枚に対し 2 秒以内の全探索を実現し、有効性とインタラクティブな検索速度を確認した。

そこで、本稿ではインターネット上の画像に対し、キーワード画像検索による絞り込みとオブジェクト画像検索を組み合わせた方式を提案し、その効果について検討する。

### 2. 複数パーツによるオブジェクト画像検索[2]

オブジェクトは、図1の「ひまわり」における構成要素「花びら」「芯」「葉」のように1つまたは複数の要素から構成される。通常、これらの各構成要素(以後、パーツと呼ぶ)は、色や模様が均一であることが多く、同一オブジェクトのパーツは異なる画像においても類似な視覚的特徴を持つ。したがって、オブジェクトを構成する複数のパーツをクエリーとし、他の画像中のパーツとの視覚的特徴量の類似度を測定することで、他の画像中に含まれる対象オブジェクトを含む画像を検索することが可能になる。

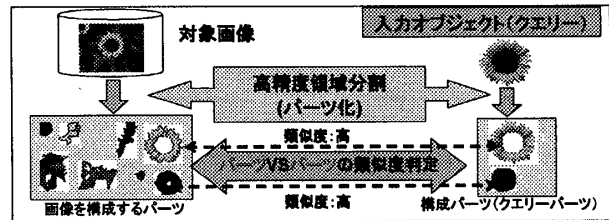


図1 複数パーツによるオブジェクト画像検索方式

### 3. キーワード+オブジェクト検索システム

インターネット上の画像からキーワードにより画像を絞り込み、オブジェクトのパーツを指定することで所望の画像を検索するシステムを構築する。

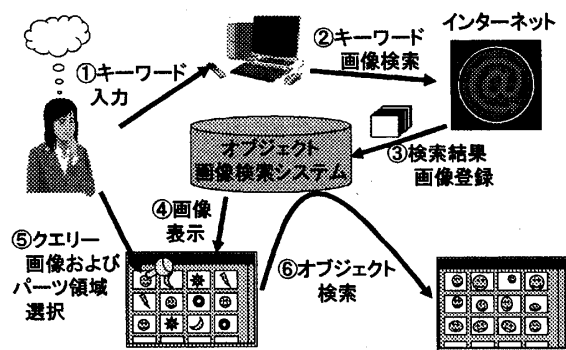


図2 キーワード+オブジェクト画像検索方式

従来のキーワードによる画像検索と同様にユーザは複数のキーワードを入力する。キーワードを基に、周辺テキストとのマッチングによるインターネット画像の検索を行い、検索結果上位の画像をデータベースに登録する。登録では、検索した全ての画像に対し領域分割を施し、パーツに分割する。また、類似度算出のために各パーツから色・形・模様に関する視覚的特徴量 [3] を抽出し記録する。

キーワードによる検索結果を上位からユーザに提示し、ユーザは所望の画像(オブジェクト)を探すキーとなるパーツを含む画像(クエリー画像)を選び、その画像から1つ以上のパーツを選択する。

選択されたパーツを用いて、登録されている画像に対しオブジェクト画像検索を行う。登録画像ごとに、選択されたパーツとの類似度の総和を統合スコアとして算出し、類似画像のランキングを求める。ランキングの結果をユーザに提示する。

ユーザのパーツ指定および結果表示手段として、図3に示すWEBベースのGUIを用いた。ユーザがクエリーとなるパーツをピンにより1つ以上選択すると、検索結果が表示

される。検索結果には、類似パーツ部分に赤い○を重ねて表示することで、対応箇所を判りやすくしている。

検索結果に応じて、ユーザはピンを移動させて別のパーツによる検索や、さらに別の位置を追加選択して複数のパーツによる検索を選択的に行うことができる。このようなGUIにより、検索結果を見ながらユーザが効果的なパーツを追加することができ、所望の画像を容易に検索することができる。

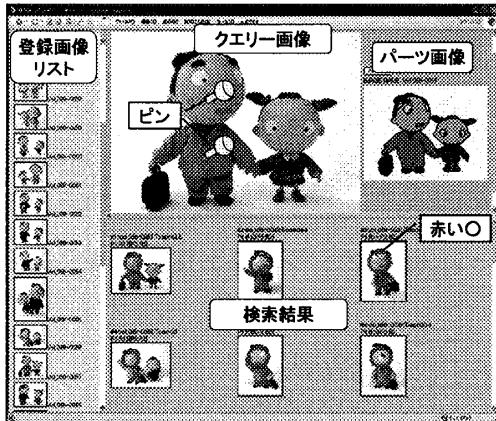


図3 インタラクティブ画像検索画面

#### 4. 実験

インターネット画像を検索エンジンにより収集し、オブジェクト検索を行うことで、上記のシステムの評価を行った。利用する画像は、表1に示す7種類のキーワード群を用いて検索エンジンにより得られた画像データである。各キーワード群において、検索エンジンにより収集した1000枚程度の画像から、キーワードを用いて得られた上位30位以内の画像と、所望のオブジェクトを含む画像を用いて、領域ベースの類似度による上位30に挙げられた画像とを比較し、所望のオブジェクトを含む画像が含まれる割合 (Precision) を抽出することで評価を行う。ただし、所望のオブジェクトを含む画像か否かの正解判定は目視で行った。

7つのキーワード群により得られた画像データにおける上位30位の正解率を図4に示す。キーワードのみの検索では、所望の画像が含まれる割合が平均17%と低いことが判る。オブジェクトを指定することにより、検索精度が平均で55%に向上することを確認した。

キーワードによる検索では、所望でない画像が30位中平均25個検索されるため、ランキングリストをそのまま見るのは非効率である。例えばある人物のスピーチの画像の検索においては、キーワードのみでは人以外の画像も多く含んでしまうため、オブジェクト検索でのパーツとして、顔領域を選択することで顔画像を表示させ、集中的に顔領域を含む画像を閲覧するだけでも効果があると考えられる。ただし、顔認識としての精度は確保できないため、別人の顔も上位に検出される。

マークやロゴなどのさまざまな画像において見え方がほぼ一定の定型オブジェクトに対して、提案手法では高い割合で抽出可能であることが判る。また、人物のスピーチの画像検索では、スピーチ台についている会社の紋章などの特徴的なオブジェクトが存在する。このようなオブジェク

トは通常メタデータとして付与されることはないが重要な意味を持つため、このような典型的なオブジェクトをパーツとして選択することによりキーワードに比べて所望の画像を検索することが可能となる。

表1 キーワードおよびオブジェクトによる検索結果

データ番号	キーワード オブジェクト	利用 枚数	キーワー ドのみ	+オブジェ クト検索
1	president+speech+us 人物の顔+スピーチ台 の紋章	906	17% (5/30)	30% (9/30)
2	space+shuttle+launch シャトル本体+噴射煙	983	27% (8/30)	93% (28/30)
3	olympic+Torino+2006 トリノ五輪のマーク	893	3% (1/30)	60% (18/30)
4	PC+linux+logo Linux ロゴ(ペンギン)	772	10% (3/30)	87% (26/30)
5	beef+mad-cow+disease 牛肉	165	20% (6/30)	60% (18/30)
6	bird+flu+china にわとり	547	27% (8/30)	33% (10/30)
7	soccer+world+cup 優勝杯	990	13% (4/30)	23% (7/30)

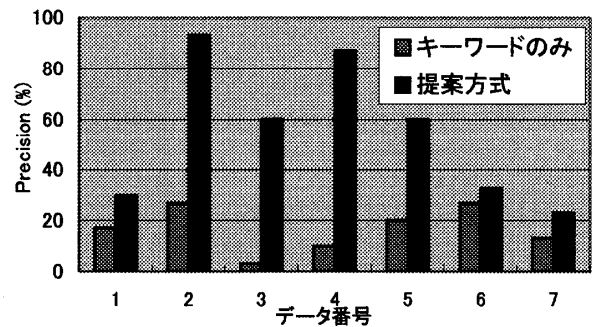


図4 上位30位中に所望の画像が含まれる割合

#### 5. むすび

本論文では、キーワードによる画像検索結果と領域を選択することによるオブジェクト検索との組み合わせによる画像検索の精度向上について評価を行った。インターネットからキーワードにより取得した1000画像程度の画像群において、上位30位までに所望の画像が含まれる割合を、ユーザが特定のオブジェクトを指定することで平均17%から55%へ向上することができた。

#### 参考文献

- [1] 平田, Mukherjer, マルチメディアウェブ検索システム AMORE, オブジェクト構成要素に基づく画像検索, 画像ラボ, Vol. 10, No. 2, pp. 21-26, 1999
- [2] 高橋, 平田, オブジェクト構成要素に基づく画像検索, 電子情報通信学会総合大会講演論文集, Vol. 2005, pp.199, 2005
- [3] K. Hirata, E. Kasutani, Y. Hara, "On Image Segmentation for Object-based Image Retrieval," Proceedings. of ICPR, vol.3, pp.1031-1034, 2002.