

# タグの共起と類似画像を用いたタグ付け支援システム

非会員 瀬崎 直人\* 正員 黄瀬 浩一\*

A System for Recommending Tags of Images  
Using Co-occurrence of Tags and Similar Images

Naoto Sezaki\*, Non-member, Koichi Kise\*, Member

In this paper, we propose a system for recommending tags of images. The proposed method presents to its user various tags with high precision by taking into account both the co-occurrence of tags and tags of similar images. Additional search with the user feedback enables us to find some new tags relevant to the image of interest. In order to test the proposed method, we compare it with a method based only on the co-occurrence of tags, as well as a method based only on the similarity of images. From the experimental results using 16002 images, we have confirmed that the proposed method is capable of finding more tags as compared to the methods for comparison. We have also confirmed that the additional search is effective to find additional tags.

キーワード: 画像検索, タグ, 共起, LSI, 類似画像, EMD

Keywords: image retrieval, tag, co-occurrence, LSI, similar image, EMD

## 1. まえがき

近年のデジタルカメラやインターネットの普及に伴い, Web 上には大量の画像が存在している。これらの画像から, ユーザが必要とするものを得る手法の一つとして, キーワードに基づいて画像を検索する手法がある。これを実現するためには, 画像に対して予めキーワードを付与しておく必要がある。このような問題を解決する一手法として, folksonomy<sup>(1)</sup>がある。この手法では, 個々のユーザが自分の視点でデータにタグと呼ばれるキーワードを付与し, それらのタグを共有することにより, 各ユーザはタグを通してそれぞれのデータにアクセスする。この手法は, flickr<sup>(2)</sup>を代表とする画像共有サイトなどで応用され, 注目を集めている。画像共有サイトでは, ユーザが画像に対してタグを付与することにより索引付け(タグ付け)を行う。

しかし, 初心者はどのようなタグをつければいいのか分からないため, 初心者がタグ付けを行った画像は十分なタグが付与されないという問題点がある。一方で, 熟練したユーザによって十分にタグが付与された画像も多く存在する。このような画像のデータベースにおいて, 互いに関連する語は, 同時に同じ画像のタグになっている(共起して

いる)ことが多い。そのため, ユーザが入力したタグと共起することが多いタグを画像の関連語として提示できると考えられる。また, 類似している画像は同じタグを含んでいることが多いので, これらのタグも関連語として提示できると考えられる。

そこで本研究では, このことを利用し, タグの共起と類似画像を利用したタグ付け支援システムを提案する。関連研究として, 画像に自動でタグ付けを行う手法がある<sup>(3)(4)</sup>。しかし, これらの手法は, 安定して特徴量が得られるタグ付き Corel 画像を対象としており, ノイズとなるタグを多く含み, 画像自体の多様性も高い Web 画像に対して, 同様の性能を得るまでには至っていない。提案システムでは, ユーザとのインタラクションを利用し, Web 上の画像に対しても, 精度よくタグ付けを行うことを目的とする。

精度のよいタグ付けを行うためには, まず, どのようなタグを付与するかが重要となる。flickr 等で付与されている画像のタグは, 大きく, 一般的な語, 固有な語, その他の語に分類される。その他の語には, 撮影した人物やカメラの名前, 日付など, 特殊な場合を除けば検索にはそれほど有効ではないタグが多く含まれる。固有な語によるタグとは, 国名, 地名, 施設名など個々の画像に特化したタグである。これらの語を画像に付与することで, 画像検索の際に対象を少数に絞込み, 所望の画像にたどりつきやすくなる。しかし, これらを正しく付与するには画像の認識が必要となるため, 現状の技術ではタグ付け支援にはかなりの

\* 大阪府立大学大学院工学研究科  
〒 599-8531 大阪府堺市学園町 1-1  
Graduate School of Engineering, Osaka Prefecture University  
Gakuen-cho 1-1, Sakai, Osaka, 599-8531 Japan

困難が伴うといえる。一方、一般的な語は、beach, flower, beautiful など、同種の画像に共通するタグである。これらは、画像検索の際に、固有な語のタグほど対象の絞り込みに有効ではないが、数多く画像に付与することで、索引が充実し、多様な検索要求にも対処できると考えられる。さらに、一般的な語は、同種の画像に共通するタグであるため、タグとして付与されている頻度が他の語よりも高い。よって、タグ付け支援のための統計的な手がかりも得やすいと考えられる。そこで、提案システムでは、タグの中でも一般的な語に焦点をあて、画像のタグ付け支援を行う。

本手法の特徴は、タグの共起に加えて類似画像の関連語も抽出することで、多様で精度の高い関連語を新しいタグの候補としてユーザに提示できる点である。また、ユーザが新たに付与したタグを入力したタグに含め、関連語の再検索を行うことで、さらに新しい関連語を抽出可能である。

タグの共起のみを用いる手法や類似画像のみを用いる手法と比較実験を行った結果、提案手法はより多くの有効な関連語を抽出できることがわかった。また、これらの有効なタグを入力タグに追加し、関連語の再検索を行うことにより、有効なタグの数をおよそ2倍に増加できることも明らかになった。

## 2. 関連研究

画像検索には、画像から抽出した色・テクスチャなどの画像の内容に基づく画像検索 (Content-Based Image Retrieval: CBIR)<sup>(5)</sup> と画像に与えられたキーワードに基づく画像検索 (Text-Based Image Retrieval: TBIR) が存在する。

CBIR では画像の色や形状などを特徴量として類似画像を検索する。このような特徴量では画像の内容を的確に表すことが難しいので、満足のいく検索性能が得られていない。この問題を解決するため、Ruiらは、関連フィードバック (relevance feedback) を用いて検索性能の向上を行う手法を提案している<sup>(6)</sup>。

TBIR では、既存の情報検索の技術を用いて画像検索を行うことが可能である。しかし、TBIRを実現するためにはデータベース中の画像に対して、テキストによるタグ付けを行う必要がある。Web ページを対象とした自動索引付け手法がいくつか提案されている<sup>(7)-(10)</sup>。しかし、Web ページは記述のされ方も様々であるため、Web 上の画像に対して安定して充実したタグ付けを行うことは容易ではない。

一方、Web には、folksonomy<sup>(11)</sup> という手法を用いてブックマークや画像を分類するサービスが存在する。この手法では、サービスを利用するユーザがデータにタグと呼ばれるメタデータを付与する。そして、付与されたタグを共有することにより、各ユーザはタグを通してデータにアクセスする。flickr はタグ付けを行う対象を画像としたサービスである。画像にタグ付けを行うことで、あるタグからの繋がりにより、新しい画像を発見することが可能となる。

タグ付けされた大量の画像から、未知の画像に対してタグ付けを支援する手法が提案されている<sup>(3)(4)</sup>。これらの手

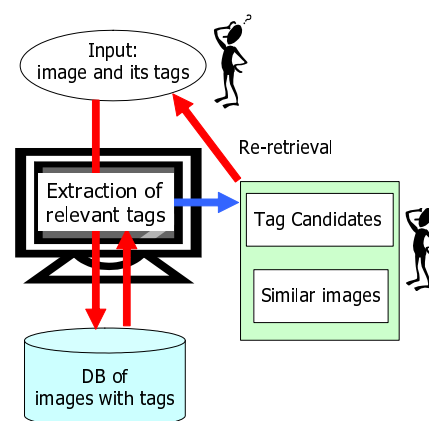


Fig. 1. System overview.

法では、画像から得られる特徴量、またはタグから得られる特徴量を用いて、画像のタグ付けを行う。しかし、これらの手法は、安定して特徴量が得られるタグ付き Corel 画像を対象としており、ノイズとなるタグを多く含み、画像自体の多様性も高い Web 画像に対して、同様の性能を得るまでには至っていない。

## 3. タグの共起と類似画像を用いたタグ付け支援システム

画像共有サイトには、熟練したユーザによって十分にタグが付与された画像が多く存在する。提案するシステムでは、このような十分にタグ付けされた画像を利用して、新たに画像の関連語を抽出し、ユーザに提示する。

3.1 システムの概要 タグ付け支援システムの全体構成を図1に示す。提案システムでは、画像と画像に付与された少数のタグのペアを入力とする。付与するタグはユーザが自由に設定する。

この画像とタグのペアを入力とし、「関連語抽出モジュール」によって、次にタグ付けすべき語を検索し、ユーザに提示する。このモジュールでは、タグが十分に付与された画像のデータベースにアクセスし、関連語を抽出する。関連語抽出は、入力したタグとデータベース中にある画像のタグの共起、また類似画像のタグを利用して行う。この処理により得られた関連語を新たなタグの候補として出力し、ユーザはそれを画像の新たなタグとして付与するかどうかを判断する。この際、タグの候補を出力すると同時に、それらが索引付けられている類似画像も出力する。これにより、ユーザは類似画像から視覚的にタグの候補を付与するかどうかを判断することができる。

さらに、このシステムでは、ユーザが新たに付与したタグを入力したタグに含め、関連語の再検索を行う。入力とするタグの数を増やし再検索を行うことで、新しい関連語をユーザに提示する。これにより、ユーザはより多くのタグを画像に付与することが可能となる。

3.2 タグの共起を利用した関連語抽出 互いに関連する語は、同時に同じ画像のタグになっている (共起し

ている)ことが多いと考えられる。この考えを利用し、入力したタグと共起することが多いタグを画像の関連語として抽出する。具体的な手順は以下の通りである。

まず、データベース中から、ノイズとなるタグを削除する。Web上から収集した画像には、撮影した人物の名前、日時といった関連語としては不適当なタグが多く存在する。これらを削除するために、データベース中での出現頻度が  $c$  回以上のタグのみを利用する。また、これらのタグから不要語 (stop word) を削除する。提案システムでは、不要語の中に、撮影したカメラの名前、国名、地名を含める。国名、地名は画像の関連語として考えられる場合があるが、本研究はより一般的に画像に付与されるタグの提示を目的としているため、これらのタグは不要語として扱う。

次に、画像に付与されているタグをタグベクトルとして表現する。データベース中に  $\alpha$  種類のタグ  $w_1, w_2, \dots, w_\alpha$  が存在するとき、画像  $I_i$  を表すタグベクトル  $t_i$  を次のように表現する。

$$t_i = [t_{i1}, t_{i2}, \dots, t_{i\alpha}]^T \quad (1)$$

ここで  $T$  は転置、 $t_{ij}$  は  $w_j$  の出現を表す。タグ  $w_j$  が画像  $I_i$  に付与されている場合は  $t_{ij}$  を 1、付与されていない場合は 0 とする。

画像  $I_i, I_j$  のタグベクトル  $t_i, t_j$  が与えられたとき、これらの間の類似度  $\text{Sim}(t_i, t_j)$  は LSI 法<sup>(11)</sup> を用いて計算する。具体的には以下の通りである。画像データベース中の各画像のタグベクトルをまとめた行列を、

$$D = [t_1, t_2, \dots, t_\beta] \quad (2)$$

とする。このとき  $D$  を次のように特異値分解する。

$$D = U \Sigma V^T \quad (3)$$

この処理で得られた  $U$  の最初の  $k$  次元の左特異値ベクトルのみから構成される行列を  $U_k$  とする。このとき  $t_i$  の  $k$  次元表現  $t_i^{(k)}$  は以下の式で与えられる。

$$t_i^{(k)} = U_k^T t_i \quad (4)$$

これを用いて、 $\text{Sim}(t_i, t_j)$  を以下の式で計算する。

$$\text{Sim}(t_i, t_j) = \frac{t_i^{(k)} \cdot t_j^{(k)}}{\|t_i^{(k)}\| \|t_j^{(k)}\|} \quad (5)$$

提案システムでは、入力とする画像とデータベース中のすべての画像に対してタグベクトルの類似度を計算する。これらの画像を類似度が高い (付与するタグが類似している) 順にソートし、上位  $a_1$  個の画像において、付与されているタグの出現回数を計算する。これらのタグを出現回数が多い順にソートし、上位  $b_1$  個のタグを関連語とみなす。

### 3.3 類似画像を利用した関連語抽出 提案システム

では、類似画像には同じようなタグが付与されていることを仮定する。例えば、類似の風景であれば、同じタグが付与されていることが多いと考えられる。この仮定のもと、類似画像に付与されているタグを、入力画像の関連語として抽出する。具体的な手順は以下の通りである。

提案システムでは画像の色の分布を用いて、類似画像を検索する。色の分布を表す画像の特徴量として color signature を用いる。color signature は色を表す代表色ベクトル  $p_i$  と色の画素数の割合  $r_{p_i}$  の組  $(p_i, r_{p_i})$  で表される。color signature 同士の距離を求めるには、Earth Movers Distance (EMD)<sup>(12)</sup> を用いる。color signature  $P = \{(p_1, r_{p_1}), (p_2, r_{p_2}), \dots, (p_m, r_{p_m})\}$ ,  $Q = \{(q_1, r_{q_1}), (q_2, r_{q_2}), \dots, (q_n, r_{q_n})\}$  と color signature の要素間の距離  $d_{ij} = d(p_i, q_j)$ , ( $1 \leq i \leq m, 1 \leq j \leq n$ ) が与えられたとき、 $P, Q$  間の EMD は次のように表現される。

$$\text{EMD}(P, Q) = \frac{\sum_{i=1}^m \sum_{j=1}^n f_{ij} d_{ij}}{\sum_{i=1}^m \sum_{j=1}^n f_{ij}} \quad (6)$$

ただし、 $f_{ij}$  は式 (6) を以下の条件で最小化する最適化問題の解である。

$$f_{ij} \geq 0 \quad 1 \leq i \leq m, 1 \leq j \leq n \quad (7)$$

$$f_{ij} \leq r_{p_i} \quad 1 \leq i \leq m \quad (8)$$

$$f_{ij} \leq r_{q_j} \quad 1 \leq j \leq n \quad (9)$$

$$\sum_{i=1}^m \sum_{j=1}^n f_{ij} = \min \left( \sum_{i=1}^m r_{p_i}, \sum_{j=1}^n r_{q_j} \right) \quad (10)$$

本手法の場合、色空間として人間の感覚に近いと言われている CIE L\*a\*b\* 色空間を用い、代表色ベクトルの距離  $d(p_i, q_j)$  としてユークリッド距離を用いる。

提案システムでは、3.2 により得られるタグの類似度が高い  $l$  個の画像に対して EMD を測定する。これはすべての画像に対して EMD を測定すると計算時間が膨大となるためである。これらの画像を EMD が小さい (色の分布が類似している) 順にソートし、上位  $a_2$  個の画像においてタグの出現回数を計算する。これらのタグを出現回数が多い順にソートし、上位  $b_2$  個を関連語とみなす。

**3.4 関連語の提示** タグの共起により抽出される関連語の集合を  $G_{\text{tag}}$ 、類似画像により抽出される関連語の集合を  $G_{\text{img}}$  とすると、最終的にユーザに提示する関連語の集合  $G_U, G_\cap$  を次のように求める。

$$G_U = G_{\text{tag}} \cup G_{\text{img}} \quad (11)$$

$$G_\cap = G_{\text{tag}} \cap G_{\text{img}} \quad (12)$$

$G_U$  は、3.2 と 3.3 の処理で抽出される、すべての関連語である。これにより、それぞれの処理だけで抽出する場合よりも、多くの関連語をユーザに提示する。また、 $G_\cap$  はタグの共起と類似画像の両方から抽出される関連語なので、タグの候補の中でも特に関連度の高いタグとしてユーザに提示される。

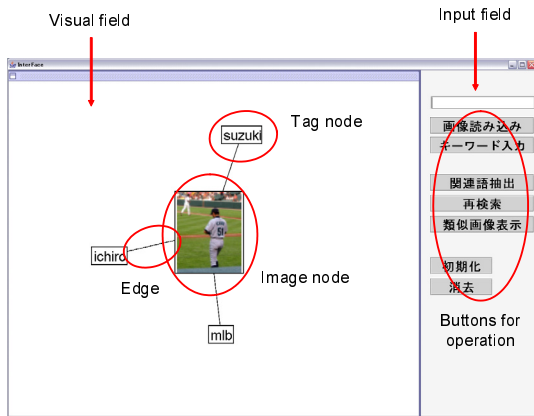


Fig. 2. GUI of the proposed system .

**3.5 関連語の再検索** 関連語を提示されると、ユーザはの中から適切なものを選択して付与する。提案システムは、新たに付与されたタグを元のタグに含め、再検索を行う。まず、元のタグに新たなタグを加え、式 (4) によりタグベクトルを求める。次に、3.2, 3.3, 3.4 の処理を再び行い、3.2 の処理から  $b_1$  個、3.3 の処理から  $b_2$  個の関連語を抽出する。ここで出力される関連語は、前の関連語抽出で得られなかった新しい関連語である。これにより、ユーザには新しい関連語を提示され、より多くのタグを画像に付与することが可能となる。

**3.6 インタフェースと実行例** 提案システムのインタフェースを図 2 に示す。以下、ユーザがこのインタフェースを使って画像にタグを付与する手順を説明する。まず、ユーザは入力フィールドでタグ付けを行う画像とそれに付与するタグを入力する。これにより、視覚フィールドには、入力された画像とタグがノードとして出力され、ユーザによる操作が可能となる。このとき画像に付与されるタグノードとその画像ノードはエッジで結ばれている。次に、関連語抽出ボタンにより、「関連語抽出モジュール」を起動する。ユーザは、これにより出力されたタグの候補ノードから画像に付与するノードを選択する。タグの候補ノードは  $G_n$  の要素が赤、それ以外 ( $G_U - G_n$ ) が青で表示される。また、類似画像表示ボタンで、関連語抽出に用いた類似画像のノードを出力することができる。さらに、再検索ボタンで再び「関連語モジュール」を起動することにより、新たなタグの候補ノードを得ることができる。

提案システムの実行例を図 3 に示す。ユーザは、入力する画像に対して、少数のタグを付与して関連語抽出を行う。その結果、視覚フィールド上に複数のタグの候補ノードが出力される。また、同時に関連語抽出に用いた類似画像も出力される。図 3 において、点線で囲まれているタグと画像が、タグの候補と類似画像である。

## 4. 実験

**4.1 実験概要** 提案手法の性能を評価するために、性能比較実験を行った。比較手法として、タグの共起のみ

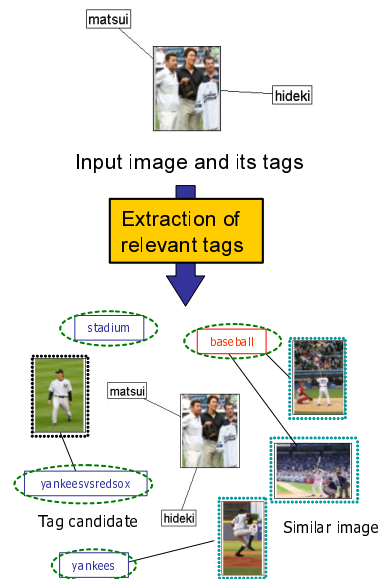


Fig. 3. A result of the extraction of relevant tags .

を用いる手法と類似画像のみを用いる手法を用いた。タグの共起のみを用いる手法は  $G_{tag}$ 、類似画像のみを用いる手法は  $G_{img}$  を関連語として提示するものである。実験に用いたパラメータを表 1 に示す。ここで、パラメータの値は、予備実験の結果に基づいて経験的に定めたものである。

まず、実験 1 では、関連語抽出の性能に関する実験を行った。画像共有サイト flickr から検索質問 33 個を用いて収集した 16002 枚の画像、それらに付与されている 18596 個のタグからノイズのタグ、不要語、低頻度語を除去した 1396 個のタグを使用した。この画像とタグの集合からデータベースを構築して関連語抽出を行い、タグの共起のみを用いる手法から得られた  $G_{tag}$ 、類似画像のみを用いる手法から得られた  $G_{img}$ 、提案手法から得られた  $G_U$ 、 $G_n$  をそれぞれ評価した。表 2 に実験に用いた検索質問を示す。評価尺度としては、結果として得られた正解タグの数  $|A|$ 、および適合率  $P = \frac{|A|}{|G_1|}$  を用いた。ここで、 $|G_1|$  は処理結果として得られたタグの数である。また、正解の判定は著者が行った。

提案システムでは、ユーザが正解タグを選択しやすいように、一度の関連語抽出で、少数のタグを提示し、その中により多く正解タグを含めることを目的とする。従って、適合率、即ち抽出したタグにどれだけ正しいものが含まれているかが重要となり、再現率、即ちすべての正解タグをどれほど抽出できるかは、さほど重要ではない。

実験 2 では、再検索の性能に関する実験を行った。実験 1 により得られた  $G_U$  から人手で正解タグを選び、入力したタグに含め、再び関連語抽出を行った。これにより新たに得られたタグについて  $G_{tag}$ 、 $G_{img}$ 、 $G_U$ 、 $G_n$  をそれぞれ評価した。評価尺度としては、結果として追加された正解タグの数  $|B|$ 、および適合率  $P = \frac{|B|}{|G_2|}$  を用いた。ここで、 $|G_2|$  は再検索により新たに得られたタグの数である。その

Table 1. Values of parameters .

$k$	$a_1$	$b_1$	$l$	$a_2$	$b_2$	$c$
500	60	10	300	60	10	10

Table 2. Queries used in the experiments .

spring	summer	autumn
ichiro	guitar	architecture
sky	leaf	cloud
matsui	lake	snow
city	winter	building
castle	animal	flower
moon	locomotive	beach
sea	coast	fruits
bird	mountain	tree
fish	waterfall	ocean
night	sunset	rock

他の実験条件は、実験 1 と同様である。

4.2 実験 1: 関連語抽出の性能に関する実験 実験 1 の結果を表 3 に示す。また、この実験により得られた正解タグの例を図 4 に示す。表 3 に示すように、提案手法により得られた  $G_U$  が、正解タグの数において最良の結果となった。これより、タグの共起だけでは抽出できなかった関連語を、類似画像も考慮することで抽出が可能であることがわかった。特に図 4 に示す「winter」、「beach」といった検索質問では、タグの共起により正解タグを抽出することは難しいが、色の分布が画像の内容をよく反映しているので、類似画像検索の精度が高くなった。このことにより、類似画像から新しい正解タグを多く抽出することができた。ここで注意すべき点は、 $G_{img}$  は色のみを用いてタグを抽出しているのではない点である。最初にユーザが付与したタグと色を同時に考慮することで、有効なタグの抽出が可能となっている。一方、検索質問「summer」、「guitar」では、類似画像から得られる正解タグの数は少なかった。色の分布だけでなく、新しい画像特徴量を用いて類似画像検索の性能を向上させれば、このような検索質問でも、正解タグを多く抽出できると考えられる。

また、提案手法で得られた  $G_\cap$  が、適合率において最良の結果となった。これより、タグの共起と類似画像の両方から得られる関連語は、画像に付与される語である可能性が高いことがわかった。検索質問「locomotive」では、類似画像だけでなくタグの共起により得られる正解タグが多かったため、両方から得られる関連語が多く抽出され、またその中に正解タグを多く含んでいた。このような場合、 $G_\cap$  を用いることで関連度の高いタグを提示できるといえる。

3.6 で述べたインタフェースでは、 $G_U$  全体を表示している。これは、どのような関連語が望ましいのかはユーザに依るところも大きいため、なるべく多くの関連語をユーザに提示すべきであると考えからである。しかし、 $G_U$  に含まれるタグの数は 20 個程度となるため、かえってユーザを混乱させる可能性もある。この問題は、関連度に応じてタグを画面上に見やすく配置することにより緩和される可

Table 3. The number of correct tags and precision obtained in the experiment 1 .

	$G_{tag}$	$G_{img}$	$G_U$	$G_\cap$
No. of correct tags	1.9	3.5	4.7	0.8
Precision	0.19	0.35	0.25	0.41

Table 4. The number of correct tags and precision obtained in the experiment 2 .

	$G_{tag}$	$G_{img}$	$G_U$	$G_\cap$
No. of correct tags	4.2	3.0	5.3	1.8
Precision	0.42	0.30	0.33	0.50

能性がある。関連度の定義には、式 (4) の  $U_k$ 、また 3.3 で用いた EMD などを用いることができるであろう。具体的な定義や評価については、今後の課題としたい。

4.3 実験 2: 再検索の性能に関する実験 実験 2 の結果を表 4 に示す。ここで示される数値は、再検索の結果、新たに得られたタグに関するものである。また、この実験により得られた正解タグの例を図 4 示す。表 4 に示すように、関連語の再検索を行うことで、タグの共起、類似画像それぞれから正解タグを追加し、画像に有効なタグの量をおよそ 2 倍に増加できることがわかった。実験 1 と比べて、タグの共起により得られる正解タグの数が大幅に増えた。これは、入力するタグが増えたことで、タグの共起による画像検索の精度が上がったためであると考えられる。検索質問「summer」、「winter」では、実験 1 よりも多くの正解タグを抽出することができた。このように、タグを追加して再検索を行うことで、少数のタグの入力では抽出できなかった関連語を抽出できると考えられる。

また、実験 1 の結果と同様に、正解タグの数においては  $G_U$ 、適合率においては  $G_\cap$  が最良の結果を得た。検索質問「architecture」では、実験 1、実験 2 とともに、このような関連語を多く抽出することができた。よって、関連語の再検索においても、提案手法は、比較手法に比べて関連語抽出の性能が優れているといえる。

一方、検索質問「guitar」は再検索を行っても、新しい正解タグを増やすことは困難であった。柔軟に関連語を抽出するために、類似画像だけでなく、タグの共起による抽出性能の向上も課題として挙げられる。

## 5. むすび

本稿では、タグの共起と類似画像を用いたタグ付け支援システムを提案した。提案システムの特徴は、タグの共起だけでなく類似画像の関連語も抽出することで、多くの関連語をユーザに提示できる点である。実験の結果、タグの共起と類似画像を用いることで、それぞれ単独で用いる手法より、多く関連語を抽出することができた。また、それらの関連語を用いて再検索を行うことで、新しい関連語を追加することができた。今後の課題としては、より多くの正しい関連語を抽出するために、タグの共起、類似画像それぞれが持つ抽出性能の向上、ならびにユーザインタフェー

スでのタグのより良い配置などが挙げられる。

## 文 献

- (1) A. Mathes: "Folksonomies - cooperative classification and communication through shared metadata", Computer Mediated Communication, UIC Technical Report (2004)
- (2) <http://www.flickr.com/>
- (3) X.-J. Wang, L. Zhang, F. Jing and W.-Y. Ma: "Annosearch: Image auto-annotation by search", Proc. of IEEE Computer Vision and Pattern Recognition, Vol.2, pp. 1483-1490 (2006)
- (4) H. Nakayama, T. Harada, Y. Kuniyoshi and N. Otsu: "Ul-

- tra high speed image annotation/retrieval method by learning the conceptual relationship between images and labels", Technical report of IEICE, PRMU2007-147, pp. 65-70 (2007)
- 中山英樹, 原田達也, 國吉康夫, 大津展之: 「画像・単語間概念対応の確率構造学習を利用した超高速画像認識・検索方法」, 電子情報通信学会技術報告, PRMU2007-147, pp. 65-70 (2007)
- (5) V. Gudivada, and V. Raghavan.: "Content-based image retrieval-systems", IEEE Comput., Vol.28, No.9, pp. 18-22 (1995)
- (6) Y. Rui, T. S. Huang and S. Mehrotra: "Relevance feedback techniques in interactive content-based image retrieval", Proc. Storage and Retrieval for Image and Video Databases (SPIE), pp. 25-36 (1998)
- (7) E.V.Munson, Y. Tsybalyenko: "To search for images on the web, look at the text, then look at the images", Proc. First International Workshop on Web Document Analysis, pp. 39-42 (2001)
- (8) H. Dehara, N. Fujimoto, H. Takeno and K. Hagihara: "A Sentence Extraction Technique Based on HTML Parsing Tree Structures around Images for WWW Image Retrieval", IEICE technical report, DC2005-30 (2005)
- 出原博, 藤本典幸, 竹野浩, 萩原兼一: 「WWW 画像検索における画像周辺の HTML 構文構造を考慮した画像説明文の抽出手法」, 電子情報通信学会技術報告, DC2005-30 (2005)
- (9) N. Aihara, W. Sunayama, and M. Yachida: "Image Labeling Using Key Sentences of HTML", The transactions of the Institute of Electronics, Information and Communication Engineers, No.2, pp. 145-153 (2004)
- 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, No.2, pp. 145-153 (2004)
- (10) K. Takeuchi and K. Kise: "Extraction of Explanatory Text for Web Images Using Similar Images and Keywords", IPSJ SIG Notes, NL-171, pp. 7-12 (2006)
- 竹内謹治, 黄瀬浩一: 「類似画像とキーワードを利用した web 画像の説明文抽出」, 情報処理学会研究報告, NL-171, pp. 7-12 (2006)
- (11) S. C. Deerwester, S. T. Dumais, T. K. Landauer, G. W. Furnas and R. A. Harshman: "Indexing by latent semantic analysis", Journal of the American Society of Information Science, Vol.41, No.6, pp. 391-407 (1990)
- (12) Y. Rubner, C. Tomasi and L. Guibas.: "The earth mover's distance as a metric for image retrieval", International Journal of Computer Vision, Vol.40, pp. 99-121 (2000)

## 謝 辞

本研究の一部は、大阪府立大学産学官連携プロジェクト、先端科学共同研究プロジェクトの支援による。

瀬崎直人 (非会員) 平成 18 年大阪府立大学工学部情報工学科卒業。平成 20 年同大学院博士前期課程修了。現在、関西電力勤務。在学中、Web 画像検索に関する研究に従事。

黄瀬浩一 (正員) 昭和 61 年大阪大学工学部通信工学科卒業。昭和 63 年同大学院博士前期課程修了。同年同大学院博士後期課程入学。平成 2 年大阪府立大学工学部助手。現在、同大学院工学研究科教授。博士 (工学)。平成 12-13 年ドイツ人工知能研究センター客員教授。文書画像解析、画像認識、情報検索などの研究に従事。電子情報通信学会、情報処理学会、IEEE、ACM などの会員。

入力する 画像とタグのペア			
		winter	summer
関連語抽出で 得られた正解タグ $G_{\cup}$	$G_{img}$	nature ocean sky	cloud beach tree
	$G_{img}$	nature sky	fog tree
	$G_{\cap}$		vacation
再検索で 得られた正解タグ $G_{\cup}$	$G_{img}$	cloud plants	leaves snow
	$G_{img}$	cloud plants	snow white storm
	$G_{\cap}$	cloud plants	snow
入力する 画像とタグのペア			
		beach	locomotive
関連語抽出で 得られた正解タグ $G_{\cup}$	$G_{img}$	cloud nature	train engine railroad steamengine diesel
	$G_{img}$	ocean sea	sunset coast water
	$G_{\cap}$		engine train railroad
再検索で 得られた正解タグ $G_{\cup}$	$G_{img}$	seashore sun	evening sand
	$G_{img}$	clouds coastal	waves sun shore
	$G_{\cap}$		travel
入力する 画像とタグのペア			
		architecture	guitar
関連語抽出で 得られた正解タグ $G_{\cup}$	$G_{img}$		landscape livemusic music
	$G_{img}$	building travel	historical landscape art
	$G_{\cap}$		landscape
再検索で 得られた正解タグ $G_{\cup}$	$G_{img}$	history historic view	tree city europe old
	$G_{img}$	city europe	historic tree
	$G_{\cap}$	city europe	historic tree

Fig. 4. Examples of recommended correct tags.