

# Eatrh Mover's Distanceに基づく Text-Based Image Retrieval

竹内 謹治<sup>†</sup> 黄瀬 浩一<sup>†</sup>

<sup>†</sup> 大阪府立大学大学院工学研究科 〒599-8531 大阪府堺市中央区学園町 1-1  
E-mail: †takeuchi@m.cs.osakafu-u.ac.jp, ††kise@cs.osakafu-u.ac.jp

**あらまし** 本稿では Earth Mover's Distance(EMD) を用いてキーワードによる画像検索 (TBIR) を行う手法を提案する。従来の TBIR 手法の性能を向上させるアプローチの一つとして、キーワードに加えて画像自体から得られる特徴量を利用することが考えられる。EMD を用いることで、本来、異質であるキーワードと画像特徴量を統一的に扱うことが可能である。提案手法の有効性を確認するため、12295 枚の画像を対象にして、LSI, PLSI などの従来手法との比較実験を行った結果、平均適合率において、最良の結果が得られた。また、EMD を用いる手法の高速化のために、近似方法を検討した結果、LSI と同程度の検索性能を保ちながら、理想的には計算量を約 8% に削減できることが分かった。  
キーワード 画像検索, TBIR, CBIR, EMD

## Text-Based Image Retrieval Using Earth Mover's Distance

Kinji TAKEUCHI<sup>†</sup> and Koichi KISE<sup>†</sup>

<sup>†</sup> Graduate School of Engineering, Osaka Prefecture University Gakuenchou 1-1, Naka-ku,  
Sakai-si, Osaka, 599-8531 Japan  
E-mail: †takeuchi@m.cs.osakafu-u.ac.jp, ††kise@cs.osakafu-u.ac.jp

**Abstract** In this report, we propose a method of Text-Based Image Retrieval(TBIR) using Earth Mover's Distance(EMD). Our approach uses EMD to combine textual and visual features into one semantic space, to improve the performance of TBIR. We compare our method with four conventional methods including LSI and PLSI. The results show our method is superior to others in average precision. We also applied an approximation in computing EMD for improving the efficiency of the proposed method. The results show that the approximate method can reduce the computational complexity in about 8%, with little loss of average precision.

**Key words** image retrieval, TBIR, CBIR, EMD

### 1. はじめに

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。これらの大量の画像に対して、ユーザが必要とする画像を得る手法の一つとして、キーワードに基づく画像検索 (Text-Based Image Retrieval: TBIR) がある。TBIR 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。このような問題の解決手段として、画像に対して自動で索引付けをする研究がな

されている [1] [2] [3] [4]。特に、Web ページから画像説明文を抽出し画像に付与する研究や [1] [2]、ユーザ自身が画像に索引語を付与する folksonomy [5] と呼ばれる技術などによって、豊富なキーワードによって索引付けされた画像が大量に存在するという環境が整いつつある。

本研究では上記の手法などによって索引付けされたキーワードと検索質問との類似度をどのように測るのかという問題に焦点を当てる。TBIR では通常、情報検索の技術を用いて類似度を測定する。しかしながら

索引付けの段階で、必ずしも正しいキーワードが画像に付与されているわけではないため、結果としてキーワードのみによる検索の精度には課題を残している。

このような課題を克服するためのアプローチとして、キーワードの他に画像自体から抽出した特徴量を利用することが考えられる。本研究ではそのようなアプローチの一手法として、Earth Mover's Distanceに基づくTBIR手法を提案する。提案手法の特徴として、キーワードと画像特徴量を統一的に扱える点が挙げられる。

実験の結果、従来手法を上回る検索精度が得られ提案手法の有効性を確認した。また、提案手法の実用化のため、わずかな検索精度の低下で高速な検索が可能な近似方法に関して検討した結果、LSI程度の検索性能を保ちながら、計算量を8%程度に削減できることが分かった。

## 2. 関連研究

画像検索には、画像の特徴量を用いて類似画像を検索する手法(Content-Based Image Retrieval: CBIR) [6]と画像に対してあらかじめ対応付けられたテキストに基づいてキーワード検索を行う手法(Text-Based Image Retrieval: TBIR)が存在する。本研究ではTBIRを取り扱う。

TBIRでは従来の情報検索技術を用いて画像を検索する。情報検索の代表的な技術として、ベクトル空間法(VSM) [7]がある。VSMではベクトルの次元と索引語の数に等しい。例えばVSMにおいては「car」と「automobile」は別物である。その結果一方の索引語から他方の索引語を持つ画像が検索できない。そのため、検索に失敗する画像が多いという問題がある。

上記のような問題点を解決する手段として、高次元空間にあるベクトルを低次元空間に射影することにより検索精度の改善を図る手法 [8] [9] [10] [11] などがある。これらの手法では射影する空間の次元の決定が実行錯誤的であるという点に課題を残している。

一般に、TBIRでは、検索精度は索引語が正しく付与されているかどうかにかかわらず依存するといえる。もし、索引語に依存しない画像特徴量を利用することが出来れば、検索性能を改善することが出来ると考えられる。そこでこのような特徴量を用いる画像検索(CBIR)について説明する。

CBIRは画像から色や模様、形などの特徴量を抽出し、特徴量が類似している画像を検索する。上記のような、画像から得られる特徴量と人間の感覚は異なるため、検索精度に課題を残している。そこで、検索精度の向上のため、画像特徴量の類似度に加え、索引語の類

似度を用いる手法も提案されている [12]。ただし、[12]の手法では上記の2つの類似度は並列的に用いられているだけであり、索引語と画像特徴量の関係については考慮されていない。

ところで、近年、CBIRの分野ではEarth Mover's Distance(EMD) [13]と呼ばれる尺度が注目されており、他の尺度より人間の感覚に近いことが知られている。EMDは需要地(供給地)の重みと需要地と供給地の間のコストを定義すればどのような問題にも適用できるため、さまざまな分野で用いられている。例えば [14]では文書間の類似度の測定に用いており、ベクトル空間法による類似度と比較して、高性能であることが実験により報告されている。

## 3. 提案手法

前述のように、EMDは需要地(供給地)の重みと、需要地と供給地の間のコストさえ定義できればどのような問題に対しても求めることができる。たとえば、需要地(供給地)に単語と画像特徴量が混在しているとしても、それぞれの重みとコストを定義することが出来ればEMDを求めることが出来る。そこで、提案手法ではこれらを定義することにより、単語と画像特徴量を統一的に扱うことの出来る画像検索の実現を目標とする。

以下ではまず、EMDについて説明する。次に、実際にEMDをどのようにTBIRに用いるのかについて説明する。

### 3.1 Earth Mover's Distance [13]

Earth Mover's Distanceは線形計画問題の1つであるヒッチコック型輸送問題の解に基づく尺度である。

ヒッチコック型輸送問題とは需要地と需要量の集合  $P = \{(p_1, r_{p1}), (p_2, r_{p2}) \cdots (p_n, r_{pn})\}$ 、供給地と供給量の集合  $Q = \{(q_1, r_{q2}), (q_2, r_{q2}) \cdots (q_n, r_{qm})\}$  及び、各需要地、供給値間の輸送コスト  $c_{ij} = \text{cost}(p_i, q_j)$ , ( $1 \leq i \leq m, 1 \leq j \leq n$ ) が定義されたとき、需要地、供給値間の総輸送コストの最小値を求める問題である。

ヒッチコック型輸送問題は以下のような最適化問題として定式化できる。

$$\text{minimize } \sum_{i=1}^m \sum_{j=1}^n f_{ij} c_{ij} \quad (1)$$

$$\text{s.t. } f_{ij} \geq 0 \quad (1 \leq i \leq m, 1 \leq j \leq n) \quad (2)$$

$$f_{ij} \leq r_{p_i} \quad (1 \leq i \leq m) \quad (3)$$

$$f_{ij} \leq r_{q_j} \quad (1 \leq j \leq n) \quad (4)$$

$$\sum_{i=1}^m \sum_{j=1}^n f_{ij} = \min \left( \sum_{i=1}^m r_{p_i}, \sum_{j=1}^n r_{q_j} \right) \quad (5)$$

このとき、P、Q間の Earth Mover's Distance は以下の式で与えられる。

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

### 3.2 EMD の TBIR への適用

EMD を TBIR に適用するには需要地 (供給地)、需要量 (供給量)、輸送コストを定義する必要がある。

まず、需要地、供給地をどのように定義するのかについて説明する。需要地としては、検索質問の各単語を、供給地としては画像の索引語、および画像特徴量を割り当てる。このように定義すると、需要地は、検索質問をベクトルで表現したときの、0 で無い要素に相当する。つまり、検索質問  $\mathbf{q}$  は

$$\mathbf{q} = [t_{q1}, t_{q2} \cdots t_{qn}]^T \quad (7)$$

と表現できる。ここで  $t_{qi}$  は単語  $w_i$  が検索質問中の語であるとき 1、それ以外の場合は 0 とする。一方、供給地は上記の  $n$  次元ベクトルに  $m$  個の画像特徴量の次元を加えたベクトルで表現する。つまり、画像  $I_i$  は

$$\mathbf{I}_i = [t_{i1}, t_{i2} \cdots t_{in}, c_{i1}, c_{i2} \cdots c_{im}]^T \quad (8)$$

という  $(n+m)$  次元のベクトルで表現する。

このように考えることにより、需要地 (供給地) の重みは、従来の情報検索の重みを用いることが出来る。以下で、具体的な重みの与え方について説明する。

まず、2進重みを与える場合について説明する。 $t_{ij}$  は単語  $w_j$  が画像  $\mathbf{I}_i$  の索引語のとき 1、それ以外は 0 とする。また、 $c_{ik}$  は、対応する画像特徴量が画像  $\mathbf{I}_i$  をよく表していれば 1、そうでなければ 0 とする。たとえば、画像特徴量として色を用いる場合では、画像  $\mathbf{I}_i$  に対して、ある色が閾値以上の割合で含まれている場合 1、そうで無い場合 0 とする。後述する実験では画像特徴量として、均等量子化により 8色に減色した色を用いた。

次に、重みとして IDF 重みを与える場合について説明する。まず、 $t_{ij}$  を以下の式で定義する。

$$t_{ij} = l_{ij} \log \frac{N}{N_j} \quad (9)$$

ここで  $l_{ij}$  は前述した 2進重みを表している。また、 $N$  はデータベースの画像の総数。 $N_j$  は  $l_{ij} = 1$  である画像  $\mathbf{I}_i$  の総数を表している。 $c_{ik}$  についても、同様の式で定義する。

今、検索質問  $\mathbf{q}$  に関しても、画像特徴量の次元が全て 0 であると考えれば  $(n+m)$  次元のベクトルとして考えられる。今後は、 $\mathbf{q}$  も  $(n+m)$  次元のベクトルで

あると考える。以下、 $\mathbf{q}$  を検索質問ベクトル、 $\mathbf{I}_i$  を画像ベクトルと呼ぶことにする。

最後に輸送コストについて説明する。輸送コストとしては、単語間あるいは単語-画像特徴量の関連度を利用する。関連度としては共起の度合いを用いる。共起は画像データベースから計算する。ここで共起とは、画像ベクトル  $\mathbf{I}_i$  において、同時に 0 で無い要素になっていることである。このとき、輸送コストを以下の式で定義する。

$$\text{cost}(w_i, w_j) = 1 - \frac{n_{ij}}{\sqrt{n_i n_j}} \quad (10)$$

ただし、 $n_{ij}$  は  $w_i$  と  $w_j$  の共起頻度、 $n_i$ 、 $n_j$  は  $w_i$ 、 $w_j$  の出現頻度である。式 (10) は単語間のコストの定義であるが、単語-画像特徴量間のコストに関しても同様の式で定義する。

さて、TBIR を実現するには、これらを用いて検索質問とデータベース中の全画像との EMD を求め、EMD が小さい画像から順にユーザに提示すればよい。

## 4. 実験

提案手法の有効性を検証するため、従来手法と性能比較実験を行った。

### 4.1 実験データおよび実験環境

まず、実験データについて示す。実験では写真共有サイト flickr から収集した 12295 枚の画像<sup>(注1)</sup>、それらに付与されている 8565 個の索引語を使用した。この画像集合からデータベースを構築し、画像検索を行い、その性能を評価した。表 1 に検索に用いた検索質問を示す。ここで、実験の評価には leave-one-out 法によって得られる平均適合率および検索時間を用いた。具体的には、14 個の検索質問を用いて最良の平均適合率が得られるパラメータを求め、残り 1 つの検索質問における性能を評価した。パラメータの詳細については後述する。

次に正解データの作成手順について説明する。まず、実験意図を知らない情報工学系の大学院生 3 人が個別に検索質問に適合する画像をデータベースから探した。次にこの結果において 3 人中 2 人が正解とした画像を検索質問に対する正解画像であるとした。

最後に実験環境について説明する。実験に使用した計算機としては CPU が AMD Opteron 2.8GHz、メモリが 16GB のものを用いた。また、プログラミング言語としては MATLAB を用いた。

(注1): 収集に使った検索質問としては  
<http://www.flickr.com/photos/tags/>に掲載されている  
 All times most popular tags 145 種類を用いた。

表 1 検索質問

beach sunset
sky clouds
flower red
moutain green tree
baby
city night
house cat
animal zoo
europa architecturure
rock concert
wedding
africa people
snow winter
japan
home party kids

#### 4.2 提案手法および比較手法の設定

実験では、比較手法として、ベクトル空間法 (以下 VSM)、擬似関連フィードバック (以下 PRF) [15]、LSI [8]、PLSI [9] を用いた。本節では提案手法および上記の比較手法のパラメータ等について説明する。

まず、全手法で共通する設定について説明する。これらの手法で用いる重み付けとしては、2進重みと IDF 重みの両方で実験を行い、良好な結果が得られた重みを採用した。その結果、提案手法は 2進重みをその他の手法では IDF 重みを採用した。ただしユークリッドノルムで正規化したものを用いた。

次にパラメータの決定方法について説明する。まず、PRF について説明する。実験では PRF を以下のように考えた。まず、検索質問拡張のための文書集合  $E$  は以下の式で定義した。

$$E = \left\{ d_j^+ \mid \frac{\text{sim}(d_j^+, \mathbf{q})}{\max_i \text{sim}(d_i, \mathbf{q})} \geq \tau \right\} \quad (11)$$

ここで、 $\mathbf{q}$  は元の検索質問ベクトルである。このとき、拡張した検索質問ベクトルを以下の式で定義した。

$$\mathbf{q}' = \frac{\mathbf{q}}{\|\mathbf{q}\|} + \lambda \frac{\mathbf{d}_s}{\|\mathbf{d}_s\|} \quad (12)$$

$$\text{s.t. } \mathbf{d}_s = \sum_{d_j^+ \in E} d_j^+ \quad (13)$$

ここで、 $\tau$  については 0.7 から 0.99 まで、0.01 刻みで実験を行い最良の結果が得られたものを採用した。また、 $\lambda$  については 1 と 2 で実験を行い、良好な結果が得られたものを採用した。

次に LSI について説明する。LSI の次元としては、200 次元から 400 次元まで 5 次元刻みで実験を行い、

表 2 実験 1 のパラメータの平均および標準偏差

	$\tau$	$\lambda$	$d_1$	$d_2$
平均	0.76	2	242	380
標準偏差	0	0	11.6	31.6

表 3 実験 1 の結果 (平均適合率)

提案手法	VSM	PRF	LSI	PLSI
0.158	0.113	0.111	0.138	0.133

表 4 実験 1 における各手法の検索時間 (秒)

提案手法	VSM	PRF	LSI	PLSI
0.18	0.0049	0.011	0.22	0.14

最良の結果が得られた次元  $d_1$  を採用した。

最後に PLSI について説明する。PLSI としては、[9] で提案されている、PLSI-Q を用いて 200 次元から 400 次元まで 50 次元刻みで実験を行い、最良の結果が得られた次元  $d_2$  を採用した。なお、PLSI は EM アルゴリズムの初期値によって結果が異なるので、10 回検索を行い、その結果の平均を結果とした。

#### 4.3 実験 1: 単語のみによる検索

まず、検索質問ベクトル、画像ベクトルとして、画像特徴量を用いずに、テキスト検索による性能を比較した。実験 1 におけるパラメータの平均、標準偏差を表 2 に示す。実験結果を表 3、表 4 に示す。

表 3 が示すように、平均適合率では提案手法が最良の結果を得た。提案手法は 15 検索質問中 13 検索質問で最良の結果が得られた。これらのことから提案手法は多くの検索質問に対して検索性能が優れているということが言える。提案手法で最良の結果が得られなかった検索質問は「house cat」、「wedding」であり、これらに対しては LSI が優れていた。

また、表 4 が示すように、検索質問あたりの平均検索時間では VSM が最も優れていた。これは、VSM においては、検索質問ベクトル、画像ベクトルが疎なベクトルであるためであると考えられる。提案手法の検索時間は 0.18 秒であり、4 番目の結果であった。

一般に、検索性能と検索時間はトレードオフの関係にある。しかしながら、実験では提案手法は LSI、PLSI とほとんど変わらない検索時間で、これらを上回る検索性能が得られるということが分かった。これは、提案手法の優れた特徴であるといえる。

#### 4.4 実験 2: 画像特徴量を考慮した検索

実験 2 では画像特徴量を考慮することによる、検索性能の向上を検証した。まず、表 5 に実験 2 における各手法のパラメータの平均、標準偏差を示す。画像特



表 5 実験 2 のパラメータの平均および標準偏差

	$\tau$	$\lambda$	$d_1$	$d_2$
平均	0.77	2	242	376.7
標準偏差	0	0	11.6	32.0

表 6 実験 2 の結果 (平均適合率)

提案手法	VSM	PRF	LSI	PLSI
0.160	0.116	0.127	0.134	0.128

表 7 実験 2 における各手法の検索時間 (秒)

提案手法	VSM	PRF	LSI	PLSI
0.20	0.0066	0.015	0.22	0.14

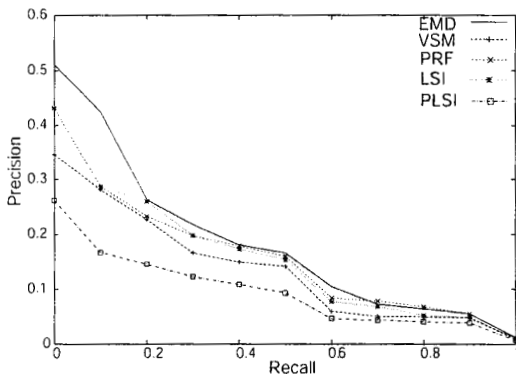


図 1 画像を考慮した場合における再現率-適合率グラフ

微量としては均等量子化によって 8 色に減色した色を用いた。具体的には、ある色が対象画像における全面素の 2 割以上を占めている場合、その色が画像をよく表していると考えられる。

平均適合率の結果を表 6 に示す。平均適合率は提案手法、VSM、PRF において、向上がみられた。一方、LSI、PLSI では平均適合率が低下した。また、実験 1 と同様に提案手法が最も良い結果だった。提案手法は 15 検索質問中 11 検索質問で最良の結果が得られた。これらのことから、画像特徴量を用いた場合においても、提案手法は検索性能が優れているということがいえる。

図 1 に各手法の再現率-適合率グラフを示す。この図が示すように、再現率 0.1 付近では、提案手法が他の手法と比べてかなり優れている。一方、高再現率では全ての手法において満足な結果が得られなかった。画像特徴量を考慮しない場合と比べて検索結果が向上したことと、この結果をふまえると、画像の色は確かに画像を表す有効な特徴であるが、識別性能としては、必

ずしも満足なものではないということがいえる。検索性能をさらに向上させるには、識別性能の高い画像特徴量を利用する必要があると考えられる。

次に、実験 2 における検索時間を表 7 に示す。この表が示すように提案手法、VSM、PRF では検索時間の増加がみられた。特に、提案手法は約 0.02 秒と、約 10% の増加がみられた。一方、LSI、PLSI においては検索時間の増加はみられなかった。提案手法の時間が他と比べて大幅に増加した理由としては、1 つの類似度を求めるのに、輸送問題を解く必要があるからであると考えられる。

ヒッチコック型の輸送問題では、需要地の数を  $m$ 、供給地の数を  $n$  としたとき、その解を求めるには、 $O((nm)^4)$  の計算量を要する<sup>(注2)</sup>。したがって、今後、検索性能を向上させるために画像特徴量を増やすことを考える場合、検索時間の増加に対処する必要がある。

#### 4.5 提案手法の近似方法に関する検討

本節では、提案手法の検索時間を短縮するため、近似方法について検討する。具体的な近似の方針については大きく分けて次の 2 つが考えられる。

- 検索対象をあらかじめ絞り込む。
- 類似度を高速に計算する。

まず、前者について考える。提案手法では検索質問と画像の索引の間に定義された単語間のコストが全て 1 であるとき、検索質問と画像の EMD は 1 になる。このような画像に対しては EMD を計算しないことで計算速度を向上させることが出来る。また、単語間のコストが 1 に近いような場合、そのコストを近似的に 1 と見なすことでさらに計算量の削減が見込める。このことを実験的に確かめる。

まず、単語間のコストを次のように定義し直す。

$$\text{cost}(w_i, w_j) = \begin{cases} 1 - \frac{n_{ij}}{\sqrt{n_i n_j}} & (1 - \frac{n_{ij}}{\sqrt{n_i n_j}} < t) \\ 1 & (1 - \frac{n_{ij}}{\sqrt{n_i n_j}} \geq t) \end{cases} \quad (14)$$

上記のコストを用いて、画像検索を行い、平均適合率、及び EMD 計算対象となる画像の割合を調べた。実験は  $t = 0.1$  から 0.1 刻みで 1 まで行った。実験結果を図 2、3 に示す。これらの図が示すように  $t = 0.8$  のとき、EMD の計算対象となる画像の割合が約 8% であるにもかかわらず、0.14 程度の平均適合率が得られた。このことは、LSI 程度の検索性能を保ちながら、計算量を理論上 8% 程度に削減できることを意味している。

次に、類似度を高速に計算する方法について検討す

(注2) : simplex 法で解く場合

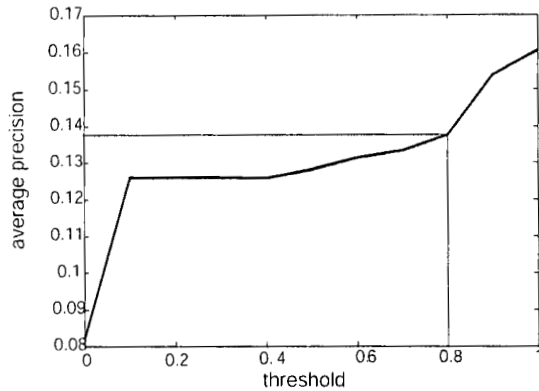


図 2 近似手法における平均適合率

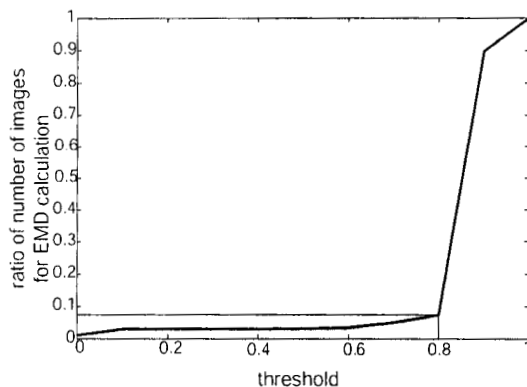


図 3 EMD の計算対象となる画像の割合

る。実験のデータでは需要地と供給地間のコストの内の 9 割程度が 1、即ち共起回数が 0 であった。このようなコストが 1 になる経路はいわばノイズとして働いていると考えられるため、もし、ヒッチコック型の輸送問題を解く代わりに、上記のようなコストが 1 の経路を無視した輸送問題を解くことが出来れば、検索時間の大幅な削減が見込める上に、検索性能もあまり変わらないのではないかと考えられる。

## 5. おわりに

本稿では Earth Mover's Distance を用いた TBIR 手法を提案した。実験の結果、従来手法と比べて最良の検索結果が得られた。また、検索対象を絞り込むことにより、わずかな性能の低下で計算量を大幅に削減できることを実験により示した。今後の課題としては、検索性能の向上のために識別性能の高い画像特徴を利用すること、および検索時間削減のために、Earth Mover's Distance の計算の高速化などが挙げられる。

## 文 献

- [1] 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, Jan., 2006.
- [2] 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,” 信学技報, DE2005-136, 2005.
- [3] 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,” 信学論 (D), vol.J87-D-1, no.2, pp.145-153, 2004.
- [4] 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,” 情処学論, vol.42, no.SIG 10(TOD 11), pp.79-91, 2001.
- [5] A. Mathes, “Folksonomies: Cooperative classification and communication through shared metadata,” Technical Report LIS590CMC, University of Illinois Urbana-Champaign, 2004.
- [6] Y. Rui, T. Huang and S. Chang, “Image retrieval: current techniques, promising directions and open issues,” Journal of Visual Communication and Image Representation, vol.10, no.4, pp.39-62, 1999.
- [7] G. Salton, A. Wong and C. S. Yang, “A vector space model for automatic indexing,” Commun. ACM, vol.18, no.11, pp.613-620, 1975.
- [8] 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.
- [9] T. Hofmann, “Probabilistic latent semantic indexing,” Research and Development in Information Retrieval, pp.50-57, 1999.
- [10] 黒岩眞吾, 柘植覚, 田仁宏典, T. Xiaoying, 獅々堀正幹, 北研二, “Simple pca を用いたベクトル空間情報検索モデルの次元削減,” 情処学 NL 研報, NL-144, pp.61-66, 2001.
- [11] 柘植覚, 獅々堀正幹, 北研二, “Non - negative matrix factorization を用いた情報検索,” 情処学 NL 研報, NL-142, pp.1-6, 2001.
- [12] S. Sclaroff, M. L. Cascia, S. Sethi and L. Taycher, “Combining textual and visual cues for content-based image retrieval on the world wide web,” Computer Vision and Image Understanding, vol.75, no.Nos.1/2, July/August, pp.86-98, 1999.
- [13] Y. Rubner, C. Tomasi and L. Guibas., “The earth mover's distance as a metric for image retrieval,” Int. J. Comput. Vision, vol.40, pp.99-121, 2000.
- [14] X. Wan and Y. Peng, “The earth mover's distance as a semantic measure for document similarity,” CIKM '05: Proceedings of the 14th ACM international conference on Information and knowledge management, New York, NY, USA, ACM Press, pp.301-302, 2005.
- [15] C. Buckley, A. Singhal, M. Mitra and G. Salton, “New retrieval approaches using smart: Trec,” Text REtrieval Conference, pp.25-48, 1996.