

# 類似画像とキーワードを利用した Web 画像の説明文抽出

竹内 謹冶<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 大阪府堺市学園町 1-1

E-mail: <sup>†</sup>takeuchi@m.cs.osakafu-u.ac.jp, <sup>††</sup>kise@cs.osakafu-u.ac.jp

あらまし 本稿では類似画像検索と共起関係に基づくキーワードの拡張を用いて Web 画像の説明文を抽出する手法を提案する。Web ページの画像は、同一ページのテキストによりなんらかの説明を受けていることが多い。このような説明文には画像に関連する語が多数含まれているため、これらを画像に対応付けることで検索索引が充実し、詳細な検索要求にも対応できると考えられる。しかしながら、言語と画像は異質であるため、簡単には対応付けられないという問題点がある。提案手法では、画像的な特徴と言語的な特徴の両方を用いることで、これらを単独で用いたときの欠点を補いつつ、高精度な説明文抽出を可能とする。手法の有効性を確認するため、995 枚の Web 画像に対して説明文を抽出した結果、F 値で 0.75 程度の結果が得られた。

キーワード 画像検索,  $\chi^2$  検定, EMD

## Extraction of Explanatory Text for Web Images Using Similar Images and Keywords

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

<sup>†</sup> Graduate School of Engineering, Osaka Prefecture University, 1-1 Gakuenchou, Sakai, Osaka, 599-8531 Japan

E-mail: <sup>†</sup>takeuchi@m.cs.osakafu-u.ac.jp, <sup>††</sup>kise@cs.osakafu-u.ac.jp

**Abstract** We propose a method of extracting explanatory text of images from HTML files based both on retrieval of similar images and expansion of keywords using co-occurrence. Images in web pages often have their explanatory text in the same pages. Such explanatory text includes important keywords that explain the corresponding image. It is therefore reasonable to put these keywords to the image for fulfilling a wide variety of needs from users. In the proposed method features extracted from the image and the text complement each other to extract explanatory text effectively. From the experimental results on 995 images on the web, the proposed method yielded F-measure of 0.75.

**Key words** image retrieval,  $\chi^2$ test, EMD

### 1. はじめに

近年、インターネットやデジタルカメラの普及に伴い、Web ページ上には大量の画像が存在している。これらの大量の画像データを活用するには、必要な画像をキーワードから検索するための画像検索技術が求められている。

このような索引付けの従来手法として alt テキストやファイル名、タグの構造などから HTML を解析し、画

像に索引付けする手法がある [1] [4] [5] [6]。しかしながら、HTML の記述は様々であるため、安定して良好な索引付けを行うことは困難であると考えられる。

我々は経験的に「類似した画像は類似したキーワードが与えられている」ということを知っている。もし、類似画像からキーワードを得ることができれば、手法の安定性が増し、多くの画像に対して、良好な索引付けが可能になると考えられる。

ところで、画像の説明文は多数のキーワードが含まれている。このような説明文を索引付けテキストとすることで、多様な要求にも対処できると考えられる。更に、説明文を読むことで画像の内容をより詳しく知ることができるという利点も考えられる。

本稿では類似画像から得られるキーワードを利用し、画像の説明文を抽出する手法を提案する。提案手法では言語的な特徴と、画像的な特徴を組み合わせることで、これらの手法を単独で用いたときの欠点を補うことができるという利点がある。言語的な特徴のみを使った [1] の手法と説明文の抽出性能に対して比較実験を行った結果、提案手法は、比較手法と比べ F 値が約 2 倍であることが分かり、提案手法の有効性が示された。

## 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) [2] 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。また、対象言語は英語とする。

TBIR では、データベース中の画像に対して、テキストによる索引付けを行う必要がある。初期の画像検索システムでは、人手により索引付けを行っていた。しかし、人手による索引付けは、コストが大きいことや、索引語が主観的であるなどの問題点が指摘されている [3]。

このような問題を解決するために、画像に対して自動的に索引付けを行う手法が研究されている。Web を対象とした、自動索引付け手法として、[4] [5] などがある。相良ら [4] は「画像からの距離による評価」、「出現頻度による評価」、「Web ページ中の重要文による評価」などを組み合わせることにより索引語を抽出する手法を提案している。Munson [5] らは HTML のメタデータを利用する手法を提案している。文献 [5] では、IMG タグの alt テキストは画像と関連していることが多いが、alt テキストが記述されている画像が少ないと言う事が報告されている。

従来の自動索引付け手法によって得ることのできるキーワードは語数が非常に限られている。そのため、既存の画像検索システムではキーワードを増やすと、画像数が急激に減少するために、所望の画像にたどり着けないということがしばしば起こる。

この問題を解決する手段として、HTML テキストから画像の説明文を選択し、画像に対応付ける手法がある [1] [6]。画像の説明文には、画像に関連するキーワードを多く含んでいるために、これらに対応付けることで検索索引が充実し、多様な検索要求にも応えられるということが考えられる。また、その画像と同時に説明文を提

示することで、ユーザが画像の内容を詳しく知ることが出来るという利点がある。

竹内ら [1] は alt テキストと関連のある語 (キーワード) を  $\chi^2$  検定により抽出し、それらを含む文を説明文とする手法を提案している。[1] の手法は画像の alt テキストが記述されていない場合には適用できないという欠点がある。

また、出原ら [6] は HTML のタグの構造を利用することで画像の説明文を抽出する手法を提案している。HTML は使われるタグの種類や、階層構造が安定しないため、安定性に課題があると考えられる。

## 3. 提案手法

我々は経験的に「類似した画像は類似したキーワードが与えられている」ということを知っている。提案手法では、この事実を利用し、類似画像のキーワードを利用し、[1] の手法により説明文を抽出する。

以下、3.1 で [1] の手法で用いられている alt テキストを利用したキーワード抽出方法、3.2 で画像の類似度の測定方法、3.3 で提案手法のアルゴリズムについて説明する。

### 3.1 alt テキストを利用したキーワード抽出

一般に、2 つの語がなんらかの関係があるとき、それらは文中で同時に出現する (共起する) ことが多い。このような特徴を利用し、alt テキスト中の語と共起することが多い語をキーワードとして抽出する。

#### (1) 準備:

Google のテキスト検索の検索結果の Web ページをダウンロードする。

次に各ページを解析し、本文および、画像の alt テキストの抽出をする。そして抽出した本文に stemming を施す。また、JPEG 以外の画像を不要画像と見なし排除する。JPEG 画像についても、アスペクト比に基づいて、横長あるいは縦長の画像を排除する。

#### (2) 単語列の抽出:

Web ページの本文の集合から、[7] の手法を用いて、出現回数 3 回以上の 4-gram までの単語列を全て抽出する。

#### (3) キーワードの抽出:

各単語列  $w$  に対して次のような帰無仮説を立て  $\chi^2$  検定を行う。「alt テキスト中の単語列  $g \in G$  と  $w$  は文中で独立に出現する」、 $\chi^2$  値を求めるには次の式を用いる。

$$\chi^2(w) = \sum_{g \in G} \frac{(\text{freq}(w, g) - n_w p_g)^2}{n_w p_g} \quad (1)$$

ここで、 $p_g$  は  $g$  が文に出現する確率、 $\text{freq}(w, g)$  は  $w$  と  $g$  が文中に同時に出現する回数、 $n_w$  は  $w$  と  $g \in G$

のどれかが文中で同時に出現する回数である。  $p_g$  はその和が 1 になるように正規化する。

ここで、  $n$  を  $G$  の要素数としたとき、自由度  $n-1$ 、有意水準  $\alpha$  の  $\chi^2$  値、  $\chi^2_{[n-1]}(\alpha)$  を用いて検定を行う。もし、  $\chi^2(w) > \chi^2_{[n-1]}(\alpha)$  であれば、その単語列  $w$  を画像に対するキーワードとみなす。

この手法では入力が必要しも alt テキストである必要はない。そこで、今後この手法の入力に用いる語をシードと呼ぶことにする。

### 3.2 画像の類似度の測定

提案手法では、画像特徴量として 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  $r_i$  の組  $(p_i, r_i)$  で表される。

Color signature 同士の距離を求めるには、Earth Movers Distance(EMD) [8] を用いる。EMD は線形計画問題の一つであるヒッチコック型の輸送問題の解に基づいて計算される。一方の Color signature を供給地の集合とそれぞれの供給量、もう一方の Color signature を需要地の集合とそれぞれの必要量と見立てて、供給地と需要地間の輸送コストが与えられた時に、需要地の需要を満たすように供給地から需要地に輸送する最小輸送コストが EMD による距離となる。

ここで、EMD の距離関数としてユークリッド距離を、色空間として人間の感覚に近いと言われている CIEL\*a\*b\* 色空間を用いる。

### 3.3 アルゴリズム

提案手法のアルゴリズムは以下の通りである。

提案手法では類似画像検索を色の分布の類似度だけで行っている。そのため、意味的にまったく異なる画像を類似画像として認定してしまう場合がある。その結果、キーワードに画像と関係のない語が含まれる可能性がある。しかし、以下の (3) を行うことによりそのような語を排除することが出来ると考えられる。

- (1) alt テキストを含んでいる画像に対して、alt テキストをシードとし、3.1 の手順でキーワードを抽出する。
- (2) 全ての画像間に対して EMD を測定する。そして、EMD が閾値  $t$  をこえない画像を類似画像とする。
- (3) 画像の存在する Web ページ中の単語列と類似画像のキーワードを比べ、両方に存在する単語列を画像のシードとする。
- (4) 3.1 の手法でキーワードを抽出する。
- (5) (3) から (4) の処理を  $i$  回繰り返す。
- (6) キーワードを  $k$  個以上含んでいる文を画像の説明文であるとする。

表 1 検索質問毎の画像数

検索質問	画像数
Abu Mena	50
Aksum	27
Basilica di san Pietro	41
Big Ben	50
Buckingham Palace	124
Cliff of Bandiagara	107
Colossi of Memnon	28
Diamond Head	71
Eiffel Tower	49
Golden Gate Bridge	93
Grand Canyon	71
Greenwich Observatory	45
Haleakala Volcano	76
Kairouan	59
Tower Bridge	104
合計	995

## 4. 実 験

提案手法の性能を評価するため、性能比較実験を行った。比較手法としては [1] を用いた。

実験はまず、提案手法で用いる類似画像検索の性能について調べた。次に、キーワードの適合率と再現率、F 値を調べる「キーワードに関する実験」そして、説明文の適合率と再現率、F 値を調べる「説明文に関する実験」を行った。

### 4.1 実験条件

実験に用いた Web ページは、Google のテキスト検索エンジンで、検索質問を「Greenwich Observatory」など世界の観光地 15カ所を入力した結果のそれぞれ上位 20 件である。実験対象とした画像は上記の Web ページに存在する 995 枚の説明文付き画像である。検索質問毎の画像数を表 1 に示す。

実験で用いた各パラメータについて示す。 $\chi^2$  検定の有意水準  $\alpha$  の値は [1] で用いられている  $\alpha = 0.05$  に設定した。また、説明文抽出の際のキーワードの数の閾値は、予備実験の結果  $k = 1$  より  $k = 2$  の方が良好な結果が得られ、 $k > 2$  では  $k = 2$  の時と結果に差が出なかったため  $k = 2$  に設定した。そして、繰り返し回数は  $i = 1$  に、類似画像検索の閾値は  $t = 3000$  に設定した。

各実験の正解判定は第一筆者が行った。キーワードに関する実験の正解判定は、画像の説明文中に存在するキーワードかどうかを基準とした。また、説明文に関する実験の正解判定は、画像とその説明文が関連しているかどうかを基準とし、検索質問に関連しているかどうかは問わなかった。

表 2 類似画像検索の性能

検索質問	再現率	適合率	F 値
Abu Mena	0.58	0.37	0.43
Aksum	0.86	0.21	0.32
Basilica di san Pietro	0.56	0.14	0.21
Big Ben	0.73	0.39	0.48
Buckingham Palace	0.45	0.1	0.17
Cliff of Bandiagara	0.6	0.08	0.14
Colossi of Memmon	0.42	0.19	0.26
Diamond Head	0.43	0.29	0.35
Eiffel Tower	0.7	0.95	0.8
Golden Gate Bridge	0.77	0.49	0.6
Grand Canyon	0.59	0.23	0.33
Greenwich Observatory	0.33	0.07	0.11
Haleakala Volcano	0.73	0.1	0.18
Kairouan	0.48	0.4	0.44
Tower Bridge	0.52	0.76	0.62
平均	0.58	0.31	0.36

#### 4.2 実験 1:類似画像検索の性能

実験 1 では, Color signature を使った類似画像検索の適合率を調べた. 実験は各検索質問毎に得られた画像集合を用いた. 各画像集合において, 無作為に 5 枚の問い合わせ画像を選択し, 類似画像検索を行った.

検索結果の評価としては, 再現率  $R = |C|/|A|$ , 適合率  $P = |C|/|B|$ ,  $F$  値  $= \frac{2RP}{R+P}$  を用いた. ここで  $A$  は問い合わせ画像と類似している画像の数,  $B$  は結果として得られた画像の数,  $C$  は結果として得られた類似画像の数である.

実験結果を表 2 に示す. 表 2 から分かるように, 「Eiffel Tower」と「Tower Bridge」ではほとんどの画像が意味的に類似した画像であったため適合率が高かった. しかし, その他の検索質問においては適合率の低さが目立った. 一方, 再現率は平均で 0.58 となり適合率ほどは低くなかった.

この結果から, Color signature を用いた類似画像検索で得られる結果には, ある程度, 類似画像を含んでいるが無関係な画像が多いということが分かった.

#### 4.3 実験 2:キーワードに関する実験

キーワードに関する実験の評価には説明文に関する実験の評価には再現率  $R = |C|/|A|$ , 適合率  $P = |C|/|B|$ ,  $F$  値  $= \frac{2RP}{R+P}$  を用いた. ここで,  $|A|$  は画像の説明文中のキーワードの数,  $|B|$  は結果として得られたキーワードの数,  $|C|$  は結果として得られた正しいキーワード数である. また, キーワードが存在する画像から, そのキーワードを全く抽出できなかった場合は, 再現率, 適合率ともに 0 であるとした.

実験結果を表 3, 表 4 に示す. この表が示すように, 提案手法は比較手法と比べて再現率, 適合率ともに 2 倍以

表 3 比較手法で抽出したキーワードの評価

検索質問	再現率	適合率	F 値
Abu Mena	0.12	0.06	0.08
Aksum	0.47	0.04	0.08
Basilica di san Pietro	0.18	0.14	0.16
Big Ben	0.22	0.28	0.24
Buckingham Palace	0.42	0.22	0.29
Cliff of Bandiagara	0.07	0.05	0.06
Colossi of Memmon	0.09	0.08	0.09
Diamond Head	0.22	0.15	0.18
Eiffel Tower	0.14	0.04	0.06
Golden Gate Bridge	0.27	0.08	0.12
Grand Canyon	0.25	0.04	0.07
Greenwich Observatory	0.55	0.09	0.15
Haleakala Volcano	0.27	0.27	0.27
Kairouan	0.15	0.04	0.06
Tower Bridge	0.19	0.1	0.13
平均	0.24	0.11	0.14

表 4 提案手法で抽出したキーワードの評価

検索質問	再現率	適合率	F 値
Abu Mena	0.24	0.08	0.12
Aksum	0.82	0.15	0.25
Basilica di san Pietro	0.6	0.35	0.44
Big Ben	0.74	0.46	0.57
Buckingham Palace	0.88	0.44	0.58
Cliff of Bandiagara	0.33	0.1	0.15
Colossi of Memmon	0.44	0.46	0.45
Diamond Head	0.65	0.18	0.28
Eiffel Tower	0.72	0.22	0.34
Golden Gate Bridge	0.57	0.33	0.42
Grand Canyon	0.53	0.06	0.1
Greenwich Observatory	0.82	0.07	0.13
Haleakala Volcano	0.58	0.31	0.4
Kairouan	0.42	0.26	0.32
Tower Bridge	0.51	0.3	0.38
平均	0.59	0.25	0.33

上高いという結果を得た. しかし, 提案手法の再現率, 適合率は良いとは言えないレベルに留まっている.

この結果から,  $\chi^2$  検定を使ったキーワード抽出では画像のキーワードを正しく且つ十分に得ることは困難であるということが分かった.

#### 4.4 実験 3:説明文に関する実験

説明文に関する実験の評価には再現率  $R = |C|/|A|$ , 適合率  $P = |C|/|B|$ ,  $F$  値  $= \frac{2RP}{R+P}$  を用いた. ここで,  $|A|$  は Web ページ内で画像を説明している文の数,  $|B|$  は結果として得られた文の数,  $|C|$  は結果として得られた正解文の数である. また, 説明文が存在する画像から, その説明文を全く抽出できなかった場合は, 再現率, 適合率ともに 0 であるとした.

表 5 比較手法で抽出した説明文の評価

検索質問	再現率	適合率	F 値
Abu Mena	0.23	0.14	0.17
Akuseum	0.48	0.18	0.26
Basilica di san Pietro	0.81	0.52	0.64
Big Ben	0.33	0.34	0.34
Buckingham Palace	0.65	0.65	0.65
Cliff of Bandiagara	0.20	0.19	0.19
Colossi of Memnon	0.13	0.06	0.08
Diamond Head	0.39	0.36	0.37
Eiffel Tower	0.36	0.35	0.35
Golden Gate Bridge	0.78	0.43	0.55
Grand Canyon	0.25	0.25	0.25
Greenwich Observatory	0.70	0.70	0.70
Haleakala Volcano	0.63	0.46	0.53
Kairouan	0.41	0.26	0.32
Tower Bridge	0.44	0.43	0.43
平均	0.45	0.35	0.39

表 6 提案手法で抽出した説明文の評価

検索質問	再現率	適合率	F 値
Abu Mena	0.85	0.40	0.55
Akuseum	0.83	0.47	0.60
Basilica di san Pietro	1.00	0.64	0.78
Big Ben	0.91	0.65	0.76
Buckingham Palace	0.95	0.74	0.83
Cliff of Bandiagara	0.78	0.39	0.52
Colossi of Memnon	0.97	0.94	0.95
Diamond Head	0.90	0.56	0.65
Eiffel Tower	0.88	0.52	0.65
Golden Gate Bridge	0.89	0.44	0.59
Grand Canyon	0.95	0.88	0.92
Greenwich Observatory	1.00	0.99	1.00
Haleakala Volcano	0.93	0.74	0.82
Kairouan	1.00	0.63	0.77
Tower Bridge	0.94	0.86	0.90
平均	0.92	0.65	0.75

実験結果を表 5, 表 6 に示す. この結果が示すように, 提案手法は比較手法と比べて, 再現率, 適合率ともかなり高いという結果を得た. この要因として, 提案手法の適用範囲の広さが挙げられる. 比較手法は画像の alt テキストが記述されていることが前提であるため, alt テキストが記述されていない画像に対して適用できない. それに対し提案手法では alt テキストが無い場合にも適用することが出来る. 実際に, 比較手法では実験で使った画像全体の 55%にしか適用できなかったのに対し, 提案手法では 95%の画像に対して適用することが出来た. 提案手法が適用出来なかった残り 5%の画像は, alt テキストが記述されてなく, 類似画像からキーワードを得ることが出来なかった画像である.

また, 4.3, 4.4 の実験により, 少数の正しいキーワードを得ることが出来れば, それらを含む文を抽出することによって, 正しく説明文を抽出することが出来る場合が多いということが分かった.

提案手法の成功例を図 1 に示す. 図 1 の例では「open」以外は正しいキーワードを得ることが出来た. また, その結果得られた説明文も正しいものであった.

## 5. 考 察

### 5.1 類似画像検索およびキーワード抽出の性能と提案手法の性能

類似画像検索の性能が低いにもかかわらず, 説明文抽出の段階では, 適合率の平均が 0.75 という結果を得た. この要因として

- (1) Google のテキスト検索による Web ページの選別
- (2) 3.3(3) の処理による画像に無関係な語の排除

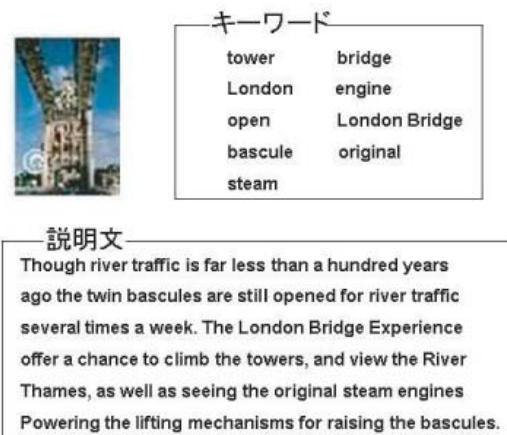


図 1 提案手法の成功例

の 2 つの言語的処理によって類似画像検索の性能の低さを補っていることが挙げられる.

まず, (1) によって, 検索質問と関連のある Web ページの集合を得ることが出来る. このような Web ページ中に存在している画像には, その検索質問に関係ある画像が多く含まれる. 実際に各検索質問ごとの画像集合では, 検索質問に何らかの関係のある画像が 90%以上含まれているが多かった.

また, (2) によって, 画像と意味的に全く異なる語を排除することが出来る. 実際に 4.3 の実験では, このような語はほとんど見受けられなかった.

このように, 言語的な特徴量を用いることで, 従来の画像特徴量だけを用いた類似画像検索で生じる意味的なギャップを埋めることができると考えられる.

### 5.2 誤りの解析と今後の課題

画像とその説明文が存在する Web ページはつぎの 3 つ

## Tower Bridge



### PROJECT OVERVIEW

In London, England is one of the most recognized bridges in the world, the Tower Bridge. Designed by Sir Horace Jones and engineered by Sir John Wolfe-Barry, the construction of the Tower Bridge began in 1896. After Jones' death in 1897, Barry was able to invoke greater artistic freedom than he would have under the original designers' supervision. The project was completed in 1894. The towers, covered in ornate Victorian masonry, stand 206 ft., 63 m. tall and are constructed of a steel frame. The structure has a main span of 200 ft., 61 m., with a bascule roadway and two steel walkways which are 140 ft., 43 m. high. Popularity has increased since the bridge was first opened, and it is now one of London's best loved landmarks.

図 2 1 枚の画像に説明文がある例

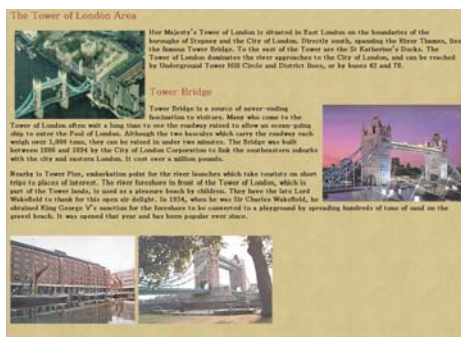


図 3 複数の画像に対する説明文が 1 つある例

に分類できる。

- (1) 画像 1 枚とその説明文がある
- (2) 複数の画像に共通の説明文がある
- (3) 複数の画像の各々に説明文がある

それぞれの Web ページの例を図 2 から図 4 に示す。提案手法では (1), (2) の Web ページに関しては、ほとんどの場合、正しく説明文のみを抽出することが出来た。しかし、(3) の Web ページに関しては誤抽出が多かった。特に、図 4 のような画像と説明文のペアが繰り返し出現するような Web ページのを対象にしたときの誤抽出が目立った。

このような Web ページに存在する画像は意味的に類似している画像が多く、説明文に含まれているキーワードも類似しているために、1 枚の画像に対してすべての画像の説明文を抽出していることが多く見受けられた。このような誤りが誤抽出全体の 69% を占めていた。

このような構造の Web ページ中には、多くの画像が存在する。その各々の説明文の適合率がかなり低いため、これらの説明文が、適合率の平均を大きく引き下げている。

このような Web ページから正しく説明文を抽出するための手段として、Web ページのレイアウト情報が有効であると考えられる。



図 4 複数の画像の各々に説明文がある例

## 6. まとめ

本稿では、HTML の alt テキストと類似画像から得られるキーワードを利用して、Web 画像の説明文を抽出する手法を提案した。提案手法では言語的な特徴と、画画的な特徴を組み合わせることで、これらの手法を単独で用いたときの欠点を補うことができるという利点がある。実験の結果、alt テキストを使う手法と比べ、F 値が約 2 倍であることが分かった。

今後の課題として、画像と説明文の繰り返し構造が存在するような Web ページに対して、説明文の適合率を上げるための工夫があげられる。その手段として Web ページのレイアウト情報の利用などを考えている。

## 文 献

- [1] 竹内謹治, 黄瀬浩一, 松本啓之亮: “語の共起の統計情報を用いた画像と説明文の対応付け”, 電子情報通信学会全国大会, D-5-1 (2005).
- [2] V. Gudivada. and V. Raghavan: “Content-Based Image Retrieval-Systems”, IEEE Comput., Vol.28, No.9, pp18-22 (1995).
- [3] A.A. Goodrum, “Image information retrieval: An overview of current research”, Inf. Sci., Vol.3, pp.63-66 (2000).
- [4] 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, Vol.J87-D1, No.2, pp.145 - 153 (2004).
- [5] E.V. Munson and Y. Tsybalenko: “To search for images on the web, look at the text, then look at the images”, Proc. First International Workshop on Web Document Analysisism pp.39-42 (2001).
- [6] 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, DE2005-136 (2005).
- [7] J. Fürnkranz: “A Study Using N-grams Features for Text Categorization”, Technical report, Austrian Research Institute for Artificial Intelligence, OEFAITR-98-30 (1998).
- [8] Y. Rubner, C. Tomasi, and L.J. Guibas. “The earth mover's distance as a metric for image retrieval”. International Journal of Computer Vision, Vol.40, No2, pp.99-121 (2000).