

D-006

## ブログへの自動画像付与のためのソーシャルアノテーションの有用性の検証 A Study on Effectiveness of Using Social Annotations for Automatic Image Insertion for Weblogs

今野 智明†  
Tomoaki Konno

明堂 絵美†  
Emi Myodo

高木 幸一†  
Koichi Takagi

川田 亮一†  
Ryoichi Kawada

### 1. はじめに

ブログ投稿者は、日々の具体的な出来事に関する記事に対し、しばしば自分で撮影した写真を付与する。また、記事に出現する具体的な事物を表す画像の付与を支援するシステムも存在する[1][2]。

一方、ポエムのようなブログ記事に対し、その雰囲気を表す画像を付与する場合も考えられる。雰囲気に適した画像をブログに付与するためには、ブログからの印象抽出と印象による画像検索の2つの課題が考えられる。前者は、人間が印象を判断すれば容易に抽出可能であるが、後者は適した画像を手動で作成または検索する必要があるためコストは大きい。本稿では、ポエムのようなブログ記事に対してその印象を表す言葉（印象語）に基づき画像を簡易に取得する方式に焦点を当てる。

印象語のような抽象語による検索の精度は、具体語と比べて低いとされている[3]。そこで、そのようなシステムの実現のために、画像共有サイトのタグやコメント（ソーシャルアノテーション）の利用が考えられる[4]。タグには印象語が少ない[5]が、コメントには多くの印象語が含まれるため、印象語と画像の対応付けにはコメントの方が有用であると予想できる。

さらに、印象語に適した画像をブログへ付与する際には、次のような問題が考えられる。例えば、「犬」の記事を書いていて、その記事の雰囲気が「かわいい」であるとする、その雰囲気を「かわいい猫」の画像で表すのは望ましくない。すなわち、具体的なものを表す画像（具体画像）より、抽象的な画像（抽象画像）の方が有用であると考えられる。

本稿では以上に鑑み、印象を表す画像を画像共有サイトから集めるに際し、ソーシャルアノテーションを活用する手法の有用性を定量的に比較評価する。

### 2. 印象語による画像取得

図1は、タグとコメントをそれぞれ利用した場合の画像取得の枠組みを示す。コメントを利用して構築した画像データベースとタグを利用して構築した画像データベースをそれぞれ用いた場合を比較する。

コメントを利用した場合、印象語と画像の関連度を、当該印象語を含むコメント数として計算し[4]、その値の大きい順に選択する（Comment A方式）。一方で、タグを利用した場合、その関連度は大小が定義できないため、該当する画像をランダムに選択する（Tag方式）。

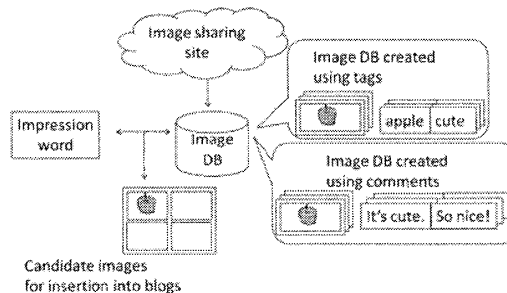


図1 印象語による画像取得の枠組み

### 3. 抽象画像取得方式

汎用的に利用できる抽象画像を取得するために、Comment A方式の改良について検討する（Comment B方式）。印象語と関連の強い単語（頻出共起単語）をコメントまたはタグに含むような画像は具体画像であると仮定し、この具体画像を検索対象から除外することで、抽象画像を得るというアプローチをとる。

まず、以下のステップで、各印象語に対して頻出共起単語のリストを作成する。

- Step 1. 印象語をクエリとして web テキストを検索
- Step 2. 取得した web ページから、共起単語を抽出
- Step 3. 印象語と共起単語の組み合わせで、Phrase 検索
- Step 4. 検索ヒット数順に頻出共起単語をソート

次に、このリストを利用して、以下のステップで、印象語から抽象画像を取得する（図2）。

- Step 1. 印象語に対応する頻出共起単語を N 個抽出
- Step 2. 印象語と頻出共起単語をクエリとして、画像データベースに対して検索
- Step 3. 印象語をコメントに含み、頻出共起単語をコメントまたはタグに含まない画像のみを取得

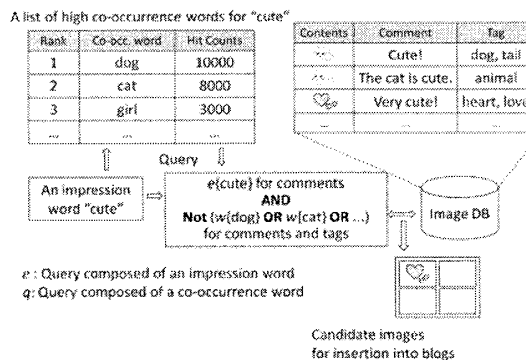


図2 抽象画像取得の方式

† 株式会社KDDI研究所, KDDI R&D Laboratories Inc.

## 4. 評価実験

### 4.1 実験条件

印象語と選択された画像群との関係性について評価実験を行った。被験者 21 名(男性 11 名, 女性 9 名)に対して, 印象語とそれに対応する 10 枚 1 組の画像セットを提示して, 適合度を評価した。Comment A, Comment B, Tag の 3 つの方式によって得られた画像セットを利用した。評価尺度は 1-5 (低-高) の連続値である。印象語は, 文献[6]中の 160 語を利用し, 実験ではそのうちの 10 語を評価に用いた。画像は Flickr[7]から取得し, 2004/10 から 2010/05 までの間に投稿された約 100 万枚の画像に対するコメントおよびタグを利用した。印象語をコメントに含む画像数は約 2 万枚で, タグに含む画像数は約 1 万枚であった。

B 方式においては, 頻出共起単語の抽出のために Yahoo! ウェブ検索[8]を利用した。上位 30 個の頻出共起単語を利用した。その際, Word-Net [9]で名詞と判定された語のみを対象とし, さらに“画像”や“Web”に関する語は, 印象語に関わらず頻出するため共起単語から除外した。

さらに, 5 名の被験者に対して, A 方式および B 方式によって得られた画像セットの対を提示し, 1 人当たり 10 対の画像セットに関して, どちらがより抽象的であるかについて一対比較法によって評価した。

### 4.2 実験結果

図 3 は, 各 3 つの方式を利用した場合の, 印象語と画像セットの適合度を示す。

A 方式と Tag 方式を比較すると, A 方式のほうがほとんどのケースで適合度が高いことがわかる。A 方式では平均適合度 3.5 で Tag 方式よりも 0.6 程度高く, その差は有意であった ( $t$ -test,  $p < 0.01$ )。B 方式についても Tag より高い適合度が得られている。コメントを利用した方式が有効であることが確認できる。また, A 方式と B 方式を比較すると, ほぼ同等である。A 方式と B 方式の違いは得られた画像の抽象度のみであるため, これは想定される結果である。

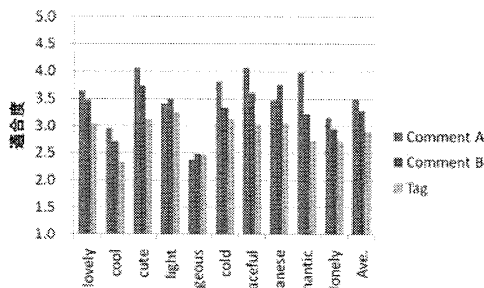


図 3 印象語と画像セットの適合度

表 1 一対比較法による抽象度評価結果

|        | lovely | light | Japanese | ... | A の個数 | B の個数 |
|--------|--------|-------|----------|-----|-------|-------|
| Rater1 | A      | B     | B        | ... | 5     | 5     |
| Rater2 | B      | B     | B        | ... | 3     | 7     |
| Rater3 | A      | B     | B        | ... | 7     | 3     |
| Rater4 | B      | B     | B        | ... | 5     | 5     |
| Rater5 | B      | B     | B        | ... | 3     | 7     |

表 2 抽出された頻出共起単語の一部

| Rank | lovely | cool   | cute | ... |
|------|--------|--------|------|-----|
| 1    | people | water  | pet  | ... |
| 2    | lady   | people | baby | ... |
| 3    | wife   | news   | girl | ... |

また, 表 1 は抽象度の評価結果を示す。表中の“A”および“B”は, それぞれ A 方式, B 方式で得られた画像のほうが抽象的であると判断されたことを意味する。A 方式が選択された回数の評価者平均は 4.6, B 方式では 5.4 であり, 明確な違いは見られなかった。各印象語について見てみると, “light”や“Japanese”については, 全評価者が B 方式で得られた画像の方が抽象的であると判定した。これらに関しては, 30 個程度の頻出共起単語を含む画像を除去すればより抽象的な画像が取得できるということがわかる。一方, “lovely”に関しては, 依然として具体画像が残っているため, 評価のばらつきが出ており, さらなる改良が必要であると考えられる。

B 方式の効果が明確に見られなかったさらなる原因として, 頻出共起単語の抽出精度が考えられる。表 2 に抽出した具体物を表す頻出共起単語を示す。例えば, “cute”に対して“pet”や“baby”が抽出されているため, 印象語と関連のある具体語を意図通り抽出できていることがわかる。一方で, “cool”に対して“news”という抽象的な単語が抽出されている。こういった単語を含む画像を除外すると抽象的な画像も除外してしまう可能性がある。すなわち, 印象語に関連があり, かつより具体的な単語を抽出する必要があり, これも今後の課題とする。

## 5. まとめ

本稿では, ブログへの自動画像付与のために, 印象語による画像取得におけるタグとコメントの有用性について定量的に比較評価を行った。コメントを利用した場合, タグを利用した場合よりも 5 段階主観評価で 0.6 ほど適合度が高く, 有用であることを確認した。さらに, コメント利用方式において, 抽象的な画像を取得するための検討を行った。一定の効果が確認されたが, さらなる洗練化が必要である。

## 参考文献

- [1] 佐藤圭太ら, “ブログ記事からのイベント文抽出によるシーンの生成”, JSAI2007, 2F4-5, 2007
- [2] Zemanta, <http://www.zemanta.com>
- [3] Kato et al., “Can Social Tagging Improve Web Image Search?,” Proc. of WISE, 2008
- [4] 中村聡史ら, “ソーシャルアノテーションに基づく動画検索手法”, DEIM, D6-1, 2009
- [5] Bischoff et al., “Can All Tags be Used for Search?,” Proc. of CIKM, 2008
- [6] 南雲治嘉, “配色イメージチャート”, グラフィックスクス社, 2000
- [7] Flickr : <http://www.flickr.com>
- [8] Yahoo! デベロッパーネットワーク, <http://developer.yahoo.co.jp/webapi/search/websearch/>
- [9] Miller et al., “Introduction to Word-Net: An On-line Lexical Database.,” International Journal of Lexicography Vol.3, No.4, 2004