

画像検索における多言語クエリの効果

高橋 有礼[†] 菱山 玲子[†]

早稲田大学理工学術院 創造理工学部経営システム工学科[†]

1 研究背景と目的

Web上で蓄積されるマルチメディアコンテンツの増加に伴い、これらのマルチメディア情報の取得に様々な手法が提案されている。またその1つの手法として、タグ付けによるコンテンツ内容の記述からマルチメディア情報を取得する手法がある。しかしタグの質が低いためコンテンツ内容の十分な補足はなされない。そのためユーザの意図を反映させた検索を行うことが困難であり、それがマルチメディア情報取得の新たな課題として挙げられている [1]。またコンテンツ内容の記述はもともと多言語で付与されており、単言語によるタグ情報を利用するだけではコンテンツの適切な取得が困難と予測される。

その他にも詳細であるため記述されないクエリや、ユーザが書き起こせないクエリが存在する。ゆえにユーザによるクエリ作成の難しさが課題として挙げられる。

本研究ではマルチメディアコンテンツの1つである画像に焦点を当て、多言語クエリを利用し、テキストおよび画像に含まれる情報の効果的利用から正解画像に至るまでのプロセスを改良することで、より適切な画像の入手を実現する。

2 関連研究

Zakariaら [1] はメタデータに依存することなく、ユーザ主体のタグ付けや地理参照メタデータによって生成された集約的な知識と組み合わせることでコンテンツ内容記述の充実化を図った。だがユーザの意図を直に反映させることはできない。

ユーザは言語横断検索時は他の言語を利用しなくてはならない。そこで Kristenら [2] はクエリと文章の双方を翻訳することで検索精度の向上を提案している。

従来のフィードバック方法では選択した画像に対して十分に満足している場合のみを考慮していた。そこで中島ら [3] はユーザの意図をより強く反映させる差異増幅型適合フィードバックを提案している。この方法はユーザが選択したものとしなかったものの差異を見ることでユーザの意図をより強く反映させている。

3 提案

本研究ではユーザが期待する画像を容易に取得できるように、ユーザの意図を反映しながら正解画像を入手する過程を改良する。この目的を達成するには、多くの適合画像を簡単な選択で入手することが必要である。そのためにユーザが検索結果に対してフィードバックをか

ける。しかしフィードバックをかけるためには、選ぶ画像群の中に適合画像が必要であり、適合画像は言語によっては記述されていないものが存在する。そのため画像検索時にクエリを翻訳することにより、コンテンツ内容の記述を補填するように、検索にかかるタグの量を増やすことで検索精度を向上させた。本研究の流れは以下のようになる。

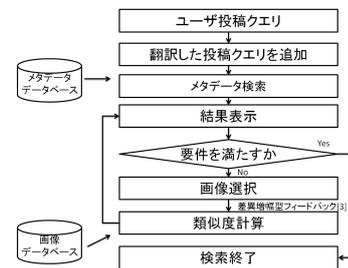


図1. 提案フロー

初めにユーザが入力したクエリをもとに検索結果画像を表示する。要件を満たさない場合、ユーザは提示された画像の中から任意の画像を選択しフィードバックをかける。そして選択された画像とデータベース内の画像に対して類似度計算をする。そして検索結果画像が再度表示される。

差異増幅型フィードバック [3] で用いる関数 f_{DA} は正事例 S_{pi} 、正事例の周辺画像 $neighbor(S_p)$ 、正事例の周辺画像数 $N_{S_{pi}}$ で定義される。 α は差異増幅の増幅係数であり、 ϵ は差異を増幅させるかの閾値である。 S_{pi} に基づく次の検索質問 Q_{i+1} を以下の式 (1) と定義する。

$$Q_{i+1} = (1 - \beta) \times Q_i + \beta \times f_{DA}(S_p, neighbor(S_p)) \quad (1)$$

ただし

$$f_{DA}(S_{pi}, neighbor(S_{pi}))$$

$$= \begin{cases} S_{pi} + \alpha \times \frac{\sum_{neighbor(S_{pi})} (S_{pi} - S)}{N_{S_{pi}}} & \left(\left| \frac{\sum_{neighbor(S_{pi})} (S_{pi} - S)}{N_{S_{pi}}} \right| \geq \epsilon \right) \\ S_{pi} & \left(\left| \frac{\sum_{neighbor(S_{pi})} (S_{pi} - S)}{N_{S_{pi}}} \right| < \epsilon \right) \end{cases}$$

4 実験

提案を評価するための実験は、各国語によるクエリによる検索結果比較 (実験1) と各国語によるクエリの組み合わせによる検索結果比較 (実験2) により行った。

4.1 実験1の設定

実験1は、次の2種の実験からなる。実験1.1では検索対象を「見上げた角度、赤色、エッフェル塔」を満たすものとし、対象の画像は Flickr [4] からエッフェル塔関連のタグがつけられた画像700枚ほどを取得した。実験1.2で

*An effect of multilingual query in image retrieval

[†]Arinori TAKAHASHI, Reiko HISHIYAMA

[†]School of Creative Science and Engineering, Waseda University

は検索対象を「ロゴを読むことができる、赤色、アリアンツアリーナ」を満たすものとし、対象の画像は Flickr[4] からアリアンツアリーナ関連のタグがつけられた画像 700 枚ほどを取得した。また取得した画像に付与されたタグも加えることでデータベースを構築した。検索成功条件は適合率が 1/3 以上となった時とし、検索失敗条件は適合率が 0 の時とした。画像の選択基準は検索結果最上位の画像とした。ただし同じ画像が 2 回続いた場合は検索結果 2 番目の画像としている。実験は以下の単言語クエリで検索を行う。

実験 1.1

- Tour+Eiffel+Rouge(仏)
- Eiffel+Tower+Red(英)
- エッフェル塔+赤(日)

実験 1.2

- Allianz+Arena+Rot(独)
- Allianz+Arena+Red(英)
- アリアンツアリーナ+赤(日)

4.2 実験 2 の設定

実験 2 では実験 1 のクエリを現地言語に翻訳したものと、英語に翻訳したものをクエリに追加した 2 種類 3 パターンの多言語クエリで検索を行う。実験は以下の多言語クエリで検索を行う。

実験 2.1

- Eiffel+Tower+Red+Tour+Eiffel+Rouge(英+仏)
- エッフェル塔+赤+Tour+Eiffel+Rouge(日+仏)
- Eiffel+Tower+Red+エッフェル塔+赤(英+日)

実験 2.2

- Allianz+Arena+Red+Allianz+Arena+Rot(英+独)
- アリアンツアリーナ+赤+Allianz+Arena+Rot(日+独)
- アリアンツアリーナ+赤+Allianz+Arena+Red(日+英)

5 結果と考察

5.1 結果

各実験の、開始時と終了時の適合率を表 1, 表 2, 表 3, 表 4 に示す。

表 1. 実験 1.1 結果

	開始時適合率	終了時適合率
仏	27	47
英	10	33
日	0	0

表 2. 実験 1.2 結果

	開始時適合率	終了時適合率
独	27	57
英	30	53
日	0	0

表 3. 実験 2.1 結果

	開始時適合率	終了時適合率
英+仏	27	47
日+仏	27	47
日+英	13	43

表 4. 実験 2.2 結果

	開始時適合率	終了時適合率
英+独	27	53
日+独	30	53
日+英	36	36

5.2 考察

表 1, 表 2 では現地言語では開始時で多くの適合画像が入手できるが、日本語では 1 件も出力されなかった。そのため検索クエリの言語は開始時の検索結果に影響を及ぼす。表 3, 表 4 ではユーザのクエリを現地言語に翻訳し追加することで追加前は適合画像が 1 枚もなかった言語でも、追加後は多数の適合画像が入手できている。実験 2.2 において英と独が適合率に差が余り見られないのは、対象の表記名が英、独で変わらないため言語による差があまり出なかったからである。

以上のことから適合画像は、ユーザの使用言語でタグ付けされたものよりも検索対象の現地の言語でタグ付けされたものに多く存在する。そしてクエリに現地の言語を翻訳したものを追加することで、開始時に多くの適合画像を得ることができる。

6 まとめと今後の課題

本研究ではユーザが求める画像を入手する過程を改良することを目的としている。検索クエリの言語は検索結果に影響を及ぼす。そして画像に対しフィードバックをかけることは、ユーザの意図を簡単に検索結果に反映させる。以上よりクエリを多言語化しユーザが画像に対してフィードバックをかけることで適合画像を容易に取得することができる。

今後の課題としてはクエリが複雑で翻訳が困難な場合を考慮する必要がある。

参考文献

[1] Lailatul Qadri Zakaria, Wendy Hall, and Paul Lewis:Modelling image semantic descriptions from web 2.0 documents using a hybrid approach, *Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services*, pp.306-312, 2009.

[2] Parton Kristen, Kathleen R. McKeown, James Allan, and Enrique Henestroza:Simultaneous multilingual search for translanguing information retrieval, *Proceedings of the 17th ACM conference on Information and knowledge management*, pp.719-728, 2008.

[3] 中島伸介, 木下真一, 田中克己:差異増幅型適合フィードバックに基づく画像データベース検索, 電子情報通信学会論文誌 D-1, Vol.87, No.2, pp164-174, 2004

[4] Flickr, URL: <http://www.flickr.com/>, (2013 年 12 月 10 日アクセス)