意味ベクトルによる画像検索

田中理恵子*¹, 芥子育雄*², 池内洋*²

Image retrieval using semantic vectors

TANAKA Reiko, KESHI Ikuo, IKEUCHI Hiroshi

[著者抄録] 従来のキーワード検索では、キーワードを思い出せないと検索できない、検索結果を順位付けすることができないなどの問題があった。これらの問題を解決し、キーワード1個からでも連想検索ができる検索手法として意味ベクトルによる連想検索技術を開発中である。この技術画像検索へ応用した画像データベースを試作した。風景写真など約170枚の画像と、それぞれに付与した説明文からなる画像データベースを構築し、意味ベクトルを用いた検索を行ったところ、素人によって数語程度の説明文をつけただけでも、かなりの検索ができることがわかった。

[著者付与キーワード] 意味ベクトル, 連想検索, 画像データベース, 画像検索

[Author Abstract] In this paper we discuss how to create an image database with a good user interface using semantic vectors. The interface realizes content-based associative retrieval without using a thesaurus, and hypertext-like retrieval without linking terms explicitly. Semantic vectors, which represent one-layered semantic networks as fixed-dimensional vectors, are constructed from encyclopedia texts. The database includes about 170 images and each of them are attached by the caption which describes the image in several words. We emphasize the effectiveness of indexing and retrieving images using semantic vectors through some experiments.

[Keywords by Author] semantic vectors, associative retrieval, image database, bootstrap system

-
- * 1 シャープ(株) 技術本部ソフトウェア研究所 (〒632 天理市櫛本町2613-1) Tel. 07436(5)0987
Software Research Laboratories, Corporate Research and Development Group, Sharp Corporation
(2613-1, Ichinomoto-cho, Tenri-shi, 632)
- * 2 シャープ(株) 技術本部情報技術研究所 (〒632 天理市櫛本町2613-1) Tel. 07436(5)2468
Information Technology Research Laboratories, Corporate Research and Development Group, Sharp Corporation (2613-1, Ichinomoto-cho, Tenri-shi, 632)

1. はじめに

絵画や写真などの画像をデータベースに登録し、思い通りの画像を必要に応じて取り出すことができれば、産業分野のみならず、人々の文化活動においても大きな意味があるだろう。しかし、今日、画像データベースを必要とする人々やその理由は多岐にわたってきており、それらすべての要求に答えるようなシステムを提供することは至難の技である。

例えば、従来のテキストデータベースで使用されているキーワード検索を画像データベースに応用した場合、あらゆるユーザの要求を満たすように、キーワードを選択して付与する必要がある。しかも、ユーザごとに感性は異なるため、「明るい、優雅」といった印象語による主観的な検索をすることができないという問題がある。

このような問題を解決し、主観的な検索を行うために、例示画をキーとして類似画検索を行う手法¹⁾や、印象語による感性検索の手法²⁾が提案されている。

本論文で述べる検索方法は、加藤の分類によれば、文字・画像型による検索に当たる³⁾。ユーザは、自然言語で質問文を入力する。それに対してシステムは、その質問文の内容にふさわしい画像を検索する。質問文と画像の類似度を比較するために、それぞれを特徴ベクトルによって表現する。質問文の特徴ベクトルには、意味ベクトル⁴⁾を用いる。画像の特徴ベクトルは、「時計のある建物。街灯。橋の上の人。」のような、画像に付与された説明文から、意味ベクトルを作成して、それを特徴ベクトルとする。

ベクトルを用いた連想検索方式については、柴田らの研究がある⁵⁾。この研究では、画像に付与するキーワードはあらかじめ登録されたものを使用する。これらのキーワードのシソーラスを作成し、これに基づいて人がキーワードの重み付けをして画像をベクトル表現する。そのため、キーワードが限られるため広範な分野に適用できない、シソーラスを作成する負担が大きい、キーワードの

重み付けが試行錯誤となり適切な重み付けが難しいなどの問題がある。

これに対し、本論文で提案する手法は、百科事典を元に生成した10万単語以上の意味ベクトル辞書を用いるため、キーワードの制限がなく、文中での出現頻度に基づいた重み付けを自動的に行うことができる。また、シソーラスではなく百科事典を元にしたことにより、意味の重み付けにバイアスがないという特徴がある。

2. 意味ベクトルによる検索

意味ベクトルとは、単語の意味のベクトル表現であり、FEATURE（求める画像の特徴を記述する要素）との関係をベクトル表現したものである⁴⁾。すなわち、 n 個の概念分類をFEATUREとし、各次元が一つのFEATUREに対応した n 次元のベクトル空間上の一点で、意味を表現するものである。例えば、「人間、悲しい、芸術、科学、興奮、政治」の6個のFEATUREを選択した場合を考えると、単語「パイロット」の6次元の意味ベクトルは、 $[1, 0, 0, 1, 1, 0]$ となる。

百科事典テキストデータベースに応用した例⁴⁾では、128個のFEATUREによる128次元の意味ベクトルを使用している。FEATUREには、百科事典のすべての内容を分類できるように、128種類を選択した。例えば、政治や経済の分野からは、「法律、政治、経済、福祉、税金」など、文明や社会の分野からは「コンピュータ、ソフトウェア、子供、大人、老人、住宅、犯罪」など、行動の分野からは「所有、開発、生産、書く、話す」などを選択した。また、形容詞としては「大きい、小さい、楽しい、悲しい」などが含まれている。

百科事典中の単語すべてについて、128個のFEATUREとの関連を付けるのは現実的ではない。そこで、ブードストラップ学習と呼ぶ手法を開発した⁴⁾。これは、まず、コア単語と呼ぶ約3,500語に人手によりFEATURE付けを行い、意味ベクトルを作成してコア単語辞書を作成し、次に、このコア単語辞書の意味ベクトルを元に、説明文

中に表る約10万単語の意味ベクトル辞書を自動生成するものである。

こうして作成した意味ベクトル辞書から、百科事典テキストデータベースを作成する。百科事典中の各項目について、説明文中に含まれる単語の意味ベクトルの加重和を求めて、その項目の意味ベクトルとする。検索の際は、質問文に含まれる単語の意味ベクトルの加重和を求めて、質問文の意味ベクトルとする。そして、質問文の意味ベクトルとの距離が近いものを検索する。

意味ベクトルを利用した検索の利点は次のようにまとめることができる。

- キーワードを思い出せなくても検索できる。
- キーワード1個などの短い質問文でも連想検索ができる。
- 検索結果をランク付けすることが可能である。
- 検索結果を次の質問文とすることで類似検索ができる。

3. 意味ベクトルによる画像検索

本章では、今回我々が試作した意味ベクトルによる画像検索システムについて述べる。

3.1 試作システムの概要

本システムは、自然言語文による質問文を入力とし、この質問文の意味ベクトルに近い特徴ベクトルを持つ画像を検索して、その結果をランク付けして表示するものである。画像の特徴ベクトルには、画像に付与された説明文から作成した意味ベクトルを使用する。

データベースに登録した画像は、自然の風景や都市の景観などの風景写真⁶⁾と、農業や漁業などに従事する人々やビルの建設現場といった各種の産業のイメージ写真⁷⁾の約170枚である。今回利用したこれらの写真には、あらかじめ簡単なタイトルが付けられている。風景写真の場合は主に地名が、産業のイメージ写真の場合は、「ビジネスイメージ」というタイトルが複数の写真につけられているものもあれば、「工業イメージ」の例のよう

に「送電線」や「ダム」などのサブタイトルがつけられているものもある。

これらの写真に、1枚につき数語程度の説明文を付与した。説明文は、単語の羅列の場合もあれば、短い文章となっているものもある。説明文は、写真を見て思いついた言葉を付与することとし、特別なルールを設けなかった。例えば、写真に写っているものすべてを列挙する、あるいは、他の写真との区別が付きやすいような特徴を取り上げるなどのルールを設けなかった。したがって、青い空や海が印象的な風景写真に「青い空」や「青い海」という説明がついている場合があっても、夕焼けの風景の写真に、「オレンジ色」などの説明がついているとは限らない。このような方法で、説明文／語を付与するのに要した時間は、入力も含めて1枚の画像につき約1分という極めて短い時間である。

次に、前述の百科事典テキストデータベースの構築の際に作成した意味ベクトル辞書を上述の方法で付与した説明文に適用させてから、意味ベクトルを作成し、これを画像の特徴ベクトルとして登録するとともに、画像データの登録も行い、画像データベースを作成した。

図1に、本システムの動作例を示す。左側は、本システムのユーザインタフェース部分のウィンドウである。図中の「キーワード文を入力して下さい。」と表示されているウィンドウで、ユーザが質問文の入力を行う。ユーザが質問文を入力すると、その質問文の意味ベクトルを作成し、登録されている画像の検索を行う。

検索結果は、質問文の意味ベクトルとの距離が近いものから順に、検索結果表示 WINDOW に表示される。このウィンドウの左端に見える数字が順位である。その横の点数は質問文の意味ベクトルと画像の特徴ベクトルの近さを100点満点で表したものである。順位と点数に並んで表示されているのは、各画像に付けられたタイトルである。

本システムでは、索引用の小さな画像は登録していないため、このように、画像に付けられたタイトルで検索結果をリスト表示する。順位付けさ

れた検索結果のリストから、表示させたいものをマウスクリックによって選択すると、新たにウィンドウが開き、画像が表示される。

画像にどのような説明文が付けられているかを確認するために、図1左下に示すように、検索文章の内容表示ウィンドウを設けている。このウィンドウ中の一番上の行は、システムの内部で使用するデータであり、画像の説明として付与したものではない。2行目に、この画像のタイトルが、その下に、説明文が表示される。この例では、タイトルが「ケニアアンボセリ」で、説明文が「縞馬の群。広い乾燥した草原。晴れた空。動物の親子。」である。画像の特徴ベクトルの作成には、説明文だけでなく、タイトルのテキストデータも使用している。

3. 2 試作システムでの画像検索

本システムを用いて画像検索の実験を行った。本節では、実験結果のいくつかを紹介する。

図1は、「広い空間」という質問文で検索を行い、検索の結果、第1位となった画像「ケニアアンボセリ」を表示したものである。この画像に付けられた説明文は、「縞馬の群。広い乾燥した草原。晴れた空。動物の親子。」であり、「広い」という単語が入っている。全データ中、「広い」という単語が説明文に入っているのは、これだけである。

次に、同じ質問文で第2位となった「工業イメージ(製図)」の画像を図2に示す。この画像の説明文は図2の左下のウィンドウに示されるように、「製図、設計、定規、直線」であり、「広い」や「空間」等の単語は含まれていない。しかし、製図版の平面が写真の外へと広がるこの写真から、「広い空間」を感じることができる。第3位以降には、第1位と同じような自然の風景が検索されている。

別の検索例を図3に示す。これは、「眺めの良い所」という質問文で検索を行った結果である。表示しているのは、第2位の画像である。「眺め」や「良い」などの単語が含まれていないにもかかわらず、検索されている。この写真の印象として、「眺めの良さ」が感じられるのではないだろうか。

この画像以外は、都市を高い位置から撮影したものが上位に選ばれている。第1位となった「ブラジル リオデジャネイロ」の説明文は「山の中の首都。高層ビル。遠景。」である。その他の画像の説明文にも、「眺め」や「良い」などの単語は含まれていない。

さらに、「研究」という質問文で検索し、第3位となった画像を図4に示す。この画像のタイトルは「工業イメージ(電波観測)」であり、説明文は「電波、観測、レーダー、宇宙」である。「研究」という言葉そのものは含まれていないが、世界各地の風景写真を多く含む本画像データベース中で、「研究」の持つイメージに近いものをうまく「連想」して検索できていると思われる。

質問文との類似度は100点満点としているが、図5のように、100点を超える場合もある。これは、意味ベクトルの圧縮に伴う誤差のためである。この検索例においても、「作物の収穫」という質問に対し、「農業イメージ」というタイトルが付けられた一連のシリーズが検索されており、質問文と画像の説明文のどちらもが非常に短い場合でも、妥当な連想検索ができることを示している。

このように、特別なルールを設けず、対象とする画像に関して専門的な知識を持たない素人が付与した簡単な説明を、画像の特徴ベクトルとしても、意味ベクトルによる連想検索によって、妥当な画像を検索することができた。画像データベースの作成が容易にできるという点、ならびに、画像データベースの検索に発見的要素が入るという点で、本手法が有効であると言える。

4. お わ り に

意味ベクトルによる検索技術を画像データベースに応用し、その有効性を確かめた。簡便な方法で付与した説明文から画像の特徴ベクトルを作成しても、妥当な連想検索が可能であることを示した。

今回利用した単語の意味ベクトル辞書は、百科事典を元に作成したものである。百科事典の場合は、全般的な事象について記述されており、本シ

(写真提供)イメージライブラリー

図1 “広い空間”で検索の結果、第1位に検索された画像

(写真提供)イメージライブラリー

図2 “広い空間”で検索の結果、第2位に検索された画像

(写真提供(株)イメージライブラリー)
図3 “眺めの良い所”で検索の結果、第2位に検索された画像

(写真提供(株)イメージライブラリー)
図4 “研究”で検索の結果、第3位に検索された画像

(写真提供(株)イメージライブラリー)

図5 “作物の収穫”で検索の結果、第4位に検索された画像

システムのように風景などを中心とした写真画像データベースの連想検索を行う場合には、ブートストラップ学習を、検索対象のテキストに対して繰り返し行うことで、各単語により適切なFEATUREがつくようになり、検索の精度が上がることを確認している。

また、今回は、被写体など画像の内容について

の専門的知識のないものが説明を付与したが、専門家が説明文を付与することで、より精度の高い検索ができるものと期待される。

現在、各ユーザの主観を反映できるように、意味ベクトルの適応的修正技術を開発中である。また、大規模なシステムの構築とその上での評価を行っていく予定である。

参 考 文 献

- 1) 栗田多喜夫, 下垣弘行, 加藤俊一. 主観的類似度に適応した画像検索. 情報処理学会論文誌, 31(2)227-237(1990)
- 2) 加藤俊一, 栗田多喜夫. 画像の内容検索—電子美術館への応用—. 情報処理, 33(5)466-477(1992)
- 3) 加藤俊一. 類似検索機能を持った画像データベースとその設計法—その1—. O plus E, (142)141-152(1991)
- 4) 芥子育雄, 池内洋, 小淵保司. 意味ベクトルによる百科事典テキストデータベースの構築. Proceedings of Advanced Database System Symposium '93, 227-234(1993)
- 5) 柴田正啓, 井上誠喜. 画像データベースの連想検索方式. 電子情報通信学会論文誌 D-II, J73-D-II(4)526-534(1990)
- 6) 世界の風景総集編. Vol. 3. 株式会社イメージライブラリー. (ImageCD)
- 7) ビジネス・産業総集編. Vol. 7. 株式会社イメージライブラリー. (ImageCD)