

意味の数学モデルの画像データベース検索への適用

3G-6

中村 恭子[†] 北川 高嗣[‡] 清木 康[‡][†]お茶の水女子大学 理学部 [‡]筑波大学 電子・情報工学系

1 はじめに

画像データを対象としたデータベースシステムの実現においては、画像データの定義、表現、および、検索のための方式設計が重要な課題である。従来の画像検索方式は、部分的なパターンマッチングによる直接的な方法と、画像の抽象情報による間接的な方法の2つに大別できる。本稿では、後者の方法を対象とした画像データの検索方式を提案する。この方式は、文脈あるいは状況に応じて動的に変化するデータ間の意味的な関係を計算するモデルとして我々が提案している意味の数学モデル [1, 2] に、パターンマッチングの機能を組み合わせて拡張した新しい画像データベースの検索方式である。

本方式では、まず、パターンマッチングによる検索を用いることにより、大量の画像データから検索対象を絞り込む。そして、意味の数学モデルを適用することにより、絞り込まれた画像データ群の中から、ユーザの印象および画像の内容による意味的連想検索を実現する。

本稿では、この検索方式における意味の数学モデルの適用部分について述べる。この意味的連想検索方式では、正規直交空間をイメージ空間として形成する。その空間上に、ユーザの印象や画像の内容を指定するためのキーワード群を座標点として写像し、また、画像自身を表現するメタデータ群を各々座標点として写像する。そして、キーワード群を用いて、印象あるいは内容を説明する文脈を与え、イメージ空間内において、その文脈に対応する部分空間を選択する。その部分空間内における画像の位置により、文脈に応じた動的な画像検索を行う。本モデルで表現できる文脈の様相の数は、ほぼ無限（現在、約 2^{800} 通り）である。

2 概要

2.1 基本構成

意味の数学モデルは、本来、言葉の意味を扱うためのモデルである。ここでは、印象や画像の内容を表現して

Application of the Mathematical Model of Meaning to Semantic Search for Image Databases

Kyoko Nakamura[†], Takashi Kitagawa[‡], and Yasushi Kiyoki[‡]

[†]Faculty of Science, Ochanomizu University, [‡]Institute of Information Sciences and Electronics, University of Tsukuba

いる文脈に対応した画像を選び出すために、拡張した意味の数学モデルの基本構成を示す。

(1) イメージ空間 I の生成：

m 個の基本データ（“空間生成用メタデータ”）が m 行 n 列の行列（以下、“データ行列”）の形で与えられる。 m 個の基本データはそれぞれ、 n 個の特徴によって特徴づけられている。このデータ行列から正規直交空間（以下、“イメージ空間”） I を生成する。

(2) 画像データの n 次元ベクトル表現：

(1) で用いた n 個の特徴と同一の特徴を用いて、個々の画像、および、ユーザの印象、画像の内容を表すキーワードを n 次元ベクトルで表現する（それぞれ、“画像のメタデータ”、“キーワードのメタデータ”）。

(3) イメージ空間 I 上への写像：

n 次元ベクトルで表現された3種類のメタデータをイメージ空間上へ写像する。これにより、1つの空間上に言葉と画像が同時に配置されることになり、言葉と画像の動的な関係を空間上での距離として計算することが可能になる。

(4) 意味的連想検索 (semantic associative search)：

印象や画像の内容を決定づけるキーワードや文脈を与えられたとする。このとき、文脈として与えられた言葉の列により、イメージ空間から動的に固有部分空間（以下、“意味空間”）を選択する。そして、この部分空間上で、文脈に適した画像を選び出す。

2.2 画像のメタデータ作成方式

ここで、画像から受けるユーザの印象、および、画像に描かれている内容を合わせて、以下、オブジェクトと呼ぶ。

(1) 画像からオブジェクトの抽出

t 個のオブジェクト o_1, o_2, \dots, o_t から成る画像 P を次のように定義する：

$$P = \{o_1, o_2, \dots, o_t\}.$$

ここで、各オブジェクト o_i は、データ行列 A の特徴と同一の特徴を用いて表現されている特徴付ベクトル

$$o_i = (o_{i1}, o_{i2}, \dots, o_{in})$$

である。

(2) 画像 P のベクトル表現

オブジェクト $\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_t$ の和演算子 \oplus を次のように定義する:

$$\bigoplus_{i=1}^t \mathbf{o}_i \equiv (\text{sign}(o_{\ell_1 1}) \max_{1 \leq i \leq t} |o_{i1}|, \text{sign}(o_{\ell_2 2}) \max_{1 \leq i \leq t} |o_{i2}|, \dots, \text{sign}(o_{\ell_n n}) \max_{1 \leq i \leq t} |o_{in}|)$$

ただし、 $\text{sign}(a)$ は、“ a ”の符号（正または負）を表し、 $\ell_k (k = 1, \dots, n)$ は、 $|o_{ik}|$ の最大値を与える添字 $i (i = 1, \dots, t)$ を表している。

3 検索方式

3.1 方法1: キーワードとの距離

選択された意味空間上における解釈の対象となるキーワード（以下、“検索キーワード”）と画像データとの距離を求めることにより、与えられた文脈に対応した画像の検索を行う。その文脈において、検索キーワードとの距離が近い画像を抽出する。

キーワードと画像データ間の文脈に依存した動的距離 $\rho(\mathbf{x}, \mathbf{y}; s_\ell)$, $\mathbf{x}, \mathbf{y} \in \mathcal{I}$ を、以下のように定める:

$$\rho(\mathbf{x}, \mathbf{y}; s_\ell) = \sqrt{\sum_{j \in \Lambda_{s_\ell}} \{c_j(s_\ell)(x_j - y_j)\}^2}$$

ここで、 $c_j(s_\ell)$ は、文脈を決定する言葉の列 s_ℓ に依存して決まる重みを表す [1, 2, 3]。

3.2 方法2: 文脈との相関距離

イメージ空間に写像された画像のメタデータの意味空間内におけるノルムを求めることにより、文脈に応じた画像の検索を行う。意味空間におけるノルムの大きい画像ほど、その文脈に強く関連しているものとする。

画像のメタデータ $\mathbf{x} \in \mathcal{I}$ のノルム $\rho'(\mathbf{x}; s_\ell)$ を、3.1節で定めた動的距離を変形させて、以下のように定義する:

$$\rho'(\mathbf{x}; s_\ell) = \sqrt{\sum_{j \in \Lambda_{s_\ell}} \{c_j(s_\ell)x_j\}^2}$$

4 実現システム

提案方式を実際に、Sun SPARC システム上で C 言語により実現した。

データ行列、および、イメージ空間の生成には、文献 [4]（以下、“BD”）と文献 [5]（以下、“LD”）の2冊の辞書を使用した [2, 3]。BD、LDとも、すべての単語は基本語によって説明されている。BDの基本語に合成語を加え、いくつかの品詞を取り除いた単語群を特徴とした。また、LDの基本語に特徴の単語群を加え、特徴を決めるときに取り除いた品詞と同じ品詞を取り除いた単語群（以下、“見出し語”）を、その説明文とともにBDから取り出した。

この特徴と見出し語をもとに、データ行列を以下に示す方法で作成した。データ行列の各要素の値は、見出し語の説明文中で基本語が肯定の意味に用いられていた場合“1”、否定の場合“-1”、使用されていない場合、“0”とした。さらに、見出し語自身が特徴に含まれるときは、その特徴の要素を“1”とした。

画像のメタデータは、イメージ空間へ写像する画像を表現している。この実現では、2.2節で示した方法によってメタデータの作成を行った。

5 おわりに

本稿では、パターンマッチングと意味的連想検索の両者を組み合わせた画像データの検索方式を提案し、意味的連想検索に意味の数学モデルを適用した場合の実現方式を示した。この方式により、ユーザの印象や画像の内容による画像検索が可能となった。

今後は、画像のメタデータに関する他の生成方法の開発、パターンマッチングとの融合による画像データベース検索システムの構築、そして、動画や音声といった画像以外のマルチメディアデータの検索への適用を行っていく予定である。

参考文献

- [1] T. Kitagawa and Y. Kiyoki, “A Mathematical Model of Meaning and its Application to Multidatabase Systems,” Proceedings of 3rd IEEE International Workshop on Research Issues on Data Engineering: Interoperability in Multidatabase Systems, pp.130-135, April 1993.
- [2] Y. Kiyoki, T. Kitagawa and Y. Hitomi, “A Fundamental Framework for Realizing Semantic Interoperability in a Multidatabase Environment,” International Journal of Integrated Computer-Aided Engineering (John Wiley & Sons) (to be published), 1995.
- [3] Y. Kiyoki, T. Kitagawa and T. Hayama, “A Metadata System for Semantic Image Search by a Mathematical Model of Meaning,” ACM SIGMOD Record, Vol. 23, No. 4, pp.34-41, Dec. 1994.
- [4] Ogden, C.K., “The General Basic English Dictionary,” Evans Brothers Limited, 1940.
- [5] “Longman Dictionary of Contemporary English,” Longman, 1987.