

テキスト情報検索手法を利用したマルチメディアデータ検索手法

鈴木 優[†] 大野 和久[†] 光川 正弘[†] 川越 恭二[†]

[†] 立命館大学大学院理工学研究科 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: †{suzuki,oono,mitsukawa,kawagoe}@coms.ics.ritsumeai.ac.jp

あらまし 本稿では、テキスト情報検索手法を利用したマルチメディアデータの検索手法を提案する。現在、単語の出現頻度や文書頻度を利用したテキスト情報検索手法が利用されている。この手法の目的は、テキスト文書に含まれる特定のフレーズを探し出すことではなく、テキスト文書に含まれる文書の内容が問合せキーワードに関連があるかどうかを確かめることである。この考え方を画像検索手法、音楽検索手法に応用することによって、特定のオブジェクト、特定のメロディを含む検索ではなく、問合せ画像、音楽に関連するマルチメディアデータを検索することが可能となると考えられる。このとき、テキスト情報検索手法において必要である単語という概念がマルチメディアデータには存在しないため、テキスト情報検索手法をそのままマルチメディアデータに応用することは困難である。そこで本論文では、マルチメディアデータから単語に相当する単位を抽出することを目的としている。評価実験を行うことによって、マルチメディアデータを高い精度で検索することが可能であることを示す。

キーワード テキスト情報検索, マルチメディア検索, ベクトル検索モデル

A Multimedia Document Retrieval Method using Textual Information Retrieval Techniques

Yu SUZUKI[†], Kazuhisa ONO[†], Masahiro MITSUKAWA[†], and Kyoji KAWAGOE[†]

[†] Graduate School of Information Science, Ritsumeikan University

1-1-1 Noji-Higashi, Kusatsu, Shiga 5258577, Japan

E-mail: †{suzuki,oono,mitsukawa,kawagoe}@coms.ics.ritsumeai.ac.jp

Abstract In this paper, we propose a multimedia data retrieval method using textual information retrieval techniques. Recently, various information retrieval techniques which are based on term frequency and document frequency are used by many information retrieval systems. Especially, many IR systems deal with vector space model for retrieving intuitively relevant documents. Therefore, when we deal vector space model with image retrieval systems, music retrieval systems, and other multimedia data retrieval systems, these retrieval systems can retrieve intuitively relevant documents. However, in multimedia data, there is no term. Therefore, in this paper, we extract information units as terms, and we calculate similarity values between two multimedia data using the information units. We evaluate our proposed method for image and music data. In our experimental evaluation, we confirm that we can extract intuitively relevant image and music data using our proposed method.

Key words Textual information retrieval, Multimedia information retrieval, Vector space model

1. はじめに

本研究では、マルチメディアデータ全体の傾向に注目した類似度を算出するために、テキスト情報検索手法を利用する手法の提案を行う。

現在、World Wide Web 上にはテキストデータの他に画像データや音楽データなどのマルチメディアデータが、様々な用途に利用され流通している。ところが、これらのデータは主にメタ

データを利用し検索されることが多い。これは、マルチメディアデータの検索目的が利用者にとって直感的に類似した検索対象を検索することであるためである。

一方、テキスト文書検索において一般的に利用されている検索モデルの一つとしてベクトル検索モデルが挙げられる。このモデルは、検索対象に含まれる単語の出現頻度をベクトルとして表現し、ベクトル間類似度を算出することにより、利用者の検索意図に適合した検索対象を検索する手法である。ベクトル

検索モデルの特徴は、利用者にとって直感的に類似した文書を検索することである。ブーリアン検索モデルのように、問合せと検索対象の部分一致が目的ではないことが重要である。つまり、検索対象全体における単語の出現傾向が類似しているかどうかを特徴として抽出することが、部分一致検索よりも利用者の直感的な類似に合致していると考えた。

そこで、著者らはベクトル検索モデルをマルチメディアデータを検索するために利用することを考える。既存のマルチメディアデータ検索システムでは多くの場合、利用者の指定したパターンに一致する部分を検索する手法が用いられていた。本研究では、それらとは異なり、検索対象全体の傾向を考慮して類似していると考えられる検索対象を抽出することを目的としている。例えば画像の場合であれば、風景画のように画像全体で一つの景色を表現しているが、特定の対象物が写っていないような問合せ画像に類似した画像を検索する場合である。

ベクトル検索モデルにおいては、テキスト検索エンジンにおける単語に相当する部分をマルチメディアデータから抽出することが必要である。本研究では、テキストにおける単語に相当する単位を **情報単位** (Information Unit) と呼ぶ。マルチメディアデータには辞書に相当するものが存在しないため、検索対象集合から情報単位を抽出する必要がある。

そこで本研究では、情報単位とは検索対象群に共通して含まれる部分であると仮定した。これは、未知の言語で書かれたテキスト文書をベクトル検索モデルを用いて検索する際に利用される手法と同一である。もし未知の言語で記述されたテキスト文書相互の類似度を算出する際には、検索対象に共通して含まれる文字列は単語であると推測することができ、それらの文字列を利用して類似度を算出することができる。そこで、この考え方をマルチメディアデータへ応用し、検索対象群に共通して含まれる部分を情報単位とすることとした。例えば、風景画像の場合であれば雲を表現している白色の部分などが、情報単位となると考えられる。また音楽データであれば、多くの音楽で利用されているフレーズが情報単位となると考えられる。

2. 関連研究

現在、画像情報検索や音楽情報検索において文書検索手法を適用した手法が提案されている。Doraisamy ら [1] は、楽曲に含まれる音符の列を文字列へ変換し、 n -gram を適用することによって検索を行う手法を提案している。音符の列を文字列に置き換えて検索を行う点は、本提案と類似する手法である。ところが、索引付けの方法として n -gram を利用しているため、部分的に一致している検索対象を検索する場合に有効である点が本提案との差異である。本提案では、部分的に一致していることを考えるのではなく、検索対象全体の傾向が類似した検索対象を検索することを目的としている。

一方、音声学に基づいて音楽構造と文書構造の対応関係に着目した研究が Reed [2] らによって行われている。この研究では、音楽を可変の区間へ区切り、その区間を一つの単語としている。可変の区間を楽曲の特徴とする点において、本研究における情報単位抽出手法と類似している部分がある。ところが、楽曲に含まれる全ての区間を情報単位としているため、その情報単位の連続によって音楽データの構成を表現している。この点において、複数のマルチメディアデータ間における類似部分からマルチメディアに含まれる情報単位を抽出しようとしている点、つまり情報単位の定義が本研究と異なる点である。

3. 文書検索手法を利用したマルチメディアデータ検索手法

本研究で提案しているシステムでは、次のような手順で検索が行われる。

(1) 利用者はシステムへ問合せを入力する。

- 問合せは Query by example で行われる。つまり、必要なマルチメディアデータに類似したデータをシステムへ入力する。

(2) 問合せから情報単位を抽出する。

(3) 問合せから得られた情報単位を基に、問合せ特徴ベクトルを抽出する。

(4) 検索対象特徴ベクトル群と問合せ特徴ベクトルの比較を行い、類似度を算出する。

(5) 最も類似度の高い検索対象から順に、利用者へ検索結果を提示する。

ここで、あらかじめ検索対象から情報単位を抽出しなければならない。そこで 3.1 節では情報単位の抽出方法について述べる。また、手順 (3) において問合せ特徴ベクトル、手順 (4) においては検索対象特徴ベクトルを作成する必要がある。そこで 3.2 節では 3.1 節で抽出した情報単位を基にマルチメディアデータから特徴ベクトルを抽出する手法について述べる。さらに 3.3 節では、特徴ベクトル間の類似度をコサイン相関値によって算出する手法について述べる。

3.1 情報単位の抽出

本節では、検索対象群から情報単位を抽出する手法について述べる。

まず、 N 個の検索対象 D_k ($k = 1, 2, \dots, N$) が与えられているとする。このとき、全ての検索対象は同一のメディアであるとする。また、各検索対象 D_k から、メディアに応じた特徴量を抽出し、配列として表現する。 D_k の特徴量 $R(D_k)$ は次の (1) 式のように表現される。

$$R(D_k) = \begin{pmatrix} f_{D_k}(1,1) & f_{D_k}(1,2) & \cdots & f_{D_k}(1,J(D_k)) \\ f_{D_k}(2,1) & f_{D_k}(2,2) & \cdots & f_{D_k}(2,J(D_k)) \\ \vdots & \vdots & \ddots & \vdots \\ f_{D_k}(I(D_k),1) & f_{D_k}(I(D_k),2) & \cdots & f_{D_k}(I(D_k),J(D_k)) \end{pmatrix} \quad (1)$$

ここで、 $f_{D_k}(i, j)$ ($i = 1, 2, \dots, I(D_k), j = 1, 2, \dots, J(D_k)$) はマルチメディアデータの部分における特徴量である。つまり、一つの行が一つの検索対象の特徴ベクトルを表現している。また、 $R(D_k)$ に含まれる要素はそれぞれ独立である必要がある。

これらを各検索対象に対して行い、 $F(D_k)$ の算出を行う。算出方法はメディアや特徴量抽出手法によって異なるが、例えば画像における輝度を特徴量とした場合には、 i, j はそれぞれ画像における x, y 座標に相当し、 $f_{D_k}(i, j)$ は輝度情報に相当する。 $f_{D_k}(i, j)$ はスカラ値の場合とベクトル値である場合がある。また、音楽データは一般的に時系列データであると考えられるため、 $R(D_k)$ は $J(D_k) \times 1$ の大きさの配列になる。

ここで、二つの検索対象 D_α, D_β ($\alpha, \beta = 1, 2, \dots, N, \alpha \neq \beta$) から情報単位を抽出する手法を述べる。 $R(D_\alpha)$ と $R(D_\beta)$ に含まれる行列のうち、類似している部分行列を取り出す。このとき、複数の重なり合う部分行列が類似している場合には、最も大きな部分行列を情報単位として取り出す。ここで、取り出された情報単位を集合 $T = \{t_1, t_2, \dots, t_L\}$ へ格納する。

以上の手順を全ての検索対象組合せによって行い、 T を算出する。このようにして得られた部分行列群を、情報単位として定義する。

3.2 特徴ベクトルの作成

本節では、3.1 節において得られた情報単位を基に、問合せと検索対象に対して特徴ベクトルを作成する手法について述べる。問合せを Q 、検索対象を D_k とおくと、簡単のために $D_0 = Q$ と定義し、 $k = 0, 1, \dots, N$ における特徴ベクトルを定義する。つまり、問合せと検索対象について同様の方法で特徴ベクトルの定義を行う。

まず、検索対象 D_k の特徴ベクトル $\vec{V}(D_k)$ を次のように定義する。

$$\vec{V}(D_k) = [w(t_1, D_k), w(t_2, D_k), \dots, w(t_L, D_k)] \quad (2)$$

ここで、 $w(t_i, D_k)$ は、検索対象 D_k における情報単位 t_i に関する重みである。 $w(t_i, D_k)$ の算出方法として、本研究ではベクトル検索モデルにおいて一般的に利用されている TFIDF [3] を利用する。つまり、情報単位 t_i が検索対象 D_k に出現する回数を $tf(t_i, D_k)$ とし、情報単位 t_i が含まれる検索対象の数を $df(t_i)$ とすると、 $w(t_i, D_k)$ は次のように算出される。

$$w(t_i, D_k) = tf(t_i, D_k) \cdot \left(-\frac{df(t_i)}{N} \right) \quad (3)$$

以上の方法によって、問合せおよび検索対象に対して特徴ベクトルの算出を行う。

3.3 類似度の算出

本節では、3.2 節において抽出された、問合せ特徴ベクトルと検索対象特徴ベクトルとの類似度を算出する。

類似度算出の方法は、テキスト検索におけるベクトル検索モデルにおいて一般的に利用されているコサイン相関値 [3] を用いる。つまり、問合せ特徴ベクトル $\vec{V}(Q)$ と検索対象特徴ベクトル $\vec{V}(D_k)$ の類似度 $S(D_k)$ は次のように求めることができる。

$$S(D_k) = \frac{\vec{V}(Q) \cdot \vec{V}(D_k)}{|\vec{V}(Q)| \cdot |\vec{V}(D_k)|} \quad (4)$$

以上のような手法を用いることによって、ベクトル検索モデルに基づくマルチメディアデータの類似度を定義することができた。

4. おわりに

本稿では、マルチメディアデータにおける検索を行う際に、テキスト情報検索手法の一つであるベクトル検索モデルを利用する手法の提案を行った。提案手法を利用することにより、検索対象全体の傾向に基づいて問合せとの類似度を算出することができるため、検索対象の一部と問合せの一部が類似している場合ではなく、全体の傾向が類似している検索対象、つまり直感的に類似していると考えられる検索対象を検索することが可能となった。

本研究の今後の課題を述べる。まず、特徴部分の出現位置に関する検討についてである。本研究における特徴ベクトル算出においては、情報単位が出現した位置を反映しておらず、その結果直感的に異なると考えられる検索対象が検索されてしまう場合がある。一方、テキスト文書検索においては情報単位の出現密度が考慮されている場合があるため、これらの手法をマルチメディアデータ検索手法として利用する手法について検討している。

さらに、実際の評価実験の結果を基に、情報単位の数を削減することによる計算量の削減について検討している。提案している手法をそのまま画像や音楽に適用した場合、利用者にとって目視できない程度の小さな領域や聴き取りが困難な程度の短い音楽が情報単位として出現する可能性がある。このような場合には、Zipf の法則を利用することによって過度に小さい、または過度に大きい情報単位を削減することが考えられる。ところが、この手法を用いるためには情報単位とするかどうかを決定するためのパラメータを適切に設定する必要があり、この手法についての検討を行っている。

謝辞 本研究の一部は、文部科学省科学研究費補助金（若手研究 (B), 40388111）によります。ここに記して感謝を表します。

文 献

- [1] S. Doraisamy and S. Ruder: “A Polyphonic Music Retrieval System using n-grams”, Proc. of International Conference on Music Information Retrieval, pp. 204 – 209 (2004).
- [2] J. Reed and C.-H. Lee: “A Study on Music Genre Classification Method Based on Universal Acoustic Models”, Proc. of International Conference on Music Information Retrieval, pp. 104 – 109 (2006).
- [3] C. D. Manning, P. Raghavan and H. Schütze: “Introduction to Information Retrieval”, Cambridge University Press (2008).