

XML データベース上における動画データベースへの文検索システム

白井 宏一[†] 椎名 広光[†]

[†] 岡山理科大学総合情報学部 〒700-0005 岡山県岡山市理大町 1-1

E-mail: †f-shirai@nifty.com, ††shiina@mis.ous.ac.jp

あらまし 本論文では動画ファイルと動画ファイルの説明文を XML サーバ上に構築し、検索文と類似する動画ファイルの提示を格フレーム構造の類似性を用いることで検索を行うシステムを作成した。

キーワード XML サーバ, 動画検索, 文書検索

A sentence search system to a movie database on an XML database

Khoichi SHIRAI[†] and Hiromitsu SHIINA[†]

[†] Faculty of Informatics, Okayama University of Science

Ridai-cho 1-1, Okayama-shi, Okayama, 700-0005 Japan

E-mail: †f-shirai@nifty.com, ††shiina@mis.ous.ac.jp

Abstract In this paper, we construct a movie file and its explanation on an XML database, and propose similarity between the explanation of a movie file and a search sentence by using the case frame.

Key words XML server, movie search, document search

1. ま え が き

近年、デジタルカメラやカメラ付き携帯電話の普及してきている。また、コンピュータのハードディスクドライブの大容量化やネットワーク回線の高速化によって、個人でも動画をデータとして持ち、それらを検索して再生することも多くなってきている。実際、大容量の動画データを送信し配信するサービスも出はじめている。そこで自分で作成した動画を Web サーバに保存したり、その中から必要な動画を検索し再生するためのサービスが今後もっと必要になってくると考えられる。

また、Web 上における情報検索 [4] では、単語頻度をベースに実現される例が多く、画像検索においても画像が保存されている HTML に対する単語を手がかりに検索結果を提示している。文の構造を用いるには文章の大きさが不均衡であったりし、Web 検索に用い難いと考えられる。

本研究では、XML データベース上に動画ファイルとそれを説明をする文を一組にして登録し、それを検索するシステム構築することを想定している。その検索には文（以下、検索文）を採用し、説明文と検索文の間の類似を格フレーム [3] をもとにして計算手法し、動画ファイルの検索に用いている。また、格フレーム間の類似度の計算方法には、動画ファイルと説明文を一組としており、その説明文の述部が動画ファイルをよく説明していると想定した方法を用い、格フレームの階層ごとの比較を行っている。

2. システムの構成

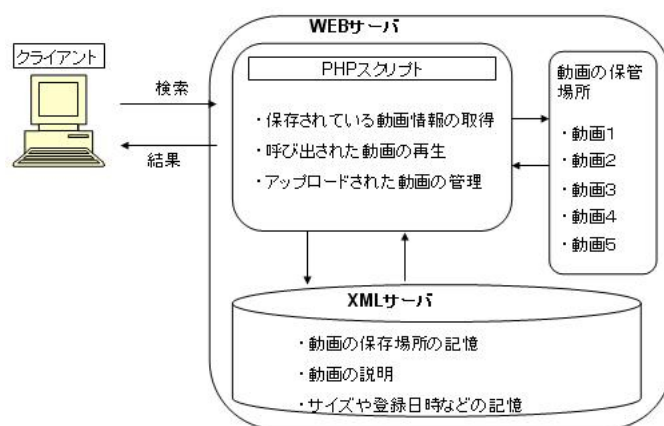


図1 動画サーバの構成図

本研究では、システムを Web サーバから CGI スクリプトを通じて XML サーバ (Xpiori [1]) に問い合わせを行う形式で構成し (図1)、結果は XML サーバから取得できた情報を HTML に CGI スクリプトで編集した上で提示している。

XML サーバでは、動画ファイルの説明文、説明文を解析したデータ、動画の登録場所を管理しており、実際の動画ファイル自体は XML サーバに格納しない方式をとっている。また、検索の提示は CGI スクリプトが XML サーバに問い合わせた

結果からファイルの場所を取り出し、ユーザに動画再生を提供する方式である。XML サーバで管理している情報については、第 3 章で述べる。

システムの機能としては、下記に示すと 3 つの機能を作成している。

- (1) 動画ファイルの登録・削除
- (2) 動画ファイルに対するコメントの修正
- (3) 動画ファイルの日本語による文検索

動画一覧

[更新](#) | [動画の追加](#)

動画検索

検索ワードを入れてください:

再生リストへ	動画名	説明文
<input type="button" value="追加"/>	飛び跳ねる	ジェームズが飛び跳ねる。
<input type="button" value="追加"/>	走る	ジェームズが走る。
<input type="button" value="追加"/>	蹴飛ばす	ジェームズが蹴飛ばす。
<input type="button" value="追加"/>	踊り	太郎が腹筋をする。
<input type="button" value="追加"/>	腹筋をする。	ジェームズがうまい踊りを踊る。
<input type="button" value="追加"/>	ジェームズのバクテン	ジェームズがバクテンをする。
<input type="button" value="追加"/>	人体浮遊	胡坐のままジェームズが浮遊する。
<input type="button" value="追加"/>	ジェームズありがとう	ジェームズが手話で「ありがとう」をする。
<input type="button" value="追加"/>	花子の手話	花子が手話で「これはうまいりんごです」を話します。
<input type="button" value="追加"/>	ジェームズの手話「言う」	ジェームズが手話で「言う」を話します。

図 2 動画一覧の画面

なお、検索に関する仕組みについては次のとおりである。

動画ファイルを登録するときに、その動画ファイルを説明を格納しておくことで、XML サーバから検索を行う際には、説明を手がかりに動画ファイルを提示する。説明文を手がかりにするにあたって説明文をあらかじめ形態素解析、係り受け解析をおこない、そのうえで説明文を格フレームで登録する。リンクを除いた格フレーム形式は、XML データ形式とは等価に扱うことができると考えられるので、説明文の格フレーム形式と説明文の XML データを同じものとして取り扱う。

検索では、文を用いて検索を行い、検索文と類似する動画を提示する。検索文と動画ファイルとの類似については、検索文も形態素解析、係り受け解析を行い格フレーム形式に変換後、検索文の格フレームと動画ファイルの説明文の格フレームを間の類似性と計算して、検索文からみて最も高い類似した動画を提示する。

2.1 システムの内部処理

システムの機能に対して、内部処理を CGI スクリプトで切り分けて XML サーバへの問い合わせを行っている。動画に対する XML 情報の登録・削除・コメント修正は、Xquery から XML サーバの XML ノードを登録・削除・タグ内の文の修正で実行される。

それに対して検索は、Xquery の検索結果から説明や説明から作成した格フレームを取り出して検索文との格フレームの類似性を求め、類似性の高い動画ファイルを提示する。

3. XML データベース上における動画の説明文の格フレーム

システムでは、XML データベース上における説明文を格フレームデータ形式に変換して保持している。説明文に使われる日本語の文は、係り受け関係や形態素に分割できる。例えば、下記の文 1 に対する係り受け関係を図 3 のように表すことができる。本研究では、係り受け関係や形態素への分割に Cabocha [2] を利用して、その結果の一部を加工して取り出すことで格フレームデータを作成している。以下では、Cabocha の XML 出力と格フレーム変換を説明する。

文 1: ジェームスがうまい手話で話します。

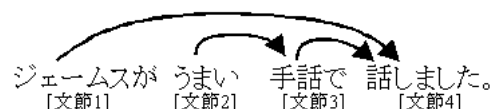


図 3 文 1 に対する係り受け

(1) Cabocha による XML 出力

文 1 の場合、Cabocha によって作成される情報のうち本研究で必要となるものを図 4 に示す。

```
<sentence>
<chunk id="0" link="3">
<tok id="0" base="ジェームズ" pos="名詞-固有名詞-人名-名">ジェームズ</tok>
<tok id="1" base="が" pos="助詞-格助詞-一般">が</tok>
</chunk>
<chunk id="1" link="2">
<tok id="2" base="うまい" pos="形容詞-自立">うまい</tok>
</chunk>
<chunk id="2" link="3">
<tok id="3" base="手話" pos="名詞-一般">手話</tok>
<tok id="4" base="で" pos="助詞-格助詞-一般">で</tok>
</chunk>
<chunk id="3" link="1">
<tok id="5" base="話す" pos="動詞-自立">話し</tok>
<tok id="6" base="ます" pos="助動詞" ctype="特殊-マス">ます</tok>
<tok id="7" base="。" pos="記号-句点">。</tok>
</chunk>
</sentence>
```

図 4 文 1 に対する Cabocha による XML 形式の出力

(2) 格フレームへの変換

格フレームへの変換は、係り受け関係のうち係り受けを受ける側の節を述部と、それ以外を主部とする。主部・述部の中でそれぞれの節に係る節もある。また、主部に直接係る節を第一階層として格助詞や係助詞の節が取り出すことができる。

文 1 の例では、「話します」が述部にあたり、それ以外は主部に当たる。また 4 文節のうちの「うまい」は格助詞-での文節「手話で」に係る。第一階層の文節は、「話します」、「ジェームズが」、「手話で」となる。

各文節に対しては、単語の基本形となる基本単語と品詞を取

り出し、係る側の文節を受ける側の文節の中の格フレームとして作成することで、Cabocha による XML 出力から格フレームを得る。

文 1 に対する格フレームを図 5 に示す。なお「ジェームスが」と「手話で」が「話します」に係っている図 5 のように、同じ文節に係っている文節の順序は、類似度の計算に影響がないため意味を持たせていない。

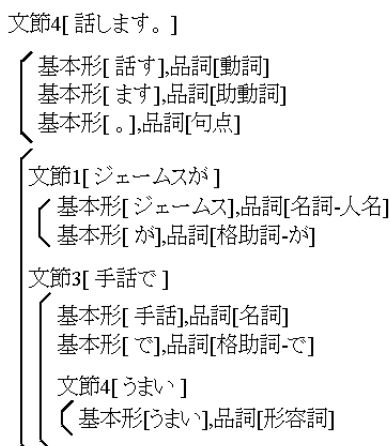


図 5 文 1 に対する格フレームデータ

また、図 5 に対する本研究で作成したシステムの XML 形式を図 6 に示す。

```

<sentence>
<chunk>話します。
<tokid="5" base="話す" pos="動詞-自立">話し</tok>
<tokid="6" base="ます" pos="助動詞" ctype="特殊・マス">ます</tok>
<tokid="7" base="。" pos="記号-句点">。</tok>
</chunk>
<chunk pos="助詞 格助詞-が">ジェームスが
<tokid="0" base="ジェームス" pos="名詞-固有名詞-人名-名">ジェームス</tok>
<tokid="1" base="が" pos="助詞 格助詞-一般">が</tok>
</chunk>
<chunk pos="助詞 格助詞-で">手話で
<tokid="3" base="手話" pos="名詞-一般">手話</tok>
<tokid="4" base="で" pos="助詞 格助詞-一般">で</tok>
<chunkid="1" pos="形容詞-自立">うまい
<tokid="2" base="うまい" pos="形容詞-自立">うまい</tok>
</chunk>
</chunk>
</sentence>
  
```

図 6 文 1 に対する XML 形式による格フレームデータ

4. 検索文と動画ファイルの説明文との類似性

検索文と動画ファイルの説明文との類似性を計算は、格フレームに変化した手法の有効性を確かめるために、簡単なテストセットとして次の表 1 を用いて比較する。なお、類似性の計算には、以下に述べる (1) 検索文の単語の動詞・名詞・形容詞の所属数方式、(2) 格フレーム間の類似度方式を 2 種類を用いた。

4.1 検索文と動画の説明文の単語比較における類似性

まず、検索文を形態素に分割後、動詞、名詞、形容詞、副詞の基本単語が動画の説明文に含まれている個数の最大値を利用する方法である。言い換えると助詞や助動詞などの意味を持た

表 1 XML データベース上の動画情報

番号	説明文
M1	ジェームスが飛び跳ねる。
M2	ジェームズが走る。
M3	ジェームズが蹴飛ばす。
M4	太郎が腹筋をする。
M5	ジェームズがうまい踊りを踊る。
M6	ジェームズがバクテンをする。
M7	胡坐のままジェームズが浮遊する。
M8	ジェームズが手話で「ありがとう」をする。
M9	ジェームズが手話で「ありがとう」を言います。
M10	花子が手話で「これはうまいりんごです」を話します。
M11	ジェームズが手話で「言う」を話します。

ない単語をあらかじめ削除しておく方式といえる。

形態素に分割後 (Cabocha による XML 形式の出力の一部、図 4)、動詞、名詞、形容詞、副詞を取り出すと「ジェームス」、「手話」、「うまい」、「話し」と表 1 の動画データベースの各項目の類似度は表 2 のように示される。なお、検索文の単語と一致している単語は太字で示している。

表 2 単語比較による類似度

番号	説明文	類似度
M1	ジェームズが飛び跳ねる。	1
M2	ジェームズが走る。	1
M3	ジェームズが蹴飛ばす。	1
M4	太郎が腹筋をする。	0
M5	ジェームズがうまく踊りを踊る。	2
M6	ジェームズがバクテンをする。	1
M7	胡坐のままジェームズが浮遊する。	1
M8	ジェームズが手話で「ありがとう」をする。	2
M9	ジェームズが手話で「ありがとう」を言います。	2
M10	花子が手話で「これはうまいりんごです」を話します。	3
M11	ジェームズが手話で「言う」を話します。	3

また、表 2 では、単語の一致個数を用いているので番号 M10、M11 の二つが最大値となり類似度としては同じとなる。また、最大値でない場合でも手話を用いることや話しをする事に関係のない番号 M5 が一致個数 2 となり目的の動画を探すときには不要なものを取り出してしまう可能性がある。

4.2 検索文と動画の説明文間の格フレームによる類似度

本研究では、動画は動きが主に説明されるため、述部に対して重きを置くことことで類似性を計算しやすくなると想定している。この想定を実現するために動画ファイルへの説明文を主部と述部に分けて考え、述部に対するパラメータを主導的に計算するモデルを提案している。また、説明文はある文節が複数の他の文節から係っているため、係っている文節は均等な説明を係る側に与えていると類似度を定式化している。

本節では、第3章の格フレームで述べたように検索文とXMLデータベースの上における説明文の両方ともに格フレーム形式に変換後、格フレーム間の類似度を計算方法について説明する。ここで類似度は、検索文 S 、動画ファイルの説明文 M に対して定義し、 $Sim(S,M)$ で表し、検索文 S の述部、主部の文節のそれぞれからみて、動画ファイルの説明文 M に一致するものがあるかどうか調べ、係数を掛けて後の総和で表すものとする。ただし、検索文の主部内の文節に再び別の文節に係る場合(例えば、文1の「うまい」が「手話で」に係り、図7では文節C31が文節C3に係る)、動画ファイルの説明文にも同じ係り受けをする文節があるかどうか調べ、主部と述部に分ける第一層目の係数に反映する。

なお、以下の説明では検索文 S 、動画ファイルの説明文 M の格フレームは図7に示すように格フレームの項目のデータとする。図7では、格フレームの同じ項目表すために項目の前に文の種類とドット(.)をつけ、文節を C 、基本単語を B 、品詞を P で表している。例えば、検索文 S の文節 C_i を $S.C_i$ 、文節 C_i 中の基本単語 $B_{i,j}$ を $S.B_{i,j}$ 、品詞 $P_{i,j}$ を $S.P_{i,j}$ で表す。また、文節 C_i にかかる文節を C_{ik} で表すとすると、 C_{ik} 中の基本単語 $B_{ik,j}$ を $S.B_{ik,j}$ 、品詞 $P_{ik,j}$ を $S.P_{ik,j}$ で表す。

格フレーム間の類似度 $Sim(S,M)$ の計算

Step 1: S の述部 $S.B_0$ の基本単語が m_1 個あり、基本単語 $S.B_{0,1}, \dots, S.B_{0,m_1}$ とする。その基本単語中に M の述部 $M.B_0$ の基本単語に対して一致する単語数を m_1^* とする。この時、 b_0 を次の式で計算する。

$$b_0 = \frac{m_1^*}{m_1}$$

Step 2: $i = 1$ を初期値として i が検索文の主部の文節個数となるまで繰り返す(検索文の主部の文節個数分繰り返す)。繰り返しが終わったら Step3 へ。

Step 2.1: S の文節 $S.C_i$ の品詞 $S.P_{i,j}$ と一致する M の文節 $M.C_k$ の品詞 $M.P_{k,l}$ を探す。一致するものが見つければ Step 2.2 に進む。見つからなければ $X_i = 0$ 、 $i = i + 1$ として Step 2 に戻る。

Step 2.2: S の文節 $S.C_i$ 中の品詞が助詞である基本単語 $S.B_{i,j}$ と M の文節の基本単語 $M.B_{k,l}$ が一致するなら $X_i = 1$ とする。

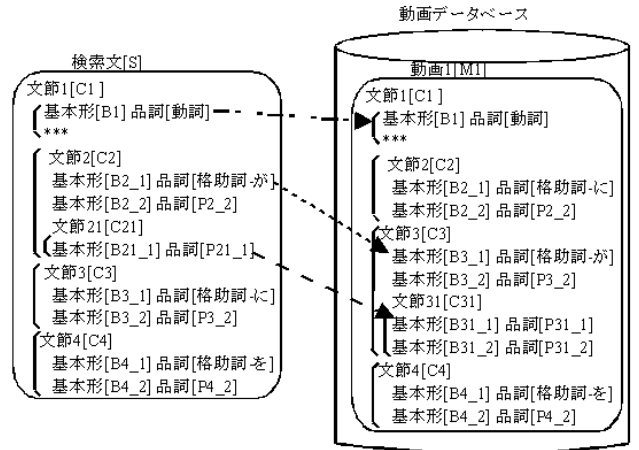


図7 格フレーム間の類似度

Step 2.3: 第一層目の文節 $S.C_i$ に対して m_i と m_i^* と計算する。

$$m_i = \text{文節 } S.C_i \text{ 中の基本単語数} + \text{文節 } S.C_i \text{ に係る文節数}$$

$$m_i^* = \text{文節 } M.C_j \text{ 中の基本単語と一致する文節 } S.C_i \text{ 中の基本単語数} + \text{文節 } M.C_j \text{ 中の文節と一致する文節 } S.C_i \text{ に係る文節数}$$

次に単語と係り受けの節の一致割合を b_i として次の式で計算し、 $i = i + 1$ として Step 2 に戻る。

$$b_i = \frac{m_i^*}{m_i}$$

Step 3: Step 2 までで計算した変数 b_i に重み a_1 をつけた和で類似度を定義する。

$$Sim(S, M) = a_0 \cdot b_0 + \sum_{i=1}^{S \text{ の主部の文節数}} a_1 \cdot b_i \cdot X_i$$

$$\text{ただし、} a_1 = \frac{1 - a_0}{S \text{ の主部の文節数}} \text{ とする。}$$

上記の類似度 $Sim(S,M)$ の Step 3 にあるように述部の動詞が一致した場合に、係数 a_0 、 $0 < a_0 < 1$ 、を掛けて、それ以外の主部は a_0 を除く係数を均等に割り当てている。よって係数 a_0 の値によって類似度の値は大きく変化する。

文1を検索文として、表1の動画情報との類似度の計算例を示す。

検索文の解析: 文1の第一層目の文節数は3(図6)で、主部の文節は2となる。品詞[格助詞-で]の文節は基本単語「手話」と文節「うまい」が係っている。よって文節の中で文節が係っている場合で、基本単語「手話」と係り側の文節「うまい」の一致を計算する必要がある。

(1)M1 との類似度: 検索文 S からみて、品詞[格助詞-が]のみ一致する場合である。

$$Sim(\text{文 } 1, M1) = \frac{1 - a_0}{2}$$

(2)M10 との類似度: 検索文 S からみて, 述部 S.C1 が一致し, 主部の品詞 [格助詞-で] の文節の基本単語「手話」しか一致しない場合である.

$$Sim(\text{文 } 1, M10) = a_0 + \frac{1 - a_0}{2} \cdot \frac{1}{2}$$

(3)M11 との類似度: 検索文 S からみて, 述部 S.C1, 品詞 [格助詞-が] の文節 S.C2 が一致する. また, 品詞 [格助詞-で] の文節 S.C2 の基本単語「手話」が一致し, 品詞 [格助詞-で] に係る基本単語「うまい」が一致しない場合である.

$$Sim(\text{文 } 1, M11) = a_0 + \frac{1 - a_0}{2} + \frac{1 - a_0}{2} \cdot \frac{1}{2}$$

$a_0 = 0.5$ とした場合の類似度を表 3 に示す.

表 3 格フレームによる類似度

番号	説明文	類似度 $a_0 = 0.5$
M1	ジェームズが飛び跳ねる。	0.25
M2	ジェームズが走る。	0.25
M3	ジェームズが蹴飛ばす。	0.25
M4	太郎が腹筋をする。	0.00
M5	ジェームズがうまく踊りを踊る。	0.25
M6	ジェームズがバクテンをする。	0.25
M7	胡坐のままジェームズが浮遊する。	0.25
M8	ジェームズが手話で「ありがとう」をする。	0.375
M9	ジェームズが手話で「ありがとう」を言います。	0.375
M10	花子が手話で「これはうまいりんごです」を話します。	0.625
M11	ジェームズが手話で「言う」を話します。	0.875

表 2 では番号 M10 と M11 は等しい類似度となるのに対して, 表 3 では番号 M11 が類似度最大となる. 単語の一致個数に対して, 動詞と形容詞の係り方に影響を類似度に反映することができると考えられる.

5. ま と め

本研究では, XML サーバ上に動画ファイルを説明する XML データベースを作成し, それと検索文との類似性を考慮して動画ファイルを提示をするシステムを作成した. 類似性に関して文間の文字や単語の一致度による類似よりも, 文を形態素解析・係り受け解析をした格フレームによる構造を考慮した手法が, 有効であることを示した. しかしながら, 構造の考慮に関しては係り受け解析から得られる同列の文節間の単語間の一致度を計算しているため, 類義語を考慮していない. また, 類似度 Sim の述部に対する係数 a_0 の値は決め方は, 検索文によって変化させるなどの手法が考えられ, 今後の課題となる. 加えて, 動画データの分量が少ないので, テストセットを充実させる必要がある.

なお, 動画ファイルを手話の単語程度に限定した場合, 説明文を文ではなく単語ベースで作成することができる. この場合, 検索文を単語分割したあとに, それに対応する動画を XML データベースから取り出す方法で作成した. 検索文から作られる形態素ごとの基本単語の単語列と XML データベース内の XML ノードごとに格納されている単語との一致度を計算し, 最大値となる動画ファイルを再生する手話教示システムを作成している.

文 献

- [1] 山田, XML データベース入門, 翔泳社, 2006
- [2] CaboCha, <http://chasen.org/taku/software/cabocha>
- [3] 長尾編, 自然言語処理, 岩波書店, 1996
- [4] 北, 津田, 獅々子, 情報検索アルゴリズム, 2002
- [5] 奥村, 難波, テキスト自動要約, オーム社, 2005
- [6] 川崎, 梶山, 椎名, XML サーバ上でのモーションデータの加工システム, FIT2006 第 5 回情報科学技術フォーラム I-7, pp7-10, 2006