

ニュース映像アーカイブの構造化に基づくキー画像の取得 Key Image Extraction Based on the Structure of a News Video Archive

孟 洋[†] 山岸 史典[‡] 井手 一郎^{*} 片山 紀生[†] 佐藤 真一[†] 坂内 正夫[†]
Hiroshi MO Fuminori YAMAGISHI Ichiro IDE Norio KATAYAMA Shin'ichi SATOH Masao SAKAUCHI

1. はじめに

近年の蓄積装置の大容量化、低価格化は、ビデオアーカイブの構築を現実のものとし、映像情報の幅広い活用を可能にしつつある。なかでも、社会的な記録とみなせるニュース映像の利用価値は高く、過去から現在までのニュース映像へのアクセスを可能とするニュース映像アーカイブは非常に有用である。大量に蓄積されたニュース映像を有効に活用するためには、意味構造に基づいた検索やブラウジングの実現など、効率的なアクセス手段が必要となる。このため、我々は、ニュース映像アーカイブの構築とともにニュース映像を意味に基づき構造化する手法 [1] の検討を行い、トピックに基づくニュース映像の閲覧を可能とした。しかし、そのインターフェースにおけるトピック内容の表示は、キーとなる画像の選定の難しさから、文字放送テキストの表示に頼るところが大きかった。

本発表では、ニュース映像アーカイブの可視化を目指し、ニューストピックを代表するキー画像の取得について検討を行ったので報告する。

2. 意味構造に基づくキー画像の取得の枠組

日々放送されるニュース映像は、複数の話題（トピック）から構成され、同一あるいは関連する話題を継続的に取り上げている。このため、複数のニュース映像の間には、話題の分岐や連続性など、単一のニュース映像からは得られない話題のつながり（トピック構造）を考えることができる。

このようなニュース映像の内容を提示する場合、多くのシステムでは、ニュース映像毎にその映像を構成するトピックをリスト形式で表示し、各トピックはそのトピックを構成する最初あるいはすべてのショットの画像を表示する、といった方法が取られてきた。しかし、このような方法では、キャスターショットの画像などトピック内容と関連性の低い画像が多く表示され、トピックの概要を把握するには不十分であった。

一方、ニュース映像は、画像的には様々なショットから構成されているが、同一あるいは関連する話題を継続的に取り上げているため、長期間のニュース映像を考えた場合、天気予報などの定型的なショットとともに、それらの話題に関連した重要なショットや貴重なショット、あるいは資料的なショットは繰り返し利用される傾向がある（図 1）。このため、ニュース映像アーカイブ内で複数回出現するショット（同一映像ショット）の画像は、ニューストピックを代表するキー画像の一つと考えることができる。

そこで、我々は、ニューストピックを代表するキー画

像の一つとして、トピックに含まれる同一映像ショットを抽出する方法を提案する。ただし、ニューストピックの概要を把握できるようにするためには、対象のトピックのみならず、そのトピックに関連するトピックの内容も同時に表現できることが望ましい。また、同一映像ショットの数は多くないため、単一のトピックに含まれる同一映像ショットが抽出できない場合も考えられる。そこで、キー画像の取得に際しては、トピック構造において、対象のトピック（主トピック）から派生するトピック群（トピッククラス）に含まれる同一映像ショットを抽出することとした（図 2）。



図 1: 繰り返し利用される映像ショットの例

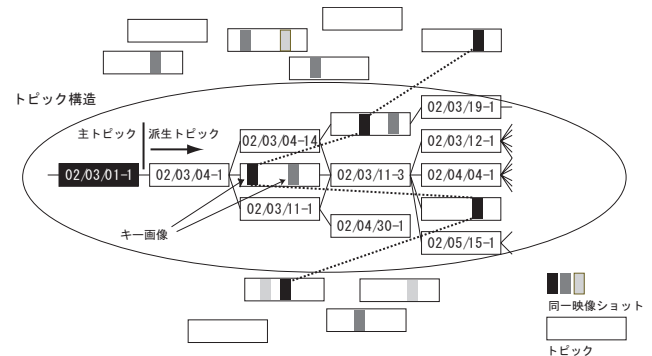


図 2: 意味構造に基づくキー画像取得の枠組

3. 意味構造の抽出とキー画像の取得

3.1 トピック構造の抽出

トピックへの分割は、文字放送テキストを用いて、以下の手順で行う。

1. 文字放送テキストの形態素を解析し、名詞列を抽出
2. 名詞列の語義属性（一般、人物、場所・組織、時相）を解析 [2] し、文単位で属性別の出現頻度付きキーワードベクトル $(\vec{k}_g, \vec{k}_p, \vec{k}_l, \vec{k}_t)$ を作成
3. 各文の境界で前後 w 文を結合した語義属性別のキーワードベクトルを作成し、隣接する文間の類似度（ \cos 距離）を算出
4. 各 w に対する語義属性別の類似度の重み付き和を計算し、閾値 θ_{seg} を用いてトピック境界を検出

[†] 国立情報学研究所, National Institute of Informatics

[‡] 東京大学大学院情報理工学系研究科, The University of Tokyo

^{*} 名古屋大学大学院情報科学研究科, Nagoya University

- トピック毎のキーワードベクトルから隣接するトピック間の類似度を計算し、閾値 θ_{cat} を用いてトピックを統合

話題のつながり具合の判定は、分割された全てのトピック相互間の類似度を評価して行う。トピック相互間の類似度が、ある閾値 θ_{trk} を超えた場合、これらのトピックは関連するとみなし、以下の手順を再帰的に実行することで、関連するトピックを階層的に構造化する。

- あるトピック（親）に関連する全てのトピック（子）を関連付ける
- ある子トピックに対して、兄弟や子孫で関連するトピックがあれば、時系列的に最も近いものの子として関連付けをし直す

関連したトピックを時系列に連鎖させることで、話題の分岐や連続性を構造化し、特定のニューストピックに関する情報を抽出しやすくする。

3.2 同一映像ショットの抽出

同一映像ショットの抽出は、画像的にほとんど同一のフレームを検出 [3] し、輝度情報を用いたショット分割の結果とあわせ、それらをショット単位へ統合することにより実行する。“同一”の判定には、全画面での精密な照合を行う必要があるため、輝度情報を用いたフレーム間の正規化相互相関 (NCC) を利用した。NCC は、-1 から 1 までの値をとり、全く同一の画像同士では $NCC = 1$ となる。本研究では、相関値がきわめて 1 に近い場合 (0.9 以上) を“同一”であると判定する。

3.3 キー画像の取得

ニューストピックを代表するキー画像を、そのトピックから派生するトピッククラス内に含まれる同一映像ショットの画像として取得する。同一映像ショットの種類を $IVS_i (i = 1, \dots, I)$ 、同一映像ショット IVS_i に属する各ショットを $VS_{ij} (j = 1, \dots, J_i)$ とする。また、トピッククラスの種類を $TPC_p (p = 1, \dots, P)$ 、トピッククラス TPC_p に属する各トピックを $TP_{pq} (q = 1, \dots, Q_p)$ とする。このとき、トピックを代表するキー画像は、トピッククラス TPC_p に属する各トピック TP_{pq} と時間的に重なりを持つ、同一映像ショット IVS_i に属する各ショット VS_{ij} の画像として求めることができる。

得られたキー画像がどの程度トピックを代表するかを示す指標として、トピッククラス TPC_p における同一映像ショット IVS_i の重要度 $S(p, i)$ を次式のように定義する。

$$S(p, i) = M_{ip} \log(N_{IVS} / J_i)$$

ここで、 $M_{ip} (M_{ip} \leq J_i)$ は、同一映像ショット IVS_i に属する各ショット VS_{ij} のうちトピッククラス TPC_p に属するショット数、 N_{IVS} は、映像アーカイブ内の同一映像ショットの総数である。

これは、あるトピックに対し重要なショットあるいは貴重なショットの出現頻度は、映像アーカイブ全体での割合は低いとそのトピックでは高くなる、逆に、あるトピックに対し密接に関係しない資料的ショットあるいは

定型的なショットの出現頻度は、映像アーカイブ全体での割合は高いがそのトピックでは低くなる、との考えに基づいたものである。

4. 実験結果

2001年9月1日から2003年2月5日までにNHKで放送された「ニュース7」のニュース映像を用いてトピッククラスに対するキー画像の取得に関する実験を行った。約500日分のデータから、オープニングやエンディングのショットを除いて、1547種類、それぞれ2~46ヶ所で見出している同一映像ショットを抽出することができた。

「拉致被害者」に関するあるニューストピック (2002年10月16日トピック#1) から派生するトピッククラス (3階層) を代表すると判定されたキー画像、上位8画像を図3に示す。この例では、トピッククラスは36トピックから構成され、関係する同一映像ショットは57種類であった。キー画像としては、帰国した拉致被害者らの画像を中心に、被害者の親族、北朝鮮核施設などに絡む画像が得られた。



図3: 拉致被害者に関するあるトピッククラスのキー画像

5. おわりに

本稿では、ニュース映像アーカイブへの効果的なアクセスを実現するため、ニューストピックを代表するキー画像の取得について検討を行った。キー画像の取得は、文字放送テキストから得られるトピック構造と、繰り返し利用される同一映像ショットの情報を利用することで実現した。キー画像の取得実験では、個々のトピックの内容を具体的に示すまでの画像は得られていないが、トピックの主題を判断するには有効と思われる画像を得ることができた。今後は、ショットの重要度とともに、話題の分岐や連続性などのトピック構造を考慮したキー画像の選択、提示について検討を行っていく予定である。

参考文献

- 井手 一郎, 孟 洋, 片山 紀生, 佐藤 真一, “大規模ニュース映像コーパスの意味構造解析,” 信学技報, PRMU2003-97, 2003.
- 井手 一郎, 浜田 玲子, 坂井 修一, 田中 英彦, “テレビニュース字幕の語義属性解析のための辞書作成,” 信学論 (D-II), vol.J85-D-II, no.7, pp.1201-1210, 2002.
- 山岸 史典, 佐藤 真一, 浜田 喬, 坂内 正夫, “大規模放送映像アーカイブにおける映像断片照合の提案と高速化,” 信学技報, PRMU2002-166, 2002.