

ビデオ・マンガ要約を用いたインタラクティブなビデオ閲覧

内橋 真吾

富士ゼロックス株式会社

IT メディア研究所

〒 259-0157 神奈川県足柄上郡中井町境 430

{shingo.uchihashi@fujixerox.co.jp}

概要

ビデオの内容を大きさの異なるキーフレームを使って視覚的に捉えやすい形態で表示するビデオ要約を用い、インタラクティブにビデオの再生を行うインターフェースを構築した。テキストの埋め込みや詳細表示などの機能を備え、希望のビデオやビデオ中の特定箇所を効率良く探し出すことが出来る。

1. はじめに

ビデオには内容を簡単に閲覧することや見たい場所を即座に探すことが困難という性質があり、ライブラリ中の大量のビデオから希望のものを探すことは容易ではない。仮にキーワードなどの付加情報が十分に備わっていて検索により候補を絞り込めても、その中から特定の箇所を取り出すには内容を一本づつ確認する必要がある。本研究では、こうした煩わしさを軽減するために、ビデオの内容をビデオ中の主要な場面を漫画のように大きさの異なるキーフレームで並べて表示して一目でわかる形式に自動的に要約するビデオ・マンガ要約を提案した。さらに作成した要約を用いて各キーフレームが示す箇所からのビデオ再生や詳細確認をインタラクティブに行えるユーザー・インターフェースを構築した。

2. 関連研究

ビデオ要約で二種類の手法がよく用いられる。一つは、ビデオをカットで区切り、各カットから一枚のキーフレームを抽出して並べて表示する方法 [2] で、もう一つは、ビデオ中の各フレームをグループ分けし、各グループから一枚づつキーフレームを選び表示する方法である [3]。前者は表示画像の量が多く内容も冗長になるため一目で内容を把握出来る要約にはならない。後者は画面の乱れなどで瞬間的にだけ大きく見えた目が変わった画像を選ぶ傾向があり出力がビデオ全体を代表しない懸念がある。第三の手法として、ビデオの各セグメントについて重みを計算し、主要なもののみからキーフレームを抽出する方法が提案された [6]。本研究のビデオ・マンガ要約もこれに属する。

3. ビデオ・マンガ要約

本研究で提案するビデオ要約法は大きく分けて三つの段階から成る。詳しくは [4] で述べているのでここでは簡単に概要を説明する。

3.1. ビデオのセグメント化

まずビデオ中の連続して似たフレームが出現する部分をセグメントとし、ビデオをセグメント単位に分割する。我々はビデオ中の各フレームを階層クラスタリ

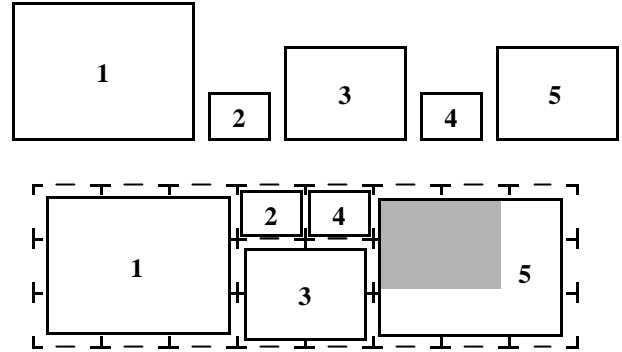


図 1: キーフレーム詰め込み模式図

ング法によりクラスタに分類し、同一クラスタに属する連続したフレームをセグメントとした。この方法では、セグメント分けした時点で元のクラスタ情報よりビデオ中での類似箇所の分布がわかる利点がある。

3.2. キーフレーム選択

次にビデオ中の各セグメントに対してセグメント長と類似セグメントの量より判断する希少性とから重要度を計算する。我々の提案する重要度は、長さのみを評価基準に用いた [6] と比較してビデオ内容をよりよく反映する [4]。重要度上位のセグメントを任意の数だけ選び、各セグメントにつきキーフレームを抽出する。今回は時間軸上でセグメントの中心に最も近いフレームを選んだ。さらに選び出したキーフレームに対し重要度に応じて大きさを割当てる。

3.3. レイアウトの決定

最後に選ばれた大きさの異なるキーフレームを並べて 1 ページの領域に表示する。領域全体をまず小領域に分割し、元のキーフレーム列の大きさを可能な限り保存して各小領域を効率的に詰めるキーフレーム列を探すことによりレイアウトを決定した。レイアウト決定例を図 1 に示す。この例では 5 個のフレームを左から右に順序を追えるよう配置し、5 番目のフレームの大きさを変更して領域を埋めている。

4. インタラクティブ・ビデオ要約

前節で述べたマンガ要約を使ってビデオをインタラクティブに操作するインターフェースを Java アプレットとして実装した。図 2 のように、通常のウェブ・ブラウザを用いてビデオ内容を確認出来る。映画ポスターを模倣して時間軸上の順序と無関係に重要シーンを配置する [6] の要約法と異なり、ビデオ再生しなくても流れを読取れる表示となっている。図 3 に示すようにポインタを各キーフレーム上に動かすと、そのフレームが色枠によって強調され、同時に時間軸上での位置も視覚的にわかりやすいよう



図 2: ブラウザ上で表示したビデオ要約



図 3: インターフェースの各要素



図 4: キーフレームからのビデオ再生

ロックで表示される。キーフレーム右下隅の四角枠に C の記号はそこに文字情報が埋め込まれていることを表し、ポインタをそのフレーム上に動かすと文字情報が吹き出し状の枠中に表示される。この文字情報には、例えばビデオ音声の書起し、脚本、視聴者のメモ等様々なものが考えられる。図 3 の例では、ビデオ音声から書起したテキストを [5] 記載の手法により区表現要約したものを使用した。このインターフェースは2ボタンマウスによる操作を想定しており、キーフレームを左ボタンでクリックすると図4のように該当する箇所からビデオが再生される。また、右ボタンでクリックすると要約表示で省略



図 5: 詳細を表示させた例

された場面を一段掘下げて図5のように最大4フレームの詳細を表示する。これらの機能を使ってユーザーはビデオ要約で全体の内容を確認しながら見たい箇所を絞り込んで選択的に見ることが出来、効率の良いビデオ閲覧が可能となる。

5. 考察

本研究で構築したビデオ要約システムは FX Palo Alto Laboratory, Inc., CA, USA の研究用ビデオ・ライブラリ MBase(<http://www.fxpal.xerox.com/mbase/>) 上に実装し動作検証を行い、登録された総計 100 時間を越す各種ビデオに対し有効性を確認した。比較実験より本稿のビデオ・インターフェースはユーザーにとって好ましいものであるという評価も得た [1]。

6. 参考文献

- [1] Boreczky, J., Girgensohn, A., Golovchinsky, G. and Uchihashi, S., "An Interactive Comic Book Presentation for Exploring Video," in *Proc. ACM CHI 2000*, pp. 185-192, 2000.
- [2] Shararay, B. and Gibbon, D. C., "Automated Authoring of Hypermedia Documents of Video Programs," in *Proc. ACM Multimedia 95*, pp. 401-409, 1995.
- [3] Taskiran, C., Chen, J.-Y., Bouman, C.A., and Delp, E.J., "Compressed Video Database Structured for Active Browsing and Search," in *Proc. ICIP '98*, vol. 3, pp 133-137, 1998.
- [4] Uchihashi, S., Foote, J., Girgensohn, A., and Boreczky, J., "Video Manga: Generating Semantically Meaningful Video Summaries," in *Proc. ACM Multimedia 99*, pp. 383-392, 1999.
- [5] Ueda, Y., Oka, M., Koyama, T., and Miyauchi, T., "Toward the "At-a-glance" Summary: Phrase-representation Summarization Method," in *Proc. COLING-2000*, pp. 878-884, 2000.
- [6] Yeung, M. and Yeo, B.-L., "Video Visualization for Compact Presentation and Fast Browsing of Pictorial Content," in *IEEE Trans. Circuit and Sys. for Video Technology*, Vol. 7, No. 5, pp. 771-785, Oct. 1997.