

複数講演コンテンツのダイジェスト自動生成のための シーン重要度算出手法の評価

呉 怡[†] 渡辺 陽介^{††} 横田 治夫^{†,††}

[†] 東京工業大学大学院 情報理工学研究科 計算工学専攻

^{††} 東京工業大学 学術国際情報センター

1 まえがき

近年、講演の様子を記録した動画アーカイブが大量に蓄積されてきた。例えば、日本データベース学会 [1] では、過去に行われた研究シンポジウムにおける発表の様子を収録した 1,200 件以上の講演コンテンツを会員向けに配信している。しかし、長時間にわたる動画の中から欲しい情報を見つけ出すことは多くの時間を要する。

このような背景により、我々は講義・講演コンテンツの中から、キーワードに関連する動画シーンを検索するシステム UPRISE を提案した [2]。そして、ユーザが短時間で、興味のあるトピックに関連する複数の講演コンテンツの概要を把握できるように、キーワードに合致する複数の講演の中から、重要なシーンを抽出して、ユーザが指定した時間長に収まるダイジェストを提供するシステムを構築している [3]。

しかし、講演ダイジェストとして、内容の概要を提供するだけでなく、講演の面白さをユーザに伝えることも必要である。これまでの手法では主にテキストをベースにシーンの重要度を算出していたため、図やグラフを用いて説明するような視覚的効果の高いシーンへの評価が常に低く、動画ならではの説明部分を含むダイジェストの生成が困難であった。

そこで本稿では、論文 [4] で提案している視覚的効果に基づく重要度算出式を改良した色分布とアニメーション使用回数を考慮する算出手法を提案し、被験者実験によって評価を行なう。

2 動画ならではのシーンの抽出

ダイジェスト自動生成までの処理の流れは図 1 で示すように、まず入力キーワードに基づき、全講演に対してランク付けをする (STEP1)。次に、単語の出現状況やそのほかの情報を用いて、各講演に含むシーンの重要度を算出する (STEP2)。最後に、ユーザが指定した時間長に収まるように、時間配分をして、ステップ 2 で算出したシーン重要度に基づき、ダイジェストを生成する (STEP3)。ダイジェスト自動生成に用いているキーワードに関連する講演コンテンツの検索および

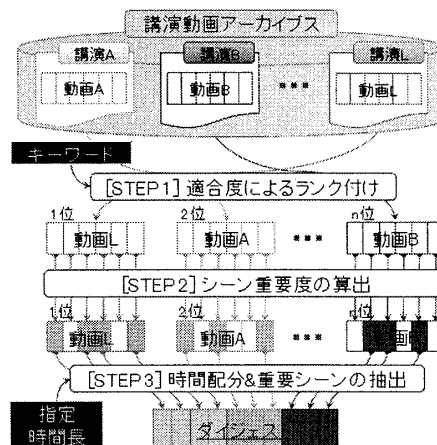


図 1: ダイジェストの自動生成

時間配分の手法については、論文 [3] を参照されたい。

ここでは、講演コンテンツにおける図などを使って説明するような動画ならではのシーンを含むダイジェストを生成するため、色分布とアニメーション使用回数を用いた重要度算出手法について述べる。

2.1 色分布とアニメーション使用回数を用いた重要度

動画ならではの説明部分は、図やグラフを用いてわかりやすく説明している場面であると考えられる。そのようなシーンには色遣いがカラフルで、アニメーションの使用回数が比較的に多い特徴があるため、ここでは、シーン c に対して、色 i ($1 \leq i \leq R$) のシーン c における画素値を r_i とし、色が多く使用され、かつより特別な色遣いであるシーンを重要と仮定して、色分布に基づく重要度 I_{color} を以下のように定義する。

$$I_{color} = \frac{1}{\sum_{i=1}^R r_i} * \left(\sum_{i=1}^R r_i * \log\left(\frac{C}{cf(i)} + 1\right) \right) \quad (1)$$

ただし、 C は講演コンテンツが持つシーンの枚数とし、 $cf(i)$ は講演コンテンツの中、色 i が使われているシーンの枚数とする。

また、シーン c におけるアニメーションの使用回数を N とし、アニメーション使用回数に基づく重要度 $I_{animation}$ を以下のように定義する。

$$I_{animation} = \log(N + 1) \quad (2)$$

さらに、パラメータ γ を用いて、 I_{color} と $I_{animation}$

Evaluation of Importance Calculation Methods for Automatic Digesting of Multiple Presentation Contents

Yi WU[†], Yousuke WATANABE^{††} and Haruo YOKOTA^{†,††}

[†]Department of Computer Science, Graduate School of Information Science and Engineering, Tokyo Institute of Technology

^{††}Global Scientific Information and Computing Center, Tokyo Institute of Technology

[†]{goi,watanabe}@de.cs.titech.ac.jp

^{††}yokota@cs.titech.ac.jp

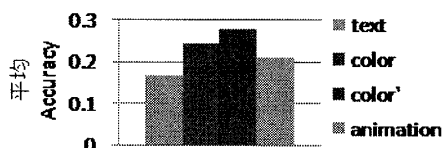


図 2: I_{text} , I_{color} , I'_{color} と $I_{animation}$ の実験結果

を組み合わせた重要度 I_{ac} を以下のように定義する。

$$I_{ac} = I_{color}^{\gamma} * I_{animation}^{1-\gamma} \quad (3)$$

以下では、視覚的効果の低い黒とグレーを重要度算出に使用しない場合の重要度を I'_{color} と I'_{ac} で表す。

3 評価実験

実験では、34 件の講演コンテンツの中から、2 種類のキーワードに関連する講演のダイジェストを自動生成する。

手法の有用性を確認するために、被験者の協力を得て、印象的なシーン・興味を引くシーンを重視して各関連講演コンテンツのシーンに点数を付けた。その点数の和とダイジェストとして抽出されるシーンの枚数によって、正解セットを作成する。

各重要度算出手法を 2 種類のキーワードに適用して生成したダイジェストにあるシーンと正解シーンと比較し、Accuracy の平均で評価を行なう。

$$Accuracy = \frac{|正解セット \cap \text{ダイジェスト}|}{|\text{ダイジェスト}|} \quad (4)$$

3.1 実験 1

実験 1 では、黒とグレーを含む色分布情報を考慮した I_{color} と、黒とグレーを色分布として考慮しない I'_{color} と、アニメーション使用回数を考慮した $I_{animation}$ を用いて、実験を行なった。比較対象として、これまで概要となるダイジェストを生成するためのテキストベースの重要度 I_{text} を用いた。

図 2 で示すように、提案手法の全てがテキストを中心とした手法より、よい結果を得られた。動画ならではのシーンの抽出においては、色分布とアニメーション使用回数を考慮することが有効である。特に、黒とグレーを考慮しない I'_{color} が最もよい結果となった。

3.2 実験 2

実験 2 では、 γ を 0.0 から 1.0 まで 0.1 刻みで変化させ、 I_{ac} と I'_{ac} における最適なパラメータを調べた。その結果を図 3 で示すように、色分布とアニメーション使用回数を組み合わせることにより、Accuracy の改善が見られた。黒とグレーの 2 色を考慮した I_{ac} では結果に変化はなかったが、黒とグレーを考慮しない I'_{ac} では $\gamma = 0.0$ から 0.8 にかけて、平均 Accuracy が増加し、 $\gamma = 0.7$ と 0.8 で最大値の 0.472 となった。

3.3 実験 3

実験 3 では、異なる時間長のダイジェストを生成する時の提案手法の有用性を確認するために、ダイジェストの時間長を 180 秒から 900 秒まで、180 秒刻みで変化させた時の平均 Accuracy を算出し、その結果を図 4 にまとめた。グラフからわかるように、 $D = 900$ 秒のときでは、 I_{ac} と I'_{ac} が同じ結果となったが、そのほ

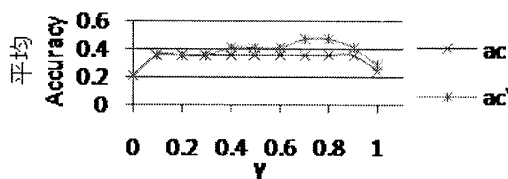


図 3: I_{ac} と I'_{ac} の実験結果

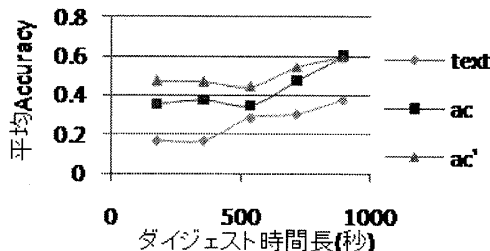


図 4: 実験 3 の結果

かの場合では、黒とグレーの 2 色を考慮しない色分布情報とアニメーション使用回数を組み合わせる手法のほうが良かった。

また、全体において、色分布情報とアニメーション使用回数を用いた手法はテキストを中心としたこれまでの手法よりも高い Accuracy となり、動画ならではのシーンの抽出に有効性を示した。

4 まとめと今後の課題

本稿では、ユーザの興味を引き出すダイジェストの提供を目的とし、動画ならではの説明部分をダイジェストに含めるように、スライド内の色分布とアニメーション使用回数を考慮したシーン重要度算出手法を提案し、被験者実験によって評価を行なって、手法の有用性を確認した。

今後、まずキーワードの種類を増やして実験を行う予定である。また、これまで提案した手法の統合や、音声情報の利用なども今後の課題である。

謝辞

なお、本研究の一部は文部科学省科学研究費補助金特定領域研究 (#21013017) の助成により行われた。

参考文献

- [1] 日本データベース学会. DBSJ Archives. <http://www.dbsj.org/>.
- [2] Haruo Yokota, Takashi Kobayashi, Taichi Muraki, and Satoshi Naoi. UPRISE: Unified Presentation Slide Retrieval by Impression Search Engine. *IEICE Trans Inf Syst*, Vol. E87-D, No. 2, pp. 397-406, 2004.
- [3] 呉怡, 渡辺陽介, 横田治夫. 講演動画アーカイブスからキーワードに合致する講演ダイジェストの自動作成. DEIM フォーラム 2009 論文集, pp. E9-2, 2009.
- [4] 呉怡, 渡辺陽介, 横田治夫. 講演コンテンツにおける視覚的効果に基づくダイジェストの自動生成. DEIM フォーラム 2010 論文集, 2010.