

ネットワーク情報配信プラットフォームとしての高自由度映像方式

－ ハイパーリンクムービーの実現 － (051301002)

Highly flexible network information delivery platform

－ End-user ready hyperlink movie implementation －

研究代表者

長谷山 美紀 北海道大学情報科学研究科

Miki HASEYAMA School of Information Science and Technology, Hokkaido University

研究分担者

山本 強

荒木 健治

土橋 宜典

伊藤 敏彦

Tsuyoshi YAMAMOTO

Kenji ARAKI

Yoshinori DOBASHI

Toshihiko ITO

北海道大学情報科学研究科

School of Information Science and Technology, Hokkaido University

研究期間 平成 17 年度～平成 19 年度

本研究開発の概要

インターネットとテレビ映像を融合した、新しいコンセプトで実現されるネットワーク情報配信プラットフォームとしての映像方式「ハイパーリンクムービー」を提案する。提案方式は、従来の映像メディアには無い制御軸（ネットワーク情報アクセス）を加えることによって、高い自由度をユーザに提供するだけでなく、ネットワーク上に存在する情報からユーザが希望する情報を自動的に検出する機能の実現を試みる。具体的には、映像を動画像と音響信号の統合データとして捉え、音響信号については、音楽と音声を異なるモデルで表現することで、映像の類似度を定義し、動画像信号については、撮像物体の認識に大量の学習データを要するシステム構造を避け、実用化に簡便なシステムとした。さらに、音声について、既存のテキスト変換手法により得られる結果を利用することで、音声認識精度が低い場合でも、それを利用することで、映像の統合データとしての性質に着眼した高精度な映像検索システムが実現された。実際に、ニュース映像の推薦システムを試作し、インターネット上に存在するニュース映像とのハイパーリンクの実現性が示された。

Abstract

A novel information delivery platform, hyperlink movie, fusing the Internet and television is presented in this article. This platform provides network access of multimedia contents and realizes fully automatic and highly flexible information retrieval on users' demand. The hyperlink system for network access is established based on a multimodal integration of visual and audio analysis techniques combined with speech-to-text ones. A prototype recommendation system has been implemented to evaluate the practicability of the hyperlinks among multimedia contents on the Internet.

1. まえがき

近年の記憶媒体の大容量化やインターネットによる映像配信サービスの普及に伴い、ユーザは大量の映像を容易に手にすることとなった。しかしながら、ユーザが希望する内容の映像を入手する際には、映像名や製作元等の予め手動で付与されたメタデータを手がかりにして検索が行われている。映像データの爆発的増加とその視聴形態の多様化が進む現在において、手動によるメタデータに頼らず、映像の内容を理解した検索技術の開発は急務である。

このような要求の下、ここ数年、コーパスベースのメディア解析手法が提案され、映像情報の意味理解技術は実用に向けて大きく前進した。コーパスベースの手法は、意味理解に有効であるが、使用したコーパスを用いるかにより、適用可能な映像の種類が限定され、大量のコーパスが必要であるなどの問題がある。そこで、申請者は、ユーザの多様な映像の好みに応えることを可能とすべく、インターネットとテレビ映像を融合したネットワーク情報配信プラットフォーム「ハイパーリンクムービー」の提案に至った。実現のために構築した映像処理基盤技術と高自由度映像アクセス基盤技術について各々以下に説明する。

2. 研究内容及び成果

2.1 映像処理基盤技術構築

2.1.1 個人が望む映像コンテンツの自動検索

申請時に記述したように、個人が望むコンテンツの自動検索の可能性について検討を行い、「個人の思考を模擬した情報収集システム」の実現のために、次の2つの手法を提案した。

(1) 個人が希望する映像を自動的に選択する学習手法

ユーザの映像の視聴履歴を学習データとして強化学習により、自動的に個人が好む映像を特定する手法を実現した。実現されたシステムを図1に示す。

(2) 個人が希望する映像を自動的に選択するための音楽シーン検索手法

(1)だけでは、多様なユーザのニーズに応えることが困難であるため、更なる拡張を目指し、音楽シーン検索を可能とする手法を提案した。従来の採譜を可能とする音楽信号の特徴解析手法と異なり、ノイズに影響を受けずに、類似音楽が存在する映像を提示する。実現したシステムを図2に示す。

2.1.2 個人の多様なニーズに応えるコンテンツ自動生成

TVで放送されたサッカー映像を映像ソースとして、個人が仮想試合を体験する手法を提案した。サッカー映像に動的輪郭モデルの一つとして知られる、アクティブネットに基づく新たな手法を提案することで、サッカー映像からパスコースを自動抽出する。この手法を実際の映像から選

手位置を変更した映像に適応することで、実試合とは異なる仮想試合を擬似的に体験することができる。実現されたシステムを図3に示す。

2.2 高自由度映像アクセス基盤技術構築

2.1で検討した、完全に自動化されたメタデータ付与手法により、蓄積メディア上に存在する映像データとインターネット上のデータとの利用の連携を可能とするシステムを実現した。以下、実現した各技術について記載する。

クエリ映像が明確に存在しない場合であっても、蓄積メディア上の映像とWebデータを活用することで、ユーザが希望する映像にたどり着くことができるインタフェースの実現を試みた。提案システムでは映像間の類似度に基づいて、3次元の空間上に映像を配置することで、空間の距離によって映像の類似性を理解することが可能なインタフェースとした。尚、本システムにおいて個人の好みをモデル化するための『プリファランスボード』を準備し、それらから得られる操作履歴により学習を行うシステムを導入した。図4に提案システムによる映像検索の様子を示す。(a)には、クエリ映像を与えた直後の様子が示されており、その後の映像検索の結果が(b)に示されている。複数の映像がデータベース中から選択されていることが確認できる。ユーザはプリファランスボードを通して“音”と“絵”の両者に自由に重みを設定し、検索を繰り返すことで、希望の映像にたどり着く。明確なクエリを持ち合わせていない場合でも、このような連想型の検索により、ユーザが望む映像を獲得する。また、本研究開発により実現されたシステムには、音響信号の分類機能が備わっており、それにより音声部分と識別された信号部分に、既存の音声のテキスト変換を適用し、インターネットのコンテンツとのハイパーリンクを実現している。組み込まれた現状の音声認識器の脆弱性を映像処理の結果により補っている。

3. むすび

本研究開発が示す検索結果の可視化も含めた研究のアプローチは、映像が持つ曖昧さを許容できるだけでなく、個人や視聴形態の変化によって適応的なビジュアライゼーションを提供するインタフェースの実現という新しい研究として位置付けられる。

【誌上発表リスト】

- [1] Naoki Nitanda and Miki Haseyama: “Audio-Based Shot Classification for Audiovisual Indexing Using PCA, MGD and Fuzzy Algorithm,” IEICE Trans. Fundamentals, vol. E90-A, no. 8, pp. 1542-1548 (2007年8月1日)
- [2] Takahiro Ogawa and Miki Haseyama: “POCS-Based Texture Reconstruction Method Using Clustering Scheme by Kernel PCA,” IEICE Trans. Fundamentals, vol. E90-A, no. 8, pp. 1519-1527 (2007年8月1日)
- [3] Hirofumi Kon and Miki Haseyama: “Players Clustering Based on Graph Theory for Tactics Analysis Purpose in Soccer Videos,” IEICE Trans. Fundamentals, vol. E90-A, no. 8, pp. 1528-1533 (2007年8月1日)

【申請特許リスト】

- [1] 長谷山美紀、音響・動画像信号の特性に基づいた類似映像の検索・分類システムおよび検索・分類結果の可視化システム、日本、2008.3.19
- [2] 長谷山美紀、画像分類装置および画像分類プログラム、

日本、2007.12.3

- [3] 長谷山美紀、類似画像検索装置、日本、2007.11.6

【受賞リスト】

- [1] 土橋宜典、情報処理学会グラフィクスとCAD研究会優秀研究発表賞、“事例ベースアプローチによる動的な雲の多重散乱を考慮した効率的レンダリング”、2007年7月7日

【報道発表リスト】

- [1] “画像から「好み」測定、キーワード不要の検索 ネット通販などに応用”、朝日新聞朝刊、2007年10月5日掲載
- [2] “写真が導く 風景検索、道内5社と北大 実証サイト公開”、北海道新聞朝刊、2008年2月1日掲載
- [3] “特集「Googleを超えろ！国産検索エンジン」”、ワールドビジネスサテライト（テレビ東京）、2007年7月25日放映

【本研究開発課題を掲載したホームページ】

<http://www-lmd.ist.hokudai.ac.jp>



図1 Easy Finder -映像検索-

(a)ユーザの視聴希望映像（左縦4つの映像）。希望映像を特定することで、その映像と類似の映像を含む映像が検索される（中央3つの映像）。(b) (a)で自動的に提示された映像中に含まれる類似映像シーン（番組のコーナー）を自動的に検出する（右縦2映像の内、上部が検出された映像。下部はユーザが視聴を希望する映像）。中央部は提案手法における映像の特徴量。CEATEC 2006 (Cyber Space Navigator -次世代情報アクセス-)にて出展。



図2 Easy Finder-音楽検索-

左縦に録画映像を上部、横に音楽データベースを表示。ユーザが視聴した映像の音楽と類似の音楽を提示する。CEATEC 2006にて出展。

図3 Sim-Soccer

ユーザが選択したサッカー映像シーン（上段左）から、パスコースを抽出する（上段右）。ユーザが選手位置を変更することで（下段左2番目）、異なる選手の位置関係におけるパスコースが提示され（下段右）ユーザが仮想試合を体験できる。CEATEC 2006にて出展。



図4 (a) クエリ映像を与えた際の表示、(b) クエリ映像に対して映像検索を行った結果。CEATEC2007 (Cyber Space Navigator -映像の次世代検索システム-)にて出展。