

平成21年 6月 5日現在

研究種目：若手研究 (B)
 研究期間：2007～2008
 課題番号：19700126
 研究課題名 (和文) 動画閲覧行動からの知識抽出による閲覧スタイルの共有・編集環境の構築
 研究課題名 (英文) Sharing and Composing Video Viewing Style by Extracting Knowledge from Habitual Video Viewing Behaviors
 研究代表者
 高嶋 章雄 (TAKASHIMA AKIO)
 東京工科大学・コンピュータサイエンス学部・助教
 研究者番号：80421999

研究成果の概要：

本研究課題では動画閲覧スタイルの共有・編集が可能な環境の構築を目指し、(a)ユーザーによる動画閲覧操作と、動画の特徴とを関連付けるための理論構築、(b)蓄積された動画閲覧スタイルを新たな動画閲覧時に適用するための方法論の確立を行った。作成したシステム、Video Viewing Experience Reproducer を利用した評価実験において、(1) 動画閲覧スタイルの再利用を、人間が自分の閲覧スタイルを再現するのと同程度の精度で再現可能であること、(2) 動画閲覧スタイルの組み合わせにおける利点と欠点、とを明らかにした。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007 年度	1,700,000	0	1,700,000
2008 年度	1,400,000	420,000	1,820,000
年度			
年度			
年度			
総計	3,100,000	420,000	3,520,000

研究分野：インタラクシオンデザイン

科研費の分科・細目：情報学・知能情報学

キーワード：動画閲覧スタイル, Active Watching, 知識抽出

1. 研究開始当初の背景

計算機性能の向上とインターネットの発展に伴い、世界に散在する知識をエンドユーザーが収集して利用することができるようになり、さらには編集・流通させることで新たな知識を生み出すことが可能になりつつある。このような状況で扱われる知識とは、CGM (Consumer Generated Media) に代表されるメディアコンテンツや、Web 上のサービス、ソフトウェアの機能などであり、言語や数式で明確に形式化できるものが対象知識とされてきた。一方で、暗黙知としても知られる

ようなノウハウやスキルなど、経験的に獲得される知識の編集・流通に関する研究は少なく、十分な成果が無いのが現状である。

研究代表者はこれまでに、動画に対して人間がどのようにインタラクシオンを行うか調査してきており、特に視線追跡装置を利用した実験では、同じ動画でも、着目する部分、および早送りやスロー再生などの操作方法が異なる場合があることを確認している。さらに、地上波デジタル放送の開始や、汎用 DVD プレイヤーにハイライト再生の機能が搭載され始めていることなどからも、動画が

より双方向的なメディアとなることは想像に難くない。しかしながら、動画閲覧に関する既存の研究においては、映像のドメイン知識を利用し、例えばサッカーゲームのシュートシーンの抜き出しや、ニュース番組のトピック分割など、映像分析に基づいた要約やハイライト抽出などが主流であり、閲覧者がどのような閲覧方法を行うかに注目した研究は極めて少ない。つまり、経験的に獲得される動画閲覧スタイルを知識として扱い、編集・流通させるような取り組みはなされておらず、世界中に散在するであろう多種多様な閲覧スタイルをもとに、動画閲覧経験を共有したり、新たな閲覧スタイルを創出したりすることは不可能である。

2. 研究の目的

人間がある状況下で似通った行動を選択する場合、その行動は個々人の経験的知識に基づき選択されたものと考えられる。本研究課題では特に、動画閲覧における個々人の閲覧スタイルを題材とする。例えば、サッカーの試合を閲覧する場合、サッカーチームの監督が試合を分析するために閲覧する方法と、スタープレイヤーのファンが映像を閲覧する方法とが異なる場合が少なくない。早送りや、スロー再生などの行動に一貫性がある場合、その動画閲覧行動を選択する知識を抽出することで、他者の動画閲覧経験を追体験したり、個々人に特化した動画要約をしたり、さらには複数の閲覧スタイルを組み合わせた新たな閲覧方法を創出したりすることが可能となる。動画閲覧はこれまで、例えば映画のように「作品」としての動画を、閲覧者が「消費」という一方的な関係が主であったが、監視カメラの記録分析、ユーザ観察実験の記録分析、スポーツの試合記録分析など、映像分析の方法によって獲得できる情報が変化するという双方向メディアとしての利用が注目されており、自分や他者の動画閲覧スタイルを再利用することで、効率的、あるいは発見的な情報獲得が期待される。このように、本研究課題では動画閲覧スタイルの共有・編集が可能な環境の構築を目指す。経験的な知識を正確に記述することは困難であるため、本研究課題では、知識そのものを取り扱うのではなく、経験的知識に基づいて行われたであろう行動と、行動が選択された状況とを関連付けて取り扱うというアプローチをとる。

具体的には、以下の2点を目的とする。

- (a) ユーザによる動画閲覧操作と、動画の特徴とを関連付けるための理論構築
- (b) 蓄積された動画閲覧スタイルを新たな動

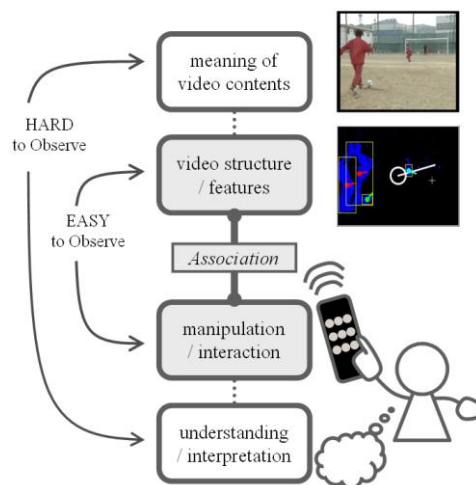


図 1. 状況と行動の関連付け

画閲覧時に適用するための方法論の確立
 上記(a)は、動画閲覧時に行われる操作を、どの動画特徴量と関連付けて閲覧スタイルとするかを明確にするものであり、(b)では、蓄積された様々な閲覧スタイルを、新たな状況にどのように適用してユーザに動画閲覧を追体験させるかを定めるものとなる。

3. 研究の方法

動画をどのように閲覧するかは、動画の意味的な内容や、閲覧者がその動画をどのように理解するかが少なからず関連する。しかしながらそれらを明確に定義し、観察することは非常に困難であるため、本研究課題では外部から観察可能な動画特徴と閲覧行動とを関連付けることで閲覧スタイルを表現する(図1)。つまり(a)で述べたように、動画閲覧における「状況」を動画の特徴量、「行動」をユーザによる早送りや巻き戻しなどの動画閲覧操作とする。例えばサッカーの試合映像の場合、ゴールシーンが重要であるといった固定観念を押し付けるのであれば、ドメイン知識を詰め込み、ゴールシーンを抽出する既存の研究成果などを利用することが可能であるが、ユーザがどのようなシーンに関して特徴的な行動をとるかは多岐に及び、全てを網羅するドメイン知識の準備は不可能である。そのため、映像の色分散や動きのあるオブジェクトのオブティカルフロー、カメラワークなど、ローレベルな動画特徴とユーザの動画閲覧行動とを関連付けることで、様々なユーザの閲覧スタイルへの対応を可能にする。

ユーザの動画閲覧行動については、既存の動画ブラウザで可能な操作を想定しており、閲

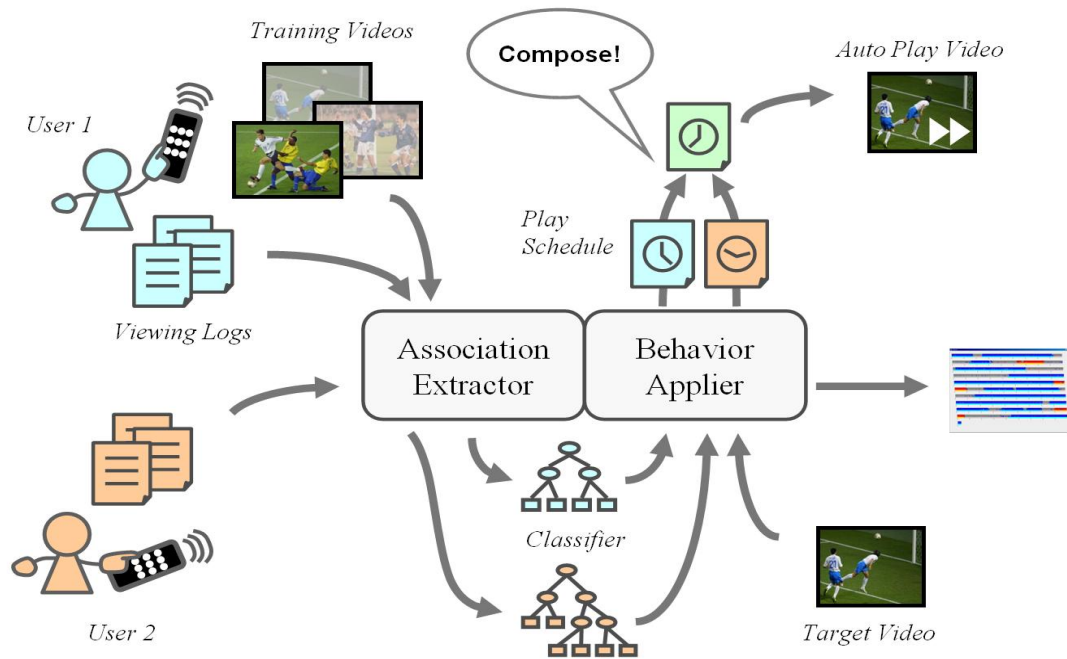


図 2. Video Viewing Experience Reproducer 概観

覧時の操作は多岐に渡らないため、操作の組み合わせかたのパターンを何らかの意図を伴う操作として考慮する。例えば、あるシーンを何度も早送りすれば、スキップするという行動であり、巻き戻しのあとスロー再生を行えば、あるシーンを再度詳察するという行動であると考えられる。本研究では、意図を伴うと考えられる操作パターンをあらかじめ設定し（操作パターンを利用するのみで、意図を明示するものではない）、特徴的な行動と動画特徴との関連付けを行う。

(b)では(a)の関連付け方法と連携しながら、蓄えられた閲覧スタイルの適用方法を探る。あるユーザの閲覧スタイルを利用するだけでなく、複数のユーザの閲覧スタイルを比較し頻度の高い閲覧方法をユーザに提供したり、頻度の低いインタラクションを集めてユーザに新たな動画閲覧の視点を与えたりするなど、閲覧スタイルの組み合わせを行う応用方法についても理論を確立する。

4. 研究成果

上述のアプローチに基づき、ユーザの動画閲覧スタイルを抽出し動画閲覧経験を提供するシステム、Video Viewing Experience Reproducer (以下 VVER) を構築した (図 2)。VVER は Association Extractor と Behavior Applier からなる。Association Extractor では、まず動画の特徴量を抽出し、どのような動画像特徴量を持つ映像フレームがどのように操作されるかを分類する分類器を作成する。Behavior Applier は、Target Video の各映像フレームの動画特徴量を算出し、そ

の値を分類器にかけ、それぞれのフレームをどのように再生するかを決定する。以下に詳細を述べるとともに、システム利用プロセスについて説明する。

Association Extractor では、ユーザによる複数の Training Video の閲覧ログと各ビデオの動画特徴量とを用いた教師付き学習により分類器が作成される。

動画特徴量としては、動画の各フレームにおける以下の 5 種類の特徴量抽出を行い、計 84 次元の特徴量を利用している。

- フレーム内に現れる色の統計値
- 最頻色の空間的な中心や分散
- オプティカルフローデータ
- 動作中のオブジェクトの数
- 音域の分布情報

フレーム内に現れる色の平均値や分散は、最もシンプルな動画フレームの情報と考えられ、HSV 表色系での値を利用している。最頻色の空間的な中心や分散は、例えばサッカーのフィールドがどの程度映像に含まれるかを知る一手段となる。現状では、最頻色の指定は、色空間における特定の範囲 (e. g. $30 < H <= 50, 80 < S <= 90, 75 < V <= 95$) を手動で設定しているが、あらかじめ学習に使う全動画における最頻色を計算し利用することで自動化することが可能である。オプティカルフローデータは、まず映像フレームを 12 分割し、各領域に含まれる特徴点を抽出する。その点が次フレームにおいてどう変化するかを算出し、各領域における変化量と方向の平均や分散値を特徴量として利用している。動作オブジェクトの数は、CAMSHIFT 法を用いて

抽出した動作オブジェクトを、その大きさに応じて3種類に分類してカウントした。音域の分布は、対象フレームにて再生される音などの音域（e.g. 0-32Hz, 32-64Hz, ..., 8000-10000Hz）に分布しているかを測定している。一方ユーザ操作としては、簡単のため、1倍速よりも早い速度で再生するSkip, あるシーンを巻き戻して1倍速以下のスピードで見直すRe-examineの2種類に限定し、通常速度で再生するNormalを合わせて計3種類とした。分類器作成にはデータマイニングツールのWEKAシステムを使用し、上述の84次元の動画特徴量を説明変数、3種類のユーザ操作を目的変数とした。

Behavior Applierは、Target Videoの動画特徴量をAssociation Extractorと同様に計算し、それを分類器にあてはめることで各フレームをどのような速度で再生するかを記したPlay Scheduleを作成する。Target VideoはPlay Scheduleに基づいて自動的に速度を変えながら再生される。実装上は、Skipと判断されたフレームを5倍速で、Re-examineと判断されたフレームを0.5倍速で、それ以外のフレームを1倍速で再生するようマッピングしている。また、Behavior Applierは、Play Scheduleにおけるノイズ（Skipと判断された連続するフレームの中に1フレームだけRe-examineと判断されたフレームが存在する場合など）を除去する機能と、Play Scheduleを確認する可視化機能を有している。

このようにVVERでは、Training Videoを用いてあらかじめ分類器を作成し、それに基づきTarget Videoを自動的に再生することで、ユーザに動画閲覧経験を与えている。

本研究では、特定の動画閲覧スタイルに基づいた動画閲覧経験をユーザに与えるだけでなく、他の動画閲覧スタイルに基づく動画閲覧経験を複数組み合わせることで、新たな動画閲覧経験を創出する環境を提供する。実際には、動画の自動再生時における速度変化を記したPlay Scheduleを組み合わせることで、新たな動画閲覧経験をユーザに与えるものである。組み合わせ方法としてIntersection $A \cap B := \{x | x \in A \text{ and } x \in B\}$, Complement $A \setminus B := \{x | x \in A \text{ and } x \notin B\}$, Union $A \cup B := \{x | x \in A \text{ or } x \in B\}$ という3種類を定義する。ここでA, Bは、あるビデオに含まれる全動画フレームがそれぞれどのようなユーザ操作（Skip, Re-examine, Normal）に関連付けられるかを表す集合であり、AとBふたりのユーザの集合を示している。xはそれらの集合に含まれる特定のフレームにおける関連を示す。これらを利用することで、一般的に広

く行われる動画閲覧方法や、自分は知らないが一般的に行われがちな閲覧方法などを創出し、ユーザに体験させることが可能となる。

システム評価実験として、ユーザの動画閲覧スタイルの再利用に関するユーザ観察、複数ユーザの閲覧スタイルの組み合わせに関するユーザ観察を行った。

動画閲覧スタイルの再利用に関するユーザ観察においては、被験者は、図2に示すTraining Videoを閲覧した後日、自分自身の閲覧スタイルに基づき自動的に再生されるTarget videoを閲覧した。閲覧後に、「自分の閲覧スタイルに基づき自動再生される動画像を見てどう感じたか」「その他の参考意見」について質問をしたところ、4名の被験者全員が自分の行った閲覧スタイルに近いと感じたと述べた。被験者へのインタビューに加え、閲覧スタイルがどの程度再現されているかを確認する定量的実験を行ったところ、システムによる閲覧スタイルの再現結果が、人間が意図的に閲覧スタイルを再現する結果と、同程度の精度を持つことが確認された。ドメイン知識を一切使わず、ユーザにとってある程度満足いく結果が得られた点は、動画の利用方法が多様化していくであろう今後の情報環境において、重要な意味を持つと考えられる。

複数ユーザの閲覧スタイルの組み合わせに関するユーザ観察においては、新たな動画閲覧経験が、ユーザにどのような印象を与えるかを調査した。結果として、一部の被験者が、サッカーの試合映像がダイジェスト版のように再生される閲覧スタイルに興味を抱いたり、普段の閲覧スタイルからは見出せなかった事象を確認できたことに興味を抱いたりする様子が観察された。一部の被験者において思いがけず興味のもてる経験をした様子が伺えた一方で、自分自身の閲覧スタイルとは異なる閲覧経験が提示された際、ネガティブな印象を抱くことがあった。これは組み合わせる場合だけでなく、他者の閲覧スタイルに基づく経験をそのまま享受するときにも起こりうるものである。動画閲覧経験がユーザにとって思いがけず面白いものと感じられるか、苛立ちを覚えさせるものになるかを事前に判断することは容易ではないが、ユーザがある程度それを把握できる仕組みが必要と考える。VVERにおけるPlay Scheduleを可視化する機能などを利用し、動画閲覧経験をユーザが把握できるような仕組みを取り入れることで、ユーザが閲覧経験をイメージしながら組み合わせを試行錯誤することが可能になると考えられる。

本研究では、動画閲覧のスキルやノウハウといった暗黙的な知識を、「動画閲覧スタイルの再利用」および「動画閲覧経験の組み合わせ」によってユーザが活用するための環境を構築した。

本研究課題の特色は、動画像に対する習慣的な体験方法である閲覧スタイルを、共有・再利用の対象として捉え、知識活用の支援を試みる点にある。人間の行動とその行動に至った状況とを細かく対応付けることができれば、逆に、ある状況に直面した人間に対し、どのような行動をとるべきか、どのような行動の選択肢が可能であるかを提示することができる。行動主義の立場でユーザの要求を完璧に予測し満たすことは容易ではないが、ユーザの意思決定を補助したり、新たな経験を与えることでユーザの興味を広げたりすることが可能となるであろう。

ストレージデバイスやセンサ類の発展、ユビキタス環境実現の取り組みを背景に、日常生活における人間の様々な行動履歴を蓄えておくことが可能になりつつあり、さらに今後は行動履歴だけでなく、その行動にいたった状況をもセットで蓄積することが可能となると考えられる。つまり状況と行動との関連付けを利用することで、たとえば料理のコツや、マシンの修理など、日常生活の様々な行動の元となる知識を活用することが可能となり、知識管理の新たなアプローチとなり得ると期待できる。本研究はそのアプローチを動画閲覧というドメインで行う意義深いものである。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計1件)

①高嶋章雄, 田中讓, 習慣的な動画閲覧行動の再利用による動画閲覧経験の拡張, 情報処理学会論文誌「21世紀の情報環境」特集号, Vol.49 No.7, pp.2589-2597, 2008 (査読あり)。

[学会発表] (計4件)

① Akio Takashima, Reproduction and Management of Video Viewing Experience, International Workshop on Knowledge Media Science 2008 (KMS2008), Leipzig, Germany, 2008.

② Akio Takashima, Enhancing Video Viewing Experience, Workshop on Perspectives of Intelligent Systems' Assistance 2007

(PISA2007), Sapporo, Japan, 2007.

③ Akio Takashima, Yuzuru Tanaka, Sharing Video Browsing Style by Associating Browsing Behavior with Low-level Features of Videos, Human-Computer Interaction, HCI Intelligent Multimodal Interaction Environments, Jacko, Julie A. (Eds.), pp.518-526, Beijing, China, 2007.

④ Akio Takashima, Yuzuru Tanaka, Reusing and Composing Habitual Behavior in Video Browsing, Proceedings of the 17th European Japanese Conference on Information Modelling and Knowledge Bases (EJC2007), pp134-141, Pori, Finland, 2007.

[図書] (計3件)

① Akio Takashima, Yuzuru Tanaka, Sharing and Composing Video Viewing Experience, A chapter in "Human-Computer Interaction", ISBN 978-953-7619-26-8, IN-TECH, 2009 (in print).

② Akio Takashima, Enhancing Video Viewing Experience, A chapter in "PISA (Perspectives of Intelligent Systems' Assistance) Book", IGI Global, 2009 (in print).

③ Akio Takashima, Yuzuru Tanaka, Reusing and Composing Habitual Behavior in Video Browsing, Information Modelling and Knowledge Bases XIX., H. Jaakkola, Y. Kiyoki and T. Tokuda. (Eds.), Frontiers in Artificial Intelligence and Applications Series, pp134-141, 408 pages, IOS Press, 2007.

6. 研究組織

(1) 研究代表者

高嶋 章雄 (TAKASHIMA AKIO)

東京工科大学・コンピュータサイエンス学部・助教

研究者番号 : 80421999