

平成23年 6月 3日現在

機関番号：32692

研究種目：若手研究 (B)

研究期間：2009～2010

課題番号：21700176

研究課題名 (和文) 動画像の時空間解像度操作を利用した閲覧知識活用環境の構築

研究課題名 (英文) Reusing Video Viewing Experience as Habitual Behavior on Distorting the Resolutions of Time and Space

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 (和文)：

本研究課題では、動画像の時空間解像度操作を利用した閲覧知識活用環境の構築を目的とし、具体的には、(1)高解像度動画像とのインタラクションを可能にするブラウザの作成 (2) ユーザの操作と動画像特徴量との関連付けによる知識抽出 (3) 抽出された知識と動画像特徴量に基づく動画像解析をサブゴールとして研究を実施した。本研究の取り組みから、動画記録方法が現在よりも空間的・時間的に高解像度になった際のインタラクションを、知識として抽出し流用する可能性を示すことができた。

研究成果の概要 (英文)：

This research work aims to develop an environment for reusing video-viewing-styles as habitual behavior. The work (1) presents the browser which allows users to watch and interact with spatially and temporally high resolution video, (2) extracts associations between the user's manipulation of videos and the low-level features of these video, and then (3) reproduces the video viewing experience by using the associations. As a result from an informal user study, the system suggests the potential of using video interaction as knowledge.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	2,000,000	600,000	2,600,000
2010年度	1,000,000	300,000	1,300,000
年度			
年度			
年度			
総計	3,000,000	900,000	3,900,000

研究分野：インタラクションデザイン

科研費の分科・細目：情報学・知能情報学

キーワード：インタラクション、動画閲覧知識、時間解像度

1. 研究開始当初の背景

計算機の普及に伴い、情報の理解のために動画像が利用される機会が増えてきている。これまでは、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

一方向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

いる。加えて、保存できる動画像の空間的、時間的高解像度化が進んでいる。空間解像度の向上は、単純に映像が美しくなるだけでなく、動画像中の空間的な一部分を拡大しても情報を読み取ることができる利点がある。時間解像度の向上は、高速度カメラに代表されるように、人間の目では通常とらえることができない事象を分析できるという利点がある。

このようにリッチな動画データを閲覧し、情報を理解する作業や、動画像とのインタラクションについては、十分に研究がなされていない。情報の理解においては、その全体像(概観)を知るためにそれを構成する部分を理解する必要がある一方で、部分(詳細)を理解するためには全体を知っておく必要があるとされる(解釈学的循環)。空間に広がる情報の把握については、情報可視化分野において、Focus+Contextなどの概観と詳細を把握する技術が広く研究されてきた。一方で時間軸における情報の概観と詳細の表現やインタラクションについては十分に議論されておらず、特に時間解像度に焦点を当てて情報の概観と詳細を表現、操作する報告は少ない。

2. 研究の目的

本研究課題では、動画像の時空間解像度操作を利用した閲覧知識活用環境の構築を目的とする。具体的には、

- (1) 高解像度動画像とのインタラクションを可能にするブラウザの作成
- (2) ユーザの操作と動画像特微量との関連付けによる知識抽出
- (3) 抽出された知識と動画像特微量に基づく動画像解析をサブゴールとして研究を進める。

3. 研究の方法

研究代表者はこれまでに、動画像に対して人間がどのようにインタラクションを行うかを、視線追跡装置を利用するなどして調査してきており、本研究課題では、時空間高解像度動画像を変更することによる動画像とのインタラクション方法を探り、リッチに記録された動画像というメディアの利用可能性について論じる。

加えて本研究では、ユーザが動画像に対して行った操作履歴を活用し、動画像内容の半自動的な分析を行う。研究代表者はこれまでに、早送りや巻き戻しといった単純な動画閲覧操作と動画像特微量とを関連付けることにより、未知の動画像を自分に合った閲覧スタイルで自動再生する方式を確立した。経験的に獲得される動画閲覧のスタイルを知識として扱い、それを共有・流通の対象として捉えている。本研究課題では、これをさらに発展させ、上述の高解像度動画像に対する概

観と詳細を把握する操作と、動画像特微量とを関連付ける。さらに概観を把握するための操作が行われた箇所や、詳細まで閲覧された箇所を統計的に処理し、多人数が同様の操作をした箇所に意味付けを行うことで、動画像の半自動的な解析を行う。

4. 研究成果

上述した目的(1)(2)(3)に対する成果をそれぞれ以下に述べる。

(1)のブラウザの要件として、動画像の持つ情報の概観と詳細を行き来して理解支援を行うことを挙げた。ここでの概観と詳細とは、空間的な意味においては、対象の表示サイズ(サッカーを例に挙げると、フィールド全体が映るようズームアウトした状態を概観、個々の選手にズームインしている状態が詳細、など)を、時間的な意味においては、再生速度(早送りで動きの流れを見る状態が概観、スローモーションで細かい動きを確認することが詳細)を、概観と詳細を表すものとした。

スポーツ指導経験者1名に対するインフォーマルなインタビューに基づき、早送りや巻き戻しといった一般的なビデオ閲覧環境に加えて、空間的および時間的な概観と詳細を変更する機能を、ブラウザに実装した。一例を図1~図4に示す(図中の白線で示す枠は、実際には表示されない)。図1は、空間的な一部分の時間スケールを変更することで、通常スピードで再生する映像の空間的な一部分を、スローモーションで再生する例である。閲覧者が、通常のスピードで映像を再生し、興味をひかれる、あるいは詳細を詳しく見たいと思う空間的な部分が現れた時点(t_0)で、その部分をマウスクリックすると、クリック位置から一定半径内の領域の時間がスローモーションで再生される。つまりクリック後の時刻(t_1)においては、2つの時点の映像フレームが、ひとつの画面に混在して表示される。一定時間経過後、スローモーションで再生された領域内の時間スケールが、徐々に通常スピードに近づき、映像フレームの時点のズレが解消される。

図2は、図1の機能を発展させたもので、上述のマウスクリックを連続して行うことで、任意形状の領域の時間をずらして表示することを可能にするものである。

図3は、通常再生時に画像の空間解像度を意図的に落とし、やや粗い状態で映像が再生されるものである。図1に述べた、クリックした空間的な部分の時間スケールを変更することに加え、同領域の空間解像度を本来の解像度まで高め、ユーザが映像の詳細を確認できるようにするものである。

図4は、図1の機能に加え、指定された領域の表示倍率を拡大するものである。この

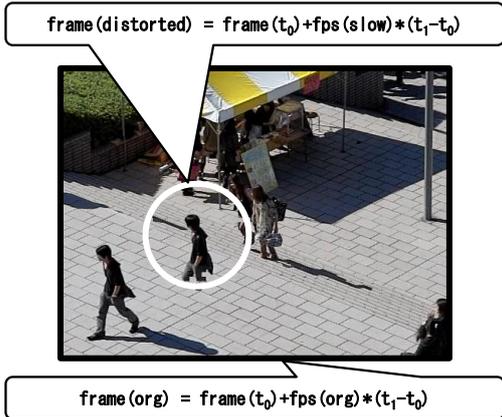


図 1 特定領域の時間スケール変更
(クリック位置から一定半径内の領域)

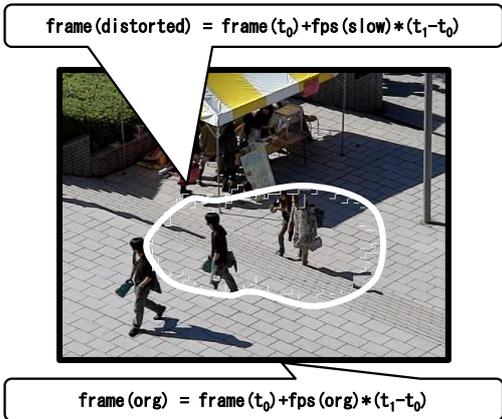


図 2 特定領域の時間スケール変更
(連続クリックによる任意領域)

ため、ユーザが映像領域の一部をクリックすると、その部分がスローで、かつ拡大されて再生されることとなる。これによりユーザが着目する点を詳細まで確認することが可能となる。

上述した4つの機能と、早送りや巻き戻しなどの従来からのビデオ閲覧機能を加えたブラウザを作成した。

(2) の知識抽出においては、ユーザの操作と動画像特徴量との関連付けを行った。ここでの閲覧知識または閲覧スタイルは、あるフレームが表示された際にユーザがどのような行動をとる傾向があるかというマッピングの集合を意味する。ユーザ操作としては、

(1) のブラウザに対して行った早送り、巻き戻し、クリック操作を用い、操作が実施された際に表示されていた映像フレームとセットで記録しておく。スポーツインストラクターなどが、特定の目的のために繰り返し行う特徴的な閲覧方法を、習慣的な行動(habitual behavior)として捉え、動画像の持つ意味内容を含まない特徴量のみで、ユーザ操作との対応付けを行っている。映像フ



図 3 特定領域の時間スケールおよび空間解像度の変更



図 4 特定領域の時間スケールおよび空間表示倍率の変更

レームの画像特徴量としては、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。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が

どの音域に分布しているかを測定している。

マッピングに関しては、動画像の特徴量と、そのフレームがどの速度で再生されたか、また、クリックされたかどうかを関連付けるため、データマイニングツール WEKA による分類器を作成し利用した。閲覧知識を抽出する際には、ユーザに分類器を学習させるためのトレーニングビデオ（同ジャンルの複数の動画）を用い、閲覧時のユーザ操作と画像特徴量とのマッピングを、閲覧知識とした。

(3) においては、(2) で抽出された閲覧知識に基づき、ターゲットとなる動画像に対して、あたかもユーザが操作を行ったかのように、自動で映像を変化させて表示する。ここでは、ターゲットビデオの動画特徴量を解析し、その結果を分類器にかけることで、あるフレームに対するユーザが実施しそうな操作を割り当てる。連続する映像フレームに対して同様のユーザ操作（たとえば早送りや、映像の一部に現れる人物の動きをスローモーションにするなど）が割り当てられた場合、それらのフレームに対して、ユーザが操作を行うのと同様の変化を加えて表示を行う。これにより、習慣的な動画閲覧行動を持つユーザは、初見の動画像に対しても、普段と同様の映像変化が自動で行われる映像を閲覧することができる。

(1) (2) (3) で述べた機能を持つシステムを、スポーツ指導経験者 1 名に利用してもらい、ユーザ観察を実施した。トレーニングビデオ、ターゲットビデオともに、サッカーの映像を題材とした。

得られた結果として、早送りや、巻き戻し後のスロー再生など、画像全体に対する時間操作に対しては、ユーザの閲覧知識を抽出し、ターゲットビデオをユーザ操作に近い操作で表示することができた。一方で、空間の一部分に対する時間操作や拡大率変更などについては、十分にユーザの操作を再現できていない。また、今回このような閲覧知識を再利用する対象ユーザとして、スポーツ指導経験者を選択したが、ユーザ観察の際に十分な人数を確保できていない。複数人が同様に行う閲覧方法の再現に向けて、さらなるユーザ観察実験を検討したい。

汎用的に利用可能な閲覧方法を提示することは十分ではないが、特定の人物の習慣的な行動を限定的に再現することが可能となった。本研究の取り組みから、動画記録方法が現在よりも空間的・時間的に高解像度になった際のインタラクションを、知識として抽出し流用する可能性を示すことができた。

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)

[図書] (計 1 件)

- ① Akio Takashima, Enhancing Video Viewing Experience, A Chapter in Interdisciplinary Advances in Adaptive and Intelligent Assistant Systems: Concepts, Techniques, Applications, and Use, IGI Global, 124-152, 2010

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