

PC 上でのユーザの行動履歴に基づく映像検索システムの検討

寺中 晶郁 長田 秀信 森本 正志

日本電信電話株式会社 NTT サイバーソリューション研究所

1. はじめに

近年、ブロードバンドの普及、ストレージの大容量化、デジタルカメラや携帯電話等のデジタル撮影機器の普及により、映像配信サービス、映像ブログ、動画共有サイトをはじめとするサービスが広く利用され、個人が PC 上で映像を扱う機会が増加している。これに伴い、大量の映像中から所望の映像を効率的に検索することが重要な課題となっている。

我々はこれまでに、カットやカメラワークに代表される映像中の物理的な特徴区間を用いて、PC 内に保存されている映像ファイルをブラウジングにより効率的に検索するシステム[1]や、テキストによる内容検索に必要なメタデータの入力支援[2]について検討してきた。本稿では、映像を検索する場面として、個人が過去に参照した映像を効率よく探す場面を想定し、映像中の物理的な特徴区間の他に個人が映像を参照する際に取った行動履歴を検索に用いる方法を提案するとともに、視聴行動の取得方法について検討を行う。

2. 関連研究

過去に参照した情報を再度参照する場面を想定し、PC 上で扱われるファイルや、Web ページの効率的な検索方法を検討している例がある[3][4][5]。これらの研究では、PC 上のファイルや Web ページの参照時刻及び参照時刻前後のユーザの操作履歴を活用して検索を行う。時刻情報や操作履歴を用いることで、参照したいものの自身についてのキーワードがわからなくても検索が可能となる。例えば、探しているファイルのファイル名以外に、参照した時刻や参照前後の操作履歴から求めるファイルを効率的に見つけることが可能となる。

しかし、これらの研究は、参照時の情報を用いている一方で、ファイル、Web ページの管理を目的としており、検索結果は、ファイルやページの概要とファイルアイコンまたはサムネイルで表示される。検索結果をブラウジングする際、文書の場合は概要を、画像の場合はサムネイルを頼りに探すことが可能である。しかし、映像の場合、テキストでの内容記述は期待できない。また、映像は時間的な長さを持っていることからファイルアイコンや Web ページのサムネイル 1 枚からその内容を判断することは困難である。

3. 提案手法

本章では、本稿で提案する、行動履歴を用いた映像検索システムについて述べる。本検索システムは、参照時刻や参照した映像のファイル名等、映像参照時の情報を検索キーとして絞込みを行い、その結果をブラウジングし求める映像を検索する。

3.1. 検索に使用する履歴

検索に使用する履歴情報として、表 1 に示す 4 種類の情報を用いる。

表 1 検索に使用する履歴情報

	例 インターネット上にある映像ファイルをダウンロードした場合
映像自身の情報	movie.mpg
参照時刻	2006/12/01 12:23:45
参照場所	http://website.com
参照方法	ダウンロード

4 種類の情報の中で、参照方法とはユーザの能動的な行動のことであり、映像の内容と関連付いてユーザの記憶に残る可能性が高いと考えられる。参照方法としては、下記の様なものが挙げられる。

- 視聴
- 編集
- PC - ネットワーク間でのファイル操作
(ダウンロード、アップロード)
- ローカル PC 上でのファイル操作(コピー・移動)

ユーザが映像を再度参照する理由として、映像内容への興味が増えられる。参照方法の中で映像の内容理解と最も関連が深いのは視聴であることから、視聴区間の情報は再参照したい映像を検索する際に最も有効であると考えられる。視聴履歴の中で映像再生時刻及びファイル名は再生プレイヤーのログ等から取得可能である。しかし、現在 PC 上で視聴される映像のフォーマット及び再生プレイヤーは多岐にわたっており取得は容易ではない。また、視聴した映像に関する情報としては、ファイル名よりも実際に目にしていた画像の方が記憶に残りやすい。そこで、ブラウジング時にユーザの記憶を支援することが可能な視聴区間の画像情報を汎用的に取得する。取得方法については、4 章で述べる。

3.2. 検索結果の表示方法

本研究で提案する映像検索システムでは、ブラウジング時の内容理解を支援するため、検索結果を図 1 の様に提示する。中心に映像のサムネイルを表示し、その両側には、映像参照前後のアクティブウィンドウ(前面に表示されているウィンドウ)のキャプチャー画像を表示する。この様に、映像参照時の様子を、その前後の行動も含めて見せることによって、映像の内容に関する記憶が想起され、検索結果の一覧から効率よく目的の映像を探し出せる。なお、映像のサムネイルは、ユーザが参照を行った区間(例:視聴区間、編集区間)を優先的に表示する。



図 1 検索結果表示の例

4. 視聴区間の抽出

本章では、3.1 で述べた視聴区間の推定方法について述べる。ここでは、ユーザが能動的に映像を選択し再生する場合を想定し、アクティブウィンドウ内で映像が表示・再生されている場合に視聴しているとみなした。

4.1. 視聴区間の推定方法

映像が表示・再生されている区間を推定するために、画面情報に着目した。図 2 は一定間隔 (Δt) で取得されたアクティブウィンドウのキャプチャーにおいて、時刻 t におけるキャプチャー画像と、 $t + \Delta t$ におけるキャプチャー画像の各ピクセルの色差分を求め、色差分のあったものを動きあり、なかったものを動きなしとした 2 値化画像の例である。

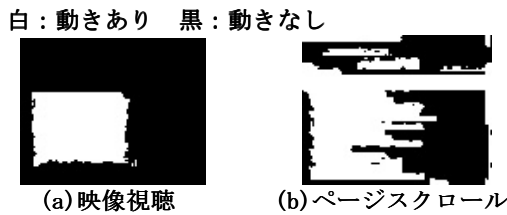


図 2 アクティブウィンドウの色差分 2 値化画像

図 2 から、映像視聴時では動きのある部分が一箇所 (図 2(a) の白い矩形部分) に集中し、映像視聴以外では動きのある部分が分散する傾向が定性的に認められる。これは、ユーザが映像視聴時にアクティブウィンドウをあまり動かさないという行動特性によるものと考えられる。

上記の傾向に基づき、視聴区間を自動推定するため、図 2 で示した 2 値化画像における動きの有無の集中度合いを表す次の指標 $C_m(t)$ を定義する。

$$C_m(t) = \sum_i \sum_j \frac{v_m(i, j, t)}{N(m, t)}$$

ただし、 m は動きの有無を表す変数 (1 または 0)、 t は時刻、 $N(m, t)$ は時刻 t における 2 値化画像において、ピクセルの値が m であるピクセル数である。また、 $v_m(i, j, t)$ は、時刻 t における 2 値化画像の (i, j) の位置にあるピクセルの値 $P(i, j, t)$ と、このピクセルに隣接する周囲のピクセルの値との差分を表す指標であり、次式で定義される。

$$v_m(i, j, t) = \begin{cases} 1 & \text{if } P(i, j, t) = m \text{ and } \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} |P(k, l, t) - P(i, j, t)| \leq Th \\ 0 & \text{otherwise} \end{cases}$$

ただし、 Th は閾値である。

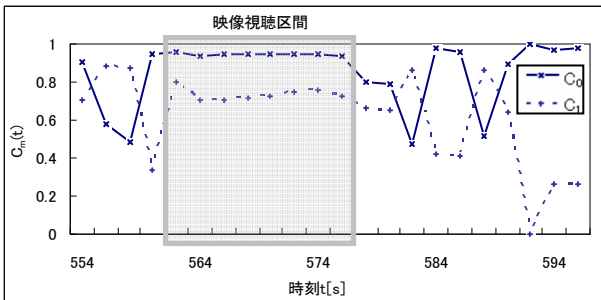


図 3 C_m の時系列変化 (抜粋)

上記の指標 $C_m(t)$ に関する予備実験を行った。実験データには、約 30 分間のアクティブウィンドウのキャプチャー画像を用いた。

結果を抜粋したものを図 3 に示す。図 3 より、映像視聴区間では $C_m(t)$ の値が一定以上の大きさを持ち、ほぼ変化しないことが認められる。本指標によって、視聴区間の推定を行う。

4.2. 実験及び結果

4.1 で示した指標を用いて視聴区間抽出の実験を行った。実験データは、約 6 時間半分のアクティブウィンドウのキャプチャー画像を用いた。なお、画像は 2 秒間隔で取得した。主に行っていた行動は、Web 閲覧 (映像の視聴を含む)、メール閲覧・作成、ファイル操作 (検索、コピー、移動)、表計算ソフトの操作、プレゼンテーション作成ソフトの操作であった。

パラメータ及び判定条件は、予備実験より下記のように設定した。

- $Th=1$
- $C_m(t) > 0.7 \quad (m=0,1)$
- $|C_m(t+1) - C_m(t)| < 0.2 \quad (m=0,1)$

結果は、再現率:0.74、適合率:0.11 となり、適合率の向上のためさらなる絞込みが必要であることがわかった。映像視聴と間違えられたものには、分割された画面の一部をスクロールしているもの、ページ切り替えを頻繁に行っているもの等、ページの一部分が動くまたは全画面が一度に変更されるものが多く見られた。これらは、キー・マウス操作やページの切り替わりの情報を併用することで省くことができると考えられる。また、見つけることのできなかった映像は動きの少ない映像であった。これらは、画像情報のみでは、動きがないため推定が困難であるが、映像視聴時は操作をあまり行わないユーザの行動特性から、キー・マウス操作情報を活用することにより再現率を改善できると考えられる。

5. おわりに

本稿では、過去にユーザが参照した映像を検索する方法について提案し、過去の参照行動のうち最も映像の内容と関連があると考えられる視聴区間の推定方法を検討した。動きの有無に基づく 2 値化画像を解析し、視聴区間の推定に有効な指標を求めることができた。

今後の予定としては、視聴履歴取得の精度向上及び、取得した履歴を検索時に用いた場合の検索効率の向上について検討を進めていく。

参考文献

- [1] Nagata et al., "VIDEOPOT: INDEXING-BASED DESKTOP VIDEO SEARCH SYSTEM," In Proc. ICME, pp.373-376, 2006.
- [2] Kuwano et al., "SceneCabinet/Live!: Real-Time Generation of Semantic Metadata combining Media Analysis and User Interface Technologies," Proc. IBC2005, pp.253-260, 2005.
- [3] 大澤 他, "俺デスク: ユーザ操作に基づく参照履歴検索ソフトウェア," インタラクシオン 2006, pp.219-220, 2006.
- [4] Morita et al., "Desktop search system based on the action-oriented algorithm," In Proc. APSITT, pp.204-207, 2005.
- [5] Chirita et al., "Activity Based Metadata for Semantic Desktop Search," In Proc. ESWC, pp.439-454, 2005.