

ユーザ提示画像を基にしたクリックابل動画自動生成システムの設計と実装

三樹一貴[†] 佐藤喬[†] 多田好克[†]

[†]電気通信大学 大学院情報システム学研究科

1 はじめに

インターネットの発展に伴い、今日では個人での Web ページ、ブログの開設や閲覧、SNS の利用が広く普及している。最近の傾向として Youtube、google video といったインターネット上の動画投稿サービスが積極的に利用されるようになってきており、個人の Web ページやブログ、SNS でも動画の掲載が可能となっている。

このような状況の下、動画視聴者が動画中で興味を持った対象をクリックした際、即座に関連性のある詳細な情報に当たれば有益である。確かに現在でも、特定箇所をクリックすることにより視聴者をリンク先に導く Web 上の動画もいくつかは存在する。しかし、それらは動画のフレームごとに人手で位置を指定し、リンク先をタグ付けしているのが現状である。一般の人々がそのような事をするには、ある程度の計算機スキルが必要であり、何より繁雑で面倒である。

そこで本研究では、Web 上において、動画投稿者の計算機スキルに関わらず、クリックابل動画を自動で生成するシステムの設計と実装を行う。ここでいうクリックابل動画とは、動画中の任意の時点、位置にクリックابلオブジェクトを含む動画のことである。

また、視聴者がその動画中に表れたクリックابلオブジェクトを容易に認識、クリック可能なインタフェースの設計と実装も同時に行う。

2 関連技術

BML(Broadcast Markup Language) はデータ放送用記述言語で、デジタル放送のコンテンツ製作で用いられる。この技術を用いれば、視聴者をコンテンツ製作者の意図したリンク先へ導くことが可能である。しかしあくまで技術仕様であり、一般の人々がコンテンツを製作するとは考えにくい。また、フルセグデータ放送が視聴可能な計算機は今現在、広く普及している

“A Design and Implementation of Automatic Generating System of Clickable Movie using an Image and Information Offered by a User”

Kazutaka Mitsugi[†], Takashi Satou[†], Yoshikatsu Tada[†]

[†]Graduate School of Information Systems, The University of Electro-Communications

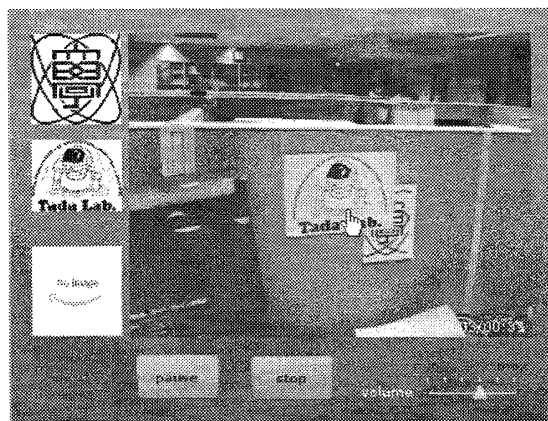


図 1: クリックابل動画インタフェース

とはいえない。

3 概要

自身の Web ページなどにクリックابل動画インタフェースを掲載したい場合、動画とイメージ、リンク先の URL を、Web アプリケーションとして実装された本システムに投稿する。すると、本システムは動画中に表れるイメージをクリックابلオブジェクトとしたクリックابل動画を生成し、タグを出力する。投稿者が出力されたタグを Web ページなどに貼り付ければ、クリックابل動画インタフェースが掲載される(図 1)。なお、本稿ではイメージの例としてロゴや広告看板、関連情報として URL を使用する。

4 設計

この章では、システムの設計を投稿者サイド、視聴者サイドのそれぞれに分けて記述する。また、本システムは、Web ブラウザからインターネットを介して利用可能な Web アプリケーションとして設計した。

4.1 投稿者サイド

この節では、システムの処理の流れについて記述する。ここで、投稿者が投稿する動画ファイル形式及び画像ファイル形式は問わない。また、アップロードするイメージの総数について内部的な制限は設けていないが、インタフェースの関係から 3 つまでとする。

システムは、以下に記述するように段階的に処理を

行う。step3において、認識対象が見つかった場合はstep2に戻り、そのフレーム前後のキャプチャ時間間隔を短くして再度キャプチャを行う。

step1. 対象動画及びイメージ、関連情報の收受

step2. 対象動画のキャプチャ

step3. キャプチャ画像解析

step4. 解析ファイル生成

step5. ブラウザヘタグを表示

step3ではSIFT[1]を利用して、ユーザ提示画像とキャプチャ画像の対応に用いる特徴量を抽出する。SIFTを選択した理由は、一枚の画像であっても、ある程度の視点変化や隠れに対処できるためである。

4.2 視聴者サイド

インタフェースは画面、アイコン表示部、ボタン、音量調節部によって構成する(図1)。

動画中において、ユーザ提示画像が認識された時点でアイコン表示部にアイコンを一定時間表示し、動画中のクリッカブルオブジェクトを視聴者に知らせる。その際、動画中のクリッカブルオブジェクト近傍と共に、アイコンもクリック可能とする。画面の大きさとの兼ね合いで、アイコン数の上限は3つとする。

5 実装

この章では、システムの実装について述べる。実装環境はLinux(Fedora Core 5)、Webサーバ(Apache2.2.2)である。

5.1 投稿者サイド

PHP5.1.6、及びC言語を用いて実装を行った。

step1. 対象動画及びイメージ、関連情報の收受

投稿者はWebブラウザを通じて本システムに対象動画、画像、関連情報をアップロードする。アップロードされた日時を基にシステムはディレクトリを作成し、以後の処理を同ディレクトリ内で進める。

step2. 対象動画のキャプチャ

まず、対象動画から単位時間ごとにキャプチャを行う。step3でユーザ提示画像をキャプチャ画像内に認識したら、その前後の時間刻みを $\frac{1}{2}$ ずつ細かくしていき、認識対象を更に追う。一般の人々が撮影する動画に激しいカメラワークが伴うとは考えにくいいため、終了条件を時間間隔0.2秒以下とする。

step3. キャプチャ画像解析

ユーザ提示画像とキャプチャ画像の特徴量[1]を比較することにより対応点を得る。その際、ロバスト推定

を用いることにより誤対応の除去を行う。更に、認識対象が見つかるごとに、キャプチャ画像中の認識領域の特徴量を取り除くことで、同画面中に同じ認識対象が複数存在しても、それぞれの位置が得られる。

step4. 解析ファイル生成

キャプチャ画像解析に基づき、動画中のいつ、どの範囲にユーザ提示画像が出現したのかを記述したxml形式の解析ファイルを生成する。

step5. ブラウザヘタグを表示

システムでの一連の処理が終わると、投稿者のブラウザのアップロード画面を、解析ファイルを含んだタグ表示画面に変える。

5.2 視聴者サイド

ActionScript3.0を用いて実装を行った。

画面では、マウスカーソルがクリック可能領域に入ったら、ハンドカーソルに変化させる。クリック可能領域は、クリッカブルオブジェクトの真上だけでなく、その周りに画面の大きさに応じてあそびの部分をつけ加えることで、視聴者が対象を容易にクリックできるようにした。クリックイベントを捕まえたらリンク先に導き、その間動画は一時停止状態とする。もし、クリッカブルオブジェクトが一部重なっている箇所をクリックした場合には面積の小さい方を優先させる。また、画面右下には経過時間と動画の総時間を表示する。

音量調節部はスライダーで0~10まで操作でき、デフォルトは6.5とする。

アイコン表示部では、画面にクリッカブルオブジェクトが表れた時点で、デフォルトアイコンから、ユーザ提示画像を加工したアイコンに変化させ、一定時間表示する。アイコンもクリック可能であり、クリックされた場合は、リンク先へと視聴者を導く。アイコンは各々がタイマを持ち、一定時間表示された後にデフォルトアイコンに戻る。一時停止ボタンが押されたり、視聴者がリンク先のWebページに飛んだ場合は、アイコンのタイマも一時停止状態とする。

動画はplay(pause)、stopのボタンで操作する。

6 おわりに

本稿では、計算機スキルに関わらず、クリッカブル動画を自動生成できるシステムと、その視聴インタフェースについての設計と実装を述べた。今後の課題として、画像アルゴリズム処理の並列計算などによる処理時間の高速化が挙げられる。

参考文献

- [1] D.Lowe: "Distinctive Image Features from Scale-Invariant Keypoints," Int.Journal of Computer Vision, Vol.60, No.2, pp.91-110,2004.