

時間及びフレーム内座標情報を利用した(2+1)次元映像アノテーションシステムの研究

青木 哲郎† 兼子 正勝†  
 電気通信大学情報理工学研究科総合情報学専攻†

1 はじめに

Web 上の動画コンテンツ数の増加に伴い、動画メタデータの活用が重要になってきている。動画メタデータの有効な活用法を検討することで、動画検索の精度向上や、動画の利用形態の多様化などが期待できる。

より精細な動画メタデータを取得するため、Web 上で配信する動画の時間軸やフレーム内座標を対象としたコメントアノテーションを利用することが考えられている。動画のフレーム内座標を対象としたアノテーションシステムとして、視聴者が動画シーン中の任意の矩形範囲を対象に自由にアノテーションを付与できる長尾らの「Synvie」[1][2]や、動画の投稿者が動画中の任意のフレーム中の1点にアノテーションを付与できる「YouTube」[3]が挙げられる。しかし動画中の1点や矩形範囲を対象としたアノテーションでは扱える内容がオブジェクトの説明などに限定される。

そこでフレーム中の2点間を結ぶ線分を対象としたアノテーションを提案する。このようなアノテーションを導入することで、動画内のオブジェクト同士の関係、オブジェクトの運動、カメラワークや構図といった高度な内容も扱うことが可能になり、動画に関する情報をより精細にすることができると考えられる。

本研究では Web 上で配信する動画に対し、フレーム内の任意の「線分」を対象としたコメントアノテーションを付与し活用するためのシステムとして「(2+1)次元アノテーションシステム」を構築する。「(2+1)次元」とは線分がフレーム内の座標情報(2次元)と動画内の時間軸情報(1次元)を併せ持っていることを意味する。

2 システム概要

ユーザーは Web ブラウザ上で動画を視聴し、コメントの投稿や閲覧を行う。システムにはコメント投稿モードとコメント閲覧モードの2つのモードが存在し、モードの切り替えボタンによってモードを切り替えることができる。

コメントの投稿はコメント投稿モードで行う。投稿モード中に動画フレーム上をドラッグすることでコメントを付与する線分を選択できる。この状態で

コメントを入力フォームに記述し、投稿ボタンを押すと選択した座標と選択時の動画タイムコードを対象にコメントが投稿される。また、プレイヤーに備えられたベクターボタンを押すことで、線分に方向性を持たせ矢印にすることもできる。線分・矢印の導入によりオブジェクト同士の関係や運動、カメラワークなどを扱ったアノテーションが可能になる。

投稿されたコメントの閲覧はコメント閲覧モードで行う。動画の再生時に、時間軸に合わせてコメントが付与された線分・矢印が動画上に重なるように表示され、カットの終了時点まで表示され続ける。表示された線分・矢印をクリックすることで、その線分に投稿されたコメントの一覧がコメント表示部に表示される。また、累積表示ボタンを押すことで指定した時間範囲内のコメント付与線分を全て表示することができる。これによってユーザーの興味の推移などを可視化できる。



図 2.1 システム UI

3 システム実装

HTML5 の video タグで動画を配信し、JavaScript によって動画の制御・タイムコード及び座標の取得、サーバーへのコメント送信などの処理を行う。投稿されたコメントとそれに付随する情報はサーバー側にて PHP で処理し MySQL のデータベースに保管する。

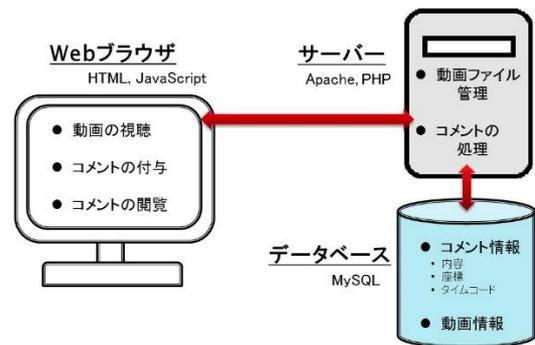


図 3.1 システム図

(2 + 1) Dimensional Video Annotation System using coordinates in frame and time code  
 Tetsuro Aoki†, Masakatsu Kaneko†  
 †Dept. of Informatics, Faculty of Informatics and Engineering, the University of Electro-Communications

### (1) 動画の配信

動画視聴ページは PHP ファイルで記述する。ユーザーからのリクエストに応じて video タグを生成することで、サーバー上に用意した mp4 ファイルを HTML 中に埋め込み動画を配信する。動画プレイヤーとしての機能は JavaScript によって実装した。

### (2) 選択座標・タイムコードの取得

JavaScript によって画面上でマウスを押し込んだときの点と離れたときの点のフレーム内座標を取得する。また、マウスの押下時にタイムコードを取得する。

### (3) コメントの投稿

コメント内容、選択座標、タイムコード、方向性の有無(矢印か否か)を Ajax によって非同期にサーバーへ送信する。

### (4) コメント付与座標の整理

各ユーザーが任意の座標を対象にコメントを付与していくと、同一の対象を扱っているコメントが複数の点や線分に分散されてしまう。そこで、コメントの投稿時にサーバサイドで以下のように「グリッド座標」を算出し、データベースに保管する。

- i) フレームをグリッド状に分割したと仮定
- ii) グリッド上の各交点に座標番号を与え「グリッド座標」とする
- iii) 付与座標に最も近いグリッドの交点を算出
- iv) 算出した交点のグリッド座標をそのコメントのグリッド座標とする

上記の処理により各線分はグリッド交点上にまとめられる。なお、グリッドによる分割の細かさは動画ごとに最適と思われる値を設定し、動画情報を保管するデータベースに保管しておく。

### (5) 閲覧モードでの線分・矢印表示

動画視聴ページの読み込み時に、リクエストされた動画に対するコメントの情報をデータベースから参照し、JavaScript の連想配列に格納しておく。閲覧モード時に、video 要素の timeupdate イベントが発生するたびにこの連想配列を参照し、次の条件を同時に満たすコメントを探す。

- 再生位置がコメントタイムコードを超過
- 現在のカットとコメントのカットと一致
- まだ描画されていない

上記条件を満たすコメントのグリッド座標を参照し、グリッド座標からフレーム上での座標を算出してその位置に直線・矢印を描画する。描画はコメントごとに canvas 要素を生成して行う。

### (6) コメント内容の閲覧

(5)の処理で画面上に複数の canvas 要素が重ねられているため、以下のようなアルゴリズムで「クリックした線分・矢印に付与されたコメントの一覧を表示する」処理を実現した。

- i) クリックした点に重ねられている要素を全て検出する。
- ii) その中に線分描画のための canvas 要素があるかを上から順に調べる。
- iii) あった場合、クリックされた点が canvas 要素内の線分が描画されている領域であるかを調べる。  
なかった場合、処理を終了する。
- iv) 描画されている領域であった場合、その canvas 要素に対応するコメントの一覧を表示する。

描画されている領域出なかった場合、下に重ねられている canvas 要素に対し 2)~4)の処理を行う。

## 4 運用実験

システムの構築後、20代の男女8人を対象に約1週間の運用実験を行った。運用実験の目的は、

- システムが正常に動作するかを確認すること
- システムやユーザーインターフェースについてのユーザーの反応を見ること

の2つである。また、実験の参加者にシステムに関するアンケートに回答してもらった。

運用実験の結果、システムは正常に動作することが確かめられた。

アンケートの結果、「線分や矢印によるコメントを面白いと思うか」という質問に対し「とても思う」と回答した人が7人、「どちらでもない」と回答した人が1人、「線分や矢印によってコメントの幅が広がると思うか」という質問に対し、「とても思う」と回答した人が5人、「そう思う」と回答した人が3人といずれも高評価であった。そのため、実運用の価値はあると考えられる。

収集したコメントの中にはオブジェクト同士の関係や運動を話題としたものも見られた。一方で構図やカメラワークに関するコメントは得られなかった。このようなコメントを収集するためには映像の研究者などに利用してもらう必要があると考えられる。

## 5 まとめ

本研究では動画フレーム内座標を対象とした、高度な内容を扱った映像アノテーションを可能にするために、線分形式でのアノテーションが可能な「(2+1)次元アノテーションシステム」を構築し、線分によるアノテーションの有用性を確認した。

## 6 参考文献

- [1] Project Synvie. <http://synvie.net/>
- [2] 山本 大介, 増田 智樹, 大平 茂輝, 長尾 確: "映像を話題としたコミュニティ活動支援に基づくアノテーションシステム", 情報処理学会論文誌, Vol.48, No.12, pp.3624-3636(2007).
- [3] YouTube. <http://www.youtube.com/>