

インデックスの重み補正のための位置情報取得

Location Detection for Compensating Significance of Indexes

佐藤 史子 寺口 正義 益満 健 越後 富夫†
Fumiko Satoh Masayoshi Teraguchi Ken Masumitsu Tomio Echigo

1. はじめに

近年の FTTH や ADSL の普及など、ネットワーク環境の高速化に伴い、インターネットにおけるビデオコンテンツの有効な利用法が求められている。その一つとして、オリジナルのビデオコンテンツから、各ユーザーの嗜好に合わせたダイジェストの生成・配信が挙げられる。

すでに開発した MPEG-7 オーサリングシステム[1] (以下、システムとする) では、ユーザーの嗜好を反映したプロファイルに基づいてダイジェストを生成・配信するため、コンテンツに内容を表すインデックスを付加する。例えば、サッカーなどスポーツビデオの場合、“シュート”、“ゴール”などのイベントに対し、イベントの発生時刻を手入力すると、そのイベントの開始時間を中心とする重み分布関数を持つインデックスを自動生成する。以下本研究では、対象としてサッカーのビデオコンテンツを用いる。

現在のシステムでは、各インデックスが持つ重み分布関数はガウス関数で表される。そこに含まれるパラメータは、例えば、“パス” イベントよりも“ゴール” イベントの方が重要度が高い、といった一般的な認識のもとに、イベントごとにあらかじめ一意に決められている。したがって、同一イベントであれば、イベントが発生した状況や時刻に関わらず同じ重み分布関数を持つ。

しかし、実際にはインデックスの重みはイベントによって一意に決定されるものではなく、試合の流れやプレー位置によって変化するものである。例えば、「前半の追加点ゴール」と「後半終了間際の逆転ゴール」を比較した場合、同じ“ゴール” イベントであっても、試合の流れを考慮すると、後者のイベントの方が重要であると考えられる。現在のシステムでは、このような試合の流れを反映した重みを持つインデックスを付加することができなかった。

より正確なダイジェストを作成するためには、試合の流れを反映して各インデックスの重みを変化させることが必要である。試合の流れを状態遷移としてとらえると、あるイベントによって状態が遷移する場合の重要度が高くなることを考えることができる。また、相手チームのゴール付近でのイベントは重要度が高く、ゴールから遠ざかるにつれて重要度は低くなるということもできる。従って、インデックスの重みを変化させるための指標の一つとして、プレー位置を取得する必要があると考えた。

しかし、プレー位置をインデックスと同時に手入力するとインデックス付加にかかる労力が増大してしまう。そこで本研究では、インデックス付加の労力を上げることなく試合の流れを反映したインデックスを作成することを目的として、プレー位置をフレーム画像から自動取得する方法を考察する。また、実際にサッカーのビデオコンテンツを

用いて実験を行い、本方法による位置情報取得の性能を評価した。

2. 位置情報取得

2.1 位置情報取得条件

本研究での位置情報の自動取得は、最終的にシステムに組み込んで利用することを前提としている。したがって、位置情報の取得方法では、以下の点を重視する必要がある。

まず、インデックス付加の労力を上げないようにするためには、位置情報を取得してもシステムの実行速度に大きな影響がないようにする必要がある。しかし、実際に要求される実行時間はシステムへの組み込み方に大きく依存すると考えられるため、まず今回は、位置情報取得をコンテンツの実時間程度で行うことを目標とした。

また、すべてのフレームに関して位置情報を取得するのは非効率であるため、試合の流れがつかめる程度の時間毎に位置情報を取得する。今回は 1 秒毎に取得することにした。

上記事項を考慮して、取得する位置情報を定義する。プレーの正確な位置情報を 1 フレーム画像に映っている情報だけから求めるには、計算量・誤りが多くなると考えられる。また、今回は実行速度が重要であり、試合の流れが得られればよいので、プレー位置の詳細を取得する必要はない。従って位置情報は、フレームに映っているフィールド位置として取得する。フィールド位置 ID (以下、位置 ID とする) は、フレーム画像中にセンターラインがどこに映っているかを基準に 5 種類定義した。画像の横方向を x 軸、 $width$ をフレーム画像の幅として、画像中のセンターラインの x 座標 x_c の値によって分類する。

位置 ID の定義を図 1. および表 1. に示す。

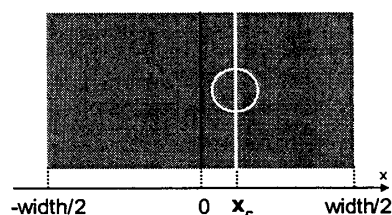


図 1. フレーム画像の座標定義

位置 ID	定義
Location 1	$-\text{width}/4 < x_c < \text{width}/4$
Location 2	$x_c < -\text{width}/4$
Location 3	$\text{width}/4 < x_c$
Location 4	Location2 より左側
Location 5	Location3 より右側

表 1. 位置 ID 定義

† (株) 日本 IBM 東京基礎研究所
〒242-8502 神奈川県大和市下鶴間 1623 番 14 号

2.2 位置情報取得方法

位置情報取得は、フレーム画像からフィールドの白線を抽出し、そのパラメータから位置 ID を取得することで行う。白線抽出には Hough 変換を用いるが、実行速度を重視し、[2]のアルゴリズムを用いた。また、白線の端点を取得すると多くの処理が必要となる上、全フレームで安定した結果を得るのは難しい。従って、白線のパラメータとしてはフレーム画像中の位置 (ρ) と傾き (θ) のみを用いた。Hough 変換により、1 枚のフレーム画像から複数の白線が抽出されるが、各白線の優先度も考慮した。今回は、ある閾値以上の優先度をもつ白線のうち、最も優先度の高い白線のパラメータを利用した。

ただし、フレーム画像に対しそのまま Hough 変換を行うと、フィールドと観客席の間や、会場の柱など無関係な情報を抽出してしまうことが少なくない。そこで、フレーム画像中のフィールド部分に対してのみ Hough 変換を行うために、処理するフレーム画像ごとにフィールド部分を表すマスクを色情報から作成し、マスク内の画像に対して Hough 変換を行った。

このように取得した白線のパラメータ (ρ , θ) を用いて、処理したフレームの位置 ID を取得する。あらかじめサンプルデータから位置 ID が Ln の場合のパラメータ (ρ (Ln), θ (Ln)) を求めておき、(ρ (Ln), θ (Ln)) と各フレームのパラメータ (ρ , θ) との比較により位置 ID を取得する。

全処理フレーム画像を処理した後、位置 ID が取得できなかったフレームや誤認識したフレームに対応するため、フレームの前後関係から各フレームの位置 ID を修正する。例えば、前後のフレームで位置 ID が大きく離れている場合や、あるフレームの前後に同じ位置 ID が続くにも関わらず、そのフレームだけ異なる位置 ID を持つ場合などはその位置 ID を強制的に修正する。

図 2. に位置情報取得の流れを示す。

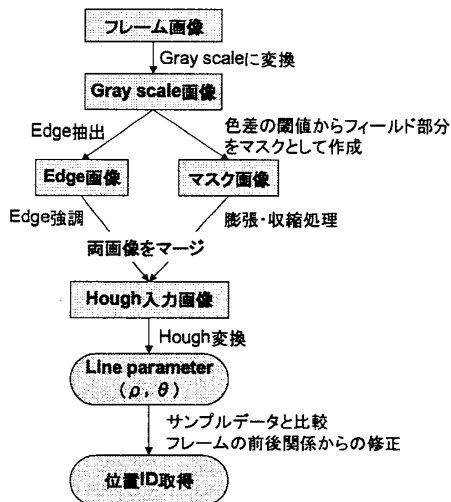


図 2. 1 枚のフレーム画像からの位置情報取得の流れ

3. 実験結果

3.1 実験概要

上記の方法を用いて、位置情報取得の性能について評価を行った。実験用コンテンツには、サッカーの試合の前半

部分を用いた。コンテンツは 352×240pixels の MPEG-4 ファイルで、フレーム数は 84030frame、フレーム率は 29.97fps、再生時間は 46.72min である。実験用マシンには、CPU Pentium4 2GHz、メモリ 768MB の Windows 2000 Professional OS を用いた。

本実験では、前半開始から前半終了までの全フレームに対し、1 秒 (30frame) 毎にフレーム画像から位置 ID を決定した。ただし、プレーヤーが大きく映っているフレームなど本方法で位置 ID が決定できないフレームに関しては除くものとした。

3.2 実験結果

本実験結果を表 2. に示す。

全フレーム数	84030 frame
処理対象フレーム数	2801 frame
処理フレーム数	1912 frame
カットフレーム数	889 frame
正解率	94.5 %
位置情報取得フレーム数	588 frame
位置情報取得精度	95.2 %
実行時間	17.3 min

表 2. 実験結果

ここで、処理対象フレーム数とは、コンテンツ 1 秒ごとに処理したフレームの数であり、処理フレームとは、処理対象フレームのうち、前述したような位置 ID が決定できないフレーム (カットフレーム) を除いたものである。正解率は、全処理フレームに対し正しく位置 ID を取得した確率を表し、位置情報取得精度は、インデックスが付加されたフレーム (位置情報取得フレーム) に対する正解率を表す。最終的には位置情報はインデックスが付加されているフレームのみが必要とするので、位置情報取得精度が最も重要である。

各処理フレームの正解の位置 ID は、あらかじめ目視で決定しておき、自動取得した位置 ID との比較により正解率・位置情報取得精度を求めた。表 2. のように正解率は 94.5%、位置情報取得精度は 95.2% を得た。実行時間は 17.3min であり、実行時間の目標は十分に達成できた。

5. まとめ

本位置情報取得方法では、コンテンツ 1 秒ごとにフレーム画像から位置 ID を取得した場合、コンテンツの実時間以内で実行でき、位置情報取得精度も 95% 以上を達成した。この実験結果より、短時間の簡単な処理でも高い精度で位置 ID を取得できることがわかったため、既存のオーサリングシステムの実行に大きな影響を及ぼさず位置情報の取得部分を組み込むことができると考えている。

今後は、実際のシステムへの組み込み方を検討するとともに、取得した位置 ID を反映した重みを持つインデックスを生成し、それを元に作られるダイジェストの評価とインデックスの有効性の検証を行っていく。

参考文献

[1] 寺口, 益満, 越後, 関口, 栄藤, “時区間インデックス作成によるパーソナルビデオダイジェスト”, 信学技報, Vol. 101, No. 302, PRMU2001-91, pp. 43-50, 2001-09.

[2] 沼田, 輿水, “Gradient 型超高速 Hough 変換アルゴリズム”, 情処研報, CV51-2, 1987.