

状況変化に基づく内容推定による

3K-7

ビデオ・データベースのインデキシング支援*

広部一弥[†] 牛尼剛聡[†] 渡邊豊英[†]名古屋大学大学院工学研究科[‡]

1 まえがき

我々は、野球中継を対象として、利用者の視点に基づくシーン検索が可能なビデオ・データベース・システム STRIKE を開発中である [1]。STRIKE では、イベント・アクション・モデルに基づいてシーン検索を行う。イベント・アクション・モデルではビデオ中の実体の特性変化と実体間の相互作用をイベントとしてモデル化し、ビデオの内容をイベント系列として表現する。イベントの例としては、「投手がボールを投げる」、「打者がボールを打つ」等が挙げられる。イベント系列を構成するイベントをビデオ・フレームにインデックスとして対応付けることにより、利用者はイベント系列の部分系列を指定することで、シーン検索が可能である。

イベント・アクション・モデルに基づくシーン検索のためには、データベース作成者が一つの野球の試合に対して数千個のイベントをインデックスとして付加する必要がある。インデキシングの労力を軽減するために、我々は STRIKE におけるインデキシング作業を支援する機構を開発中である。本稿では、画像処理により抽出した情報を利用してビデオ中で発生したイベント候補を推定することによりインデキシングを支援する機構を提案する。

イベントは、対象世界のオブジェクトの運動に基づいて発生するため、画像処理を利用してオブジェクトを認識し、イベントを自動的に抽出することが考えられる。しかし、一般に画像処理を用いて抽出可能な情報は抽象化レベルが低いいため、困難である [2]。野球中継では、投手の投球する場面などで試合の状況を表すテロップが出現する。このテロップは、画像処理によって比較的容易に抽出可能である [3]。本稿では、テロップが状況を表すことを利用して、テロップから観測可能な状況の変化から、イベント候補を推定する手法を提案する。なお、本稿ではテロップは正しく認識可能であることを前提とする。

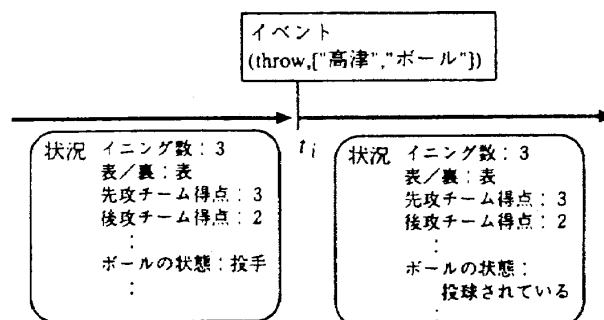


図 1: イベントと状況の変化の例

2 状況とテロップ

2.1 状況

イベント・アクション・モデルでは、ビデオ中の実体の特性の変化に着目するため、任意の時刻における実体の状態を陽に表現していない。そこで、任意の時刻の実体の状態を表現するために状況の概念を導入する。状況は実体間の 2 項関係の集合 $\{(en_{11}, en_{12}), (en_{22}, en_{22}), \dots, (en_{n1}, en_{n2})\}$ である。ここで、 $en_{ij} (1 \leq i \leq n, 1 \leq j \leq 2)$ は実体である。状況 s の要素 $(en_1:en_2)$ に対して、 en_1 を状況 s の属性と呼び、 en_2 を状況 s における属性 en_1 の値と呼ぶ。属性の例としては、「アウト・カウント」、「ボールの状態」等が挙げられる。

状態内の属性の値はイベントによって変化する。図 1 に、状況が変化する例を示す。図 1 では選手「高津」がボールを投げたことを表すイベント $(throw, [\"高津\", \"ボール\"])$ によって、状況の属性「ボールの状態」の値が投手から「投球されている」に変化している。

2.2 テロップ

テロップは、編集者がビデオに付加した記号である。ビデオ中には様々なテロップが表示されるが、本研究では、以下のテロップを対象とする。

- 先攻チームの得点, 後攻チームの得点
- イニング数, 表/裏
- アウト・カウント, ストライク・カウント, ボール・カウント

*Support of Indexing on Video Database Based on Inferring Video Contents

[†]Kazuya HIROBE, Taketoshi USHIAMA and Toyohide WATANABE[‡]Graduate School of Engineering, Nagoya University

- 各塁の走者の有無

テロップはビデオが記録した対象世界の状況を表す。しかし、状況が変化する時刻とテロップが変化する時刻が同一になるとは限らない。テロップが変化するとき、前回のテロップの変化から今回の変化までの時区間で状況が変化している。状況はイベントが発生したときに変化するため、テロップの変化に挟まれた時区間で、イベントが発生していると考えられることができる。

3 イベント候補推定とインデキシング支援

3.1 イベント候補推定

s_1 である時区間で、対象世界上では状況の変化には直接関与していないイベントが発生している場合がある。これは、状況を変化させるイベントの発生が、他のイベントに依存している場合である。したがって、状況を変化させるのは単一のイベントではなく、イベント系列であると考えられることができる。イベント系列 es が状況 s_1 を状況 s_2 に変化させるとき3項 (es, s_1, s_2) を状況遷移パターンと呼ぶ。ここで s_1 を事前状況と呼び、 s_2 を事後状況と呼ぶ。

状況遷移パターンを利用して、テロップから観測される状況の変化をもたらしたイベント候補を推定する。いま、テロップによって観察される状況が s_1 から s_2 に変化したとする。このとき、状況 s_1 を事前条件とする状況パターンの事後状況を s'_2 をとしたとき、以下の手順に従ってイベント候補を推定する。

- 1) s'_2 が s_2 と一致するなら、状況遷移を止める。
- 2) s'_2 が s_2 と一致しないなら、状況 s'_2 を事前条件とする状況遷移パターンについて考える。
- 3) s'_2 よりも時間的に後であり、 s_2 に一致しない状況になった場合、無限に遷移を続けることを防ぐため遷移を止める(この処理には、任意の状況 s_i, s_j 間に時間的な前後関係を定義する必要がある。野球の場合はカウントの順序によってこの関係を決定できる)。

以上の手順で求められた状況の遷移は、状況をノードとする有向グラフで表現できる。この有向グラフでは、状況 s_i から状況 s_j までのパスに存在するイベント系列が状況 s_i, s_j 間で発生したイベント系列であると推定できる。

3.2 インデキシング支援

インデキシング支援機構は次のモジュールから構成される。

- ビデオ表示モジュール
- イベント系列の候補表示モジュール

- イベント-フレーム対応付けモジュール
- イベント候補推定モジュール
- データベースへのイベント出力モジュール
- テロップ抽出のための画像処理モジュール
- 情報入力モジュール

利用者は、情報入力モジュールで状況の属性や値の範囲、状況遷移パターン等の対象ビデオに関する情報を入力する。次に、システムはテロップ抽出のための画像処理モジュールで画像処理を用いてテロップを抽出する。イベント推定モジュールでテロップに基づいてイベント系列の候補を推定し、イベント系列の候補表示モジュールで利用者に提示する。利用者はビデオ中の対応する部分を観測し、正しいイベント系列を選択する。さらに、利用者はイベント-フレーム対応付けモジュールでイベントをフレームに対応付ける。システムはデータベースへのイベント出力モジュールでイベントをインデックスとしてデータベースに格納する。

4 おわりに

本稿では、ビデオテロップに基づいたイベント候補の推定により STRIKE のインデキシングを支援する機構を提案した。状況とそこで発生するイベント系列を状況遷移パターンとして表現する。テロップから観測可能な状況の変化に対して、その変化を満足する状況遷移パターンの組合せを導出することにより、テロップが変化するまでの時区間で発生したと考えられるイベント候補を推定し、インデキシングを支援可能である。

この機構では、推定によって求められたイベント系列の候補を評価しないため、実際の対象世界で起きる頻度の高いイベント系列と起きる頻度の低いイベントが同列に利用者に提示されるという問題点がある。これを解決するためには、イベント系列候補を評価する関数が必要であり、イベント系列の難型との比較によりイベント系列候補を評価する手法を現在開発中である。

参考文献

- [1] 牛尾剛聡, 広部一弥, 渡邊豊英: “イベント-アクション・モデルに基づく動画像検索システム STRIKE”, 情報処理学会研究報告, 98-DBS-115, pp. 79-85 (1998).
- [2] 美濃導彦: “知的映像メディア検索技術の動向”, 人工知能学会誌, Vol. 11, No. 1, pp. 4-9 (1996).
- [3] 桑野秀豪, 倉掛正治, 小高和己: “映像データ検索のためのドロップ文字抽出法”, 信学技報 PRMU96, pp. 39-46 (1996).