

定性時空間推論の枠組みを用いた動画からのイベント推論について

On Detecting an Event from a Video Data in a Framework of Qualitative Spatial Reasoning

阪井田 政樹*¹ 高橋 和子*²

Masaki Sakaida Kazuko Takahashi

関西学院大学大学院 理工学研究科

School of Science & Technology, Kwansai Gakuin University

*²関西学院大学

Kwansai Gakuin University

We describe a method of detecting an event occurring in a video data using qualitative spatial reasoning. We extract objects from a spatial data in a rectangle form and represent their relative positional relationship. We adopt 9box as a basic framework but also use RCC for some objects. An event is defined as a time sequence of spatial relationships. We detect an event in the time sequence of qualitative relationships obtained from a video data. We discuss this method by applying it to a video of soccer game as a case study.

1. はじめに

定性空間推論は、空間データに対してそこにあるオブジェクト同士の相対的な位置情報、方向、大きさなど定性的な情報に着目し、それらを数値ではなく述語で表現し、その上で成り立つ性質や、時間的な変化を推論する手法である [Cohn 08]. 定性空間推論に関する研究は非常にたくさんあるが、2次元の静止画を対象としているものが多い。一方計算機やネットワークの性能向上に伴い多くの動画データを扱う機会が増えており、これらの動画データの分類をするためにそこで起こっているイベントを効率よく抽出する必要がある。動画からイベントを抽出する研究もいくつか行われているが、扱えるパターンは種類は少ない [Sridhar 11].

一般に動画解析は定量的なデータを扱うため、多大な計算量を必要とする。しかし、いつも正確かつ詳細なデータが必要だということではなく、そこで生じているイベントや大まかな流れを把握するためには定性的なデータがあれば十分であり、定性的なデータが扱えれば計算量は少ない。

著者らは定性空間推論の枠組みの1つ RCC [Randell 92] を使ってバスケットボールの動画でイベント推論を行った [Takahashi 12]. しかし、オブジェクトの切り出しが非常に困難な上、RCC では表現力が不足することがわかった。本研究では、この経験をもとに、動画からイベントを抽出する場合、イベントの定性的な表現としてどのような表現力が必要かを明確にすることを目的とする。事例としてサッカーの動画を取り上げ、シュートなどのイベントを自動的に抽出して推論するための方法について考察する。具体的には RCC と 9-ボックスという2つの枠組みを組み合わせるものを提案する。

本発表の構成は以下の通りである。第2節では、RCC と 9-ボックスの記述方法について述べる。第3節では、本研究での領域の扱いについて述べる。第4節では、イベントの定義について述べる。第5節では、結論と今後の課題について述べる。

2. 記述方法

2.1 RCC

Region Connection Calculus (RCC) は、2次元上の2つの領域の相対的な位置関係を、記号を使って表現する定性空間推論の一手法である [Randell 92].

2つの領域 $r1, r2$ の関係は RCC では以下の8つの基本表現のいずれかで表される (図1).

$DC(r1, r2)$ 離れている。

$EC(r1, r2)$ 外接している。

$PO(r1, r2)$ 交わっている。

$EQ(r1, r2)$ 完全に一致する。

$TPP(r1, r2)$ $r1$ が $r2$ に内接している。

$TPPi(r1, r2)$ $r1$ が $r2$ に内接されている。

$NTPP(r1, r2)$ $r1$ が $r2$ に含まれていて接していない。

$NTPPi(r1, r2)$ $r1$ が $r2$ を含んでいて接していない。

これらは互いに排反であり、また任意の2つの領域はこれらのいずれかの関係になっている。

2つの領域間の関係の時間的推移を考えると、直接遷移可能なものとそうでないものがある。例えば、 DC から PO に遷移するためには必ず EC という状態を通過する必要があり、 PO から直接可能な遷移は $EC, TPP, EQ, TPPi$ の4つである。

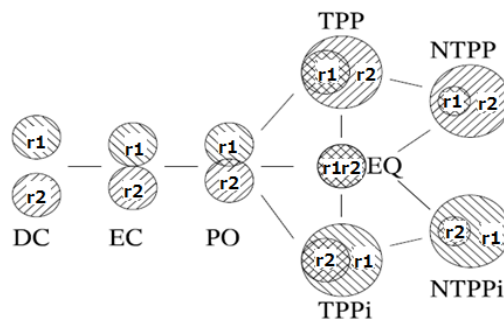


図1: RCC8 の基本表現

連絡先: 阪井田 政樹, 関西学院大学大学院 理工学研究科 高橋和子研究室, 兵庫県三田市学園2丁目1番地, 079-565-8391, bfd76635@kwansai.ac.jp

2.2 9-ボックス

9-ボックスは、やはり2つの領域の相対的な位置関係に着目した表現方法である [Cohn 12]。9-ボックスでは単位となる領域を矩形に限定し、これらの閉包を9つの部分枠(ボックス)に分割する。9つの枠に対して各領域がそこに存在しているかどうかを記述することで、2つの矩形領域の位置関係を表す(図2)。RCCとは異なり2つの領域がどの部分、位置で接している、交わっているかなどの方角まで区別でき、これらの関係の変化を詳しく扱うことができる。また、数値データを付加して相対的な大きさや距離も区別できるように拡張した形も考えられている。図2は離れている領域A,Bを9-ボックスで表現したものであるが、A,Bは離れているということ以外に、Aの方がBよりも上下に長いことがわかる。

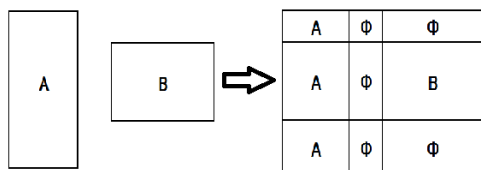


図2: 9-ボックスの例 (Φのボックスには領域が存在しない)

3. 領域間の関係の記述

ここでは事例としてサッカーの試合を取り上げるが、試合中に起こるイベントを記述するという観点でRCCと9-ボックスを比較する。

RCCは9-ボックスと比べるとデータ量も少なく領域同士の相対的な位置関係はわかるが、相対的方向や共有部分の各領域内での位置は表現できない。そのためプレイヤーが体のどの部分でボールに触れているかが表現できない。例えば、プレイヤーがボールをキックしたかヘディングしたかを区別する場合、RCCではボールとプレイヤーの領域が接したか交わったかまでしかわからずどちらを行ったか判断できない。しかし9-ボックスならば、ボールの領域とプレイヤーの領域の下部分が交わっていればキックであり、上の部分が交わっていればヘディングだと判断できる。従ってプレイヤーやボールの関係は9-ボックスを使って記述する。

一方9-ボックスは矩形領域が基本のため、動画を撮る方向によっては意図した領域が正確に切り取れない。2次元上の写像ではゴールの領域が矩形にならず、無理にゴールを矩形で切り出すと、ゴールでない部分もゴールの領域と判断されてしまう。そのため、ゴールに関しては矩形ではなくゴールラインを領域 GF_1 、奥のゴールポストを領域 GF_2 、ゴールの上2辺の領域を GB 、ゴールラインを含めたゴール全体を領域 G として扱う(図3)。この時 GF_1, GF_2, GB は極めて細い領域とする。ただし、ボールが当たってゴールネットが歪んだ場合、領域 G もそれに対応して歪むものとする。これらの領域は矩形ではないので、他の領域との関係についてはRCCを用いて記述する。

4. イベントの定義

例としてシュートというイベントを定義する。ただし、ボールの軌跡が図4のように角度が急な弧を描く場合は極めて低いことから今回は除外して考える。

まずボールがラインを越えるイベントを定義する。この時ラインは極めて細長い領域と考える。ボールの領域 Ba 、ライン

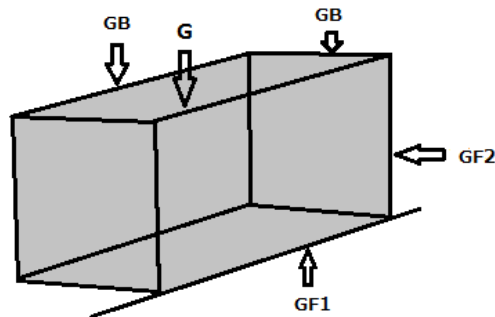


図3: ゴールの領域

の領域 L 、越えた後に至る領域 Do の関係をRCCで記述する。 Ba と L の関係は外接から交わり再度外接という流れになる。また、 Ba が Do に含まれるか否かを記述することでラインを越える方向を示す。このイベントはRCCを使って以下のような時系列として定義できる。

$$EC(Ba, L) \wedge \neg(TPP(Ba, Do))$$

↓

$$PO(Ba, L)$$

↓

$$EC(Ba, L) \wedge TPP(Ba, Do)$$

↓

$$NTPP(Ba, Do)$$

このイベントを $PL(Ba, L, Do)$ と表す。

次に述語 PL を使ってシュートのイベントを定義すると、 $PL(Ba, GF_{1(2)}, G)$

↓

$$\neg PO(Ba, GB) \wedge \neg PL(Ba, GF_{2(1)}, \bar{G})$$

となる。これはボールの領域 Ba が領域 $GF_{1(2)}$ を越えてゴールの領域 G に入り、その後領域 GB と交わらず、 $GF_{2(1)}$ を越えて G 以外の領域 \bar{G} に至らないことを表す。最後の条件は3次元を2次元に写像しているため必要なものである。たとえば図5では $PL(Ba, GF_2, \bar{G})$ が満たされているためシュートが失敗、またはシュートではなかったと考えられる。

ボールがプレイヤーに当たり返されたというイベントはプレイヤーの領域を P とすると、

$$PL(Ba, GF_1, G) \vee PL(Ba, GF_2, G)$$

↓

$$EC(Ba, P) \vee PO(Ba, P)$$

↓

$$PO(Ba, GB) \vee PL(Ba, GF_1, G) \vee PL(Ba, GF_2, G)$$

と記述できる。

5. おわりに

本研究ではサッカーの試合を事例として定性時空間推論の枠組みを用いて動画からそこで起こっているイベントの推論手法について考察した。しかし、この手法では現在、2つの領域のどちらが視点から見て前面にあるかといった関係や、部分的に隠れている領域の補完などを人間が決定しているため一般的とは言えない。今後はこれらの扱いをどのようにするかを決める必要がある。

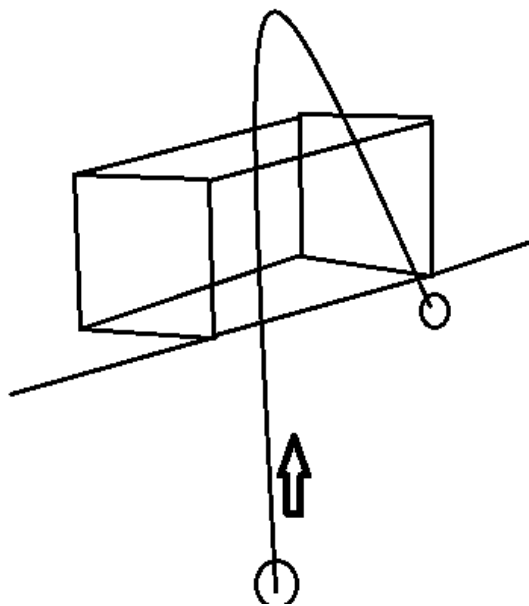


図 4: あまり起こりえない軌跡

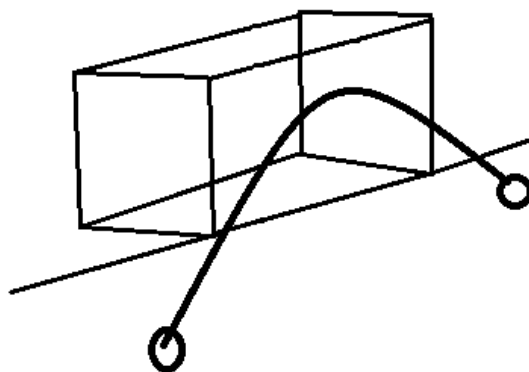


図 5: GF_1 を越えた後に GF_2 を越える場合

参考文献

- [Cohn 08] Cohn G. A. and Renz J., “Qualitative spatial representation and reasoning”, Handbook of knowledge Representation, Chapt. 13, pp. 551-596, F. van Harmelen, V. Lifschitz, and B. Porter, Eds., Elsevier, 2008.
- [Cohn 12] Cohn G. A., Renz J. and Sridhar M., “Thinking Inside the Box: A Comprehensive Spatial Representation for video Analysis”, Proceedings of 13th International Conference on Principles of Knowledge Representation and Reasoning, 2012.
- [Randell 92] Randell A. D., Cui Z. and Cohn G. A., “A spatial logic based on regions and connection”, pp.165-176. Proc.of KR92, 1992.

[Sridhar 11] Sridhar M., Cohn G. A., Hogg D, C, “From Video to RCC8: exploiting a Distance Based Semantics to Stabilise the Interpretator of Mereotopological Relations” Proc. COSIT, 2011.

[Takahashi 12] Takahashi K., “Reasoning about Relative Relationships in 3D Space for Objects Extracted from Dynamic Image Data”, ECAI-2012 Workshop on Spatio-Temporal Dynamics(SteDy 2012), 2012.