

映像解析による将棋対局の自動棋譜生成システムの試作

小森 匠† 林 貴宏†

関西大学 総合情報学部†

1. はじめに

将棋の上達には、棋譜の作成が有効とされている。これまでにコンピュータ支援による棋譜作成を目的として、カメラで撮影された将棋盤の画像から局面情報を認識する手法が開発されている[1]。この手法では撮影された一つの盤面に対し、盤上の各位置に対して駒の認識を行う。しかし、一箇所の誤認識が局面全体の認識に与える影響は無視できない。さらに、入力された盤面のみでの再現であるため、対局全体を作成するには手間がかかる。本研究では、将棋の対局を録画した映像から、可能な指し手(以下、合法手)を考慮することで、認識対象位置を絞り込み、画像認識モデルの使用を最小限化するとともに対局全体の棋譜を生成する手法を提案する。

2. 提案手法

2.1 概要

提案手法の概要を図1に示す。提案手法では、盤面の認識を行うために、映像から将棋盤領域を特定し、透視変換を適用することで盤面のみの画像に変換する。次に、盤上での変化を得るため、連続する画像間でフレーム間差分を取り、変化があったマス特定する。最後に、変化を検出できた各マスをCNNにかけ、駒の種類を特定することで、指された手を特定する。本研究では、対局映像は、先手から見て将棋盤の右横側から撮影されたものであり、カメラ、将棋盤が固定されているものと仮定する。

2.2 将棋盤領域の検出

映像中のフレーム画像から、将棋盤の領域を検出し、将棋盤の四隅の座標を取得する。図2は検出の流れを表したものである。

まず、フレーム画像をグレースケール変換し、適応的閾値処理(近傍領域の重み付き平均値を閾



図1. システムの構成

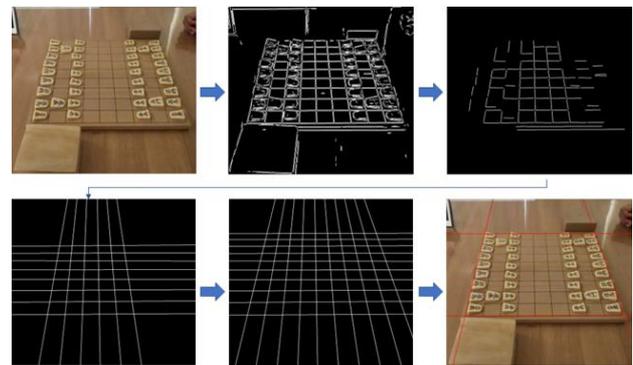


図2. 将棋盤領域特定の流れ

値として処理)で二値化し、LSD[2]を用いて線分検出を行う。次に、各線分を、傾きを用いて水平方向、垂直方向に分類する。このとき、盤上から外れた位置の線分を検出して除外するため、線分の midpoint の X 座標、Y 座標の Z スコアの絶対値を計算し、両方とも閾値 1.5 以下を満たす線分のみを残す。同一直線状にある線分同士は同一の直線であるとみなすことで枠を構成する線分を得る。

9x9 のマスを構成するには、10x10 の線分の集合が必要であるため不足している線分を補完、追加する必要がある。線分を追加する際、線分を上下左右方向それぞれに何本追加するのか考慮する必要がある。例えば、水平方向の線分が 8 本取得できた場合、2 本の線分を上下方向に追加するパターンは、上に 2 本、上に 1 本下に 1 本、下に 2 本である。そのため、それぞれのパターンの線分を描画した画像をパターンの枚数分作成し、元の線分検出された画像との一致度が最も高いパターンを正しく補完できた線分であると判断する。

Development of an automatic game score generation system for shogi games using video analysis

†Takumi Komori, Takahiro Hayashi

†Kansai University

表 2. ネットワーク構造

| Layer | Output Shape |
|--------------|---------------|
| Conv2D | (62, 62, 32) |
| MaxPooling2D | (31, 31, 32) |
| Conv2D | (29, 29, 64) |
| MaxPooling2D | (14, 14, 64) |
| Conv2D | (12, 12, 128) |
| MaxPooling2D | (6, 6, 128) |
| Flatten | (4608) |
| Dense | (128) |
| Dense | (1) |

最後に、得られた線分から将棋盤の四隅の座標を特定し、その座標を用いて映像の各フレーム画像に透視変換を適用することで盤面の画像を得る。

2.3 駒の画像認識

盤上の画像を 9x9 のマスに分割し、各マスでフレーム間差分(各画素の濃度差の合計)を計算する。差分が閾値以上あればそのマスにおいて変化があったと判断できる。変化があったマスを後述する画像認識モデルに入力することで、各位置にある駒の種類を特定する、現在の局面における合法手、駒の種類と変化があったマスの位置によって、指された手を特定できる。

駒の画像認識モデルは、CNN(Convolutional Neural Network)を用いる。ネットワークの構造を表1に示す。CNNは、64x64pixelの盤面を9x9に分割したマスの画像を入力とし、特定の駒かそれ以外の駒かを判定する二クラス分類モデルとして、それぞれの駒ごとに構築される。駒14種類、駒なしを含めて15種類のモデルが構築される。出力層はSigmoid関数を用いた。駒の画像認識において、合法手と変化したマスの位置を考慮することで、可能性のある駒に対応するモデルにのみ入力すればよいため、画像認識モデルの使用を最小限に留めることができる。

本機能の実装において、合法手の取得にはpython-shogiライブラリを用いた。学習で使用するデータセットは、実際の将棋映像から切り抜いた各駒の画像をそれぞれ6枚ずつ、背景画像12枚を用いた。駒画像を±15°のランダムな角度で、明るさやコントラストをランダムに変化させて背景画像に重ね合わせることで生成した。各クラス1万枚の画像を生成した。

3. 評価実験

実際の対局映像を使用して、提案手法の有効性を確認する実験を行った。実験で使用した対局映像は、対局途中の22手分を切り出した映像である。

提案手法による局面認識の結果、22手中14手に対して正しい手を出力することができた。出力に失敗した例として、将棋盤領域の検出に失敗したケース、画像認識モデルが誤判定したケースが確認された。将棋盤領域の検出に失敗したケースでは、線分の補完処理を精度よく行うことができなかった。また、画像認識モデルが誤判定したケースは、学習で用いた各駒の学習用画像が不足している点が問題であると考えられる。これらの対策として、将棋盤領域検出手法の改良、学習用画像の拡充が必要である。また、両対局者の手が同時に写りこんでいるために、一手分の正確な盤面画像の差分を得ることができず、認識に失敗した例も見られた。この場合は数手動いた後の盤面との差分から、複数の手を推測する必要がある。

4. おわりに

対局映像から棋譜を生成するために、差分情報と画像認識モデルを使った手法を提案した。実験により提案手法の有効性と限界点を確認した。実験により確認された課題として、将棋盤領域検出手法の改良、画像認識モデルに使用したデータセットの拡充、正確に盤面画像の差分が取得できない場合を考慮した処理の追加が残されている。さらに、実際の対局映では様々な角度から撮影される、カメラ、将棋盤がずれることが予想されるため、将棋盤領域検出において、四隅の座標を毎フレーム高い精度で検出する、もしくはトラッキングするといった処理を加える必要がある。

参考文献

- [1] 佐藤旭, 森住啓, 清水郁子, 中川正樹, 「スマートフォンで撮影された将棋局面画像の認識」, 電子情報通信学会技術研究報告, pp. 11-15, 2019.
- [2] R. G. V. Gioi, J. Jakubowicz, J. M. Morel and G. Randall, “LSD: a Line Segment Detector”, Image Processing On Line, pp. 35-55, 2012.