

TCG プレイ動画におけるカード特定のための SURF 特徴量を用いたカード認識*

本園 遥[†]
高知工科大学[†]

栗原 徹[‡]
高知工科大学[‡]

1 はじめに

トレーディングカードゲーム (TCG) とは各プレイヤーがコレクションしたカードの中からルールにのっとって組み合わせたカードの束であるデッキを持ち寄り、2人以上で対戦を行うゲームである。身近な人と対戦はもちろんのこと、全国大会や世界大会などの多くの大会も開催されている。今日では、インターネット、動画配信サービスの普及によりこれらで行われた対戦動画などを配信するようになり、より多くの人が対戦の様子を楽しむようになった。しかし、TCG のカードは種類が多いうえにカードの名前や効果が書かれた文章が小さく、動画ではどのカードが利用されているのか分かりにくい状態が多々ある。

TCG にはさまざまな種類が存在するが、今回は遊☆戯☆王オフィシャルカードゲーム (遊戯王 OCG)[1] を対象とし、そのカードを提示した動画に対し、SURF 特徴量 [2] を用いたカード認識システムを提案する。

2 提案手法

本提案手法において、以下の処理手順で行う。まず、前処理として (1)(2) において使用する全テンプレート画像の SURF 特徴量から Bag-of-Features の作成とカードのタイプ特定を行う。次に、動画 1 フレームごとに (3)-(5) の処理を行い、カードの認識をする。

カードのタイプとは表 1 に示すように枠の色とカードの裏表を用いたカード分類のことである。遊戯王 OCG のカードは裏である場合カードの中心が黒であり、表であった場合は枠の色を用いることで 9 種類に分けることができる。これらを用いた 10 種類の分類をカードのタイプとする。

- (1) 各テンプレート画像から SURF 局所特徴量を抽出し、Bag-of-Features を用いて各テンプレート画像のヒストグラムを作成する。
- (2) 各テンプレート画像のカードのタイプを特定し、ヒストグラムをタイプ別に分類する。
- (3) 撮影した動画からカード部分をトリミングする。
- (4) トリミングしたカード画像から SURF 局所特徴量を抽出し、Bag-of-Features を用いてヒストグラムを作成する。
- (5) トリミングしたカード画像のカードのタイプを特定する。その結果と同じタイプの各テンプレート画像のヒストグラムと比較し、類似度が一番高いテンプレート画像を認識画像とする。

2.1 Bag-of-Features

Bag-of-Features[3] は多数の局所特徴量をベクトル量子化し、ヒストグラムとして表す手法である。これを用いて各テンプレート画像と動画内のカード画像のヒスト

グラムを作成する。

全テンプレート画像の SURF 局所特徴量を抽出する。抽出した局所特徴量に対し、k-means 法を用いてクラスタリングを行い、visual words を作成する。このときのクラスタ数は 15 個とした。作成した visual words をもとに各テンプレート画像と 1/2, 1/4, 1/8 にリサイズした各画像のヒストグラムを作成する。これにより動画内でのサイズの違いによつての誤認識を減らす。また、動画から抽出したカード画像のヒストグラムも visual words をもとに作成する。

表 1 表カードのタイプと色

タイプ	枠の色
シンクロ	彩度 50 以下
エクシーズ	明度 60 以下
通常	色相 16 以下
効果	色相 16 以上 25 以下
儀式	色相 100 以上 120 以下
融合	色相 120 以上 150 以下
ペンデュラム	色相 16 以上 25 以下 色相 80 以上 100 以下 (上下で色が違うため)
魔法	色相 80 以上 100 以下
罨	色相 150 以上 170 以下

2.2 カードのタイプ特定

カードのタイプを特定することで、比較する数を少なくし、誤認識を減らす。図 1 は裏のカード、図 2 は表のカード例である。カードのタイプ特定には図 3 の黒い四角部分 3 か所における画素の HSV 値を利用する。中央のタイプ特定領域はカード横幅 1/6 の正方形領域、左側上下のタイプ特定領域はカード横幅 1/20、縦幅 1/3 または 2/3 部分の 3×3pixel である。まず、裏表の判別を行う。中央のタイプ特定領域内における HSV の分散値を計算し、明度の分散値が 10 以下の場合を裏のカード、それ以外を表のカードとする。次に表のカードと判定されたものに対して表 1 のどのタイプに対応するカードであるか特定を行う。左側上下のタイプ特定領域における HSV の平均値を計算し、先に表 1 の値よりシンクロまたはエクシーズに当てはまるカードがあるか調べる。これに当てはまらない場合は他の 7 種類のどれに当てはまるのか表 1 の値より調べ、タイプを特定する。このとき、上下の画素を利用する理由として、上下の枠の色が違うタイプがあるためである。

2.3 カードの抽出

動画の 1 フレーム画像からカード部分だけを抽出する。画像に対して輪郭抽出を行い、カード部分だけを抽出する。抽出した画像に対し、射影変換を行い、傾いて

* Card recognition using SURF in TCG play movie

[†] 「Haruka Motozono · Kochi University of Technology」

[‡] 「Toru Kurihara · Kochi University of Technology」



図1 裏



図2 表



図3 カードのタイプ特定に利用する画素

置かれた画像であってもカード部分だけをトリミングするようにする。

2.4 ヒストグラムの比較

トリミングしたカードのタイプと同じ各テンプレート画像のヒストグラムとトリミングしたカード画像のヒストグラムを交差法を用いて比較をする。この結果により類似度が一番高いヒストグラムのテンプレート画像を認識結果とする。

3 実験

3.1 実験環境

本実験では、Web カメラを用いて図4のようにカードが机の上に1枚出された動画を撮影し、それに対してカードの特定を行う。今回使用したカードは図5である。撮影にはiBUFFALO製のBSW50KM02BKのWebカメラを用いる。動画サイズは1920×1080、フレームレートは30fpsである。テンプレート画像として370枚のスキャンした遊戯王 OCG カード画像を利用する。テンプレート画像は708×1038pixelである。処理にはpython, OpenCV, scikit-learnを利用した。出力結果として類似度結果の上位20位を表示し、正しいテンプレート画像がどの順位であるかを調べる。

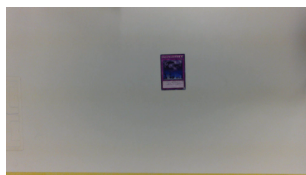


図4 動画の1フレーム

3.2 実験結果

撮影した動画の総フレーム数85フレームのうち、カードが手から離れ、机に置かれているフレーム数は35フレームであった。そのいずれにおいてもカードのトリミングを行うことができ、すべてにおいて正しいカードのタイプに分類を行えた。図6はその中の1フレームにおけるトリミングされたカード画像である。カードが机に置かれている各フレームにおける上位20位の類似度を図7に示す。図における横軸は総フレーム中カードが机に置かれているフレーム番号を、縦軸は類似度の値である。ひし形のマーカーが正しいテンプレート画像の類似度を、丸いマーカーは上位20位までの類似度を表して

おり、ひし形のマーカーより上部に丸いマーカーが現れているフレームは誤認識となる。35フレームのうち、正しいテンプレート画像の類似度が1位であったものは8フレームであり、最低順位は12位であった。



図5 動画内で使用したカード



図6 トリミングしたカード画像

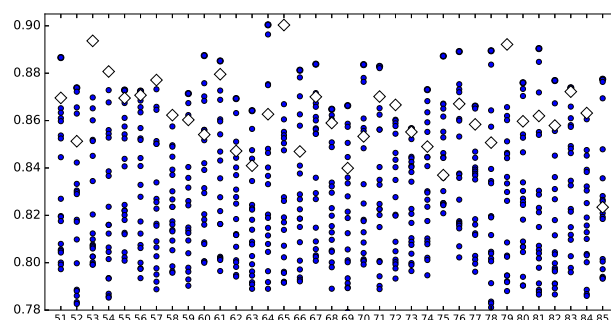


図7 各フレームにおける上位20位の類似度

4 考察

実験結果からカードがトリミングされた35フレームすべてにおいて、類似度順位12位以内に正しいテンプレート画像が認識されたが、1位であるフレームは35フレーム中8フレームのみであった。これは、SURF局所特徴量の特徴点の位置を考慮しなかったからではないかと考えられる。今後の課題として、特徴点の位置を利用し、さらなる認識率の向上を行う必要がある。

5 おわりに

本稿ではSURF特徴量を用いて動画における遊戯王 OCG カードの認識を行った。結果として、総フレーム数85フレームのうちカードが手から離れ机に置かれたフレーム数35フレームすべてにおいてカードのトリミングを行うことができ、カードのタイプも正しく特定することができた。カードが検出された35フレーム中において正しいテンプレート画像の類似度順位は12位以内であったが、類似度順位が1位であるフレームは8フレームのみであった。今後は、SURF局所特徴量の特徴点の位置を利用することでさらなる認識率の向上を行う予定である。

参考文献

- [1] 高橋和希 スタジオ・ダイス/集英社 企画・制作 /KONAMI, 遊戯王オフィシャルカードゲーム, <http://www.yugioh-card.com/japan/>.
- [2] H.Bay A.Ess T.Tuytelaars and L.Gool, "Speed-up robust features(SURF)", Computer Vision and Image Understanding, 110, pp.346-359, 2008.
- [3] 岡田貴之ら, "コンピュータビジョン最先端ガイド 3", アドコム・メディア株式会社, 2010.