

カスケード型分類器を用いた照明環境変化にロバストなサッカーボール認識

A Robust Soccer Ball Recognition against Lighting Fluctuation by Using Cascade Classifier

森のどか，會澤要，鈴木拓央，小林邦和

Nodoka MORI, Yo AIZAWA, Takuo SUZUKI and Kunikazu KOBAYASHI

愛知県立大学 大学院 情報科学研究科

Graduate School of Information Science and Technology, Aichi Prefectural University

kobayashi@ist.aichi-pu.ac.jp

Abstract

現在ロボカップサッカーの標準プラットフォームリーグ (SPL) では，白地に黒色の五角形のテクスチャ模様がある，いわゆるサッカー柄のボールを使用し，自然光が差し込むフィールドでの精度が高いボール認識が要求されている．本稿では，そのような照明環境が変動する環境に対してロバストかつ，リアルタイムにボール認識を行うため，画像の特徴量によるカスケード型の分類器を用いた手法を提案する．そして，最良の精度が得られる特徴量を調べるため，複数の特徴量を用いて比較検証を行った．

オレンジ色のボールを使用していたときは，画像内のオレンジ色の領域を探索すればよかったが，白地に黒色の模様のボールに変わり，単にボールの色を探索するだけでは認識できなくなった．しかし，その手法ではボールの影や，内部に生じるエッジによりズレが生じたり，NAOの関節などをボールと誤検出してしまう (例：図 3,4) ．



図 3: 影によって誤認識した図

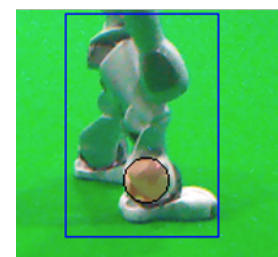


図 4: NAO をボールと誤認識した図

1 はじめに

ロボカップサッカーの標準プラットフォームリーグ (SPL) は，全チームが同一のロボットを用いてプログラミングの優劣を競うサッカーリーグである [1]．現在は Soft-Bank Robotics 社のヒューマノイド型ロボット，NAO が用いられている．また，SPL のルールは年々変更されており，人間がサッカーを行う環境に近づき，難易度も増している．その中でもボールは，以前はオレンジ色のボール (図 1) を用いられていたが，ルール変更により白地に黒色の五角形のテクスチャ模様のボール (図 2) に変わり，フィールド環境も自然光の中で行うよう変更された．

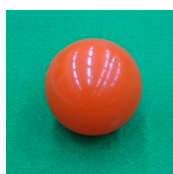


図 1: 以前使用していたオレンジ色のボール



図 2: 現在使用している白地に黒色の模様のボール

また，NAO は計算能力が限られており，その中でボールや NAO の認識，自己位置の計算等をリアルタイムで行わなければならない．

近年，物体認識の手法として，機械学習を用いる手法が注目されている．この手法は，人が物体の特徴を設定したりする必要がなく，複雑な物体でも認識することが可能である．有名なものとして，Deep Learning を用いた手法が挙げられる．これは，現在物体認識において，認識精度が最も高い手法であり，2012 年の画像認識コンテストで 2 位の手法と圧倒的な差をつけて優勝したことにより注目されるようになった．しかし Deep Learning を用いた手法は，計算量が莫大であるため，NAO の CPU ではリアルタイム処理が難しい．そこで本研究では，機械学習の内，Boosting されたカスケード型分類器を用いることで，NAO でリアルタイムなボール認識を実現させた．

2 先行研究

2.1 Boosting

ブースティングとは、弱分類器を組み合わせることで、より精度の高い分類器を作り上げる手法である (図 5)[2]。

本研究では、複数ある Boosting の手法の内、AdaBoost という手法を用いた。AdaBoost とは、Adaptive Boosting の略であり、学習過程で、分類器の認識率に対して適応的に重み付けをし再学習することで精度の良い分類器を作る手法である [3]。

図 5 において、 $h_T(x)$ は、 T 個目の特徴量を指し、 α は重みを指す。

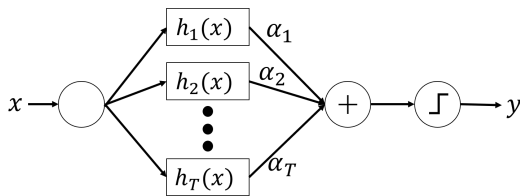


図 5: AdaBoost

2.2 画像特徴量

カスケード型分類器は、学習時に特徴量を抽出しそこから分類器を作成する。今回は以下の 3 つの特徴量について比較した。

- Haar-like 特徴量
- Local Binary Pattern (LBP) 特徴量
- Histogram of Oriented Gradients (HOG) 特徴量

次項からそれぞれの特徴量の説明を記す。

2.2.1 Haar-like 特徴量

Haar-like 特徴量とは、画像の輝度の差を利用したものである [4][5]。この手法は、予め Haar-like 特徴と呼ばれる矩形領域のパターンを複数用意する (図 6)。

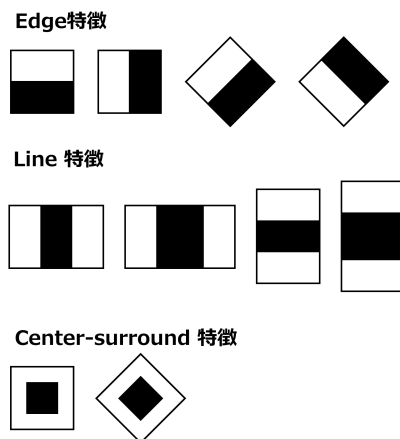


図 6: Haar-like 特徴

次に、入力画像上の一部分を切り取り、そこにパターンをそれぞれ当てはめ、式 (1) のように計算し、その値を特徴量とする。

$$H(r_1, r_2) = S(r_1) - S(r_2) \quad (1)$$

ここで、 $S(r_1)$ と $S(r_2)$ がそれぞれ用意したパターン中の白と黒の 2 つの領域の平均輝度値である。

この手法は、顔認識によく用いられており、年齢、人種関係無い特徴を得ることができる。

2.2.2 LBP 特徴量

LBP 特徴量とは画像のヒストグラムを利用したものである [6][7]。この手法は回転変動などには弱いですが、画像の照明変化の影響を受けにくく、また高速計算が可能である。

LBP の計算は 3×3 ピクセルサイズの画素領域に着目し行う。まず中心の輝度値と周辺 8 近傍の画素の輝度値を比較する。その 8 近傍の内、輝度値が中心の輝度値以上のとき 1、それ以外を 0 とする。ここにマスクを掛け合わせ、その総和を求めこの値が中心画素の輝度値と置き換える。マスクとは、左上から時計回りに 2^n の重みを割り振ったものである (図 7 中、右から 2 番目の図)。この作業を全画素に対して行い、できた画像を LBP 画像と呼ぶ。こうして求めた LBP 特徴量を用いて物体を認識する。

図 7 の場合、上記のように計算し、式 (2) より LBP 特徴量は 241 となる。

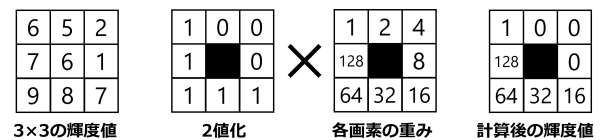


図 7: LBP の計算方法

$$LBP = 1 + 16 + 32 + 64 + 128 = 241 \quad (2)$$

2.2.3 HOG 特徴量

HOG 特徴量とは、局所領域の輝度の勾配方向をヒストグラム化したものである [8]。HOG 特徴量の求め方は、まず局所領域の輝度の勾配方向と強度を算出し、ヒストグラムを作成する。そして局所領域ごとで正規化を行うことで求めることができる。

この特徴量は、画像内の対象物体の大きさ変化や回転には弱いですが、対象の物体が一定の枠内に収まっている画像等では、照明の変化などにも強く、人物検出として用いられている。

2.3 Attention Cascade

Attention Cascade とは、強分類器を複数連結させた構造をしている (図 8)[9]。こうすることで、上位で関係ない領域を早めに削除していくため、効率よく認識するこ

とができる．各ステージの強分類器は AdaBoost で学習させている．



図 8: カスケードと学習器の構成



図 10: ポジティブ画像 (屋外)

3 提案手法

本研究は，画像の特徴量を用いてカスケード型分類器を作成し，ボール認識を行った．カスケード型分類器は，画像群を学習データとして学習させ作成する．用いる学習データセットとして，ボールが写っている画像 (ポジティブ画像) と，ボールが写っていない画像 (ネガティブ画像) をそれぞれ用意した．ポジティブ画像には，室内で撮影したものと屋外で撮影したものを (例：図 9,10)，ネガティブ画像に，画像内にボールが写っていない画像を用いた (例：図 11)．これらの画像データセットは公開しており，誰でも利用可能である¹．

また，機械学習で精度を向上させるためには大量のデータが必要となるため，同じ SPL チームの SPQR が公開しているデータセットも用いた [10]²．これらの画像データを用いて，カスケード型分類器を作成する．



図 9: ポジティブ画像 (室内)



図 11: ネガティブ画像 (屋内)

また，画像全体を探索すると処理に時間がかかってしまうため，まずボールの候補領域を抽出し，その領域候補を分類器に適用する．このボール候補領域の抽出には，SPL のチーム，B-Human のコードリリース 2016 内の『Ball-Region』というモジュールを使用した [11]³．このモジュールでは，まず画像全体を垂直方向に探索し，白色の領域をボールの候補として検出している．

本研究では，このモジュールにより得られたボール候補をカスケード型分類器に入力し，1 つに絞った．

4 実験

本研究では，前節で説明したそれぞれの特徴量のカスケードを作成し，性能を比較する．評価するポイントは次の 3 点である．

- ボールの認識率
- 照明環境の変化に対するロバスト性
- 処理速度

¹<http://www.ist.aichi-pu.ac.jp/lab/robocup-spl/en/publications.html>

²<http://www.dis.uniroma1.it/~labrococo/?q=node/459>

³<https://github.com/bhuman/BHumanCodeRelease/tree/coderelease2016>

実験条件は表 1 の通りである．実験で使ったプログラムコードは github で公開している⁴．今回，ステージ数を 14 にしたのは，それぞれの特徴量でカスケードを作成した際に，最も少ないステージ数で候補を絞れたカスケードのステージ数に揃えたためである．

表 1: 実験条件

画像処理ライブラリ	OpenCV 2.4.13.1
Boosting type	Gentle AdaBoost
プログラミング言語	C++
画像サイズ	640×480
カスケードのステージ数	14

また，精度を比較するにあたり，認識率の値とボールではない部分をボールだと誤認識した数を用いた．誤認識の判断は，ボール全体を囲っていない，又は，囲った内側とボールの間に隙間がないものとした．

4.1 実験用カスケード作成

まず，実験を行うために Indoor カスケードと Outdoor カスケードの 2 種類のカスケードを作成する．本研究では，カスケードの作成に OpenCV (Open Source Computer Vision) のライブラリを用いた．

学習に用いるデータ画像は，NAO に備わっているカメラで撮影した．NAO のカメラは額の中心と口元の 2 箇所にあり，この内，額の中心にあるカメラを用いた．この画像サイズは 640×480 であり，フレームレートは 30fps である．

Indoor カスケードは屋内で撮影した画像のみで学習させ作成する．学習に用いたデータの内訳は表 2 に示す．

表 2: Indoor カスケードの作成に用いた画像枚数

ポジティブ画像	700 枚
ネガティブ画像	300 枚

Outdoor カスケードは，屋外で撮影した画像と屋内で撮影した画像を両方用いて学習し作成した．学習に用いたデータの内訳は表 3 に示す．

表 3: Outdoor カスケードの作成に用いた画像枚数

ポジティブ画像	屋内画像	700 枚
	屋外画像	210 枚
ネガティブ画像		300 枚

4.2 ボールの認識精度の比較

屋内で撮影したボール 100 個分の画像群を用いて，それぞれの特徴量における認識精度を比較する．この実験では Indoor カスケードを使用した．実験の結果は表 4 の通りとなった．

表 4: ボールの認識精度

特徴量	認識率 [%]	誤認識数 [個]
Haar-like	95	44
LBP	97	197
HOG	86	5

この結果から，LBP 特徴量は認識率が高いが誤認識した数も多く，HOG 特徴量は他の特徴量と比べると，認識率は低いが誤認識した数は少ないことが分かった．

4.3 照明環境の変化に対するロバスト性の比較

次は，屋外の画像群をボール 100 個分用いて，それぞれの特徴量における精度を比較する．今回の実験では，Indoor カスケードと Outdoor カスケードそれぞれで精度比較を行う．実験の結果は表 5 の通りとなった．

表 5: 照明環境の変化に対するロバスト性

カスケード	特徴量	認識率 [%]	誤認識数 [個]
Indoor	Haar-like	41	78
	LBP	52	405
	HOG	63	11
Outdoor	Haar-like	83	191
	LBP	60	1184
	HOG	85	17

この結果から，照明環境の変化に対して最もロバスト性が高いのは HOG 特徴量となった．そして，全ての特徴量に対して屋内の画像のみで作成したカスケードよりも，屋外の画像も合わせて作成したカスケードの方が認識精度が向上したことも分かった．

4.4 処理速度の比較

最後に，Indoor カスケードを用いて，それぞれの特徴量における処理速度を比較する．今回はネガティブ画像を 2000 枚用いて処理速度を計測し誤認識した数を計 3 回ずつ測定した．実験の結果は表 6 の通りとなった．

この結果から，Haar-like 特徴量が最も処理時間が早く，逆に HOG 特徴量が一番時間がかかることが分かった．しかし HOG 特徴量は誤認識した数が最も少なく，LBP 特徴量が誤認識した数が最も多かった．

⁴<http://github.com/CamelliaDragons/CamelliaDragonsCodeRelease>

表 6: 処理速度

特徴量	処理速度 [ms]	誤認識数 [個]
	8.78×10^4	
Haar-like	8.08×10^4	11902
	8.79×10^4	
LBP	10.14×10^4	39352
	10.09×10^4	
	10.16×10^4	
HOG	19.00×10^4	1687
	18.90×10^4	
	18.20×10^4	

4.5 考察

4.2~4.4節の結果から、それぞれの特徴量に対し次のことが分かった。

Haar-like 特徴量

Indoor カスケードで屋外でのボール認識の精度が良くないが、Outdoor カスケードでは屋外でのボール認識の精度は高いため、学習データに対応する環境のデータを加えることで対応できることが分かる。また、処理速度が3つの特徴量の中で最も早い。

LBP 特徴量

誤認識した数が多く、また Indoor カスケードを用いて屋内でボール認識を行ったときは認識率は一番高いが、Outdoor カスケードを用いたときの屋外のボール認識精度が低いため、より複雑な環境下での精度を高めるには大量の学習データやカスケードのステージ数を増やす必要があると考えられる。

HOG 特徴量

照明変化に強く、誤認識した数も低いが、処理速度が他の特徴量を用いた場合よりも圧倒的に遅い。

これらのことを考えると、SPL におけるリアルタイムボール認識には、3つの特徴量の中では Haar-like 特徴量が適していると考えられる。

最後に、実際に Haar-like 特徴量を用いてボール認識を行った結果を示す。以下に示す画像において、黒い輪郭で赤いフィルターでマークされたものがボールと認識されたものである。

まず屋内にて、3m 離れたところにボールを置いた場合と、4.5m 離れたところにボールを置いた場合でも図 12、13のように正確に認識することができた。



図 12: 3m 離れたボールの認識



図 13: 4.5m 離れたボールの認識

次にボールをフィールドの白線上に置き、認識が困難な場合でも、ボールを正確に認識することができている(図 14)。

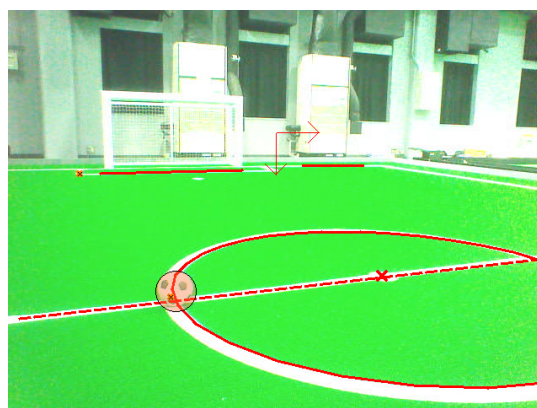


図 14: 白線上のボールの認識

また自然光が当たる屋外における環境においても、ボールの認識が正確にできている(図 15)。

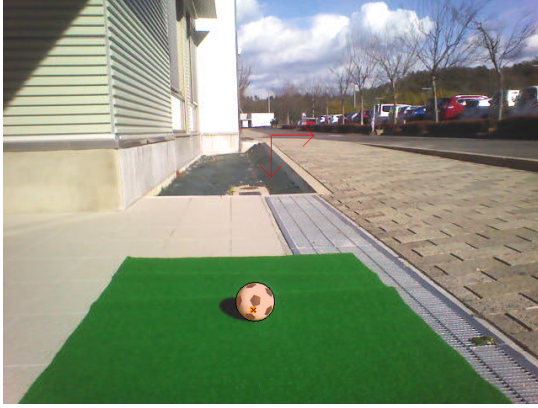


図 15: 屋外でのボールの認識

5 おわりに

本稿では、画像特徴量を弱識別器とした Adaboost を用いたカスケード型分類器によるボールのリアルタイム認識を行った。機械学習により、物体の特徴を自動で抽出できるため、人間にとって特徴を定義するのが困難な場合でも特徴を捉え識別できることが分かった。また前節の結果より、SPL においては、Haar-like 特徴量を用いた物体認識が適していることが分かった。

今後の課題として、静止画像に対して分類器を作成したため、動いているボールの認識精度はあまり高くない。強豪チームはキックしたボールの速度が速いため、このような状況への対処方法も検討していきたい。

参考文献

- [1] Robocup standard platform league. <http://spl.robocup.org/>.
- [2] 藤吉弘亘. 局所特徴量と統計学習手法による物体検出. <http://www.vision.cs.chubu.ac.jp/cvtutorial/pdf/03objectdetection.pdf>.
- [3] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.*, Vol. 55, No. 1, pp. 119–139, 8 1997.
- [4] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. pp. 511–518, 2001.
- [5] Rainer Lienhart and Jochen Maydt. An extended set of haar-like features for rapid object detection. In *IEEE ICIP 2002*, pp. 900–903, 2002.
- [6] 長谷川修. Local binary pattern, a review. 研究報告グラフィクスと CAD (CG), Vol. 2012, No. 3, pp. 1–6, nov 2012.

- [7] 寺島裕貴, 喜田拓也. 勾配情報を用いた local binary pattern の改良. 第 6 回データ工学と情報マネジメントに関するフォーラム論文集, Vol. 2014, , 3 2014.
- [8] 藤吉弘亘. Hog 特徴量と boosting を用いた人検出. http://www.vision.cs.chubu.ac.jp/joint_hog/pdf/HOG+Boosting_LN.pdf.
- [9] 皆川卓也. Opencv による物体検出徹底解説. <https://www.slideshare.net/takmin/20141008-40019861>.
- [10] Domenico Bloisi, Francesco Del Duchetto, Tiziano Manoni, and Vincenzo Suriani. Machine learning for realisticball detection in robocup SPL. *CoRR*, Vol. abs/1707.03628, , 2017.
- [11] Thomas Röfer, Tim Laue, Jonas Kuball, Andre Lübken, Florian Maaß, Judith Müller, Lukas Post, Jesse Richter-Klug, Peter Schulz, Andreas Stolpmann, Alexander Stöwing, and Felix Thielke. B-Human team report and code release 2016, 2016. Only available online: <http://www.b-human.de/downloads/publications/2016/coderelease2016.pdf>.