DEIM Forum 2013 P4-5

色情報によるイラスト画像データの人物キャラクター分類

寺島　裕貴†　　喜田　拓也†

† 北海道大学大学院情報科学研究科　〒 060–0814 北海道札幌市北区北１４条西９丁目
E-mail: †{tera,kida}@ist.hokudai.ac.jp

あらまし　今日では，プロ，アマチュアを問わず，多くの漫画家やイラストレイター，アニメーターと呼ばれる人々が，日々多くのイラストをインターネット上に公開している．本稿では，人物キャラクターが描かれたイラスト画像データに対して，そのキャラクターの判別を行う手法について検討する．具体的には，Local Binary Pattern (LBP) を用いてキャラクターの顔部分の検出を行い，カラーヒストグラムおよび色順位に基づき類似度を測ることでキャラクターの分類を行う．また，各要素技術について評価実験を行い，本手法の有効性について議論する．
キーワード　物体認識, 顔検出, イラスト, Local Binary Pattern, カラーヒストグラム

Illustrated Characters Classification with Color Information

Hiroki TERASHIMA† and Takuya KIDA†

† Graduate School of Information Science and Technology, Hokkaido University
Kita 14jo, Nishi 9, Sapporo, 060–0814 Japan
E-mail: †{tera,kida}@ist.hokudai.ac.jp

Abstract
Key words　object recognition, face detection, illustration, Local Binary Pattern, color histogram

1. はじめに

1.1 研究背景とその目的

近年のインターネットの普及に伴い，我々の日常生活において様々なデジタル化された画像，及び動画が存在する．特に，近年では安価でかつ高性能なペイントソフトの登場やソーシャル・ネットワーキング・サービスの充実，多くのメディアの存在によりプロ，アマチュア問わず多くの人が絵を描くようになった．その影響により，アニメや漫画，イラストといったメディアが急速に発展してきた．よって，今後これらのメディアで用いられる画像を解析することは非常に重要な問題であると考えられる．これらの画像は，人手によって描かれた絵であり，実世界の画像とは異なる．すなわち，スケールの変化，照明の変化，オクルージョンなどのアピアランスの変化の他に，クリエイターの特徴，色の塗りなどといった変化が増え，実世界の画像に比べ同一の対象物に対して非常に大きくアピアランスが変化する．本稿では，このような画像を**イラスト画像**と呼び，画像内に描かれた人型の絵を**キャラクター**と呼ぶ．そして，本稿では既存の画像認識に用いる手法を用いてキャラクターを判別する手法を提案する．

提案は，入力されたイラスト画像に対し物体検出を行い，画像内のキャラクター部分を検出し出力する．その後，画像分類の手法を用いてキャラクターを判別し分類していく．提案システムにおける物体検出ではテクスチャ解析の一般的な手法である **LBP (Local Binary Pattern)** [5] を用た．LBP は計算コストが少なく輝度情報に基づいた特徴量であり，濃度変化に不変な性質を持っている．

提案システムにおけるイラスト画像の画像分類では，キャラクターの持つ色情報はキャラクターを認識し分類するための最良な特徴であると判断し**カラーヒストグラム**および色順位による特徴量を用いた．

実験では，LBP を用いてキャラクター顔検出を行いその性能について評価し，考察する．また，カラーヒストグラムと色順位特徴量を用いてキャラクターを分類する実験を行いその性能を求め，提案するキャラクターを判別する手法の有効性について述べる．

結果の概要として LBP を用いたキャラクター顔検出実験では誤検出が少なく精度が良い一方で検出できない顔が多々存在した．また，カラーヒストグラムと色順位に基づいた特徴量を用いたキャラクター分類実験では用いた類似度からでは分類率が 40% 弱に留まった．

本稿の構成は，2 節で関連研究を紹介し，3 節で LBP についての説明を行い，4 節で提案するキャラクター判別システム及び，カラーヒストグラムに基づいた色順位特徴量についての説明を行う．そして，5 節でキャラクターを判別する手法についての実験を行い性能を述べる．

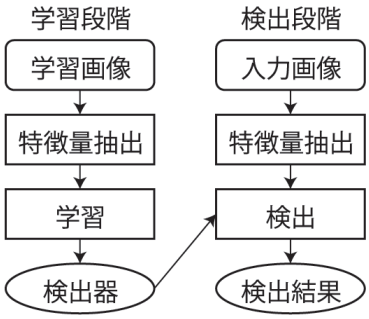

図 1　物体検出の流れ

2. 関 連 研 究

物体検出では，積分画像を用いることで画像内の特定領域の輝度を高速に求める **Haar-like** [7] や局所領域の勾配情報に基づきエッジの特徴を掴んだ **HOG (Histogram of Oriented Gradient)** [2] が有名である．また，画像分類において **Bag-of-Feature** [1] というアプローチが近年において非常に成功している．この手法には，**SIFT (Shift Invariant Feature Transform)** [4] が用いられる．また，アニメや漫画，イラストを扱った先行研究として Lab 表色系を用いてキャラクターを特徴づける [8] がある．

3. 準　　備

3.1 物 体 検 出

物体検出とは，画像認識における 1 つの問題であり，画像中から目的の物体が含まれている領域を特定する問題である．物体検出の流れを図 1 に示す．

特定物体検出では，まず検出したい物体の学習画像から特徴量を抽出する．そして，特徴量を統計的機械学習法によって学習するにより検出したい物体の検出器を作成する．この検出器を基に，入力されたクエリ画像から検出したい物体を探し出す．実際に，画像の中から物体を探し出すにはクエリ画像の中を可変の矩形の窓を走査させる．窓を動かすごとに，指定された画像領域において特徴量を抽出し，検出器によって目的の物体か否かを判断する．

近年では，物体検出に用いる特徴量として Haar-like や HOG の出現により，人間の顔や動物の顔，あるいは車など様々な物体を検出できるようになった．

本論文では，キャラクターの顔という物体を検出するための特徴量として LBP を，統計的機械学習手法として **Ada Boost** [3] を用いている．

3.2 Local binary pattern

LBP とはテクスチャ解析で用いられる手法であり，ある 1 つの画素に注目し，その周囲に位置する画素との輝度情報の差を 2 進表現で表したものをヒストグラム化し，特徴量として記述したものである．実際には，図 2 のように注目画素と，その周囲の 8 画素との輝度情報を比べ，注目画素より周囲の画素が高ければ 1 を，注目画素より周囲の画素が低ければ 0 を割り当てる．カラー画像の場合は，グレースケール画像に変換して輝度情報を調べる．すべて割り当て終わると注目画素の周囲の画素のうち注目画素の左隣の画素から反時計回りに，割り当てられた数を連結する．最終的には注目画素の周囲の輝度情報を 8 ビットの 2 進表現で表すことができる．これを，**LBP pattern** と呼び、LBP pattern を 10 進数の値に直したものを **LBP value** と呼ぶ．

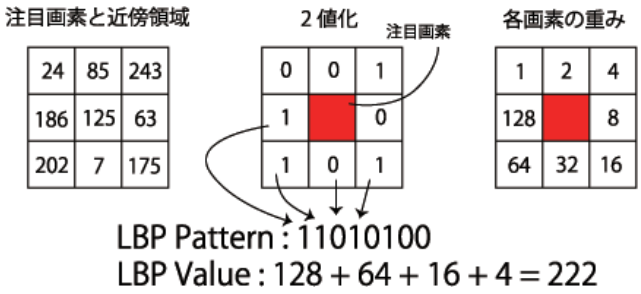

図 2　LBP の計算法

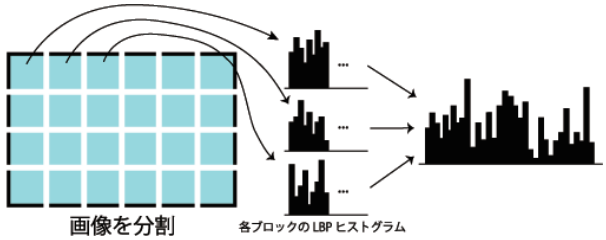

図 3　LBP の記述法

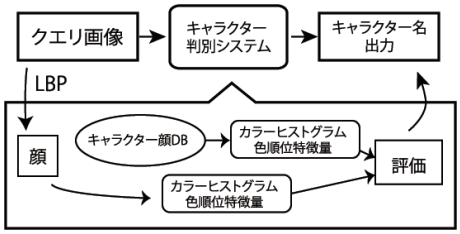

図 4　キャラクター判別システムの概要図

この操作を図 3 のように，M 個のブロックに分割した画像の各ブロック内の各画素に対し行う．ブロック毎に LBP pattern の頻度からヒストグラムを作成する．そして，すべてのブロックのヒストグラムを結合とすると $M \times 256$ 次元のベクトルとなり，これを特徴量とする．

4. 提 案 手 法

4.1 キャラクター判別問題について

先に述べたようにイラスト画像中のキャラクターは対象のキャラクターが同じであっても実世界の画像以上にアピアランスの変化が大きい．それは近年では，筆，水彩筆，クレヨンなどの画材をデジタルで再現することが可能となり，さらにクリエイター各々よっても表現の違いが存在するためである．そこで，キャラクターの持つ目や髪の色はキャラクターを分類する上で非常に重要な手掛かりとなると考え，提案システムではカラーヒストグラムと色情報に基づく特徴量 (**色順位特徴量**) を用いた．

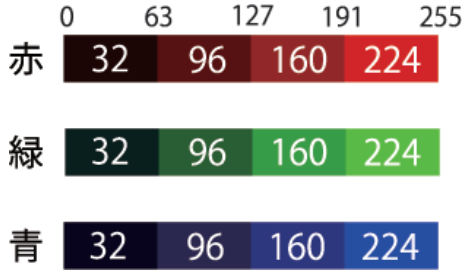

図 5　ポスタリゼーションを行い 64 色に減色する

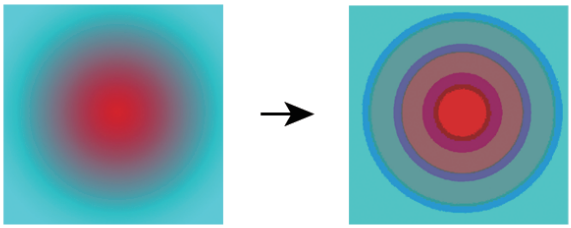

図 6　画像のポスタリゼーション

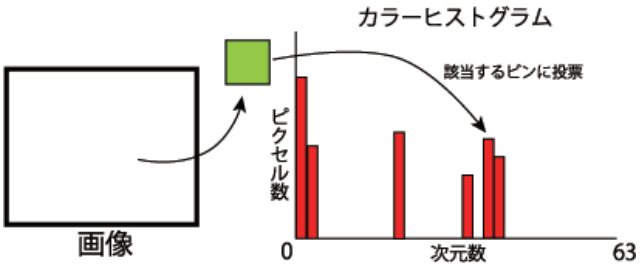

図 7　カラーヒストグラムの記述法

4.2　提案手法の概要

キャラクター判別システムは，図 4 のようにクエリとなる画像を入力すると LBP から作られたキャラクターの顔検出器からクエリ画像中の一人あるいは複数のキャラクターの顔部分を切り出す．そして，カラーヒストグラムと色順位基づく特徴量を用いてキャラクターの顔データベースからキャラクターの分類を行う．そして最終的にクエリ画像中に写っているキャラクターの名称を出力する．

4.3　カラーヒストグラム

カラーヒストグラムとは，画像分類に用いられる手法で画像中の各画素の色情報を調べ，同色の色がどの程度存在するかをヒストグラムで表現したものである．通常では RGB 画像，すなわち赤，緑，青成分の 3 トラックの画像の場合，表現できる色の種類は各色につき 256 通り存在するので $256 \times 256 \times 256 = 16777216$ 通りの色が存在する．本稿では，ポスタリゼーション (階調変更) を行い各トラックを 4 通りの色で表現することで，$4 \times 4 \times 4 = 64$ 通りの色に減色しカラーヒストグラムを 64 次元の特徴量で表現する．以下にカラーヒストグラムの構成方法について述べる．

画像中の各画素値を図 5 のように減色する．つまり各成分について値が 0〜63 なら 32，64〜127 なら 96，128〜191 なら 160，192〜255 なら 224 に値を置き換える．これにより図 6 の左側の画像をポスタリゼーションすると図 6 の右側のように

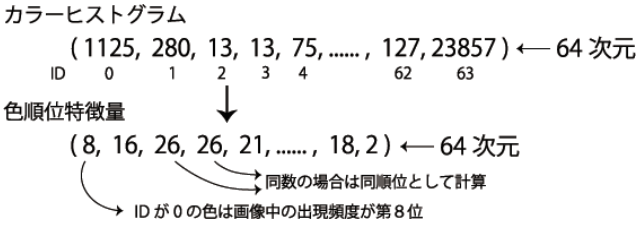

図 8　色順位特徴量の記述法

なる．

次にヒストグラムを計算していく．まず，減色された各色について色を識別するために ID を付ける．ヒストグラムの計算は図 7 のように減色された画像中の各画素について色情報を求め，ID の番号から該当するヒストグラムの要素を追加していく．

4.4　色順位特徴量

本論文では提案システムにおいてカラーヒストグラムの他に画像中に存在する代表的な色を評価するため画像中に存在する色を順位付けした特徴量を用いて画像分類を行う．

特徴量の記述法には，カラーヒストグラムの特徴量を用いる．図 8 のように，カラーヒストグラムの特徴量の各要素の値を見て大きい順に順位付けをしていく．要素数が同じものに関しては同順位として計算する．これらの操作を 64 色すべてについて行う．したがって色順位特徴量もまた 64 次元のベクトルとなる．

5.　評 価 実 験

本節では，LBP を用いた顔検出の精度，そしてカラーヒストグラムを用いたキャラクター分類のための評価法について実験し述べる．

5.1　キャラクター顔検出とその性能

キャラクター判別システムの実現のため，まず LBP を使用したキャラクター顔検出の実験とその性能について評価する．実験では，様々な人によって描かれた一人．もしくは複数人のキャラクターが描かれているイラスト画像 2635 枚を使用した．また．実験に使う検出器として lbpcascade_animeface.xml (注1) を使用した．この検出器ではキャラクターの顔を検出するために正例 20000 枚，負例 20000 枚の画像が学習に使用されており，LBP の特徴量の次元は 256 次元となっている．

また，性能を評価する指標として **F 値 (F-measure)** を用いた．F 値は，**精度 (precision)** と呼ばれる正解と判断されたもののうち正解している割合と，**再現率 (recall)** と呼ばれる実際に正解であるもののうち検出された割合によって求められる．

R を正解と検出されたもののうち実際に正解である数，N を正解，不正解問わず検出された数，C を実際の正解数とすると，精度，再現率および F 値は以下の式によって求められる．

$$precicion = \frac{R}{N}$$

(注1)：lbpcascade_animeface.xml - http://anime.udp.jp/data/lbpcascade_animeface.xml

表 1　キャラクター顔検出の性能

指標	結果
精度	95.1%
再現率	51.7%
F 値	67.0%

$$recall = \frac{R}{C}$$

$$F - measure = \frac{2R}{N + C}$$

結果は表 1 のようになった．表精度は，約 95% という非常に高い結果となった．これは，誤検出が少ないことを意味しており提案するキャラクター判別システムにおいて非常に望ましい結果であると言える．精度に反して，再現率は約 52% と半分程度にとどまった．これは，イラスト画像が実世界の画像以上にアピアランスの変化が大きいことが原因であると考えられる．

5.2　キャラクター分類の性能

カラーヒストグラム，および色順位特徴量の性能についての実験し評価を行う．

実験には animeface-character-dataset (注2) を使用した．animeface character dataset は計 14990 枚の 175 のキャラクターの顔のイラスト画像からなっており，これらの画像からカラーヒストグラムと色順位特徴量を抽出しそれをデータベースとした．

以下に実験で使用した類似度について述べる．

カラーヒストグラムではヒストグラムの距離を求め類似度を決める **Histogram Intersection** [6] を用いた．

2 つのヒストグラムをそれぞれ H_1，H_2 とし，ヒストグラム H の i 番目の要素を $H[i]$ とすると Histogram Intersection は以下の式で表せられる．

$$Histgram\ Intersection = \frac{\sum_{i=0}^{63} min(H_1[i], H_2[i])}{\sum_{i=0}^{63} H_1[i]}$$

Histogram Intersection は 0.0 から 1.0 までの値をとり，値が高いほど類似度が大きいことを意味する．また，値が 1.0 のときヒストグラムは一致している．

色順位特徴量ではクエリ画像に存在する第 10 番目までの色がデータベース内の画像において何番目に現れるかを調べた 2 つの類似度を用いた．

クエリ画像の特徴量を C_1，データベース内の画像の特徴量を C_2 とし，ヒストグラム C の第 i 番目の色を $R_C[i]$ とすると類似度 sim_1，sim_2 以下の式で表せられる．

$$\text{方法 1：}\ sim_1(C_1, C_2) = \sum_{i=0}^{9} (63 - |R_{C_1}[i] - R_{C_2}[i]|)$$

$$\text{方法 2：}\ sim_2(C_1, C_2) = \sum_{i=0}^{9} (63 - R_{C_1}[R_{C_2}[i]])$$

方法 1 ではクエリ画像の第 10 番目までに現れる色がデータベース内の画像の画像において何番目に現れるかを調べて差の絶対値を類似度から類似度を測った．方法 2 ではクエリ画像の第 10 番目までに現れる色がデータベース内の画像の画像において何番目に現れるかを調べてデータベース内の順位から類似度を測った．

(注2)：nimeface-character-dataset - http://www.nurs.or.jp/~nagadomi/animeface-character-dataset/

表 2　キャラクター分類の性能

類似度	結果
Histogram Intersection	39 / 100
方法 1	22 / 100
方法 2	37 / 100

実験では，クエリ画像 100 枚から抽出したカラーヒストグラムおよび色順位特徴量とデータベース内の各画像の特徴量から類似度を計算して 175 のキャラクターそれぞれにおいて類似度の平均を求めキャラクター名で順位付けをし，クエリ画像と同名のキャラクターが第 10 番目以内に現れていれば分類成功とした．

実験結果は表 2 となった．Histogram Intersection は 39%，方法 1 は 22% そして方法 2 では 37% と良く分類できなかった．原因は animeface-character-dataset はキャラクターの顔が人の手によって切り出された切り出されたものであるため，各々の画像におけるキャラクター以外の背景領域の大きさが異なるため色情報のみを用いるカラーヒストグラムや色順位特徴量では大きく背景の影響受けてしまうことが考えられる．そしてカラーヒストグラムと色順位特徴量では人間の視覚上似たような色でも異なる色と評価されてしまう可能性があることなどが考えられる．

6.　おわりに

本論文では，LBP を用いたキャラクター顔検出とカラーヒストグラムと色順位特徴量を用いたキャラクター分類を組み合わせたキャラクター判別システムを提案した．LBP を用いたキャラクター顔検出実験ではある程度提案システムに有効であろう結果が得られたが，キャラクター分類実験では良い結果が得られなかった．今後は，顔検出の再現度を上げること，そしてキャラクター分類に有効な方法を見つけることが重要なテーマとなる．

文　献

[1] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In *Workshop on statistical learning in computer vision, ECCV*, Vol. 1, p. 22, 2004.

[2] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, Vol. 1, pp. 886–893. IEEE, 2005.

[3] Y. Freund and R. Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In *Computational learning theory*, pp. 23–37. Springer, 1995.

[4] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, Vol. 60, No. 2, pp. 91–110, 2004.

[5] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on fea-

tured distributions. *Pattern recognition*, Vol. 29, No. 1, pp. 51–59, 1996.

[6] M.J. Swain and D.H. Ballard. Color indexing. *International journal of computer vision*, Vol. 7, No. 1, pp. 11–32, 1991.

[7] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, Vol. 1, pp. I–511. IEEE, 2001.

[8] 高山耕平, J.Henry, 西田友是. 特徴抽出によるアニメキャラクターの顔認識とキャラクター検索. Visual Computing/グラフィクスと CAD 合同シンポジウム予稿集 (CD-ROM), 2012.