

人物画像における衣服領域の推定と色情報の抽出

Clothing Region Estimation and Color Information Extraction of Street Snap

松田 直樹[†] 横田 翔皇[†] 清水 郁子[†] 柳井 啓司[‡] 柴原 一友[§] 藤本 浩司[§]

Naoki Matsuda Shoko Yokota Ikuko Shimizu Keiji Yanai Kazutomo Shibahara Koji Fujimoto

1. まえがき

近年, SNS, 画像共有サイトやブログの発達により, 全世界の誰しもが簡単に個人写真などをアップロードすることができるようになった. これらの写真には衣服の着こなしを撮影したファッション画像も大量に含まれている. なかでも, ストリートスナップは街で見かける若者たちの服装や着こなしを撮影したスナップ写真のことであり, ファッションの流行のトレンドをいち早く反映しているものであると考えられる.

我々は流行を把握するために膨大な数の画像情報から衣服の分析を行い, ファッション情報を抽出することを目指す. ファッション情報といっても衣服の形や, 衣服の種類 (カーディガン, ワンピース等) などの情報があるが, 本研究では衣服の色情報について着目する. つまり本研究ではファッション画像を分析し, 衣服の色を抽出することを目的とする. 本稿では, ファッション画像を与え画像中の人物の衣服が上半身は白色の衣服, 下半身は黒色の衣服というように文章でファッション情報を得るものとする. これによって画像情報からテキスト情報を抽出することができるようになり, ファッション情報について検索や分析がスムーズに行われることが期待される.

2. 関連研究

人物画像における色情報を扱った先行研究としていくつかの研究が行われている. Kang らはファッション推薦システムを提案している [2]. この手法では, 衣服の色をファッションの一要素と捉えて使用しているが, 画像情報から抽出しているわけではなく元々与えられているものとしている. Gallagher らの研究では衣服の色情報から同一人物を認識している [1]. 同一の日に撮影された複数の写真において同一人物は同じ服を着ているものとして個人の特定を行っている. この手法では衣服の色情報を使用しているものの, 衣服が何色なのかという点には着目していない. 一方で, 相田らは画像中の衣服の色を認識する手法を示している [3]. この手法では, 画像を領域分割し, 衣服領域を決定してその色を認識している. しかし, 顔検出の頑健性や色の照明による影響が考慮されていなかった.

3. 提案手法

本研究では人物画像を入力すると画像中の人物の上半身, 下半身それぞれに着用している色名を返す手法を提案する. なお, 入力されるストリートスナップは一人の人物が映っているものに限定する.

本研究で提案する色情報抽出の全体の概要を以下に示す.

1. 画像入力: 処理を行うストリートスナップの画像を与える.
2. 顔検出: 画像中の顔を検出する. このとき, 色や位置に制約を与え, 誤検出を低減する.
3. 色補正: 色温度に基づき, 照明の影響があると判断された画像に色補正を行う.
4. 領域分割: 顔の位置から衣服の位置を推定する. GrabCut による領域分割を行い, 失敗した場合には矩形領域を指定する.
5. 色情報分析: 衣服領域の上半身, 下半身それぞれに対し, 色を抽出する.

4. 色と位置に基づく頑健な顔検出

本研究では人物の位置, 衣服領域を推定するために入力されたストリートスナップに対して顔検出を行う.

まず, Haar-like 特徴を用いたブースティングによる分類器 [4] を利用し, 複数の識別器を用いて顔候補領域の検出を行う. そして, 顔の誤検出を減らすために肌色検出を行う. 顔候補領域の肌色割合が閾値以下なら顔でないと判断してその顔候補領域を棄却する.

さらに, ストリートスナップは人物の服装を撮影したものであるため全身が撮影されているものが一般的である. その条件から顔の位置をおおまかに推定できる. そのために顔位置に制約を与え, 誤検出を減らす. 具体的には画像の最下部, 右上, 左上に顔が存在する場合を除外する. 除外箇所を図 1 に示す.

最後に, 複数の識別器による検出の整合性を確認する. 顔領域が画像全体で一つしか検出されない場合はその領域を顔とする. 二つ以上顔候補領域が検出された場合には, 二つ以上の識別器で検出した領域のみを顔とする.

5. 色補正

5.1. HSV 色空間における色の分類

本手法では, HSV 色相系に画素値を変換して色抽出を行う. 抽出する色は一般的な色とされる 12 色とする. また本研究では人物領域を抽出するため肌色も追加して, 使用する色は以下の色となる.

[†]東京農工大学工学府, 東京都小金井市中町 2-24-16
Tokyo University of Agri. and Tech. 2-24-16 Naka-cho,
Koganei-shi, Tokyo 184-8588 Japan
[‡]電気通信大学, 〒182-8585 東京都調布市調布ヶ丘 1-5-1
The University of Electro-Communications 1-5-1 Chohugaoka,
Chohu, Tokyo
182-8585 Japan
[§]テンソル・コンサルティング株式会社, 東京都千代田区麹町
2-10-1
Tensor Consulting Co.Ltd. 2-10-1 Kouji, Chiyoda, Tokyo
102-0083 Japan

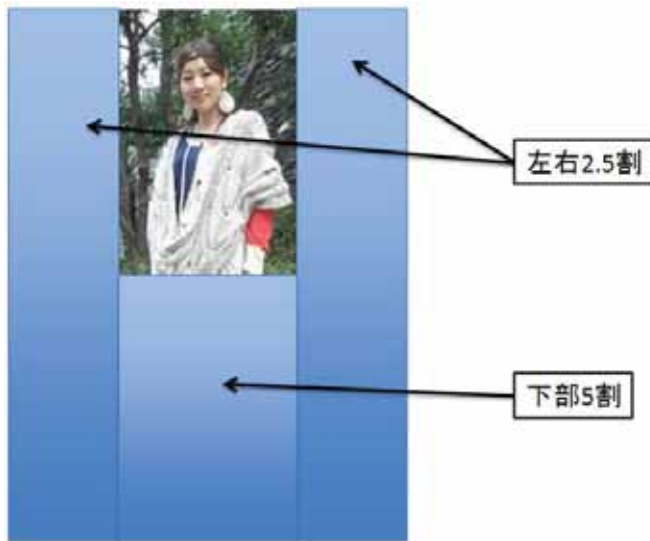


図 1: 顔位置による顔候補領域の除外箇所

表 1: 使用する 13 色一覧

赤色	橙色	黄色	緑色	水色	青色	紫色
桃色	白色	灰色	黒色	茶色	肌色	

表 1 で示した色について, HSV の値と色の対応を手動で与え, 色を分類する. まず S(Saturation: 彩度) と V(Value of Brightness: 明度) の値からまず無彩色 (黒色, 白色, 灰色) を求め, その後 H(Hue: 色相) の値により各色に分類するものとする.

5.2. 照明の推定と色補正

与えられたストリートスナップは多様な照明条件下で撮影されている. そのため照明の影響が考えられる状況でそのまま画素値から色を分類してしまうと, 人間が直感的に認識する色と異なる色が抽出される. 人間が認識する色とできるだけ近い色名を抽出するために, 色補正を行う.

与えられたストリートスナップの色補正を行うために, 色温度に着目する. 色温度とは色を温度で表したものであり, 赤っぽい色ほど低い温度, 青っぽい色ほど高い温度で表記される. 本来, 色温度とは画像に対する光源の色尺度であるので画像中から一意に求めることができない. そこで画像中の画素値から疑似色温度として疑似的に色温度を求めるものとする.

色温度を推定し, 色補正を行う手順を以下に説明する.

まず, RGB 表色から XYZ 表色へ変換する. RGB 表色では人間が認識する色の全ての色を表現することができず, 無理やり表現しようとする RGB の値がマイナスになってしまう. この問題点を解消するために XYZ 表色に変換する.

次に, XYZ 表色から Y_{xy} 表色へ変換し, 色温度を算出する. XYZ 表色からグラフに示す場合, 軸が三つ必要になってしまう. そこで X, Y, Z それぞれが合計値に占めている割合を小文字の x, y, z でそれぞれ表

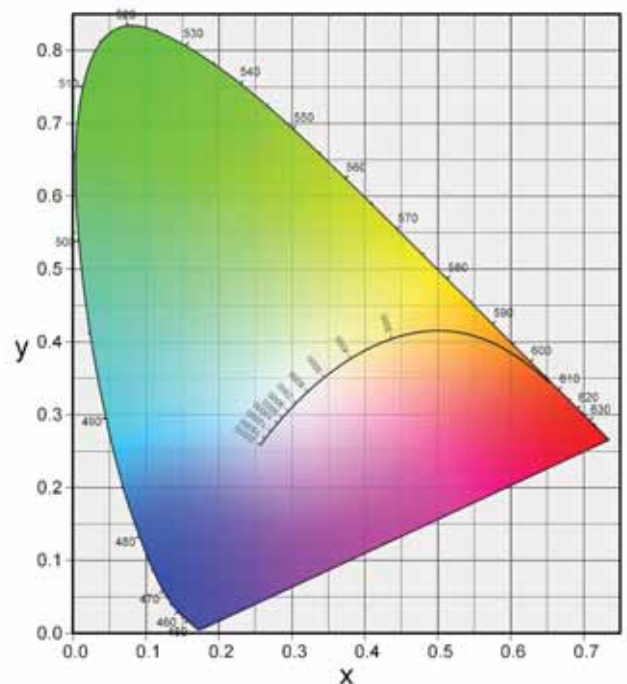


図 2: xy 色度図

すことによって, 二つの軸から色を表すことができる. Y_{xy} 表色から等色温度線を参照すると色温度を求めることができる. 参照する等色温度線を図 2 に示す.

そして, 色温度から逆色温度を算出する. 色温度は数値の差が色の違いに比例していないので, 色温度の逆数を取り 10 の六乗をかけたものを逆色温度として求める. 単位はミレッドで表される. これにより定量的な色温度の計量が可能になる.

最後に色補正を行う. 求められた逆色温度が閾値以上の場合に, 色補正を行う. しかし, 画像全体に色補正をかけると照明があたっていない領域にも補正をかけてしまうので, 疑似色温度が閾値以上の画素のみに色補正をかけるものとする.

5.3. 領域分割

本研究では人物領域とそれ以外の領域分割の手法として GrabCut を用いる [5]. GrabCut は二つのエネルギー関数を最小化することによって領域分割を行う. 二つのエネルギー関数とは隣接画素との差, 背景・前景の確率モデルである.

具体的には, 顔検出で得られた顔座標と顔半径から人物を含むと考えられる十分に大きな矩形領域を定め, 矩形領域内を前景に矩形領域外を背景としてガウス混合分布を用いて確率モデルを作成し, 領域分割を行う.

背景と服装のテクスチャが類似している場合などに, GrabCut による領域分割がうまくいかないことがある. そのような場合は分割結果が明らかに失敗であると判定できるため, 顔座標, 顔半径より上半身と下半身それぞれの領域を矩形で推定する.



図 3: 色補正の実行結果

5.4. 色情報分析

領域分割で決定された上半身領域，下半身領域内の画素を全走査して各色画素数を数えることにより色抽出を行う。抽出された結果，最も多い色を上半身，下半身ごとに第一要素とする。

6. 実験

6.1. 顔検出の実験結果

ファッションサイトから収集した 10000 枚のストリートスナップに対して実験を行った。顔検出の評価実験では OpenCV で提供されている顔検出と提案手法を比較して実験を行う。

表 2: 顔検出の精度

	正検出	誤検出	検出率	適合率
OpenCV	3763 枚	2215 枚	59.8 %	62.9 %
提案手法	4877 枚	164 枚	50.9 %	96.7 %

OpenCV のみを使用した場合とそれに様々な制約を加えた提案手法を比較すると提案手法では検出率は下がっているものの，適合率は大幅に向上している。Web 上には大量の画像が存在しているので，本研究では顔検出における検出率を上げるより適合率を上げることにより誤検出を減らすことが重要である。本手法はそのような用途に適しているといえる。

6.2. 色補正の実行結果

色補正の実行結果を図 3 に示す。図 3 は左から元画像，元画像に対して減色処理を実行した結果，色補正を行った後に減色処理を実行した結果をそれぞれ表す。

図 3 では，白いコートを来た女性が映っている。電球色の照明が当たっているため，単純に画素の明るさから色を判定するとオレンジ色と判定されているが，照明の影響を低減するため色補正を行うことにより，上半身の衣服の色が白色になっていることが分かる。

6.3. 色分析の実行結果

領域分割と最終的な色分析の結果の例を図 4 に示す。入力されたストリートスナップに対し，上半身と下半身の衣服の色をそれぞれ抽出できていることが分かる。

さらに，ファッションサイトから収集した 10000 枚のストリートスナップに対し，顔検出が成功した画像に



図 4: 色分析の実行結果例

対して色を分析する実験を行った。本手法によって抽出した第一要素の色が正しいかを手動で確認し，一致率を調べた。上半身，下半身ごとにランダムに選んだ 250 枚のストリートスナップに対して求めた。約 70% 程度が一致していることが分かった。

表 3: 色抽出の精度

	一致数	不一致数	一致率
上半身	185 枚	65 枚	74.0 %
下半身	176 枚	74 枚	70.4 %

7. おわりに

本研究では，ストリートスナップを用いて人物画像における衣服の色分析に関する研究を行った。そのために顔検出，色補正，領域分割，色情報分析について述べた。本研究によって，大量のファッション画像から色情報の自動抽出を行うことができたのでファッション流行の分析に役立つと思われる。

参考文献

- [1] A. Gallagher, T. Chen, Multi-Image Graph Cut Clothing Segmentation for Recognizing People, IEEE Conference on Computer Vision and Pattern Recognition 2008
- [2] Hanhoon Kang, Seong Joon Yoo, SVM and Collaborative Filtering-based Prediction of User Preference for Digital Fashion Recommendation Systems, IEICE Transactions on Information and Systems E90-D-12, pp.2100-2102, 2007,
- [3] 相田優, 柳井啓司, 柴原一友, 藤本浩司, “服飾画像マイニングのための衣類領域からの色情報抽出”, 電子情報通信学会技術報告, vol. 111, no. 478, IE2011-173, pp. 235-240, 2012

- [4] Paul Viola, Michael J. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision, vol.57, no.2, pp.137–154, 2004
- [5] Carsten Rother , Vladimir Kolmogorov , Andrew Blake , "GrabCut" — Interactive Foreground Extraction using Iterated Graph Cuts , ACM Transactions on Graphics (TOG) , 2004