

髪の特徴に基づく顔画像の印象類似検索

藤 賢大¹ 福里 司² 増田 太郎² 平井 辰典¹ 森島 繁生³

概要: Web上の人物画像検索は、名前だけを知っている人物の顔を調べたいときや、同一人物の他の画像を見てみたいといった状況で有用である。人物を検索するには、その人物であるかどうかはもちろんのこと、外見の印象、特に髪の影響は重要な要素である。しかし、既存の人物画像検索システムの多くはテキスト情報に基づいて検索しているため、人物の外見の特徴や印象を十分に考慮できていない。そこで、本研究では人物の外見の印象を大きく左右する髪に注目した画像検索手法を提案する。髪の特徴を定量化する手法は今までに提案されているが、人間がどういった髪特徴に注目するかについて明確にされていない。そこで本研究では、ユーザの主観評価の結果に基づいて髪特徴量を組み合わせて評価関数を設計する。それを基にユーザの注目する髪の特徴を考慮した人物画像検索手法を提案する。

1. はじめに

近年、カメラの小型化・高性能化によりスマートフォンで写真が撮られることが多くなった。さらには、データ通信の高速化やSNSの利用者の増加により、手軽にWeb上に画像をアップロードできるようになった。そのため、Web上の画像コンテンツ数は爆発的に増大しており、今後増え続けていくと予想される。また、そのような莫大な画像の中から効率的に検索するシステムやアプリが注目されている。

既存の画像検索システムとして、Web上の画像検索エンジンであるGoogle画像検索[1]などが挙げられる。これはWebページに含まれるキーワードやリンク情報に基づいた検索方法である。これにより検索に要する時間が短くなり、手軽に利用することができる。このようなWeb上の画像検索は、ユーザが見たことはないが名前だけを知っている人物/物/場所を興味で見たいとき、同じ人物/物/場所の他の画像を見てみたいといった状況のときに利用されている。その中でも、特にWeb上では人物の画像が検索されていることが多い。特にYahoo画像検索やgoo検索でよく検索されているキーワードは、人物名と言われている[2][3]。さらに、人物を検索するにあたり、外見から判断する印象、特に髪に対する印象は最も重要な要素といえる。その理由として、髪型や髪色は、個人を認識するため

にもっとも重要な役割を果たしていることが、心理学的に証明されている[4][5][6]。また、髪色や髪型は時代の流行を表しており、美容院やヘアカタログの存在から検索技術が必要とされているといえる。しかし、現状ではGoogle画像検索などのようにテキスト情報のみでは、髪などの人物の外見の特徴を考慮することはできない。それは、人物の髪の特徴がテキスト情報として画像に付与されていることは少なく、また髪の特徴を表す適切なキーワード自体が存在しないことも多いためである。そこで、本研究では、画像情報を用いて外見の印象を決定づける髪の特徴を考慮した画像検索を目的とする。

従来研究として、人物の画像から髪の長さや色などの特徴を定量化し、類似度を計算する手法が提案されている。しかし、髪型と髪色を組み合わせた場合については考察されておらず、髪全体としての印象の定量化はできていない。

そこで本研究では、人物画像から髪色や髪型等の特徴を取得し、髪色や髪型の特徴の重要度について考察する。重要度については、ユーザ(認識者)の感性によって変動があると考えられるため、主観評価実験の結果を学習した上で推定を行う。本技術により、Web上でヘアカタログのように髪の検索が実現可能となり、さらにユーザの検索効率を向上させるためにインタラクティブな機能を実装する。

2. 関連研究

人物の髪を考慮した検索をするにあたって、髪の領域の情報が必要となる。また、髪は人の顔とは大きく異なり、色や形状による個人差が大きく、髪領域抽出手法やその応用手法が数多く提案されている。Chaiらは人物画像1枚

¹ 早稲田大学
Waseda University

² 早稲田大学/JST
Waseda University/JST

³ 早稲田大学理工学術院総合研究所/JST
Waseda Research Institute for Science and Engineering/JST

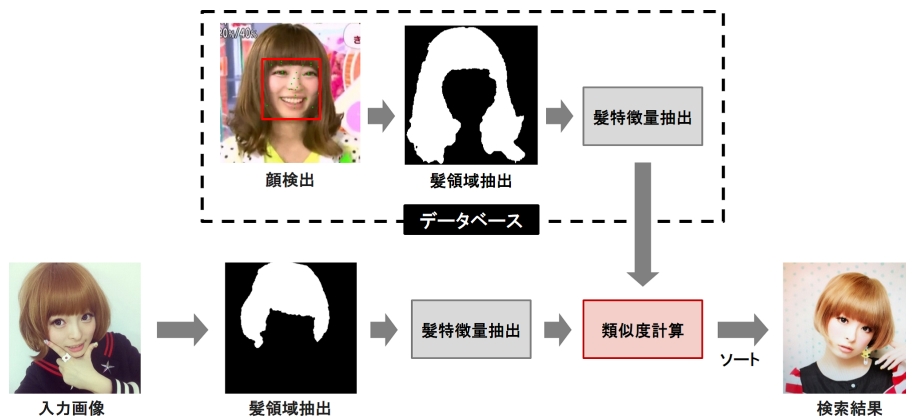


図 1 提案手法のフロー

から髪モデリングを行う手法を提案している [7]. 紐ベースの 3 次元頭髪モデルを生成することにより, 顔向きを変えることや人物画像の髪の入替えが可能となっている. そのため髪領域を正確に抽出することができていると考えられるが, 髪領域のアノテーションはユーザが与えた正解ラベルをもとに行っているため, 自動で髪領域の抽出を行うことができない. また, 色情報を用いた領域成長による髪領域の自動抽出法 [8] や, 色と周波数成分を用いた髪領域の自動抽出法 [9] が存在するが, 色や形状の個人差, 撮影環境の影響を大きく受けてしまう.

髪の特徴を考慮した研究としては, 髪色を顔の属性の 1 つとして顔画像検索する研究 [10] や, 髪型で顔画像検索を行う研究もある [11]. また, 髪の特徴の定量化に関する研究として, Yacoob らは 8 つの特徴量を提案し, 各特徴量における画像間の類似度を可視化した [8]. しかし, 提案された特徴量自体が外見的印象に関連するかどうか, 個人を認識する際に上記の特徴量のどれを優先しているか (特徴量の重要度) の言及がされておらず, 髪検索に適用した際, ユーザが満足する画像を見つけることは困難といえる.

以上のように, 従来の髪の研究では, 髪領域の抽出や髪特徴の数値化を重視しているため, 人間がどのような特徴に着目して髪を認識しているのかについては未考察である. そこで本研究では, 各ユーザが主観的にどのような髪が似ていると認識しているかを学習し, 関連性の高さと髪の特徴量を組み合わせた評価関数をもとに, 人物画像の検索手法を提案する. また髪領域の抽出手法や髪の特徴量の定量化にも着手した.

3. 提案手法

本手法では, 検索対象とする全ての顔画像 (以降, データベース画像群と呼ぶ) に対して髪領域を抽出し, その髪領域内の特徴を抽出する. 一方, ユーザが入力した顔画像に対しても同様に髪特徴を抽出し, データベース画像群との類似度を算出する. その値をもとに, 入力した顔画像の髪印象と類似した顔画像の検索を行う.

提案手法のフローを図 1 に示す. まず, 事前計算としてデータベース画像群の髪特徴を抽出しておく. その過程で, データベースの全画像に対し, 顔向きと大きさの正規化を行う. さらに, 正規化された顔画像に対して髪領域を抽出し, その領域内で髪特徴量を計算する. 次に, ユーザの入力画像の髪特徴を同様に抽出し, データベース画像群との類似度を計算する. 最後に, 類似度をもとにデータベース画像群を提示する検索システムを実現する. 以下に提案手法の詳細を述べる.

3.1 顔画像の正規化

顔画像の正規化および顔領域と髪領域の位置を推定するために顔特徴点検出を行う. 顔特徴点検出は, Active Structure Appearance Model (ASAM) による顔領域のグローバルなフィッティングと顔器官毎のローカルなフィッティングを統合させた階層的なフィッティングによる顔検出方法を用いる [12]. ASAM では表情変化や顔向きの変動に対して誤検出を起しやすいためという問題があった. そこで, Irie らは ASAM に顔器官ごとのローカルなフィッティングを組み合わせることによって, 表情変化や顔向きの変動に対してロバストな顔特徴点検出を実現している. この手法によって検出された顔特徴点の配置を図 2(a) に示す.

入力やデータベースの画像はそれぞれ異なる環境で撮影されたものであり, 顔の大きさや顔向き等が異なる. そこで, 全ての顔画像から出来る限り同じような条件のもとで特徴量を抽出するために, 顔画像の大きさと顔向きの正規化を行う. 図 2 のように両目内側の 2 点を結んだ線が水平かつ長さが 50 ピクセルになるようにアフィン変換を行う.

3.2 髪領域抽出

本節では, 人物の髪領域を自動抽出する手法について述べる. 第一段階として, 画像全体から髪領域の候補を推定するため, 細井らの手法を用いる [13]. まず 3.1 節で検出された顔特徴点から図 (3) のように肌と髪の矩形領域 (以降, 肌サンプル領域, 髪サンプル領域と呼ぶ) を取得する.

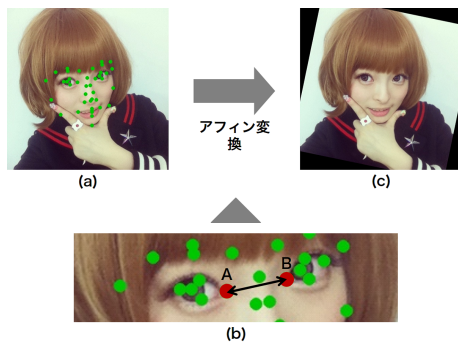


図 2 (a) 検出した顔特徴点, (b) 両目内側の 2 点を結んだ直線, (c) 顔向き正規化画像

取得した肌サンプル領域で各画素の RGB 値を $L^*a^*b^*$ 表色系に変換し, 領域内の $L^*a^*b^*$ 値の平均・標準偏差をそれぞれ算出する. その平均・標準偏差をもとに全ての画素に対して式 (1) を用いて肌色強度 I を計算する. ただし, $I_{i,j}$ は (i,j) における色強度, $L_{i,j}$, $a_{i,j}$, $b_{i,j}$ は (i,j) における $L^*a^*b^*$ 値, μ_L , μ_a , μ_b は領域内の $L^*a^*b^*$ 値の平均, σ_L , σ_a , σ_b は領域内の $L^*a^*b^*$ 値の標準偏差である.

$$I_{i,j} = \exp \left[- \left\{ \left(\frac{L_{i,j} - \mu_L}{\sigma_L} \right)^2 + \left(\frac{a_{i,j} - \mu_a}{\sigma_a} \right)^2 + \left(\frac{b_{i,j} - \mu_b}{\sigma_b} \right)^2 \right\} \right] \quad (1)$$

各画素に対して色強度を計算すると, 図 4(a) のように肌色強度分布が得られる. この肌色強度分布で閾値処理を行い, ラベリングをすることによっておおよその肌領域が得られる (図 4(b)(c)).

髪色成分も同様の計算方法で推定する. 図 3(b) の髪サンプル領域から肌領域を除外した領域で, 各画素の RGB 値を $L^*a^*b^*$ 表色系に変換し, 領域内の $L^*a^*b^*$ 値の平均・標準偏差をそれぞれ算出する. その平均・標準偏差をもとに全ての画素に対して式 (1) を用いて髪色強度 I を計算する. 各画素に対して色強度を計算すると, 図 5(a) のように髪色強度分布が得られる. この髪色強度分布で閾値処理を行い, クロージング処理とラベリング処理を行うことによっておおよその髪領域が得られる (図 5(b)(c)).

次に髪領域の候補から, 最適な髪領域を推定する手法として Graph-Cut Segmentation を行う [14]. Graph-Cut 法は, エネルギー最小化問題を解くことで画像を領域分割する手法であり, 指定した前景・背景シード (正解ラベル) をもとに領域分割を行う. 前景・背景シードと近い色の画素はそれぞれ前景・背景となるように, なおかつ隣接画素の色の差が大きい部分に領域の境界となるように, 最適な領域分割が計算される. この Graph-Cut 法の前景・背景シードとして, これまでに取得した髪領域・肌領域を与える. それにより, 髪領域が前景となり, 肌領域が背景となるような領域分割が行われる. その領域分割された結果に対してラベリング処理を行い, 最終的な髪領域抽出結果を図 6 に示す.

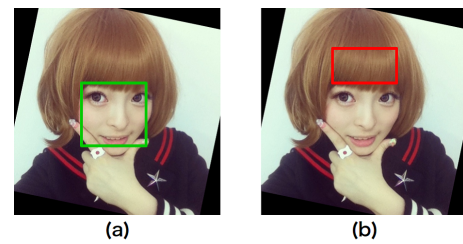


図 3 (a) 肌サンプル領域, (b) 髪サンプル領域



図 4 (a) 肌色強度分布, (b) 2 値化処理後, (c) クロージング処理とラベリング処理後

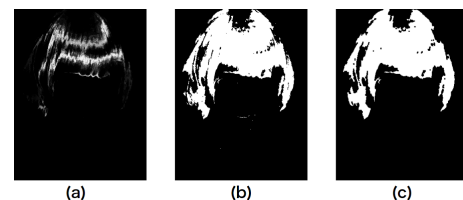


図 5 (a) 髪色強度分布, (b) 2 値化処理後, (c) クロージング処理とラベリング処理後



図 6 (a) 抽出された髪領域, (b) 髪マスク画像

3.3 髪特徴量

本章では髪特徴量の計算方法について述べる. 特徴量の計算は髪領域 (図 6(b)) において行う. Yacob らは, 1 枚の顔画像から色/分け目/ボリューム/長さ/表面積/対称性/髪形状/テクスチャの 8 つの特徴量を抽出している. これらの特徴量のうち, (1) 色, (2) 長さ, (3) ボリューム, (4) テクスチャの 4 つを数値化する. 残りの 4 つの特徴量に関しては, 分け目や髪の形状など, 選択した特徴量と類似しており, 本稿の目的である外見的印象を考慮した画像検索において必要ないと判断した.

まず, 髪色の特徴量の計算方法について述べる. 髪色の特徴量として, $L^*a^*b^*$ カラーヒストグラムを用いる. $L^*a^*b^*$ 表色系では色を L , a , b の 3 値で表し, L が明度指数, a と b がクロマティックネス指数と呼ばれ色相と彩度を同時に表している. 計算方法として, まず画像の各ピクセル

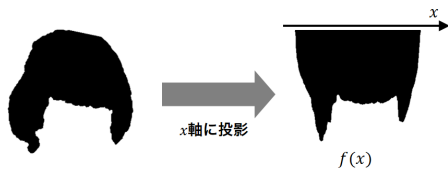


図 7 長さ特徴量

ルの RGB 値を $L^*a^*b^*$ 値に変換する. そして, L, a, b の値をそれぞれ 5 等分, 10 等分, 10 等分に量子化して 500 色に減色し, 画像の髪領域内で各色のピクセル数を数え, 500 次元の色特徴量とする.

次に, 髪の長さの特徴量の計算方法について述べる. 髪の長さは, まず図 7 のように髪領域を x 軸に投影した. これにより, 髪の長さのヒストグラムを作ることができ, このヒストグラムの x 軸をビンとして y 軸方向のピクセル数を数え, その値を長さの特徴量として用いた. なお, ヒストグラムの比較が可能となるように, 両目の中心の x 座標を基準に各ビンに値を格納した. また, この特徴量の次元は x 軸のサンプリング数 n_x となる (今回は経験的に $n_x = 600$ とした).

続いて, 髪のボリュームの特徴量の計算方法について述べる. 髪のボリュームは, 図 8 のように両目の中心を中心に左右に分割し, 髪領域を y 軸に投影して, 左右でボリュームのヒストグラムを作成した. それぞれのヒストグラムの y 軸をビンとして x 軸方向のピクセル数を数え, その値をボリュームの特徴量として用いた. 左右に分割することによって, Yacoob らが提案する髪形状と分け目を同時に考慮することができる. なお, ヒストグラムの比較が可能となるように, 顎の特徴点の y 座標を基準に各ビンに値を格納した. また, この特徴量の次元は y 軸のサンプリング数 n_y となる (今回は経験的に $n_y = 600$ とした).

最後に, 髪のテクスチャの特徴量の計算方法について述べる. 髪のテクスチャは, LBP 特徴量のヒストグラムを用いた. LBP 特徴量は, 照明変化にロバストで, 局所領域の輝度値パターンを用いてテクスチャを表現する特徴量である. 3×3 ピクセル領域の中央画素値と近傍 8 画素値の大小を比較し, 大きければ “1” を, そうでなければ “0” を割り振る. これによって得られた 8 桁の 2 進数を 10 進数に変換することにより, 256 種類の輝度値パターンを表現することができる. 今回は髪領域で LBP 特徴量を各画素ごとに計算し, そのパターンの頻度をテクスチャの特徴量とした.

3.4 類似度計算

ここまでで得られた髪特徴量をもとに類似度を計算する. 入力画像とデータベースの類似度計算方法として, まず入力画像とデータベースの各特徴量の距離を計算する. 今回用いた特徴量はカラーヒストグラム, ピクセル数, LBP ヒ

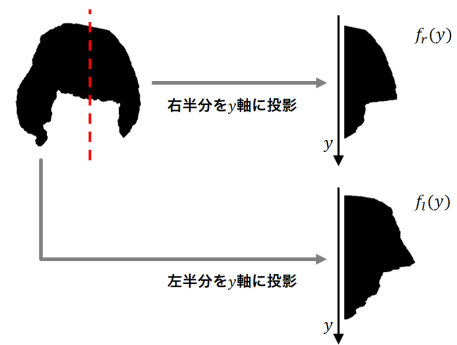


図 8 ボリューム特徴量

ストグラムとスケールが異なるデータを扱っているため, 正規化を行いスケールを合わせる必要がある. そこで, 距離尺度はデータベースの共分散行列によって計算されるマハラノビス距離を用いた. 共分散行列が Σ であるようなデータベースの特徴量ベクトル \vec{x} と入力画像の特徴量ベクトル \vec{y} のマハラノビス距離 d_* は以下のように表される.

$$d_* = \sqrt{(\vec{x} - \vec{y})^T \Sigma^{-1} (\vec{x} - \vec{y})} \quad (2)$$

ここで, 式 (3) のように各特徴量の距離の重み付け線形和を計算する. ただし, c は色, l は長さ, v はボリューム, t はテクスチャ, w はそれぞれの重み付け係数を表す.

$$d = w_c d_c + w_l d_l + w_v d_v + w_t d_t \quad (3)$$

この線形和 d が小さいものほど類似度が高いとし, データベース画像群から d が小さい順に画像を提示する.

3.5 髪特徴の重み付け

本節では, 3.4 節の重み付け係数の推定方法について述べる. 最適なパラメータは個人の感性によって変動すると考えられるため, ユーザの主観評価を学習することによって行う. ランダムに選択した顔画像 n 枚を顔画像 X と比較し, n 枚の顔画像に対してユーザの主観で髪の類似順位をつける (今回は $n=10$ とした). ここで, 3.4 節で求めた各特徴量の距離について小さい方から順位をつける. その順位とユーザが評価した順位との間のスピアマン順位相関係数を求め, 各髪特徴の重み付け係数とした. また, 相関係数の値の範囲は -1 から 1 であるが, マイナスの値のときは相関がないものとし, 重み付け係数を 0 とした.

4. 実験

4.1 実験結果

本手法を用いて髪の特徴を考慮した画像を検索する実験を行った. 実験に用いるデータとして, Google で「きゃりーぱみゅぱみゅ」をテキスト検索して得られた画像 100 枚を用いた. 比較手法として Yacoob らと同様に, 定量化した髪の各特徴量で検索した場合, 重み付け係数を一律に



図 9 入力画像 (a) に対する検索結果の上位 (色)



図 15 入力画像 (b) に対する検索結果の上位 (色)



図 10 入力画像 (a) に対する検索結果の上位 (長さ)



図 16 入力画像 (b) に対する検索結果の上位 (長さ)



図 11 入力画像 (a) に対する検索結果の上位 (ボリューム)



図 17 入力画像 (b) に対する検索結果の上位 (ボリューム)



図 12 入力画像 (a) に対する検索結果の上位 (テクスチャ)



図 18 入力画像 (b) に対する検索結果の上位 (テクスチャ)



図 13 入力画像 (a) に対する検索結果の上位 (全て一律)



図 19 入力画像 (b) に対する検索結果の上位 (全て一律)



図 14 入力画像 (a) に対する検索結果の上位 (最適な重み)



図 20 入力画像 (b) に対する検索結果の上位 (最適な重み)

した場合、最適な重み係数の場合の3つの手法を用いた。図9の入力画像(a)に対する検索結果の上位3枚を図9, 10, 11, 12, 13, 14に、図15の入力画像(b)に対する検索結果の上位3枚を図15, 16, 17, 18, 19, 20に示す。ただし、重み付け係数は表1の通りである。

4.2 考察

まず、髪色で検索した結果(図9, 15)について考察する。どちらの入力画像に対しても検索結果の上位に髪色の近い

表 1 各検索結果における重み付け係数

特徴量	重み付け係数			
	w_c	w_l	w_v	w_t
色 (図 9, 15)	1.0	0.0	0.0	0.0
長さ (図 10, 16)	0.0	1.0	0.0	0.0
ボリューム (図 11, 17)	0.0	0.0	1.0	0.0
テクスチャ (図 12, 18)	0.0	0.0	0.0	1.0
全て一律 (図 13, 19)	0.25	0.25	0.25	0.25
最適な重み (図 14, 20)	0.04	0.42	0.40	1.4

画像が得られた。しかし、図 15 の 1 位のように部分的に髪色が異なる画像も上位に入っているため、今後は位置情報も考慮する必要がある。さらにその際には、照明環境によっては髪にハイライトが乗ってしまい、色特徴量にもハイライトの成分が含まれてしまうことが考えられるため、この点についても考慮する必要がある。

二つ目に、髪の長さで検索した結果(図 10, 16)について考察する。入力画像(a)に対しては検索結果の上位に髪の長さの近い画像が得られた。しかし、図 16 の 2,3 位を見ると、入力画像より髪の短い画像である。これは長さの特徴量のヒストグラムの位置合わせが正しくできていないためであると考えられる。

三つ目に、髪のボリュームで検索した結果(図 11, 17)について考察する。どちらの入力画像に対しても検索結果の上位にボリューム感の近い画像が得られた。また、上位の多くに髪の長さの近い画像が入っていたため、髪の長さと同様に考慮した特徴量になっている可能性がある。そのため、今後は長さと同様に考慮した特徴量に改善する必要がある。

四つ目に、髪のテクスチャで検索した結果(図 12, 18)について考察する。図 12 の 1 位は入力と近いものであったが、他の画像はあまり近いものが得られなかった。その原因として、LBP 特徴量が局所領域でのテクスチャを表すものであるためと考えられる。また、部分的にパーマがかかっている髪を表すためには、位置情報を考慮した特徴量にする必要がある。

最後に、全ての重みが一律の結果と最適な重みの結果(図 13, 図 19 と図 14, 図 20)について考察する。入力画像(a)に対する結果は、一律の重みと最適な重みの場合で 1 位は同じとなった。一律の重みの場合には色の寄与が少なく、最適な重みの場合には長さと同様にボリュームの寄与が大きいといえる。それに対して、入力画像(b)に対する結果は、一律の重みと最適な重みの場合で上位 3 位に入る画像は同じとなった。最適な場合には一律の場合に比べて色の重要度が低くなっている。また、どちらの入力画像に対しても一律の重みの場合と最適な重みの場合で結果があまり変動しなかった。そのため、お互い寄与しないような特徴量にすることや、特徴量の正規化などが今後の課題として挙げられる。

5. おわりに

本稿では、髪の特徴に基づく顔画像の検索手法を提案した。ユーザが似ていると判断する上で重視する特徴を考慮した画像検索を実現した。

本研究の今後の課題として、髪領域の抽出精度の向上が挙げられる。髪の形状は、人の印象と関連性が高く、画像検索の精度に大きく影響する。今後は周波数成分やエッジ成分を用いて髪の流れを考慮した手法を検討していく必要

がある。また本稿では、Yacoob らの提案している特徴量のうち 4 種類をもとに行ったが、今後はさらに多くの特徴量の中からどの成分が寄与しているかを定量化する必要がある。

最後に、髪以外に外見的な特徴や印象として挙げられる表情、化粧、服装なども考慮することによって、画像検索の多様性・効率化を図りたい。

謝辞 本研究の一部は JST CREST「OngaCREST プロジェクト」の支援を受けた。

参考文献

- [1] Google: Google 画像検索, <http://images.google.com/>
- [2] Yahoo: サービス別 画像 - Yahoo!検索データ, http://searchranking.yahoo.co.jp/image_buzz/
- [3] goo: 画像検索 - goo 画像・動画・音楽検索, <http://bsearch.goo.ne.jp/>
- [4] 小松佐穂子, 箱田裕司: 髪の色が人物印象に与える効果: 平凡な顔と個性的な顔における違い, 電子情報通信学会技術研究報告 ヒューマン情報処理, pp. 7-12 (2001).
- [5] 金子智栄子, 門脇幹雄: 外見の印象: 髪型が性格のイメージに及ぼす影響, 日本教育心理学会総会発表論文集, Vol.43, No.35, (2001).
- [6] P. Sinha and T. Poggio: United We Stand: The Role of Head-Structure in Face Recognition, Perception, Vol. 31, No. 1 (2006).
- [7] M. Chai, L. Wang, Y. Yu, Y. Weng, B. Guo, and K. Zhou: Single-View Hair Modeling for Portrait Manipulation, ACM SIGGRAPH 2012, Papers, No. 116 (2012).
- [8] Y. Yacoob, and L. Davis: Detection and analysis of hair, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No.7, pp. 1164-1169 (2006).
- [9] C. Rousset, and P.Y. Coulon: Frequential and color analysis for hair mask segmentation, IEEE International Conference on ICIP 2008, pp. 2276-2279 (2008).
- [10] N. Kumar, P. Belhumeur, and S. Nayar: FaceTracer: A search engine for large collections of images with faces, In Computer Vision-ECCV 2008, pp. 340-353 (2008).
- [11] N. Wang, and H. Ai: Hair style retrieval by semantic mapping on informative patches, Asian Conference on Pattern Recognition, pp. 110-114 (2011).
- [12] A. Irie, M. Takagiwal, K. Moriyama, and T. Yamashita: Improvements to Facial Contour Detection by Hierarchical Fitting and Regression, Asian Conference on Pattern Recognition, pp. 273-277 (2011).
- [13] 細井聖, 田畑尚弘, 秋間正道, 川出雅人: 髪型の認識と合成, 電子情報通信学会技術研究報告 PRMU, 97.387, pp. 25-32 (1997).
- [14] Y.Y. Boykov, and M-P. Jolly: Interactive graph cuts for optimal boundary and region segmentation of objects in ND images, IEEE International Conference on ICCV (2001).