

7S-4 複数の動的な網のモデルの協調による顔部品の抽出

船山竜士^{†‡}

横矢直和[†]

岩佐英彦[†]

竹村治雄[†]

[†]奈良先端科学技術大学院大学 情報科学研究科

[‡]シャープ株式会社 情報商品開発研究所

1 はじめに

顔部品を抽出する問題は、個人同定やマン・マシンインターフェイスなどへの広範囲な応用が考えられ、従来から様々な手法が試みられてきた [1][2]。

投影やテンプレートマッチングを用いる手法 [1] はアルゴリズムが簡単であるが、照明や背景、回転などの影響を受けやすく、制御された環境における画像に対してでないとは有効でない。また、Deformable Template による顔部品の抽出法 [2] は、照明や回転の影響を受けにくいといった特徴があるが、テンプレートの初期位置を従来のテンプレートマッチングのような手法を用いて設定する必要がある。

本稿では、複数の動的な網のモデル [3] の協調により、顔輪郭および目、口唇といった顔部品を抽出する手法について述べる。エネルギー最小化原理に基づく動的な網のモデルは、抽出する領域の形状やテクスチャの情報などを盛り込むことができる。この動的な網のモデルを協調して動作させることにより、画像中の顔位置や大きさ、回転、照明条件などに対してロバストで、網を特定の初期位置に設定せず顔部品を抽出することが可能となる。

2 色情報による領域確率画像の生成

前処理として、色情報を用いた顔領域および口唇領域確率画像の生成を行なう。図 1 は、24 人のサンプルデータより導出した顔皮膚領域および口唇領域の確からしさを示す 2 次元正規密度関数である。左側の軸が色相、右側の軸が彩度で高さが確からしさを示す。この領域確率密度関数を図 2 の入力画像に適

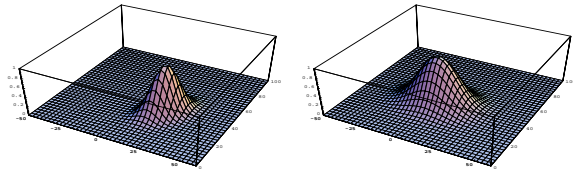


図 1: 顔領域 (左)、口唇領域 (右) の確からしさを示す密度関数
用いたものが、それぞれ顔領域確率画像、口唇領域確率画像である。明度が大きい点ほど、顔領域あるいは口唇領域である可能性が高いことを示している。



図 2: 領域確率画像の生成

3 動的な網のモデルによる顔領域の抽出

色情報を用いて生成された顔領域確率画像に対して動的な網のモデルを作用させ、顔領域を抽出する。ここで用いるエネルギー関数は以下の通りである。

$$E_{net} = \int_0^1 \int_0^1 (E_{int}(v(p, q)) + E_{image}(v(p, q))) dpdq$$

E_{int} は網の内部歪みエネルギーであり、網を収縮させかつ滑らかに保とうとする性質を持つ。また、 E_{image} は網と画像の適合性エネルギーであり、顔領域の特徴を反映して決定される。

図 3 は、顔領域確率画像に対して動的な網のモデルを作用させ、網が顔領域に向かって収束する様子を示している。内部格子点は顔領域に向かい、最外郭格子点は顔領域の外に留まろうとする様にエネルギーを定義しているため、網は顔領域を囲むように収束している。

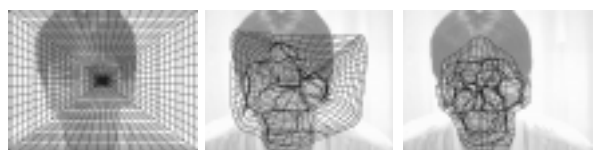


図 3: 動的な網のモデルによる顔領域の抽出

Extraction of Facial Components Using Cooperative Active Net Models

Ryuji Funayama, Naokazu Yokoya, Hidehiko Iwasa and Haruo Takemura

[†]Nara Institute of Science and Technology (NAIST), 8916-5 Takayama, Ikoma, Nara 630-01, Japan

[‡]Information Systems Product Development Laboratories, SHARP Corporation, 492 Minosho, Yamatokoriyama, Nara 639-11, Japan

4 複数の網のモデルの協調

顔領域を囲む網が収束すると、その網を初期配置として右目、左目、口唇領域に収束する網の動作を開始させる。これらの4つの網が同時に存在することによって生ずるエネルギー E_{person} は、個々の網のエネルギーと、網の位置関係によって生ずるエネルギー $E_{balance}$ の線形結合によって表される。

$$E_{person} = E_{face-net} + E_{mouth-net} +$$

$$E_{left-eye-net} + E_{right-eye-net} + E_{balance}$$

内部エネルギーにはそれぞれの部品の形状、大きさを制約条件として加え、画像エネルギーについては抽出する部品それぞれの特徴を反映するように定義する。例えば、口唇領域を抽出するための網は、内部格子点は図2の口唇領域確率画像とそれに対応する顔領域確率画像の明度の差の大きい点に引き寄せられ、最外郭格子点は顔領域確率画像の明度の大きい点に向かう。そして内部エネルギーは網を、口唇の近似形である楕円の形状と顔幅から推定される口唇の平均的なサイズへと収縮させる。 $E_{balance}$ は、個々の網の位置関係を表すエネルギーであり、次のように定義する。

$$E_{balance} = E_{mouth-gravity} + E_{left-eye-gravity} +$$

$$E_{right-eye-gravity} + E_{inter-ocular-distance} +$$

$$E_{eye-mouth-height} + E_{face-angle} + E_{eye-mouth-angle}$$

$E_{mouth-gravity}, E_{left-eye-gravity}, E_{right-eye-gravity}$: それぞれ、口唇領域、左目領域、右目領域抽出用網の最外郭格子点の重心が、網が囲む領域の重心に引き寄せられるようなエネルギー

$E_{inter-ocular-distance}$: 左右の目用網が囲む領域の重心間の距離が、顔幅に対する特定の比率になろうとするエネルギー

$E_{eye-mouth-height}$: 左右の目用網が囲む領域の重心同志を結ぶ線分の中点と口唇用網が囲む領域の重心との距離が、顔幅に対する特定の比率になろうとするエネルギー

$E_{face-angle}$: 左右の目用網が囲む領域の重心同志を結ぶ線分の中点と口唇用網が囲む領域の重心を結ぶ軸および、顔の傾きを表す軸とがなす角が0になろうとするエネルギー

$E_{eye-mouth-angle}$: 左右の目用網が囲む領域の重心同志を結ぶ線分および、その線分の中点と口唇用網が囲む領域の重心を結ぶ軸とがなす角が直角になろうとするエネルギー

これらのエネルギー項の線形結合によって表される $E_{balance}$ が最小となるように、それぞれの網を移動させる。4つの網がそれぞれの領域に向かって変形、

収縮し、網の位置関係が一定の平衡状態を保つように移動することにより、それぞれの網は最終的に個々の領域を囲むように収束する。図4に、入力画像に対して本手法を適用した結果を示す。上から順に、口唇領域、左目領域、右目領域用の網が各領域を抽出しているのがわかる。なお、ここでは見易いように各網の動きを別々に示している。

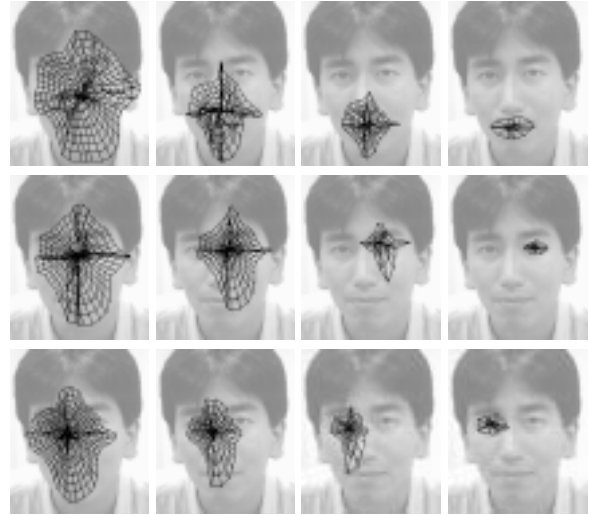


図4: 複数の網のモデルの協調による顔部品の抽出

5 まとめ

複数の動的な網のモデルの協調により顔部品を抽出する手法について述べた。本手法の特徴としては、

- 画像中の顔位置や大きさ、回転、照明条件などに対してロバストである
- Deformable Template のように他手法を用いた初期位置の設定などを行なう必要がない

といったことが挙げられる。

今後は、よりロバストな顔部品抽出のためのエネルギー関数、エネルギー項の重み付け係数の検討などを進める予定である。

参考文献

- [1] R. Brunelli and T. Poggio, "Face recognition: Features versus templates", IEEE Trans. PAMI, Vol.15, No.10, pp.1042-1052, 1993.
- [2] A. Yuille, D. Cohen and P. Hallinan, "Feature extraction from faces using deformable templates", Proc. CVPR'89, pp.104-109, 1989.
- [3] 坂上, 山本, "動的な網のモデル Active Net とその領域抽出への応用", テレビジョン学会誌, Vol.45, No.10, pp.1155-1163, 1991.