

# 複数の動的な網のモデルの協調とその顔部品抽出への応用

船山竜士<sup>†‡</sup> 横矢直和<sup>†</sup> 岩佐英彦<sup>†</sup> 竹村治雄<sup>†</sup>

<sup>†</sup>奈良先端科学技術大学院大学 情報科学研究科

〒630-01 奈良県生駒市高山町 8916-5

電話番号：07437-9-9211(EX.5296) FAX 番号：07437-2-5299

<sup>‡</sup>シャープ株式会社 情報商品開発研究所

〒639-11 奈良県大和郡山市美濃庄町 492

電話番号：07435-3-5521(代表) FAX 番号：07435-2-1659

E-mail:{ryuuji-f, yokoya, iwasa, takemura}@is.aist-nara.ac.jp

あらまし 本研究では、あらかじめ相対的位置関係がわかっている複数の領域を抽出する問題において、グローバルな制約により、複数の動的な網のモデルを位置関係が一定の平衡状態を保つように制御することにより領域を安定に抽出する手法を提案する。この複数の動的な網のモデルの協調を顔画像からの顔部品抽出に適用する手法は、従来のテンプレートマッチングや投影を用いた手法、あるいは Deformable Template のようなエネルギー最小化原理に基づいた他の手法に比べ、画像中の顔位置や大きさ、回転、テンプレートの初期位置の変化などに対してロバストである。本発表では、提案手法の概要とカラー顔画像を用いた実験結果について報告する。

キーワード 顔部品抽出, 動的な網のモデル, 協調動作, エネルギー最小化原理

## Cooperative Active Net Models and Its Application to Facial Component Extraction

Ryuji Funayama<sup>†‡</sup> Naokazu Yokoya<sup>†</sup> Hidehiko Iwasa<sup>†</sup> Haruo Takemura<sup>†</sup>

<sup>†</sup>Nara Institute of Science and Technology (NAIST)

8916-5 Takayama, Ikoma, Nara 630-01, Japan

TEL : (+81)7437-9-9211(EX.5296) FAX : (+81)7437-2-5299

<sup>‡</sup>Information Systems Product Development Laboratories, SHARP Corporation

492 Minosho, Yamatokoriyama, Nara 639-11, Japan

TEL : (+81)7435-3-5521 FAX : (+81)7435-2-1659

E-mail:{ryuuji-f, yokoya, iwasa, takemura}@is.aist-nara.ac.jp

**Abstract** This paper describes cooperative active net models for extracting multiple regions whose relative spatial configurations are known. Applying this method, facial components can be successfully extracted from a face image. Multiple active nets are cooperatively deformed by employing global constraints which specify spatial configurations of facial components. This facial component extraction method has been experimentally shown to be robust to variations of facial size, position and orientation compared with existing methods such as template matching, projection-based method and deformable templates.

**key words** Facial component extraction, Active net model, Cooperative behavior, Energy minimization principle

## 1 はじめに

近年、顔画像処理はコンピュータビジョンにおける非常に重要な研究分野を構成しており、これまでに多くの研究報告がなされてきた。対象が動画像、静止画像のいずれであるかに関わらず、画像中のどこに顔があるのか、それは誰なのか、また、どのような表情をしているのかといったことがこの研究分野の中心的課題となっている [1]。

これらを解決するための手法には、様々なアプローチが考えられるが、その中でも目・鼻・口といった顔部品の抽出は、個人同定を実現するための重要な足掛かりとなる。また、ビジョン分野以外でも、位相情報通信や表情生成などのアニメーション、視線方向の検出などといったアプリケーションが想定され、顔部品抽出は、最も基本的かつ重要な課題であると考えられる。

したがって、これまでの研究で顔部品の抽出に関する様々な手法が提案されてきた。代表的なものに、投影を用いる手法とテンプレートマッチングを用いる手法がある [2]。前者はアルゴリズムが簡単で高速処理が可能であるという利点があるものの、照明条件や背景などの影響を受けやすい。後者は、濃淡画像を用いて比較的安定した結果が得られるものの、同じく照明や回転、背景などの影響を受けやすいという欠点がある。

一方、エネルギー最小化原理に基づく Deformable Template による顔部品の抽出法 [3] が提案されているが、初期テンプレートの位置を、抽出しようとする顔部品の近傍に設定しなければならないため、対話的な方法やヒューリスティックな手法に頼らざるを得ないという欠点がある。この方法では、テンプレートの初期位置が適当でないと、誤抽出を起こしやすいこともわかっている。

本報告では、エネルギー最小化原理に基づく動的な網のモデル [4] を用いて、入力されたカラー人物画像より顔領域および目、口唇といった顔部品領域を抽出する手法について述べる。複数の網を同時に協調して動作させることにより、単独で動作させた場合に網が局所解へと収束して誤った領域を抽出してしまうという問題を回避することができる。

本手法は、投影やテンプレートマッチングなどによる従来の顔部品抽出法に比べ、画像中の顔位置や大きさ、回転などに対してロバストであり、また Deformable Template のように、初期位置を他の手法によって設定する必要もないといった特長が挙げられる。

## 2 動的な網のモデル

本研究では、顔領域および顔部品領域の抽出に動的な網のモデル [4] を適用する。以下に、動的な網のモデルについて簡単に説明する。

エネルギー最小化原理に基づく領域抽出の手法として、Kass、Witkin らが提案した動的輪郭モデル (Snakes) がある [5]。これは、画像上に閉曲線で表される輪郭を想定し、その輪郭自身が持つエネルギー、輪郭と画像との間で生ずるエネルギー、および外部からの強制力によるエネルギーを定義し、そのエネルギーが小さくなるように輪郭を変形していくことによって、領域の抽出を行なうというものである。

積山、坂上らの提案した動的な網のモデルは、この Snakes を 2 次元的な面モデルに拡張したものであり、領域内部のテクスチャ情報なども利用して、よりロバストな領域抽出を可能にする。網のエネルギー  $E_{net}$  は、網自身の収縮力や平行性を保つ力を与える内部エネルギー  $E_{int}$ 、網が画像の特徴的な領域へ適合するよう作用する画像エネルギー  $E_{image}$ 、そして外部からの強制力に対応する外部エネルギー  $E_{con}$  の 3 つのエネルギー項の線形結合として式 (1) のように表される。

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

ここで、 $v(p, q)$  は網の格子点を表し、周囲 4 近傍の格子点を接続することによって網が構成される。

## 3 顔領域の色分布解析

本研究では、入力画像より顔領域を抽出するために、顔の色情報を利用する。以下に、顔領域の色分布を解析し、それを網の画像エネルギー定義に用いる確率画像を生成する方法について述べる。

まず、実験用に撮影した人物画像より、顔肌領域と口唇領域を対話的に切り出す。この切り出した顔肌領域および口唇領域それぞれについて、色分布の解析を行ない、各領域を特徴付ける指標を導出する。用いる指標は、照明の変化や個人差による肌の明るさの違いなどに対してロバストであることが望ましいため、ここでは HSV 表色系における H (色相) 成分と S (彩度) 成分を用いる。色相成分は、照明などによる反射や陰影の影響を受けにくく、また、顔領域は色相値がほぼ一定であるため、肌の特徴を記述するのに向いていると考えられる。また、室内背景に比べ、人間の顔は相対的に彩度の高い値を示すた

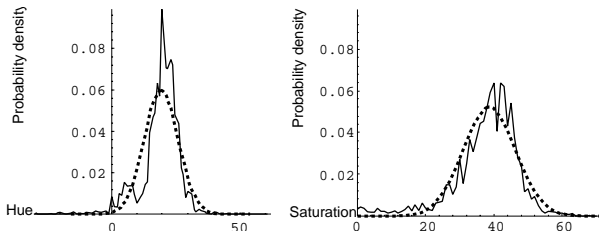


図 1: 顔肌領域の色相分布 図 2: 顔肌領域の彩度分布

め、背景と分離するための指標として彩度を用いることも妥当であると考えられる [6]。

サンプル画像データより、顔肌の色相 ( $-180^\circ \sim 180^\circ$ ) と彩度 ( $0 \sim 100$ ) の度数分布を調べると、図 1、図 2 のように顔肌の色相は狭い範囲に分布し、彩度も極端に高い値や低い値は少ないということがわかる。また、両分布とも図の点線のように正規分布で近似できることが確認できる。したがって、両値を引数とした 2 次元正規分布の確率密度関数を導出し、これを顔肌領域の確からしさを表す指標とする。

導出された確率密度関数をプロットすると、図 3 のようになる。左側の軸が色相、右側が彩度で、高さが確からしさを表している。口唇領域についても、同様の関数が得られる。このようにして求めた関数を入力画像に適用すると、図 4 のような確率画像が得られる。ここで、顔肌領域確率画像とは顔肌領域の確からしさを画素の輝度値で表したものであり、口唇領域確率画像についても同様である。

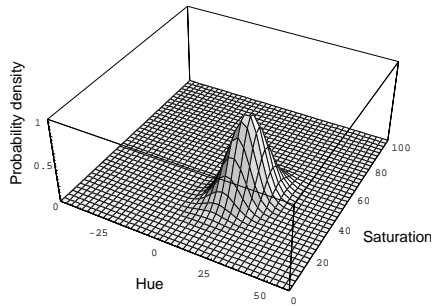
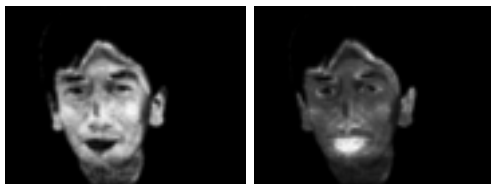


図 3: 顔肌領域の確からしさを表す確率密度関数



入力画像



顔肌領域確率画像 口唇領域確率画像

図 4: 顔肌領域および口唇領域を表す確率画像

## 4 動的な網のモデルによる顔領域および顔部品領域の抽出

入力画像から、顔領域および顔部品領域を抽出するための処理の流れを図 6 に示す。まず、画像全体を覆う網を初期値として設定し、その網が顔領域を囲むように網を変形、収縮させる。そして、顔領域を覆って停止した網を、今度は左右の目と口唇領域を抽出するための 3 つの網の初期値とし、それらの網を並列に協調して動作させる。

以下に、網のモデルを用いた顔領域および顔部品領域の抽出法について述べる。

### 4.1 動的な網のモデルによる顔領域の抽出

最初のステップとして、動的な網のモデルを用いて顔領域を抽出する手法について述べる。顔領域を抽出するための網のエネルギーは、内部エネルギーと画像エネルギーについてのみ定義し、外部エネルギーは考慮しない。

#### 4.1.1 網のモデル化

まず、本研究で用いる網のモデルを定義する。抽出の対象となるのは顔領域であり、その形状はおおよそ楕円として近似できる。したがって、図 5 のような円形の網を定義する。

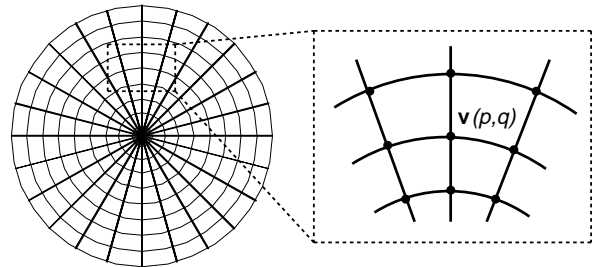


図 5: 網の定義

$v$  は網の格子点を表し、 $p$  と  $q$  の 2 つのパラメータによって、網を構成する任意の格子点を指定する。 $p$  が円周方向、 $q$  が中心から放射状に伸びる線分方向を表す。

#### 4.1.2 内部エネルギー

内部エネルギーは、式 (2) のように定義する。

$$E_{face-int} = (\alpha(|v_p|^2 + |v_q|^2) + \beta(|v_{pp}|^2 + 2|v_{pq}|^2 + |v_{qq}|^2))/2 \quad (2)$$

ただし、 $v_p = \partial v / \partial p$ 、 $v_q = \partial v / \partial q$ 、 $v_{pp} = \partial^2 v / \partial p^2$ 、 $v_{pq} = \partial^2 v / \partial p \partial q$ 、 $v_{qq} = \partial^2 v / \partial q^2$  である。ここで 1 次偏微分の項  $|v_p|$ 、 $|v_q|$  は、注目格子点とその隣接格子点までの距離に相当し、2 次偏微分の項  $|v_{pp}|$ 、 $|v_{pq}|$ 、 $|v_{qq}|$  は、網の平行性と直線性を表す。これら

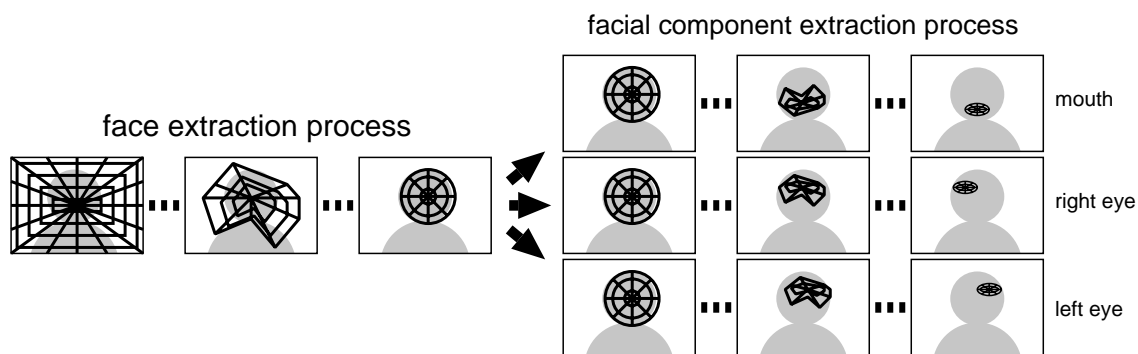


図 6: 顔部品領域を抽出するための処理の流れ

の項が小さくなるように網を変形させるということは、網が円形状を保ちながら収縮することに相当する。

#### 4.1.3 画像エネルギー

顔領域を抽出するための網の画像エネルギーは、入力画像より生成された顔肌領域確率画像 (図 4左) を用いて定義する。格子点  $v$  における、顔肌領域確率画像の画素の輝度値を  $I_{face}(v)$  とすると、画像エネルギー  $E_{face-image}$  は式 (3) のように表される。

$$E_{face-image} = \begin{cases} -w_f I_{face}(v) & (v \text{ が内部格子点の時}) \\ w_f I_{face}(v) & (v \text{ が最外郭格子点の時}) \end{cases} \quad (3)$$

ここで、 $w_f$  は正符号を持つ係数である。この式で表されるエネルギーが小さくなるように網を変形させるということは、次のことを意味する。すなわち、内部格子点が顔肌領域確率画像の輝度値の高い画素上にある場合、画像エネルギーの値は小さくなる、つまり、内部格子点は顔肌である確からしさの高い領域に引き寄せられることになる。一方、最外郭格子点での画像エネルギーは、内部格子点でのエネルギーと符号が逆転しているの、逆に最外郭格子点は顔肌領域の外に留まろうとする力が働く。

#### 4.1.4 顔領域の抽出

顔領域を抽出するための網のエネルギー  $E_{face-net}$  は、式 (4) のように、上で述べてきた内部エネルギーと画像エネルギーの 2 つのエネルギー項の線形結合によって定義される。

$$E_{face-net} = E_{face-int} + E_{face-image} \quad (4)$$

内部エネルギー  $E_{face-int}$  の作用により、網は一定の形状を保ちながら収縮し、画像エネルギー

$E_{face-image}$  の作用により網が顔領域を囲ったところで停止する。これにより、入力顔画像中の顔領域を抽出することが可能となる。

図 7 に、入力画像より動的な網のモデルを適用して顔領域を抽出する様子を示す。初期値として画像全体を覆っていた網が、徐々に収縮していき、最終的に顔領域を抽出しているのがわかる。

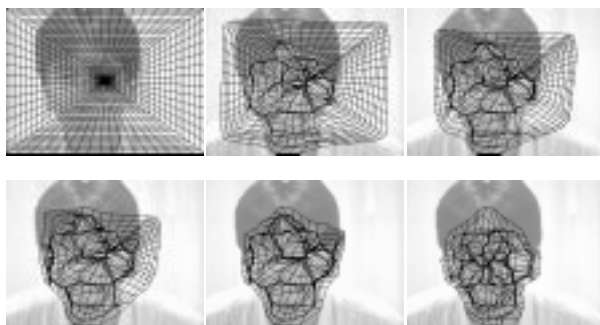


図 7: 動的な網のモデルによる顔領域の抽出

## 4.2 動的な網のモデルによる顔部品領域の抽出

### 4.2.1 内部エネルギー

ここで顔部品とは、口唇領域および左右の目領域のことをいう。各顔部品領域を抽出するための内部エネルギー  $E_{mouth-int}$ 、 $E_{le-int}$ 、 $E_{re-int}$  の定義は、顔領域を抽出するための網の場合と全く同じである。

### 4.2.2 画像エネルギー

格子点  $v$  における口唇用網の画像エネルギー  $E_{mouth-image}$  の定義は、式 (5) のようになる。

$$E_{mouth-image} = \begin{cases} -w_m I_{mouth}(v) & (v \text{ が内部格子点の時}) \\ w_m I_{mouth}(v) - w_{me} Edge(v) & (v \text{ が最外郭格子点の時}) \end{cases} \quad (5)$$

ここで  $I_{mouth}(v)$  は、格子点  $v$  における口唇領域画像の輝度値である。口唇領域画像とは、図4の口唇領域確率画像から顔領域確率画像を引き、輝度値が負になった画素の値を0にしたものであり、明るい大きな領域が口唇を表している(図8)。また  $Edge(v)$  は、格子点  $v$  での入力画像のエッジ強度を表す。 $w_m$ 、 $w_{me}$  は正符号を持つ係数であり、 $E_{mouth-image}$  が小さくなるように網を変形させるということは、内部格子点が口唇領域画像での明るい領域へ引き寄せられ、最外郭格子点が口唇領域画像での暗い領域かつ、口唇の境界付近に留まるということを意味する。

目領域用の網も同様に、目領域画像の輝度値を用いて定義する。目領域画像は、顔および口唇領域確率画像の2つの画像を重ね合わせた時に、対応する画素の明るい方をとった画像で、顔領域中の暗い領域が目と眉毛を表す(図9)。



図 8: 口唇領域画像



図 9: 目領域画像

#### 4.2.3 外部エネルギー

ここでは、2つの外部エネルギーを定義する。形状制約エネルギー  $E_{form}$  は顔部品の形状に関する先見的な制約を反映したもの、対称制約エネルギー  $E_{sym}$  は目の形状の対称性を反映したものである。まず、網の2つの重心について定義し、そこから導き出される顔の傾きおよび顔幅の推定法について述べる。そして  $E_{form}$ 、 $E_{sym}$  について定義する。

##### (1) 輪郭重心と領域重心

輪郭重心  $g$  を式(6)、領域重心  $G$  を式(7)のように定義する。

$$g = \frac{\sum_{i=1}^n v_i}{n}, \quad n = \text{最外郭格子点数} \quad (6)$$

$$G = \frac{\sum_{i=1}^N I(p_i)p_i}{N}, \quad N = \text{領域内の画素数} \quad (7)$$

$v_i$  は最外郭格子点であり、 $I(p_i)$  は画像  $I$  における点  $p_i$  での輝度値である。すなわち、 $g$  はすべての最外郭格子点座標値の平均であり、 $G$  は網が囲む領域内にあるすべての画素の座標値を、その画素の輝度値で重み付けして平均したものである。したがって、網が囲む領域に輝度値の高い領域が存在する場合、

領域重心  $G$  は輪郭重心  $g$  よりも、その輝度値の高い領域の近くに存在することになる(図10)。

##### (2) 顔の傾きと顔幅の推定

顔領域を覆って停止した網の輪郭重心  $g_f$  は、顔のほぼ中央に位置することになる。一方、顔領域を初期値として口唇領域を抽出するための網が動作し、口唇領域に向かって収縮していくが、図8における口唇領域画像上でこの網の領域重心  $G_m$  は、 $g_f$  よりも口唇領域の近くに位置することになる。したがって、 $g_f$  と  $G_m$  を結ぶ直線を、顔の傾きを表す軸とみなす。また、この軸に垂直で  $g_f$  を通る直線が網を横切る距離を顔幅の推定値とする(図11)。

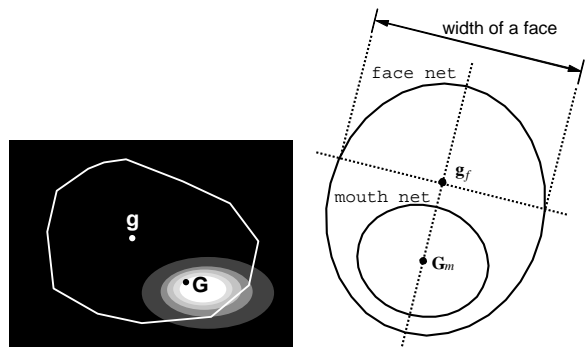


図 10: 輪郭重心  $g$  と領域重心  $G$  図 11: 顔の傾きと顔幅の推定

##### (3) 形状制約エネルギー

人間の口唇の大きさと形状は、顔の大きさからおおよそ推定することができる。そこで、推定された顔幅に対して一定の比率を乗ずることにより得られる口唇の幅、高さをそれぞれ長径、短径とし、顔の傾きだけ回転させた楕円を、口唇の推定形とする。その楕円の中心を口唇領域重心  $G_m$  に重ねて置き、その楕円領域に向かって網が収縮する力を形状制約エネルギーとして式(8)のように定義する。

$$E_{mouth-form} = \begin{cases} w_{mf} |F_{mouth}(v)|^2 & (v \text{ が最外郭格子点か、} \\ & \text{内部格子点かつ推定領域の} \\ & \text{外にある時)} \\ 0 & (v \text{ がそれ以外の格子点の時)} \end{cases} \quad (8)$$

$F_{mouth}(v)$  は、格子点  $v$  から網の領域重心方向に伸ばした、楕円との交点までのベクトルを表す(図12)。 $w_{mf}$  は正符号を持つ係数であり、 $E_{mouth-form}$  を小さくするように網を変形させることは、推定された口唇領域に向かって網を収縮させることに相当する。

左右の目領域を抽出する網の形状制約エネルギー ( $E_{le-form}$ 、 $E_{re-form}$ ) も同様に定義できる。

#### (4) 対称制約エネルギー

対称制約エネルギーは左右の目領域を抽出するための網に関するエネルギーで、一方の目を左右反転すると、もう一方の目とほぼ等しい形状になる、という先見的な制約に基づいている。

左目の対称制約エネルギー  $E_{le-sym}$  は、式 (9) のように定義する。右目  $E_{re-sym}$  についても同様。

$$E_{le-sym} = \begin{cases} w_{les} |S_{le}(v)|^2 & (v \text{ が最外郭格子点の時}) \\ 0 & (v \text{ が内部格子点の時}) \end{cases} \quad (9)$$

$S_{le}(v)$  は、顔の傾きを表す軸に対して右目用の網を左右反転させたものと、左目用の網とを領域重心を重ねて表示した時に、左目用網の最外郭格子点  $v$  から領域重心方向に伸ばした、反転した右目用網との交点までのベクトルを表す (図 13)。 $w_{les}$  は正符号を持つ係数であり、したがって、 $E_{le-sym}$  (および  $E_{re-sym}$ ) を小さくするように左右の目の網を変形させることは、顔の傾きを表す軸を対称として 2 つの網が同じ形状となる力を与えることに相当する。

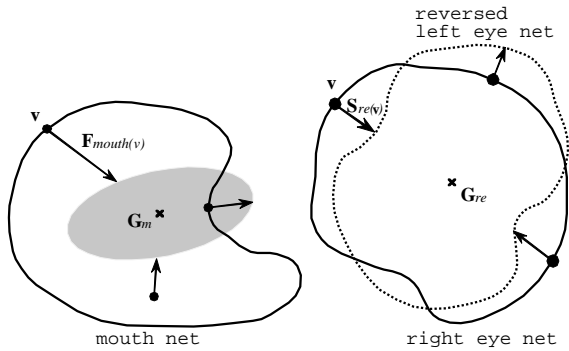


図 12: 形状制約エネルギー 図 13: 対称制約エネルギー

#### 4.2.4 顔部品を抽出するためのエネルギー

顔部品を抽出するためのエネルギーの定義についてまとめると、次のようになる。

$$E_{mouth-net} = E_{mouth-int} + E_{mouth-image} + E_{mouth-form} \quad (10)$$

$$E_{le-net} = E_{le-int} + E_{le-image} + E_{le-form} + E_{le-sym} \quad (11)$$

$$E_{re-net} = E_{re-int} + E_{re-image} + E_{re-form} + E_{re-sym} \quad (12)$$

#### 4.3 複数の動的な網のモデルの協調

これまで各顔部品領域を抽出するための網のエネルギーについて定義してきたが、これらの網を単純に顔画像に適用するだけでは、顔部品領域を正しく抽出することができない。網のエネルギーが局所的な最適解に落ち込み、誤った領域が抽出されてしまうことが多い。局所解は無数に存在し、その解は網の初期値に大きく依存する。したがって、通常は網の初期位置を目的の領域の近傍に設定することになる。

しかし、顔画像中の左右の目と口唇というように、個々の領域の位置関係に一定の規準が存在する場合、その制約をエネルギーとして定義することが可能である。そのエネルギーが与える制約にしたがって、個々の領域を抽出するための網の位置関係を制御し、協調動作させることにより、誤抽出を防ぎ、安定して顔部品を抽出することが可能となる。

ここでは、複数の網の協調動作を実現するために新しいエネルギー項  $E_{balance}$  を定義する。これは、口唇および左右の目用の網の位置関係などによって生ずるエネルギーであり、このエネルギーが小さくなるように 3 つの網を移動させて位置関係を制御する。 $E_{balance}$  は式 (13) のように定義する。

$$E_{balance} = E_{mg} + E_{leg} + E_{reg} + E_{iod} + E_{emh} + E_{ema} + E_{fa} \quad (13)$$

最初の 3 つの項  $E_{mg}$ 、 $E_{leg}$ 、 $E_{reg}$  は、それぞれ口唇領域、左目領域、右目領域の、各輪郭重心と領域重心との距離の自乗で与えられる。すでに述べたように、輪郭重心よりも領域重心の方が、抽出しようとする領域に近い場所に位置する。したがって、これらのエネルギー項が小さくなるように網を移動することは、網が抽出しようとする領域に向かって移動することに相当する (図 14)。

残りの 4 つのエネルギー項については、図 15 を用いて説明する。人間の顔部品の配置は、いくつかの先見的な制約条件に基づいている。顔の大きさに対して両目間の距離と、両目の中央-口唇の高さは、おおよその値が推定できるため、顔幅に対してそれぞ

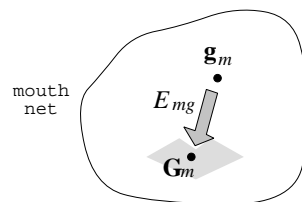


図 14: エネルギー項  $E_{mg}$

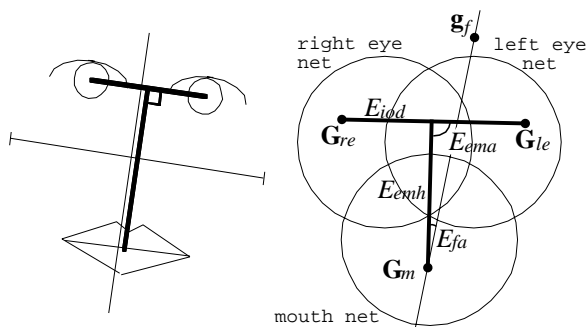


図 15: 顔部品の位置関係と対応する網のエネルギー

れ一定の比率を乗じた値を、それぞれの推定値とする。また、それらの2つの線分、すなわち両目を結ぶ線分と両目の中央-口唇を結ぶ線分は、通常は直角を構成し、後者の線分は顔の傾きを表す軸と一致する。

このような顔部品の配置に関する制約を、そのまま網の位置関係を制御するエネルギー項として定義する。各網の領域重心をその顔部品の位置とすれば、すべての網が収束して各顔部品領域が抽出された時、各領域重心の位置関係は、真の顔部品の位置関係と同じになる。したがって、各エネルギー項を以下のように定義する。

$E_{mg}, E_{leg}, E_{reg}$ : それぞれ、口唇領域、左目領域、右目領域抽出網の輪郭重心と領域重心との距離を縮めるためのエネルギー

$E_{iod}$ : 左右の目用網の領域重心間の距離が、顔幅に対する特定の比率になろうとするエネルギー

$E_{emh}$ : 左右の目用網の領域重心同志を結ぶ線分の中点と、口唇用網の領域重心との距離が、顔幅に対する特定の比率になろうとするエネルギー

$E_{fa}$ : 左右の目用網の領域重心同志を結ぶ線分の中点と、口唇用網の領域重心とを結ぶ軸および、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

$E_{ema}$ : 左右の目用網の領域重心同志を結ぶ線分および、その線分の中点と口唇用網の領域重心とを結ぶ軸とがなす角が直角になろうとするエネルギー

顔部品抽出のための網は、それぞれの網に定義したエネルギーを最小化するように変形、収縮する一方、各網の位置関係が最適の状態になるよう制御される。各網は、その位置関係が一定の平衡状態を保ちつつ、個々の領域に向かって収縮、変形していくため、顔部品領域を安定して抽出することができる。

## 5 実験と考察

今回の実験では、網の変形アルゴリズムとして、グリーディアルゴリズムを用いた局所的な最適化の手法を採用した。注目格子点の近傍点を調べ、最

もエネルギーが小さい点を、その格子点の新しい移動先として記憶する。そして、すべての格子点について同じ処理を施した後に、格子点位置を更新し、網を変形させる。その時のすべての格子点でのエネルギーの和を、網のエネルギーとし、同じプロセスで網を変形していき、エネルギーの変化がなくなったところで網を停止させる。

図4に示した入力画像に対して提案アルゴリズムを適用した結果が図16である。上から順に、口唇領域、左目領域、右目領域が抽出される様子を示している。最初に顔全体を覆っていた網が、次第に各領域に向かって収縮していく様子が見られる。図17は、同じ入力画像を回転させて生成した画像に対して実験したもので、同様に各顔部品領域が抽出されていることがわかる。

この実験で、提案したアルゴリズムが顔部品領域を抽出するのに有効で、顔の回転に対してもロバストであることが示された。エネルギー最小化原理に基づいた領域抽出手法の多くは、対象領域の近傍に初期値を設定しなければならなかったが、本稿で示したように、複数のモデルを協調して作用させることによって、局所解に落ち込むことなく安定した領域抽出が可能となる。

## 6 まとめ

複数の動的な網のモデルの協調動作と、その顔部品抽出への応用について述べた。今回提案した手法では、画像中の顔の回転に対してロバストな顔部品領域の抽出が確認できた。複数の網を協調動作させるような今回の手法は、他の様々な正則化の手法と組合せて適用し、処理のロバスト性を向上させることが可能であると思われる。

今後は、エネルギー関数についてさらに検討し、抽出精度とロバスト性の向上を目指す。また、現在はエネルギーの重み付け係数を経験的に設定しているが、これについても何らかの解決策を検討していく予定である。

謝辞 第一著者に対して本研究の機会を与えて頂きましたシャープ(株)情報商品開発研究所坂田安男所長および斗谷充宏主任研究員に深く感謝します。なお、本研究の一部は、新情報処理開発機構からの受託研究による。

## 参考文献

- [1] R. Chellappa, C. L. Wilson and S. Sirohey, "Human and machine recognition of faces: A survey", Proc. IEEE, Vol.83, No.5, pp.705-740, 1995.



- [2] R. Brunelli and T. Poggio, "Face recognition: Features versus templates", IEEE Trans. PAMI, Vol.15, No.10, pp.1042-1052, 1993.
- [3] A. Yuille, D. Cohen and P. Hallinan, "Feature extraction from faces using deformable templates", Proc. CVPR'89, pp.104-109, 1989.
- [4] 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.
- [5] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active countour models", International Journal of Computer Vision, 1, 4, pp.321-331, 1988.
- [6] 松橋, 藤本, 中村, 南, "顔領域抽出に有効な修正 HSV 表色系の提案", テレビジョン学会誌, Vol.49, No.6, pp.787-797, 1995.
- [7] 船山, 横矢, 岩佐, 竹村, "複数の動的な網のモデルの協調による顔部品の抽出", 情報処理学会第 51 回全国大会, 7S-4, 1995.

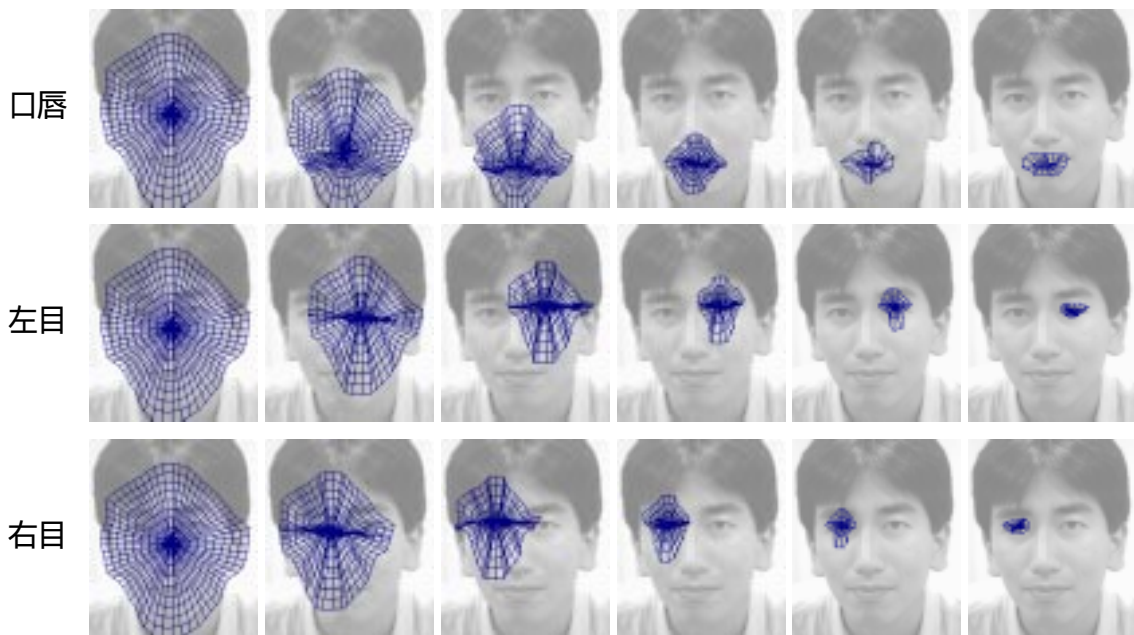


図 16: 動的な網のモデルの協調による顔部品の抽出

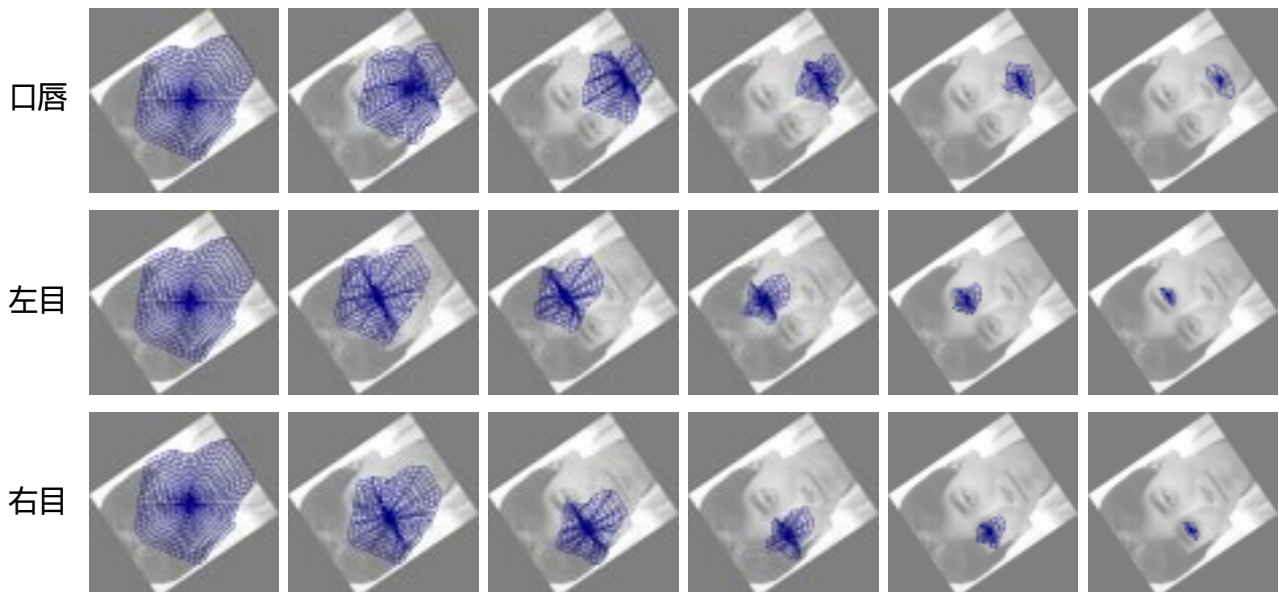


図 17: 回転を含んだ画像における顔部品の抽出