

Active Appearance Model 探索による目領域構造の自動抽出

森山 剛[†] 金出 武雄^{††} Jeffrey F. Cohn^{††} 小沢 慎治[†][†] 慶應義塾大学理工学部情報工学科

〒 223-8522 神奈川県横浜市港北区日吉 3-14-1

^{††} Robotics Institute, Carnegie Mellon University

5000 Forbes Avenue, Pittsburgh, PA 15213-3891, U.S.A.

E-mail: †{moriyama,ozawa}@ozawa.ics.keio.ac.jp, ††{tk,jeffcohn}@cs.cmu.edu

あらまし 目領域画像の上眼瞼(うわまぶた), 下眼瞼, 眼瞼裂の各部分領域について, 目領域の見えからその解剖学的な構造を自動的に抽出する手法を提案する. 目の高精度な運動解析や解剖学的解析は, 医学, バイオメトリクス, 心理学で応用される重要な技術である. 著者らは, 目の構造と運動を分けてモデル化し, 目の運動追跡を行う方法を提案したが, 目領域の見えは, それが平静の見え(元々の構造)であるか, 表情(運動)を伴った見えであるかは区別できないために, 構造について手動で与えていた. 提案する手法は, 目領域の構造に関する情報をあらかじめ付与した学習サンプルから Active Appearance Model (AAM) を構築し, AAM 空間において未知入力に最も近い学習サンプルを見つけることで, その学習サンプルの構造情報から間接的に入力目領域の構造を自動取得するものである. 上眼瞼領域については眉毛及び目の両端点の計4点, 眼瞼裂及び下眼瞼の各領域については目の両端点を固定する制約付き AAM を, 69名の平静(開眼, 瞳中央)な見え状態での目領域画像で学習し, 学習外の33名の目領域画像を AAM で近似する実験を行った. その結果, 目の各領域の多様な見えに対して AAM が高い表現能力を有することが示された. さらに, 目領域の構造取得に関しても良好な結果を得た.

キーワード 画像処理, 目, 表情解析, active appearance model

Automated Individualization of Eye Region Model
using Active Appearance Model SearchTsuayoshi MORIYAMA[†], Takeo KANADE^{††}, Jeffrey F. COHN^{††}, and Shinji OZAWA[†][†] Dept. of Information and Computer Science, Keio University

3-14-1 Hiyoshi, Kouhoku-ku, Yokohama, 223-8522 Japan

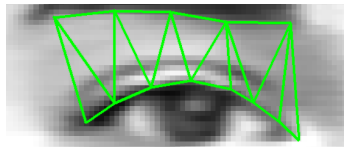
^{††} Robotics Institute, Carnegie Mellon University

Forbes Avenue 5000, Pittsburgh, PA 15213, U.S.A.

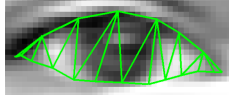
E-mail: †{moriyama,ozawa}@ozawa.ics.keio.ac.jp, ††{tk,jeffcohn}@cs.cmu.edu

Abstract This paper proposes automated parameterization of eye region image by searching the nearest active appearance model (AAM). The authors have developed a generative eye region model that parameterized both the structure and the motion of the eye region. It automatically tracked eye motion after manually individualizing the structure in the first frame of the input image sequence. The proposed method automates the individualization. It divides the eye region into three subregions (the upper eyelid, the palpebral fissure (the eye aperture), and the lower eyelid), and constructs the AAM for each using training samples that are accompanied with the structure parameters such as the size and color of the iris, the width, boldness, and complexity of the eyelids, the width of the bulge below the eye, and the width of the illumination reflection on the bulge. Given an input eye region image, the system searches the nearest AAM for each subregion, and copies the attached structure parameters to the eye region model that is applied to the input for motion tracking later.

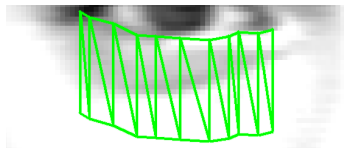
Key words computer vision, eye, facial expression, active appearance model



(a) 上眼瞼 (眉毛下端 5 点と上眼瞼下端 8 点の計 13 点)



(b) 眼瞼裂 (上眼瞼下端 8 点と下眼瞼上端 11 点の計 19 点)



(c) 下眼瞼 (下眼瞼上端 11 点とそれに平行な点 11 点の計 22 点)

図 1 各部分領域における特徴点と三角パッチ化

1. はじめに

目の高精度な運動解析や解剖学的解析は、医学、バイオメトリクス、心理学の様々な分野で応用される重要な技術である。画像を用いたこれらの技術では、カメラに対する向き、照明条件、さらに顔に特有の表情に関して、その影響を吸収することが研究の主眼とされてきた。中でも表情を伴う目は、それが平静の見え(元々の構造)であるか、表情(運動)を伴った見えであるかが区別できないために困難である。

著者ら [1] は、目領域画像の現象を、目の静的個人的特徴である構成成分と、その動的特徴である運動成分とに分けてモデル化し、生成型の目領域モデルを用いて目の運動追跡を実現した。著者らの方法では、結果として目の構造と運動の両方を得ることができるが、運動追跡が自動で行われていたのに対し、構造モデルの初期化は手動で行われなければならない。

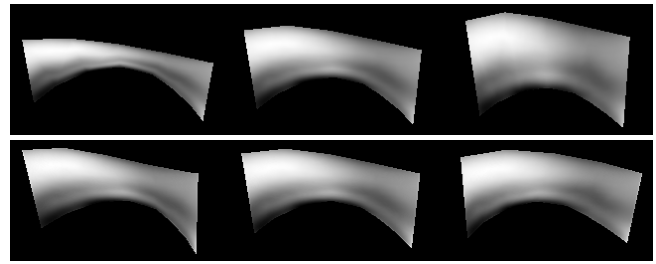
Active Appearance Model (AAM) [2] は、顔や臓器といった変形を伴う対象の画像を低次元で表現し、さらに、学習サンプルに十分近い新しい画像を生成できる。本論文では、AAM で多様な目領域画像の見えを表現し、AAM 空間で未知の目領域画像に最も近い学習サンプルを見つけることで、間接的に入力構造情報を自動取得する方法を提案する。

2. AAM による目領域画像のモデル化

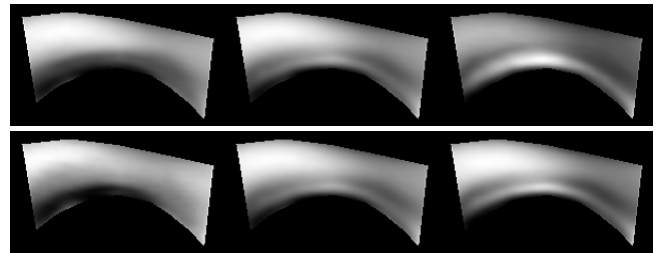
AAM は、画像上で平均的な形状から様々に変形する物体(例えば顔)に関して、その形状と内部の明度分布を同時に低次元で表現することのできる統計モデルである。

2.1 AAM の構築

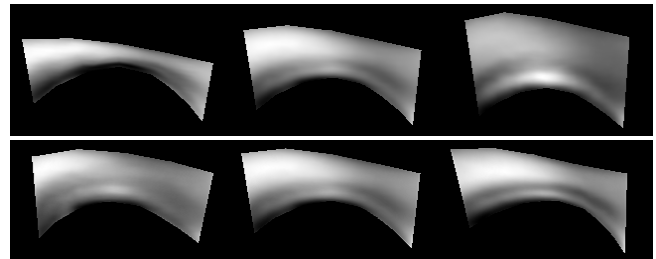
画像中の物体の輪郭を構成する点の座標がすべて与えられた



(a) (上から) 第 1, 2 形状パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)



(b) (上から) 第 1, 2 明度分布パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)



(c) (上から) 第 1, 2 appearance パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)

図 2 上眼瞼領域について構築した AAM

学習サンプル群があるとする。各点の座標を並べたベクトルを x で表すと、 x は形状を表し、これを全ての学習サンプルについて平均して得られる \bar{x} を平均形状と呼ぶ。 x は、 \bar{x} からの偏差を主成分分析して得られる固有ベクトル P_s を用いて、

$$x = \bar{x} + P_s b_s \quad (1)$$

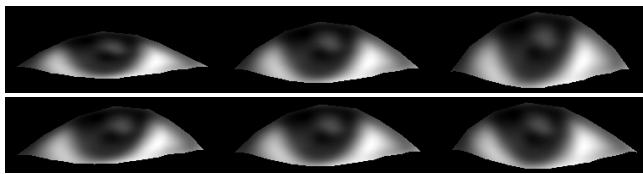
と表すことができる。ここで、主成分得点ベクトル b_s を形状パラメータと呼ぶ。

次に、すべての学習サンプルを共通の形状(例えば平均形状)に変形すると、形状について正規化された明度分布 g を得る。 g についても、それらを平均した平均明度分布 \bar{g} からの偏差を主成分分析して得られる固有ベクトル P_g を用いて、

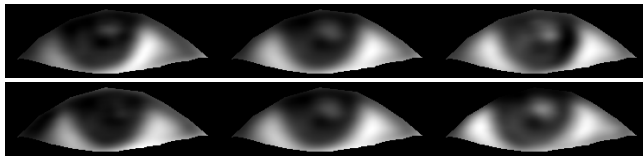
$$g = \bar{g} + P_g b_g \quad (2)$$

と表すことができる。ここで、主成分得点ベクトル b_g を明度分布パラメータと呼ぶ。

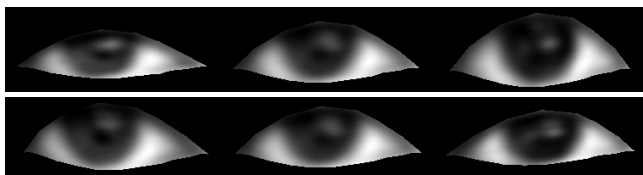
最後に、形状とテクスチャ間の単位を正規化する(形状パラメータを単位変動させた時の明度分布の変動量を与える)行列 W_s を用いて、式 (1) と (2) より、



(a) (上から) 第 1, 2 形状パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)



(b) (上から) 第 1, 2 明度分布パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)



(c) (上から) 第 1, 2 appearance パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)

図 3 眼瞼裂領域について構築した AAM

$$\mathbf{b} = \begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} = \begin{pmatrix} \mathbf{W}_s \mathbf{P}_s^T (\mathbf{x} - \bar{\mathbf{x}}) \\ \mathbf{P}_g^T (\mathbf{g} - \bar{\mathbf{g}}) \end{pmatrix} \quad (3)$$

を作る。これを主成分分析すると、物体画像中の位置とその明度分布の間の相関関係から、

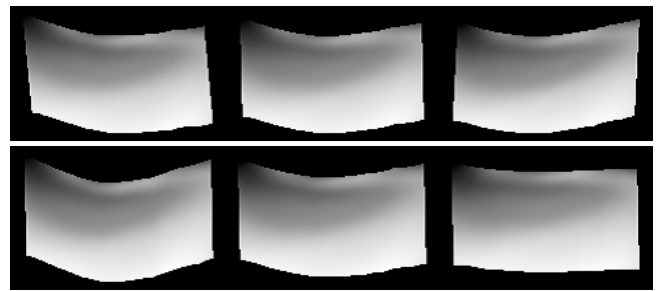
$$\mathbf{b} = \mathbf{Q} \mathbf{c} \quad (4)$$

と表すことができる。ここで、 \mathbf{Q} は固有ベクトルであり、 \mathbf{c} は、物体の画像上の形状と明度分布の両方を表し、appearance パラメータと呼ぶ。

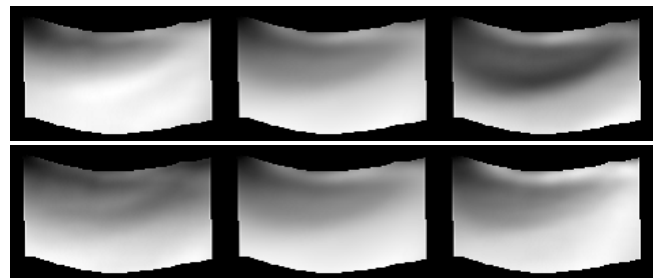
構築した AAM を用いて、与えた appearance パラメータから式 (4) で共通形状の明度分布を求め、それをさらに式 (1) で最終形状へ逆変形することで物体の画像を合成することができる。また、入力として与えた物体の画像に近い AAM を探索することもできる。この場合、最急降下法による近傍探索が尤もらしい解に収束するように、何らかの方法で、最終形状に十分近い初期形状を与える必要がある [2]。AAM が未知の入力を良く復元するためには、学習データが十分な多様性を有する必要がある、さらに、モデルの柔軟性を保ちつつオーバーフィッティングを回避するように、式 (1) と (2) における主成分数を決定する必要がある [3]。

2.2 AAM の目領域画像への適用結果

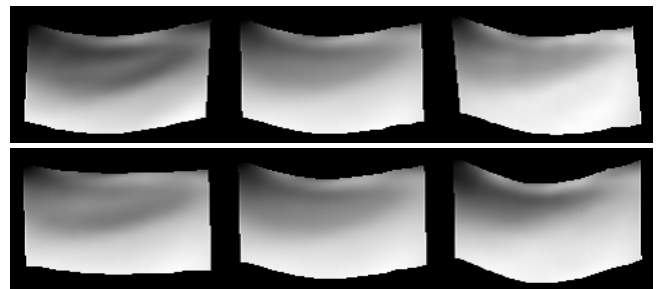
本手法では、目の各部位の微細な多様性について、十分に多様な学習サンプルを確保するために、目領域を上眼瞼、眼瞼裂、そして下眼瞼の 3 つの部分領域に分割し、それぞれ独立に AAM の構築及び後の処理を行うこととする (図 1)。また、顔はカメラに正対しているものとし、手動もしくは自動の方法に



(a) (上から) 第 1, 2 形状パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)



(b) (上から) 第 1, 2 明度分布パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)



(c) (上から) 第 1, 2 appearance パラメータ ($-3\sigma \sim 0 \sim +3\sigma$)

図 4 下眼瞼領域について構築した AAM

よって眉毛及び目の両端点が与えられているものとする。

2.2.1 用いる顔画像データベース

Cohn-Kanade AU-Coded Facial Expression Database [4] からの 102 名 (18 歳から 50 歳の男性 35 名、女性 67 名) のうち、69 名 (白人 59 名、黒人 8 名、他 2 名、男性 25 名、女性 44 名) を用いた。本データベースには、被験者が実験室内で指示されて表出した AU (Action Units) [5] を撮影した画像シーケンスが含まれており、AAM の構築には、被験者ごとに一つの画像シーケンスの第一フレーム (開眼で瞳が中央に位置) を用いた。

2.2.2 目の部分領域への特徴点の付与と三角パッチ化

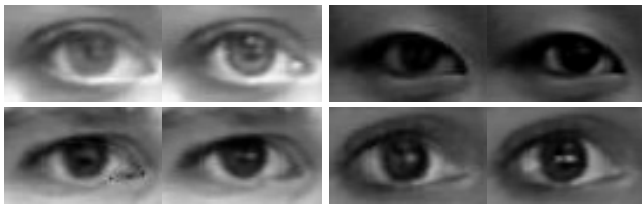
上眼瞼領域は、まゆげの下端 5 点と上眼瞼下端 8 点の計 13 点 (図 1(a))、眼瞼裂領域は、上眼瞼下端 8 点と下眼瞼上端 11 点の計 19 点 (図 1(b))、下眼瞼領域は、下眼瞼上端 11 点とそれらから下眼瞼上端の幅の半分の長さ下りた点 11 点の計 22 点 (図 1(c)) でそれぞれ定義する。2.2.1 節の各画像についてこれらの点の座標を付与し、これを学習サンプルとして用いる。

表 1 累積寄与率が 90%を超えた AAM の次元数

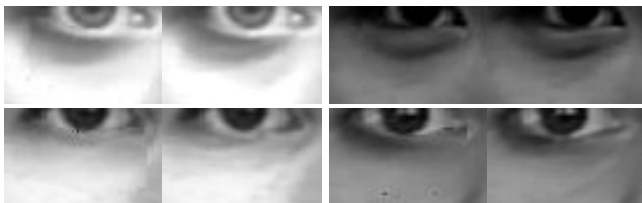
基底	上眼瞼領域	眼瞼裂領域	下眼瞼領域
形状パラメータ	4	2	5
明度分布パラメータ	12	16	14
appearance パラメータ	17	19	20



(a) 上眼瞼領域，一重/二重，種々の二重を表現できている



(b) 眼瞼裂領域，瞳内部の種々の模様を表現できている



(c) 下眼瞼領域，膨らみや陰影を表現できている

図 5 AAM 探索による近似結果の例 (左: AAM, 右: 入力)

2.2.3 AAM の構築

3つの部分領域それぞれについて，AAM を構築した。この際，眉毛及び目の両端点，計 4 点を固定する Constrained AAM [6] を用いた。その結果，固有値の累積寄与率 90%以上となる次元数は表 1 に示す通りになった。形状パラメータ b_s ，明度分布パラメータ b_g ，そして appearance パラメータ c のうち，寄与率の大きい順に上位 2 パラメータをそれぞれ $\pm 3\sigma$ 変化させた時の AAM 合成画像を図 2~図 4 に示す。

2.2.4 AAM 探索による目領域画像の近似

次に，構築した AAM を用いて目領域画像を近似することを考える。目領域画像 I に対して，AAM で合成された画像 I_m とすると， c を変えて差分ベクトル $\delta I = I - I_m$ を最小化することで，入力を最も良く近似する AAM を求めることができる [2]。そこで，平均 AAM を与えられた眉毛及び目の両端点の座標を元に目の近傍に配置し，あらかじめ求めておいた c の変分 δc と δI の関係から appearance パラメータ c を更新し，未知 (学習外) 目領域画像を近似する AAM を探索する実験を行った。結果の例を図 5 に示す。4 例それぞれにおいて，右の入力に対して収束した AAM を左に示す。様々な形状と明度

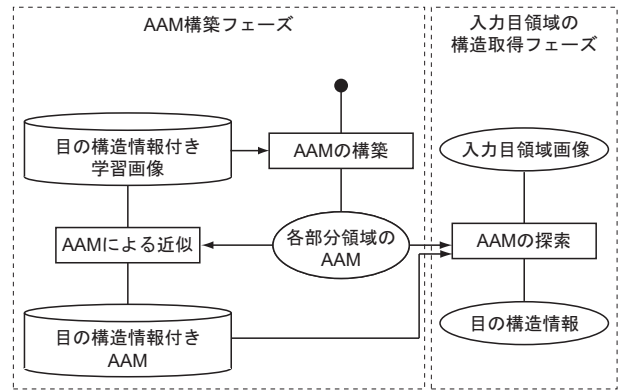


図 6 AAM 探索による目領域構造の自動抽出

分布について，AAM が高い表現能力を有していることがわかる。ここで注意すべきなのは，これらは単にそれぞれの領域の輪郭と内部の明度分布が再現できたに過ぎず，目の各部位の構造 (例えば瞳の位置やサイズ) が抽出できたわけではないことである。これを行う一つの方法として，学習サンプルに目領域の構造をあらかじめ埋め込んでおき，AAM 空間で入力に最も近い学習サンプルを見つけることによって，間接的に入力の目の構造を決定する手法を提案する。

3. AAM 探索による目領域構造の自動抽出

図 6 に示すように，2.2 節の AAM 構築フェーズでは，AAM を構築すると共に，各学習サンプルについて AAM 探索を行って，appearance パラメータを求めておく。入力目領域の構造取得フェーズでは，構築した AAM を用いて，入力目領域画像を最も良く近似する AAM を探索し，appearance パラメータ空間内で，入力に最も近い学習サンプルを探し，学習サンプルに付与された目領域構造に関する情報を入力目領域構造情報として出力する。

3.1 上眼瞼の構造とそのモデル

上眼瞼 (upper eyelid) は，眼瞼裂 (palpebral fissure) の上部の皮膚であり，眼輪筋 (orbicularis oculi muscle) が皮下脂肪で覆われた構造をしている。眼輪筋は，上から眼窩部 (orbital orbicularis)，隔膜前部 (preseptal orbicularis)，瞼板前部 (pretarsal orbicularis) の三つの部分に分かれ，さらに瞼板前部が上部と下部に分かれる。開瞼 (目を開くこと) の際には，瞼板前部の上部と下部の境界で折りたたまれて筋 (crease, 上眼瞼溝; sulcus suprapalpebralis) ができる。

この上眼瞼溝 (二重のしわ) の位置が低い場合，東アジア人に特徴的な一重まぶた (もしくは奥二重: 東アジア人全体のおよそ 50%) になり，位置が高い場合は幅の広い二重まぶたになる。上眼瞼溝の深さは，上眼瞼下部の瞼板 (superior tarsal plate) と皮膚の間の結合強度と，結合に関与しない上部の皮膚と皮下脂肪の量によって決まる。結合強度がより強く，かつ皮膚と皮下脂肪の量が多い場合，上眼瞼挙筋 (levator muscle) が上眼瞼を上へ引き上げる際，下部の皮膚，眼輪筋及び瞼板が一つの塊としてより高く引き上げられ，より多くの皮膚がその上にせり出して，上眼瞼溝が深く“濃い”二重となる。

すなわち、上眼瞼の見えの多様性を与える構造には、上眼瞼溝の位置と濃さの二つがあると考えられる。著者らが提案した目領域モデル [1] では、上眼瞼の下辺と上眼瞼溝を共に折れ線でモデル化した上で、目の運動追跡を精度良く行う上で必要なパラメータとして、両線分間の距離 (d_u) としわの濃さ (f ; 上眼瞼溝の太さと、両線分で囲まれた領域の明るさ) の二つのパラメータを提案している。そこで提案法においても、これら二つのパラメータ $\{d_u, f\}$ を上眼瞼の構造情報とする。

3.2 下眼瞼の構造とそのモデル

下眼瞼 (lower eyelid) は、下眼瞼上辺の瞼板 (inferior tarsal plate) が、目の両端点の腱 (medial canthal tendon 及び lateral canthal tendon) を支点にして、眼瞼裂の下部を覆う皮膚を吊り下げたハンモックようになっており、その瞼板は、眼窩隔膜 (orbital septum) によって眼窩 (眼球を入れる頭蓋骨のくぼみ) の縁 (orbital rim) につながっている。皮下の眼輪筋の長さが収縮して皮膚が盛り上がり、加齢によって弛んだりすると、下眼瞼に膨らみが生じ、眼窩下溝 (頬の上にある斜めのしわ, infraorbital furrow) がより深くなる。また、膨らみの上下幅には生まれつきの個人差もある。

すなわち、下眼瞼の見えの多様性を与える構造には、この膨らみの幅や高さ (奥行き方向) があると考えられる。著者らのモデルでは、下眼瞼の上辺と眼窩下溝を共に折れ線でモデル化した上で、両線分間の距離 (d_b) をパラメータとしている。また、著者らのモデルでは、照明がこの膨らみに反射して周辺に比べてより明るくなる場合も考慮し、その領域の下辺を線分でモデル化して、この線分と下眼瞼の上辺の距離 (d_r) も考慮し、これによって目の運動追跡の精度が向上することを示している。従って提案法においても、膨らみの幅と照明反射による明るい領域の幅の二つのパラメータ $\{d_b, d_r\}$ を下眼瞼の構造情報とする。

3.3 眼瞼裂の構造とそのモデル

眼瞼裂 (palpebral fissure) は、上眼瞼と下眼瞼に挟まれた開口部である。眼瞼裂から見える部位は、大きく二つの部分、すなわち虹彩 (iris) 及び強膜 (白目; sclera) に分けられ、強膜の鼻側の内端点には涙器による赤褐色の領域がある。虹彩の大きさは、眼球の大きさと共に年齢差及び個人差がある。虹彩の中心には、黒色の空洞である瞳孔 (pupil) があり、網膜 (retina) に届く光の量によって大きさが変化する。虹彩の色は、色素の量によって異なり、色素が最も多い場合は茶になり、最も少ない場合は青になる。また色素が少ない場合には、虹彩の様相がより明瞭に見える。さらに虹彩には、角膜 (cornea) 及び水晶体 (lens) の反射像 (プルキンエ像) が映る。

すなわち、眼瞼裂の見えの多様性を与える構造には、主に虹彩の大きさと内部の様相の二つの要因があると考えられる。著者らのモデルでは、虹彩内部の様相については考慮せず、虹彩を一つの明度値を持つ円盤としてモデル化した上で、その大きさ (r_i : 半径) と明度値 (I_{r7}) の二つのパラメータを提案している。提案法においても、これら二つのパラメータ $\{I_{r7}, r_i\}$ を眼瞼裂の構造情報とする。

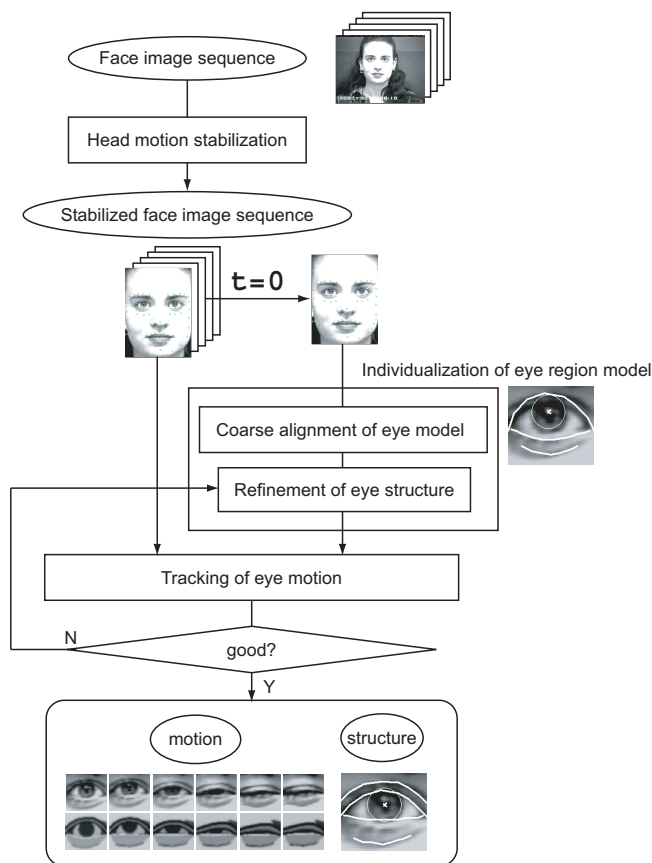


図7 著者らの生成型目領域モデルによる目の運動追跡法 [1]

3.4 AAM 探索による構造情報の取得

2.2.1 節のデータベースのうち、残りの 33 名 (白人 28 名, 黒人 4 名, 東アジア人 1 名, 男性 10 名, 女性 23 名) をテストサンプルに用いた。結果の例として、図 5 と同じ入力画像に対して AAM 空間内で最も入力に近いものとして選択された学習サンプル, そしてそれらに付与されていた構造パラメータ, さらにその構造パラメータで初期化した著者らの目領域モデルを表 2 に示す。選択された学習サンプルの列は、各々上から順に上眼瞼, 眼瞼裂, 下眼瞼に対して選択された学習サンプルを、入力との比較のために手で切り出して表示した。

表 2(a), (c), (d) の上眼瞼についてみると、入力画像に非常に良く似た学習サンプルが選択されているが、表 2(b) は一重まぶたの入力に対して二重のものが選ばれている。これは学習サンプルに一重まぶたの被験者が含まれなかったためである。しかし二重である点を除けば、学習サンプルのうちで最も入力に近いものであった。用いるデータベースに一重まぶたの被験者を導入することで対処できると考えられる。

その他については、各部分領域とも、AAM 空間内で appearance パラメータの類似性を評価することによって、入力に似た学習サンプルを選択することができており、本手法によって、未知の目領域の目領域の構造を取得することが可能であることが示されている。

3.5 目の運動追跡システムへの応用

図 7 に示すように、著者らの目の運動追跡システム [1] は、頭部運動を正規化した入力顔画像列の最初のフレームに対して、

表 2 AAM 探索による構造抽出結果の例

	入力目領域画像	選択された学習サンプル	取得した構造パラメータ
(a)			$d_u = 0.5$ $f = 1.0$ $I_{r,7} = 113$ $r_i = 10$ $d_b = 0.66$ $d_r = 0.1$
(b)			$d_u = 0.7$ $f = 0.6$ $I_{r,7} = 77$ $r_i = 10$ $d_b = 0.66$ $d_r = 0.1$
(c)			$d_u = 0.8$ $f = 1.0$ $I_{r,7} = 50$ $r_i = 10$ $d_b = 0.0$ $d_r = 0.15$
(d)			$d_u = 0.5$ $f = 1.0$ $I_{r,7} = 50$ $r_i = 10$ $d_b = 0.1$ $d_r = 0.1$

目領域モデルの構造パラメータを手動で入力に合わせ (individualization), 残りのフレームについて, 前のフレームを初期値として, 現フレームの運動パラメータを逐次最急降下法により求めるといふ, 目の運動をモデルパラメータのドメインで追跡する (tracking of eye motion) 手法である.

提案手法を用いて, 構造パラメータの初期化を自動化し, 目の運動追跡をした結果の例を表 3 に示す. 構造の取得が概ね良好だった例のうち, 表 3(c), (d) では, 手動で初期化した場合と同様の良い結果となった. 表 3(a) は, 目を見開く表情 (AU5) が強くなるに従って, 上眼瞼溝の位置が上昇し, 二重の幅が大きく拡大する場合に相当するが, さらに上眼瞼が照明を反射して明るくなり, 運動パラメータの解の空間が一層不安定になったと考えられる. 構造パラメータの初期化に関する提案手法の精度を検討する必要がある. 表 3(b) では, 初期化時に一重の入力に二重のモデルを与えてしまったため (表 2), 本来の上眼瞼に二重側がマッチし, それがモデル全体を押し下げたために下方へずれる結果となった.

表 3 提案手法と手動の間の目の運動追跡結果の比較

	第 1 フレーム	最終フレーム	手動初期化	提案法初期化
(a)				
(b)				
(c)				
(d)				

4. ま と め

表情を伴う目は, それが平静の見え (元々の構造) であるか, 表情 (運動) を伴った見えであるかが区別できないために困難であった. 本論文では, 目領域を, 上眼瞼, 下眼瞼, そして眼瞼裂の 3 つの部分領域に分け, その見えからそれぞれ解剖学的な構造を自動的に取得する方法を提案した. 本手法は, あらかじめ解剖学的な構造の情報を付与した学習サンプルから Active Appearance Model (AAM) を構築する. そして未知の入力目領域画像を近似する AAM を探索し, AAM 空間内で最も入力に近い学習サンプルから, 入力目領域の構造に関する情報を間接的に取得するものである. 実験により, 構築した AAM が目領域の微細な構造を良く表現でき, 入力と構造の似通った学習サンプルから, 目領域の構造情報を取得できることを示した. 今後は, 目の運動追跡システムにおける構造モデル自動初期化モジュールの観点から, 構造取得の性能を定量評価する.

謝 辞

本研究は, 米国国立精神保健研究所補助金 R01 MH51435, 文部科学省科学技術振興調整費「環境情報獲得のための高信頼性ソフトウェアに関する研究」の支援により行った.

文 献

- [1] T. Moriyama, T. Kanade, J. Xiao and J. F. Cohn: "Meticulously detailed eye region model and its application to analysis of facial images", IEEE Transactions on Pattern Analysis and Machine Intelligence, **28**, 5, pp. 738–752 (2006).
- [2] T. F. Cootes, G. J. Edwards and C. J. Taylor: "Active appearance models", Proc. of the 5th European Conference on Computer Vision (Ed. by H. . B. Neumann), Vol. 2, Springer, pp. 484–498 (1998).
- [3] J. L. Horn: "A rationale and test for the number of factors in factor analysis", Psychometrika, **30**, pp. 179–185 (1965).
- [4] T. Kanade, J. F. Cohn and Y. Tian: "Comprehensive database for facial expression analysis", Proc. IEEE Face and Gesture '00, pp. 46–53 (2000).
- [5] P. Ekman and W. Friesen: "Facial Action Coding System", Consulting Psychologists Press, Palo Alto, CA (1978).
- [6] T. F. Cootes and C. J. Taylor: "Constrained active appearance models", International Conference on Computer Vision 2001, Vol. I, Vancouver, Canada, pp. 748–754 (2001).