

Haar ウェーブレットを用いた顔認識

松本 雄大[†], 岩井 儀雄[†], 谷内田 正彦[†]

[†] 大阪大学大学院基礎工学研究科システム創成専攻

〒 560-8531 大阪府豊中市待兼山町 1-3

y-matsu@yachi-lab.sys.es.osaka-u.ac.jp, iwai@sys.es.osaka-u.ac.jp,

yachida@sys.es.osaka-u.ac.jp

顔認識の特徴量には様々なものが利用されているが、本稿では、計算量が少なく、簡便に処理が行える Haar ウェーブレット特徴量を用いた顔認識システムを提案する。提案手法では、正規化された顔領域に対し Haar ウェーブレット変換を行うことにより特徴量を抽出し、事前に AdaBoost により生成された個人毎の分類器によって認識を行う。また、抽出した特徴量のうち、分類精度の高いものだけを選択して利用することで、計算コストの削減を行う。顔認識実験を行い、提案システムの有効性を確認した。

Face recognition by using Haar wavelets

Yuta MATSUMOTO[†], Yoshio IWAI[†], Masahiko YACHIDA[†]

[†] Graduate School of Engineering Science, Osaka University

1-3 Machikaneyama, Toyonaka, Osaka 560-8531 Japan

y-matsu@yachi-lab.sys.es.osaka-u.ac.jp, iwai@sys.es.osaka-u.ac.jp,

yachida@sys.es.osaka-u.ac.jp

Many image features have been used for face recognition in recent decades. We propose a method for face recognition with Haar wavelet for low computation cost. This method extracts image features by performing Haar wavelet on a normalized face image, and then recognizes a face by each classifier learned by AdaBoost in advance. The computation cost for face recognition is significantly reduced by selecting image features with high confidence. We show the performance of the proposed method by experimental results.

1 はじめに

近年、犯罪の凶悪化、巧妙化が進み、セキュリティに対する社会的関心が高まっている。従来の個人認証には鍵、カード、パスワードなどが用いられてきた。このうち、鍵、カードを利用した認証では、それ自体を保持しておく必要があり、紛失や盗難等の危険性がある。さらには偽造、貸与による本人以外の不正使用といった問題がある。パスワードを利用した認証では、覚えやすく簡単な文字列をパスワードとして設定する場合が多く、また、文字列のパターン総数には限界があるため、第三者に推測されるという問題がある。

そこで現在、生体情報を利用したバイオメトリクス認証が注目されている。生体情報は個人毎に異なっているため、バイオメトリクス認証は従来の個人認証手法と比べなりすましにくく、認証に必要な情報

を本人が記憶したり、所持しておく必要がないため利便性も高い。バイオメトリクス認証の代表的なものとして、指紋、血管パターン、虹彩、顔を用いたものが挙げられる。このうち顔を用いた認証は、我々人間が他人を認識する際に普段から行っている方法であり、最も自然な認証方法である。また、非接触・非拘束で認証を行えるため、ユーザビリティが高い。さらに、顔画像が記録されることで、犯罪の抑止効果が期待できる。本研究では、これらの特性をもつ顔認証に焦点をあてる。

顔認識には様々な研究が行われてきた。主成分分析 (PCA) により特徴次元数の圧縮を行う Eigenface[1], PCA を局所表現にした局所特徴分析 (LFA) を用いたもの [2], 独立成分分析 (ICA) を用いたもの [3], 判別分析を用いた Fisherface[4], 隠れマルコフモデル (HMM) を用いたもの [5], SVM を用いたもの [6], Gabor ウェーブレット変換を用いたもの [7] などが挙

げられる。Gabor ウェーブレット変換を用いた方法では、顔認識の特徴量として Gabor フィルタと入力信号の内積を用いており、これによってパターンの局所的な周期性と方向性を抽出することができる。この Gabor フィルタは、生体の初期視覚系におけるパターン処理を近似したモデルであり、画像の照明変動の影響を受けにくいという利点があるため、顔認識の分野で広く利用されている。しかし、各種ウェーブレット (Gabor, Haar, French hat, Mexican hat, Daubechies, Symlet, Coiflet, O-spline) の認識性能を比較した内海の研究 [8] により、Haar ウェーブレットが最も良い認識率を示すという結果が得られた。さらに、照明変化にもロバストである上、複雑な計算が不要なので、計算コストの面でも優れた性能を示した。

よって本研究では、代表的な顔認識の特徴量であった Gabor ウェーブレットに替わって、Haar ウェーブレットを利用し、高速顔認識システムを構築することを目的とする。正規化された顔領域に対し、Haar フィルタを任意のサイズ、位置に配置することで特徴量を抽出し、AdaBoost によって学習した個人毎の分類器によって認識を行う。また、学習された分類器のうち、分類精度の高いものだけを利用することで、計算コストの削減を行う。

2 顔認識システム

本研究で提案する顔認識システムの概略を図 1 に示す。本章では、提案する顔認識システムの主要なモジュールについて詳細に述べる。まず、2.1 節で顔領域の特定方法、2.2 節では顔認識に用いる特徴量の抽出方法、2.3 節では認識方法について説明を行う。

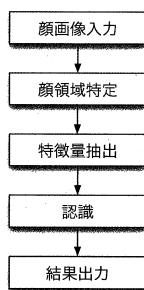


図 1: 顔認識システムの概略

2.1 顔領域特定

本システムでは、図 2 に示すように、画像上の鼻の頭の位置を中心とし、左目と右目の幅 L の二倍 $2L$ を一辺の長さとした正方形領域を顔領域とする。画像上の顔領域の大きさは撮影状況、対象人物によって異なるので、切り出された顔領域が一定の大きさになるよう正規化を行う。また、顔のエッジ部分により認識を行うため、RGB 値を輝度値に変換しグレースケール化を行う。

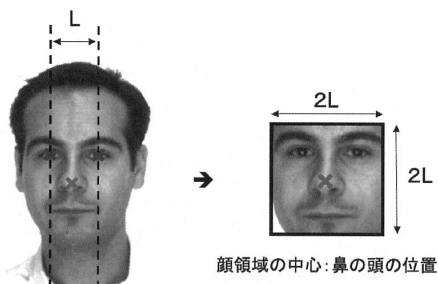


図 2: 顔領域の特定

2.2 特徴量抽出

本研究では、目、鼻、口といった顔を構成する部位の形状や配置の個人差に着目し、顔画像の局所的な特徴を認識に利用する。そこで、顔の局所的な特徴量抽出にウェーブレット変換を用いる。

本節では、Haar ウェーブレット変換について説明し、本研究で用いる Haar ウェーブレット特徴量について述べる。

2.2.1 Haar ウェーブレット変換

時間 t における入力信号 $f(t)$ のウェーブレット変換は、基本ウェーブレット $\psi_{a,b}(t)$ を用いて、

$$W_\psi = |a|^{-1/2} \int f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (1)$$

と定義される。ここで、 $\overline{\psi(\cdot)}$ は $\psi(\cdot)$ の複素共役、 a はスケールパラメータ、 b はシフトパラメータを示し、 $a, b \in \mathbb{R}, a > 0$ である。また、パラメータ a, b をそれぞれ $a = 2^j, b = 2^j k, (j, k \in \mathbb{Z})$ とおくと、

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k) \quad (2)$$

が得られ、ウェーブレット関数の時間軸方向での並行移動および拡大縮小が離散化される。これを離散ウェーブレットと呼ぶ。

次に、この離散ウェーブレットにシフトに関する直交性とスケーリングに関する直交性を持たせる。シフトに関する直交性は、

$$\langle \psi(t-k), \psi(t-n) \rangle = \begin{cases} c & n=k \\ 0 & n \neq k \end{cases} \quad (3)$$

で与えられる。また、スケーリングに関する直交性は、

$$\langle 2^{-m/2}\psi(2^{-m}t), 2^{-n/2}\psi(2^{-n}t) \rangle = \begin{cases} 1 & m=n \\ 0 & m \neq n \end{cases} \quad (4)$$

で与えられる。式(3)、(4)を満たす最も単純なウェーブレットがHaar ウェーブレットである。式と波形をそれぞれ式(5)、図3に示す。

$$\psi(t) = \begin{cases} 1 & 0 \leq t < \frac{1}{2} \\ -1 & \frac{1}{2} \leq t < 1 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

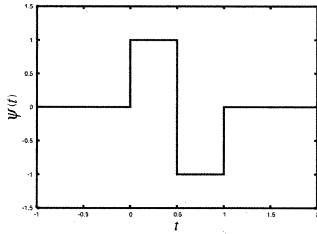


図 3: Haar ウェーブレット

2.2.2 Haar ウェーブレット特徴量

Haar ウェーブレット変換を行うことにより、特徴量が得られる。画像は二次元信号であるので、顔画像から特徴点を中心とした $n \times n$ [pixel] の範囲を切り出し、各行について横方向の一次元 Haar ウェーブレット変換を行った後、変換された信号の各列に対して縦方向の一次元 Haar ウェーブレット変換を行う。これは、図4に示す2値(黒が-1,白が1)をもつ矩形フィルタを画像上の特徴点に適用することと同値である。本研究では、矩形フィルタを $0, \pi/8, \pi/4, 3\pi/8$ [rad] 回転させた4種類を使用する。これらの矩形フィルタを、正規化された顔領域の任意の位置に、任意のサイズで配置し、特徴量を抽出する。1枚の顔画像 i か

ら得られた t 個目の特徴量 f_t^i を順に並べたベクトルを、Haar 特徴ベクトル $\mathbf{x}_i = (f_1^i \cdots f_T^i \cdots f_T^i)^T$ と表す (T は特徴量の個数)。

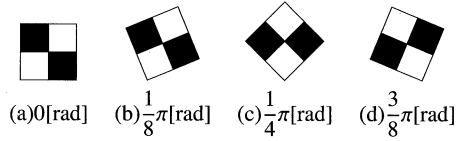


図 4: 矩形フィルタ

2.3 認識

本研究では、AdaBoost[9]により多数のサンプルを学習させた強分類器を個人毎に構成し、顔認識を行う手法を提案する。本手法の流れを図5に示す。

以下の節では、それぞれの処理の詳細を述べる。まず、AdaBoostの説明と顔認識への適応、次に得られた分類器による認識方法について説明する。

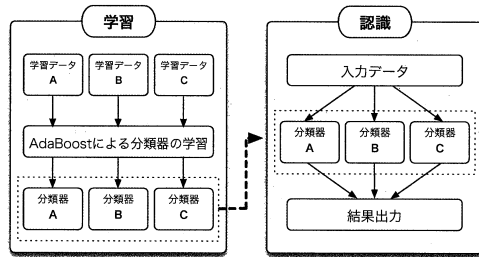


図 5: 認識の流れ

2.3.1 AdaBoost による分類器の学習アルゴリズム

本研究では、AdaBoost をクラス分類器の学習に用いる。AdaBoost は、分類器が誤分類するサンプルに指数的に更新される重みをかけることで再学習を行うアルゴリズムである。本研究で提案する分類器の学習アルゴリズムを以下に示す。

学習には、クラス k に属するポジティブサンプルを m 個、属さないネガティブサンプルを n 個使用し、サンプル i における Haar 特徴ベクトル $\mathbf{x}_i = (f_1^i \cdots f_T^i \cdots f_T^i)^T$ と、その属性 y_i の組 $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_{m+n}, y_{m+n})$ を入力データとする。なお、サンプル i がポジティブサンプルであれば $y_i = 1$ 、ネ

ガティブサンプルであれば $y_i = -1$ とする。学習済みの分類器 $H^k(\mathbf{x})$ は、弱分類器 $h_t^k(f_i)$ の重み付き線形和で表される。

$$H^k(\mathbf{x}) = \sum_{t=1}^T \alpha_t^k h_t^k(f_i) \quad (6)$$

ここで α_t^k は、弱分類器 $h_t^k(\cdot)$ の分類精度を表す重みであり、信頼度と呼ぶ。弱分類器 $h_t^k(f_i)$ は、Haar 特徴ベクトル \mathbf{x} の各要素 f_i ごとに用意し、特徴量 f_i と閾値 θ_t^k を用いた関数で表される。

$$h_t^k(f_i) = \begin{cases} 1 & f_i < \theta_t^k \\ -1 & \text{otherwise} \end{cases} \quad (7)$$

弱分類器 $h_t^k(f_i)$ の閾値 θ_t^k は、学習データにおける特徴量 f_i の最小値から最大値の範囲で、弱分類器 $h_t^k(f_i)$ の重み付きエラー ϵ_t^k が最小となる θ_t^k を閾値として設定する。AdaBoost の学習アルゴリズムをまとめたものを、図 6 に示す。なお、 t 回更新されたサンプル \mathbf{x}_i の重みを $w_t(i)$ とする。

2.3.2 認識方法

認識対象の顔画像から抽出した Haar 特徴ベクトル $\mathbf{x}_i = (f_1^i \cdots f_n^i \cdots f_T^i)^T$ を学習済みの全ての分類器で評価し、それらの中で最大値をとる分類器のクラスを認識結果とする。なお、全ての分類器の出力が負である場合は、入力サンプルは未登録者であるとみなす。

また、AdaBoost によって学習された弱分類器のうち、識別精度の高いものだけを利用することで計算コストを削減を行う。弱分類器の選択には各弱分類器 $h_t^k(\cdot)$ の分類精度を表す信頼度 α_t^k を利用するが、その選択方法は、実験結果とともに 3.2 節において述べる。

3 実験

これまで説明した手法がどのような認識率を示すか、評価実験を行った。3.1 節で実験環境について述べ、3.2 節で AdaBoost によって生成された分類器の性能評価および、分類に有効な矩形フィルタの検討を行う。

入力: $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_{m+n}, y_{m+n})$

$$(\mathbf{x}_i = (f_1^i, \dots, f_n^i, \dots, f_T^i), y_i \in \{+1, -1\})$$

重みの初期化: $w_1(i) = \frac{1}{m+n} \quad (i = 1, \dots, m+n)$

for $t = 1, \dots, T$

1. 閾値 θ_t^k の探索

for $\theta_t^k = \min_{f_i}, \dots, \max_{f_i}$

$$\epsilon_t^k = \sum_{i=1}^{m+n} w_t(i) \times |h_t^k(f_i^i) - y_i| = \sum_{i: h_t^k(f_i^i) \neq y_i} w_t(i)$$

$\theta_t^k \leftarrow \epsilon_t^k$ が最小となる θ_t^k

2. 信頼度 α_t^k の決定

$$\alpha_t^k = \frac{1}{2} \ln\left(\frac{1 - \epsilon_t^k}{\epsilon_t^k}\right)$$

3. 重みの更新

$$\begin{aligned} w_{t+1}(i) &= \frac{w_t(i) \exp(-\alpha_t^k y_i h_t^k(f_i^i))}{Z_t} \\ &= \frac{w_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t^k} & \text{if } h_t^k(f_i^i) = y_i \\ e^{\alpha_t^k} & \text{if } h_t^k(f_i^i) \neq y_i \end{cases} \end{aligned}$$

$Z_t : \sum w_{t+1}(i) = 1$ とする正規化定数

$t \leftarrow t + 1$

出力: 弱分類器の閾値 θ_t^k , 信頼度 α_t^k

図 6: AdaBoost による分類器の学習アルゴリズム

3.1 実験環境

本システムの実装は、MacBookPro(Mac OS X 10.5.1, CPU Intel Core 2 Duo 2.4GHz, MEMORY 4GB) 上で行った。撮影に用いたカメラの解像度は 640×480 [pixel] であり、各画素は 24bit の RGB 値をもつ。被験者 20 人に対し、各人 45 枚、合計 900 枚の無表情、正面顔を撮影した。AdaBoost による個人毎の分類器の学習には、一つのカテゴリにつき、ポジティブサンプル 30 枚、ネガティブサンプル 285 枚、合計 315 枚を使用した。特徴量抽出のための矩形フィルタのサイズは、一辺 8, 16, 24, 32, 40, 48, 56, 64[pixel] の 8 種類を使用した。正規化された顔領域 64×64 [pixel] に対して矩形フィルタを 4[pixel] ずつ移動させ、2720 個の特徴量を抽出した。本システムの性能評価として、認識率、FRR(false reject rate), FAR(false acceptance rate) を用いた。

3.2 実験結果と考察

表 1 に、抽出した全ての特徴量を使用して顔認識を行った結果を示す。なお、学習に使用した 600 枚の顔画像をトレーニングサンプル、学習に使用していない 300 枚をテストサンプルとする。表 1 に示したとおり、トレーニングサンプルに対して全てのサンプルを正しく分類でき、テストサンプルに対しても認識率 99.3% と、非常に高精度で分類できた。よって、AdaBoost により学習した分類器が顔認識に有効であることが明らかになった。

表 1: 分類精度

	認識率 (%)	FRR (%)	FAR (%)
トレーニングサンプル	100	0	0
テストサンプル	99.3	0.66	0

次に、分類に有効な矩形フィルタの選択について考える。AdaBoost によって学習された分類器 $H(\mathbf{x})$ は、分類精度を表す信頼度 α_t によって重み付けされた弱分類器 $h_t(f_t)$ の線形和であった。ここで、信頼度 α_t をもとに、分類精度の高い矩形フィルタから順に使用し、認識結果がどのように変化するか調べた。信頼度 α_t の値は、個人毎、各矩形フィルタ毎によって大きく異なる。そこで、以下に示す三つの指標 A_t, B_t, Γ_t によって矩形フィルタの選択を考える。学習済みの分類器の数を K 個、抽出した特徴数を T 個とする。

- $A_t = \sum_{k=1}^K \alpha_t^k$
- $B_t = \sum_{k=1}^K \beta_t^k$ ただし $\beta_t^k = \frac{\alpha_t^k}{\alpha_{max}^k}, \alpha_{max}^k = \max_t \alpha_t^k$
- $\Gamma_t = \sum_{k=1}^K \gamma_t^k$ ただし $\gamma_t^k = \frac{\alpha_t^k}{\sum_{t=1}^T \alpha_t^k}$

A_t は、全ての分類器の t 番目の矩形フィルタの信頼度 α_t の和であり、 t 番目の矩形フィルタの絶対的な分類精度を表す。 β_t^k は、クラス k における信頼度 α_t^k の最大値 α_{max}^k によって信頼度 α_t^k を正規化したものであり、クラス k における t 番目の矩形フィルタの重要度を表す。 B_t は、全ての分類器の t 番目の矩形フィルタの β_t^k の和である。 γ_t^k は、クラス k における信頼度 α_t^k の和によって信頼度 α_t^k を正規化したもの

であり、 Γ_t は、全ての分類器の t 番目の矩形フィルタの γ_t^k の和である。

テストサンプルにおいて、 A_t, B_t, Γ_t の大きい矩形フィルタから順に使用していった場合の認識率の変化を図 8, 9, 10 に示す。実験の結果、特徴数 150 個において A_t を指標とすると認識率 99.6%, B_t では 98.6%, Γ_t では 99.3% という結果が得られた。つまり、2720 個の矩形フィルタのうち、分類精度の高い上位 5.5% の矩形フィルタだけで高精度に分類が行えることが分かった。これにより、認識の際に抽出する特徴量が大幅に削減されるので、処理の高速化が図れる。また、 A_t, B_t, Γ_t のうち、特徴数の少ない場合では Γ_t を指標として矩形フィルタを選択した場合が最も良い分類性能を示した。図 11 に、 Γ_t の上位 30 個の矩形フィルタを示す。使用した矩形フィルタのうち、分類精度の高い矩形フィルタは目や鼻、口の周辺に多く見られた。これは、顔を構成する部品の形状や配置の個人差を Haar ウェーブレット変換によってうまく抽出できたことを示している。矩形フィルタのサイズは 8 種類を用いたが、16, 24, 32, 40, 48 [pixel] の矩形フィルタが分類に有効であることが分かった。64×64 [pixel] の顔領域に対し、56, 64 [pixel] といった大きい矩形フィルタを用いると、顔を構成する部位の形状や配置といった局所的な特徴を抽出できず、得られた特徴量に個人差が十分に生じないため認識には向かない。一方、8 [pixel] といった小さい矩形フィルタでは、画像上のノイズや正規化を行う際に生じる誤差の影響を受けやすいと考えられる。これらの影響への対策として、顔の傾き補正やヒストグラム平坦化、ラプラシアン・ガウシアンフィルタの適用等の前処理を行うことで、各矩形フィルタの分類精度は改善されると考えられる。

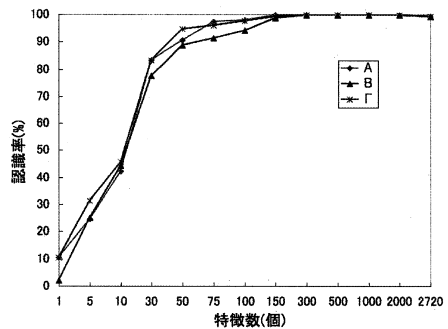


図 7: 特徴数と認識率の関係

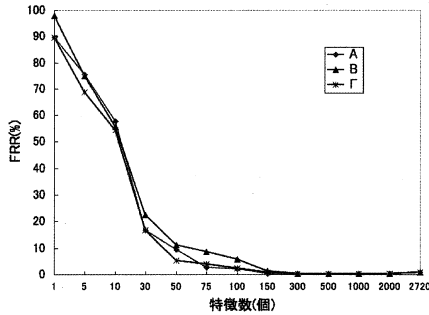


図 8: 特徴数と FRR の関係

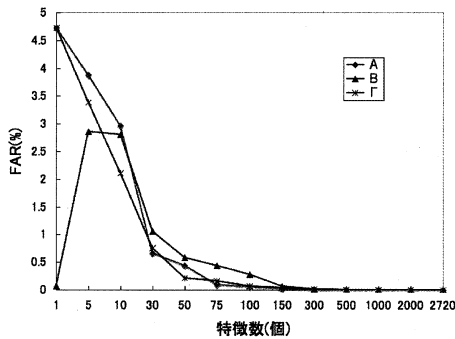


図 9: 特徴数と FAR の関係

4 おわりに

本論文では, Haar ウェーブレット特徴量を用いた顔認識システムを提案した. 実験の結果, 抽出した全ての特徴量を認識に使用した場合には, テストサンプル 300 個に対し認識率 99.3%, FRR 0.666%, FAR 0%となり, 非常に高精度な認識を行うことができた. さらに, 使用した矩形フィルタ 2720 個のうち, 分類精度の高い矩形フィルタ上位 150 個によって認識を行った場合には, 認識率 99.3%, FRR 0.666%, FAR 0.035%となり, 全ての特徴量を使用した場合と同等の認識結果が得られることが確認できた.

今後は, 個人毎の分類器を統合した認識木生成アルゴリズムの検討および, 瞳の位置検出による顔領域特定の自動化, 照明変動への対応, 顔の傾き補正処理を行う.

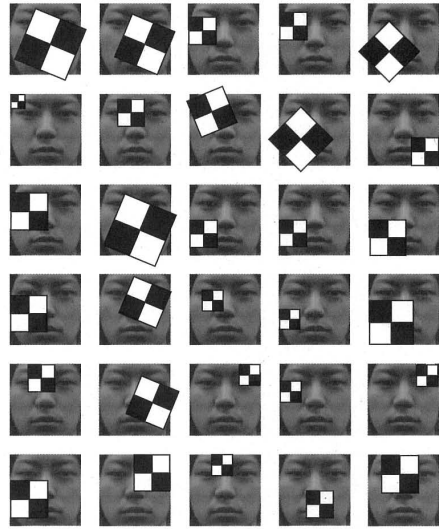


図 10: 分類精度上位 30 個の矩形フィルタ

参考文献

- [1] M. Turk and A. Pentland, "Eigenface for recognition", *Journal of Cognitive Neuroscience*, Vol. 3, No.1, pp. 71–86, 1991.
- [2] P. Penev and J. Atick, "Local feature analysis : A general statistical theory for object representation", *Netw., Comput. Neural Syst.*, pp. 477–500, 1996.
- [3] M. S. Bartlett, J. R. Movellan and T. J. Sejnowski, "Face Recognition by Independent Component Analysis", *IEEE Trans. Neural Networks*, Vol. 13, No. 6, pp. 1450–1464, 2002.
- [4] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces : Recognition Using Class Specific Linear Projection", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 19, No.7, pp. 711–720, 1997.
- [5] R. Samaria and S. Young, "HMM based architecture for face identification", *Proceedings of Image and Computer Vision*, Vol. 12, pp. 537–543, 1994.
- [6] M. Pontil and A. Verri, "Support vector machines for 3D object recognition", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 20, No. 6, pp. 637–646, 1998.
- [7] L. Wiskott, J. M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 19, No. 7, pp. 775–779, 1997.
- [8] 内海ゆづ子, "顔認識のためのウェーブレット特徴量の評価", 大阪大学大学院基礎工学研究科修士学位論文, 2007.
- [9] J. Friedman, T. Hastie and R. Tibshirani, "Additive Logistic Regression: a Statistical View of Boosting", *The Annals of Statistics*, Vol. 28, No. 2, pp. 337–374, 2000.