

顔の確率分布推定を用いた高速顔検出手法

吉永 智明 長屋 茂喜 軽部 勲

日立製作所 組込みシステム基盤研究所

概要 顔検出では、画像中にある任意の顔を検出するために、検出する顔のスケールを変えながら画像上の全位置に顔・非顔の判定を行う必要があり、多くの処理時間を要する。この課題に対して、顔検出における多くの研究は顔・非顔判定を高速化することで処理時間の低減を図っている。これに対して本報告では、画像全体に対して効率的に顔を探査することで、高速に顔検出を行う手法を提案する。顔検出に用いるカスケード型識別器から顔の確率分布を求め、左記の確率分布に基づいて簡易探索と詳細探索を組み合わせることにより、顔・非顔判定処理を行う回数を削減し高速化を実現する。実験の結果、検出率を落とすことなく、従来の5倍以上に高速化できることを確認した。

Fast Face Detection by Using Estimation of Face Probability Distribution

Tomoaki YOSHINAGA Shigeki NAGAYA and Isao KARUBE

Embedded System Platform Research Laboratory, Hitachi, Ltd.

Abstract Face detection requires high computational complexity, because face-classification(classify face or non-face) is executed at multiple scale and multiple locations on image. Most previous researches on fast face detection are aiming at face-classification algorithm. In this paper, we propose fast face detection by using the efficient face search algorithm. The proposed algorithm can reduce the number of face-classification operations by using the combination search; rough search and fine search based on face probability distribution, which is calculated using cascade face detector. Experiments show that the proposed algorithm is 5 times faster than non-optimized search, and improves detection rate by 1.1%.

1. はじめに

顔検出は、入力画像中にある人物の顔を検出する技術であり、デジタルカメラや監視システムなど様々な用途に必要な技術である。近年では、顔認証や表情認識など、検出した顔に対するアプリケーションの研究開発も盛んになっている。このため、顔検出の重要性は更に高まっており、組込み機器上での動作のために更なる高速化が求められている。

顔検出の処理は、「顔判定」と「顔探索」の大きく2つに分けることができる。本報告において、顔判定とは、ある一定サイズの画像が顔なのか非顔なのかを識別器を用いて判定する処理と定義する。一方、顔探索とは、顔判定の処理を入力画像全体に対して行い、画像中に存在する任意の顔を見つける処理と定義する。

顔検出の多くの研究は、顔判定に対するものであり、識別器の高精度化と高速化によって顔検出の性能向上を提案している。Rowley[3]らは、ニューラルネットワ

ークを用いた識別器を提案している。また、Viola と Jones は AdaBoost によって選別した Haar-like 特徴とカスケード構造の識別器によって高速かつ高精度な顔検出を実現している[1]。また、三田らは複数の Haar-like 特徴量の共起性を用いて特徴量の個数を削減することで高速化を実現している[5]。Wu らは Real AdaBoost を使い、従来2値であった弱識別器の出力を多値化することで必要な弱識別器の個数を削減することで高速化を実現している[4]。

こうした多くの顔判定の研究における顔探索の処理は、画像上の全ての位置に対して、識別器や画像のスケールを変更しながら顔判定を行うという極めてシンプルな手法である。顔判定以外の処理による高速化については、前処理において肌色情報や目などの顔特徴を見つけることで、顔候補領域の絞込みを行う手法が提案されている[6-7]。しかし、これらの手法では顔候補除去のために顔判定とは異なる追加の演算が必要で

ある上、入力画像の絵柄によっては十分に顔候補を除去できない場合がある。

そこで、報告者らは効率的な顔探索を行うことで、様々な画像に適用可能な高速顔検出手法を提案する。提案手法では、カスケード型識別器による顔判定の結果から得られる顔確率分布に基づき、探索間隔を間引くことで処理時間を削減する。

まず2章で、カスケード型識別器における強識別器のクリア数と顔確率分布の関係に関して調査結果を示す。そして、3章では顔確率分布の特性に基づいた探索手法を提案し、4章で実験結果を述べる。

2. カスケード型識別器における強識別器のクリア数と顔確率分布の関係

本章では、カスケード型識別器を用いた顔判定の結果が、顔周辺の位置・スケールに対してどのような特性を示すかを調査し、その特性から顔探索に要する演算処理の低減を検討した。

2.1. カスケード型識別器の概要

本研究では、顔か非顔かの識別処理を行う識別器として AdaBoost によって学習されたカスケード型識別器を用いる[1]。カスケード型識別器は、複数の弱識別器によって構成される強識別器を1つの段として、複数個の段(強識別器)を従続接続する構成の識別器である(図1)。各強識別器において、ある一定サイズの画像に対して顔(True)か非顔(False)かの判定を行い、画像が顔と判定された(強識別器をクリアした)場合は次の強識別器による判定処理を行う。また、非顔と判定された場合はその段で判定処理を打ち切る。左記処理の結果、全ての強識別器をクリアした場合に、顔と判定される。カスケード型識別器では、画像中のほとんどの領域を占める非顔の領域を少ない演算回数で打ち切り判定ができるため、高速に顔判定が行えるという利点がある。

2.2. カスケード型識別器の出力

カスケード型識別器による顔判定では、顔との類似度が低い画像ほど少ない強識別器で判定が終了し、顔との類似度が高い画像ほど多くの強識別器を用いて判定されることになる。つまり、顔判定においてクリアした強識別器の数は、顔である確率を示していると言える。このクリアした強識別器の数を、本報告では「クリア数」と呼ぶことにし、画像上の各位置で得られるクリア数の分布を「顔確率分布」と呼ぶことにする。

上記の考え方に関して、高塚らは、AdaBoost 学習で

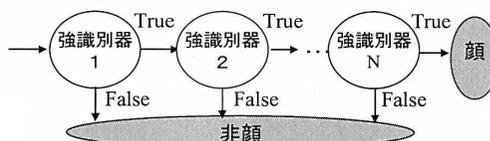
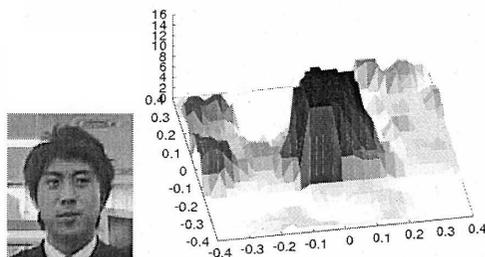


図1 カスケード型識別器



(a)顔画像 (b)クリア数として得られる顔確率分布

図2 顔周辺位置における顔の確率分布

作成した500個の弱識別器の出力値の総和を「顔らしさ」と定義し、この顔らしさの分布を用いて高精度化する手法を提案している[2]。高塚らは、顔らしさが、顔周辺の位置・スケールに対して分布する様子が顔と非顔で異なることを利用することで誤検出を削減し、検出精度を向上させている。

これに対し、カスケード型識別器におけるクリア数は、打ち切りのある顔・非顔判定の演算結果として直接得られ、かつ上記の「顔らしさ」を示すものと考えられる。クリア数から得られる顔確率分布の位置・スケールに対する特性を調査した。

2.3. 位置に対する顔確率分布の特性

まず、顔周辺の位置に対するカスケード型識別器のクリア数を実験により調査した。図2(a)の顔画像に対してカスケード型識別器による顔判定を行い、顔周辺の各位置におけるクリア数から顔確率分布を求めた。本実験ではカスケード型識別器の段数 $N=15$ とした。集計したクリア数をプロットしたグラフを図2(b)に示す。図2(b)の x, y 軸は、それぞれ画像の x, y 座標を示しており、顔画像の中心からの変位量を正解値の顔サイズで正規化した値(顔画像の中心からの相対座標÷顔サイズ)を表している。z軸はクリア数を表している。顔確率分布は正解値の顔座標(図2(b)の $x=0, y=0$)のみでクリア数が大きくなるようなインパルス形状を取るのではなく、正解値を中心としたなだらかな山のような形状を取ることがわかる。具体的には、図2(b)では正解値の顔座標から ± 0.1 程度離れた周辺の座標においても途中の強識別器までクリアしている。

2.4. スケールに対する顔確率分布の特性

次に、図 2(a)の顔画像に対して識別器のスケールを変更して顔判定を行い、クリア数から顔確率分布を求めた。識別器のスケールは、正解値となる顔のスケールに対して0.7倍から1.4倍まで0.1刻みで変化させた。図 3(a)~(h)に各スケールで識別処理を行った際の顔確率分布のグラフを示す。グラフでは、スケール=1.0倍(図 3(d))のときに顔確率分布が最も広い範囲で現れており、スケールが1.0倍から離れる、即ち正解値となるスケールから遠ざかるに従って顔の確率分布が小さくなっていくことがわかる。つまり、スケールに対する顔確率分布についても2.3節で述べた位置に対する顔確率分布と同様に、正解値のスケールから多少離れたスケールにおいても顔判定において途中の強識別器までクリアすることがわかる。

2.5. 顔確率分布を用いた演算処理削減

カスケード型識別器のクリア数は、顔判定の途中演算結果から得ることができる。つまり、あらかじめ制限したM段($M < N$)までの強識別器を用いて顔判定を行えば、少ない演算量で大まかな顔確率分布(顔確率分布が示す山の裾野部分)の形を推定できる。更に、顔確率分布は位置・スケールに対して相関があるため、この大まかな顔確率分布の推定において顔判定の間隔を間引くことが可能となる。以上の2つの演算数削減によって、少ない処理時間で顔確率分布を推定し、得られた顔確率分布に対して詳細に顔判定を行うことで顔を検出する。

3. 提案する高速顔検出手法

本章では、提案する顔検出手法について述べる。提案手法では、カスケード型識別器から得られる顔確率分布の形状を鑑み、分布の裾野を捉える簡易探索と、分布の頂点を見つける詳細探索を組合せて行うことで、顔検出に必要な顔判定の演算回数を削減する。

3.1. 提案手法の概要

提案手法における顔探索の概要を図 4 に示す。図 4 では、比較のため、従来手法と提案手法における顔探索の概念を、それぞれ図 4(a)(b)に示す。

図 4(a)の2つのグラフは上側がスケール= s に対する探索処理を、下側がスケール= $s+1$ に対する探索処理を示すものであり、どちらも横軸が画像のx座標、縦軸がクリア数を、矢印は各座標位置に対して行う顔判定処理を示している。矢印が短いほど少ない強識別器で

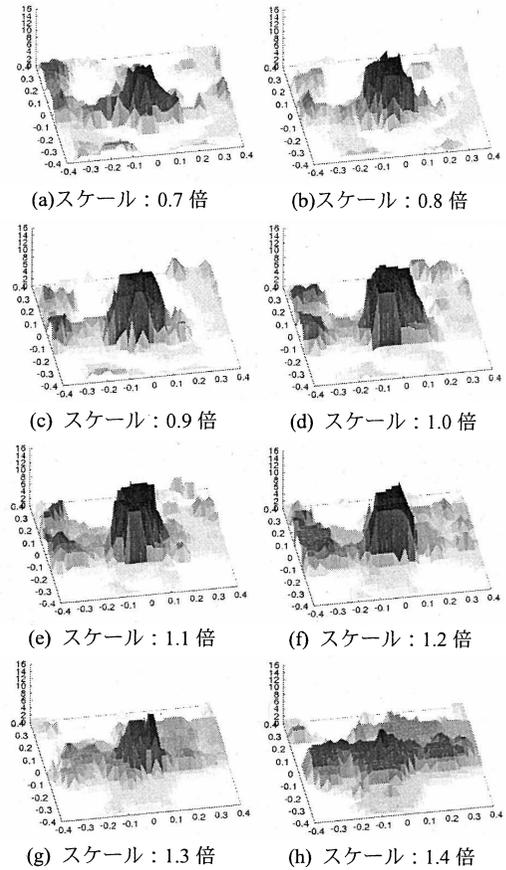


図 3 スケール毎の顔確率分布

判定が終了し、矢印が長いほど多くの強識別器を用いて顔判定を行うことを示している。

従来の顔探索では、まず、あるスケール= s においてx座標上に一定間隔毎に顔判定していくことで、クリア数がNとなる位置を顔の位置として探していく。次に、識別器のスケールを $s \rightarrow s+1$ に変えて、同様に一定間隔毎に顔判定を実行していき、異なるサイズの顔を見つけていく。このような従来の顔探索方法では、探索を行うスケールと位置の間隔を大きくしすぎると探索位置が山の顔確率分布の頂点からずれてしまい、クリア数がNとなる座標を検出できなくなる可能性がある。このため検出漏れが起きないようにするには、十分に細かい探索間隔で顔判定処理を行う必要があり、多くの処理時間を要する。

これに対して、提案手法における顔探索の概念は図 4(b)のようになる。2章で述べた顔確率分布の調査結果から、途中までの強識別器をクリアする領域は広く分

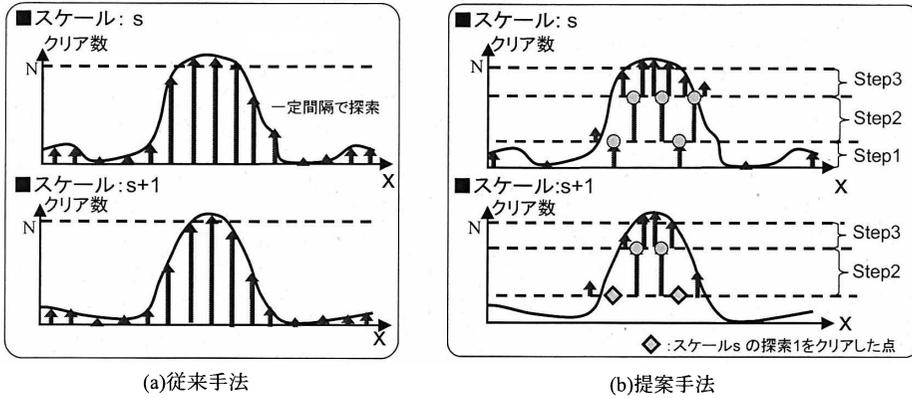


図4 顔探索の処理概念

布する。これを利用し、提案手法では最初に位置・スケールに対して大きく間引いて、途中の強識別器までを用いた顔判定を行うことで、顔確率分布の山がある位置を限定する。次に、最初の探索をクリアした領域の周囲を詳細に探索する。顔確率分布の推定において、顔判定を行う位置を大きく間引くことで演算数を削減する。

提案手法における顔探索は、以下の3つの探索ステップで構成される。

Step1: 分布位置の推定

スケール s の識別器を用いて広間隔で識別処理を行うことで、顔確率分布の山の位置を推定する。

Step2: 分布スケールの推定

Step1で見つけた顔確率分布の山の位置に対して、隣り合うスケール $s-1$, $s+1$ の顔識別器を用いてそれぞれ顔判定を行う。

Step3: 顔位置の決定

Step2をクリアした周辺位置に対して詳細な間隔で顔判定を行い、顔を検出する。

以上の流れで顔探索を行うことで、顔判定を行う回数を大きく削減する。顔判定の処理回数を示す図4の矢印の長さの総和を比べると、従来手法に比べて識別処理の回数が大きく減少できることがわかる。

3.2. 提案手法における具体的な処理の流れ

提案手法における顔探索の具体的な処理の流れを図5に、実際の画像上での処理の様子を図6に示す。

提案手法では、隣り合う3つのスケール毎に顔探索を行う。step1は隣り合う3つの真ん中のスケールに対してのみ行われ、探索2からはスケール毎に独立して顔探索が行われる。なお、隣り合うスケールで識別器の窓幅は1.2倍ずつ変えていく。また、各探索ス

顔検出を行う全てのスケール0から S に対して以下の流れで顔を探る。

●For scale=0,1,..., S/3

■exe_scale=scale*3+1としてStep1を実行
識別器の窓幅=スケール0の窓幅×1.2^{exe_scale}

Step 1

1. x , y 方向とも $4w$ 毎の間隔で強識別器 $1 \sim \alpha$ を用いて顔判定。
2. 1の顔判定をクリアした点 (x_1, y_1) とその周囲4点 $(x_1 \pm 2w, y_1 \pm 2w)$ に対して強識別器 $\alpha + 1 \sim \beta$ を用いて顔判定。

■For exe_scale=scale*3, scale*3+1, scale*3+2
識別器の窓幅=スケール0の窓幅×1.2^{exe_scale}

Step 2

探索1をクリアした座標に対して、強識別器 $\beta + 1 \sim \gamma$ を用いて顔判定。

Step 3

探索2をクリアした点 (x_2, y_2) とその上下左右の4点 $(x_2 \pm w, y_2)$, $(x_2, y_2 \pm w)$ に対して強識別器 $\gamma + 1 \sim N$ を用いて顔判定。

●End

図5 提案手法における探索処理の流れ

テップにおける探索位置の間隔は、 $4w$ (Step1-1)から w (Step3)まで探索ステップが進むに連れて小さくなるように制御する。ここで、各探索ステップにおける顔判定は、それぞれ異なる強識別器を用いる。図5に示した各探索ステップで用いる強識別器番号 α , β , γ は、 $0 < \alpha < \beta < \gamma < N$ という関係にある。この α , β , γ の値は、実験により経験的に決定した。

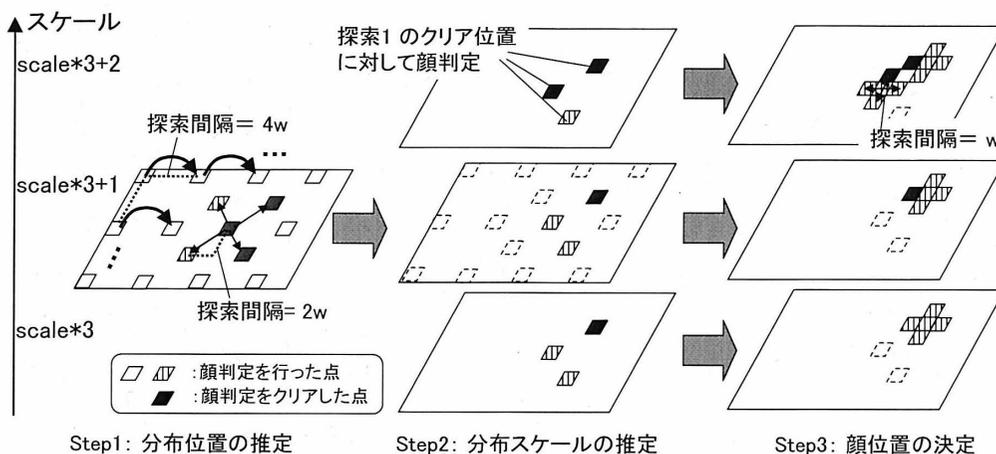


図 6 提案手法における探索動作例

4. 評価実験

4.1. 実験概要

提案手法の性能評価のため、複数の探索方法を用いて顔検出の実験を行った。本実験で用いた顔検出における顔探索方法は以下の4つである。

1. 通常探索

画像上の全位置・全スケールに対して一定間隔で識別処理を行い、顔を検出する。このときの位置とスケールの間隔は、viola らの論文[1]と同様に、スケールの間隔を 1.25 倍、位置の間隔 $w = [\text{mul_scale} * \Delta]$ と定義した。ここで、 $[\]$ は丸め関数を、 mul_scale は基準の識別器の窓幅に対するスケール倍率を、 Δ はステップ幅を決めるパラメータを表しており、 $\Delta = 1.0$ とした。

本実験では、後述のように窓幅 20x20 画素を基準スケールとしている。このため、基準スケールでの顔探索は 1 画素間隔、次の探索スケールである窓幅 25x25 画素の顔探索は、1.25 画素間隔で顔判定処理を行うことになる。

2. 位置間隔 2 倍

顔探索における探索位置の間引きの影響を評価するため、通常探索の 2 倍となる $\Delta = 2.0$ として顔探索を行う。

3. スケール間隔 1.4 倍

スケールの間引きの影響を評価するため、スケール間隔を 1.4 倍に拡大して顔探索を行う。このスケール間隔はスケールによる顔確率分布の変化から、検出漏れることが無いように経験的に決定した。

4. 提案手法による探索

前節で紹介した提案手法に基づき顔探索を行う。

実験では、全ての探索方法で同じカスケード型識別器を用いて顔検出を行った。このカスケード型識別器は、窓幅 10x10 画素サイズの識別器を AdaBoost 学習により作成し、これを 2 倍に拡大して 20x20 画素にした識別器を用いた。学習は、14000 枚の顔画像と 7000 枚の非顔画像を用いて行い、強識別器の段数 $N = 20$ のカスケード型識別器を作成した。評価用画像には、MIT+CMU の正面顔画像テストセット (画像 117 枚、顔 507 個) を用いた。評価用画像に対して顔検出を行った結果得られた検出率と計算量を比較することで、各探索方法の性能評価を行う。計算量の評価のため、全画像に対する処理時間と、弱識別器の演算回数を計測した。

4.2. 実験結果

実験結果を表 1 に示す。全画像の顔検出に要する弱識別器の演算回数は、通常探索で 878M 回だったのに対して提案手法では 109M 回と 1/8 以下に削減でき、処理時間も 61.4 秒から 11.3 秒となった。さらに、このときの顔検出率は通常探索では 84.3% だったのに対して、提案手法では 85.4% に向上した。一方、単純に位置とスケールの探索間隔を拡大した探索方法 2 と 3 では、間引かれた間隔に比例して処理時間を削減できているが、どちらも提案手法の数倍の処理時間を要する。さらに、顔検出率は、両探索方法とも 80% 以下に劣化し、提案手法を下回る結果となった。以上から、提案手法による段階的な間引き探索が処理時間においても検出精度においても有効であることがわかった。提案手法

表1 探索方法による処理時間と検出率の比較 (誤検出数=100 のとき)

探索方法	1. 通常探索	2. 位置間隔 2 倍	3. スケール間隔 1.4 倍	4. 提案手法
演算回数(回)	878M	220M	642M	109M
処理時間(秒)	61.4	22.4	46.6	11.3
検出率 (%)	84.3	79.0	77.1	85.4

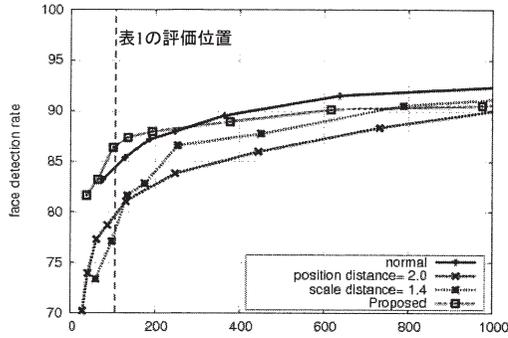


図7 探索方法による検知精度の比較

を用いて顔検出を行うことで、通常探索に対して処理時間を約 1/5 以下に削減でき、なおかつ検出率を 1% 向上できることを確認した。

各探索方法における検出精度の推移の様子を図7に示す。図7は、横軸が誤検出数、縦軸が顔検出率を示している。グラフ上にプロットした各点は、カスケード型識別器のある強識別器までを使った際の検出精度を示している。まず、探索間隔を広げることで処理時間を削減した探索方法2, 3では、誤検出数が減少するに従って通常探索よりも大きく検出率が下がっていることがわかる。これは、探索間隔を広くしたことで顔判定を行う位置が実際に顔のある位置から離れてしまうことがあり、この場合に非顔と判定してしまうため、検出率が劣化したものと考えられる。

これに対して提案手法では、誤検出数が減少して行っても検出率が大きく減少することはない。提案手法では顔確率分布から顔のある位置を推定し、その周辺位置を詳細に探索することで、実際に顔のある位置に対して顔判定が行われるため検出率を維持できると考えられる。また、提案手法では誤検出数 200 以下において通常探索よりも検出精度が向上している。これは、提案手法の顔確率分布に基づいた間引き探索により、顔確率分布が局所的にピークを持つ非顔領域を、非顔と判定することができるためと考えられる。これにより、通常探索よりも誤検出を削減することができ、検出精度が向上したのと考えられる。

5. まとめ

本報告では、顔確率分布に基づいた効率的な探索により高速に顔を検出する手法を提案した。カスケード型識別器のクリア数に着目することで、一部の強識別器だけを用いた顔判定の途中演算結果から顔確率分布を得ることができる。更には、この顔確率分布に基づいて、位置・スケールに対して間引いた簡易探索と詳細探索を組み合わせることで、顔判定を行う回数を大きく削減することができる。

提案手法を用いた顔検出の評価を行ったところ、検出率を落とすことなく、処理速度を従来の5倍以上に高速化できることを確認した。

文 献

- [1] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," Proc. on CVPR, 2001.
- [2] 高塚皓正, 田中正行, 奥富正敏, "顔らしさの評価値分布を利用した顔検出の提案," 情報処理, Vol.48, No.SIG16, pp.51-54, 2007.
- [3] H. Rowley, S. Baluja, and T. Kanade. "Neural network-based face detection," In *IEEE Patt. Anal. Mach. Intell.*, volume 20, pp 22-38, 1998.
- [4] Bo Wu, Haizhou Ai, Chang Huang, Shihong Lao, "Fast Rotation Invariant Multi-View Face Detection Based on Real Adaboost," FG2004, pp79-84.
- [5] 三田雄志, 金子敏充, 堀修, "顔検出に適した共起に基づく Joint Haar-like 特徴," 電子情報通信学会論文誌, J89-D, pp.1791-1801, 2006.
- [6] 勞世竝, 小鶴俊幸, 岡本卓也, 田畑弘, 川出雅人, "360 度回転対応高速顔検出システム," SSI03 論文集, pp.511-514, 2003.
- [7] 鈴木哲明, 細井利憲, 櫻井和之, 佐藤敦, "リングフィルタを用いた高速顔検出処理の開発," 信学会総合大会, D-12-52, p213, 2003.
- [8] 岩井儀雄, 勞世竝, 山口修, 平山高嗣, "画像処理による顔検出と顔認識," 情報処理学会研究報告 CVIM, Vol.2005, No.38, pp.343-368, 2005.