

最大クリーク抽出に基づく向きの変化に依存しない人物の顔検出法

堀田 一弘[†] 富田 悦次[†] 高橋 治久[†]

本論文では、最大クリーク抽出に基づく向きの変化に依存しない人物の顔検出法を提唱する。検出対象である人物の顔と入力画像をグラフで表現し、グラフを変形させながらマッチングを行えば、向きの変化に依存しない顔検出が可能になると考えられる。しかし、大局的なグラフ構造を保ったまま変形させる場合、その組合せは膨大な数になり、現実的でない。そこで、モデルグラフを枝ごとの局所部分グラフに分解して変形させながらマッチングを行い、類似度が高い場合のみ直積空間内の対応する節点間に枝を張っていく。この操作を繰り返すことにより、直積空間内にいくつかのクリークが生成される。直積空間内のクリークの節点数はモデルグラフとの類似度を表しているため、顔検出の問題が直積空間からの最大クリーク抽出問題に置換できる。直積空間はモデルグラフと入力グラフの全節点の組合せを保持しているため、クリーク抽出の際に各々のマッチング結果を大局的に統合、評価していることになる。入力画像に複数人の顔が含まれている場合には、直積空間内に複数のクリークが生成されるだけであるため、1回のクリーク抽出処理により画像中に含まれる複数人の顔を同時に検出することができる。HOIP 顔画像データベースを用いた実験により、モデルグラフの顔の向きから ± 30 度位までの向きの変化に対して 80%以上の検出率が得られることを確認した。さらに、側面顔のモデルグラフを作成し、正面と側面の2つのモデルグラフを併用することにより広範囲な向きの変化に依存しない顔検出を実現した。

A View-invariant Human Face Detection Method Based on Maximum Cliques

KAZUHIRO HOTTA,[†] ETSUJI TOMITA[†] and HARUHISA TAKAHASHI[†]

This paper presents a view-invariant human face detection method based on maximum cliques. The independence on the view changes is realized by distorting the graph when human faces and an input image are represented by graphs. However, the number of the combination of distortion becomes too large when the graph is distorted globally. To avoid this problem, the model graph is decomposed into local subgraphs for every edges. Two vertices are connected in the Cartesian product space when the model graph and an input graph have similar local subgraphs. As a result, the matching result of two graphs forms some cliques in the Cartesian product space. Face detection problem turns to a maximum clique extraction problem, because the size of cliques stands for the similarity with the model graph. Since the Cartesian product space has all the combination of edges between the model graph and an input graph, the clique extraction corresponds to an evaluation of the integrated matching results of local subgraphs. In the case that several target human faces are included in an input image, the corresponding number of cliques are generated in the Cartesian product space. The whole target human faces can be detected by a clique extraction process. True positive rate is over 80% while the view of faces in the images are changed up to an angle of ± 30 degrees. We have realized a face detection method which is independent on the wide view changes by using the model graphs of a frontal face and a profile face.

1. はじめに

画像から人物の顔を見つけ出す顔検出は、コンピュータによる顔の自動認識の第1ステップとして重要であり、近年さかんに研究されている^{1),2)}。今日では、学習サンプルの能動的な収集法³⁾、Support Vector

Machine (SVM)のような汎化能力の高い識別器の利用^{4),24)}、認識に利用する特徴の選択^{13),14)}等により、顔の向きが限定されている状況下ではかなり安定した顔検出が可能となってきた^{1),2)}。しかし、実環境下では人間が必ず正面を向いているとは限らないので、向きの変化に依存しない認識手法が必要となる。このような現状から、本論文では向きの変化に依存しない顔検出の問題を扱う。

近年、向きの変化を考慮に入れたいくつかの顔検出

[†] 電気通信大学

The University of Electro-Communications

法が提案されている^{6)~11)}．その中には色情報を用いることにより向きの変化の影響を受けにくくする顔検出法も含まれるが^{9)~11)}，本論文では他の対象への応用も視野に入れ，濃淡画像からの顔検出に課題を限定する．濃淡画像から向きの変化に依存せずに顔を検出する方法もいくつか提案されている^{6)~8)}．それらの手法では向きごとのモデル(検出器)を複数個用意することにより，向きの変化に対応していた．しかし，特定の向きの顔検出でもすでに難しい問題であるので，当然各向きのモデルは特定の向きへの依存度が大きく，向きの変化の影響を受けやすい．したがって，このアプローチで高精度な検出を行うためには多くのモデルが必要となり，計算時間等の問題が生じる．もし，1つのモデルで対応できる向きの範囲を広げることができれば，少ない数のモデルだけで高精度な検出が可能になると考えられる．

1つのモデルによる向きの変化に依存しない認識法として，Ladesらが提案したグラフを用いた個人識別法がある¹⁵⁾．この手法では認識対象と入力画像をグラフで表現し，グラフの形状と類似度から定義されるエネルギーを最小化するように入力画像にあてはめたグラフを変形させる．入力グラフを変形させることにより，入力画像中の顔の多少の向きの変化や表情変化に頑健な個人識別を実現している．しかし，この方法はエネルギー最小化により入力画像にあてはめたグラフを変形させるため，そのままの枠組みでは，入力画像から複数人の顔を検出しなければならない顔検出の問題に適用することは難しい．グラフの持つ柔軟な特性を生かしたまま顔検出の枠組みにあてはめることができれば，向きの変化に依存しない顔検出が可能になると考えられる．

人物の顔と入力画像をグラフで表現し，グラフを変形させながら探索を行えば，向きの変化に依存しない顔検出が可能になると考えられる．しかし，グラフの大局的な構造を保ったまま変形させる場合，変形の組合せは膨大な数になる．顔検出では顔がどこにあるか，どれくらいの大きさであるか，いくつ含まれるのか等が事前には分からないので，大きさと位置を変えながら入力画像中を探索しなければならない^{1),2)}．したがって，グラフの変形の組合せ数が膨大であるマッチング法は顔検出問題の解法として適さない．本論文では，グラフの変形マッチングを高速に行うために，モデルグラフを局所部分グラフ(最も単純には2節点間の枝)に分解してマッチングを行い，最後に各々のマッチング結果を大局的に統合する方法を用いる．モデルグラフを局所部分グラフに分解することにより変

形の自由度が減るので(1本の枝の場合，2節点間の位置関係のみ)，高速な変形マッチングが可能となる．

人物の顔をグラフによりモデル化し，入力画像を画素を節点とするグラフであると考えれば，2つのグラフの関係はすべての節点間の組合せを保持した直積空間内で表現できる．モデルグラフを局所部分グラフに分解して入力グラフとの類似度を測り，類似度が高い場合にのみ直積空間内の対応する節点間に枝を張っていくと，直積空間内にクリークが生成される．類似した部分グラフがある場合にのみ直積空間内に枝を張るので，節点数が多いクリークほど類似度が高いことを意味する．つまり，顔検出の問題が直積空間内の(最大)クリーク抽出の問題に置き換えられる．しかも，局所部分グラフのマッチングの時点では他の部分グラフとの関係やマッチングの結果を考慮する必要はなく，クリークを抽出する際に初めて他の局所部分グラフのマッチング結果と統合，評価される．直積空間ではモデルグラフと入力グラフのすべての節点間の組合せを保持しているため，クリーク抽出の際にグラフの大局的な変形の様々な組合せを厳密に評価していることになる．また，入力画像に複数人の顔が含まれる場合には直積空間内に複数のクリークが生成されるだけであるので，1回のクリーク抽出処理により複数人の顔を検出することができる．これが本論文で提唱する最大クリークに基づく向きの変化に依存しない顔検出法である．

最大クリーク抽出に基づくマッチング法は，これまでも様々な分野で応用されている．たとえば，Ishitaniはドキュメント画像からの名前や所属等の自動抽出を行う際，局所的にマッチングした結果を大局的に判断してエラーを除去するために最大クリーク抽出を利用している²¹⁾．また，Ogawaは星座のマッチングに応用している²²⁾．Bahadurらは二次元電気泳動画像のマッチングに最大クリーク抽出を利用している²³⁾．

提案手法では検出結果が直積空間内のクリークとなるので，直積空間内の枝の張り方が検出精度に影響を与える．そこで，枝を張るか否かの判断(局所部分グラフのマッチング)に従来の顔検出法で利用されているパターン認識手法(特徴量や識別手法)を適用する^{1),2)}．具体的には，顔検出¹³⁾や対象識別¹⁵⁾に有効であるガボール特徴を利用する．モデルグラフから切り出した局所部分グラフの各節点周辺で抽出したガボール特徴を1つの特徴ベクトルとして扱い，入力グラフ中の部分グラフ(入力画像中のある領域)から抽出した特徴ベクトルとの類似度を測る．以下の実験では，最も単純な場合として特徴ベクトル間の距離を類似度

として定義し、距離が閾値以下である場合のみ直積空間内の対応する節点間に枝を張る。

実験では、画像から抽出したすべての特徴を用いる従来のマッチング法や大局的なグラフ構造を用いたマッチング法と比較し、局所部分グラフに分解して変形マッチングを行い、クリーク抽出により各々のマッチング結果を大局的に統合、評価する提案手法の有効性を示す。また、HOIPの顔画像データベースを用いた実験により、モデルグラフの顔の向きから ± 30 度位の向きの変化がある場合にも 80%以上の検出率が得られることを示す。さらに、側面顔のモデルグラフを作成し、正面と側面の2つのモデルグラフを併用することにより広範囲な向きの変化に依存しない顔検出を実現する。

以下、2章では最大クリーク抽出に基づく顔検出法について説明する。3章では提案手法がどの程度の向きの変化まで対応できるかを評価する。また、正面と側面の2つのモデルグラフを併用した向きの変化に依存しない顔検出の結果も示す。最後に、まとめと今後の課題を4章で述べる。

2. 最大クリーク抽出に基づく顔検出法

本章では、最大クリーク抽出に基づく顔検出法について説明する。提案手法の概念図を図1に示す。図1(a)はグラフにより人物の顔をモデル化したモデルグラフの例であり、白い四角が節点を表し、黒線が枝を表している。モデルグラフのあてはめ方については3.1.1項で説明する。図1(b)は入力画像にあてはめたグラフの節点の例であり、節点を丸で示してある(ここでは分かりやすさのために意図的に節点数を少なくしている)。入力画像へのグラフのあてはめ方については2.1節で説明する。図1(c)はモデルグラフ(4節点)と入力グラフ(8節点)の直積空間である。横軸がモデルグラフの4節点を表し、縦軸が入力グラフの8節点を表している。提案手法ではモデルグラフを局所部分グラフに分解してマッチングを行い、直積空間に枝を張っていく。以下の実験では、最も単純な場合として局所部分グラフを2節点間の枝とする。つまり、モデルグラフ内の任意の2節点を結ぶ1本の枝が1つの識別器に対応し、入力グラフ中で類似した特徴量(局所的パターン)を持つ2節点間の枝を探す。具体的に説明すると、モデルグラフ内の任意の節点 p と節点 q の2節点から抽出した特徴量と入力グラフ内の任意の節点 s と節点 t から抽出した特徴量の類似度を測り、もし類似度が閾値以上であれば直積空間内の節点 (p, s) と節点 (q, t) 間に枝を張る(類似度の測り方は



図1 最大クリーク抽出に基づく顔検出の概念図

Fig. 1 Human face detection based on maximum cliques.

2.2節で述べる)。この操作を繰り返すことにより、モデルグラフと入力グラフで類似度が高い節点どうしに枝が張られ、直積空間内にクリークが生成される。これにより顔検出の問題が直積空間内でのクリーク抽出の問題に置換され、最大クリークを与える部分グラフが最も顔らしい領域となる。図1では1人の顔だけが含まれている例を示したが、画像中に複数人の顔が含まれる場合には直積空間内で複数のクリークが生成されるだけであるので、各人の顔に対応する部分グラフが同じ節点数の(最大)クリークとなっていれば、1回の最大クリーク抽出処理により画像中のすべての人の顔を検出することができる。

提案手法は、入力画像へのグラフのあてはめ(節点選択)、モデルグラフと入力グラフのマッチングによる直積空間内の枝張り、直積空間内のクリーク抽出の3つのプロセスからなる。最も単純な入力グラフのあてはめ方は入力画像の全画素を節点とすることである。しかし、この場合にはモデルグラフとのマッチング回数が多くなり、顔検出に時間を要する。そこで、高速化のためにHarrisの特徴点抽出¹⁹⁾を用いて入力グラフの節点選択を行う。Harrisの特徴点抽出法を用いた入力グラフの節点選択について2.1節で説明する。モデルグラフのあてはめについては3.1.1項で説明する。

次に、モデルグラフと入力グラフのマッチング(直積空間内の枝張り)について2.2節で説明する。検出結果は直積空間内のクリークとなるので、検出精度は直積空間内の枝の張り方に依存する。安定した枝張りを行うために、従来手法で利用されたパターン認識手法を適用する。ここでは、顔検出¹³⁾や対象識別¹⁵⁾で実績のあるガボール特徴に基づくマッチング法を利用する。ガボールフィルタとその特性を2.2.1項で説明する。

最後に、直積空間内の最大クリーク抽出に利用するTomitaらの提案した分枝限定アルゴリズム²⁷⁾を2.3節で説明する。

2.1 Harris の特徴点抽出による入力グラフの節点選択

本論文では、顔検出の高速化のために特徴点抽出法を利用して入力画像から特徴的な場所だけを選択し、それを入力グラフの節点とする。節点の選択により顔の特徴点を候補から除外してしまわないように、環境変化の影響を受けにくい特徴点抽出法を利用したい。Schmid らは変化に対する安定性と抽出された特徴点の情報量の観点からいくつかの特徴点抽出法の評価を行った²⁰⁾。その結果、Harris らの特徴点抽出法が変化に対する安定性と情報量の観点から最も良い方法であるという報告をしている。そこで、本論文では Harris らの特徴点抽出法¹⁹⁾を利用して入力画像から節点の選択を行う。以下に、Harris らの特徴点抽出法を説明する。

基本となる評価基準は、画像上の点 (a, b) を中心とした局所領域を横軸、縦軸に各々 x, y だけ動かしたときの変化量

$$E_{x,y}(a,b) = \sum_{u,v \in L(a,b)} w_{u,v} |I_{u+x,v+y} - I_{u,v}|^2, \quad (1)$$

である。ここで、 $L(a,b)$ 、 $I_{u,v}$ 、 $w_{u,v}$ は、点 (a,b) を中心とする局所領域内の点の集合、画像上の点 (u,v) の輝度値、ガウシアン重みを表している。つまり、現在評価している領域が平らであれば動かしたときの変化量が小さく、エッジであればエッジに沿った方向に動かしたときだけ小さな変化量となり、角のような特徴点であればどの方向に動かしても変化量が大きいという評価基準である。

$I_{u+x,v+y}$ を Taylor 展開し、2 次以上の項を無視すると、式 (1) は

$$E_{x,y}(a,b) = \sum_{u,v \in L(a,b)} w_{u,v} (x(\partial I/\partial x) + y(\partial I/\partial y))^2, \quad (2)$$

となり、点 (a,b) を中心とする局所領域を x, y だけ変移させたときの変化量は、

$$E_{x,y}(a,b) = Ax^2 + 2Cxy + By^2, \\ = (x \ y)M(x \ y)^T, \quad (3)$$

となる。ここで、 $A = (\partial I/\partial x)^2 \otimes w$ 、 $B = (\partial I/\partial y)^2 \otimes w$ 、 $C = (\partial I/\partial x)(\partial I/\partial y) \otimes w$ 、 $M = \begin{pmatrix} A & C \\ C & B \end{pmatrix}$ で

ある。 \otimes は畳み込み積分を表している。式 (3) から $E_{x,y}(a,b)$ は行列 M に依存することが分かる。したがって、行列 M の 2 つの固有値を調べれば領域が平ら、エッジ、特徴点であるかの判定が可能となる。2



図2 Harris らの特徴点抽出による入力グラフの節点選択
Fig. 2 Point selection of an input graph by Harris operator.

つの固有値がともに小さい値であれば平ら、片方だけ大きな値であればエッジ、2 つとも大きな値であれば特徴点と判定できる。Harris らは固有値問題を解かなくても画像上の点 (a,b) の特徴点らしさを判定できる指標

$$R(a,b) = \text{Det}(M) - \alpha \text{Tr}(M)^2, \quad (4)$$

を提案した。 $\text{Det}(M)$ と $\text{Tr}(M)$ はそれぞれ行列 M の行列式と対角和であり、これらは固有値問題を解かなくても計算できる。特徴点を抽出するには、画像上の点 (a,b) を中心とする領域の特徴点らしさ $R(a,b)$ を計算し、その値が閾値以上であれば特徴点として選択すればよい。上記の特徴点らしさ $R(a,b)$ はエッジの場合に負の大きな値を出す。エッジの部分も節点として選択した方がよいと考えられるので、以下の実験では上記の指標 $R(a,b)$ の絶対値を用いた。また、 $\alpha = 0.04$ とし、点 (a,b) を中心とした局所領域を 5×5 とした。

図 2 に顔を含む画像に Harris らの特徴点抽出を適用した例を示す。図中の白点が特徴的な点として選択された場所である。図から分かるように、目、鼻、口等は特徴的なパターンであるので、入力グラフの節点として選択されている。この例では入力グラフの節点数が全画素数から約半分に削減された。

2.2 モデルグラフと入力グラフのマッチング

直積空間での枝張りのためのモデルグラフと入力グラフのマッチングについて説明する。本論文では、モデルグラフを枝ごとの局所部分グラフに分解し、枝を構成する 2 節点から抽出したガボール特徴を基にマッチングを行う。つまり、モデルグラフの任意の 2 節点間の枝が 1 つの識別器に対応し、入力グラフ中の類似したガボール特徴（局所的な見えのパターン）を持つ 2 節点間の枝を探す。識別方式として、各々の節点で抽出したガボール特徴を基に SVM^{4),24)}等を利用することも可能である。しかし、識別方式は本論文の主張

ポイントではないので、提案手法の有効性(グラフ表現の導入と局所部分グラフに分解して変形マッチングを行い、クリーク抽出により大局的に評価すること)を明確にするために複雑な識別方式の利用を避け、最も単純なモデル特徴との距離によるマッチングを用いる(識別方式の変更は容易である)。具体的には、入力グラフ中の1本の枝を構成する2節点(s と t)で独立にガボール特徴を抽出し、それらを縦に並べてできる特徴ベクトル $\mathbf{y}_{s,t}$ とモデルグラフの2節点(p と q)から抽出したモデル特徴ベクトル $\mathbf{x}_{p,q}$ との距離を測り、 $\|\mathbf{x}_{p,q} - \mathbf{y}_{s,t}\|^2 \leq \text{threshold}$ ならば直積空間内の節点(p, s)と節点(q, t)間に枝を張る。しかし、モデルグラフの枝を構成する2節点から抽出したガボール特徴だけでは情報が少なくマッチングが不安定になる可能性もあるので、節点を中心とした局所領域内で抽出したガボール特徴を利用してマッチングを行う。これにより、節点周辺の局所的な形状情報も利用できるため、マッチングが安定する。以下の実験では、各節点を中心とした 5×5 画素の局所領域で抽出したガボール特徴を用いてマッチングを行う。

また、マッチング時の局所部分グラフ(枝)の変形に関してであるが、すべての変形を許すと組合せ数が増大し、通常ではありえない変形での誤識別が増える可能性があるため、変形の範囲を限定した。以下の実験では、個人性の違いや向きの変化に十分対応できるように、モデルグラフの2節点間の元の位置関係から4画素以内の変形を許すことにした。

2.2.1 ガボールフィルタ

生体の視覚皮質(第一次視覚野)には、局所的な方向だけに選択的に反応する神経細胞があることが知られている。また、猫の視覚皮質の単純型細胞の受容野特性は、ガボールフィルタでうまく近似されることが知られている¹⁶⁾。ガボールフィルタは、

$$\psi_{\mathbf{k}}(\mathbf{x}) = \frac{\mathbf{k}^2}{\sigma^2} \exp\left(\frac{-\mathbf{k}^2 \mathbf{x}^2}{2\sigma^2}\right) \cdot \left[\exp(i\mathbf{k}\mathbf{x}) - \exp(-\sigma^2/2)\right], \quad (5)$$

のように定義される¹⁵⁾。ここで、 $\exp(-\sigma^2/2)$ は直流成分を0にするための項である。式(5)中のパラメータは、それぞれ、 $\mathbf{x} = (x, y)^T$ 、 $\mathbf{k} = k_\nu \exp(i\phi)$ 、 $k_\nu = k_{max}/f^\nu$ 、 $\phi = \mu \pi/8$ である。以下の実験では、 $\nu = \{0\}$ 、 $\mu = \{0, \dots, 7\}$ 、 $\sigma = \pi$ 、 $k_{max} = \pi/2$ 、 $f = \sqrt{2}$ とした。また、抽出した8次元(方向)のガボール特徴はノルムを1に正規化してから利用した。これは、場所ごとに反応する方向を明確にする効果がある。

ガボールフィルタの出力は、スパースコーディング

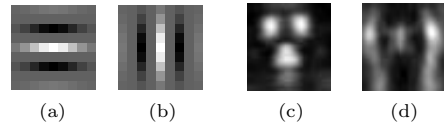


図3 ガボールフィルタとその適用例

Fig. 3 Examples of Gabor filters and Gabor features.

(sparse coding) になっていることが知られている。自然画像から切り出した局所領域のスパースさを最大にするような制約条件を用いて自己組織的にフィルタを構成したところ、ガボールフィルタに類似したフィルタが得られたことが報告されている¹⁷⁾。ここで、スパースコーディングは、近年、脳内での情報表現として注目を集めており、小数の鋭い反応選択性を持つ細胞の発火パターンによる情報表現のことである。

正面の顔画像(図1(a)の顔画像)にガボールフィルタを適用したときの結果を図3に示す。図3(a)、(b)は $\mu = 0$ 、 $\mu = 4$ の場合のガボールフィルタを表し、(c)、(d)はそれぞれ(a)、(b)のガボールフィルタを適用したときの結果である。この図から、ガボールフィルタの出力はほとんどが0であり、ある特定の方向成分が入力されたときだけ選択的に反応することが分かる。

ガボール特徴が顔や3次元対象の認識に有効であることはすでに示されている^{13),15)}。その要因の1つは、ガボールフィルタの出力がスパースな情報表現だからであるといえる。また、自然画像の独立成分分析を行うと、ガボールフィルタに類似したフィルタが得られる¹⁸⁾ことから、反応する方向が異なるガボールフィルタの出力は確率的独立性も高いと考えられる。フィルタ出力の確率的独立性が高いこともこの特徴が認識に有効である要因の1つである。

2.3 最大クリーク抽出アルゴリズム

本節ではモデルグラフと入力グラフのマッチングにより生成された直積空間からの最大クリーク抽出方法について説明する。無向グラフ中の最大クリーク抽出はNP困難なクラスに属する問題であるが、Tomitaらは分枝限定法を用いて厳密解を高速に得るアルゴリズムMCQを提案している²⁷⁾。入力画像に含まれる複数人の顔を検出するためには同じ節点数のすべての最大クリークを抽出する必要があるが、アルゴリズムMCQは1個の最大クリークだけを抽出する方法である。そこで、同じ節点数のすべての最大クリークを抽出できるように拡張し、直積空間内のすべての最大クリーク、すなわち入力画像からの検出結果を得る。以下に最大クリーク抽出アルゴリズムMCQを説明する

が、詳細は文献 27) を参照していただきたい。

クリークを探索するには、ある時点で保持しているクリークにそのクリーク中の全節点に隣接している節点を新たに付け加え、より大きなクリークとする操作を深さ優先探索で進める。つまり、最初に全節点集合 (= 最初の候補節点集合) を根とし、ある時点における候補節点集合 R を頂点として、 R 中の 1 つの節点 p に隣接するすべての節点を R 中から抽出して新たな集合 R_p を作成し、これを R の子として枝で結ぶことにより、探索木として表現できる。この探索を分枝限定法により効率良く実行する。

隣接している節点どうしに異なる色を付けていくと、クリークをなす節点どうしは必ず異なる色となるので、全節点集合の彩色数は最大クリークの節点数以上の値となる。つまり、彩色数は最大クリークの上界となる。したがって、候補節点集合の彩色数を高速かつ精度良く計算できれば、クリーク探索の効率化につながる。なぜならば、現在の候補節点集合におけるクリークの節点数にこの集合の彩色数を加えた数が現在までに得られている最大クリークの節点数未満であるならば、現在探索中の探索木を探索しても最大クリークに到達することがないので、それ以降の探索を省くことができるからである。これにより、高速かつ厳密な最大クリーク抽出を可能としている。

アルゴリズム MCQ では、分枝限定を効率的に行うために、全節点集合を次数 (隣接節点の数) で整列し、次数の小さい順に探索を行う。次数の小さい順に探索を行うと、次数の大きい節点に対しては分枝限定効果で探索をしなくて済む可能性が高くなり、高速化できる。これまでに提案されている手法と比較を行い、次数の小さい順に探索することにより分枝が限定され、従来手法よりも高速に最大クリークを抽出することが可能であることが確認されている²⁷⁾。

3. 向きの変化に依存しない人物の顔検出

本章では、まず、正面顔画像だけを用いてモデルグラフを作成し、どの程度の顔の向きの変化に対応できるかを調べる。次に、提案手法が対応できる向きの範囲に応じて側面顔のモデルグラフを用意し、正面と側面の 2 つのモデルグラフにより広範囲な向きの変化に依存しない顔検出を実現する。

以下、3.1 節では正面顔画像から作成したモデルグラフの向きの変化に対する頑健性の評価実験について述べる。3.2 節では側面顔のモデルグラフを作成し、側面顔のモデルグラフが対応できる向きの範囲を評価した後、正面顔のモデルグラフと併用して向きの変化



図 4 顔画像の例

Fig. 4 Examples of face images.

に依存しない顔検出を行う。

3.1 向きの変化に対する頑健性の評価

まず、本節の実験に用いる画像データベースおよびモデルグラフについて 3.1.1 項で説明する。3.1.2 項では向きの変化に対する頑健性の評価方法について説明する。評価実験の結果を 3.1.3 項に示す。

3.1.1 画像データベースとモデルグラフ

本実験ではモデルグラフ作成と直積空間内での枝張りの際の閾値決定のために、HOIP 顔画像データベースの 300 人の正面顔画像と、MIT と CMU の顔検出のためのテストデータベース⁵⁾に含まれる正面顔画像を利用した。これらの画像に含まれる顔の大きさや位置はそれぞれ異なるので、両目、鼻、口の位置を用いてアフィン変換により 38×36 画素の大きさに正規化した。ただし、顔の大きさや位置が明らかにずれているものは除外し、計 936 枚の正面顔画像を得た。顔画像の例を図 4 に示す。図から分かるように、様々な環境下で撮影された顔画像を含んでいる。これらの顔画像をランダムにモデル特徴作成用の 136 枚と閾値選択用の 800 枚に分けて利用した。

次に、モデルグラフのあてはめについて説明する。モデルグラフの節点の位置や個数等は、未学習サンプルに対する誤識別率等を用いて決定した方が安定すると考えられるが、本論文では局所部分グラフ (以下の実験では 2 節点間の枝) に分解して変形させながらマッチングを行い、クリーク抽出の際に各々のマッチング結果を大局的に統合することにより向きの変化に依存しない顔検出が可能となるか否かを確認することに主眼を置き、ガボール特徴の情報量¹²⁾が多く、入力グラフの節点選択に用いる Harris らの特徴点抽出により安定して抽出される両目、鼻、口をモデルグラフの節点とした。本実験で利用するモデルグラフを図 5 に示す。図内の白い四角が節点を表しており、黒線は枝を表している。黒線で結ばれる 2 節点 (枝) が 1 つの識別器を構成しており、入力グラフ中で類似した特徴量を持つ 2 節点 (枝) を探すことになる。したがって、直積空間内で節点数 4 のクリークを与える部分グラフがこのモデルグラフとの最も高い類似度を持つ領域となる。顔検出では画像中に含まれるすべての顔を検出しなければならないため、すべての可能性を高速に調べる必要がある。したがって、画像中に含まれる

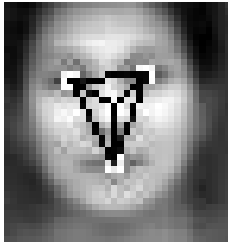


図 5 正面顔のモデルグラフ

Fig. 5 A model graph of a frontal face.

複数人の顔（直積空間内の複数の最大クリーク）を高速に検出するために、本論文で利用するような高速な最大クリーク抽出アルゴリズムが必要となる。

上記 136 枚の正面顔画像を用いて各局所部分グラフのモデル特徴を作成した。各局所部分グラフのマッチングにおける閾値はモデル特徴作成に用いていない顔画像と顔以外の画像に対する誤識別率が小さくなるように決定する。ただし、提案手法では局所部分グラフに分解してマッチングを行い、最後にクリーク抽出という形で結果を統合して評価するので、統合したときに誤識別率が低くなるように各々の閾値を決めた。閾値決定のための画像として、800 枚の正面顔画像、顔が含まれていない画像からランダムに選択した 17,550 枚の顔以外の画像とグラフを変形させない場合にモデルグラフとの類似度が高い 1,000 枚の顔以外の画像を用いた。以下の実験では、各節点の周辺 5×5 の領域で抽出した 8 方向のガボール特徴を利用する。したがって、各局所部分グラフ（枝）のマッチングで用いるガボール特徴の次元数は $400 = 5 \times 5 \times 2$ （節点） $\times 8$ （方向）である。

3.1.2 評価方法

提案手法の向きの変化に対する頑健性を評価するための方法について説明する。顔検出は顔と顔以外の 2 クラスの識別問題であるので、その精度は顔と顔以外の 2 つの識別率を用いて評価しなければならない。一般に、顔（対象）を正しく識別できた正解率である True Positive Rate (TPR) と顔以外（対象以外）を顔（対象）であると誤識別する False Positive Rate (FPR) を用いる²⁴⁾。各々の指標の性質から分かるように、FPR が低く、TPR が高い検出器ほど優れていることになる。ここでは、FPR を評価するために、PBIC 画像データベース²⁵⁾から得た 100 枚の画像を利用する。提案手法ではグラフの変形を許すため、元画像のままでは FPR の評価が難しい。そこで、100 枚の元画像から 10,000 枚の顔を含まない画像をランダムに切り出し、FPR の評価に用いた。これら

の画像の大きさはモデルグラフ作成用の顔画像と同じ 38×36 画素である。提案手法では最大クリークの節点数がモデルグラフと入力グラフのある領域との類似度を表しているの、最大クリークの節点数と閾値処理により識別率を計算する。以下の実験では、最大クリークの節点数が 4、すなわちモデルグラフとの最大類似度であれば顔と判定し、それ以外は顔でないと判定して評価を行った。FPR の評価の際、提案手法ではモデルグラフの各節点の元の位置関係から 4 画素以内の変形を許し、直積空間内で節点数 4 のクリークが 1 つでも生成されれば不正解とした。

もう一方の評価基準である TPR の評価には HOIP 顔画像データベースに含まれる顔の向きが 15 度、30 度、45 度の 300 人の顔画像を用いる。これらの顔画像に顔検出を適用し、顔の領域が正しく検出されるか否かで評価する。顔検出を行う際、画像中の顔の大きさの変化に対処するために、入力画像の大きさを $1/1.1$ の割合で変化させながらマッチングを行った¹³⁾。Harris らの特徴点抽出は対象の大きさの変化に影響を受けやすい²⁰⁾ので、入力画像の大きさを変えるごとに Harris らの特徴点抽出を適用し、入力グラフを作成した。ただし、画像の大きさを変化させた場合でも元画像の大きさの直積空間に対応づけて枝を張ることができるので、元画像の大きさで生成された直積空間を用いて枝張りおよびクリーク抽出を行う。

提案手法の有効性（1：グラフ表現の導入、2：枝ごとの局所部分グラフに分解して変形させながらマッチングを行い、クリーク抽出という形で各部分グラフのマッチング結果を大局的に統合すること）を示すために、グラフ表現を用いない従来のマッチング法、局所部分グラフに分解しない大局的なグラフマッチング法との比較を行う。この比較実験は提案手法の上記 2 つの有効性を明らかにすることが目的である。つまり、グラフ表現を利用しないものと比較することにより、グラフ表現を用いることの有効性を明らかにし、局所部分グラフに分解して変形マッチングを行い、クリーク抽出により各々の結果を統合することの有効性を評価するために、大局的なグラフ構造を保ったままマッチングする方法と比較する。第 1 の比較手法である従来型のマッチング法では 38×36 画素の 136 枚の顔画像から抽出した 6,720 ($= 30 \times 28 \times 8$) 次元の平均ガボール特徴をモデルとし、入力画像から抽出した特徴量との距離を用いてマッチングを行う。第 2 の比較手法である大局的なグラフ構造を用いたマッチング法では、提案手法と同様に両目、鼻、口の 4 節点のモデルグラフを利用し、各節点の周辺 5×5 の領域で抽出し

たガボール特徴を用いる．各節点の周辺で抽出した全ガボール特徴を $800 (= 5 \times 5 \times 8 \times 4$ (節点)) 次元の1つの縦ベクトルとして距離を測る．ただし，このマッチング法では，提案手法のようにグラフを変形させながらマッチングを行うことが難しい．なぜなら，大局的なグラフ構造を保ったまま変形させる場合，その組合せ数が膨大になるからである．たとえば，図5に示す節点数4のモデルグラフで各節点を上下左右に1画素だけ変形をさせた場合，1つの領域に対してマッチングの回数が $5^4 = 625$ となる．顔検出では，通常，位置と大きさを変えながらマッチングを行うので，1つの領域に対してこれだけの組合せがあると現実的でない．しかも，これは上下左右に1画素だけ移動した場合の例であり，この変形だけで向きの変化に対処するのは難しい．一方，提案手法ではモデルグラフを枝ごとの局所部分グラフに分解してマッチングを行う．この場合，各々のマッチングは2節点間であるので，変形が容易である．しかも，直積空間はモデルグラフの全節点と入力グラフの全節点の組合せを保持しているため，マッチング時には2節点間の変形しか扱っていないにもかかわらず，直積空間内でのクリーク抽出の際にはグラフの大局的な変形を厳密に評価していることになるのである．これにより，実際にモデルグラフを大局的に変形させてマッチングすることなく，大局的に変形させた場合のマッチング結果を高速に評価できる．これが本提案手法の最大の利点である．

3.1.3 評価実験

以下に実験結果を説明するが，各識別方法でFPRとTPRが異なる場合には評価が難しくなる．そこで，提案手法のFPRを計算し，比較手法のFPRが提案手法のそれと同じ値となるように閾値を決定し，そのときのTPRで評価を行うことにした．FPRの評価の際，提案手法では各節点の元の位置関係から4画素以内の変形を許したマッチングを行うが，2つの比較手法では1枚の画像に対して1回のマッチングのみで評価した．また，公平な評価を行うために，FPRの評価ではHarrisの特徴点抽出を利用せずに画素を節点とする入力グラフとした．PBICの100枚の画像から切り出した10,000枚の画像に対する提案手法のFPRは11/10,000であった．この時の結果(TPR)を表1に示す．全特徴量を用いたマッチング法では顔の向きが15度の場合ですでに精度が低い．これは，画像から抽出したすべての特徴量を用いることにより，向きの変化の影響を受けやすくなっていることを示している．このマッチング法では，顔の向きが30度，45度の場合にほとんど検出ができていないことが分かる．

表1 向きの変化に対する頑健性の評価
Table 1 Evaluation of robustness to view changes.

	TPR(15°)	TPR(30°)	TPR(45°)
全特徴マッチング	98/300	3/300	0/300
大局的グラフ	289/300	228/300	86/300
提案手法	276/300	259/300	146/300
提案手法(3×3)	278/300	261/300	164/300
提案手法(5×5)	279/300	263/300	166/300

一方，大局的なグラフ構造を用いたマッチングの場合には，両目，鼻，口の位置関係が大きく変化しない程度の向きの変化には対応できる．顔の向きが15度の場合にはほとんど検出ができており，この場合には提案手法よりも精度が高い．しかし，顔の向きが30度となると，マッチング時に変形を行う提案手法の方が高い精度となる．顔の向きが45度になると，この差がさらに広がるのが分かる．これは，提案手法がモデルグラフを枝ごとの局所部分グラフに分解して変形マッチングを行い，クリーク抽出の際にグラフを大局的に変形させた場合についても厳密に評価しているからであり，大局的なグラフ構造を保ったままマッチングする方法では変形の組合せ爆発のためグラフを変形させながらマッチングすることができないからである．これが30度と45度の場合のTPRの差になったと考えられる．

提案手法の結果を見ると，顔の向きが45度の場合にTPRが低下していることが分かる．これは向きの変化により見えにくくなった片方の目の識別に失敗し，最大クリークの節点数が3となることが多くなるからである．本論文では提案手法の有効性を明確にするために，各局所部分グラフの識別方法として最も単純なモデル特徴との距離によるマッチングを用いたが，SVM等^{4),24)}を用いてこのような場合を学習させることにより精度はさらに向上すると考えられる．

この実験に用いたHOIPの15度，30度，45度の顔画像に対して提案手法を適用したときの結果を図6に示す．図6(a)，(b)，(c)はそれぞれ15度，30度，45度の向きの顔画像に対する検出結果である．図中の白線は直積空間で最大クリークを与える部分グラフに対応する入力画像中の画素間に枝を張ったものである．近い領域で同じ節点数の複数のクリークが生成されたので，白い線が重なっているのが分かる．この結果から分かるように，モデルグラフを枝ごとの局所部分グラフに分解して変形マッチングを行うことにより，向きの変化に影響を受けにくい顔検出が可能となった．

この評価実験では節点数が4のクリーク，すなわちモデルグラフの全局部分グラフのマッチング結果が

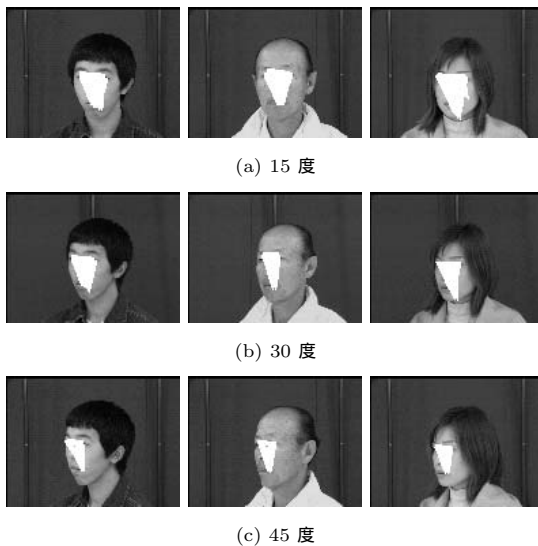


図6 HOIP顔画像に対する検出結果

Fig. 6 Examples of face detection from HOIP images.

顔であるときのみ顔として識別される。提案手法では入力グラフの1節点が1画素に対応しているので、1画素もずれることなくすべての局所部分グラフとのマッチングに成功しなくてはならないのである。これを多少緩めることにより検出精度がさらに向上すると考えられる。そこで、直積空間における入力グラフの1節点を1画素ではなく、 3×3 画素または 5×5 画素と多少大きめに対応づけた。 3×3 画素または 5×5 画素が1節点となるので、その領域内で局所部分グラフマッチングにより顔と判定されれば、対応画素がずれていても直積空間内のその節点に枝を張ってよい。この方法は検出精度を向上させるだけでなく、直積空間をコンパクトにし、最大クリーク抽出の計算をさらに高速化するという利点もある。この場合の結果を表1の下2列に示した。入力グラフの1節点を 3×3 画素または 5×5 画素に対応づけることにより、TPRがさらに向上することが分かる。顔の向きが45度の場合には、大局的なグラフマッチングの約2倍の正答率が得られた。一方、この場合のFPRは、それぞれ18/10,000と19/10,000であった。FPRも多少増えているが、その割合以上にTPRが向上していることが分かる。

この実験により、画像から抽出したすべての特徴を用いるよりも余分な情報を切り落としたグラフ構造を用いた方が向きの変化の影響を受けにくくなることを確認した。さらに、グラフ表現を用いる場合には、グラフを変形させながらマッチングを行うことにより、向きの変化の影響を受けにくくなること、すなわち提

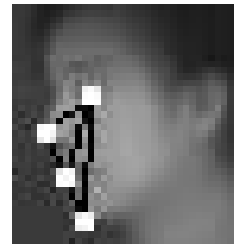


図7 側面顔のモデルグラフ

Fig. 7 A model graph of a profile face.

案手法の有効性を確認できた。

3.2 側面顔のモデルグラフとの併用による向きに依存しない顔検出

前節の実験結果から分かるように、提案手法は正面から ± 30 度位までの向きの変化に対して十分に対応できている。もし、提案手法がモデルの向きによらず ± 30 度位の向きに対応できるのであれば、顔の向きが60度の側面顔モデルグラフを保持することにより、90度位までの向きの変化にも対応できることになる。そこで、本節では、60度の向きの側面顔モデルグラフを作成し、それがどの程度の向きの変化に対応できるかを調べる。そして、正面顔のモデルグラフと併用し、広範囲な向きの変化に依存しない顔検出を実現する。

まず、側面顔のモデルグラフ(モデル特徴)作成と直積空間内での枝張りの際の閾値を決定するために利用する画像データについて説明する。ここでは、HOIP顔画像データベースに含まれる顔の向きが60度の300人の顔画像とPIE顔画像データベース²⁶⁾に含まれる光源方向を変えながら撮影した56人の顔画像を用いる。これらの顔画像の大きさは正面顔画像と同じ 38×36 画素である。これらの顔画像をモデル作成用と閾値決定用の2つの画像セットに分割して利用する。モデル作成には150人のHOIP顔画像と28人 \times 2枚のPIE顔画像を用いる。一方、閾値決定には150人のHOIP顔画像と28人 \times 10枚のPIE顔画像を用いる。閾値は正面顔のモデルグラフの場合と同様に顔画像と顔以外の画像に対する誤識別率が低くなるように決定する。閾値決定に用いる顔以外の画像として、正面顔のモデルグラフの閾値決定に用いた17,550枚と側面顔のモデルグラフを変形させない場合に類似度の高い1,000枚の顔以外の画像を用いた。

次に、側面顔のモデルグラフについて説明する。ここでは、ガボール特徴の情報量¹²⁾が多く、Harrisらの特徴点抽出により安定して抽出される目、鼻、口、顎を節点とするグラフをあてはめた。図7に顔の向きが60度の側面顔のモデルグラフを示す。図中の白

表 2 側面顔のモデルグラフの向きの変化に対する頑健性の評価
Table 2 Evaluation of robustness to view changes of a profile model graph.

	TPR(45°)	TPR(75°)	TPR(90°)
提案手法 (5 × 5)	241/300	298/300	257/300

い四角が節点を表しており、黒線は枝を表している。黒線で結ばれる 2 節点 (枝) が 1 つの識別器を構成している。モデルグラフが 4 節点から構成されているので、直積空間内で節点数 4 のクリークを与える部分グラフがこのモデルグラフとの最も高い類似度を持つ領域となる。マッチングの際は正面顔のモデルグラフと同様に各節点の周辺 5 × 5 画素で抽出したガボール特徴を利用する。また、枝の変形に関しても正面顔のモデルグラフの場合と同様にモデルグラフの 2 節点間の元の位置関係から 4 画素以内の変形を許す。

次に、側面顔のモデルグラフがどの程度の向きの変化に対応できるかを調べる。正面顔のモデルグラフの評価と同様に FPR と TPR により評価する。FPR の評価には正面顔のモデルグラフの評価に用いた PBIC の 100 枚の画像からランダムに切り出した 10,000 枚の画像を用いる。一方、TPR の評価には、正面顔のモデルグラフが 30 度の向きまで十分に対応できることから HOIP 顔画像データベースに含まれる顔の向きが 45 度、75 度、90 度の顔画像を用いた。これらの画像に対して実際に顔検出を適用し、その検出率を TPR とした。本実験では、正面顔のモデルグラフの評価実験の際に最も高い精度が得られた入力グラフの 1 節点を 5 × 5 画素に対応づける方法で評価を行った。側面顔モデルグラフを用いたときの FPR は 5/10,000 であった。そのときの TPR を表 2 に示す。表から、顔の向きが 60 度の側面顔のモデルグラフが 45 度から 90 度の向きに十分に対応できていることが分かる。つまり、正面顔のモデルグラフが 0 度から ±30 度位の向きまで対応でき、60 度の側面顔のモデルグラフが 45 度 (30 度以降) から 90 度位までの向きに対応できることが分かった。したがって、この 2 つのモデルグラフを併用することにより、広範囲な向きの変化に依存しない顔検出が可能になると考えられる。-30 度から -90 度に関しては、入力画像の鏡面画像を生成し、マッチング時に鏡面画像から抽出したガボール特徴を利用することにより、1 つの側面顔のモデルグラフだけで対応できる。

最後に、正面と側面の 2 つのモデルグラフを併用して向きに依存しない顔検出を行う。2 つのモデルグラフを併用する場合には、図 1 (c) の横軸を 8 節点 (正

面顔のモデルグラフの 4 節点 + 側面顔のモデルグラフの 4 節点) とする直積空間を用いればよく、1 回の最大クリーク抽出により、向きによらず画像中の複数人の顔を同時に検出できる。以下の実験では、直積空間内で節点数が 4 のクリーク、すなわちモデルグラフとの最大類似度となる部分グラフのみを顔として検出した。複数人の顔が含まれる画像からの顔検出結果の例を図 8 に示す。図中の白線は直積空間内で節点数 4 のクリークとなる部分グラフのすべての枝を画像上に表示したものである。図 8 から向きの変化があるにもかかわらず、安定した顔の検出ができていることが分かる。

これらの結果をふまえて提案手法の能力および限界について考察する。まず第 1 に、提案手法ではグラフを局所部分グラフに分解して変形マッチングを行い、クリーク抽出という形で各々の結果を大局的に統合することにより、グラフ間の変形マッチングを容易にし、向きの変化の影響を低減した。HOIP の顔画像データベースを用いた実験により、モデルグラフの向きによらずモデルグラフの顔の向きから ±30 度位までの向きの変化に対して 80% 以上の検出率が得られることを確認した。さらに、±30 度位までの向きの変化に対応できる正面と側面の 2 つのモデルグラフを併用することにより、広範囲な向きの変化への対応が可能となった。このことは評価実験や図 8 に示す結果から分かる。また、提案手法では画像中の人数に依存しない顔の同時検出が可能である。図 8 の結果から画像中の人数に依存せずに顔の検出ができることが分かる。画像中に複数人の顔が写っている場合、提案手法では直積空間内の異なる場所に複数のクリークが生成されるだけであるので、1 人の顔だけが写っている場合と同じく 1 回の最大クリーク抽出処理ですべての顔を検出できる。それは、モデルグラフと入力グラフのマッチング、すなわち直積空間内の枝張りが完了した後に直積空間からの最大クリーク抽出を行うからである。もちろん、何らかの影響 (顔にかかる影や隠れ等) により、ある人の顔に対応するクリークの節点数が小さくなった場合には、1 回の最大クリーク抽出処理で画像中のすべての顔を同時に検出することはできない。このような場合には、各局所部分グラフのマッチングで得られた類似度を枝の重みと見なして枝重みつき最大クリーク抽出²⁸⁾を適用すれば、影響を受けていない枝の重みが影響を受けた部分を補い、顔として検出できる可能性がある。これは今後の課題である。次に、部分的な隠れの影響について考察する。提案手法では、モデルグラフの節点となっている部分が隠れた場合に

は直積空間内でのクリークの節点数が小さくなり、検出されなくなる。しかし、それ以外の部分の隠れには対応できる。たとえば、図8の一番下の例では1人の顔の一部分が前の人物の頭で隠れているが、正面顔のモデルグラフの節点である両目、鼻、口が隠れていないため、正しく検出されている。モデルグラフの節点となっている部分が隠れた場合への対応については今後の課題であるが、上記の枝重みつき最大クリーク抽出を用いることにより対応できる可能性がある。また、画像中の顔の大きさの変化に関しては、3.1.2項で述べたようにマッチング時に入力画像の大きさを変化させることにより対応している。図8の結果から画像中の顔の大きさに影響を受けていないことが分かる。

以上の実験結果から、提案手法が画像中の顔の向きや大きさの変化の影響を受けにくく、画像中の複数人の顔を同時に検出することが可能であること、すなわち提案手法の有効性を確認できた。

4. おわりに

検出対象である人物の顔と入力画像をグラフで表現し、モデルグラフを枝ごとの局所部分グラフに分解して変形させながらマッチングを行い、直積空間に枝を張っていくと、直積空間にいくつかのクリークが生成される。直積空間に生成されたクリークの節点数はモデルグラフとの類似度を表しているの、画像からの顔検出の問題が直積空間での最大クリーク抽出問題に置換できる。しかも、直積空間はモデルグラフと入力グラフの全節点の組合せを保持しているの、マッチング時には局所的な変形しか扱っていないにもかかわらず、直積空間内でのクリーク抽出の際にはグラフの大局的な変形を厳密に評価していることになる。これにより、画像中に含まれる複数人の顔を向きの変化に依存せずに検出することが可能となった。

実験では、画像から抽出したすべての特徴を用いる従来のマッチング法、大局的なグラフ構造を用いたマッチング法と比較し、提案手法の有効性を確認した。また、モデルグラフの顔の向きによらずモデルグラフの顔の向きから ± 30 度位までの向きの変化に対して80%以上の検出率が得られることをHOIP顔画像データベースを用いた実験により確認した。さらに、 ± 30 度位までの向きの変化に対応できる正面と側面の2つのモデルグラフを併用することにより広範囲な向きの変化に依存しない顔検出を実現した。また、提案手法では画像中の人数によらずに複数人の顔を同時に検出できることを示した。

本論文では最大クリーク、すなわちクリークをなす

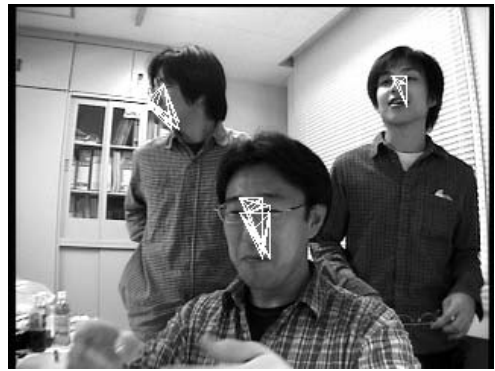


図8 顔検出の例

Fig. 8 Examples of face detection.

部分グラフの最大節点数を評価基準としているが、各局所部分グラフのマッチングの時点で類似度が得られているので、それらを積極的に利用することができれば精度はさらに向上すると考えられる。各局所部分グラフのマッチングで得られる類似度を枝の重みと見なせば、枝重みつきクリーク²⁸⁾が適用でき、類似度が最大となる部分グラフをより精密に決めることができると考えられる。また、部分的な隠れの影響等でクリークの節点数が小さくなった場合、枝重みつきクリークを利用すれば、影響を受けていない枝の重みが隠れの影響を補い、顔として検出できる可能性もある。

また、現在の枠組みでは各局所部分グラフのマッチング(識別器)が独立に機能しているが、各識別器間で Adaboost³⁰⁾のような学習を行い、枝重みつきクリーク抽出を利用すれば、枝重みつきクリーク抽出処理が各々の識別器で得られた結果の重みつき統合を行うことになり、識別精度がよりいっそう向上すると考えられる。

謝辞 本研究に対して有益なご議論をいただきました京都大学化学研究所バイオインフォマティクスセンター阿久津達也教授に感謝いたします。また、直積空間内の最大クリーク抽出アルゴリズムの開発に協力していただいた電気通信大学大学院生関友和君に感謝します。なお、本研究は電気通信大学研究・教育活性化支援システムの支援を受けている。

また、本論文に使用した HOIP 顔画像データベースは、財団法人ソフトピアジャパン研究開発部地域結集型共同研究推進室から使用許諾を受けたものである。権利者に無断で複写、利用、配布等を行うことは禁じられている。

参 考 文 献

- 1) Hjelmas, E. and Low, B.K.: Face detection: A survey, *Computer Vision and Image Understanding*, Vol.83, No.3, pp.236-274 (2001).
- 2) Yang, M.-H., Kriegman, D. and Ahuja, N.: Detecting faces in images: A survey, *IEEE Trans. Pattern Anal. & Mach. Intell.*, Vol.24, No.1, pp.34-58 (2002).
- 3) Sung, K. and Poggio, T.: Example-based learning for view-based human face detection, *IEEE Trans. Pattern Anal. & Mach. Intell.*, Vol.20, No.1, pp.39-51 (1998).
- 4) Osuna, E., Freund, R. and Girosi, F.: Training support vector machines: An application to face detection, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.130-136 (1997).
- 5) Rowley, H.A., Baluja, S. and Kanade, T.: Neural network-based face detection, *IEEE Trans. Pattern Anal. & Mach. Intell.*, Vol.20, No.1, pp.23-38 (1998).
- 6) Féraud, R., Bernier, O.J., Viallet, J-E. and Collobert, M.: A fast and accurate face detector based on neural networks, *IEEE Trans. Pattern Anal. & Mach. Intell.*, Vol.23, No.1, pp.42-53 (2001).
- 7) Schneiderman, H. and Kanade, T.: A statistical method for 3D object detection applied to faces and cars, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.746-751 (2000).
- 8) Li, Y., Gong, S. and Liddell, H.: Support vector regression and classification based multi-view face detection, *Proc. 4th IEEE International Conference on Automatic Face and Gesture Recognition*, pp.300-305 (2000).
- 9) Wu, H., Chen, Q. and Yachida, M.: Face detection from color images using a fuzzy pattern matching method, *IEEE Trans. Pattern Anal. & Mach. Intell.*, Vol.21, No.6, pp.557-563 (1999).
- 10) Yang, J. and Waibel, A.: A real-time face tracker, *3rd IEEE Workshop on Applications of Computer Vision*, pp.142-147 (1996).
- 11) Xu, G. and Sugimoto, T.: Rits eye: A software-based system for realtime face detection and tracking using pan-tilt-zoom controllable camera, *Proc. 14th International Conference on Pattern Recognition*, pp.1194-1197 (1998).
- 12) Hotta, K., Kurita, T., Umeyama, S. and Mishima, T.: Face matching through information theoretical attention points and its applications to face detection and classification, *Proc. 4th IEEE International Conference on Automatic Face and Gesture Recognition*, pp.34-39 (2000).
- 13) 堀田一弘, 三島健稔, 栗田多喜夫: 未知の画像に対する識別率を用いた顔検出のための特徴点の順序づけ, *信学論 D-II*, Vol.84-D-II, No.8, pp.1781-1789 (2001).
- 14) Serre, T., Heisele, B., Mukherjee, S. and Poggio, T.: Feature selection for face detection, A.I.Memo No.1697 (2000).
- 15) Lades, M., Vorbrüggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Würtz, R.P. and Konen, W.: Distortion invariant object recognition in the dynamic link architecture, *IEEE Trans. Comput.*, Vol.42, No.3, pp.300-311 (1993).
- 16) Jones, J.P. and Palmer, L.A.: An evaluation of the two-dimensional Gabor filter model of sim-

- ple receptive fields in the cat striate cortex, *J. Neurophysiology*, Vol.58, pp.1233-1258 (1987).
- 17) Olshausen, B.A. and Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images, *Nature*, Vol.381, No.13, pp.607-609 (1996).
- 18) Bell, A.J. and Sejnowski, T.J.: Edges are the 'independent components' of natural scenes, *Vision Research*, Vol.37, No.23, pp.3327-3338 (1997).
- 19) Harris, C. and Stephens, M.: A combined corner and edge detector, *Proc. Alvey Vision Conference*, pp.147-151 (1988).
- 20) Schmid, C., Mohr, R. and Bauckhage, C.: Evaluation of interest point detectors, *International Journal of Computer Vision*, Vol.37, No.2, pp.151-172 (2000).
- 21) Ishitani, Y.: Model-based information extraction and its applications for document images, *Proc. Workshop on Document Layout Interpretation and its Applications* (2001).
- 22) Ogawa, H.: Labeled point pattern matching by Delaunay triangulation and maximal cliques, *Pattern Recognition*, Vol.19, No.1, pp.35-40 (1986).
- 23) Bahadur, K.C.D., Akutsu, T., Tomita, E., Seki, T. and Fujiyama, A.: Point matching under non-uniform distortions and protein side chain packing based on efficient maximum clique algorithms, *Genome Informatics*, No.13, pp.143-152 (2002).
- 24) Heisele, B., Serre, T., Pontil, M. and Poggio, T.: Component-based face detection, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.657-662 (2001).
- 25) Pedestrian and Bicycle Information Center Image Library. <http://www.pedbikeimages.org/> Dan Burden.
- 26) Sim, T., Baker, S. and Bsat, M.: The CMU Pose, Illumination, and Expression (PIE) Database, *Proc. 5th IEEE International Conference on Automatic Face and Gesture Recognition*, pp.53-58 (2002).
- 27) Tomita, E. and Seki, T.: An efficient branch-and-bound algorithm for finding a maximum clique, *Proc. 4th International Conference on Discrete Mathematics and Theoretical Computer Science*, LNCS 2731, pp.278-289 (2003).
- 28) 鈴木純一, 富田悦次, 関友和: 枝重み最大クリーク抽出アルゴリズムと実験の評価, 情報処理学会研究報告, MPS-42-12, pp.45-48 (2002).
- 29) 堀田一弘, 富田悦次, 関友和, 高橋治久: 最大クリーク抽出に基づく対象検出, 情報処理学会

研究報告, MPS-42-13, pp.49-56 (2002).

- 30) Schapire, R.E., Freund, Y., Bartlett, P. and Lee, W.S.: Boosting the margin: a new explanation for the effectiveness of voting methods, *Annals of Statistics*, Vol.26, No.5, pp.1651-1686 (1998).

(平成 15 年 4 月 17 日受付)

(平成 15 年 6 月 4 日採録)



堀田 一弘 (正会員)

昭和 50 年生。平成 9 年埼玉大学工学部情報工学科卒業。平成 11 年同大学院理工学研究科情報工学専攻博士前期課程修了。平成 14 年同大学院理工学研究科情報数理科学専攻博士後期課程修了。博士(工学)。同年電気通信大学電気通信学部情報通信工学科助手, 現在に至る。平成 11 年~平成 14 年日本学術振興会特別研究員。パターン認識, コンピュータビジョンの研究に従事。電子情報通信学会, 日本顔学会, IEEE Computer Society 各会員。



富田 悦次 (正会員)

昭和 17 年生。昭和 41 年東京工業大学理工学部電子工学科卒業。昭和 46 年同大学院理工学研究科電子工学専攻博士課程修了。工学博士。同年東京工業大学工学部電子物理工学科助手を経て情報工学科助手。昭和 51 年電気通信大学電気通信学部通信工学科助教授。昭和 61 年同教授。昭和 62 年電子情報学科教授を経て, 現在情報通信工学科教授。オートマトン・言語理論, 計算論的学習理論, 組合せ最適化問題等の研究に従事。著訳書『オートマトン・言語理論』(森北出版, 共著), 『情報処理』(電気書院, 共著), 『コンピュータ基礎理論ハンドブック』(丸善, 共訳)等。昭和 46 年度電子通信学会米澤賞, 平成 15 年船井情報科学振興賞各受賞。電子情報通信学会フェロー。本会会誌編集委員会主査, 数理モデル化と問題解決研究会主査等を経て, 現在本会コンピュータサイエンス領域委員会委員長。ACM, EATCS, 人工知能学会, 神経回路学会各会員。



高橋 治久

昭和 27 年生 . 昭和 50 年電気通信
大学電気通信学部通信工学科卒業 .
昭和 52 年同大学院修士課程修了 .
昭和 55 年大阪大学大学院工学研究
科博士後期課程修了 . 工学博士 . 同
年豊橋技術科学大学助手 . 昭和 61 年電気通信大学講師
を経て現在同教授 . 現在 , 形式ニューラルネットワ
ーク , 学習等の研究に従事 . 昭和 59 年度電子情報通信
学会学術奨励賞受賞 . 電子情報通信学会 , 国際ニュー
ラルネットワーク学会各会員 .
