

最大クリーク抽出に基づく画像からの対象検出

堀田一弘, 富田悦次, 関友和, 高橋治久

電気通信大学 〒182-8585 東京都調布市調布ヶ丘 1-5-1

Email: { hotta, tomita, seki-t, takahasi }@ice.uec.ac.jp

本論文では、最大クリーク抽出に基づく対象検出法を提案する。提案手法では、検出対象のモデルと入力画像を共にグラフで表現する。モデルグラフと入力グラフで類似した2節点間の関係(枝)があった場合、直積空間内でその節点同士に枝を張っていくと、2つのグラフ間のマッチング結果は直積空間内でのクリークとなる。つまり、対象検出の問題が直積空間内でのクリーク抽出となる。この方法では直積空間内の枝の張り方が検出精度に影響を与えるので、枝を張るか否かの判断に節点の周辺領域から抽出したガボール特徴に基づくパターンマッチングを利用する。マッチングの精度を向上させるために認識に利用する特徴の選択も行う。マッチングの際、2節点間(枝)の伸縮や変形を許すことにより、向きの変化やゆがみにロバストになると期待される。提案手法を顔検出に応用し、正面の顔画像だけからモデルグラフを作成したにも関わらず、正面から30度位の向きの変化に対してロバストな顔検出が可能であった。

Object Detection Method Based on Maximum Cliques

Kazuhiro HOTTA, Etsuji TOMITA, Tomokazu SEKI, Haruhisa TAKAHASHI

The University of Electro-Communications

1-5-1 Chofugaoka, Chofu, Tokyo, 182-8585 Japan

This paper presents an object detection method based on maximum cliques. Object model and input image are represented by graph. The two vertices are connected in direction product space when model and input graph have similar relation of two vertices. As a result, the matching result of two graphs forms cliques in direction product space. Therefore, object detection problem turns into the maximum clique extraction problem. Since the accuracy of detection depends on the edges, pattern matching which is based on Gabor features extracted from two vertices is used to decide whether two vertices should be connected or not. Feature selection is also applied to improve the accuracy of pattern matching. Although model graph is created by frontal faces, face detection which is robust to the view changes of about 30 degrees is realized by allowing the modification of edges.

1 はじめに

画像から特定の対象を見つける対象検出は、コンピュータによる対象の自動認識の第一ステップとして重要であり、近年盛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]、認識に利用する特徴の選択 [7, 8] 等により、対象の向きが限定されている状況下では安定した対象検出が可能となってきている [1, 2]。しかし、人間が生活する実環境下では、対象が必ず正面を向いているとは限らない。このような現状から、本論文では向きの変化

にロバストな対象検出の問題を扱うことにした。

向きの変化を扱った従来手法では複数の向きモデルを用意することにより対処していた [5]。しかし、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残る。従って、1つのモデルが対応できる向きの範囲を広げることができれば、全体として少ないモデル数で高精度な検出が可能になると考えられる。

Ladesらはグラフを用いた向きや表情の変化にロバストな個人識別法を提案した [9]。この手法ではモデルをグラフで表現し、さらに、入力画像にもグラフを当てはめる。そして、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当てはめたグラフを変形させることにより、多少の向きや表情変化にロバスト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当てはめたグラフを変形させるため、そのままの枠組では画像から複数の対象を検出しなければならない対象検出に応用することは難しい。従って、グラフの持つ柔軟な特性を生かしたまま対象検出の枠組に当てはめることができれば、柔軟な対象検出が可能になると考えられる。

検出対象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格子状のグラフであると考え、2つのグラフの関係は全ての節点間の組合せを保持した直積空間内で表現される。モデルグラフと入力グラフで互いに類似した2節点間の関係(枝)があった場合、直積空間内でその節点間に枝を張っていくと、この2つのグラフ間のマッチング結果は直積空間内のクリークとなる。入力画像に複数の対象が含まれる場合、この方法では直積空間内に同じ節点数のクリークが複数できることになる。つまり、入力画像からの対象検出が直積空間内のクリーク抽出となるのである。これが本論文でのグラフを用いた対象検出の基本型である。

最大クリーク抽出を用いたマッチング法は、これまでも様々な分野で応用されている。例えば、IshitaniはDocument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[15]。また、Ogawaらは星座のマッチングに応用している[16]。Bahadurらはタンパク質側鎖パッキングに最大クリーク抽出を利用している[17]。

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ことになる。そこで、枝を張るか否かの判断に従来の対象検出法で利用されている統計的パターン認識の方法を利用する[1, 2]。具体的には、顔検出[7]や対象識別[9]に有効であるガボール特徴を利用し、マッチングの精度を向上させるために認識に利用する特徴の選択を行った[7]。モデルグラフと入力グラフのマッチングの際、枝の伸縮や変形を許すことにより、向きの変化やゆがみにロバストになると考えられる。

本論文では、サンプルを比較的容易に収集できることと実現した時の応用範囲の広さから人間の顔を認識対象とした。実験では、正面顔画像だけからモデルグラフを作成したにも関わらず、正面から30度位の向き

の変化に対してロバストな顔検出が可能であることを確認した。

以下、2節では最大クリーク抽出に基づく対象検出法について説明する。3節では提案手法を顔検出に応用した結果を示す。最後に、まとめと今後の課題について4節で述べる。

2 最大クリーク抽出に基づく対象検出法

本節では、最大クリーク抽出に基づく対象検出法の概略を説明する。提案手法の概念図を図1に示す。図の左はグラフにより対象をモデル化した例であり、白い四角が節点を表し、黒線が枝を表している。図の中央は入力画像に当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。(ここでは説明を簡単にするために節点数を少なくしている。)図の右はモデルグラフ(4節点)と入力グラフ(8節点)の直積空間である。縦軸が入力の8節点を表し、横軸がモデルの4節点を表している。モデルグラフの節点 p と節点 q の2節点間の関係(枝)と入力グラフの節点 s と節点 t の関係(枝)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) に枝を張る。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同士に枝が張られ、直積空間内にクリークができる。これにより対象検出の問題が直積空間内のクリーク抽出となり、最大クリークを与える部分グラフが最も対象らしい領域となる。図1では1人の顔の例を示したが、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できることになる。

本節では、提案手法を効率的に行うために利用するいくつかの方法を説明する。まず、Harrisの特徴点抽出法[13]を用いた入力画像に当てはめるグラフの節点選択について2.1節で説明する。入力画像の全画素を節点とすると、モデルとの組合せ数が膨大になり、対象検出に多大な時間を要する。そこで、Harrisの特徴点選択を用いて入力画像から節点の選択を行う。モデルグラフの当てはめについては3節の実験の所で説明する。

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

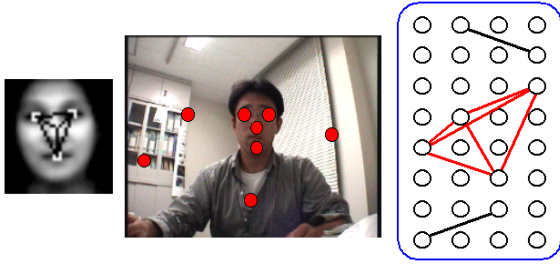


図 1: 最大クリーク抽出に基づく対象検出の概念図

依存する。安定した枝張りを行うために、枝を張るか否かの判断にパターン認識の方法を用いる。ここでは、顔検出 [7] や対象識別 [9] で実績のあるガボール特徴を利用し、さらに、識別能力を向上させるために変数選択を利用して認識に不要な情報の削減も行った [7]。ガボールフィルタと変数選択を利用した不要な情報の削減法はそれぞれ 2.2.1 節と 2.2.2 節で説明する。

最後に、直積空間内の最大クリーク抽出に利用する関らの提案した分枝限定法を用いて厳密解を高速に得る MCQ アルゴリズム [18] を 2.3 節で説明する。

2.1 Harris の特徴点抽出法

本論文では、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。節点の選択により対象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 [14]。その結果、Harris 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 の特徴点抽出法 [13] を利用して入力画像から節点の選択を行う。以下に、Harris の特徴点抽出法の概略を説明する。

基本となる評価基準は、ある点を中心とした局所領域を少し動かした時の変化量

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

である。つまり、現在評価している領域が平であれば動かした時の変化量が小さく、エッジであればエッジに沿った方向に動かした時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基準である。ここで、 $w_{u,v}$ はガウシアン重みである。

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

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

となり、局所領域を (x, y) だけ変移させた時の変化量は、

$$\begin{aligned} E_{x,y} &= Ax^2 + 2Cxy + By^2, \\ &= (x \ y)M(x \ y)^T, \end{aligned} \quad (3)$$

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。ここで、 \otimes は畳み込み積分を表している。式 (3) から $E_{x,y}$ は行列 M に依存することが分かる。従って、行列 M の 2 つの固有値を調べれば領域が平、エッジ、特徴点であるかの判定が可能となる。2 つの固有値が共に小さい値であれば平、片方だけ大きな値であればエッジ、2 つ共大きな値であれば特徴点と判定できる。Harris は固有値問題を解かなくても特徴点らしさを判定できる指標

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

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、ある点を中心とする領域の特徴点らしさ R 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良い。以下の実験では、 $\alpha = 0.04$ とした。

2.2 モデルグラフと入力グラフのマッチング

本節では、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。ここでは、枝を構成する 2 つの節点から抽出したガボール特徴を基に識別を行う。つまり、モデルグラフの任意の 2 節点間の枝が 1 つの識別器となる。識別方式としては SVM 等を利用することも可能であるが、提案手法ではマッチングの回数が多くなるので、計算が簡単なモデル特徴との距離によるマッチングを用いる。具体的には、入力グラフ中の 1 つの枝を構成する 2 節点で独立にガボール特徴を抽出し、それらを縦に並べてで

きる特徴ベクトルとモデルグラフの2節点から抽出したモデル特徴ベクトルとの距離を測り、閾値処理で枝を張るか否かを決定する。だが、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と考えられる。本論文では、識別能力をさらに向上させるために、変数選択を利用した認識に不要な情報の削減も行った [7]。これにより、識別能力の向上と高速化が併せて可能となる。

また、マッチング時のグラフの変形に関しては、全ての変形を許すと組合せ数が増大したり、通常ではあり得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、それ以上の変形は扱わないこととした。これにより、マッチングの候補数を削減でき、高速化にもつながる。

まず、ガボールフィルタについて 2.2.1 節で説明し、次に、不要な情報の削減による識別能力の向上法について 2.2.2 節で説明する。

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 [10]。ガボールフィルタは、

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

のように定義される [9]。ここで、 $\exp(-\sigma^2/2)$ は直流成分 0 にするための項である。式 (5) 中のパラメタは、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、各点で抽出した 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、各点毎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 (sparse coding) になっていることが知られている。例えば、自然画像から切り出した局所領域のスパースさを

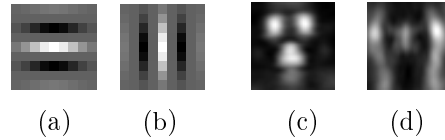


図 2: ガボールフィルタとその適用例

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 [11]。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。

顔画像にガボールフィルタを適用したときの結果を図 2 に示す。図 2(a),(b) は $\mu = 0, 4$ 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時だけ選択的に反応していることが分かる。この図から、方位の異なるガボールフィルタの出力がスパースな情報表現であることが分かる。

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既に示されている [7, 9]。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であるからだと言える。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 [12] 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。

2.2.2 変数選択による不要な特徴の削減

本論文では、マッチングの安定化のために、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する。しかし、2 節点を中心とした局所領域内のガボール特徴には、認識に有効な情報と有効でない情報が含まれていると考えられる。そこで、高速化と識別能力の向上を目的として、変数選択を利用した認識に不要な情報の削減を行う。ここでは、未学習サンプルに対する誤識別率を選択の基準とした。また、変数選択の方法として、全ての特徴量を利用する場合からスター

トし、認識に不要な情報を1変数ずつ取り除いて行く Sequential Backward Selection(SBS) を利用した。

以下に、SBS を用いた不要な情報の削減法を説明する。まず、モデルとなる特徴ベクトルを作成し、対象と対象以外の未学習サンプルから特徴を抽出しておく。次に、モデルとなる特徴ベクトルの中からランダムに1変数を取り除く。未学習サンプルの特徴ベクトルからも同じ変数を取り除き、その時の誤識別率を計算する。この処理をモデル特徴ベクトル内の全ての変数に対して独立に行い、取り除いた時に最も誤識別率が高くなる変数を不要な変数として取り除く。取り除いた変数は二度と利用しない。この操作を取り除く変数がなくなるまで繰り返すことにより、不要な情報が取り除かれ、認識するのに必要最小限の変数の組が得られる。変数選択はモデルグラフの全ての2節点間(枝)の識別器で独立に行い、未学習サンプルに対して最も誤識別率が低くなる変数の組を選択する。選択された場所等の具体的な例は3節に示す。

2.3 最大クリーク抽出アルゴリズム

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、関らは分枝限定法を用いて厳密解を高速に得る MCQ アルゴリズムを提案している [18]。本論文では、関らの MCQ アルゴリズムを用いて直積空間内の最大クリーク、即ち検出結果を得る。

以下に MCQ アルゴリズムを簡単に説明するが、詳細は文献 [18] を参照して頂きたい。

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れば良い。つまり、全節点集合を根とし、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 に隣接する全ての節点を R 中から選択して新たな集合 Q を作成し、これを R との親子関係として枝で結ぶことにより、探索木として表現できる。この探索木の枝を実際に探索することなく除去することができれば高速化につながるのである。これを行うのが分枝限定法である。

隣接している節点同士に異なる色をつけていくと、クリークをなす節点同士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みな

すことができる。従って、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、クリーク探索の高速化につながる。なぜなら、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足し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れば、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できないので、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が可能となるのである。

MCQ アルゴリズムでは、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、高速化できるのである。これまでに提案されている手法や次数の大きい順に探索する場合と比較を行い、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クリーク抽出が可能となることが確認されている [18]。

3 顔検出問題への応用

本論文ではサンプルが比較的容易に収集できると実現した時の応用範囲の広さから人間の顔を認識対象とする。正面顔だけを用いてモデルグラフを作成し、顔の向きの変化に対して柔軟に対処できるか否かを確認する。

本実験では、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 [6] 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と位置はそれぞれ違っているので、目、鼻、口の位置を用いてアフィン変換により 46×44 画素の大きさに正規化した。ただし、顔の大きさや位置が目で見えていないものは除外し、計936枚の正面顔画像を得た。これをランダムにモデル特徴作成用の136枚と変数選択用の800枚に分けて利用した。

次に、モデルグラフの当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などは、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、本論文ではグラフを用いることにより柔軟な検出が可能となるか否かを確認することに主眼を置き、Harris 特徴点抽出により安定して抽出される左目、右目、鼻、口をモデルグラフの節点とした。

以下の実験で利用するモデルグラフを図3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。黒線で結ばれる節点同士が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節点の関係(枝)を探ることになる。従って、直積空間内の節点数4のクリークを与える部分グラフがこのモデルと最も類似度が高い領域となる。

次に、マッチングの高速化と識別能力の向上のために変数選択を用いて不要な情報を除去した結果を示す。上記136枚の正面顔画像でモデル特徴を作成し、800枚の正面顔画像と顔が含まれていない画像からランダムに選択した17,550枚の顔以外の画像に対する誤識別率を選択の基準とした。ここでは、 11×11 の大きさの2節点から抽出した8方向のガボール特徴を利用するので、特徴の次元数は $11 \times 11 \times 2 \times 8 = 1,936$ である。全ての特徴量を用いた場合からスタートし、1変数ずつ減らしていく。左目と右目の2節点で抽出した特徴を用いた識別器において変数選択を行った時の誤識別したサンプル数の変化を図4に示す。図から、不要な変数を削減することにより識別能力が向上することが分かる。少ない特徴量で認識できるので、高速化も併せて実現できたことになる。実際にマッチングを行う際には、各2節点の特徴に基づく識別器において未学習サンプルに対する誤識別率が最も低くなる変数の組を用いた。図5にモデルグラフ内の各識別器で実際にマッチングに利用する場所、即ち変数選択により選択された変数を示す。図中の白点が認識に有効であるとして選択された変数である。ここでは、見やすくするために方向成分を無視して場所のみを表示したが、ガボール特徴は8方向あるので、図中の白点の場所で抽出した全ての方向成分を利用するわけではない。

次に、向きの変化への対応に関してであるが、モデルグラフの節点(両目、鼻、口)が充分に見える範囲であれば、グラフの変形により対処することができると考えられる。しかし、任意の変形を許すと、組合せ数が増加してマッチング回数が増え、通常ではありえない変形での誤識別が起こる可能性もあるので、変形の範囲を限定し、それ以上の変形は扱わないことにした。以下の実験では、モデルグラフの2節点間の関係から距離が4画素以内の変形をマッチングの候補とし、それ以外は候補から省いた。

実際に入力画像から顔検出を行う際は、入力画像内の顔の位置や大きさが未知であるので、大きさと位置

を変えながらマッチングを行う。また、Harrisの特徴点抽出法は大きさの変化に影響を受けやすい[14]ので、入力画像の大きさを変える毎に特徴点抽出を行った。以下の実験では、入力画像の大きさを1/1.1倍づつを変化させた。

正面顔画像から作成したモデルを用いて、向きの変化にどれくらい対応できるかを調べるために、HOIPの顔画像データベースに含まれる顔の向きが15度、30度、45度の場合の各々300人の顔画像を用いて検出実験を行った。ここでは、入力画像内の顔の位置に対応する部分グラフが入力画像中の最大クリークとなっていれば同じ節点数のクリークで誤検出があっても正解とし、正解率とそのクリークの節点数により評価した。

顔の向きが15度の時は、300人中299人の顔が正しく検出され、その内の296人の顔が節点数4のクリークであった。ここで、節点数4のクリークはモデルグラフと最も類似していることを意味している。顔の向きが30度の時も300人中299人の顔が正しく検出され、その内の268人の顔が節点数4のクリークであった。顔の向きが45度の時は300人中264人の顔に対応する部分グラフが最大クリークとなったが、節点数が4のクリークとなったのはその内の109人だけであった。顔の向きが45度になると、片方の目を逃すことが多く、節点数が3の最大クリークが増加する。この実験では、最大クリークの節点数が4の場合には誤検出が少なかったが、最大クリークの節点数が3になると誤検出を含むことが多くなった。

この実験結果をまとめると、30度までの向きの変化では節点数が4のクリークにより安定した検出が可能であった。しかし、45度になると節点数が3のクリークが最大となることが多くなり、誤検出が増えた。つまり、提案手法(ここで利用したモデルグラフ)では、正面顔だけからモデルグラフを作成したにもかかわらず正面から30度の向きの変化までは充分に対応できたと言える。

最後に、向きを変化させた状況下で撮影した顔画像に対する顔検出結果を図6に示す。ここでは、Harrisの特徴点抽出で入力画像から節点の選択を行っているが、入力画像が大きくなると節点数が多くなり、メモリや計算時間の問題が起こる。そこで、Harrisの特徴点抽出で選択された点に対し、その周辺 5×5 の領域から抽出したガボール特徴を用いてマッチングを行い、閾値以上の類似度を与える場所だけを入力画像の節点とした。これにより選択された節点だけに対して 11×11

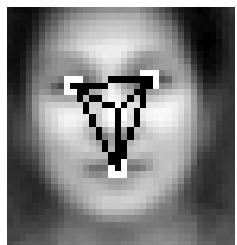


図 3: モデルグラフ

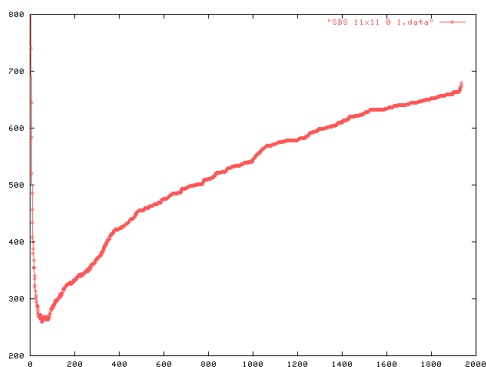


図 4: モデルグラフの左目と右目の 2 節点で抽出した特徴を用いた識別器における変数選択

の識別器を適用した。ここで、 5×5 の識別器に関しても 11×11 の識別器と同様に不要な情報を削減している。

図 6(a) は 1 人の顔の向きを上下左右に変化させた時の結果である。図中の白線は直積空間内の枝に対応しており、最大クリークを形成する部分グラフを全て表示している。線が複数重なっているのは、近い領域で同じ節点数のクリークが複数得られているからである。モデルグラフの節点である両目、鼻、口が充分に見える範囲では、安定した検出ができていのが分かる。複数人の顔が含まれる画像からの顔検出結果を図 6(b) に示す。図から、顔の向きの変化があるにも関わらず、4 人の顔が正確に検出されていることが分かる。

4 おわりに

対象をグラフでモデル化し、入力画像にもグラフを当てはめると、画像からの対象検出は直積空間でのクリーク抽出となる。これによりグラフの利点を生かし

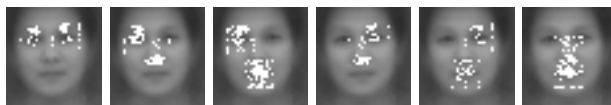


図 5: モデルグラフ内の各識別器でマッチングに利用する変数

つつ画像内に含まれる複数の対象を検出することが可能となる。実験では正面顔のモデルしか持っていないにもかかわらず、グラフの 2 節点間 (枝) の変形を許すことにより、30 度位の向きの変化に充分対応できることを確認した。

本論文では最大クリーク、即ち全ての 2 節点間の関係が同時に成り立てば対象であると判断しているが、本来はマッチングの時点で類似度が得られているので、それを無視する必要はない。各識別器で得られる類似度を枝に付いた重みと見なせば、重みつきクリーク [19] が適用でき、類似度が最大となる部分グラフを一意に決めることができると考えられる。

また、現在の枠組では、各々の識別器が独立に機能しているが、識別器間で Adaboost[20] のような学習を行い、重みつきクリークを利用すれば、クリーク抽出が各々の識別器で得られた結果の統合を行うことになり、識別能力がより一層向上すると考えられる。

謝辞

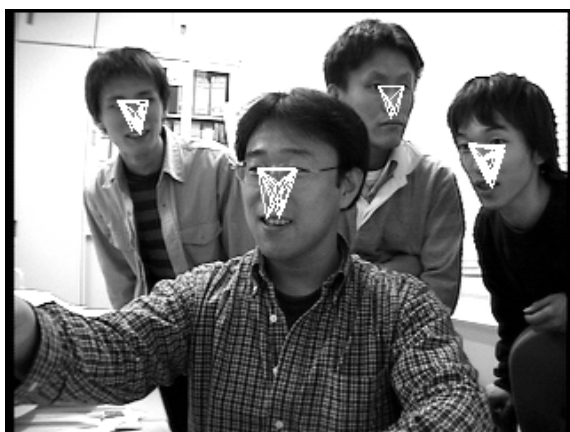
本論文に使用した顔画像データは、財団法人ソフトウェアジャパン研究開発部地域結集型共同研究推進室から使用許諾を受けたものです。権利者に無断で複製、利用、配布等を行うことは禁じられています。

参考文献

- [1] E.Hjelmas and B.K.Low, "Face detection: A survey," Computer Vision and Image Understanding, Vol.83, No.3, pp. 236-274, 2001.
- [2] M.-H.Yang, D.Kriegman, and N.Ahuja, "Detecting faces in images: A survey," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.24, No.1, pp. 34-58, Jan., 2002.
- [3] K.Sung and T.Poggio, "Example-based learning for view-based human face detection," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.20, No.1, pp. 39-51, 1998.



(a)



(b)

図 6: 顔検出例

- [4] E.Osuna, R.Freund, and F.Girosi, "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] H.Schneiderman and T.Kanade, "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.
- [6] H.A.Rowley, S.Baluja, and T.Kanade, "Neural network-based face detection," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.20, No.1, pp. 23-38, 1998.
- [7] 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, Aug., 2001.
- [8] T.Serre, B.Heisele, S.Mukherjee, and T.Poggio, "Feature selection for face detection," A.I.Memo No.1697, Sep., 2000.
- [9] M.Lades, J.C.Vorbrüggen, J.Buhmann, J.Lange, C.von der Malsburg, R.P.Würtz, and W.Konen, "Distortion invariant object recognition in the dynamic link architecture," IEEE Trans. Computer, Vol.42, No.3, pp. 300-311, Mar., 1993.
- [10] J.P.Jones and L.A.Palmer, "An Evaluation of the two-dimensional Gabor filter model of simple receptive fields in the cat striate cortex," J. Neurophysiology, Vol.58, pp. 1233-1258, Dec., 1987.
- [11] B.A.Olshausen and D.J.Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," Nature, Vol.381, No.13, pp. 607-609, 1996.
- [12] A.J. Bell and T.J.Sejnowski, "Edges are the 'independent components' of natural scenes," Vision Research, Vol.37, No.23, pp. 3327-3338, 1997.
- [13] C.Harris and M.Stephens, "A Combined Corner and Edge Detector," Proc. Alvey Vision Conference, pp. 147-151, 1988.
- [14] C.Schmid, R.Mohr, and C.Bauckhage, "Evaluation of Interest Point Detectors," International Journal of Computer Vision, Vol.37, No.2, pp. 151-172, 2000.
- [15] Y.Ishitani, "Model-Based Information Extraction and its Applications for Document Images," Proc. Workshop on Document Layout Interpretation and its Applications, 2001.
- [16] H.Ogawa, "Labeled Point Pattern Matching By Delaunay Triangulation and Maximal Cliques," Pattern Recognition, Vol.19, No.1, pp. 35-40, 1986.
- [17] D.Bahadur K.C., T.Akutsu, E.Tomita, T.Seki, and A.Fujiyama, "Point matching under non-uniform distortions and protein side chain packing based on efficient maximum clique algorithms," Genome Informatics, No.13(in press), 2002.
- [18] 関 友和, 富田 悦次, "分枝限定法を用いた最大クリーク抽出アルゴリズムの効率化," 信学技報, COMP2001-50, pp. 101-108, 2001.
- [19] 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的評価," 第 42 回 数理モデル化と問題解決研究会, 2002.
- [20] R.E.Schapire, Y.Freund, P.Bartlett, and W.S.Lee, "Boosting the margin: a new explanation for the effectiveness of voting methods," Annals of Statistics, Vol.26, No.5, pp. 1651-1686, 1998.