

# 局所特徴量と近似最近傍探索を用いた 大規模データベースに対する高速顔認識

内海 ゆづ子<sup>2,a)</sup> 坂野 悠司<sup>1,b)</sup> 前川 敬介<sup>2,c)</sup> 岩村 雅一<sup>2,d)</sup> 黄瀬 浩一<sup>2,e)</sup>

概要：本稿では、犯罪捜査支援を目的とした大規模データベースに対する高速な候補の絞り込み手法を提案する。犯罪捜査支援を目的とした顔認識に必要とされる条件は、高速かつ高精度な絞り込みが可能であり、低解像度画像であっても、精度を落とすことなく絞り込み可能であることである。そこで、局所特徴量と近似最近傍探索に基づく特徴量の投票処理を組み合わせた手法を用いることで、高速で高精度な顔認識を実現する。また、低解像度画像に対する認識精度向上のため、Dense sampling を用いた特徴抽出を行う。1000 万顔画像データベースを作成し、顔認識実験を行った結果、精度 98.5 [%] で 185 [ms] の探索が可能であった。また、低解像度画像に対して Dense sampling を用いて特徴抽出した場合、10 万枚顔画像データベースに対して精度 97.7[%] で 53[ms] の探索が可能となり、提案手法が低解像度画像に対しても有効に働くことが示された。

## 1. はじめに

監視カメラが商店街、学校、駅などの公共施設に取り付けられるようになって久しい。犯罪が起きた際、監視カメラ画像を犯人逮捕の手がかりにすることも珍しくなく、今後、監視カメラ画像が犯罪捜査に用いられる機会がますます増えていくと考えられる。現在、監視カメラ画像を犯罪捜査に用いる場合、画像中から犯人と思われる人物の特定、特定した人物の顔画像の切り出し、顔画像の犯罪者データベース上での検索といった作業はすべて人手により行われている。このうち、データベースに対する検索では、データベースの規模が大きい場合、捜査の大きな負荷となる。検索の際に、自動的にデータベースから候補者の絞り込みができれば、人手と捜査時間の大幅な減少につながると考えられる。

そこで、本稿では、犯罪捜査の支援を目的とし、大規模な顔画像データベースから犯罪者の候補を絞り込む手法の提案を行う。犯罪捜査の支援を行う場合、まず必要とされる条件は、高速かつ高精度な絞り込みである。続いて、

データベースにデータが追加された場合でも、簡単にデータベースが拡張できる手法がであることが望ましい。これは、犯罪者データベースには、頻繁にデータが追加されることが考えられるため、データベースの拡張を容易にできる方が捜査に有利となるからである。また、犯罪者の顔画像は、1 人あたり数枚の画像しか保存されておらず、データベースにある画像が 1 人あたり 1 枚であっても認識のできる手法でなければならない。さらに、監視カメラ画像の解像度は一般に低く、低解像度画像であっても、精度を落とすことなく絞り込みが可能であることが望ましい。

これらの条件を満たすため、我々は局所特徴量と投票を組み合わせた顔認識手法を提案する。提案手法では、局所特徴量のマッチングに最近傍探索を用いる。このことにより、新たにデータベースにデータが登録された場合でも、学習をせずに認識を行うことができる。また、最近傍探索に近似を行うことで、高速にマッチングをすることになり、認識にかかる時間を大幅に減らすことができる。最近傍探索による特徴量 1 つ 1 つのマッチングの精度はあまり高くないが、これらのマッチング結果の統合に投票を用いることで、精度良く絞り込みができる。最後に、低解像度画像では、局所特徴量検出で得られる特徴量が少ないため、認識率が低下する。そこで、画像上から特徴点の Dense sampling を行い特徴抽出をすることで、低解像度画像であっても精度良く絞り込みを行う。犯罪捜査に用いられるデータベースの規模を 1000 万人と想定し、その中から候補を 1000 人に絞り込むこととした。実験の結果、1000 万

<sup>1</sup> 大阪府立大学 工学部 知能情報工学科 〒 599-8531 大阪府堺市中区学園町 1-1

<sup>2</sup> 大阪府立大学 大学院工学研究科 〒 599-8531 大阪府堺市中区学園町 1-1

a) yuzuko@cs.osakafu-u.ac.jp

b) sakano@m.cs.osakafu-u.ac.jp

c) maekawa@m.cs.osakafu-u.ac.jp

d) masa@cs.osakafu-u.ac.jp

e) kise@cs.osakafu-u.ac.jp

顔画像データベースに対して，提案手法は画像1枚あたり平均185 [ms]での探索が可能となり，1000位累積認識率は98.5[%]を達成した．

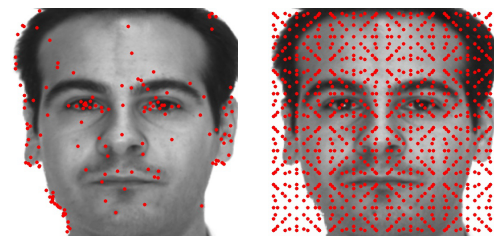
以降，2節ではこれまでの顔認識手法や公開データベースについて述べ，3節で本稿で局所特徴量として用いたPCA-SIFT特徴量の画像からの抽出手法，4節では認識に用いた投票の手法と近似最近傍探索について説明する．5節では提案手法の大規模データベースに関する評価実験と低解像度画像に対する評価実験の結果について述べたのち，6節で本稿をまとめる．

## 2. 関連研究

顔認識はコンピュータビジョンの分野で重要な研究課題の1つであり，これまで，数多くの顔認識手法が提案されてきた．代表的なものに，線形判別分析 [5] や独立成分分析 (ICA) [28] といった多変量解析を用いたものや，SVM [7] や多様体 [9] を用いた認識手法などが提案されている．これらの手法は，学習に1人あたり複数枚の画像を用いることを前提としており，本稿で想定しているような1人あたり1枚しかデータベースに画像がない場合，適用が困難となる．また，これらは認識の精度をもとに手法の評価をしており，学習や認識にかかる計算時間に対する評価があまりされていない．よって，データベースのスケールが拡大した場合，認識を高速に行うことができるか不明である．大規模なデータベースに対して高速に認識を行う手法も提案されている [26]．この手法では，木構造を用いて，認識を高速に実現している．しかし，新しくデータがデータベースに追加された場合，木構造を最初から作り直す必要がある．犯罪捜査で利用されるデータベースのようにデータベースの更新が頻繁に行われると，学習にかかる時間が増加する問題が起こる．

そこで，本研究では，認識に最近傍探索に基づいた投票を行うことで，1人あたり1枚しかデータがない場合でも検索を実現する．また，最近傍探索は事前の学習が必要ないため，新たにデータがデータベースに追加された場合，特徴をデータベースに登録するだけで，認識が可能となる．

顔認識の研究に合わせて，顔認識の評価のベンチマークとして，顔データベースが様々な企業や研究機関から発表されるようになった．代表的なデータベースに含まれる画像の枚数，人物の人数，画像の種類をまとめたものを表1に示す．表1から，データベースに含まれている画像の枚数は数百枚から数十万枚程度である．そのため，1000万枚規模のデータベースに対する精度や処理時間の評価には，新しく顔画像データベースを作成しなければならない．そこで，公開されている顔画像データベースとweb上からダウンロードされたデータベースを用いて新たに1000万顔画像データベースを作成し，評価を行った．



(a) 特徴点検出 (b) Dense sampling

図1 特徴点の例．赤い点が特徴点を示す．

## 3. 特徴抽出

本稿では，顔画像からPCA-SIFT [18] 特徴量を抽出し，顔の候補絞り込みに用いる．PCA-SIFTは特徴点のまわりの画素のx, y方向のエッジ強度を計算し，それを直交基底に写像して，次元圧縮をしたものである．PCA-SIFT特徴量は回転，スケール変化，照明変化などに頑健な特徴量であると言われ，物体認識で広く用いられている [1]．

PCA-SIFTは特徴点を検出する段階と得られた特徴点から特徴量を抽出する段階の2段階に分かれている．以下の節では，それぞれについて詳しく説明する．

### 3.1 特徴点の検出

本稿では，PCA-SIFTで一般的に用いられている特徴点を自動で検出する手法と，特徴点をDense samplingする手法の2つを用いた．まず，特徴点を自動で検出する手法について説明する．特徴点を検出する顔画像に対して，複数のスケールのガウス関数で平滑化を行い，スケールの異なる平滑化画像の差分を求め，差分画像から極値を検出し，特徴点の候補とする．最後に特徴点候補中から，エッジ上の点とノイズを削除することで安定して特徴点検出される点を取り出す．特徴点を自動検出した例を図1(a)に示す．図1(a)から，顔の目や鼻，口といった，エッジ情報の多い特徴的な部分から特徴点検出されていることがわかる．

顔画像の解像度が十分高い場合，特徴点検出により特徴点が大量に得られる．しかし，解像度が低い場合，特徴点検出により得られる特徴点の数が減少し，認識率に影響を与える．そこで，低解像度画像からも一定の特徴点を得るため，特徴点をDense samplingする．Dense samplingでは，画像上からある一定の間隔で格子状に特徴点を配置する．Dense samplingを用いた場合の特徴点を図1(b)に示す．

### 3.2 特徴量抽出

PCA-SIFT特徴量は，特徴点の周りの画素値のエッジ強度をもとに計算される．特徴抽出をする場合はまず，特徴量の抽出は，得られた特徴点の周り $39 \times 39$  [pixel]のx, y方向のエッジ強度を計算する．得られた $39 \times 39 \times 2 = 3042$

表 1 主なデータベースと含まれる画像の枚数

データベース名称	画像の枚数	人数	顔画像の種類
AT & T The Database of Faces[10]	400	40	表情変化
The University of Oulu Physics-Based Face Database[11]	2,000	125	顔向き変化, 照明変化, 表情変化
The AR Face Database[12]	4,000	126	表情変化, 照明変化, アクセサリ
The Yale Face Database B[13]	5,760	38	照明変化
The Color FERET Fatabase[14]	14,126	994	顔向き変化
PIE Database[15]	41,386	68	顔向き変化, 照明変化, 表情変化
CAS_PEAL Face Database [16]	99,594	1,040	顔向き変化, 表情変化, アクセサリ, 照明変化
Multi PIE[17]	750,000	337	顔向き変化, 照明変化, 表情変化

次元の特徴量をあらかじめ他の画像から計算していた 36 次元の直交基底空間に写像する．この圧縮された 36 次元の特徴量が PCA-SIFT 特徴量である．本稿では，一般画像から自動的に検出された特徴点から得られたエッジ強度の特徴量に PCA をかけることで，36 次元の直交基底を計算した．

#### 4. 検索候補の絞り込み

得られた特徴量をもとに，データベースから候補の絞り込みを行う．本稿では，特徴量のマッチングにユークリッド距離に基づく最近傍探索を行い，マッチングの上位  $k$  個の特徴量を用いて投票を行う．特徴量のマッチングには，特徴量を抽出した特徴点の位置をもとに，マッチングに用いる特徴量を限定した．また，絞り込みの高速化を図るため，マッチングに近似最近傍探索を適用した．以下，マッチングに用いる特徴点の限定手法， $k$  近傍を用いた投票処理，マッチングを高速化する近似最近傍探索について説明する．

##### 4.1 特徴量の限定

特徴量のマッチングに，データベースのすべての特徴量を用いた場合，特徴量が顔の別の部位と誤対応する可能性がある．具体的には，鼻のまわりから抽出された特徴量を目のまわりから抽出された特徴量に誤対応する場合である．このような誤対応は認識率の低下の原因となる．そこで，検索対象の特徴量を抽出した特徴点をもとに，特徴量のマッチングに用いる特徴量を限定した．特徴点を自動検出した場合は，図 2(a) に示すように，検索画像から検出された特徴点の座標から，半径  $a$  以内の特徴点のみを用いる．また，Dense sampling を用いた場合は，図 2(b) に示すように，同じ座標から抽出された特徴量のみをマッチングに用いる．

##### 4.2 $k$ 近傍探索による投票処理

特徴量のマッチングは，検索画像の特徴量とデータベースの特徴量のユークリッド距離に基づいて最近傍探索により行う．そして，検索画像の特徴量から  $k$  近傍の特徴量の顔画像 ID に投票する．投票の際に，特徴点からの距離の

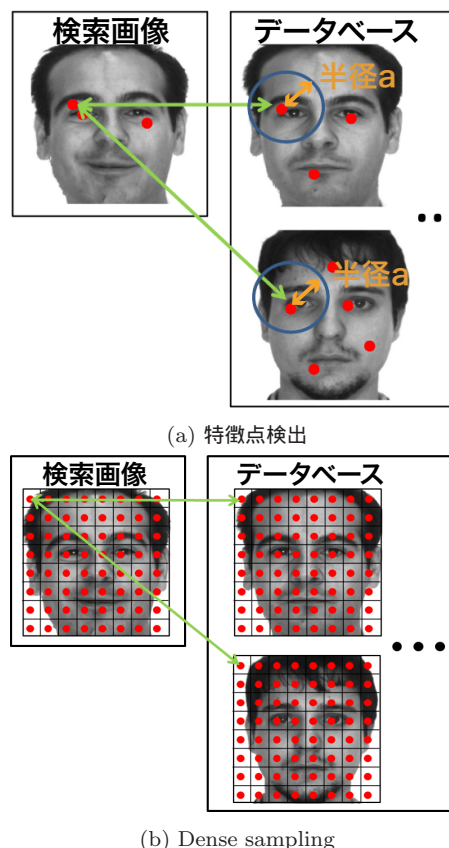


図 2 マッチングに用いる特徴量の限定手法

逆比  $\frac{1}{d}$  の重みをつけて投票を行う．このことにより，正解候補の絞り込みの精度を向上させることができる．マッチングと投票を検索画像から得られたすべての特徴量で行い，最終的に得票の多い ID 上位  $n$  個を正解画像候補とする．

特徴量の最近傍探索の精度はあまり高くない．しかし，投票と組み合わせることで，認識精度を向上させることが可能である．特徴量の検索が失敗した場合，投票がランダムに行われると仮定すると，複数の特徴量が同一の顔画像 ID に誤投票される確率は，正解の顔画像 ID に複数の特徴量から投票される確率よりもはるかに低くなる．よって，複数の特徴量から投票がなされ，他の顔画像 ID よりも得票数が多くなるものは，正解の確率が高いといえる．よって，投票による認識は，一つ一つの特徴量の検索精度が低くても，高精度に認識可能である．



### 4.3 近似最近傍探索による高速化

特徴量の探索において、探索対象の特徴量すべてに探索処理を行うと、探索対象すべての特徴量とのユークリッド距離を計算することになり、データベースの規模に比例して探索時間が増加する。そこで、探索を高速化するために  $k$  近傍探索に近似を用いる。

本稿では、Bucket Distance Hashing (BDH) [4] を用いることで探索を高速化する。BDH は計算時間のかかる処理を回避しつつ、高い確率で真の最近傍点を見つけるように設計されているため、大規模なデータベースに適応可能である。最新の従来手法 [27] と比較すると、半分の計算時間で、同等の最近傍点の検出精度を達成することが実験により示されている。この手法により、高速な特徴量の探索を行い、高速な正解候補の絞り込みを行う。

## 5. 実験

提案手法の大規模なデータベースに対する正解候補の絞り込みの精度と認識にかかる計算時間を評価するため、実験を行った。また、低解像度画像に対する認識精度の評価を行った。

### 5.1 実験条件

提案手法の正解候補の絞り込み精度と計算時間の評価を行うため、1000 万枚顔画像データベースを構築した。1000 万顔画像データベースは、AR Face Database[12] の Set 1 132 枚、CAS-PEAL[16]、FERET[14]、The ORL Database of Faces[19]、Georgia Tech Face Database[20]、Surveillance Cameras Face Database[21]、Indian Face Database[22] のデータベース用顔画像と、web 上からダウンロードした顔画像から構成される。web 上からダウンロードした画像は、重複したものを除いており、全く同じ画像は存在しない。



図 3 web からダウンロードした顔画像の例.

データベース画像はまず、画像から顔検出 [23] を行い、

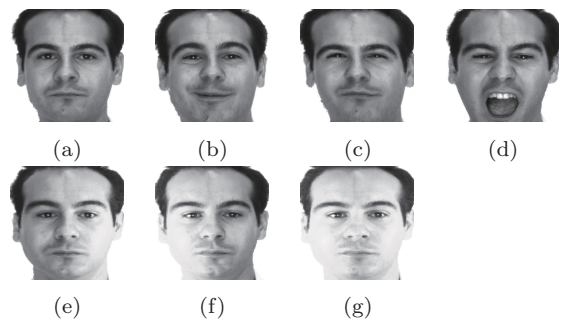


図 4 AR Face Database の画像例 . (a) Set 1: neutral , (b) Set 2: smile , (c) Set 3: anger , (d) Set 4: scream , (e) Set 5: left side light on , (f) Set 6: right side light on , (g) Set 7: all side light on .

検出領域から顔特徴点 14 点を自動で検出する [24] . 得られた特徴点に基づいて顔向きを正規化を行い [25] , 画像の切り出しを行う . 顔の向きを正規化するため、顔の向きを正規化のために検出された特徴点はすべての画像で同じ位置となっている . 画像は 8 bits グレースケールに変換され、解像度は  $512 \times 512$  [pixel] である . データベース中には、顔の切り出しや正規化に失敗した画像も含まれている . 切り出し、正規化を行ったデータベースの顔画像の一部を図 3 に示す .

精度と計算時間の評価のため、AR Face Database の Set 2–Set 7 の画像 792 枚 (132 枚  $\times$  6 sets) をテスト画像として用いた . AR Face Database の Set 1–Set 7 の例を図 4 に示す . Set 1–7 はそれぞれ、neutral , smile , anger , scream , left side light on , right side light on , all side light on に分けられており、照明条件や表情が変化した画像となっている . これらの画像もデータベース画像と同様、顔向き、画像の解像度を正規化した .

実験に使用した計算機は、CPU が AMD Opteron (tm) Processor 6174 (2.2GHz) , メモリは 256GB であり、OS は Debian GNU/Linux wheezy/sid を搭載している . 本稿の実験での処理時間は特徴量の検索にかかった時間のみを測定し、画像の正規化や特徴抽出の時間は含まない . 予備実験により、投票には 30 近傍を用いることとした .

### 5.2 大規模データベースに対する評価

提案手法の大規模なデータベースに対する検索時間と絞り込み精度の評価を行った . 特徴点は、自動的に検出を行う手法を用いて決定した . 1000 万枚顔画像データベースに対する 1000 位累積認識率と平均の処理時間を表 2 に示す . 表 2 の Detector-no-restriction は、マッチングに用いる特徴点の絞り込みを行わなかった場合であり、Detector-coord は、マッチングに用いる特徴点の絞り込みを行った場合である . 表には、BDH のパラメータを変え、最も良い認識精度で最も処理時間の平均が短いものを示している . 表 2 から、Detector-no-restriction, Detector-coord とともに scream



表 2 1000 万顔画像データベースに対する認識率と処理時間

クエリの種類		smile	anger	scream	left side light on	right side light on	all side light on
Detector-no-restriction	認識率 [%]	95.5	99.2	30.3	98.5	97.0	21.2
	処理時間 [ms]	191	180	120	216	195	247
Detector-coord	認識率 [%]	98.5	99.2	39.4	98.5	97.7	28.8
	処理時間 [ms]	193	106	230	185	228	214

と all side light on 以外では, 97[%] 以上の認識率が得られた. left side light on では, Detector-no-restriction の場合で認識率 98.5[%], 処理時間 185 [ms] とテスト画像に用いたデータセットの中では最も高速に認識した. 表 2 の Detector-no-restriction と Detector-coord を比較すると, マッチングに用いる特徴量を絞り込んだ方がよりよい認識率を示した. これは, 検索対象を限定することで, 特徴点の誤対応を減らすことができたからだと考えられる. また, 表 2 から, マッチングに用いる特徴量を絞り込むことで, 計算コストを低減することができ, より高速に認識可能なことがわかる. 一方で, scream, all side light on では認識率が低かった. Natural, scream, all side light on の特徴点検出例を図 5 に示す. 図 5(a), (b) を比較すると, natural で特徴点検出している点は, scream では顔の表情の変化により, 特徴点の周りのテクスチャが変化している. このことにより特徴量のマッチングに失敗したため, 多くの画像が認識に失敗したと考えられる. また, 図 5(a), (c) を比較すると, all side light on の特徴点は natural と比べて少なく, 照明条件の変化により, 特徴点があまり検出されなかったことで認識率が低下したものと考えられる.

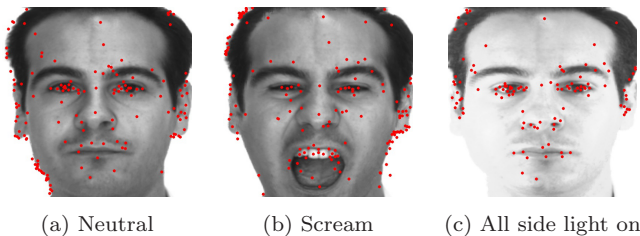


図 5 自動検出により得られた特徴点の例

また, データベースの枚数と処理時間の評価を行った. 特徴点のマッチングには, データベースを絞り込む手法 Detector-coord を用いた. 認識精度は 1000 位累積認識率とし, データベースの画像の枚数を 10 万枚, 20 万枚, 50 万枚, 100 万枚, 200 万枚, 500 万枚, 1000 万枚と変化させ, AR Face Database の left side light on に対して 98.5% 以上の認識率のときの最速の処理時間を測定した. データベースの枚数を変化させたときの処理時間の変化を図 6 に示す. 図 6 から, データベースが 700 万枚までは, データベースの数が大幅に増加しても, 計算時間はあまり増加しなかった. しかし, データベースの枚数が 1000 万枚になったとき, 計算時間が大幅に増加した. これは, データベースの数が増えたことにより特徴量の弁別能力が低くなった

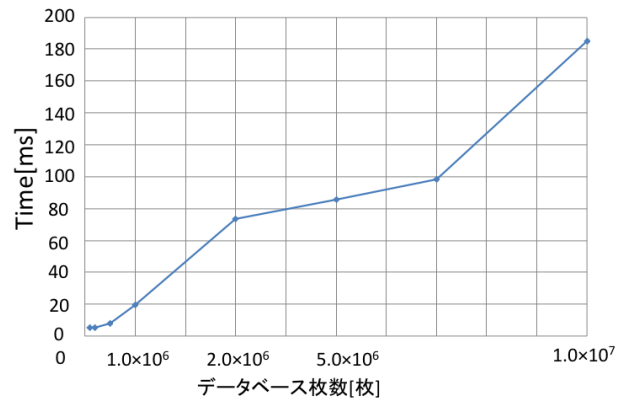


図 6 データベースの枚数と処理時間



図 7 低解像度画像の例. (a) Set 2: smile, (b) Set 3: anger, (c) Set 4: scream, (d) Set 5: left side light on, (e) Set 6: right side light on, (f) Set 7: all side light on.

ためだと考えられる. 同じ認識精度を保つためには, 700 万枚のデータベースと比較して広い範囲を探索する必要があり, 検索時間が大幅に増加したものと考えられる.

### 5.3 低解像度画像での評価

提案手法の低解像度画像に対する認識精度の評価を行った. 低解像度画像には, AR Face Database の set 2-set 7 をバイキュービック法により  $25 \times 25$  [pixel] に縮小した後,  $512 \times 512$  [pixel] に拡大したものをを用いた. クエリとして用いた低解像度画像の例を図 7 に示す. データベースは, AR Face Database set 1 と, 1000 万顔画像の一部を用いて 10 万枚としたものをを用いた. 特徴点は自動検出と Dense sampling を用いて決定し, 特徴抽出を行った. Dense sampling では, 特徴点のサンプル間隔を 40, 80, 120,

表 3 低解像度画像の 10 万顔画像データベースに対する認識率と処理時間 .

クエリの種類		smile	anger	scream	left side light on	right side light on	all side light on
Detector	認識率 [%]	97.0	98.5	41.0	84.1	68.9	25.8
	処理時間 [ms]	65	105	66	141	148	90
Dense sampling	認識率 [%]	97.0	97.7	26.5	98.5	97.0	77.2
	処理時間 [ms]	82	53	82	135	92	127

160, 200 [pixel] とし, 880 点の特徴点を用いた. 特徴抽出する際, orientation は 0 とし, PCA-SIFT 特徴量を抽出する領域をを 40 [pixel] から 200 [pixel] に変化させて特徴抽出を行った. 低解像度画像をテスト画像に用いた際の 1000 位累積認識率と処理時間を表 3 に示す. 表 3 より, left side light on, right side light on, all side light on no といった照明変動がある画像では, Dense sampling を用いて特徴点を決定した方が認識率が大幅に改善した. 低解像度画像に対して自動で特徴点検出した場合, 得られた特徴点は平均で 25 点であった. これは, Dense sampling の 880 点と比較して大幅に少なく, 認識精度に大きく影響していたと考えられる. 一方, Dense sampling では, 常に一定の特徴量が得られるため, 画像の解像度が低い場合でも, 認識が可能であったと考えられる.

#### 5.4 実環境で撮影されたデータでの評価

実環境で撮影された低解像度画像に対する提案手法の性能評価を行った. 評価を行うため, 高解像度画像のデータベース画像と低解像度画像のテスト画像を撮影した. 撮影した画像には 15 人分の画像が含まれている. データベース画像は Logicool 社製 C270 で撮影を行い, もとの画像の解像度は 640×480 [pixel] である. テスト画像は Panasonic 社製 BB-HCM715 で撮影を行い, もとの顔画像の解像度は 1280×960 [pixel] である. これらの画像は 5.1 節のデータベースと同様, 顔検出, 顔向き正規化を行い, 512×512 [pixel] に正規化して実験に用いた. 用いた画像のもとの画像と正規化後の画像を図 8 に示す. データベース画像には, 撮影したデータ 15 枚に加え, 1000 万顔画像の一部を追加して, 1 万枚の顔画像データベースを作成した. 特徴点は, Dense sampling を用いて決定し, 880 点から特徴抽出を行った. 1 万枚顔画像データベースに対して, 1000 位累積認識率で評価を行った結果, 15 人中 14 人が認識でき, 平均の計算時間は 109 [ms] であった. このように, 実環境で撮影された画像に対しても, 提案手法は高精度に絞り込み可能であることがいえた.

## 6. まとめ

本稿では, 犯罪捜査支援を目的とし, 局所特徴量の 1 つである PCA-SIFT 特徴量と特徴量の最近傍探索に基づいた投票処理を用いた高速な顔画像候補の絞り込みを行った. 1000 万枚の顔画像データベースを用いた実験の結果, AR

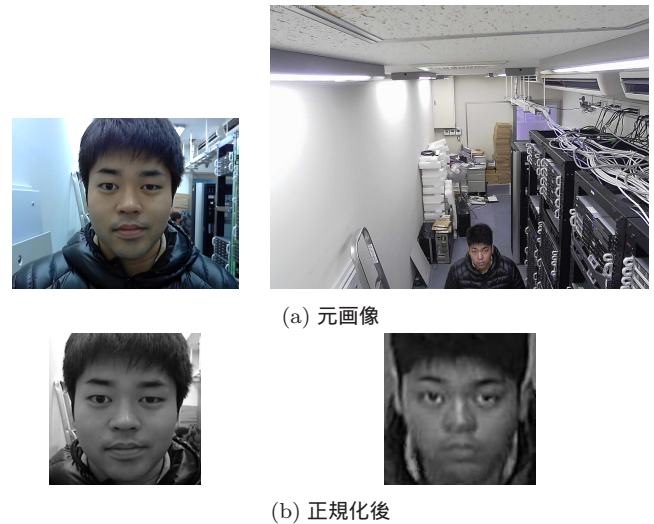


図 8 実環境で撮影されたデータの一例. 左がデータベース用の画像, 右がクエリ用の画像.

Face Database の left side light on に対して 1000 位累積認識率が 98.5 [%] で 185 [ms] で認識が可能であった. マッチングに関して, 探索の際に座標の近い特徴点のみを検索の対象とすることでより計算コストが削減され, かつ高い認識率を達成した. 照明変動がある低解像な画像を検索に用いた場合, 特徴点検出をする手法では, 検出される特徴点の数が少なく認識精度が低下するが, Dense sampling を用いると, 一定の特徴量を抽出することができ, 認識精度が改善した. 今後の課題として, 更なる認識率の向上のため, 照明条件の変動や表情変化による顔画像の見えの変化に対応することが挙げられる.

謝辞 本研究は科学技術戦略推進費「安全・安心な社会のための犯罪・テロ対策技術等を実用化するプログラム」の補助を受けた. ここに記して感謝する.

#### 参考文献

- [1] W. Zhang, H. Deng, T.G. Dietterich, and E.N. Mortensen, "A hierarchical object recognition system based on multi-scale principal curvature regions," Proceedings of 18th International Conference on Pattern Recognition, vol.1, pp.778-782, 2006.
- [2] S. Zickler and A. Efros, "Detection of multiple deformable objects using PCA-SIFT," Proceedings of the 22nd National Conference on Artificial Intelligence, vol.2, pp.1127-1132, 2007.
- [3] 野口和人, 黄瀬浩一, 岩村雅一, "近似最近傍探索の近似と多段階化による物体の高速認識," 画像の認識・理解シンポジウム (MIRU2007) 論文集, pp.111-118, 2007.
- [4] T. Sato, M. Iwamura, and K. Kise "Fast and memory effi-

- cient approximate nearest neighbor search with distance estimation based on space indexing,” IEICE Technical Report, PRMU2012-142, 2013.
- [5] T.-K. Kim, B. Stenger, J. Kittler, and R. Cipolla, “Incremental linear discriminant analysis using sufficient spanning sets and its applications,” *International Journal of Computer Vision*, vol.91, pp.216–232, Sept. 2010.
- [6] J. guoWang, Y. shengLin, W. kouYang, and J. yuYang, “Kernel maximum scatter difference based feature extraction and its application to face recognition,” *Pattern Recognition Letters*, vol.29, pp.18320–1835, 2008.
- [7] K. Hotta, “Robust face recognition under partial occlusion based on support vector machine with local gaussian summation kernel,” *Image and Vision Computing*, vol.26, no.11, pp.1490–1498, Nov. 2008.
- [8] P.J. Phillips, “Support vector machines applied to face recognition,” *Advances in Neural Information Processing Systems*, vol.11, pp.803–809, 1998.
- [9] M. Nishiyama, O. Yamaguchi, and K. Fukui, “Face recognition with the multiple constrained mutual subspace method,” *Proceedings of 5th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA '05)*, pp.71–80, 2005.
- [10] “AT & T The Database of Faces,” <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
- [11] “The University of Oulu Physics-Based Face Database,” <http://www.cse.oulu.fi/CMV/Downloads/Pbfd>
- [12] A. Martínez and R. Benavente, “The AR face database,” *Technical Report 24*, Computer Vision Center, Bellaterra, 1998.
- [13] “The Yale Face Database B,” <http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html>
- [14] J.P. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, “The feret evaluation methodology for face-recognition algorithms,” *IEEE Trans. on PAMI*, vol.22, no.10, pp.1090–1104, 2000.
- [15] “PIE Database,” [http://www.ri.cmu.edu/research\\_project\\_detail.html?project\\_id=418&menu\\_id=261](http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261)
- [16] W. Gao, B. Cao, S. Shan, D. Zhou, X. Zhang, D. Zhao, and S.S.E. Al, “The CAS-PEAL large-scale Chinese face database and evaluation protocols,” *Technical report*, Joint Research & Development Laboratory, CAS, 2004.
- [17] “The CMU Multi-PIE Face Database,” <http://www.multipie.org/>
- [18] Y. Ke and R. Sukthankar, “PCA-SIFT: A more distinctive representation for local image descriptors,” *CVPR'04*, vol.2, pp.506–513, 2004.
- [19] “ORL Face Database,” <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
- [20] “Georgea Tech Face Database,” [http://www.anefian.com/research/face\\_reco.htm](http://www.anefian.com/research/face_reco.htm)
- [21] M. Grgic, K. Delac, and S. Grgic, “Scface - surveillance cameras face database,” *Multimedia Tools and Applications*, pp.1–17, 2009.
- [22] V. Jain and A. Mukherjee, “The indian face database,” 2002. [http://vis-www.cs.umass.edu/\\$sim\\$vidit/\\{I}ndian{F}ace{D}atabase/](http://vis-www.cs.umass.edu/$sim$vidit/\\{I}ndian{F}ace{D}atabase/)
- [23] T. Mita, T. Kaneko, B. Stenger, and O. Hori, “Discriminative feature co-occurrence selection for object detection,” *IEEE Trans. on PAMI*, vol.30, pp.1257–1269, 2008.
- [24] M. Yuasa, T. Kozakaya, and O. Yamaguchi, “An efficient 3D geometrical consistency criterion for detection of a set of facial feature points,” *IEICE - Trans. Inf. Syst.*, vol.E91-D, pp.1871–1877, 2008.
- [25] T. Kozakaya and O. Yamaguchi, “Face recognition by projection-based 3d normalization and shading subspace orthogonalization,” *FGR '06*, pp.163–168, 2006.
- [26] W. R. Schwartz and H. Guo and L. S. Davis, “A robust and scalable approach to face identification,” *Proceedings of European Conference on Computer Vision*, pp.476–489, 2010.
- [27] A. Babenko and V. Lempitsky, “The inverted multi-index,” *Proceedings of 2012 IEEE International Conference on Pattern Recognition*, pp.3069–3076, 2012.
- [28] J. Kim, J. Choi, J. Yi, and M. Turk “Effective representation using ICA for face recognition robust to local distortion and partial occlusion,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol.27, no.12, pp.1977–1981, 2005.