

2. 指紋認識データベースの調査

2.1. データベースの調査

各国の指紋認識用データベースを検索し、表1に示す研究機関、一般企業のホームページから必要な情報を抽出した。

表1 指紋認識用データベース

No.	指紋認識用データベース	国名	大学、研究機関または企業名
(1)	NIST [1]	アメリカ	米国商務省標準化技術研究所
(2)	FVC2000 [2]	アメリカ	指紋照合コンテスト
(3)	East Shore Technologies [3]	アメリカ	一般企業

2.2. NIST

NIST (National Institute of Standards and Technology)は、米国の技術標準化をつかさどる官庁で、米国商務省標準化技術研究所と呼ばれている。生体情報の標準化も実施している。

NIST セキュリティポリシーとして、NIST ホームページの訪問者の個人情報を訪問者の許可なく収集することはしない。但し、訪問者のドメイン名、IP アドレス、ブラウザのタイプ、アクセス日時、ページ、リンク元の Web サイトは自動的に収集し、蓄積している。

2.3. NIST 指紋データベース

NIST では、8 つの指紋認識用のデータベースを保持している。

(1) NIST Special Database 4 (NIST 8-Bit Gray Scale Images of Fingerprint Image Groups(FIGS))

本データベースは、2000 個の 8 ビットグレースケール指紋画像データのペアを持っている。それぞれの指紋画像は、下側に 32 行の白空白を持つ 512 × 512 画素で、下記の 5 つのタイプに分けられている。

A=アーチ、L=左ループ、R=右ループ、T=テント状アーチ (Tented Arch)、W=渦巻き
各タイプは 400 個の指紋のペアで構成されている。

画像は JPEG ロスなし圧縮で、約 636MB で圧縮され格納されている。元データは 1.1GB である。

各指紋のペアは、同じ指紋から取った完全に異なる回転の 2 つの指紋である。

価格 : 90 \$

画像数 : 2000 個

解像度 : 19.7 画素/ミリ

画素数 : 512 × 512 画素

色階調 : 8 ビットグレースケール

フォーマット : JPEG

DB サイズ : 636MB

(2) NIST Special Database 9 (NIST 8-Bit Gray Scale Images of Mated Fingerprint Card Pairs(MFCP) Volumes 1-5)

本データベースは、現在5ボリュームがリリースされている。各ボリュームは、3枚のディスクで構成され、各ディスクは8ビットグレースケール指紋画像で、90枚のカード指紋から構成されている。各ディスクには900指紋画像がある。

区切られた各画像は、832×768画素で、FBIから与えられた National Crime Information Center (NCIC:国立犯罪情報センタ)のクラスでクラス分けされている。

画像はJPEGロスなし圧縮で、約630 - 660MBで圧縮格納されている。元データは1.0 - 1.2GBである。

価格 : 90 \$

画像数 : 13500 個

解像度 : 19.7 画素/ミリ

画素数 : 832×768 画素

色階調 : 8ビットグレースケール

フォーマット : JPEG

DB サイズ : 630 - 660MB

(3) NIST Special Database 10 (NIST Supplemental Fingerprint Card Data(SFCD))

本データベースは補足指紋カードデータ(SFCD)に関するデータベースである。本データベースは、NIST Special Database9の中で、自然発生頻度が低く、変化しやすい指紋のクラスの指紋パターンをより多く含んだサンプルを提供する。

NIST Special Database10は3枚のCDで構成されている。最初のCDは2160画像、あとの2つはそれぞれ1680画像を含んでいる。それぞれの区分された画像は832×768画素で、FBIによって与えられた国立犯罪情報センター(NCIC)のクラスを使用して分類されている。

画像は、JPEGロスなし圧縮で圧縮されている。

最初のCDがおおよそ690MBの大きさと2番目と3番目のCDがおおよそ590MBの大きさである。

NIST Special Database10には、以下の特徴がある。

- アーチ、テント状アーチ、低い隆線数のループや渦巻きだけを集めた補足データの552のペアなしカード
- FBIによって与えられたNCIC分類

価格 : 90 \$

画像数 : 5520 個

解像度 : 19.7 画素/ミリ

画素数 : 832×768 画素

フォーマット : JPEG

DB サイズ : 590 - 690MB

(4) NIST Special Database 14 (NIST Mated Fingerprint Card Pairs 2 (MFCP2))

本データベースは3枚のCDから構成されており、それぞれのCDは9,000組の画像を持っている。大きさは圧縮されておよそ670MB、元データは11.5GBである。

それぞれの区分された画像は、832×768画素で、FBIによって与えられたNCICのクラスを使用して分類されている。

NIST Special Database14には、以下の特徴がある。:

- 2万7000組の区分された8ビットのグレースケール指紋画像
- WSQ(wavelet scalar quantization)で圧縮された画像
- FBIによって与えられたNCIC分類

価格 : 90 \$

画像数 : 27000 個

解像度 : 19.7 画素/ミリ

画素数 : 832 × 768 画素

色階調 : 8 ビットのグレースケール

フォーマット : WSQ

DB サイズ : 670MB

(5) NIST Special Database 24 (NIST Digital Video of Live-Scan Fingerprint Database)

本データベースは、指紋データの圧縮された動画(MPEG2)を含んでいる。このデータベースは、指紋照合システムの開発及びテストで利用するために配布される。

CD2枚で構成され、1つは10秒間(300フレーム、720×480画素)のプラスチックで歪んだ指紋データ、もう1つは、10秒間の色々な回転角度の指紋データである。データベースは、10人(男5人、女5人)のすべての手を含んでおり、CDにはそれぞれ合計100個のMPEG2ファイルがある。それぞれのMPEG2ファイルは6.25MBに圧縮されている。

価格 : 90 \$

画像数 : 100 × 2

画素数 : 720 × 480 画素

被験者数 : 10 人

フォーマット : MPEG2

DB サイズ : 625MB × 2

(6) NIST Special Database 27 (Fingerprint Minutiae from Latent and Matching Tenprint Images)

本データベースは事件現場に残されている指紋とそれらの照合用回転指紋画像を含んでいる。全部で、258の残留ケースがある。各ケースは残留画像、照合用の10指の指紋画像及び残留画像検査官の専門チームによって有効になった4セットのマニューシャ(特徴点)を含んでいる。1番目の

セットは残留指紋画像のすべてのマニューシャを含んでいる。2番目のセットは10指の指紋のペアの片方のすべてのマニューシャを含んでいる。他の2セットは残留指紋と10指の指紋のペアの片方と共通するマニューシャを含んでいる。照合残留指紋と共通な5460個のマニューシャを含め、10指の指紋の集合全部で27426個のマニューシャが記録されている。それぞれの画像は800×768画素の大きさで、19.69画素/ミリ(500PPI)で撮影され、256階調グレー、非圧縮形式で格納されている。

価格 : 90 \$
画像数 : 27426 個
解像度 : 19.69 画素/ミリ(500PPI)
画素数 : 800×768 画素
色階調 : 256 グレー階調
DB サイズ : 317MB

(7) NIST Special Database 29 (Plain and Rolled Images from Paired Fingerprint Cards)

本データベースは216個の回転及び平面の10指の指紋カードの組で構成されている。画像は19.7画素/ミリ。カードは10個の回転指紋と4グループの平面指紋(左指、右指、左親指、右親指)に区分けされている。全画像はWSQ圧縮されている。

2160個の回転指紋画像のうち、757個はNIST4で使用している指紋と同じである。

価格 : 90 \$
画像数 : 2160 個
解像度 : 19.7 画素/ミリ
フォーマット : WSQ

(8) NIST Special Database 30 (Dual Resolution Images from Paired Fingerprint Cards)

本データベースは指紋圧縮及び指紋照合システムの開発用、テスト用のデータベースである。19.7画素/ミリ(500DPI)と39.4画素/ミリ(1000DPI)の解像度で取得された指紋画像用のアルゴリズムなら本データベースの利用が可能である。データは19.7及び39.4画素/ミリで取得された回転及び平面指紋画像を持っている36組の10指指紋の印刷されたカードで構成されている。カードは10個の回転指紋と4グループの平面指紋(左指、右指、左親指、右親指)に区分けされている。全画像はロスなしJPEG圧縮で圧縮されている。

360個の回転指紋画像のうち、316個はNIST4で使用されている指紋と同じものである。

データベースは4枚のCDで構成され、1枚目は解像度19.7画素/ミリの画像データで他の3枚は解像度39.4画素/ミリの画像データである。

価格 : 90 \$
画像数 : 360 個
解像度 : 19.7 及び 39.4 画素/ミリ
フォーマット : JPEG

2.4. FVC2000 指紋データベース

本データベースは指紋照合アルゴリズムのコンテストで利用したものである。

指紋データベースには、下記4つの異なったデータベースがある。

- DB1: 安価な光学式センサ (KeyTronic 製 SecureDesktopScanner) で収集したデータベース
- DB2: 安価な静電容量式センサ (STMicroelectronics 製 TouchChip) で収集したデータベース
- DB3: 光学センサ (IdenticatorTechnology 製 DF-90) で収集したデータベース
- DB4: 合成指紋データベース

これらのデータベースには、110 指が収集されており 1 つの指に対して 8 画像収集しているので合計 880 個の指紋がある。101 番目から 110 番目の指は、コンテスト参加者のアルゴリズム提供前に、パラメータ調整を許す目的で、関係者が利用出来るように作成されている。そのため、ベンチマーク用のデータベースは 1 番目から 100 番目の指で構成される。

(1) DB1 : 安価な光学式センサで収集したデータベース

本データベースの特徴は、下記の通りである。

- ・ 画像サイズ 300 × 300
- ・ 解像度は 500dpi
- ・ 指紋は主に 20 歳から 30 歳の学生で半数は男性である。
- ・ ボランティアの方の左右の人差し指、中指の最大 4 指の指紋が収集されている。
- ・ 画像は訓練していない人から 2 回の異なったセッションで収集し、最低の獲得品質を保証するための努力はしていない。
- ・ 同じ個人の全ての画像は、1 回ごとに指の順番を代えて収集している。例えば、左人差し指の 1 回目、右人差し指の 1 回目、左中指の 1 回目、右中指の 1 回目、左人差し指の 2 回目のように収集している。
- ・ 指紋の正しい中心位置あわせのチェックに注意を払っていないので、指紋のコアとデルタの存在は保証しない。
- ・ センサの表面は自動的にクリーンされない。
- ・ 収集した指紋は、最大回転角度が約 - 15 度から 15 度の範囲になることを保障するため手動で解析されている。更に、同一指の指紋のペアは、指紋領域のどこかはオーバーラップしている。

(2) DB2 : 安価な静電容量式センサで収集したデータベース

本データベースの特徴は、下記の通りである。

- ・ 画像サイズ 256 × 364
- ・ 解像度は 500dpi
- ・ 指紋は主に 20 歳から 30 歳の学生で半数は男性である。
- ・ ボランティアの方の左右の人差し指、中指の最大 4 指の指紋が収集されている。
- ・ 画像は訓練していない人から 2 回の異なったセッションで収集し、最低の獲得品質を保証する

ための努力はしていない。

- ・ 同じ個人の全ての画像は、1 回ごとに指の順番を代えて収集している。例えば、左人差し指の1 回目、右人差し指の1 回目、左中指の1 回目、右中指の1 回目、左人差し指の2 回目のように収集している。
- ・ 指紋の正しい中心位置あわせのチェックに注意を払っていないので、指紋のコアとデルタの存在は保証しない。
- ・ センサの表面は自動的にクリーンされない。
- ・ 収集した指紋は、最大回転角度が約 - 15 度から 15 度の範囲になることを保障するため手動で解析されている。更に、同一指の指紋のペアは、指紋領域のどこかはオーバーラップしている。

(3) DB3 : 光学センサで収集したデータベース

本データベースの特徴は、下記の通りである。

- ・ 画像サイズ 448 × 478
- ・ 解像度は 500dpi
- ・ 指紋は 5 歳から 73 歳のボランティアから収集し 55%は男性である。
- ・ ボランティアの 3 分の1は 55 歳以上、3 分の1は 18 歳以下、6 分の1は 7 歳以下である。
- ・ 6 指(左右の親指、人差し指、中指)のうちの 2 画像は各セッションで各ボランティアの方から連続して収集されている。
- ・ 各ボランティアは4セッション実施するが、1 日ではたった 2 セッションだけである。
- ・ 最初と最後のセッションの間隔は、ボランティアに依存するが、少なくとも3日、長くて3ヶ月である。
- ・ センサの表面は画像収集ごとに計画的にクリーンされる。
- ・ 各ボランティアの1つのセッションでは、指はアルコールで磨いて綺麗にして、乾燥させる。
- ・ 各画像でコアの一部は明白になっているが、1つのセッションの間に連続して取られた画像の間で完全なオーバーラップは避けるよう注意して収集された。
- ・ 収集した指紋は、最大回転角度が約 - 15 度から 15 度の範囲になることを保障するため手動で解析されている。更に、同一指の指紋のペアは、指紋領域のどこかはオーバーラップしている。

(4) DB4 : 合成指紋データベース

本データベースの特徴は、下記の通りである。

- ・ 画像サイズ 240 × 320
- ・ 解像度は 約 500dpi
- ・ 合成ジェネレータのパラメータは、獲得領域が小さい安価なセンサと同じとなるように調整されている。
- ・ 最大の回転、置き換え、および皮膚のひずみは、他のデータベース(DB1、DB2、および DB3)における場合と同じになるように調整される。

2.5. East Shore Technologies 指紋データベース

East Shore Technologies は米国の一般企業である。

East Shore 社では、それぞれ 10 万以上の指紋を持つ 2 つの指紋データベースが利用可能である。1 つは完全に平坦に捺印された指紋、もう 1 つは回転された指紋のデータベースである。平坦な指紋データベースの少なくとも 80% の指は、回転指紋データベースの回転指紋と同じ指である。

全ての指紋画像は、500dpi、256 グレースケールの画像である。圧縮モードで記録された指紋画像は存在しない。

指紋画像は 10 個の指紋が印刷された指紋カードから収集されている。指紋カードは、非常に品質の悪い指紋から優良な指紋まで色々な品質の指紋がミックスされた見本となることを保障するため、ランダムに選択されている。

全ての 10 個の回転指紋はそれぞれ 1 万枚の指紋カードから作成されている。平坦な指紋は、回転指紋が記録されている同じ指紋カードから抜粋されている。平坦な指紋は事前に関連できる、多少は識別できそうな隆線の構造や沢山の識別可能なマニューシャを含んだ平坦な画像だけが記録されている。

10 万個の画像の一部には、300 個の照合画像があり、元指紋画像と照合指紋画像は異なった期間に印刷されている。300 個の照合指紋画像はテスト目的のためなら、特別に識別できる。

平坦指紋は 22 枚の CD、回転指紋は 36 枚の CD に記録されている。

2.6. 被験者及び指紋収集の要件に関する考察

上述のように、指紋認識用のデータベースとして、NIST で 8 個、FVC2000 で 4 個、East Shore 社で 2 個の合計 14 個を調査したが、ここでは指紋データベースの被験者及び指紋収集の要件について考察する。

(1) 被験者の要件

収集した指紋の被験者の要件としては、下記が考えられる。

- ・ 性別
- ・ 年齢
- ・ 職種
- ・ 習熟度

性別、年齢、職種などの構成は、統計的な偏りを少なくするため一般的な人口構成と同様であることが望ましいとされているが、今回調査した 14 個のデータベースでは、性別、年齢、職種などを詳述しているものはなかった。

性別に関して記述があったデータベースは 4 個、年齢に関しては 3 個、職種に関して記述されているものはなかった。また、人口構成を意識して収集しているということはどのデータベースにも記述がなかった。

また、収集に利用した装置に対する指紋入力方法の指導を受け、入力操作に習熟している方が、利用に適した指紋画像を収集できると考えられるが、習熟度に関して記述のあったデータベースは

2 個であった。しかしながら、2 つともに“訓練していない人から収集した”という記述があるのみであり、“操作を 1 日かけて指導”, “マニュアルを配布して事前に学習を促す”などの具体的な習熟方法を記述しているものではなかった。

以上の点から考察すると、今回調査した指紋データベースでは、収集時における被験者の要件に関する意識が十分に高いとは言えないと考えられる。

(2) 指紋収集の要件

収集した指紋の要件としては、下記が考えられる。

- ・ 被験者数
- ・ 被験者一人に対して利用した指数
- ・ 利用した指の種類
- ・ 一指から収集した画像数
- ・ 指紋の型
- ・ 被験者の肌質
- ・ 指紋の回転、ゆがみ
- ・ 収集時の気温、湿度、季節
- ・ 収集期間、収集時間
- ・ 型式、メーカーなどの装置情報
- ・ 指紋データファイルと被験者の関係

指紋収集の要件としては、上記のような多数の項目が考えられるが、今回調査した 14 個のデータベースでは、上記項目を全て記述しているものは無かった。

特に被験者の肌質、収集時の気温、湿度、季節に関する記述は全く無かった。

指紋の回転、ゆがみに関しては 7 個、指の種類に関しては 5 個、指紋の型に関しては 4 個のデータベースで記述があり、指紋の回転、ゆがみに対する意識が他より高いことがわかった。

また、指紋の経時変化を考慮して、時間変化、季節変化、経年変化を意識して収集することが望ましいが、指紋の収集期間や収集手順に関して記述があったのは 3 個であった。

またファイル名から被験者が特定できるような指紋データファイルと被験者の指紋との関係が記述されていたものはなかった。

今回調査した殆どのデータベースの中で、指紋データのサイズ(画素数)、解像度、色階調に関しては記述があり、また指紋の回転、ゆがみ、指紋の型などデータそのものに関する記述が多かった。これらの結果から、指紋データそのものに対する意識の方が収集する環境(被験者構成、年齢、性別、収集期間、肌質、季節)より高いことが推測される。

また、指紋の照合精度は、同一の指紋入力装置を用いた場合でも、指紋の型や肌の質および指紋データの経時変化によって左右されると言われているが、今回調査したデータベースでは被験者の肌質、収集時の気温、湿度、季節に関する記述がないことから、指紋の収集要件に対しても意識が高いとは言えないと考えられる。

2.7. 個人情報管理に関する考察

生体認証データベースの取り扱いに関するガイドラインが、「指紋認証システムの精度評価方法 (JIS - TR)」[12]に記述されている。それによれば、生体認証データは、生体認証データ提供者に対する人格尊重の理念の下に慎重に取り扱われるべきものであり、生体認証データベースを取り扱うすべての者は、以下に掲げる条項を遵守する必要がある。

下記に概要を抜粋する。

- a) 利用目的による制限
“ 生体認証データは、採集時あるいはそれに先立ってその利用目的を明確にしなければならない。また、明確化された利用目的以外のために利用してはならない。”
- b) 適正な方法による取得
“ 生体認証データは、生体認証データ提供者の同意を得た上で、適法かつ適正な方法によって取得されること。”
- c) 内容の正確性の確保
“ 生体認証データは、その利用目的の達成に必要な範囲内において正確 (accurate) かつ完全 (integrity) であり、最新なものに保たれること。”
- d) 安全保護措置の実施
“ 生体認証データは、適切な安全保護措置を講じた上で取り扱われること。”
- e) 透明性の確保・提供者の権利保証
“ 生体認証データベースに関わる開発、運用、および方針については、一般に公開されなければならない。また、生体認証データの提供者に対する権利を保証しなければならない。”
- f) 生体認証データベース管理者の責任
“ 生体認証データベース管理者は、上記 a) から e) の各項を実施するための措置に従うこと。”

今回調査したデータベースでは、データベースの利用目的に関して記述があったのは3個、適正な方法による取得と推測できるもの(被験者がボランティアであることなど)を明記したものが3個あった。それ以外では上記条項に対して明確に記述したものは見当たらなかった。

従って、今回調査したデータベースに関しては、被験者のプライバシーを保護するという個人情報の管理に関しては記述が少なく、個人情報保護の意識が高くはないと考えられる。

2.8. 指紋データベースに関する考察

指紋データベースは、その指紋データをどのような被験者からどのようにして収集したのか、また収集した指紋データの管理をどのように実施しているかを、明確に規定して、それに従って運用する必要があると考えられるが、今回調査した範囲ではそれらを十分に満たしているものは無かったと考えられる。今後、日本において標準的な指紋データベースを構築する場合には、被験者の要件、指紋収集の要件、個人情報保護の条項に従いデータを収集し、管理することが望まれる。