論題 Title	第 1 部　モダリティ別の技術動向
他言語論題 Title in other language	Part1 Trends in Research and Technology by Modality
著者／所属 Author(s)	大阪大学
書名 Title of Book	生体認証技術の動向と活用：科学技術に関する調査プロジェクト（Current Trends in Biometrics）
シリーズ Series	調査資料 2018-6　（Research Materials 2018-6）
編集 Editor	国立国会図書館 調査及び立法考査局
発行 Publisher	国立国会図書館
刊行日 Issue Date	2019-3-29
ページ Pages	1-35
ISBN	978-4-87582-839-6
本文の言語 Language	日本語（Japanese）
キーワード keywords	生体認証、バイオメトリクス、モダリティ
摘要 Abstract	生体認証のモダリティは、身体的な特性・特徴を利用したものと、行動的な特性・特徴を利用したものに分けられる。前者として指紋、DNA、虹彩、静脈、顔、後者として歩容、発話、署名等を紹介する。

第 1 部
モダリティ別の技術動向

第１部　モダリティ別の技術動向

【要　旨】

生体認証（「バイオメトリクス」と呼ばれる。）は、人の身体的な特性・特徴や行動的な特性・特徴に基づいて、その人物を自動的に確認・識別することであり、その技術を生体認証技術という。生体認証には様々なモダリティ（様式、分類）が提案・研究・活用されているが、それらは身体的な特性・特徴を利用したものと、行動的な特性・特徴を利用したものに分かれる。第１部では、前者について、指紋、DNA、虹彩、静脈、顔の各種の認証手法を取り上げ、それらに加えて後者として歩容、発話、署名などを用いた方法を紹介する。

生体認証では、人物から生体情報を取得し、事前に登録してある情報と比較して、確認や識別を行うが、取得される生体情報は、同じ人物から取得されたものであっても、様々な要因により同一とはならないために、スコアによって２つの情報の類似度を評価する。生体認証には２種類の認証方法があり、文字どおり２つの生体情報を比較し、その２つの生体情報が同じ人物あるいは同じ部位から取得されたものかどうかを確認する「一対一認証」（確認）と呼ばれるものと、入力された生体情報が、登録されている多数の生体情報のどれに類似しているのか見つけてくる「一対多認証」（識別）と呼ばれるものである。一対一認証のスコアは、あらかじめ決めたしきい値によって、「同じ」か「異なる」かが判断される。同一人物・同一部位間のスコアであっても、異なると誤って判定される確率を「本人誤拒否率」（False Rejection Rate: FRR）、同一人物・同一部位間ではないにも関わらず、同じと誤って判定される確率を「他人誤受入率」（False Acceptance Rate: FAR）と呼ぶ。FRR と FAR はしきい値によって変化し、かつ両者は、片方を下げるともう片方が上がるというトレードオフの関係にある。

指紋認証は、センサによって指先の紋様が取得され、認証が行われる。人間の DNA は、ほとんどの部分が共通しているが、共通でない部分、特に STR と呼ばれる反復配列が認証に利用されている。虹彩認証は、眼球の色のついた部分から黒い瞳孔部分を取り除いた部分の領域、すなわち黒目の周りの模様を用いて認証が行われる。静脈認証は、データ取得には近赤外光が用いられ、静脈パターンが取得され認証される。顔認証は、コンピュータビジョンやパターン認識の技術を用いて、顔検出・顔の位置合わせ（正規化）・特徴抽出・識別処理を経て、確認・識別に至る。大量の学習データと深層学習の導入により，顔認識の精度は飛躍的に向上している。歩容認証も顔認証と同様、コンピュータビジョンやパターン認識の技術を用いて、入力画像の取得・前処理・特徴抽出を経て、確認・識別に至る。近年の深層学習による手法により、様々な変動要因に対する頑健性が格段に向上している。行動的な認証でも、例えば、声紋認証の場合は、マイクなどによって取得された音声信号から、個人性が現れる特徴を抽出し、その特徴を比較することでスコア化し、比較する。

それぞれのモダリティには、正確性の問題だけでなく、経時変化やなりすましといった、モダリティごとに異なる解決すべき課題がある。

Ⅰ　生体認証（バイオメトリクス）とは

１　はじめに

生体認証（「バイオメトリクス」（Biometrics）[1]と呼ばれる。）は、人の身体的な特性・特徴[2]や行動的な特性・特徴に基づいて、その人物を自動的に識別・確認することであり、その技術を生体認証技術という。人の身体には、人それぞれで異なる情報を持つ部位が存在している。例えば、指先の紋様や、虹彩のしわ模様は、人によって異なることが知られている[3]。また人の行動の中には、例えば歩き方や筆記方法など個人の癖や個人性が現れる動作がある。前者のような情報を特性・特徴ととらえて個人認証に利用するものが身体的な特性・特徴を用いた認証技術であり、後者の個人性が現れる動作を特性・特徴ととらえて認証するものが行動的な特性・特徴を用いた認証技術である。また、認証時に利用される、人物から直接取得される情報は生体情報と呼ばれる。

生体認証を理解するには、従来の手法と比較をするとよい。１つの例を考える。例えば、事前に許可された限られた人物しか入室してはいけない部屋があり、許可された人物はいつでも部屋に入ることができ、許可されていない人物は入ることができないようにしたい。この時、その部屋に入ろうとする人が、許可されている人物かどうかを判断する必要がある。１つの方法は、事前に許可された人物には入室許可証などを配付しておき、入室の際に、その許可証を提示してもらい確認することで、その人物が許可された人物であるかどうかを判断する方法である。この方法は、事前に配付した物を介して、その人物を確認する方法であるため「所有物による方法」と呼ばれる。他の方法としては、事前に許可されている人物間で合言葉を共有し、入ろうとする際に、その合言葉を言えるのか確認することで、その人物が許可された人物であるかどうかを判断する方法である。この方法は、その人物が知っている情報を介してその人物を確認する方法であるため「知識による方法」と呼ばれる。所有物による方法や知識による方法を用いることで、事前に許可された人物はその部屋に入ることができ、逆に事前に許可されていない人物はその部屋に入ることができなくなったように思えるが、実はこれは正しくない。例えば、入室を許可された人物が許可証を失くしてしまった場合や、合言葉を忘れてしまった場合には、その許可された人物は部屋に入ることができない。逆に、入室を許可されない人物が許可証をなんらかの方法で入手しその現場に現れ、その許可証を提示したとすると、この許可証は本物であるため、本来入室できない人物であるにも関わらず入室できてしまう。また、合言葉を推測し、その推測した合言葉が正しかった場合には、本来入室できない人物であって

＊　本稿におけるインターネット情報の最終アクセス日は、2019 年 1 月 31 日である。

(1)　「バイオメトリクス」（Biometrics）という単語は「Biology」（生物学）と「Metrics」（測定）を組み合わせて作られた単語である。一般的に、生体認証に用いられる情報を指す場合と、生体認証や生体認証技術のことを指す場合がある。本稿で扱うのは後者である。両者を扱う国際会議については「バイオメトリクス」と記す。

(2)　アニル・ジャイン（Anil Jain）ミシガン州立大学教授らは、バイオメトリクス（Biometrics）を “biometrics refers to the automatic recognition of individuals based on their physiological and/or behavioral characteristics”（バイオメトリクスは人物の身体的あるいは行動的な特性・特徴を用いて自動的に認識する技術を言う）と説明している。（Anil K. Jain, et al., “An introduction to biometric recognition,” *IEEE Transactions on Circuits and Systems for Video Technology,* Vol.14 No.1, 2004.1, pp.4-20.）本稿では、この “characteristics” を「特性・特徴」と訳す。生体認証に関する文献では “characteristics” を「特徴」と訳しているものが多いが、生体認証が関係するパターン認識分野において特徴というと、一般的に “feature” が用いられるため、これらを区別するために「特性・特徴」という語を用いた。

(3)　より正確に言うと、人それぞれ異なるのではなく、指紋であれば指によって異なるし、虹彩であれば左右の目によって異なる。

も、入室が可能となる。このような問題が生じる理由は、これら2つの方法が、その人物が誰なのかを直接確認をしているのではなく、物や知識を介して確認する方法であるためである。それに対し、生体認証は、その人物が誰なのかを直接確認しようとする方法[4]であり、その確認をする際に、人物から直接、生体情報を取得するため、紛失したり、忘れたりするリスクは極めて低い[5]。また自分の身体的あるいは行動的特性・特徴を書き換えたり、変更したりすることは容易ではない[6]。そのため、生体情報を用いることでより便利で確かな人物の識別・確認が可能になると考えられる[7]。

2　生体認証における基本処理

生体認証では、人物から生体情報を取得し、取得した生体情報と事前に登録してある情報を比較して、識別や確認を行う。事前に登録した情報と比較を行う点では、暗証番号やパスワードを用いた方法と同じではあるが、判定を行う方法に大きな違いがある。暗証番号やパスワードは事前に登録されている情報と「一致」か「不一致」を判定するものであるが、生体認証の場合は、取得される生体情報は、同じ人物から取得されたものであっても、様々な要因により完全に一致することはない。例えば、指紋であれば、センサに入力する角度や場所、また接触の強さなどの影響で、取得されるデータには位置ずれや回転、ひずみなどが生じ、取得するごとに異なるデータとなる。そのため、生体認証では、生体情報が「一致」か「不一致」を評価するのではなく、2つの情報がどの程度似ているのか、あるいは似ていないのか、を評価するスコアを算出する。算出されるスコアの代表的なものが類似度や確率、相違度や距離といった指標[8]である。生体認証の基本的な考え方は、2つの比較する生体情報間でスコアを計算した上で処理を行う。

生体認証の基本的処理内容を説明する前に、生体認証における2つの異なる認証方法について説明する。1つ目は「一対一認証」[9]と呼ばれるもので、もう1つは「一対多認証」[10]と呼ばれるものである。一対一認証では、文字どおり2つの生体情報を比較し、その2つの生体情報が同じ人物あるいは同じ部位から取得されたものかどうかを確認するものである。したがって、期待される認証結果は、「同じ」か「異なる」の2つの選択肢のどちらかである。こ

(4)　英語では、生体情報を用いた認証を表現する際に “something you are” と表現されることがある。それに対して所有物による方法は “something you have” 知識による方法は “something you know” と表現されることがある。

(5)　例えば、指をどこかに置き忘れたり、なくしたりすることは通常の生活においては考えにくいが、ケガにより一時的に使えなくなることや、事故等により指を失うこともあり得るため、その可能性がゼロではない。

(6)　外科的手術により、自分の顔や指紋情報などを変更することは可能である。また、変更できないことは長所である反面、短所にもなる。変更できないことが長所となるのは、人物認証が安定して実施できることである。短所となるのは、もしも何らかの理由によりその特徴が利用できなくなった場合に、取替が効かないということである。例えば、パスワード等は問題が出た場合には何度でも変更ができるが、生体特徴はそういうわけにはいかない。

(7)　ただし、生体認証は完璧な手法ではなく、解決すべき課題が存在する。例えば、生体認証は、取得情報は確認対象人物から直接取得されていることを前提としている。取得される情報が対象人物以外（例えば、事前に準備した本人の特性・特徴を持つ人工物）から取得されたものである場合には、正しく認証することができないため、対策が必要である。これらの対策としては生体検知技術等が挙げられる。

(8)　類似度や確率は、値が大きいほど似ていることを示す。一方、相違度や距離は、値が小さいほど似ていることを示す。

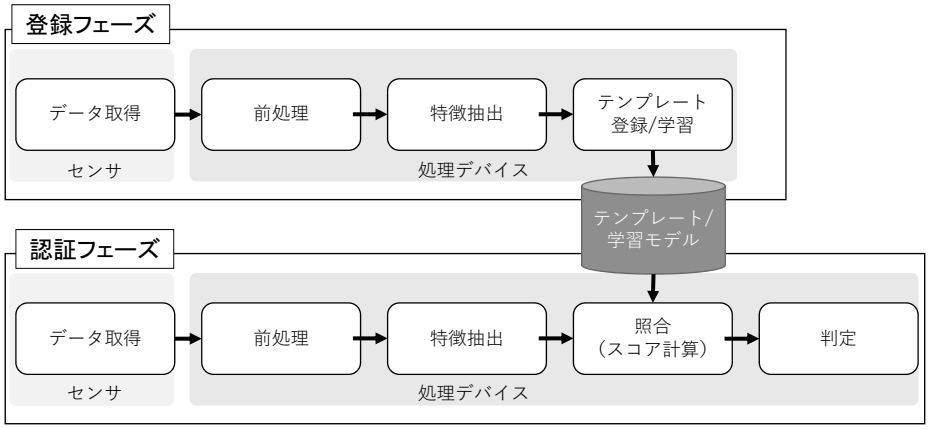
(9)　ID 付き認証とも呼ばれる。英語では “Verification” の単語が用いられ、日本語では「確認」と訳される

(10)　ID なし認証とも呼ばれる。英語では “Identification” の単語が用いられる。日本語では「識別」と訳される。なお、一対 N 認証と呼ばれることもあり、この場合の「N」は「多数」を意味している。後に出てくる Rank-N の「N」とは意味が異なる。

の認証は、2つの生体情報間のスコアを計算した後、そのスコアを事前に設定したしきい値を比較することで行われる。一方で、一対多認証は入力された生体情報が、登録されている多数の生体情報のどれに類似しているのか見つけてくるものである。したがって、期待される認証結果は、類似している登録生体情報を順位付きで返すことである。この認証は入力された生体情報と、登録されている多数の生体情報の各々とのスコアを算出し、似ている度合いが高い順に並べ替えることで行われる。どちらの認証も、2つの生体情報間のスコアを算出することが基本となる。そこで基本的な処理を、一対一認証を想定して説明する。

図1に一対一認証の基本的な処理[11]を示す。生体認証は2つのフェーズ、登録フェーズと認証フェーズから構成される。登録フェーズでは、人物と生体情報をひも付け、認証できるように準備をするフェーズ[12]であり、認証フェーズは、入力される生体情報を基に、事前に登録されているテンプレートと比較をすることで、スコアを算出し、判定を行うフェーズである。各フェーズはデータ取得、前処理、特徴抽出、テンプレート登録／学習、照合（スコア計算）、判定等の処理から構成される。このうち、データ取得はセンサにより行われ、センサにより取得された情報がデータとして処理デバイスに提供され、以降の処理は処理デバイスによって行われる。処理デバイスでは、センサから提供されたデータに様々な前処理を施した後に、そのデータから認証に有効な特徴を抽出する。登録フェーズではこの特徴をテンプレートとして登録し、認証フェーズでは、抽出された特徴と、事前に登録されているテンプレートを照合することで、スコアを算出し、算出されたスコアに基づき、判定を行う。

図1　生体認証（一対一認証）の基本的な処理



（出典）筆者作成。

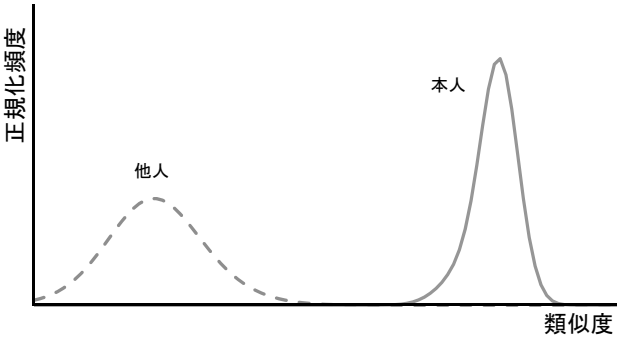
3　生体認証の精度評価

同一人物・同一部位[13]から取得された生体情報であっても、それらは完全に同一とはならない。また、生体情報間のスコアは、生体情報を取得するたびに異なる。そのため、同一人物・同一部位から取得された生体情報間のスコアは、あるばらつきを持って分布することになる。この本人の生体情報間のスコアの分布を本人スコア分布という[14]。一方で、異なる人物

(11)　近年注目されている深層学習に基づく処理は、特徴抽出などを明示的に行わない場合がある。
(12)　登録フェーズでID（人物ごとに設定される番号等）とともに事前に登録されているデータは「テンプレート」や「参照データ」などと呼ばれる。
(13)　指紋は各指で異なるため、同じ指から取得された生体情報間スコアを本人スコアとし、同一人物の指であっても異なる指から取得される生体情報間のスコアは他人スコアとする。
(14)　このばらつきは「個人内変動」と呼ばれる。

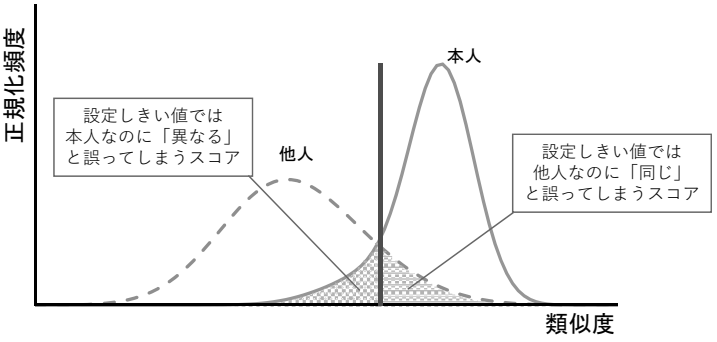
から取得された生体情報間のスコアも、ばらついて分布することになり[15]、この分布を他人スコア分布という。図2、図3に、類似度を考えた場合のスコア分布の概念図を示す。横軸はスコアの値であり、縦軸は正規化した頻度である[16]。理想的な状態は、本人スコア分布と他人スコア分布が一切重なりを持たないものである（図2）。しかしながら実際には、2つの分布を完全に分離することは難しく、結果として図3に示すように、重なりを持ったものとなる。一対一認証では、計算されたスコアに基づき、「同じ」か「異なる」の判断を行う。類似度に基づくスコア分布の場合には、スコアがしきい値以上の場合には「同じ」と判定し、しきい未満の場合には「異なる」と判定するが、図3に示すように、同一人物・同一部位間のスコアであっても、異なると誤って判定されるものもあれば、同一人物・同一部位間ではないにも関わらず、同じと誤って判定されるものもある。前者の誤りを「本人誤拒否」（False Reject）、後者の誤りを「他人誤受入」（False Accept）と呼び、それらの誤り率をそれぞれ「本人誤拒否率」（False Rejection Rate: FRR）、「他人誤受入率」（False Acceptance Rate: FAR）と呼ぶ。FRRとFARはしきい値によって変化し、かつ2つの誤りはトレードオフの関係にある。そのため、FRRとFARのトレードオフ関係について、縦軸にFRR、横軸にFARをプロットした「検出誤りトレードオフ（Detection Error Trade-off: DET）曲線」[17]（図4）により評価したり、FRRとFARの値が同じとなるしきい値における誤り率を「等誤り率」（Equal Error Rate: EER）として計算し、評価したりする。DET曲線は、曲線が左下の原点に近いほど精度が高いと考えることができる。また、特定のFAR（あるいはFRR）を想定した場合に、どの程度のFRR（あるいはFAR）が実現できるかにより評価する目的にも利用される。一方、一対多認証の場合には、入力される生体情報と事前登録されている人物の生体情報間のスコアをそれぞれ計算し、その上でスコアに基づき入力生体情報に似ているN人の人物を抽出してくる。この上位N人の中に、正解の人物が含まれていれば、正解とする。これを多数の入力に対して実施し、上位N位以内に入る割合を算出したものが「N位認証率」であり、「Rank-N認証率」ともいわれる。例えば、N=1の場合は1位認証率であり、正解人物の生体情

図2　本人スコア分布と他人スコア分布
（重なりがない理想的な場合）



（出典）筆者作成。

図3　本人スコア分布と他人スコア分布
（一般的な生体認証の場合）



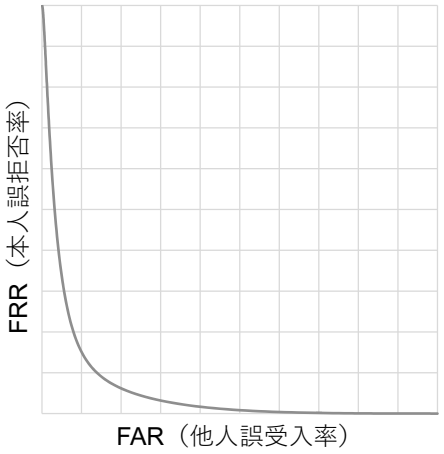
（出典）筆者作成。

[15]　このばらつきは「個人間変動」と呼ばれる。
[16]　正規化した頻度とは、スコア分布の面積が1となるように、各分布における各値での頻度を、各分布の合計頻度で割ったものである。
[17]　生体認証の分野では、この曲線を「受信者操作特性（Receiver Operation Characteristic: ROC）曲線」と呼ぶことが多いが、厳密にはROC曲線とは縦軸と横軸の取り方が異なる。

報が入力生体情報に最も似ている割合を示す。認証率はNの値を大きくすれば大きくなるため、評価の際には横軸にNを、縦軸に対応する認証率をプロットした「累積識別精度特性（Cumulative Match Characteristic: CMC）曲線」（図5）により評価する。一対多認証の評価ではRank-1認証率がよく用いられる。CMC曲線では、曲線が上に位置するほど精度が高いと言える。また、認証率が100％に到達するNを求めて評価することもある（この際に100％に到達するNの値が小さいほど精度が高い、と評価される）。

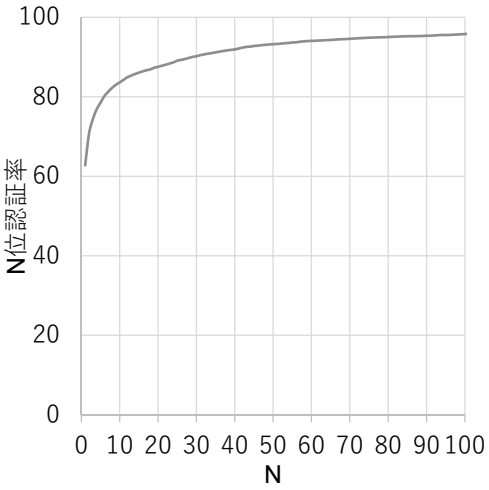
近年、生体認証分野においては、なりすまし攻撃のリスクが指摘され、なりすまし攻撃に対する研究が盛んに行われている[18]。生体認証は、人の特性・特徴が、人によって異なることを利用しており、入力されるものが、認証対象人物の正しい特性・特徴であることを前提としている。しかしながら、センサに対し、人の特性・特徴を再現した人工物などを提示することで、認証システムを攻撃することができる。このような攻撃に対する評価指標としては、本物（bona fide）と攻撃（attack）の判別を考え、攻撃を誤って本物と判定してしまう誤り（Attack Presentation Classification Error Rate: APCER）と本物を攻撃と誤って判定してしまう誤り（Bona Fide Presentation Classification Error Rate: BPCER）などがあり、それらの平均（Average Classification Error Rate: ACER）が用いられたりする。

図４　DET（検出誤りトレードオフ）曲線



（出典）筆者作成。

図５　CMC（累積識別精度特性）曲線



（出典）筆者作成。

４　生体認証に用いられるモダリティ

生体認証は、人物の様々な部位や行動に注目した手法が存在している。表1及び図6に代表的なモダリティを示す。モダリティとは、認証を行う際に注目する特性・特徴のことであり、仮に同じ部位に着目したものであっても、注目する特性･特徴が異なるものは異なるモダリティである。例えば、手に注目した生体認証には、手の形状に注目する掌形、掌の紋様に注目する掌紋、掌（手のひら）内部の静脈形状に注目する掌静脈、また指先の紋様に注目する指紋、指内部の静脈形状に注目する指静脈など、複数のモダリティが存在する。身体的な特性・特徴や行動的な特性・特徴であれば、全てが生体認証に利用できるわけではなく、認証での利用可能性は表2の項目等を考慮しつつ検討される。理想的にはより多くの項目を満たすことが好まし

[18]　例えば、指紋認証について、指状にしたゼラチンの表面に対象者の指紋を再現した「グミ指」が有名である。Tsutomu Matsumoto et al., “Impact of artificial "gummy" fingers on fingerprint systems,” *Proceedings of SPIE,* Vol.4677, 2002.4.19, pp.275-289.

いが、全ての項目を完全に満たすことは困難であるため、利用環境や利用方法などを考慮しつつ利用されている。なお、各モダリティは単独で用いられるだけではなく、複数を組み合わせて用いられることもある。複数のモダリティを組み合わせて認証を行う手法はマルチモーダル認証と呼ばれ、様々な研究がなされている。

以降の章では、いくつかのモダリティを用いた認証技術を紹介するが、その際に、どのような特性・特徴を、どのように取得し、どのように処理をするのか、という点に注目すると、それらの違いが分かりやすい。具体的には、身体的特徴として第Ⅱ章では指紋、第Ⅲ章ではDNA、第Ⅳ章では虹彩、第Ⅴ章では静脈、第Ⅵ章では顔について概説する。また、第Ⅶ章では行動的特徴である歩容について概説し、第Ⅷ章では行動的特徴である声紋と署名について概説する。各モダリティでは、手法の説明に加え、データベース、精度評価、コンペティションなどについてもまとめる。生体認証の研究を進める上では、どのようなデータベースがあるのかを知ることは、研究実施に不可欠なことである。また、どのようなデータによって各モダリティが評価されているのかを知ることは、各モダリティの認証精度を比較する上で非常に重要となる。どの程度のデータサイズ（被験者数、データ数）で評価されているのかは、認証精度と切り離して議論できない。また、認証精度は、データベースを用いて評価されることが一般的であるが、データベースを用いた評価では、そのデータベースに対してのみ高い精度を達成している可能性があるため、より現実的な環境での評価が望まれる。コンペティションでは、手法開発者が過去に利用したことがない新しいデータにより評価されるのが一般的であるため、より現実的な環境に近い状況での精度や傾向を知る手がかりとなる。

表１　生体認証の代表的なモダリティ

身体的 / 行動的	部位	モダリティ
身体的	顔	顔
身体的	眼	虹彩
身体的		網膜
身体的	耳	耳介
身体的・行動的	口（喉）	声紋
身体的	手（全体）	掌形
行動的		署名（筆記）
行動的		キーストローク（キーボードによる入力動作）
身体的	手（掌）	掌紋
身体的		掌静脈
身体的	手（指先）	指紋
身体的		指静脈
身体的		関節紋様
行動的	全身	歩容
身体的	その他	DNA

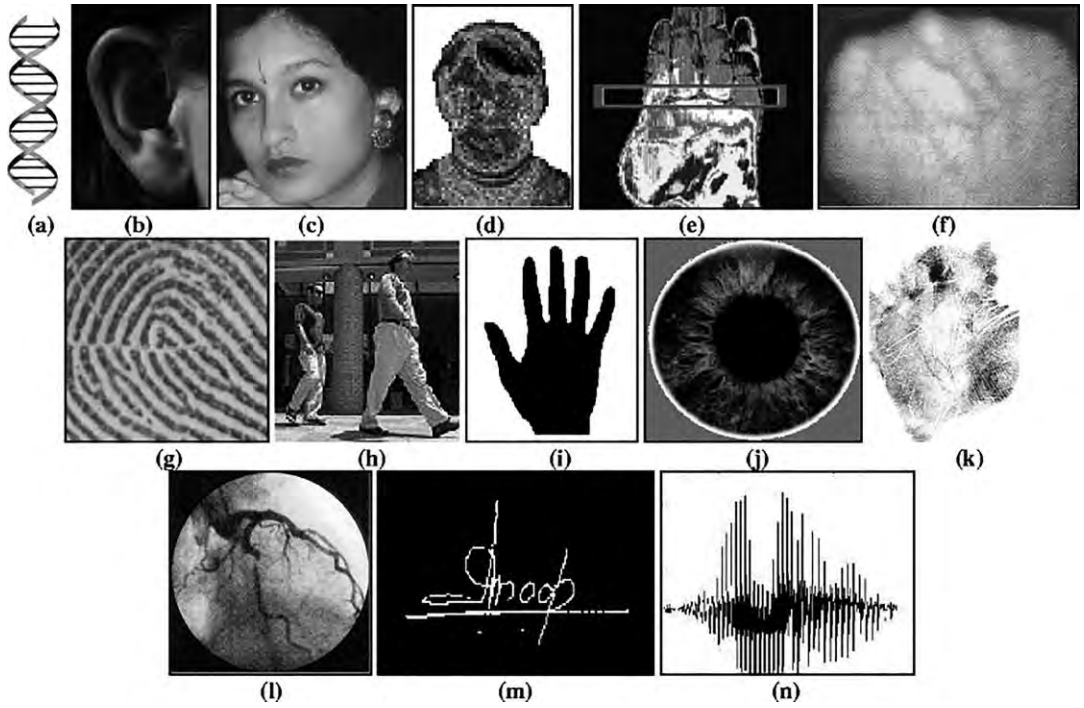
（出典）筆者作成。

表２　モダリティ（特性・特徴）の認証における利用可能性の判断項目

	項目	概要
1	普遍性（Universality）	全ての人が持っていること
2	唯一性（Uniqueness）	万人が異なる特徴を持っていること
3	永続性（Permanence）	終生特徴が変わらないこと
4	収集性（Measurability）	データとして取得可能であること
5	受容性（Acceptability）	データ取得・利用に嫌悪感がないこと
6	精度（Performance）	高い認証精度が実現できること
7	脅威耐性（Circumvention）	なりすまし攻撃等にも強い耐性をもつこと
8	更新可能性（Renewability）	情報が漏洩した場合に変更ができること

（出典）筆者作成。

図６　生体認証の代表的なモダリティ



（注）(a)DNA、(b) 耳、(c) 顔、(d) 顔のサーモグラム、(e) 手のサーモグラム、(f) 手の静脈、(g) 指紋、(h) 歩容、(i) 掌形、(j) 虹彩、(k) 掌紋、(l) 網膜、(m) 署名、(n) 声紋

（出典）Copyright 2004 IEEE. Reprinted, with permission, from Anil K. Jain, Arun Ross and Salil Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.14 No.1, 2004.1, p.8. (Figure 3) <https://doi.org/10.1109/TCSVT.2003.818349>

Ⅱ　指紋

１　技術の概要

人の指先にみられる紋様を指紋といい、線状に隆起している線（隆線）によって形成されている。この指紋パターンは、指によって異なり、また永続性があることから、個人識別に有効な特性・特徴として用いられている。指紋は表出したものであり、人が素手で物に触れた場合に指紋が残るため、主に犯罪捜査[19]などにおける犯人特定目的に利用されてきたが、近年はスマートフォン（以下「スマホ」）に指紋認証機能が搭載されたものが多く発売され、身近に利用されているモダリティである。指紋は利用しやすいモダリティである反面、様々な攻撃が想定されるため、近年は想定される攻撃に対する耐性の研究が多く行われている。

２　認証の流れと特徴表現

指紋認証を実現するためには、まず指紋情報をデータとして読み込む必要がある。指紋は３次元物体である指先のパターンであるため、本来は３次元的なパターンであるが、多くの生体認証アルゴリズムは、指紋パターンを画像（２次元）として取得して処理を行う[20]。

指紋取得には様々なセンサがある。例えば、エリアセンサとラインセンサとに分類され、エ

⑲　日本の警察においては、隆線特徴点 12 点による一致を科学的に同一指紋と判定する基準としているという。齋藤保『弁護人のための指紋鑑定』現代人文社, 2013, pp.26-27.

⑳　３次元指紋を取得し認証を行う手法も提案されている。Javier Galbally et al., "Full 3D touchless fingerprint recognition: Sensor, database and baseline performance," *2017 IEEE International Joint Conference on Biometrics (IJCB)*, 2017, pp.225-233.

リアセンサは、指が静止した状態で指紋を取得するのに対し、ラインセンサはセンサに対し指をスライドさせることで指紋情報を取得する。また指紋情報の読み取り方法も様々であり、光学方式や静電容量方式、電界強度方式、感圧方式、感熱方式などがある。多くのセンサは、指紋が指の凹凸によって形成されている点に注目し、センサに指を接触させることで指紋情報を取得しているが、非接触で指紋を取得する方法も提案されている。センサの目的は指紋のパターンを正確に読み取りデータ化することであるが、乾燥やしわなどの影響等により、実際にデータ化される指紋パターンは、途切れやかすれといったノイズの影響を受ける。また、指の摩耗により指紋パターン取得が困難な場合もある[21]。

図7に指紋のサンプル画像を示す。指における隆線が画像上では濃い色で表示されている。

図7　指紋のサンプル

（注）(a) 登録データ、(b) 入力データ

（出典）Copyright 2004 IEEE. Reprinted, with permission, from Anil K. Jain, Arun Ross and Salil Prabhakar, “An introduction to biometric recognition,” *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.14 No.1, 2004.1, p.8. (Figure 5) <https://doi.org/10.1109/TCSVT.2003.818349>

指紋パターンがデータ化された後は、画像処理アルゴリズムを適用してデータをきれいにした後に、比較する2つの指紋データ間でのスコアを計算し、認証を行う。様々なアルゴリズムが提案されているが、大別すると指紋の特徴的な点（「マニューシャ」という。）に注目する方法と、 指紋のパターンを直接比較する方法とに分けられる。前者の方法では、 特徴的な点として、隆線の終了点である端点や、隆線部分が枝分かれする分岐点等を抽出し認証に利用する。これらの特徴点の種類や位置関係を用いて認証をしたり、それらに加えてマニューシャ間の隆線数を特徴としてとらえ、認証に用いたりする。一方、後者の方法は、指紋の対応する領域パターンの画像を比較することで類似度をスコアとして計算し、認証に用いる。

3　コンペティション、精度評価、データベース

ボローニャ大学（イタリア）、サンノゼ州立大学（米国）、ミシガン州立大学（米国）等により、2000年に国際的な指紋認証コンペティションFVC（Fingerprint Verification Competition）2000が開催され、その後FVC2002、FVC2004、FVC2006が開催されている。同組織が、2009年に開設したFVC-OnGoing[22]というWebベースで指紋認証アルゴリズムの精度評価をするサイトでは、2019年2月13日現在、1,478人が参加し、6,697のアルゴリズムの精度評価がなされている。

データベースとしては、香港理工大学（中国）は、接触型、非接触型センサで取得した2次

[21]　取得された指紋データの質は品質尺度（取得されたデータの良さを評価する指標）として評価することがある。

[22]　FVC-OnGoingは2009年の国際会議で発表がなされ、2009年6月からアルゴリズムの評価を開始している。コンペティション結果は複数の国際会議において報告されている。Bernadette Dorizzi et al., “Fingerprint and On-Line Signature Verification Competitions at ICB 2009,” Massimo Tistarelli and Mark S. Nixon eds., *Advances in Biometrics: Third International Conferences, ICB 2009, Alghero, Italy, June 2-5, 2009, Proceedings*, Berlin:Springer, 2009, pp.725-732; FVC-onGoing website <https://biolab.csr.unibo.it/FVCOnGoing/UI/Form/Home.aspx>（過去のコンペティションFVC2000、FVC2002、FVC2004、FVC2006へのリンクもある。）

元指紋データベースや低解像度指紋、3次元指紋データベースなどを研究用に公開している[23]。元来、指紋認証の精度は、生体から取得された指紋情報が、同一指から取得されたものかどうかを正しく判定する能力として評価されていたが、近年はなりすまし攻撃に対するコンペティションなどが盛んに行われており、それらの結果が報告されているとともに、指紋認証アルゴリズムの精度評価や耐攻撃性評価等の研究に必要なデータベースが提供されている。例えば、カリアリ大学（イタリア）が開催した国際コンペティション LivDet（Liveness Detection Competitions）2017では、17のアルゴリズムがコンペティションに参加し、指紋認証における生体検知精度が評価されている[24]。

指紋認証用のデータベースには米国国立標準技術研究所（National Institute of Standards and Technology: NIST）が提供するデータベースがあったが、現在はその配布が中止されている。また犯行現場などに残された遺留指紋（Latent fingerprint）の評価にはウエストヴァージニア大学（米国）のデータベースがよく用いられている[25]。

指紋は、生体認証における主要なモダリティであるため、複数のマルチモーダルデータベースでもそのデータが提供されている。例えば、マドリッド自治大学（スペイン）による MCYT[26]や欧州の非営利機関（Association BioSecure）による BioSecureDS2[27]、山東大学（中国）による SDUMLA-HMT[28]に指紋データが含まれている。

4　残された課題と対策

攻撃への対策が重要な課題である。2018年9月にドイツで開催されたバイオメトリクスに関する国際会議 BioSig（International Conference of the Biometrics Special Interest Group）[29]、翌10月に米国で開催された IEEE のバイオメトリクスに関する国際会議 BTAS（IEEE International Conference on Biometrics: Theory, Applications and Systems）[30]では、指紋や顔、虹彩を用いた認証への攻撃に関する基調講演があった。また、2018年2月に開催されたパターン認識国際連盟（International Association for Pattern Recognition: IAPR）のバイオメトリクスに関する国際会議 ICB2018でも、指紋認証への攻撃の検知に関するチュートリアル[31]が行われており、この課題が注目されていることを示している。

23　“Databases.” Ajay Kumar (Associate Professor at Hong Kong Polytechnic University) website <https://www4.comp.polyu.edu.hk/~csajaykr/database.php>

24　Valerio Mura et al., “LivDet 2017 Fingerprint Liveness Detection Competition 2017,” 2018.3.15. arXiv.org website <https://arxiv.org/pdf/1803.05210.pdf>

25　このデータベースは、インターネット上で直接ダウンロードできる形で公表されていないが、ウエストヴァージニア大学の担当者に連絡することで取得可能とされている。Kai Cao and Anil K. Jain, “Automated Latent Fingerprint Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.3.22, p.10.

26　“MCYT-Fingerprint-100 database.” Biometrics and Data Pattern Analytics Lab at Universidad Autónoma de Madrid website <http://atvs.ii.uam.es/atvs/mcyt100f.html>

27　“Biosecure Database.” Association BioSecure website <https://biosecure.wp.tem-tsp.eu/biosecure-database/>

28　指紋のほか、虹彩や指静脈、歩容、顔を含む。“SDUMLA-HMT Database.” Machine Learning and Data Mining Lab Department of Computer Science and Technology Shandong University website <http://mla.sdu.edu.cn/info/1006/1195.htm>

29　“BIOSIG 2018.” Fachgruppe Biometrik und elektronische Signaturen, Gesellschaft für Informatik website <http://fg-biosig.gi.de/biosig-2018>

30　“BTAS 2018: IEEE 9th International Conference on Biometrics: Theory, Applications and Systems,” Information Science Institute website <https://www.isi.edu/events/btas2018/home>

31　“Tutorials.” The 11th IAPR International Conference on Biometrics website <https://icb2018.org/tutorials/> TUTORIAL#3の項。

また、近年は遺留指紋に関する研究が盛んに行われている。遺留指紋とは、物に付着して残っている指紋のことであり、犯罪捜査などに用いられる。認証用に取得される指紋と比較すると、そのデータの品質が非常に低く、認証が難しい。BTAS2018 では 9 件の指紋に関する発表があり、そのうち 5 件が遺留指紋に関する発表であった。

近年、認証手法として、深層学習[32]を用いた指紋認証手法も多数提案されている。また、興味深い研究としては、BioSig2018 で発表された、取得された指紋の品質尺度に関する研究が挙げられる。その内容は、指紋の品質尺度が年齢によってかなり異なることを示すものであった。これは年齢によって指紋認証の精度が変わることを示すものではないが、既存の指紋認証アルゴリズムが年齢という要因に影響を受ける可能性を示している。

Ⅲ　DNA

1　技術の概要

DNA（deoxyribonucleic acid. デオキシリボ核酸）とは、デオキシリボース（糖）、リン酸、塩基が 1 つずつ結合したヌクレオチドという単位が、1 本の鎖状に結合した高分子である。ヌクレオチドに含まれる塩基には A（Adenine. アデニン）、T（Thymine. チミン）、G（Guanine. グアニン）、C（Cytosine. シトシン）の 4 種類がある。DNA の A/T/G/C の並び順（塩基配列）が遺伝情報となり、生物の体の形成や生命活動を行う際の設計図として働く。

ヒトの 1 個の細胞の核にある全 DNA をつなぐと、約 60 億塩基分の長さになる。このうち 99.9% は全人類に共通である。共通でない残りの部分には、個人ごとに少しずつ異なる形で、塩基の挿入や欠失、塩基の置換、塩基の反復回数の変化などが見られる。このような共通でない部分が、DNA による個人識別に利用される。

2　認証の流れと特徴表現

個人ごとの DNA の相違のうち、現在、個人識別に最もよく用いられるものは、STR（Short Tandem Repeat）と呼ばれる反復配列である[33]。ヒトの染色体 DNA には STR が 2 万箇所ほど存在し、各 STR に、D8S1179 や TH01 などの名称が付けられている。例えば、D8S1179 という STR は、TCTA という 4 塩基の単位が TCTATCTATCTA…のように反復しているもので、日本

---

[32]　人の神経網を模したニューラルネットワークを用いた学習のうち、特に、多くの層を用いた深いネットワーク構造によるものを指す。

[33]　個人ごとの DNA の相違には、塩基の反復回数が異なる場合と塩基配列そのものが変化している場合がある。塩基の反復回数が異なる場合は、さらに、マイクロサテライトとミニサテライトとに分けられる。マイクロサテライトとは、数塩基の長さの単位で配列が反復するもので、例えば 3 塩基単位の反復なら ATCATCATCATC…などとなる。マイクロサテライトは、STR とも呼ばれる。一方、マイクロサテライトより長い反復単位をもつものをミニサテライトと呼び、例えば 10 塩基以上の配列を単位として反復する。次に、塩基配列そのものが変化している場合とは、1 ～数塩基の置換や挿入、欠失などである。有名なものとして、1 塩基が別の塩基と置換した SNP（1 塩基多型。スニップと発音する。）がある。例えば…ATCG…という配列をもつ人と、C が G に置き換わった…ATGG…をもつ人が、一定の割合で存在するようなとき、C/G の箇所を SNP と呼ぶ。DNA によって個人識別が可能であるという世界初の報告は、1985 年、英国の遺伝学者アレック・ジェフリーズ（Alec J. Jeffreys）らによる「DNA 指紋」（DNA fingerprinting）と呼ばれる手法で、ミニサテライトを用いていた。（A.J. Jeffreys et al., “Individual-specific ‘fingerprints’ of human DNA,” *Nature*, Vol.316, 1985.7.4, pp.76-79.）ミニサテライトは、全体で非常に長い配列になることがあるため、個人識別には、劣化していない DNA 試料が多量に必要になることがある。

人では13回反復するタイプを持つ人が一番多く(34)、次が14回、その次が15回となり、7回から20回までの計12タイプが知られている（8回や19回反復は報告がない）(35)。また、TH01では、6回反復から10回反復までの、5タイプが知られている。反復回数を知るには、反復する配列の部分だけをPCR（Polymerase Chain Reaction. ポリメラーゼ連鎖反応）という手法で増幅し、その長さを電気泳動法で検出する。

個人識別においては、問題とする試料と、あらかじめ本人（あるいは血縁者）のものと明確にわかっている試料のSTRを検査して比較し、タイプ（反復回数）が一致するかどうかを調べる。複数のSTRでタイプを調べ、それらを組み合わせると、ある個人にしか存在しないパターンになる(36)。現在は、異なる20前後のSTRを組み合わせ、個人を識別できるようにした検査キットが、米国企業などから複数販売されている(37)。また日本の警察庁も、上記市販のキットに用いられるのと同様のSTRを用いた、個人識別システムを採用している(38)。

3　データベース、精度評価

よく利用されるSTRでは、日本人や民族などの集団におけるそれぞれのタイプの出現頻度が計測され、公開されている(39)。

一方、統計情報ではなく、実際に誰がどのタイプのSTRを持つかという個人の情報については、犯罪者や容疑者を対象に、国などが収集して、犯罪捜査のために氏名などとともにデータベース化していることがある(40)。例えば、日本では、警察庁によって、犯罪者等のSTRを登録するDNA型データベースが運用されている(41)。他方、犯罪者や容疑者ではない人については、民間事業者、特にインターネットを介して遺伝子検査キットを販売する遺伝子検査サービス会社が、顧客の情報としてDNAのタイプを収集している(42)。これらの事業者は米国に多く存在し、STRに限らず、その他の多様なDNAの部分について、DNAのタイプを検査しているが、同時に、そのデータを蓄積して、家系図作成や祖先検索等のサービスを展開している。

ところで、STRのタイプが、刑事や民事での鑑定に用いられるときは、確率計算が必要となることがある。例えば、刑事鑑定では、2つの試料の間で各STRの全タイプが一致した場合、そのタイプを持つ人の存在確率を計算する(43)。15のSTRを判定に用いるのなら、15のSTRのそれぞれのタイプの日本人一般における出現頻度を掛け合わせたものが、存在確率となる(44)。

---

(34)　染色体は父方と母方から受け継いで2本1組の1対になっているので、片方の染色体のSTRは13回反復でもう片方では14回反復であるなど、2つのタイプの組合せになる。(13, 14) などと表記する。

(35)　勝又義直『最新DNA鑑定―その能力と限界―』名古屋大学出版会, 2014, pp.267-272.（付録2：日本人集団における各STRアリールの出現頻度）

(36)　15程度のSTRを組み合わせると、全世界の人口を超える数の異なるタイプの組合せができる。したがって、他人同士での偶然の一致は非常に珍しいことになる。

(37)　例えば、“PowerPlex® 21 System.” Promega website <https://www.promega.jp/products/genetic-identity/genetic-identity-workflow/str-amplification/powerplex-21-system/?catNum=DC8902>

(38)　警察庁編『警察白書 平成28年版』警察庁, 2016, p.89. <https://www.npa.go.jp/hakusyo/h28/pdf/pdf/06_dai2syo.pdf>

(39)　勝又　前掲注(35)

(40)　英国の国家DNAデータベース（National DNA Database: NDNAD）や米国のCODIS（Combined DNA Index System）などがある。

(41)　勝又　前掲注(35);「DNA型記録取扱規則」（平成17年国家公安委員会規則第15号）

(42)　Yaniv Erlich et al., “Identity inference of genomic data using long-range familial searches,” *Science*, Vol.362, Issue 6415, 2018.10.11, pp.690-694.

(43)　勝又　前掲注(35), p.95.

(44)　ただし、両親から受け継いだ1対のSTRのタイプのどちらが母親由来でどちらが父親由来かという点は調べないので、それぞれのSTRで2を掛けておく。

もしそれが 1% となるなら、容疑者が 1,000 人いれば、10 人の同タイプの人がいることになり、STRのみで犯人と判定することはできない。ただし、実際の検査では、15ものSTRを用いれば、頻度の高いタイプの組合せでも存在確率は数億分の 1 といった非常に小さな値になる。よって日本人全体などの大きな集団中でも、偶然に 15 の全タイプが同じという人が現れることはほぼないと考えられる。また、STR のタイプが一致しないときにも、別途の確認が必要なことがある。例えば、STR 部分を増幅する PCR において、その両末端を決定するプライマーに相当する DNA 部位に突然変異が生じることがあり、その場合には、STR が存在しても PCR で増幅されず、消失したように見えることになる。このような場合、確認や条件の付加といった対応が必要である。人の人生に大きな影響を及ぼす刑事や民事の鑑定、あるいは 2001 年 9 月 11 日の米国同時多発テロなどのテロ事件や大規模災害による身元確認などに利用される場合には、慎重な判定が行われている。

４　残された課題と対策

STR による個人識別は、キットが市販されており、 判定までの作業がほぼ自動化されているため、一度に大量の試料を処理することができる。しかし、試料採取からタイプの判定までに最低でも数時間かかるため、この時間が支障にならない分野での利用に限られる。例えば、犯罪捜査での容疑者の識別や親子鑑定では、STR がよく用いられている。一方、STR のタイプを個人に固有な数字の列として使い、サイバー空間の本人認証にも応用できるのではないかと言われているが、判定に時間がかかることが課題の 1 つとされ、現在まだ実用には至っていない[(45)]。

これまで、識別の目的や、試料の種類や状態に合わせて、識別の手法が工夫されてきた。今後も、より精確で安定的に、効率良く識別ができるよう、手法の検討や開発が続けられると考えられる。例えば、STR は短めの塩基配列であるため、偶然他人が同じタイプである可能性は皆無ではない。血縁鑑定では血縁関係が離れるにつれ、STR を複数組み合わせても判定が難しいことがある。また、一卵性双生児では、同じ DNA を持つため STR のみで識別はできない。核 DNA が残っていないような非常に古い試料では、ミトコンドリア DNA が利用されることがあるが、ミトコンドリアは母親からしか子供に伝わらないため、父子関係は分からないといった問題がある。

最後に、DNA は身体のほぼ全ての部位から取得できるため、本人が知らないうちに他人が勝手に試料を取得することによる、なりすましを防ぐ方策が必須である。使用済みの歯ブラシや落ちた毛髪の取得、あるいは眠っている人の口腔粘膜を綿棒で擦り取るといった方法で、他人の DNA を本人の了承なしに入手することが可能である。また、日本の「個人情報の保護に関する法律」（平成 15 年法律第 57 号。「個人情報保護法」）では、個人情報の盗み取りなどは規律できても、DNA などの物質自体は対象とされていないことから、個人の遺伝情報が DNA 試料から勝手に盗み取られた場合、個人情報漏洩や不正利用として規律することが難しいと考えられる。DNA による個人識別が広く実用化されるときには、情報を含む個人の細胞や組織をどのように保護するかの検討も必要と考えられる。

---

(45)　バイオメトリクスセキュリティコンソーシアム編『バイオメトリックセキュリティ・ハンドブック』オーム社, 2006, pp.203-214.

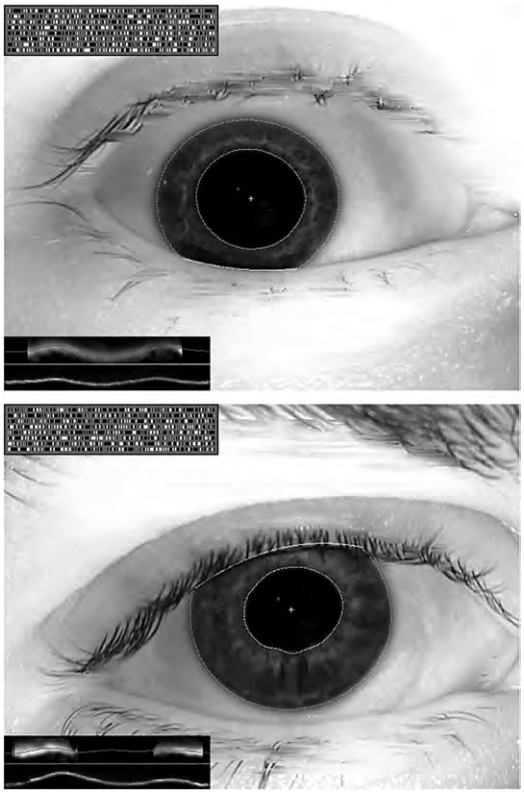
Ⅳ　虹彩

１　技術の概要

虹彩は眼球の色のついた部分から黒い瞳孔部分を取り除いた部分であり、この領域のしわ模様が人によって異なること[46]、また生涯不変で永続性があることから[47]、モダリティとして利用されている。黒目の周りの模様（図６(j)、図８）を用いて認証を行うものが虹彩認証である。

虹彩についてもなりすまし攻撃が報告されており、それらの評価なども行われている。

図８　虹彩とアイリスコード（各写真の左上）



（出典）Copyright 2008 IEEE. Reprinted, with permission, from John Daugman and Cathryn Downing, “Effect of Severe Image Compression on Iris Recognition Performance,” *IEEE Transactions on Information Forensics and Security*, Vol.3 No.1, 2008.3, p.54. (figure 2) <https://doi.org/10.1109/TIFS.2007.916009>

２　認証の流れと特徴表現

虹彩認証ではジョン・ドーグマン（John Daugman）ケンブリッジ大学教授の提案したアルゴリズム[48]が最も有名である。認証では目の領域をカメラで撮影することでデータを取得する。データ取得には可視光や近赤外光などが使われる。センサで虹彩のみを撮影することはできないため、データ取得後虹彩領域を特定し、虹彩領域部分のみを切り出す前処理を行う。この際にまぶたやまつ毛などの影響がある領域を特定し、認証に利用しないようにする。その後、虹彩領域を同心円の８層に分け、フィルタを適用した上で「アイリスコード」（図８）と呼ばれるデータを作成し、これをテンプレートとして保存する。認証時には同じく入力データからアイリスコードを生成し、登録されているアイリスコードと入力データから生成されたアイリスコードの相違度を計算し、認証に利用する。

３　データベース、精度評価、コンペティション

虹彩認証のコンペティションは米国の NIST により IRIS Challenge Evaluation (ICE) 2005 や ICE 2006 などが開催され、最近ではクラークソン大学（米国）が 2017 年に Liveness Detection-Iris Competition 2017（LivDet-Iris2017）[49] を開催している。

---

(46)　同一人物でも左右でその模様は異なる。

(47)　虹彩の紋様は２歳くらいまでに形成され、その後変化しないとされる。虹彩に関する解剖学的、生理学的、臨床的な詳細については、次が詳しい。F. H. Adler, *Physiology of the eye: Clinical Application*, fourth ed., London: The C.V. Mosby Company, 1965; Anil K. Jain et al., eds., *BIOMETRICS: Personal Identification in Networked Society*, Berlin: Springer, 2006.

(48)　J John G. Daugman, “High Confidence Visual Recognition of Persons by a Test of Statistical Independence,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.15, No.11, 1993.11, pp.1148-1161.

(49)　David Yambay et al., “LivDet iris 2017: Iris liveness detection competition 2017,” *2017 IEEE International Joint Conference on Biometrics (IJCB)*, 2017, pp.733-741.

虹彩認証には、多数のデータベースが公開されている。例えば、中国科学院自動化研究所により CASIA-IrisV1 ～ V4 などが提供されている[50]。また、ノートルダム大学（米国）は ICE2005 や ICE2006 で利用された虹彩のデータセット（ND-IRIS-0405 Iris Image Dataset）を始め、LivDet-Iris2017 で利用された模様付きコンタクトレンズ装着有無のデータセットなどを多数公開している[51]。このほか、マルチモーダル DB には虹彩が含まれているものが多数存在する。

４　残された課題と対策

模様のついたコンタクトレンズを装着したなりすまし攻撃への対応などが現在学会等で議論されており、今後の課題である[52]。また、虹彩認証では、高解像度の画像が必要となるため、近距離で撮影された場合にのみ適用が可能であった。その欠点を補うため、近年は顔の眼を含む部分領域の情報を使った認証手法などが提案されている[53]。

Ⅴ　静脈

１　技術の概要

静脈認証は、血管パターンを利用して個人を認証しようとするものである。現状利用されている部位は手と眼であり､手を対象とする掌静脈認証や指静脈認証などが広く利用されている。血管パターンは体内にあるため、静脈認証を行うためには通常のカメラではなく、近赤外カメラなどを用いて生体情報を取得する。静脈認証は日本国内において複数の銀行 ATM などに導入されている。例えば､掌静脈認証は三菱 UFJ 銀行や大垣共立銀行の ATM に導入されており、指静脈認証は三井住友銀行やみずほ銀行の ATM などで導入されている。

２　認証の流れと特徴表現

体内にある血管パターンを特徴として取得する。データ取得には近赤外光が利用される。静脈中に存在する還元ヘモグロビンは近赤外光をよく吸収するため、近赤外光を照射し、撮影することで、還元ヘモグロビンが存在する静脈部位が暗く映り、静脈パターン情報を非接触で取得することができる。パターンが取得できれば、その後は前処理を施した後に、静脈パターンの類似度をスコアとして計算し、しきい値で比較することで認証を実現する。図 9 に掌静脈の例を、図 10 には指静脈の例を示す。図 9 及び図 10 のいずれにも実際の生体から取得したデータとともに、なりすまし攻撃を想定したデータも併せて示す。

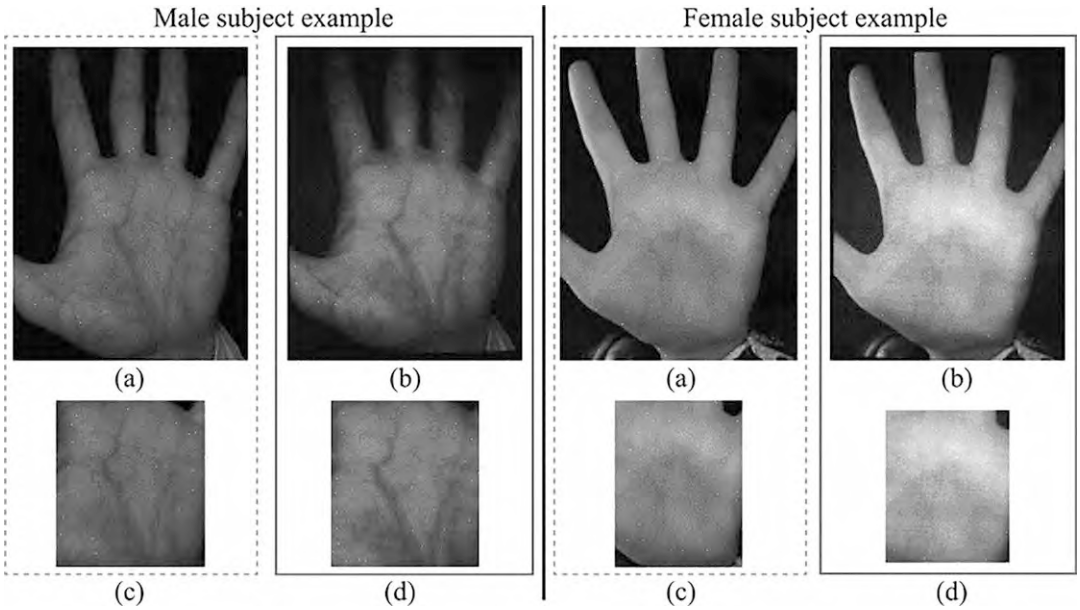
⑸0　“Databases.” Center for Biometrics and Security Research website <http://www.cbsr.ia.ac.cn/china/Iris%20Databases%20CH.asp>　同研究所のパターン認識国家重点実験室の研究グループの１つとして、バイオメトリクス・セキュリティ研究センター（Center for Biometrics and Security Research: CBSR）が置かれている。

⑸1　“Datasets.” Computer Vision Research Lab, University of Notre Dame website <https://cvrl.nd.edu/projects/data/>

⑸2　例えば、BioSig2019 では、ウエストヴァージニア大学の大学院生研究員（Graduate Research Assistant）ドメニック・ポスター（Domenick Poster）らによる“Deep Sparse Feature Selection and Fusion for Textured Contact Lens Detection”という発表があった。

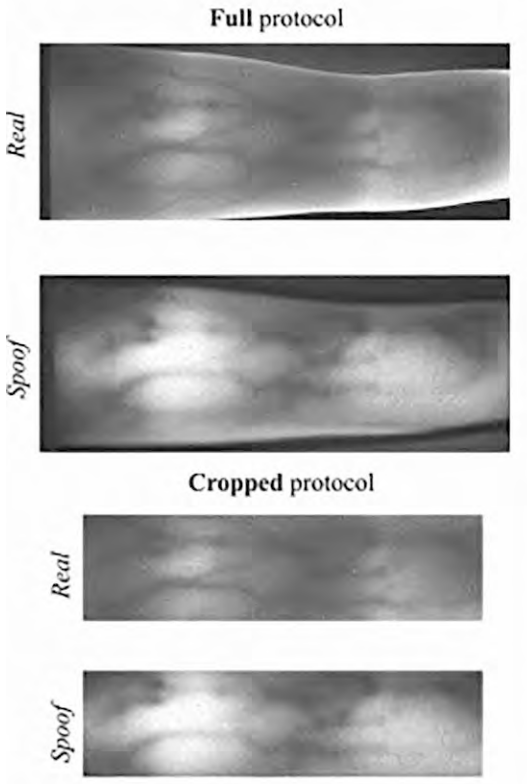
⑸3　例えば、Zijing Zhao and Ajay Kumar, “Improving Periocular Recognition by Explicit Attention to Critical Regions in Deep Neural Network,” *IEEE Transactions on Information Forensics and Security*, Vol.13 No.12, 2018.12, pp.2937-2952.

図９　掌静脈



（注）左は男性、右は女性の例。
(a)、(c) は本物の手をかざしてセンサにより取得した画像。
(b)、(d) は印刷した画像をセンサにより再度撮影して取得した攻撃を想定した画像。
（出典）Copyright 2015 IEEE. Reprinted, with permission, from Pedro Tome and Sébastien Marcel, “On the vulnerability of palm vein recognition to spoofing attacks,” *2015 International Conference on Biometrics (ICB)*, 2015, p.321. (Figure 2) <https://doi.org/10.1109/ICB.2015.7139056>

図１０　指静脈



（注）Real は本物を、Spoof はなりすまし攻撃の画像データである。上２つの図は取得した指静脈のデータ全体を示しており、下２つの図はそこから認証に使う部分だけを切り出したデータである。
（出典）Copyright 2015 IEEE. Reprinted, with permission, from P. Tome et al., “The 1st Competition on Counter Measures to Finger Vein Spoofing Attacks,” *2015 International Conference on Biometrics (ICB)*, 2015, p.517. (Figure 1) <https://doi.org/10.1109/ICB.2015.7139067>

3　データベース

静脈にかかわるデータベースを表3にまとめる。公開されているデータベースのほとんどは指静脈に関するものである。掌静脈については掌側から撮影されたデータがほとんどであるが、手の甲側から撮影されたデータも一部存在する[54]。

表3　指静脈、掌静脈のデータベース

データベース名	機関名（国・地域名）	人数	画像数	部位
THU-FVFDT	清華大学深圳大学院（中国）	610	6540	指（2本）*1
SDUMLA-HMT	山東大学（中国）	106	3816	指（6本）
HKPolyU-FID	香港理工大学（中国）	156	3132	指（2本）*2
UTFVP	トウェンテ大学（オランダ）	60	1440	指（6本）
MMCBNU_6000	全北大学（韓国）	100	6000	指（6本）
FV-USM	マレーシア科学大学（マレーシア）	123	5940	指（4本）
VERA FingerVein	Idiap研究所（スイス）	110*3	440	指（2本）
VERA PalmVein		110	2200	掌（左右）
PROTECT Multimodal DB	PROTECT研究コンソーシアム（EU）*4	20	240	指（4本）
		31*5	920	掌
PLUSVeinDorsal-Palmar Finger Vein	ザルツブルク大学（オーストリア）	60	7200	指（6本）*6

（注）*1 THU-FVFDT1、THU-FVFDT2、THU-FVFDT3 の 3 つに分けて提供している。
*2 静脈のみならず指表面のテクスチャ情報も取得している。
*3 現在は 50 人分のデータのみ利用可能である。
*4 全体のとりまとめは、レディング大学（英国）による。
*5 31 人の被験者からデータを取得しているが、ウェブサイトでダウンロードできるのは 20 名分のみである。
*6 指静脈を表側、裏側両方から取得しており、また 2 種類のスキャナで取得している。

（出典）“Tsinghua University Finger Vein and Finger Dorsal Texture Database (THU-FVFDT).” Graduate School at Shenzhen, Tsinghua University website <http://www.sz.tsinghua.edu.cn/labs/vipl/thu-fvfdt.html>; “SDUMLA-HMT Database.” Machine Learning and Data Mining Lab, Shandong University website <http://mla.sdu.edu.cn/info/1006/1195.htm>; “The Hong Kong Polytechnic University Finger Image Database (Version 1.0).” Ajay Kumar (Associate Professor at Hong Kong Polytechnic University) website <http://www4.comp.polyu.edu.hk/~csajaykr/fvdatabase.htm>; “Finger Vein.” Services, Cybersecurity and Safety researchgroup, University of Twente website <https://scs.ewi.utwente.nl/downloads/show,Finger%20Vein/>; “Finger Vein Database: MMCBNU_6000.” Multimedia Lab, Chonbuk National University website <http://multilab.jbnu.ac.kr/MMCBNU_6000>; “Finger Vein USM (FV-USM) Database.” Bakhtiar Affendi Rosdi (Associate Professor at Universiti Sains Malaysia) website <http://drfendi.com/fv_usm_database/>; “The Idiap Research Institute VERA Fingervein Database.” Idiap Dataset Distribution Portal website <https://www.idiap.ch/dataset/vera-fingervein>; “PROTECT Multimodal Dataset.” PROTECT consortium website <http://projectprotect.eu/dataset/>; “PLUSVein-FV3 Finger Vein Data Set - Download Page.” Multimedia Signal Processing and Security Lab, University of Salzburg website <http://www.wavelab.at/sources/PLUSVein-FV3/> 等を基に筆者作成。

4　残された課題と対策

BioSig2018 では複数件の静脈認証に関する発表があった。そこで取り上げられた課題の 1 つとして、指の回転による認証精度劣化の問題がある。これについて、様々な角度から静脈データを撮影し、複数のアルゴリズムにより精度を検証したところ、実際の指静脈システムなどで、10 度以上回転すると精度が劣化するという報告があった[55]。これは、指静脈システムにおいても指の置き方に関する精度劣化に改善の余地があることを示している。

静脈は体内情報であるため、そのパターンを入手することは容易ではないが、一度パターン

(54)　掌側から取得する場合と手の甲側から取得する場合では、取得できる静脈特徴が異なるため、同じモダリティを用いた異なる特徴として価値がある。

(55)　Bernhard Prommegger et al., “Longitudinal Finger Rotation in Finger-Vein Recognition,” *2018 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2018. <http://wavelab.at/papers/Prommegger18b.pdf>

が流出すれば、なりすまし攻撃のリスクもある。静脈においても人工物によるなりすまし攻撃への対応はまだ解決しているわけではない。表3で示したデータベースにおいても、なりすましを想定したデータが収集され、攻撃耐性評価や、なりすまし攻撃に頑健なアルゴリズム構築のために利用されている。

Ⅵ　顔

1　技術の概要

顔は、人の体の部位の中で、人が最もよく見ている部位であり、人と人とのコミュニケーションにおいても、顔から読み取ることのできる情報、例えば、性別・年齢・表情・感情・意図が活用されている。中でも、個人性の情報は重要であり、実際に、人が個人を認識する場合に最もよく用いられるモダリティは顔であるといっても過言ではない。これらの顔情報を計算機によって自動的に解析することにより、顔検出によるデジタルカメラのオートフォーカス、年齢推定による商品販売の可否判定、個人認証によるアクセスコントロールや科学捜査等への幅広い応用が期待される。

計算機による自動的な顔画像認識は、1970年代の黎明期の研究[56]を皮切りにして、1980年代以降から現在に至るまで、コンピュータビジョン[57]やパターン認識[58]の分野における主要なトピックとして数多くの研究がなされている。また、顔画像から読み取ることのできる情報が個人性にとどまらないことから、顔画像による性別識別・年齢推定・表情認識・魅力度推定・感情推定・意図推定等、非常に多岐にわたる研究がなされている。本章では、顔画像による個人認証に焦点を当て、顔認識技術を概説する[59]。

2　認証の流れと特徴表現

（1）顔認識の流れ

顔認識は、コンピュータビジョンやパターン認識の技術によって実現するものであり、入力画像に対して、顔検出、顔の位置合わせ（正規化）、特徴抽出、識別処理を経て、個人を認証する。以下、顔認識の流れを概説する。

(a) 入力画像の取得

多くの場合、カメラにより単一の入力画像を取得するが、個人内変動を含めるために時系列画像（ビデオ）を入力とすることもある。また、高いセキュリティレベルを要する場所では、顔写真によるなりすまし攻撃への耐性を得るために顔の3次元形状を解析することもあるた

56　Takeo Kanade, “Picture Processing System by Computer Complex and Recognition of Human Faces,” Doctoral dissertation, Kyoto University, November, 1973. <https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/162079/2/D_Kanade_Takeo.pdf>

57　コンピュータビジョン（computer vision）とは、計算機に人間の視覚機能を与えることを目指す研究分野であり、画像から物体やシーンを認識することや、実世界に存在する物体の3次元的な幾何構造や光学的な性質を理解することなどが挙げられる。

58　対象を定量的なパターンで表現して、認識する学問領域であり、画像のみならず音声信号･加速度センサの信号といった様々なメディアを対象とした認識問題を扱う。

59　顔認識の詳細な解説については次を参照されたい。 Rama Chellappa et al., “Human and machine recognition of faces: a survey,” *Proceedings of the IEEE*, Vol.83 No.5, 1995.5, pp.705-741; 岩井儀雄ほか「画像処理による顔検出と顔認識」『情報処理学会研究報告』 Vol.2015, No.38, 2005.5, pp.343-368. <http://id.nii.ac.jp/1001/00052376/>; Mei Wang and Weihong Deng, “Deep Face Recognition: A Survey,” 2018.4.18. <https://arxiv.org/pdf/1804.06655.pdf>

め[60]、複数カメラによる多視点顔画像や距離センサによる顔の奥行き画像を利用することもある。

(b) 顔検出

画像全体から顔の位置・領域を検出する。顔領域を含む外接矩形[61]として検出することが一般的である。

(c) 顔の位置合わせ（正規化）

口や目などの顔器官や輪郭を含むランドマーク点[62]を検出して、顔画像の位置合わせや顔向きの正規化を行う。

(d) 特徴抽出

検出された顔領域から特徴を抽出する。外接矩形内の画像全体から大局的な特徴を抽出する方法に加えて、検出された顔器官やランドマーク点に基づいて局所的な特徴を抽出する方法がある。

(e) 識別

抽出された特徴間の相違度を算出する。一般的な識別手法として、最近傍識別[63]、k 近傍識別[64]、判別分析[65]、サポートベクターマシン[66]等が用いられる。特に近年は、深層学習による枠組みが主流になりつつあり、特徴抽出と合わせて識別器を学習する[67]事例が増加している。さらには、深層学習を用いて顔検出・位置合わせ・特徴抽出・識別を同時に行う手法も提案されている。

（2）顔認識における個人内変動

人物の顔は、非剛体[68]の 3 次元形状を持つ、すなわち、一定程度変形する 3 次元の形状であることから、同一人物であっても観測される画像は変動する。特に、顔画像における個人内変動の主要 3 要素である、顔の向き（Pose）、照明条件（Illumination）、表情（Expression）の変動は、認証精度に大きな影響を与えることから、英訳の頭文字を取って、PIE と総称される。加えて、経年変化、サングラスやマスクによる顔の部分隠蔽、撮影距離の違いによる画像解像度の変化等も、顔画像における個人内変動として挙げられる。

---

(60)　Alize Scheenstra et al., “A survey of 3D face recognition methods,” Takeo Kanade eds., *Audio- and Video-Based Biometric Person Authentication 5th International Conference, AVBPA 2005, Hilton Rye Town, NY, USA, July 20-22, 2005. Proceedings*, Berlin: Springer, 2005, pp.891-899.

(61)　対象物体を取り囲むように設定する矩形（長方形）をいう。

(62)　画像を特徴付ける目印となる点。

(63)　データベースに登録されている複数の人物の顔画像に対して、入力の顔画像に最も似ている顔画像を選択して、選択した画像に対応する人物として認識する手法。

(64)　データベースに登録されている複数の人物の顔画像に対して、入力の顔画像に最も似ている k 枚の顔画像を選択して、選択した k 枚の画像に対応する人物セットから多数決などにより人物を決定する手法。

(65)　同一人物内のばらつきである個人内変動を小さくしつつ、人物間の差である個人間変動を大きくするような特徴空間を用いて識別する手法。

(66)　2 つのクラス間のマージン（識別境界付近でいずれのクラスの学習サンプルも存在しない領域の幅）が最大になるような識別平面を用いて識別する手法。

(67)　識別器は、入力されたデータを各クラスに識別するためのアルゴリズム。一般的に識別のためのパラメータが存在するため、学習データを用いてそのパラメータを求めることを指す。

(68)　変形しない物体である剛体に対して、変形する物体のことを指す。

（３）顔検出

顔検出処理は、その精度が後段の顔認識に大きな影響を与えることから、顔認識にとって極めて重要な前処理となる。顔検出の流れは、まず入力画像中に特定の位置と特定の大きさの外接矩形を設定して、その外接矩形によって切り取られる部分画像が顔かどうかを判定する。この作業を、画像全体で様々な位置と大きさに外接矩形を設定することで、画像中の顔を検出する。このように、顔検出は、画像全体を探索する処理であり、その計算量は非常に大きいものとなることから、高速なアルゴリズムが望まれる。

顔検出の黎明期においては、顔輪郭や顔器官のエッジ（線）に基づく手法[69]、顔器官の輝度が周囲と比較して低くなることに基づく輝度分布による手法[70]、顔器官検出に基づく手法[71]、顔の肌色検出に基づく手法[72]等、ルールベースの手法が数多く提案された。しかし、これらの手法は、各手法が想定する状況とは異なる場合にうまくいかないといった欠点がある。

これに対して、学習データを用いた統計的機械学習による顔検出手法も数多く提案されている。いずれも、部分画像（画像全体から一部分の画像を切り取ったもの）を入力として、それが顔であるか否かを判定する手法となる。例えば、部分画像を主成分分析[73]や線形判別分析[74]によって低次元な部分空間に射影[75]して判定する手法[76]や、部分画像に対して直接ニューラルネットワークやサポートベクターマシン等の識別器を適用する手法[77]が提案されている。特に、Viola と Jones による、Haar-like 特徴[78]と AdaBoost[79]を用いたカスケード型[80]の顔検出器（図 11）[81]は、極めて高速に動作し、かつ高精度に顔を検出できることから、長らく顔検出の標準的な枠組みとして用いられており、また、その派生手法も多数提案されてきた[82]。

一方、近年では、物体検出を始めとする画像認識タスクにおいて深層学習が成功を納めたこ

⑹⑼　T. Sakai et al., “Line extraction and pattern detection in a photograph,” *Pattern Recognition*, Vol.1 No.3, 1969.3, pp.233-236.

⑺⓪　Guangzheng Yang and Thomas S. Huang, “Human face detection in a complex background,” *Pattern Recognition*, Vol. 27 No.1, 1994.1, pp.53-56.

⑺⑴　T.K. Leung et al., “Finding faces in cluttered scenes using random labeled graph matching,” *Proceedings of IEEE International Conference on Computer Vision*, 1995, pp.637-644.

⑺⑵　Jean-Christoph Terrillon et al., “Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images,” *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (AFGR 2000)*, 2000.3.

⑺⑶　画像を始めとする高次元のデータを、情報の損失ができるだけ小さくなるように低次元のデータに変換する手法。

⑺⑷　画像を始めとする高次元のデータを、各クラスの判別が容易な低次元データに線形に変換する手法。

⑺⑸　多数の要素を持つ高次元空間のある部分を表現する空間（例えば、体積が表現される 3 次元空間に対して、面積しか表現されない 2 次元空間）を部分空間という。ここでは、高次元空間にあるデータを、限られた要素数で表現される部分空間上のデータに変換することを表す。

⑺⑹　Kah-Kay Sung and Tomaso Poggio, “Example-based learning for view-based human face detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.20 No.1, 1998.1, pp.39-51.

⑺⑺　Henry A. Rowley et al., “Neural network-based face detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.20 No.1, 1998.1, pp.23-38.

⑺⑻　画像中の局所領域の明暗パターンに着目した特徴。

⑺⑼　個々では必ずしも識別性が高くはない弱識別器を組み合わせて用いることによって、識別性の高い強識別器を構成することで識別する手法。

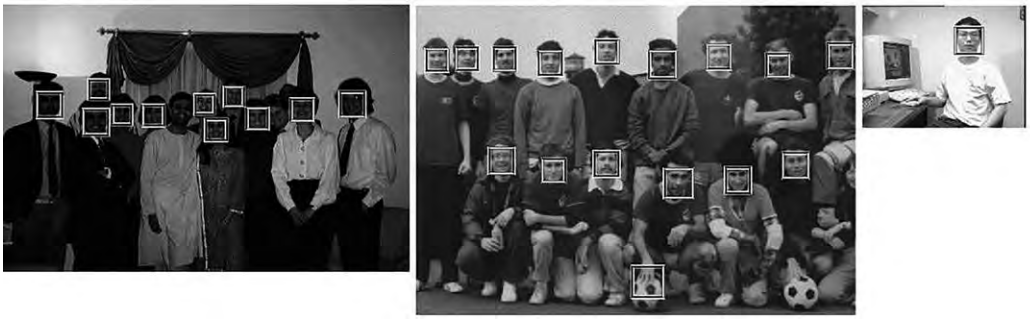
⑻⓪　複数の顔検出器を多段階に接続し、明らかに顔ではないサンプルを早期に棄却することで高速化を図る枠組み。

⑻⑴　Paul Viola and Michael Jones, “Rapid object detection using a boosted cascade of simple features,” *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)*, Vol.1, 2001, pp.511-518.

⑻⑵　Takeshi Mita et al., “Joint Haar-like features for face detection,” *Tenth IEEE International Conference on Computer Vision (ICCV 2005)*, Vol.1, 2005.

とから[83]、顔検出についても、特徴抽出と識別器の設計を同時に行う深層学習が用いられる例が増えている[84]。

図１１　ViolaとJonesによる顔検出手法の結果の例



（注）一部、検出漏れの顔（左）や、誤検出しているサッカーボール（中央）があるが、おおむね正しく顔を検出できていることが分かる。
（出典）Copyright 2001 IEEE. Reprinted, with permission, from Paul Viola and Michael Jones, “Rapid object detection using a boosted cascade of simple features,” *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001),* Vol.1, 2001, p.518. (Figure 8) <https://doi.org/10.1109/CVPR.2001.990517>

（４）顔の位置合わせ（正規化）

顔認識手法の中には、顔画像そのものを特徴として識別するものもあることから、顔器官や輪郭を含めて、事前に顔の位置合わせをすることが重要となる。そのために、顔器官や輪郭上で定義されたランドマーク点（例えば、口や目の端）を検出して、それらを標準的な顔モデルに合うように変形する。より具体的には、顔画像上に適当に初期配置したランドマーク点に対して、顔器官や輪郭のエッジ等の画像特徴への当てはまり具合を計算することで、徐々に顔器官や輪郭に収束するように動かしていくことが一般的である。

そのような枠組みとして、形状の変形を効率良く表現する Active Shape Model（ASM）[85] を用いた手法がある。ASM では、ランドマーク点の学習データを多数集めて主成分分析を適用し、平均顔と少数の基底ベクトルの線形和により顔の変形を表現する。これにより、各ランドマーク点を独立に動かす場合と比較して、初期配置のランドマーク点を顔器官や輪郭へ効率良く収束させることが可能となる（図 12）。また、形状に関するモデルである ASM に、アピアランス（テクスチャ情報）を付加した Active Appearance Model（AAM）[86] も提案されている。他にも、顔器官の端点を節点（ノード）としてそれらを結ぶことで構築するグラフを用いたグラフマッチングによる手法[87] 等が提案されている。

---

(83)　Joseph Redmon and Ali Farhadi, “YOLO9000: Better, Faster, Stronger,” *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2017, pp.6517-6525.

(84)　Sachin Sudhakar Farfade et al., “Multi-view Face Detection Using Deep Convolutional Neural Networks,” *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval (ICMR 2015),* 2015, pp.643-650; Xudong Suna et al., “Face detection using deep learning: An improved faster RCNN approach,” *Neurocomputing,* Vol.299, 2018.7, pp.42-50.

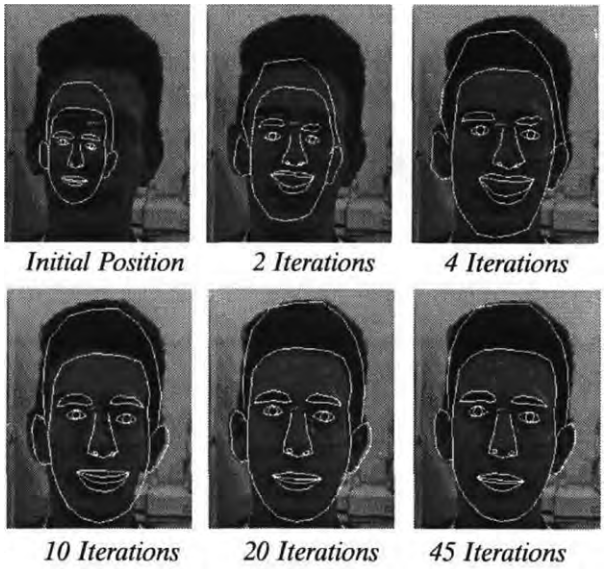
(85)　T.F. Cootes et al., “Active Shape Models: Their Training and Application,” *Computer Vision and Image Understanding,* Vol.61 No.1, 1995.1, pp.38-59.

(86)　T.F. Cootes et al., “Active Appearance Models,” Burkhardt H. and Neumann B. eds., *Computer Vision - ECCV’98: 5th European Conference on Computer Vision, Freiburg, Germany, June 2-6, 1998, Proceedings,* Vol.2, Berlin: Springer, 1998, pp.484-498.

(87)　Laurentz Wiskott et al., “Face Recognition by Elastic Bunch Graph Matching,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol.19 No.7, 1997.7, pp.775-779.

これらは２次元的な顔画像の位置合わせ手法となるが、顔向きの変化（横向きや上向き）に対する位置合わせ（正規化）を行うためには、３次元的な位置情報が必要となる。また、顔の３次元形状が分かると、顔の立体形状によって生じる陰影の付き方も分かることから、照明条件の正規化にも有効である。そのため、３次元的な顔の変形モデル[88]を用いて、顔向きの変化並びに照明変化への頑健性を実現している手法もある[89]。さらに、３次元顔形状へのマッピングを介して、顔の位置合わせと認識を同時かつリアルタイムに行う手法も提案されている[90]。

図１２　ＡＳＭによる初期配置からの顔器官や輪郭への当てはめ結果



（注）初期配置（左上）から当てはめを繰り返すことで、顔器官の正確な位置の把握に近づいていく。
（出典）Timothy F. Cootes et al., “Active Shape Models: Evaluation of a Multi-Resolution Method for Improving Image Search,” Edwin R. Hancock ed., *Proceedings of the British Machine Conference*, [Guildford]: BMVA Press, 1994, p.332. (Figure 5)

（５）特徴抽出・識別

顔認識における特徴抽出と識別は密接に関連していることから、本節ではそれらをまとめて概説する。大別すると、顔画像そのもの、ないしは顔画像全体から抽出された特徴を用いる大局的な手法と、顔器官周辺などの局所領域から抽出される特徴を用いる局所的な手法になる。ただし、局所的な手法[91]の利用は限定的であり、大局的な手法が主流となっている。

---

[88] V. Blanz and T. Vetter, “A morphable model for the synthesis of 3D faces,” *Proceedings of the 26th annual conference on Computer graphics and interactive techniques (SIGGRAPH '99)*, 2009, pp.187-194.

[89] Pascal Paysan et al., “A 3D Face Model for Pose and Illumination Invariant Face Recognition,” *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2009, pp.296-301.

[90] Yuki Oka and Takeshi Shakunaga, “Real-time Face Tracking and Recognition by Sparse Eigentracker Augmented by Associative Mapping to 3D Shape,” *Image and Vision Computing*, Vol.30 No.3, 2012.3, pp.147-158.

[91] 注目点の近傍の画素値の共起関係を表す高次局所自己相関（Higher-order Local Auto-Correlation: HLAC）を用いた手法（T. Kurita et al., “A face recognition method using higher order local autocorrelation and multivariate analysis,” *Proceedings:11th IAPR International Conference on Pattern Recognition*, Vol.2, 1992, pp.213-216.）、局所２値パターン（Local Binary Pattern: LBP）を用いた手法（Timo Ahonen et al., “Face Recognition with Local Binary Patterns,” Tomas Pajdla and Jiri Matas eds., *Computer Vision - ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part IV*, Berlin: Springer, 2004, pp.469-481.）、SIFT特徴量を用いた手法（Cong Geng and Xudong Jiang, “Face recognition using sift features,” *2009 16th IEEE International Conference on Image Processing (ICIP)*, 2009, pp.3313-3316.）などが提案されている。

大局的な手法として、1990 年代から 2000 年代前半にかけて盛んに研究されたのが、部分空間の解析に基づく手法[92]である。

また、2000 年代後半になると、少数の基底で顔画像を表現する代わりに、多数の基底（辞書）とそれに対する係数（重み）ベクトルを用意し、その係数ベクトルの非零要素の個数を小さく抑える疎性や、データ構造の本質的な部分が少ない次元で表現可能であるという低ランク性に基づく手法[93]が数多く提案されるようになった。

さらに、2010 年代になると、深層学習を用いた手法が主流となっている。特に、DeepFace と呼ばれる手法[94]は、実環境下で撮影された顔画像データベース Labeled Face in the Wild（LFW）に対して、約 97% の認証精度を達成しており、人の目視による認証精度が同じく約 97% であったことから、 人と同等程度の認証性能が得られたということで注目を集めた（図 13）。その後も、より複雑な構造を持つネットワーク構造、例えば、VGG[95]、GoogleNet[96]、ResNet[97]等を用いた顔認識手法が提案され、更なる認証精度向上へと繋がっている。

図１３　DeepFaceのネットワーク構造の概要

（出典）Copyright 2014 IEEE. Reprinted, with permission, from Yaniv Taigman et al., “DeepFace: Closing the Gap to Human-Level Performance in Face Verification,” *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, p.1704. (Figure 2) <https://doi.org/10.1109/CVPR.2014.220>

⑼2　代表例として、顔画像に対して主成分分析を適用することで、平均顔と少数の基底、すなわち、固有顔で表現する手法（Matthew A. Turk and Alex P. Pentland, “Face recognition using eigenfaces,” *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991, pp.586-591.）や、線形判別分析を適用することで生成される基底（フィッシャー顔とも呼ばれる。）として用いる手法（Peter N. Belhumeur et al., “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19 No.7, 1997.7, pp.711-720.）が挙げられる。上記が単一画像に対する識別を対象としているのに対して、一人の人物に対して、個人内変動を含む複数画像を登録（複数カメラによる多視点画像の取得や、動画像による時系列画像などが考えられる。）して、それらの画像セットから抽出される部分空間（基底）そのものを特徴として見なす部分空間法を用いた手法（Y. Ariki et al., “Extraction and recognition of facial regions by subspace method,” *ACCV 1995*, Vol.3, 1995, pp.733-742.）、入力側も複数画像とする相互部分空間法を用いた手法（Osamu Yamaguchi et al., “Face recognition using temporal image sequence,” *Proceedings of the Thirrd IEEE International Conference on Automatic Face and Gesture Recognition*, 1998.）も提案されている。

⑼3　特に部分隠蔽に対して頑健な顔認証手法の代表例として、各登録人物に対する疎性に加えて、サングラスやスカーフといった隠蔽によって生じるノイズについても疎性を考慮する手法（John Wright et al., “Robust Face Recognition via Sparse Representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.31 No.2, 2009.2, pp.210-227.）や、各登録人物に対する低ランク性とノイズに対する疎性を考慮する手法（Chih-Fan Chen et al., “Low-rank matrix recovery with structural incoherence for robust face recognition,” *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp.2618-2625.）などが提案されている。

⑼4　Yaniv Taigman et al., “DeepFace: Closing the Gap to Human-Level Performance in Face Verification,” *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp.1701-1708.

⑼5　Omkar M. Parkhi et al., “Deep Face Recognition,” Xianghua Xie et al., eds., *Proceedings of the British Machine Vision Conference (BMVC)*, BMVA Press, 2015, pp.41.1-41.12.

⑼6　Florian Schroff et al., “FaceNet: A unified embedding for face recognition and clustering,” *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp.815-823.

⑼7　X. Zhang et al., “Range Loss for Deep Face Recognition with Long-Tailed Training Data,” *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp.5419-5428.

３　データベース、精度評価、コンペティション

顔画像データベースは、学習や性能評価の面で重要な役割を果たす。顔画像データベースに必要な要素としては、被験者数の多さや共変量（個人内変動）の大きさやその種類の多様性が挙げられる。表4に主要な顔画像データベースの一覧を示す。初期の顔画像データベースとしては、被験者数が数百名までで、かつ、実験室環境で撮影されたものが大半であった。それに対して、実環境下で撮影された5,000人規模の顔画像データベースLFWが登場すると、実環境下での顔認識の研究が加速した。また、それ以来、数千から数十万人規模の被験者数、百万枚規模の画像を含む顔画像データベースも構築され、大量の学習データを必要とする深層学習とも相まって、顔認識の研究の流れが更に加速している。実際に、LFWに対する認証精度は、固有顔の約60%から、DeepFaceの約97%へと飛躍的に向上しており、VGG、GoogleNet、ResNet等の更に深いネットワーク構造を持つ手法によって、99%以上の精度が達成されており、すでに認証精度が飽和状態となっている。

また、顔認識のコンペティションも数多く開催されており、中でもNISTが開催しているFace Recognition Technology（FERET）やFace Recognition Vender Test（FRVT）系列のコンペティションは、1994年から現在に至るまで、数年おきに継続的に開催されている大規模なものであり、大学のみならず民間企業も数多く参画している。それ以外にも、各種データベースの公開元が主催するコンペティションが、コンピュータビジョン・パターン認識・生体認証に関する国際会議の併設ワークショップとして開催されるなどしている。

表４　顔画像データベース

名称	被験者数	画像枚数 / 人*	撮影条件・共変量
The Yale Face Database	15	11	照明・表情
PIE Database	68	多数	向き・照明・表情
AR Face Database	126	26	照明・表情
BANCA (Biometric Access Control for Networked and e-Commerce Applications) Database	208	12	-
XM2VTSDB (eXtended Multi Modal Verification for Teleservices and Security applications project DataBase)	295	8	-
FERET (Facial Recognition Technology) Database	1,209	2~	向き
NIST MID (Mugshot Identification Database)	1,573	2	向き（正面・横）
FG-Net (Face and Gesture Recognition Working group)	82	12.2	経年変化
Morph	13,618	4.1	経年変化
LFW (Labeled Faces in the Wild database)	5,000	1 / 2.3 / 530	Uncontrolled **
MS-Celeb (Microsoft Celebrity) 1M (Challenge 1)	100,000	100	Uncontrolled **
MS-Celeb 1M (Challenge 2)	20,000	1 / - / 100	Uncontrolled **
MS-Celeb 1M (Challenge 3)	80,000	-	Uncontrolled **
CASIA (Chinese Academy of Sciences, Institute of Automation) WebFace	10,575	2 / 46.8 / 804	Uncontrolled **
UMD (University of Maryland) Faces-Videos	3,107	-	Uncontrolled **
MegaFace	672,057	3 / 7 / 2,469	Uncontrolled **
IJB (IARPA Janus Benchmark) -A	500	11.4	Uncontrolled **
IJB -B	1,845	36	Uncontrolled **
VGG (Visual Geometry Group) Face	2,622	1,000	Uncontrolled **
VGG Face2	9,131	87 / 362.6 / 843	Uncontrolled **
CelebFaces+	10,177	19.9	Uncontrolled **
Google	10M+	50	Uncontrolled **
Facebook	4,000	800 / 1,100 / 1,200	Uncontrolled **

（注）＊スラッシュ区切りの複数数値は、最小 / 平均 / 最大を表す。
＊＊撮影条件・共変量における「Uncontrolled」は、撮影条件・共変量を制御していない条件下で収集されたデータベースであることを意味しており、様々な撮影条件・共変量が含まれているため、より挑戦的な課題となっていることを表す。

（出典）I. Kemelmacher-Shlizerman et al., “The MegaFace Benchmark: 1 Million Faces for Recognition at Scale,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp.4873-4882; “Databases.” Face Recognition Homepage <http://www.face-rec.org/databases/> 等を基に筆者作成。

４　残された課題と対策

大量の学習データと深層学習の導入により、顔認識の精度は飛躍的に向上しており、データベースによっては、人の目視による認識精度を上回る例も散見される。しかしながら、例えばLFW データベースでは、顔の向きが様々ではあるものの比較的正面に近いものが多く、また、解像度も比較的高いことから、実環境で真に観測されるような、大きな顔向きの変化や画像圧縮に伴うノイズを含む低解像度画像を対象とした精度評価が必要である。加えて、前段の顔検出や顔の位置合わせの精度が、後段の識別精度に大きく影響することから、検出・位置合わせ・識別を全体として最適化するような深層学習の枠組みについても、更に研究の余地があると思われる。

Ⅶ　歩容

１　技術の概要

歩容（又は、歩様）は、動物の歩行パターンを表す用語であり、生体運動学や制御工学の分野で古くから研究対象とされてきた。特に、四足歩行の動物の典型例として、馬の歩容が古くから研究されており、速さが増すにつれて、常足（Walk）・早足（Trot）・駆足（Cantor / Gallop）と変化し、四足の着地の順番等が変化する。一方で、二足歩行であるヒトの歩容は、基本的には左右の足を順について歩く（スキップを除く。）ことから、四足歩行の動物と比較すると歩容の変化は限られている。しかしながら、二足歩行であるヒトの歩容には、依然として多様な情報が含まれることが知られている。古くは、シェイクスピアの『テンペスト』の一節にも「女王陛下であらせられる偉大なるジュノーが来る、それは私が歩き方で彼女だと分かったからである。」（High'st Queen of state, Great Juno comes; I know her by her gait）(98) とあるように、多くの人が、遠方にいる知人をその歩容で識別できた経験を持つものと想像される。実際に、精神生理学の研究では、歩容から友人を認識することやその人の性別を識別することが可能であるということが報告されている (99)。また、点光源を人体の関節部に付けて表示するという、単純化された歩容の生体運動情報を提示した場合に、個人識別や感情の属性推定が可能であるということも報告されている (100)。このような歩き方の個性に基づく個人認証を歩容認証 (101) と呼び、行動的生体情報の１つとして認識されている。歩容認証は、顔認証が利用できないような遠方からの撮影・後方からの撮影や、顔をヘルメットや目出し帽で隠した場合にでも利用できることから、防犯カメラ映像に基づく新たな個人識別法 (102) として、犯罪捜査等への利用の期待も高まっている。

本章では、これまでに提案されている歩容認証手法を概説するとともに、性能評価や学習に用いる歩容データベース、歩容認証を困難にする要因について紹介し、今後の展望を述べる。

２　認証の流れと特徴表現

（１）歩容認証の流れ

歩容認証は、コンピュータビジョン (103) やパターン認識 (104) の技術によって実現するものであり、その特徴表現としては、人体モデルの当てはめにより関節角 (105) の系列といった歩容のパラメータを抽出するモデルベースの表現と、モデル当てはめを行わずに画像の見えの情報を直接的に利用する見えに基づく表現に分類される。本節では、歩容認証の流れを図 14 を参照

(98)　"All speeches (Lines) for Ceres in "Tempest"" Open Source Shakespeare website <https://www.opensourceshakespeare.org/views/plays/characters/charlines.php?CharID=Ceres&WorkID=tempest>

(99)　L.T. Kozlowski and J.E. Cutting, "Recognizing the sex of a walker from a dynamic point-light display," *Perception & Psychophysics*, Vol.21 No.6, 1977.11, pp.575-580.

(100)　N.F. Troje, "Decomposing biological motion: A framework for analysis and synthesis of human gait patterns," *Journal of Vision*, Vol.2 No.5, 2002.9, pp.371-387. <https://doi.org/10.1167/2.5.2>; S. Stevenage et al., "Visual Analysis of Gait as a Cue to Identity," *Applied Cognitive Psychology*, Vol.13 No.6, 1999.12, pp.513-526.

(101)　M.S. Nixon et al., *Human Identification Based on Gait,* New York: Springer, 2005; Y. Makihara et al., *Gait Recognition: Databases, Representations, and Applications*, John Wiley & Sons, Inc., 2015, pp.1-15.

(102)　「平成 26 年度警察白書　概要」p.13. 警察庁ウェブサイト <https://www.npa.go.jp/hakusyo/h26/youyakuban/youyakuban.pdf>

(103)　前掲注(57) を参照。

(104)　前掲注(58) を参照。

(105)　人の関節を構成する 2 つのリンクの成す角。例えば、膝の関節角は、太ももとすねの成す角となる。

しつつ概説する。

(a) **入力画像の取得**　多くの場合、単一のカラーカメラにより入力画像を取得する。高いセキュリティレベルを要する場所においては、人物の3次元的な歩容解析[106]を目的として、複数台の同期カメラや距離センサによって入力画像を取得することもある。

(b) **前処理**　対象人物の検出や追跡を行う。加えて、特に、後述の見えに基づく表現を用いる場合は、背景差分法等により、対象人物の領域（シルエット）を抽出する。

(c) **特徴抽出**　歩行周期を検出して、人体モデル当てはめや、人物領域の平均化処理[107]・周波数解析[108]等を行い、歩容特徴を抽出する。

(d) **識別**　抽出された特徴間の相違度を算出する。ユークリッド距離[109]による最近傍識別[110]から、主成分分析や判別分析による次元削減、サポートベクターマシン[111]による識別等、各種パターン認識技術を用いて個人を識別する。特に近年は、深層学習[112]による枠組みが主流になりつつあり、特徴抽出部分と合わせて学習する事例も見受けられる。

図１４　歩容認証の流れ（見えに基づく手法の場合）

（出典）筆者作成。

（2）モデルに基づく表現

人体は関節物体であることから、腕・脚・胴体といった体の部位をリンク[113]で表現し、それらを関節で接続したモデルを用いることが一般的である。また、人体モデルの当てはめ結果に基づいて、リンクの大きさといった静的な特徴や、関節角の系列といった動的な特徴を抽出して認識に用いる。

人体モデルの例としては、腰・膝・足下の関節点を質点[114]とした振り子モデル[115]や、フーリエ記述子[116]による関節角の動きモデル[117]が提案されている。また、人体のリンクの大き

⑽⑹　Gregory Shakhnarovich et al., “Integrated Face and Gait Recognition from Multiple Views,” *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001),* 2001, Vol.1, pp.439-446; Hozuma Nakajima et al., “Depth-based Gait Feature Representation,” *IPSJ Transactions on Computer Vision and Applications,* Vol.5, 2013, pp.94-98.

⑽⑺　人物領域の時系列画像に対して、画素ごとに時間方向に平均を計算する処理。

⑽⑻　人物領域の時系列画像に対して、画素ごとの時系列信号を抽出して、その周期や各周波数に対する強度を求める処理。

⑽⑼　多次元空間における距離尺度の1つであり、ピタゴラスの定理（三平方の定理）に基づいて計算される。3次元空間においては、実際の距離と同一である。

⑽⑽　前掲注⑹③を参照。

⑽⑾　前掲注⑹⑥を参照。

⑽⑿　前掲注⑶②を参照。

⑽⒀　隣り合う関節点同士を接続するもの。

⑽⒁　本来大きさを持つものを、大きさの無い点に全ての質量が集まっていると見なすもの。

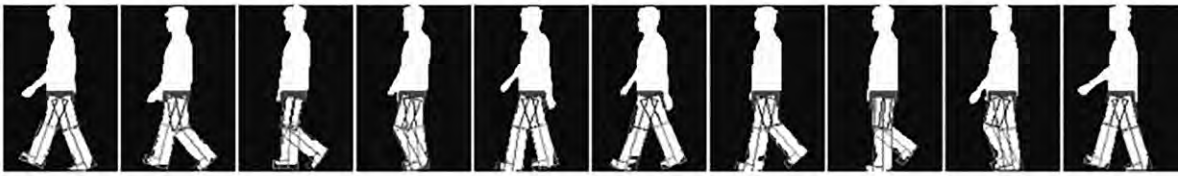
⑽⒂　ChewYean Yam et al., “Automated Person Recognition by Walking and Running via Model-based Approaches,” *Pattern Recognition,* Vol.37 No.5, 2004.5, pp.1057-1072.

⑽⒃　周波数解析手法であるフーリエ級数展開を適用した際に得られる各周波数成分に対する係数のこと。

⑽⒄　Imed Bouchrika and Mark S. Nixon, “Exploratory Factor Analysis of Gait Recognition,” *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition,* 2008.

さを考慮した方法として、楕円リンクによる表現[118]や、台形リンクによる表現[119]（図15）がある。これらはいずれも２次元モデルの当てはめであるのに対して、人体のリンクを３次元の楕円体により近似するモデル[120]も提案されている。さらに、人体のリンクを弾性体[121]として表現する人体モデル[122]も提案されている。これらモデルベースの特徴表現は、人物モデル当てはめに要する計算コスト（計算に要する時間やメモリ）が大きい等の問題があることから、後述の見えに基づく手法が主流となっている。

図１５　シルエット系列に対する台形モデル当てはめの例



（出典）Copyright 2010 IEEE. Reprinted, with permission, from Akira Tsuji, Yasushi Makihara and Yasushi Yagi, “Silhouette transformation based on walking speed for gait identification,” *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, p.719. <https://doi.org/10.1109/CVPR.2010.5540144>

（３）見えに基づく表現

初期の研究では、歩行者と背景の境界面の時空間解析[123]による特徴[124]や、オプティカルフロー[125]の空間分布解析[126]による特徴[127]が提案されている。その後、現在まで服装の色やテクスチャによる影響を受けないよう、人物のシルエットを用いた特徴表現が主流となっている。例えば、シルエット系列を特徴とする方法[128]や、固有空間[129]に投影したシルエット系列を用いる方法[130]が提案されている。また、シルエットの輪郭に着目した手法として、フーリエ記述子による表現[131]や、 シルエット中心から各輪郭点までの距離系列による表現[132]が提案されている。

118　L. Lee and W. E. L. Grimson, “Gait analysis for recognition and classification,” *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition (FGR 2002)*, 2002.

119　Akira Tsuji et al., “Silhouette Transformation based on Walking Speed for Gait Identification,” *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp.717-722.

120　Raquel Urtasun et al., “Priors for People Tracking from Small Training Sets,” *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)*, Vol.1, 2004.

121　伸び縮みしない剛体に対して、バネのように伸び縮みするもののこと。

122　Gunawan Ariyanto and Mark S. Nixon, “Marionette mass-spring model for 3D gait biometrics,” *2012 5th IAPR International Conference on Biometrics (ICB)*, 2012, pp.354-359.

123　時間１次元と画像空間２次元の合計３次元の空間を時空間と言い、その空間においてデータ解析をすること。

124　Sourabh A. Niyogi and Edward H. Adelson, “Analyzing and recognizing walking figures in XYT,” *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp.469-474.

125　画像中の各点の速度ベクトルのこと。

126　画像空間において、対象となるデータがどのように分布しているかを解析すること。

127　James J. Little and Jeffery E. Boyd, “Recognizing People by Their Gait: The Shape of Motion,” *Videre: Journal of Computer Vision Research*, Vol.1 No.2, 1998, pp.1-32.

128　Sudeep Sarkar et al.,“The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.27 No.2, 2005.2, pp.162-177.

129　元の特徴空間に対して、データの分布を考慮して少数の基底を用いて表現した低次元空間のこと。

130　Hiroshi Murase and Rie Sakai, “Moving Object Recognition in Eigenspace Representation: Gait Analysis and Lip Reading,” *Pattern Recognition Letters*, Vol.17 No.2, 1996.2, pp.155-162.

131　Stuart D. Mowbray and Mark S. Nixon, “Automatic Gait Recognition via Fourier Descriptors of Deformable Objects,” *Proceedings of the 1st IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2003, pp.566-573.

132　Liang Wang et al., “Silhouette analysis-based gait recognition for human identification,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.25 No.12, 2003.12, pp.1505-1518.

これらは、時系列情報を直接照合する手法であるのに対して、時系列情報を圧縮した特徴表現も提案されている。代表的な特徴が、左右 2 歩分から成る 1 歩行周期でシルエットを平均化した歩容エネルギー画像（Gait Energy Image: GEI　又は平均シルエットと呼ばれる。）[133] である（図 16 (a)）。歩容エネルギー画像においては、一歩行周期を通して人物領域が占めている画素が白、背景領域が占めている画素が黒で表現され、これらより主に体型成分を見て取ることができる。一方、一歩行周期中に人物領域と背景領域の両方が現れる画素は中間色である灰色で表現され、そこから動き成分を見て取ることができる。これらは極めて単純な表現方法ではあるものの、その実装の容易さ、計算コストの低さ、認証精度の高さなどの点で、歩容認証の研究において最も幅広く用いられている特徴表現となっている。また、その派生系として、フーリエ解析に基づく周波数領域特徴（Frequency-Domain Feature: FDF）[134]（図 16(c)）、シルエット輪郭に歩容位相の情報を付加した時間歩容画像（Chrono-Gait Image: CGI）[135]（図 16 (d)）等が提案されている。また、GEI に対する変換処理を施す手法も提案されており、動き成分を強調する歩容エントロピー画像（Gait Entropy Image: GEnI）[136]（図 16 (b)）、動きの少ない部分をマスクした GEI [137] 等が提案されている。

図１６　見えに基づく歩容特徴の例

(a) GEI　(b) GEnI　(c) FDF　(d) CGI

（出典）Reprinted with permission from Y. Makihara, D. S. Matovski, M. S. Nixon, J. N. Carter and Y. Yagi, “Gait Recognition: Databases, Representations, and Applications,” J. G. Webster ed., *Wiley Encyclopedia of Electrical and Electronics Engineering,* John Wiley & Sons, Inc., 2015. <https://doi.org/10.1002/047134608X.W8261>

これらの見えに基づく特徴表現は、計算コストが低く、幅広く用いられている一方、観測方向や服装変化等の条件変化による影響を受けやすいという問題点もある。よって、条件変化への頑健性を高めるため、識別時には機械学習手法と合わせて用いられることが多い。

３　データベース、精度評価、コンペティション

歩容データベースは、学習や性能評価の面で重要な役割を果たす。歩容データベースに必要な要素としては、被験者数の多さや共変量（個人内変動）の大きさやその種類の多様性が挙げ

[133] S. Sarkar et.al., *op.cit.*[128]; Ju Han and Bir Bhanu, “Individual Recognition Using Gait Energy Image,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.28 No.2, 2006.2, pp.316-322.

[134] Yasushi Makihara et al., “Gait Recognition Using a View Transformation Model in the Frequency Domain,” *Computer Vision - ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006, Proceedings, Part III*, Berlin: Springer, 2006, pp.151-163.

[135] Chen Wang et al., “Human Identification Using Temporal Information Preserving Gait Template,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.34 No.11, 2012.11, pp.2164-2176.

[136] Khalid Bashir et al., “Gait recognition using gait entropy image,” *3rd International Conference on Imaging for Crime Detection and Prevention (ICDP 2009)*, 2009.

[137] Khalid Bashir et al., “Gait recognition without subject cooperation,” *Pattern Recognition Letters*, Vol.31 No.13, 2010.10, pp.2052-2060.

られる。表5に主要な歩容データベースの一覧を示す。静止画である顔や指紋等の他の生体情報と比較すると、動画を収集する必要がある歩容データベースは、その収集に手間がかかることから、100人程度の比較的小規模なデータベースが多い。しかしながら、近年は被験者数が60,000人を超える大規模歩容データベースも公開されており、大量の学習データが必要とされる深層学習の枠組みも十分に適用可能な状況となっている。実際に、これらの大規模歩容データベースを深層学習の学習データに用いた歩容認証手法により、観測方向が同じ場合と最も観測方向が異なる場合（90度の観測方向差）とで、本人認証における等誤り率がそれぞれ1.1%と4.6%、約5,000人の登録者に対する個人識別におけるRank-1認証率がそれぞれ89.5%と17.3%であると報告されている[138]。一方で、歩容認証は他の生体認証に対して比較的新しい技術であることから、コンペティションが実施された例はない。しかし、近年、歩容データベースの整備と共に、歩容認証に特化したワークショップ[139]やチュートリアル[140]が開催されていることから、コンペティションについても開催の機運は高まっている。

---

(138) Noriko Takemura et al., "On Input/Output Architectures for Convolutional Neural Network-Based Cross-View Gait Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, 2017.

(139) "ACCV2014 Workshop, Human Gait and Action Analysis in the Wild: Challenges and Applications." Associate Professor Jian Zhang at University of Technology Sydney website <http://www.multimediauts.org/ACCV2014WSGaitAction/>

(140) "ECCV 2016 Tutorial, Human Identification at a Distance by Gait and Face Analysis." Watrix website <http://www.watrix.ai/2018/05/02/human-identification-at-a-distance-by-gait-and-face-analysis-a-tutorial-for-eccv2018/>

表５　歩容データベース

名称	被験者数	シーケンス数	共変量（視点以外）の有無	視点数	屋内（I）/ 屋外（O）
CMU MoBo (Carnegie Mellon University, Motion and Body dataset)	25	600	有	6	I (Treadmill) *
Georgia Tech	15	268	有	-	O
	18	20	有	-	-
HID-UMD (Human Identification Dataset, University of Maryland)	25	100	無	1	O
	55	222	有	2	O
SOTON (University of Southampton) Small Database	12	-	有	3	I
SOTON Large Database	115	2,128	有	2	I/O
SOTON Multimodal	>300	>5,000	有	12	I
SOTON Temporal	25	2,280	有	12	I
USF (University of South California) HumanID	122	1,870	有	2	O
CASIA (Chinese Academy of Science, Institute of Automation) A	20	240	有	3	I
CASIA B	124	1,240	有	11	I
CASIA C	153	1,530	有	1	O
OU-ISIR (Osaka University, the Institute of Scientific and Industrial Research), Treadmill A	34	612	有	1	I (Treadmill) *
OU-ISIR, Treadmill B	68	2,764	有	1	I (Treadmill) *
OU-ISIR, Treadmill C	200	200	有	25	I (Treadmill) *
OU-ISIR, Treadmill D	185	370	無	1	I (Treadmill) *
OU-ISIR, LP (Large Population)	4,007	7,842	無	2	I
OU-ISIR, LP-Age	63,846	63,846	有	1	I
OU-ISIR, LP-Bag	62,528	178,018	有	1	I
OU-ISIR, MV (Multi-View) LP	10,307	277,358	無	14	I
TUM (Technical University of Munich) - IITKGP (Indian Institue of Technology Kharagpur)	35	850	有	1	O
TUM - GAID (Gait from Audio, Image and Depth)	305	3,370	有	1	O
WOSG (WVU (West Virginia University) Outdoor SWIR (Short-wave infrared) Gait)	155	684	有	8	O

（注）* 歩行器上での歩行データであることを表す。

（出典）Y. Makihara et al., *Gait Recognition: Databases, Representations, and Applications*, John Wiley & Sons, Inc., 2015, pp.1-15; 槇原靖ほか「大規模歩行映像データベースの構築とその歩行映像解析への応用」『画像ラボ』Vol. 30 No. 3, 2019.3, pp.11-17 等を基に筆者作成。

４　残された課題と対策

歩容認証は、防犯カメラ映像による犯罪捜査支援が主な応用先として考えられていることから、制約なく無意識に歩く人物を対象とすることが想定される。そのため、防犯カメラの取り付け角度や対象人物の歩行方向変化による観測方向変化を始めとして、服装変化、靴の変化、荷物変化、歩行速度変化、経時変化など、歩容認証を困難にする要因が存在する。これらの問題に対しては、特徴表現を工夫することでこれらの変動要因に対する不変性を獲得する研究[141]や、変動要因を含む学習データを与えることで、学習の段階で変動要因に対する頑健性を獲得

[141] Chi Xu et al., “Speed Invariance vs. Stability: Cross-Speed Gait Recognition Using Single-Support Gait Energy Image,” Shang-Hong Lai et al., eds., *Computer Vision – ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part II*, Springer, 2016, pp.52-67.

する研究[142]等が挙げられる。特に、近年の深層学習による手法により、これらの変動要因に対する頑健性が格段に向上している。大規模歩容データベースが利用可能となり、深層学習に基づく手法の研究開発が加速していくことで、歩容認証の適用可能範囲はますます加速していくことが予想される。現状では、世界各国の科学捜査の場面における歩容認証の鑑定利用[143]が進められているが、今後は、カメラ渡りの人物追跡[144]や自動での指名手配犯の検出など、更なる適用範囲の拡大が期待される。

Ⅷ　行動

１　技術の概要

行動的な特性・特徴を用いた認証技術としては、前述した歩容認証のほかに、例えば発話に着目した声紋認証、筆記動作に着目した筆者認証、キーボードを利用する動作に着目したキーストローク認証などがある。例えば、顔が見えなくても、我々は知人であれば声から、その人物が誰なのかを判断することができる。また、文字を見れば、その文字を書いた人物が誰なのかを推定することができる。こういった発話や筆記を対象とした認証技術が声紋認証であり、筆者認証である。また、キーボードを利用して入力する動作にも個人の違いがみられ、この個人性を利用するのがキーストローク認証である。これらは、発声する、筆記する、キーボードで入力する、という行動に基づき生成されるデータを用いた認証手法であるため、行動的特徴を用いた手法と分類される[145]。

発話に着目した手法や筆記動作に注目した手法は、テキスト依存型、テキスト独立型、テキスト提示型、などに分類される。テキスト依存型は、事前に決められた同じ発話や筆記をしてもらい認証を行うもので、発話を用いた認証であれば、例えば「OK Google」[146]という決まったフレーズを言い、そのフレーズを発した人物を認証するものである。テキスト依存の方法では、全員が同じフレーズを使うものもあれば、個人ごとに異なるフレーズを使う場合もある。テキスト独立型は、発話内容を固定しない方法であり、テキスト提示型は、認証システム側から発話内容を指示するものである。

筆記動作に基づくテキスト依存の手法は署名認証がこれに該当する。署名認証は、自分で設定をした署名を筆記することで、自分であることを示す技術であり、他のモダリティと比較して、筆者の意志を考慮でき、筆記内容を変更することで、特性・特徴が変更できるという特長がある。また、最近はスマホやタブレットデバイスが普及したことにより、指先で筆記する方法も考えられる。一方、キーストロークは、キーボードを打つタイミングを利用して認証を行

142　Raul Martin-Felez and Tao Xiang, “Uncooperative gait recognition by learning to rank,” *Pattern Recognition*, Vol.47 No.12, 2014.12, pp.3793-3806.

143　Imed Bouchrika et al., “On Using Gait in Forensic Biometrics,” *Journal of Forensic Sciences*, Vol.56 No.4, 2011.7, pp.882-889; N. Lynnerup and P.K. Larsen, “Gait as evidence,” *IET Biometrics*, Vol.3 No.2, 2014.6, pp.47-54; Haruyuki Iwama et al., “Gait Verification System for Criminal Investigation,” *IPSJ Transactions on Computer Vision and Applications*, Vol.5, 2013, pp.163-175.

144　視野を共有しない複数のカメラ間で同一人物を追跡すること。防犯カメラ映像による犯人や被疑者の足取り捜査における重要なステップであり、現在は、主に捜査員の目視確認によって行われている。

145　声は発声器官を通して作られるため、声の個人性には発声器官の違いが反映される。そのため、声紋認証は身体的と行動的両方の性質を有している。

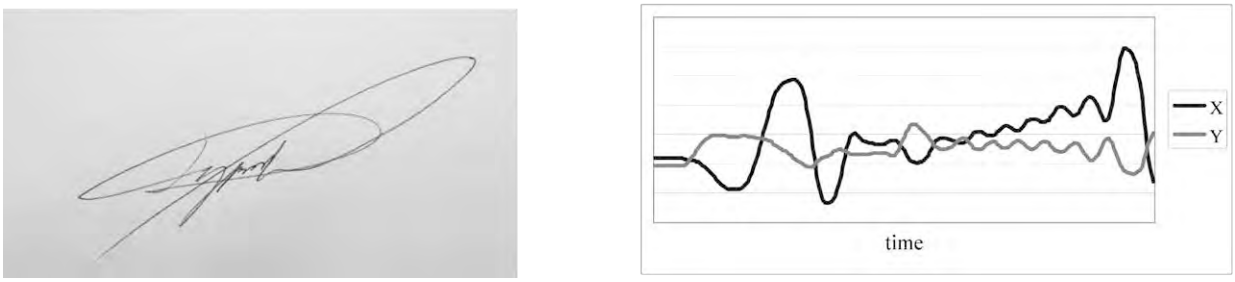
146　Georg Heigold et al., “End-to-end text-dependent speaker verification,” *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp.5115-5119.

うものであり、パソコン等でタイピングを行っている間、継続的に人物を認証できるという特長を持っている。

２　認証の流れと特徴表現

声に基づく手法は、発声をマイクにより取得し、音の信号として処理をする。近年は動画撮影が一般的になってきているため、音声だけでなく、動画として映像と音声が一緒になったデータが対象になることもある。実際、動画データベースも構築されている[147]。マイクなどによって取得された音声信号（図 6(n)）から、個人性が現れる特徴を抽出し、その特徴を比較することでスコア化し、比較するのが声紋認証である。処理においては、音声信号から、時間ごとの周波数特性を解析し、それを特徴として利用する[148]。筆記動作に基づく手法はオフラインとオンラインの 2 つの手法がある。オフラインは筆記などの動作が完了した後に、カメラやスキャナなどを用いて、筆記内容を 2 次元のデータとして取得し処理する。一方オンラインでは、筆記動作を行っている間にデータを取得するものであり、 例えばタブレットとペンにより、ペンの軌跡情報が取得できるため、筆記動作を時系列データとして取得し処理する（図 17）。以下では主にオンライン署名認証について説明する。

図１７　オフライン署名（左）とオンライン署名（右）の違い



（注）グラフは、署名を書き進めるに伴う、ペンの左右（X）及び上下（Y）方向の位置変化を示す。
（出典）筆者作成。

オンライン署名認証では、取得される情報は例えば筆記の際に用いるペン先の移動軌跡や筆圧である。オンライン署名認証の難しさは、筆記という行動に伴うデータが対象となるため、同一人物のデータであってもばらつきが生じ、また他人がその署名を真似するなりすまし攻撃[149]も考えられることにある。つまり、本人分布はある広がりを持ったものになるのに対し、他人分布は、なりすましにより本人分布に近づいてしまう。したがって、特徴抽出においては、本人内では安定している特徴で、かつ他人が真似しにくい特徴を抽出し、利用することが重要である。

一方で、キーストローク認証は、人がキーボードを用いてデータを入力する際のキーを打つ行動に注目したもので、入力動作中の各キーを押すタイミングや離すタイミングを記録し、これらから計算される時間特徴を認証に用いる。

⑿⑺　VoxCeleb A large scale audio-visual dataset of human speech <http://www.robots.ox.ac.uk/~vgg/data/voxceleb/>
⑿⑻　メル周波数ケプストラム係数（Mel-Frequency Cepstrum Coefficients: MFCC）などが用いられる。
⑿⑼　本人が筆記した署名は「真筆」、なりすましを想定した署名は「偽筆」と呼ばれる。なりすましの強さは、なりすましを行う人物が持っている情報や練習度合い、スキルに依存するため、署名認証では偽筆をなりすましの強さに応じて分類している。

3　データベース、精度評価、コンペティション

オンライン署名のコンペティションは2004年に初めてSVC2004が開催され、その後BMEC2007、BSEC2009、SigComp2009、ESRA2011、SigComp2011、SigWiComp2013、SigWIcomp2015などが開催されている。また、オンライン署名認証の主なデータベースは表6のとおりである。

表6　オンライン署名のデータベース

名称	取得デバイス	被験者数	真筆取得セッション数*	各署名当たりの署名（筆記）数	
				真筆	偽筆
Caltec vision Gr. 1	Camera	56	3	25	10
Caltec vision Gr. 2	Camera	50	3	30	10
BIOMET	Pen tablet	84	3	15	12
MCYT	Pen tablet	330	1	25	25
SVC2004 Task1	Pen tablet	40	2	20	20
SVC2004 Task2	Pen tablet	40	2	20	20
SUSIG Blind	Pen tablet	100	1	8 or 10	10
SUSIG Visual	LCD display tablet	100	2	20	10
BioSecurID	Pen tablet	400	4	16	16
BioSecure DS2 **	Pen tablet	672	2	30	20
BioSecure DS3	PDA	713	2	30	20
SigComp2009***	Pen tablet	5/100	1	12/12	31/24
SigComp2011 C ***	Pen tablet	10/10	-	平均23/（平均9+ 平均12）	12/平均11
SigComp2011 D ***	Pen tablet	10/54	-	24/（12+12）	12/平均11
SigWiComp2013***	Tablet PC	11/20	3	42	36

（注）* セッション数が複数の場合は、複数の日にまたがり真筆を収集している。
** 有償である。
*** 学習用データと評価用データを明確に分けて提供しているものについては、表中で（学習用／評価用）と記載している。例えば、SigComp2009では、学習用被験者数5、評価用被験者数100なので5/100と記載しており、SigComp2011 Dでは、真筆には学習用として24個の署名、評価用には12個の参照データと12個の評価用真筆が提供されているため24/(12+12)、偽筆には学習用12個、評価用12個が提供されているため12/12と記載している。参照データとは、事前に認証システムに、該当被験者の署名として登録されるデータ（真筆）であり、認証のために入力された署名は、この登録データに基づきスコアが計算され、判定が行われる。なお、SigComp2011の署名数は被験者によってその個数にばらつきがあり、特にSigComp2011 C（Chinese）はそのばらつきが非常に大きいため、平均値を記している。

（出典）“Research: Data Sets.” Computational Vision at Caltech website <http://www.vision.caltech.edu/mariomu/research.html>; Sonia Garcia-Salicetti et al., “BIOMET: A Multimodal Person Authentication Database Including Face, Voice, Fingerprint, Hand and Signature Modalities,” *Audio- and Video-Based Biometric Person Authentication: 4th International Conference, AVBPA 2003 Guildford, UK, June 9–11, 2003 Proceedings*, New York: Springer, 2003, pp.845-853; “MCYT-Signature-100 database.” Biometrics and Data Pattern Analytics Lab at Universidad Autónoma de Madrid website <https://atvs.ii.uam.es/atvs/mcyt100s.html>; “Download.” First International Signature Verification Competition (SVC 2004) website <https://www.cse.ust.hk/svc2004/download.html>; “SUSIG: An On-line Handwritten Signature Database, Associated Protocols and Benchmark Results.” Sabanci University Biometrics Research Group website <https://biometrics.sabanciuniv.edu/susig.html>; “BiosecurID-SGlobalLocalFeat DB.” Biometrics and Data Pattern Analytics Lab at Universidad Autónoma de Madrid website <https://biosecure.wp.tem-tsp.eu/biosecure-database>; “BioSecure Database.” Association BioSecure website <https://biosecure.wp.tem-tsp.eu/biosecure-database>; “ICDAR 2009 Signature Verification Competition (SigComp2009).” International Association for Pattern Recognition Technical Committee Number 11 (IAPR-TC11) website <http://www.iapr-tc11.org/mediawiki/index.php/ICDAR_2009_Signature_Verification_Competition_(SigComp2009)>; “ICDAR 2011 Signature Verification Competition” IAPR-TC11 website <http://www.iapr-tc11.org/mediawiki/index.php/ICDAR_2011_Signature_Verification_Competition_(SigComp2011)>; “Signature Verification and Writer Identification Competitions for On- and Offline Skilled Forgeries.（SigWiComp2013）” IAPR-TC11 website <http://tc11.cvc.uab.es/datasets/SigWiComp2013_1> 等を基に筆者作成。

４　残された課題と対策

行動に基づく認証手法の大きな課題は、経時変化の問題である。様々なモダリティにおいて、経時変化対策は重要な課題であるが、行動に基づく手法の場合には、行動の癖が時間とともに変わる可能性が高いため、他のモダリティよりも時間経過が精度に与える影響が大きい。声紋認証については、他のモダリティと同様になりすまし攻撃への対策が必要である。近年、音声合成や動画合成技術が著しく発展しているが、それらの技術を用いたなりすまし攻撃は非常に大きな脅威となる。署名認証やキーストロークといった行動に基づく手法は、個人によりその精度が大きく変わってしまうという問題が存在する。個人による精度差の低減は今後も解決すべき重要な課題である。また、行動の真似に対する耐性も重要となる。

執筆：大阪大学産業科学研究所　准教授　村松（むらまつ）　大吾（だいご）（Ⅰ、Ⅱ、Ⅳ、Ⅴ、Ⅷ）
大阪大学データビリティフロンティア機構　特任講師　山本（やまもと）　奈津子（なつこ）（Ⅲ）
大阪大学産業科学研究所　　准教授　槇原（まきはら）　靖（やすし）（Ⅵ、Ⅶ）