6. バイオメトリクス認証技術における精度評価の動向

（株）日立製作所　システム開発研究所
瀬戸 洋一
三村 昌弘

オープンネットワークを前提とした電子的商取引が立ち上がりつつある．この場合非対面であるため個人認証技術が重要となる．個人（本人）認証技術としてバイオメトリクス（生体情報）の利用が注目されている．バイオメトリクス認証装置の性能は，利用環境条件，処理パラメータなどにより統計的に決まるため，評価条件が公開されていない現在，各社から提供される性能値を同等に比較できない．このため，導入時にユーザに混乱を与えるおそれがある．健全な市場形成のためには，精度評価の標準化が必要である．本稿では，タイプ1エラー（本人拒否率），タイプ2エラー（他人受入率）など認証精度の考え方，および米国，EUにおける標準化動向などを交え精度評価の最新の動向を紹介する．

❑ バイオメトリクス認証技術の普及のために必要なこと ❑

本特集で紹介された指紋，虹彩，署名，声紋などのバイオメトリクス（biometrics）認証技術は，各社で製品化されているが，1997年時点の国内市場は約15億円程度であり，また，海外においても150億円程度の市場規模である．適用先も，施設管理（入退出管理）および犯罪捜査という分野に限定されている．ただし，今後さらに認証装置の低価格化，小型化および高精度化が進むと，グループウェア，シングルサインオンなどの情報リソースへのアクセス管理，電子商取引EC（Electronic Commerce）など本格的な本人認証への展開が可能となり，2005年に認証装置の市場規模は現在の倍になるという予測もある[1), 2)]．

バイオメトリクスを用いた本人認証装置を導入するにあたり，よくある質問は，(1) どの程度の精度か？（技術的），(2) 投資コストとセキュリティ上の利点は？（ビジネス的），(3) 社会および顧客に受け入れられるか？（社会的）などである[3)]．後者の2つの質問はベンダおよびユーザがともに考えねばならない問題と考える．本稿では，技術的な質問に関する精度評価方法を紹介する．

現在，何社からかバイオメトリクス認証装置が販売されているが，精度評価条件が開示されていないため，同種の認証装置を製品化するA社とB社の装置を正しく比較できない問題がある．また，精度評価条件によって精度は大きく変化する可能性がある．たとえば，メーカのカタログ値と公的な機関で評価した結果とでは，メーカ提示値の方が2桁以上よい精度が示されているという報告もある[4)]．これらは，ユーザあるいはシステムインテグレータが認証装置を導入する際の障害になっている．このため，健全な市場を形成するためには認証装置の精度評価の標準化が急務である．

以上を背景に本稿では，精度評価における基本的な考え方，国内外で進む標準化の動向，および米国National Biometrics Test Centerで採用されている精度評価方法について紹介する．

❑ バイオメトリクス認証技術における精度の定義 ❑

認証モデル

認証における精度について，本人認証の代表的な方法であるパスワードとの比較で問題点を述べる．**図-1**に示すように，(a) パスワードモデルにおける認証は，キーボードからのデータの入力と事前に登録したパスワードの文字（数）列との比較により行う．パスワードモデルにおける誤差要因としては，入力時における勘違いやタイプミスがある．判定は入力されたデータと蓄積パスワードとの

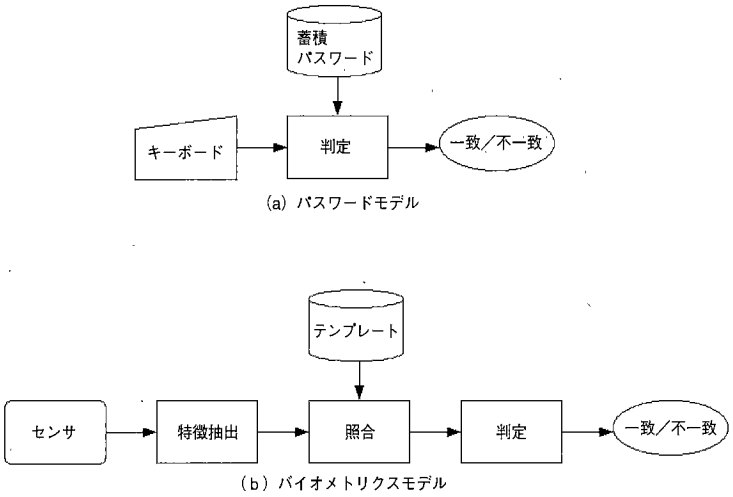

図-1　認証モデルの例

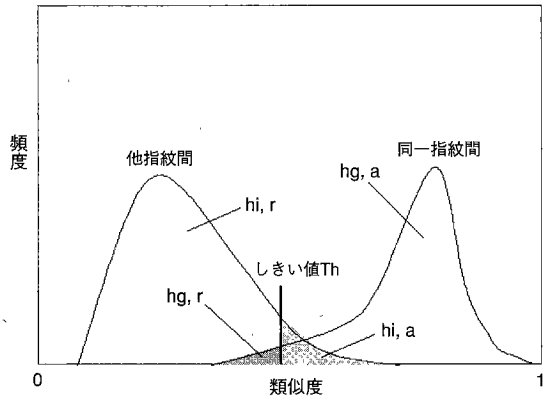

図-2　2つの誤差（本人拒否および他人受入誤差）

文字列判定で行われる．したがって誤差は，いくつかの文字が一致しない場合に生じる確定的なものである．

一方，(b) バイオメトリクスモデルにおける認証は，センサからのデータ入力，特徴抽出などの前処理の後，事前に登録しておいたバイオメトリクスデータ（テンプレートデータという）との照合処理により類似度を算出する．類似度とは入力データがテンプレートデータにどれだけ似ているかを表す特徴空間での尺度である．類似度が，事前に設定したしきい値以上の場合は一致，以下の場合は不一致と判定する．

バイオメトリクスによる認証は，1次元（たとえば声紋）あるいは2次元データ（たとえば指紋）の入力データに対するパターンマッチング処理が基本であり，これに起因する統計的な誤差が生じる．たとえば入力装置において，入力における環境条件，つまり，人間の身体的（たとえば，指の湿気具合）もしくは行動的な変化（たとえば，風邪をひいた時の声質の変化），特徴抽出においては，入力データに対するアルゴリズム対応性（たとえば声紋において，どの程度の周辺ノイズに対応可能か）に起因する誤差，照合判定においては，設定するしきい値により，たとえ同一人物が入力した場合でも，結果が同じになることは保証できない．

精度の表現

認証装置の精度は，どのようなバイオメトリクスを用いても，一般的に次のような有意性検定法により定義できる[3), 4)]．

指紋による本人認証装置の精度を例に述べる．**図-2**の横軸は照合処理における類似度，縦軸は頻度を表す．ここでは指紋による本人認証の例を用いた．2つの分布は，それぞれ同一のデータを照合した場合と，異なるデータを照合した場合の類似度分布を示す．類似度は右にいくほど大きくなる．これは，比較する2つのバイオメトリクス特徴量が一致している度合いが増えることを意味する．

2つの類似度分布曲線が重ならず，しきい値を重なりのないところに設定すれば，原理的に誤差はゼロになるが，現実には重なり合うことが多い．このため認証誤差が生じる．

統計的に独立に採取された指紋において，同一指紋同士を照合した場合の類似度分布をhg，異なる指紋同士を照合した場合の類似度分布をhiとし，しきい値Thによる判定処理を行うと，以下の4つのケースが存在する．

(1) hgが受け入れられる場合，その分布をhg, aとする．
(2) hgが却下される場合，その分布をhg, rとする．
(3) hiが却下される場合，その分布をhi, rとする．
(4) hiが受け入れられる場合，その分布をhi, aとする．

同一指紋同士を照合し，類似度分布hgが却下される場

表-1　代表的な海外の精度標準化機関

機関	NBTC	ICSA	BioTest
組織の概要	・米国国務省設立のBiometric Consortiumの精度評価機関 ・1997年よりSanJose州立大で運用	・バイオメトリクスを含むセキュリティ製品の認定を行う組織 ・1989年より運用	・欧州ESPRIT出資の精度評価プロジェクト ・1996年から1998に推進
活動内容	・政府，ベンダの依頼で精度評価を行う ・本人認証装置の評価 米国，カナダ，フィリピンのAFIS評価で実績あり ・標準指紋データをCD-ROMで配布，精度評価の標準化を推進	・ベンダの依頼により製品認定を行う．精度評数値は非公開 ・Miros社（顔），MYTEC社（指紋）などの製品認定実績あり	・精度評価手法の開発を目的とする実験プロジェクト ・活動内容は非公開 ・指紋，掌形などを評価
評価の方法	・照合アルゴリズムと判定ポリシーを区別，また，統計的信頼性を考慮し評価サンプル数を決定	・判定ポリシー，経時変化を考慮した評価を行う	・非公開のため詳細不明．NBTCと同様の精度評価を実施した模様

NBTC: National Biometric Test Center
ICSA: International Computer Security Association

合，分布hg, rに相当する値を本人拒否誤差（あるいは本人拒否率）FRR（False Reject Rate）と呼ぶ．また，異なる指紋同士を照合し，類似度分布hiが受理される場合，分布hi, aに相当する値を他人受入誤差（あるいは他人受入率）FAR（False Accept Rate）と呼ぶ．本人拒否率FRRは有意性検定におけるタイプ1エラー，他人受入率FARはタイプ2エラーに相当する．

{h}を分布hの頻度の積分値とすると，それぞれの誤差は，次式で表せる．

$$FRR = \{hg, r\} / (\{h\ g, r\} + \{hg, a\}) \quad (1)$$

$$FAR = \{hi, a\} / (\{hi, r\} + \{hi, a\}) \quad (2)$$

タイプ1エラー（本人拒否率）が高いと利用者はフラストレーションを引き起こし，タイプ2エラー（他人受入率）が高くなると詐欺を引き起こす．タイプ2エラーはタイプ1エラーに比べ1桁から2桁小さくするのが一般的である．この2つのエラーのトレードオフを調整することが重要であり，運用上のノウハウとなる．

客観的な精度を求めることはやさしいことではなく，欧米では精度評価の標準化を目的としたテストセンタを設立している．本件に関しては次章で述べる．

□ 各国における精度評価標準化への取り組み □

欧米では標準化プロジェクトおよび精度評価を行うテストセンタの設立が盛んである．標準化組織として一番有名なものは，バイオメトリクコンソーシアム（Biometric Consortium）である（http://www.biometrics.org）．米国政府関係機関で利用するバイオメトリクス技術の標準化を目的に，国防総省が1992年に設立した．具体的にはバイオメトリクス技術の改善，テストセンタの設立，政府，産業界および学界間の情報交換，バイオメトリクス技術の安全性，法律上および倫理上の問題への取り組みなど多方面に渡っている．

表-1に代表的な精度評価機関を示す．National Biometric Test Center（NBTC）はSan Jose州立大学（センタ長はWayman教授）に設置され，精度評価手法の開発および装置に対する評価を行っている．NBTCの活動例として，フィリピン政府が導入するAutomatic Fingerprint Identification System（AFIS）の性能評価を行った実績がある．また，最近の情報によれば，NBTCは評価用指紋データをCD-ROMなどにて配布している[5]（http://www.engr.sjsu.edu/biometrics/）．

米国のInternational Computer Security Association（ICSA）はセキュリティ製品の認定を行っている（http://www.icsa.net/）．ICSAは評価手順（ICSA Commercial Biometric Certification Program V2.0）を策定し，認定を依頼された製品を評価している．本人拒否率および他人受入率が既定値以内であればICSA準拠製品として認定する．認定の合否のみの公表であり精度値は公開していない．運用を考慮した評価を行っていることから，ユーザの立場から製品を認定しているといえる．Miros社（顔），MYTEC社（指紋），iNTELiTRAK社（声）などの製品認定の実績がある．

欧州では1996年よりBioTestプロジェクトにおいて，

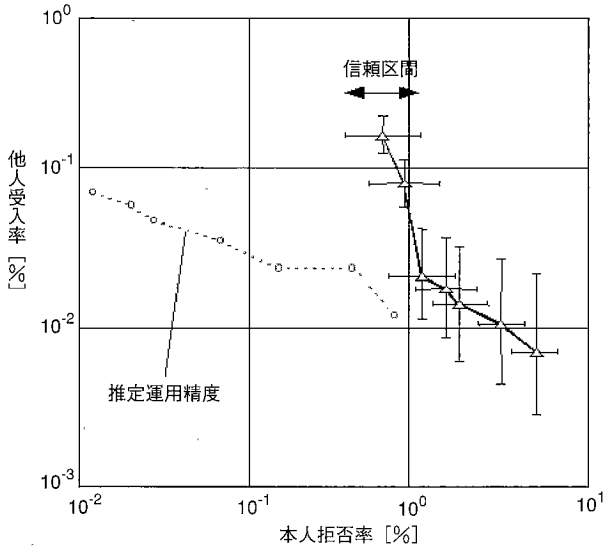

図-3　指紋を用いた本人認証装置の精度評価の一例

メーカ自身が標準的な方法で自社製品を評価できるよう評価技術を開発している．スポンサはESPRIT☆1である．指紋2件，掌形1件の精度評価を行った実績がある．

この他に，米国エネルギー省管轄のSandia National Laboratoriesがある（http://www.sandia.gov）．Sandia National Laboratoriesでは施設管理の観点から，自主的に1986年から3年間認証装置の性能評価を行った．指紋，掌形，虹彩，声紋，署名などの認証装置について，80人の職員による3カ月間の運用評価を行った実績がある．

日本においては，ECOM（電子商取引実証推進協議会）（http://www.ecom.or.jp）のWG6（本人認証ワーキンググループ）でバイオメトリクス技術の本人認証ガイドラインが検討された．ワーキンググループでは認証技術に関し，精度だけでなく利便性や利用者受容性など多面的な評価方法を検討し，ガイドラインを策定した[6]．

上記を整理すると精度評価手法の開発はNBTC，評価手順の明確化はICSAで行われている．BioTestにおいても基本的にはNBTCなどと精度評価方法は同じである．運用を考慮し評価を行うか，客観的な数値を求めることができる照合アルゴリズムに力点を置くか，あるいは，ベンダ的な観点かユーザ的な観点かなどの相違がそれぞれの機関におけるスタンスの違いとなっている．

□ 精度評価の具体的な方法 □

精度評価における課題

客観的な精度評価を行うためには，評価データベースの整備および評価方法の標準化が必要である．本稿では，データベースに関しては，誌面の都合上，認証装置のデータ対応性（対応率）および評価サンプルと精度の関係を言及するのにとどめる．

独立に採取されたデータが多量にあれば，精度評価の信頼性は向上するが，現実には限られたデータ数で評価を行うしかない．精度評価における問題としては，(1) どのような精度をユーザあるいはシステムインテグレータに提示すれば有益か（Receiver Operating Curve），(2) 現在所有するデータで評価した精度値はどの程度の信頼性があるか，また，どの程度のデータがあればどの程度の信頼ある精度を求められるか（サンプル数と信頼度），(3) 実際に収集したデータの中に，精度を著しく劣化させる特異なデータの扱いをどうするか（対応率），以上の3点がある．以下，これらの課題に対する考え方を述べる．

Receiver Operating Curveについて

評価方法として有名なのは，NBTCで使用されているROC（Receiver Operating Curve）である．

バイオメトリクスを用いた本人認証において，事前に登録したテンプレートと認証のために入力したデータの類似度を算出し，類似度があるしきい値よりも大きければ同一人物，小さければ他人と判定する．ROCとは，そのしきい値をパラメータとして他人受入率FARと本人拒否率FRRをプロットしたグラフである．

図-3にROCの一例として指紋による本人認証装置の評価結果を示す．図中のポイントはしきい値を変化させた場合のFRR（横軸）およびFAR（縦軸）を示す．各ポイントにおいては，信頼度95％での測定データの信頼区間を示す．

ROCの長所は，要求するFARとFRR，そのときのしきい値を容易に把握できる点，および，信頼区間も考慮した表示が可能な点にある．

実運用で，どの程度の精度が得られるか，照合アルゴリズムの評価結果から類推できることが望ましいが，ROCを用いれば，入力のリトライの回数（認証が失敗した場合の再入力許可数）による精度の見積もりが可能な長所もある．

たとえば，本人拒否が起こっても3回までリトライを許す場合，すなわち3回のうち1回でも照合されれば結果として本人として認める場合，本人認証装置の運用上の本人拒否率は，3回連続して本人を拒否する確率に等しく，照合アルゴリズムで得た値を3乗することにより推定できる．また，他人受入率は，3回のうち1回でも他人を受け入れる確率に等しく，元の値を3倍することにより推定できる（図-3における点線）．

信頼度と対応率について

精度評価の結果をROCで表すと，本人拒否率FRRと他

☆1 ESPRIT（European Strategic Programme R&D in Information Technology）：EUの情報技術研究および技術競争力の向上と産業育成を目的とするプロジェクト．

人受入率FARの関係をビジュアルに把握できることを述べた．次に対応率およびサンプル数と信頼度の考え方を述べる．

(1) 対応率

認証装置が対応できないデータを含むサンプルで評価を行うと，評価結果が著しく悪くなり，代表的なデータで処理を行った場合の平均的な精度が求まらないため，評価結果の信頼性を損なうおそれがある．たとえば，特定の業種に従事する著しく擦り減った指紋パターン，糖尿病をわずらっている人の眼底パターンなどがこれに相当する．ECOMワーキンググループでは，このような装置（あるいはアルゴリズム）が対応できないデータを装置未対応（アルゴリズム未対応）として評価に用いるサンプルから除き，全サンプル中の対応可能なサンプルの割合を算出し，それを対応率として精度評価項目に加える提案をしている[6]．

現在，メーカで行われている精度評価は，対応可能なデータのみを用いた精度評価が多く，どのようなデータが対応できないのか，対応率はいくらなのか明記されていない．対応率の情報を開示することは有益と考える．

(2) サンプル数と信頼度

どれだけの評価用サンプル（テンプレートおよび照合サンプル）があれば，どの程度の精度の評価が可能かは現実の問題として重要である．

たとえば，NBTCによれば信頼度95％において，誤差pの照合アルゴリズム評価に必要な最低のテンプレート数と照合用サンプルの組数（照合組数）Nminは，

$$N_{min} \fallingdotseq 3/p \qquad (3)$$

と表せる[3]．

表-2は指紋を用いた場合の照合アルゴリズムの精度と被験者数の関係を示す．

式（3）より1％（$p = 0.01$）の誤差を持つアルゴリズムの評価には最低300の照合組数が必要であることが分かる．

本人拒否率を算出する場合，同一指のテンプレートと照合用サンプルを300組照合するので，300組のテンプレートと照合用サンプルが必要となる．1人の被験者から6指のテンプレートと照合サンプルを収集するのであれば，50人（300÷6）の被験者が必要となる．

一方，他人受入率を算出する場合，異なる指同士のテンプレートと照合用サンプルを300組照合するので，指数×（指数－1）÷2≧300となる指数，すなわち25指，5人の被験者が必要となることが分かる．

同様に0.01％の本人拒否率を評価するには30,000の照合組数，5,000人の被験者，0.01％の他人受入率を評価するには，41人の被験者が必要となる．

一般に評価データの収集はコストがかかる．したがって，以上のように，求めるべき誤差とそれに必要な評価サンプル数の関係を定量化できることは非常に有益である．

表-2　精度誤差とサンプル数の関係

誤差	1％		0.01％	
認証精度	本人拒否	他人受入	本人拒否	他人受入
照合組数（組）	300	300	30,000	30,000
被験者数（人）	50	5	5,000	41

□ 今後の標準化の展開 □

今後，バイオメトリクスを用いた本人認証装置のグローバルな市場展開が予想できる．したがって，健全な技術および市場の育成のためには，精度評価は国際標準の方向に進める必要がある．

日本におけるバイオメトリクス認証装置の評価基準に関しては，本文で述べたように，平成8年度から2年間，ECOM（電子商取引実証推進協議会）の本人認証ワーキンググループで検討された．現在，タスクフォースにて継続検討を行っている．また，情報処理振興事業協会（Information-technology Promotion Agency）が次世代ディジタル応用基盤技術開発事業にて国内標準化の作業を進めている．本件に関しては成果が明確になった時点であらためて紹介したい．

参考文献

1) Newham, E. et al.: The Biometrics Report '99, SJB Services (1998).
2) 瀬戸: バイオメトリクスを用いた本人認証技術, 計測と制御, Vol.37, No.6, pp.395-401 (1998).
3) Wayman, J. L.: A Scientific Approach To Evaluating Biometric Systems Using a Mathematical Methodology, The World Premier Card and Security Technology Conference, Proc. CardTech/SecurTech '97, pp.385-395 (1997).
4) 林: 個人識別技術とそのニーズおよび期待, 計測と制御, Vol.25, No.8, pp.1-5 (1986).
5) Wayman, J. L.: Biometric Technology: Testing, Evaluation, Results, The World Premier Card and Security, Technology Conference, Proc. CardTech/SecurTech '99, pp.393-401 (1999).
6) 本人認証技術検討WG: 本人認証技術検討WG報告書, 評価基準第0.5版 (1997) および評価基準第1版 (1998).

(平成11年7月27日受付)