情報処理学会論文誌 Vol. 52 No. 9 2631–2640 (Sep. 2011)

Renyiエントロピーを用いた虹彩情報の情報量評価手法

披田野 清良[†1] 赤尾 直彦[†1]
小松 尚久[†1] 高橋 健太[†2,†3]

生体認証の識別性能の評価尺度としては，従来より，FNMR や FMR などがある．しかし，これらの尺度は，システムの性能を評価するものであり，一般的に識別性能や安全性を情報量で評価するパスワードや暗証番号などの他の認証手段との比較に用いることができない．また，近年，テンプレート保護型生体認証の安全性を情報量の概念に基づき証明する試みもある．そこで，本論文では，生体情報の情報量評価尺度として，2 次の Renyi エントロピー（以下，Renyi エントロピー）を採用し，Renyi エントロピーを用いた生体情報の情報量評価手法を提案する．Renyi エントロピーは，2 つの生体情報が一致する可能性を情報量で表現した尺度であり，生体情報のサンプルを用いた他人間照合実験を通して導出できる．そこで，本論文では，Renyi エントロピー評価の一例として，虹彩情報の情報量をシミュレーションを交えて定量的に評価する．

An Evaluation Method of Iris Information Content Using a Renyi Entropy

Seira Hidano,[†1] Naohiko Akao,[†1] Naohisa Komatsu[†1]
and Kenta Takahashi[†2,†3]

Identification performance of biometric authentication is evaluated using measures such as FNMR and FMR. However, these measures cannot be utilized to compare biometric information with other authentication methods such as passwords and personal identification numbers (PINs) because these measures evaluate the performance of the system but not information content that biometric information has in its own. Furthermore, in the discussion on the template protection-based biometric authentication, there have been attempts to prove the security on the basis of information content. Therefore, we propose an evaluation method of biometric information content using a Renyi entropy of order 2. The Renyi entropy is the possibility that two biometric samples coincide expressed in information content and can be derived through the experiment for interpersonal matching using a set of biometric samples. In this paper, we show how to evaluate it, and evaluate iris information content through simulations.

1. まえがき

生体認証は，記憶，所持の煩わしさから解放されるという利便性があり，入退室管理やネットワークアクセスなどのアクセスコントロール，ネットワークバンキングなどのフローコントロール，サーベランスシステムなどのトラッキングへの展開が期待されている．しかし，生体情報は，環境条件の違いやセンサの取得誤差，経時変化などにより，同一の生体から取得される場合でもわずかに異なる情報となる．また，異なる生体から取得される場合でも，生体情報間の相関性により，類似の情報となる可能性が高い．したがって，生体認証の識別性能を定量的に評価することは，生体情報を個人認証に適用するうえで必要不可欠といえる．

生体認証の識別性能に関する評価尺度としては，従来より，同一の生体の情報を不一致と判定する誤り確率（以下，FNMR）と異なる生体の情報を一致と判定する誤り確率（以下，FMR）がある．これらの誤り確率は照合判定閾値に依存して変化するため，閾値を変化させたときの FNMR と FMR の関係の軌跡（以下，ROC カーブ）を用いて生体認証の認証精度を表現する方法が標準化されている[1]．しかし，これらの尺度は，システムの性能を評価するものであり，一般的に識別性能や安全性を情報量で評価するパスワードや暗証番号などの他の認証手段との比較に用いることができない．異なるモダリティ間の比較においても，ROC カーブに交わりがある場合，どちらの生体情報が優位であるかを端的に表現できない．

また，生体認証の識別性能を生体情報の情報量で評価する試みもある[2)–4)]．情報量で評価することにより，生体認証の識別性能を直観的に理解でき，他の認証手段との比較や，異なるモダリティ間の比較が容易になると考えられる．しかし，これまでに提案されている生体情報の情報量評価尺度は，ある特定のモダリティを対象としているものや，実際に測定する

†1 早稲田大学理工学術院
Faculty of Science and Engineering, Waseda University
†2 株式会社日立製作所横浜研究所
Yokohama Research Laboratory, Hitachi, Ltd.
†3 東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology, The University of Tokyo

ことが困難なものであり，任意の生体情報に対して適用可能で標準的な評価尺度は存在しない．

さらに，近年，生体情報を暗号技術などにより解読不可能な状態に変換し，生体情報の秘匿性を保ちつつ認証を可能とするテンプレート保護型生体認証が注目されている[5),6)]．これらの安全性に関する議論でもまた，情報量の概念に基づきテンプレート保護技術の安全性を証明する試みがある[7)]．しかし，それらの議論では，生体情報の情報量が十分に大きいことを前提としており，生体情報間の相関性により当該情報量が減少する可能性については言及していない．このような安全性評価では，テンプレート保護型生体認証が実用化した際に，要求される安全性のレベルを十分に達成できず，漏洩したテンプレートから生体情報が復元され，システムへのなりすましなどの危険性が生じる．したがって，生体情報の情報量を正確に評価することは，テンプレート保護技術の安全性を評価するうえでも重要な課題となる．

そこで，本論文では，生体情報の情報量評価尺度として，生体情報間の距離に着目した 2 次の Renyi エントロピーを採用し，Renyi エントロピーを用いた生体情報の情報量評価手法を提案する．Renyi エントロピーは，2 つの生体情報が一致する可能性を情報量で表現した尺度であり，生体情報のサンプルを用いた他人間照合実験を通して導出できる．また，本論文では，虹彩認証を例に本評価手法の適用方法について詳述し，シミューレションを交えて虹彩情報の情報量を定量的に評価する．その結果，虹彩認証の識別性能は環境要因の影響により大きく変化するという知見を得ている．

2. 関 連 研 究

生体認証の識別性能や安全性を生体情報の情報量で評価する試みとして，以下に代表的な 3 つの研究事例を示す．

まず，Daugman による虹彩情報の情報量を実験的に評価する試みがあげられる[2)]．虹彩情報を虹彩コード $\{0,1\}^{2048}$ で記述した場合，瞼や睫毛，反射光などの環境条件の違いやビット間の相関により，照合の際に識別に有効ではないビットが混入する．そこで，本検討では，虹彩情報のサンプルを用いた他人間照合実験を通して，識別に有効なビット数を実験的に評価し，虹彩の普遍的な識別性能を評価することを目的としている．具体的には，まず，識別に有効なビット数を $\hat{n}$，各ビットの一致確率を θ とし，虹彩情報間の距離分布を二項分布 $Bi(\theta,\hat{n})$ でモデル化する．このとき，距離の平均と標準偏差は，それぞれ $\mu=1-\theta$，$\sigma=\sqrt{\theta(1-\theta)/\hat{n}}$ で与えられる．そして，他人間照合実験を通して得られる平均と標準偏差を用いて，$\hat{n}$ の値を推定する．Daugman の実験では，θ がほぼ理想的に 1/2 となることから，$\hat{n}$ を識別エントロピーと呼び，これを虹彩情報の識別性能を示す評価尺度としている．本評価尺度は，環境要因の影響やビット間の相関など，実世界で虹彩が受ける影響を考慮した尺度であり，また標準的な生体認証の精度評価方法[1)] と同様の手順で実験的に導出できるため，有用かつ実用的である．なお，4 章で詳述するが，本論文で提案する Renyi エントロピーを用いた情報量評価手法は Daugman の評価方法を理論的に一般化した手法といえる．

次に，Adler らによる生体情報の個人内変動を考慮して情報量を実験的に評価する試みがあげられる[3)]．生体認証では，本人内スコア分布と他人間スコア分布の重なりが小さいほど，照合判定閾値の設定により FNMR や FMR を抑えることができ，個人識別性能は高くなる．一方，分布間の重なりが大きいほど，個人識別性能は低くなる．そこで，本検討では，生体情報の情報量を生体情報が個人識別に関して持つ情報量と定義し，生体情報の全体分布と個人分布の相対性を利用して当該情報量を実験的に評価することを目的としている．具体的には，生体情報 B の全体に関する確率関数を $p_B(b)$，ある特定の個人 U に関する確率関数を $p_{B|U}(b|u)$ とし，次式に示す Kullback Leibler 情報量（以下，KL 情報量）を評価尺度とする．

$$D_{KL}(p_{B|U} \parallel p_B) = \sum_b p_{B|U}(b|u) \log_2 \frac{p_{B|U}(b|u)}{p_B(b)} \tag{1}$$

ただし，本検討では，B が連続値をとる場合についても言及している．$p_B(b)$ と $p_{B|U}(b|u)$ が一致するとき，$D_{KL}(p_{B|U} \parallel p_B) = 0$ となり，$p_B(b)$ と $p_{B|U}(b|u)$ が異なるほど，$D_{KL}(p_{B|U} \parallel p_B)$ は大きい値を示す．全体の分布とある個人の分布が同一であれば，生体情報が個人に関する情報を持たないことは明らかであり，本評価尺度は個人識別性能の意味で自然な性質を持つ．しかし，$D_{KL}(p_{B|U} \parallel p_B)$ を導出するためには，$p_B(b)$，$p_{B|U}(b|u)$ を推定をする必要がある．Adler らは，B をユークリッド空間上のベクトルで表現できる場合のみを対象として，$p_B(b)$，$p_{B|U}(b|u)$ を多次元正規分布でモデル化し，B のサンプルからそれぞれの分布を実験的に推定することを提案している．ここで，分布推定の精度を高めるためには，一般的に，B の次元数に対して十分に多くのサンプルが必要となる．このため，B が高次元の場合，$p_{B|U}(b|u)$ を推定するためには，ある特定の個人 U から多くのサンプルを取得しなければならない．また，指紋情報をマニューシャ情報の集合で記述した場合，それぞれのマニューシャ情報は順序を持たないため，ユークリッド空間上のベクトルで表現できない．したがって，測定の困難さと任意のモダリティへ適用できないことから，本評価尺度の実用性は低い．

最後に，Takahashi らによる生体認証システムが出力する照合スコアのみを用いて情報量を実験的に評価する試みがあげられる[4]．本検討では，まず，識別情報の情報量をある個人 U がある識別情報 X に関して持つ情報量，すなわち次式に示す U と X の相互情報量で定義する．

$$I(U;X) = H(X) - H(X|U) \tag{2}$$

ただし，$H(X)$ は X の Shannon エントロピー，$H(X|U)$ は U の観測後に残る X の条件付き Shannon エントロピーとする．X を生体情報 B とした場合，$I(U;B)$ を導出するためには B の全体分布や個人分布を推定する必要があるが，上述したように B が高次元の場合，それは容易ではない．また，U の識別は，B の観測のみでは行えず，あらかじめ登録されているテンプレートとの照合により実現する．そこで，本検討では，生体情報の情報量を生体認証システムが個人識別に関して持つ情報量と定義し，X を B とテンプレートとの照合スコアとして本人内スコア分布と他人間スコア分布の相対性により近似的に $I(U;X)$ を評価することを目的としている．具体的には，本人内照合スコア X_G の確率関数を $p_{X_G}(x)$，他人間照合スコア X_I の確率関数を $p_{X_I}(x)$ とし，次式に示す KL 情報量を評価尺度とする．

$$D_{KL}(p_{X_G} \parallel p_{X_I}) = \sum_x p_{X_G}(x) \log_2 \frac{p_{X_G}(x)}{p_{X_I}(x)} \tag{3}$$

ただし，本検討では，X_G，X_I が連続値をとる場合についても言及している．$p_{X_G}(x)$ や $p_{X_I}(x)$ は，標準的な生体認証の精度評価方法と同様の手順により推定できる．また，本評価尺度は照合スコアのみから導出でき，生体認証システムをブラックボックスとして扱うことができる．このため，情報量評価を第三者機関に委託することもでき，実用的な評価尺度といえる．しかし，得られた情報量から B の Shannon エントロピーや Renyi エントロピー[8] などの B 自身の情報量について言及することは難しく，すでに Dodis らの報告[7] により B の情報量に基づき定式化されているテンプレート保護技術の安全性の評価尺度としては必ずしも有用ではない．

3. 生体情報の情報量評価手法

本章では，Renyi エントロピーを用いた生体情報の情報量評価手法を提案する．まず，生体情報の情報量を評価する生体認証モデルついて述べる．次いで，生体情報の情報量評価尺度として Renyi エントロピーを定義し，Shannon エントロピーとの比較を行う．最後に，他人間照合実験を通した Renyi エントロピーの評価手順を示す．

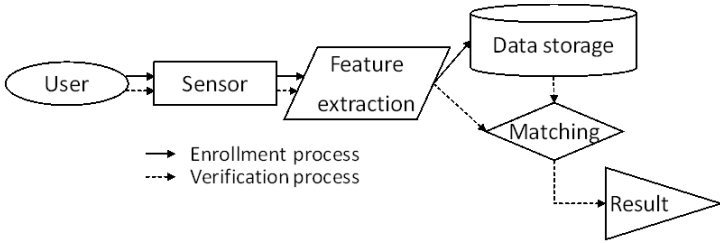

図 1　生体認証モデル
Fig. 1　Biometric authentication model.

3.1 生体認証モデル

本論文では，**図 1** に示す生体認証モデルにおける生体情報の情報量を評価する．

登録過程では，まず，ユーザからセンサを通して生体に関する原情報を取得する．そして，それらの情報からより識別性の高い情報を抽出し，その情報をテンプレートとしてシステムに保管する．本論文では，システムに保管するテンプレートと同形式の情報を生体情報 B とする．ただし，B は離散確率変数とする．照合過程では，登録過程と同様にユーザが提示する原情報から B と同形式の情報 B' を抽出し，スコア関数 $g(B,B')$ により B と B' の間の距離 D もしくは類似度 S を算出する．D および S は，B と同様に離散確率変数とする．スコア関数としては，たとえば，ハミング距離や差集合距離などがある．

ハミング距離の場合，B は長さ n の q 元符号 $\mathcal{Q}_q^n$ の元として与えられる．ハミング距離は，2 つの生体情報 $B = (B_1, \ldots, B_n)$，$B' = (B'_1, \ldots, B'_n)$ を用いて，$g_{HD}(B,B') = |\{i \mid B_i \neq B'_i, i = 1, \ldots, n\}|$ で表せる．ただし，$|\cdot|$ は集合の要素数を示す．ハミング距離をスコア関数に持つ認証モデルの例としては，Daugman の虹彩コードを用いた虹彩認証があげられる[2]．

差集合距離の場合，B はある普遍集合 Ω の部分集合の全集合 2^Ω の元として与えられる．差集合距離は，2 つの生体情報 B，B' を用いて，$g_{SDD}(B,B') = |\{x \mid x \in B, x \notin B'\}|$ で表せる．差集合距離をスコア関数に持つ認証モデルの例としては，指紋情報をマニューシャ情報の集合で記述した指紋認証があげられる[9]．

3.2 Renyi エントロピー

ある確率分布に従う確率変数の観測により得られる情報量は，一般的に，Shannon エントロピーで定義される[10]．たとえば，ユーザが選択するパスワードは，記憶しやすい英語の文字列などが利用されるため，ランダムに選択されたパスワードに比べて推測が容易とな

る．Shay らは，その推測困難性の評価尺度として Shannon エントロピーを採用し，統計的にパスワードの情報量を算出している[11]．生体情報 B の Shannon エントロピー $H(B)$ は，B のとりうる値の集合を $\mathcal{B}$，確率関数を $p_B(b)$，$b \in \mathcal{B}$ とすると，次式で表せる．

$$H(B) = -\sum_{b \in \mathcal{B}} p_B(b) \log_2 p_B(b) \tag{4}$$

$p_B(b)$ が既知であれば，$H(B)$ は容易に導出できる．しかし，生体情報は，同一の生体でも取得ごとにわずかに異なる情報となることや，まったく異なる生体から類似の情報が取得されることがあるため，それらの複雑な相関性を考慮して $p_B(b)$ を理論的にモデル化することは困難である．また，実験的に $p_B(b)$ を推定することも，B は一般的に高次元の情報であり，B の次元数に対して十分に多くのサンプルが必要となるため，容易ではない．したがって，式 (4) より $H(B)$ を導出することは現実的に難しく，生体情報の情報量評価尺度としては不適といえる．

そこで，本論文では，次式に示す 2 次の Renyi エントロピー $H_2(B)$ を生体情報の情報量評価尺度として定義する．

$$H_2(B) = -\log_2 \sum_{b \in \mathcal{B}} p_B(b)^2 \tag{5}$$

$H_2(B)$ は次式に示す α 次の Renyi エントロピー $H_\alpha(B)$ の $\alpha = 2$ のときの特殊形として与えられる[8]．

$$H_\alpha(B) = \frac{1}{1-\alpha} \log_2 \sum_{b \in \mathcal{B}} p_B(b)^\alpha \tag{6}$$

ただし，$\alpha \geq 0$，$\alpha \neq 1$ とする．Renyi エントロピーは，2 つの生体情報が一致する可能性，すなわち B の衝突困難性に関する情報量の評価尺度である．2 章で示した Adler らや Takahashi らの評価尺度が本人と他人の相対性を利用して個人識別性能を評価する尺度であるのに対し，$H_2(B)$ は，2 章で示した Daugman の評価尺度と同様に，識別性能の限界を評価する尺度である．ただし，Daugman が虹彩の普遍的な識別性能を評価することを目的としているのに対し，本論文では，3.1 節で示した B の記述形式とスコア関数が定義された特定の認証モデルにおける識別性能の限界を評価することを目的とする．また，$\mathcal{B}$ のサイズが大きいことは，B の安全性に関する要件の 1 つでもあるため[12]，$H_2(B)$ は安全性の評価尺度としても有用と考えられる．ここで，式 (5) の $\sum_{b \in \mathcal{B}} p_B(b)^2$ は，2 つの生体情報 B，B' が同一の値 $b \in \mathcal{B}$ をとる確率を示しており，B，B' が同一の値をとるとき，B と B' の間の距離は 0 となるため，$H_2(B)$ は，生体情報間の距離 D の確率関数 $p_D(d)$，$d \in \mathbb{R}$ を用いて，次式で表せる．

$$H_2(B) = -\log_2 p_D(0) \tag{7}$$

$p_D(0)$ は，$p_D(d)$ が既知であれば容易に導出できる．$p_D(d)$ は，B のサンプルを用いた他人間照合実験を通して得られる D のサンプルを学習データとして推定でき，標準的な生体認証の精度評価方法[1] に従いサンプルの収集および照合を行うことで，分布推定の信頼性が高められる．したがって，$H_2(B)$ は式 (7) より現実的に評価でき，生体情報の情報量評価尺度として実用的な尺度といえる．

ここで，$H_2(B)$ と $H(B)$ の関係について述べる．$H(B)$ は α を 1 に近づけたときの $H_\alpha(B)$ の極限値であり，$H_\alpha(B)$ は α に関する広義の単調減少関数である．このため，$H_2(B)$ と $H(B)$ の関係は次式で表せる．

$$H_2(B) \leq H(B) \tag{8}$$

等号は $p_B(b)$ が一様分布の場合に成立する．

さらに，近年，Renyi エントロピーを用いて式 (8) より厳密に Shannon エントロピーの上界と下界を推定する方法が報告されている[13),14]．これらの報告により，$H_2(B)$ を評価することで，B の衝突困難性だけでなく，$H(B)$，すなわち B の推測困難性に関する評価も可能となる．また，ある確率変数の Shannon エントロピーのみからその確率変数の Renyi エントロピーの上界と下界を推定できることは同様に明らかであり，たとえば，パスワードの Renyi エントロピーは Shay らが算出したパスワードの Shannon エントロピーの推定値より算出できる．したがって，Renyi エントロピーを生体情報の情報量評価尺度として採用することにより，異なるモダリティ間や他の認証手段との間で，推測困難性および衝突困難性の両点から識別性能や安全性の比較が容易になると考えられる．

3.3　Renyi エントロピー評価手順

図 2 に生体情報 B の Renyi エントロピー $H_2(B)$ の評価手順を示す．

まず，B のサンプルを用いて他人間照合実験を行い，生体情報間の距離 D もしくは類似度 S のサンプルを取得する．ただし，Renyi エントロピー評価の信頼性を高めるために，この手順は 3.2 節で述べたように生体認証の標準的な精度評価方法に従う．

次に，取得した D のサンプルより D の確率関数 $p_D(d)$ を推定する．2 つの生体情報が一致する確率，すなわち D が 0 となる確率 $p_D(0)$ は，きわめて小さい値になると考えられるため，D のサンプルから直接推定することは現実的ではない．したがって，$p_D(0)$ は $p_D(d)$ の推定分布より間接的に導出することが望ましい．ただし，Daugman の虹彩認証モデル[2] のように $p_D(d)$ の形状が従来の研究において十分に検討された生体情報であれば，$p_D(d)$

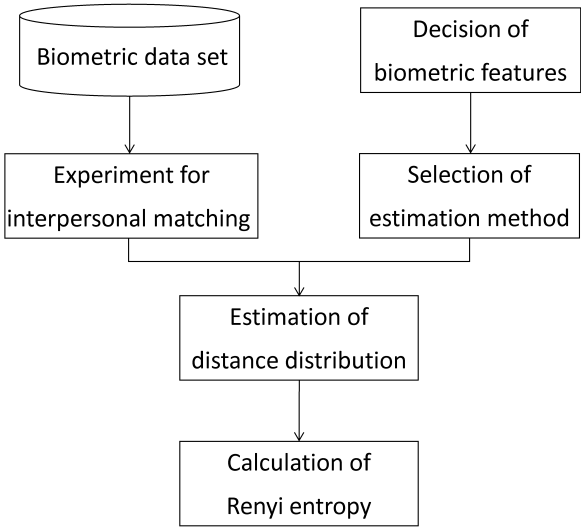

図 2　Renyi エントロピー評価手順
Fig. 2　Evaluation procedure of Renyi entropy.

は計算の容易性から D のサンプルを用いてパラメトリックに推定する[15]．一方，$p_D(d)$ の形状が未知でモデル化できない場合は，分布の形状を仮定せずにデータに依存して推定するノンパラメトリックな手法を用いる．ノンパラメトリックな手法としては，核関数に基づく方法などがあり，近年，$H_2(B)$ を算出するうえで重要な $p_D(0)$ において誤差を抑えることができる特殊な核関数が報告されている[16]．S の場合も同様に，そのサンプルより S の確率関数 $p_S(s)$，$s \in \mathbb{R}$ を推定し，S が最も大きい値をとる確率を導出する．

最後に，$p_D(0)$ もしくは S が最も大きい値をとる確率の推定値を用いて式 (7) より $H_2(B)$ を算出する．

なお，本論文では，4 章で，$p_D(d)$ の形状が二項分布としてよく知られている Daugman の虹彩認証モデルを例に，本評価手法の適用方法について詳述し，虹彩情報の情報量を定量的に評価する．また，代表的な認証モデルの 1 つであるマニューシャ情報を用いた指紋認証においては，Pankanti らにより指紋情報間の $p_S(s)$ のモデルが検討されている[17]．そこで，付録 A.1 で，本評価手法の指紋認証への適用例として，Pankanti らによる $p_S(s)$ のモデル化について紹介するとともに，そのモデルに基づき $p_S(s)$ をパラメトリックに推定した場合の指紋情報の Renyi エントロピーの評価方法を示す．

4. 虹彩情報の情報量評価

本章では，Renyi エントロピー評価の一例として，3.3 節の評価手順に従い，虹彩情報の情報量を定量的に評価する．まず，対象とする虹彩認証モデルについて述べ，虹彩情報の Renyi エントロピーの評価方法を示す．そして，実際に虹彩画像のデータベースを用いて他人間照合実験を行い，虹彩情報の情報量を算出する．

4.1 虹彩認証モデル

本論文では，Daugman の虹彩認証モデルを採用する[2]．虹彩情報 C は，虹彩画像より生成可能な長さ n の虹彩コード $\{0,1\}^n$ で記述する．また，虹彩は瞼や睫毛，反射光などの環境要因の影響を受ける可能性があるため，C の生成と同時に，環境要因の影響の有無を示す C と同形式のマスクコード W を生成する．W の各ビットは，対応する C のビットに環境要因の影響がある場合は 0 を返し，それ以外は 1 を返す．

2 つの虹彩情報 C と C' の間の距離 D_C は，3.1 節で示したハミング距離 g_{HD} とそれぞれのマスクコード W，W' を用いて，次式で表せる．

$$D_C = \frac{g_{HD}(C \cap W \cap W', C' \cap W \cap W')}{\| W \cap W' \|} \tag{9}$$

ただし，$\| \cdot \|$ は 1 を返すビットの個数を示す．また，C は，センサへの正視の仕方により，回転の影響を受ける可能性がある．このため，Daugman の認証モデルでは，虹彩画像から抽出した虹彩領域を回転させて複数の異なる C を生成し，それらの中で D_C が最小となる場合の結果を照合スコアとして採用している．以下，回転補正を行う照合をベストマッチ，補正を行わない照合を非ベストマッチと呼ぶ．

4.2 虹彩情報の Renyi エントロピー

虹彩情報 C の Renyi エントロピー $H_2(C)$ の評価方法について述べる．

2 つの虹彩情報間の距離 D_C の確率関数 $p_{D_C}(d)$ は，C の各ビットの一致確率を θ，C の中で識別に有効なビットの数を $\hat{n}$ と仮定した場合，次式に示す二項分布でモデル化される[2]．

$$p_{D_C}(d) = \frac{\hat{n}!}{(\hat{n}d)!(\hat{n}(1-d))!}\theta^{\hat{n}(1-d)}(1-\theta)^{\hat{n}d} \tag{10}$$

D_C の期待値 $E(D_C)$ と分散 $V(D_C)$ はそれぞれ次のように表せる．

$$E(D_C) = 1 - \theta \tag{11}$$

$$V(D_C) = \theta(1-\theta)/\hat{n} \tag{12}$$

式 (10) より，2 つの虹彩情報が一致する確率は $p_{D_C}(0) = \theta^{\hat{n}}$ となり，$H_2(C)$ は次式で表せる．

$$H_2(C) = -\log_2 \theta^{\hat{n}} \tag{13}$$

$\hat{n}$ は，C のすべてのビットにおいて環境要因の影響がなく，またビット間の相関もない場合，C のコード長 n と一致する．しかし，上述したように，虹彩は瞼や睫毛，反射光などの環境要因の影響を受ける可能性があり，C の中に環境要因の影響を受けたビットがある場合，マスクコード W の対応するビットは 0 となる．照合時は，2 つの虹彩情報のマスクコードを W，W' とすると，式 (9) より，$\| W \cap W' \|$ 個のビットのみが識別に用いられ，$n - \| W \cap W' \|$ 個のビットは利用されない．そして，$\| W \cap W' \|$ 個のビットの中でも，W により環境要因を検出できなかったビットや，ビット間に何らかの相関がある場合，$\hat{n}$ は $\| W \cap W' \|$ よりさらに減少する．したがって，環境要因の影響や C の相関性により，$\hat{n}$ は n とは一致せずに大きく減少すると考えられる．そこで，$\hat{n}$ は，C のサンプルを用いた他人間照合実験を通して得られる D_C の平均および分散と式 (11)，式 (12) より推定する．

2 章で述べたように，Daugman は，他人間照合実験を通して得られる D_C の平均と分散を用いて $\hat{n}$ を推定し，これを識別エントロピーと呼び，虹彩情報の識別性能を示す評価尺度としている．識別エントロピーは，他人間照合実験を通して評価できるという実用的な利点がある一方，Shannon エントロピーなどの一般的な情報量評価尺度との関係が不明であることが問題であった[18]．Daugman の実験では，θ がほぼ理想的に 1/2 となることから，式 (13) より，識別エントロピーは Renyi エントロピーと解釈できる．しかし，Daugman は，$\theta = 1/2$ とならない場合の識別エントロピーについては必ずしも言及しておらず，またこの場合，$\hat{n}$ は直接 Renyi エントロピーや Shannon エントロピーを意味しない．そこで，本論文では，識別エントロピーを Renyi エントロピーの観点から一般化した $H_2(C)$ を虹彩情報の情報量評価尺度として採用する．これにより，任意の θ の値に対して，一般的な情報量評価尺度である Renyi エントロピーの概念に基づく評価が可能となり，さらに 3.2 節で示したように Shannon エントロピーとの関係も明白となる．また，Daugman の評価では，虹彩の普遍的な識別性能を評価することを目的としているため，非ベストマッチを採用し，認証モデルになるべく依存しない条件で識別エントロピーを推定している．しかし，環境条件や対象とする認証モデルが異なれば，識別性能もまた大きく変化すると考えられる．このため，4.3 節では，実際の虹彩認証の際に適用されるベストマッチを採用し，異なる環境条件の下で $H_2(C)$ を定量的に評価する．

表 1　虹彩 DB 諸元
Table 1　Specifications of iris DB.

DB	Interval	Lamp	Twins
人数	249	411	200
虹彩数	396	819	400
画像数	2,655	16,213	3,183
特徴	環境要因無	環境要因有	双子

4.3　情報量評価実験

4.2 節で示した評価方法に従って，虹彩画像のデータベースを用いて他人間照合実験を行い，虹彩情報の情報量を算出する．

他人間照合実験を行うに際して，使用した虹彩画像のデータベース（以下，虹彩 DB）の諸元を**表 1** に示す．本実験では，虹彩 DB として，CASIA-IrisV3[19] に収録されている 3 つの異なる虹彩 DB Interval，Lamp，Twins を使用した．3 つの虹彩 DB はそれぞれ異なる特徴を持つ．Interval は環境要因の影響が少ない虹彩画像のみを収録しており，Lamp は異なる照明状況下で撮影された虹彩画像を収録している．Twins は，双子の虹彩画像を収録しており，環境要因の有無については明記されていないため，照明などの影響を受けた画像を含む可能性がある．他人間照合実験では，それぞれの虹彩 DB ごとに異なる虹彩から 2 つの画像をランダムに選択し，100,000 回の照合を行った．ただし，虹彩コードの生成および照合に Telecom & Management Sud Paris が公開している OSIRIS version 2.01[20] を使用した．OSIRIS は，4.1 節で述べた Daugman の虹彩認証モデルを実装しており，虹彩情報 C としてコード長 $n = 2016$ の虹彩コードを生成し，ベストマッチにより照合を行う．

図 3 に，一例として，Interval を用いた際の他人間照合実験結果を示す．水平軸は 2 つの虹彩情報間の距離 D_C を表し，垂直軸は D_C の各値の出現頻度を表す．このとき，D_C の平均は 0.463，分散は 0.0004 であった．式 (11)，式 (12) より，各ビットの一致確率 θ の推定値は 0.537，識別に有効なビット数 $\hat{n}$ の推定値は 607 となり，虹彩情報の Renyi エントロピー $H_2(C)$ は 545 bit と算出できる．

異なる虹彩 DB を用いた際の D_C の平均および分散，$\hat{n}$，$H_2(C)$ の値を**表 2** に示す．表 2 より，$H_2(C)$ は虹彩 DB の特徴に応じて大きく異なることが分かる．ここで，マスクコード W，W' における $\| W \cap W' \|$ の値に着目すると，各虹彩 DB の $\| W \cap W' \|$ の平均は，Interval が約 1,670 個，Lamp が約 1,422 個，Twins が約 1,544 個であった．Interval が環境要因の影響が少ない虹彩 DB であるのに対し，Lamp は照明の影響を受けた虹彩 DB であ

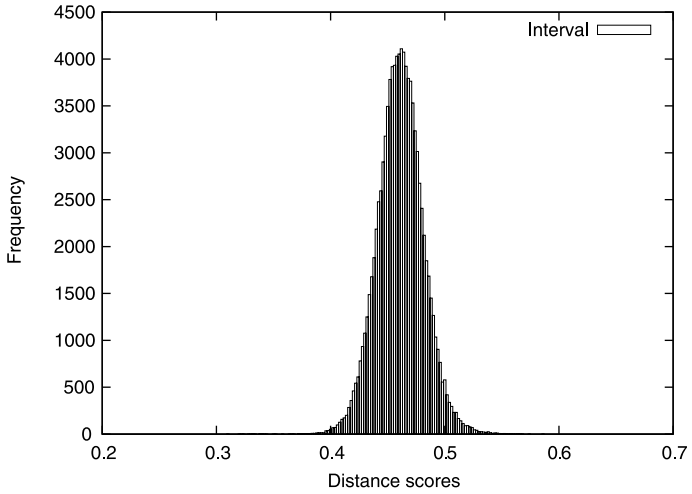

図 3　他人間照合実験結果（Interval）
Fig. 3　Distribution of distance scores.

表 2　虹彩情報の Renyi エントロピー
Table 2　Renyi entropy of iris information.

DB	Interval	Lamp	Twins
平均	0.463	0.438	0.45
分散	0.0004	0.0012	0.0009
$\hat{n}$	607	212	325
$H_2(C)$	545	177	239

ることから，環境要因の影響が大きいほど，W により識別に用いられるビットが少なくなることが分かる．また，Twins も，Lamp より程度は低いが，環境要因の影響を受けている可能性がある．次に，表 2 の $\hat{n}$ に注目すると，どの虹彩 DB においても $\hat{n}$ は $\| W \cap W' \|$ より大きく減少しており，4.2 節で述べたように，C のビット間に強い相関があることがうかがえる．また，環境要因の影響が大きい虹彩 DB ほど，その減少率も大きくなることから，4.2 節で述べたように，C が本来持つビット間の相関だけでなく，W により環境要因を検出できなかったビットも影響していると考えられる．最後に，表 2 の D_C の平均に注目する．ベストマッチを採用した場合，補正する回転量が大きいほど，D_C の平均はより小さくなることが期待される．本実験では，OSIRIS により回転量は固定されているが，$\hat{n}$ の値が小さいほど，相対的に回転の影響を大きく受けた可能性があり，表 2 における D_C の平均が Interval，Twins，Lamp の順に小さくなったと考えられる．したがって，式 (13) における θ，$\hat{n}$ の両パラメータの面から考えても，環境要因の影響が大きい場合に $H_2(C)$ が減少することは自然な結果といえる．

以上より，Renyi エントロピーを用いて虹彩情報の情報量を定量的に評価することで，ベストマッチを用いた際の虹彩認証の識別性能が，環境要因の影響により大きく変化するという知見を得た．また，近年，テンプレート保護型生体認証が注目されていることから，テンプレート保護技術に適用可能で安全性の向上を目的とした量子化手法が多く検討されている[12),21)]．これらの検討では，生体情報は，虹彩情報と同様にコード長 n の符号語 $\{0,1\}^n$ で記述され，またハミング距離により照合を行う．したがって，量子化された生体情報においても，4.2 節で示した虹彩情報と同様の方法により Renyi エントロピーを導出できると考えられる．

5. むすび

本論文では，生体情報の情報量評価尺度として，2 次の Renyi エントロピーを採用し，Renyi エントロピーを用いた生体情報の情報量評価手法を提案した．Renyi エントロピーは，2 つの生体情報が一致する可能性を情報量で表現した尺度であり，標準的な生体認証の精度評価方法に従って導出できる．また，本評価尺度により，生体情報の Shannon エントロピーの上界と下界を定めることができる．さらに，本論文では，本評価手法を Daugman の虹彩認証モデル[2)] に適用し，虹彩情報の Renyi エントロピーが Daugman の識別エントロピーの一般形であることを明らかにした．そして，Renyi エントロピーを用いて虹彩情報の情報量を定量的に評価し，ベストマッチを用いた際の虹彩認証の識別性能が，環境要因の影響により大きく変化するという知見を得た．

今後は，まず，4.3 節で述べたように，虹彩情報の Renyi エントロピーと同様の方法により評価が可能な量子化された生体情報[12),21)] の情報量を定量的に評価し，識別性能や安全性について虹彩認証との比較を行う．次いで，付録 A.1 に示す評価方法により，マニューシャ情報の集合で記述された指紋情報の情報量を定量的に評価し，指紋認証の識別性能や安全生について明らかにしていく．

参 考 文 献

1) Mansfield, A.J. and Wayman, J.L.: Best Practices in Testing and Reporting Performance of Biometric Devices: Version 2.01 (2002).
2) Daugman, J.: The importance of being random: Statistical principles of iris recognition, *Pattern Recognition*, Vol.36, No.2, pp.279–291 (2003).
3) Adler, A., Youmaran, R. and Loyka, S.: TOWARDS A MEASURE OF BIOMETRIC INFORMATION, *Proc. Canadian Conference on Electrical and Computer Engineering* (*CCECE 2006*), pp.210–213 (2006).
4) Takahashi, K. and Murakami, T.: A Metric of Information Gained through Biometric Systems, *Proc. 20th International Conference on Pattern Recognition* (*ICPR 2010*), pp.1184–1187 (2010).
5) Ratha, N., Connell, J. and Bolle, R.: Enhancing security and privacy in biometrics-based authentication systems, *IBM Systems Journal*, Vol.40, No.3, pp.614–634 (2001).
6) Juels, A. and Sudan, M.: A Fuzzy Vault Schem, *Proc. IEEE International Symposium on Information Theory* (*ISIT 2002*), p.408 (2002).
7) Dodis, Y., Ostrovsky, R., Reyzin, L. and Smith, A.: Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data, *SIAM J. Comput.*, Vol.38, No.1, pp.97–139 (2008).
8) Renyi, A.: ON MEASURES OF ENTROPY AND INFORMATION, *Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability*, Vol.1, pp.547–561 (1961).
9) Maltoni, D., Maio, D., Jain, A.K. and Prabhakar, S.: *Handbook of Fingerprint Recognition*, Springer (2003).
10) Shannon, C.E.: A Mathematical Theory of Communication, *Bell System Technical Journal*, Vol.27, pp.379–423, 623–656 (1948).
11) Shay, R., Komanduri, S., Kelley, P.G., Leon, P.G., Mazurek, M.L., Bauer, L., Christin, N. and Cranor, L.F.: Encountering Stronger Password Requirements: User Attitudes and Behaviors, *Proc. 6th Symposium on Usable Privacy and Security* (*SOUPS2010*), pp.1–20 (2010).
12) Chen, C. and Veldhuis, R.: Binary Biometric Representation through Pairwise Polar Quantization, *Proc. 3rd International Conference on Advances in Biometrics* (*ICB 2009*), pp.72–81 (2009).
13) Harremoes, P. and Topsoe, F.: Inequalities Between Entropy and Index of Coincidence Derived From Information Diagrams, *IEEE Trans. Information Theory*, Vol.47, No.7, pp.2944–2960 (2001).
14) Zyczkowski, K.: Renyi Extrapolation of Shannon Entropy, *Open Systems and Information Dynamics*, Vol.10, No.3, pp.297–310 (2003).
15) Bishop, C.M.: *PATTERN RECOGNITION AND MACHINE LEARNING*, Springer (2006).
16) Kokonendji, C.C., Kiesse, T.S. and Zocchi, S.S.: Discrete triangular distributions and non-parametric estimation for probability mass funtction, *Journal of Nonparametric Statistics*, Vol.19, pp.241–254 (2007).
17) Pankanti, S., Prabhakar, S. and Jain, A.K.: On the Individuality of Fingerprints, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.24, No.8, pp.1010–1025 (2002).
18) Li, S.Z. and Jain, A.K.: *Encyclopedia of Biometrics*, Springer (2009).
19) Institute of Automation of the Chinese Academy of Sciences: CASIA-IrisV3, available from ⟨http://www.cbsr.ia.ac.cn/IrisDatabase.htm⟩.
20) Telecom & Management Sud Paris: Open Source for IRIS version 2.01, available from ⟨http://biometrics.it-sudparis.eu/BMEC2007⟩.
21) Xu, H. and Veldhuis, R.N.J.: Binary Representations of Fingerprint Spectral Minutiae Features, *Proc. 20th International Conference on Pattern Recognition* (*ICPR 2010*), pp.1212–1216 (2010).
22) Zhang, H.: A Note About Maximum Likelihood Estimator in the Hypergeometric Distribution, *Revista Comunicaciones en Estadistica, UNIVERSIDAD SANTO TOMAS*, Vol.2, No.2 (2009).

付　　録

A.1　指紋情報の Renyi エントロピー

指紋認証モデルとしてマニューシャマッチングモデル[9] を採用し，指紋情報間の類似度分布を Pankanti らの報告[17] に基づきモデル化した場合の指紋情報の Renyi エントロピーの評価方法を示す．

指紋情報 F は，指紋画像から取得可能な T 個のマニューシャ情報 M で記述し，M はマニューシャの位置を示す座標情報 $L = (L_x, L_y)$ と隆線ベクトルの方向を示す角度情報 R の組とする．F および M はそれぞれ次のように表せる．

$$F = \{M_1, \ldots, M_i, \ldots, M_T\} \tag{14}$$

$$M_i = (L_i, R_i) \tag{15}$$

2 つの指紋情報 F と F' の間の類似度 S_F は，3.1 節で示した差集合距離 g_{SDD} とそれぞれの指紋情報を構成するマニューシャ情報数 T，T'（$T < T'$）を用いて，次式で表せる．

$$S_F = T - g_{SDD}(F, F') \tag{16}$$

S_F の確率関数 $p_{S_F}(s)$ は，L と R が独立に一様分布に従うと仮定した場合，超幾何分布と二項分布の混合分布でモデル化される[17]．まず，2 つの指紋情報間において R は考慮せずに L のみで照合を行った際の一致マニューシャ情報数 S_L に着目すると，S_L の確率関数 $p_{S_L}(t)$，$t \in \{0, \ldots, T\}$ は，L のとりうる値の全パターン数 N_L を用いて，次式に示す超幾何分布でモデル化される．

$$p_{S_L}(t) = \frac{{}_{T'}\mathrm{C}_t \ {}_{N_L-T'}\mathrm{C}_{T-t}}{{}_{N_L}\mathrm{C}_T} \tag{17}$$

ただし，${}_iC_j$ は i 個の情報から j 個の情報を選択する組合せ数を示す．S_L の期待値 $E(S_L)$ は次式で表せる．

$$E(S_L) = \frac{TT'}{N_L} \tag{18}$$

次に，S_L 個のマニューシャ情報が同一と判定された場合の S_F の条件付き確率関数 $p_{S_F|S_L}(s|t)$，$s \in \{0, \ldots, S_L\}$ は，L が同一のマニューシャ情報間において R が一致する確率 ρ を用いて，次式に示す二項分布でモデル化される．

$$p_{S_F|S_L}(s|t) = {}_t\mathrm{C}_s \rho^s (1-\rho)^{t-s} \tag{19}$$

式 (17) および式 (19) より，$p_{S_F}(s)$ は次式に示す $p_{S_L}(t)$ と $p_{S_F|S_L}(s|t)$ の混合分布でモデル化される．

$$p_{S_F}(s) = \sum_{t=s}^{T} p_{S_L}(t) p_{S_F|S_L}(s|t) \tag{20}$$

式 (20) より，2 つの指紋情報が一致する確率は $p_{S_F}(T) = ({}_{T'}C_T / {}_{N_L}\mathrm{C}_T)\rho^T$ となり，F の Renyi エントロピー $H_2(F)$ は次式で表せる．

$$H_2(F) = -\log_2 \frac{{}_{T'}C_T}{{}_{N_L}\mathrm{C}_T} \rho^T \tag{21}$$

Pankanti らは，$p_{S_F}(s)$ の形状に関する各パラメータの値を指紋の構造から理論的に決定しているが，本論文では，3.3 節の手順に従い，他人間照合実験を通して得られる統計量を利用した評価方法を示す．N_L は，L のみに着目した他人間照合実験を通して得られる S_L の平均と式 (18) より推定する[22]．ただし，T，T' は，Pankanti らの報告と同様に，他人間照合実験で使用する指紋画像のデータベースを利用して，1 つの指紋情報に含まれるマニューシャ情報数の平均を調査し，それを近似値とする．また，ρ は，他人間照合実験を通して L が一致するマニューシャ情報数と L，R の両方が一致するマニューシャ情報数の割合を調査することで推定する．そして，それらの推定値と式 (21) より，$H_2(F)$ を算出する．

(平成 22 年 11 月 30 日受付)
(平成 23 年 6 月 3 日採録)

披田野清良（学生会員）

2007 年早稲田大学理工学部コンピュータ・ネットワーク工学科卒業．2009 年同大学理工学術院基幹理工学研究科修士課程修了．同博士後期課程在学中．2010 年日本学術振興会特別研究員（DC）．2011 年早稲田大学理工学術院基幹理工学部助手．生体認証のテンプレート保護技術に関する研究に従事．

赤尾　直彦

2011 年早稲田大学理工学術院基幹理工学部情報理工学科卒業．同大学基幹理工学研究科修士課程在学中．テンプレート保護型生体認証の安全性評価に関する研究に従事．

小松　尚久（正会員）

1979 年早稲田大学理工学部電子通信学科卒業．1981 年同大学院理工学研究科修士課程修了．同年 NTT に入社．1987 年早稲田大学理工学部助手．1989 年同専任講師．1996 年同教授．工学博士．セキュリティと品質を考慮したヒューマン/ネットワークインタフェースに関する研究に従事．特に，生体認証技術とその適用に興味を持つ．電子情報通信学会，画像電子学会，IEEE 各会員．

高橋　健太（正会員）

1998 年東京大学理学部情報科学科卒業．2000 年同大学院理学系研究科修士課程修了．同年（株）日立製作所入社．以来，同システム開発研究所（現，横浜研究所）にて生体認証技術および情報セキュリティ技術の研究開発に従事．平成 13 年情報処理学会高度交通システム研究会優秀論文賞受賞．平成 20 年度情報処理学会論文賞受賞．電子情報通信学会，IEEE 各会員．