

生体情報計測による感性可視化のための機械学習的クラスタリング手法の開発

富山県立大学工学部電子・情報工学科 指導教員：奥原浩之
1415004 麻生 到

1 はじめに

バイオメトリック認証とは、生体情報として指紋、顔、血管パターン、虹彩などの身体的特徴、あるいは筆跡、音声などの行動的特徴を用いて個人認証を行う技術である [1]。

特に、指紋は認証性能が高く、デバイスのログイン手法として、すでに iPhone などで普及されている。しかし、認証に必要となる生体情報が常に外部にさらされているため、生体情報を入手されやすく、実際に指紋認証を偽装した事件も起きている。

一方、脳波を生体情報として利用したバイオメトリック認証は偽装される危険性が非常に小さいと考える。脳波は個人により違った特徴を示すが、同一人物でもイメージするものによって変化する。この特徴を利用することで、生体情報であっても、イメージするものを変化させることで、生体情報を変化させることが可能となる [2]。

従来の脳波を用いた個人認証の研究では、単チャンネルにおけるウェアラブル脳波計を用いた研究であったり、多チャンネルを用いた研究では有線の脳波計を使用していた。

しかし、実際には多チャンネルを用いた脳波の個人認証は脳波を計測する上で計測される人とコストの面においても負担が大きく、あまり実用的ではないと考えている。また、他チャンネルの脳波計であると、個人によって大きく反応する部位が異なったり、効果的な部位がどこであるかを検討することは不可能である。

そこで本研究では、実用性の高い脳波を用いた個人認証システムの検証を行う。8 チャンネルのウェアラブル脳波計を使用した個人認証システムを作成と効果的な計測部位の検討を行う。

本研究によって提案された必要最低限の測定箇所を用いることで、脳波個人認証専用やコストのかからない認証デバイスのシステム作成を目指す。

2 従来研究

2.1 実験概要

単チャンネルと 16 チャンネルの多チャンネルを用いた認証システムの検証を行っていた。

また、認証システムを作成するにあたり被験者には深呼吸 (breathing)、指を上下に動かすイメージ (finger)，好きなスポーツをするイメージ (sport)，好きな歌を歌うイメージ (song)，機械音を聞き，その後一点を注視 (audio)，赤・緑・青・黄から一色選び，表示された画像の中に選択した色が含まれている個数をカウント (color)，好きなパス思考をイメージ (pass) のように 7 つのタスクを行ったときの脳波を計測していた。

2.2 特徴抽出

θ 波から γ 波までの 4-40Hz の脳波を使用した。

3Hz 以下の脳波には EOG(眼球運動) が多く含まれているため、解析では使用しない。 θ 波帯域を 2 領域、 α 波帯域を 3 領域、 β 波帯域を 5 領域、 γ 波帯域を 3 領域、計 13 領域に分割し、データごとに各領域のパワースペクトル平均を求めて、これを特徴量として利用した。

3 特徴抽出とクラスタリング

3.1 特徴抽出

本研究では、基礎律動と呼ばれる α 波と覚醒脳波である β 波を特徴量として用いる [3]。これらの脳波は周波数ごとに検出することが可能であるため、高速フーリエ変換による周波数解析を行い、特徴量を抽出する。

特徴量となるパワースペクトル PS は、高速フーリエ変換により得ることができる実部 Re と虚部 Im から以下の式で表される。

$$PS = \sqrt{Re^2 + Im^2} \quad (1)$$

また、それぞれの周波数領域を図 1 のようにいくつかに分割し、それぞれの領域での正規化し平均値を特徴量とした。

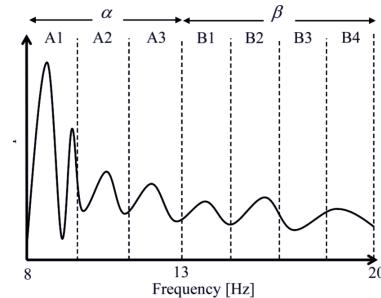


図 1 α 波と β 波の分割例

それぞれの領域におけるパワースペクトルの個数 i の合計をそれぞれのパワースペクトルで割ったものを正規化したパワースペクトルとした。

3.2 クラスタリング

本研究では、クラスタリングの手法としてランダムフォレストを活用した。

ランダムファレストは、複数の決定木を構成して識別などを行う機械学習アルゴリズムである。個々の決定木の性能は高くないが、それらを複数用いてそれぞれの結果を補うことで高い性能を得ることができる [4]。

データからランダムにサンプリングされたサブセットのデータに対して、これまたランダムに選択された特徴量などの属性を用いて複数の決定木を作成する。この複数の決定木による識別結果から最終的な結果を算出する。決定木は、属性によってデータ集合を部分集合に分けていき、最終的には各クラスごとの集合になるように分割することを目的としている。決定木によるデータの分割の指標として、エントロピーやジニ係数などが挙げられる [5]。エントロピーを用いた情報量 E は以下の式 (2) で表される。

$$H(S) = - \sum_{i=1}^n P_i \log 2P_i \quad (2)$$

P_i はクラス i のデータ数の割合を示す。ジニ係数を用いた情報量 E は以下の式 (3) で表される。

$$H(S) = 1 - \sum_{i=1}^n P_i^2 \quad (3)$$

これらの情報量を用いて、情報利得 I は以下の式 (4) に表すことができる。

$$I = H(S) - \sum_{i=1}^m \frac{S_i}{S} H(S_i) \quad (4)$$

S は学習する全体のデータ、 S_i は注目しているデータ、 m は分割するノード数、決定木が二分木の場合は $m = 2$ となる。

分類条件は、情報利得 I を図 2 のように最大となるよう分割していく。

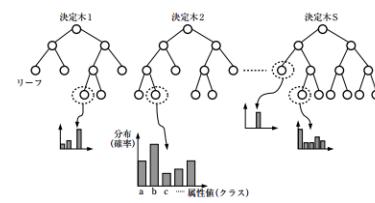


図 2 ランダムフォレストの例

4 脳波個人認証システムの構築

4.1 被験者ごとの脳波個人認証システムの提案

本研究では、じゃんけんの画像を利用した脳波個人認証システムを提案する。

提案する脳波個人認証システムの GUI を図 3 に示す。



図 3 GUI

データ収集では 2人の被験者に図 4 のようなそれぞれ 2秒間ずつ表示されるスライドを見て指示に従い、じゃんけんをしてもらう。最初のスライドは勝ち・負け・引き分けの 3種類、次のスライドはグー・チョキ・パーの 3種類のスライドの合計 9パターンを 1セットとした。

チェックは、取得したデータの判別が可能かどうかを確認する項目である。チェックのボタンを押すとじゃんけんゲームがはじまり、取得したデータをもとに被験者がグー・チョキ・パーのどれをだすかを予測してスライドを表示する。認証開始では、実際に被験者を認証できていればログイン成功、認証できなければログイン失敗とラベルに表示する。



図 4 スライドの例

被験者には、スライドの指示に従ってもらい、グー・チョキ・パーの画像が表示されてから 1秒間の脳波データを取得した。

4.2 解析手順

本研究の脳波計には OpenBCI を利用した。この脳波計では、5Hz から 50Hz のバンドパス処理されたデータを取得することができる。

取得したデータを高速フーリエ変換を行い、分割した周波数帯ごとに正規化したデータの平均を特徴量として、解析する。特徴量をランダムフォレストに適用して解析結果を取得する。その際、ランダムフォレストのパラメータはクロスバリテーションにより求める。最終的な結果は図 5 のように 7 個の分類器の F 値による重み付き多数決により算出する。



図 5 重み付き多数決

5 数値実験結果ならびに考察 ランダムフォレストのパラメータはクロスバリテーションによる学習により求めた。求めたパラメータは、決定木の数と木の深さ、データを分割指標である。探索範囲は、指標はエントロピーかジニ係数、深さは 1 から 6、決定木の数は 10,20,50,80,100,200,1000 の中に探索して精度の良くなるパラメータを用いた。探索結果を表 1 に示す。

表 1 探索結果

周波数帯	指標	木の深さ	決定木の数
α_1	ジニ係数	2	80
α_2	ジニ係数	2	20
α_3	ジニ係数	5	100
β_1	ジニ係数	2	200
β_2	ジニ係数	1	10
β_3	エントロピー	4	100
β_4	ジニ係数	1	50

7 個の分類器による重み付き多数決の結果は、個人識別においては 2人のためどちらかの判別は F 値が 1.0 という結果だったが、グー・チョキ・パーの判別まで細かく識別しようとすると F 値は 0.389 という結果となつた。

5.1 チャンネル数と測定部位の検討

ランダムフォレストによる説明変数の重要度を求めて、チャンネル数と測定部位の検討を行う。

本研究では、図 6 のような 10-20 法に従って脳波を測定した。

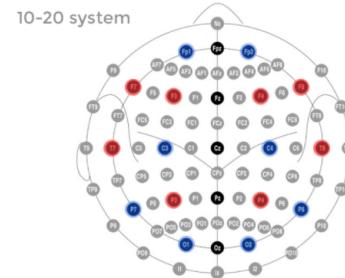


図 5 10-20 法

表 2 に測定部位の重要度を示す。

表 2 各チャンネルの重要度

	Fp1	Fp2	C3	C4	P7	P8	O1	O2
α_1	0.101	0.158	0.174	0.115	0.064	0.118	0.154	0.136
α_2	0.196	0.144	0.124	0.083	0.073	0.150	0.080	0.149
α_3	0.203	0.095	0.153	0.009	0.008	0.128	0.120	0.117
β_1	0.165	0.111	0.129	0.119	0.101	0.131	0.123	0.117
β_2	0.197	0.130	0.154	0.050	0.077	0.131	0.099	0.165
β_3	0.089	0.142	0.202	0.077	0.084	0.162	0.124	0.116
β_4	0.122	0.131	0.167	0.120	0.089	0.132	0.102	0.132

6 まとめと今後の課題

6.1 今後の方針

ジャンケンゲームによるチェックを行う。

学習データを増やすことで分類精度の向上を目指す。

参考文献

- [1] 小松尚久, 内田薰, 池野修一, 坂野鋭, バイオメトリクスのおはなし, 日本規格協会, 2008.
- [2] 石川由羽, 吉田智奈美, 高田雅美, 城和貴, 脳波個人認証のための特徴抽出, 情報処理学会研究報告, 2014
- [3] 鈴木良次, 佐藤俊輔, 池田研二, 吉川昭, 生体信号一計測と解析の実際一, コロナ社出版
- [4] 波部齊 ランダムフォレスト, 情報処理学会, 2012
- [5] 山岡啓介 知っておきたいキーワードランダムフォレスト