

# 長岡技術科学大学工学部課題研究報告書

## 題 目

骨伝導デバイスによる個人認証に関する研究

指導教員 准教授 杉田 泰則

著者 電気電子情報工学課程

21103591 小池 礼

令和 6 年 2 月 14 日 (金)

# 概要

情報処理技術の中には、生体認証と呼ばれる、人間の生体的な特徴を利用して個人を識別する技術が存在する。しかし最近のカメラやマイクの性能向上に伴い、従来の生体認証手段では容易に複製や模倣が可能になってきているという課題がある。

そこで新たな生体認証技術として骨導伝達特性に着目した研究が注目を集めている。この特性は、他人によって容易に入手できない上に、模倣が困難であるため、生体認証において非常に有望な特徴とされる。

本研究では、骨伝導デバイスを用いて行う骨伝導特性での個人認証が可能なものの調査を行った。また個人認証の精度についての調査をデータ処理の方法を変化させながら行い、最適な手法を検討した。

骨伝導デバイスの脱着によって個人識別の精度は低下するという結果が得られた。またフィルタバンクと DCT を用いて処理を行った。フィルタバンクは 50～1000 の間で調査を行ったが精度は全ての場合で悪くなる結果が得られた。DCT は 10～400 の間で調査を行ったが DCT:10 以外の場合で精度が良くなった結果が得られた。研究全体で最も精度が良かった結果はフィルタバンク:なし、DCT:400 の場合で平均正答率は 97.5%となった。また分散の結果から個人の特徴は低周波帯域に大きく依存している可能性が高いことが分かった。今後の展望として個人識別に重要な周波数帯域の調査を行い、不要な情報を除去することでさらなる精度向上が期待される。また音楽や非可聴音での認証について調査が検討される。

# 目次

第1章	はじめに	1
1.1	研究背景	1
1.2	研究目的	2
1.3	本報告書構成	2
第2章	理論的背景	3
2.1	フィルタバンク	3
2.2	離散コサイン変換	4
2.3	ユークリッド距離	4
2.4	t-SNE	5
第3章	実験	6
3.1	実験条件	6
3.2	評価指標	8
3.3	実験結果	8
3.3.1	デバイス脱着での精度の変化	8
3.3.2	フィルタバンク、DCTによる精度の変化	11

3.3.2	周波数領域での個人差の調査	14
3.3.3	メルフィルタバンクでの調査	15
第4章	むすび	17
	謝辞	17
	参考文献	18

# 第1章 はじめに

## 1.1 研究背景

情報処理技術の中には、生体認証と呼ばれる、人間の生体的な特徴を利用して個人を識別する技術が存在する。近年ではパスワードの代わりに様々な場面でこの生体認証技術が多く使われている。例えば、スマートフォンのロック解除、金融取引の認証、さらには公共交通機関での支払いシステムなど、多くの用途に活用されている。生体認証には指紋[1][2]や声紋[3]、静脈[4]、虹彩[5]、顔[6]などを使用したものがあり、これらは時間が経過しても変化が少なく、長期間にわたり利用できるという利点がある。また、生体特徴は物理的なパスワードやカードのように紛失する心配がなく、さらに記憶を必要としないため、パスワードを忘れるといった問題を回避できる点も大きな特徴である。

しかし最近のカメラやマイクの性能向上に伴い、従来の生体認証手段では容易に複製や模倣が可能になってきているという課題がある。生体認証の信頼性を高めるためには、用いる生体特徴が他人によって入手困難であり、模倣が極めて難しいものである必要がある。このような課題を背景に、新たな生体認証技術として骨導伝達特性に着目した研究が注目を集めている。骨導伝達特性とは、音が頭蓋骨を通じて伝わる際の特性である。この特性は、他人によって容易に入手できない上に、模倣が困難であるため、生体認証において非常に有望な特徴とされる。既存研究によって、骨導伝達特性が個人特有である[7]ことはある程度確認されているものの、実際にそれを用いて個人を認証できるかどうか、また認証時にどの程度の精度が得られるかについては、十分に調査が行われていない。これらの課題を解決するためには、より詳細な分析や実験的な検証が求められている。骨導伝達特性を用いた生体認証の可能性を探ることは、今後のセキュリティ分野において重要な研究課題の一つとなると考えられる。

## 1.2 研究目的

本研究では骨伝導デバイスを用いて行う骨伝導特性での個人認証が可能なのかの調査を行う。また個人認証の精度についての調査をデータ処理の方法を変化させながら行い、最適な手法を検討する。

## 1.3 本報告書の構成

本報告書の構成を以下に示す。

第2章では、フィルタバンク、離散コサイン変換、ユークリッド距離、t-SNE についての説明をする。

第3章では、骨伝導デバイスを用いた個人認証の実験結果について評価する。

第4章では、本報告書のまとめと今後の展望について述べる。

## 第2章 理論的背景

### 2.1 フィルタバンク

フィルタバンクが応用されている例として音声信号処理、画像処理、通信システムなどが主にある。フィルタバンクとは、入力信号を複数の帯域に分解しそれぞれの成分を抽出するためのフィルタの集合のことである。フィルタバンクには主に2種類に分類され、周波数帯域を均等に分割するものと周波数帯域を不均一に分割するものがある。フィルタバンクを使うメリットとしては、人が認識しにくい周波数成分を除去できるためデータ圧縮ができることや、目的の信号のみを抽出しノイズ抑制ができるなどがある。

本研究ではフィルタバンクにより周波数帯域を均等に分割することによって、計算コストの削減やノイズの削減することを目的として使用している。また図 2.1 のような三角形の形状のフィルタがオーバーラップすることで周波数成分間の連続性が確保されるため急激な変化を抑えられる。これが長方形のフィルタであるとオーバーラップがない状態となり特定の周波数成分が1つのフィルタにのみ影響を与えるため、局所的なノイズの影響を受けやすくなってしまう。

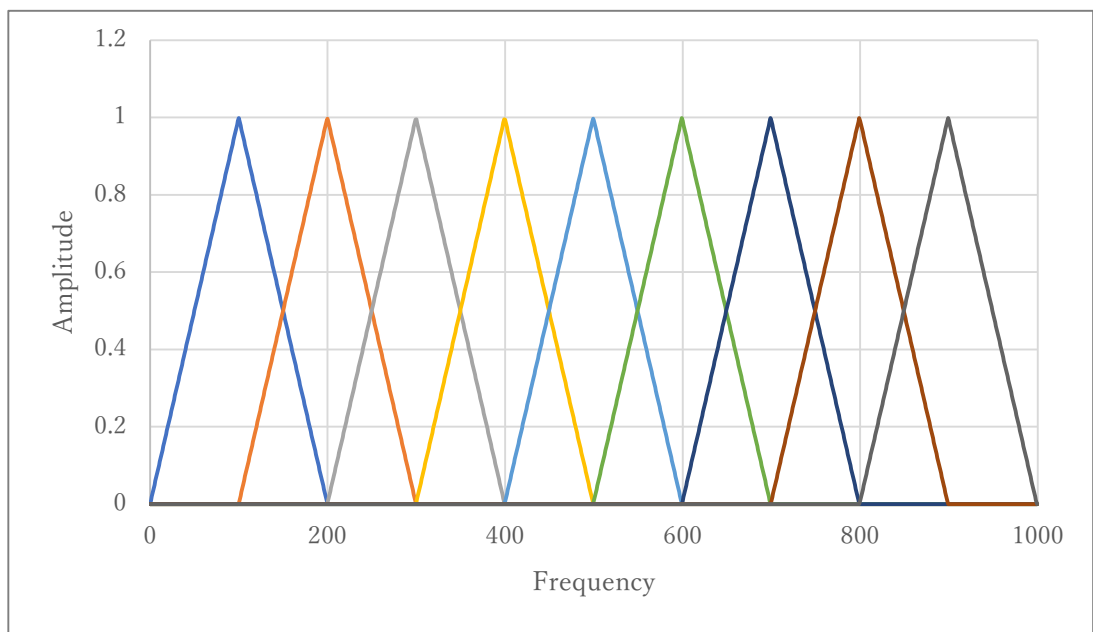


図 2.1 フィルタバンク例

## 2.2 離散コサイン変換

離散コサイン変換(DCT)は信号処理やデータ圧縮の分野で広く使用されている。特に JPEG 画像圧縮や MPEG 動画圧縮において重要な役割を果たしデータ圧縮効率向上や特徴抽出に向いている。また DCT は信号や画像の重要な情報を少数の係数に集中させる特徴がある。これによりデータ圧縮時に重要な部分を保持しつつ、ノイズなどの不要な部分を削除することが可能である。

DCT はフーリエ変換と類似した周波数解析手法であるが異なる点として基底関数が実数のコサイン波のみで構成されるといった特徴がある。DCT は主に以下の数式で定義される。

長さ $N$ の離散信号 $x_n$ に対して DCT は次式となる。

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right]$$

ここで、 $X_k$ は変換後の DCT 係数である。

## 2.3 ユークリッド距離

ユークリッド距離は、数学及びデータ分析の分野で広く使用されている距離の計算方法である。空間内の2点間の直線的な距離を計算する方法で2点間のユークリッド距離は次のように定義される。

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

ここで $\mathbf{x}$ と $\mathbf{y}$ は $n$ 次元空間ベクトルを示し、 $x_i$ と $y_i$ はそれぞれの座標軸における位置を表します。距離は、各次元の差の2乗を合計した後、平方根を取ることで求められる。



## 2.4 t-SNE

t-SNE は次元削減手法であり、高次元データを低次元空間に可視化するために広く用いられる。t-SNE の基本的な考え方は、高次元空間におけるデータ点の類似性を低次元空間に適切に埋め込むことである。これを実現するために、t-SNE は以下の処理を行う。

### 1. 高次元空間での類似度計算

各データ点 $\mathbf{x}_i$ に対して他のデータ点 $\mathbf{x}_j$ との類似度を条件付き確率として定義する。高次元空間における点 $\mathbf{x}_i$ の周囲に点 $\mathbf{x}_j$ が存在する確率 $p_{j|i}$ は、ガウス分布を用いて次のように定義される。

$$p_{j|i} = \frac{\exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|\mathbf{x}_i - \mathbf{x}_k\|^2 / 2\sigma_i^2)}$$

### 2. 低次元空間での類似度計算

高次元空間の類似度分布 $\mathbf{P}$ を低次元空間に対応する類似度分布 $\mathbf{Q}$ に近づけるように埋め込みをする。低次元空間では、ユークリッド距離に基に t 分布を用いて類似度を定義する。

$$q_{ij} = \frac{(1 + \|\mathbf{y}_i - \mathbf{y}_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|\mathbf{y}_k - \mathbf{y}_l\|^2)^{-1}}$$

t 分布を用いることで、低次元空間においてデータ点が遠くにある場合にも適切な類似度を持つようにできる。

### 3. コスト関数の最適化

高次元空間と低次元空間の類似度分布 $\mathbf{P}$ と $\mathbf{Q}$ の間の情報損失を最小化するために勾配降下法を用いて最適化する。

$$C = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}$$

## 第 3 章 実機実験

### 3.1 実験条件

実験条件を表 3.1 に示す。また使用機器については表 3.2 に示す。

骨伝導イヤホンと振動ピックアップの位置関係については図 3.1 に示すように、骨伝導イヤホンに振動ピックアップを固定し相対位置が変わらないようにしている。また図 3.2 のように耳の前方で骨を振動させ耳の後方で振動を測定した。

表 3.1 実験条件

被験者数	14 名
試行回数	脱着あり/なし 各 30 回
サンプリング周波数	48kHz
切り出したインパルス応答のデータ数	8193

表 3.2 使用機器

骨伝導イヤホン	AFTERSHOKZ AEROPEX
振動ピックアップ	小野測器 NP-2000 シリーズ
オーディオインターフェイス	Roland OCTA-CAPTURE UA-1010
ヘッドホンアンプ	Audio-technica AT-HA21
チャージアンプ	小野測器 CH-1200A



図 3.1 実験デバイス



図 3.2 装着イメージ

## 3.2 評価指標

インパルス応答のデータ 30 回分のうち 10 回分を学習データ、20 回分をテストデータとした。インパルス応答は FFT により周波数領域で処理を行った。個人を判別する際にはユークリッド距離を利用してテストデータがどの被験者に近いか判別し、実際にその被験者のデータに該当するのかが判断したのち正答率として結果を得た。

また t-SNE を用いて学習データを 2 次元で可視化し、個人のデータが他人のものと類似性を見た。

## 3.3 実験結果

### 3.3.1 デバイス脱着での精度の変化

デバイスの脱着によって生じる測定位置の些細な変化がどれほど正答率に影響を及ぼすかを確認するために実験を行った。テストデータがどの学習データに近いかを判別しその正答率を以下の混同行列にまとめ、t-SNE での可視化結果についても示した。脱着なしでは全体の平均正答率は約 99.6% となったが、脱着有では平均正答率は約 89.6% であった。t-SNE で可視化した結果でも脱着ありのほうがばらつきの大きい結果が得られた。また t-SNE の結果からも脱着ありのほうが個人ごとのばらつきが大きいことが見られる。これらの結果はデバイスの脱着による装着位置のずれから生じるものだと考えられる。

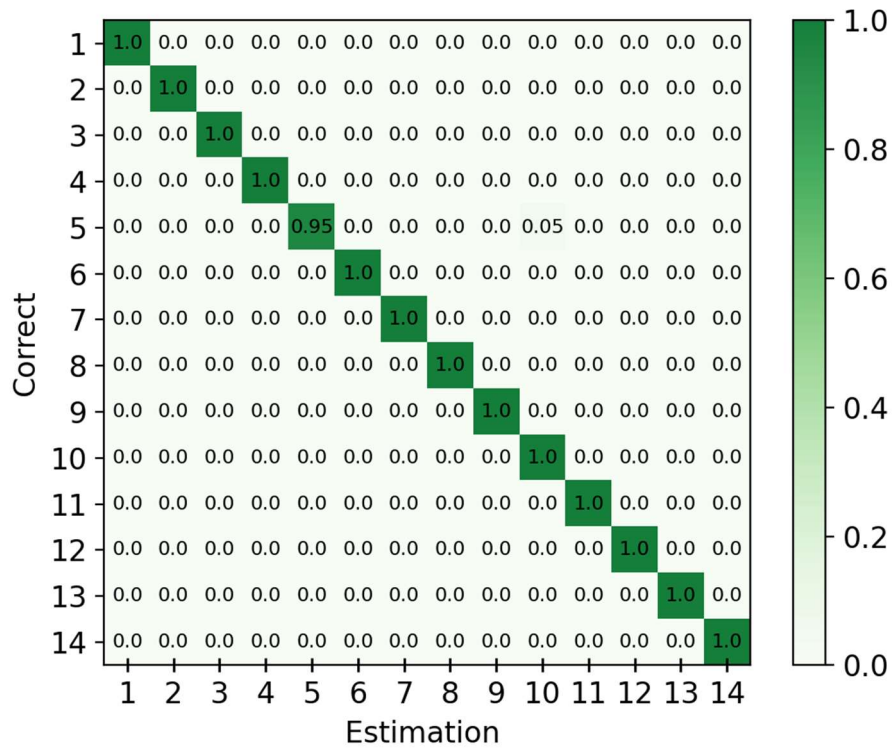


図 3.3 脱着なしでの正答率

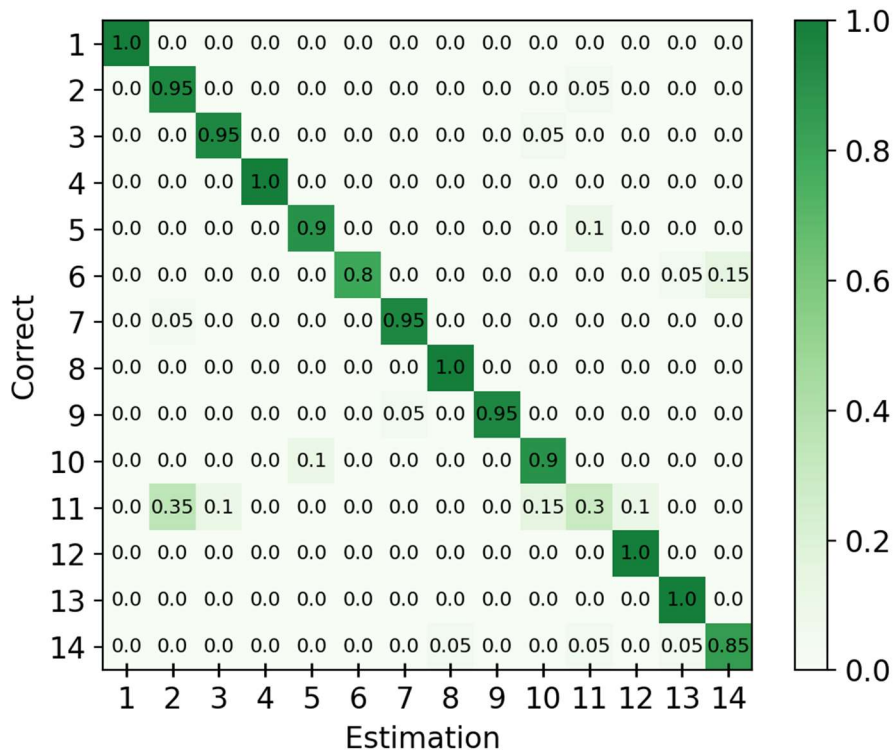


図 3.4 脱着ありでの正答率

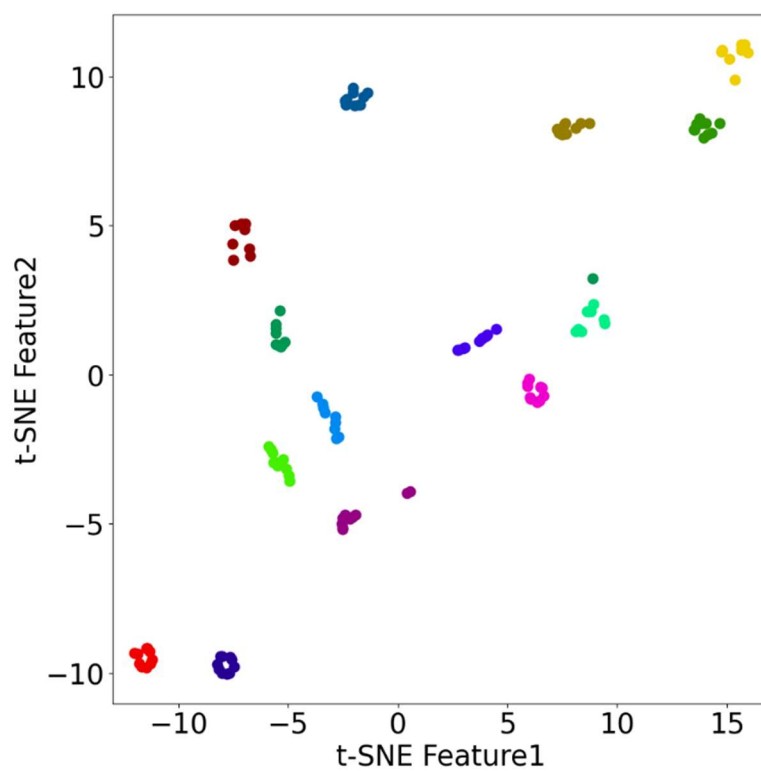


図 3.5 t-SNE(脱着なし)

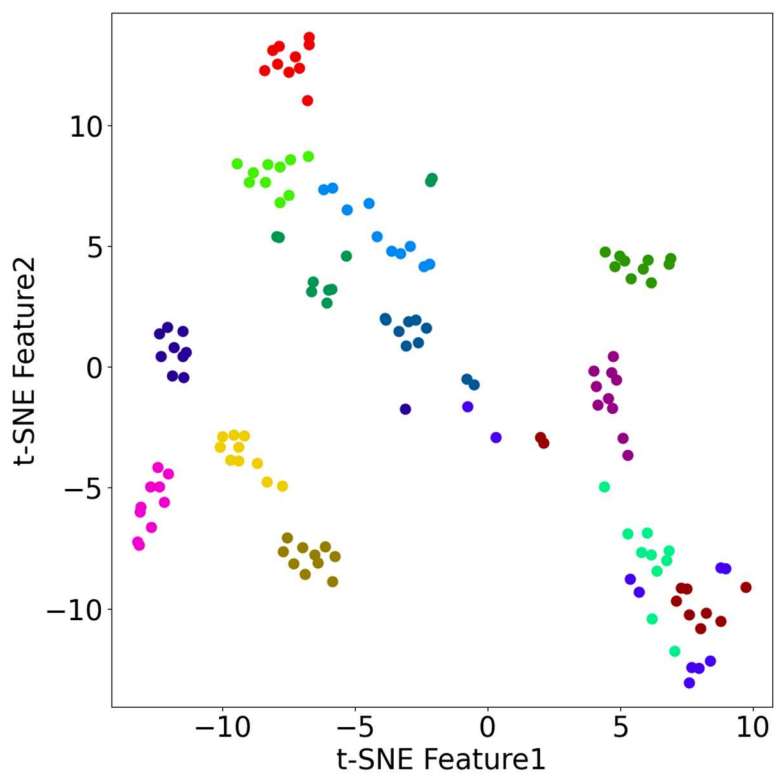


図 3.6 t-SNE(脱着あり)

### 3.3.2 フィルタバンク、DCT による精度の変化

デバイスの脱着ありの場合でフィルタバンクと DCT をかけて、精度の比較を行った。フィルタバンクは 50~1000 の間で変化をさせ、DCT は 10~400 の間で変化をさせた。これらの正答率をまとめたグラフ、混合行列、t-SNE での可視化結果を以下に示す。混合行列ではフィルタバンク:1000 のみの場合と DCT:400 のみの場合の結果であり、フィルタバンクを使用した場合には平均正答率が約 88.2%、DCT では平均正答率が 97.5%であった。t-SNE での可視化ではフィルタバンク:1000 の結果より個人の集合が他人のものと重なる部分があるため、ほかの結果に比べると個人を分類することが難しいことが分かる。フィルタバンクでの正答率が低くなった理由としてはフィルタリングによって個人差が現れる情報を損失させてしまっていることが考えられる。また DCT では低周波帯域が信号の主要な情報を保持するため個人識別に重要な情報が多く残り精度が良くなったと考えられる。また高周波帯域のノイズが抑制されたことも精度が高くなった要因であると考えられる。

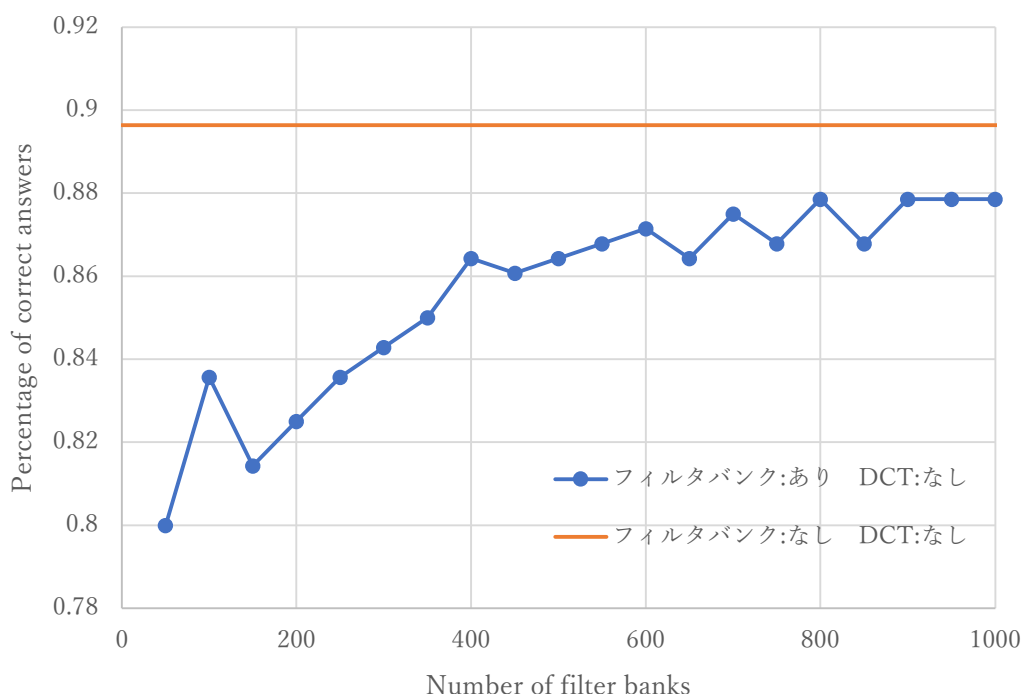


図 3.7 フィルタバンクありでの正答率

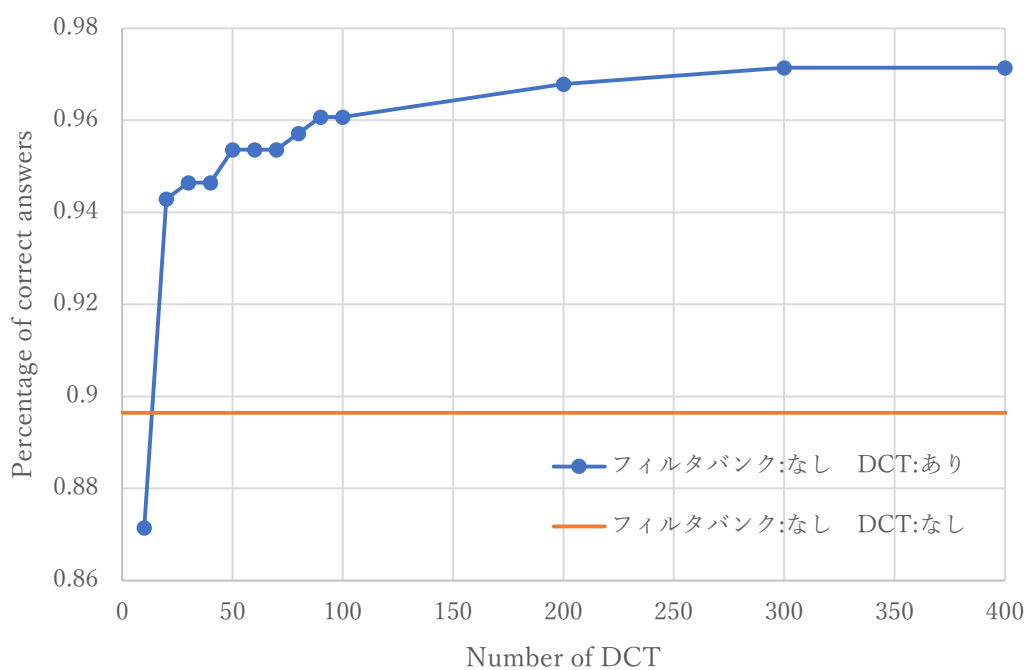


図 3.8 DCT ありでの正答率

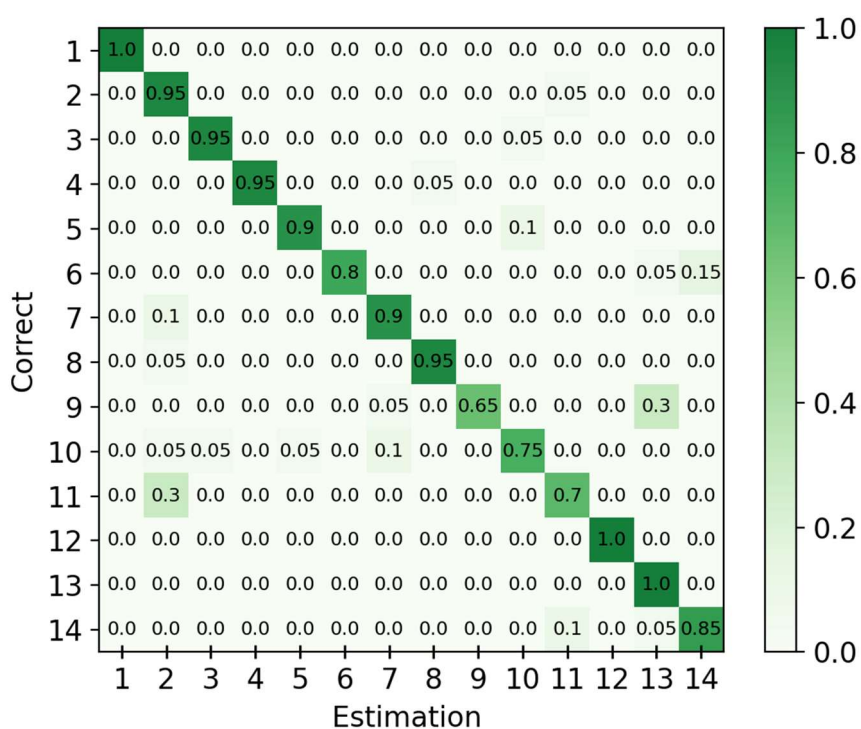


図 3.9 フィルタバンク:1000(脱着あり)での正答率



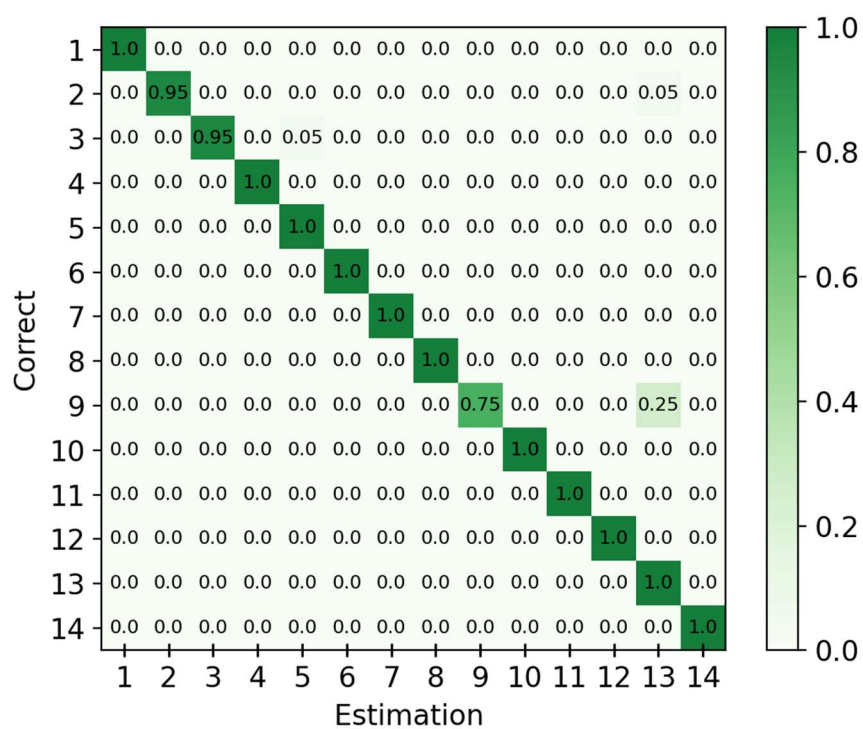


図 3.10 DCT:400(脱着あり)での正答率

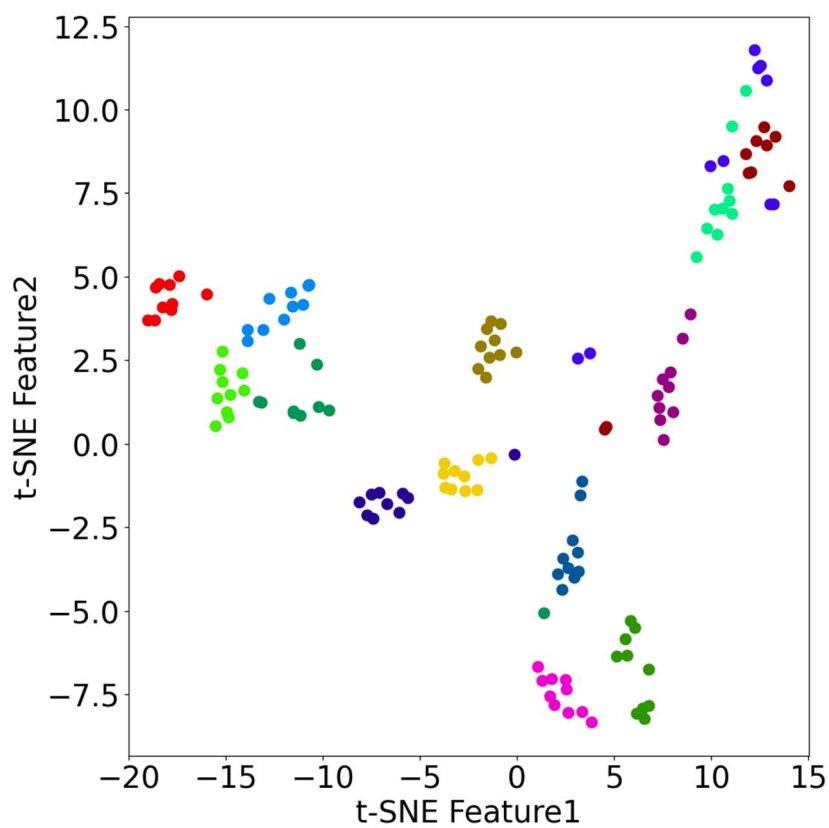


図 3.11 フィルタバンク:1000(脱着あり)での t-SNE

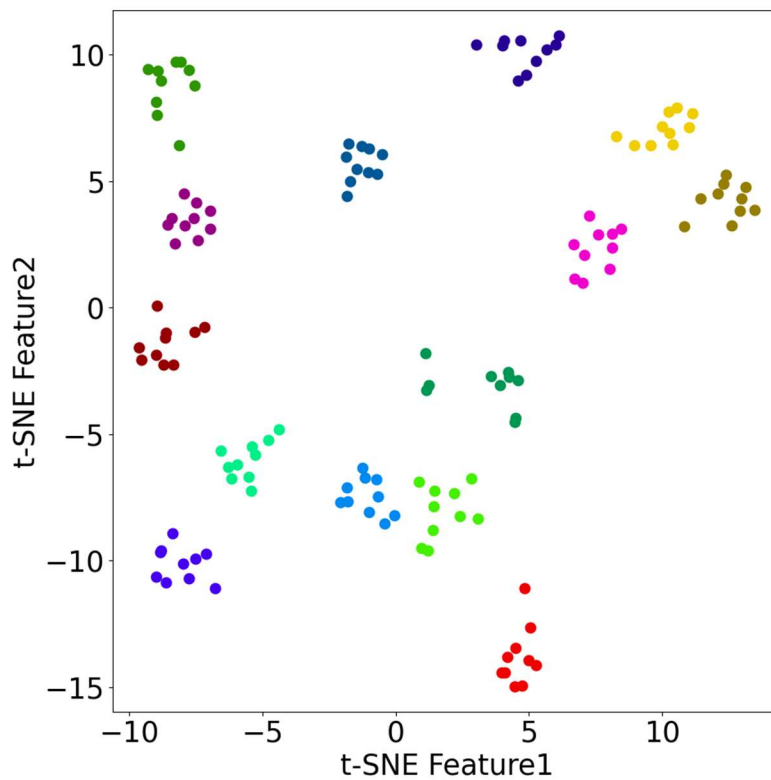


図 3.11 DCT:400(脱着あり)での t-SNE

### 3.3.2 周波数領域での個人差の調査

周波数領域ですべての被験者に対する分散をとった結果を以下に示す。グラフの横軸は周波数[Hz]を示している。

分散が大きく表れたのは低周波帯域であり個人ごとの特徴はこの部分に大きく依存していると考えられる。

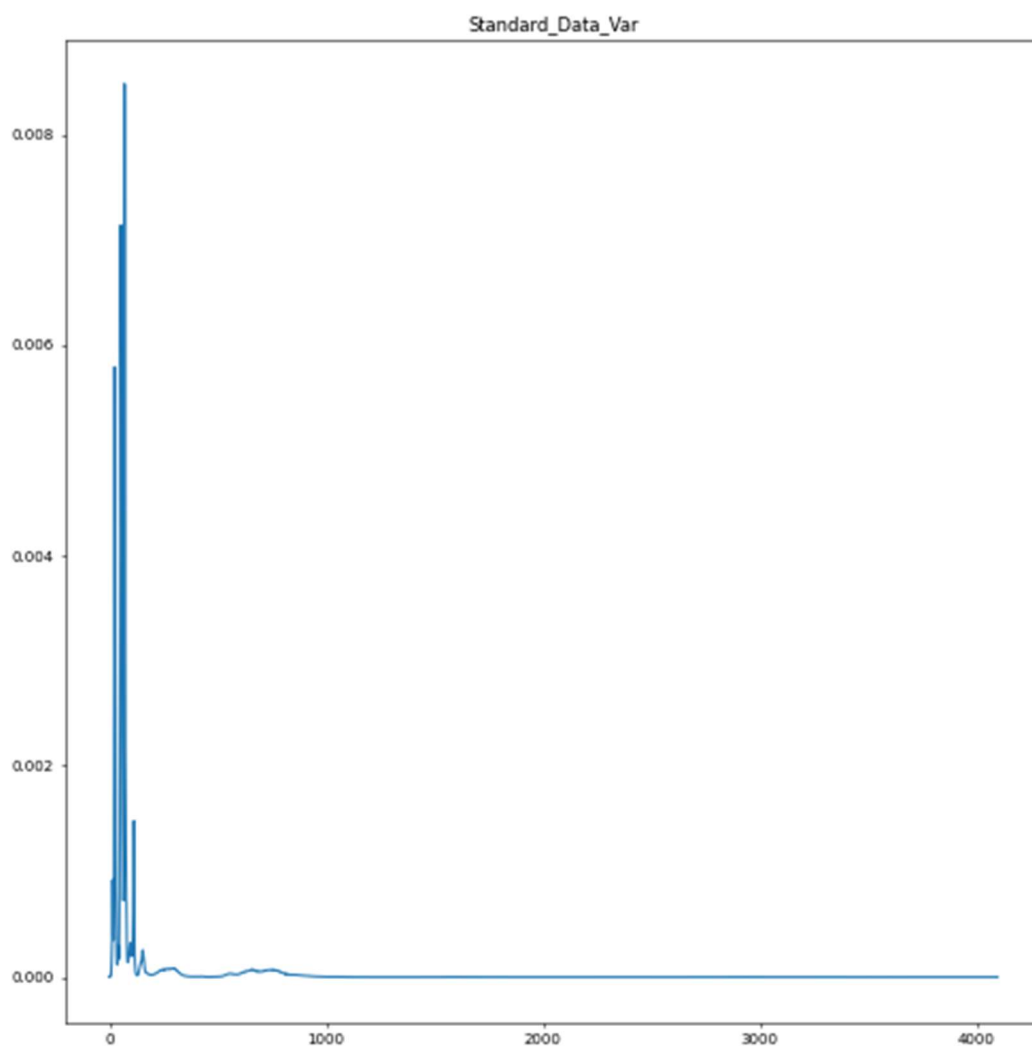


図 3.12 全被験者の分散

### 3.3.3 メルフィルタバンクでの調査

人間の聴覚特性を模倣した周波数スケールであるメル尺度に元図いてフィルタバンクを生成し、それを使用して処理を行った結果を以下に示す。

メルフィルタバンクを用いた場合であっても正答率は、メルフィルタバンクなしより低くなってしまいう結果となった。また平均正答率は約 85.3% となりかなり低い結果となった

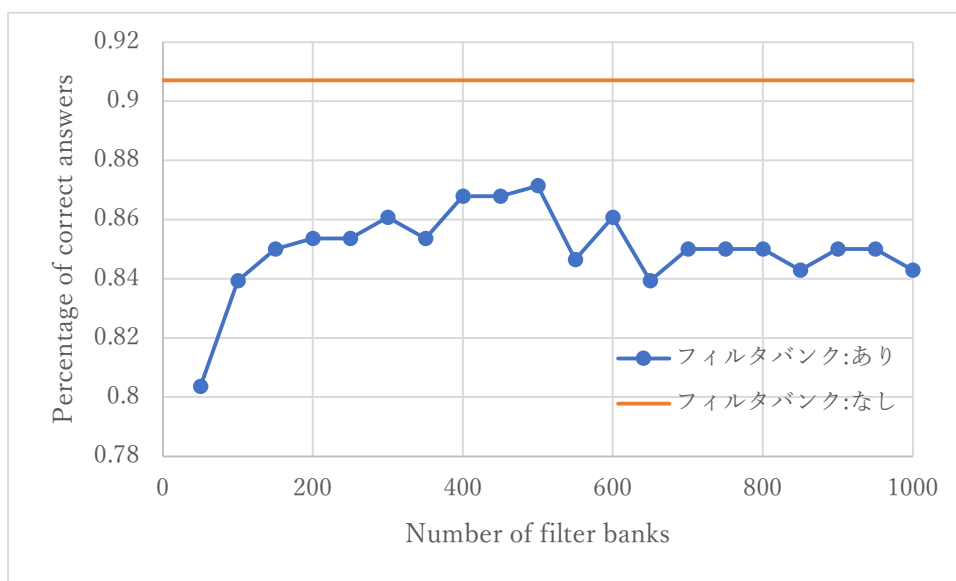


図 3.13 メルフィルタバンクでの正答率

## 第4章 むすび

本報告書では、骨伝導デバイスによる生体認証が可能なのか調査、また制度についても調査を行った。既存研究では骨導伝達特性には個人特有のものがあることが分かっている。そこで本研究で実際に骨伝導デバイスを用いてインパルス応答を測定し個人識別を行った。骨伝導デバイスの脱着によって個人識別の精度は低下するという結果が得られた。またフィルタバンクと DCT を用いて処理を行った。フィルタバンクは 50～1000 の間で調査を行ったが精度は全ての場合で悪くなる結果が得られた。DCT は 10～400 の間で調査を行ったが DCT:10 以外の場合で精度が良くなった結果が得られた。研究全体で最も精度が良かった結果はフィルタバンク:なし、DCT:400 の場合で平均正答率は 97.5%となった。また分散の結果から個人の特徴は低周波帯域に大きく依存している可能性が高いことが分かった。

今後の展望として個人識別に重要な周波数帯域の調査を行い、不要な情報を除去することでさらなる精度向上が期待される。また音楽や非可聴音での認証について調査が検討される。

## 謝辞

本研究を進めるにあたり、日ごろから様々なご指導ご鞭撻を賜りました、杉田泰則准教授に深く感謝いたします。また、本報告書の審査において貴重なご意見やお時間を賜りました、圓道知博教授に御礼申し上げます。そして本研究に関してご指導、ご意見、ご協力下さった信号処理応用研究室の皆様には感謝の意を表します。

## 参考文献

- [1] R. S. Germain, A. Califano, S. Colville, "Fingerprint Matching using Transformation Parameter Clustering", IEEE Computational Science and Engineering, Vol. 4, No. 4, pp. 42-49, 1997.
- [2] A. K. Jain, S. Prabhakar, L. Hong, "A Multichannel Approach to Fingerprint Classification", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 21, No.4, pp.348-359, 1999.
- [3] M. H. George, R.A. King, "A Robust Speaker Verification Biometric", Proceedings of IEEE 29th Annual 1995 International Carnahan Conference On Security Technology, pp. 41-46, 1995.
- [4] N. Miura, A. Nagasaka, T. Miyatake, "Feature Extraction of Finger-vein Patterns based on Repeated Line Tracking and Its Application to Personal Identification", Machine Vision and Applications, Vol. 15, No. 4, pp. 194-203, 2004.
- [5] Y. Zhu, T. Tan, Y. Wang, "Biometric Personal Identification based on Iris Patterns", Proceedings of 15th International Conference on Pattern Recognition, Vol. 2, pp. 801-804, 2000.
- [6] A. M. Martinez, H. Jia, "Support Vector Machines in Face Recognition with Occlusions", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Neural Networks, Vol. 14, pp. 795-803, 2001.
- [7] 貝沼 遼太郎, 矢野 昌平, 和久井 直樹 (2022) 骨伝導型個人認証デバイスによる個人毎の差異に関する検討 電気情報通信学会誌,101(3),123-130
- [8] van der Maaten, L., & Hinton, G. (2008). "Visualizing Data using t-SNE". Journal of Machine Learning Research, 9, 2579-2605.
- [9] 山下 幸彦 (2013) 画像符号化のための線形変換 映像情報メディア学会誌 Vol.67, No.2, pp.131-135
- [10] Van der Maaten, L. (2008). *Dimensionality reduction: A comparative review*. Journal of Machine Learning Research, 9, 1-50. Retrieved from