

ウェーブレット変換を用いた音声認証の実現可能性についての検討

田中 英之*

笠井正三郎†

Study on Realization of Voice Authentication Using Wavelet Transform

Hideyuki TANAKA*

Shozaburo KASAI†

ABSTRACT

Recently, biometric authentication is spreading to strengthen security. The voice authentication is a kind of biometric authentication and extracts the individual feature from the voice using Fourier Transform or Short-Time Fourier Transform(STFT). Continuous Wavelet Transform(CWT) is a method of analyzing signal such as Fourier transform. CWT is similar to STFT. But CWT has changeable resolution of frequency and time which STFT doesn't. In this paper, we show a result of analysing voices using Wavelet transform. At first, because the the image of transformed voice is periodic, we compare a cycle image of stable part(middle of voice). And, we compared a contour of whole image. But we can't find the individual feature. So, we use Wavelet Packet Analysis instead of CWT. Wavelet Packet Analysis is a kind of Wavelet Transform. As a result, we find the individual feature like STFT. Finally, we think the voice authentication is able by using Wavelet Packet Analysis.

Keywords : biometrics authentication, wavelet transform

1. 緒論

現在、認証方法として最も多く普及しているのはパスワード認証である。パスワード認証にはとても安価にシステムを構成することができるなどの利点があるが、パスワードが盗まれたり、本人が忘れてしまうなどの欠点も存在する。近年では、このような欠点を突いて他人が本人になりすます犯罪も増えてきており、それを防止するために、セキュリティの強化が見込めるバイオメトリクス認証を採用する機会が増えてきている。バイオメトリクス認証とは個人の身体的特徴から認証を行う方法である。身体的特徴は他人に利用されにくいいため、バイオメトリクス認証はパスワード認証よりも強固な認証システムであるといえる。

本研究ではバイオメトリクス認証のひとつである音声認証に着目した。音声認証は、システム構築に必要なハードウェアがマイクだけでよい、電話越しの認証が行える、指紋などに比べて個人データの登録に対する心理的抵抗が少ないなどのハードルの低さから、バイオメトリクス認証技術の中では比較的導入しやすいとされている。現在使用されている音声認証の多くはフーリエ変換や短時間フーリエ変換を用いて、周波数スペクトル、後者ではさらにその時間変化も抽出し、本人確認が行なわれている。

フーリエ変換などと同じ信号解析の手法として、ウェーブレット変換が存在する。ウェーブレット変換は、信号解析の手法としては比較的新しい技術であり、これに関する研究はフーリエ変換と比べればまだあまり行なわれていない。ウェーブレット

変換は、短時間フーリエ変換と同じく、周波数と時間を同時に軸にとって解析を行うことができる。しかし、短時間フーリエ変換と全く同じというわけではなく、分解能が可変である点などの利点がある。ウェーブレット変換はその性質上、短時間信号や時間軸の情報が重要な意味を持つ信号の解析に有利であると考えられる。本研究ではウェーブレット変換を用いた音声認証の実現可能性について検討した。本論文ではその検討方法、検討結果を報告する。

2. ウェーブレット変換⁽¹⁾⁽²⁾

ウェーブレット (Wavelet) の語源は「波 (wave)」と「小さな (let)」の複合語であり、「小さなさざなみ」を意味する。三角関数や対数関数のような特定の決まった関数ではなく、局在する波を表すさまざまな関数の総称であり、「時間的にも周波数的にも局在性のある関数」のことである。

ウェーブレットの一例として、メキシカンハット基底 $\psi(t)$ を図 1 に示す。このようなウェーブレットを拡大縮小、および平行移動したものをを用いて、観測パターンとの相似性により、時

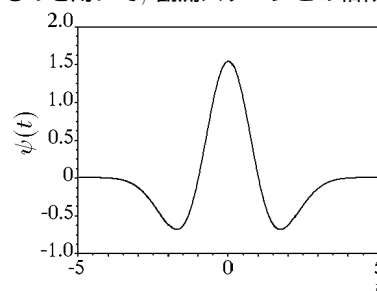


図 1 メキシカンハット基底

*電気電子工学専攻 2 年

†電子工学科 教授

間周波数特性の解析を行うものをウェーブレット解析と呼ぶ。

2.1 連続ウェーブレット変換 連続ウェーブレット変換 (continuous wavelet transform) は、2乗可積分区間の空間 $L^2(R)$ に属する関数 $\psi(t)$ を基底として、 $L^2(R)$ に属する任意の信号を表現する手段である。このとき ψ を、全体の積分値が0で時間的に(原点 $t=0$ の周りに)局在する関数と考える。この関数 $\psi(t)$ をマザーウェーブレット、あるいは単にウェーブレットと呼ぶ。これを拡大縮小(スケール)、およびシフトすることにより、式(1)のように基底 $\psi_{a,b}(t)$ を生成する。

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (1)$$

ここで、 a はスケール、 b はシフトのパラメータであり、 $1/a$ が周波数、 b が時間シフトに対応する。また、 $1/\sqrt{a}$ は正規化のための係数である。ウェーブレット変換は式(1)と信号 $f(t)$ との内積であるため、式(2)で表される。

$$(W_{\psi}f)(b,a) = \frac{1}{\sqrt{a}} \int_R f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (2)$$

ここで、 $\overline{\psi(t)}$ は $\psi(t)$ の複素共役を表している。連続ウェーブレット変換では、式(2)において、パラメータ a, b が連続的な値をとる。基底としては既に様々なものが考案されており、条件を満たす関数ならば、新しい基底として定義することもできる。そのため、解析目的、解析対象に適した基底を選択することができる。

3. 解析方法

解析対象となる音声信号はフリーソフトの SoundEngine Free を用いて、モノラル wave ファイルとして録音した。10人から母音である「あ」行を2回ずつ、いずれも自然に発声したものを録音した。「あ」は音高が変わるように発声した場合など追加録音したものもあるため37個、他の音に関しては全て20個、計117個のサンプルを録音した。サンプリング周波数と量子化ビットは電話音質を意識し、8kHz、8bitを採用しようと考えたが、電話で話す際に相手が誰か判別できない事もあるため、サンプリング周波数はより高く、11.025kHzとした。量子化ビットについては、8bit、16bitの2つのサンプルをウェーブレット変換して比べた結果、特に差異はなかったため、8bitを採用した。本報告書にて解析を行った結果は主に「あ」に対してであるため、音声について何も明示されていない場合は「あ」についての解析結果である。また、解析には Matlab, scilab 及び、それぞれ Toolbox である WaveletToolbox を用いた。

3.1 変換における基底の選択 ウェーブレット変換において、解析に適した基底の選択は重要である。そこで、同じ音声に対して、いくつかの基底を用いてウェーブレット変換を施し、その結果から用いる基底を Coiflets 基底(以下 coif と略す)に決定した。coif 以外の基底の詳細については参考文献(3)を参照されたい。MexicanHat 基底や、Haar 基底、Daubechies 基底(以下 db 基底と略す)なども検討を行ったが、それぞれ次のような問題があったため使用しなかった。MexicanHat 基底の結果は約100以下の比較的低いスケールに情報が偏った結果となっていた。スケールの刻み幅とスケールの幅を半分にするれば

無駄な結果を省くことができるが、他の基底と同じ条件で比較すると高周波成分などの特徴が把握しにくいという事が考えられた。Haar 基底の結果では値の大きい部分がばらついているため、どのスケールの成分が強いのか把握しにくいという事が考えられた。db 基底は coif 基底と同じく複数のバリエーションが存在しており、その変換結果も coif 基底に近いものがあったが、coif を用いた場合の変換結果の高スケール(低周波)部分を歪めたような結果となったため、coif 基底を使用する事にした。

3.1.1 Coiflets 基底 (coif 基底) Coiflets 基底は高次の消滅モーメントを持ち、式(3)を満たすように考案された基底である⁽⁴⁾。

$$\int x^l \psi(x) dx = 0, \quad l = 0, \dots, L-1 \quad (3)$$

ここで、 L は Coiflets の次数と呼ばれる。次数が高い基底ほど高い周波数成分を含むため、次数は高い方が高周波の解析に適していると考えられる。図2に $L=5$ である coif5 基底の波形、図3(a)に変換対象の「あ」の波形と同図(b)に coif5 を用いた変換結果を示す。ただし、変換対象と変換結果を全て表示すると形がわかりにくくなるため、それぞれ部分的に示している。図3の横軸は時間軸でデータの数が250個、(a)の縦軸は

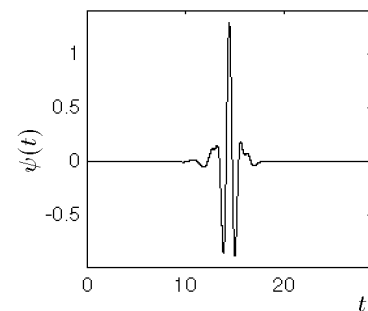
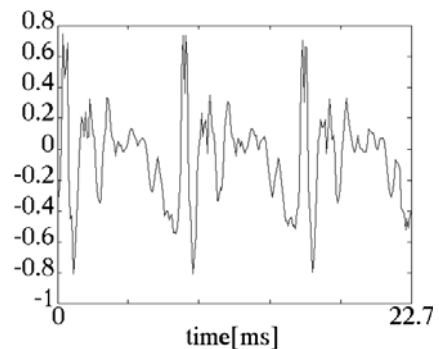
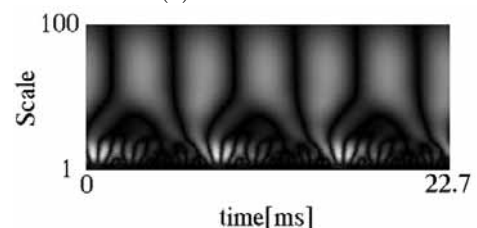


図2 Coiflets5 基底



(a) 信号強度波形



(b) coif5 によるウェーブレット変換結果

図3 「あ」音の信号

信号の強度, (b) の縦軸はスケールの範囲で 1 ~ 100 の目盛りを取る. 横軸は 1 目盛り $90.7\mu s$ で, 最大約 $22.7ms$ となる.

4. 音声信号のウェーブレット解析

4.1 音声の中間部に対する検討 音声をウェーブレット変換した結果は図 3(b) を見てもわかるように, もとの音声の周期的なために周期的な画像となった. そこで, 変換画像において 1 周期を定義し, 比較する事で認証が行なえるか検討する. ただし, 発声開始や, 終了付近では安定した特性が得られないため, もっとも安定していると考えられる中間に存在する 1 周期を比較対象とした.

4.1.1 1 周期の定義 本研究における 1 周期の定義は複数の変換画像から次のように決定した. まず, 変換画像のスケール 70 までの範囲に注目する. 変換画像を閾値 0 で 2 値化すると図 4 のようになる. 次にラベリングを行い, スケール 70 に存在する山に注目する. この時注目した山と同一のラベルをもつ最も右に存在する画素と, 注目した山より 1 つ左の山と同一のラベルをもつ最も右に存在する画素の間を 1 周期とする.

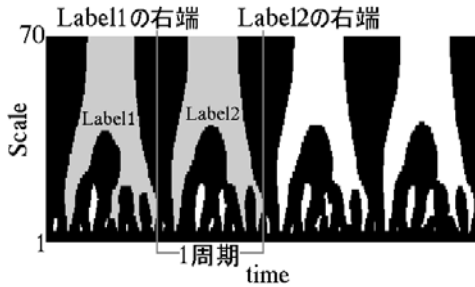


図 4 スケール 1 から 70 までの 2 値化変換画像

4.1.2 相関係数 相関係数とは 2 つの変数集合 ($\{x_i\}, \{y_i\}$) の間の類似性を示す指標である. 相関係数は -1 から 1 の範囲に存在し, 1 に近い場合は 2 つの変数の間に正の相関があるといい, 逆に -1 に近い場合には 2 つの変数の間に負の相関があるという. また, 0 に近い場合は 2 つの変数に類似性はないということになる. 2 次元データにおける相関係数 R は式 (4) で表せる.

$$R = \frac{\sum_{y=0}^{H-1} \sum_{x=0}^{W-1} (f(x, y) - \bar{f})(g(x, y) - \bar{g})}{\sqrt{\sum_{y=0}^{H-1} \sum_{x=0}^{W-1} (f(x, y) - \bar{f})^2} \sqrt{\sum_{y=0}^{H-1} \sum_{x=0}^{W-1} (g(x, y) - \bar{g})^2}} \quad (4)$$

$f(x, y)$ と $g(x, y)$ が位置 (x, y) での画素値であり, H と W は画像の高さ方向と幅方向のピクセル数を, \bar{f} と \bar{g} はそれぞれの平均値を示している. 本章では, これを用いて定量的に比較を行う.

4.1.3 比較結果 ここでは, 変換画像から 1 周期を取得して, 輝度を高さに変換し, 立体画像として比較する. これは輝度の違いよりも高さの違いの方が, 違いがわかりやすいためである. ここでは 3 人の音声サンプルの解析結果を示し, 比較考察を行う. 図 5 に人物 A, 人物 B, 人物 C それぞれにたいする解析結果を示す. 1 周期の長さは同一人物でも多少違っていたが, 1 周

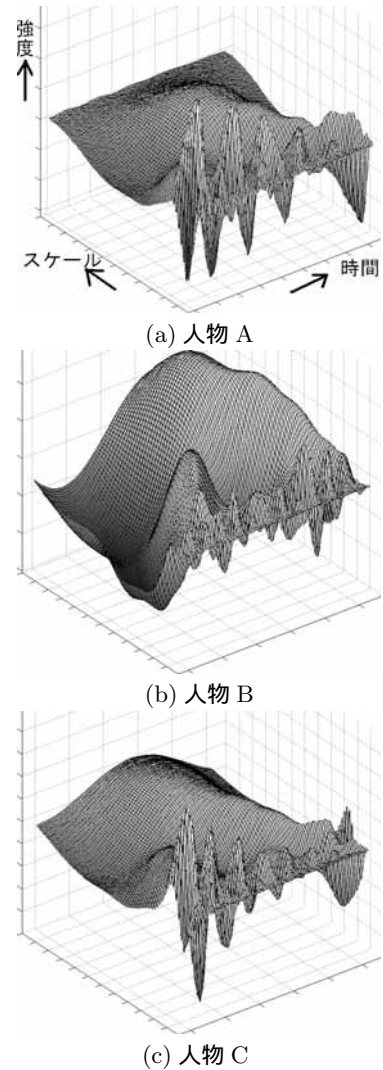
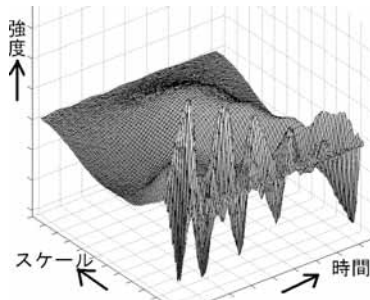


図 5 1 周期画像の比較

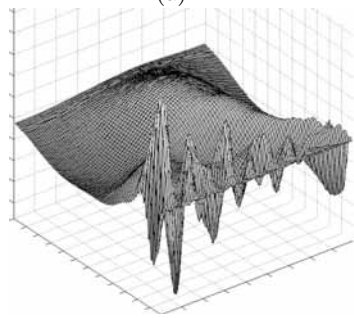
期の形に注目するため, 相関係数を用いる時は画像サイズが同じになるように修正した.

人物 A と人物 C は似た結果となっているが, 人物 B はこれらに比べて大きく異なった結果となっている. 人物 A と人物 C でも, 高スケール (低周波) における成分の強さは人物 C の方が高く, 個人の違いは見られる. 同一人物であれば同じような結果となるのを確認するため, 図 5(a) を A-1 とし, これとは別に録音した人物 A の音声の抽出結果を A-2 として図 6 に示す. 低スケール (高周波) の挙動が若干異なった結果となっているが, 図 5 の人物 B や人物 C よりも, やはり人物 A に近い結果となっている. 表 1 にこれらそれぞれについて相関係数を求めた結果を示す.

図 5 の結果で人物 A, 人物 C と人物 B は大きく異なっていたため, 相関係数もそれに応じた結果となっている. しかし, 人物 A 同士よりも, 人物 A と人物 C の相関が高いという結果も現れている. このことから, 相関係数ではある程度の類似は検出できるが, 細部の違いにまでは対応できていないことがわかる. この他にも組み合わせを変えて比較したが, 同一人物の方が似た結果とはなりやすいが, 他人同士でも似た結果となってしまう事も少なくなかった. また, ウェーブレット変換の特徴であ



(a)A-1



(b)A-2

図6 同一人物の比較

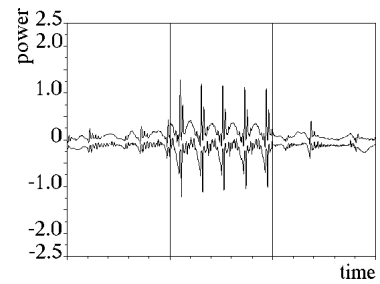
表1 相関結果

組み合わせ	相関係数
人物 A と人物 B	0.4371
人物 A と人物 C	0.8039
人物 B と人物 C	0.4382
人物 A 同士	0.7822

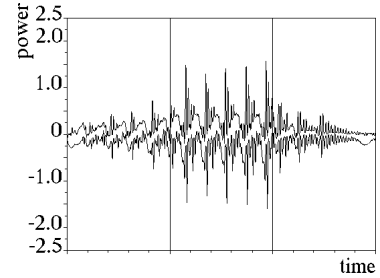
る時間的特徴もみられるという事を考えると、1周期は平均して約9msという短い期間のため、この方法ではウェーブレット変換の特徴を活かせていない可能性があり、全体的に変換画像をみる必要がありそうである。

4.2 発声開始/終了部に対する検討 変換画像全体を一度に把握することは難しいため、発声開始、中間部、発声終了の各部分を等間隔(約45ms)に取得し、時間軸方向からみた場合の強度の輪郭(以下、強度輪郭と示す)を比較に用いる。図3(b)を見ればわかるように、連続ウェーブレット変換の結果は、高スケール(低周波)部分は凹凸の変化が緩やかであり、低スケール(高周波)部分は凹凸の変化が激しい。したがって、強度輪郭の変化が緩やかである場合は高スケールが支配的であり、逆に変化が激しい場合には低スケールが支配的であるという事がわかる。この事に注目して比較を行う。図7に人物A、人物B、人物C、それぞれの各部における強度輪郭を示す。縦軸は強度、横軸は時間を示している。

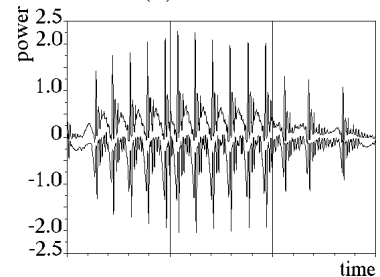
人物Aの終了部に注目すると、変化が緩やかであることから低周波成分が強くなっていることがわかる。逆に人物Bや人物Cでは変化が激しいため、高周波成分が強くなっていることがわかる。また、人物Cの開始部を見てみると、中間部が安定した状態だとすれば、人物Aや人物Bに比べ安定するのが速いことがわかる。なおかつ、人物Cの終了部では、人物Bの終了部が比較的緩やかに減衰しているのに対して、部分的な突起が目立っている事がわかる。図8に人物Cの別の音声から同じように比較した結果を示す。



(a) 人物 A



(b) 人物 B



(c) 人物 C

図7 強度輪郭の比較

人物Cの特徴であると考えられた、「安定するのが速い」、「終了部において部分的な突起が目立つ」などの特徴が見られる。これらのことから、時間的特徴があるとすれば、音声の開始や、終了部に存在する可能性がある。しかし、時間的な特徴というのは、発声の仕方への依存も大きいことが予想される。現在までに使用したのは、自分が一番楽な発声の仕方での発声してもらったものである。必然的に、人によって発声方法が固定されてくることになると考えられるので、音程が変わるように発声をかえた2つの音声の強度輪郭を比較した。図9に人物Dによる、発声を変えた場合の同一人物の比較結果を示す。

この結果から、個人の特徴であると考えていた「終了部には高周波成分が大きく現れている」などの特徴は、発声の仕方ひとつで大きく変わってしまうことがわかった。このことから、時

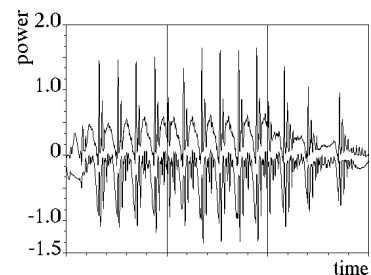
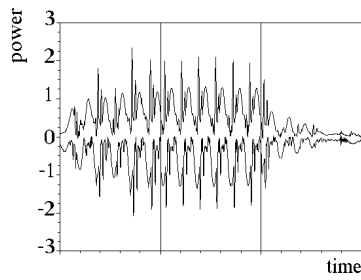
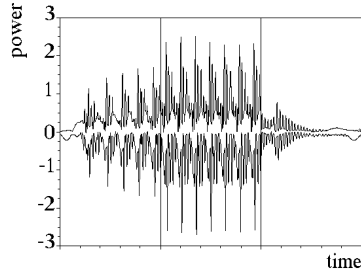


図8 人物Cの別音源の強度輪郭



(a) 人物 D-1



(b) 人物 D-2

図 9 発声を変えた場合の同一人物の比較

間軸側から見た強度の輪郭は個人の特徴というより、発声の仕方の特徴である可能性があり、音声認証には使用できないという事がわかった。

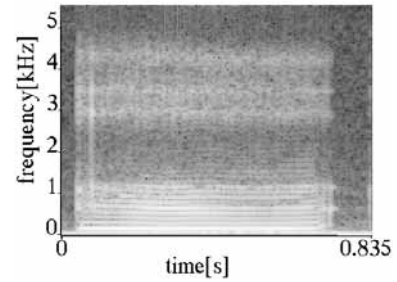
4.3 短時間フーリエ変換の利用 前節までは連続ウェーブレット変換結果のみに着目して個人の特徴について検討してきたが、有用な特徴は得ることができなかった。そこで、短時間フーリエ変換を用いた場合に得られる特徴について考察し、その特徴が連続ウェーブレット変換の結果に現れているかどうか検討する。ここで使用する短時間フーリエ変換では窓関数として約 23.2ms のハミング窓を使用し、その窓関数をずらす間隔は約 2.9ms とした。短時間フーリエ変換の結果から音声認証を行う場合はさらに特徴抽出を行うが、今回は個人の特徴をある程度把握できればよいので短時間フーリエ変換結果の比較を行う。

4.3.1 短時間フーリエ変換による特徴抽出 周波数的な特徴を抽出するため、同一人物から新たに 3 つのサンプルを採取した。音声は全て「あ」だが、全て音高をかえて発声したものを録音した。楽器用チューナーでそれぞれの音高を調べると、表 2 のような結果となった。また、それぞれの短時間フーリエ変換の結果を図 10 に示す。

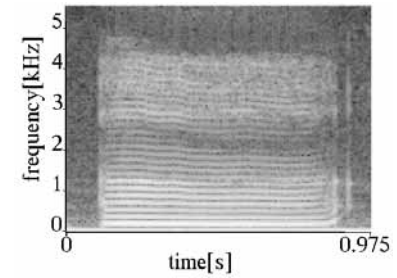
表 2 の周波数成分は個人の意思で変える事ができるため、特徴として使用できない。また、「あ」と発音した場合、1100Hz と 800Hz 付近の周波数成分が含まれることが分かっており⁽⁶⁾、これも個人の特徴としては使用できない。この事を踏まえて、それぞれの短時間フーリエ変換結果について考察する。それぞれの変換結果を見てみると、いずれにも 0Hz ~ 約 1500Hz、約 2500Hz ~ 約 4200Hz の間に周波数成分が含まれている事がわ

表 2 サンプルの音高

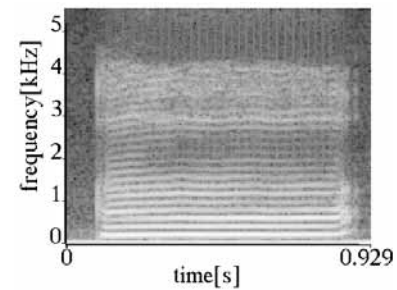
番号	およその基本周波数 [Hz]
サンプル A1	130
サンプル A2	162
サンプル A3	190



(a) サンプル A1



(b) サンプル A2

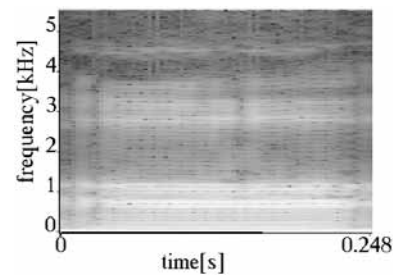


(c) サンプル A3

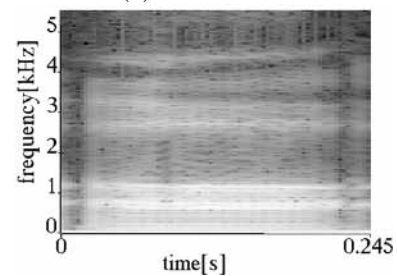
図 10 短時間フーリエ変換結果

かる。しかし、0Hz ~ 1100Hz にはさきほど説明したように個人の特徴としては見ることができない成分も多く含まれているため、それより高い周波数成分に着目し、他のサンプルと比較して個人の特徴となるかどうか検証する。

図 11 は別の人物 B から得たサンプルを短時間フーリエ変換



(a) サンプル B1



(b) サンプル B2

図 11 短時間フーリエ変換結果 2

した結果である。人物 A と同じく 0Hz ~ 約 1000Hz までに強い周波数成分があるが、次に周波数成分の多い部分は約 2500Hz ~ 約 3400Hz と、人物 A よりも範囲が狭くなっている。また、それより少し周波数の高い位置や、4000Hz ~ 4500Hz の間に孤立した成分が存在している事がわかる。これは人物 A にはなかった特徴である。さらに、約 1000Hz ~ 約 2500Hz の成分が多少ではあるが人物 A よりも多いように思える。この結果から、1000Hz より高い周波数の強度とその時間変化は音声認証に使用できる可能性がある事がわかった。

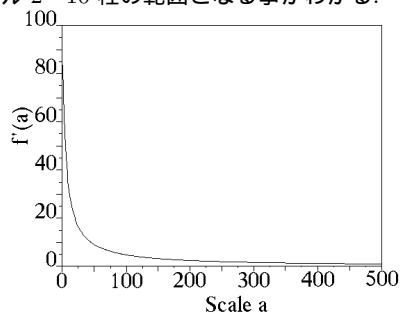
4.3.2 連続ウェーブレット変換において 1000Hz 以上に対応する部分 前節で個人の特徴は 1000Hz より高い成分とその時間変化を見ることが有用であると考えられた。ここでは連続ウェーブレット変換の結果ではその周波数成分はどこに対応するか調べた結果を示す。ただし、個人の特徴が考えていたよりも高周波に存在していたため、今までのような一定間隔のスケール範囲ではなく、周波数に近い軸 $f'(a)$ を式 (5) のように定義する。

$$f'(a) = \frac{F'A}{F'a + A} \quad (5)$$

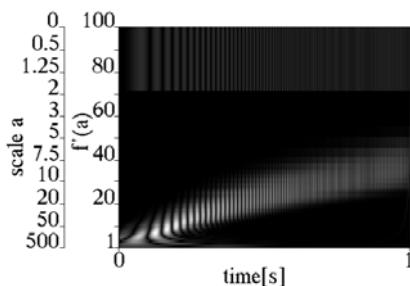
ここで、 F' は $f'(a)$ の最大値、 a はスケール、 A はスケールのおよその最大値を示している。 A を 500、 F' を 100 とすると図 4.3.2(a) のような範囲となる。このスケール範囲を用い、式 (6) のように時間に比例して周波数 0Hz ~ 1000Hz まで変化する sin 波の変換結果を図 4.3.2(b) に示す。

$$Sign(t) = \sin(2000\pi \frac{t^2}{T}) \quad (6)$$

ここで、 T は t の最大値である。変換元の sin 波から考えて、スケール 2 以下に現れているのは異常な結果であるという事がわかるが、この異常な部分を無視すると、基底を coif5 、サンプリング周波数 11.025kHz としている場合、1000Hz より高い部分はスケール 2 ~ 10 程の範囲となる事がわかる。



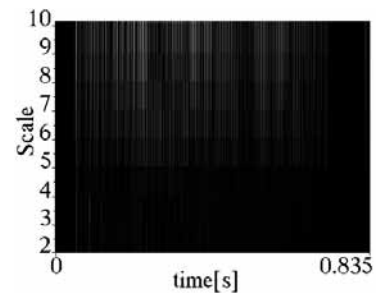
(a) スケール変換



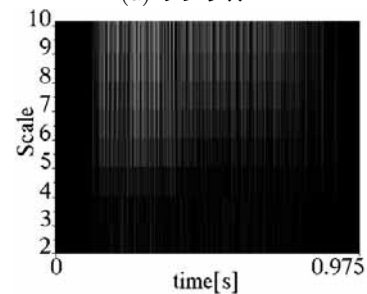
(b) 新しいスケールでのウェーブレット変換

図 12 スケール変換を行なった場合のウェーブレット変換

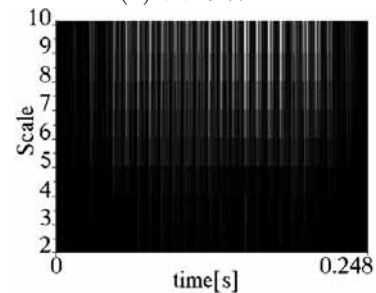
次に、2 ~ 10 (0.1 刻み) のスケール範囲を用いて、図 10 の A1 と A2、図 11 の B1 を連続ウェーブレット変換した結果を示す。図 13(a)(b) が A1, A2, (c) が B1 の変換結果である。短時間フーリエ変換の時に見られたような A1 と A2 の共通点や違いは結果からは把握できないものとなっている。この事から、1000Hz より高い成分を詳しく解析するには、サンプリング周波数やウェーブレット変換時の基底の再検討を行う必要があるということがわかった。サンプル録音時のサンプリング周波数を上げた場合の比較も行ったが、画像の粗さが緩和されただけで、短時間フーリエ変換で得られたような特徴は得られなかった。そのため、低周波における特徴や時間的な特徴など新たな指標の提案が必要であると考えられる。



(a) サンプル A1



(b) サンプル A2



(c) サンプル B1

図 13 スケール 2 ~ 10 の連続ウェーブレット変換結果

4.4 ウェーブレットパケット解析⁽⁵⁾ ある信号を離散ウェーブレット変換した場合、その信号は高周波と低周波に分解される。得られた低周波に対して離散ウェーブレット変換を行い、またそこで得られた低周波に対して離散ウェーブレット変換を行う、といった手順を、決めた回数だけ繰り返すのが離散ウェーブレット変換の基本的な用法だが、ウェーブレットパケット解析の場合は高周波に対しても分解を行うため、より高度な解析が可能となっている。繰り返し分解した回数はレベルといい、分解の深さを表している。様々なレベルで音声のウェーブレットパケット解析を行い、比較した結果、レベルは 7 が最も適して

いと判断し、使用した。また、基底は引き続き `coif5` を使用している。図 14 に 4.3.1 節で使用したサンプル A1, B1 をウェーブレットパケット解析し、周波数 - 時間表示した図を示す。連続ウェーブレット変換で得た結果とは異なり、1000Hz 以上の成分も解析できている事がわかる。人物 B の 4000Hz ~ 4500Hz の間に存在する孤立した成分に関してはやや把握しにくい結果となっているものの、解析結果の大部分が短時間フーリエ変換の結果と類似しているため、ウェーブレットパケット解析を用いる事で、ウェーブレット変換を用いた音声認証は実現可能であると考えられる。

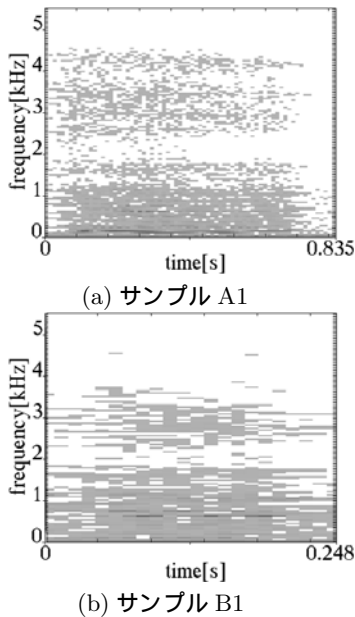


図 14 ウェーブレットパケット解析

5. 結論

連続ウェーブレット変換を用いた場合、抽出できる個人的な音声の特徴は、発声の中間部に対する検討においても、発声開始や発声終了部に対する検討においても個人を識別できるだけの十分なものではなかった。そこで、短時間フーリエ変換を利用して個人の特徴がどのように現れるか調べると、考えていたよりも高周波に特徴が存在する可能性がある事がわかった。つまり、中間部などにおいて得られた変換結果は大部分が 1000Hz 以下の解析で、個人の特徴として使えるものではなかった可能性がある。そこで、1000Hz より高い成分を重視してみるために変換結果において対応する部分を調べたが、現在使用している基底、サンプリング周波数でのウェーブレット変換結果ではその部分を詳しく解析できないことがわかった。対策としてサンプリング周波数を 44.1kHz まであげてサンプルを録音し、解析を行なったが、画像の粗さが改善されただけで、短時間フーリエ変換で見られたような高周波における特徴は得られなかった。また、本論文には載せていないが、基底を変えたとしても、変換結果の低スケール（高周波）部分で大きな変化は見込めない事がわかった。これらの事から、短時間フーリエ変換で得られたような高周波における特徴ではなく、低周波における特徴や時間的な特徴が存在すれば、連続ウェーブレット変換を用いて行う音声認証の実現は可能であると考えられる。しかし、低

周波における特徴や時間的な特徴は、中間部に対する検討でも、発声開始や発声終了部に対する検討でも得る事ができなかったため、連続ウェーブレット変換を用いての音声認証の実現は難しいと判断した。

一方、ウェーブレットパケット解析を用いて周波数 - 時間解析を行った結果、短時間フーリエ変換で得られたような特徴がある程度得る事ができた。本研究では主に連続ウェーブレット変換を用いる場合を考えていたため、ウェーブレットパケット解析について十分な検討を行うことができなかったが、解析結果の大部分が短時間フーリエ変換の結果と類似しているため、ウェーブレットパケット解析を用いた音声認証は実現可能である判断した。また、 N をデータの長さとする、ウェーブレットパケット解析は計算量が $O(N)$ である離散ウェーブレット変換を元としており、 $O(N \log N)$ の計算を要する高速フーリエ変換を元としている短時間フーリエ変換よりも処理が速いなどの利点も考えられる。

参考文献

- (1) B.B. ハバード 山田道夫/西野 操 訳：“ウェーブレット入門 - 数学的道具の物語 -”，朝倉書店，2003 .
- (2) C.Sidney Burrus, Ramesh A.Gopinath, Haitao Guo : “Introduction to Wavelets and Wavelet Transforms”, Prentice-Hall Inc ., 1998 .
- (3) Michal Misiti, Yves Misiti, Georges Oppenheim, Jean-Michel Poggi, サイバネットシステム株式会社 訳：“ウェーブレットツールボックス ユーザーズガイド”，サイバネットシステム株式会社，pp.1-27 ~ pp.1-32, pp.6-48 ~ pp.6-58, 1996 .
- (4) I. ドブシー 山田道夫/佐々木文夫 訳：“ウェーブレット 10 講”，シュプリンガー・フェアラーク東京，pp.100 ~ pp.103, 2003 .
- (5) Michal Misiti, Yves Misiti, Georges Oppenheim, Jean-Michel Poggi, サイバネットシステム株式会社 訳：“ウェーブレットツールボックス ユーザーズガイド”，サイバネットシステム株式会社，pp.5-1 ~ pp.5-4, 1996 .
- (6) 古井 貞熙：“音響・音声工学”，近代科学社，pp.100 ~ pp.103, 1992 .