

# 波形を用いた人の声の特徴について

川田美都紀 神野芽衣 中村奈央

## 【概要】

人の声は様々な周波数で構成されており、人によってその混ざり方に個性があることを知った。我々は、その個性を視覚的に見つけることができれば誰の声なのかということ特定することができるかと推測した。そこで数人の声を録り、比較することで人物の特定をすることができるかを研究した。さらに、声の調子を変化させることにより、波形や周波数に変化があるかどうかを見た。女性 16 人、男性 13 人の声の波形を録ったところ周期性が 4 パターンあり、それぞれ特徴があることがわかった。また、固定電話の回線を通して聞こえる声と通さない声の波形および MAP 機能を用いて比較し、どのような違いがあるかを見た。

An outline of the research. We found out that people's voice is composed of various frequencies and has individuality with the individual. We guessed that we can determine whose voice it is if we identify the characteristic visually. So, we studied how accurate we can specify the owner of the voice by recording the voice of a certain number of people and comparing them. Also, we studied whether there are changes of wave or frequencies when we changed the tone of the voice. The result of the experiment of the samples of 15 females and 15 males showed that there are two patterns of periodicity and also some features were found. Furthermore, we made another experiment by MAP to see if there would be some differences between the normal voices and the ones through the land-lined phone.

## 【研究動機と目的】

音の研究で何か実用性があり、社会に役立つ研究はないのかと考えていた。近年オレオレ詐欺など電話から詐欺の被害につながっていることがわかった。オレオレ詐欺の被害件数は、平成 20 年の 7,615 件、被害額約 155 億円 1,928 万円のピークから平成 25 年には 698 件、被害額約 18 億 5,552 万円。振り込め詐欺の被害件数は、平成 20 年の 20481 件、被害総額約 275 億 9436 万円のピークから平成 25 年 3 月末現在 1876 件、被害総額約 49 億 306 万円とどちらも減っているがまだまだなくなっていない。被害者は、年齢 60 歳以上の方が全体の約 92% を占めているそうだが、60 歳以下の子供を持つ両親も被害者に相当する。最近では警察官になりすましたものなど巧妙な手口が増えてきている。手口が変わってくると自分たちの祖父や祖母も被害に遭いやすくなると思った。高校の物理の授業で

音は基音と倍音で成り立っていて、人の声は管楽器などと比べても多数の倍音が混ざって構成されている、ということ学んだ。PC ソフト(サウンドモニター(FFT(高速フーリエ変換)、WAVE))の MAP 機能を使えば周波数や声の大きさなどが分かるそうだが、高校の物理の授業では波形について学んだので、WAVE 機能を用いて人の声の特徴を見つけないかと思ったからだ。

## 【倍音】

基音の整数倍の周波数を持つ音。100hz の音があったとすると、2 倍の 200hz が倍音である。このように 3 倍、4 倍、5 倍…も同様となる。倍音がかなり少なくなつてきたものが、おんきである。



図-1 おんさ

### 【FFT(高速フーリエ変換)について】

周期現象には基本振動があり、その振動数の整数倍の振動数を持つ周期現象が重なりあって様々なパターンを作り出す。基本振動数で繰り返すサイン・コサインあるいは複素指数関数と、整数倍の振動数で繰り返す関数がどれだけ混ざっているかを表すのがフーリエ展開(フーリエ級数への展開)であり、各振動数の寄与を振幅として表すのがフーリエ係数である。下の公式を使い計算したもの。

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\Delta\omega t}, \quad c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-in\Delta\omega t} dt$$

図-2 公式

### 【仮説】

通常状態で男性の声をとった場合、男性は女性に比べ比較の声が低いことから振動数が少ない波形が得られると思った。また声は人によって様々なので、波形にも個人差が表れ、共通する点はないのではないかと考えた。電話回線を通した実験では、同一人物であっても直接聞く声と電話回線を通して聞く声は個人差があるが、多少異なって聞こえるため、得られる波形にも違いが表れると思った。また表れた波形は人によって通常状態より振動数が多かったり、少なかったりなど異なった特徴が出ると推測した。

### 【実験器具と音源】

- ・ノートパソコン
- ・パソコン用スタンドマイク
- ・PCソフト(サウンドモニター(FFT(高速フーリエ変換)、WAVE))
- ・固定電話(2台)
- ・USB
- ・おんさ
- ・女性の声(16人)
- ・男性の声(13人)

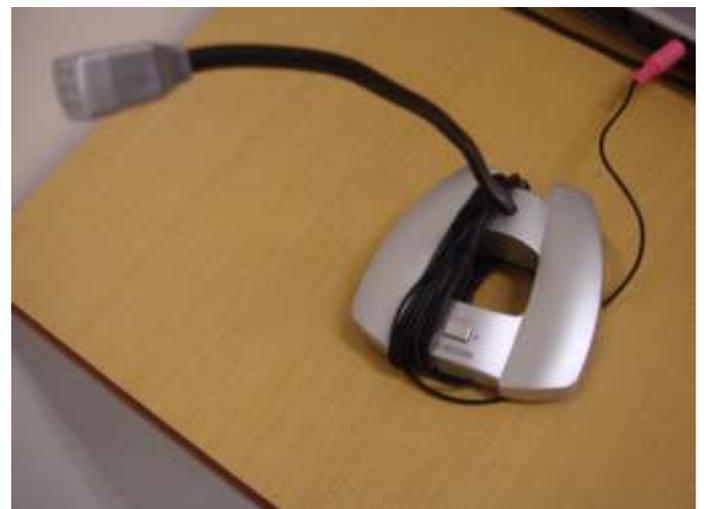


図-3 スタンドマイク

### 【予備実験】

パソコンに、サウンドモニター(FFT(高速フーリエ変換)、WAVE)をダウンロードしておく。

(1) USB、スタンドマイクを接続し、『サウンドモニター(FFT(高速フーリエ変換)、WAVE)』を開く。

(2) メニューバーの『表示(V)』→『セッテイ(S)』の順に開く。

(3) 『WAVE表示』を選択

①取り込時間に合わせる(T)のチェックを外す

②横軸表示範囲(10～取り込み時間)(H)を20～25に設定

(※音の高さによって調整する。)

(1) 図-4の『取り込み時間』を選択  
297secに設定

(2) 図-4の『録音』を開く

- ①『テンポラリの設定(T)』を開く
- ②『リムーバブルディスク(F:)』を選択

(3)録音操作

- ①録音ボタンを押す
- ②マイクの電源を ON にして、マイクに向かっておんさを鳴らす
- ③データがとれたら停止ボタンを押す

【本実験】条件として録る声は、全て『あー』という音を録るものとする。予備実験と同様に操作し、女性の声(16人)、男性の声(13人)を次の4パターンで録る。

- ①女性 (16人)
- ②男性 (13人)
- ③女性 鼻をつまんだ声 (4人)
- ④男性 鼻をつまんだ声 (2人)

【追加実験】

録る声は、全て『あー』という音を録るものとする。予備実験と同様に操作し、おんさ、女性の声(3人)、男性の声(3人)を固定電話の回線を使用して、次の5パターンで録る。

- ① おんさ
- ② 女性 通常状態
- ③ 男性 通常状態
- ④ 女性 鼻をつまんだときの声
- ⑤ 男性 鼻をつまんだときの声

【結果と考察】

＜本実験の結果＞

本実験で女性の声(16人)、男性の声(13人)の声をサウンドモニターFFT WAVEのWAVE機能を用いて波形を録った。『あー』と4秒のばしてもらい、波形が安定してきた(一定になってきたところ)で波形を周期ごとにくぎる。その部分の波形の形を見て同じ形にあるものをパターン分けした。その結果、主に4つのパターンに分けられた。

1. パターン1

- ・ 女性9人、男性6人 3:2 の割合でみられた。
- ・ 1波長内では、前半が大きく振れ後半にかけて段々と小さくなる特徴がみられた。
- ・ 女性の声は全てこの型に当てはまった。
- ・ 声の低い男性に多く、女性にはみられなかった。

2. パターン2

- ・ 女性5人、男性1人 5:1 の割合でみられた。
- ・ 1は町内では、前半部分にはやく小さくなり後半部分でまた大きくなっていく特徴がみられた。
- ・ 女性に多くみられた。

3. パターン3

- ・ 女性2人、男性2人 1:1の割合でみられた
- ・ 前半部分ではやく小さくなり、真ん中部分で再度大きくなってから小さくなり、後半部分でまた大きくなる特徴がみられた。
- ・ 他のパターンよりもこのパターンに当てはまった人数が少なかった。

4. パターン4

- ・ 女性0人、男性4人
- ・ 区切り方によりパターン2と似ている
- ・ 小さい振れが段々下がっていき、大きく振れたところからパターン2のような形の波形になり、上がっていきながら小さくなっていく。
- ・ 女性はこのパターンにひとりも当てはまらず、他のパターンと比べて男女の偏りがみられた。
- ・

《全体を通しての結果》

- ・ 女性の方が1波長の幅が小さく、男性の方は幅が大きくみられた。

- ・ 全体としてパターン1が最も多くみられた。
- ・ 本実験、追加実験のどちらも声の調子を変えた(鼻をつまんだとき)場合には、通常状態とそれほど変化が見られなかった。要因として、雑音が多く混ざってしまったことが挙げられる。

#### <追加実験の結果>

本実験とは別の女性3人、男性3人の声をMAP機能とWAVE機能を用いて比較した結果、女性男性ともに同姓間での違いはさほど見られなかった。

#### 【今後の課題】

1つ目は波形を調べるだけでは声を特定するまでには至らなかったため、周波数スペクトルを三次元波形で表示する〔FFT 三次元波形表示〕、〔FFT 波形表示コマンド〕という別の機能も用いてより正確な人の声の特徴を見つける。2つ目は波形を文字に変えることができるかもしれないという情報を得たので実験し、会話を解析する。

研究動機で書いた実用性について、個人の声データをデータ戸籍に登録するようにして数年置きに録る。そして、このデータを使って大手の電話会社や電機メーカーに協力してもらい、固定電話にそのデータを搭載する。(日本人全員のものを各家庭の固定電話にインプットすることは不可能なので、家族、親戚、仲のよい友人、交流のある人など特定の人をピックアップしておく。) そうすると例えば風邪で声が変わっていたとしても正確に誰からかかってきたのかわかるようになる。現在の日本の技術でこれを実行するのは難しいかもしれないが、将来声紋の技術が更に発展すれば不可能ではないと思う。これを実現するため更に研究を進めていきたいと思う。

#### 【参考資料】

書物：音律と音階の科学 —ドレミ…はどのようにして生まれたか—

ホームページ：

振り込め詐欺：警視庁

<http://www.keishicho.metro.tokyo.jp/seian/koreisagi/koreisagi.htm>

警察庁振り込め詐欺対策HP

[http://www.npa.go.jp/safetylife/seianki31/1\\_hurikome.htm](http://www.npa.go.jp/safetylife/seianki31/1_hurikome.htm)

声紋鑑定専門のオフィシャルサイトです。

<http://www.kantei-senmon.com/voiceprint.html>

法科学鑑定研究所

<http://www.e-kantei.org/voice/>

音声認識のしくみ

<http://recognition.web.fc2.com/>