

電子メールとWeb技術を用いた公衆型音声分析システム*

◎熊谷勝, 三輪 譲二 (岩手大学 工学部)

1. はじめに

音声分析というと、専門家以外の人達にとって、難しく近寄りがたいイメージがあるように思われる。

しかし、最近のパソコンとインターネット^[1]の著しい普及に伴い、電子メールやWWWが盛んに利用されるようになり、パソコンの操作にある程度慣れている人ならば、パソコンのマルチメディア機能により、自分の音声を録音して電子メールに添付することは比較的容易なことである。

そこで、電子メールとWeb技術を利用した公衆型音声分析システム(PASTE:Public Analysis System for speech using Techniques of E-mail and web)を構築した。本システムは、音声波形データを添付した電子メールをPASTEシステムに送信するだけで分析結果をWWWで閲覧することが可能である。

これにより、パソコン、電子メール、インターネットを利用できる環境にある人ならばだれでも気軽に音声分析に参加することが可能となる。

2. 公衆型音声分析システム

2.1 システム構成

本システムは、利用者から送信された電子メールに添付された音声波形データを分析し、その分析結果を画像にしてWebページに貼り付ける。利用者はインターネットを用いて分析結果のWebページを見ることにより、分析結果を見ることができる。システムの全体構成を図1に、音声分析処理の流れを図2に示す。

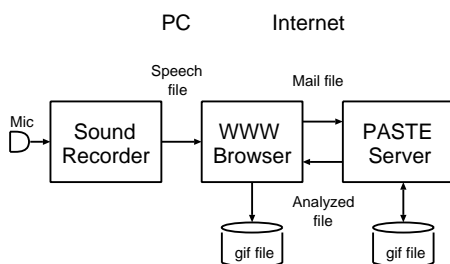


図1. 公衆型音声分析システムの全体構成

なお、分析結果を見るにはURLの他に、音声分析システムより与えられるIDが必要となるため、分析結果を他人に見られることがなく、また、一度分析し終わった音声データなら、IDがあれば指定された期間内は何度でも見ることができる。

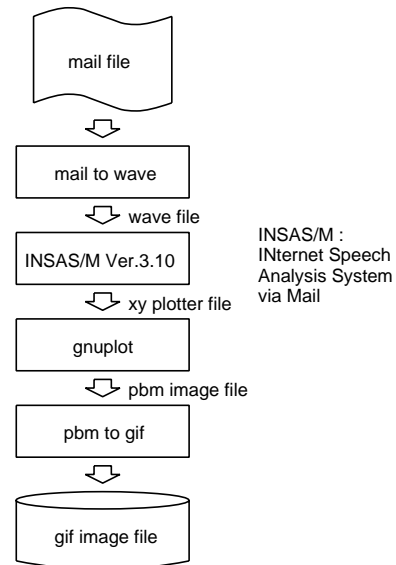


図2. PASTE処理過程

また、電子メールに添付されるファイルは、base64という形式で符号化される。本システムではbase64での符号化を想定しているため、添付ファイルの符号化にbase64を用いるメーラーによって音声波形データを添付して送信する必要がある。このため、ブラウザとして、例えばIE4.0を使用できる。

さらに、分析可能なデータの形式を表1に示す。このデータ形式のファイルは、Windowsのサウンドレコーダー等を用いて作成することができる。

表1. 分析可能なデータ形式と分析条件

ファイル形式	WAV
サンプリング周波数	11.025kHz
チャンネル数	1
量子化ビット, 形式	16bits, 線形
フレーム周期	10ms
分析窓	ハミング窓 (30ms)
分析	13次LPC

2.2 分析機能

本分析システムでは、現在のところ、INSAS/M^[2] Ver.3.1として、表1の条件によるLPC分析の軌跡表示とセクション表示の機能を備えている。

2.2.1 LPC分析の軌跡表示

音声データ全体を10msおきにLPC分析し、音声波形と、基本周波数、音声パワー、ゼロクロス数、ホル

* Public Analysis System for Speech Using Techniques of E-mail and Web.

By Masaru Kumagai and Jouji Miwa (Faculty of Engineering, Iwate University).

マント周波数、ホルマント周波数の帯域幅の特徴の時間軌跡をグラフィック表示する。この例を図3に示す。

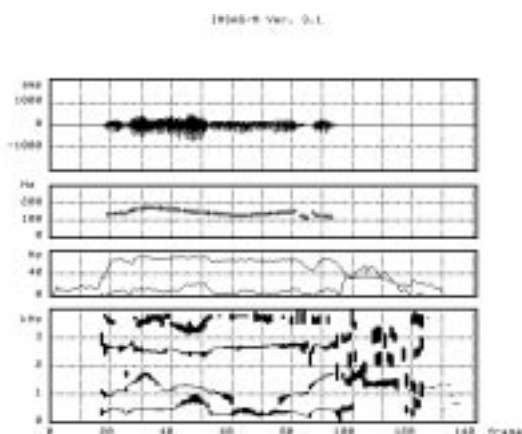


図3. LPC分析の軌跡表示の例

2.2.2 LPC分析のセクション表示

音声波形データの1フレームをLPCセクション分析する。音声波形、窓関数、パワースペクトル、平滑化パワースペクトル、音声パワー、ゼロクロス数、ホルマント周波数、帯域幅をグラフィック表示する。この結果の例を図4に示す。

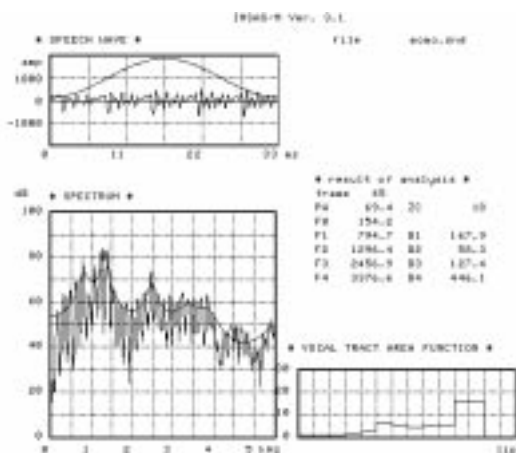


図4. LPCセクション分析表示の例

2.3 使用手順

2.3.1 音声の切り出し

音声を分析するにはまず、Windowsのサウンドレコーダー等によって、表1に示した形式の音声波形データファイルを作成する。

2.3.2 電子メール送信

作成した音声波形データ添付してをPASTEシステムのメールアドレスに送信すれば分析が行われる。このメールを送信する際のサブジェクトに分析コマンドを入力することにより分析の機能、分析するフレーム位置を変更することができる。分析コマンドは分析機

能とオプションで構成され、分析機能で分析の種類を設定し、LPC軌跡とLPCセクションがある。なお、LPCセクション分析においてはオプションで分析するフレームを変更できる。例えば、サブジェクトをsc45とした場合、データの45フレーム目をLPCセクション分析し、scとフレーム数を指定しない場合、音声波形の値が一番大きい点をLPCセクション分析する。図3は”Subject: lc”、図4は”Subject: sc45”とした場合である。

表2. 分析コマンドと分析の対応

分析機能	
lc	LPC分析の軌跡(locus)表示
sc	LPCセクション(section)分析
無し	LPC分析の軌跡表示
scのオプション	
整数値	指定したフレームを分析
無し	音声波形の最大値点を分析

2.3.3 分析結果閲覧

分析が終了すると、システムから分析終了と分析結果を見るために必要なIDを知らせるメールが送られてくる。分析システムのホームページにアクセスしてIDを入力すると、分析結果の画像を見ることができる。また、画像をダウンロードして自分のパソコンに保存することもできる。

3. 考察

本システムは、電子メールにより音声データを送信し、分析結果をWebを用いて見るという、誰にでも取り組みやすいシステムになっている。また、分析結果を見るにはIDが必要となるため、データの保護の点においても安全であると言える。

しかし、現在のところ分析機能が2つしかないため、分析機能の追加が必要であると思われる。

4. まとめ

本報告では、構築した、電子メールとWeb技術を用いた公衆型音声分析システムの構成について報告した。今後の課題として、分析機能の多様化等が挙げられる。

なお、本システムは、以下のURLで公開する予定である。

<http://sp.cis.iwate-u.ac.jp/sp/paste/>

参考文献

- [1] 三輪, 熊谷, 山本: “インターネット・マルチプラットフォーム対応の対話型音声分析システム”, 音講論, 2-p-27, pp.347-348 (March 1997).
- [2] 三輪: “パソコン音声処理”, 昭晃堂(1991).