

TR - A - 0136

UNIX 上の音声研究用ツール
- *SpeechTools* -

天白 成一

1992. 3. 9

ATR 視聴覚機構研究所

〒619-02 京都府相楽郡精華町光台 2-2 ☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411

Facsimile: +81-7749-5-1408

も く じ

第 1 章	はじめに	1
第 2 章	コマンドの使い方	2
2. 1	使い方 (基本編)	2
2. 2	使い方 (応用編)	5
[例題	1 : DFT 分析]	6
[例題	2 : LPC 分析]	7
[例題	3 : ローパスフィルタリング]	8
[例題	4 : ピッチ抽出]	9
[例題	5 : データの切り出し]	10
[例題	5 : バイナリィファイルの作成]	11
第 3 章	ユーティリティ	12
3. 1	acat / bcat	12
3. 2	cv	12
3. 3	merge / separate	12
3. 4	bcalc / smath	13
第 4 章	ライブラリ	14
第 5 章	まとめ	15

付録.

第1章 はじめに

概要

音声・聴覚の研究においては、音響信号波形をA/D変換して計算機上に取り込み、その波形やスペクトルあるいは信号処理結果を表示したり、波形やスペクトルを加工／編集したり、加工／編集した信号をD/A変換して聴いたりすることが必要であり、このような一連の作業を効率良く行えるか否かは大きな問題である。そこで、このような作業を支援するために音声研究用ソフトウェアツール(SpeechTools)開発してきた。

SpeechToolsは、広くUNIXをOSとする計算機上で利用することができ、音声分析・合成に関する各種のアルゴリズムを単一のインタフェースを通して使用できるコマンドと種々のデータ形式の変換等を行うユーティリティとそれらを支えるライブラリによって構成されている。

SpeechToolsの対象分野としては音声の研究全般を考え、音声分析・合成等に関する確立された技術を提供することを主目的とした。そして、SpeechToolsは大きな一つのプログラムではなく、小さな機能ごとに個別に分けられたコマンドの集合体として設計されており、コマンドを連結することでより多くの目的に利用できるようになっている。また、データ変換などを行うユーティリティとしては、acat(アスキーデータからバイナリデータへの変換)、bcat(バイナリデータからアスキーデータへの変換)、bcalc(2項演算を2つバイナリデータファイルに実行)、cv(各種バイナリデータのフォーマット変換)、merge(アスキーデータの列の併合)、separate(アスキーデータの列の分割)、smath(アスキーデータの科学算術演算)の7個があり、ライブラリとしては、libcommand.a(ファイルの入出力、コマンドのインタフェース等を含む)、libmx.a(行列演算を含む)、libst.a(音響分析・合成、DFT、高次方程式の解法等を含む)、libtlist.a(簡易なリスト処理を含む)の4個がある。プログラムは、すべてC言語で記述されており、総行数は約2万行である。また、マニュアルは、すべてroff形式で記述されており、オンラインで読むことができる。

利用方法

SpeechToolsは、通常/usr/local/stにインストールされ、コマンドとユーティリティは/usr/local/st/binにインストールされるので、各自のホームディレクトリのファイル'.login'や'.cshrc'のコマンドのサーチパスに、/usr/local/st/binを加えなければならない。また、ライブラリとヘッダファイルは、それぞれ/usr/local/st/libと、/usr/local/st/includeにインストールされており、SpeechToolsのライブラリを用いて新たにコマンドを開発する場合は、これらのディレクトリに注意しなければならない。同様に、SpeechToolsのマニュアルは/usr/local/st/manにインストールされている。

本稿の構成

本稿は、5章から構成されている。第2章では、SpeechToolsのコマンドで実現されている一貫した使用方法を基本編で説明し、応用編ではシェルスクリプトによる利用例について説明する。第3章では、SpeechToolsのユーティリティに関して、第4章では、SpeechToolsのライブラリに関して簡単にそれぞれ説明する。第5章では、まとめとしてSpeechToolsの問題点と今後の方向性等について述べる。そして、付録にはマニュアルの目次とマニュアルの一部を付けた。

第2章 コマンドの使い方

この章では、SpeechToolsの多くのコマンドで実現されている一貫した使用方法を、基本編と応用編に分けて説明する。基本編では、一貫した使用方法について、応用編では、利用例を通じて使い方を説明する。

2.1 使い方(基本編)

このSpeechToolsのコマンドの多くは、一貫した同じ方法で使うことができる。以下の説明では、下線部がユーザが入力している部分で、% はUNIXのプロンプトでありユーザの入力を促している。

①UNIXのプロンプトの出ている状態でコマンド名を入力する。

```
% fft_run
Calculate DFT running spectra.
usage :: fft_run filename
        fft_run -o arguments

Defaults are as follows.
SAMPLING FREQUENCY : 20.0 kHz
WINDOW LENGTH      : 30 msec
WINDOW TYPE        : HANNING
FFT LENGTH         : 1024
FRAME PERIOD       : 5 msec
PREEMPHASIS       : 0.98
INPUT FILE NAME    : TMP.DAT float
OUTPUT FILE NAME   : TMP.DFT float
```

コマンドは、簡単な説明と使い方とパラメータの既定値(default value)を画面に表示する。

```
Calculate DFT running spectra.
これが、コマンドの簡単な説明である。
usage :: fft_run [options] filename
        fft_run [options] -o arguments
```

これが、コマンドの使い方である。

各コマンドは、パラメータを持ち、各パラメータには既定値が設定されている。コマンドを使うときは、パラメータを指定して使わなければならない。

コマンドには、大きく分けて2通りの使い方がある。コマンドの引数にファイル名を与える方法と、-oを与え、以下、パラメータをコマンドラインの引数に与える方法である。これらの方法を通じてパラメータの値を変更し、コマンドを実行する。

以下に、パラメータの例を示す。

```
SAMPLING FREQUENCY : 20.0 kHz
WINDOW LENGTH      : 30 msec
WINDOW TYPE        : HANNING
FFT LENGTH         : 1024
FRAME PERIOD       : 5 msec
PREEMPHASIS       : 0.98
INPUT FILE NAME    : TMP.DAT float
OUTPUT FILE NAME   : TMP.DFT float
```

パラメータは、以下の形式で指定する。

パラメータ名 : 値 単位 # コメント

たとえば、

```
SAMPLING FREQUENCY : 20.0 kHz
```

では、SAMPLING FREQUENCY がパラメータ名で、20.0 が値で、kHz が単位である。コメントは、記号 (#) から改行までであり、コマンドを実行するときは、コメントは無視される。

② パラメータをファイルに書き込む。

```
% fft_run > envfile
Calculate DFT running spectra.
usage :: fft_run [options] filename
        fft_run [options] -o arguments
```

Defaults are as follows.

パラメータをファイルに一旦取り込むにはリダイレクション (redirection) を使う。リダイレクションによって、標準出力装置 (stdout) に出力されていたものがファイルに書き込まれる。ここで注意すべきは、'Defaults are as follows.' 以下の行だけが標準出力装置に出力されているということである。

確認のために、cat を使ってファイルの内容を表示させると次のようになる。

```
% cat envfile
SAMPLING FREQUENCY : 20.0 kHz
WINDOW LENGTH      : 30 msec
WINDOW TYPE        : HANNING
FFT LENGTH         : 1024
FRAME PERIOD       : 5 msec
PREEMPHASIS        : 0.98
INPUT FILE NAME    : TMP.DAT float
OUTPUT FILE NAME   : TMP.DFT float
```

これを、エディタを使って編集して望みの値を設定する。

```
% fft_run envfile
```

```
.....
!120
```

多くのコマンドは、実行すると記号 '.' が、標準出力装置に表示され、最後に記号 '!' と数値を表示する。数値は、記号 '.' の数に等しく大きなループの回数を表わしている。

③ -o オプションを使う。

第1引数に -o を指定すると、コマンドに、ファイルを使用せずに既定値を変更することができる。この使い方は、特にシェルスクリプトを使用してバッチ処理を行なうときに有効である。

```
% fft_run -o "FRAME PERIOD : 10 msec" "FFT LENGTH : 4096"
```

注意事項

- ・ 指定されないパラメータは、その既定値が使われる。
- ・ 指定するパラメータの順序は問わないが、2回以上与えると後の方が有効となる。
- ・ 全ての文字列では、大文字と小文字の区別を行なってる。
- ・ パラメータの既定値を変更するには、~/ . s t r c を利用する。

変数名について

各コマンドの変数名は、おおよそ同じなので `lpc_run` を例にして変数名について説明する。`lpc_run` の変数を表1にまとめた。

表1 `lpc_run` の変数

変数名	意味	既定値	単位	
SAMPLING FREQUENCY	標本化周波数	20.0	kHz	①
WINDOW LENGTH	分析窓長	30.0	msec	②
WINDOW TYPE	分析窓の種類	HANNING		③
FFT LENGTH	DFTの長さ	1024		④
FRAME PERIOD	分析フレームの周期	5.0	msec	⑤
PREEMPHASIS	高域強調のパラメータ	0.98		⑥
ORDER OF LPC	線形予測係数の次数	16		⑦
INPUT FILE NAME	入力のデータファイル名	TMP.DAT	float	⑧
OUTPUT LPC FILE NAME	出力のLPCデータのファイル名	TMP.LPC	float	⑨
OUTPUT ALPHA FILE NAME	出力の線形予測係数のファイル名	TMP.ALP	float	⑩

①には、音声データの標本化周波数を指定する。

②には、分析窓長を指定する。

③には、以下のものから選択できる。

RECTANGULAR
HANNING
HAMMING
BLACKMAN
BARTLETT
SINC

④は、 2^n の正の整数である。ただし、標本化周波数×分析窓長よりも長くなくてはならないが、上限はない。

⑤には、分析のフレーム周期を指定する。もし、分析窓長よりも長い場合は、分析フレームは、重なりがないことになる。

⑥は、高域強調に関する係数で、0～1の値を取る。0を与えると高域強調しない。

⑧⑨⑩には、ファイル名を与えるが、単位に示されたデータ型のバイナリファイルである。通常、音声データに関してはshortを用いるが、コマンドの単位がfloatである場合、単位の部分に望みの単位を指定すると自動的に変換してファイルを読み書きする機能がある（詳しくは、各コマンドのマニュアルを参照すること）。また、ファイル名の指定に際しては、絶対パスで指定することができるが、特殊文字（'～', '.', '...', '）を用いて相対パスで指定することもできる。なお、ワイルドカードを表す特殊文字（'*', '?'）は、使えない。

2. 2 使い方 (応用編)

例えば、多くのデータを一度に同じ方法で分析したり、あるいは、ある一定の条件で異なったパラメータを用いて合成音を作ったりと、定型的な作業を連続して行なうときがよくある。この様な一連の定型的な処理のことをバッチ(batch)処理と言う。

UNIXでは対話的に処理を行なうことが基本となっているが、バッチ処理を行なう場合にはシェルスクリプト(shell script)を用いるのが簡単な方法である。シェルとはUNIXの対話的なユーザーインターフェースそのものであり、ユーザーからの要求を解釈して実行するものである。

他の多くのオペレーティングシステムとは異なり、UNIXでは、シェルは、一つのコマンドであり、そのため、UNIXには多くのシェルが存在する。代表的なものとして、sh (Bourneシェルまたは単にシェルという)とcsh (Cシェルという)がある。シェルスクリプトとは、シェルの文法に従って記述するプログラムのことである。

このマニュアルでは、シェルにはcshを用いる。以下、例題にしたがっていくらかのcshの説明を加えるが、詳しくは正著[Anderson & Anderson, 1986]にゆずる。例えば、ファイルを編集して、シェルを実行するには、ファイルに実行許可の指定を行なわなければならないことなどは、ここでは説明しない。

ここで述べる例は極めて簡単なものであるが、これらのものは基本的なものであり、組み合わせることによって多くの複雑なことができる。よって、ここで取り上げた例を修得することを勧める。

参考文献

Anderson, G. and Anderson, P. (1986) : "THE UNIX C SHELL FIELD GUIDE",
訳 落水浩一朗、大木敦雄、パーソナルメディア(株)

[例題 1 : DFT分析]

```
[dft.csh]
#! /bin/csh -f
set DATA = TEST
fft_run -o\
    "SAMPLING FREQUENCY : 20.0 kHz"\
    "WINDOW LENGTH      : 30 msec"\
    "WINDOW TYPE        : HANNING"\
    "FFT LENGTH         : 1024"\
    "FRAME PERIOD       : 5 msec"\
    "PREEMPHASIS        : 0.98"\
    "INPUT FILE NAME    : $DATA.DAT short"\
    "OUTPUT FILE NAME   : $DATA.SPC float"
```

c s h の起動

まず、第1行目では、c s hを使用することを宣言している。記号'#'は、通常、注釈(コメント)を表わすが、ファイルの一番最初の行に'#! /bin/csh -f'と記述すると以下このファイルを、c s hの文法に従って記述でき、かつ、実行されることを意味する。c s hを使うときは必ず第1行目にこれを書かなければならない。もし、この行がファイルの第1行目がないときはc s hの記法は使用できないので、文法的なエラーが生じる。

'-f'は'csh'のオプションで、素早くc s hを起動することを意味する。c s hは、通常、ユーザのホームディレクトリにある'.cshrc'ファイルを読んだ後、起動されるが、'-f'オプションを付けておくと'.cshrc'ファイルを読まないで起動される。このことによって、'.cshrc'ファイルに記述されたエイリアス(alias)等の定義は、有効にならない。しかし、環境変数(定義済シェル変数)は、プロセス生成時に継承されるので、例えば、環境変数の1つである'path'の指定は、継承される。

シェル変数を定義するには

'set'を使って変数を定義するには、

```
set 変数名 = 値
```

とする。値には、数値、文字列、リストがある。リストは、記号'('と記号')'によって作られる。ここでは、

```
set DATA = MAU10018
```

'DATA'という変数名に'MAU10018'という文字列が代入されている。変数を使用するときには変数名の前に記号'\$'を付ける。'\$DATA.AD'と'\$DATA.DFT'は、それぞれ、'MAU10018.AD'と'MAU10018.DFT'に対応している。

一行に収めるには

記号'\ 'は、バックスラッシュである。c s hでは、バックスラッシュによって便宜的に行を続けることができる。'-o'オプションを使うときは、1行に全ての既定値を書かなければならないので、バックスラッシュによって一つの行としなければならない。

なお、コマンドに'-x'オプションを与えると、シェルスクリプトを書くのが便利のようにバックスラッシュ着きでパラメータが標準出力装置(stdout)に出力される。

[例題 2 : LPC分析]

```
[lpc.csh]
#!/bin/csh -f
foreach FILE ( $* )
    set DATA = $FILE:r
    lpc_run -o \
        "SAMPLING FREQUENCY      : 20.0 kHz" \
        "WINDOW LENGTH           : 30 msec" \
        "WINDOW TYPE              : HANNING" \
        "FFT LENGTH               : 1024" \
        "FRAME PERIOD             : 5 msec" \
        "PREEMPHASIS              : 0.98" \
        "ORDER OF LPC              : 16" \
        "INPUT FILE NAME           : $DATA.DAT short" \
        "OUTPUT LPC FILE NAME      : $DATA.LPC float" \
        "OUTPUT ALPHA FILE NAME    : $DATA.ALP float"
end
```

コマンド引数の渡し方

コマンド行からファイル名等を渡したいときがよくある。'\$*'は、コマンド行のリスト全てになる。'\$*'は、cshでは、コマンド行のリストである。具体的に述べると、コマンド行で、以下の様に入力されると

```
% lpc.csh FSU10018.AD FKN10018.AD MAU10018.AD MHT10018.AD
```

'\$1'は、FSU10018.AD、'\$2'は、FKN10018.AD、'\$3'は、MAU10018.AD、'\$4'は、MHT10018.ADになる。

なお、'\$1'と'\$argv[1]'ならびに'\$*'と'\$argv[*]'は、表記が違うが同じ意味である。'argv'という変数は、cshの方であらかじめ定義されている変数リストである。このように、cshの方であらかじめ定義された変数が幾つかある。

一般的に、リストのn番目を参照するときは、'\$変数名[n]'とし、リストの長さは、'\$#変数名'とすれば良い。

例えば、以下のようなことが可能である。

```
% set list = ( a b c d )
% echo $list
a b c d
% echo $#list
4
% echo $list[2]
b
```

繰り返し実行の仕方

繰り返し実行させるときには、'foreach'を使うのが1つの方法である。'foreach'は、'end'までを繰り返し実行する。

'foreach FILE (\$*)'では、'FILE'は仮変数であり、リストの先頭に順次、単一化されて行く。リストが空になった時点でこのループは終了する。

[例題 3 : ローパスフィルタリング]

```
[lpf.csh]
#!/bin/csh -f
set LIST = `cat TEMPLATE`
foreach DATA ( $LIST )
    lpf1 -o\
        "SAMPLING FREQUENCY : 20.0 kHz"\
        "CUT OFF FREQUENCY : 100.0 Hz"\
        "INPUT FILE NAME : $DATA short"\
        "OUTPUT FILE NAME : $DATA.LPF float"
end
```

外部ファイルからの入力

外部ファイルからファイル名のリストを読み込んで実行するには、

```
set LIST = `cat TEMPLATE`
```

とする。ここで、記号``は、バッククォートであり、バッククォートで囲まれた文字列は、cshによって評価される。例ではファイル`TEMPLATE`を出力する。出力された内容が、変数`LIST`に代入される。

また、外部コマンドを実行して変数に代入するには、以下の様にすれば良い。

```
set LIST = `ls *.AD`
```

[例題 4 : ピッチ抽出]

```
[pitch.csh]
#!/bin/csh -f
if ( $#argv != 3 ) then
    echo "usage: "
    echo "pitch.csh filename min-freq[Hz] max-freq[Hz]"
    exit -1
endif
#
set DATA = $1
set MIN = $2
set MAX = $3
#
pitcher -o \
    "SAMPLING FREQUENCY : 20.0 kHz" \
    "WINDOW LENGTH      : 30 msec" \
    "WINDOW TYPE        : HANNING" \
    "FRAME PERIOD       : 5 msec" \
    "MINIMUM FREQUENCY  : $MIN Hz" \
    "MAXIMUM FREQUENCY  : $MAX Hz" \
    "THRESHOLD POWER    : 40 dB" \
    "POLARIZATION       : ON" \
    "INPUT FILE NAME    : $DATA short" \
    "PITCH FILE NAME    : $DATA.PIT float" \
    "POWER FILE NAME    : $DATA.POW float"
```

目的に応じて

しばしば、コマンドラインから条件を変えて分析をしたいときがある。この例では、分析対象のファイル名とピッチ抽出の最大と最小の基本周波数を指定している。

忘れるということを忘れないように

cshやC言語などを使ってプログラムを作るとき、少し時間がたち日もたつと、使い方を忘れてしまったり、どんなことをしてくれるプログラムなのかを忘れてしまったりする 때가よくある。引数リストの順番は、必ずとっていいほど忘れてしまう。もし要求される引数の数が違ったら、実行しないようにするよう心掛けると良い。また、引数の数が間違っていたら使い方を簡単に説明するようにプログラムするのが賢明な方法である。

[例題 5 : データの切り出し]

```
[cut.csh]
#!/bin/csh -f
echo -n "Please input a file name : "
set DATA = $<
/bin/ls -l $DATA
if ( $status ) then
    exit -1
endif
echo -n "Offset [byte] : "
set OFFSET = $<
echo -n "Length [byte] : "
set LENGTH = $<
#
subset -o\
    "SOURCE FILE NAME : $DATA"\
    "RESULT FILE NAME : $DATA.CUT"\
    "SIZE OF ARRAY      : 1 byte"\
    "OFFSET             : $OFFSET"\
    "LENGTH             : $LENGTH"
```

会話的に使用するには

'echo -n'は、メッセージを出力した後に改行しないことを表わす。そして、'\$<'によって入力を待つ。

```
set DATA = $<
```

入力された文字列は変数 'DATA' に代入される。

失敗を発見する

このような方法で、cshでも会話的に処理を行なうことができるが、会話的に処理を行なうとき注意しなければならないのは入力ミス等である。cshでは、入力ミス等が発見するための機能が、あまり用意されてはいない。

直前のコマンド等が正常に終了したかどうかを確かめるには、変数 '\$status' を参照する。

```
/bin/ls -l $DATA
if ( $status ) then
    exit -1
endif
```

'bin/ls -l \$DATA'によってファイル名をリストし、かつファイルの長さを利用者に提示している。もしファイルが存在しなかったときは、'\$status'には0以外の数字が入っている。UNIXの多くのコマンドは、正常終了したときは0を、異常終了したときは0以外の数字を返すようになっている。

本マニュアルで説明するコマンドも多くは、この慣例に従って、正常終了したときは0を、異常終了したときは-1を返すようになっている。

[例題 5 : バイナリィファイルの作成]

```
[atobi.csh]
#! /bin/csh -f

cat << EOF > TMP.ASC
  0.0  120.0
1000.0 100.0
EOF
#
atobi -o\
      "INTERPOLATION METHOD : LINEAR"\
      "TOTAL LENGTH       : 1000.0 msec"\
      "FRAME PERIOD       : 5.0 msec"\
      "TIME COLUMN        : 1"\
      "DATA COLUMN        : 2"\
      "INPUT FILE NAME    : TMP.ASC ascii"\
      "OUTPUT FILE NAME   : TMP.OUT float"
#
/bin/rm -f TMP.ASC
```

ヒアドキュメントの利用

ヒアドキュメント（埋め込み文書）とは、c s h 中に文書を埋め込む方法である。例では、

```
cat << EOF > TMP.ASC
  0.0  120.0
100.0  100.0
EOF
```

この部分が、ヒアドキュメントである。記号'<<'はインラインでの入力の切り替えを指定する。'cat << EOF'は、文字列 'EOF' がくるまでを読み込んで、それを出力する。この例では、'cat << EOF > TMP.ASC' となっているので、出力はリダイレクションによってファイル 'TMP.ASC' に出される。

コマンド 'atobi' は、ファイル 'TMP.ASC' を読み込んで、線形補間によってバイナリィデータを作成する。ファイル 'TMP.ASC' の第1列目は時刻 [単位 : m s e c] であり、第2列目は、なんらかの値（例えば基本周波数の値）である。

3. 4 b c a l c / s m a t h

b c a l c
s m a t h

バイナリデータ間の四則演算
アスキーデータに対する算術演算

b c a l c は、2つのバイナリファイルを読み込んで四則演算 (+, -, ×, ÷) をほどこした後、結果をバイナリデータとしてファイルに書込むか、または、アスキーデータとして標準出力装置(stdout)に出力する。

一方、s m a t h では、指数・対数演算、三角関数等の算術演算を行うことができ、標準入力装置(stdin)からアスキーデータを読み込んで、算術演算を入力したデータにほどこした後、演算結果を標準出力装置(stdout)に出力する。

第4章 ライブラリ

SpeechToolsのライブラリには、以下の4種類があり、C言語で記述されたプログラムから利用することができる。

libcommand.a	ファイルの入出力、コマンドのインタフェース等
libst.a	音響分析・合成、DFT、高次方程式の解法等
libmx.a	行列演算、連立方程式の解法等
libtlist.a	簡易なリスト処理等

通常、SpeechToolsのライブラリとヘッダファイルは、それぞれ/usr/local/st/libと、/usr/local/st/includeにインストールされており、SpeechToolsのライブラリを用いて新たにコマンドを開発する場合は、これらのディレクトリに注意しなければならない。このようなときには、`stmkmf` (SpeechTools Make Makefile)を利用するとよい。これは、シェルスクリプトであり、コンパイルとリンクに必要なMakefileと`main.c`を作成するものである。`stmkmf`によって作成されるMakefileにはライブラリやヘッダファイルのディレクトリ情報が記述されており、コンパイルやリンク等に必要なフラグが設定されている。また、`main.c`にはSpeechToolsのコマンドで実現されている統一化されたインターフェースを使ってコマンドを作成するために必要な変数の受け渡し等が記述されている。利用者は、これらのファイルを目的にしたがって変更することによって、独自のコマンドを作成することができる。

なお、SpeechToolsのライブラリは、UNIXの数学ライブラリである`libmx.a`を利用してしているので、新たにコマンドを開発するときは、このライブラリとリンクする必要がある。

第5章 まとめ

各種の計算機上で利用できる汎用的な音声研究用ツール(SpeechTools)の利用方法について詳しく述べてきたが、SpeechToolsは、以下に示すような特徴を持っている。

・単一化されたインターフェース

従来のUNIX上の音声研究用ツールでは、分析条件等のパラメータをコマンドの引数としてUNIXのコマンドラインから与える方式を採っているために各コマンドごとにどのようなオプションがあるのか、また、どのような意味を持つのかをマニュアル等で参照しなければならなかった。さらに、各コマンドごとに使い方が微妙に異なっている場合が多く、使い勝手が悪かった。これに対して、SpeechToolsでは、ユーティリティコマンド以外の約50個の音声分析などを行うコマンドの使い方が全て同じであり、コマンド自身が簡単な説明を行うということである。

・機能ごとに分れたコマンド

SpeechToolsは大きな一つのプログラムではなく、小さな機能ごとに個別に分けられたコマンドの集合体として設計されており、コマンドを連結することでより多くの目的に利用できるようになっている。また、データ変換などを行うユーティリティを備え、既存のUNIXコマンドとの連結を支援している。

・ファイルを介したデータの受け渡し

SpeechToolsでは、コマンドの連結に際してファイルを利用することである。UNIXにおいては、コマンドを連結する場合にパイプ機能を利用することができるが、パイプによるコマンドの連結はテキストデータを扱う場合には良いが、音声研究で利用するデータを1バイトのデータ列とするにはあまりにも複雑すぎ、また、ある種の分析手法では1つの入力データから2つ以上の分析結果を出力する場合があるのでパイプを使ってコマンドを連結するのは非常に困難である。そこで、SpeechToolsでは一部のユーティリティコマンドを除いてデータを受け渡しする方法としてファイルを用いている。

・煩雑なデータの変換の支援

音声研究においては、各種データ型のファイルを扱う必要があり、データ変換の作業は、必要不可欠である。そこで、SpeechToolsには、7個のユーティリティコマンドがあり、煩雑なデータ変換の作業を行えるようになっている。また、コマンドには、ファイルを読み込む際にデータフォーマットを自動的に変換する機能を持っているものもある。

・多数のライブラリ

SpeechToolsのライブラリには既存の音声分析/合成等に必要なアルゴリズムやファイルの入出力、データ変換等のプログラムの部品となる関数が数多く集められている。

・各種計算機上での利用が可能

現在、SpeechToolsは、以下の計算機上で動作することが確認されている。

FX/80	Alliant Computer Systems Co.
HP9000	Hewlett-Packard Company
MC6000	Concurrent Computer Co.
R6000/R3000	MIPS Computer Systems Inc.
SUN3/SUN4	Sun Microsystems Inc.
NEWS	Sony Co.
NeXT	NeXT Computer, Inc.
VAX8600	Digital Equipment Co.

今後の課題

一方、各機種上で利用するためにSpeechToolsには各機種ごとの特別なチューニングが行われていないので、各機種固有の機能であるA/D・D/Aやグラフィック等は含まれていないのと、対話的なインタフェースを持っていないのが欠点である。

現在、A/D・D/Aに関しては、MC6000(Concurrent Computer Co.)上でA/D・D/AコンバータMD8000(パベック電子開発)を、SUN3/SUN4(Sun Microsystems Inc.)上でDSPボードDSP56K(M. I. システムズ)をそれぞれ利用することができ、SpeechToolsとリンクしたコマンドを使用している。また、SpeechToolsのライブラリを一部取り入れて、SUN4上で波形表示/編集などの機能を持ったツールをDSP56KとX-Windowを利用して作成中である。これらのものは、SpeechToolsと密接には関係しておらず、それぞれ独立に動作するツールである。今後は、このようにSpeechToolsの周辺ツールとして、各種の装置や多様なウィンドウシステム上でA/D・D/Aや波形等をグラフィック表示するツールを個別に増やしていく必要があるが、SpeechToolsは、どのような状況下にもでも利用可能なコマンドやライブラリを多数備えているものと確信している。

筆者としては、特にNeXT上で動作するA/D・D/A、波形表示/編集、各種音響パラメータの表示/加工などの機能を持ったSpeechToolsのインタフェースを開発していきたいと考えている。

【謝辞】

SpeechTools作成に関して多くの方より御助言と御助力を頂きました。聴覚研究室に滞在していたAlain de Cheveigne氏や京都工芸繊維大学小林豊氏からは有益な御助言を頂戴しました。ここに深く感謝いたします。また、本稿を執筆する際して、草稿段階より多くの御助言を頂きました聴覚研究室の東倉洋一氏、平原達也氏(現NTT)、山田玲子氏、小林範子氏、小原和昭氏ならびに各位に感謝いたします。

付 録 .

SpeechTools(1)	SpeechTools – a powerful workbench system for the speech researcher.
Utility(1)	acat, bcat, cv, bcalc, merge, separate, smath – utility commands of SpeechTools
acat(1)	acat – convert ascii data (from 'stdin') to binary format file.
alp_fmt(1)	alp_fmt – Formant data from the alpha parameters of LPC method.
alp_spc(1)	alp_spc – Calculate spectrum envelope from LPC parameters.
atobi(1)	atobi – Convert the ASCII table to binary data.
autocorr(1)	autocorr – Calculate auto-correlation.
autoseg(1)	autoseg – Automatic segmentation.
bcalc(1)	bcalc – calculate a basic arithmetic expression (+ , - , * /) on binary data.
bcat(1)	bcat – convert binary format files to ascii data ('stdout').
bpx(1)	bpx – Band pass filter to eliminate noise components
cep_dist(1)	cep_dist – LPC Cepstrum Distance.
cv(1)	cv – convert between binary data formats.
dcep(1)	dcep – Calculation of spectrogram movement.
dft_bark(1)	dft_bark – Convert the frequency axis from [Hz] scale to [Bark] scale.
dft_cep(1)	dft_cep – Calculate DFT cepstrum smoothed envelope.
dft_forward(1)	dft_forward – Transfer time-domain data to complex data.
dft_inverse(1)	dft_inverse – Transfer complex data to time-domain data
dft_mel(1)	dft_mel – Convert the frequency axis from [Hz] scale to [Mel] scale.
dft_one(1)	dft_one – DFT one frame data
differential(1)	differential – Differential calculation.
emphasis(1)	emphasis – Self differential filtering.
euclid_dist(1)	euclid_dist – Calculate Euclid distance.
fft_run(1)	fft_run – Calculate the DFT running spectra.
fileshuffle(1)	fileshuffle – Shuffle file lists.
find_zerocrs(1)	find_zerocrs – Find the zero crossing points.
fmt_smooth(1)	fmt_smooth – Median smoothing for formant data.
ftrack(1)	ftrack – Formant Tracking by LPC method.
hpf1(1)	hpf1 – IIR high-pass filter.
iir_response(1)	iir_response – Calculate frequency response of IIR filter.
integral(1)	integral – Integral calculation.
lpc_cepst(1)	lpc_cepst – Calculate LPC CEPSTRUM.
lpc_rev(1)	lpc_rev – Calculate moving-average parameters.
lpc_run(1)	lpc_run – Calculate LPC running spectra.
lpf1(1)	lpf1 – IIR low-pass filter.
lstsq_smooth(1)	lstsq_smooth – Smoothing by least square method.
make_pair(1)	make_pair – Making Pairs.
median_smooth(1)	median_smooth – Median Smoothing.

merge(1)	merge – merge ascii streams
noise_wave(1)	noise_wave – Generating a band noise.
npulse(1)	npulse – Generating a pulse/noise source.
parcor(1)	parcor – Calculate PARCOR parameters.
peak_pick(1)	peak_pick – Peak-Picking for formants.
pitcher(1)	pitcher – Pitch extraction by auto-correlation method.
power(1)	power – Calculation of power with window normalization.
pulse(1)	pulse – Pulse generation
residual(1)	residual – Calculate residual error.
separate(1)	separate – split an ascii stream
sig_wave(1)	sig_wave – Generating a simple signal waveform.
smath(1)	smath – calculate arithmetic functions using the standard input/output stream.
spc_dist(1)	spc_dist – Calculate Spectrum Distortion.
subset(1)	subset – Cutting a subset of the file.
syn_alpha(1)	syn_alpha – Synthesizing with alpha parameters
syn_cascade(1)	syn_cascade – Cascade formant synthesizer.
syn_parcor(1)	syn_parcor – PARCOR Synthesizer.
syn_pole(1)	syn_pole – Single formant filter.
syn_zero(1)	syn_zero – Single anti-formant filter.
taper(1)	taper – Tapering the data.
transpose(1)	transpose – Transpose array (X \Leftrightarrow Y).
winfun(1)	winfun – Generating time-window functions.
zerocrs(1)	zerocrs – Counting the zero cross points.

Utility

acat(1)	acat – convert ascii data (from 'stdin') to binary format file.
bcat(1)	bcat – convert binary format files to ascii data ('stdout').
cv(1)	cv – convert between binary data formats.
merge(1)	merge – merge ascii streams
separate(1)	separate – split an ascii stream
bcalc(1)	bcalc – calculate a basic arithmetic expression (+ , - , * /) on binary data.
smath(1)	smath – calculate arithmetic functions using the standard input/output stream.

Analysis

peak_pick(1)	peak_pick – Peak-Picking for formants.
pitcher(1)	pitcher – Pitch extraction by auto-correlation method.
power(1)	power – Calculation of power with window normalization.
zerocrs(1)	zerocrs – Counting the zero cross points.

Axis

dft_bark(1)	dft_bark – Convert the frequency axis from [Hz] scale to [Bark] scale.
dft_mel(1)	dft_mel – Convert the frequency axis from [Hz] scale to [Mel] scale.

DFT

dft_cep(1)	dft_cep – Calculate DFT cepstrum smoothed envelope.
dft_forward(1)	dft_forward – Transfer time-domain data to complex data.
dft_inverse(1)	dft_inverse – Transfer complex data to time-domain data
dft_one(1)	dft_one – DFT one frame data
fft_run(1)	fft_run – Calculate the DFT running spectra.

Distortion

cep_dist(1)	cep_dist – LPC Cepstrum Distance.
euclid_dist(1)	euclid_dist – Calculate Euclid distance.
spc_dist(1)	spc_dist – Calculate Spectrum Distortion.

Edit

autoseg(1)	autoseg – Automatic segmentation.
subset(1)	subset – Cutting a subset of the file.
taper(1)	taper – Tapering the data.
transpose(1)	transpose – Transpose array (X \Leftrightarrow Y).

Experiment

fileshuffle(1)	fileshuffle – Shuffle file lists.
make_pair(1)	make_pair – Making Pairs.

Filter

bpfx(1)	bpfx – Band pass filter to eliminate noise components
emphasis(1)	emphasis – Self differential filtering.
iir_response(1)	iir_response – Calculate frequency response of IIR filter.
hpf1(1)	hpf1 – IIR high-pass filter.
lpf1(1)	lpf1 – IIR low-pass filter.

	syn_pole(1)	syn_pole – Single formant filter.
	syn_zero(1)	syn_zero – Single anti-formant filter.
Find		
	find_zeroocrs(1)	find_zeroocrs – Find the zero crossing points.
Interpolation		
	atobi(1)	atobi – Convert the ASCII table to binary data.
LPC		
	alp_fmt(1)	alp_fmt – Formant data from the alpha parameters of LPC method.
	alp_spc(1)	alp_spc – Calculate spectrum envelope from LPC parameters.
	dcep(1)	dcep – Calculation of spectrogram movement.
	ftrack(1)	ftrack – Formant Tracking by LPC method.
	lpc_cepst(1)	lpc_cepst – Calculate LPC CEPSTRUM.
	lpc_rev(1)	lpc_rev – Calculate moving-average parameters.
	lpc_run(1)	lpc_run – Calculate LPC running spectra.
	parcor(1)	parcor – Calculate PARCOR parameters.
Misc		
	autocorr(1)	autocorr – Calculate auto-correlation.
	differential(1)	differential – Differential calculation.
	integral(1)	integral – Integral calculation.
Smoothing		
	fmt_smooth(1)	fmt_smooth – Median smoothing for formant data.
	lstsq_smooth(1)	lstsq_smooth – Smoothing by least square method.
	median_smooth(1)	median_smooth – Median Smoothing.
Synthesis		
	npulse(1)	npulse – Generating a pulse/noise source.
	pulse(1)	pulse – Pulse generation
	residual(1)	residual – Calculate residual error.
	syn_alpha(1)	syn_alpha – Synthesizing with alpha parameters
	syn_cascade(1)	syn_cascade – Cascade formant synthesizer.
	syn_parcor(1)	syn_parcor – PARCOR Synthesizer.
Waveform		
	noise_wave(1)	noise_wave – Generating a band noise.
	sig_wave(1)	sig_wave – Generating a simple signal waveform.
	winfun(1)	winfun – Generating time-window functions.

NAME

SpeechTools – a powerful workbench system for the speech researcher.

DESCRIPTION

SpeechTools offers a collection of speech processing tools, both as stand-alone commands and as a library of routines. *SpeechTools* has three categories: **commands**, **utilities**, and **libraries**. **Commands** are stand-alone commands, which share the same usage (described later). **Utilities** are also stand-alone commands, which are concerned with conversion data format. **Libraries** are library function routines, which support a C-language interface.

[Commands]

All *SpeechTools* commands share the same usage and syntax. Basically, **commands** accept three usages:

```
% command
or
% command parameter_file
or
% command -o parameters...
```

The first form prints a list of the parameters that the **command** needs to run, together with their default values. This list can be redirected to a parameter file:

```
% command > parameter_file
```

After that, the *parameter_file* can be edited, and fed back to the **command** using the second form:

```
% command parameter_file
```

The **command** then runs using these parameters.

In the third form:

```
% command -o parameters...
```

parameters are specified on the command line and the **command** runs using these parameters. In any case, parameter entries obey the following format:

```
PARAMETER NAME : VALUE UNIT # COMMENT
```

Each parameter entry must be on a separate line. Spaces are allowed in the *PARAMETER NAME*, however, spaces are not allowed in the *VALUE* field. The *UNIT* field may be required for some parameters, in which case they are checked for validity when the **command** runs. Mismatching *UNIT* field causes a warning message. Because some **commands** support an automatic converting when the command reads/writes data from/to files. The *COMMENT* field is introduced by a '#' sign, and finishes at the end of the line. The *COMMENT* field is an optional note and is ignored when the **command** runs. Maximum line length for a parameter entry is fixed at 128 characters.

Example:

```
SAMPLING FREQUENCY      : 20 kHz
INPUT FILE NAME         : TMP.SRC short # short integers binary file
OUTPUT FILE NAME        : TMP.OUT float # floating point binary file
```

Using the third form, the same parameters specified on the command line or in a shell script:

```
command -o \
"SAMPLING FREQUENCY      : 20 kHz\"
"INPUT FILE NAME         : TMP.SRC short\"
"OUTPUT FILE NAME        : TMP.OUT float"
```

Parameters may be specified in any order. Some or all parameter entries may be missing. Missing parameters are given default values. Each command has its own default values and also, each user can set the values using a startup file in the user's home directory (*~/strc*), which she or he must own. File formats of *~/strc* is equal to those of *parameter_file*. If a parameter is repeated several times in the *~/strc* file, in the parameter file or on the command line, the last specification overrides.

In addition, **commands** allow three options:

- d enable debug mode
- s disable warning message
- w enable warning message

and special usages:

- i parameters specified from *stdin* in order to fork process
- x printing a command name and a list of the parameters in order to help writing shell scripts
- h printing help messages

File path names relative to current or home directory are allowed ('~', '.', '..'). But wildcard specifications ('*', '?', etc.) are not supported.

Note that the input and output of *SpeechTools* commands are always to or from a file. Standard input (*stdin*) and output (*stdout*) and pipes are not allowed. The reasons for this restriction are explained below. Exceptions to this rule are the following seven utilities.

[Utilities]

- acat** convert ascii data (from *stdin*) to binary format file.
- bcac** convert binary format files to ascii data (*stdout*).
- cv** convert between binary data formats.
- bcalc** calculate a basic arithmetic expression (+, -, *, /) on binary data.
- merge** merge ascii streams.
- separate** split an ascii stream.
- smath** arithmetic functions that operate on an ascii stream.

For a full explanation of the utilities, see the manual page for each specific utility and see also the manual page of **Utility(1)**.

[Libraries]

The *SpeechTools* routines are available in four libraries:

- libcommand.a** routines for feeding parameters to programs
- libst.a** routines for speech processing
- libmx.a** routines for matrix handling
- libtlist.a** routines for list handling

SpeechTools libraries support a C-language interface. For a full explanation of the libraries, see the manual page for each specific library function and see also the manual page of **Library(3)**.

PHILOSOPHY

The parameter mechanism of the *SpeechTools* commands is self-documenting, and is designed to give all commands a uniform usage style. The overhead may seem slightly heavy for simple commands, but it pays off quickly for commands that require more complex parameter specifications.

SpeechTools follows the convention from/into files for input and output of numerical data, to the exclusion of *stdin* and *stdout*. The reason for this restrictive convention is consistency. Many commands need to process several input channels, or produce multiple channels of data. Others require random access to all the data at once. Allowance of the standard input and output in some cases but not in others would seem arbitrary to the user, and would be difficult to document. Redirection of input and output and pipes are, therefore, not allowed: chains of commands must use temporary files. This should not cause too great a performance penalty, as modern file systems are well buffered.

INSTALLATION

The *SpeechTools* commands and utilities are normally installed in the `/usr/local/st/bin` directory. This directory should be in the path environment variable specified in `.login` or `.cshrc`. There are routine libraries and include files in the `/usr/local/st/lib` and `/usr/local/st/include`, respectively. When you want to make a new application using *SpeechTools* libraries, pay attention to the location of these directories. In this case, `stmkmf` (= *SpeechTools Make Makefile*) can help you. The `stmkmf` is a C-shell script, which makes a new directory and creates a *Makefile* and a *main.c* under the new directory. The *Makefile* involves information of routine libraries and include files for compiling and linking. The *main.c* is a template file for making a new command with *SpeechTools* libraries. Manual pages are also installed in the `/usr/local/st/man` directory. *SpeechTools* manual pages are written in *roff* format (with *eqn*). In order to read the manual pages on terminals (such as *tty* terminals), use `stman`, which is a C-shell script, instead of `man` command.

SpeechTools can be used on wide range of UNIX-based workstations. *SpeechTools* currently runs on the following systems:

FX/80	Alliant Computer Systems Co.
HP9000	Hewlett-Packard Company
MC6000	Concurrent Computer Co.
R6000/R3000	MIPS Computer Systems Inc.
SUN3/SUN4	Sun Microsystems Inc.
NEWS	Sony Co.
NeXT	NeXT Computer, Inc.
VAX8600	Digital Equipment Co.

It should be relatively easy to build the software on a variety of other UNIX systems.

FILES

`~/strc` user's startup file

AUTHOR

SpeechTools Copyright (C) 1990, 1991, 1992
 ATR Auditory and Visual Perception Research Laboratories,
 2-2 Hikoridai Seika-cho Soraku-gun Kyoto 619-02 Japan
 developed by Seiichi TENPAKU

NAME

acat, bcat, cv, bcalc, merge, separate, smath – utility commands of SpeechTools

DESCRIPTION

SpeechTools has seven utilities;

acat convert ascii data (from *stdin*) to binary format file.
bcat convert binary format files to ascii data (*stdout*).
cv convert between binary data formats.
bcalc calculate a basic arithmetic expression (+ , - , * , /) on binary data.
merge merge ascii streams.
separate split an ascii stream.
smath arithmetic functions that operate on an ascii stream.

They are very convenient for importing into *SpeechTools* from another tools or for exporting from *SpeechTools* into another tools, because most *SpeechTools* commands read or write binary files. For instance, if you have sound data as ascii format, you can convert the data into short integer binary data by using **acat** or by using **smath** and **acat**. In *SpeechTools* sound data format must be 16 bit bipolar, when your sound data is different from that, **smath** helps to convert the data. Then the data can be fed into *SpeechTools* commands.

merge and **separate** do it easy to make ascii tables of of data from *SpeechTools* outputs. For example, **pitcher** and **ftrack** produce F0 (fundamental frequency) and formant data in single precision floating point binary data. To produce an ascii table, you first apply **bcat**, then **merge** and **separate** as needed. In this case, **bcat** helps to get ascii data from any binary data in *SpeechTools*. If you just need F1, F2 and F3, **separate** can pick up F1, F2 and F3 from all formant frequency data by using following style;

```
example% bcat -p 8 -t "%4.1f" TMP.FRQ
961.6 1070.5 2076.4 3400.5 4695.9 5978.2 7253.1 8530.4
970.5 1043.2 2044.5 3383.2 4684.0 5969.4 7246.8 8526.7
974.8 1032.1 2008.9 3361.9 4667.8 5956.2 7236.5 8520.6
978.4 1026.5 1990.7 3351.4 4659.9 5949.9 7231.6 8517.7
example% bcat -p 8 -t "%4.1f" TMP.FRQ | separate -s '1 2 3'
961.6 1070.5 2076.4
970.5 1043.2 2044.5
974.8 1032.1 2008.9
978.4 1026.5 1990.7
```

Then you can make an ascii table by using **merge**. Of course, **awk** or other UNIX commands can do these operations, but they have many functions and they are more complicated to use than **merge** and **separate**.

smath supports many arithmetic functions [e.g. sin, cos, log, ...]. **bcalc** has four basic arithmetic functions: they are add, subtract, multiply and divide. **smath** calculates the four basic arithmetic functions, too. But the most difference point between **bcalc** and **smath** is input and output. Since **smath** reads input data from *stdin* and prints out the result to *stdout*, **smath** can be used in a *pipe*. When you need to calculate dB, use following sequence:

```
example% cat data.ascii
1
2
3
example% cat data.ascii | smath log10 | smath mul 10
0.000000
3.010300
4.771210
```

On the other hand, **bcalc** reads input data from two binary files and writes the result to a file or *stdout*,

so that `bcalc` can modify binary data files. If you need to square data, in a data file named 'foo', which is a single floating binary file:

```
example% bcat foo
1.000000
2.000000
3.000000
example% bcalc +m foo foo
1.000000
4.000000
9.000000
```

Anyway, there are many cases to use these seven utilities. Please see each manual page carefully. They can help solve intricate *conversion* problems.

SEE ALSO

`acat(1)`, `bcat(1)`, `cv(1)`, `bcalc(1)`, `merge(1)`, `separate(1)`, `smath(1)`

AUTHOR

Seiichi TENPAKU

Library(3)	libcommand.a, libst.a, libmx.a, libtlist – library routines of SpeechTools
alp_cep(3)	alp_cep_float, alp_cep_double, – conversion of linear prediction parameters into cepstrum parameters.
alp_par(3)	alp_par_float, alp_par_double, – conversion of linear prediction parameters into PARCOR parameters.
average(3)	average_float, average_double, average_short, average_int – calculate the average of value in an array.
axis(3)	MelToFreq, FreqToMel, FreqToBark, BarkToFreq, StageToBark, StageToFreq, critical_bandwidth, critical_frequency – frequency axis conversion.
cepstrum(3)	cepstrum, cepstrum_envelope – DFT cepstrum analysis.
column(3)	columnGet – get a string from a column string.
command(3)	CommandInit – command interface initializer for SpeechTools.
complex(3)	cx_zero, cx_add, cx_sub, cx_mul, cx_div, cx_conj, cx_exp, cx_pow, cx_root, cx_tocomplex, cx_abs, cx_arg, cx_print – complex function utilities.
convert(3)	convertArray, convertRead, convertWrite – convert data type of array.
cor_alp(3)	cor_alp_float, cor_alp_double – conversion of autocorrelation coefficients into linear prediction parameters.
correlation(3)	correlation_float, correlation_double – calculate correlation coefficients.
count(3)	countLine, countChar, countWord – calculate a number of lines, words and characters in file.
critical(3)	critical_damp – second order critically damped model.
differential(3)	differential_float, differential_double, differential_short, differential_int – calculate a the differential of an array.
dka(3)	aberth_init, dka_method, dka_solve – solving a high order polynomial expression.
dot(3)	dot_float, dot_double, dot_short, dot_int – calculate an inner product.
fft(3)	fft, ifft – discrete Fourier transform.
file(3)	fopenp, get_file_length, openRead, openWrite, binRead, binStore – file utilities of SpeechTools.
filter(3)	filter_lpf1, filter_hpf1, filter_dif1 – IIR filter function utilities.
find(3)	find_float_zero, find_float_peak, find_float_max, find_float_min, find_float_absmax, find_float_absmin, find_double_zero, find_double_peak, find_double_max, find_double_min, find_double_absmax, find_double_absmin, find_short_zero, find_short_peak, find_short_max, find_short_min, find_short_absmax, find_short_absmin – finding utilities.
gain(3)	gain_control_float, gain_control_double, gain_control_short – amplify the value of array.
getabspath(3)	getabspath – return the absolute pathname.
gethome(3)	gethome – get pathname of user's home directory.
getpwd(3)	getpwd – get the current working directory pathname.
interface(3)	IntrfcLineGet, IntrfcFindUnit, IntrfcPktFind, IntrfcPktAset, IntrfcPktRead, IntrfcPktWrite, IntrfcFileRead, IntrfcFileWrite, IntrfcPktStr, IntrfcStrPkt, IntrfcPktGet, IntrfcPktSet, – command interface utilities of SpeechTools.

- interpol(3)** interpol_gettype, interpol_linear, interpol_lagrange, interpol_blend_init, interpol_blend, interpol_spline_init, interpol_spline – interpolate function utilities.
- lagwin(3)** lagwin_float, lagwin_double – calculate a binomial window coefficients.
- lip(3)** lip_inverse_float, lip_inverse_double – eliminate lip radiation characteristics.
- lsp(3)** lsp_calc - calculate LSP (line spectrum pair) coefficients.
- lstsq(3)** lstsq_float, lstsq_calc_float, lstsq_double, lstsq_calc_double – least square method.
- median(3)** median_float – get the median value fo array.
- message(3)** debugMsgSet, debugMsg, errorMsgSet, errorMsg, warningMsgSet, warningMsg – message function utilities of SpeechTools.
- meter(3)** meterSwitch, meterSet, meterInc, meterEnd – meter function utilities of SpeechTools.
- mxalloc(3)** mx1alloc, mx2alloc, mx3alloc, mx1free, mx2free, mx3free – matrix allocation function utilities.
- mxcholesky(3)** mx_cholesky – matrix solve function utilities with Cholesky method.
- mxhouse(3)** mx_householder_solve – matrix solve function utilities with Householder method.
- mxmul(3)** mx_trans_mul, mx_mx_mul, mx_vc_mul – matrix multiply function utilities.
- mxsolve(3)** mx_ldua, mx_solve – matrix solve function utilities.
- mxtrans(3)** mx_transpose – transpose matrix.
- par_alp(3)** par_alp_float, par_alp_double, – conversion of PARCOR parameters into linear prediction parameters.
- pole_alpha(3)** pole_alpha - calculate the parameters of transfer function from the pairs of frequency and bandwidth.
- pole_zero(3)** pole_zero, pole_zero_fast - calculate the pairs of frequency and bandwidth from the parameters of transfer function.
- random(3)** random_var, random_num, random_seed – pseudo-random number generator.
- sig_cor(3)** sig_cor_float, sig_cor_double, – calculate autocorrelation coefficients.
- signal(3)** set_signal_handler – set the signal handlers of SpeechTools.
- smooth(3)** smooth_moving_ave – smoothing.
- stat(3)** stat_arithmetic_mean, stat_correlation, stat_variance – statistics functions.
- string(3)** strsave, strrev, string_compare, string_partition – string utilities of SpeechTools.
- swap(3)** swap_float, swap_double, swap_int, swap_long, swap_short, swap_char – swap data.
- synthe(3)** syn_direct, syn_twomul, syn_pole, syn_zero – speech synthesis function utilities.
- taper(3)** taper_gettype, taper_linear, taper_second, taper_sine, taper_sinc – taper function utilities.
- tlist(3)** tlist_read, tlist_getword, tlist_search, tlist_parent, tlist_next, tlist_print – tiny list utilities of SpeechTools.

winfun(3) winfun_gettype, winfun_set_func, winfun_set_short_float,
winfun_set_short_double, winfun_set_float_float, winfun_set_float_double,
winfun_generate_float, winfun_generate_double, – windowing function utilities.

zplane(3) ZpToFB – converts Z-plane data to frequency domain data.

NAME

libcommand.a, libst.a, libmx.a, libtlist – library routines of SpeechTools

DESCRIPTION

The *SpeechTools* routines are available in four libraries:

libcommand.a routines for feeding parameters to programs. See following manual pages: column(3) command(3) convert(3) count(3) file(3) getabspath(3) gethome(3) getpwd(3) interface(3) message(3) meter(3) signal(3) string(3)

libst.a routines for speech processing. See following manual pages: alp_cep(3) alp_par(3) average(3) axis(3) cepstrum(3) complex(3) cor_alp(3) correlation(3) critical(3) differential(3) dka(3) dot(3) fft(3) filter(3) find(3) gain(3) interpol(3) lagwin(3) lip(3) lsp(3) lstsq(3) median(3) par_alp(3) pole_alpha(3) pole_zero(3) random(3) sig_cor(3) smooth(3) stat(3) swap(3) synthe(3) taper(3) winfun(3) zplane(3)

libmx.a routines for matrix handling. See following manual pages: mxalloc(3) mxcholesky(3) mxhouse(3) mxmul(3) mxsolve(3) mxtrans(3)

libtlist.a routines for list handling. See following manual pages: tlist(3)

SpeechTools libraries support a C-language interface. When you want to make a new application using *SpeechTools* libraries, pay attention to the location of routine libraries and include files. In this case, **stmkmf** (= SpeechTools Make Makefile) can help you. The **stmkmf** is a C-shell script, which makes a new directory and creates a *Makefile* and a *main.c* under the new directory. See the *Makefile* and the *main.c*. Since *SpeechTools* libraries needs the **libmx.a**, which is a mathematical library in each UNIX system, the new application must be linked to the **libmx.a**.

LIST OF LIBRARY FUNCTIONS

Function Name	Manual Page	Library Name
BarkToFreq	axis(3)	libst.a
CommandInit	command(3)	libcommand.a
FreqToBark	axis(3)	libst.a
FreqToMel	axis(3)	libst.a
IntrfcFileRead	interface(3)	libcommand.a
IntrfcFileWrite	interface(3)	libcommand.a
IntrfcFindUnit	interface(3)	libcommand.a
IntrfcLineGet	interface(3)	libcommand.a
IntrfcPktAset	interface(3)	libcommand.a
IntrfcPktFind	interface(3)	libcommand.a
IntrfcPktGet	interface(3)	libcommand.a
IntrfcPktRead	interface(3)	libcommand.a
IntrfcPktSet	interface(3)	libcommand.a
IntrfcPktStr	interface(3)	libcommand.a
IntrfcPktWrite	interface(3)	libcommand.a
IntrfcStrPkt	interface(3)	libcommand.a
MelToFreq	axis(3)	libst.a
PointToBark	axis(3)	libst.a
PointToFreq	axis(3)	libst.a

ZpToFB	zplane(3)	<i>libst.a</i>
aberth_init	dka(3)	<i>libst.a</i>
alp_cep_double	alp_cep(3)	<i>libst.a</i>
alp_cep_float	alp_cep(3)	<i>libst.a</i>
alp_par_double	alp_par(3)	<i>libst.a</i>
alp_par_float	alp_par(3)	<i>libst.a</i>
average_double	average(3)	<i>libst.a</i>
average_float	average(3)	<i>libst.a</i>
average_int	average(3)	<i>libst.a</i>
average_short	average(3)	<i>libst.a</i>
binRead	file(3)	<i>libcommand.a</i>
binStore	file(3)	<i>libcommand.a</i>
cepstrum	cepstrum(3)	<i>libst.a</i>
cepstrum_envelope	cepstrum(3)	<i>libst.a</i>
columnGet	column(3)	<i>libcommand.a</i>
convertArray	convert(3)	<i>libcommand.a</i>
convertRead	convert(3)	<i>libcommand.a</i>
convertWrite	convert(3)	<i>libcommand.a</i>
cor_alp_double	cor_alp(3)	<i>libst.a</i>
cor_alp_float	cor_alp(3)	<i>libst.a</i>
correlation_double	correlation(3)	<i>libst.a</i>
correlation_float	correlation(3)	<i>libst.a</i>
countChar	count(3)	<i>libcommand.a</i>
countLine	count(3)	<i>libcommand.a</i>
countWord	count(3)	<i>libcommand.a</i>
critical_bandwidth	axis(3)	<i>libst.a</i>
critical_damp	critical(3)	<i>libst.a</i>
critical_frequency	axis(3)	<i>libst.a</i>
cx_abs	complex(3)	<i>libst.a</i>
cx_add	complex(3)	<i>libst.a</i>
cx_arg	complex(3)	<i>libst.a</i>
cx_conj	complex(3)	<i>libst.a</i>
cx_div	complex(3)	<i>libst.a</i>
cx_exp	complex(3)	<i>libst.a</i>
cx_mul	complex(3)	<i>libst.a</i>
cx_pow	complex(3)	<i>libst.a</i>
cx_print	complex(3)	<i>libst.a</i>
cx_root	complex(3)	<i>libst.a</i>

cx_sub	complex(3)	<i>libst.a</i>
cx_tocomplex	complex(3)	<i>libst.a</i>
cx_zero	complex(3)	<i>libst.a</i>
debugMsg	message(3)	<i>libcommand.a</i>
debugMsgSet	message(3)	<i>libcommand.a</i>
differential_double	differential(3)	<i>libst.a</i>
differential_float	differential(3)	<i>libst.a</i>
differential_int	differential(3)	<i>libst.a</i>
differential_short	differential(3)	<i>libst.a</i>
dka_method	dka(3)	<i>libst.a</i>
dka_solve	dka(3)	<i>libst.a</i>
dot_double	dot(3)	<i>libst.a</i>
dot_float	dot(3)	<i>libst.a</i>
dot_int	dot(3)	<i>libst.a</i>
dot_short	dot(3)	<i>libst.a</i>
errorMsg	message(3)	<i>libcommand.a</i>
errorMsgSet	message(3)	<i>libcommand.a</i>
fft	fft(3)	<i>libst.a</i>
filter_dif1	filter(3)	<i>libst.a</i>
filter_hpf1	filter(3)	<i>libst.a</i>
filter_lpf1	filter(3)	<i>libst.a</i>
find_double_absmax	find(3)	<i>libst.a</i>
find_double_absmin	find(3)	<i>libst.a</i>
find_double_max	find(3)	<i>libst.a</i>
find_double_min	find(3)	<i>libst.a</i>
find_double_peak	find(3)	<i>libst.a</i>
find_double_zero	find(3)	<i>libst.a</i>
find_float_absmax	find(3)	<i>libst.a</i>
find_float_absmin	find(3)	<i>libst.a</i>
find_float_max	find(3)	<i>libst.a</i>
find_float_min	find(3)	<i>libst.a</i>
find_float_peak	find(3)	<i>libst.a</i>
find_float_zero	find(3)	<i>libst.a</i>
find_short_absmax	find(3)	<i>libst.a</i>
find_short_absmin	find(3)	<i>libst.a</i>
find_short_max	find(3)	<i>libst.a</i>
find_short_min	find(3)	<i>libst.a</i>
find_short_peak	find(3)	<i>libst.a</i>

find_short_zero	find(3)	<i>libst.a</i>
fopenp	file(3)	<i>libcommand.a</i>
gain_control_double	gain(3)	<i>libst.a</i>
gain_control_float	gain(3)	<i>libst.a</i>
gain_control_short	gain(3)	<i>libst.a</i>
get_file_length	file(3)	<i>libcommand.a</i>
getabspath	getabspath(3)	<i>libcommand.a</i>
gethome	gethome(3)	<i>libcommand.a</i>
getpwd	getpwd(3)	<i>libcommand.a</i>
ifft	fft(3)	<i>libst.a</i>
interpol_blend	interpol(3)	<i>libst.a</i>
interpol_blend_init	interpol(3)	<i>libst.a</i>
interpol_gettype	interpol(3)	<i>libst.a</i>
interpol_lagrange	interpol(3)	<i>libst.a</i>
interpol_linear	interpol(3)	<i>libst.a</i>
interpol_spline	interpol(3)	<i>libst.a</i>
interpol_spline_init	interpol(3)	<i>libst.a</i>
lagwin_double	lagwin(3)	<i>libst.a</i>
lagwin_float	lagwin(3)	<i>libst.a</i>
lip_inverse_double	lip(3)	<i>libst.a</i>
lip_inverse_float	lip(3)	<i>libst.a</i>
lsp_calc	lsp(3)	<i>libst.a</i>
lstsq_calc_double	lstsq(3)	<i>libst.a</i>
lstsq_calc_float	lstsq(3)	<i>libst.a</i>
lstsq_double	lstsq(3)	<i>libst.a</i>
lstsq_float	lstsq(3)	<i>libst.a</i>
median_float	median(3)	<i>libst.a</i>
meterEnd	meter(3)	<i>libcommand.a</i>
meterInc	meter(3)	<i>libcommand.a</i>
meterSet	meter(3)	<i>libcommand.a</i>
meterSwitch	meter(3)	<i>libcommand.a</i>
mx1alloc	mxalloc(3)	<i>libmx.a</i>
mx1free	mxalloc(3)	<i>libmx.a</i>
mx2alloc	mxalloc(3)	<i>libmx.a</i>
mx2free	mxalloc(3)	<i>libmx.a</i>
mx3alloc	mxalloc(3)	<i>libmx.a</i>
mx3free	mxalloc(3)	<i>libmx.a</i>
mx_cholesky	mxcholesky(3)	<i>libmx.a</i>

mx_householder_solve	mxhouse(3)	<i>libmx.a</i>
mx_ldua	mxsolve(3)	<i>libmx.a</i>
mx_mx_mul	mxmul(3)	<i>libmx.a</i>
mx_solve	mxsolve(3)	<i>libmx.a</i>
mx_trans_mul	mxmul(3)	<i>libmx.a</i>
mx_transpose	mxtrans(3)	<i>libmx.a</i>
mx_vc_mul	mxmul(3)	<i>libmx.a</i>
openRead	file(3)	<i>libcommand.a</i>
openWrite	file(3)	<i>libcommand.a</i>
par_alp_double	par_alp(3)	<i>libst.a</i>
par_alp_float	par_alp(3)	<i>libst.a</i>
pole_alpha	pole_alpha(3)	<i>libst.a</i>
pole_zero	pole_zero(3)	<i>libst.a</i>
pole_zero_fast	pole_zero(3)	<i>libst.a</i>
random_num	random(3)	<i>libst.a</i>
random_seed	random(3)	<i>libst.a</i>
random_var	random(3)	<i>libst.a</i>
set_signal_handler	signal(3)	<i>libcommand.a</i>
sig_cor_double	sig_cor(3)	<i>libst.a</i>
sig_cor_float	sig_cor(3)	<i>libst.a</i>
smooth_moving_ave	smooth(3)	<i>libst.a</i>
stat_arithmetic_mean	stat(3)	<i>libst.a</i>
stat_correlation	stat(3)	<i>libst.a</i>
stat_variance	stat(3)	<i>libst.a</i>
string_compare	string(3)	<i>libcommand.a</i>
string_partition	string(3)	<i>libcommand.a</i>
strrev	string(3)	<i>libcommand.a</i>
strsave	string(3)	<i>libcommand.a</i>
swap_char	swap(3)	<i>libst.a</i>
swap_double	swap(3)	<i>libst.a</i>
swap_float	swap(3)	<i>libst.a</i>
swap_int	swap(3)	<i>libst.a</i>
swap_long	swap(3)	<i>libst.a</i>
swap_short	swap(3)	<i>libst.a</i>
syn_direct	synthe(3)	<i>libst.a</i>
syn_pole	synthe(3)	<i>libst.a</i>
syn_twomul	synthe(3)	<i>libst.a</i>
syn_zero	synthe(3)	<i>libst.a</i>

taper_gettype	taper(3)	<i>libst.a</i>
taper_linear	taper(3)	<i>libst.a</i>
taper_second	taper(3)	<i>libst.a</i>
taper_sinc	taper(3)	<i>libst.a</i>
taper_sine	taper(3)	<i>libst.a</i>
tlist_getword	tlist(3)	<i>libtlist.a</i>
tlist_next	tlist(3)	<i>libtlist.a</i>
tlist_parent	tlist(3)	<i>libtlist.a</i>
tlist_print	tlist(3)	<i>libtlist.a</i>
tlist_read	tlist(3)	<i>libtlist.a</i>
tlist_search	tlist(3)	<i>libtlist.a</i>
warningMsg	message(3)	<i>libcommand.a</i>
warningMsgSet	message(3)	<i>libcommand.a</i>
winfun_generate_double	winfun(3)	<i>libst.a</i>
winfun_generate_float	winfun(3)	<i>libst.a</i>
winfun_gettype	winfun(3)	<i>libst.a</i>
winfun_set_float_double	winfun(3)	<i>libst.a</i>
winfun_set_float_float	winfun(3)	<i>libst.a</i>
winfun_set_func	winfun(3)	<i>libst.a</i>
winfun_set_short_double	winfun(3)	<i>libst.a</i>
winfun_set_short_float	winfun(3)	<i>libst.a</i>