# Documentation for matlab_speech_features

This is documentation for matlab_speech_features, a library for speech feature extraction. Code is available at https://github.com/jameslyons/matlab_speech_features. If you find any errors, feel free to make a pull request or leave a comment at the bottom of the page.

 Download matlab_speech_features.zip

## msf_mfcc - Mel Frequency Cepstral Coefficients

```
function feat = msf_mfcc(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes Mel frequency cepstral coefficients for each frame. For a tutorial on MFCCs, see MFCC tutorial.

- `speech` – the input speech signal, vector of speech samples
- `fs` – the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` – length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` – step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'nfilt'` – the number filterbanks to use. Default: 26
- `'lowfreq'` – the lowest filterbank edge. In Hz. Default: 0
- `'highfreq'` – the highest filterbank edge. In Hz. Default: fs/2
- `'nfft'` – the FFT size to use. Default: 512
- `'ncep'` – the number of cepstral coeffients to use. Default: 13
- `'liftercoeff'` – liftering coefficient, 0 is no lifter. Default: 22
- `'appendenergy'` – if true, replaces 0th cep coeff with log of total frame energy. Default: true

Example usage:

```
mfccs = msf_mfcc(signal,16000,'nfilt',40,'ncep',12);
```

## msf_lpc - Linear Prediction Coefficients

```
function feat = msf_lpc(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes Linear Prediction Coefficients for each frame.

- `speech` – the input speech signal, vector of speech samples

- `fs` – the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` – length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` – step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'order'` – the number of coefficients to return. Default: 12

Example usage:

```
lpcs = msf_lpc(signal,16000,'order',10);
```

## msf_rc - Reflection Coefficients

```
function feat = msf_rc(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes Reflection Coefficients for each frame.

- `speech` – the input speech signal, vector of speech samples
- `fs` – the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` – length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` – step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'order'` – the number of coefficients to return. Default: 12

Example usage:

```
rcs = msf_rc(signal,16000,'order',10);
```

## msf_logfb - Log Filterbank Energies

```
function feat = msf_logfb(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes log filterbank energies for each frame.

- `speech` – the input speech signal, vector of speech samples
- `fs` – the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` – length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` – step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'nfilt'` – the number filterbanks to use. Default: 26
- `'lowfreq'` – the lowest filterbank edge. In Hz. Default: 0
- `'highfreq'` – the highest filterbank edge. In Hz. Default: fs/2
- `'nfft'` – the FFT size to use. Default: 512

Example usage:

```
logfbs = msf_logfb(signal,16000,'nfilt',40,'ncep',12);
```

## msf_filterbank - return a mel-spaced filterbank

```
function fbank = msf_filterbank(nfilt,fs,lowfreq,highfreq,nfft)
```

returns a mel-spaced filterbank for use with filterbank energies, mfccs, sscs etc.

- `nfilt` - the number filterbanks to use.
- `fs` - the sample rate of 'speech', integer
- `lowfreq` - the lowest filterbank edge. In Hz.
- `highfreq` - the highest filterbank edge. In Hz.
- `nfft` - the FFT size to use.

Example usage:

```
lpcs = msf_filterbank(26,16000,0,16000,512);
```

## msf_lsf - Line Spectral Frequencies

```
function feat = msf_lsf(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes Line Spectral Frequencies for each frame.

- `speech` - the input speech signal, vector of speech samples
- `fs` - the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` - length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` - step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'order'` - the number of coefficients to return. Default: 12

Example usage:

```
lsfs = msf_lsf(signal,16000,'order',10);
```

## msf_lpcc - Log Area Ratios

```
function feat = msf_lar(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes Log Area Ratios for each frame.

- `speech` - the input speech signal, vector of speech samples
- `fs` - the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` – length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` – step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'order'` – the number of coefficients to return. Default: 12

Example usage:

```
lars = msf_lar(signal,16000,'order',10);
```

## msf_framesig - break a signal into frames

```
function win_frames = msf_framesig(signal, frame_len, frame_step, winfunc)
```

Takes a 1 by N signal, and breaks it up into frames. Each frame starts *frame_step* samples after the start of the previous frame. Each frame is windowed by wintype.

– to specify window, use e.g. @hamming, @(x)chebwin(x,30), @(x)ones(x,1), etc.

- `signal` – the input signal, vector of audio samples
- `frame_len` – length of window in samples.
- `frame_step` – step between successive windows in seconds. In samples.
- `winfunc` – A function to be applied to each window.

Example usage with hamming window:

```
frames = msf_framesig(speech, winlen*fs, winstep*fs, @(x)hamming(x));
```

## msf_lpcc - Linear Prediction Cepstral Coefficients

```
function feat = msf_lpcc(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes Linear Prediction Cepstral Coefficients for each frame.

- `speech` – the input speech signal, vector of speech samples
- `fs` – the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` – length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` – step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'order'` – the number of coefficients to return. Default: 12

Example usage:

```
lpccs = msf_lpcc(signal,16000,'order',10);
```

## msf_ssc - Spectral Subband Centroids

```
function feat = msf_ssc(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes Spectral Subband Centroids for each frame.

- `speech` – the input speech signal, vector of speech samples
- `fs` – the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` – length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` – step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'nfilt'` – the number filterbanks to use. Default: 26
- `'lowfreq'` – the lowest filterbank edge. In Hz. Default: 0
- `'highfreq'` – the highest filterbank edge. In Hz. Default: fs/2
- `'nfft'` – the FFT size to use. Default: 512

Example usage:

```
sscs = msf_ssc(signal,16000,'nfilt',40,'ncep',12);
```

## msf_powspec - Compute power spectrum of audio frames

```
function pspec = msf_powspec(speech,fs,varargin)
```

given a speech signal, splits it into frames and computes the power spectrum for each frame.

- `speech` – the input speech signal, vector of speech samples
- `fs` – the sample rate of 'speech', integer

optional arguments supported include the following 'name', value pairs from the 3rd argument on:

- `'winlen'` – length of window in seconds. Default: 0.025 (25 milliseconds)
- `'winstep'` – step between successive windows in seconds. Default: 0.01 (10 milliseconds)
- `'nfft'` – the FFT size to use. Default: 512

Example usage:

```
lpcs = msf_powspec(signal,16000,'winlen',0.5);
```