

Moments: batch speech spectrum moments analysis

Paul H. Milenkovic *

Department of Electrical and Computer Engineering
University of Wisconsin-Madison
1415 Johnson Drive
Madison, Wisconsin 53706

Copyright 1999
All Rights Reserved

The DOS batch command `moments` analyses one or more waveform files for spectrum moments and places the results in text files. Moments are measures of the speech spectrum that are useful in classifying consonant sounds as discussed by K. Forrest, G. Weismer, P. Milenkovic, and R. N. Dougall (1988): Statistical analysis of word-initial voiceless obstruents: Preliminary data, *Journal of the Acoustical Society of America* Volume 84, pp. 115-123.

At the DOS prompt invoke

```
moments *.wav
```

to compute moments for all waveform files with the `.wav` extension in the current directory, placing the results in ASCII files with `.mmt` extension. Waveform files in a variety of formats are acceptable, including SPHERE (TIMIT database) and RIFF (Microsoft Windows `.wav` format) as well as the formats produced by Kay CSL, CSpeech, CSRE, ILS, and Sensimetrics Speech Station.

The default analysis window length is 20 ms, and the default time step is 10 ms. The time label for the moments analysis is at the center of the analysis window. It is common practice to save a short waveform file containing a speech event for moments analysis. You may want to center the first moment analysis on some acoustic event like a burst. Make sure you save the waveform with 10 ms of waveform leader before the burst to ensure that the automatic analysis centers the first window on that burst.

The batch file `moments.bat` invokes the moments filter `cmmt.exe`, which has the following inputs, outputs, and parameter settings.

*Please refer inquiries to Paul Milenkovic, 118 Shiloh Dr., Madison, Wisconsin 53705-2433, U.S.A., 608/833-7956

cmmt

Input Signal waveform.

Output Text (ASCII) file with `.mmt` extension. Each analysis frame produces the line of output

$$t \ dB \ L1 \ L2 \ L3 \ L4 \ B1 \ B2 \ B3 \ B4 \ \tau=t \ fname$$

where t is the time of the center of the analysis frame in ms, and dB is the decibel value for that analysis frame. The value t appears in two places, giving maximum flexibility in the use of the DOS filters `sort` and `find` to organize these data. The entry $fname$ is the name of the file from which the data originated. The numeric values $L1-L4$ are the linear frequency scale spectrum moments, and $B1-B4$ are the Bark scale spectrum moments. Moment 1 is the mean while Moment 2 is the standard deviation. For the linear scale Moments 1 and 2 have units of kHz) while for the Bark scale the units are Bark. Moments 3 and 4 are dimensionless versions of the higher moments which contain the second moment information in the normalization used to make them dimensionless. Moment 3 is the coefficient of skewness and Moment 4 is the coefficient of kurtosis.

Parameters

/W: w Analysis window w in ms: default is 20 ms (maximum limited by 1024 samples times dt between waveform samples – at a sampling rate of 22 kHz this is 1024 times .045 ms or 46 ms)

/T: t Time step between each moments computation t in ms: default is 10 ms (default of 20 ms window with 10 ms steps produces 50 percent overlap).

/I: i Initial spectrum frequency in kHz for the moments calculation. Default is zero.

/F: b Final spectrum frequency in kHz for the moments calculation. Default is half the sampling frequency.