

NOIZEUS: A noisy speech corpus for evaluation of speech enhancement algorithms

A noisy speech corpus (NOIZEUS) was developed to facilitate comparison of speech enhancement algorithms among research groups. The noisy database contains 30 IEEE sentences (produced by three male and three female speakers) corrupted by eight different real-world noises at different SNRs. The noise was taken from the AURORA database and includes suburban train noise, babble, car, exhibition hall, restaurant, street, airport and train-station noise. This corpus is available to researchers free of charge.

The description of this corpus was published in the following paper, which we ask that you cite when using NOIZEUS:

Hu, Y. and Loizou, P. (2007). "Subjective evaluation and comparison of speech enhancement algorithms," *Speech Communication*, 49, 588-601. [[pdf](#)]

The NOIZEUS corpus was also used by our lab to evaluate the correlations of common objective measures used in speech enhancement. This work was reported in the following papers:

Hu, Y. and Loizou, P. (2008). "[Evaluation of objective quality measures for speech enhancement](#)," *IEEE Transactions on Speech and Audio Processing*, 16(1), 229-238. [[Matlab code](#)]

Ma, J., Hu, Y. and Loizou, P. (2009). "Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions", *Journal of the Acoustical Society of America*, 125(5), 3387-3405 [[pdf](#)]

Speech Material

Thirty sentences from the IEEE sentence database [1] were recorded in a sound-proof booth using Tucker Davis Technologies (TDT) recording equipment. The sentences were produced by three male and three female speakers. The IEEE database (720 sentences) was used as it contains phonetically-balanced sentences with relatively low word-context predictability. The thirty sentences were selected from the IEEE database so as to include all phonemes in the American English language. The list of sentences recorded for NOIZEUS are given in [Table 1](#). The sentences were originally sampled at 25 kHz and downsampled to 8 kHz.

Filtering

To simulate the receiving frequency characteristics of telephone handsets, the speech and noise signals were filtered by the modified Intermediate Reference System (IRS) filters used in ITU-T P.862 [1] for evaluation of the PESQ measure. The filter frequency response is shown in Figure 1.

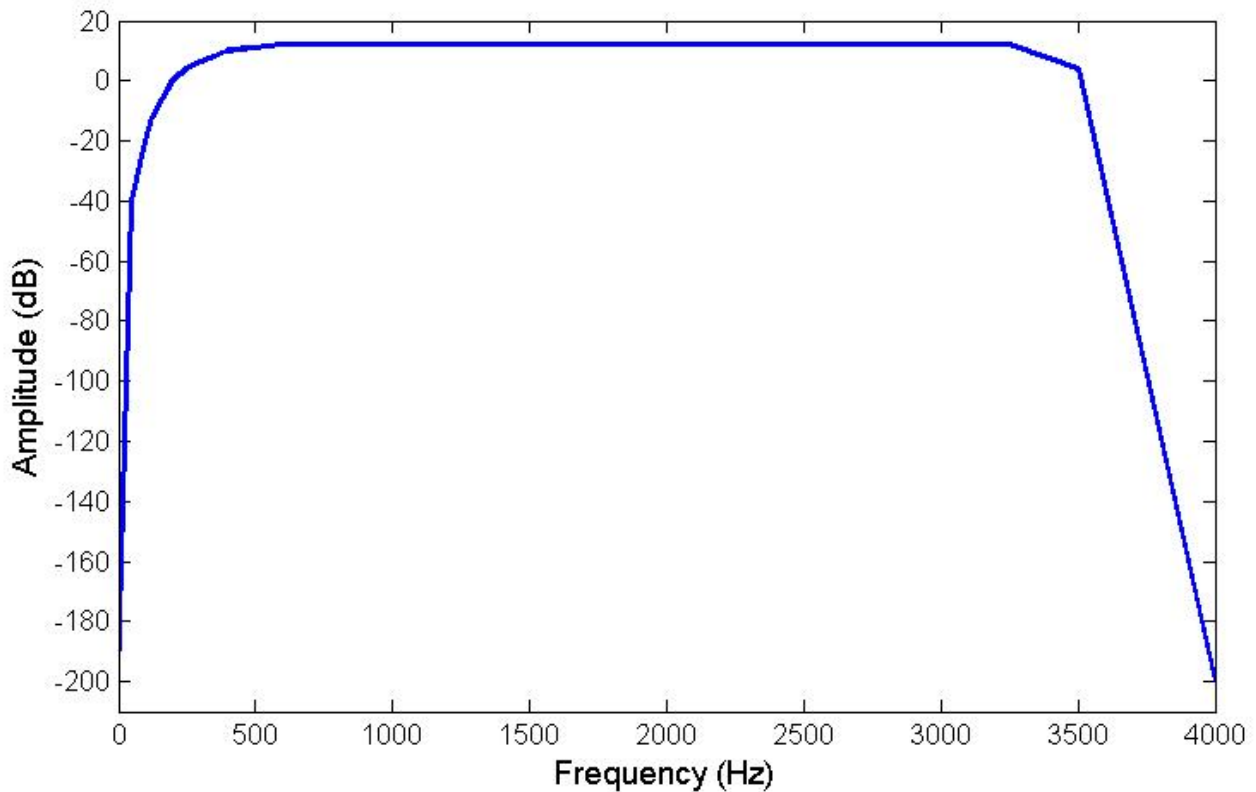


Figure 1: Frequency response of IRS filter.

Adding Noise

Noise is artificially added to the speech signal as follows. The IRS filter is independently applied to the clean and noise signals. The active speech level of the filtered clean speech signal is first determined using the method B of ITU-T P.56 [3]. A noise segment of the same length as the speech signal is randomly cut out of the noise recordings, appropriately scaled to reach the desired SNR level and finally added to the filtered clean speech signal.

Noise signals were taken from the AURORA database [4] and included the following recordings from different places:

- Babble (crowd of people)
- Car
- Exhibition hall
- Restaurant
- Street
- Airport
- Train station
- Train

The long-term spectra of the above noises are given in [4]. The noise signals were added to the speech signals at SNRs of 0dB, 5dB, 10dB, and 15dB.

Download files

All files were saved in Windows wav format (16 bit PCM, mono).

- **Clean files** [Clean.zip](#)

Train noise

SNR (dB)	0	5	10	15
Noisy files (.zip)	train_0dB	train_5dB	train_10dB	train_15dB

Babble noise

SNR (dB)	0	5	10	15
----------	---	---	----	----

Noisy files (.zip)	babble_0dB	babble_5dB	babble_10dB	babble_15dB
--------------------	----------------------------	----------------------------	-----------------------------	-----------------------------

Car noise

SNR (dB)	0	5	10	15
Noisy files (.zip)	car_0dB	car_5dB	car_10dB	car_15dB

Exhibition hall noise

SNR (dB)	0	5	10	15
Noisy files (.zip)	exhibition_0dB	exhibition_5dB	exhibition_10dB	exhibition_15dB

Restaurant noise

SNR (dB)	0	5	10	15
Noisy files (.zip)	restaurant_0dB	restaurant_5dB	restaurant_10dB	restaurant_15dB

Street noise

SNR (dB)	0	5	10	15
Noisy files (.zip)	street_0dB	street_5dB	street_10dB	street_15dB

Airport noise

SNR (dB)	0	5	10	15
Noisy files (.zip)	airport_0dB	airport_5dB	airport_10dB	airport_15dB

Train station

SNR (dB)	0	5	10	15
Noisy files (.zip)	station_0dB	station_5dB	station_10dB	station_15dB

References

- [1] IEEE Subcommittee (1969). IEEE Recommended Practice for Speech Quality Measurements. *IEEE Trans. Audio and Electroacoustics*, AU-17(3), 225-246.
- [2] ITU P.862 (2000). Perceptual evaluation of speech quality (PESQ), and objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs. *ITU-T Recommendation P. 862*
- [3] ITU-T P.56 (1993). Objective measurement of active speech level.
- [4] H. Hirsch, and D. Pearce (2000). "The Aurora Experimental Framework for the Performance Evaluation of Speech Recognition Systems under Noisy Conditions." *ISCA ITRW ASR2000*, Paris, France, September 18-20.

Questions or Feedback about NOIZEUS

Any inquiries or suggestions about NOIZEUS, please direct them to Dr. Philip Loizou (email: loizou@utdallas.edu)

[Home](#)