ABSTRACT

In the last twenty years, much has been done on speech enhancement but in most cases, methods have been tested with additive white Gaussian noise and sometimes are only derived in this case.

This paper deals with real noise processing and essentially speech enhancement in a car environment. For full hand-free radiocommunications there are three needs, two for transmission (enhancement for speech to be transmitted and echo cancelling) and one for vocal dialing (good recognition in noise). Unfortunately car noise is not white and methods have to be tried on this kind of noise. In this paper only real car noise and speech recordings (1) are tested.

First, methods that have been tested for transmission are presented. With regard to echo cancelling the classical method of LMS is tested and derived for the case of car noise. Two methods of speech enhancement for transmission have been studied: the MPLPC coding system with an ARMA estimation (2) and a method based on Forward-Backward adaptive filtering (3) that, at same performances, doesn't require pitch period estimation as Sambur's approach (4). The segmental signal to noise ratio between clean speech, speech corrupted by white noise, speech corrupted by car noise are computed. For speech recorded in a car, the only way to characterize the improvements is to listen.

Second, the goal is to increase the recognizer accuracy: this can be done by modifying different parts: robust AR estimation, robust distance and adaptation of the codebook. Three methods are implemented: the Short-Time Modified Coherence representation (5), the Capstral Projection Measure (6), the noise-adapting codebook (7). Their performances are compared in terms of word accuracy.

These comparative studies show that improvements obtained with additive white Gaussian noise may be less important for real applications. Nevertheless, we can improve performances of algorithms by preprocessing the noisy signal.

For each function of a whole hand-free radio telephone with vocal dialing, best approaches are retained and adapted. The whole system is implemented on a hardware based on ATT DSP 32C.

References

(1) I Lecomte, M Lever, J Boudy, A Tassy, "Car noise processing for speech input", to be published in Proc. IEEE ICASSP 89.
(7) D B Roe, "Speech Recognition with a noise-adapting codebook", Proc IEEE ICASSP87