عنوان مقاله:

New VAD Algorithm using Sparse Representation in Spectro Temporal Domain

نویسندگان:

Mohadese Eshaghi - *Department of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.*

Farbod Razzazi - *Department of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.*

Alireza Behrad - *Department of Electrical and Electronic Engineering, Shahed University, Tehran, Iran*

خلاصه مقاله:

This paper proposes two algorithms for Voice Activity Detection (VAD) based on sparse representation in spectro-temporal domain. Spectral-temporal components which, in addition to the frequency and time dimensions, have two other dimensions of the scale and rate. Scale means spectral modulation and the rate means temporal modulation. On the other hand, using sparse representation in learning dictionaries of speech and noise, separate the speech and noise segment to be better separated. The first algorithm was made using two-dimensional STRF (Spectro-Temporal Response Field) space based on sparse representation. Dictionaries with different atomic sizes and two dictionary learning methods: NMF (non-negative matrix factorization) and the K-SVD (k-means clustering method), were investigated in this approach. This algorithm revealed good results at high SNRs (signal-to-noise ratio). The second algorithm, whose approach is more complicated, suggests a speech detector using the sparse representation in four-dimensional STRF space. Due to the large volume of STRF s four-dimensional space, this space was divided into cubes, with dictionaries made for each cube separately by NMF (non-negative matrix factorization) learning algorithm. Simulation results were presented to illustrate the effectiveness of our new VAD algorithms. The results revealed that the achieved performance was 90.11% and 91.75% under -5 dB SNR in white and car noise respectively, outperforming most of the state-of-the-art VAD algorithms

کلمات کلیدی:

Speech Processing; Voice Activity Detector; VAD; Spectro-Temporal Domain Representation; Sparse Representation, NMF, K-SVD.

لینک ثابت مقاله در پایگاه سیویلیکا:

https://civilica.com/doc/993171