

Improved design of dtw and gmm cascaded arabic speaker verification./ Shuoshuo Chen, Junbo Zhao, Ruiqi Yang

Shuoshuo Chen, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20448193&lokasi=lokal>

Abstrak

In this paper, we discuss about the design, implementation and assessment of a two-stage Arabic speaker recognition system, which aims to recognize a target Arabic speaker among several people. The first stage uses improved DTW (Dynamic Time Warping) algorithm and the second stage uses SA-KM-based GMM (Gaussian Mixture Model). MFCC (Mel Frequency Cepstral Coefficients) and its differences form, as acoustic feature, are extracted from the sample speeches. DTW provides three most possible speakers and then the recognition results are conveyed to GMM training processes. A specified similarity assessment algorithm, KL distance, is applied to find the best match with the target speaker. Experimental results show that text-independent recognition rate of the cascaded system reaches 90 percent.

Dalam paper ini, kami membahas desain, implementasi dan penilaian sistem pengenalan dua tahap untuk penutur Bahasa Arab, yang bertujuan untuk mengenali target penutur Bahasa Arab di antara beberapa orang. Tahap pertama menggunakan algoritma improved DTW (Dynamic Time Warping) dan tahap kedua menggunakan SA-KM berbasis GMM (Gaussian Mixture Model). MFCC (Mel Frequency Cepstral Coefficients) dan variasi perbedaannya, seperti fitur akustik, diekstrak dari sample suara. DTW menyediakan tiga penutur yang paling mungkin dan kemudian hasil pengenalan diteruskan ke proses pelatihan GMM. Sebuah algoritma penilaian kesamaan yaitu KL distance, diaplikasikan untuk menemukan pasangan yang paling cocok dengan penutur sasaran. Hasil penelitian menunjukkan bahwa tingkat pengenalan teks-independen dari sistem mencapai 90 persen.