

Modul parametric filtering untuk segmentasi sinyal ucapan suku kata bahasa Indonesia

Aloysius Nugroho S., author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=71746&lokasi=lokal>

Abstrak

Ada tiga tahap utama dalam pengenalan ucapan yaitu tahap segmentasi untuk menemukan batas-batas fonem, tahap ekstraksi untuk mendapatkan nilai atas segmen-segmen sinyal dan tahap klasifikasi untuk perbandingan nilai segmen dengan nilai acuan. Di antara ketiganya, segmentasi merupakan tahap yang paling menentukan, karena jika keluarannya salah maka dua tahap berikutnya menjadi sia-sia.

Berbagai metode dipergunakan dalam teknik segmentasi. Salah satunya Adapted Local Trigonometric Transforms (ALIT) yang mendekomposisi sinyal suara menjadi partisi-partisi trigonometris dan orthonormal dengan pemanfaatan DCT (Discrete Cosine Transform) yang memang memiliki keunggulan dalam hal kompresi sinyal [7]. Meski demikian, teknik ini juga memiliki kekurangan yaitu sensitif terhadap pergeseran sinyal ucapan.

Teknik lain yang lebih tahan terhadap pergeseran sinyal ucapan dan bahkan derau adalah Parametric Filtering (PF) [8]. Teknik ini menggabungkan parametric filter bank dengan sebuah analisa otokorelasi selisih-satu (lag-one autocorrelation) untuk menghasilkan fungsi karakterisasi yang baru sama sekali dan tahan terhadap perubahan frekuensi dan derau. Otokorelasi selisih-satu merupakan korelasi sinyal terhadap sinyal itu sendiri setelah mengalami penundaan sebesar satu sekuensi. Fungsi ini mampu menampilkan dan menangkap perubahan spektral dan energi walau dalam domain waktu.

Sinyal di-filter dalam sebuah filter bank yang berupa Parametric HR filter all-pole orde satu, lalu dikolerasikan dan didemodulasi untuk diambil yang diperlukan. Sambil dilakukan pengukuran distorsi untuk mencari nilai batas segmen, sinyal dapat ditampilkan sehingga mempermudah analisa. Tesis ini dilakukan dengan mensegmentasikan suku-kata bahasa Indonesia secara hampir real time dengan penundaan maksimum 10 detik.

Analisa dilakukan terhadap performansi sistem dari sisi akurasi, waktu proses dan stabilitas. Diharapkan ini akan menjadi langkah awal yang tepat bagi tahap-tahap pengenalan ucapan selanjutnya dalam merintis sistem Voice To Text yang sesuai untuk suku-kata bahasa Indonesia.

<hr>

There are three major steps in speech recognition which are accordingly segmentation for end-point detection, extraction to get the values upon the segmented signals and classification that yielding the corresponding references. Among the three, segmentation is the most important step since once it's output is incorrect then the next two will be nothing.

There are methods used in this step. Among others are the so-called Adapted Local Trigonometric

Transforms (ALIT) which decompose signals into orthonormal-trigonometrical parts using Discrete Cosine Transform (DC?) that is good in signal compressing [7]. Anyway, this method lacks of sensitivity over unstable signals.

Another method that is robust to such sigwls as well as noise is Parametric Filtering (PF) [8]. This method combines a parametric filter bank and a lag-one autocorrelation analysis to produce a new characterisation function that overcome the frequency shift and noise. Lag-one autocorrelation means correlating the signal toward itself after one-sequence delayed. It also displays and catches spectral and energy changes even in time domain.

Signal is filtered in a filter banks which are all-pole first order parametric IIR filter, after which it is correlated by lag-one autocorrelation. Next, signal is demodulated to yield what is needed. While being processed by distortion measures to find end-points, the signal can be displayed to make the analysis visible. Research done by segmenting Indonesia syllabels in almost real time with ten-second maximum delay.

There are analysis upon system performances in accusation, time consuming and stability point of view. Hopefully, this will become first step for the next speech recognition steps and lead to the suitable VoiceToText system for Indonesia syllabels.