

Optimasi Parameter Splitting dan Asimilasi dalam Automatic Indonesian Speech Segmentation dengan Metode Algoritma Genetika

Riksa Meidy Karim¹, Suyanto²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung
¹riksameidy@students.telkomuniversity.ac.id, ²suyanto@telkomuniversity.ac.id

Abstrak

Segmentasi merupakan sub sistem dari *Speech Recognition*. Proses segmentasi suara pada Bahasa Indonesia lebih cocok dibuat ke dalam silabel sebagai satuan dasarnya. Penelitian sebelumnya telah menerapkan segmentasi pada Bahasa Indonesia dan menambahkan proses *splitting* dan asimilasi untuk meningkatkan akurasi dan menurunkan *insertion error* dan *deletion error* namun parameter yang digunakan untuk proses *splitting* dan asimilasi menggunakan observasi manual sehingga hasilnya belum dapat optimal. Penelitian ini membahas optimisasi parameter *boundary detection*, *splitting* dan asimilasi pada *Automatic Indonesian Speech Syllable Segmentation* dan memperkenalkan dua modifikasi *post-processing* yaitu *iterative splitting* dan *iterative assimilation*. Optimisasi dilakukan menggunakan algoritma genetika, dan masing-masing tiga metode dioptimalkan menggunakan algoritma genetika tersendiri. Parameter yang dioptimalkan menyebabkan *iterative assimilation* berhasil mengurangi *insertion error* dan membuat segmen yang diharapkan tidak tergabung secara berlebihan. *Iterative splitting* dilakukan lebih baik dibandingkan dengan *splitting*, memberikan 1.9% akurasi yang lebih baik dan 2.15% *deletion error* yang lebih sedikit, sedangkan *iterative assimilation* memberikan 4.29% *insertion error* yang lebih sedikit daripada asimilasi. *Iterative splitting* yang dipadukan secara berurutan dengan *iterative assimilation* memberikan akurasi terbaik sebesar 88.57% dan *deletion error* terkecil sebesar 14.55% dalam 110 *Indonesian speech dataset* ini.

Kata kunci : *speech recognition, syllable segmentation, short term energy, boundary detection, iterative splitting, iterative assimilation, parameter optimization, genetic algorithm*

Abstract

Segmentation is a sub system of *Speech Recognition*. The process of sound segmentation in Bahasa Indonesia is more suitable to be made into syllables as its basic units. Previous research has applied segmentation in Bahasa Indonesia and added splitting and assimilation process to improve accuracy and decrease insertion error and deletion error but the parameters used for splitting and assimilation process using manual observation so that the result can not be optimal. This paper discuss the optimization of the parameters of boundary detection, splitting and assimilation in *Automatic Indonesian Speech Syllable Segmentation* and introduced two new proposed post-processing method which are the iterative splitting and iterative assimilation. The optimization carried out using genetic algorithm, and each 3 methods are optimized using each own separate genetic algorithm. The optimized parameters caused the iterative assimilation to successfully reduced the insertion and makes the expected segments are not over merged. The iterative splitting performed better than splitting, giving 1.9% better accuracy and 2.15% less deletion error, while the iterative assimilation gives less 4.29% insertion than the assimilation. The sequentially combined iterative splitting gives 88.57% accuracy, and 14.55% deletion error in this 110 Indonesian speech dataset.

Keywords: *speech recognition, syllable segmentation, short term energy, boundary detection, iterative splitting, iterative assimilation, parameter optimization, genetic algorithm*

1. Pendahuluan

Pada bagian pendahuluan memuat beberapa sub bab. Sub bab tersebut diantaranya latar belakang, perumusan dan batasan masalah, tujuan, dan organisasi penulisan. Latar belakang mengemukakan masalah dan rencana penelitian, perumusan dan batasan masalah mengemukakan permasalahan dan batasan pada penelitian, sub bab tujuan mengemukakan tujuan yang dicapai pada penelitian, dan organisasi penulisan mengemukakan struktur penulisan pada laporan ini.

Latar Belakang

Segmentasi merupakan sub sistem dari *Speech Recognition*. Segmentasi suara pada awalnya dilakukan secara manual yaitu dengan mengobservasi gelombang pada sinyal suara menggunakan spektogram namun membutuhkan waktu yang lama dan keakuratannya sangat rentan. Seiring dengan kemajuan teknologi, segmentasi suara dapat dilakukan secara otomatis (*Automatic Speech Segmentation*). Suara dapat disegmentasikan ke dalam satuan dasar, antara lain kata, silabel (suku kata) dan fonem (satuan bunyi terkecil) [1].