

INTISARI

KLASIFIKASI LOGAT BALI, JAWA, DAN MADURA BERDASAR UCAPAN BAHASA INDONESIA MENGUNAKAN MFCC+DWT DAN LPC+DWT

Oleh

Bagas Eko Prasetyo

17/409368/PA/17675

Bahasa memiliki karakteristik yang berbeda tergantung dari asal daerah Bahasa tersebut. Karakteristik tersebut biasanya dinamakan dengan logat. Karakteristik logat setiap daerah dapat terlihat ketika penutur menggunakan Bahasa Indonesia. Karakteristik logat setiap daerah dapat menunjukkan asal daerah dari penuturnya. Penelitian ini mengembangkan sistem pengenalan logat daerah yang ada di Indonesia berdasar ucapan Bahasa Indonesia. Logat yang digunakan pada penelitian ini adalah Logat Bali, Jawa dan Madura.

Penelitian ini menggunakan beberapa metode yaitu *Discrete Wavelet Transform* (DWT) digabungkan dengan *Mel Frequency Cepstral Coefficient* (MFCC) dengan koefisien 13 dan 20. Kemudian dilakukan analisis menggunakan *Principal Component Analysis* (PCA). Selain itu, *Discrete Wavelet Transform* digabungkan dengan *Linear Predictive Coding* (LPC) dengan jumlah koefisien 13. Hasil dari ekstraksi ciri lalu diklasifikasi dengan menggunakan Jaringan Syaraf Tiruan.

Hasil dari pengembangan sistem pengenalan logat Bali, Jawa dan Madura memberikan hasil tertinggi sebesar akurasi 95,4 %, presisi 95,87 %, *True Positive Rate* (TPR) 95,56 % dan *False Discovery Rate* (FDR) 4,13 %. Hasil yang terlalu baik memungkinkan sistem memberikan hasil yang tidak sesuai ketika digunakan data pengujian lain dikarenakan jumlah dataset yang digunakan kurang banyak.

Kata kunci — MFCC, LPC, DWT, logat

ABSTACT

CLASSIFICATION OF BALINESE, JAVANESE, AND MADURA ACCOUNTS BASED ON INDONESIAN SPEECH USING MFCC+DWT AND LPC+DWT

by

Bagas Eko Prasetyo

17/409368/PA/17675

Language has different characteristics depending on the area of origin of the language. These characteristics are usually called accents. The accent characteristics of each region can be seen when speakers use Indonesian. The accent characteristics of each region can indicate the regional origin of the speakers. This research develops a regional accent recognition system in Indonesia based on Indonesian speech. The accents used in this study are Balinese, Javanese, and Madurese accents.

This study used several methods, namely Discrete Wavelet Transform (DWT) combined with Mel Frequency Cepstral Coefficient (MFCC) with coefficients 13 and 20. Then analysis was carried out using Principal Component Analysis (PCA). In addition, the Discrete Wavelet Transform is combined with Linear Predictive Coding (LPC) with a coefficient of 13. The results of feature extraction are then classified using an Artificial Neural Network.

The results of the development of the Balinese, Javanese, and Madurese accent recognition systems gave the highest accuracy of 95.4%, 95.87% precision, 95.56% True Positive Rate (TPR), and 4.13% False Discovery Rate (FDR). Results that are too good allow the system to provide inappropriate results when other test data are used because the number of datasets used is not large.

Keywords — MFCC, LPC, DWT, accent