

INTISARI

KLASIFIKASI UMUR MANUSIA BERBASIS SINYAL SUARA DENGAN FITUR MFCC *VOICED* DAN *UNVOICED SPEECH*

Oleh

Ihsanul Hajid

15/381059/PA/16739

Sistem klasifikasi dapat diterapkan pada berbagai bidang, baik keamanan, pendidikan, maupun hiburan. Ada berbagai macam cara klasifikasi, salah satunya melalui suara. Salah satu klasifikasi berbasis sinyal suara adalah klasifikasi umur manusia. Pada penelitian ini digunakan metode *K-Nearest Neighbor* (KNN) dengan tiga buah ekstraksi ciri untuk membedakan tiga kelas suara manusia berdasarkan umur, yaitu *Mel Frequency Cepstral Coefficient* (MFCC), *Zero Crossing Rate* (ZCR), dan frekuensi dasar. Pada frekuensi dasar dan ZCR data sinyal suara dikelompokkan ke dalam *voiced speech* dan *unvoiced speech*. Validasi data penelitian ini menggunakan *10-fold cross validation* untuk mengetahui kinerja dari sistem. Sistem klasifikasi umur manusia dengan menggunakan ekstraksi ciri MFCC dihasilkan rata-rata akurasi tertinggi sebesar 88.65% tanpa pemisahan *voiced/unvoiced*, 62.63% dengan *unvoiced speech*, dan 48.95% dengan *voiced speech*. Klasifikasi menggunakan ekstraksi ciri frekuensi dasar dihasilkan rata-rata akurasi tertinggi sebesar 64.64% tanpa pemisahan *voiced/unvoiced*, 63.50% dengan *voiced speech*, dan 52.83% dengan *unvoiced speech*. Ketika ciri digabung didapat rata-rata akurasi tertinggi sebesar 85.70%. Pengujian *online* dengan tiga kali percobaan dihasilkan akurasi sebesar 61.53%.

Kata kunci : Klasifikasi Umur, *Mel Frequency Cepstral Coefficient*, ZCR, Frekuensi Dasar, *K-Nearest Neighbor*.

ABSTRACT

SPEECH-BASED HUMAN AGE CLASSIFICATION WITH MFCC VOICED AND UNVOICED SPEECH

By

Ihsanul Hajid

15/381059/PA/16739

Classification system can be applied to many fields such as security, education, even entertainment. There are various methods to classify, one of it is through speech. One of classification based on speech signals is classification of human age. In this study, the K-Nearest Neighbor (K-NN) method is used with three feature extractions to distinguish three classes of human voices based on ages, which are Mel Frequency Cepstral Coefficient (MFCC), Zero Crossing Rate (ZCR), and fundamental frequency. For the fundamental frequency and ZCR the voice signal data is grouped into voiced speech and unvoiced speech. Ten-fold cross validation is applied in our experiment. The human age classification system using MFCC feature extraction produced the highest average accuracy of 88.65% without voiced / unvoiced separation, 62.63% with unvoiced speech, and 48.95% with voiced speech. Classification using the extraction of basic frequency characteristics produced an average accuracy of 64.64% without the separation of voiced / unvoiced, 63.50% with voiced speech, and 52.83% with unvoiced speech. When the characteristics are combined, the highest average accuracy is 85.70%. However, the accuracy was dropped to 61.53% when the system was tested in the real environment.

Keywords: Age Classification, Mel Frequency Cepstral Coefficient, ZCR, Fundamental Frequency, K-Nearest Neighbor.