

2.3 Teori Pendukung

Berikut adalah teori yang melandasi pelaksanaan tugas akhir yang berkaitan dengan judul

2.3.1 MFCC

MFCC merupakan salah satu jenis metode ekstraksi ciri yang digunakan sebagai vektor ciri yang baik untuk merepresentasikan suara manusia dan sinyal musik. Analisis suara pada Mel-Frequency didasarkan pada persepsi pendengaran manusia, karena telinga manusia telah diamati dapat berfungsi sebagai filter pada frekuensi tertentu. MFCC digambarkan dalam skala mel-frequency yang merupakan frekuensi linier dibawah 1000Hz dan logaritmik diatas 1000Hz. Adapun beberapa tahap MFCC yakni :

2.3.1.1 Frame Blocking

Sinyal suara manusia merupakan sinyal yang tidak stabil. Namun dapat diasumsikan sebagai sinyal yang stabil pada skala waktu 10-30ms. Oleh karena itu dibutuhkan frame blocking untuk memotong sinyal menjadi bagian yang lebih kecil sehingga didapat karakteristik suara yang stabil.

2.3.1.2 Windowing

Proses windowing dilakukan untuk mengurangi kebocoran spektral atau aliasing yang merupakan efek dari timbulnya sinyal baru yang memiliki frekuensi yang berbeda dari frekuensi aslinya. Hal tersebut dapat terjadi karena rendahnya sampling rate atau karena proses framing yang menyebabkan sinyal menjadi diskontinu.

2.3.1.3 FFT

Untuk mendapatkan sinyal dalam domain frekuensi salah satu metode yang dapat digunakan adalah DFT (Discrete Fourier Transform). DFT dilakukan terhadap semua frame yang telah di-windowing sebelumnya. Untuk mempercepat proses DFT dan mengurangi beban komputasi maka dilakukanlah FFT (Fast Fourier Transform) yaitu metode transformasi Fourier dengan lebih cepat. Berikut persamaan dari Fast Fourier Transform (FFT) :

$$F(w)=\sum_{n=0}^{N-1} Y(n)e^{-j2\frac{\pi}{N}wn}$$

Keterangan :

N = jumlah frame pada suatu data sinyal suara

Y(n) = sinyal input frame ke-n pada tahap ini

W = panjang DFT, dimana $0 < w < N-1$

$F(w)$ = Hasil FFT pada DFT ke- w

2.3.1.4 Mel-Frequency Wrapping

Pada proses ini dilakukan pengubahan skala frekuensi menjadi skala Mel. Skala Mel-Frequency adalah frekuensi linier dibawah 1KHz dan Logaritmik diatas 1KHz.

2.3.1.5 Cepstrum

Selanjutnya dilakukan DCT (Discrete Cosine Transform) untuk memperoleh sinyal dalam domain frekuensi. Hasil keluaran dari proses DCT ini disebut Mel-Frequency Cepstral Coefficient (MFCC). Pada sistem pengenalan suara biasanya hanya diambil 12 atau 13 koefisien pertama dari MFCC untuk mendapatkan hasil ekstraksi yang baik.

2.3.2 Joystick

Sebuah joystick pada dasarnya terdiri dari dua buah potensio, dimana potensio pertama akan berputar saat joystick digerakkan ke kiri dan kanan dan potensio kedua akan berputar saat joystick digerakkan ke atas dan bawah. Keluaran dari potensio menghasilkan tegangan analog yang berubah sebanding dengan gerakan putaran potensio yang merupakan hasil gerakan joystick