

BAB II

TEORI PENUNJANG

Mel-Frequency Cepstral Coefficients

Dalam pengenalan suara, terdapat dua proses yaitu proses ekstraksi dan identifikasi. Hal ini juga terjadi dalam identifikasi nada yang dihasilkan oleh suara alat musik. Beberapa metode yang sering digunakan dalam ekstraksi sinyal suara adalah Mel-Frequency Cepstrum Coefficients [3].

Metode MFCC memiliki beberapa tahapan, diantaranya :

1. *Preprocessing*.

Preprocessing pada MFCC meliputi *framing* dan *windowing*. Framing berfungsi untuk memotong sinyal suara dengan durasi yang panjang menjadi durasi yang lebih pendek, agar didapatkan karakteristik sinyal suara yang lebih stabil. Sedangkan proses *windowing* bertujuan untuk mengurangi terjadinya kebocoran spektral atau aliasing yang merupakan suatu efek dari timbulnya sinyal baru yang memiliki frekuensi yang berbeda dengan sinyal aslinya.

2. DFT

Untuk mendapatkan sinyal dalam domain frekuensi dari sebuah sinyal diskrit, salah satu metode transformasi fourier yang digunakan adalah Discrete Fourier Transform (DFT). DFT dilakukan terhadap masing-masing frame dari sinyal yang telah di-*windowing*. DFT dilakukan setiap 10ms pada sinyal suara tersebut.

3. *Mel-Frequency Wrapping*

Skala Mel-Frequency adalah frekuensi yang linier di bawah 1 kHz dan logaritmik di atas 1 kHz. Skala Mel dapat diperoleh dengan pendekatan persamaan :

$$B(f) = 1125 * \ln(1 + f/700)$$

Di mana B adalah skala Mel-Frequency dan f adalah frekuensi linier.

4. Cepstrum

Mel-Frequency Cepstrum kemudian didapat dari DCT (Discrete Cosine Transform) untuk mendapatkan kembali sinyal dalam domain waktu. Hasilnya disebut sebagai Mel-Frequency Cepstral Coefficient (MFCC). MFCC bisa didapat dari pendekatan persamaan :

$$C_i = \sqrt{\frac{2}{N}} \sum_{j=1}^p m_j \cos\left(\frac{\pi i}{N} (j-0.5)\right)$$

Di mana adalah hasil akumulasi dari kuadratik magnitudo DFT yang dikalikan dengan Mel-Filter Bank. Setelah itu didapatkan MFCC. Pada sistem pengenalan suara, biasanya hanya 13 cepstrum koefisien pertama yang digunakan[4].

Ekstraksi merupakan representasi parametric terbaik dari sinyal akustik untuk menghasilkan kinerja pengenalan yang lebih baik. Mel Frequency cepstral coefficients (MFCC) adalah salah satu ekstraksi fitur yang paling berhasil dalam pengenalan suara, dan koefisien yang diperoleh melalui filter bank analysis. Langkah-langkah yang dilakukan dalam ekstraksi adalah pre-emphasis, frame blocking, windowing, filter bank analysis, logarithmic compression dan discrete cosine transformation [5].

Support Vector Machine (SVM)

SVM pertama kali diperkenalkan oleh Boser, Guyon, Vepnik, yang dipresentasikan untuk pertama kalinya pada tahun 1992 di Annual Workshop on Computational Learning Theory. konsep dari svm merupakan kombinasi harmonis dari konsep komputasi yang sudah ada puluhan tahun sebelumnya, seperti hyperplane (Duda dan Hart, 1973, cover 1965, Vapnik, 1964). Kernel diperkenalkan oleh Aronszajn (1950) dan demikian dengan konsep-konsep lainnya[7].

Prinsipnya SVM bekerja secara linear, dan dikembangkan untuk dapat diterapkan pada masalah non-linear. Dengan menggunakan metode kernel trick yang mencari hyperplane dengan cara mentransformasi dataset ke ruang vektor yang berdimensi lebih tinggi (feature space), kemudian proses klasifikasi dilakukan pada feature space tersebut. Penentuan fungsi kernel yang digunakan akan sangat berpengaruh terhadap hasil prediksi[8].

Tensor Flow

Tensor Flow Merupakan sebuah framework komputasional yang biasa digunakan sebagai *machine learning* dalam sebuah penelitian ataupun percobaan ilmiah. Tensor Flow menyediakan berbagai macam *Tool Kit* yang dapat digunakan untuk membuat model pada tingkat abstraksi yang memungkinkan untuk percobaan dan penelitian. Modul ini dapat digunakan sebagai API dengan tingkat yang rendah dalam membuat model dengan menentukan operasi aritmatika.