BAB II

TINJAUAN PUSTAKA

2.1 Kursi Roda

Kursi roda adalah alat bantu untuk individu dengan keterbatasan mobilitas dalam menjalani aktivitas sehari-hari. Keterbatasan mobilitas ini bisa disebabkan oleh cacat fisik, cedera, atau penyakit yang diderita. Awalnya, kursi roda digerakkan secara manual menggunakan tenaga tangan atau bantuan dari orang lain, namun kini telah berkembang menjadi kursi roda elektrik dengan penambahan motor sebagai penggerak dan joystick sebagai pengendali roda[11].

Penggunaan joystick sebagai pengendali kursi roda elektrik belum sepenuhnya memenuhi kebutuhan penggunanya. Menurut para ahli klinis, setiap tahunnya terdapat 85% pasien yang tidak dapat menggunakan kursi roda elektrik karena keterampilan yang terbatas dalam mengoperasikan kursi roda bermotor, masalah kekuatan fisik dan penglihatan yang kurang tajam, serta pergerakan motor yang kasar dan tidak akurat sehingga tidak mencapai target yang diinginkan[7].

2.2 Pengenalan Suara

Pengenalan suara adalah proses mengidentifikasi suara berdasarkan kata-kata yang diucapkan dengan cara mengubahnya menjadi sinyal akustik yang ditangkap oleh perangkat audio. Teknologi ini juga merupakan sistem yang dirancang untuk mengenali perintah kata dari suara manusia dan menerjemahkannya menjadi data yang dapat dipahami oleh komputer. Teknologi ini memungkinkan perangkat untuk mengenali dan memahami kata-kata yang diucapkan dengan mendigitalkan kata-kata tersebut dan mencocokkan sinyal digitalnya dengan pola tertentu yang tersimpan dalam perangkat[12].

Pengenalan suara adalah teknologi yang membantu manusia dalam melakukan aktivitas sehari-hari dengan lebih efisien. Teknologi ini memungkinkan pengguna untuk mencocokkan suara mereka dengan sampel suara yang telah divalidasi sebelumnya, sehingga memverifikasi kecocokan suara pengguna untuk tujuan identifikasi[13].

2.3 Frekuensi Suara Manusia

Efektivitas sistem pengenalan ucapan diukur berdasarkan kejelasan ucapan, yaitu seberapa baik ucapan dapat dimengerti oleh pendengar. Kejelasan ucapan bergantung pada energi ucapan dalam rentang frekuensi 1000–6000 Hz, meskipun rentang daya ucapan menyebar dari 80 hingga 8000 Hz, dengan konsentrasi utama pada 200–600 Hz. Energi di atas 1000 Hz, yang merupakan 5% dari total daya ucapan, menyumbang sekitar 60% terhadap kejelasan, sementara energi di bawah 250 Hz hanya berkontribusi sekitar 2%. Konsonan yang memiliki energi frekuensi tinggi (di atas 1500 Hz) sering kali sulit didengar, sementara vokal yang memiliki energi lebih besar pada frekuensi rendah hingga menengah biasanya lebih mudah dikenali[14].

Untuk mencapai kejelasan ucapan yang baik, diperlukan rentang frekuensi yang lebih luas, yaitu antara 200 hingga 7000 Hz. Faktor teknis seperti lebar pita transmisi, jenis dan tingkat kebisingan latar belakang, serta osilasi dan gema juga memengaruhi kejelasasn ucapan. Dalam kondisi ideal, rentang frekuensi hingga 7000 Hz diperlukan untuk mencapai kejelasan ucapan lebih dari 95%, sedangkan rentang yang lebih sempit, seperti 300 hingga 3400 Hz pada telepon, dapat menyebabkan kejelasan yang lebih rendah[14].

2.4 *Mel-frequency Cepstral Coefficients (MFCC)*

Untuk memodelkan persepsi pendengaran manusia dibutuhkan Metoda salah satunya *MFCC*[15]. *MFCC* adalah Metoda pengambilan fitur yang menggunakan transformasi Fourier. *MFCC* merupakan Metoda yang paling sering digunakan dalam pemrosesan ucapan karena dianggap sangat efektif dalam menangkap fitur dan sinyal. Cara kerja *MFCC* dalam menangkap suara dan sinyal frekuensi mirip dengan cara manusia mengenali karakteristik suara[16]. Berikut proses untuk *MFCC*:

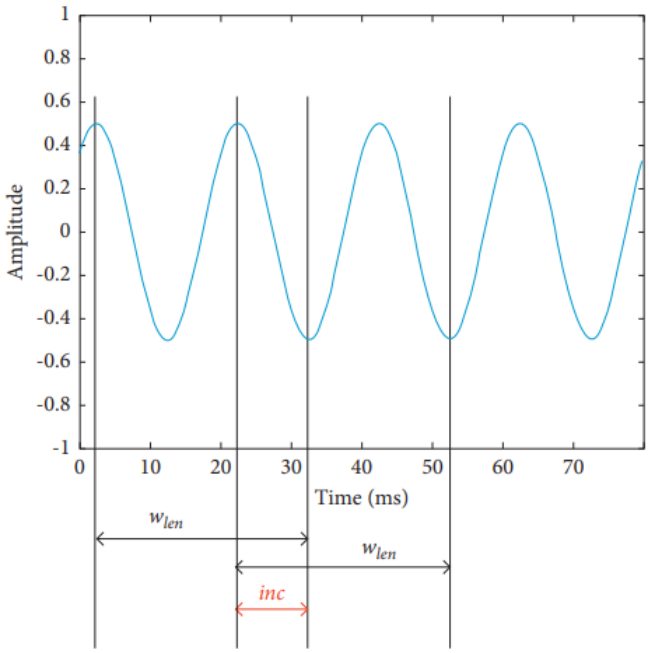
1. *Pre-emphasis*

   Komponen frekuensi dari sinyal suara terletak pada frekuensi tinggi, sehingga *Pre-emphasis* diterapkan untuk meratakan spektrum magnitudo dan menyeimbangkan komponen frekuensi tinggi dan rendah dari sinyal suara. Selain itu, *Pre-emphasis* digunakan untuk mengurangi rentang dinamis yang tinggi dari bentuk gelombang suara dan meningkatkan rasio sinyal terhadap noise[17]. Untuk menyelesaikan hal tersebut, sinyal ucapan dilewatkan melalui filter *Finite Impuls Response (FIR) High-pass*. *Pre-emphasis* memastikan bahwa dalam domain frekuensi, semua format sinyal ucapan memiliki amplitudo yang serupa sehingga mereka mendapatkan perhatian yang sama pentingnya dalam tahap pemrosesan selanjutnya[18].

2. *Framing*

   Sinyal suara adalah sinyal yang tidak stabil, sehingga tidak memungkinkan untuk melakukan ekstraksi ciri secara langsung[19]. ucapan dapat dianggap sebagai sinyal kuarasi-stasioner yang hanya stabil

untuk periode singkat, yang berarti ketika sinyal dianalisis dalam periode waktu singkat (5-100 ms), sinyal tersebut relatif stabil. Ini berarti bahwa sinyal dibagi atau diblok dalam frame-frame dengan durasi biasanya 20-30 ms. Dalam hal ini, frame-frame yang berdekatan umumnya tumpang tindih satu sama lain sebesar 30-50%, hal ini dilakukan agar tidak kehilangan informasi penting dari sinyal ucapan saat penggunaan *Windowing*[18].

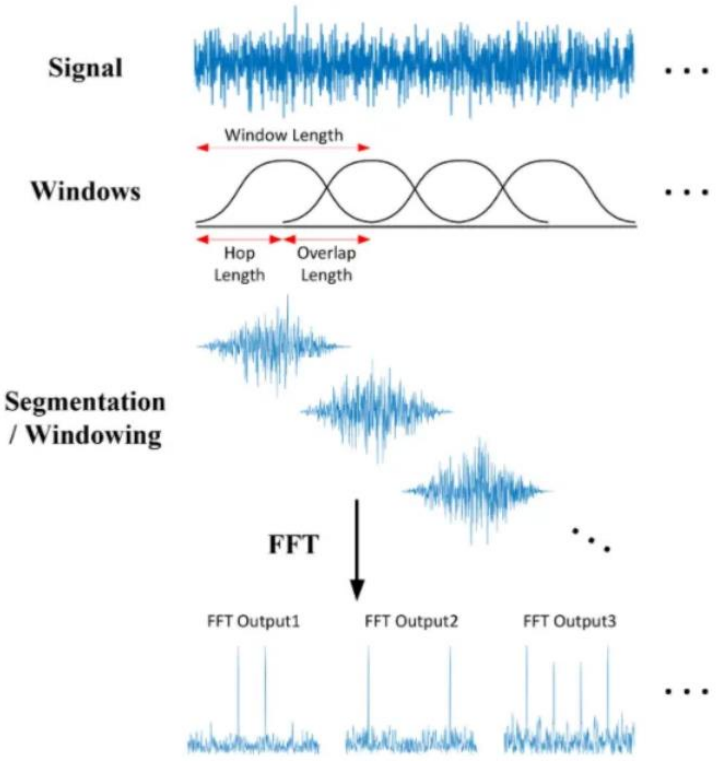


Gambar 2. 1 *Framing* pada sinyal[19]

3. *Windowing*

Penerapan *Windowing* bertujuan untuk mengurangi gangguan pada sinyal yang bersifat tidak kontinu di bagian awal dan akhir setiap frame[20]. Konsep *Windowing* secara utama digunakan untuk membantu menyempurnakan sinyal dan menghindari ketidakkontinuan sinyal yang dihasilkan oleh distorsi spektral selama tahap *Framing*, dengan mengurangi sinyal di awal dan akhir setiap frame menjadi nol. Namun,

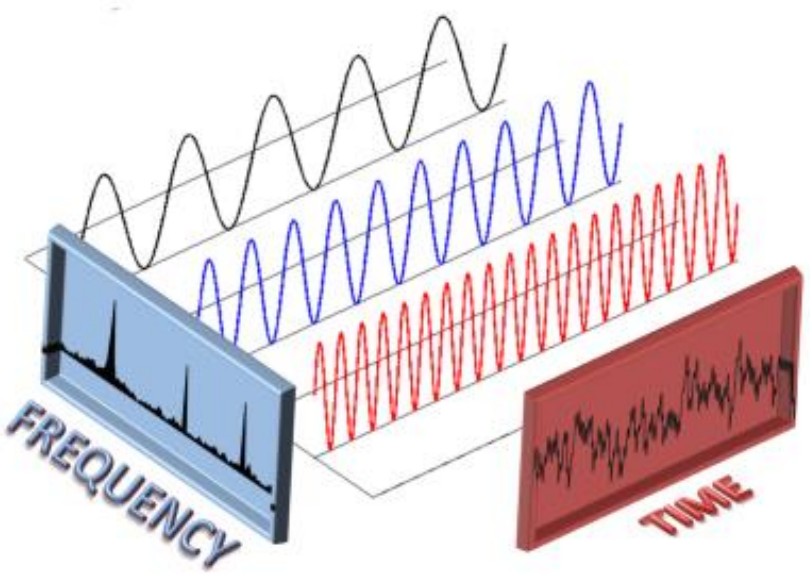
penggunaan teknik jendela waktu kadang-kadang memiliki dampak pada perubahan sinyal di awal dan akhir frame. Mengalikan gelombang ucapan dengan fungsi jendela memiliki dua dampak utama. Pertama, untuk menghindari perubahan tajam di ujung-ujung frame, fungsi ini secara bertahap mengurangi amplitudo di kedua ujung interval waktu tersebut. Dampak kedua adalah mengkonvolusi spektrum ucapan dengan transformasi *Fourier* dari fungsi jendela. Oleh karena itu, pemilihan fungsi jendela yang akan diterapkan untuk mengurangi distorsi spektral harus memenuhi dua kondisi: resolusi frekuensi tinggi dan kebocoran spektral rendah. Dalam analisis *Automatic Speech Recognition* (*ASR*), disarankan untuk menggunakan jendela pendek dengan panjang frame antara 20 hingga 30 ms, dengan tumpang tindih antara 5 hingga 10 ms[17].



Gambar 2. 2 Penerapan *Windowing* pada sinyal[21]

4. *Fast Fourier Transform* (*FFT*)

Fast Fourier Transform (*FFT*) adalah algoritma yang digunakan untuk menghitung *Discrete Fourier Transform (DFT)* secara efisien. *FFT* memungkinkan analisis sinyal dalam domain waktu dan frekuensi dengan cara yang cepat dan mengurangi waktu komputasi secara signifikan. Algoritma ini mengambil sampel data diskrit dari sinyal waktu dan mengubahnya menjadi representasi frekuensi, yang berguna dalam berbagai aplikasi seperti analisis spektral digital, simulasi filter, dan pengenalan pola. *FFT* secara luas digunakan dalam banyak bidang teknik karena kemampuannya untuk memproses data dengan cepat dan efisien[22].

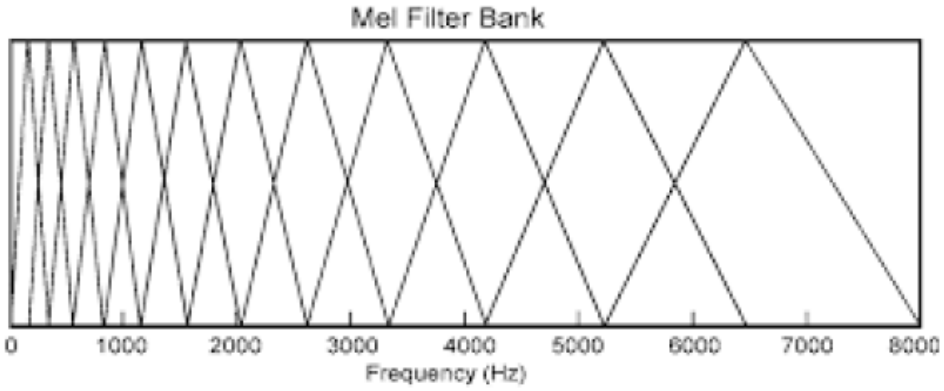


Gambar 2. 3 *Fourier Transform*[23]

5. *Mel Frequency Wrapping*

Telinga manusia mendengar frekuensi secara nonlinier. Frekuensi diukur secara linier hingga 1 kHz. Kerangka kerja terkait suara audio

diskalakan oleh frekuensi yang disebut frekuensi skala mel. Sinyal suara untuk setiap frame diperbolehkan melalui filter pita frekuensi skala mel untuk mencerminkan persepsi pendengaran audio. Oleh karena itu, untuk setiap nada dengan frekuensi nyata f, yang diukur dalam Hz, nada subyektif diukur pada skala yang disebut skala mel. Skala frekuensi mel adalah pembagian frekuensi linier di bawah 1000 Hz dan pembagian logaritmik di atas 1000 Hz. Frekuensi mel berbanding lurus dengan logaritma frekuensi linier, meniru efek serupa dalam tingkat persepsi pendengaran manusia. Contoh di bawah ini menunjukkan hubungan antara frekuensi mel dan frekuensi linier[24].



Gambar 2. 4 *Mel frequency triangular filter*[25]

Alasan penggunaan filter *bandpass* segitiga adalah untuk menghaluskan spektrum magnitudo dari sinyal suara dan mengompres koefisien cepstral dari sinyal tersebut. Dengan demikian, harmonik yang terkompresi diperoleh[24].

6. *Discrete Consine Transform* (*DCT*)

Pada langkah ini, spektrum mel log dikonversi kembali ke domain waktu. Representasi cepstral dari spektrum suara memberikan

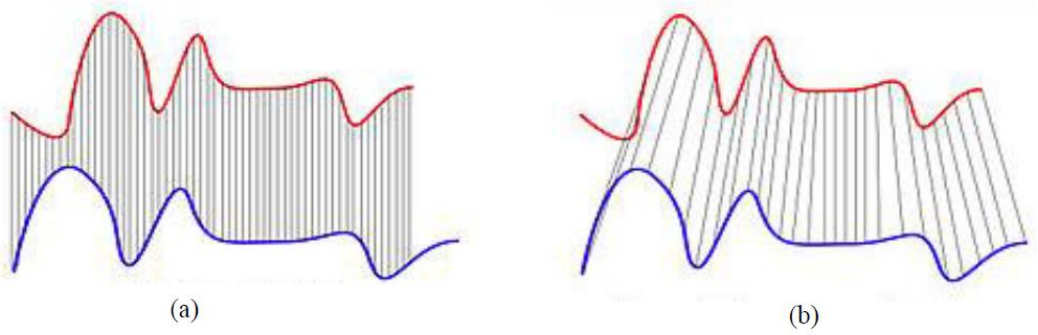
representasi yang baik terhadap sifat spektral lokal dari sinyal untuk analisis frame yang diberikan. Karena koefisien spektrum mel (dan logaritmanya) adalah bilangan real, kita dapat mengubahnya ke domain waktu menggunakan *Discrete Consine Transform* (*DCT*). Pada langkah akhir ini, spektrum mel log dikonversi kembali ke domain waktu. Hasilnya disebut *Mel-Frequency Cepstral Coefficients* (*MFCC*). Kumpulan koefisien ini disebut vektor akustik. Oleh karena itu, setiap ucapan yang masuk diubah menjadi urutan vektor akustik[25].

2.5 *Dynamic Time Warping (DTW)*

DTW adalah algoritma yang menentukan jalur terbaik antara dua data [12]. Algoritma *DTW* menggunakan teknik pemrograman dinamis untuk mengukur kesamaan antara dua deret waktu yang mungkin memiliki variasi dalam durasi atau kecepatannya. Algoritma ini dapat diterapkan untuk menemukan bagian yang cocok antara dua deret waktu atau untuk menilai kesamaan di antara keduanya[26].

Pengguna yang sama bisa mengucapkan dua kata yang sama pada waktu yang berbeda. Sebagai contoh, kata "dua" bisa diucapkan sebagai "dua" atau "duuaa". *DTW* menyelesaikan masalah ini dengan menyelaraskan kata-kata tersebut secara tepat dan menghitung jarak minimum antara kedua kata itu. Perbedaan dalam penyelarasan waktu ucapan adalah masalah utama dalam pengukuran jarak pada pengenalan ucapan. Pergeseran kecil dapat menyebabkan identifikasi yang salah. *DTW* adalah Metoda yang efektif untuk mengatasi masalah penyelarasan waktu ini, sehingga lebih realistis untuk digunakan dalam mengukur kesamaan pola (pattern/template matching). Data yang diproses selalu berada dalam zona waktu,

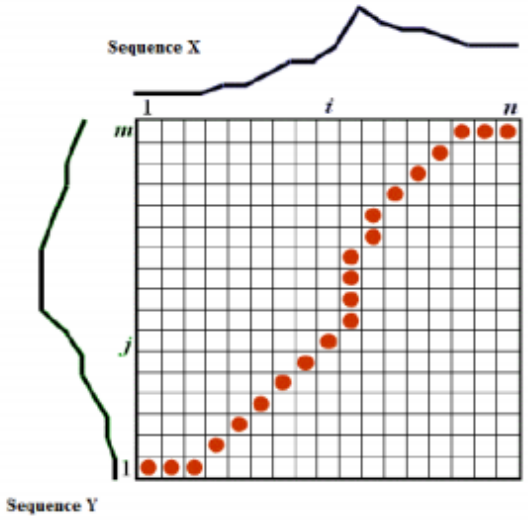
sehingga urutan data yang kita miliki dianggap berbeda-beda terhadap waktu[12]. Ilustrasi pencocokan menggunakan Metoda *DTW* dapat dilihat pada **Gambar 2.5**.



Gambar 2. 5 Perbandingan Metoda *Euclidean Distance* (a) dan *DTW* (b)[12]

Algoritma *DTW* bertujuan untuk menyelaraskan dua deretan vektor dengan mengubah sumbu waktu berulang kali sampai tercapai kesesuaian optimal antara kedua deretan tersebut. Algoritma ini berfungsi sebagai pemetaan linier sumbu untuk menyinkronkan kedua sinyal[12]. Misalkan terdapat dua deretan vektor dalam ruang berdimensi n:

$x = [x1\ x2\ \dots\ xn]$ dan $y = [y1\ y2\ \dots\ yn]$



Gambar 2. 6 Diagram sebar dari urutan Pencocokan[12]

Di setiap sel, jarak diukur dengan membandingkan elemen yang sesuai dari dua urutan. Jarak antara dua titik dihitung menggunakan jarak *Euclidean*[12].

$$Dist\,(x,y)=|\,x-y\,|=\sqrt{[\,(\mathrm{x1}-\mathrm{y1})^2+(\mathrm{x2}-\mathrm{y2})^2+\cdots+\,(\mathrm{xn}-\mathrm{yn})^2]} \qquad (2.1)$$

Jarak antara dua *DTW* dihitung dari jalur pembengkokkan optimal (optimal warping path)[21]. Jarak *DTW* dihitung dengan rumus:

$$D(U,V) = \mathrm{y}(m,n) \qquad (2.2)$$

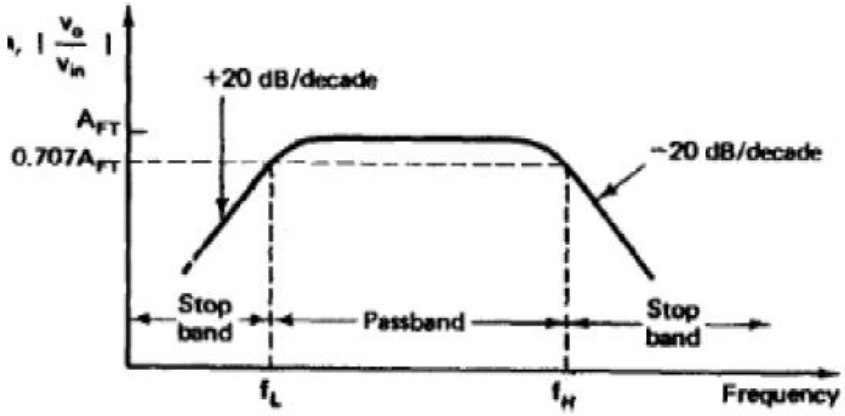
$$\mathrm{y}(m,n) = dbase(u_{\mathrm{i}},v_{\mathrm{j}})+ \min \begin{cases} \gamma(\mathrm{i}-1,\mathrm{j}) \\ \gamma(\mathrm{i}-1,\mathrm{j}-1) \\ \gamma(\mathrm{i},\mathrm{j}-1) \end{cases} \qquad (2.3)$$

2.6 Bandpass Filter

Band Pass Filter (BPF) adalah jenis rangkaian yang berfungsi untuk menyaring sinyal-sinyal berdasarkan frekuensi. *BPF* memungkinkan sinyal dengan frekuensi dalam rentang tertentu, yang berada antara dua titik frekuensi cut-off, untuk melewatinya, sementara meredam atau menolak sinyal dengan frekuensi yang berada di luar rentang tersebut. Fungsi utama *BPF* adalah untuk menjaga agar hanya sinyal dalam rentang frekuensi yang diinginkan yang diteruskan, sehingga sinyal yang berada di bawah atau di atas batas frekuensi ini akan diblokir atau dilemahkan. Dengan cara ini, *BPF* membantu dalam menghilangkan noise atau sinyal yang tidak relevan yang dapat mengganggu sinyal utama yang diinginkan[27].

BPF pada dasarnya merupakan kombinasi dari dua filter, yaitu *highpass filter* dan *lowpass filter*. *Highpass filter* berfungsi untuk memblokir frekuensi rendah dan hanya meneruskan frekuensi yang lebih tinggi dari titik cut-off tertentu, sedangkan *lowpass filter* berfungsi sebaliknya, yakni memblokir frekuensi tinggi dan meneruskan frekuensi yang lebih rendah dari titik cut-off yang lain. Ketika kedua filter ini digabungkan, mereka menciptakan *BPF* yang hanya meneruskan frekuensi yang berada di antara dua titik cut-off tersebut, sering disebut sebagai frekuensi median. Oleh karena itu, *BPF* sangat berguna dalam berbagai aplikasi, seperti komunikasi dan pengolahan sinyal, di mana diperlukan untuk memfilter sinyal agar sesuai dengan rentang frekuensi tertentu yang relevan untuk pengoperasian sistem yang diinginkan[27].

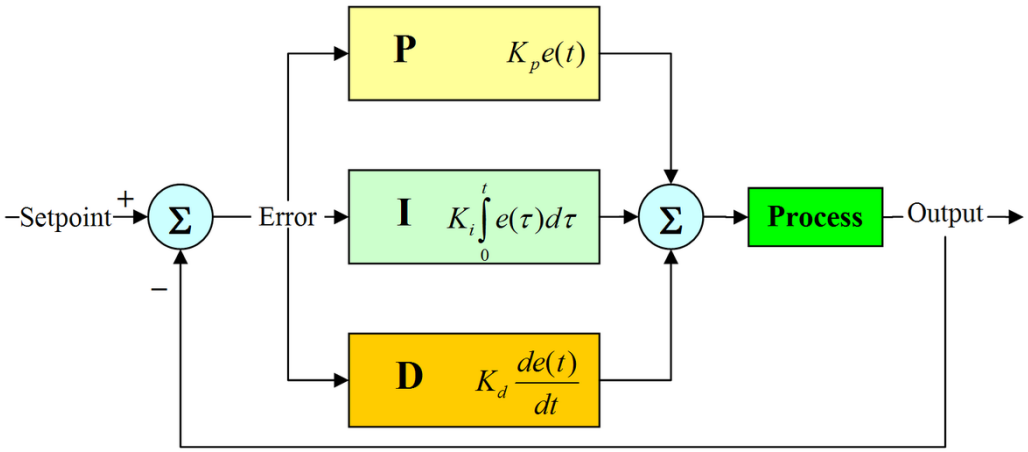
Penggunaan *BPF* dengan rentang frekuensi 300-3400 Hz sangat sesuai untuk pengenalan suara manusia. Dengan membatasi frekuensi pada rentang 300-3400 Hz, *BPF* secara efektif menyaring sinyal yang relevan, mengurangi noise yang tidak diinginkan, dan mempertahankan komponen frekuensi yang paling penting untuk pengenalan suara, sehingga meningkatkan akurasi dan efisiensi proses pengenalan suara[14].



Gambar 2. 7 *Bandpass Filter*[27]

2.7 *Proportional Integral Derivative (PID)*

Sistem Kontrol *PID* adalah pengendali yang digunakan untuk menentukan ketepatan suatu sistem instrumentasi dengan fitur umpan balik pada sistem tersebut (Feedback)[28]. Kontrol *PID* adalah Metoda kontrol yang umum diterapkan pada motor DC. Metoda ini dikenal karena kemampuannya yang akurat dalam mengatur kecepatan motor DC, sehingga motor dapat berputar dengan stabil sambil menjaga kecepatan yang diinginkan. Penyesuaian konstanta parameter *PID* dapat mempengaruhi kinerja motor DC dan hasil yang dicapainya[29]. Terdapat tiga parameter pada *PID* yang ditampilkan pada **Gambar 2.7**.



Gambar 2. 8 Blok diagram *PID*[30]

1. *Proporsional*

Proporsional adalah hasil kali antara konstanta proporsional dengan nilai kesalahannya. Perubahan pada nilai input akan menghasilkan sinyal output yang besarnya sebanding dengan konstanta pengalinya[31]. Pada kontrol *proporsional*, output u(t) dirumuskan sebagai[30]:

$$u(t) = k_p e(t) \tag{2.4}$$

Dimana:

U(t) = sinyal output yang dikirim ke aktuator

Kp = konstanta *proporsional*

e(t) = eror pada waktu t

2. *Integral*

Kontrol *integral* memiliki karakteristik untuk mengurangi waktu naik, menambah *overshoot* dan waktu turun, serta menghilangkan *steady state eror*. Kontrol integral membantu meningkatkan respons sehingga menghasilkan output yang diinginkan[31].]. Pada kontrol *integral*, output u(t) dirumuskan sebagai[30]

$$u(t) = k_i \int_0^t e(t)dt \tag{2.5}$$

Dimana:

U(t) = sinyal output yang dikirim ke aktuator

Ki = konstanta *integral*

e(t) = eror pada waktu t

$\int_0^t e(t)dt$ = *integral* dari sinyal kesalahan selama waktu 0 hingga t

3. *Derivatif*

Kontrol *derivatif* tidak bisa digunakan secara mandiri, karena kontroler ini hanya aktif selama periode transisi. Selama periode transisi, kontrol derivatif membantu meredam sistem, sehingga mengurangi lonjakan. Seperti kontrol *proporsional*, kontrol *derivatif* juga tidak bisa

menghilangkan *offset*[31]. Pada kontrol *derivatif*, output u(t) dirumuskan sebagai[30]

$$u(t) = k_d \frac{de(t)}{dt} \tag{2.6}$$

Dimana:

U(t) = sinyal output yang dikirim ke aktuator

Ki = konstanta *derivatif*

e(t) = eror pada waktu t

$\frac{de(t)}{dt}$ = turunan pertama dari sinyal kesalahan terhadap waktu

2.8 Arduino UNO

Arduino Uno adalah papan mikrokontroler berbasis ATmega328 yang memiliki 14 pin input/output digital (6 di antaranya dapat digunakan sebagai output PWM), 6 input analog, resonator keramik 16 MHz, konektivitas USB, jack daya, header ICSP, dan tombol reset. Desainnya mendukung mikrokontroler agar dapat berfungsi dengan baik, cukup dengan menghubungkannya ke komputer menggunakan kabel USB atau memberinya daya melalui adaptor AC ke DC atau baterai. Arduino Uno tersedia dalam dua versi, yaitu Arduino Uno dan Genuino Uno, dengan perbedaan tergantung pada wilayah. Arduino Uno berbeda dari papan sebelumnya karena tidak menggunakan chip driver USB-ke-serial FTDI. Nama "Uno" yang berarti "satu" dalam bahasa Italia diberikan untuk menandai rilis Arduino 1.0 yang akan datang. Arduino Uno adalah model referensi terbaru untuk platform Arduino[32].

2.9 Sensor Ultrasonik HC-SR04

Modul sensor ultrasonik HC-SR04 memiliki kemampuan pengukuran jarak non-kontak antara 2 cm hingga 400 cm dengan tingkat akurasi mencapai 3 mm. Modul ini terdiri dari pemancar dan penerima ultrasonik, serta rangkaian kontrol. Cara kerjanya adalah: pertama, memberikan sinyal pemicu tinggi selama minimal 10μs melalui IO. Kedua, modul akan mengirim delapan gelombang ultrasonik dengan frekuensi 40 kHz dan mendeteksi keberadaan sinyal pantul. Ketiga,jika sinyal pantul terdeteksi, durasi sinyal tinggi pada output IO menunjukkan waktu yang dibutuhkan untuk pengiriman dan penerimaan sinyal ultrasonik. Jarak diukur dengan rumus: (waktu sinyal tinggi × kecepatan suara (340 m/s) / 2). Energi listrik diubah menjadi gelombang suara untuk mengirim pulsa, dan gelombang suara yang kembali diubah menjadi sinyal listrik, sehingga perbedaan waktu antara pengiriman dan penerimaan digunakan untuk menghitung jarak ke objek. Sensor harus dipasang dengan jarak yang cukup untuk menghindari gangguan antar sensor atau "*crosstalk*", dengan target dipasang secara tegak lurus terhadap sumbu sensor[32].

2.10 Motor Driver BTS7960

BTS7960 berasal dari modul H-Bridge IBT-2, yang memiliki tegangan kerja maksimum 27 V dan arus maksimum 43 A. Modul ini memiliki port kontrol dan port daya. Port kontrol mencakup pin daya, pin enable, dan pin PWM. Pin daya dihubungkan ke 5 V dan ground mikrokontroler, sedangkan pin enable harus selalu berada dalam kondisi tinggi. Pin PWM terhubung ke output PWM Arduino. Port daya terdiri dari pin input dan output, dengan pin output M+ dan M- yang umumnya digunakan untuk menggerakkan motor DC. Tegangan pada pin output ini bisa dikendalikan secara mandiri berdasarkan sinyal PWM, misalnya, M+ diatur ke 5 V

dan M- ke 3 V, sehingga memungkinkan untuk mengendalikan sinyal tiga fasa untuk mengoperasikan motor BLDC[33].

2.11 Speed Sensor FC-03 (Photo Interrupter)

Photo-interrupter, atau dikenal juga sebagai opto-interrupter, adalah sensor foto tipe transmisi yang menyatukan komponen penerima dan pemancar optik dalam satu paket perangkat. Sensor ini terdiri dari sebuah rumah plastik yang berisi LED inframerah (IR) yang menghadap ke fototransistor dalam sebuah celah. Setiap objek yang berada di dalam celah tersebut akan mengganggu pancaran sinar IR, sehingga menyebabkan fototransistor menyala atau mati. Alat ini memiliki kecepatan yang tinggi dan cocok digunakan untuk menghitung, mengukur waktu, atau mendeteksi, serta banyak diaplikasikan dalam sakelar optik, mesin ATM, mesin penjual otomatis, deteksi tepi dan posisi, serta peralatan otomatisasi kantor. Prinsip kerja photo-interrupter mirip dengan optocoupler, namun lebih ditujukan untuk mendeteksi adanya hambatan antara pemancar dan penerima. Jika tidak ada hambatan, cahaya inframerah dari LED IR (pemancar) dapat ditangkap oleh fototransistor (penerima), sehingga output-nya adalah logika HIGH atau "1". Sebaliknya, jika terdapat hambatan, cahaya inframerah tidak dapat diterima oleh fototransistor, sehingga output-nya menjadi logika LOW atau "0"[34].