

SKRIPSI

**RANCANG BANGUN MOBILE ROBOT DISINFEKTAN
MENGUNAKAN UVC (*ULTRAVIOLET C*) DENGAN FITUR *SPEECH
RECOGNITION CONTROLLER* BERBASIS IOT (*INTERNET OF THINGS*)**

Disusun dan Diajukan oleh :

JOVAN LOARDI

D021 171 328



DEPARTEMEN TEKNIK MESIN

FAKULTAS TEKNIK

UNIVERSITAS HASANUDDIN

GOWA

2022

SKRIPSI

**RANCANG BANGUN MOBILE ROBOT DISINFEKTAN
MENGUNAKAN UVC (*ULTRAVIOLET C*) DENGAN FITUR *SPEECH
RECOGNITION CONTROLLER* BERBASIS IOT (*INTERNET OF THINGS*)**

Disusun Dan Diajukan Oleh:

JOVAN LOARDI

D021171328

**Merupakan salah satu syarat untuk memperoleh gelar Sarjana Teknik
Mesin pada Fakultas Teknik Universitas Hasanuddin**

DEPARTEMEN TEKNIK MESIN

FAKULTAS TEKNIK

UNIVERSITAS HASANUDDIN

GOWA

2022

LEMBAR PENGESAHAN

Tugas akhir ini diajukan untuk memenuhi salah satu persyaratan Mengikuti Ujian Akhir guna memperoleh gelar Sarjana Teknik Mesin pada Departemen Mesin Fakultas Teknik Universitas Hasanuddin

JUDUL:

**RANCANG BANGUN MOBILE ROBOT DISINFEKTAN
MENGUNAKAN UVC (ULTRAVIOLET C) DENGAN FITUR SPEECH
RECOGNITION CONTROLLER BERBASIS IOT (INTERNET OF THINGS)**

JOVAN LOARDI

D021171328

Telah diperiksa dan disetujui oleh:

Gowa, 26 Desember 2022

Dosen Pembimbing I

Dosen Pembimbing II



Dr. Eng. Andi Amijoyo Mochtar, ST., MSc

NIP. 19760216 201012 1 002



Dr. Muhammad Syahid, ST., MT

NIP. 19770707 200501 1 001

Mengetahui,

Ketua Departemen Teknik Mesin Fakultas Teknik
Universitas Hasanuddin



Dr. Eng. Jalaluddin, ST., MT

NIP. 19720825 200003 1 001

PERNYATAAN KEASLIAN

Saya yang bertanda tangan dibawah ini:

Nama : Jovan Loardi
NIM : D021171328
Program Studi : Teknik Mesin
Jenjang : S-1
Judul Skripsi : Rancang Bangun Mobile Robot Disinfektan
Menggunakan UVC (*Ultraviolet C*) dengan Fitur
Speech Recognition Controller Berbasis IoT
(*Internet of Things*)

Menyatakan dengan sesungguhnya bahwa Skripsi ini merupakan hasil penelitian, pemikiran, dan pemaparan asli dari saya sendiri. Saya tidak mencantumkan tanpa pengakuan bahan-bahan yang telah dipublikasikan sebelumnya atau ditulis oleh orang lain atau sebagai bahan yang pernah diajukan untuk gelar atau ijazah pada Universitas Hasanuddin atau perguruan tinggi lainnya.

Apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik sesuai dengan peraturan yang berlaku di Universitas Hasanuddin.

Demikian pernyataan ini saya buat

Gowa, 26 Desember 2022

Yang membuat pernyataan



Jovan Loardi

ABSTRAK

Mobile robot disinfektan menggunakan Ultraviolet C (UVC) merupakan alat yang digunakan untuk membunuh mikroorganisme, salah satunya COVID-19. Dengan menggunakan mobile robot dengan sistem kontrol berbasis IoT dan DIGITALOCEAN sebagai server, kita dapat mengurangi dan menghindari bahaya UVC bagi tubuh manusia. Dengan sistem kontrol suara teknologi saat ini menjadi lebih mudah digunakan bersama, sistem kontrol suara adalah salah satu bidang kecerdasan mesin yang berkembang pesat, hal ini ditandai dengan hampir semua perangkat teknologi saat ini yang dilengkapi dengan perintah suara. Hampir semua orang di luar sana menggunakan perangkat Smartphone, dan Android adalah salah satu jenis yang paling banyak digunakan. Dalam penelitian ini, pembuatan aplikasi Android menggunakan MIT App Inventor yang terhubung dengan server online.

Hasil penelitian ini menunjukkan bahwa jarak penerimaan suara yang efektif dari sumber suara ke mikrofon adalah 300 mm. Untuk hasil penelitian, efektivitas penerimaan suara terhadap perbedaan aksent menunjukkan bahwa aksent Jawa gagal 3/35, aksent Papua gagal 7/35, dan aksent Batak gagal 10/35. Untuk hasil penelitian, kecepatan respons robot mobile menggunakan speech recognition menunjukkan rata-rata waktu aksent Batak 0,814 detik, aksent Jawa 0,8 detik, dan aksent Papua 0,642 detik.

Kata kunci: UVC, COVID-19, mobile robot, IoT, speech recognition, MIT App Inventor

ABSTRACT

Mobile robot disinfectant using Ultraviolet C (UVC) is a tool used to kill microorganisms, one of which is COVID-19. By using a mobile robot with a control system based on IoT and DIGITALOCEAN as a server, we can reduce and avoid UVC hazards to the human body. With today's technology voice control systems becoming easier to use together, voice control systems are one of the rapidly growing areas of machine intelligence, it is characterized by almost all today's technological devices equipped with voice commands. Almost everyone out there uses Smartphone devices, and Android is one of the most widely used types. In this study, the creation of an Android application using the MIT App Inventor which is connected to an online server.

The results of this study show that the distance of effective sound reception from the sound source to the microphone is 300 mm. For the results of the study, the effectiveness of sound acceptance of differences in accents showed that the Javanese accent failed 3/35, the Papuan accent failed 7/35, and the Batak accent failed 10/35. For the results of the study, the response speed of mobile robots using speech recognition showed an average Batak accent time of 0.814s, Javanese accent 0.8s, and Papuan accent 0.642s.

Keywords: UVC, COVID-19, mobile robot, IoT, speech recognition, MIT App Inventor

KATA PENGANTAR

Puji dan syukur senantiasa kita panjatkan ke-hadirat Tuhan Yang Maha Esa karena atas rahmat dan berkat-Nya penulis dapat menyelesaikan skripsi dengan judul “Rancang Bangun Mobile Robot Disinfektan Menggunakan UVC (*Ultraviolet C*) dengan Fitur *Speech Recognition Controller* Berbasis IoT (*Internet of Things*)” yang mana merupakan salah satu syarat untuk memperoleh gelar sarjana teknik pada Departemen Teknik Mesin Fakultas Teknik Universitas Hasanuddin. Selama proses pengerjaan skripsi ini penulis menerima begitu banyak bantuan dari berbagai pihak. Untuk itu peneliti ingin mengucapkan terimakasih kepada:

1. Kedua orang tua tercinta Bapak dan Ibu yang selalu mendampingi, memberi semangat dan mendoakan.
2. Dr.Eng. Jalaluddin, ST., MT. sebagai Ketua Departemen Teknik Mesin Fakultas Teknik Universitas Hasanuddin beserta seluruh staf Departemen Teknik Mesin Fakultas Teknik Universitas Hasanuddin atas segala bantuan dan kemudahan yang diberikan
3. Dr.Eng. Andi Amijoyo Mochtar, ST., MSc, dan Ir. Mukhtar Rahman, MT. dan Dr. Muhammad Syahid, ST., MT sebagai Dosen Pembimbing I dan Dosen Pembimbing II yang telah memberikan waktu, arahan, dan saran selama proses pengerjaan skripsi ini.
4. Fauzan, ST., MT., Ph.D dan Dr.Ir. Nasruddin Azis, M.Si selaku penguji yang telah memberikan saran-saran selama proses pengerjaan skripsi.
5. Ir. Baharuddin Mire, MT. sebagai Dosen Pembimbing Akademik yang telah memberikan ilmu dan nasehat sejak menjadi mahasiswa baru.
6. Bapak/Ibu dosen Departemen Teknik Mesin Universitas Hasanuddin yang telah memberikan ilmu, nasehat dan pengalaman kepada penulis selama menempuh studi di dunia perkuliahan
7. Teman-teman Teknik Mesin angkatan 2017 / ZYNCROMEZH'17 dan KMKT-UH yang senantiasa mendukung dan berjuang bersama sejak mahasiswa baru hingga saat ini.

8. Juan Daniel Tambunan Teknik Pertambangan angkatan 2019, Rahmat Dwi Hadyan Teknik Mesin angkatan 2017, dan Tri Gloria Bamba Akuntansi angkatan 2017 yang telah membantu proses pengambilan data.
9. Serta seluruh pihak yang telah membantu yang tidak bisa disebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna walaupun telah menerima bantuan dari berbagai pihak. Apabila terdapat kesalahan-kesalahan dalam skripsi ini sepenuhnya menjadi tanggung jawab penulis dan bukan para pemberi bantuan. Kritik dan saran yang membangun akan lebih menyempurnakan skripsi ini.

Gowa, Desember 2022

Penulis

DAFTAR ISI

SAMPUL	i
HALAMAN JUDUL	ii
LEMBAR PENGESAHAN	iv
PERNYATAAN KEASLIAN	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
NOMENKLATUR	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	3
1.4. Batasan Masalah	3
1.5. Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
2.1 Mobile Robot	5
2.2 Internet of Things (IoT)	8
2.3 UVC (Ultraviolet C)	10
2.4 Speech Recognition	12
2.5 Analisis Audio Recognition	16
BAB III METODOLOGI PENELITIAN	26
3.1 Waktu dan Tempat Penelitian	26
3.2 Desain Mobile Robot	26
3.3 Desain Perangkat Elektronik	27
3.4 Desain Perangkat Lunak	27
3.5 Alat dan Bahan	28
3.6 Metode Pengambilan Data	33
3.7 Flow Chart Penelitian	35

3.8 Rencana dan Jadwal Penelitian	36
BAB IV HASIL DAN PEMBAHASAN	37
4.1. Desain dan Pembuatan Mobile Robot Disinfektan UVC	37
4.2. Uji Coba Speech Recognition Controller	43
BAB V PENUTUP	63
5.1 Kesimpulan	63
5.2 Saran	64
DAFTAR PUSTAKA	65
LAMPIRAN	67

DAFTAR GAMBAR

Gambar 2.1 Orientasi Posisi Mobile Robot Dalam Sistem Koordinat Cartesia...	6
Gambar 2.2 Ilustrasi pergerakan robot berdasar titik pusat rotasi	7
Gambar 2.3 Pinout ESP32	7
Gambar 2.4 Lampu UVC	10
Gambar 2.5 Spektrum lampu UVC dan cara kerja UVC terhadap DNA	11
Gambar 2.6 Ilustrasi Speech Recognition	14
Gambar 2.7 Langkah-langkah pemrosesan dalam Analisis Fourier waktu-diskrit dari sinyal waktu-kontinu.....	17
Gambar 2.8 Sebuah sinyal suara komposit (hitam) yang terdiri dari beberapa sinyal-sinyal murni (biru) dengan frekuensi yang berbeda (merah)	18
Gambar 2.9 Diskritisasi sinyal waktu.....	19
Gambar 2.10 Membagi sinyal menjadi dua sinyal baru	19
Gambar 2.11 a) Dengan bobot jaringan yang terlatih dan tanpa noise gambar tambahan, hasilnya menunjukkan fitur yang terintegrasi dengan baik dari gaya dan konten. b) Dengan bobot acak, pengaruh gaya sulit dideteksi dan konten terdengar berisik. c) Menambahkan noise ke gambar awal menghasilkan suara yang memiliki fitur amplitudo kasar dari konten dan kayu berisik yang hampir tidak dapat diidentifikasi dengan sumber gaya d) Bobot acak dan noise gambar tambahan menyebabkan hilangnya rasa konten dan gaya	23
Gambar 2.12 Sinyal epilepsi (EEG 1) diambil dari pasien berusia 1,5 tahun	24
Gambar 2.13 Perbandingan spektrum periodogram dan AR dari sinyal EEG+1	25
Gambar 3.1 Skema Desain Mobile Robot Disinfektan UVC	26
Gambar 3.2 Skema Rangkaian Elektronik Mobile Robotic	27
Gambar 3.3 Sasis dasar robot	28
Gambar 3.4 Motor DC	28
Gambar 3.5 ESP-32	29
Gambar 3.6 ESP-32CAM	29
Gambar 3.7 Lampu UVC	29
Gambar 3.8 Baterai Li-ion 18650	30
Gambar 3.9 Motor Driver L298N.....	30
Gambar 3.10 Modul XH-M603	30

Gambar 3.11 Modul BMS	31
Gambar 3.12 Sensor ML8511	31
Gambar 3.13 Inverter 40W	31
Gambar 3.14 Modul LM2596	32
Gambar 3.15 Modul Relay	32
Gambar 3.16 Modem	32
Gambar 3.17 Android	33
Gambar 4.1 Tampilan perangkat lunak Fusion 360	37
Gambar 4.2 Metode pencetakan salah satu body menggunakan Printer 3D	38
Gambar 4.3 Skema Rangkaian Elektronik Mobile Robot	39
Gambar 4.4 Tampilan dashboard Digitalocean	40
Gambar 4.5 Tunneling EPS-32 CAM melalui Ngrok	40
Gambar 4.6 Tampilan software Arduino IDE	41
Gambar 4.7 Tampilan aplikasi kontrol Android.....	42
Gambar 4.8 Tampilan halaman codeblock pada MIT App Inventor.....	43
Gambar 4.9 Karakteristik suara pada perintah “maju”	44
Gambar 4.10 Karakteristik suara pada perintah “mundur”	47
Gambar 4.11 Karakteristik suara pada perintah “kanan”	50
Gambar 4.12 Karakteristik suara pada perintah “kiri”	53
Gambar 4.13 Karakteristik suara pada perintah “berhenti”	55
Gambar 4.14 Karakteristik suara pada perintah “on”	57
Gambar 4.15 Karakteristik suara pada perintah “off”	60

DAFTAR TABEL

Tabel 2.1 Efektivitas Dosis Ultraviolet terhadap Coronavirus.....	12
Tabel 4.1 Data kecepatan respon input speech recognition controller	62

NOMENKLATUR

I	Intensitas ultraviolet	W/m^2
D	Dosis UV	J/m^2
t	Waktu	s
A	Amplitudo	m
f	Frekuensi	Hz
P	Penguatan daya/power	dB

BAB I PENDAHULUAN

1.1. Latar Belakang

Sejak Februari 2020 wabah corona virus yang disebabkan oleh virus SARS-Cov2 telah menyebar ke Indonesia. Diyakini bahwa penyebaran virus corona pertama kali terjadi di pasar makanan laut di Wuhan dan terus menyebar melalui kontak orang ke orang.

Salah satu cara virus corona dapat menyebar yaitu manusia melalui droplet atau partikel yang disebabkan oleh batuk atau bersin. Selain itu, media lain seperti logam, kertas, kaca, dll juga bisa menjadi tempat penyebaran virus corona. Namun, karena virus corona sensitif terhadap panas, disinfektan yang mengandung klorin dan pelarut lemak pada suhu 56°C selama 30 menit, eter, alkohol, asam perasetat, formalin, oksidan dan kloroform dapat membunuh virus corona.

Selama wabah virus corona, penggunaan disinfektan aerosol yang disemprotkan pada manusia atau lingkungan di mana virus dapat menyebar sangat populer. Kementerian Kesehatan RI juga telah melarang penggunaan disinfektan aerosol secara terus-menerus karena dapat mengancam kesehatan manusia, terutama bahaya pada kulit dan selaput lendir. Meskipun berbahaya, penggunaan disinfektan dapat digunakan untuk sterilisasi ruangan.

Mobile robot merupakan jenis robot yang memiliki ciri khas mempunyai akuator berupa roda sebagai penggerak utamanya, sehingga robot tersebut dapat berpindah dari satu titik ke titik yang lain. Robot beroda yang menggunakan sistem gerak differential drivenya terdiri dari dua buah roda yang dipasangkan pada bagian kiri dan kanan robot untuk memungkinkan robot berputar di tempat dengan memutar motor ke arah yang berlawanan.

Internet of Things (IoT) adalah salah satu sistem yang dapat mengontrol perangkat elektronik yang terhubung ke Internet. *Internet of Things* telah mengubah dunia teknologi dimana objek-objek tradisional diaplikasikan sistem ini menjadi perangkat pintar melalui teknologi yang semakin berkembang, peralatan yang ada, teknologi komunikasi, jaringan sensor,

protokol dan jaringan internet. Metode ini memungkinkan seseorang mengontrol suatu objek (dalam hal ini adalah robot disinfektan) dari jarak jauh menggunakan koneksi jaringan internet, sehingga memudahkan kita untuk memonitoring dan mengendalikan robot disinfektan di manapun kita berada selama terhubung dengan internet.

Speech Recognition merupakan teknologi yang digunakan untuk menerjemahkan kata yang diucapkan oleh manusia menjadi tulisan. Sistem yang menerapkan *speech recognition* mampu mengerti instruksi yang diucapkan oleh manusia dan bertindak sesuai instruksi yang diberikan. *Speech recognition* adalah salah satu bidang kecerdasan mesin yang sedang berkembang pesat, hal itu ditandai oleh hampir semua *device* teknologi dilengkapi oleh *voice command*.

Proses mengubah pola suara ke dalam teks terlihat sederhana, namun kenyataannya melalui proses transkripsi yang rumit. Salah satu masalah utama dalam pengenalan dan pemahaman ucapan komputer adalah mengatasi banyak ruang pencarian. Keakuratan sebuah pengenalan pola suara didasari pada banyaknya jumlah kata yang terdapat dalam sebuah sistem, jika jumlah kosa kata meningkat maka ukuran ruang pencarian juga meningkat. Seiring bertambahnya ukuran ruang pencarian berakibat pada kinerja pengenalan pola suara menurun. Masalah lain dalam pengenalan suara adalah kompleksitas yang sangat besar yang terlibat dalam menganalisis input ucapan. Variasi dalam pengucapan, aksen, fisiologi pembicara, penekanan dan karakteristik lingkungan akustik biasanya menghasilkan ratusan klasifikasi fonem yang berbeda untuk setiap suara.

Salah satu metode yang dapat digunakan untuk ekstraksi suara adalah Mel Frequency Cepstral Coefficients (MFCC) karena metode ini dapat dengan baik dalam mempresentasikan sinyal. Sedangkan untuk pengenalan pola suara, metode yang dapat digunakan Vector Quantitation karena metode ini mampu mengklasifikasi sinyal suara dari hasil ekstraksi suara. Vector Quantitation adalah sebuah proses dalam memetakan vektor dari ruang vektor yang besar menjadi bentuk terbatas dalam ruang tersebut.

1.2. Rumusan Masalah

Berdasarkan latar belakang penelitian, maka dapat dirumuskan beberapa permasalahan yang dapat dibahas dalam penelitian ini untuk membatasi apa yang akan dibahas dalam inti penelitian ini. Beberapa masalah tersebut sebagai berikut.

1. Bagaimana mendesain pembuatan mobile robot disinfektan menggunakan UVC dengan *speech recognition controller* berbasis IOT?
2. Berapa jarak penerimaan suara efektif dari sumber suara ke *microphone*?
3. Bagaimana efektivitas penerimaan suara masuk dengan perbedaan aksent?
4. Bagaimana kecepatan respon mobile robot menggunakan *speech recognition*?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah di atas, ada beberapa tujuan dari penelitian ini antara lain :

1. Merancang dan mendesain mobile robot disinfektan menggunakan UVC dengan *speech recognition controller* berbasis IOT.
2. Menghitung jarak penerimaan suara efektif dari sumber suara ke *microphone*.
3. Menganalisa efektivitas penerimaan suara masuk dengan perbedaan aksent.
4. Menganalisa kecepatan respon mobile robot menggunakan *speech recognition*.

1.4. Batasan Masalah

Supaya penelitian ini dapat mengarah pada tujuan dan menghindari terlalu banyaknya permasalahan yang muncul, maka penulis memberikan batasan batasan masalah yang sesuai dengan judul penelitian ini. Adapun batasan masalah tersebut yaitu :

1. Perancangan mobile robot ini menggunakan ESP32 sebagai kontrol utamanya.
2. Menggunakan *software* MIT App Inventor.
3. Menggunakan bahasa pemrograman C++.
4. Tidak memperhitungkan material baik kekuatan bahan maupun sifat dari bahan yang digunakan.
5. Ukuran dari rancangan mobile robot disinfektan menggunakan UVC berbasis IOT yaitu panjang 220 mm, lebar 230 mm, tinggi 390 mm, diameter roda 67 mm.
6. Penerima suara menggunakan Handphone Android.

1.5. Manfaat Penelitian

1. Diharapkan Dapat Dijadikan Sebagai Pedoman Pembuatan Alat Sterilisasi Ruang Publik Yang Lebih Modern Dengan System IOT Dan Speech Recognition.
2. Memberikan solusi terhadap masalah sterilisasi ruangan saat ini.

BAB II TINJAUAN PUSTAKA

2.1. Mobile Robot

Mobile Robot adalah kombinasi dari berbagai komponen fisik (perangkat keras) dan komputasi (perangkat lunak). Dalam hal komponen perangkat keras, mobile robot dapat dianggap sebagai suatu kumpulan subsistem seperti:

- a. *Locomotion*: Bagaimana robot bergerak melalui lingkungannya
- b. *Sensing*: Bagaimana robot mengukur sifat dirinya dan lingkungannya
- c. *Reasoning*: Bagaimana Robot memetakan pengukuran ini menjadi tindakan
- d. *Communication*: Bagaimana robot berkomunikasi dengan perangkat lain

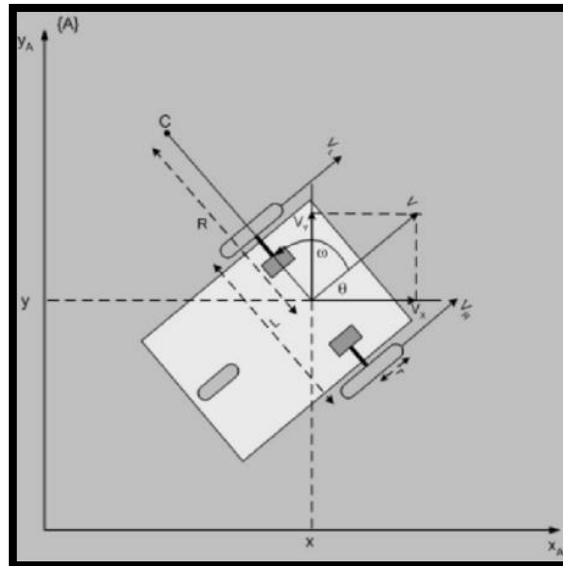
Secara umum mobile robot dapat diartikan sebagai robot yang memiliki ruang kerja yang luas. Dikatakan luas, karena bagian dasar dari robot ini diletakkan alat untuk bergerak, seperti roda atau kaki. Selain itu mobile robot dapat dibedakan juga berdasarkan kendalinya, yaitu Autonomous Robot, Teleoperated Robot, dan Semi Autonomos. (Erni D, 2009)

2.1.1. Kinematika Mobil Robot Penggerak Diferensial

Salah satu jenis penggerak robot mobile yang umum digunakan terutama untuk dioperasikan dalam ruangan adalah robot mobil adalah dengan sistem penggerak diferensial (differential drive). Alasan utamanya karena relatif lebih fleksibel dalam melakukan manuver serta kemudahan dalam pengontrolannya. (Erni D, 2009)

Secara teknis, robot jenis ini pada dasarnya memiliki dua roda utama yang masing-masing digerakan oleh penggerak tersendiri (umumnya berupa motor DC magnet permanent dengan gear-pereduksi yang berfungsi untuk memperkuat torsi motor), selain itu robot ini dilengkapi juga dengan satu atau dua buah roda castor yang ditempatkan dibagian belakang robot yang berfungsi sebagai penyeimbang. **Gambar 2.1** memperlihatkan arsitektur

robot dilihat dari bagian atas: Jika kedua roda penggerak tersebut berputar dengan kecepatan yang sama maka robot tersebut akan bergerak dengan arah yang lurus, sedangkan jika kecepatan salah satu roda lebih lambat maka robot akan bergerak membentuk kurva dengan arah lintasan menuju salah satu roda yang bergerak lebih lambat.



Gambar 2.1 Orientasi dan Posisi Mobile Robot Dalam Sistem Koordinat Cartesian (Erni, 2009)

Untuk panjang jari-jari roda r , serta kecepatan rotasi masing-masing roda kanan dan roda kiri berturut-turut ω_r dan ω_l , maka kecepatan linear roda kanan dan roda kiri dapat dicari dengan menggunakan persamaan berikut:

$$V_r(t) = r\omega_r(t) \quad (2.1)$$

$$V_l(t) = r\omega_l(t) \quad (2.2)$$

Ketika robot melakukan gerak memutar (berotasi) sesaat dengan panjang jari-jari R yang diukur dari pusat rotasi (ICC) dan titik pusat kedua roda seperti pada **Gambar 2.2** (Erni, 2009). Maka kecepatan rotasi disetiap titik robot tersebut akan selalu sama (robot adalah sistem mekanis yang rigid), sehingga persamaan berikut ini berlaku untuk menghitung kecepatan rotasi dari robot tersebut.

$$V_R = \omega(R+d) \quad (2.3)$$

$$V_L = \omega(R-d) \quad (2.4)$$

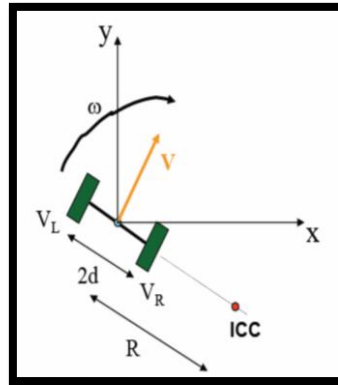
kemudian,

$$\omega = (V_R - V_L) / 2d \quad (2.5)$$

$$R = d(V_R + V_L) / (V_R - V_L) \quad (2.6)$$

Sehingga kecepatan robot :

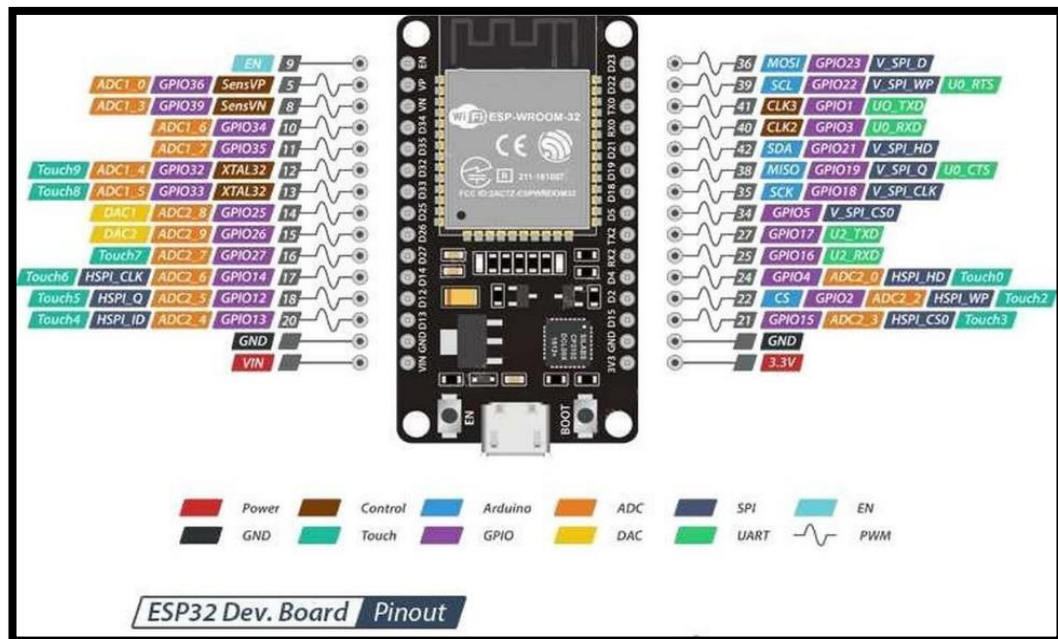
$$V = \omega R = (V_R + V_L) / 2 \quad (2.7)$$



Gambar 2.2 Ilustrasi pergerakan robot berdasar titik pusat rotasi (Erni D, 2009)

2.1.2. Mikrokontroler ESP32

ESP 32 adalah rangkaian sistem berdaya rendah untuk mikrokontroler chip yang mengintegrasikan *WiFi* dan *Bluetooth* mode ganda. Mikrokontroler ini sudah memiliki modul *WiFi* pada chipnya, yang membuatnya sangat berguna untuk membuat sistem aplikasi *Internet of Things*. (Grieco L., 2014)



Gambar 2.3 Pinout ESP32

Terdapat 36 pin GPIO (**Gambar 2.3**) yang bisa difungsikan sebagai berikut :

- *Analog to Digital Converter* (ADC) : 16 kanal SAR ADC 12 bit. Rentang ADC bisa diatur di dalam program, apakah 0-1 V, 0-1.4 V, 0-2V atau 0-4V.
- *Digital to Analog Converter* (DAC) : terdapat DAC 8 bit yang bisa menghasilkan tegangan analog.
- *Pulse Width Modulation* (PWM) : 16 kanal PWM yang bisa digunakan untuk mengendalikan LED atau motor.
- *Touch Sensor* : 10 GPIO memiliki kemampuan pengindera kapasitif yang dapat digunakan sebagai 10 tombol buttonpad.
- UART : 2 pin antarmuka UART. Salah satunya digunakan untuk mendownload program secara serial.
- I2C, SPI, I2S : Terdapat dua antarmuka I2C dan 4 antarmuka SPI untuk mengakses sensor dan perangkat ditambah lagi 2 antarmuka I2S.
(Grieco L., 2014)

2.2 Internet Of Things (IoT)

IoT merupakan segala aktifitas yang pelakunya saling berinteraksi dan dilakukan dengan memanfaatkan internet. Dalam penggunaannya *Internet of Things* banyak ditemui dalam berbagai aktifitas, contohnya banyaknya transportasi online, *e-commerce*, pemesanan tiket secara online, *live streaming*, *e-learning* dan lain-lain bahkan sampai alat-alat untuk membantu dibidang tertentu seperti *remote temperature sensor*, *GPS tracking*, dan sebagainya yang menggunakan internet atau jaringan sebagai media untuk melakukannya. (F, 2018)

2.2.1 Konsep dan Cara Kerja Internet of Things

Cara kerja dari IoT yaitu setiap benda harus memiliki sebuah alamat Internet Protocol (IP). Alamat IP adalah sebuah identitas dalam jaringan yang membuat benda tersebut bisa diperintahkan dari benda lain dalam jaringan yang sama. Alamat IP dalam benda-benda tersebut akan dikoneksikan ke jaringan internet.

Saat ini koneksi internet sudah sangat mudah didapatkan. Dengan demikian pengguna dapat memantau benda bahkan memberi perintah kepada benda lain dengan koneksi internet. Setelah modul *wireless* yang dilengkapi sistem IoT memiliki alamat IP dan terkoneksi dengan internet, pada modul tersebut juga dipasang sebuah sensor. Sensor pada module memungkinkan modul tersebut memperoleh informasi yang dibutuhkan. Setelah memperoleh informasi, modul tersebut dapat mengolah informasi itu sendiri, bahkan berkomunikasi dengan benda-benda lain yang memiliki alamat IP dan terkoneksi dengan internet juga. (Kurniawan, 2018)

2.2.2 Rekomendasi Perangkat Internet of Things

Untuk membuat sebuah sistem IoT, diperlukan sebuah perantara untuk menghubungkan pengguna dengan mikrokontroller. Perantara yang digunakan adalah API atau *Application Programming Interface*.

API dapat memudahkan *programmer* dalam mengendalikan mikrokontroller IoT yang dimiliki. Beberapa API yang bisadigunakan yaitu: Agnosthings, Evothings, Telegram, Blynk, dan lain sebagainya. Dari sekian banyak API, yang paling mudah untuk para pemula adalah Telegram. (A. Yahya, 2019)

2.2.3 VPS (Virtual Private Server)

Untuk membangun sistem IoT dibutuhkan server untuk menjalankan API, salah satu jenis server yang digunakan adalah VPS. VPS (Virtual Private Server) adalah teknologi virtualisasi server. Server fisik dibagi menjadi beberapa server pribadi virtual, dengan masing-masing VPS bertindak seperti server mandiri yang sebenarnya. Setiap VPS memiliki pengaturan sendiri seperti akses root penuh, sistem operasi, dan skrip inisialisasi termasuk sumber daya server seperti CPU dan RAM yang berdiri sendiri, pengguna, pemrosesan, dan sistem file. Tidak seperti shared hosting, yang menggunakan sumber daya server untuk saling memengaruhi, proses yang berjalan di VPS tidak memengaruhi VPS lain di satu server. (Ricky et al., 2010)

VPS memungkinkan Anda untuk menjalankan beberapa sistem operasi secara bersamaan pada satu sistem server fisik. Ini dapat dilakukan tanpa

mempartisi ulang partisi dan me-reboot. VPS yang ditentukan menjalankan sistem operasi sesuai kebutuhan. Dengan cara ini, pengguna dapat mem-boot sistem operasi (misalnya, Linux) dan menjalankan sistem operasi lain. Sistem operasi yang berjalan pada sistem operasi host disebut sistem operasi tamu.. (Eckart, Chris. 2008)

2.3 UVC (Ultraviolet C)

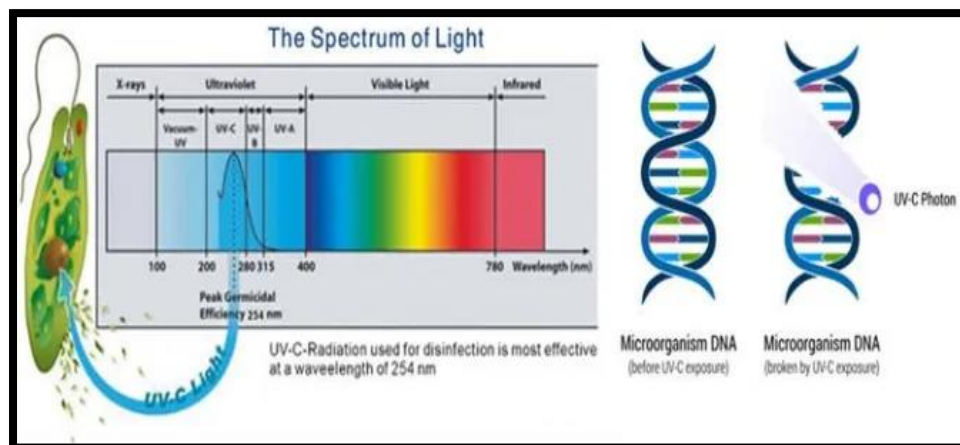
Radiasi ultraviolet telah dipelajari sebagai kemungkinan pembersih patogen lingkungan, termasuk desinfeksi permukaan, instrumen, dan udara. Sinar ultraviolet menghancurkan ikatan molekul dalam DNA pada panjang gelombang tertentu untuk memainkan efek bakterisida. Panjang gelombang spesifik sinar UVC adalah antara 200 dan 270 nanometer (biasanya 254 nanometer), dan panjang gelombang ini berada di bagian kuman dari spektrum elektromagnetik (200 hingga 320 nanometer). Studi tentang efek radiasi ultraviolet harus mempertimbangkan interaksi antara beberapa parameter yang berbeda, secara tidak langsung, waktu paparan, posisi lampu relatif terhadap permukaan yang diterangi, dan hambatan antara sumber cahaya dan permukaan target. Kekuatan cahaya yang dipancarkan, kecepatan dan aliran pergerakan udara. Semua ini mempengaruhi efek keseluruhan radiasi UVC pada permukaan. (Cahyo, 2020)



Gambar 2.4 Lampu UVC

Sumber cahaya ultraviolet buatan umumnya berasal dari lampu fluorescent khusus, seperti lampu merkuri bertekanan rendah dan lampu

merkuri bertekanan sedang. Lampu merkuri bertekanan sedang dapat menghasilkan keluaran radiasi ultraviolet yang lebih tinggi daripada lampu merkuri bertekanan rendah. Namun, lampu merkuri bertekanan rendah mengkonsumsi listrik lebih efisien daripada lampu merkuri bertekanan sedang. Panjang gelombang maksimum radiasi yang dihasilkan oleh lampu merkuri bertekanan rendah adalah 253,7 nm, yang berakibat fatal bagi mikroorganisme, protozoa, virus, dan ganggang. Pada saat yang sama, panjang gelombang radiasi yang dipancarkan oleh lampu merkuri tekanan sedang adalah 180 hingga 1370 nm. Sterilisasi UV adalah proses fisik di mana energi elektromagnetik ditransfer dari sumber (lampu) ke bahan seluler tubuh (protein dan asam nukleat). Sinar ultraviolet merusak DNA mikroorganisme dengan membentuk dimer timin. Dimer ini mencegah mikroorganisme dari menyalin dan mereplikasi DNA yang pada akhirnya menyebabkan kematian sel. Mekanisme kerusakan DNA UV berdasarkan Alcamo (1984) ditunjukkan pada gambar berikut.



Gambar 2.5 Spektrum lampu UVC dan cara kerja UVC terhadap DNA

Penggunaan dosis radiasi ultraviolet telah terbukti efektif melawan bakteri dan virus di air tanah tanpa menyebabkan pembentukan produk samping yang beracun. Tingkat inaktivasi mikroba tergantung pada dosis UV yang digunakan. Dosis UV adalah produk dari intensitas UV dan waktu paparan, seperti yang ditunjukkan pada **Persamaan 2.8** (Metcalf, Eddy, 2003).

$$D = I \cdot t \quad (2.8)$$

dimana :

D = Dosis UV ($\text{mW} \cdot \text{s}/\text{cm}^2 = \text{mJ}/\text{cm}^2$)

I = Intensitas ultraviolet (mW/cm^2)

t = Waktu paparan (sekon)

Berikut beberapa data tentang dosis yang dapat menonaktifkan virus SARS Coronavirus berdasarkan data dari beberapa sumber.

Tabel 2.1 Efektivitas Dosis Ultraviolet terhadap Coronavirus

Jenis Virus	Dosis (J/m^2)	Efektivitas	Sumber
SARS-CoV-2 (Italy-INMI1)	12.3	90%	Bianco 2020
SARS Coronavirus (Frankfurt 1)	16.4	90%	Eickmann 2020
SARS Coronavirus (CoV-P9)	40.0	90%	Duan 2003
SARS-CoV-2 (SARS-CoV-2/Hu/DP/Kng/19-027)	41.7	90%	Inagaki 2020
SARS Coronavirus (Hanoi)	133.9	90%	Kariwa 2004
SARS-CoV-2	237.1	99.99%	Anna et al 2021

Dari **Table 2.1** dapat diketahui bahwa dosis $237.1 \text{ J}/\text{m}^2$ dapat menonaktifkan coronavirus tipe SARS-CoV-2 dengan efektivitas sebesar 99.99%. Dengan demikian berdasarkan data-data dosis UV terhadap efektivitas menonaktifkan virus corona bisa dijadikan dasar untuk melakukan penelitian ini.

2.4 Speech Recognition

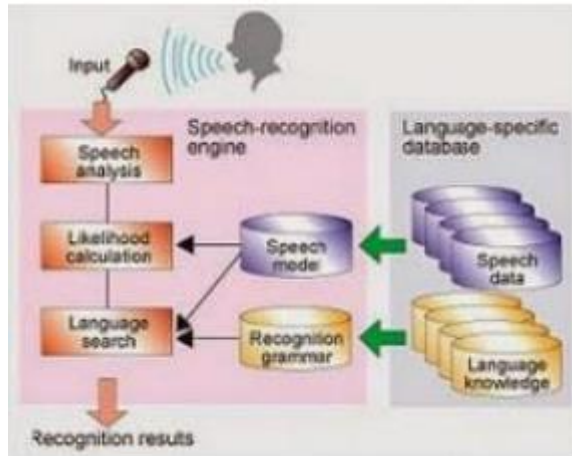
Voice recognition sendiri dibagi menjadi dua jenis, yaitu *speech recognition* dan *speaker recognition*. Berbeda dengan *speaker recognition* yang merupakan pengenalan identitas yang diklaim oleh seseorang dari suaranya (khusus dapat berupa intonasi suara, tingkat kedalaman suara, dan sebagainya), *speech recognition* adalah proses yang dilakukan komputer untuk mengenali kata yang diucapkan oleh seseorang tanpa mempedulikan

identitas orang terkait. Implementasi *speech recognition* misalnya perintah suara untuk menjalankan aplikasi komputer. (J. Dioses, 2020)

Speech Recognition, juga dikenal sebagai *Automatic Speech Recognition* (ASR), pengenalan ucapan komputer, atau ucapan-ke-teks, adalah kemampuan yang memungkinkan program untuk memproses ucapan manusia ke dalam format tertulis. Meskipun umumnya bingung dengan pengenalan suara, pengenalan suara berfokus pada terjemahan ucapan dari format verbal ke teks sedangkan pengenalan suara hanya berusaha mengidentifikasi suara pengguna individu. (Dioses, 2020)

Suara adalah bentuk dasar dari metode komunikasi yang efisien bagi orang untuk berinteraksi satu sama lain. Saat ini teknologi komunikasi umumnya tersedia untuk berbagai tugas yang terbatas namun menarik. Teknologi ini memungkinkan perangkat keras merespons suara manusia dengan benar dan andal serta menyediakan layanan yang bermanfaat dan berharga. Karena berkomunikasi dengan komputer lebih cepat menggunakan suara daripada keyboard. Komunikasi di antara manusia didominasi oleh bahasa lisan, oleh karena itu wajar bagi orang untuk mengharapkan antarmuka suara dengan komputer. Hal ini dapat dicapai dengan mengembangkan sistem pengenalan suara (*voice recognition system*). Sistem pengenalan suara adalah sebuah sistem yang mengubah suara yang ditangkap melalui microphone ke dalam sebuah kata atau kalimat. (Dioses, 2020)

Speech Recognition bekerja dengan cara merubah sinyal analog menjadi sinyal digital oleh *analog to digital converter* (ADC). Pada Gambar 2.6 ilustrasi *Speech Recognition* merupakan skema dari yang menjelaskan proses untuk mengenali dan memahami kata-kata yang diucapkan oleh satu orang atau banyak orang yang diinputkan melalui mikrofon (ASR). Kemudian memprosesnya sehingga dapat mengenali apa yang diucapkan oleh penggunanya. (Rizki Aditya, 2020)



Gambar 2.6 Ilustrasi *Speech Recognition*

Beberapa faktor dapat menyebabkan kesalahan dalam proses verifikasi dan identifikasi suara antara lain:

- 1) Kesalahan dalam pengucapan (misspoken) dan pembacaan (misread) frasa.
- 2) Keadaan emosional yang ekstrim (misalnya stres).
- 3) Pergantian penempatan mikropon (intrasession atau intersession).
- 4) Kekurangan atau ketidak-konsistenan akustik dari ruangan (misalnya multipath dan noise).
- 5) Sakit (misalnya flu yang dapat merubah vocal tract).
- 6) Aging (model vocal tract dapat berubah berdasarkan usia).

Faktor-faktor ini secara umum di luar area algoritma atau koreksi yang lebih baik dengan pengertian lain dari algoritma, misalnya mikropon yang lebih baik. Faktor-faktor ini penting, bagaimanapun, karena tidak ada materi sebaik sebuah algoritma dari speech recognition, human error (misalnya misspeaking atau misreading) pada akhirnya mengurangi performa pengenalan suara. (Rizki Aditya, 2020)

2.4.1 Kelebihan dan Kekurangan Aplikasi Pengenalan Suara

Adapun kelebihan aplikasi pengenalan suara, yaitu:

- 1) Pada bidang komunikasi:
 - a. Sebagai komando suara pada komputer untuk melakukan perintah seperti membuka, menyimpan, menutup file dan sebagainya.

- b. Sebagai alat pendiktean untuk membuat laporan dimana aplikasi akan menuliskan teks sesuai dengan yang diucapkan pembicara.
- 2) Pada bidang kesehatan, alat pengenalan ucapan digunakan untuk membantu para penyandang cacat agar dapat beraktivitas dengan memerintahkan alat-alat bantu melalui suaranya.
- 3) Pada bidang militer, aplikasi alat pengenalan ucapan digunakan pada pengatur lalu lintas udara yang dikenal dengan Air Traffic Controllers (ATC). Alat ini dipakai untuk memberi keterangan mengenai keadaan lalu lintas udara seperti radar, cuaca, dan navigasi dengan kata lain sebagai pengganti operator yang memberikan informasi kepada pilot dengan cara berdialog.
- 4) Secara umum, pengenalan ucapan banyak digunakan karena mudah yakni hanya menggunakan suara dan cepat prosesnya.

Adapun kekurangan aplikasi pengenalan suara, yaitu:

- 1) Rawan terhadap gangguan sinyal suara lain terutama di tempat yang ramai.
- 2) Kata-kata yang diucapkan dapat sulit dikenali karena cara pengucapan yang berbeda walaupun oleh pembicara yang sama. Intonasi, logat, dan kecepatan pengucapan sangat mempengaruhi.
- 3) Bahasa lisan seringkali diucapkan tidak sesuai dengan kaidah tata bahasa yang baku.
- 4) Jumlah kata yang dapat dikenali masih terbatas.

(Rizki Aditya, 2020)

2.4.2 *Speech Recognition Service* pada Perangkat *Android*

Pengembangan aplikasi pengenalan suara dimungkinkan dengan adanya framework speech recognition yang sudah disediakan dalam SDK Android. Speech Recognition menggunakan suara yang masuk melalui perangkat mikrofon dan dideteksi oleh perangkat lunak yang menafsirkan suara tersebut dan menerjemahkannya ke dalam kata atau kalimat. Speech recognition services dapat digunakan untuk perintah pendek seperti kalimat tanpa jeda, namun belum mampu untuk kalimat yang relatif panjang. SDK Android 2.1

dan SDK Android yang lebih tinggi dapat mendukung Speech Recognition sampai batas tertentu. Android Speech Recognizer Intent digunakan untuk meluncurkan perekam suara. File suara akan dikirim ke server Google kemudian akan diproses, hasil dari pengolahan speech recognition akan menghasilkan beberapa pilihan kata yang keluar dari pengolahan server, sehingga fitur ini memerlukan jaringan internet yang mendukung. (Vaipai, 2016)

2.5 Analisis Audio Recognition

2.5.1 Fourier Transform

Deret Fourier adalah ide matematika yang berlaku untuk hampir semua hal di dunia nyata kita. Transformasi Fourier memberi kita cara yang ampuh untuk melihat bentuk gelombang ini. Transformasi Fourier dianalogikan dengan memisahkan cat campuran [2]. Jika kue adalah bentuk gelombang, resepnya adalah Transformasi Fourier, daftar bahan dan cara menggabungkannya. Aplikasi utamanya termasuk memisahkan sinyal dari kebisingan dalam data yang dikumpulkan untuk astronomi, kedokteran, genetika dan kimia. Anggota keluarga deret Fourier yang terkait dengan pemrosesan sinyal digital adalah Discrete Fourier Transform (DFT) yang beroperasi pada domain waktu sampel yang periodik. Setiap suara digital, f adalah jumlah suara murni pada frekuensi yang berbeda [3]. Pertimbangkan suara digital $x[n]$ yang periodik dengan periode N , sehingga $x[n] = x[n+rN]$ untuk setiap nilai bilangan bulat dari n dan r . Urutan tersebut dapat diwakili oleh deret Fourier yang sesuai dengan jumlah deret eksponensial kompleks yang terkait secara harmonis. Eksponen kompleks periodik ini berbentuk:

$$e_k[n] = e^{j(2\pi/N)kn} = e_k[n+rN] \quad (2.9)$$

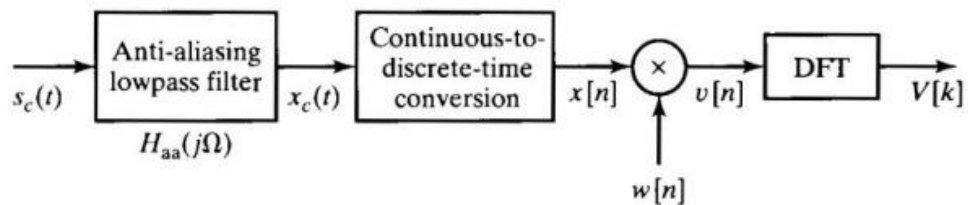
di mana k adalah bilangan bulat, dan representasi Deret Fourier memiliki bentuk

$$x[n] = 1/N \sum_k X[k] e^{j(2\pi/N)kn} \quad (2.10)$$

Fourier Series memungkinkan eksperimentalis untuk mengidentifikasi sumber suara. Dalam aplikasi kita, untuk menghilangkan sumber suara, kita harus menggunakan Inverse Fourier Transform. Secara umum, tujuan dari

Inverse Fourier Transform adalah untuk mengambil $x[n]$ dan menggunakannya untuk merekonstruksi sinyal aslinya.

Langkah-langkah dasar dalam menerapkan Transformasi Fourier Diskrit (DFT) ke sinyal waktu kontinu ditunjukkan pada gambar. Filter antialiasing digabungkan untuk menghilangkan atau meminimalkan efek aliasing ketika sinyal waktu kontinu diubah menjadi urutan. $x[n]$ dikalikan dengan ukuran jendela $w[n]$, untuk panjang terbatas yang diperlukan dalam DFT. Dalam menganalisis musik, $x[n]$ panjangnya tak berhingga. Oleh karena itu, jendela durasi-terhingga $x[n]$ diterapkan ke $x[n]$ sebelum penghitungan DFT. (Cao, 2019)



Gambar 2.7 Langkah-langkah pemrosesan dalam Analisis Fourier waktu-diskrit dari sinyal waktu-kontinu

Konversi $x_c[t]$ ke barisan sampel $x[n]$ direpresentasikan dalam domain frekuensi dengan replikasi periodik dan normalisasi frekuensi yaitu:

$$X(e^{j\omega}) = \frac{1}{T} \sum_{r=-\infty}^{\infty} X_c \left(j \frac{\omega}{T} + j \frac{2\pi r}{T} \right), \quad (2.11)$$

Seperti yang ditunjukkan, urutan $x[n]$ dikalikan dengan jendela durasi-hingga $w[n]$, karena input ke DFT harus durasinya berhingga. Ini menghasilkan urutan panjang hingga $v[n]=w[n]*x[n]$. Efek dalam domain frekuensi adalah konvolusi periodik, yaitu:

$$V(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\theta})W(e^{j(\omega-\theta)})d\theta. \quad (2.12)$$

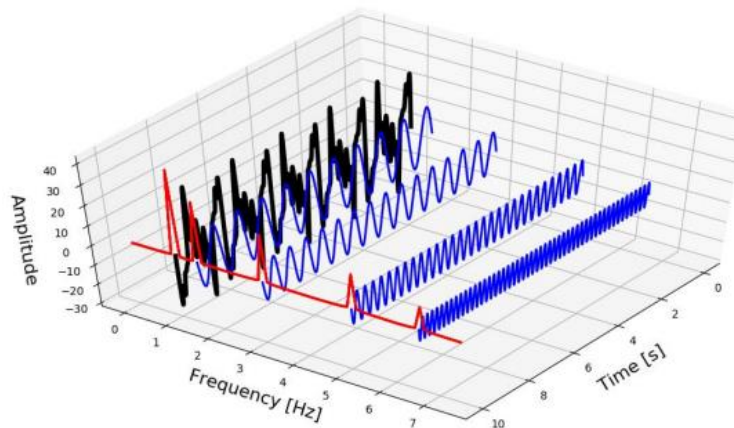
Konvolusi $W(e^{j\omega})$ dan $X(e^{j\omega})$ akan cenderung menghaluskan puncak tajam dan diskontinuitas pada $X(e^{j\omega})$. DFT dari barisan berjendela $V[n]=w[n]*x[n]$ adalah,

$$V[k] = \sum_{n=0}^{N-1} v[n]e^{-j(2\pi/N)kn}, \quad k = 0, 1, \dots, N-1, \quad (2.13)$$

Penganalisis spektrum Audacity didasarkan pada teori ini. Cara ini berguna untuk membersihkan rekaman – jika ada rekaman suara desis dan pop in, bisa dibersihkan dengan melakukan analisis spektrum. Catatan individu yang tidak diinginkan dapat secara luas diklasifikasikan ke:

1. Menghapus catatan yang tidak diinginkan

Misalkan kita memiliki suara digital yang memiliki nada tinggi yang mengganggu yang ingin kita edit. Suara datang sebagai fungsi dari berbagai intensitas dari waktu ke waktu (sinyal biru) yang membentuk gelombang suara komposit (gelombang hitam pada gambar di bawah). Transformasi Fourier dari gelombang suara akan menunjukkan lonjakan tajam yang sesuai dengan amplitudo dari masing-masing nada. Kita dapat menyaring nada yang tidak diinginkan dengan menghancurkan spike ke bawah. Kami kemudian menerapkan Transformasi Fourier terbalik ke semua nada untuk membentuk suara digital yang diedit.



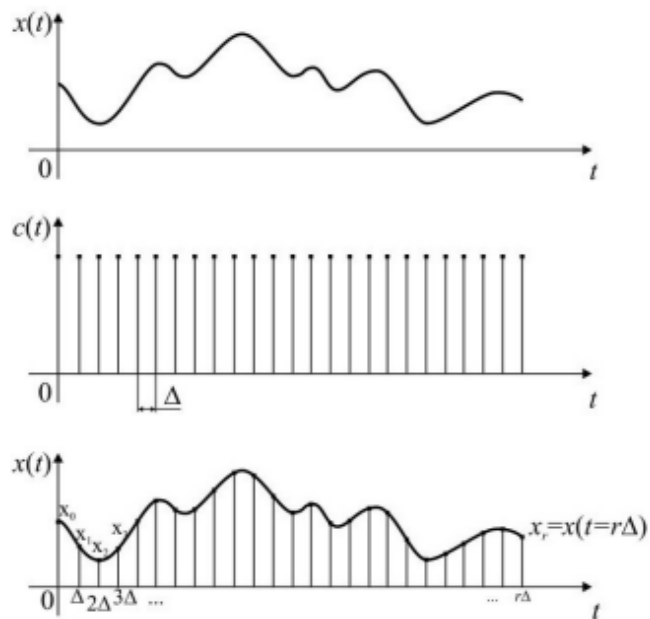
Gambar 2.8 Sebuah sinyal suara komposit (hitam) yang terdiri dari beberapa sinyal-sinyal murni (biru) dengan frekuensi yang berbeda (merah)

2. Menghilangkan Kebisingan

Diskrit Fourier Transform juga digunakan dalam menghilangkan suara. Di sini komponen kecil berfrekuensi tinggi dibuang. Setelah rekonstruksi bentuk gelombang suara asli, suara menjadi bebas noise. Transformasi Fourier juga dapat digunakan untuk mengekstrak sinyal yang terkubur dalam kebisingan sehingga telinga manusia tidak akan menyadarinya. Ini memungkinkan komunikasi dengan daya yang sangat rendah. (Fauzi, 2020)

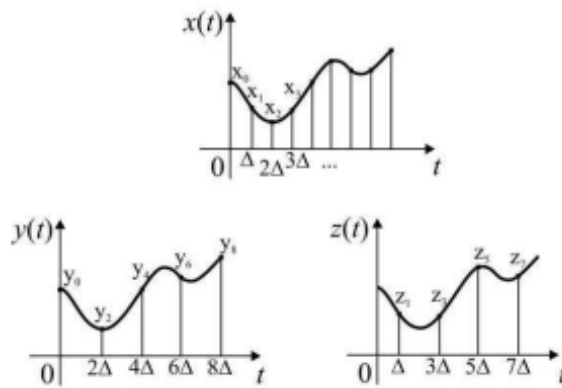
2.5.2 Fast Fourier Transform

Fast Fourier Transformation (FFT) adalah algoritma yang efektif dari Discrete Fourier Transformation (DFT) yang mengurangi waktu penghitungan dari N^2 ke $M\log_2 N$ di mana N adalah jumlah sampel dari sinyal diskrit. Fakta itu berarti penghematan waktu yang sangat besar dibandingkan dengan DFT. Diskritisasi sinyal waktu yang dibutuhkan untuk Transformasi Fourier Diskrit ditunjukkan pada **Gambar 2.9**.



Gambar 2.9 Diskritisasi sinyal waktu

Algoritma FFT didasarkan pada kenyataan bahwa setiap transformasi Fourier diskrit dengan N sampel dapat dibagi menjadi dua transformasi Fourier, masing-masing dengan $N/2$ sampel (pertama dengan sampel genap dan kedua dengan sampel ganjil), **Gambar 2.10**. (Fauzi, 2020)



Gambar 2.10 Membagi sinyal menjadi dua sinyal baru

Transformasi Fourier kemudian merupakan jumlah dari dua transformasi Fourier baru:

$$\begin{aligned}
 X_k &= \sum_{r=0}^{N-1} x_r e^{-j\frac{2\pi kr}{N}} = \\
 &= \sum_{r=0}^{\frac{N}{2}-1} x_{2r} e^{j\frac{2\pi k(2r)}{N}} + \sum_{r=0}^{\frac{N}{2}-1} x_{(2r+1)} e^{j\frac{2\pi k(2r+1)}{N}} = \\
 &= \sum_{r=0}^{\frac{N}{2}-1} x_{2r} e^{j\frac{2\pi kr}{N/2}} + e^{-j\frac{2\pi k}{N}} \cdot \sum_{r=0}^{\frac{N}{2}-1} x_{(2r+1)} e^{j\frac{2\pi kr}{N/2}}
 \end{aligned} \tag{2.14}$$

dimana r adalah nomor sampel. Kami memiliki dua transformasi Fourier baru dalam persamaan (2.14) sehingga kami dapat mendefinisikan variabel nyata

$$Y_k = \sum_{r=0}^{\frac{N}{2}-1} x_{2r} e^{j\frac{2\pi kr}{N/2}} \tag{2.15}$$

$$Z_k = \sum_{r=0}^{\frac{N}{2}-1} x_{(2r+1)} e^{j\frac{2\pi kr}{N/2}} \tag{2.16}$$

Dan variabel kompleks

$$W = e^{-j\frac{2\pi k}{N}} \tag{2.17}$$

Berkenaan dengan persamaan (2.15), (2.16) dan (2.17) kita dapat menulis persamaan (2.14) sebagai

$$X_k = Y_k + W^k Z_k \quad (2.18)$$

Persamaan (2.18) termasuk dalam sebagian besar perangkat lunak pengolah sinyal digital yang menggunakan FFT. (Fauzi, 2020)

2.5.3 Audio Spectrogram

Sampel audio mentah bersifat *lossless* dan dapat dengan mudah diubah menjadi audio. *WaveNet* [van den Oord et al., 2016], adalah jaringan konvolusi dalam (tidak berulang) yang menggunakan sampel audio mentah sebagai input dan dilatih untuk memprediksi sampel berikutnya yang paling mungkin secara berurutan. Selama fase generatif, setiap sampel yang diprediksi dimasukkan ke dalam urutan yang digunakan untuk memprediksi sampel berikutnya. Dengan informasi "pengkondisian" (seperti fonem mana yang diucapkan) bersama dengan input, kontrol parametrik yang menarik pada waktu sintesis dimungkinkan. Implementasi *WaveNet* berjalan sedalam 60 lapisan, dan audio mentah biasanya diambil sampelnya dengan kecepatan mulai dari 16K hingga 48K per detik, sehingga sintesis lambat pada beberapa menit pemrosesan audio per detik. (Helen, 2010)

Spektrum magnitudo juga dapat digunakan untuk aplikasi generatif yang diberikan teknik untuk menurunkan fase dari sifat spektrum magnitudo untuk merekonstruksi sinyal audio. Teknik rekonstruksi fase yang paling sering digunakan berasal dari Griffin dan Lim [1984], yang diimplementasikan di perpustakaan *Librosa* [McFee et al., 2015]. Namun, ini melibatkan banyak iterasi maju dan mundur. permukaan kesalahan yang terkadang mencegah rekonstruksi berkualitas tinggi. Penelitian terbaru telah menghasilkan metode yang secara teoritis dan dalam praktik real time [Zhu et al., 2007] [Pruřsa dan Søndergaard, 2016]; metode yang dapat menghasilkan transien yang sangat meyakinkan (kejadian padat sementara) [Pruřsa, 2017]; dan metode non-iteratif dengan kualitas wajar yang secepat penghitungan STFT tunggal [Beauregard et al., 2015].

Spektogram adalah gambar 2D yang mewakili urutan spektrum dengan waktu di sepanjang satu sumbu, frekuensi di sepanjang sumbu lainnya, dan kecerahan atau warna yang mewakili kekuatan komponen frekuensi pada setiap kerangka waktu. Dengan demikian representasi ini setidaknya

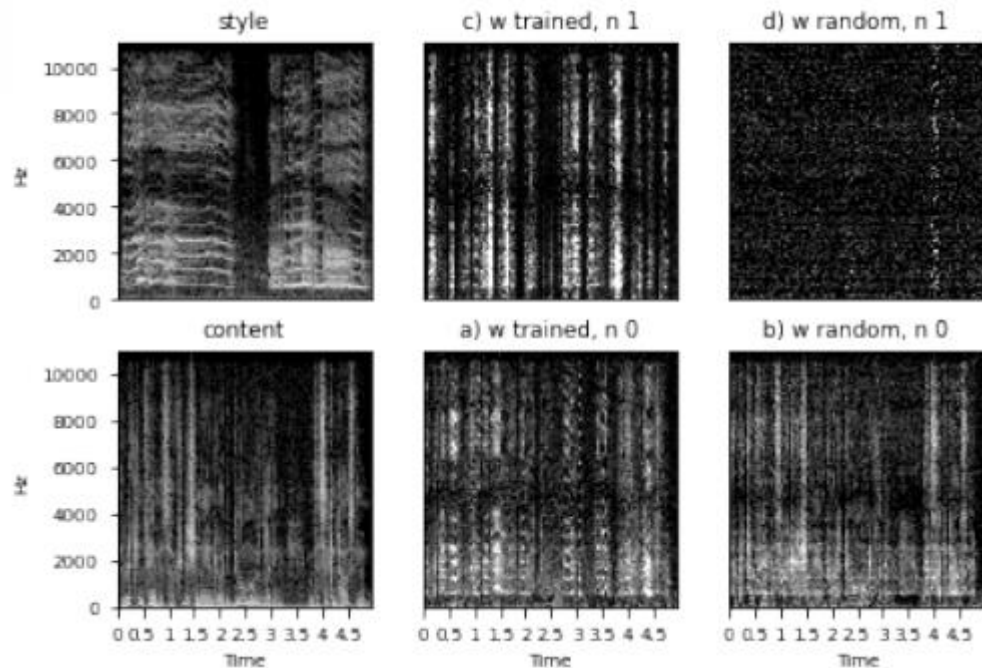
menunjukkan bahwa beberapa arsitektur jaringan saraf convolutional untuk gambar dapat diterapkan langsung ke suara.

Dmitry Ulyanov Ulyanov dan Lebedev [2016] melaporkan dalam posting blog tentang penggunaan jaringan saraf konvolusi dengan cara yang berbeda untuk transfer gaya audio. Dia menggunakan spektogram, tetapi alih-alih mewakili nampan frekuensi sebagai dimensi y dalam gambar, dia menganggap frekuensi yang berbeda sebagai yang ada pada titik yang sama dalam representasi 1D sebagai tumpukan "saluran" dengan cara yang sama 3 saluran untuk merah, hijau, dan biru ditumpuk di setiap titik dalam gambar visual 2D. Seperti dalam aplikasi gambar, kernel konvolusi mencakup seluruh dimensi saluran; tidak ada kernel konvolusi bobot bersama kecil yang bergeser di sepanjang dimensi saluran seperti halnya dalam dimensi spasial. Jumlah saluran audio, biasanya 256 atau 512, jauh lebih besar daripada 3 saluran yang digunakan untuk gambar berwarna, dan dimensi vertikal dikurangi menjadi satu.

Ada dua aspek luar biasa pada jaringan yang digunakan oleh Ulyanov untuk transfer gaya yang membedakannya dari pendekatan "klasik" yang dijelaskan oleh Gatys et al. [Gatys dkk., 2015]. Pertama, jaringan hanya menggunakan satu lapisan. Aktivasi jaringan yang mendorong pembuatan konten dan generasi gaya penggerak tersebut berasal dari satu set bobot yang sama. Perbedaan antara konten dan gaya dengan demikian tidak berasal dari kedalaman lapisan, tetapi hanya dari perbedaan antara ukuran aktivasi orde pertama dan orde kedua. Kedua, jaringan tidak dilatih sebelumnya, tetapi menggunakan bobot acak. Posting blog mengklaim pendekatan tidak intuitif ini menghasilkan hasil sebaik yang lain, dan contoh suara yang diposting memang menarik. (Laksono, 2018)

Sonogram yang dihasilkan dengan bobot dan kondisi kebisingan yang berbeda ditunjukkan pada **Gambar 2.11**. Target konten adalah ucapan dan target gaya adalah ayam jantan yang berkokok. Studi ini menunjukkan perbedaan yang signifikan antara bobot acak dan bobot pra-terlatih. Selain itu, jaringan yang dilatih untuk klasifikasi audio tidak memperkenalkan artefak yang dapat didengar dari jenis yang kami temukan menggunakan

jaringan gambar terlatih. Meskipun transfer gaya tidak bekerja tanpa memperhatikan bobot hanya berdasarkan pada urutan pertama dan urutan kedua konten dan strategi pencocokan gaya, jaringan yang dilatih untuk klasifikasi audio tampaknya menghasilkan sintesis konten dan gaya yang lebih terintegrasi. (Nada, 2019)



Gambar 2.11 a) Dengan bobot jaringan yang terlatih dan tanpa noise gambar tambahan, hasilnya menunjukkan fitur yang terintegrasi dengan baik dari gaya dan konten. b) Dengan bobot acak, pengaruh gaya sulit dideteksi dan konten terdengar berisik. c) Menambahkan noise ke gambar awal menghasilkan suara yang memiliki fitur amplitudo kasar dari konten dan kayu berisik yang hampir tidak dapat diidentifikasi dengan sumber gaya d) Bobot acak dan noise gambar tambahan menyebabkan hilangnya rasa konten dan gaya

2.5.4 Audio Periodogram

Metode periodogram yang digunakan untuk menentukan rapat daya komponen frekuensi dalam sinyal didasarkan pada konversi Fourier. Deret Fourier dapat mewakili sinyal periodik dan sekali lagi bentuk gelombang dapat diperoleh dengan menggunakan koefisien Fourier. Untuk mendapatkan spektrum daya sinyal EEG dengan metode periodogram, sinyal EEG dibagi menjadi frame 64, 128, 256, dengan pangkat 2. Untuk meningkatkan resolusi,

dilakukan proses zero-padding, karena resolusi FFT berbanding terbalik pada saat data sampel disajikan. (Helen, 2010)

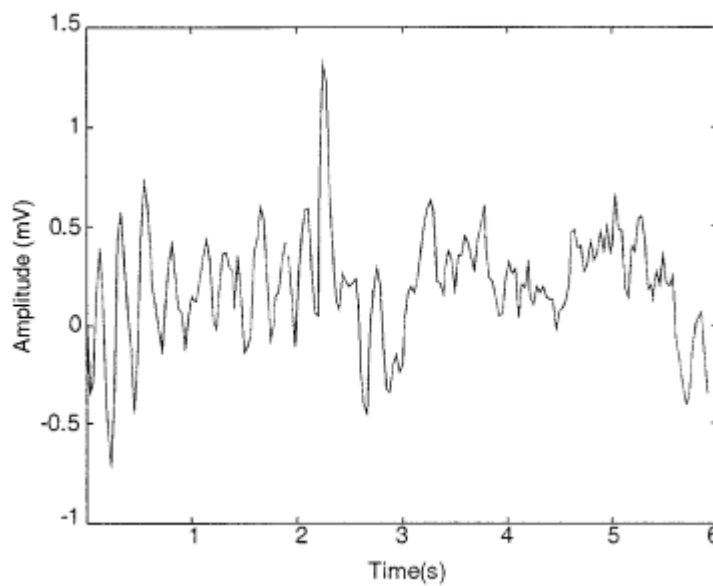
Sinyal berjendela dengan jendela yang tepat karena tidak mungkin bekerja dengan sinyal yang diperpanjang tak terbatas secara praktis. Beberapa jendela telah dikembangkan dengan studi yang dilakukan, dan jendela terbaik yang sesuai untuk data yang digunakan, dan informasi yang diperlukan untuk diperoleh telah disarankan. Ketika sinyal biomedis diperiksa, terutama lebih baik jendela yang sesuai adalah persegi panjang dan jendela Hanning pada analisis spektral sinyal ini. (Helen, 2010)

Transformasi Fourier Diskrit dari sinyal $X[n]$ waktu diskrit didefinisikan sebagai berikut.

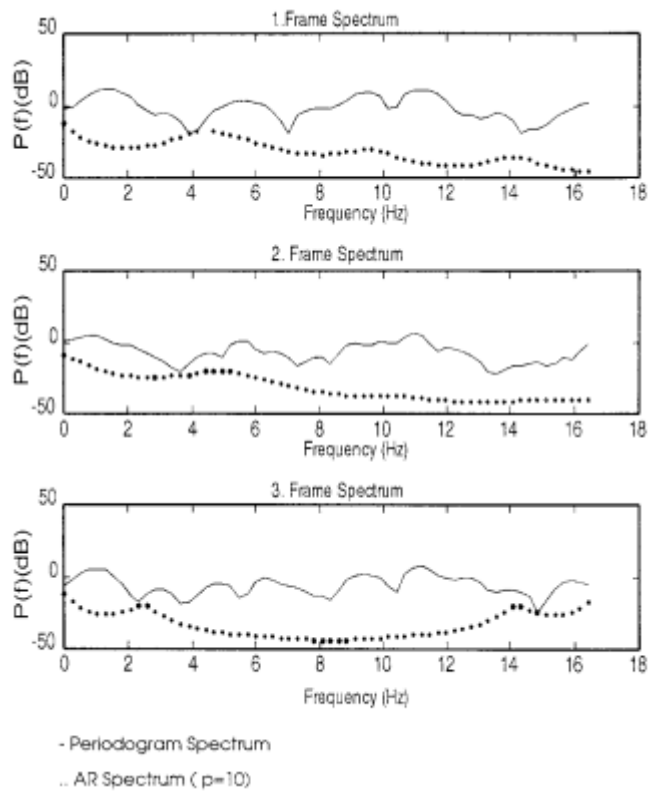
$$X[m] = \sum_{n=0}^{N-1} x[n] \exp\left(\frac{-j2\pi mn}{N}\right) \quad m = 0, 1, 2, 3, \dots, N-1 \quad (2.19)$$

Untuk mencari spektrum daya sinyal ini, digunakan persamaan berikut.

$$\hat{P}_{per}(f_m) = \frac{1}{N\Delta t} |X[m]|^2 \quad \Delta t = \frac{1}{f_s} \quad (2.20)$$



Gambar 2.12 Sinyal epilepsi (EEG 1) diambil dari pasien berusia 1,5 tahun



Gambar 2.13 Perbandingan spektrum periodogram dan AR dari sinyal EEG 1

Dimana, N , Δt , dan f_s masing-masing menunjukkan nilai estimasi, jumlah sampel dalam sinyal, periode sampling, dan frekuensi sampling. (Helen, 2010)