**BAB III**

**METODOLOGI**

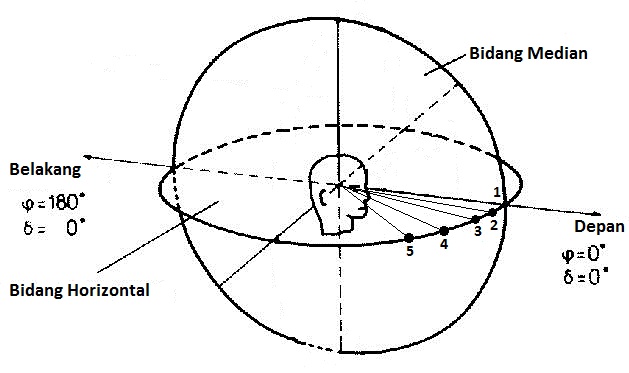
Pada bab ini akan dijabarkan metode yang disusun dan digunakan dalam penyelesaian permasalahan seperti pada gambar berikut.

**Gambar 3.1** Ilustrasi Metodologi Penelitian

Pada Gambar 3.1 menjelaskan bagaimana penelitian ini berjalan. Pada penelitian ini, terdapat dua proses yang dilakukan yaitu proses *data training* dan proses pemisahan. Pada proses *data training* terdapat lima percobaan yaitu pertama simulasi sinyal *binaural* sebagai sinyal suara masukan yang akan digunakan. Kedua *auditory periphery* yaitu simulasi sistem pendengaran manusia pada telinga bagian tengah hingga suara disalurkan ke otak. Ketiga melakukan perhitungan *binaural cue.* Keempat adalah perhitungan *relative strength* sinyal target terhadap sinyal masker. Pada tahap pertama sampai ketiga dilakukan 25 kali percobaan untuk memperoleh data yang maksimal. Kelima perhitungan *probability density estimation* antara *binaural cue* dan *relative strength.* Pada proses pemisahan terdapat lima percobaan yaitu pertama simulasi sinyal *binaural* sebagai sinyal suara masukan yang akan digunakan. Pada tahap yang pertama ini hanya dilakukan satu kali percobaan. Kedua *auditory periphery* yaitu simulasi sistem pendengaran manusia pada telinga bagian tengah hingga suara disalurkan ke otak. Pada tahap ini pencampuran sinyal suara dilakukan. Ketiga melakukan perhitungan *binaural cue* dari suara tercampur. Keempat proses estimasi *binary mask* berdasarkan *probability density* yang diperoleh dari proses *data training*. Pada tahap ini sinyal target direkonstruksi. Kelima evaluasi hasil pemisahan dengan metode subjektif dan objektif.

* 1. **Proses *Data Training***
     1. **Simulasi *Spatial Hearing***

Tahap pertama pada proses *data training* adalah simulasi *spatial hearing.* Proses ini sangat penting dalam *binaural processing* karena pada tahap ini sinyal suara yang pada awalnya tidak memiliki data spasial atau sinyal *monaural*, yaitu posisi sumber suara terhadap kedua telinga, akan diberi data spasial dengan cara mengkonvolusikan sinyal *monaural* tersebut dengan *head related transfer function* (HRTF) sesuai dengan persamaan 2.1.



**Gambar 3.2 Koordinat Sistem Pendengaran Spasial**

Data HRTF yang digunakan adalah CIPIC *(Center for Image Processing and Integrated Computing)* HRTF *Database* yang merupakan hasil pengukuran *head-related impulse response* (HRIR) pada telinga kiri dan kanan dari KEMAR manikin dalam ruang kedap *(unechoic room)* dengan jarak 1.4 m terhadap titik tengah kepala manusia pada bidang horisontal atau azimut φ dan bidang median atau elevasi (φ, δ) adalah titik 1 (0o, 0o), titik 2 (5o, 0o), titik 3 (10o, 0o), titik 4 (20o, 0o) dan titik 5(30o, 0o) seperti pada Gambar 3.2.

Untuk konfigurasi percobaan yang dilakukan pada tahap ini, dapat dilihat pada tabel berikut.

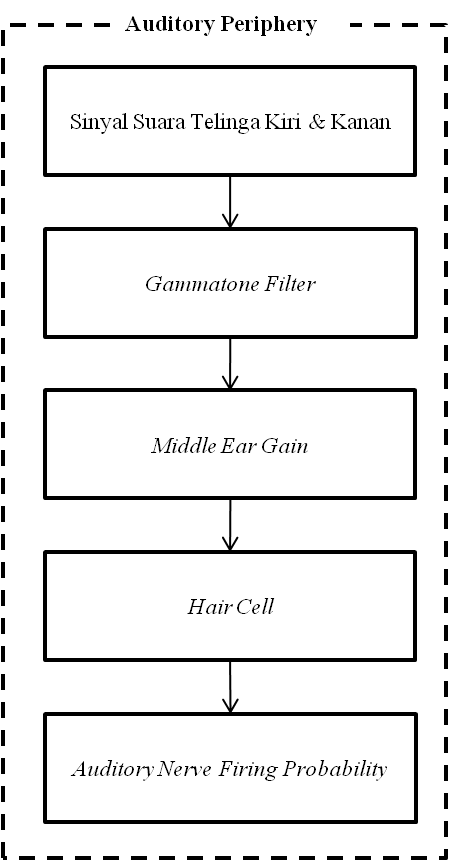
**Tabel 3.1** Konfigurasi Sinyal Target dan Masker pada Proses *Data Training*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sinyal Target | Azimut dan Elevasi | Sinyal Masker | Azimut dan Elevasi | *Signal to Interference Ratio* (SIR) |
| T1 | (0o, 0o) | M1 | (30o, 0o)  (20o, 0o)  (10o, 0o)  (5o, 0o) | 10 dB  5 dB  0 dB |
| T2 | M2 |
| T3 | M3 |
| T4 | M4 |
| T5 | M5 |

Sinyal suara yang digunakan dalam pada percobaan ini adalah data rekaman suara percakapan laki-laki sebanyak 5 kalimat sebagai sinyal target (T1, T2, T3, T4, T5) dan suara percakapan perempuan sebanyak 5 kalimat sebagai sinyal masker (M1, M2, M3, M4, M5) berbahasa Indonesia dengan frekuensi sampling 16.000 Hz. Konfigurasi yang dilakukan untuk sinyal target adalah selalu pada posisi yang sama yaitu (0o, 0o) dengan sinyal masker bergantian dari posisi (30o, 0o) hingga (5o, 0o) serta SIR yang berubah dari 10 dB hingga 0 dB. Sehingga total percobaan yang dilakukan adalah 5 x 1 x 5 x 4 x 3 = 75 percobaan.

* + 1. ***Auditory Periphery***

Suara target dan masker pada telinga kiri dan telinga kanan yang telah memiliki data spasial hanya memodelkan telinga bagian luar saja. Oleh sebab itu, untuk memodelkan suara yang diterima telinga semirip mungkin dengan aslinya maka pada tahap ini dilakukan proses *auditory periphery.*

**

**Gambar 2.3** Diagram Alir Proses *Auditory Periphery*

Pada Gambar 2.3,proses *auditory periphery* terdapat tiga tahap. Tahap pertama adalah *gammatone-filter*, yaitu suara target dan masker hasil simulasi *binaural* pada pada kedua telinga difilter menggunakan *gammatone-filter* orde-4 sesuai dengan persamaan 2.5 dengan *center frequency* sebanyak 128 yang terdistribusi secara merata dalam skala ERB 80 Hz hingga 5000 Hz (Roman,2003). Tahap kedua *middle ear gain* yaitu memberikan *gain* untuk mensimulasikan efek *middle-ear* berdasarkan fungsi *equal-loudness* dari BS3383, *Normal Equal-Loudness Level*. Dan yang terakhir adalah *hair cell* yaitumemodelkan aktifitas dari *auditory-nerve* yang berupa *firing-probability* pada telinga bagian dalam (Meddis,1986).

* + 1. **Perhitungan Binaural Cue**

Perhitungan *binaural cue* untuk *interaural time difference* (ITD) menggunakan *cross-correlogram* dari persamaan 2.3. *Cross correlogram* dilakukan antara *firing-probability* dari sinyal masker pada telinga kanan dan kiri pada setiap kanal frekuensinya dengan nilai *lag* yang terdistribusi antara -1 ms dan 1 ms dan panjang *window* sebesar 20 ms dengan nilai *shifting* sebesar 5 ms. Selanjutnya, pengambilan *binaural cue* untuk *interaural level difference* (ILD) yang dihitung antara *firing-probability* dari telinga kiri dan telinga kanan pada seluruh kanal frekuensi. Keluaran dari masing-masing *binaural cue* merupakan unit *time-frequency.*

* + 1. ***Relative Strength.***

Perhitungan nilai *Relative Strength* dari sinyal target terhadap suara tercampur dilakukan dalam unit *time-frequency* dengan ukuran yang sama terhadap *binaural cue*. Nilai *relative-strength* dari sinyal target dapat dihitung sebagai ratio antara *firing probability* dari sinyal target terhadap sinyal tercampur dengan menggunakan persamaan.

(3.1)

Dimana R adalah *relative strength* pada *center frequency* i, s adalah sinyal suara target dan n adalah sinyal suara masker

* + 1. ***Probability Density Estimation***

Pada tahap ini dilakukan analisa pola antara kedua *binaural cue* yaitu *interaural time difference* (ITD) dan *interaural level difference* (ILD) terhadap *relative strength* dari sinyal target terhadap suara tercampur dalam unit *time-frequency*. Analisa pola yang dilakukan menggunakan *Kernel Density Estimation* untuk memperoleh nilai *probability density* ) yang merupakan hubungan statistik antara ketiga nilai tersebut (*interaural time difference, interaural level difference* dan *relative strength)*.

Perhitungan *probability density* dilakukan pada dua jenis *relative strength,* yaitu nilai *probability density* saat sinyal target lebih dominan daripada sinyal masker () dan sinyal masker yang lebih dominan daripada sinyal target ().pada unit T-F dengan persamaan berikut.

(3.2)

Dimana adalah ratio dari banyaknya nilai terhadap nilai ITD dan ILD, sedangkan adalah ratio dari banyaknya nilai terhadap nilai ITD dan ILD.

* 1. **Proses Pemisahan**

Pada proses pemisahan, langkah yang dilakukan pada tahap simulasi sinyal *binaural*, *auditory periphery* dan perhitungan *binaural cue* adalah sama seperti pada subbab 3.1.1, 3.1.2 dan 3.1.3 tetapi sinyal keluaran dari auditory periphery merupakan sinyal suara tercampur antara sinyal target dan sinyal masker. Sinyal suara target yang digunakan pada proses pemisahan ini sebanyak 120 kalimat dengan sinyal suara masker yang selalu sama/tetap. Pada Tabel 3.2, proses pemisahan yang pertama dilakukan adalah pada titik 1 untuk sinyal target dan titik 5 untuk sinyal masker dengan nilai SIR 10 dB, 5 dB dan 0 dB. yang. Setelah suara target dan masker berhasil dipisah, maka akan dilakukan pemisahan untuk sinyal masker yang posisinya semakin mendekati sinyal target yaitu titik 4, titik 3 dan titik 2. Sehingga total percobaan adalah 12 kali dengan masing-masing percobaan sebanyak 120 kalimat.

**Tabel 3.2** Konfigurasi Sinyal Target dan Masker pada Proses Pemisahan

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sinyal Target | Azimut dan Elevasi | Sinyal Masker | Azimut dan Elevasi | *Signal to Interference Ratio* (SIR) |
| 120 Kalimat yang berbeda  (Suara Percakapan Laki-Laki) | (0o, 0o) | 1 Kalimat (Suara Percakapan Perempuan) | (30o, 0o)  (20o, 0o)  (10o, 0o)  (5o, 0o) | 10 dB  5 dB  0 dB |

* + 1. **Perhitungan Azimut**

Pada tahap ini, dilakukan perhitungan azimut untuk mengetahui posisi sinyal target dan sinyal masker. Dengan diperolehnya azimut sinyal target dan sinyal masker, selanjutnya sistem akan mengambil database nilai *probability density* dari *data training* dengan azimut yang sesuai. Nilai *probability density* ini kemudian digunakan untuk mengestimasi *binary mask* dalam proses pemisahan suara target.

* + 1. ***Binary Mask***

*Binaural cue* yang diperoleh pada proses pemisahan kemudian dilakukan pencocokan terhadap nilai *binaural cue* pada *probability density* dari *data training* digunakan untuk memperoleh nilai *relative strength.* Nilai *relative strength* ini kemudian digunakan untuk menentukan *binary mask* sesuai persamaan berikut.

(3.3)

**3.5 Evaluasi Subjektif**

Metode subjektif yang dipakai adalah *Speech Intelligibility (Percent Correct Word)*. Dalam metode ini, pengambilan data dilakukan dengan menggunakan manusia sebagai naracoba. Naracoba yang digunakan adalah mahasiswa Teknik Fisika ITS yang berusia rata-rata 22 tahun, diasumsikan memiliki pendengaran normal karena tidak ada catatan medis perlu ataupun pernah menggunakan "*hearing impairment*". Naracoba bersedia untuk memberikan penilaian sesuai pendengaran dan presepsi mereka dengan mengikuti prosedur yang ditetapkan.



**Gambar 3.4** Naracoba Sedang Melakukan Tes Subjektif

Pengambilan data dilakukan dengan menggunakan *headphone* HD580 yang terhubung ke komputer. Setiap naracoba akan mengevaluasi 120 kalimat yang memiliki 12 jenis suara hasil pemisahan dengan konfigurasi yang berbeda-beda sesuai dengan yang tertulis di BAB 3.

**3.5 Evaluasi Objektif**

Metode objektif yang dipakai adalah *signal to noise ratio* (SNR) dengan menggunakan persamaan (2.6). Perhitungan SNR dilakukan pada 120 kalimat kemudian dirata-rata pada setiap konfigurasi yang dilakukan sesuai dengan yang tertulis di BAB 3.1.