

ESTIMASI ARAH SUMBER SUARA BERBASIS GAUSSIAN MIXTURE MODEL**SOUND SOURCE DIRECTION ESTIMATION BASED ON GAUSSIAN MIXTURE MODEL****Irma Safitri**Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro,
Universitas Telkom
irmasaf@telkomuniversity.ac.id**Abstrak**

Estimasi arah sumber suara menjadi topik penting yang berhubungan dengan aplikasi robot, sistem sensor dan keamanan. Variasi kondisi eksperimen dalam melakukan estimasi tersebut akan menentukan nilai akurasi. Dalam penelitian ini, variasi terhadap temperatur dan waktu pantul diambil untuk dianalisa terhadap nilai akurasi estimasi arah sumber suara. Sinyal yang digunakan adalah sinyal *binaural* dengan menggunakan sinyal pengganggu *white noise* dan *human speech like (HSL) noise* untuk sudut *azimuth* bervariasi. Estimasi dilakukan dengan menggunakan metode *Gaussian Mixture Model (GMM)* untuk tipe *horizontal plane* dan *horizontal – vertical planes*. Hasil eksperimen menunjukkan sudut *azimuth* yang dekat dengan pendengar akan menyampaikan sinyal suara lebih cepat daripada sudut yang jauh, sinyal dengan durasi waktu yang panjang yaitu 2000 milidetik akan memberikan akurasi estimasi yang lebih tinggi daripada durasi sinyal yang lebih pendek: 100, 500, dan 1000 milidetik. Selain itu, akurasi estimasi lebih tinggi untuk suara dengan *white noise* daripada suara dengan *HSL noise*. Hasil lainnya adalah estimasi memiliki performansi lebih tinggi untuk *horizontal – vertical planes* daripada hanya kondisi *horizontal plane*. Estimasi mencapai 98,6% akurasi untuk *horizontal plane* dan 100% akurasi untuk *horizontal-vertical planes*.

Kata kunci : HRTF, GMM, estimasi arah sumber suara**Abstract**

Sound source direction estimation has become an attractive issue related to the robotic applications, sensoric systems, and security. Experimental variation factors in estimating sound source direction will determine the estimation accuracy. In this research, we have two variation factors: temperature and reverberation time to be analyzed for its sound source estimation accuracy. Binaural signal is used with disturbances of white noise and human speech like (HSL) noise for changing azimuth angles. Estimation is done by using Gaussian Mixture Model (GMM) method for horizontal plane and horizontal – vertical planes. Experimental results show that nearest azimuth angles to the listener will transport sound signal faster than those other angles, signal with longer duration, 2000 ms, will perform higher estimation accuracy compared to those other shorter durations: 100, 500, and 1000 ms. Besides that, estimation accuracy is higher for sound signal having white noise compared to those with HSL noise. Another experiment result is higher performance in estimation for horizontal – vertical planes compared to horizontal plane only. Estimation can have 98.6% accuracy for horizontal plane only and 100% accuracy for horizontal – vertical planes.

Keywords: HRTF, GMM, sound source direction estimation

1. PENDAHULUAN

Sistem pendengaran *binaural* manusia memiliki kemampuan untuk membentuk dunia yang berhubungan dengan *aural*, pendengaran. Pendengaran manusia memiliki kemampuan yang mencengangkan untuk mendeteksi, mengklasifikasi, dan mengidentifikasi arah sumber suara. Suara dapat dipersepsikan berada di atas, bawah, belakang, dan sisi kiri atau kanan dari pendengar. Deteksi arah sumber suara merupakan teknik yang sangat penting dan banyak digunakan secara luas, seperti di dalam bidang: *speech enhancement*, perekaman suara, dan sistem keamanan. Beberapa tahun terakhir, banyak penelitian berdasarkan metode *microphone arrays* [1 – 3] yang menjabarkan metode yang menggunakan beberapa mikrofon untuk mendapatkan performansi tinggi dalam deteksi arah suara. Namun, jumlah mikrofon tersebut dapat dikurangi untuk menghemat biaya penelitian dan pemeliharaan.

Untuk metode estimasi, *Gaussian Mixture Model* (GMM) dapat digunakan. GMM adalah fungsi densitas probabilitas secara parametrik yang direpresentasikan dalam bobot jumlah komponen densitas *Gaussian*. GMM biasanya digunakan sebagai model parametrik dari distribusi probabilitas ciri atau pengukuran kontinu terutama dalam ciri spektral sistem deteksi dan rekognisi *speaker* [4]. GMM telah menjadi standard estimasi disebabkan performansinya yang lebih baik dibandingkan dengan metode estimasi lainnya, seperti: *multi layer perceptrons* [5], *factor analysis* [6], dan *support vector machine* [7].

Arah sumber suara dapat dideteksi dengan sinyal *binaural* yang terdiri dari *interaural time difference* (ITD) dan *interaural level difference* (ILD). Sinyal *biaural* didapatkan dengan cara mengkonvolusi sinyal sumber suara dan respon impuls sinyal *biaural* ruangan (BRIR – *binaural room impulse response*). Sinyal *binaural* ini mengandung *head related transfer function* (HRTF) dan respon impuls ruangan yang merepresentasikan kondisi akustik. Kita dapat menentukan arah suara dengan dua komponen ILD dan ITD walaupun lingkungan berubah [8 – 10].

Deteksi arah sumber suara dengan kondisi lingkungan yang bervariasi menjadi topik sangat menarik untuk diteliti. Referensi [11] mendeteksi arah speaker yang mengeluarkan suara *binaural* untuk *humanoid robots*, dimana estimasi arah sumber suara menggunakan GMM dan ekstraksi ciri *Mel-Frequency Cepstral Coefficient* (MFCC) dalam kondisi lingkungan dengan *noise*. Referensi [12] mendeteksi arah sumber-sumber suara secara dua dimensi yang mengeluarkan suara secara bersamaan. Metode yang digunakan adalah variasi algoritma *expectation maximization* (EM), pemisahan suara dan deteksi arah sumber suara. Namun, penelitian yang berhubungan dengan estimasi arah sumber suara berdasarkan variasi temperatur ruang dan waktu pantul belum tergal. Untuk itu, penelitian ini mengangkat topik estimasi arah sumber suara menggunakan sinyal *binaural* berdasarkan GMM. Eksperimen yang dilakukan untuk mengobservasi BRIR yang berubah berdasarkan lama waktu datangnya suara berdasarkan sudut kedatangan suara, performansi model dalam dua situasi yaitu *horizontal plane* dan *horizontal-vertical planes*. Untuk *horizontal plane*, eksperimen dilakukan berdasarkan variasi temperatur dan waktu pantul. Sedangkan untuk *horizontal – vertical planes*, eksperimen dilakukan berdasarkan variasi waktu pantul saja.

Pengorganisasian jurnal ini adalah sebagai berikut. Bab 1 menjelaskan pentingnya deteksi arah sumber suara dan penelitian sebelumnya yang terkait. Bab 2 menerangkan metodologi penelitian yang terdiri dari parameter ciri, metode estimasi menggunakan distribusi Gaussian dan pengukuran BRIR dimana kondisi pengukuran dituliskan dengan detail. Bab 3 mengenai hasil dan pembahasan dari Bab 2. Di akhir, kesimpulan dituliskan di Bab 4.

2. METODOLOGI

2.1 Parameter Ciri

Metode evaluasi yang digunakan adalah *ILD envelope*. Untuk mendapatkan *ILD envelope*, parameter *cepstrum* dihitung dengan menggunakan prosedur berikut. Pertama – tama, penulis menggunakan $x_R(t)$ yang menunjukkan sinyal yang datang di telinga kanan dan $x_L(t)$ yang menunjukkan sinyal yang datang di telinga kiri. *Hamming window* digunakan di kedua sinyal tersebut dengan panjang *frame* l dan pergeseran *frame* l_s . Kemudian dengan menggunakan persamaan (1), *ILD* dihitung.

$$|X_{LR}(f)| = \frac{|X_L(f)|}{|X_R(f)|} \tag{1}$$

dimana f merupakan frekuensi, $X_R(f)$ adalah respon magnituda dari sinyal yang diterima oleh telinga kanan, dan $X_L(f)$ adalah respon magnituda dari sinyal yang diterima oleh telinga kiri.

Langkah selanjutnya adalah menggunakan transformasi Fourier pada logaritma *ILD* untuk mendapatkan parameter ciri. Persamaan (2) digunakan untuk mendapatkan parameter ciri yang dinotasikan dengan $c[n]$. Dengan orde *cepstrum* *ILD* yang lebih rendah, *ILD envelope* dapat dihitung. Orde *cepstrum* adalah $N=15$. Distribusi dari *cepstrum* *ILD* orde 15 didapat dengan menggunakan distribusi Gaussian.

$$c[n] = \frac{1}{N} \sum_{k=0}^{N-1} 10 \log_{10} |X_{LR}(f)| e^{j2\pi nk/N}$$

$$n = 0, 1, 2, \dots, N \tag{2}$$

2.2 Estimasi Arah Sumber Suara dengan Distribusi Gaussian

Gaussian mixture model (GMM) adalah model statistik yang merepresentasikan kombinasi linier dari distribusi Gaussian. Metode ini sering digunakan untuk *speaker recognition*, *speech recognition*, dan sebagainya.

Dalam jurnal ini, dengan menggunakan sinyal *binaural*, model Gaussian untuk setiap arah suara digunakan untuk data latih dan data uji. Prosedur untuk data latih dijelaskan berikut. Penulis menggunakan distribusi Gaussian sebagai model statistik untuk mendapatkan distribusi dari *cepstrum* *ILD*. Model statistik tersebut untuk setiap arah dituliskan dalam persamaan (3).

$$\lambda_\theta = \{ \mu_m, w_m, \Sigma_m \}$$

$$m = 1, 2, \dots, M \tag{3}$$

Algoritma maksimal ekspektasi memiliki rata-rata μ_m untuk setiap distribusi, bobot w_m dan matriks *covariance* Σ_m . Kemudian model estimasi λ_θ digunakan untuk data latih untuk setiap arah sumber suara. Metode yang digunakan adalah matriks *diagonal covariance*.

Setelah prosedur data latih dilakukan, dilanjutkan dengan estimasi arah sumber suara dengan menggunakan model Gaussian. Prosedurnya adalah sebagai berikut. *Cepstrum* *ILD* $c[n]$ dari sinyal masukan dihitung dengan menggunakan persamaan (2). Kemudian *posterior probability*

dihitung antara *cepstrum* ILD dari sinyal input dan setiap data latih model Gaussian di atas. Arah sumber suara ditentukan berdasarkan arah model yang menunjukkan probabilitas yang maksimum.

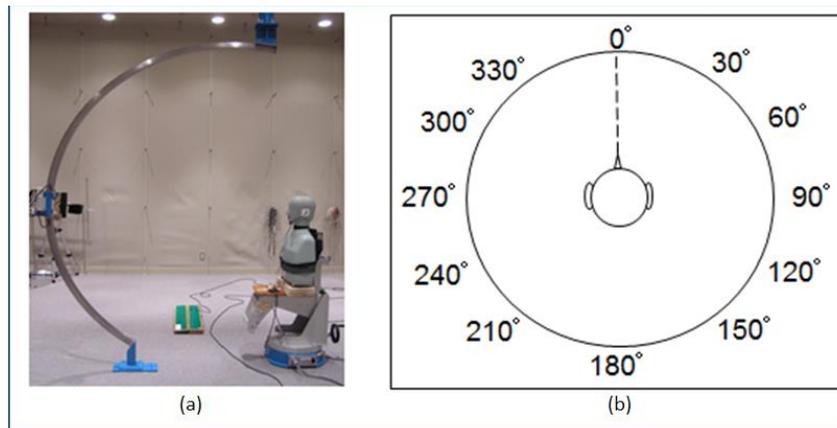
Estimasi arah sumber suara dengan variasi temperatur, tidak terlepas dari faktor sudut kedatangan suara dan temperatur dalam eksperimen nanti. Kecepatan kedatangan suara menggunakan persamaan fisika mengenai cepat rambat bunyi dituliskan pada persamaan (4) berikut.

$$c = \sqrt{\frac{\gamma \cdot R \cdot T}{M}} = \sqrt{\gamma \cdot \frac{p}{\rho}} \quad (4)$$

dimana c adalah kecepatan suara, γ adalah konstanta Boltzmann, R adalah konstanta gas molar, T adalah temperatur absolut, M adalah massa molar gas, p adalah tekanan, ρ adalah massa jenis.

2.3 Pengukuran BRIR

Sinyal *binaural* didapatkan dengan mengkonvolusi BRIR dengan data uji. BRIR diukur dengan *loudspeaker* (BOSE *Acoustimass*) dalam dua ruangan pengukuran menggunakan *head-and-torso simulator* (HATS, B&K 4128). Simulator dan kondisi eksperimen terlihat pada Gambar 1. Untuk mengukur respon impuls ruangan, penulis menggunakan sinyal *swept sine* dengan durasi 1,365 detik. Parameter akurasi deteksi lokasi sumber suara diukur untuk dua kondisi ruangan, yaitu berdasarkan variasi waktu pantul suara dan temperatur. Suara dapat terus dipantulkan dalam ruangan tertutup pada periode tertentu setelah suara tersebut keluar dari sumbernya. Waktu pantul suara untuk 60 dB didefinisikan sebagai waktu yang dibutuhkan untuk suara dalam ruangan berkurang sebanyak 60 dB setelah sumber berhenti mengeluarkan suara. Parameter ini adalah penentu utama dari lingkungan dengan karakteristik akustik. Waktu pantul optimal sangat bergantung pada jenis udara yang digunakan dalam eksperimen.



Gambar 1. Kondisi Eksperimen: (a) Simulator HATS pada *turntable* dan *arched traverse*, (b) Sudut *Azimuth* dari HATS secara *horizontal plane*

Waktu pantul pada ruangan eksperimen pertama dapat diatur dari 151 ke 459 milidetik dengan cara mengeluarkan panel penyerap suara di dinding. Eksperimen dilakukan dalam delapan kondisi waktu pantul seperti yang dicantumkan pada Tabel 1. HATS ditempatkan pada *turntable*, penyangga yang dapat berputar, dan *dodecahedral loudspeaker* ditempatkan pada *arched traverse*. Kedua benda ini dapat digerakkan dengan interval 1° , dengan akurasi 0.3° . Jarak antara *dodecahedral loudspeaker* dan pusat dari *bitrignon* adalah 1.2 m. BRIR diukur dengan 72 *azimuth*

dan 29 elevasi pada tiap kondisi waktu pantul. Total jumlah titik eksperimen adalah 16136 $((72 \times 28 + 1) \times 8)$ buah, dan dengan *sampling frequency* 48kHz.

Pada ruangan eksperimen ke dua, impuls respons diukur untuk tiga temperatur ruang yang berbeda, yaitu 18°C, 22°C, dan 26°C. Jika waktu pantul dan temperatur ruang berubah, waktu datang dari gelombang pantul juga akan berubah. Hal ini mengindikasikan bahwa BRIR bervariasi bergantung pada kedua parameter tersebut. Peralatan eksperimen lainnya sama dengan peralatan eksperimen yang berada di ruangan eksperimen pertama. Jarak antara *dodecahedral loudspeaker* dan pusat *listener* adalah 1,5 m dan BRIR diukur untuk 72 *azimuth* di *horizontal plane*. Kondisi eksperimen di ruangan ke dua ini dapat dilihat pada Tabel 2.

Tabel 1 Kondisi Pengukuran BRIR dengan Variasi Waktu Pantul

Level <i>background noise</i>	19.1 dB
Level tekanan udara	72.6 dB
<i>Azimuth</i>	0 – 355°,5° interval
Elevasi	0 – 355°,5° interval
Waktu Pantul	151, 179, 198, 224, 238, 261, 303, 459 milidetik

Tabel 2 Kondisi Pengukuran BRIR dengan Variasi Temperatur

Level <i>background noise</i>	13.2 dB
Temperatur ruang	18.0, 22.0, 26.0°C
<i>Azimuth</i>	0 – 355°,5° interval
Elevasi	0° (<i>horizontal plane</i>)
Waktu pantul	151 milidetik

Sudut *azimuth* yang diambil sebagai acuan adalah sebagai berikut: 0° untuk arah depan dari HATS, sudut negatif adalah HATS bagian kiri, sudut positif adalah bagian kanan dari HATS, dan 180° adalah sudut tepat arah belakang dari HATS. Sudut elevasi negatif menunjukkan bahwa posisi sumber suara berada di bawah *horizontal plane* dan sudut positif elevasi berada di atas *horizontal plane*. *Horizontal plane* berada di 0°.

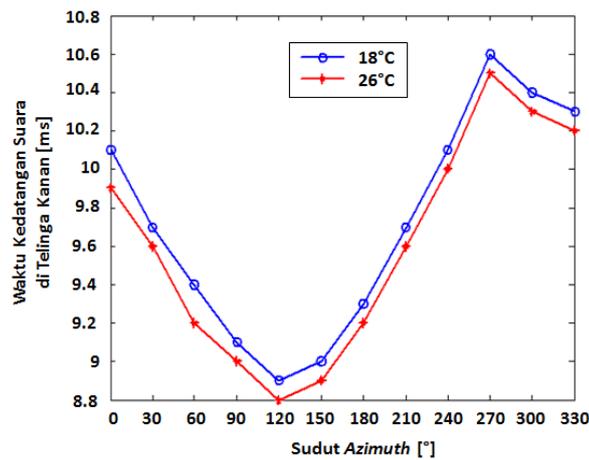
Di dalam eksperimen, *human-speech like (HSL) noise* dan *white noise* digunakan sebagai sinyal data uji. *Noise HSL* dibangkitkan dengan melakukan superposisi banyak sinyal *speech* dan karakteristik *noise HSL* ini dapat dikontrol dengan mengubah jumlah superposisi. Penulis menggunakan *noise HSL* dengan 24 superposisi dengan sinyal durasi 100, 500, 1000, dan 2000 milidetik. Durasi sinyal untuk data latih dan data uji adalah sama. Parameter yang digunakan dalam eksperimen ini adalah sebagai berikut: panjang *frame* 128, *frame shift* 32, dan jumlah ciri 1.

Eksperimen dilakukan untuk dua kondisi yaitu hanya *horizontal plane* dan kombinasi *horizontal-vertical planes*. Eksperimen di kondisi hanya *horizontal plane*, BRIR digunakan dengan waktu pantul dan temperatur ruang yang berbeda. Ada delapan kondisi untuk waktu pantul dan tiga kondisi untuk temperatur ruang. Untuk waktu pantul, tujuh kondisi digunakan sebagai data latih dan sisanya untuk data uji sedangkan untuk temperatur ruang, dua kondisi digunakan sebagai data latih dan sisanya sebagai data uji. Untuk eksperimen dengan kombinasi *horizontal-vertical planes*, BRIR digunakan dengan waktu pantul yang berbeda. Dalam eksperimen ini, BRIR diukur untuk *horizontal* dan *vertical planes* pada sudut kedatangan 30°.

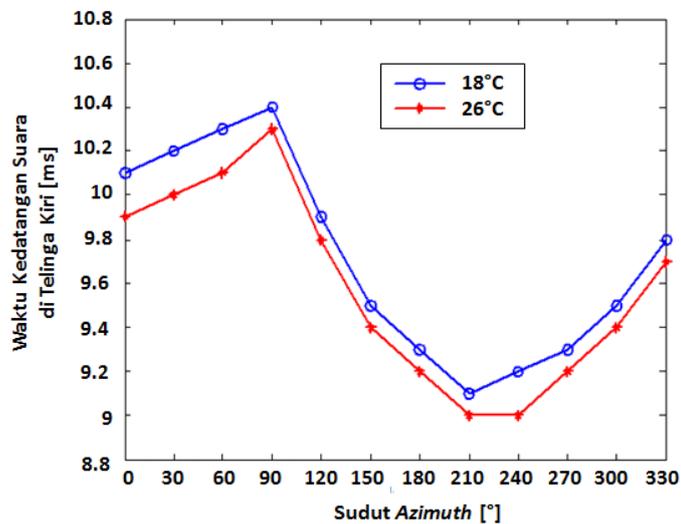
3. HASIL DAN PEMBAHASAN

Hasil eksperimen lamanya waktu datang suara dari sumber suara ke penerima dilakukan untuk berbagai sudut *azimuth* peletakan sumber suara. Untuk telinga kanan, suara diterima paling

cepat pada saat sumber suara terletak pada sudut 120° , sedangkan paling lama diterima pada saat sumber terletak pada sudut 270° . Hal ini diperlihatkan di Gambar 2. Sedangkan telinga kiri, suara paling cepat sampai saat sumber suara terletak 210° dan 240° , dan paling lama pada sudut 90° , seperti terlihat di Gambar 3. Hal ini sesuai dengan persamaan fisika mengenai cepat rambat bunyi di berbagai medium, seperti yang dituliskan di persamaan (4). Temperatur udara berbanding lurus dengan kecepatan suara, semakin tinggi suhu maka kecepatan suara akan meningkat sehingga waktu sampainya suara di telinga akan semakin cepat. Untuk telinga kanan, sumber suara yang diindikasikan terdekat berada pada sudut *azimuth* 90° , 120° , atau 150° , dan 120° menjadi sudut *azimuth* terdekatnya. Sedangkan untuk telinga kiri, sumber suara yang diindikasikan terdekat berada pada sudut *azimuth* 210° , 240° , atau 270° , dan 240° menjadi sudut *azimuth* terdekatnya.



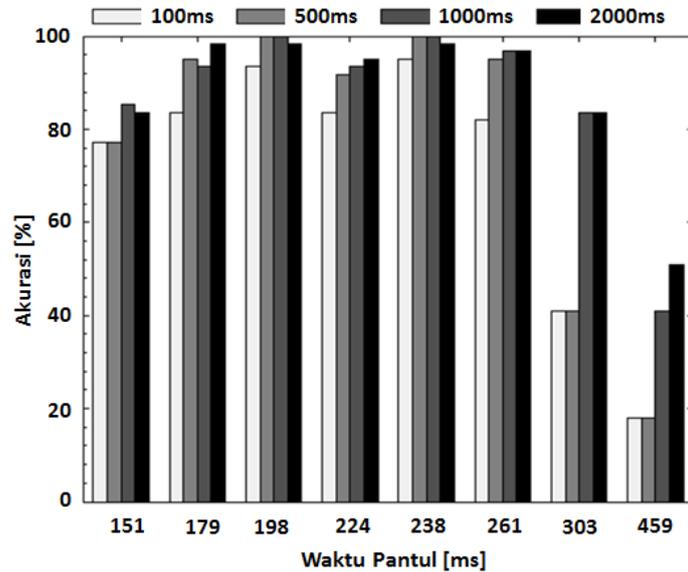
Gambar 2. Lamawaktu datang suara ke telinga kanan dari berbagai sudut *azimuth* untuk 18°C dan 26°C



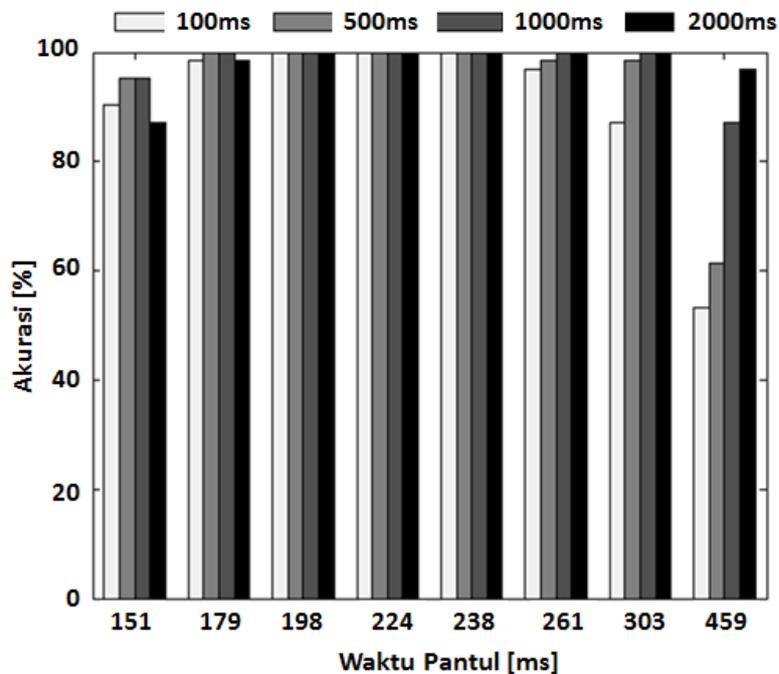
Gambar 3. Lamawaktu datang suara ke telinga kiri dari berbagai sudut *azimuth* untuk 18°C dan 26°C

Hasil estimasi pada *horizontal plane* menggunakan BRIR yang diukur pada ruang eksperimen pertama dapat dilihat di Gambar 4 dan Gambar 5. Untuk setiap variasi waktu pantul, sinyal dengan durasi yang panjang memiliki tingkat akurasi yang lebih tinggi untuk estimasi arah sumber suara daripada sinyal dengan durasi pendek. Waktu pantul optimal untuk mendapatkan akurasi tertinggi dalam eksperimen ini adalah pada range 179 sampai 261 milidetik dengan tingkat akurasi mencapai 100%. Hasil eksperimen pada Gambar 5 menunjukkan akurasi yang lebih tinggi daripada hasil di Gambar 4, disebabkan oleh tipe noise yang digunakan. Performansi estimasi

menggunakan HSL noise lebih rendah daripada white noise, dimana HSL noise [13] merupakan suara dari sumber lain yang mengganggu suara asli dan seringkali membuat performansi dalam deteksi speech menjadi berkurang, sedangkan white noise merupakan variabel random dengan distribusi normal.



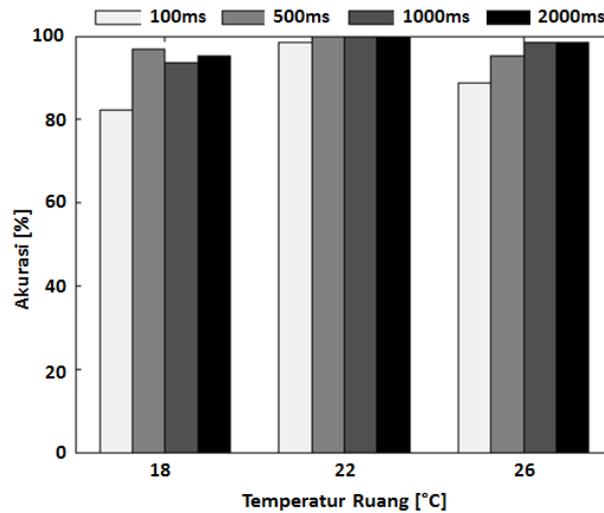
Gambar 4. Hasil Estimasi Arah Sumber Suara pada *Horizontal Plane* dengan Variasi Waktu Pantul Menggunakan *Human Speech-Like (HSL) Noise*



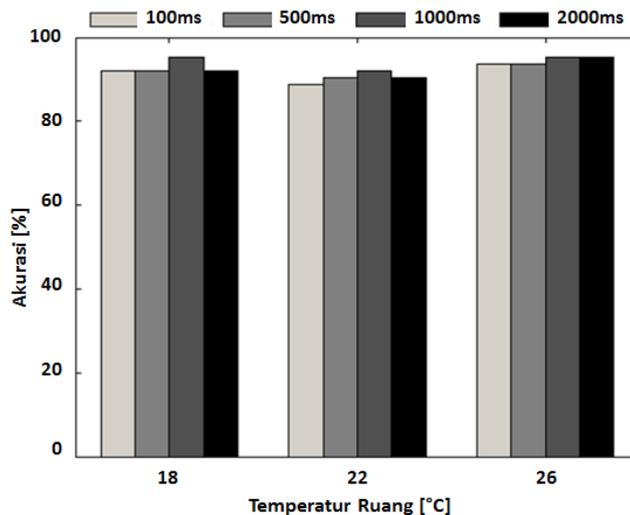
Gambar 5. Hasil Estimasi Arah Sumber Suara pada *Horizontal Plane* dengan Variasi Waktu Pantul Menggunakan *White Noise*

Gambar 6 dan Gambar 7 memberikan hasil estimasi pada *horizontal plane* dengan BRIR diukur di ruang eksperimen ke dua, berdasarkan variasi temperatur ruang. Semakin panjang durasi sinyal relatif memberikan akurasi estimasi yang semakin tinggi. Untuk HSL noise, temperatur 22°C menunjukkan akurasi estimasi yang terbaik, sementara untuk white noise, temperatur 26°C

memberikan akurasi estimasi yang terbaik.



Gambar 6. Hasil Estimasi Arah Sumber Suara pada *Horizontal Plane* dengan Variasi Temperatur Ruang Menggunakan *Human Speech-Like (HSL) Noise*

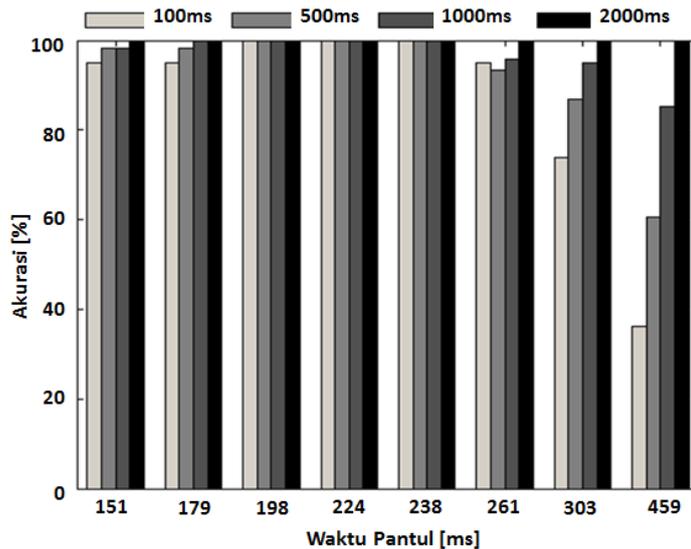


Gambar 7. Hasil Estimasi Arah Sumber Suara pada *Horizontal Plane* dengan Variasi Temperatur Ruang Menggunakan *White Noise*

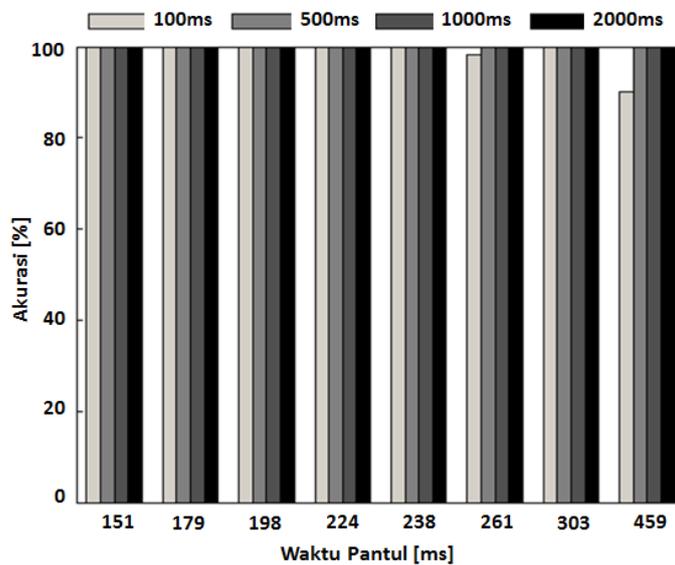
Gambar 8 dan Gambar 9 memperlihatkan hasil estimasi pada *horizontal* dan *vertical planes* dengan BRIR diukur dalam ruang eksperimen pertama, berdasarkan variasi dari waktu pantul. Semakin panjang durasi suara memberikan akurasi lebih tinggi dalam estimasi sumber suara dibandingkan dengan durasi suara yang lebih pendek, sementara *white noise* menunjukkan performansi lebih baik dalam akurasi estimasi dibandingkan dengan *HSL noise*. Waktu pantul optimal adalah antara 179 sampai 261 milidetik. Gambar 8 dan 9 ini memperlihatkan akurasi lebih tinggi dibandingkan dengan estimasi yang ditampilkan di Gambar 4 dan 5. Hal ini dikarenakan tipe *plane* yang digunakan berbeda. Gambar 4 dan 5 hanya melakukan estimasi arah sumber suara pada *horizontal plane*, sementara Gambar 8 dan 9 melakukan estimasi baik di arah *horizontal* dan *vertical planes*. Oleh karena itu, estimasi arah suara dengan kedua *horizontal* dan *vertical planes* memberi hasil performansi lebih baik daripada hanya pada *horizontal plane*. Hasil eksperimen mengindikasikan bahwa model distribusi Gaussian dapat mengestimasi arah sumber suara dan performansi lebih tinggi untuk waktu pantul didapat antara 179 sampai 261 milidetik dan pada

22.0°C. Selain itu, performansi tinggi didapat saat sinyal suara panjang, seperti untuk sinyal suara 1000 dan 2000 milidetik.

Berdasarkan semua hasil simulasi di atas, walaupun penulis tidak melakukan perbandingan metode estimasi yang lain, metode GMM ini terbukti dapat memberikan hasil akurasi estimasi yang tinggi. Hal ini sesuai dengan hasil penelitian yang lalu [5 – 7]. Dalam jurnal ini, performansi estimasi arah sumber suara dapat mencapai 98,6% akurasi untuk *horizontal plane* dan 100% akurasi untuk kombinasi *horizontal-vertical planes*.



Gambar 8. Hasil Estimasi Arah Sumber Suara pada *Horizontal* dan *Vertical Planes* dengan Variasi Waktu Pantul Menggunakan *Human Speech-Like (HSL) Noise*



Gambar 9. Hasil Estimasi Arah Sumber Suara pada *Horizontal* dan *Vertical Planes* dengan Variasi Waktu Pantul Menggunakan *White Noise*

4. KESIMPULAN

Kesimpulan dalam penelitian ini antara lain adalah:

1. Lama waktu datangnya suara di pendengar berbeda-beda tergantung dari posisi pendengar terhadap sudut *azimuth* sumber suara. Sumber dengan sudut *azimuth* terdekat

- akan memberikan waktu sampainya suara semakin cepat, waktu tercepat sampainya suara 8,8 milidetik dan yang terlamanya 10,6 milidetik.
2. Estimasi arah sumber suara pada *horizontal plane* dalam ruang eksperimen pertama bernilai akurasi semakin tinggi untuk durasi sinyal suara yang panjang dan performansi akurasi sinyal dengan *white noise* lebih baik daripada sinyal dengan *HSL noise*.
 3. Estimasi arah sumber suara pada *horizontal plane* dalam ruang eksperimen ke dua bernilai akurasi relatif lebih tinggi untuk durasi sinyal suara yang panjang daripada durasi sinyal yang pendek. Akurasi estimasi dengan *HSL noise* terbaik pada 22°C dan *white noise* pada 26°C.
 4. Pada ruang eksperimen pertama, dengan *horizontal – vertical planes*, estimasi arah sumber suara memberikan akurasi semakin tinggi untuk durasi sinyal yang semakin panjang. Performansi estimasi lebih baik untuk eksperimen menggunakan *white noise* daripada *HSL noise*.
 5. Tipe *plane* mempengaruhi tingkat akurasi estimasi. Eksperimen menggunakan *horizontal – vertical planes* akan memberikan akurasi estimasi arah sumber suara lebih tinggi daripada eskperimen hanya menggunakan *horizontal plane*.
 6. Estimasi memiliki performansi 98,6% akurasi untuk *horizontal plane* dan 100% akurasi untuk kombinasi *horizontal-vertical planes*.

Daftar Pustaka:

- [1] Thiergart O, Del Galdo G, Taseska M, dan Habets E.A.P. 2013. Geometry-Based Spatial Sound Acquisition Using Distributed Microphone Arrays. *IEEE Transactions on Audio, Speech, and Language Processing*. ISSN 1558-7916. **21:12** 2583-2594.
- [2] Chen H, Bao Y, dan Ser W. 2015. Effects of Sensor Position Errors on Farfield/Nearfield Wideband Beamformers for Microphone Arrays. *IEEE Sensors Journal*. ISSN 1530-437X. **15:9** 4812- 4825.
- [3] Pertila P, Hamalainen M. S., dan Mieskolainen M. 2013. Passive Temporal Offset Estimation of Multichannel Recordings of an Ad-Hoc Microphone Array. *IEEE Transactions on Audio, Speech, and Language Processing*. ISSN 1063-6676. **21:11** 2393-2402.
- [4] Bishop, C.. 2006. *Pattern Recognition and Machine Learning*. New York : Springer.
- [5] Cardinaux F, Sanderson C, dan Marcel S. 2003. Comparison of MLP and GMM Classifiers for Face Verification on XM2VTS. *International Conference Audio- and Video- Based Biometric Person Authentication (AVBPA)*. Springer Berlin Heidelberg. ISBN 978-3-540-40302-9.
- [6] Dehak N, Dehak R, Kenny P, dan Dumouchel P. 2008. Comparison Between Factor Analysis and GMM Support Machines for Speaker Verification. In *Proc. IEEE Odyssey: The Speaker and Language Recognition Workshop*. Institute of Electrical and Electronics Engineers (IEEE). ISBN 9781424404711.
- [7] Sen I, Saraclar M, dan Kahya Y. P. 2015. A Comparison of SVM and GMM-Based Classifier Configurations for Diagnostic Classification of Pulmonary Sounds. *IEEE Trans. Biomed. Engineering*. Institute of Electrical and Electronics Engineers (IEEE). ISSN 0018-9294. **62:7** 1768-1776.
- [8] Hongo, S, Sakamoto, S, dan Suzuki, Y. 2012. Binaural Speech Enhancement Method by Wavelet Transform Based on Interaural Level and Argument Differences. *International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4673-1534-0.

- [9] Morikawa, D. 2014. Effect of Interaural Difference for Localization of Spatially Segregated Sound. International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP). Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4799-5389-9.
- [10] Zohny, Z, dan Chambers, J. 2014. Modelling Interaural Level and Phase Cues with Student's T-Distribution for Robust Clustering in MESSL. International Conference on Digital Signal Processing (DSP). Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4799-4612-9.
- [11] Youssef K, Argentieri S, dan Zarader J. L. 2010. Binaural Speaker Recognition for Humanoid Robots. International Conference on Control, Automation, Robotics and Vision (ICARCV). Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4244-7814-9.
- [12] Deleforge A, Forbes F, dan Horaud R. 2015. Acoustic Space Learning for Sound-Source Separation and Localization on Binaural Manifolds. International Journal of Neural Systems, World Scientific Publishing. ISSN 0129-0657. **25:01**.
- [13] Okada K. 2012. Sound Category Detection Including the Human Speech like Noise after Independent Component Analysis and the Sound Trace. International Workshop on Information and Electronics Engineering (IWIEE). Procedia Engineering. Elsevier Ltd. ISBN 1877-7058.