



## Penentuan Sudut Datang Sumber Suara Menggunakan *Directional Microphone array*

Annisa Firasanti, Putra Wisnu Agung Sucipto, Mega Fatimah, M. Fadhil Dermawan  
Fakultas Teknik Universitas Islam "45" Bekasi, Jl. Cut Meutia 83 Bekasi Timur 17113  
\*E-mail korespondensi: adnis1210@gmail.com

**Abstrak.** Penentuan lokasi sumber akustik adalah teknik yang sangat penting yang dibutuhkan di banyak bidang aplikasi praktis seperti telekonferensi, komunikasi dengan mesin manusia, keamanan, navigasi, dan lainnya. *Microphone array* adalah tiruan dan pengembangan dari proses pendengaran manusia dengan dua telinga. Dalam penelitian ini akan dilakukan proses penentuan lokasi sumber suara dengan salah satu metode yang umum digunakan dalam bidang ini yaitu mengukur *Time Difference of Arrival (TDOA)*, dimana akan dilakukan *cross-correlation* dari informasi yang diterima oleh sepasang atau beberapa mikrofon. Setelah itu akan dilakukan estimasi untuk mencari nilai delay yang dapat memaksimalkan *cross-correlation* itu sendiri. Nilai tersebut kemudian akan diproses dari persamaan geometri yang dapat menghasilkan sudut sumber suara yang dihitung dari mikrofon acuan. Hasil penelitian sementara didapatkan bahwa suara yang mempunyai rentang frekuensi yang tinggi lebih mudah dideteksi daripada yang mempunyai frekuensi tunggal. Adapun sudut yang terdeteksi mempunyai error rata-rata  $2,78^\circ$ .

**Kata kunci:** Penentuan lokasi sumber suara, *microphone array*, *time difference of arrival (TDOA)*

### PENDAHULUAN

Dalam penerapan teknologi saat ini fitur penentuan lokasi suara telah banyak diteliti, dikembangkan dan diterapkan. Sensor pendengaran, khususnya pada robot atau peralatan elektronik lainnya, mempunyai kelebihan yaitu sumber suara tidak harus berada dalam satu garis lurus dengan alat. Kelebihan tersebut dapat dengan melengkapi kemampuan kamera (*vision*) dengan baik dalam rangka mencari lokasi seseorang atau kejadian menarik lainnya dalam sebuah ruangan. Penentuan lokasi sumber dengan menggunakan sensor akustik (contohnya *mikrofon* dan *hydrophone*) telah diaplikasikan ke dalam banyak bidang. Dalam pemrosesan sinyal sonar, yang menjadi fokus adalah bagaimana menemukan lokasi sumber akustik bawah air dengan menggunakan *hydrophone* baik tunggal ataupun *array* (Nosal, 2013). Pada *video conference* untuk keperluan *e-learning* maupun *distributed meeting*, *microphone array* telah dikembangkan sedemikian rupa untuk dapat menemukan posisi pembicara yang berada di dalam sebuah ruangan (Zhang et al, 2008). Pencarian sinyal akustik juga telah digunakan untuk mengestimasi lokasi kendaraan pada sebuah jaringan sensor nirkabel (Sheng & Hu, 2005) dan juga di-*install* pada sebuah robot untuk mencari lokasi operator (manusia) di dalam sebuah bangunan (Mumolo et al, 2002).

Dalam penelitian ini akan dilakukan penerapan *TDOA (Time Difference of Arrival)* yang didapatkan dari *microphone array*. Adapun sensor yang digunakan adalah *directional mikrofon* karena mempunyai sensitivitas yang tinggi terhadap suara pada arah tertentu saja, berbeda dengan *omni-directional mikrofon* yang mempunyai sensitivitas yang sama ke semua arah. Fitur ini tentunya merupakan sebuah nilai tambah untuk keperluan penentuan lokasi. Namun karena memiliki sensitivitas yang tinggi, maka suara-suara yang bukan berasal dari sumber (derau dan gema) juga akan ikut tertangkap dengan jelas. Oleh karena itu dibutuhkan sebuah proses untuk menghilangkan derau tersebut sebelum data akustik digunakan untuk kepentingan lokalisasi. Adapun tujuan dari penelitian ini adalah untuk mencari sudut dari sumber suara yang berasal dari suara terekam.

### METODE PENELITIAN

#### Metode Pencarian Lokasi Sumber Suara *Time Difference of Arrival (TDOA)*

Penentuan lokasi sumber suara ditentukan dari informasi tentang waktu tunda antar pasangan mikrofon. Estimasi lokasi dari waktu tunda sudah pasti merupakan masalah non linier. Namun dengan menggunakan beberapa pendekatan, akan dapat dicari solusi yang sederhana untuk memecahkan masalah ini.

Misalnya ada *windowed frame* dari  $N$  sampel data dengan 50% *overlap*. Untuk menentukan *delay* dari sinyal yang tertangkap oleh dua mikrofon yang berbedadiperlukan untuk menetapkan pengukuran yang koheren dan jenis pengukuran koheren yang paling umum adalah korelasi silang (*cross-correlation*) sederhana antara sinyal yang diterima oleh dua mikrofon seperti pada (1),

$$R_{ij}(\tau) = \sum_{n=0}^{N-1} x_i[n]x_j[n - \tau] \quad (1)$$

dimana  $x_i[n]$  adalah sinyal yang diterima oleh mikrofon  $i$  dan  $\tau$  adalah *correlation lag* pada sampel.  $R_{ij}(\tau)$  akan bernilai maksimal apabila  $\tau$  sama dengan *offset* antara dua sinyal yang diterima (Gay & Benesty, 2000). Permasalahan yang terdapat pada (1) adalah kompleksitasnya yang tinggi. Solusinya adalah dengan menghitung pendekatan dalam domain frekuensi dengan menghitung invers transformasi Fourier dari *cross-spectrum* untuk mengurangi kompleksitas perhitungan seperti yang terdapat pada (2),

$$R_{ij}(\tau) \approx \sum_{n=0}^{N-1} X_i(k)X_j(k)^* e^{t2\pi k\tau/N} \quad (2)$$

dimana  $X_i(k)$  adalah transformasi diskrit Fourier dari  $x_i[n]$  dan  $X_i(k)X_j(k)^*$  adalah *cross-spectrum* dari  $x_i[n]$  dan  $x_j[n]$ .

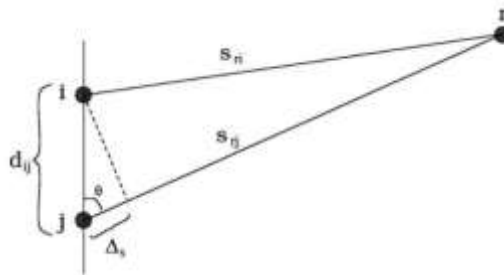
Permasalahan dalam pendekatan ini adalah bahwa korelasi tersebut sangat bergantung pada sifat statistik dari sinyal sumber, Karena kebanyakan sinyal suara mempunyai frekuensi rendah, korelasi antara sampel yang berdekatan akan bernilai tinggi sehingga akan menghasilkan *peak* yang lebar pada *cross-correlation*. Hal ini dapat diatasi dengan melakukan proses *whitening* pada spectrum sinyal seperti pada (3) yang sesuai dengan invers dari transformasi Fourier dari *cross spectrum* yang telah di normalisasi (*whitened*) (Valin et all, 2003).

$$R_{ij}^{(w)}(\tau) \approx \sum_{k=0}^{N-1} \frac{X_i(k)X_j(k)^*}{|X_i(k)||X_j(k)|} e^{t2\pi k\tau/N} \quad (2)$$

SSL (Sound Source Localization)

Misalnya ada sebuah susunan sederhana yang terdiri dari sumber akustik dan rangkaian mikrofon seperti yang digambarkan pada Gambar 1. Jika sumber akustik mengeluarkan sinyal suara  $r(t)$ , maka sinyal yang diterima oleh mikrofon ke- $i$  dapat dinyatakan dalam model matematis (4) dimana  $\alpha_i$  adalah faktor pelemahan akibat propagasi sinyal akustik yang berjalan dari sumber suara ke mikrofon,  $\tau_i$  adalah waktu propagasi dan  $n_i(t)$  merupakan gabungan dari semua derau yang ada di dalam ruangan termasuk gaung dan gema.

$$s_i(t) = \alpha_i r(t - \tau_i) + n_i(t) \quad (4)$$



**Gambar 1.** Susunan geometri untuk persoalan penentuan arah

Kemudian anggaplah  $d_{ij}$  adalah jarak antara pasangan mikrofon  $i$  dan  $j$ , maka  $\delta_{ij} = \tau_j - \tau_i$  adalah beda waktu kedatangan sinyal akustik antara kedua mikrofon. Dibawah asumsi penyederhanaan—pada dasarnya berasal dari teori *far-field* (dimana jarak sumber suara diasumsikan jauh lebih besar daripada jarak *microphone array*), yaitu  $S_{rj}, S_{ri} \gg d_{ij} \gg \lambda$  dimana  $S_{ri}, S_{rj}$  adalah jarak antara sumber  $r$  dengan mikrofon  $i$  dan  $j$ ,  $\lambda$  adalah panjang gelombang dominan dari sinyal—maka dapat ditentukan sudut  $\Psi$  kedatangan sumber akustik dari waktu tunda.

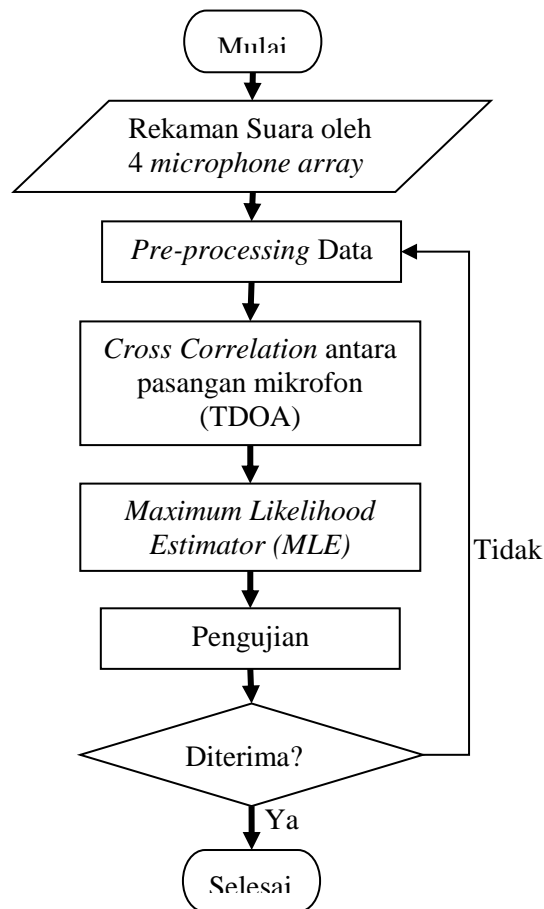
$$\Delta_s = d_{ij} \cos(\theta) \cong S_{rj} - S_{ri} = v(\tau_j - \tau_i) = v\delta_{ij} \quad (5)$$

Hal ini dijelaskan pada Gambar 1, dimana  $v$  adalah kecepatan suara, sehingga sudut  $\theta$  sangat mungkin untuk diestimasi. Dengan menggunakan persamaan geometris maka akan memungkinkan untuk memperkirakan koordinat  $(x, y)$  dari sumber akustik (Mumolo et all, 2003).

### Tahapan Penelitian

Bahan dan peralatan yang digunakan dalam penelitian ini adalah *Software Scilab*, *Software Microsoft Office 2007* dan file suara dengan format mono berekstensi *.wav*. Sebelum dilakukan proses deteksi lokasi, dilakukan pra pengolahan file suara dengan cara menghilangkan derau (*noise*) agar hasil suara yang didapat lebih baik. Tahapan penentuan lokasi sumber yang akan dilakukan dalam penelitian ini terdapat pada Gambar 2.

Data yang digunakan adalah rekaman suara dari mikrofon dengan format *.wav*. Sinyal suara yang diterima oleh mikrofon telah merupakan penggabungan dari sinyal asli yang *ter-delay* karena proses propagasi dan juga semua derau dan gaung yang ada di dalam ruangan tersebut. Oleh karena itu, perlu dilakukan *pre-processing* data sebelum data benar-benar digunakan yaitu dengan mengurangi atau bahkan menghilangkan *derau* tersebut. Pemodelan *derau* yang tepat sangat penting dalam proses ini. Idenya adalah membuat filter untuk mengurangi suara-suara lain yang bukan merupakan suara asli.



**Gambar 2.** Tahapan Penentuan Lokasi Sumber Suara

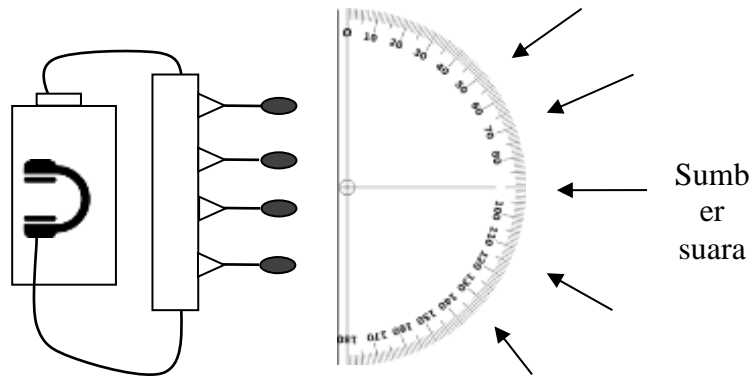
Kemudian setelah didapatkan suara yang bebas derau, dilakukan proses TDOA yaitu *cross-correlation* antar dua mikrofon. Proses ini biasanya dilakukan dengan menggunakan bantuan transformasi Fourier untuk mengurangi kompleksitas perhitungan. Setelah TDOA dilakukan, maka akan dicari nilai *delay* yang akan memaksimalkan nilai *cross-correlation* ini. Nilai tersebut kemudian diproses dengan rumus geometri sehingga dihasilkan sudut sumber suara. Hasil tersebut kemudian dibandingkan dengan sudut sumber yang asli pada saat percobaan dilakukan.

#### Metode Pengambilan Data

Pengambilan data dilakukan di 2 tempat yaitu di studio musik berukuran 8 m x 4 m, dan di ruangan kelas di Fakultas Teknik Universitas Islam Bekasi dengan lebar kelas sama dengan studio. Hal ini dilakukan dengan tujuan membandingkan antara rekaman suara yang dilakukan di tempat kedap suara dan yang tidak. Adapun peralatan yang digunakan adalah:

- 4 buah *directional microphone* Boya MV-01
- 1 buah *soundcard* Presonus FP 10
- 4 buah *converter* Rode VXL
- 1 buah *studio headphones* Samson SR850
- Software* Cubase 7
- Laptop
- Express card*
- Busur derajat (dalam bentuk spanduk)

Alat-alat tersebut dipasang dengan konfigurasi seperti pada Gambar 3.



Gambar 3. Konfigurasi pemasangan alat rekaman

### HASIL DAN PEMBAHASAN

Data-data suara yang direkam adalah data suara jentikan jari, hentakan kaki, tepuk tangan, kunci terjatuh dan sebuah *tone* (nada) dengan frekuensi 440 Hz berdurasi 3 detik yang dihasilkan oleh aplikasi android. Empat jenis suara tersebut (jentikan jari, hentakan kaki, tepuk tangan, kunci terjatuh) dipilih karena dianggap mempunyai rentang frekuensi yang tinggi sehingga lebih mudah dideteksi (Valin et al, 2003), sedangkan suara kelima (*tone*) dipilih sebagai pembanding karena suara ini mempunyai rentang frekuensi yang sedikit. Masing-masing suara dilakukan percobaan sebanyak 3 kali dengan sudut dan jarak yang bervariasi. Jarak antara *microphone* adalah 6 cm karena itu adalah jarak terbaik untuk pendeteksian suara (Fausi et al, 2015). Proses pengambilan suara dapat dilihat pada Gambar 4 dan sudut asli dari data yang diambil terdapat pada Tabel 1.

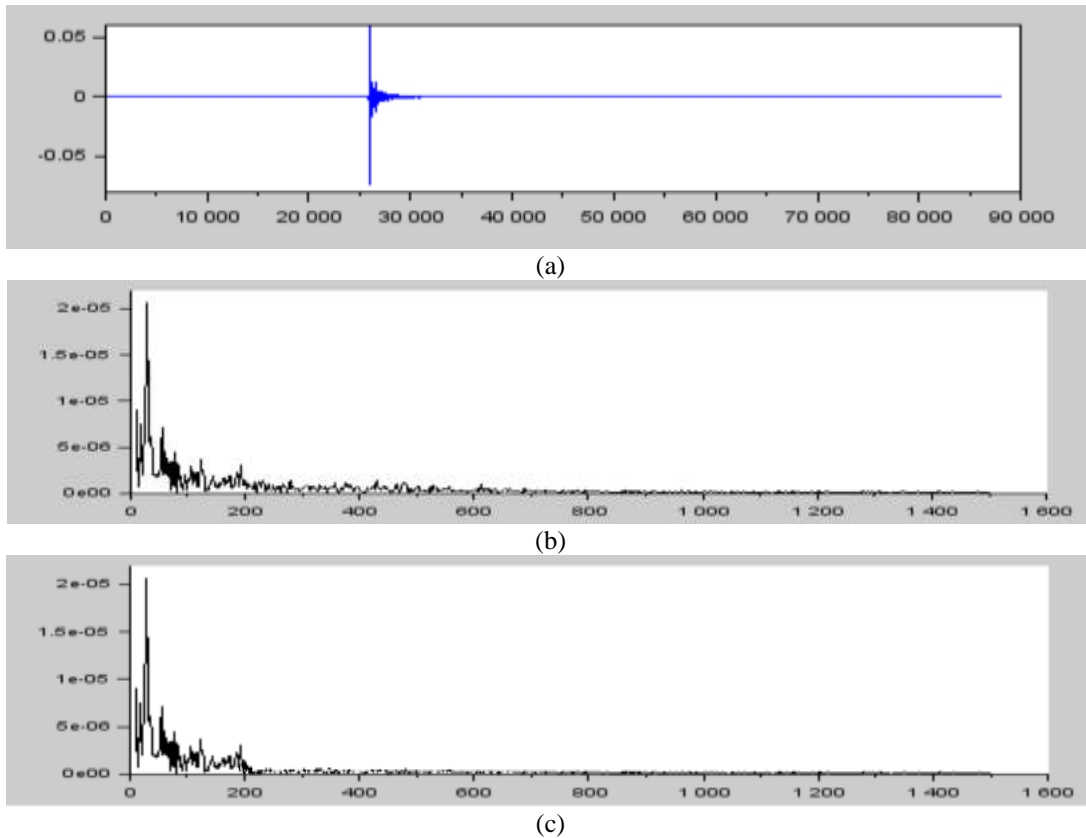
Pra-pengolahan suara dilakukan untuk mengurangi derau yang ada pada sinyal suara. Derau yang dapat dikurangi adalah derau yang tertangkap pada saat proses pengambilan data berlangsung, bukan derau yang diakibatkan oleh peralatan rekaman. Proses ini dilakukan pada domain frekuensi, sehingga langkah pertama adalah melakukan FFT (*Fast-Fourier Transform*) pada sinyal terekam. Kemudian dilakukan proses *filtering* pada sinyal suara sehingga menghasilkan suara yang mempunyai lebih sedikit derau. Hasil pengambilan data suara dan pre-processing sinyal dapat dilihat pada Gambar 5.



Gambar 4. Dokumentasi Pengambilan Data di Studio

Tabel 1. Data Pengambilan Suara

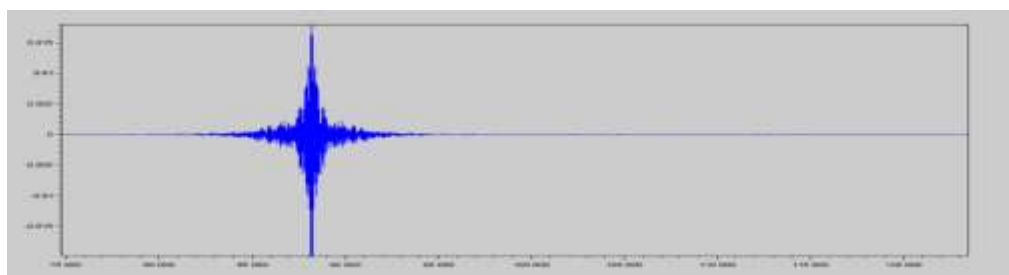
Jenis Suara	Sudut Asli (derajat)
Tepuk Tangan	0
	30
	60
Ketukan Kaki	10
	40
	80
Kunci Jatuh	20
	50
	70
Jentikan Jari	50
	0
	80
Nada 440 Hz	80
	40
	20



**Gambar 5.** Hasil pengambilan data yaitu (a) suara tepuk tangan dalam domain waktu, (b) suara tepuk tangan dalam domain frekuensi dan (c) hasil pre-processing sinyal suara

Proses pendeteksian sudut suara dilakukan dengan melakukan *cross-correlation* antara 2 sinyal. Ada 4 buah *microphone* yang digunakan dalam pengambilan data, oleh karena itu terdapat 4 buah sinyal terekam untuk satu sampel suara yang akan mengakibatkan banyaknya proses *cross-correlation* yang harus dilakukan. Gambar 6 adalah contoh salah satu proses *cross-correlation* dari dua buah sinyal yang sebelumnya telah dilakukan proses *whitening*.

Untuk mengurangi kompleksitas perhitungan tersebut, maka korelasi hanya dilakukan pada mikrofon dengan amplitude tertinggi sebagai acuan dan dengan mikrofon yang mempunyai nilai tertinggi kedua. Indeks nilai yang menyebabkan terjadinya puncak itulah yang akan diambil sebagai nilai acuan yang dianggap sebagai nilai *delay* sinyal. Nilai *delay* sinyal yang terdeteksi di *software* kemudian dibagi oleh frekuensi sampling yang akan menghasilkan nilai *delay* dalam satuan waktu. Nilai ini kemudian diproses dengan menggunakan persamaan geometri untuk mendapatkan sudut datang sumber suara.



**Gambar 6.** Salah satu proses *cross-correlation*

Tabel 2. Hasil pencarian sudut sumber suara

Jenis Suara	Sudut Asli (°)	Selisih Sudut Terdeteksi (°)	
		Studio	Kelas
Tepuk Tangan	0	2.1	1.4
	30	2.8	2.4
	60	2	4.5
Ketukan Kaki	10	3.8	4
	40	3.1	3.5
	80	3.7	4.1
Kunci Jatuh	20	1.2	2.1
	50	1.4	1.4
	70	1.1	2.5
Jentikan Jari	50	1.3	3.1
	0	1.9	2.1
	80	2.4	3.8
Nada 440 Hz	80	4.3	3.6
	40	3.6	3.5
	20	2.9	3.7
<b>Rata-rata</b>		2.506667	3.046667
		2.78	

Hasil penelitian menunjukkan bahwa secara umum hasil deteksi sudut dari suara terekam di studio mempunyai hasil yang lebih baik daripada yang direkam di kelas. Hal ini dikarenakan ruang kelas tidak mempunyai peredam dan di kelas-kelas yang ada di sebelahnya sedang dilaksanakan perkuliahan yang mengakibatkan derau yang lebih banyak daripada di studio. Namun hal ini menunjukkan bahwa sistem yang dibuat masih terpengaruh oleh *noise*.

Suara yang mempunyai rentang frekuensi yang lebar (tepu tangan, ketukan kaki dan jentikan jari) di studio mempunyai *error* yang lebih kecil daripada *tone*. Hal ini diakibatkan oleh proses *whitening* yang dilakukan di awal yang mengakibatkan setiap frekuensi mempunyai bobot yang sama terhadap proses korelasi. Artinya jika spektrum yang dianalisa adalah sebuah nada tertentu, maka hanya sedikit sekali informasi yang dapat digunakan dalam mencari sudut sumber dan *cross-correlation* yang terjadi akan didominasi oleh derau.

### KESIMPULAN

Dengan menggunakan *microphone array* yang terdiri dari 4 buah *microphone* dalam satu bidang planar peneliti telah berhasil membuat sistem yang dapat mendeteksi sudut dari sumber suara. Kesimpulan sementara yang dapat diambil antara lain yaitu suara yang mempunyai rentang frekuensi yang besar menunjukkan hasil deteksi sudut yang lebih baik. Selain itu suara yang direkam di studio juga mempunyai *error* yang lebih kecil dibandingkan suara yang direkam di kelas, yang artinya pemodelan derau yang dilakukan masih kurang tepat karena sinyal suara masih terpengaruh oleh derau sehingga perlu dilakukan perbaikan lebih lanjut. Adapun hasil sudut terdeteksi mempunyai *error* rata-rata 2,78°.

### UCAPAN TERIMA KASIH

Penulis mengucapkan banyak terima kasih kepada Direktorat Riset dan Pengabdian Masyarakat Kementerian Riset, Teknologi dan Pendidikan Tinggi yang telah membiayai penelitian ini untuk skema Penelitian Dosen Pemula (PDP) tahun anggaran 2018.

### DAFTAR PUSTAKA

- Buchholz, N.M.K.J. & Gal, G.F.J. 2005. *Sound Source Localization: Microphone array Design and Evolutionary Estimation*. IEEE International Conference on Industrial Technology (ICIT), USA, pp:286– 281.
- Fausi, Naba, dkk. 2015. *Analisis Jarak Microphone array dengan Teknik Pemrosesan Sinyal Fast Fourier Transform Beamforming*. Jurnal EECCIS, Vol. 9, No. 1, Juni 2015, hal 85-90.
- Fay, R.R. & Popper, A.N. (Eds). 2005. *Sound Source Localization*. USA: Springer Science.
- Gay, S.L. & Benesty, Jacob (Eds). 2000. *Acoustic Signal Processing For Telecommunication*. USA: Springer Science.
- Mumolo, Enzo, et all. 2003. *Algorithms for acoustic localization based on microphone array in service robotics*. *Journal Robotics and Autonomous Systems*, Vol. 42, pp 69-88.
- Nosal, Eva-Marie. 2013. *Methods for tracking multiple marine mammals with wide-baseline passive acoustic arrays*. *The Journal of the Acoustical Society of America*, Vol. 134, No. 3, USA.
- Sheng, Xiaohong & Hu, Yu-Hen. 2005. *Maximum Likelihood Multiple-Source Localization Using Acoustic Energy Measurements with Wireless Sensor Networks*. *IEEE Transactions on Signal Processing*, Vol. 53, No. 1, pp 44-53.
- Valin, Jean-Marc, et all. *Robust Sound Source Localization Using a Microphone array on a Mobile Robot*. *Proc. of 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Las Vegas, Nevada, USA, pp. 1228-1233.



Zhang, Cha, et all. 2008. *Maximum Likelihood Sound Source Localization and Beamforming for Directional Microphone arrays in Distributed Meeting*. *IEEE Transactions on Multimedia*, Vol. 10, No. 3, pp 538-548.