

DETEKSI POSISI SUMBER BUNYI MENGGUNAKAN *INTERAURAL TIME DIFFERENCE* DAN *CROSS CORRELATION*

Rahyul Amri

Jurusan Teknik Elektro Fakultas Teknik
Universitas Riau

Jl. HR.Subrantas km.12,5 Panam, Pekanbaru

rahylumri@yahoo.com

Abstrak

Interaural Time Difference (ITD) digunakan pada sistem pendengaran manusia untuk menentukan posisi sumber bunyi pada bidang *horizontal*. Implementasi *ITD* pada sistem robotik diharapkan akan membuat robot tsb mampu mendeteksi posisi dan menjejak sumber bunyi mendekati kemampuan manusia. Pada percobaan robot simulasikan dengan sebuah sistem penjejak yang mempunyai sistem pendeteksi dan sistem penggerak. Sistem pendeteksi mendeteksi posisi sumber bunyi dan nilai posisi diberikan ke sistem penggerak, selanjutnya sistem penggerak akan bergerak sesuai dengan besar nilai yang diberikan. Dari hasil percobaan yang dilakukan kemampuan sistem untuk mendeteksi posisi mempunyai akurasi rata-rata 91.4 % .

Kata kunci : Bunyi, *Interaural Time Difference*, *azimuth*, *cross correlation*, posisi .

1. PENDAHULUAN

Interaksi manusia - mesin/komputer/robot merupakan bidang yang potensial pada dewasa ini. Salah satu cara agar mesin/komputer/robot tersebut mampu berinteraksi adalah dengan mengimplementasikan kemampuan manusia pada mesin/komputer itu [01].

Penjejakan posisi sumber bunyi adalah salah satu kemampuan sistem pendengaran manusia yang dapat diimplementasikan pada mesin/komputer/robot. Dengan mengimplementasikan kemampuan ini diharapkan mesin/komputer/robot dapat merespon sumber bunyi seperti manusia.

Penjejakan pada prinsipnya adalah mendeteksi lokasi atau posisi sumber bunyi dan mengikutinya. Pada sistem robotik dengan sumber bunyi yang bergerak, dengan diketahuinya posisi sumber bunyi maka robot tersebut dapat digerakkan untuk menjejak sumber bunyi tersebut.

Pada umumnya teknik yang digunakan untuk melokalisasi lokasi atau posisi sumber bunyi adalah berdasarkan *Time Difference Of Arrival (TDOA)* [02]. Konsep ini mendasarkan prinsipnya pada perbedaan waktu datangnya sinyal pada susunan input. Susunan input pada kasus ini adalah jajaran *microphone* yang akan menerima sinyal pada waktu yang berbeda, dari sumber sinyal yang sama. Sebagai contoh, bunyi yang mencapai *microphone* yang lebih dekat akan mempunyai waktu yang lebih pendek dari *microphone* yang jauh.

Data dari perbedaan waktu datang (*time difference of arrival*) masing-masing *input* akan digunakan untuk mengetahui posisi dari sumber bunyi dengan menggunakan aturan *geometri*. Dari posisi yang diketahui ini sistem (robot) akan mengarahkan posisinya ke sumber bunyi.

2. DASAR TEORI

2.1. Bunyi

Bunyi adalah suatu gelombang mekanik bujur yang merambat melalui udara, air dan zat antara bermateri lainnya dalam rentang frekuensi antara 20 sampai dengan 20.000 Hz [03]. Bunyi adalah sangat penting didalam kehidupan manusia dan binatang ,yang mempunyai organ-organ trespesialisasi untuk menghasilkan dan mengamati gelombang-gelombang ini. Dengan menggunakan bunyi, manusia dan binatang mampu berkomunikasi dan berinteraksi satu dengan yang lainnya dan untuk memperoleh informasi tentang lingkungannya [04].

Bunyi yang merambat diudara mempunyai kecepatan (v) sebesar 345m/dtk pada suhu ruang.

Persamaan bunyi yang merambat diudara dapat dituliskan sebagai berikut :

$$v = s / t \quad (1)$$

$$s = t v \quad (2)$$

$$t = s / v \quad (3)$$

dimana

v : kecepatan bunyi merambat diudara

s : jarak yang ditempuh oleh bunyi dalam satuan meter

t : waktu sampai bunyi pada tempat yang dituju dalam satuan detik

2.2. Pendeteksian Posisi Sumber Bunyi

Dari persamaan(3) kita ketahui bahwa bunyi yang merambat pada udara akan sampai pada suatu objek pendeteksi pada waktu tertentu berhubungan dengan jarak yang harus ditempuhnya. Apabila ada dua objek pendeteksi, maka jarak waktu tempuh sumber bunyi terhadap objek pendeteksi yang satu dengan yang lain akan berbeda bila posisi relatif sumber bunyi dengan kedua objek pendeteksi tersebut berbeda atau dengan kata lain waktu sampai sumber bunyi pada kedua objek pendeteksi berbeda, hal ini dikenal sebagai perbedaan waktu sampai-*Time Difference Of Arrival (TDOA)*. *TDOA* ini pada sistem pendengaran manusia dikenal sebagai *Interaural Time Difference (ITD)* [05].

Untuk mendeteksi posisi sumber bunyi berdasarkan waktu sampai bunyi pada pendeteksi ada cara langsung dan tidak langsung. Pada cara langsung adalah berdasarkan *delay and sum of power* sinyal, sedangkan cara tak langsung adalah dengan merekam bunyi pada pendeteksi yang terpisah, memprosesnya dengan algoritma tertentu untuk mendapatkan *TDOA*-nya, kemudian dengan menggunakan aturan *geometri* akan didapatkan posisi sumber bunyi [06].

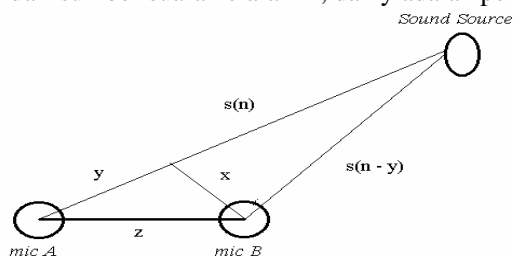
2.2.1. Time Difference Of Arrivall (TDOA)

Perhitungan *TDOA* timbul dari macam ragam bidang, termasuk *speech localization* dan pengolahan menggunakan jajaran mikropon. Jajaran mikropon yang kecil terdiri dari dua mikropon seperti gambar 1 dapat di hitung *TDOA*-nya adalah sebagai berikut :

$$t_a = s(n) \quad (4)$$

$$t_b = s(n - y) \quad (5)$$

dimana t_a dan t_b adalah waktu sampai pada *mic A* dan *mic B*, $s(n)$ adalah *referensi* sinyal dan n adalah jarak dari sumber suara ke arah A, dan y adalah perbedaan waktu sampai antara t_a dan t_b .



Gambar 1. Penentuan TDOA

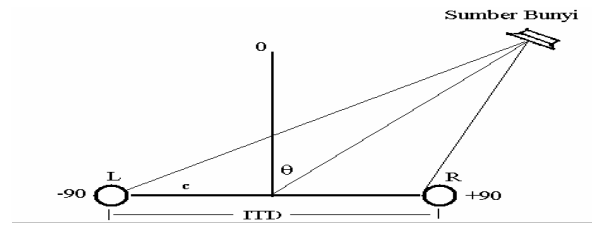
Kemudian sudut arah dari sumber suara yang terjauh dapat ditentukan sebagai berikut :

$$\theta = \cos^{-1} (y/z) \quad (6)$$

dimana y adalah perbedaan waktu diantara dua signal dan z adalah jarak antara *mic A* dan *B*.

2.2.2. Interaural Time Differences

Untuk menemukan posisi sumber bunyi dengan menggunakan *Interaural Time Differences (ITD)* kita menghitung *azimuth* – sudut dari sumber suara yang menghadap titik *referensi* 0° biasanya adalah hidung manusia pada bidang horizontal – di tunjukkan pada gambar 2.



Gambar 2. Perhitungan menentukan azimuth

$$\text{Azimuth } \theta = \arcsin \text{ITD} \times V_{\text{sound}} / c \quad (7)$$

dimana

ITD = beda waktu sumber bunyi antara L dan R

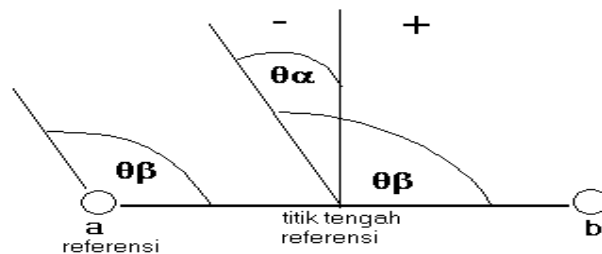
V_{sound} = kecepatan rambat suara di udara

c = jarak antara L dan R

2.2.3. Referensi Satu Microphone (RSM)

Menggunakan titik tengah sebagai titik referensi adalah konsep yang digunakan oleh manusia untuk membedakan letak sumber suara di kiri atau di kanannya, dikenal sebagai *Interaural Time Difference (ITD)*.

Untuk menjelaskan perbedaan penggunaan titik referensi untuk menentukan letak sumber suara dapat dijelaskan sebagai berikut, lihat gambar gambar 11, bila θ_α adalah sudut yang ditempuh dengan menggunakan titik referensi titik tengah antara mikrofon a dan b, dan θ_β adalah sudut yang ditempuh dengan menggunakan titik referensi *microphone* a, yang berarti memakai Referensi Satu *Microphone* (RSM).



Gambar 11. Besar sudut yang ditempuh menggunakan ITD vs RSM

maka untuk menempuh sudut yang lebih besar dari 90, besarnya θ_β menjadi :

$$\theta_\beta = 90 + (\pm \theta_\alpha) \quad (8)$$

dan

$$\theta_\alpha = \pm \theta_\alpha \quad (9)$$

dimana tanda \pm menyatakan kiri atau kanan dari titik tengah referensi.

Dari persamaan (8) dan (9) tampak bahwa bila menempuh sudut lebih dari 90 dengan memakai referensi satu mikrofon (RSM) menempuh jarak yang lebih besar, yang berarti membutuhkan waktu lebih lama untuk mencapai posisi azimuth sumber suara dibandingkan dengan menggunakan titik tengah mikrofon sebagai titik referensi.

2.2.4. Cross Correlation

Suatu operasi matematis yang menghitung korelasi dua deret sinyal dengan tujuan untuk mengukur derajat kemiripan sinyal disebut dengan *cross correlation* [07].

Fungsi *cross correlation* didefinisikan sbb :

$$\text{Corr}(g,h)_j(t) = \sum_{k=0}^{N-1} g_{j+k} h_k \quad (10)$$

Fungsi korrelasi akan bernilai maksimum ketika fungsi $g(t)$ di geser melewati fungsi $h(t)$ mencapai kemiripannya..

Cross correlation disini di gunakan untuk membandingkan dua *vector A* dan *B* (yang berisi nilai dari sinyal $g(t)$ dan $h(t)$). *Vector A* dan *B* disini merepresentasikan sinyal yang ditangkap *mic A* dan *B*. Sinyal kemudian digeser melewati masing-masing titik untuk memberikan *product vector C* yang panjangnya ditunjukkan sbb :

$$\text{Length}(\mathbf{C}) = (\text{length}(\mathbf{A}) + \text{length}(\mathbf{B})) - 1 \quad (11)$$

Untuk menghitung perbedaan waktu (*delay*) dari dua sinyal, kita merepresentasikan kiri dan kanan sebagai *vector A* dan *vector B*, *vector* ini berisi nilai amplituda dari sinyal masing-masing lokasi. Jika $g(t)$ dan $h(t)$ tidak mempunyai *delay*-perbedaan *fase* sinyal= 0° - maka nilai maksimum akan terjadi pada pertengahan titik dari *vector A* dan *B* yaitu letak titik *vector C*.

Bila ada *delay* (perbedaan waktu dengan salah satu sinyal sebagai titik referensi) pada dua sinyal tersebut maka besarnya dihitung dengan persamaan (10) dengan titik pertengahan *vector* tersebut adalah titik referensi.

$$\text{Delay} = C_{\max} - (\text{panjangVectorC}/2) \quad (12)$$

Dimana C_{\max} adalah nilai maksimum dari *cross correlation* dan panjang *Vector* didapat dari persamaan(11).

2. 3. Algoritma Penentuan Posisi *Azimuth* (\square) menggunakan *ITD* dan *Cross Correlation*

Posisi *Azimuth* (\square) dihitung dari proses sebagai berikut :

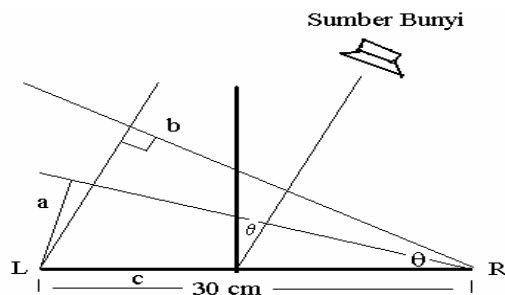
Pertama sekali metode *cross correlation* dipakai untuk mendeteksi *delay* (\square) perbedaan waktu dari dua sinyal dengan salah satu sinyal sebagai referensi. Nilai *delay* dari korelasi silang sinyal ditentukan dengan titik tengah *C* sebagai 0° , kemudian dari titik tengah ini digeser ke lokasi yang memberikan posisi maksimum *vector*. Posisi ini bisa negatif (kurang dari titik tengah) atau positif (lebih besar dari titik tengah).

Nilai *delay* ini kemudian di gunakan untuk menghitung *ITD* pada sistem. Untuk menemukan *ITD* kita perlu menghitung *variable sampling*.

Variable pertama yang kita perlukan adalah waktu antara *sampling* (\square)

$$\Delta = 1 / 44,1 \times 10^3 = 2,2676 \times 10^{-5} \text{ s} \quad (13)$$

Assumsi kita adalah bunyi datang *parallel* seperti di perlihatkan pada gambar 3.



Gambar 3. Sudut datang sumber suara

ITD didapat :

$$ITD = \Delta \times \delta \quad (14)$$

Sehingga didapat azimuth (θ) :

$$\theta = \text{Sin}^{-1} (\Delta \times \delta) \times V_{\text{sound}} / c \quad (15)$$

dimana

Δ = waktu antara *sampling*

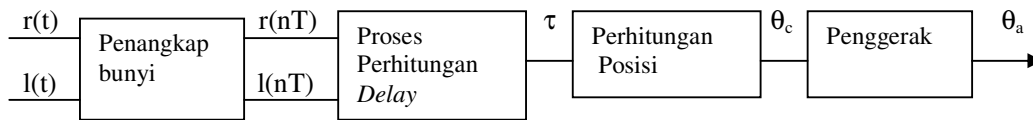
δ = nilai yang didapat dari hasil operasi *cross correlation*

V_{sound} = kecepatan rambat suara di udara

c = jarak antara L dan R

3. PERANCANGAN SISTEM

Secara keseluruhan diagram blok sistem dapat digambarkan seperti gambar 5

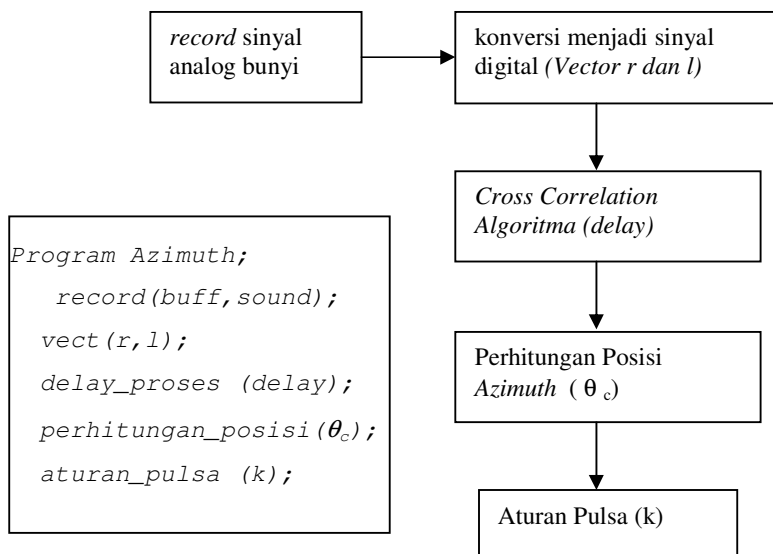


Gambar 5 Diagram blok sistem secara keseluruhan

Input dari sistem ini adalah sinyal bunyi yang ditangkap oleh dua *microphone* kiri dan kanan pada bagian penangkap bunyi berupa sinyal analog $r(t)$ dan $l(t)$. Agar bisa diolah oleh komputer sinyal analog ini diperkuat dan diubah menjadi sinyal diskrit $r(nT)$ dan $l(nT)$. Sinyal kemudian diolah oleh bagian proses perhitungan *delay* dengan menggunakan algoritma *cross correlation* yang hasilnya adalah *delay* diantara dua sinyal (τ). Setelah didapat *delay* (τ) maka akan dapat dihitung posisi *azimuth* sumber bunyi terhadap system (θ_c). Dari posisi *azimuth* yang di peroleh dari bagian perhitungan posisi, bagian penggerak akan menggerakkan sistem kearah posisi *azimuth* sumber bunyi (θ_a). Proses ini ini terus berulang untuk menjejak / mengikuti posisi *azimuth* sumber bunyi yang baru.

3.1. Perangkat Lunak

Perangkat lunak yang mengontrol seluruh sistem dapat digambarkan sebagai berikut :



Listing perangkat lunak

Software pertama sekali merekam sinyal *analog* bunyi kiri kanan dan mengkonversinya menjadi sinyal *digital* yang tersusun atas *vector r* dan *l*. *Vector r* dan *l* ini akan diolah oleh algoritma *cross correlation* untuk mendapatkan informasi *delay* dari sumber bunyi. Setelah *delay* diperoleh akan dilakukan fungsi untuk menghitung posisi *azimuth* (θ_c) sumber bunyi terhadap sistem (robot). Dari informasi posisi *azimuth* (θ_c) ini dilakukan prosedur aturan pulsa yang akan menghasilkan nilai aturan pulsa(k) (lihat tabel1). Nilai aturan pulsa (k) akan dikeluarkan melalui *interface LPT1* komputer untuk menggerakkan motor ke posisi yang dituju (θ_c).

4. HASIL DAN PEMBAHASAN

Pada bab ini akan membahas mengenai pengujian dari sistem yang telah dibuat dan pembahasan hasilnya. Pengujian dilakukan untuk :

1. Mengetahui kemampuan algoritma *cross correlation* untuk mendapatkan *sample delay* dari dua sinyal bunyi.
2. Mengetahui kemampuan sistem untuk mendeteksi posisi *azimuth* sumber bunyi.
3. Mengetahui kemampuan waktu reaksi sistem untuk menjejak/mengikuti sumber bunyi.

4.1. Pengujian algoritma *cross correlation*

Experimen ini dilakukan dengan memberikan bunyi yang telah direkam ke blok Proses Delay untuk mencoba kemampuan algoritma *cross correlation* untuk :

1. Membedakan arah sinyal yang datang terhadap titik tengah penangkap bunyi kiri dan kanan.
2. Menghitung *sample delay* bunyi .

Dari pengujian yang dilakukan didapat hasil seperti pada table 1

Tabel 1. Hasil pengujian algoritma *cross correlation*

Percobaan n	Time Delay Yang diberikan	Amplitude	Nilai Titik referensi	Vector location maksimum <i>cross correlation</i>	Sample Delay hasil <i>cross correlation</i> / Arah datang sumber sinyal
Pertama	0	Kiri,Kanan=100 %	7056	7056	0 /depan
Kedua	kanal kanan = 500 μ s	Kiri = 80 %	7056	7058	22 / kanan
Ketiga	kanal kiri = 500 μ s	Kanan = 80 %	7056	7034	-22 / kiri

Dari percobaan pertama didapat hasil nilai maksimum *cross correlation* pada sistem terletak pada lokasi *vector* bernilai 7056. Dari persamaan (12) dapat dihitung bahwa besarnya *delay* adalah 0 yang menunjukkan bahwa algoritma ini dapat menunjukkan hasil yang sesuai apabila perbedaan waktu dari sinyal adalah 0(nol). Percobaan kedua dan ketiga menunjukkan algoritma *cross correlation* pada sistem mampu membedakan apakah sinyal datang dari kiri atau kanan titik referensi. Nilai *sample delay* positif 22, berarti sumber bunyi datang dari arah kanan dengan perbedaan waktu sampai 22 sample. *Sample delay* negatif 22, berarti sumber bunyi datang dari arah kiri dengan perbedaan waktu sampai 22 sample .

Pengujian pendeteksian posisi azimuth

Pengujian ini dilakukan untuk mengetahui kemampuan perangkat lunak untuk menghitung posisi azimuth sumber bunyi dan kemampuan penggerak dari untuk menggerakkan penangkap bunyi kiri dan kanan kearah posisi sumber bunyi.

Tes ini dilakukan dengan menempatkan sudut sumber bunyi pada sudut tertentu dengan jarak 1 m dari dan melihat apakah penangkap bunyi dapat mengarah ke posisi sumber bunyi. Setiap percobaan diulang 5 kali dan diambil rata-ratanya. Akurasi dihitung dengan persamaan :

$$\text{Akurasi} = ((\text{nilai aktual} - \text{range pengukuran})/\text{nilai aktual}) \times 100\%$$

Dari percobaan yang dilakukan didapat hasil seperti yang ditunjukkan tabel 2

Tabel 2. Hasil dari pengujian pendeteksian posisi azimuth

Percobaan n	Posisi Aktual (derajat)	Posisi Sistem (rerata)	Akurasi (%)
1	-50	± 4	92
2	-40	± 3	92.5
3	-30	± 2	93.3

4	0	± 2	97.7
5	10	± 2	80
6	20	± 2	90
7	35	± 3	91.4
8	45	± 4	91.1
9	70	± 5	92.8

Dari hasil percobaan yang dilakukan dapat dilihat bahwa sistem dapat bergerak dan mengarah keposisi yang ditentukan dengan akurasi rata-rata 91,4 %.

5. SIMPULAN

5.1. Kesimpulan

Dari hasil pengujian yang telah dilakukan menunjukkan bahwa konsep *Interaural Time Difference* dapat diterapkan pada sistem robotik untuk mendeteksi posisi *azimuth* sumber bunyi dan menjejaknya. Proses untuk mendapatkan *delay* dari dua sinyal yang di korelasikan dengan menggunakan algoritma *cross correlation* menunjukkan hasil yang memuaskan. Sistem yang dirancang dan dibuat telah memenuhi :

1. Akurasi atau ketepatan *ITD* untuk mendeteksi posisi *azimuth* pada tesis ini adalah 91,1 % dengan *sampling rate* 44,1 kHz, bila dibandingkan dengan kemampuan manusia hasil ini cukup mendekati (akurasi manusia 97,5 % untuk mendeteksi posisi *azimuth* [7]).

5.2. Saran

Untuk meningkatkan waktu reaksi dari sistem sehingga mencapai *real time* ada beberapa hal yang perlu diberi perhatian khusus :

1. Penggerak yang digunakan, dimungkinkan untuk penjejakan yang lebih cepat dengan menggunakan sistim penggerak *servomechanisme*.
2. Dengan menggunakan *soundcard* yang lebih kompleks dan *bitrate* tinggi akurasi dari proses *delay* dapat ditingkatkan.
3. Komputer dapat digantikan dengan *device* elektronik *digital* lain,misalnya *PLD,FPGA* dengan mengkonversi algoritma *cross correlation* ke protokol alat tersebut yang akan menghasilkan proses yang lebih cepat.

6. DAFTAR ACUAN

- [01] Frank Maddix, "Human-Computer Interaction, Theory and Practice", EllisHorwood Limited, 1990
- [02] Pedro Julian ,G.Andreou,'A Comparative Study of Sound Localization Algorithms for Energi Aware Sensor Network Nodes", Journal IEEE Transaction On Circuit and System, Vol 51, No.4 April 2004
- [03] Alan H.Cromer, Sumartono Prawirosusanto,Koesminarto, "Fisika Untuk Ilmu-Ilmu Hayati " original edition : "Physic For Live Science Second Edition ",Gajah Mada University Press, 1994
- [04] S.S.Stevens,Fred Warshhofskey,dan para Editor Pustaka *Time Live*, edisi Indonesia oleh: Drs.Willie Koen Lic.Phill,dr.Yurita Handoyo,Drs R.Sumarsono D, "Bunyi Dan Pendengaran " edition original : " Sound And Hearing", PUSTAKA TIME-LIVE, 1983
- [05] John C.Murray,Harry Erwin, Stefan W., "Robotic Sound Source Localization And Tracking Using Interaural Time Difference And Cross Correlation ", AI Workshop on NeuroBotics, Germany, September 2004

- [06] Parham Aarabi, Alborz Mahdavi, "The Relation Between Speech Segment Selectivity And Source Localization Accuracy", Proc. IEEE Int. Conf. Acoustic, Speech and Signal Processing, 1996
- [07] Brandstein M.S, John E. Adcock, and H.F. Silverman, "A Practical Time Delay Estimator For Localizing Speech Source with a Microphone Array", Journal Computer, Speech and Language, Vol. 9:153-169, 1995