Detection of Audio Source Direction Using Autonomous Robot

ADITYA AGARWAL, NILESH GOEL, SUBHAYAN BANERJEE and HARANATH KAR
Department of Electronics and Communication Engineering, MNNIT, Allahabad- 211 004, India
Email: adityamnnit@rediffmail.com

Abstract: In this paper, we propose a system for reliable detection of the direction of the audio source that is to be installed in an embedded robot. Compared with previous researches, this system comprises simpler, faster and more accurate algorithm and a threshold level setting mode to make the system immune to external noise. This system consists of a microphone assembly of three microphones, band pass filter, microcontroller as processing unit, motor controller unit and mechanical assembly. The robot senses the audio signal lying in the frequency range (300 Hz-3400 Hz) and detects the direction of sound source using delay of arrival (DOA) in three microphones. This system is to be installed in an experimental autonomous robot named 'SODDRO' (SOund Direction Detection RObot). With this system, 'SODDRO' will move towards sound source after detecting the direction of sound.

Keywords: reliable detection, delay of arrival (DOA).

1. Introduction

The sound detection is an interesting field that has drawn the attention of many researchers in the recent years [1], [2], [5]-[8]. In [1], [2], the sound detection mechanism has been developed for human robot interaction. Sound detecting humanoid robot that can recognize speech can work in very harsh inhuman conditions like nuclear reactors where extremely high temperatures and harmful radiations prevail.

The aim of this paper is to design an autonomous robot using microcontroller which will follow its operator whenever an audio signal is received from operator.

The rest of the paper is organized as follows. The description of the system for reliable detection of the direction of the audio source is given in Section 2. In Section 3, an algorithm to track sound direction is proposed. Some illustrations to prove the viability of the proposed algorithm are presented in Section 4.

2. Proposed System

The block diagram of the proposed system for reliable detection of the direction of the audio source is shown in Fig. 1.

Figure 1: Block diagram of the proposed system.

A brief description of various modules of the proposed system is given below.

2.1 Microphone Assembly

The microphone assembly as shown in the Fig. 2 consists of three condenser microphones (M1, M2 and M3) oriented at an angle of 120° with respect to each other. The reason of choosing three microphones is to cover the whole 360° uniformly, so that effective detection of sound’s direction can be performed. The effective region of each microphone for sound detection is 120°.
The microphone works as a transducer which converts the audio signal to electrical signal. This arrangement uses minimum number of microphones for reliable detection.

2.2 Band Pass Filter

The band pass filter used in the system as shown in Fig. 3 is an active second-order band pass filter which is designed using operational amplifiers [4] (IC LM324). The pass band of this filter lies within the audio range (300Hz – 3400Hz). The purpose of using band pass filter is to eliminate environmental noise. The power spectral density of noise shows that maximum power is concentrated at low frequencies near to 50 Hz as compared to higher frequencies. Therefore, it is necessary to remove this noise.

2.3 Processing and Decision Making Unit

The processing and decision making as shown in Fig. 4 is done by ATMEL’s AVR family ATMEGA32L microcontroller [3], [9]. This microcontroller has 32k programmable flash memory and its maximum clock frequency is 8 MHz. It has an inbuilt 8 channel, 10 bit A/D converter. A/D converter is used for converting analog signal from the band pass filter to digital signal which is processed by the processing unit of microcontroller and accordingly, it will generate appropriate logic signals to drive the motors in mechanical assembly.

2.4 Motor Controller Unit

Dual full H bridge motor driver IC L298 [10] as shown in Fig. 4 is used to control the movement of motors. Each H bridge is capable of moving the motors ‘clockwise’ or ‘anticlockwise’ depending upon the direction of current flow through the circuit. Using IC L298, operator can ‘jam’ or ‘free’ the motors if required. Basically L298 [10] acts as an interface between the low power logic signal generated by microcontroller and the motor assembly which requires relatively high power for driving of motors.

2.5 Mechanical Assembly

This part mainly consists of two brushed DC motors, gear boxes, side steering control and vehicle chassis. The side steering mechanism can effectively control the motors while taking sharp turn. Motor 1 control the motion of left wheels assembly and motor 2 controls the right wheel assembly as shown in the Fig. 5.
3. Algorithm to Track Sound’s Direction

The Delay of Arrival (DOA) [1] mechanism is used for efficient detection of the direction of sound. Such mechanism uses the time delay from the sound source to each microphone. The microphones are connected as the vertices of an equilateral triangle. The electrical signal (obtained from the microphone assembly) is converted to digital signal with the help of A/D converter inbuilt in AVR microcontroller.

First, microcontroller takes some predefined number of samples from each microphone to set the threshold level that depends on the amplitude of local disturbances. This makes the system immune to local disturbances as now microcontroller recognizes only those signals that are having higher amplitude than the set threshold level. The microcontroller samples the three microphones continuously and detects to which microphone the sound comes first (having amplitude higher than the threshold level). After determining the first microphone that receives the sound, the microcontroller sets the offset angle ($0^\circ$ for M1, $120^\circ$ for M2 and $-120^\circ$ for M3) according to the orientations of microphones M1, M2 and M3 and then it only samples the rest two microphones so that problems due to echo do not arise.

The rest two microphones (say M2 and M3) are monitored continuously and the microcontroller determines to which microphone the sound comes next and accordingly it determines whether the robot has to take a clockwise or anticlockwise turn and generates the control signals for motor controller IC. After this, it only samples the remaining microphone and calculates the delay in the arrival of sound between the two microphones (M2 and M3) using the inbuilt timer. This delay $T_d$ determines the angle of deviation from the offset angle and that will be the final angle at which the robot has to rotate to reach the operator. Conversion formula is derived to convert the offset angle to corresponding time period for which motor has to run. To explain the above stated algorithm further, consider the following example.

Suppose the sampling of three microphones is being done in the sequence of M1, M2 and M3 and this sequence is repeated. The sequence entirely depends on the programming of the microcontroller.

There are six combinations in which sound can be detected by the microphones.

1. M1, M2, M3
2. M1, M3, M2
3. M2, M1, M3
4. M2, M3, M1
5. M3, M1, M2
6. M3, M2, M1

Let’s consider case one (M1, M2, M3). In this case M1 will detect the sound first which means that the sound source must be in the angular range ($\theta_1$) as shown in Fig. 6.

Let the DOA of sound between M2 and M3 is $T_d$. Depending upon the sound source
direction, the microcontroller generates the required logical signals for the time interval that is decided by the on-chip timer of AVR microcontroller that counts for $T_d$ and the offset time. The offset time is determined on the basis of the reception of sound by the microphone at which sound comes first. The logical signals generated by the microcontroller are the input for the motor controller IC L298 that controls the bidirectional motion of motors of mechanical assembly.

For reliable calculation of $T_d$, it is necessary that

$$T_d > \frac{1}{f} \quad (1)$$

where $f$ is the frequency at which the microcontroller samples the digital data received from microphones.

If condition (1) is not satisfied, i.e., the time taken by the microcontroller to sample one microphone to other microphone is more than the DOA, the microcontroller will be not able to judge that which microphone received the sound first and proposed algorithm will be not efficient. To avoid this situation, the sampling frequency, $f$, is kept as high as possible.

The maximum delay of arrival can be observed in the case when the sound source is just between two microphones, i.e., when the angle between the source and the M1 is $60^\circ$ as shown in Fig. 7. In this situation, the delay of arrival between M2 and M3 is maximum that can be given by

$$T_{d \text{max}} = \frac{3L}{2V_s} \quad (2)$$

where $V_s$ is the speed of sound and $L$ is the distance of each microphone from the center O of the microphone assembly. Using (1) and (2), we obtain

$$\frac{1}{f} < \frac{3L}{2V_s} \quad (3)$$

For the proposed algorithm, there should be a minimum width of the sound pulse generated by the sound source for the reliable calculation of DOA between two microphones. In other words, we need

$$\text{Sound pulse width} > \frac{3L}{2V_s} \quad (4)$$

3.1 Time Calculation for Control Logic Signals

From Fig. 8, it is clear that

$$d = 2L \cos 30^\circ \sin \theta$$

$$= \sqrt{3} L \sin \theta \quad (5)$$

$$\text{DOA} = \frac{d}{V_s} \quad (6)$$

For this DOA the timer will count and let the count of counter after this time delay is $N$ and clock frequency of timer is $f_T$. Therefore,

$$N = \text{DOA} \times f_T$$

$$= \frac{3L}{2V_s} \times f_T$$

Figure 7: Calculation of maximum time delay

Figure 8: Sound source at an angle
where use has been made of (5) and (6). Equation (7) can be rewritten as

\[ \theta = \sin^{-1} \left( \frac{NV_s}{\sqrt{3} f T L} \right). \]  

(8)

Total angle turned by the robot = offset value ± \( \theta \) where

\[
\text{offset value} = \begin{cases} 
0^\circ & \text{if M1 detects the sound first} \\
120^\circ & \text{if M2 detects the sound first} \\
-120^\circ & \text{if M3 detects the sound first} 
\end{cases}
\]

and ‘+’ sign is taken when the 2nd microphone which detects the sound after the first one is in the clockwise direction to the first detected microphone and the ‘–’ sign is taken if the 2nd microphone detected is in the anticlockwise direction to the first detected microphone.

Total rotation time of robot

\[ t = \frac{\text{offset value} \pm \sin^{-1} \left( \frac{NV_s}{\sqrt{3} f T L} \right)}{\omega_r} \]  

(9)

where \( \omega_r \) is the turning speed of robot.

3.2 Error Calculation

As the time taken by the microcontroller in sampling one microphone to other microphone is \( \frac{1}{f} \), so there is an error in calculation of angle or in other words it is the minimum angle (resolution of the system) that can be efficiently detected by the robot.

From Fig. 9, we obtain

\[ a = 2 L \cos 30^\circ \sin \theta \]

\[ = \sqrt{3} L \sin \theta \]  

(10)

Time Delay = \( \frac{a}{V_s} \)

\[ = \sqrt{3} L \sin \theta / V_s \]

\[ = 1/f \] (for one sampling time period)  

(11)

From (10) and (11), we have

\[ \theta_{\text{error}} = \sin^{-1} \left( \frac{V_s}{\sqrt{3} f L} \right). \]  

(12)

\( \theta_{\text{error}} \) is the maximum error or the minimum detectable angle (resolution of the system) in determining the angle.

If we generalize it for an assembly of say ‘n’ microphones then from Fig. 10

\[ a = 2 L \sin (180^\circ / n) \]

\[ = \sqrt{3} L \sin \theta \]  

(10)

Time Delay = \( \frac{a}{V_s} \)

\[ = \sqrt{3} L \sin \theta / V_s \]

\[ = 1/f \] (for one sampling time period)  

(11)

From (10) and (11), we have

\[ \theta_{\text{error}} = \sin^{-1} \left( \frac{V_s}{2 f L \sin (180^\circ / n)} \right). \]  

(13)

If \( V_s = 332 \text{ m/sec} \), \( L = 0.32 \text{ metre} \), \( f = 22,727 \text{ Hz} \), then the curve of maximum error in degree

\( \theta_{\text{error}} \)
(θ_{error}) vs. no of microphones (n) plotted with the help of MATLAB is as shown in Fig. 11:

![Graph showing θ_{error} vs. n](image)

**Figure 11: Plot of θ_{error} vs. n**

From the above curve it can be easily observed that θ_{error} increases linearly with the increase in the numbers of microphones. So it is clear that an assembly of three microphones is the best choice to reduce the maximum error (θ_{error}) and system complexity.

### 4. Illustrations

To check the viability of the algorithm some practical cases may be discussed.

Consider the case when sound first reaches M1, then M2 and then M3. In this case, there are three possibilities when the sound reaches the M1.

1. Microcontroller is sampling M1 at that time.
2. Microcontroller is sampling M2 at that time.
3. Microcontroller is sampling M3 at that time.

In the following, let us consider all the cases one by one.

When there is sound on a particular microphone, it is assumed to be 1 otherwise 0. In the following cases, the ‘1’ or ‘0’ below any microphone represent that microcontroller is sampling particular microphone at that instant of time.

**Case 1: Microcontroller is sampling M1**

<table>
<thead>
<tr>
<th>M1M2M3</th>
<th>M1M2M3</th>
<th>M1M2M3</th>
<th>M1M2M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 0 1 0 0 1 0 0 1 1 0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Timer Start**

**Timer Stop**

**Case 2: Microcontroller is sampling M2**

<table>
<thead>
<tr>
<th>M1M2M3</th>
<th>M1M2M3</th>
<th>M1M2M3</th>
<th>M1M2M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 0 1 0 0 1 0 0 1 1 0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Timer Start**

**Timer Stop**

**Case 3: Microcontroller is sampling M3**

<table>
<thead>
<tr>
<th>M1M2M3</th>
<th>M1M2M3</th>
<th>M1M2M3</th>
<th>M1M2M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 0 1 0 0 1 0 0 1 1 0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Timer Start**

**Timer Stop**

In case 1 and case 2, the count of the timer is same while in case 3 there is only a difference of one count with respect to case 1 or case 2 and this is due to the time taken by the microcontroller to sample one microphone i.e. \(\frac{1}{f}\). So at most there may be an error of only one count but this is not the limitation of this algorithm as the number of count is large. Therefore this error can be neglected.

The same fact can be proved for all the rest five combinations of M1, M2, M3 and this shows the viability of the algorithm.

### Conclusion

A system for reliable detection of the direction of the audio source has been developed. An algorithm for an autonomous robot that
detects the direction of sound source is developed. The viability of proposed algorithm has been illustrated. As the algorithm is to be implemented with microcontroller that can be operated at a very high frequency, the maximum possible error will be negligible. Moreover, a concept of threshold is also incorporated to make the system immune to local disturbances.

References


