

Article

Screen-Based Sports Simulation Using Acoustic Source Localization [†]

Sang-Woo Seo ¹, Somi Yun ², Myung-Gyu Kim ¹, Mankyu Sung ^{3,*} and Yejin Kim ^{4,*}

¹ Creative Content Research Division, Electronics and Telecommunications Research Institute, Daejeon 34129, Korea

² College of Physical Education, Kookmin University, Seoul 02707, Korea

³ Department of Game and Mobile, Keimyung University, Daegu 42601, Korea

⁴ School of Games, Hongik University, Sejong 30016, Korea

* Correspondence: mksung@kmu.ac.kr (M.S.); yejkim@hongik.ac.kr (Y.K.)

[†] This paper is an extended version of Seo, S.-W., Kim, M. and Kim, Y. Efficient Sound-source Localization Method based on Direction Estimation using 2D Microphone Array. In Proceedings of the 6th International Conference on Information and Technology: IoT and Smart City (ICIT), Hong Kong, China, 29–31 December 2018.

Received: 5 June 2019; Accepted: 23 July 2019; Published: 24 July 2019



Featured Application: Screen-Based Sports Simulation.

Abstract: In this paper, we introduce a novel acoustic source localization in a three-dimensional (3D) space, based on a direction estimation technique. Assuming an acoustic source at a distance from adjacent microphones, its waves spread in a planar form called a planar wavefront. In our system, the directions and steering angles between the acoustic source and the microphone array are estimated based on a planar wavefront model using a delay and sum beamforming (DSBF) system and an array of two-dimensional (2D) microelectromechanical system (MEMS) microphones. The proposed system is designed with parallel processing hardware for real-time performance and implemented using a cost-effective field programmable gate array (FPGA) and a micro control unit (MCU). As shown in the experimental results, the localization errors of the proposed system were less than 3 cm when an impulsive acoustic source was generated over 1 m away from the microphone array, which is comparable to a position-based system with reduced computational complexity. On the basis of the high accuracy and real-time performance of localizing an impulsive acoustic source, such as striking a ball, the proposed system can be applied to screen-based sports simulation.

Keywords: acoustic source localization; impulsive sound; delay and sum beamforming; microelectromechanical system; screen-based sports simulation

1. Introduction

The screen-based simulation is becoming popular for various sports as it requires less play time and space in any weather condition [1,2]. For example, various golf simulators have been developed mainly for pleasure [3]. Active studies have been conducted to track ball motion as it is a core technique for screen-based sport simulations. Most current commercial simulators are based on a computer vision technique using high-speed cameras [4,5] or a radar-based technique using the doppler effect [6]. However, these systems have a limited range to locate a ball position in three-dimensional (3D) space. For example, computer vision-based systems require a user to place a ball in the field of view of the cameras. This restricts the ball placement to a small area on the ground and is not suitable for active sports such as soccer or baseball. The radar-based systems require a ball flying over a long distance (i.e., several meters or more) for accurate measurement [7], which is not suitable for indoor simulation.

On the other hand, acoustic source localization provides more flexible placement and can be used to develop an indoor system for screen-based sports simulation.

The technique of acoustic location was first utilized as an air defensive device during World War I with the aim of tracking the position of enemy aircraft. In recent years, acoustic source localization has been adopted in various industrial fields such as sonar [8], car repair service [9], sports simulators [10], and musical instruments [11]. Different types of localization estimation methods have been developed for each purpose.

The methods of locating an acoustic source is divided into two categories according to the acoustic function type: parametric and nonparametric. Acoustic holography [12] is one of the nonparametric methods that detects not only the position of the acoustic source but also the characteristics of the acoustic field. However, it requires a lengthy calculation time and many microphones to locate the source. On the other hand, the parametric method finds the origin of the acoustic source faster than the nonparametric method by using the signal parameters with a small number of sensors. These methods are classified as a frequency domain approach and a delay and sum beamforming (DSBF) method using the concept of the time difference of arrival (TDOA) in a temporal domain.

Loesch et al. proposed a method that estimates the direction and distance based on the frequency domain [13]. However, it was difficult to distinguish the background noises from impulsive sources such as gunshots, clapping, or hitting a ball. On the other hand, an approach based on TDOA estimates the source position using the time difference of the sound signals from a small number of samples. Heilmann et al. proposed a technique to locate the source in 3D space by using a set of microphones arranged in a spherical form [9]. However, their system required many expensive microphones and careful placement, which made it difficult to use for a sports simulation. Recently, a deep neural network was introduced to locate an acoustic source from a noisy and reverberant environment [14,15]. However, it demanded a large amount of training data in advance and was not effective with acoustic data that has a very short duration.

The main use of the proposed system is to estimate the ball position from the impact sound in an indoor sports simulation. For this, we utilized the TDOA technique to recognize an instant length of the feature source such as the impulsive sound of striking a ball, which only lasts milliseconds. Seo and Kim adopted the TDOA-based technique to estimate the position and direction of an acoustic source using wavefront models based on the distances and degrees between the acoustic source candidates and microphones [16]. In their approach, a DSBF method with many two-dimensional (2D) planes was used to estimate an accurate source position in 3D space. Porteous et al. proposed a 3D position estimation technique by applying a dual position-based beamforming technique without using the candidate 2D planes [17]. However, their approach demanded a large number of microphones, making the system complicated for use in sports simulation. Rizzo et al. proposed a system based on a direction-based scheme [18]. In their approach, various factors such as sound pressure, calibration, sensitivity, and signal-to-noise ratio (SNR) of the microphone array were used for the position estimation process.

In this paper, we introduce a novel acoustic source localization in a 3D space, based on a direction estimation technique. Assuming the distance an acoustic source is from adjacent microphones, its waves spread in a planar form called planar wavefront. In our system, the directions and steering angles between the acoustic source and the microphone array are estimated based on the planar wavefront model using a delay and sum beamforming (DSBF) system and an array of 2D microelectromechanical system (MEMS) microphones. The proposed system is designed with a small number of input parameters such as time delays for candidate degrees between the acoustic source and the microphones, and it is implemented using a cost-effective field programmable gate array (FPGA) and a micro control unit (MCU). The experimental results show that the localization errors of the proposed system are less than 3 cm when an impulsive acoustic source is generated over 1 m away from the microphone array, which is comparable to a position-based system with reduced computational complexity.

One of the main contributions of this paper is to analyze and compare two different types of DSBF algorithms in terms of computation complexity, processing performance, and accuracy to locate the accurate 3D source location of impulsive sounds between the position-based and direction-based methods. Furthermore, we introduce a low-cost high-speed DSBF hardware engine that processes the DSBF algorithm in real time, making our system suitable for use in time critical systems such as sports simulation. On the basis of the high accuracy and real-time performance of localizing an impulsive acoustic source, such as striking a ball, the proposed system is applied to improve the user’s sports skills. As shown in the experimental results, users trained with the simulation program demonstrated noticeable improvements in their soccer kicks and baseball hits as compared with outdoor users.

The remainder of this paper is organized as follows: The direction and position-based estimation methods of an impulsive acoustic source are analyzed in Section 2. The proposed system to locate the 3D position of an impulsive acoustic source based on a geometric model is described in Section 3. An overview of proposed system and the detailed architecture of our hardware design is described in Section 4. The experimental results are presented in Section 5. We conclude this paper with ongoing improvements in Section 6.

2. Impulsive Sound Estimation

2.1. Direction-Based DSBF

Figure 1 shows an example model of the planar wavefront to estimate the direction of the acoustic source using a microphone array that is arranged in a 2D x - y plane. Sound has the property of spreading spherically from the originally generated position. However, when $|r_s|$, the distance between the acoustic source and the microphone array, is far away as compared with the distance between two adjacent microphones, it is assumed that the wavefront of the sound is planar, as shown in Figure 1.

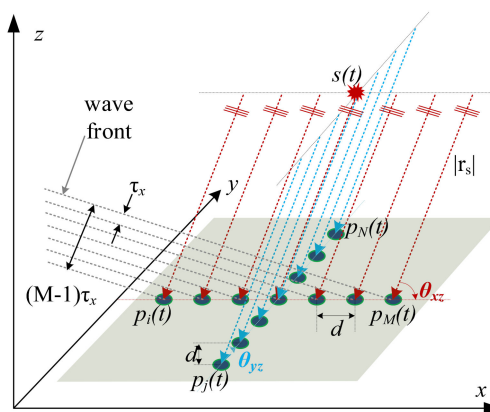


Figure 1. Planar wavefront model on a uniformly arranged microphone array in the x - y plane.

To utilize the planar model, $|r_s|$ should be satisfied as follows:

$$|r_s| > \frac{(d \times M)^2}{\lambda}, \tag{1}$$

where, d is the distance between two adjacent microphones, M is the number of microphones in a single axis, and λ is the wavelength of the sound. The time, τ , which the time it takes for sound waves to pass through microphones separated by d , is defined as $\frac{d \cdot \cos\theta}{c}$, where c is the speed of the sound waves. The measured signal at each microphone in the free acoustic field, where $i = [1, \dots, M]$, is defined as follows:

$$p(t) = s[t - (i - 1)\tau], \tag{2}$$

where, s is the source location in 3D space. If Inequality (1) is satisfied, the planar model-based beamforming output for a single axis at t with respect to candidate steering angles θ is defined as follows:

$$bf(\theta, t) = \frac{1}{M} \sum_{i=1}^M p_i [t - (i - 1)\tau]. \tag{3}$$

For the impulsive sound samples at a sampling frequency, f_s , Equation (3) is rewritten as follows,

$$bf(\theta, k) = \frac{1}{M} \sum_{i=1}^M s[s_k - (i - 1)\tau f_s], \tag{4}$$

where, s_k is the k th sound sample measured at t . The direction-based DSBF for valid impulsive sound samples with length L is expressed as follows:

$$DSBF(\theta) = \frac{1}{L} \sum_{k=1}^L bf(\theta, k). \tag{5}$$

The steering angle, θ_s , for n candidate θ between the acoustic source and the microphone array is defined as follows:

$$\theta_s = \max_n [DSBF(\theta)]. \tag{6}$$

Finally, the unit vector, \vec{v} , is expressed as follows:

$$\begin{cases} v_x = \cos \theta_{xz} \times \sin \theta_{yz} \\ v_y = \cos \theta_{yz} \\ v_z = \sin \theta_{xz} \times \sin \theta_{yz} \end{cases}, \tag{7}$$

where, θ_{xz} and θ_{yz} are the direction between the acoustic source and the microphone array in the x and y axes, respectively.

2.2. Position-Based DSBF

The acoustic source location on a 2D plane is estimated from the spherical wavefront model at the certain depth away from the microphone array. The beamforming output is defined as follows:

$$bf(P_s, s_i) = \frac{1}{M} \sum_{j=1}^M p_j \left[s_i - \frac{f_s}{c} |P_s - P_{mj}| \right], \tag{8}$$

where, s_i is the i th sample from the j th microphone, P_s is the position of the candidate acoustic source, and P_{mj} is the position of the j th microphone. This method requires many more microphones and candidate acoustic source locations than the direction-based DSBF to achieve high accuracy, which results in a higher computational cost, as the time complexity of the beamforming algorithm is defined as $M \times C$, where M is the number of microphones and C is the number of candidate acoustic source locations. This beamforming is one of the general approaches to localize an acoustic source and can be implemented in different ways [16,19].

3. Three-Dimensional Impulsive Acoustic Source Localization

As shown in Figure 2, the 3D impulsive acoustic source position is estimated from the closest point of the two vectors obtained from Section 2. Two identical DSBF models are used to estimate the direction of the impulsive acoustic source. Each model has its own unit vector \vec{v}_1 and \vec{v}_2 . P_{S1} and P_{S2}

are the center positions of each microphone array. The parametric form of the line equation for the two vectors is expressed as follows:

$$\begin{cases} l_1 = P_{S1} + h_1 \vec{v}_1 \\ l_2 = P_{S2} + h_2 \vec{v}_2 \end{cases} \quad (9)$$

where, h_1 and h_2 are the parameters to find P_{E1} and P_{E2} , respectively. They are the closest points between two vectors on each line. The shortest distance between those two straight lines is the length of the vector, \vec{w} , which is perpendicular to both lines. Thus, \vec{w} is expressed as follows:

$$\vec{w} = l_2 - l_1 = P_{S2} + h_2 \vec{v}_2 - (P_{S1} + h_1 \vec{v}_1). \quad (10)$$

Since l_1 and l_2 are perpendicular to \vec{w} , the dot product of the direction vector of each line and \vec{w} is expressed as follows:

$$\begin{cases} \vec{w} \cdot \vec{v}_1 = 0 \\ \vec{w} \cdot \vec{v}_2 = 0 \end{cases} \quad (11)$$

By substituting Equation (10) into (11), Equation (11) is rewritten as follows:

$$\begin{cases} P_{S2} + h_2 \vec{v}_2 - (P_{S1} + h_1 \vec{v}_1) \cdot \vec{v}_1 = 0 \\ P_{S2} + h_2 \vec{v}_2 - (P_{S1} + h_1 \vec{v}_1) \cdot \vec{v}_2 = 0 \end{cases} \quad (12)$$

Thus, the parameter h_1 and h_2 is obtained as follows:

$$\begin{cases} h_1 = \frac{(P_{S2} - P_{S1}) \cdot \vec{v}_1 + h_2 \vec{v}_1 \cdot \vec{v}_2}{\vec{v}_1 \cdot \vec{v}_1} \\ h_2 = \frac{-(P_{S2} - P_{S1}) \cdot \vec{v}_2 + h_1 \vec{v}_1 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} \end{cases} \quad (13)$$

By solving Equation (13), P_{E1} and P_{E2} are calculated as follows:

$$\begin{cases} P_{E1}(h_1) = P_{S1} + h_1 \vec{v}_1 \\ P_{E2}(h_2) = P_{S2} + h_2 \vec{v}_2 \end{cases} \quad (14)$$

Finally, the position of impulsive acoustic source is obtained as follows:

$$P_C = P_{E1} + \frac{P_{E2} - P_{E1}}{2}. \quad (15)$$

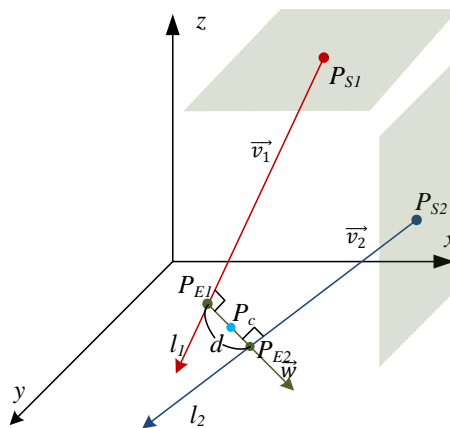


Figure 2. Geometric model to find the closest point of two vectors in a three-dimensional (3D) space.

4. System Overview

An accurate estimation of ball positions using acoustic source localization requires several preprocessing steps as follows: calibration for all MEMS microphones and impulsive sound detection and recognition to distinguish it from surrounding noises and other unwanted sounds. Further details of these steps are described in [10]. Figure 3 shows an overview of the proposed system to estimate the 3D location of an impulsive acoustic source. The proposed system is comprised of the following two steps: an estimation of two individual acoustic source directions, \vec{v}_1 and \vec{v}_2 , which are obtained by DSBF controllers using a MEMS microphone array arranged in 2D space, and an estimation of the closest point of the two vectors by the host processor.

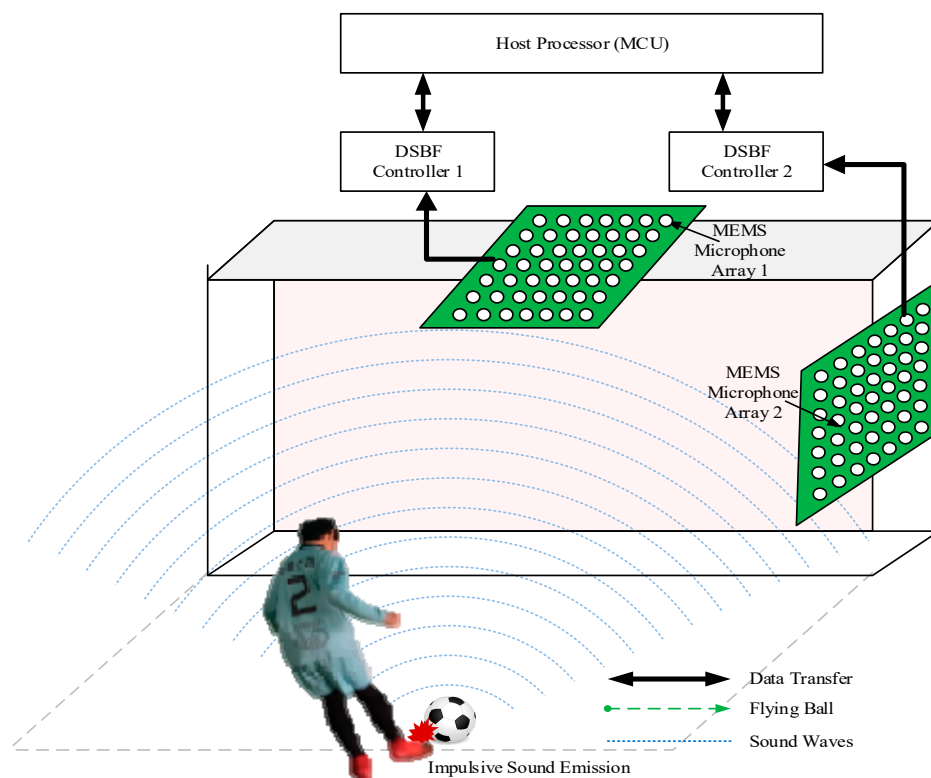


Figure 3. Overview of the proposed system: Acoustic source localization using two delay and sun beamforming (DSBF) systems.

Figure 4 shows the overview of the proposed estimation architecture for 3D acoustic source localization. When a sound is generated and its decibel exceeds a predefined threshold, the impulsive sound detector classifies it as a valid target sound based on a feedforward neural network (FFNN) [20]. If it is classified as a valid sound, micro control unit (MCU) calculates the delay between all the MEMS microphones (i.e., 49 in our system) and candidate degrees as discussed in Section 2. The MCU transfers it to the DSBF unit in the FPGA.

The delay information is calculated only once before processing the DSBF and implementing a double precision \times angle between the sound and the microphones using the input acoustic sources and the delay in parallel for all microphones as described in Figure 5. The USB controller transfers the results of DSBF to a PC to evaluate the estimated acoustic source.

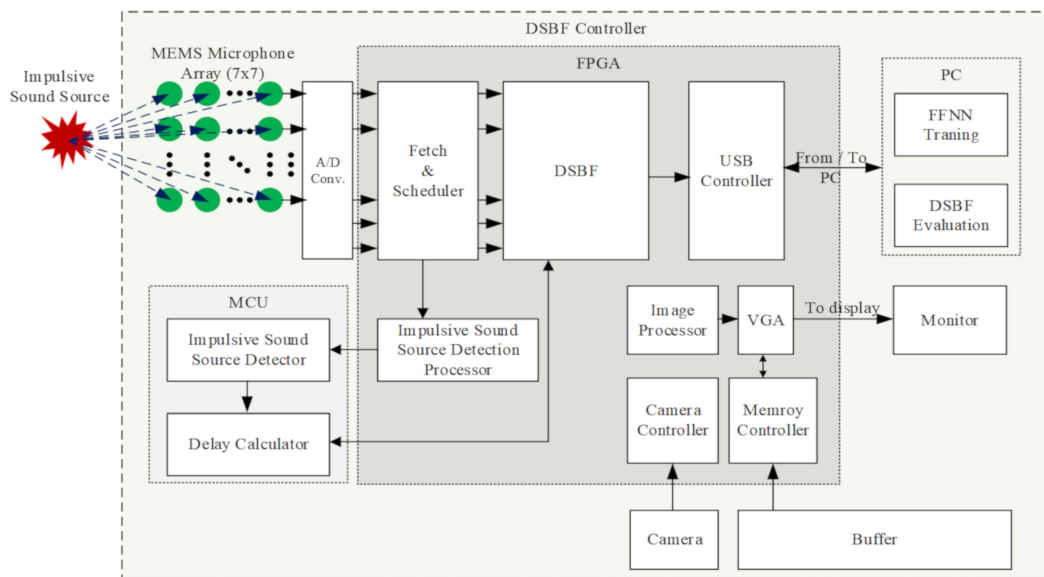


Figure 4. Overview of the proposed evaluation architecture.

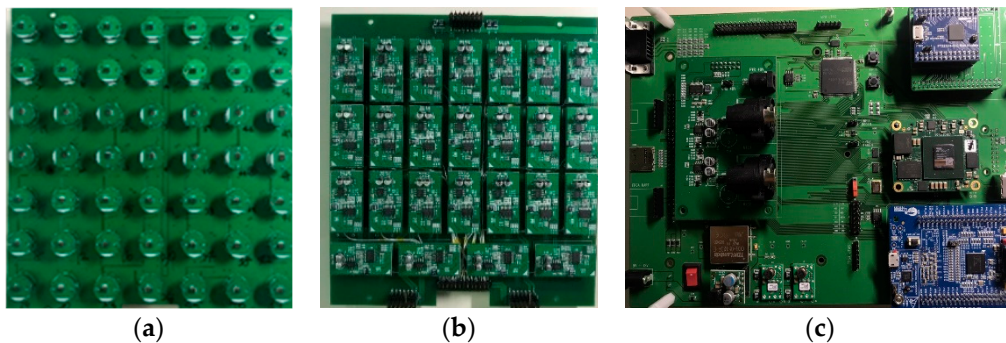


Figure 5. Implementation of the two-dimensional (2D) microelectromechanical system (MEMS) microphone array and the processor: (a) Arrangement of the MEMS microphone array, (b) A/D conversion modules, and (c) field programmable gate array (FPGA-) based controller for impulsive acoustic source localization.

5. Experimental Results

5.1. System Implementation

As shown in Figure 5, our system places 49 microphones, in a 7×7 rectangular array, with 25 mm between two adjacent microphones. An impulse sound is generated at least 1 m away from the microphone array. The distance of 25 mm is close enough to assume that the sound wavefront is planar. Although only 13 MEMS microphones (seven microphones for each axis, sharing the microphone positioned in the center) are required to estimate the direction-based beamforming algorithm in our system, all of the 49 microphones are used to compare the accuracy of the location-based beamforming technique used in [16].

All the signals of the MEMS microphones (Knowles SPH1642HT5H-1 which has dimensions of width $2.65 \text{ mm} \times$ length $3.5 \text{ mm} \times$ height 1 mm , the highest signal-to-noise ratio in the current market, and an analog top type) are inputted to 25 two-channel A/D converters (Cirrus Logic Inc. CS5351). The audio signals are sampled at a rate of 192 kHz for each microphone with a 24-bit resolution. We designed an A/D conversion board by referencing the development kit offered by the manufacturer.

Figure 5 shows the hardware-based implementation of the main controller that processes the DSBF algorithm and manages every device in our system such as 49 MEMS microphones, A/D converters, power management, DDR memory, camera, USB, and others. For a quick and robust

hardware implementation, the circuitry related to the FPGA was implemented by utilizing Trencz Electronic’s TE0713-01-200-2C, which is one of the most cost-efficient modules for implementing the proposed method.

Figure 6 shows a block diagram of the proposed FPGA-based design for processing both position and direction DSBF algorithms. To reduce the resources in the FPGA and boost the performance, the MCU in Figure 4 calculates the time delay between the microphone array and the candidate directions and positions of acoustic source in advance. This design scheme makes it easy to adjust the position and direction resolutions. Furthermore, the proposed design processes all DSBF algorithms in parallel as the DSBF processes for a single microphone are summed independently. The size of each frame, L , is set to 256 as the main target sound is mostly impulsive and lasts only milliseconds.

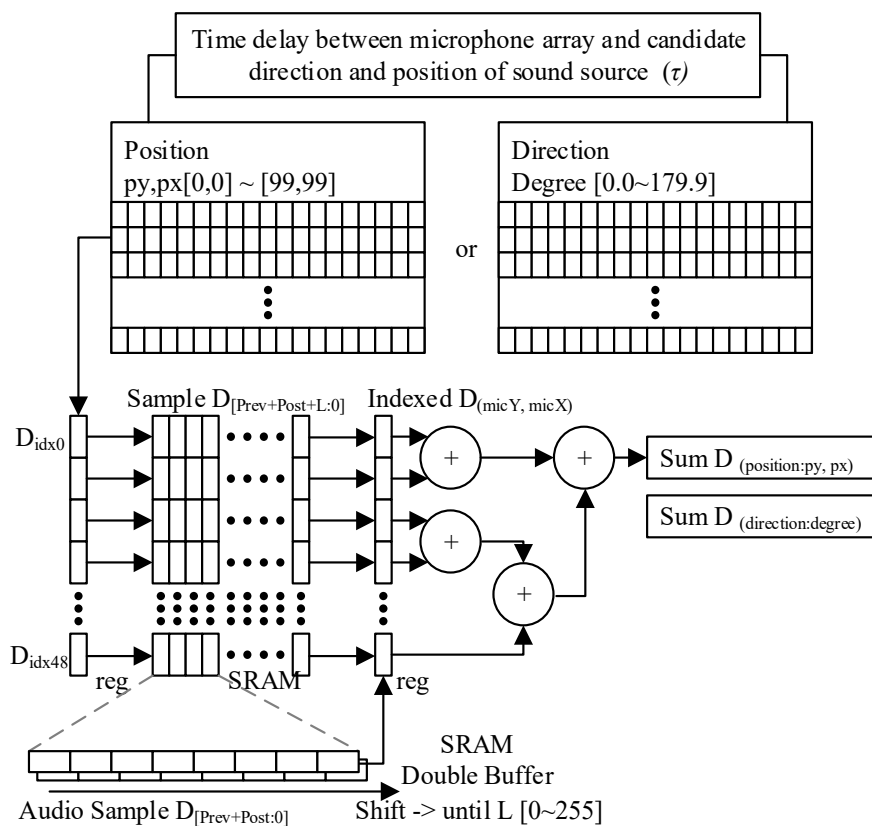


Figure 6. FPGA-based design of the DSBF processor.

5.2. Three-Dimensional Impulsive Acoustic Source Localization

Our experiments were performed in a room with dimensions of width 5 m \times length 7 m \times height 3 m and did not consider the effects of reflection and noises that were not caused by users. Figure 7 shows all input signals from the MEMS microphones array, capturing an impulsive acoustic source (kicking and hitting a ball) for a short period of time (less than 2 ms). As shown in the figure, it is noticeable that the variations of the TDOA from the single impulsive source depends on the position of the microphone. Figure 8 shows the sound pressure level, that is, the decibel sound pressure level (dBSPL), of the selected microphone (microphone 1) in the array. The peak value is approximately 100 dB and the average peak value is approximately 85 dB. Background noises that are slightly less than and equal to the average dBSPL of the estimated impulsive acoustic source were used for the experiments, as shown in Figure 8.

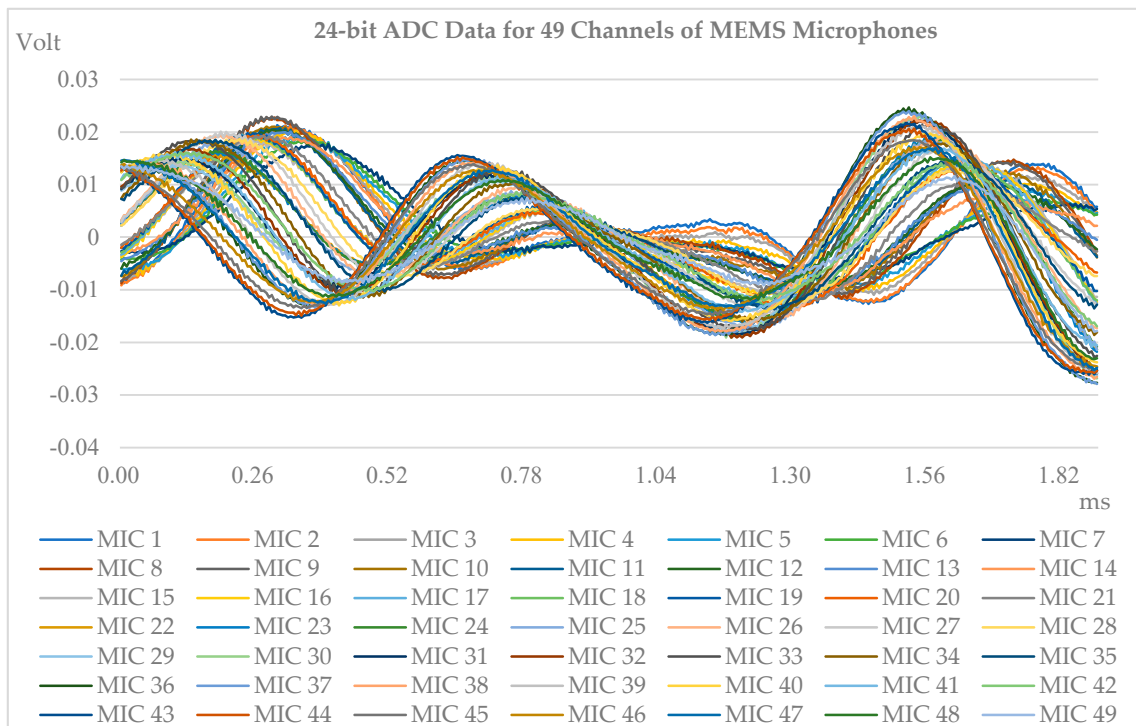


Figure 7. The 24-bit converted input signal voltages of 49 channels of MEMS microphones for an impulsive acoustic source generated by kicking a soccer ball.

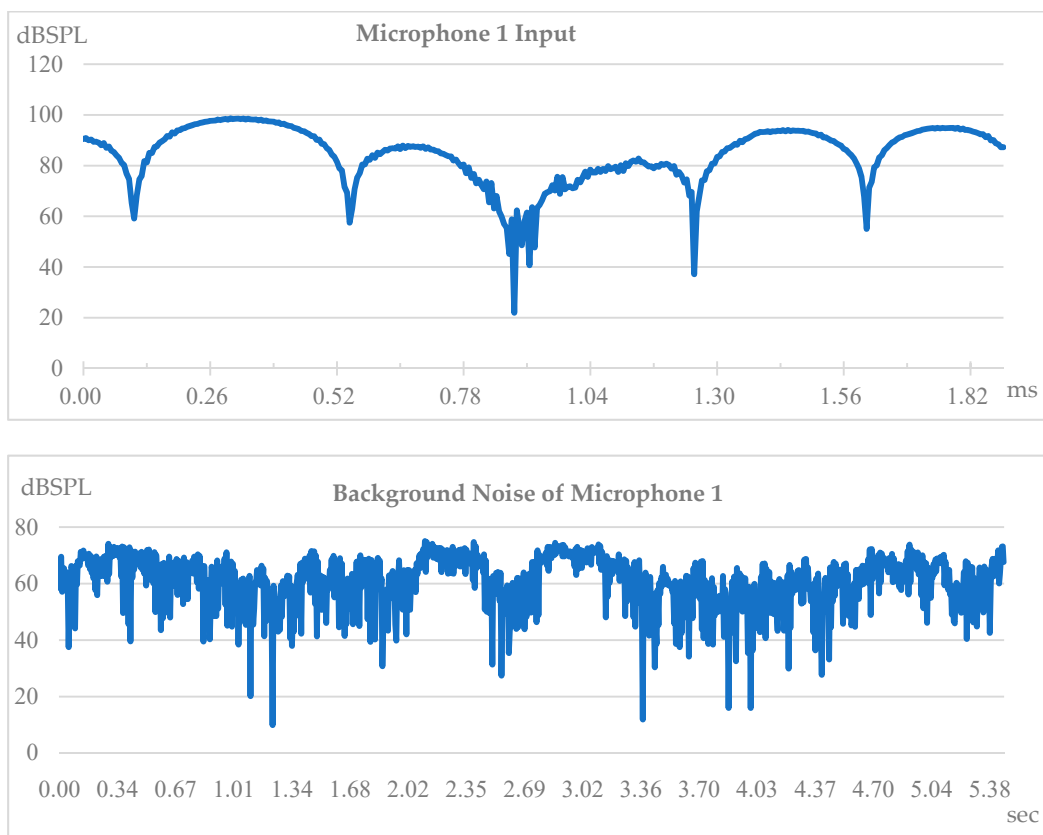


Figure 8. Acoustic source input (top) and background noise (bottom) of microphone 1.

Figure 9 shows the accuracy of the beamforming results with background noises which are ambient white noises caused by users and impulsive acoustic sources such as clapping and footstep sounds. To evaluate the beamforming accuracy, BA , Equation (5) is rewritten as follows:

$$BA = 255 \times \left(\frac{DSBF(\theta) - \min[DSBF(\theta)]}{\max[DSBF(\theta)]} \right). \tag{16}$$

It shows that even in a noisy environment, if the average level of the background sound is smaller than the average value of the ball impact sound, the noise sources do not affect the direction estimation results.

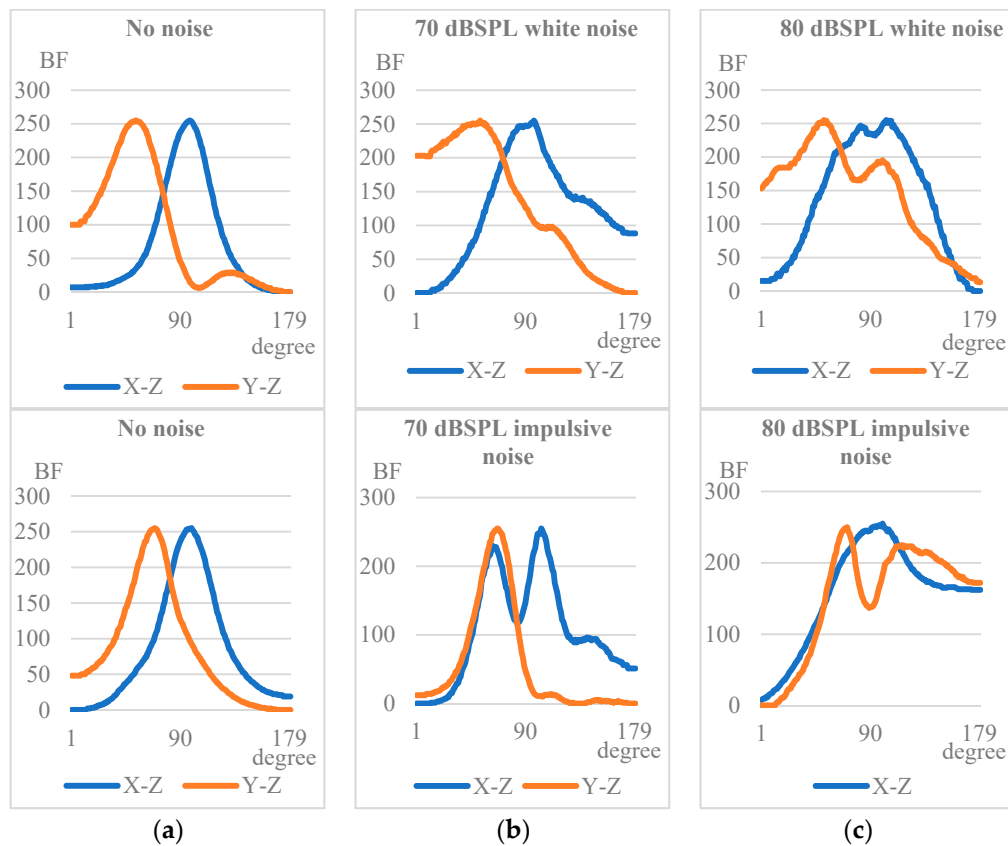


Figure 9. The results of proposed beamforming engine with background ambient white noise (top) and impulsive sound (bottom): (a) Degree of beamforming results under 40 dB background noise, (b) degree of beamforming results about 70 dB background noise, and (c) degree of beamforming results over 80 dB background noise. Here, X-Z is θ_{xz} between the x and z axis, and Y-Z is θ_{yz} between the y and z axis as shown in Figure 1.

Figure 10 shows the estimation results of the proposed 3D acoustic source localization engine for both direction- and position-based methods. The estimated position errors are evaluated by the differences between the position (P_O in Figure 11) of the true impulsive acoustic source generated by kicking a ball and the position (P_D and P_P in Figure 11) estimated by the proposed DSBF engine. For the direction-based DSBF method, the error values get smaller as the distance between the microphone array and the acoustic source becomes larger. This is because $|r_s|$ becomes much larger than $\frac{(d \times M)^2}{\lambda}$ when the distance is larger, making the wavefront more planar and resulting in a higher accuracy for our system. The spectrum of the impulsive sound by striking a ball is in the range of 100 Hz to 10 kHz. In our experiment, the right part of Inequality (1) does not exceed 0.9 m since d is 0.025 m, M is 7, and λ is 0.034 m, which is the shortest wavelength in the spectrum. Therefore, the accuracy of the estimated position is high when the impulsive sounds are generated 1 m away from the microphone

array, as shown in Figure 10. However, if the acoustic source is generated too far away from the microphone array, the SNR significantly decreases, and the errors increase. For the position-based method, the errors are estimated using 49 microphones with a given predefined depth plane. Therefore, the errors remain constant until the SNR significantly decreases.



Figure 10. Errors of the estimation results.

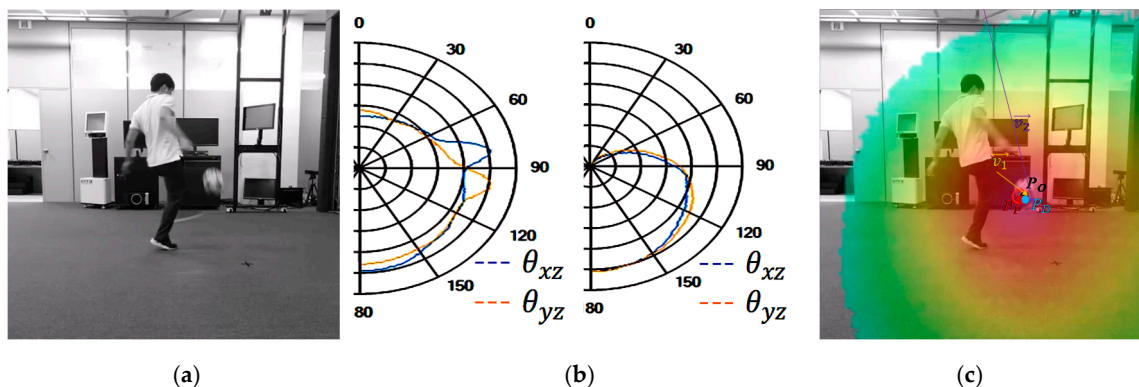


Figure 11. Results of the 3D acoustic source localization method: (a) Generation of impulsive acoustic source by kicking a ball, (b) the results of DSBF direction on the top plane (left) and on the side plane (right), and (c) the results of the conventional spherical model-based method and the proposed dual planar model-based method.

Figure 11 shows the results of our system with the generated impulsive acoustic source. Here, P_O is the true position of impulsive sound, P_P is the position estimated by spherical waterfront model-based DSBF, and P_D is the position estimated by the proposed dual planar waterfront model-based DSBF engine using \vec{v}_2 and \vec{v}_1 , which are the unit vectors calculated from the results of DSBF directions on the top and side planes, respectively.

5.3. Screen-Based Sports Simulation

The proposed method of impulsive acoustic source localization is applied to a virtual sports simulator and operates with [4] to estimate ball motion such as 3D speed, spin, and fire angle. Diverse

software content was developed to train a user’s sports skills such as soccer kicks and baseball hits, as shown in Figure 12.

Figure 13 shows a screen-based simulation for training soccer kicks. For this, a virtual simulation room was equipped with a ball motion simulator, a collision coordinate recognizer, an air cleaner, an air conditioner, and an automatic ball collector to support the training session. For evaluating the effectiveness of the proposed system, a total of 24 students in an elementary school were recruited and divided into two groups: a virtual sport (13 persons) training group and an outdoor (11 persons) training group. The physical characteristics of the participants are described in Table 1. A total of 15 training sessions were conducted for six months on the four evaluation items of soccer skills: target accuracy, kick speed, kick power, and outdoor kick accuracy.



Figure 12. Screen-based sports simulation: Software contents (top) and room (bottom) to train a user’s sports skills such as soccer kicks and baseball hits.



Figure 13. Screen-based simulation for training soccer kicks.

Table 1. Physical characteristics of the training groups

Training Group	Height (cm)	Weight (kg)	Age (years)	Body Mass Index (BMI) (kg/m ²)
Virtual sport	179.9 ± 7.5	39.0 ± 7.8	11.8 ± 1.0	17.6 ± 2.0
Outdoor	151.5 ± 9.7	45.2 ± 10.2	11.4 ± 3.4	19.5 ± 2.7

We used the repeated measure analysis of variance (ANOVA) to evaluate the target accuracy, kick speed, and the accuracy of kick speed. For the analysis of kick speed, kick power, and outdoor kick accuracy, a two-way repeated measure ANOVA was used to evaluate the interaction of virtual sport program effect and practice period.

Figure 14 shows a comparison of the differences between the pre and post experiments for the kick speed, the kick power, and the outdoor kick accuracy, which were verified by the paired *t*-test. During the training period, the target accuracy score, the kick speed accuracy, and the kick speed increased significantly for the virtual sport group, as shown in Figure 14. The training results show that the screen-based simulation provides a more effective environment where the participants focus more on improving their hitting accuracy and kick speed as compared with the outdoor training group. The experimental results for the sports simulation are available in an online video located at https://drive.google.com/open?id=12D9VFpEAS8nwsn41VOOA_d6U0IYI2Qfp. Additional results are located at <https://drive.google.com/open?id=1Mx3FO175Cw6oWHU0WzUx9czvqieQiwYE>.

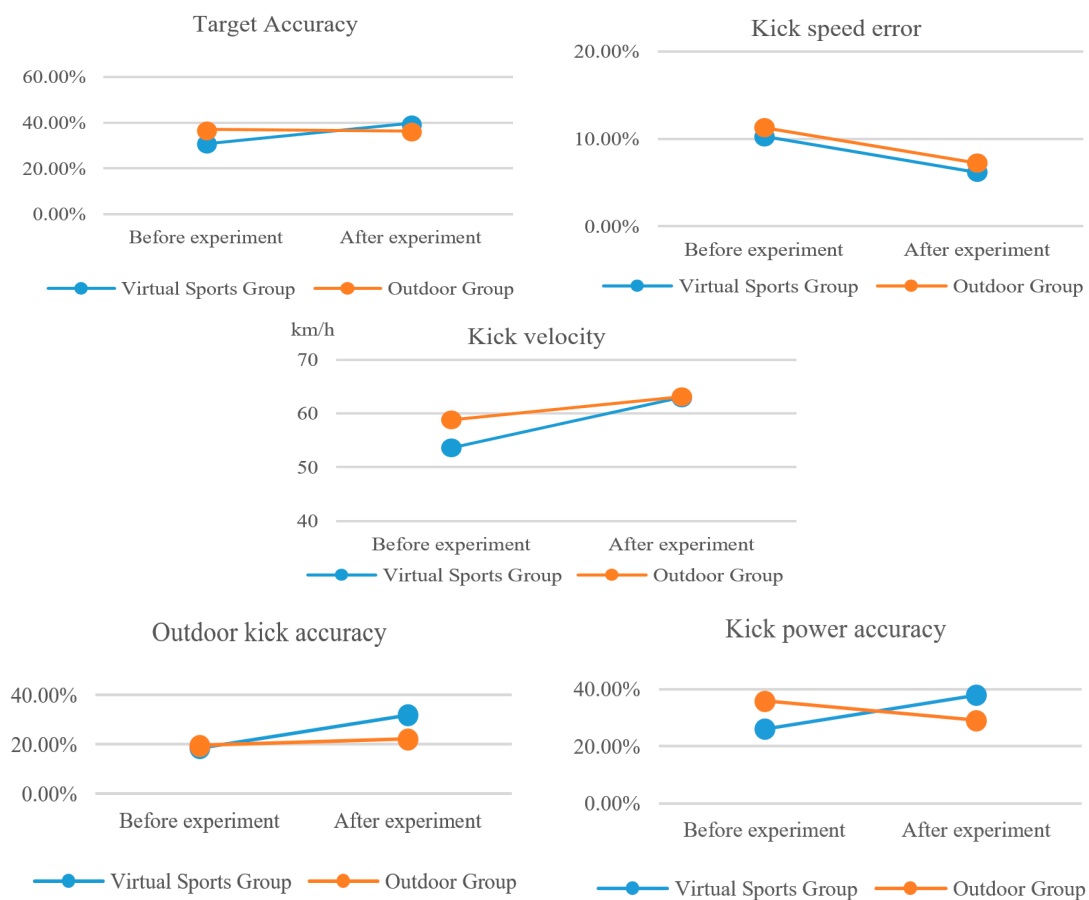


Figure 14. Comparison of training results between the virtual sport and the outdoor group.

6. Conclusions

This paper introduced an efficient method of localizing an impulsive acoustic source, which was used to simulate screen-based ball sports. In the proposed system, a small set of microphones are arranged into a 2D MEMS array (13 for each array), and the system estimates a ball position from the acoustic source in 3D space using the direction-based DSBF algorithm in the temporal domain. Compared to other spherical model-based DSBF systems, our system generates relatively small errors (i.e., less than 3 cm from an acoustic source generated over 1 m away from the microphone array) in tracking ball motion in real time. Furthermore, our system design and implementation require a small space for hardware and can be applied to a variety of real-time simulators. As demonstrated in the experimental results, our system effectively improves a user's sports skills such as soccer kicks and baseball hits in a given training period.

The current system does not consider the effects of sound reflection from the obstacles in the room. Since the effective wavelength is very short, the surrounding environment has little effect on the impulsive sound. Therefore, if there are no direct obstacles between the microphone array and the sound source, the results are expected to be similar. In addition, our system does not consider high-level noises from the environment. For example, if the background noise is louder than the acoustic source, it is difficult for the system to localize the acoustic source without a sophisticated noise filtering method. We are currently improving the accuracy of ball position estimation by integrating an active noise cancellation [21] and a room sound wave reflect-aware scheme [22] in the existing system.

Author Contributions: Conceptualization, S.-W.S., M.S., and Y.K.; data curation, S.-W.S. and S.Y.; formal analysis, S.-W.S., S.Y., M.S., and Y.K.; funding acquisition, M.-G.K.; investigation, S.-W.S., S.Y., M.S., and Y.K.; methodology, S.-W.S. and Y.K.; project administration, M.-G.K.; resources, S.-W.S. and S.Y.; software, S.-W.S.; supervision, M.-G.K., M.S., and Y.K.; validation, S.-W.S., S.Y., M.-G.K., and Y.K.; visualization, S.-W.S.; writing—original draft, S.-W.S. and Y.K.; writing—review and editing, S.-W.S., M.S., and Y.K.

Funding: This research was supported by the Sports Promotion Fund of Seoul Olympic Sports Promotion Foundation from the Ministry of Culture, Sports and Tourism and by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) (No. 2017R1C1B5017000).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kang, D.; Kim, J.-E. Fine, Ultrafine, and Yellow Dust: Emerging Health Problems in Korea. *J. Korean Med. Sci.* **2014**, *29*, 621–622. [[CrossRef](#)] [[PubMed](#)]
2. Whiting, E.; Ouf, N.; Makatura, L.; Mousas, C.; Shu, Z.; Kavan, L. Environment-Scale Fabrication: Replicating Outdoor Climbing Experiences. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI '17, Denver, CO, USA, 6–11 May 2017; pp. 1794–1804.
3. Lee, H.-G.; Chung, S.; Lee, W.-H. Presence in virtual golf simulators: The effects of presence on perceived enjoyment, perceived value, and behavioral intention. *New Media Soc.* **2013**, *15*, 930–946. [[CrossRef](#)]
4. Kim, J.; Kim, M. Smart vision system for soccer training. In Proceedings of the 2015 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, 28–30 October 2015; pp. 257–262.
5. Wang, S.; Xu, Y.; Zheng, Y.; Zhu, M.; Yao, H.; Xiao, Z. Tracking a Golf Ball With High-Speed Stereo Vision System. *IEEE Trans. Instrum. Meas.* **2018**, *68*, 2742–2754. [[CrossRef](#)]
6. Li, B.; Sun, B.; Chen, C.F.; Jiao, X.J.; Zhang, S.Y.; Wang, Y. Simulation of Golf Realtime Tracking Based on Doppler Radar. *Appl. Mech. Mater.* **2015**, *743*, 828–835.
7. Martin, J.J. Evaluation of Doppler Radar Ball Tracking and Its Experimental Uses. Ph.D. Thesis, Washington State University, Pullman, WA, USA, 2012.
8. Burguera, A.; González, Y.; Oliver, G. Sonar Sensor Models and Their Application to Mobile Robot Localization. *Sensors* **2009**, *9*, 10217–10243. [[CrossRef](#)] [[PubMed](#)]
9. Heilmann, G.; Meyer, A.; Döbler, D. Beamforming in the Time-domain using 3D-microphone arrays. In Proceedings of the XIXth Biennial Conference of the New Zealand Acoustical Society, Auckland, New Zealand, 27–28 November 2008.

10. Seo, S.-W.; Kim, M.; Kim, Y. Optical and Acoustic Sensor-Based 3D Ball Motion Estimation for Ball Sport Simulators. *Sensors* **2018**, *18*, 1323. [[CrossRef](#)] [[PubMed](#)]
11. Ishi, C.T.; Chatot, O.; Ishiguro, H.; Hagita, N. Evaluation of a MUSIC-based real-time sound localization of multiple sound sources in real noisy environments. In Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA, 10–15 October 2009; pp. 2027–2032.
12. Kim, Y.-H. Acoustic Holography. In *Springer Handbook of Acoustics*; Rossing, T.D., Ed.; Springer New York: New York, NY, USA, 2014; pp. 1115–1137. ISBN 978-1-4939-0755-7.
13. Loesch, B.; Uhlich, S.; Yang, B. Multidimensional localization of multiple sound sources using frequency domain ICA and an extended state coherence transform. In Proceedings of the 2009 IEEE/SP 15th Workshop on Statistical Signal Processing, Cardiff, UK, 31 August–3 September 2009; pp. 677–680.
14. Yalta, N.; Nakadai, K.; Ogata, T. Sound Source Localization Using Deep Learning Models. *J. Robot. Mechatron.* **2017**, *29*, 37–48. [[CrossRef](#)]
15. Suvorov, D.; Dong, G.; Zhukov, R. Deep residual network for sound source localization in the time domain. *arXiv* **2018**, arXiv:1808.06429.
16. Seo, S.-W.; Kim, M. 3D Impulsive Sound-Source Localization Method through a 2D MEMS Microphone Array using Delay-and-Sum Beamforming. In Proceedings of the 9th International Conference on Signal Processing Systems, Auckland, New Zealand, 27–30 November 2017; pp. 170–174.
17. Porteous, R.; Prime, Z.; Doolan, C.J.; Moreau, D.J.; Valeau, V. Three-dimensional beamforming of dipolar aeroacoustic sources. *J. Sound Vib.* **2015**, *355*, 117–134. [[CrossRef](#)]
18. Rizzo, P.; Bordoni, G.; Marzani, A.; Viperman, J. Localization of sound sources by means of unidirectional microphones. *Meas. Sci. Technol.* **2009**, *20*, 055202. [[CrossRef](#)]
19. Zimmermann, B.; Studer, C. FPGA-based real-time acoustic camera prototype. In Proceedings of the 2010 IEEE International Symposium on Circuits and Systems, Paris, France, 30 May–2 June 2010; p. 1419.
20. Paulraj, M.P.; Yaacob, S.B.; Nazri, A.; Kumar, S. Classification of vowel sounds using MFCC and feed forward Neural Network. In Proceedings of the 2009 5th International Colloquium on Signal Processing Its Applications, Kuala Lumpur, Malaysia, 6–8 March 2009; pp. 59–62.
21. Liebich, S.; Fabry, J.; Jax, P.; Vary, P. Time-domain Kalman filter for active noise cancellation headphones. In Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO), Kos, Greece, 28 August–2 September 2017; pp. 593–597.
22. Tervo, S.; Pätynen, J.; Lokki, T. Acoustic Reflection Localization from Room Impulse Responses. *Acta Acust. United Acust.* **2012**, *98*, 418–440. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).