Mohanad Aljanabi, Ammar O. Abdallh, Mohammed Jamal Mohammed, et al.

View Online   Export Citation

ARTICLES YOU MAY BE INTERESTED IN

Multi-camera Tracking Technique of a Speaker Based on Arduino

Mohanad Aljanabi[1, a)], Ammar.O. Abdallh[1, b)], Mohammed Jamal Mohammed[2, c)] and Bushra Majid Hameed[1, d)]

[1]AL- Mausaib, Department of Electrical Power Techniques Engineering,Technical College, Al-Furat Al- Awsat Technical University Najaf, Iraq

[2]College of Medicine, University of Al-Ameed-Karbala, Iraq.

a) Corresponding author: com.mhn@atu.edu.tr.
b) ammar.o.abdallh@atu.edu.iq
c) Eng.mohammed.j@alameed.edu.iq
d) basharn273@gmail.com

Abstract: In this aera of smart applications, multi-camera tracking technique has established itself as an emerging technology that is being popular day by day. Addition identification of the active speaker using camera assisted microphone array. The possibility of identifying the active speaker component and the camera automatically turns towards the sound source. The PZT cameras work not limited to the human voice, but it is sensitive to all the vibration of sound and works on the camera and audio recording is not observed, but it can record video conference. It has been programmed to save the information about the speaker or the sound source and the ability of displaying it at the anther time if the need arises. According to the option of night vision by adding an infrared sensor. This type of sound sensitive surveillance cameras is of a great important and urgent need, because of its benefit commensurate with all requirement. The camera tracking has developed pervious research on surveillance cameras, with an external cover suitable for environmental changes, and a few very high efficiencies. Assessment by user and independent codes suggests promising initial results and future work. We associate the projected 3D tracker against anther procedures on the all database and on an anew collected data group with sensors, which we variety presented to the investigation public. A speaker position combined with camera image protects effective and accurate tracking structure. The scheme was applied Arduino microcontroller for image processing and development of tracking system and increasing the accuracy of the tracking of living organisms.

Keywords: - Arduino, Automated camera, Different sensor, Image Signal processing, Speaker tracking.

INTRODUCTION

Automatic strategies are universal these years, with the advance of computer vision skill and complexity cameras [1].Speaker tracking is main portion of multi applications which supports us to control the speaker lines and evaluates the behaviour of speakers. Mostly there are 3 wide-ranging tracking procedures for example sensor, vision based and a speaker method. These systems get the capacity to multi- camera to save depend on the speaker. For the sensor [2,3,4] methods the speaker must be uniform IR sensor to lead electric indicators, wherever the receiver element apply the signal to detect the speaker location, this resources that an additional component around the body which is maddening and uncomfortable[5,6]. These performances get the capacity to camera to save absorbed on the speaker individual [6]. Low-cost deepness cameras have been investigated for centuries with anew obtainable systems, and the advance of software and processes for them improves their purposes, the statement of the multi-cameras lets extra and more people to love the advantage of this knowledge [7]. Passive infrared sensor (PIR) , is a searching used generally to distinguish a human program to someplace and different the sensitive red energy (IR sensor) this is for the reason that IR sensor does not comprise an IR sensor and level accept these emissions from the surrounding matters . The combination of audio- video indicators has singing an identical regulation for calculation intelligence in the arena of engineering and observation applications.

Communication that anybody with a temp. greater than zero spreads IR energy and the extra radiation rises [8]. The scheme of [9] estimates each suggestion by building a cuboid positioned on each theory and measured to estimate the graphic presence of the people. Used for apiece camera opinion, this cuboid is predictable interested in

the image, the full forefront bulk [10] within the predictable region, approached for speed by a rectangle. From figure 1. shows the inside PIR structure and figure 2. IR assemblage about the sensitive.
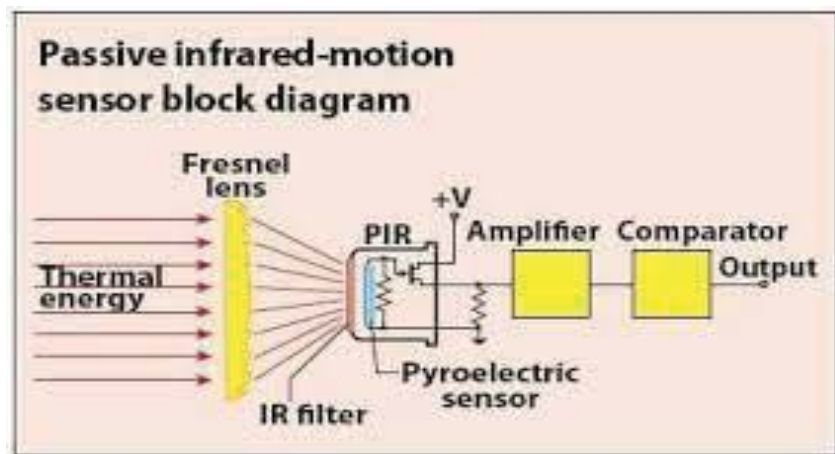

FIGURE 1. The inside PIR structure [11].
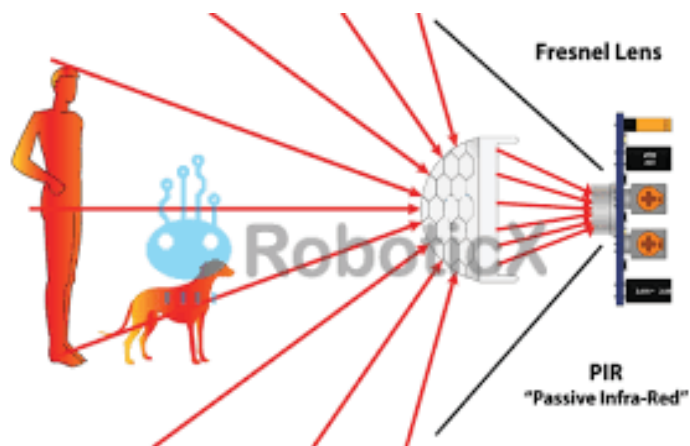
FIGURE 2. IR assemblage about the sensitive [12].

Arduino is applied mostly the strategy of electronic plans that purpose to use various conservational sensors. Arduino can be associated to various programs on a PC, and its programming depends on open language [13]. From figure 3 illustrations the inside connection of Arduino. The software design codes are alike to the C language and are measured one of the informal program designs applied to inscribe microcontroller.

FIGURE 3. Inside connection of Arduino [13].

This article was focused to the proposal of a rotating camera with the speaker voice. In this scheme have been rising the camera determination feature, it developed less possible to interfere with the other sound waves.

Program the camera and make it interaction the smart speaker voice to guide alarms when fault or an unsuitable happening arises, as a faint of the quantified person. Development of the PZT tracking cameras where it has developed up to 360°.Security control of the camera by addition ciphers, code to troubleshoot the new work of registration and tracking correctly and safety the option of storage data with all.

METHODOLOGY

At each border of 3Dmulti-camea tracker is used, and we apply the resultant 3D probabilistic representations to regulate the location of the body. These 3D article opinions are measured for example theories for the upper of the speaker, after joined with the multi-camera adjustment material, each gets a 3D ray lengthways which speaker.
We search for justification for these theories in additional assessments through calculating the shortest distance d [$r1$, $r2$] for each couple (r1, r2) of ray's various cameras [14]. Completely such couplings are estimated, under control and associated to a distance threshold (D =700mm), by the bordering competition measured first. The resulting process is applied: -
1. Begin with a pool of ray theories for apiece camera.
2. If d [$r1r2$] < D make a 3D theory at the medium (r1, r2) of the shortest D.
3. Quest the residual cameras for rays $r3$ that permit within D, change the theory to (r1, r2, r3…).
4. Supply the theory $Xi(t)$ and eliminate of 3D theory.
5. Recurrence from stage 2.
    From figure 4. the circuit diagram for details in electronics circuit and figure 5. Shows that Arduino process and feature controls classification of signal voice.
Figure 6. Illustrations that the setup of 4-camera Arduino microcontroller with 3-D tracking**.**
The steps of segmentation for multi-camera tracking technique of a speaker with algorithm below: -
  1- Read segmented image.
  2- Use the resulting process to each image with 4 cameras.
  3- /* Scan the image row by row*/.
  4- for $i = 1$ to h do.
  5- for $j = 1$ to w do.
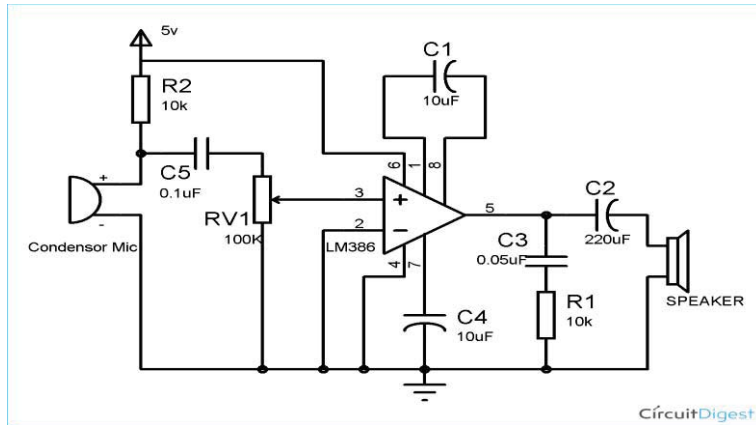 6- if $D(i, j) > Tb$ then.
  7- end for.

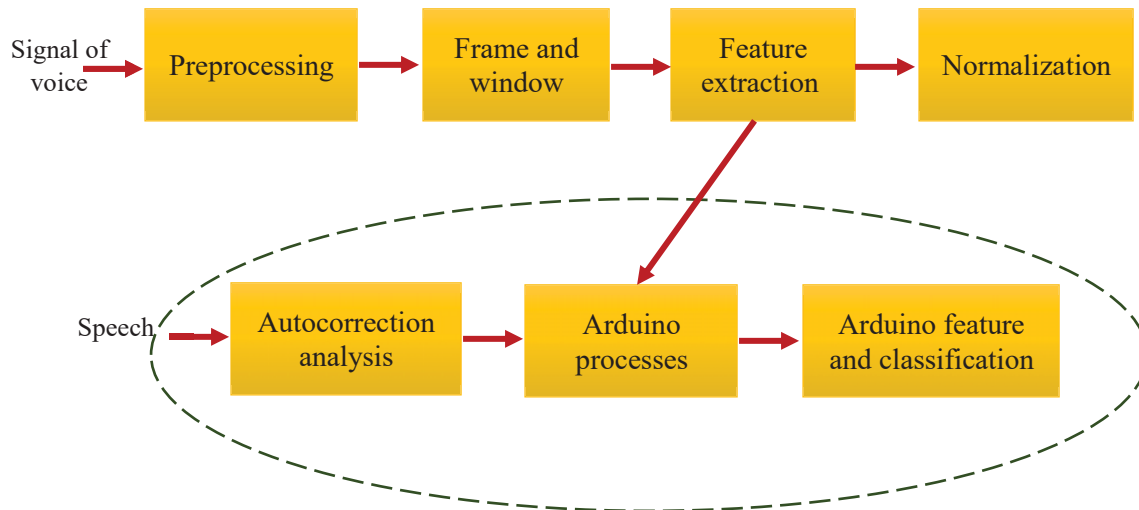**FIGURE 4**. The circuit diagram for details in electronics circuit.

FIGURE 5. Arduino process and feature controls classification of signal voice.

FIGURE 6. Setup of 4-camera Arduino microcontroller with 3-D tracking [15].

RESULTS AND DISCUSSION

The speaker tracking camera tries to detect the voice in the any position. It resolves o/p a subcategory of the original image comprising the anther sound. The discussions of each database are separated into segments for advance and assessment. Outcomes are obtainable on the assessment parts. Figure 7 shows design of the locations of multi-camera tracking applied for assortment of the examination information of optimal position of a speaker. From figure 6. design   the locations of multi-camera tracking applied for assortment of the examination information of optimal position of a speaker.
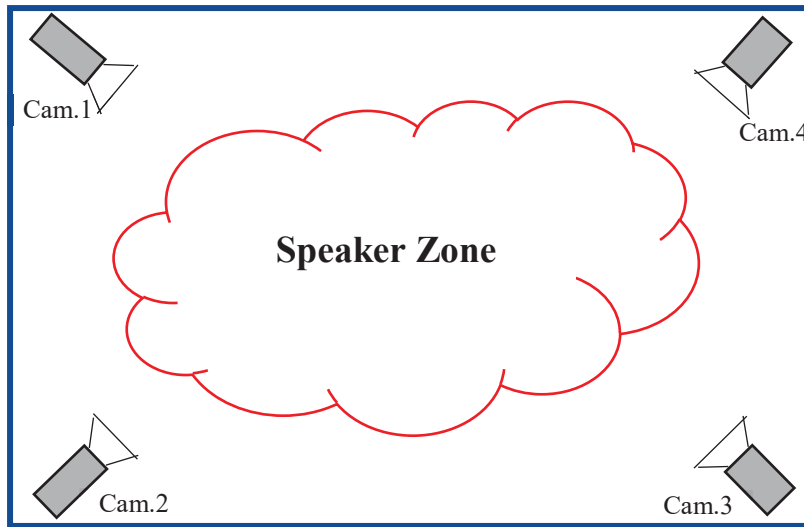

FIGURE 7. Design of the locations of multi-camera tracking applied for assortment of the examination information of optimal position of a speaker.

A healthy initialization and re-initialization process give this scheme, as a result of its exactness and its informal development to several extra degrees of freedom. A mixture, applying the Median filtering method for recognition and initialization and control-configuration for tracking achievable.

The equilibrium between put on the tracking present theories and discovering the space for original theories is complicated if the quantity of particles is to be reserved within "real-time" bounds, and the instrument for accomplishing this equilibrium wants additional work to reach good tracking outcomes. The originate position of the speaker is filtered [16]: -

$$\text{a) } x' = \alpha x * x_t + (1 - \alpha x) * x_{t-1} \quad\quad\quad (1)$$
$$\text{b) } y' = \alpha y * y_t + (1 - \alpha y) * y_{t-1} \quad\quad\quad (2)$$

The camera saves a minor buffer of where the speaker has been to set a suitable zoom side by side. If the speaker changes out of the border, despite the fact, the camera can gradually get an evolution cut to a summary shot then back to the speaker. This problem is giving a lecture using the histogram scheme designated above. Despite the fact there are various tracking processes that may be working. The purpose of this research is to advance an intelligent audio-based camera tracking scheme.

The camera rotation is measured by an exact voice, the application of a speaker signals to track persons are flattering critical, the integration of audio signals has in performance a very significant rule for addition intelligence in the field of investigation applications [14]. As an impermeable of knowledge, our scheme works. It can track a speaker within 120^0 of their position in less than 43 sec, which is fit within the assortment of the camera. it similarly makes available a good determination demonstration from the cameras and does good work of disregarding unnecessary noises, which is central in defining who is speaking.

TABLE 1. Survey results for speaker tracking quality with different parameters of overall perception of the system.

Study Session	Human Operator			Our System		
	Mean	Median	St. dv.	Mean	Median	St. dv.
Field of test camera 1 (Cam1)	3.19	3.00	0.83	2.65	2.50	0.88
Field of test Cam 2	3.50	3.50	0.53	2.87	3.00	0.83
Field o test Cam 3	3.11	3.00	0.88	2.67	3.00	1.02
Field of test Cam 4	4.00	4.00	0.53	3.00	3.50	1.20

The resulting idea we removed from consideration was the static camera scheme. It required multi-camera and microphones to be used, remote supplementary than through the spinning scheme.

CONCLUSION AND FUTURE WORK

We become to apply low-priced mechanisms which make the scheme powerless to distinguish a low volume speaker as a result of their minor signal to noise ratio also our scheme for sending all of the audio and image data to the computer includes several wires and an A/D which is slightly unworkable and expensive.

We want some technique of distinguishing as soon as there is instantaneous speaker accordingly that the camera oscillation.

Detection and recognition of continuous actions from video is a core problem to report for enabling intelligent schemes that can extract and achieve satisfied wholly automatically. A speaker position based on the energy from apiece sensor which suggests the uses of ability to speak window to control the energy the dimensions in chief such that the scheme work is unchanging for short period stoppage.

This system assures agile and easy to use surveillance model with a lot of accuracy and more scope to integrate new and innovative techniques in future.

REFERENCES

1. Johnson et Al., Computer street microphone arrays for sound transduction in large rooms, *Acoustical Society of America,* Vol. 78, No. 5, PP. 1508 -1518, 1985.
2. S. Mukhopadhyay, B. Smith., Passive Capture and Structuring of Lectures, *Proc. of ACM Multimedia*, 1999.
3. P. Kumar G. and S. M, Real Time Detection and Tracking of Human Face Using Skin Color Segmentation and Region Properties, *International Journal of Signal Processing Systems*, Vol. 2, No. 2, December 2014.
4. M. K. Leung and Y. H. Yang, , First sight: A human body outlines labeling system. *IEEE Trans. on PAMI,* 17(4):359–377,1995.
5. Q. Cai and J. K. Aggarwal, , Tracking Human Motion Using Multiple Cameras, *Computer Vision and image understanding*, volume 73, issue 3, ,9 march1999.
6. D.J. Lee, P. Zhan, A. Thomas, R. Schoenberger, , Shape-based Human Intrusion Detection, *International Symposium on Defense and Security*, vol. 5438, April 12-16, 2004.

7. Harvey F. Silverman and Stuart E. Kirtman, A two -Stage algorithm for determininning talker location from linear microphone array data, *Computer Speech and language*,6, 129- 152, 1992.

8. Chan Boon Cheng, Asral Bahari Jambek*, ,*Implementation of a camera system using noise II on the altera DE2-70 board, *Indonesian Journal of Electrical Engineering and Computer,* Vol 14, No 2, *May 2019*.

9. Kai Nickel, Tobias Gehrig, Rainer Stiefelhagen, and John McDonough, A joint particle filter for audiovisual speaker tracking. In *International Conference on Multimodal Interfaces ICMI 05*, 2005.

10. C. Stauffer and W. E. L. Grimson, Adaptive background mixture models for real-time tracking. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Fort Collins, CO,* June 23-25, pages 246–252, 1999.

11. Reda Hanifi Elhachemi Amar, Laredj Benchikh, Hakima Dermeche, Ouamri Bachir, Zoubir Ahmed-Foitih*, ,* Trajectory reconstruction for robot programming by demonstration, *International Journal of Electrical and Computer Engineering (IJECE),*Vol 10, No 3,June 2020.

12. H. Hatimi, M. Fakir, M. Chabi, M. Najimi*,* New Approach for Detecting and Tracking a Moving Object, *International Journal of Electrical and Computer Engineering* (*IJECE),*Vol 8, No 5, October 2018 .

13. H.R. Hatem, J.N. Shehab and I. Abdul-Rahman, ARDUINO Microcontroller Based Building Security System, *Engineering and Technology Journal*, Vol. 35, Part A, No. 5, pp. 532-536, 2017.

14. A.W. Senior, G. Potamianos, S. Chu, Z. Zhang, A. Hampapur, A Comparison of Multicamera Person-Tracking Algorithms, In *International Conference on Multimodal Interfaces ICMI 05,* 2015.

15. Q. Liu, Y. Rui, Anoop Gupta and J. Cadiz, Automating Camera Management for Lecture Room Environments, Special Interest Group on Computer-Human Interaction, 2001.

16. X. Zhong and A. B. Premkumar. Particle filtering approaches for multiple acoustic source detection and 2-D direction of arrival estimation using a single acoustic vector sensor. IEEE Transactions on Signal Processing, 2020.