General Sir John Kotelawala Defense University

Faculty of Computing

Department of Computer Science

Group Project Undertaken in partial fulfillment of the requirement for the

BSc Information Technology Computer Science/ Computer Engineering/ Software Engineering Degree

Intake 36

PROJECT PROPOSAL

Group Details		
Group Number	**Student Number**	**Student Name**
Group 1	D/BCE/19/0002	AMRNVB Pethiyagoda
	D/BCS/19/0003	EMHI Ekanayaka
	D/BSE/19/0009	TA Gamage
	C/BCE/18/5711	WDDA Gunawardhana
Project Details		
Project Title	Rotating Camera for Video Conferencing Based on Speaker Voice	
Supervisor	Mr. DMR Kulasekara	

Contents

PROJECT PROPOSAL 1

1. Introduction 3

1.1 Background and Motivation 3

1.2 Problem Domain 3

1.3 Aim 3

1.4 Objectives 3

2. Literature Review 4

3. Methodology 4

3.1 Hypothesis 4

3.2 Functional Requirements 5

3.3 Non-Functional Requirements 5

3.4 Technology 5

4. Conclusion 6

5. References 7

1. Introduction

1.1 Background and Motivation

As technology capabilities grow many applications that use both audio and video in both domestic and military aspects have been implemented with high efficiency and low cost, tracking of objects via audio and video signals are becoming more intelligent and essential. Nowadays modern video conferencing systems use different technologies to track a specific location of a speaker's voice (or both audio and video recording). These types of systems play a special role in conference meetings. Conventional systems bring quite a lot of disadvantage, one of which can force people to focus in the direction of recording so in order to overcome different issues we design to implement a system that tracks the speakers voice, by interfacing the camera with a motor. The interfaced camera and the motor is located at the center of the table and microphones placed at each end of the speaker so that it can turn to the present speaker. Therefore, it allows a group to carry on the normal meeting. So, the members of a meeting should be focused on who is speaking, our design system would be useful to companies that frequently use web conferencing.

1.2 Problem Domain

As we proceeded with the research of our project, we encountered some difficulties that may or cannot be resolved by a few changes. The design is implemented based on the emphasis of a speaker and to the loudest. There can be a problem when two individuals started speaking concurrently, hence it can move back and forth and distract the people in the room. There has to be a solution to shield the motor noise so that the microphones won' pick up the noise. We may have to establish security procedures when streaming the video live, to prevent against unwanted unauthorized access. We have to ensure the integrity of the system and it has to be precise. Also, there can be slight time gap between the transmission and displaying the video.

1.3 Aim

To be able to provide an efficient and enhanced video conferencing system, by automating the existing system eliminating human effort, cost, and security issues, using microphones to detect the voice and to track down the speaker's voice by using a microphone and then move the camera towards the speaker.

1.4 Objectives

Making a simple electronic device that is capable of recording a video session of a meeting or a conference. It eliminates the human effort, cost, and security issues (arising out of a conventional method of human intervention) and can be used in point to point conferencing system. In accordance with the present invention, an automatic voice tracking camera system and method of operations are provided that substantially eliminate or reduce disadvantages and problems associated with previously developed video conferencing systems. Our main objective is to create Smart Room Technologies that eliminates human effort, many workplaces involve in hiring a person to record a conference meeting, where it can lead to cost and security issues. A microphone array includes a plurality of microphones. The microphone array receives the voice of a speaker and to provide an audio signal to generate from the audio signal speaker position information. Also, necessary control actions are needed so that it automatically tracks the position of the speaker.

2. Literature Review

This section contains a literature research conducted.

[1] In the year 1992, An algorithm was developed to determine the talker location from the linear microphone data array. The quality of sound pickup in large rooms such as conference rooms, auditoriums, or classrooms is impaired by reverberation and interfering noise sources [1]. [2] These degradations can be minimized by a transducer system. The signal-seeking electrical device system is enforced as a dual-beam “track whereas scanning” array. It utilizes signal properties to differentiate between desired speech sources and intrusive noise. Using 4 microphones to locate a sound source this research and derived equations were focused on. [2] [3] A method of automatically controlling a camera to track a position of a speaker using a voice of the speaker, comprising: receiving a voice of the speaker and providing an audio signal representing the voice; processing the audio signal to generate speaker position data representing the position of the speaker as coordinates for a point in space; determining a speaker position change if changes in the coordinates exceed a predetermined amount; determining an appropriate responsive camera movement from the speaker position data based upon the speaker position change determination; generating camera control signals based upon the appropriate responsive camera movement, and providing the camera control signals to a first camera such that a view of the first camera automatically tracks the position of the speaker [3]. [4] It is recorded that by Paul C. Meuse, Harvey F.Silverman has developed Microphone array technology to introduce the concept of using wireless array of microphones in order to improve the reception of a sound and to locate the position of the speaker [4]. Since there were no wireless array of microphones available within, we decided to connect using wired media.

3. Methodology

3.1 Hypothesis

Hypothesis 1: There is a small difference between the time of capturing the video and at the receiving end, this is known as Latency.

This cannot be eliminated but can be reduced. By adjusting settings in VLC media player.

Hypothesis 2: The sensors are able to pick up the motor noise and may lead to incorrect tracking of the speaker.

The motor and sensors should be placed apart a reliable distance and should be well shielded.

Hypothesis 3: Speakers may speaker in various postures and this may can cause the sensors not to pickup the incoming signals

Therefore, error calculation should be calculated and the sensitivity should be adjusted accordingly.

Hypothesis 4: Two speakers may speak at the same time.

The system should be configured to locate the loudest incoming signal to each sensor or with further investigation necessary precautions needed to be taken.

3.2 Functional Requirements

- Should be able to track the voice of the speaker and direct the camera to that location.
- Transmit the recording video live across the network.
- Able to record at all resolutions.

3.3 Non-Functional Requirements

- Security - Protection of system is viable, unauthorized access to the network should be prevented.
- Recoverability - Data should be restorable from a backup if data has been damaged, lost or deleted.
- Reusability - The system should be portable. So, moving from one OS to other OS does not create any problem.

3.4 Technology

Software Requirements

- Raspbian Operating System.
  A Debian-based OS for Raspberry Pi.
  We will be using the Raspberry Pi 3 Model B microcontroller. Therefore, in order to operate will it there should be an OS installed.
- Thonny Python Editor
  This application software in pre-installed with the Raspbian OS therefore we decide to write our code in python language.
- VLC Media Player
  It is a free and open-source cross platform media player software; we could use this to stream the conference video live across the network.

Hardware Requirements

- Raspberry Pi 3 Model B
  A Microcontroller that will be out central control unit of the system. It collects data from the microphone sensor array, decides in which direction the speaker is located, and controls the motor control unit. Raspbian Operating System should be installed in order to work with the board.
- LC-227 Microphone Sensor
  Dynamic microphones are in wide use and their quality of reproduction is superb. They are utilized in the recording business for music and speech wherever sound reproduction is needed. In our case, we used an array of highly sensitive microphone sensors (LC-227).
- Servo Motor
  Allows rotation of 180 degrees. Servomotors, however, have inherent position feedback that permits such motors to correct any bloomer they'll create. The motor is required to turn a full 360 degrees and to complete a full 360-degree rotation in some seconds like 2 secs (velocity of 180 degrees per sec).
  e.g. SG-90, MG996R

- Raspberry Pi Camera Module v2
  We decided to use a camera module (Raspberry Pi Camera Module v2) directly compatible with Raspberry Pi 3. The weight of the camera lies well within the range that the motor can bear. The camera sits atop the motor's rotor and is connected to the raspberry pi board.
- 2A Phone charger
  2A Micro USB phone charging cable is needed to power up the raspberry pi board. Although minimum of 2A is needed it is recommended to use 3A.
- Jumper wires
  To connect the modules and servo motor to distribute power and carry signals coming from the microcontroller.
- Bread Board
  Needed to connect the sensors via jumper wires and also distribute power.
- Wireless Router
  A router with access to internet should be available to live feed the video.

4. Conclusion

With the correct adjustment, our product or we are going to be are valuable to corporations who often do internet conferencing. The largest drawback is that we have a tendency to presently use dear elements that hinder the vary of the system. Ideally, we'd like to check out the planning with microphones of excellent quality that those we've used, additionally we'd like to case our style in a very structure Also, we want a way of sleuthing once there's a concurrent speaker in order that the camera doesn't oscillate back and forth between them. Ideally, it'd simply opt for one in all the speakers to trace, the louder speaker, however, we've not puzzled out the simplest way to accomplish this. There are still detecting issues with our current style. Additionally, we have a tendency to still haven't styled an honest thanks to building our design visually appealing whereas shielding the motor noise, that is critical in order that the conference participants are that would insulate the motor noise and canopy up the electrical elements and wiring within. A quieter motor would additionally profit our detecting in order that the microphones won't be able to capture its noise. Since we are able to stream the captured video live but security is at risk so we need to look into that.

5. References

[1]. Shilpashree R N, Shruthi H O,Smitha S, Veenashree C N, Arpitha Shanrka S I, International Journal of Innovative Science, Engineering and Technology “Rotating camera based on speaker voice”

[2]. W. Durfee, A.Byres “L293 Motor Driver” Dec-2001 and Sept 2006.

[3]. “AUTOMATIC VOICE TRACKING CAMERA SYSTEM AND METHOD OF OPERATIONS” by joonyoul Maeng, Errol R.williams (forgent networksInc.) Appl No.08/509,228 july 1995.

[4]. Jiehan Yao, Yuxiao Lu,“Voice tracing video camera designed for meeting recording”Spring 2013.