

Project Abstract

Title: Audio – Visual Matching System

Team Members: Faysal Shair, Ertem Nusret Tas

Project Description: This project aims to implement a detection system that matches the voice of a source to its image reflected on a camera. The project consists of two parts. First, it derives the location of a sound source from the time delays a distinct voice wave induces while hitting microphones facing different directions. Second, it identifies a region of possible locations for the sound source on the image frame of the camera. Finally, if these two tasks can be successfully achieved, the system will be enhanced to calculate the speed of the sound source from the frequencies by using the Doppler Effect. By comparing the speed calculated from the sound waves with the changes on the camera image, the system would then be able to infer the boundaries of the sound source inside the previously identified set of possible locations. Another enhancement might be identifying a certain class of objects (i.e. faces) in those regions but due to its complexity, this task will be assumed only after the previous functionalities are implemented.