International Journal of Engineering Research and General Science Volume 4, Issue 3, May-June, 2016
ISSN 2091-2730

Real Time Gesture to Audio Conversion using FPGA Based System

Prof.A.S.Pawar[1], Shantanu Kulkarni[2], Griraj Jangilwad[3],Vaibhav Fundkar[4]

1,2,3,4-Department of Electronics and Telecommunication Engineering , Pimpri Chinchwad College of Engineering,Nigdi Pune.

Email - shantanu281994@gmail.com , Contact no.-8655805865

Abstract— Audio generation based on the real time video analysis is always been challenging and increasingly important task for embedded system design domain. Though it is the most effective way of communication, has still leftover many challenges. Dumb person make use of this system to communicate and express their opinions to the society. The intended system generates audio clip automatically with the help gestures captured in the image form from the CMOS camera module. The audio generation is carried out via three staged process: Image capture, Image comparison and audio generation. The realm of this paper is to portray the stage wise mechanism of audio generation and describing of various other features provided by this system.

Keywords— Audio Generation, Image Capture, Image Comparison, CMOS camera module, FPGA.

I.INTRODUCTION

Audio generation based on the real time video analysis is always been challenging and increasingly important task for embedded system design domain. Recently there has been a great deal of interest in the field of FPGA. Advances in field programmable gate array (FPGA) have resulted in the functionality of FPGA.For the video processing, various different technologies have been used. The software implementation of most of these technologies could not achieve the frame rate required for real time processing of video data. But FPGA provides the sufficient frame rate for the real time video processing. Hence this project uses the FPGA as a platform for the processing of real time gestures.

Main aim of this project is to design & construct an interface between the CMOS camera module i.e. OV7670 and speakers by using FPGA. In this project, we will convert the real time gesture into the audio clip by using CMOS camera module i.e. OV7670 and FPGA kit.

The design of the real time gesture to audio conversion system can be divided into 4 modules namely camera, image processing, camera control, audio generation through speaker. The hardware block diagram is shown in the Fig 2. In the hardware implementation CMOS camera module i.e. OV7670 is interfaced with the Zync Evaluation and Development Board. i.e. Zed Board using various protocols such as SCCB protocol, I^2C protocol to provide the fast image acquisition from the real time video. The camera is switched to the camera capture mode in order to capture images continuously. These picture images are stored into the inbuilt memory of the FPGA board. At the same time the captured images are subjected to analog as well as digital processing so as to obtain better resolution.A database is created as per the required application and is stored into inbuilt FPGA memory. Large no. of reference images are stored in the database. An audio clip is assigned to every image in the database.

The captured image from the camera is also stored into inbuilt memory of FPGA. This captured image is then compared with the reference image in the database. If both the images are identical then the audio assigned to the reference image is fed to speakers and this audio clip is generated by speakers.

II. IMPLEMENTATION OF SYSTEM

The diagram shows the design and construction of an interface between the CMOS camera module, FPGA, speakers. The input image is stored into the on board memory of FPGA kit.

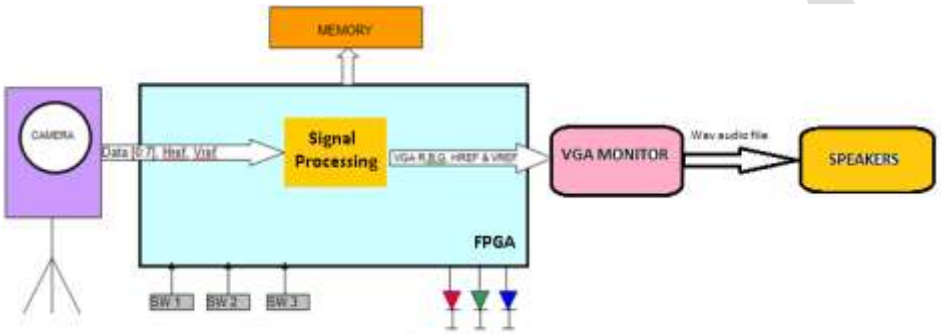


Fig. 2 System Block Diagram

The information flows in two ways: on one hand there are commands from the FPGA to the camera to change different characteristics of it, on the other hands images from the camera is processed and sent to the speakers.

A.CAMERA INTERFACING

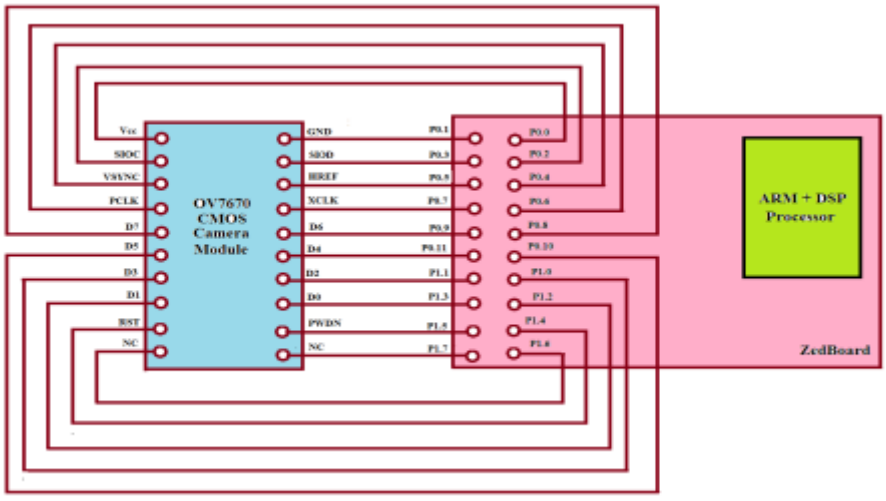


Fig. 4 Camera Interfacing

In this section we describe about the camera interfacing with the Zed Board which consisted the camera capture mode which helps to capture a gesture from real time video. We have used 2 ports of the Zed Board for camera interfacing. The interfacing diagram is shown in Fig. 4. We have used the port J_A and port J_B for this purpose. OV7670 has 20 pins, out of which first 12 pins are J_A and remaining 8 pins are assigned to J_B port. Fig. 5 shows how images are transmitted from CMOS camera module to the FPGA.

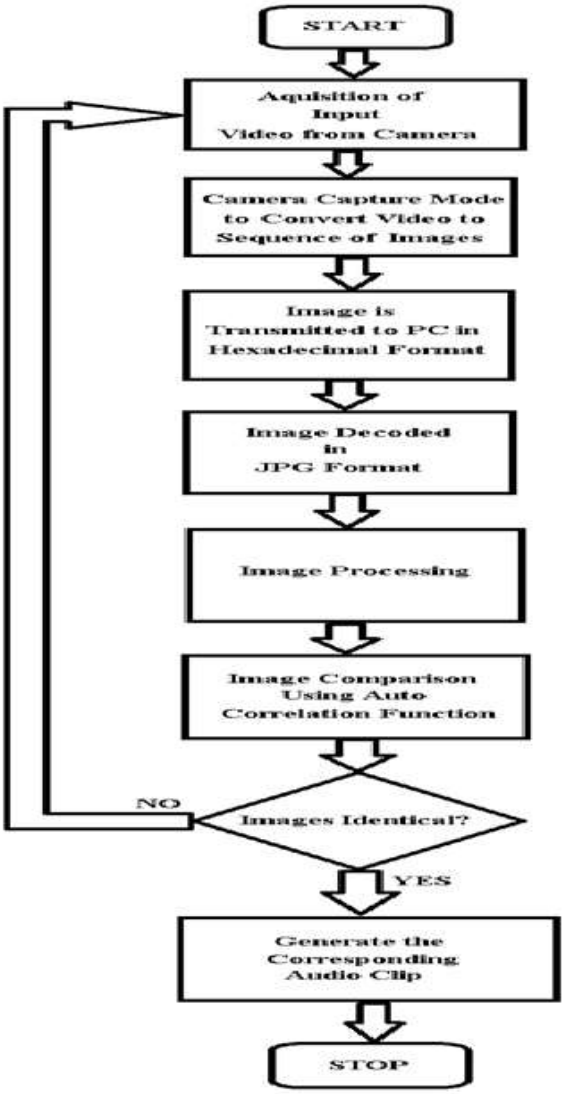


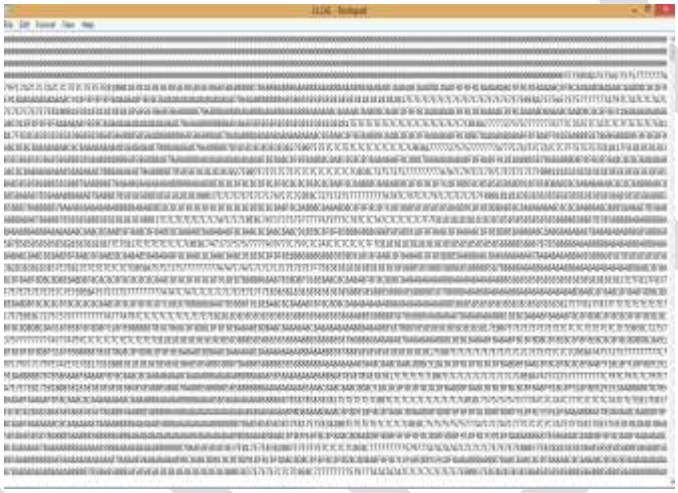
Fig. 3 System Flowchart

III. IMAGE PROCESSING

The image is received from camera on the Hyperterminal through serial communication using universal asynchronus reciever transmiter.It is in the hexadecimal format.The orignal image is reconstructed in MATLAB by decoding the log file.The decoded image is mirror image and is roatated by 270 degrees with respect to orignal one.So after decoding the mirroring and rotation of image is also carried out in MATLAB and the orignal image is extracted in MATLAB. The comparision of the extracted image and the stored image is carried out .If both the images are identical then the audio file corresponding to that gesture is generated through speakers.

IV. RESULTS:

1. **Received log file:** The image captured by camera is obtained on serial terminal of PC. The received log file is in the hexadecimal format.Each of the two nibbles corresponds to one byte of information. “00” represents the pure black level while “FF” represents the pure white level.



2. **Reconstruct image by decoding log file:** The decoding of the log file is done in MATLAB to reconstruct the orignal image shown. But this reconstructed image is the mirror image of the orignal one with rotation of -270 degrees

.

3. **270 degree rotation of decoded image:**

4. **Output of Mirroring using Matlab:** After rotating the image by 270 degrees , the mirror image of the image shown in fig.6 is taken .The algorithm for the same is built in MATLAB.

5. **Stored image:** The stored image is located in the memory of PC. It is read using MATLAB functions so as to compared with the capture image. When both the images are identical then the audio file with '.wav' extension is played through speakers.The MATLAB doer not support the .mp3 file format, hence .wav file format is used.

CONCLUSION

The end user is now moving towards a whole new path of human machine interaction. This is creating a demand for enabling gesture recognition in every facet of market. The proposed technique is simple, efficient and faster . This system has achieved its aim of fullfilling societal need by providing the assistance to the physically impaired people to express their opinion in front of the world.

REFERENCES:

[1] Rushikesh T. Bankar1 and Suresh S. Salankar "Head Gesture Recognition System Using Gesture Cam" International Journal of u- and e- Service, Science and Technology Vol.8, No. 6 (2015), pp.341-346

[2] Yean Choon Ham and Yu Shi "Developing a Smart Camera for Gesture Recognition in HCI Applications" The 13th IEEE International Symposium on Consumer Electronics (ISCE2009).

[3] S.M.Alex Raj Sreelatha G. and Supriya M.H.2 "Gesture Recognition Using Field Programmable GateArrays".

[4] Sue Han Lee1, Soon Nyean Cheong, Chee Pun Ooi, Wei Heng Siew "Real Time FPGA Implementation of Hand Gesture Recognizer System".

[5] Ignacio Bravo *, Javier Baliñas, Alfredo Gardel, José L. Lázaro, Felipe Espinosa and Jorge Garcí " Efficient Smart CMOS Camera Based on FPGAs Oriented to Embedded Image Processing" ISSN 1424-8220 Sensors 2011, 11, 2282-2303; doi:10.3390/s110302282.

[6] Mandeep Kaur Ahuja and Dr.Amardeep Singh "A Survey of Hand Gesture Recognition" Volume 3, Issue 5, May 2015 International Journal of Advance Research in Computer Science and Management Studies.

[7] Spartan-3 FPGA Family Data Sheet.

[8] CMOS OV7660 Camera Module 1/5-Inch 0.3-Megapixel Module Datasheet.

[9] www.XILINX.com

[10] www.datasheetcatalog.com

[11] http://dx.doi.org/10.14257/ijunesst.2015.8.6.34

[12] www.mathworks.com

[13] www.avnet.com.