2015

Head-Tracking Wireless Streaming Device

Christopher Foster
Virginia Commonwealth University

Kyle Arpe
Virginia Commonwealth University

Zachary Somers
Virginia Commonwealth University

ECE-10
Team Members: Christopher Foster, Kyle Arpe, Zachary Somers

Faculty Advisor: Dr. Carl Elks

ELECTRICAL AND COMPUTER

Head-Tracking Wireless Streaming Device

CAPSTONE DESIGN EXPO 2015

Introduction

- In various businesses and services, there is a need for tight integration between visual media, human response to that media, and coordination of that response. For example, emergency responders may need information from a separate perspective using robotically controlled cameras in order to improve coordination efforts. The aim of this project is to design a low cost, high performance video streaming device. The essential feature of our design is to wirelessly send a video stream from a webcam to a micro-display and remotely control the orientation of the webcam using head movements.

- The above image shows the basic components of the system which include the Zynq-7000, Raspberry Pi, Logitech webcam, and NTSC video glasses.

Image Capture and Compression

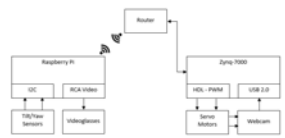
- In order to wirelessly send a frame from the camera, the frame must first be compressed. In this case, we're using JPEG compression which takes advantage of the fact that a discrete time signal, such as an image, can be stored in the frequency domain using significantly fewer amounts of data. This helps reduce the bandwidth of the video stream on the network and helps improve latency. The following transform, the 2-D Discrete Cosine Transform, is typically used for JPEG compression algorithms.

$$X_{k_1,k_2} = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x_{n_1,n_2} \cos\left[\frac{\pi}{N_1}\left(n_1+\frac{1}{2}\right)k_1\right] \cos\left[\frac{\pi}{N_2}\left(n_2+\frac{1}{2}\right)k_2\right]$$

2-D Discrete Cosine Transform

- In order to accelerate image capture and manipulation, the OpenCV (Open Source Computer Vision) libraries are being utilized.
- In order to track head movement, an accelerometer is being used to accurately determine the tilt angle. For yaw, magnetometers are being used to measure the angle relative to the Earth's poles.

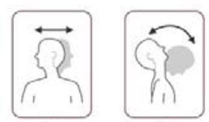
System Design



System Overview

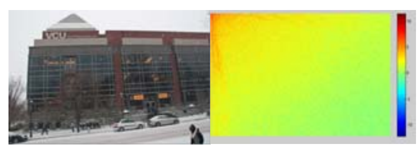
- A basic system overview is shown above. The Zynq-7000 SoC is used to acquire frames from a webcam using OpenCV. Additionally, the Zynq's FPGA was utilized to control two servo motors that change the webcam's orientation. The encoded frame captured from the webcam is sent to the Raspberry Pi, which wirelessly receives the frames, decodes the frames, and outputs the image to the micro-display. The Pi also sends tilt/yaw information back to the Zynq to control the servo motors.

$$\psi_M = \arctan\left(\frac{H_y}{H_x}\right)$$



$$\rho = \arctan\left(\frac{A_X}{\sqrt{A_Y^2 + A_Z^2}}\right)$$

- **The above picture describes the relationship between the user's heading and the accelerometer and magnetometer data. Where Ψ_M is the yaw based on magnetometer data and ρ is the pitch based on accelerometer data.**



- Shown above are images of the DCT (Discrete Cosine Transform) of a picture of the West Hall Engineering Building. The left picture is the school. On the right, is a picture of the magnitude of the spectrum. Most of the energy in the spectrum is concentrated toward lower frequencies in the upper left hand corner of the spectrum. Most of the higher frequency components have been quantized to zero.

Results

- The system captures images and transfers images at varying speeds (fps) based on the quality factor used to encode the image and the amount of data in the captured frame. An image encoded with a lower quality factor greatly reduces the amount of data that needs to be sent through the network, however the compression introduces data loss which affects the quality of the image upon reconstruction.

- The above two images show sample frames that have been reconstructed after compression. The picture on the left was encoded with a quality factor of 10. The image on the right was encoded with a quality factor of 85. As shown, the image on the left introduces noticeable compression artifacts (fuzziness).

Future

- There are many future applications for this project including target recognition, blind spot detection, robotics, human studies, and security.
- Future improvements include the utilization of a transparent screen using OLED technology, power aware computing, data overlay onto the image displayed to the user, and a more ergonomic electronic solution.

Acknowledgements

- We would like to thank Carl Elks, Ph.D., Matthew Leccadito, M.S., and Ryan Littleton for their support and knowledge throughout this design process.

Slide 1

CRE1 I would start more broadly for the introduction... Like, " In various businesses and services there is a need for tight integration between visual media, human reposnse to that media and coordination of that response. As example, emergency repsonders may need to know what a robotic team member is seeing when they are seperated in order to coodinate better.

Carl R Elks, 3/30/2015