Human Activity Recognition Using Kinect Motion Sensor

A Proposal Prepared for the Final Project of the Course EE482/582: Computer Vision & Digital Image Processing

Oct 22, 2013

Prepared by:
Rui Ma*, Andrew Swindle and Alicia Simon
rma3@crimson.ua.edu, awswindle@crimson.ua.edu, ajsimon@crimson.ua.edu
Department of Electrical and Computer Engineering
University of Alabama
Tuscaloosa, AL 35487-0286 USA

* Team Leader
Abstract

This project uses the Microsoft Kinect motion sensor to capture the depth image of the human target and to recognize the human activity with the depth information of the target based on the knowledge of computer vision and image processing.

2.1 Research Question or Problem

There are existing softwares that can read in Kinects’ video signal and extract the human model like the Xbox itself and the third party software iPi Motion Capture. However we are not able to access the codes of these softwares, and what’s more important, it is not necessary to use such an accurate model for some activity recognition. For example, it takes 25 minutes for the iPi Motion Capture to extract the human model from a 5 seconds video. So we hope to develop some simpler algorithms that can make a better compromise between the processing speed and the recognition performance.

2.2 Research Goals and Objectives

This project is going to recognize the standing, walking, sitting, squatting and falling of the single person. These activities can be described by 10 frames of video, so the recognition needs to process about 3 M pixels each time assuming that every frame is 640*480=0.3 M. As we know, it takes about 0.02 seconds for a 1G CPU to implement a fast Fourier transform upon 1 M data, so it may take 3 seconds for the recognition if the computation amount of the recognition is equal to 50 FFT upon 3 M pixels. On such premise, our goal of the project is set to implement the activity recognition within 3 seconds.
2.3 Research Design and Methods

(1) Read in the depth map images from the Kinect to MATLAB. Directly read in the depth information using the plugins or obtain the depth information from the color gradients of the image which denotes the depth of the targets.

(2) Extract human target by slicing the depth image along the depth. Calculate the correlation between these slices and the preset frames to determine the depth of the human target.

(3) Get the center of the human target and the distances between the center and the boundary of body to obtain the shape of the human target. Calculate the correlation between the tested shape and the shape used as model to determine what activity this shape belongs to.

2.4 Staffing Plan

In this section, the background and expertise of each team member should be introduced. Each member’s contribution should also be stated specifically.


Andrew Swindle: undergraduate student. Contribution: human model extraction.


2.5 Timeline

(1) Nov. 1~ Nov. 8: Read in depth image from Kinect to MATLAB

(2) Nov. 9~ Nov. 15: Human model extraction

(3) Nov. 16~ Nov. 22: Human activity recognition

(4) Nov. 23~ Nov. 29: Human activity recognition

(5) Nov. 30~ Dec. 12: Model improvements and project demo and report.
Reference


[2] Introduction to using iPi Motion Capture for Kinect video recording:
http://wiki.ipisoft.com/Main_Page

[3] MATLAB Simulink support for Kinect:
http://www.mathworks.com/matlabcentral/fileexchange/32318-simulink-support-for-kinect