Human Body Component Tracking and Object Detection Using Monocular Video Sequence

Ariyasinghe, G.N.P., Perera, D.H.L. and Wijayarathna, P.G.

Email: nishadiari@gmail.com, leshan.hira@gmail.com, gamini@kln.ac.lk

Abstract

Medical education plays a vital role in a country's education system. It is essential that a medical student should be provided with a realistic environment in order to effectively learn and practice disease diagnostics. According to medical education, initially, diseases are determined by diagnosing abnormal heart and lung sounds. Practicing such diagnostics requires a large pool of patients representing each disease which needs to be learnt. However, providing such a large number of patients for an examination session is impractical. Finding patients representing each disease to be learnt is another challenge. The current method used to practice diagnostics via heart and lung sounds is either by using a dummy or a healthy human and identifying disease according to symptoms described by the performer or the doctor/lecturer. This leads to an unrealistic examination environment for the medical student, thereby decreasing the productivity of the medical education system. Meanwhile, object detection in human body pose and component tracking from video inputs has been an active research field motivated by various applications including human computer interaction, motion capture systems and gesture recognition. One of the most important biomedical applications focuses on building simulators to carry on activities in the medical field. Most current tracking methods include multiple cameras and many markers placed on key body points. This makes the examination environment become less realistic and the methods are proven to be slow and unreliable. Furthermore, many tracking systems must be initialized by a human operator before they can track a sequence. Pose tracking using 3D Time of Flight (TOF) cameras exists. However, purchasing TOF cameras are expensive and since they only detect infrared emitting surfaces, they are difficult to be used for many applications. Several learning-based techniques have been proposed for monocular sequence view, but these rely on accurate body silhouette extraction and require relatively large number of training images. SimHaL (Hybrid Computer-based Simulator for Heart and Lung disease diagnosis to enhance medical Education) is an ongoing project which intends to build a hybrid computer-based simulator with an integrated human and computer components. Its aim is to enhance the productivity of medical education by simulating patient examination in a more realistic environment. Therefore it acts as a simulator for disease diagnosing by identifying relevant heart and lung sounds by medical student. The current state of SimHaL focuses on detecting the location where the Chest piece of a stethoscope is placed on a patient's torso. Since the major target is to build an optimal realistic examination environment for the medical student, a single camera is used to monitor the activity. The output is a monocular

video sequence which is the only source available for identifying the torso and the Chest piece as objects. The methodology focuses on object detection categorized into two approaches:

1. Detecting the chest piece of the stethoscope

2. Detecting the patient's torso

In order to identify the chest piece, a circle detection program is implemented using OpenCV. Here the monocular video sequence is divided into frames and circle is detected based on a provided range of radius value. The provided radius value range approximates the radius of the chest piece. Other circles detected in the background will be discarded if their radius value is not in the provided radius range. Next, the motion detection of the identified chest piece is obtained by computing the difference of Cartesian coordinates of circles detected in adjacent frames. Circles with differences which exceed a certain threshold value, are discarded. Currently this threshold value is set as a fixed value assumed to be the width of the patient's torso. This avoids unusual movements of any detected circle and makes sure the circle detected in the current video frame is the same circle which was detected in the previous one, but has now moved to a new location. Results of this approach consists a concurrent output of x and y Cartesian values relative to the video frame along with a video sequence with a circle drawn in each frame. Radius of the circle is the radius of the chest piece detected at the beginning and the x and y values indicates the circle's center. The current status of the research concludes identification of the chest piece. Detecting the human torso and thereby determining the location where the chest piece is placed is yet to be implemented.

Keywords: *object detection, component tracking, virtual reality*