

# Use of a Monocular Video Sequence for Human Body Component Tracking and Object Detection

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

**Abstract**— Medical education plays a vital role in a country's education system. Therefore, it is essential that a medical student be provided with a state-of-art education. Engaging with a realistic learning environment ensures an effective study and practice of disease diagnostics for a medical student, thereby enhancing the medical education system. As the practice followed in the medical field, diseases are determined initially by diagnosing abnormal heart and lung sounds. Thus, 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 a classroom learning session is impractical. Finding patients with each disease which needs to be learnt is another challenge. Simulation based medical education has currently emerged to practice diagnostics via heart and lung sounds. Either by using a dummy or a simulator patient, diseases are identified according to symptoms described by the performer or the doctor/lecturer. This leads to an unrealistic examination environment for the medical student. SimHaL (Hybrid Computer-based Simulator for Heart and Lung disease diagnosis to enhance medical Education) is a hybrid computer-based simulator with integrated human and computer components which simulates patient examination in a more realistic environment.

**Keywords**—object detection, virtual reality, simulator patient, Kinect V2 Sensor, OpenCv, SimHaL

## I. Introduction

During the past three or four decades, the practice of medicine has undergone through drastic changes. Several new inventions and ideas have resulted in dynamic and interactive disciplines. According to [11], although the modern healthcare environment has continually been changed, due to the little change of medical education since the early stages, there is a concern where medical students are not being adequately prepared to provide optimal health care in the industry. This situation has mostly been occurred within the developing countries. [11] Also states that a teacher depended learning process, uneven access to medical education and insufficiently planned and supervised clinical teaching of medical students where more emphasis is given to theoretical sessions is some of the shortfalls and reasons for an output of such students. [2] Also states that although students complete their educational programs armed with theoretical knowledge, they have proved to lack many clinical skills vital for their work. Thus, it is essential for a medical student to be provided with a realistic environment which simulates the real-world clinic where he/she can practice medical treatment along with their theoretical education.

## A. *Simulation based medical education*

In order to enhance professional practice and maintain patient safety, simulation based medical education has been identified as a valuable aide for training medical graduates. Therefore, it has become an essential part of today's medical training [8]. The term simulation has been defined variably in relation to different disciplines in the literature. However, one of the most suitable definitions regarding to medical education is stated as “a generic term that refers to an artificial representation of a real-world process to achieve educational goals through experiential learning. Simulation based medical education is defined as any educational activity that utilizes simulation aides to replicate clinical scenarios” [2]. Literature suggests that roots of simulation based medical education at least stretch to the 17th Century, where birth manikins were used in France. Provision of a risk-free learning environment where students can be allowed to practice without harming patients, learning from errors and repeating practice until mastery of skills are some of the major benefits of using a simulator for medical education.

The advances of medical simulators have been developed from plastic mannequins such as the TEE Simulator [3], [9], to trained actors and to a more advanced virtual, but realistic environment. It has been proved that the more realistic the educational environment for the medical student be, the more productive medical education he/she will be able to obtain.

According to [12], in the medical field, initially, diseases are determined by diagnosing abnormal heart and lung sounds of the patient. Medical situations such as pneumonia, heart failure, pleural effusion and many other diseases can be identified or predicted by listening to abnormal heart sounds or murmur. Practicing such diagnostics requires a large pool of patients representing each disease which needs to be learnt by the medical student. 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. In order to overcome this issue, many technologically advanced methods including the collaboration between ever advancing computer sciences and related technologies and medical education have been introduced and already in practice. However, these advanced technologies such as medical simulators have the requirement of the use of high cost computer equipment and their maintenance. For example the innovative simulation-based training program for cardio vascular specialties built on the concept of Distributed Simulation [5] requires a high cost simulation environment. Another drawback of such simulators is its functional fidelity not accurately representing the reality due to hardware and performance issues. As a developing nation, one of the most significant roadblocks we face is to secure funds and other

resources to bring-in technological advancements. The gap of functional fidelity could be rectified if a human component is incorporated into this technology. Some of the current low-cost methods used to practice diagnostics via heart and lung sounds is either by using a dummy or simulated patient and identifying disease according to symptoms described by the performer or the doctor/lecturer. Simulated patients are specially trained normal people, guided and oriented to act as patients in training medical graduates. This situation however leads to an unrealistic examination environment for the medical student, thereby decreasing the productivity of the medical education system.

## B. Object detection from video inputs

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 [15]. One of the most important biomedical applications focuses on building simulators to carry out activities in the medical field. Most current tracking methods include multiple cameras and many markers placed on key body points of the patient to be detected [15]. 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 [15], [14], [13] 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 [1], but these rely on accurate body silhouette extraction and require relatively large number of training images.

## II. SimHaL

SimHaL (Hybrid Computer-based Simulator for Heart and Lung disease diagnosis to enhance medical Education) is an ongoing collaborative research project carried out in the University of Kelaniya, Sri Lanka. This project intends to build a high-fidelity hybrid computer-based simulator with 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 cardio-respiratory physical signs (i.e. heart and lung sounds) by the medical student. It also helps to facilitate in bridging the gap of “human touch” in the existing high-fidelity simulator models.

In SimHaL, a simulated patient is used to practice heart and lung sound diagnosing. The stethoscope is improvised as a haptic device to integrate auscultative signs between the medical student and the simulator, thereby generating a realistic clinical environment where a doctor examines a patient. The current state of SimHaL focuses on detecting the location where the Chest piece of a stethoscope is placed on a patient’s torso (front or back). Since the major target is to build an optimal realistic examination environment for the

medical student, a single camera is used to monitor the whole activity. The Kinect V2 Sensor for Xbox One [7] was used for this purpose. The output gives a monocular video sequence which is the only source available for identifying the torso and the Chest piece as objects. The methodology of the research focuses on object detection categorized into the following approaches:

1. Identification of simulated patient and his/her torso.
2. Detect Doctor/Medical Student.
3. Identification of Chest Piece.
4. Determining Chest Piece location on patient’s torso.

The above objects are detected by comparing adjacent frames extracted from the video sequence output by the Kinect. Object identification was implemented using OpenCV [10] which is a library consisting of programming functions mainly aimed at real-time computer vision. This library was used along with the Java platform. Identification of the simulated patient and the medical student is obtained by the skeleton tracking feature available in the OpenCV library. Afterwards, the chest of the simulated patient is detected by building up a square shaped boundary over his/her torso. Similarly, the medical student is also detected after he/she has also entered the scene.

The Chest Piece is detected using one of the two implemented strategies which are a “circle detection program” and a “color detection program”. The reason for using the mentioned two strategies was to avoid any misdetection of the Chest Piece due to issues occurred in the simulation setup such as the background or foreground light intensity imbalances. The first approach was to detect the Chest Piece as a circle using the circle detection feature in OpenCv. However, if any of the mentioned issues occur, in order to maintain the constant flow of the process, the simulator switches to a color detection strategy. For this purpose, the camera visible side of Chest Piece had to be recolored. Results of several test cases proved that the color “Green” should be used for this purpose. The reasons to use green color for the Chest Piece was that it was easily detected by the Kinect and the color was unchangeable with regard to alternations of light intensities.

Circles detected beyond the boundary of the torso are discarded. 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. x and y coordinates determining the Chest Piece location within the torso could be obtained in each frame of the video sequence.

## III. Results and Discussion

The implemented user interaction of SimHaL is shown in Fig. 1, Fig. 2 and Fig. 3. In order to create the ideal setup, several test cases were conducted where the best parameter values were determined for constructing the simulator unit (See Appendix). Initially the application detects the simulator patient as a skeleton.

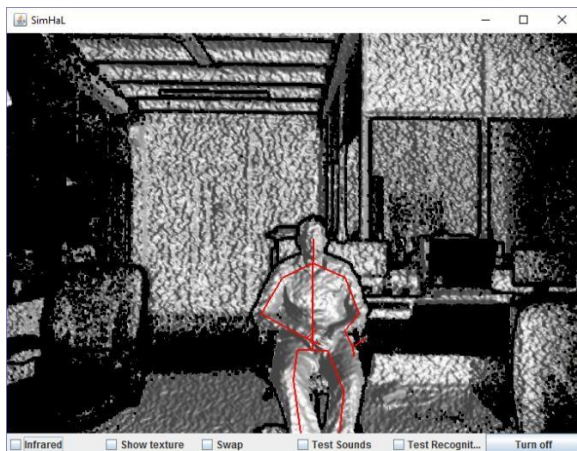


Figure 1. Skeleton Tracking

Then the patient is identified and the boundary of his torso is detected.



Figure 2. Patient Detection

The Chest Piece is detected either by using circle detection or color detection approach.



Figure 3. Chest Piece Detection

Results of this approach consists of a concurrent output of x and y Cartesian values relative to the video frame along with a video sequence, with a drawn circle depicting the Chest Piece on each frame.

There are however several limitations at the current situation while working with the simulator, as described below:

1. Simulator Patient, chair of the patient (currently disease diagnosing is done when the patient is in the seated position only) and the Kinect camera has to be still throughout the process.
2. Patient should not wear costumes in green color (color of the Chest Piece) or with patterns consisting of circles.

3. Doctor/medical student should enter the scene after patient is detected by the application.
4. Doctor/medical student must not wear a circle like accessories (Eg: wrist watch).
5. Application effectively works in normal light conditions. Should avoid direct or bright lighting towards camera and patient.
6. Doctor should hold the chest piece in the standard way, where chest piece is visible to the camera.

Software and Hardware used to implement the simulator are as follows:

TABLE I. SOFTWARE AND HARDWARE DETAILS

<b>Hardware Devices</b>	<ul style="list-style-type: none"> <li>• Kinect V2 Sensor for Xbox One [7]</li> <li>• Kinect Adapter for Windows [6]</li> </ul>
<b>Programming Language</b>	<ul style="list-style-type: none"> <li>• Java</li> </ul>
<b>IDE</b>	<ul style="list-style-type: none"> <li>• Netbeans 8.0.2 (JDK 1.8)</li> </ul>
<b>Software Libraries</b>	<ul style="list-style-type: none"> <li>• J4K Java Library [4]</li> <li>• OpenCv 2.4.13</li> </ul>

## IV. Conclusion

This paper proposes a Hybrid Computer-based Simulator for identifying cardio-respiratory physical signs by the medical student in the process of practicing disease diagnosing. In order to maintain a realistic study environment, rather than using a mannequin, a simulator patient without using any body markers is used in the application. The use of only one camera throughout the simulation process further enhances the realistic study environment. It also avoids the necessity for the use of high cost technological equipment which happens to be one of the most significant roadblock in developing countries such as Sri Lanka. The current status of SimHaL concludes identification of the chest piece and the simulated patient's torso. It also identifies the doctor or medical student if necessary. Determining the location where the chest piece is placed on the human torso and thereby building the interaction between the simulator and the practitioner through the haptic device is yet to be implemented. The simulator is currently being implemented at the Medical Faculty at University of Kelaniya, Sri Lanka. After completing the installation process in the faculty laboratories, further data analysis is to be carried out.

## Acknowledgment

The authors would like to acknowledge the Department of Medical Education at the Faculty of Medicine, University of Kelaniya, Sri Lanka (<http://medicine.kln.ac.lk/centres/mec>) who particularly supported this work at the implementation and installation process.

## References

- [1] Agarwal, Triggs B. (2004). 3D Human Pose from Silhouettes by Relevance Vector Regression. IEEE Conference on Computer Vision and Pattern Recognition.

- [2] Al-Elq A.H. (2010). Simulation-based medical teaching and learning. DOI: 10.4103/1319-1683.68787. Journal of Family and Community Medicine.
- [3] Heartworks User Manual (<http://www.inventivemedical.com/tee-simulator/>) [Accessed 24.09.2017]
- [4] Java Library for Kinect (<http://research.dwi.ufl.edu/ufdw/j4k/index.php>) [Accessed 24.09.2017]
- [5] Kelay T., Chan K.L., Ako E., Yasin M., Costopoulos C., Gold M., Kneebone R.K., Malik I.S., Fernando B. (2017). Distributed Simulation as a modeling tool for the development of a simulation-based training program for cardiovascular specialties. Advances in Simulation. DOI: 10.1186/s41077-017-0049-y.
- [6] Kinect Adaptor for Windows ([https://www.microsoftstore.com/store/msusa/en\\_US/pdp/Kinect-Adapter-for-Xbox-One-S-and-Windows-PC/productID.2233937600](https://www.microsoftstore.com/store/msusa/en_US/pdp/Kinect-Adapter-for-Xbox-One-S-and-Windows-PC/productID.2233937600)) [Accessed 24.09.2017]
- [7] Kinect V2 product description (<https://www.microsoft.com/en-us/store/d/Kinect-Sensor-for-Xbox-One/91HQ5578VKSC>) [Accessed 24.09.2017]
- [8] Lane, J., Slaviv, S., & Ziv, A. (2001). Simulation in medical education: A review. DOI: 10.1177/104687810103200302. Journal on Simulation and Gaming.
- [9] Matyal R., Mitchell J.D., Hess P.E., Chaudary B., Bose R., Jainandunsing J.S., Wong V., Mahmood F. (2014). Simulator-based trans esophageal echocardiographic training with motion analysis: a curriculum-based approach. US National Library of Medicine. DOI: 10.1097/ALN.0000000000000234.
- [10] OpenCv (<http://opencv.org/>) [Accessed 24.09.2017]
- [11] Rizk D.E.E. (2007). Medical education in developing countries: The way forward. Journal of Pakistan Medical Association.
- [12] Sarkar M., Madabhavi I., Niranjana N., Dogra M. (2015). Auscultation of the respiratory system. Annals of Thoracic Medicine. DOI: 10.4103/1817-1737.160831.
- [13] Song S., Xiao J. Sliding Shapes for 3D Object Detection in Depth Images. The 13th European Conference on Computer Vision (ECCV2014).
- [14] SwissRanger. Online Time-of-Flight Camera Information. (<http://hptg.com/product/#sensing>) [Accessed 14.08.2017]
- [15] Zhu Y., Fujimura K. (2010). A Bayesian Framework for Human Body Pose Tracking from Depth Image Sequences. Sensors - Open Access Journal. DOI:10.3390/s100505280.

About Author (s):



Ariyasinghe G.N.P.  
Faculty of Computing &  
Technology, University of Kelaniya  
Sri Lanka



Perera D.H.L.  
Faculty of Computing &  
Technology, University of Kelaniya  
Sri Lanka



Wijayarathna P.G.  
Faculty of Computing &  
Technology, University of Kelaniya  
Sri Lanka

## Appendix

Comments: circle detection is good and color detection  
 good torso detection is moderate

Several test cases were conducted in order to determine the best environmental setup for the simulator unit. Best parameter values for the camera position, the background of the unit, the simulator patient's and the medical student's physical appearance were obtained. The accuracy of the detection of the patient's torso and the accuracy of detection of the chest piece was measured. The test with best results by far is shown below:

<u>Camera</u>		
Characteristics	Parameters	Value
Camera Position	x,y,z	middle of the patient and doctor, eye line, 110cm far
Camera Angle	45, 90	

<u>Background</u>		
Characteristics	Parameters	Value
Background color	Color	brown
Light intensity	very low, low, moderate, high, very high	moderate
Direction of the light source related to patient	Front, left, right, above, behind	Front left side

<u>Patient</u>		
Characteristics	Parameters	Value
Patient's shirt color	Color	white
Shoulder size of patient	Large, small	small
Arm position of patient		Hands on legs

<u>Doctor</u>		
Characteristics	Parameters	Value
Doctor's skin tone	Dark, Fair	fair
Accessory type		No accessories
Doctor's position relative to patient	Front, Left, Right	right

<u>Results</u>	
Torso detection accuracy	95%
Chest piece detection accuracy	85%