

## CONCEPTUAL DESIGN AND IMPLEMENTATION FOR VISUAL TRACKING ANKLE REHABILITATION SYSTEM

Lim Chee Chin<sup>1</sup>, Shafriza Nisha Bin Basah<sup>1</sup>, Sazali bin Yaacob<sup>1</sup>, and Yeap Ewe Juan<sup>2</sup>

<sup>1</sup>School of Mechatronic Engineering, Universiti Malaysia Perlis,  
Pauh Putra Campus, 02600 Arau, Perlis, MALAYSIA.

Email: cheechin10@gmail.com / shafriza@unimap.edu.my

<sup>2</sup>Consultant Orthopaedic Surgeon, Sime Darby Medical Centre Park City Sdn Bhd,  
Desa Park City, 52200 Kuala Lumpur, MALAYSIA.

### ABSTRACT

The simple needs of an ankle rehabilitation system are valid for medical evaluation, user-friendly, and perform efficiently at low cost. However, most of the current ankle rehabilitation systems face a lot of problems, such as inconvenient face-to-face therapy, manual evaluation by the physiotherapist, the limited number of physiotherapists, and the high cost. Therefore, the key conceptual issues in designing and implementing an ankle rehabilitation system are identified and discussed in this article in order to overcome these problems. The aim of designing an ankle rehabilitation system is to furnish an alternative for ankle sprain patients so that they can efficiently perform rehabilitation exercises in their household surroundings. Additionally, the output data from the ankle rehabilitation system provides valuable patient information for further medical evaluation and monitoring. This article describes the conceptual design phase of an ankle rehabilitation system. It starts with a needs analysis and focuses on conceptual design. Six concept options are designed based on the needs identified. The selected concept is decided based on the system needs and characteristics of the conventional ankle rehabilitation method. Finally, the preliminary implementation result is included to demonstrate the feasibility of the selected concept for the ankle rehabilitation system.

**Keywords:** Visual tracking; ankle rehabilitation system (ARS); structured light camera (SLC); Red-Blue-Green-Depth (RBG-D) images.

### INTRODUCTION

Ankle sprain is one of the more common injuries in sport and the domestic environment, and accounts for between 15% and 20% of all sports injuries, according to [1]. Roughly 25,000 people experience this each day, according to a report from the American Academy of Orthopaedic Surgeons [2]. A sprained ankle can happen to athletes and non-athletes, children and adults. Ankle sprains happen when the foot twists, rolls or turns beyond its normal motion. It is important to have an ankle rehabilitation exercise system. If injuries do not heal properly after an ankle sprain, the ankle can become weak and unstable and may subsequently give out with only minor trauma, such as stepping off a kerb. Rehabilitation exercises can help repair and strengthen the injured ligaments. Thus, rehabilitation exercises are critical to ensure that the ankle heals completely and reinjury does not occur.

The ankle sprain recovery process needs a standard medical protocol [3] for ankle rehabilitation. The conventional rehabilitation process is done manually using repeated face-to-face therapy by physiotherapy and doctors. The conventional rehabilitation consists of four phases of rehabilitation exercises and it takes about 3 to 5 months to recover from the ankle sprain injuries. For each phase, the timing and type of rehab exercises are recommended according to the doctor's and therapist's preferences. The therapist needs to guide the patient to undergo daily rehab exercises. However, there is a limited number of therapists and doctors for the number of patients. Thus, a Visual Tracking Ankle Rehabilitation System is designed for the purpose of providing a reliable, high quality service with positive ankle recovery outcomes; overcoming the problem of a limited number of physiotherapists; and providing remote rehabilitation in the home environment rather than in a hospital due to geographical distances and the immobility of the patient. Visual tracking is ideal to create a process of locating ankle motion over time using a camera. Visual tracking is useful for identifying human ankle movement during rehabilitation exercises and for further analysis of the ankle motion. This system will be able to analyse the direction and angle of ankle motion.

The aim of this research work is to provide a systematic review of studies that investigate the effectiveness of the visual tracking therapy method and recovery of function from musculoskeletal or neurologic ankle injuries. The objective of this paper is to identify the best conceptual design for a visual tracking ARS, and to implement a suitable system set-up and camera configuration. The next section of this paper discusses the method to identify the best conceptual design. Then, an overview of implementation of the camera system set-up and configuration is presented. After that, comparative results and discussion are elaborated before the final section concludes this paper.

## **METHODOLOGY**

In this paper the main focus is on designing an Ankle Rehabilitation System (ARS). The ankle motion (angle and position) for rehab exercises is collected as input data. Thus, the system concept is designed based on the angle and position of the subject's foot and ankle during ankle rehabilitation exercises. The methodology of designing and implementing the preliminary ankle rehabilitation system is important to obtain the best input data. Hence, it is clearly shown in Figure 1. The needs identification of ARS is recognized for the purpose of some initial designs. Based on the needs identification, the conceptual design is done by giving precedence to hypothetical functions. The creation of new ideas is suggested to build an efficient ARS. These creations will generate some conceptual options for the ARS, which then need to be compared so as to select the best ARS before system implementation.

### **Needs Identification**

The needs identification is the process of collecting the users' (doctors, physiotherapists and patients) needs for the purpose of establishing a design goal and specification of ankle rehabilitation systems. The needs identification is used to overcome the problem statement that is faced in this project and guide the ankle rehabilitation system development process [4]. The problems faced by the medical experts are poor compliance by the patients, the limited number of experts and the huge work load. Besides that, there are problems faced by patients, such as rehab exercising at home

without proper clinical guidance, long distance to the rehab centre and expensive medical costs. Those problems can be overcome by the ARS needs identification.

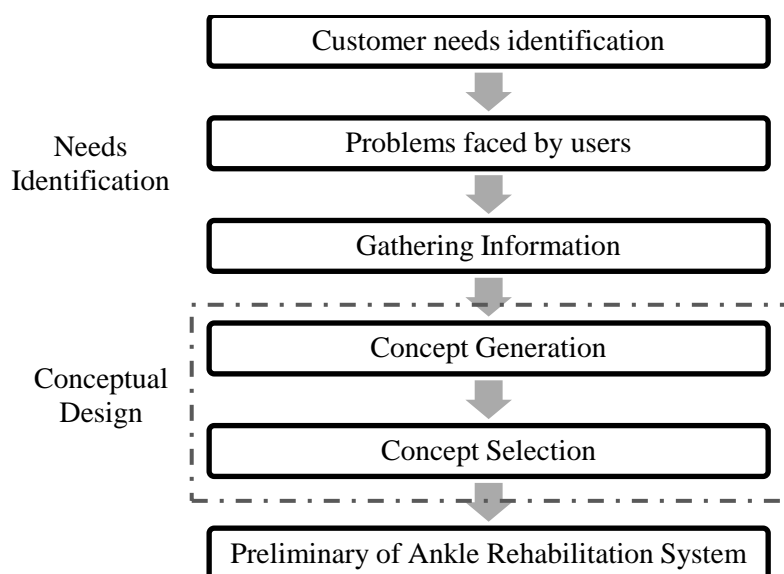


Figure 1. Methodology of design and implementation of preliminary ankle rehabilitation system (ARS)

The needs identified for the Ankle Rehabilitation System (ARS) are validity, user-friendliness, efficiency, and low cost. Validity means that the evaluation result from the system must be valid also in the doctor's evaluation. In fact, the system is able to capture and record the ankle exercises at home when recuperating from ankle sprains. Moreover, this system must be user-friendly in terms of being semi-portable, with suitable system feedback and being easy for patients, physiotherapists and doctors to use. A semi-portable ARS is designed because the patient's ankle injury makes it difficult to move. Thus, the semi-portable system can be set up in a home environment with proper clinical guidelines for users. The system must be safe for patients to use. The important issue is that the system must be simple to use so that it is applicable in the home. Thus, this system can reduce the requirement for face-to-face therapy between experts and patients. The third need identified for system performance is efficiency. For efficient system performance it must be suitably robust, convenient timing, highly accurate and semi-automated. Robustness of the ARS refers to the ability of system algorithms to accommodate change without adapting its initial stable configuration. The time taken for the computing process should be short and convenient. Therefore, the convenience of the time taken is computed starting from capturing images of the patient until the analysis results. The time taken should be about 1–2 seconds for processing each data input. The cost of the ARS should be cheaper than conventional ankle rehabilitation. Besides that, the ARS should be low maintenance and inexpensive to repair.

### Conceptual Design

Based on the needs identification for the ARS in the previous section, a number of solutions are not suitable for designing the ARS. For example, non-visual tracking

sensors (electromyogram, foot scan, accelerometer, inertial tracking sensor, etc.) are not readily required for ARS according to the needs identification. This is because of misallocation of the clinical sensor, obstructions for ankle motion, and so on. Thus, according to the needs identification, visual tracking sensors are more suitable for the ARS design than non-visual tracking sensors. The two main aspects that need to be considered in the conceptual generation are computer systems and visual tracking sensors. From these two main aspects, we found that six conceptual options are generated: - (1) laptop with RGB video camera; (2) laptop with single structured light camera (SLC); (3) laptop with multiple SLCs; (4) desktop PC with RGB video camera; (5) desktop PC with SLC; (6) desktop with multiple SLCs. The morphological method is used for creating a design matrix for the stages of presenting and evaluating the alternative conceptual options for the ARS. Based on Table 1, Pugh’s method is used for comparing and selecting the most suitable concepts for the ARS.

Table 1. Pugh’s method to select Ankle Rehabilitation System.

Options	Datum	Option (1)	Option (2)	Option (3)	Option (4)	Option (5)	Option (6)
Computer system		Laptop	Laptop	Laptop	Desktop	Desktop	Desktop
Visual tracking sensors	Conventional ankle rehab method	RGB video camera	Single SLC	Multiple SLC	RGB video camera	Single SLC	Multiple SLC
Input images data		2D RGB images	RGB-D images	RGB-D images	2D RGB images	RGB-D images	RGB-D images
Validity	Manually	-	+	+	-	+	++
Simple to use	Difficult to seek	-	++	+	-	++	+
Medical feedback	Doctor comment	-	+	++	-	+	++
Safety	Safe	+	+	+	+	+	+
Semi-portable		+	+	+	+	++	++
Accuracy	≈90%	-	+	++	-	+	++
Semi-automatic		-	+	+	-	+	+
Processing time		-	-	-	-	++	++
Data storage	No systematic data storage	-	-	-	+	++	++
System upgrades	Not applicable	-	-	-	+	+	+
Σ+		2	8	9	4	14	16
Σ-		8	3	3	6	0	0

For comparing and selecting the most suitable concepts for the visual tracking Ankle Rehabilitation System, we will use the comparative and elimination approaches. Firstly, we will compare the laptop and desktop PC in terms of the computer system. Secondly, we will eliminate the unqualified conceptual options. Thirdly, we will compare the RGB video camera and structured light camera (SLC) in terms of visual tracking sensors. One of the best visual tracking sensors will be selected. After that, we will compare using either a single camera or multiple cameras for the system set-up. Finally, the best concept option will be selected from this comparative and elimination approach. The selected concept will then undergo implementation to develop the preliminary visual tracking ARS.

## **IMPLEMENTATION**

The development of the preliminary visual tracking Ankle Rehabilitation System is installed and experimental activities are conducted to ensure that the system operates well and can be used effectively during prototyping. This implementation of the preliminary ARS is based on concept option (6) with two Kinects and a desktop computer. In terms of the computer system, desktop model Compaq Elite 8300 CMT with processor Intel® Core™ i5-3570 CPU @ 3.40 GHz, 6 MB Cache, 4 Cores is used. This computer system is interfaced with the Kinect sensors by using a USB 2.0 connection for the data interface. The USB 2.0 connection is less noisy and less subject to interruption compared to wireless communication. The graphical user interface (GUI) displays the data from the sensor using Visual Studio 2012. In terms of visual tracking sensors, a configuration of two Kinect sensors is tested to obtain the best configuration and system set-up. Thus, two types of configuration are investigated: an opposite camera and a side-by-side camera configuration [5]. These two types of configuration are shown in Figure 2 (a) and (b) and have been investigated from the camera positioning experiment.



(a) Opposite camera configuration



(b) Side-by-side camera configuration

Figure 2. Configuration of structural light camera.

## **RESULTS AND DISCUSSION**

Initially, we compare the laptop and desktop computer in terms of the computer system. The laptop model is a Dell Inspiron 15R (N5110) Direct Base with 2<sup>nd</sup> generation Intel® Core™ i7-2630QM @ 2.0 GHz base, up to 2.9 GHz, Qual Core, 8T, 6 MB L3; whereas the desktop model is an HP Compaq Elite 8300 CMT Intel® Core™ i5-3570

CPU @ 3.40 GHz, 6 MB Cache, 4 Cores/ 4 Threads/ Turbo. The costs of both models are similar. Although both computer systems are applicable with visual tracking sensors, we found that the desktop computer is the most appropriate computer to use in the ARS. This is because the performance of the desktop is higher than the laptop. The desktop processor has a higher processor speed (3.40 GHz) than the laptop (2.9 GHz). When multiple devices connect to a computer, the computer must not slow down. Thus, the desktop provides 1 TB hard drive which is able to store more image data and provides a faster data acquisition process compared to the laptop's 500GB. Although a desktop computer is semi-portable, it is suitable to apply for ankle rehabilitation exercises at home for the reason that the patient is also facing the problem of mobility. The maintenance cost of the desktop is less than the laptop because the general electronic board part for the desktop is easily available from the shops and at lower price than for the laptop. The laptop's electronic board is in a limited model chipset, so is less available and more expensive. The desktop is more easily upgraded than the laptop, because its upgrading card and RAM can be added externally. Hence, the desktop is more suitable than the laptop to use in ARS.

After comparing the desktop and laptop, options (1), (2), and (3) can be eliminated. Thus, options (4), (5) and (6) are in the focus. In terms of visual tracking sensors, we next compare the RGB video camera and structured light camera (SLC). The video camera is a normal digital camera that provides a 2D image of the patient's ankle joint as an input. The 2D image is straightforward and matches well with the human visual system's strong response to red, green and blue (RGB) images, as presented in Figure 3. However, working with RGB images presents a number of challenges. The RGB image doesn't contain any information on the distances between points on the surface skin of the ankle joint. Complex mathematical formulae need to be generated for distance calculation from multiple-view RGB images [6]. Therefore, it is very difficult and complex to reconstruct the human ankle joint from RGB images. By using RGB images, it is difficult to get an accurate depth or contour measurement. On the other hand, the human ankle joint is a non-rigid (fluidic) and articulated object. This non-rigid body sometimes has sparse features, which makes the 3D reconstruction process challenging [7]. Localizing ankle shapes in still images remains a fundamental problem in computer vision. Human ankle characteristics with no fixed arches and few corner points make it difficult to reconstruct its motion. As a result, it is difficult to recognize the ankle segments' motion. In addition, there is a problem of post-processing and non-real-time due to the high computing overhead. Due to the common error from RGB images, the non-rigid human ankle motion, the post-processing problem and some other aspects, accurate results are difficult to achieve.



Figure 3. RGB camera used to capture ankle rehabilitation exercises.

However, those challenges can be overcome by using a structured light camera. One successful implementation of an SLC system is in the recently popular game console, the Kinect sensor for Microsoft Xbox. A Kinect was developed to directly determine both depth information and image data – in terms of RGB-D images (Red-Blue-Green-Depth images) [8]. The depth information is the distance between the camera and the subject's foot and ankle and can be directly measured without a complex computation process such as triangulation [9]. Thus, the 3D reconstruction process can be easily done by using RGB-D images [10, 11] and the system can operate in real-time and get better performance. The possibility of producing the depth information will help to distinguish the ankle motion from the background elements [12]. Hence, our concept selection is narrowed down to options (5) and (6).

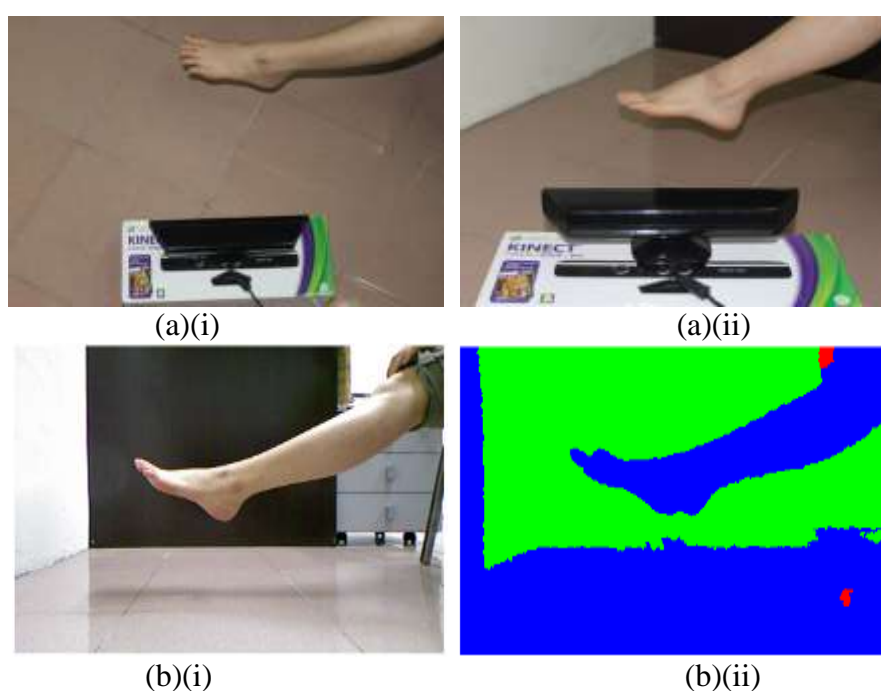


Figure 4. Single Kinect: (a) positioning of Kinect to subject; (i) top view; (ii) side view; (b) input images; (i) RGB images; and (ii) depth images.

Concept option (5) uses a single Kinect which is easy to handle. This concept avoids any overlapping region from multiple Kinects causing interference between them, and avoids also the depth miscalculation shown in Figure 4 (b) (ii). This system can focus on the distance between the Kinect and patient in order to get a highly accurate image. Although the shutters can be used to allow different Kinects to capture data at different times, this reduces the frame rates and exposure time, which also reduces data quality [13]. However, the single Kinect produces limited data for 3D reconstruction of the ankle rehabilitation. There will be overlapping of the ankle medial view, causing limited data for 3D reconstruction, as shown in Figure 4. Thus, this concept is not very suitable to use for ARS. On the other hand, concept option (6) uses multiple Kinects as shown in Figure 5. Using multiple Kinects will present a problem in terms of the interference issue. Thus, only two Kinects are used to minimize the problem with interference. In addition, the interference problems can also be reduced by using certain techniques [14, 15] such as motion-based interference-reducing techniques or colour image de-blurring. Multiple Kinect sensors are used to capture the foot to get

the lateral view and medial view at the same time. Thus, multiple Kinects provide a lot of data and sufficient data for the 3D reconstruction process compare to a single Kinect. As a result, the best concept design for ARS is Option (6), using a desktop with multiple [3] structured light cameras and it provides RBG-D images as input data.

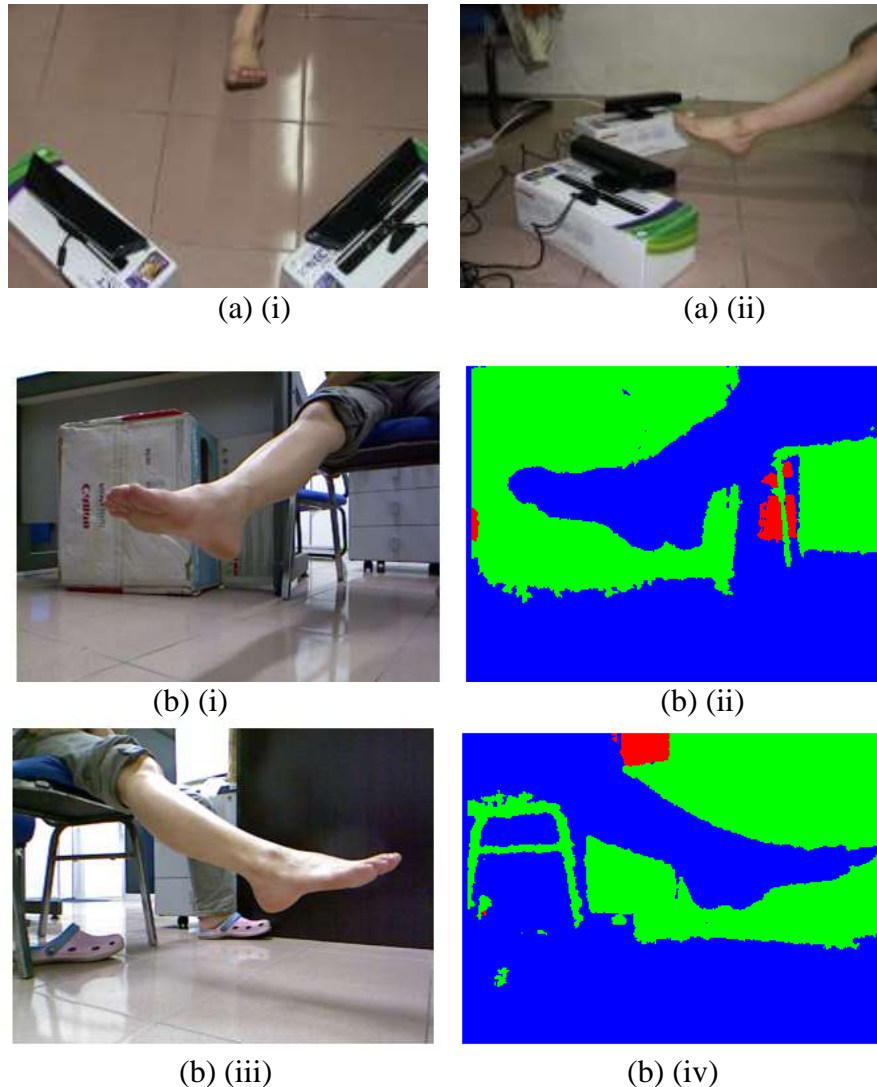


Figure 5. Two Kinects: (a) positioning of Kinect to subject: (i) top view; (ii) side view; (b) Input Images: (i) RGB images from 1<sup>st</sup> Kinect; (ii) depth images from 1<sup>st</sup> Kinect; (iii) RGB images from 2<sup>nd</sup> Kinect; (iv) depth images from 2<sup>nd</sup> Kinect.

In the implementation phase, the configuration of two Kinect sensors is tested to obtain the best configuration and system set-up. The camera configuration and measurement set-up are mainly related to the lighting condition and the imaging geometry. The lighting condition influences the correlation and measurement of disparities. In strong light the laser speckles appear in low contrast in the infrared image, which can lead to outliers or gaps in the resulting point cloud [16]. The imaging geometry includes the distance to the object and the orientation of the object surface relative to the sensor. The operating range of the sensor is between 0.5 m and 5.0 m according to the specifications, and, as we will see in the following section, the random error of depth measurement increases with increasing distance to the sensor [17]. Two



types of configuration are investigated: opposite camera and side-by-side camera configuration [5]. The results of the two types of configuration are shown in Figure 6 (a) and (b). The opposite camera configuration produces more interference than the side-by-side camera configuration, due to incorrect infrared absorption from its opposite camera. The Kinect camera angle should not be adjusted for a negative degree due to the reflection from the shining floor and nearest objects. Thus, the Kinect camera is adjusted for a positive degree to reduce the reflection problem and reduce noise, as shown in Figure 6(a). Hence, the visual tracking ARS features a side-by-side camera configuration with positive degrees of camera angle adjustment, as the best camera configuration, which provides good quality input data.

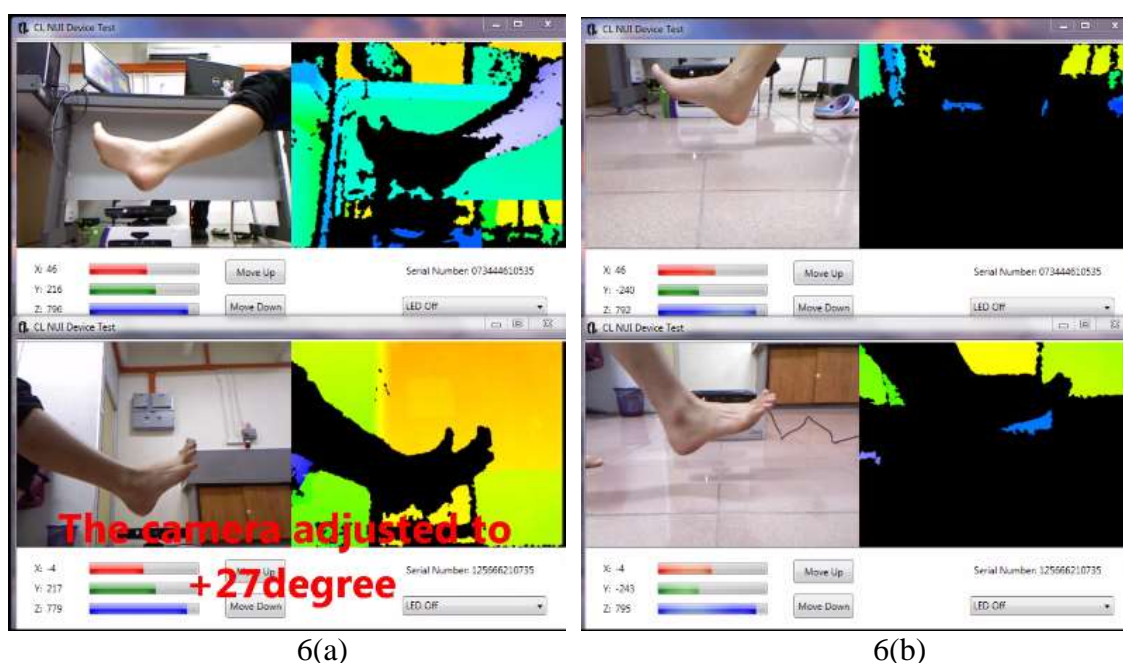


Figure 6. Two Kinect cameras configuration:  
(a) side-by-side camera configuration; (b) opposite camera configuration.

## CONCLUSIONS

In this paper, we presented a novel method for the conceptual design and implementation of a preliminary visual tracking Ankle Rehabilitation System (ARS). From the methodology steps (needs identification and conceptual design), the best conceptual design was identified and found to be two Kinects with a desktop computer. The reasons for selecting two Kinects with the desktop are: the RGB-D image input has similar features to real objects; fewer interference problems; higher processor speed; larger storage; faster data acquisition process; semi-portable; sufficient image processing data compared to the other conceptual option. The camera configuration and measurement set-up for the preliminary visual tracking ARS were implemented. Thus, the best ARS configuration is the side-by-side camera configuration, which gives clear ankle motion analysis data. Future works will improve the ankle rehabilitation system, which includes the calibration of depth values from the depth sensors, 3D reconstruction of the human ankle joints, reducing interference and testing validation of the system.

## ACKNOWLEDGEMENTS

The authors thank University Malaysia Perlis and the Ministry of Higher Education for Research Acculturation Grant Scheme (RAGS).

## REFERENCES

- [1] Lamb SE, Marsh JL, Hutton JL, Nakash R, Cooke MW. Mechanical supports for acute, severe ankle sprain: a pragmatic, multicentre, randomised controlled trial. *The Lancet*. 2009;373:575-81.
- [2] AAOS (American Academy of Orthopaedic Surgeons). Sprained ankle. American Orthopaedic Foot and Ankle Society; 2012.
- [3] University Health Network. Ankle rehabilitation protocol. Toronto Western Hospital; 2013.
- [4] Ulrich KT, Eppinger SD. Product design and development. 5 ed: McGraw-Hill New York; 2007.
- [5] iPiSoft Wiki. User guide for dual depth sensor configuration. 2013.
- [6] Miyahara M, Yoshida Y. Mathematical Transform Of (R, G, B) Color Data To Munsell (H, V, C) Color Data. *Proc SPIE 1001, Visual Communications and Image Processing '88: Third in a Series*. 3 ed. Cambridge, MA1988. p. 650-7.
- [7] Aggarwal JK, Cai Q, Liao W, Sabata B. Articulated and elastic non-rigid motion: a review. *Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects, 1994, . Austin, TX1994*. p. 2-14.
- [8] Hossny M, Filippidis D, Abdelrahman W, Zhou H, Fielding M, Mullins J, et al. Low cost multimodal facial recognition via kinect sensors. *LWC 2012: Potent land force for a joint maritime strategy: Proceedings of the 2012 Land Warfare Conference*. Australia: Commonwealth of Australia; 2012. p. 77-86.
- [9] Shotton J, Sharp T, Kipman A, Fitzgibbon A, Finocchio M, Blake A, et al. Real-time human pose recognition in parts from single depth images. *Commun ACM*. 2013;56:116-24.
- [10] Wiredchop. Kinect biomechanics, engineering sport.
- [11] Henry P, Krainin M, Herbst E, Ren X, Fox D. RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments. In: Khatib O, Kumar V, Sukhatme G, editors. *Experimental Robotics: Springer Berlin Heidelberg*; 2014. p. 477-91.
- [12] Hg RI, Jasek P, Rofidal C, Nasrollahi K, Moeslund TB, Tranchet G. An RGB-D Database Using Microsoft's Kinect for Windows for Face Detection. *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS)*,. 8 ed2012. p. 42-6.
- [13] Gang P, Shi H, Zhaohui W, Yueming W. 3D Face Recognition using Mapped Depth Images. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2005 CVPR Workshops*. San Diego, CA, USA2005. p. 175.
- [14] Maimone A, Fuchs H. Reducing interference between multiple structured light depth sensors using motion. *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*. University of North Carolina at Chapel Hill2012. p. 51-4.
- [15] Butler DA, Izadi S, Hilliges O, Molyneaux D, Hodges S, Kim D. Shake'n'sense: reducing interference for overlapping structured light depth cameras.

- Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, USA: ACM; 2012. p. 1933-6.
- [16] Khoshelham K. Accuracy analysis of kinect depth data. ISPRS workshop laser scanning. University of Twente, Netherlands 2011. p. W12.
- [17] Khoshelham K, Elberink SO. Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. Sensors. 2012;12:1437-54.