ORIGINAL ARTICLE SPECIAL ISSUE: Intelligent Robotics (1/2) Guest edited by Yi-Hung Liu

Design of an Intelligent Soldier Combat Training System

Zen-Chung Wang[1, 2], Ching-Chih Tsai[1, *], and Ming-Chen Chien[2]

[1] *Department of Electrical Engineering, National Chung Hsing University, Taiwan*

[2] *Chung-Shan Institute of Science & Technology, Armaments Bureau, Ministry of National Defense, Taiwan*

(**Received** 5 October2012; **Accepted** 25 October 2012; **Published on line** 1 December 2012)

*Corresponding author: cctsai@nchu.edu.tw

DOI: 10.5875/ausmt.v2i4.164

Abstract: This paper presents an intelligent solider combat training (ISCT) system using X-Box Kinect and augmented virtual reality (AVR) to improve soldier training performance through AVR-based simulation combat fields. In this human-machine interactive system, a systematic design and mechatronic approach is used to construct the overall system structure. The Kinect detects soldiers' positions and postures, situating them in interactive combat soldier training courses. Experimental results demonstrate the effectiveness and applicability of the proposed method.

Keywords: Augmented virtual reality; Human-Robot Interaction; Kinect; soldier combat training

---

1. Introduction

Most modern militaries train their soldiers in highly-realistic simulated battle environments. This realism is widely recognized as increasing the effectiveness of training, but it is expensive to produce. In 1980, the U.S. army introduced a Multiple Integrated Laser Engagement System (MILES) [1] to simulate live fire battlefields for training new recruits. Since then, the role of virtual battlefields in training has expanded and integrated new technologies (the current version of MILES features global positioning systems (GPS) for outdoor training). Despite its effectiveness, MILES is expensive to operate [2]. However, the success of MILES has prompted researchers to apply gaming technologies and/or augmented virtual reality (AVR) to develop new military training simulation systems [3]. For example, in 2011 the U.S. army launched the Dismounted Soldier Training System (DSTS) [4-6], designed to provide training in squad tactics. A single DSTS installation can simultaneously train 9 soldiers, each equipped with a helmet-mounted display (HMD), a head tracking sensor, stereo speakers, a microphone headset, a 3D display processor (image generator), and an instrumented weapon (rifle). In the virtual environment, each trainee controls an avatar, and all trainees on a mission can see and interact with each other's avatars [5].

The effectiveness of this training approach lies in its ability to elicit a wide range of authentic responses and behaviors. The immersive system design in DSTS uses the head tracking and body motion sensors to detect trainee posture and project it onto the avatar in the virtual battle field. Through their head-mounted displays, trainees view a vividly-depicted virtual world through the eyes of their avatars, and interact with the avatars of virtual enemies [3]. DSTS is almost fully way, and its battle scenarios can be designed, edited and deployed through an accompanying suite of software tools. These features make DSTS training programs highly flexible, and reduce the cost and time required to design and execute training programs. However, DSTS is only suitable for familiarizing military units with special missions or specific working environments, or to simulate working procedures. DSTS providers tout the system's many advantages including the ability recreate limitless, lifelike virtual combat scenarios, but system application is restricted by the design of its accessories. Specifically, the HMD blocks the user's eyes making it impossible for the trainee to manually sight his rifle. In addition, trainees must carry a significant amount of gear including a CPU (3D display processor), and head tracking and body motion sensors, making it difficult for trainees to dodge and roll, or engage in other realistic combat movements.

In 2010, Microsoft introduced the Kinect as a motion-sensing input extension of its Xbox gaming console. The Kinect augments the interactivity of games through the detection and recognition of humans and objects, gesture recognition, and target localization. Unlike conventional object recognition approaches, such as SIFT [7] and SURF [8], which work by simultaneously processing images from a stereo camera, the Kinect projects infrared light (IR) and calculates the distance and positions of objects according to the reflected light gathered by an IR-sensitive camera, thus providing more accurate positioning data.

Zen-Chung Wang was born in Taichung County, Taiwan, ROC, in 1972. He received his M.S. degeree from Chung Cheng Institute of Technology, Taoyuan, Taiwan. Currently, he is a Ph.D. candidate in the Department of Electrical Engineering, National Chung-Hsing University, Taichung, Taiwan. Mr. Wang also works as an engineer for Chung-Shan Institute of Science & Technology, Taiwn. His research interest is in electronic control system design, especially in training simulators.

Ming-Chen Chien was born in Nantou County, Taiwan, in 1972. He received B.S., M.S., and Ph.D. degrees in electrical engineering from National Central University, Taiwan, in 1994, 1996 and 2011 respectively. From 1996 to 2000, he was a software engineer of Accton Technology Corporation, where he developed network software for Ethernet switch hub. From 2001 to 2010, he was a lecturer of the Electronic Engineering Department in Asia-Pacific Institute of Creativity, Taiwan. In 2011, he was a lecturer of Multimedia and Game Development Department in the same Institute. He is currently an assistant researcher of Chung-Shan Institute of Science and Technology, Taiwan. His research interests include image processing, game development, video coding, and embedded system design.

Ching-Chih Tsai was born in Taichung County, Taiwan, ROC, in 1961. He received the Diplomat in Electrical Engineering from National Taipei Institute of Technology, Taipei, Taiwan, the M.S. degree in Control Engineering from National Chiao Tung University, Hsinchu, Taiwan and the Ph.D. degree in Electrical Engineering from Northwestern University, Evanston, IL, USA, in 1981, 1986 and 1991, respectively. Currently, he is a distinguished Professor in the Department of Electrical Engineering, National Chung-Hsing University, Taichung, Taiwan. Dr. Tsai has published over 80 journal papers, more than 250 conference papers and seven patents in the fields of control theory, systems technology and applications. Dr. Tsai is respectively the recipients of the Outstanding Automatic Control Engineering Award in 2008, Chinese Automatic Control Society (CACS), and the Outstanding Engineering Professor Award in 2009, the Chinese Institute of Engineers in 2009, and the 2012 IEEE Most active SMC TC award in 2012 . He is an IET Fellow, a CACS Fellow and an IEEE Senior Member. His current interests include advanced control methods, mobile robotics, intelligent service robotics, mechatronics, intelligent learning and methods and their applications to industrial processes and machines.

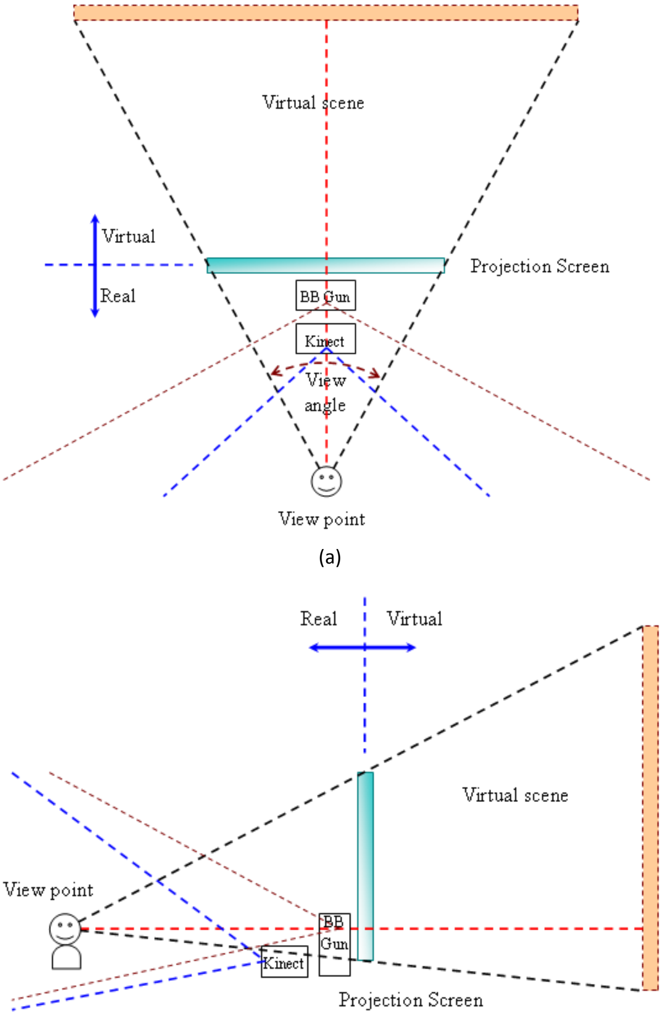
This paper aims to develop a novel intelligent solider combat training (ISCT) system using the Kinect and take the advantages of both MILES and DSTS training systems. The proposed design presents three principal contributions. First, instead of the head tracking and motion sensors used in DSTS, the Kinect sensor is used to detect and analyze trainee position and posture. Second, an overhead projector is used with a large screen to replace the helmet mounted display in DSTS, thereby reducing the equipment trainees must carry and freeing them to engage in more authentic movements. Third, to increase the authenticity of virtual enemies, shots from BB gun mounted on a two-dimensional rotating platform simulates enemy gunfire.

The rest of the paper is organized as follows. Section 2 describes the hardware configuration and software of the proposed intelligent solider combat training system. In Section 3, the Kinect is used to estimate trainee postures and positions. Section 4 is describes the construction of an interactive BB gun control system using the trainee's on-line position information. Experimental results are provided in Section 5 to demonstrate the effectiveness of the proposed methods. Section 6 concludes this paper.

2. Description of the Proposed ICST System

This section describes the hardware configuration and software of the proposed ISCT system. Unlike MILES and DSTS, in the proposed system, the Kinect locates the trainee's skeleton joints in 3D space. These data are then passed to the host computer to analyze the trainee's posture and position, and to predict trainee actions. Based on these analyses and predictions, the virtual enemy depicted on the screen engages in real-time reaction by dodging, attacking, abusing, threatening or provoking. Information on trainee position and posture is also used to control a BB gun mounted on a 2D motion platform which shoots pellets at the trainee to simulate fire from the on-screen enemy. The fire conditions of the BB gun shooting control system are triggered by the virtual enemy that decides which trainee would be the biggest intimidator.

As shown in Figures 1 and 2, the proposed ICST system integrates a large projection screen, a Kinect sensor, a motion platform and a BB gun. As opposed to the wearable head tracking and motion detectors employed in DSTS, the proposed ICTS system uses the Kinect sensor to determine trainee position and posture, and uses a projector/screen setup rather than HMD. Thus, under the proposed system the trainee does not need to carry any special equipment, allowing him to focus more closely on combat training. To increase the authenticity of the enemy threat, an interactive BB gun shooting control system mimics shooting from the enemy, requiring trainees to keep under cover as appropriate. Thus, the proposed system combines the advantages of MILES combat skill training and the rich DSTS virtual environment.



(a)

(b)

Figure 1. Conceptual layout of the Kinect sensor, BB gun, projection screen and virtual scene: (a)top view; (b) side view.

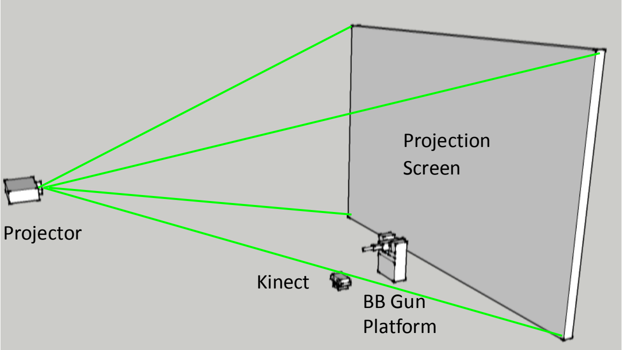


Figure 2. Physical layout of the Kinect sensor, BB gun and projection screen in the ISCT system.

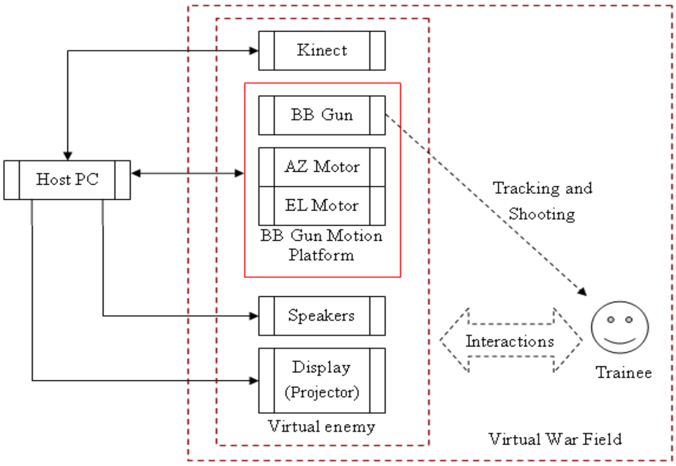


Figure 3. Hardware configuration of the proposed ISCT system.

2.1 Proposed Hardware Configuration

Figure 3 displays the hardware configuration of the proposed ISCT system. The proposed hardware system is consists of a PC-based host computer, a Kinect sensor, a interactive BB gun shooting module, two speakers, a projector and a display module. The Kinect is used to acquire the RGB images, voice data, and trainee depth and 3D position information which is then passed to the host computer. The BB gun is mounted on a two-dimensional motion platform controlled by two independent PI feedback control loops, and can be easily rotated to any desired azimuth and elevation angles. The speaker produces virtual battlefield sounds and enemy voices. The projector and projection screen show the environmental images of the virtual battlefield and enemy images produced by the host PC. The PC-based host computer are used for the following eight functions: (i) analyze depth information from the Kinect, (ii) control the azimuth and elevation angles of the BB gun motion platform, (iii) generate the environmental images of the virtual battlefield, (iv) generate images of the virtual enemy, (v) make decisions or computations for the virtual enemy's movement and behavior, (vi) analyze the trainee's posture and motion, (vii) produce environmental sounds and (viii) produce the virtual enemy's voice.

2.2 Proposed Software

The main software environment and function of the proposed ISCT system consists of (i) Windows XP Operating system, (ii) OPENNI for Kinect data processing, (iii) PrimeSense to drive the Kinect and provide sample NITE codes, (iv) Unity 3D - game development software used for for editing virtual scenes and characters, (v) 3d

Max to create 3D environment model, (vi) Visual Studio 2008 Pro C++. The proposed software architecture is shown in Figure 4.

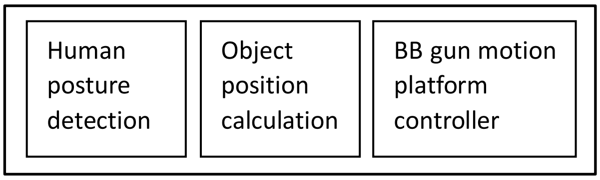


Figure 4. Software modules of the proposed ISCT system.

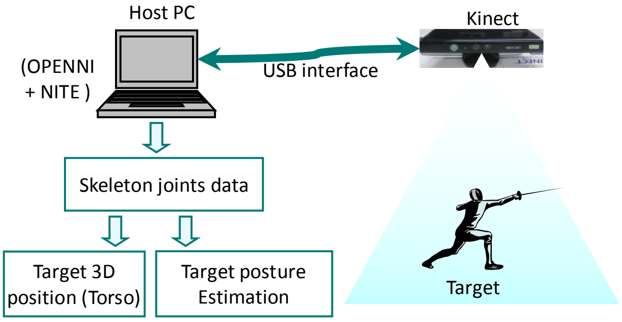


Figure 5. Trainee posture and position estimation via Kinect.

3. Trainee Posture and Position Estimation via Kinect

This section describes how the Kinect can be used to determine trainee position and posture in the virtual combat field. Figure 5 shows the flow process chart for trainee posture and position estimation. In principle, in this limited training environment the Kinect can be used to obtain depth and 3D-positioning data for any recognizable target. After the position data stream is acquired by the host computer, PrimeSense NITE and the OPENNI development environment are directly applied to process the trainee's outlook and appearance, thereby obtaining the trainee's 3D skeleton joint position data. Table 1 lists the fifteen skeleton joints directly processed by the Kinect.

Table 1. Skeleton joints estimated by Kinect.

XN_SKEL_HEAD	**XN_SKEL_LEFT_SHOULDER**	**XN_SKEL_RIGHT_SHOULDER**
XN_SKEL_NECK	**XN_SKEL_LEFT_ELBOW**	**XN_SKEL_RIGHT_ELBOW**
XN_SKEL_TORSO	**XN_SKEL_LEFT_HAND**	**XN_SKEL_RIGHT_HAND**
XN_SKEL_RIGHT_KNEE	**XN_SKEL_LEFT_HIP**	**XN_SKEL_RIGHT_HIP**
XN_SKEL_LEFT_KNEE	**XN_SKEL_LEFT_FOOT**	**XN_SKEL_RIGHT_FOOT**

The Kinect has an optimal detection range of between 1.2 to 3.5 meters due to the limited visual range of the Charge-Coupled Device (CCD) camera and the limited energy of the infrared light. Within this range, the Kinect can detect any three-dimensional target with an accuracy of a few mm. It provides reasonable good detection at distances ranging from 3.5 m to 4.0 m, but at distances greater than 4.5 m, detection accuracy is unreliable due to energy decay of the nonparallel IR light. However, since the combat training system does not require a very high degree of accuracy to determine the trainee's position (judged by the position of the trainee's torso, which is the most easily recognizable feature), the simulated training environment takes up about 5 square meters of space.

The trainee's posture can be easily recognized by his or her joint information. In combat, soldiers adopt a range of generic postures, such as standing, crouching or prone, but also engage in certain motions unique to combat including armed movement, creeping, aiming and shooting. The Kinect acquires information from fifteen joints, and uses the relationship between the joints to infer the trainee's current posture. Table 2 presents logic conditions used to identify standing and crouching postures.

For more advanced posture recognition, three-dimensional coordinates are used with the rotational vector information of each joint to establish more advanced logic conditions to distinguish more sophisticated postures or actions.

Table 2. Logic of two postures

Posture	***Joint Relationship***	***Condition***
Standing	**XN_SKEL_TORSO、XN_SKEL_RIGHT_HIP、XN_SKEL_LEFT_HIP**	**If the y-axis coordinate of XN_SKEL_TORSO is larger than that of YXN_SKEL_RIGHT_HIP or XN_SKEL_LEFT_HIP**
Crouching	**XN_SKEL_TORSO、XN_SKEL_RIGHT_HIP、XN_SKEL_LEFT_HIP**	**If y-axis coordinate XN_SKEL_TORSO is less than that of XN_SKEL_RIGHT_HIP or XN_SKEL_LEFT_HIP in the standing status.**

4. Interactive BB Gun Shooting Control System

This section describes the interactive BB gun shooting control system with two proportional-integral (PI) feedback control loops. The system comprises a

Kinect, a BB gun, a motion platform and a PC-based controller performing two PI feedback control laws. The Kinect is used to locate the three-dimensional positions of the trainee's fifteen skeleton joints. The PC-based controller then analyses potential targets, usually avoiding the trainee's head and neck area for safety. The trainee's position and movements can result in overlapping joints, creating a degree of ambiguity for targeting. This ambiguity can be avoided by defining clear attack priorities: top priority is given to a circle with a radius of 200 mm centered on the middle of the torso, followed by a semi-circular area composed of the right and left shoulders, and another semi-circular area composed of the right and left hips (see Figure 6).

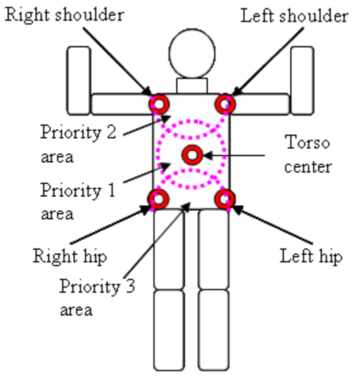


Figure 6. Three prioritized attack targets.

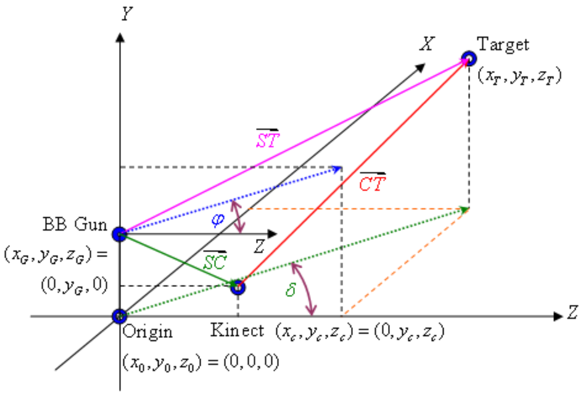


Figure 7. Target position coordinates in the world frame.

The PC-based host controller follows these priorities in selecting a target. The target location is represented by converting its local position from the Kinect into a global position in the world frame, and the global position is used to calculate the azimuth and elevation commands for the interactive BB gun shooting control system. The calculation procedure is described in Figure 7.

As shown in Figure 7, the three-dimensional XYZ Cartesian coordinates in the world frame are used to represent the positions of the Kinect senor, the motion platform and the trainee's attackable targets. In the three-dimensional XYZ Cartesian coordinates, X, Y and Z respectively denote the coordinates of the trainee moving left or right, up or down, and far or near. For example, the distance from the Kinect to the trainee is measured along the Z axis. Therefore, the locations of the Kinect sensor, motion platform and target are denoted by (x_c, y_c, z_c) and (x_T, y_T, z_T). To simplify computation, the motion platform is installed at position (x_G, y_G, z_G)=(0, y_G, 0). To maximize the viewing angle, the Kinect is mounted at (x_c, y_c, z_c)=(0, y_c, z_c). Once the Kinect has detected the attackable target, it computes the target position based on (x_t, y_t, z_t), indicated by $\overrightarrow{CT}$ in Figure 7. Since the relative position between the motion platform and the Kinect is given, the target position with respect to the motion platform is easily obtained from the following relation:

$$\begin{aligned}(x_T, y_T, z_T) &= (x_t + x_c, y_t + y_c, z_t + z_c) \\ &= (x_t, y_t + y_c, z_t + z_c)\end{aligned}. \tag{1}$$

The azimuth and elevation angle commands for the interactive BB gun shooting system are calculated using the known positions of the target (x_T, y_T, z_T) and motion platform (x_G, y_G, z_G) as follows:

$$\delta = \tan^{-1}\{(x_T - x_0)/(z_T - z_0)\} = \tan^{-1}\{x_T / z_T\}, \tag{2}$$

$$\varphi = \tan^{-1}\{(y_T - y_G)/(z_T - z_G)\} = \tan^{-1}\{(y_T - y_G)/z_T\}. \tag{3}$$

Due to limitations of the simulation environment, the azimuth angle command δ is restricted within a range from -150° to +150°, and the elevation angle command φ is limited within a range from -45° to +45°. In fact, both ranges of the azimuth and elevation commands are less than the maximum allowable angles to prevent injuries or accidental damage to equipment.

To validate both derived Equations (1-3), MS Excel was used to compute the azimuth and elevation commands for the BB gun. Table 3 depicts the computed results, thus verifying the correctness of Equations (2-3).

Table 3. Calculated azimuth and elevation commands.

Target-Kinect Position			Kinect Position		
x_t	-600.885	mm	x_c	0	mm
y_t	98.614	mm	y_c	1020	mm
z_t	2719.519	mm	z_c	600	mm
BB Gun Position			Target Position		
x_G	0	mm	x_T	-600.885	mm
y_G	720	mm	y_T	1118.614	mm
z_G	0	mm	z_T	3319.519	mm
BB Gun Angle					
azimuth angle command δ				-10.26034	degrees
elevation angle command φ				6.847403	degrees

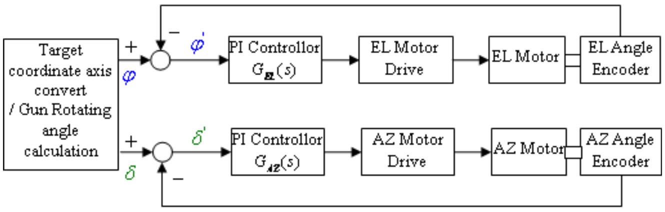


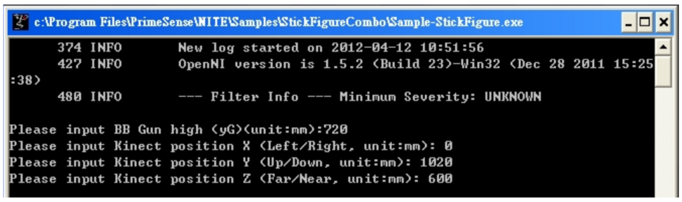
Figure 8. Block diagram of the interactive BB gun shooting control system with two PI feedback control loops.

Figure 8 depicts the block diagram of the proposed interactive BB gun shooting control system with two PI feedback control loops. This system accepts the computed azimuth and elevation commands from Equations (2-3), and executes two independent PI feedback controllers to achieve precise command tracking. The BB gun is fired once the control system has reached the desired azimuth and elevation within the allowable margin of error.

5. Experimental Results and Discussion

This section examines the effectiveness and applicability of the proposed system by conducting three experiments: one to determine the trainee's position, the second to estimate the trainee's posture, and the third to evaluate the performance of the interactive shooting control system.

Prior to the experiments, relevant settings are given as follows. The Kinect sensor was mounted at the IR camera's reference position of both (x_c, y_c, z_c) = (0 mm, 1020 mm, 600 mm); the BB gun motion platform is located at (x_c, y_c, z_c) = (0 mm, 720 mm, 0 mm) at the intersection of the azimuth and elevation rotation axes. Notice that the locations of both devices are expressed in mm because the Kinect sensor's measurements are calculated in mm and mm is used as the basic unit in the computer programs for the sake of simplicity. Figure 9 shows the start-up phase of the experiment with the input of the locations of the Kinect sensor and the motion platform.

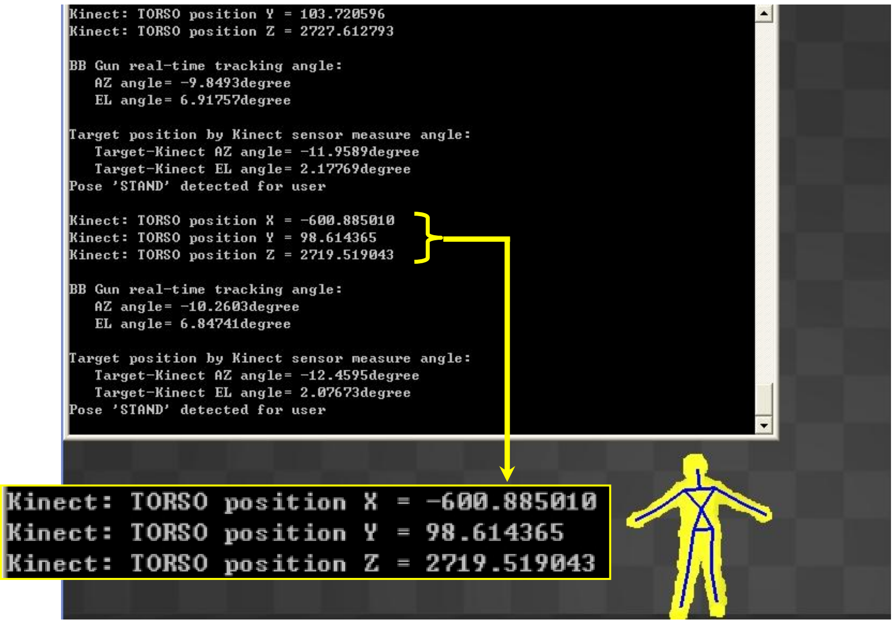


```
     374 INFO       New log started on 2012-04-12 10:51:56
     427 INFO       OpenNI version is 1.5.2 (Build 23)-Win32 (Dec 28 2011 15:25
:38)
     480 INFO       --- Filter Info --- Minimum Severity: UNKNOWN

Please input BB Gun high (yG)(unit:mm):720
Please input Kinect position X (Left/Right, unit:mm): 0
Please input Kinect position Y (Up/Down, unit:mm): 1020
Please input Kinect position Z (Far/Near, unit:mm): 600
```

Figure 9. Set position parameters of the Kinect sensor and BB gun motion platform.

Experiment 1: Position determination

The first experiment was conducted to determine the trainee's position in the system. As mentioned in Section 4, the center of the trainee's torso can be easily determined using the Kinect's IR camera and processing the measured serial data. However, the calculated position of the torso center is based on the IR camera frame. This local position information must be transformed into a global position in the world frame. According to Equation (1), the trainee's global position in the world frame is determined by combining the local position measured by the Kinect sensor and the coordinates of the Kinect sensor in the world frame. As shown in Figure 10, the center position of the trainee's torso as measured by the Kinect is expressed as X = -600.88510 mm, Y = 98.614365 mm, Z = 2719.519043 mm.



```
Kinect: TORSO position Y = 103.720596
Kinect: TORSO position Z = 2727.612793

BB Gun real-time tracking angle:
   AZ angle= -9.8493degree
   EL angle= 6.91757degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= -11.9589degree
   Target-Kinect EL angle= 2.17769degree
Pose 'STAND' detected for user

Kinect: TORSO position X = -600.885010
Kinect: TORSO position Y = 98.614365
Kinect: TORSO position Z = 2719.519043

BB Gun real-time tracking angle:
   AZ angle= -10.2603degree
   EL angle= 6.84741degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= -12.4595degree
   Target-Kinect EL angle= 2.07673degree
Pose 'STAND' detected for user
```

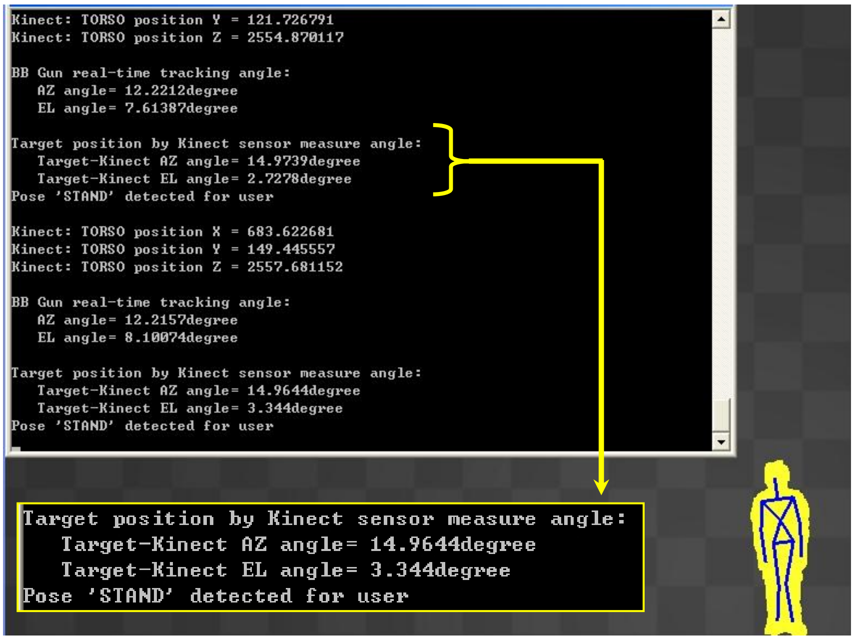
```
Kinect: TORSO position X = -600.885010
Kinect: TORSO position Y = 98.614365
Kinect: TORSO position Z = 2719.519043
```

Figure 10. Kinect measurements of the position of the trainee's central torso.

Experiment 2: Posture estimation

The second experiment was performed to show how the Kinect is used to estimate the trainee's posture. As discussed in Section 4, posture estimation is mainly performed based on the position of the joints. According to the conditions shown in Table 2, two postures can be easily determined according to the y-axis coordinates of the trainee's computed torso center position. Figure 11 depicts the experimental results of a "standing" trainee whose torso center position was measured by the Kinect as (683.6 mm, 149.4 mm, 2557.7 mm), while Figure 12 displays the experimental results of a "crouching" trainee whose torso center position was measured as (705.4 mm, -319.7 mm, 2542.4 mm) . Note that both measurement data sets reveal that the distance from the trainee to the BB gun is about 2560 mm, and there is no obvious difference between the x-axis coordinates of either posture.

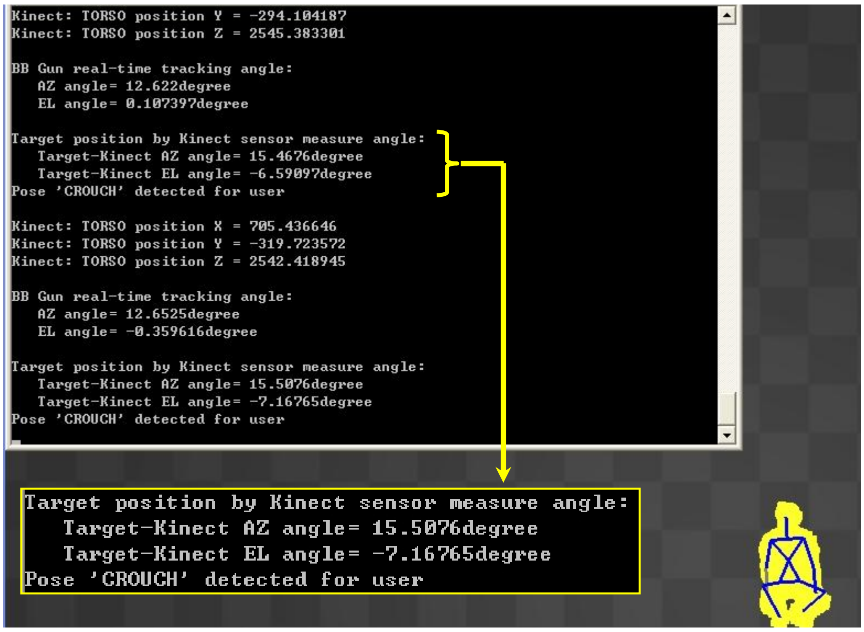
The results in Figures 11 and 12 clearly indicate that the torso center positions for "standing" and "crouching" are significantly differentiated by a 469.1 mm difference in the y-axis coordinate.



```
Kinect: TORSO position Y = 121.726791
Kinect: TORSO position Z = 2554.870117

BB Gun real-time tracking angle:
   AZ angle= 12.2212degree
   EL angle= 7.61387degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= 14.9739degree
   Target-Kinect EL angle= 2.7278degree
Pose 'STAND' detected for user

Kinect: TORSO position X = 683.622681
Kinect: TORSO position Y = 149.445557
Kinect: TORSO position Z = 2557.681152

BB Gun real-time tracking angle:
   AZ angle= 12.2157degree
   EL angle= 8.10074degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= 14.9644degree
   Target-Kinect EL angle= 3.344degree
Pose 'STAND' detected for user
```

```
Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= 14.9644degree
   Target-Kinect EL angle= 3.344degree
Pose 'STAND' detected for user
```

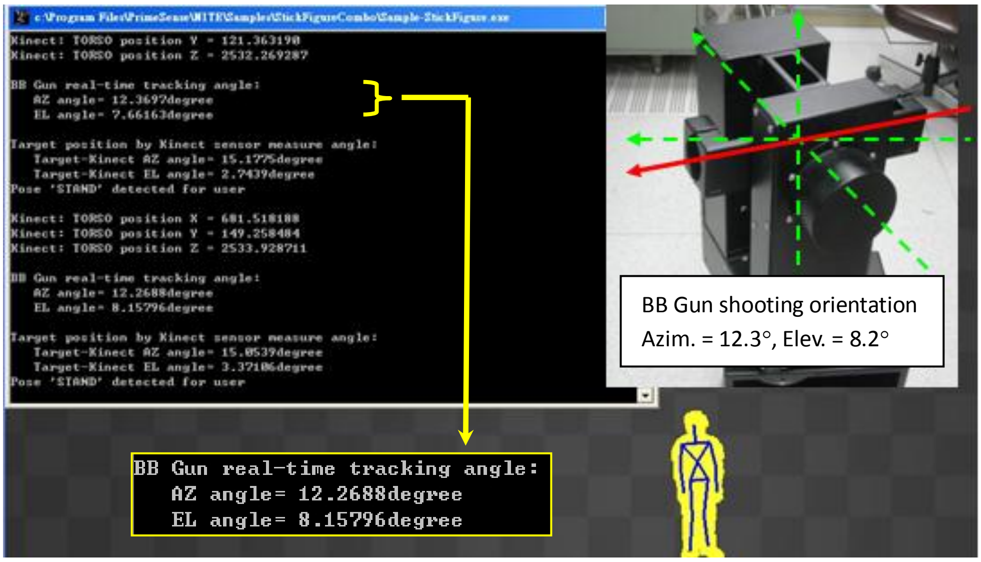
Figure 11. Posture analysis of trainee shown result in 'STAND' status.



```
Kinect: TORSO position Y = -294.104187
Kinect: TORSO position Z = 2545.383301

BB Gun real-time tracking angle:
   AZ angle= 12.622degree
   EL angle= 0.107397degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= 15.4676degree
   Target-Kinect EL angle= -6.59097degree
Pose 'CROUCH' detected for user

Kinect: TORSO position X = 705.436646
Kinect: TORSO position Y = -319.723572
Kinect: TORSO position Z = 2542.418945

BB Gun real-time tracking angle:
   AZ angle= 12.6525degree
   EL angle= -0.359616degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= 15.5076degree
   Target-Kinect EL angle= -7.16765degree
Pose 'CROUCH' detected for user
```

```
Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= 15.5076degree
   Target-Kinect EL angle= -7.16765degree
Pose 'CROUCH' detected for user
```

Figure 12. Posture analysis of trainee shown result in 'CROUCH' status.

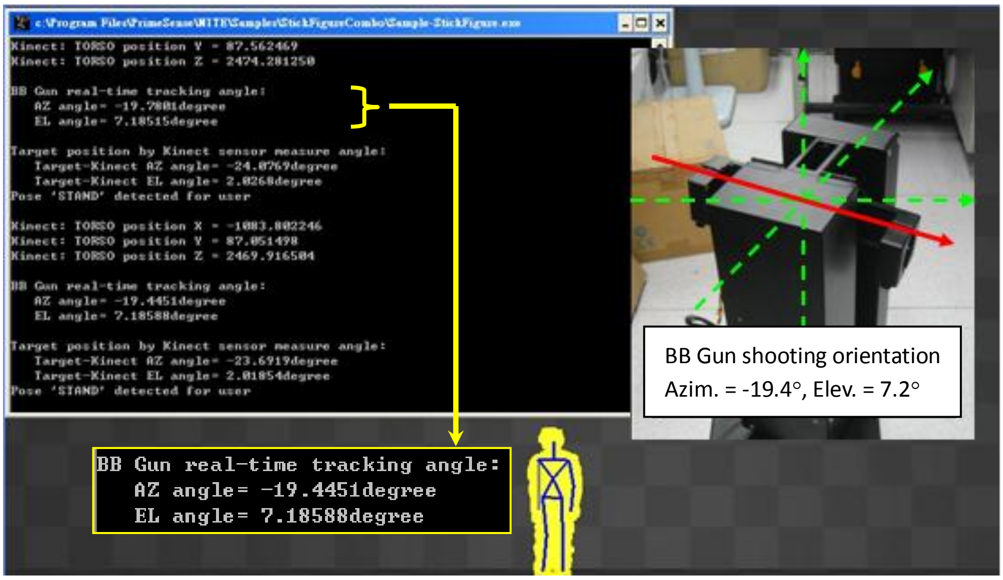


```
Kinect: TORSO position Y = 121.363190
Kinect: TORSO position Z = 2532.269287

BB Gun real-time tracking angle:
   AZ angle= 12.3697degree
   EL angle= 7.66163degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= 15.1775degree
   Target-Kinect EL angle= 2.7439degree
Pose 'STAND' detected for user

Kinect: TORSO position X = 681.518188
Kinect: TORSO position Y = 149.258484
Kinect: TORSO position Z = 2533.928711

BB Gun real-time tracking angle:
   AZ angle= 12.2688degree
   EL angle= 8.15796degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= 15.0539degree
   Target-Kinect EL angle= 3.37106degree
Pose 'STAND' detected for user
```

BB Gun shooting orientation
Azim. = 12.3°, Elev. = 8.2°

```
BB Gun real-time tracking angle:
   AZ angle= 12.2688degree
   EL angle= 8.15796degree
```

Figure 13. Illustration of BB gun turning right and moving toward the desired shooting orientation: Azimuth= 12.3° and Elevation= 8.2°.



```
Kinect: TORSO position Y = 87.562469
Kinect: TORSO position Z = 2474.281250

BB Gun real-time tracking angle:
   AZ angle= -19.7801degree
   EL angle= 7.18515degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= -24.0769degree
   Target-Kinect EL angle= 2.0268degree
Pose 'STAND' detected for user

Kinect: TORSO position X = -1083.802246
Kinect: TORSO position Y = 87.051498
Kinect: TORSO position Z = 2469.916504

BB Gun real-time tracking angle:
   AZ angle= -19.4451degree
   EL angle= 7.18588degree

Target position by Kinect sensor measure angle:
   Target-Kinect AZ angle= -23.6919degree
   Target-Kinect EL angle= 2.01854degree
Pose 'STAND' detected for user
```

BB Gun shooting orientation
Azim. = -19.4°, Elev. = 7.2°

```
BB Gun real-time tracking angle:
   AZ angle= -19.4451degree
   EL angle= 7.18588degree
```

Figure 14. Illustration of BB gun turning left and moving toward the desired shooting orientation: azimuth = -19.4° and elevation = 7.2°.

Experiment 3: Interactive shooting control

The third experiment was carried out to verify the effectiveness of the proposed interactive shooting control system, namely that the motion platform is controlled by the azimuth and elevation commands calculated from Equations (2-3). Figures 13-14 show the azimuth and elevation commands computed from Equations (2-3) and the Kinect's output. As shown in Figure 13, the trainee's torso center was positioned at (681 mm, 149.3 mm, 2533.9 mm), and the azimuth and elevation commands obtained for this position were 12.3° and 8.2°, respectively. After receiving the azimuth and elevation commands, the motion platform with the BB gun first turned right toward the desired orientation, as indicated by the red arrow in Figure 13.

Similarly, when the trainee moved to his left, Figure 14 depicts his new torso center position as (-1083.8 mm, 87.0 mm, 2469.9 mm), with new corresponding azimuth and elevation commands at -19.4° and 7.2°, respectively. Once both commands were inputted to the interactive shooting control system, the PI-controlled motion platform turned left toward this orientation, as indicated by the red arrow in Figure 14. During the movement in Figure 14, the elevation angle was lowered slightly from 8.2° to 7.2°. This angle change was too small to be detected by the human eye, but the moving trajectories of the elevation angle were recorded by the encoders mounted inside the platform.

6. Conclusion

An intelligent solider combat training system was developed using the Kinect interactive gaming device and augmented virtual reality (AVR) to enhance training performance for solders in AVR-based simulated combat scenarios. This human-machine interactive system was constructed through a synthesis of mechatronics, controls and software. Trainee positions and postures in the simulated battlefield are detected by the Kinect, controlling the audio and visual depictions of a the virtual battlefield and enemies, along with the interactive shooting control of a mounted BB gun which fires pellets to simulate enemy fire. Experimental results indicate that the proposed system, combined with the proposed interactive shooting control, provides an immersive and authentic interactive training environment. Future research can investigate the interactive behavior design of the virtual attacker, and to apply Kalman filtering to predict future positions of moving soldiers.

References

[1] Wikipedia, *Multiple integrated laser engagement system*. [Online]. Available: http://en.wikipedia.org/wiki/%20Multiple_Integrated_Laser_Engagement_System [Accessed: April 6, 2012]

[2] *Multiple integrated laser engagement system (MILES)*, Federation of American Scientists, 1999, [Online]. Available: http://www.fas.org/man/dod-101/sys/land/miles.htm [Accessed: April 6, 2012]

[3] M. Macedonia, "Games soldiers play," *IEEE Spectrum,* vol. 39, no. 3, pp. 32-37, 2002. doi: 10.1109/6.988702

[4] *Quantum3D and intelligent decisions to demonstrate U.S. Army dismounted solider training system at AUSA*, Quantum3D, 2011, [Online]. Available: http://www.quantum3d.com/media/news/2011/20111006_Quantum3D.ID%20AUSA%20Press%20Release_Final.pdf [Accessed: April 9, 2012]

[5] *U.S. Army to use first-ever, fully immersive virtual simulation training for American soldiers*, Intelligent Decisions, 2011. [Online]. Available: http://www.intelligent.net/?q=news/press-releases/us-army-use-first-ever-fully-immersive-virtual-simulation-training-american [Accessed: April 9, 2012]

[6] T. Eshel, *Virtual reality prepares soldiers for the real fight*, Defence Update, 2011. [Online]. Available: http://defense-update.com/20111031_virtual-reality-prepares-soldiers-for-the-real-fight.html [Accessed: April 9, 2012]

[7] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision,* vol. 60, no. 2, pp. 91-110, 2004. doi: 10.1023/b:visi.0000029664.99615.94

[8] H. Bay, A. Ess, T. Tuytelaars, and L. van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding,* vol. 110, no. 3, pp. 346-359, 2008. doi: 10.1016/j.cviu.2007.09.014