Prokiwii: A Projected Portable Interface for Enriching Object Scenarios with Kinect and Wii Remote

Mohamed M. Abdel Aty, Ahmed R. Fahmy, Omar M. Elsayed, Ahmed H. Mohamed, Hassan M. Abd El-Azeem, Ahmed I. Rihan, and Ayman Atia

***Abstract*—In this paper we present an intuitive interaction technique for everyday objects. We propose a system that allows the user to interact with everyday objects, and use them to make smart scenarios. We enrich the objects with a portable projected interface that is supported by 3D accelerometer in order to build novel user object interaction scenarios. Moreover, we present double crossing as an interaction technique for manipulating finger interface. We conducted experiment to test the applicability of the proposed system with different applications. Our results showed that the proposed system has a lot of potentials to enrich different domains of applications like object data storage, education Montessori cubes and smart shopping goods.**

***Index Terms*—Object gestures, handheld projector, ubiquitous environments and portable interface interaction.**

I. Introduction

Interaction with real objects is essential in our daily life and could be enriched by portable devices. These devices will make the interaction with the object easier and more interesting. In addition, interaction with real objects is one of the most challenging areas in the scope of ubiquitous environment [1], [2].

In modern life handheld devices are widely used anywhere. They allow many users to communicate, display and store data on it. Nowadays everyone needs an easy way to interact with these handheld devices. Our research is motivated by two main fields; the first field is the vision of using handheld devices to present a dynamic shared interactive space. Portable projectors are small to be held by one hand and can be moved easily [3]. The second field is the interaction with everyday objects using object gestures. We used object gestures instead of hand gestures, because of some social acceptance issues [4]. On the contrary, using the object gestures is more applicable for interaction, as we can use some natural gestures capture from object movements.

We present an interactive system, which combines Kinect sensor that captures the object gestures. Handheld projector is used to display interface over object surface. A 3D accelerometer attached to the displaying surface for more interaction techniques and object gestures tracking.

The proposed system improves objects interaction with several functions as storing and sharing data between objects, configuring specific actions and scenarios for grouped objects or over the object itself. The proposed system is capable of distinguishing between gestures of each object.

II. Related Work

We classified the related work according to object detection, gesture recognition, surface calibration and portable interface interaction.

Object detection depends on types of algorithms used for detection. Different algorithms of detection were discussed by Kandil, H., and Atwan, A. [5]. They made a Comparative study between the Speeded Up Robust Features (SURF) algorithm and the Scale-Invariant Feature Transform (SIFT) algorithm. They found that SURF-Tracker has proven to be more accurate and reliable than SIFT-Particle tracker.

Khan, N. Y. [6] Summarized the performance of two robust feature detection algorithms (SIFT) and (SURF) on several classification data sets. They stated that SURF had also been observed to give good classification results on different data sets. Object scenarios were shown by Karl D.D. Willis *et al.* [7] they designed a system for multi-users interaction using a handheld projector. They made scenarios between two handheld projectors, and then applied these scenarios in different applications like Mobile Content Exchange (file transfer and contact exchange), games (Boxing and Cannon), and Education, Also their system is not tied to a fixed location.

An important aspect of determining the movement in our system is gesture recognition and tracking. Gesture recognition was presented by Ayman Atia *et al.* [8] discussed different types of gestures that based on context and object's shapes. They found that the tapping gesture could be easily achieved because of its fixed pattern. F. Niaki [9] discusses gesture-based interface which using Hand and arm gestures for desktop tasks. Their experiment showed that using gestures on a big-screen display is more natural and pleasant than using a mouse/keyboard in HCI (Human Computer Interaction). Frances *et al.* [10] illustrated two systems designed for 3D gesture user interaction using the Wii remote [11] and Kinect. The study reveals that Wii remote presents several advantages over desktop and mouse. Different types of gesture recognition algorithms were discussed here Andrew D. Wilson *et al* [12] compared $1 recognizer, Dynamic Time Warping, and the Rubine classifier on user supplied gestures. They found that $1 obtains over 97% accuracy with only 1 loaded template and 99% accuracy with 3+ loaded templates. Youngbum Lee *et al.* [13] discussed

Manuscript received June 9, 2013; revised September 30, 2013. This work was supported by Information Technology Industry Development Agency (ITIDA).

The authors are with HCI-LAB, Department of Computer Science, Faculty of Computers and information, Helwan University, Cairo, Egypt (e-mail: ayman@fcih.net).

another type of gesture recognition. They use 2D accelerometers with wireless motion network and use it in fall detection. They state that their system can be applied to patients and elders for activity monitoring, fall detection and sports athletes exercise measurement. Yang Li *et al.* [14] designed a tool that allows a quick access various data items on a mobile phone by drawing gestures on its touch screen. Moreover, they found that gesture enabled fast and easy access to mobile data in their day to day lives.

Li, Y [15] they discussed uWave, an efficient recognition algorithm for such interaction using a single three-axis accelerometer. They used Dynamic Time Warping (DTW) to measure similarities between two-time series of accelerometer readings. Their experiment shows that uWave achieves 98.6% accuracy starting with only one training sample. Zhou Ren *et al* [16] designed a system called (Finger-Earth Mover's Distance (FEMD)). FEMD uses Kinect with Open NI to detect and recognize the human hand. They found that it has the potential to be applied in many other HCI applications that adopt hand as the interface. Furthermore, Chris Harrison *et al* [17] described a system application that allows interactive multi-touch on every day with finger tracking using Kinect. They found that the errors in interacting with a projected surface can be misaligned in the projector and camera calibration, inaccuracy in the fingertip estimation and user inaccuracy when clicking targets.

Using surface calibration with handheld projectors provides dynamic interaction for the user on any surface. Benko, H. [18] presents a set of interaction techniques for supporting collaboration with multiple handheld projectors, and discuss application scenarios enabled by them like casual communication, exchanging contacts and games, scheduling a meeting and music sharing. They stated that handheld projectors provided interesting design challenges.

Willis, K. D. [19] presents the Motion Beam metaphor for character interaction with handheld projectors. With their prototype system, users interact and control projected characters by moving and gesturing with the handheld projector itself. They found that the major challenges when dealing with handheld projectors is to develop interaction techniques that accommodate movement.

Yoshida, T. [20] proposes a novel interface called twinkle for interacting with an arbitrary physical surface using a handheld projector and a camera. The handheld device recognizes the features of the physical environment and displays images and sounds according to the user's motion and collisions of projected images with objects. They applied their system in two different applications named Gaming interface and Music interface. Huber, J. [21] talks about Pico projectors, which have lately been investigated as mobile display and interaction devices. They turned everyday objects sojourning in a beam into dedicated projection surfaces and tangible interaction devices. They designed prototype with the use of Kinect depth image in tracking algorithm. They also analyze the optical flow of detecting objects within the RGB image, to detect if an object has been rotated. Yuji Oyamada *et al*. [22] aims to solve the problem of the calibration object by integrating a planner marker tracking; algorithm that can track its target market even with partial occlusion. They found that more discreet images lead to more accurate camera parameters; continuous frames in a video sequence provide less information than the ones of discredited image. Furthermore, Benko *et al*.[23] designed an interactive system that merges real and virtual worlds into a single spatially registered experience on top of a table. They use contact depth camera and projector to present 3D visualizations to a single user, also They could project their system on any surface of any shape and color. Zhou *et al*.

Furthermore, different types of user interface were shown in the following papers. Nikola Banovic *et al.* N. Banovic [24] investigates how to design context menus for efficient manual multi-touch use. They used the pie menu to interact easily, and they showed that pie menu supports fast menu selection. Hisamatsu and Shizuki *et al.* [25] used a laser pointer as an input device instead of mouse and keyboard. They have designed and implemented a new technique using moving gestures. The system includes two moving operations, crossing and encircling. They found that crossing operation and encircling proved to be easier and effective than the holding technique. Nakamura.T. [26] described a new type of interaction for hand gesture with large displays, which called double-crossing. They made a prototype that based on double crossing, and they also compared single crossing and double crossing in their experiment, and they found that the selection time and error rate for double crossing were better than those for single crossing. So this motivated us to use it in our experiment, and compare its performance with the finger pause technique performance.

III. System Overview

We build a system that provides user by smart interaction techniques to interact with everyday objects. Our target is to improve interaction with objects and create dynamic scenarios that can be reconfigured by the user, instead of static ones as shown in related work.

As shown in Fig. 1 the proposed system consists of Kinect sensor, Hand-held projector and 3D accelerometer attached to any portable surface.

The Kinect is detecting objects, hands and marked surfaces using SURF, OpenNI [27] and TUIO library [28] respectively.

The 3D accelerometer used for tracking the movements of the portable surface. It enables the user to perform several gestures like tilting and tapping, in addition to the object gestures. The Handheld Projector is used for displaying information on the detected object surface according to its location. These ingredients divided the system into three main components:

- Recognition Engine.
- User Interface.
- Calibration Method.

IV. Recognition Engine

We used two devices, Kinect camera and motion sensor. We retrieve the captured frames data and extract it to get the information which will be suitable for the recognition engine. Fig. 2 shows the object detection using SURF algorithm. After recognizing objects, the system starts to get the interaction scenarios which assigned to group of objects or the object itself.

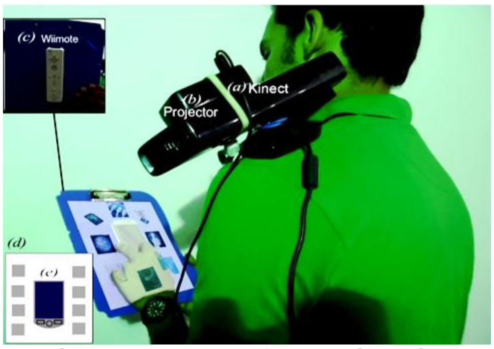


Fig. 1. Proposed system components: (a). Kinect. (b). Pocket projector. (c). Wii remote attached to plan sheet. (d). User interface. (e) Object (PDA).

We used Dynamic Time Wrapping classifier with the Wii remote to track the surface's gestures. We implemented tilting and tapping gestures to control the interface with the movement of the surface. We used the Wii remote to trigger the start and the end of the gesture without using any button to recognize the movement of the surface. The recognition engine using Kinect also detects and tracks the user's finger. There are two techniques to interact with interface controls using finger (Finger Pause, and Finger Crossing).

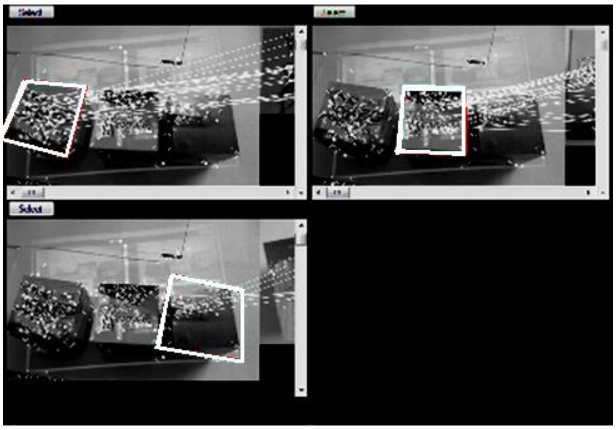


Fig. 2. Multi object detection.

V. User Interface

We designed an interface consists of five forms; user can navigate between forms by swiping gesture. Each frame consists of some dynamic control which is resizable according to surface size and location, this process is called calibration process. Controls content on each form can be dynamically changed relative to any action from users and these controls can be selected or clicked by finger crossing through the edges of this control.

VI. Calibration Method

We used two calibration techniques to fit the display of the projector on the target surface to make the projection dynamic on any surface. We use two types of surface calibration.

- The first technique: calibrating object surface using the output of SURF detection to locate the object and start projection on its surface.
- The second technique: using TUIO (Tangible User Interface Object) to calibrate other portable surfaces by marking the corners of the surface to detect its edges, and control the size of the user interface depending on the distance between the projector and the surface.

VII. Evaluation Experiments

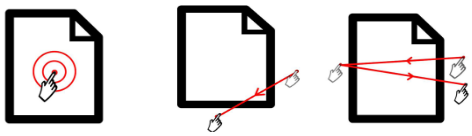
Our setup consists of a 3D accelerometer attached to a plan sheet, Kinect sensor and handheld projector. The Kinect and the handheld projector are fixed upon the user's shoulder. The handheld projector displays the user interface on the plan sheet with resolution of 800 pixels width and 600 pixels height. The experiment performed with 15 users; there ages were between 18 and 21 years old. Each user puts an object in front of the Kinect camera, and then the system will recognize the object and display data on its surface or on the plan sheet.

Each user starts to interact with the projected interface by his finger. We will evaluate two methodologies for user interaction with the projected interface.

A. *Experiment One:*

Measuring the accuracy and the time of selecting items by finger from the projected interface:-

- Selecting by finger pause: user pauses his finger on a specific item for 1 second to be selected.
- Selecting by crossing: user crosses his finger over the edges of a specific item to be selected.



(a) Finger pause (b) Single crossing (c) Double crossing
Fig. 3. Interaction techniques with user interface.

B. *Experiment Two:*

Measuring the accuracy and the time of searching for a specific item in many forms:-

- Searching using finger: user crosses his finger over right and lift edges of the form to swipeit.
- Searching using 3D accelerometer: user tilts the plan sheet then taps on its right or left edge to navigate between forms.

VIII. Results

We performed the experiments in two standing positions. Fig. 4 shows the average time that users took to select an item by finger using the crossing technique or finger pause technique. In each position the user tries to select 4 items. The results show that selecting by the crossing technique is faster than selecting by finger pause technique, as well as, the position of sitting down is easier for the users in selection operation and takes less time.

Fig. 5 shows the average time for users to search for a specific item from 5 forms by swiping using either the finger or the 3D accelerometer gestures.

We observed that the swiping using the 3D accelerometer is faster than the swiping using finger. Moreover, the position of standing up is easier for the users in using 3D accelerometer gestures.

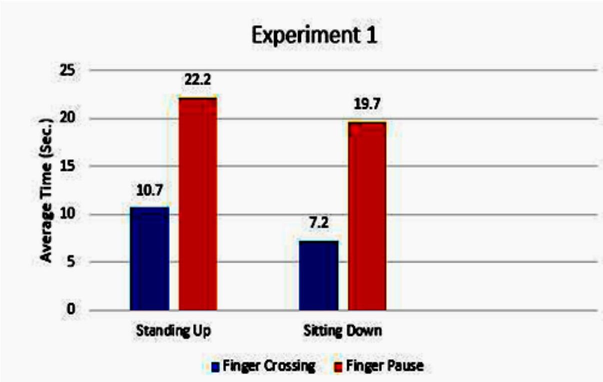


Fig. 4. Experiment 1 results.

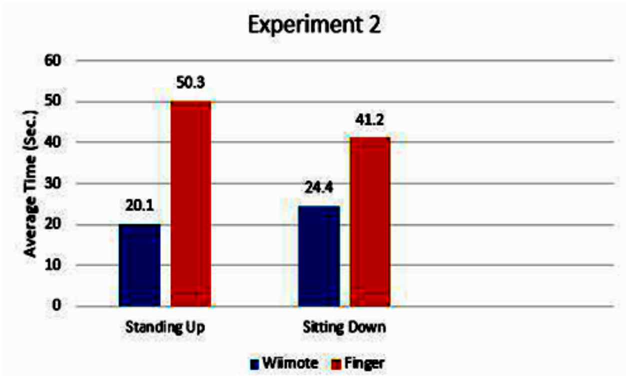


Fig. 5. Experiment 2 results.

IX. Applications

We built some applications which represent some smart scenarios for interaction with objects.

A. Portable Data Storage

In the portable data-storage application, the user can store data (Pictures, Music, Movies, etc...) on real objects virtually as the system detects the object and displays the data stored on it, in its surface then the user can transfer or modify this data.

B. Smart Montessori Cubes

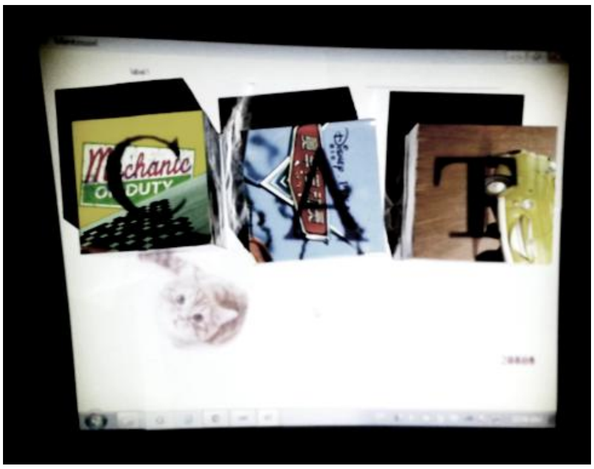


Fig. 6. Smart montessori cubes.

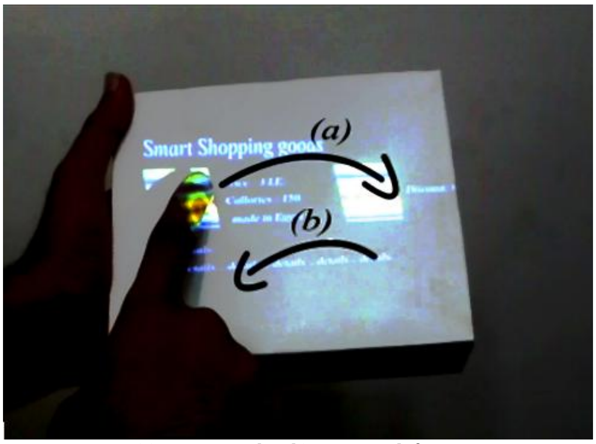
(Cubes used in classrooms for learning process)

Adding new techniques for interaction with Montessori cubes will help to increase its functionality and use less number of cubes to make more scenarios.

In this application, we display letters and drawings on solid cubes surfaces to make children interact with these cubes and manipulate it to get meaningful information through it as shown in Fig. 6.

C. Smart Goods

As shown in Fig. 7 the customer puts the shopping item in front of the Kinect camera then the Kinect detects it and displays information on its surface. Customer can attach the 3D accelerometer sensor to the item and do some gestures using it tilting and tapping to display more information on the shopping item's surface.



(a)Swiping right (b) Swiping left
Fig. 7. Smart goods.

X. Conclusion

We have presented a novel interaction technique for real objects with handheld projector. Through use of Kinect camera and 3D accelerometer, we have shown the ability of using this system to provide various objects' gestures for the user to make smart scenarios. Unlike previous systems which have limitations with the social acceptance by relying on body gestures, our system allows the user to make a various gestures by daily life objects related to its ordinary usage as well as the ability to program any object to make dynamic scenarios.

XI. Future Work

We plan to involve this system in different types of learning process and involve it in some security issues.

References

[1] F. Kawsar, K. Fujinami, and T. Nakajima, "Augmenting everyday life with sentient artefacts," in *Proc. the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-Aware Services: Usages And Technologies*, *SOC-EUSAI* '05, New York, NY, USA, 2005. ACM, pp. 141–146.

[2] L. Console, G. Biamino, F. Carmagnola *et al.*, "Interacting with social networks of smart objects for sharing and promoting cultural heritage,", in *Workshop and Poster Proc. the 20th Conference on User Modeling, Adaptation, and Personalization*, 2012, vol. 872.

[3] E. Rukzio, P. Holleis, and H. Gellersen, "Personal projectors for pervasive computing," *IEEE Pervasive Computing*, vol. 11, no. 2, no. 30–37, April 2012.

[4] P. Y. Chi, J. H. Chen, S. Y. Liu, and H. H. Chu, "Designing smart living objects: enhancing vs. distracting traditional human-object

interaction," in *Proc. the 12th International Conference on Human-Computer Interaction: Interaction Platforms and Techniques, HCI'07*, Berlin, Heidelberg, Springer-Verlag, 2007, pp. 788–797.

[5] H. Kandil and A. Atwan, "A comparative study between sift- particle and surf-particle video tracking algorithm," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 5, no. 3, pp. 12, 2012.

[6] N. Y. Khan, B. McCane, and G. Wyvill, "Sift and surf performance evaluation against various image deformations on benchmark dataset," in *Proc. 2011 International Conference on Digital Image Computing Techniques and Applications (DICTA), IEEE*, 2011, pp. 501–506.

[7] K. D. D. Willis, I. Poupyrev, S. E. Hudson, and M. Mahler, "Side by Side: Ad-hoc multi-user interaction with handheld projectors," in *Proc. the 24th annual ACM Symposium on User Interface Software and Technology, UIST '11*, ACM, New York, NY, USA, 2011, pp. 431–440.

[8] A. Atia, S. Takahashi, and J. Tanaka, "Smart gesture sticker: Smart hand gestures profiles for daily objects interaction," in *Proc. the 2010IEEE/ACIS 9th International Conference on Computer and Information Science, ICIS '10, IEEE Computer Society*, Washington, DC, USA, 2010, pp. 482–487.

[9] F. F. Niaki, R. G. Aghaei, and A. Arya, "Empirical study of a vision-based epth-sensitive human-computer interaction system," in *Proc. the 10th Asia Pacific Conference on Computer Human Interaction, APCHI '12*, ACM, New York, NY, USA, 2012, pp. 101–108.

[10] R. Francese, L. Passero, and G. Tortora, "Wiimote and kinect: gestural user interfaces add anatural third dimension to hci," in *Proc. the International Working Conference on Advanced Visual Interfaces, AVI '12*, New York, NY, USA, 2012. ACM, pp. 116–123.

[11] T. Schlomer, B. Poppinga, N. Henze, and S. Boll, "Gesture recognition with a wiicontroller," in *Proc. the 2nd International Conference on Tangible and Embedded Interaction, TEI'08*, ACM, New York, NY, USA, 2008, pp. 11–14.

[12] J. O. Wobbrock, A. D. Wilson, and Y. Li, "Gestures without libraries, toolkits or training: a $1recognizer for user interface prototypes," in *Proc. the 20th Annual ACM Symposium on User Interface Software and Technology, UIST '07*, New York, NY, USA, 2007. ACM, pp. 159–168.

[13] Y. Lee, J. Kim, M. Son, and M. Lee, "Implementation of accelerometer sensor module and fall detection monitoring system based on wireless sensor network," in *Proc. 29th Annual International Conference on Engineering in Medicine and Biology Society* of *the IEEE, EMBS 2007*, 2007, pp. 2315–2318.

[14] Y. Li, "Gesture search: a tool for fast mobile data access," in *Proc. the 23nd annual ACM Symposium on User Interface Software and Technology*, *UIST '10*, ACM, New York, NY, USA, 2010, pp. 87–96.

[15] J. Liu, L. Zhong, J. Wickramasuriya, and V. Vasudevan, "Uwave: Accelerometer-based personalized gesture recognition and its applications*," Pervasive Mob. Comput*., vol. 5, no. 6, pp. 657–675, December 2009.

[16] J. Zhou, L. Wang, A. Akbarzadeh, and R. Yang, "Multi-projector display with continuous self-calibration," in *Proc. the 5th ACM/IEEE International Workshop on Projector camera systems, PROCAMS '08*, ACM, New York, NY, USA, 2008, pp. 3:1–3:7.

[17] C. Harrison, H. Benko, and A. D. Wilson, "Omnitouch: wearable multitouch interaction everywhere," in *Proc. the 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11*, ACM, New York, NY, USA, 2011, pp. 441–450.

[18] X. Cao, C. Forlines, and R. Balakrishnan, "Multi-user interaction using handheld projectors," in *Proc. the 20th Annual ACM Symposium on User Interface Software and Technology*, ACM, 2007, pp. 43–52.

[19] K. D. D. Willis, I. Poupyrev, and T. Shiratori, "Motion beam: a metaphor for character interaction with handheld projectors," in *Proc. the SIGCHI Conference on Human Factors in Computing Systems, CHI '11*, ACM, New York, NY, USA, 2011, pp. 1031–1040.

[20] T. Yoshida, Y. Hirobe, H. Nii, K. Kawakami, and S. Tachi, "Twinkle: Interacting with physical surfaces using handheld projector," in *Proc. Virtual Reality Conference (VR) 10*, 2010, pp. 87–90.

[21] J. Huber, J. Steimle, C. Liao, Q. Liu, and M. M ̈uhlh ̈auser, "Light beam: nomadic pico projector interaction with real world objects," *CHI '12 Extended Abstracts on Human Factors in Computing Systems, CHI EA '12*, pp. 2513–2518, New York, NY, USA, 2012. ACM.

[22] Y. Oyamada, P. Fallavollita, and N. Navab, "Single camera calibration using partially visible calibration objects based on random dots marker tracking algorithm," in *Proc. ISMAR*, Atlanta, USA, November 2012.

[23] H. Benko, R. Jota, and A. Wilson, "Miragetable: freehand interaction on a projected augmented reality tabletop," in *Proc the SIGCHI Conference on Human Factors in Computing Systems, CHI '12*, ACM, New York, NY, USA, 2012, pp. 199–208.

[24] N. Banovic, F. C. Y. Li, D. Dearman, K. Yatani, and K. N. Truong, "Design of unimanual multi-finger pie menu interaction," in *Proc. the ACM International Conference on Interactive Tabletops and Surfaces, ITS '11*, ACM, New York, NY, USA, 2011, pp. 120–129.

[25] T. Hisamatsu, B. Shizuki, and J. Tanaka, "A novel click-free interaction technique for large-screen interfaces," in *Proc. the 7th Asia-Pacific Conference on Computer-Human Interaction (APCHI2006)*, October 2006.

[26] T. Nakamura, S. Takahashi, and J. Tanaka, "Double-Crossing: A new interaction technique for hand gesture interfaces," in *Proc. the 8th Asia-Pacific Conference on Computer-Human Interaction, APCHI '08*, Berlin, Heidelberg, Springer-Verlag, 2008, pp. 292–300.

[27] Andrew D. Wilson, "Using a depth camera as a touch sensor," in *Proc. ACM International Conference on Interactive Tabletops and Surfaces, ITS '10*, ACM, New York, NY, USA, 2010, pp. 69–72.

[28] M. Kaltenbrunner, T. Bovermann, R. Bencina, and E. Costanza, *TUIO A Protocol for Table-Top Tangible User Interfaces*, 2005.

Mohamed M. Abdel Aty received the B.Sc of computer science from Helwan University at Egypt. He is a member of HCI lab at Helwan University, a member at IEEE, the vice chairman of IEEE FCIH Student Branch 2012, and Committee head at IEEE FCIH SB 2013.

Ayman Atia is an assistant professor at HCI-LAB, Department of computer Science, Faculty of computers and information systems, Helwan University, Egypt. He received his Ph.D. from the university of Tsukuba, Japan 2011. He received the BS and MS degree from the Department of Computer Science, Helwan University, Egypt in 2000 and 2004, respectively. He is a current member of the IEEE computer society.