The ISC Int'l Journal of Information Security

August 2019, Volume 11, Number 3 (pp. 91–96)

http://www.isecure-journal.org

Selected Paper at the ICCMIT'19 in Vienna, Austria

# Hand Gestures Classification with Multi-Core DTW☆

Ayman Atia [1,*], and Nada Shorim [2]

[1] Misr International University, Faculty of computers and information, HCI-LAB, Helwan University

[2] Faculty of Computer Science, Misr International University

ARTICLE INFO.

*Keywords:*
Gesture recognition, DTW, Multi-core processing, Hand gesture classification

**Abstract**

Classifications of several gesture types are very helpful in several applications. This paper tries to address fast classifications of hand gestures using DTW over multi-core simple processors. We presented a methodology to distribute templates over multi-cores and then allow parallel execution of the classification. The results were presented to voting algorithm in which the majority vote was used for the classification purpose. The speed of processing has increased dramatically due to using multi-core processors and DTW.



## 1 Introduction

The number of people owning smart-phones has increased rapidly in the last few years. This gives a chance to build applications that detects gestures to use it easily anywhere and anytime. Gestures are defined as the movement of the hand holding the mobile/smartwatch in certain directions to draw shapes. It is easy to control and interact with a device using gestures, especially for people who are disabled [1].

Hand gestures were used to control IOT devices [2], TVs and to recognize daily activities [3]. Hand gestures were used in medical health applications and hospitals to control surrounding devices or ask for help [4]. This type of gestures is the most suitable way for those people as it does not require touching a device or seeing it clearly.

The main problem for any gesture recognition system is the time consumed to recognize a gesture in a real-time system. Algorithms for machine learning that depends on time signal comparison such as Dynamic Time Warping (DTW) [5] are good candidates for classification of gestures in real time. It was used explicitly for classifying hand gestures in sign language [6], object gestures [7], and abnormal driving behaviour [8].

However, one of the main problems facing these algorithms is the time taken to compare templates with input gestures. DTW proved to work efficiently in several related works with small number of templates, yet for large amount of templates things are different [9]. Probably, DTW, KNN, and SVM took so much time for comparisons with templates [10]. Some trials (e.g., [11]) try to speed up DTW with distributed grid computing for OCR detection. Many trials (e.g., [12–14]) have been done in literature to speed up the DTW. Even though the performance of DTW was progressive, sometimes it fails to run with the big amount of continues time series data online. Hence, there is a need for processing gestures continuously with an acceptable response time.

AirPincher [15] was introduced as a device that

☆ The ICCMIT'19 program committee effort is highly acknowledged for reviewing this paper.

* Corresponding author.

Email addresses: ayman.ezzat@miuegypt.edu.eg, Nada.ayman@miuegypt.edu.eg

ISSN: 2008-2045 

can replace restrictions of spatial space and wearable devices to control virtual objects using finger gestures. Trials for multi-core DTW was presented in [16], and they show windowing approach for real-time hand gestures classification. The idea presented in their work suffers from delay latency around 280 ms between successive gestures classifications.

In this paper, we aim to present a modification to the Dynamic Time Warping algorithm. The algorithm heavily depends on constructing cost path matrix, and hence, tries to align two-time series with each other. The DTW algorithm will consume much time doing a set of comparisons between input gestures and archived templates giving a complexity of $O(n^2)$.

The approach used in this paper is recognizing the gesture using the DTW algorithm and the values of the X-axis, Y-axis, and Z-axis. The values were captured from wearable accelerometer sensors used for hand gesture classifications on multi-core processors.

The rest of the paper is organized as follows; Section 2 explains the related work and other approaches. Section 3 explains our proposed approach. Section 4 shows the experiments and results in terms of measures. Section 5 is the discussion of our system. finally, we conclude our work in Section 6.

## 2 Related Works

Krishna et al. [17] proposed a gesture recognition system based on machine learning that utilizes accelerometer and generates a set of genetic features that can operate in three different modes (User Dependent, Mixed User, and User Independent) according to the end user's choice. To test the system, two public datasets composed of accelerometer-bases gestures are used - uWave and Sony. Krishna et al. stated the best classifier in each category - Efficiency and Classification Time - for a gesture sample, in all three modes after trying several classifiers. Extremely Randomized Trees performed the best in terms of accuracy achieving 98.63% in Mixed User Mode whereas Ridge Classifier achieved the least classification time 0.0013 seconds using small training data.

Carmona and Climent [18] proposed a comparative study between HMM and DTW to know which is better in gesture recognition. Carmona and Climent constructed their own data set using Kinect. Different set of experiments were applied and DTW obtained average accuracy of 98.8% and HMM obtained average accuracy of 96.46%.

Assaleh et al. [19] proposed a system to recognize Arabic sign language using hand gesture and sensor glove. They made a comparison between K-nearest neighbors and DTW. K-nearest neighbors achieved

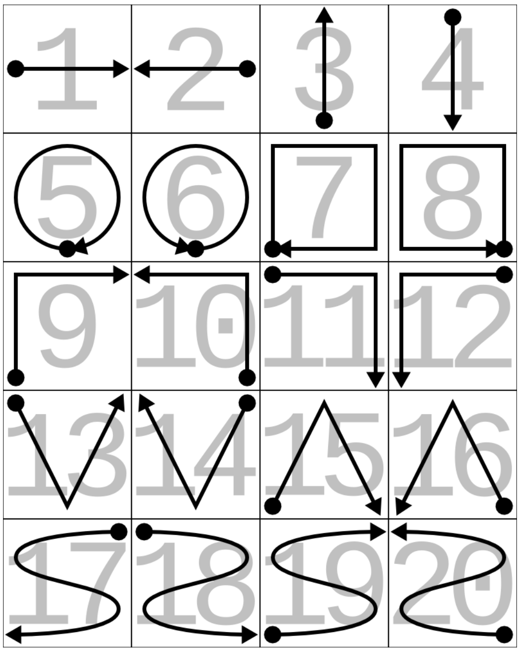


**Figure 1**. Watch hand gestures captured by [20]

an accuracy of 92.5% and 95.3% in user independent and user dependent modes with an average time of 0.27 seconds while DTW achieved an accuracy of 95.1% and 97.5% in user independent and user dependent modes with an average time 4.9 seconds. DTW achieved better accuracy but with higher computation time.

## 3 Proposed System

This research proposes a model in which runs the comparison between the testing and the templates on multi-core processors using DTW. Figure 2 shows the system detailed implementation. Data is collected using accelerometer sensors of a mobile-phone/smartwatch device. The values of X-axis, Y-axis and Z-axis are recorded sequentially in a comma-separated string to be sent for the prepossessing purpose. The data passes by four phases of preprocessing before being passed to the classifier. The cost for DTW was calculated using Euclidean distance as shown in Equation 1 and minimum distance is calculated with Equation 2.

$$Q = \sqrt{(A_x - A_x)^2 + (B_y - B_y)^2 + (C_z - C_z)^2} \quad (1)$$

$$\begin{aligned} Dist(i,j) = min(&Dist(i-1,j)+1, \\ &Dist(i,j-1)+1, \\ &Dist(i-1,j-1)+Q) \end{aligned} \quad (2)$$

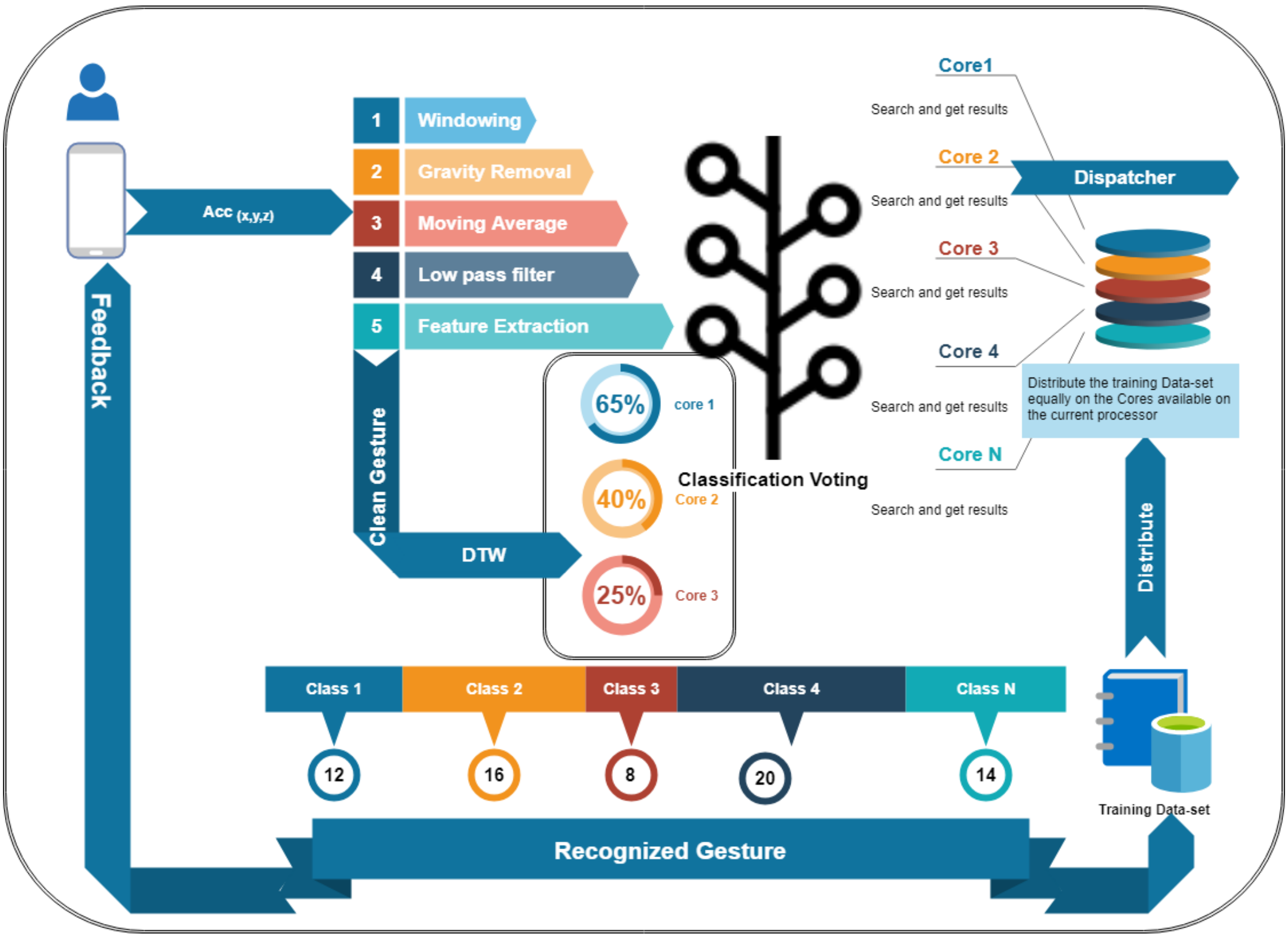


**Figure 2**. System Overview for Multi-core DTW system

### 3.1 Preprocessing

The signal is sent from the mobile phone as a stream of ACC(x,y,z) using a window size of 120 points per second [21] [16]. After the stream received on the server component, the gravity is removed and the linear acceleration is obtained [22]. We apply a moving average function used to smooth the signal [23]. The signal acquired from the accelerometer must pass by low pass filters [21]. Figure 4 shows the values of x, y, and z after applying low pass filter. After removing the gravity value, the values of x,y,z were normalized, then extract a magnitude value to get pattern shape for the signal. The Preprocessed gesture is then passed to the multi-core DTW distributor for processing.

### 3.2 Processing

Dynamic Time Warping was first introduced in [24] and then used in many signal processing applications. It was used for voice recognition [25], Hand and Object gestures classifications [7, 26], and driving behaviour [8]. All templates are distributed equally on a number of available cores.

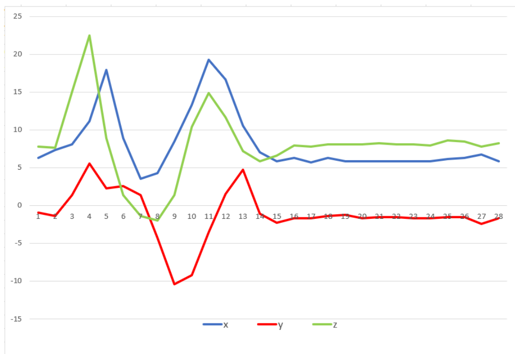


**Figure 3**. Before Preprocessing

**Table 1**. Single core classifiers accuracy and time results

| Classifier | Accuracy (%) | Time (Seconds) |
|---|---|---|
| DTW | 99.2 | 43.7 |
| SVM | 89.2 | 41.1 |
| KNN | 79 | 45 |

The dispatcher is a module that assures that each core receives appropriate templates for loading by dividing the templates over the cores. We have used

**Table 2**. Multi-core results with different processors, time and accuracy

| Processor / Gesture N | $g_1$ | $g_5$ | $g_7$ | $g_{11}$ | $g_{14}$ | $g_{17}$ |
|---|---|---|---|---|---|---|
| I36100-3.7Ghz-4 cores | 0.3 | 0.29,81% | 0.29,86% | 0.28,78% | 0.29,77% | 0.29,88% |
| I73770-3.4Ghz-8 Cores | 0.17 | 0.163,92% | 0.16,79% | 0.16,77% | 0.16,79% | 0.16,80% |
| I74710HQ-2.5Ghz-8 Cores | 0.77 | 0.77 | 0.76 | 0.78s | 0.75 | 0.75 |
| I77500-2.70Ghz-4 Cores | 0.36, 100% | 0.37, 98% | 0.036, 100% | 0.40, 93% | 0.36, 90% | 0.37, 100% |

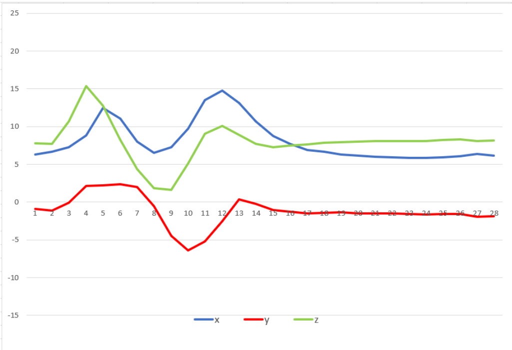


**Figure 4**. After Preprocessing

the dataset presented in [20] for all our experiments. We run the cross-validation between all gestures of the eight users with all 20 shapes. The highest accuracy achieved was 98.2% selecting gesture profiles from all users as templates no (1,9,4,17). Hence, each core will be loaded by the data for user x and the 20 gesture shapes using only profiles mentioned before. It means each core is loaded with 80 template shapes and is ready for applying DTW searching.

The dispatcher checks regularly the cloud data storage for a new gesture that was inserted. After that, the dispatcher starts running the N cores for getting results for DTW per each core. The cores' results are then subjected for counting frequencies and maximum voting is collected. The final decision is then recorded in the data store and a feedback is sent to the user with the appropriate recognized gesture.

## 4 Experiments and Results

The objective of the first experiment is to test the performance of DTW using single core compared to SVM, KNN classifiers.

Figure 1 shows the sampled gestures used in the paper [20]. Hence, each user have 4 templates for each gesture.

### 4.1 Single Core

Different classifiers are build in order to find which will have more accurate result running on single core. Using the dataset in [20] four classes are used in training and the other sixteen classes used in testing. Table 1 shows the result achieved using different classifiers, the time shown in the table is the time taken in training and testing all samples for single core classifications and their accuracy, respectively. The device used in this experiment is equipped with 8.00 GB RAM and CPU core-i7. According to results achieved when using the single core mode, we get the results of DTW in 43.7 seconds

### 4.2 Multi-Core DTW Experiments

In this experiment we measured the effect of multi-core DTW on the time results. We have conducted an experiment to recognize certain gestures; we run the experiment 100 times to measure the average time to classify gesture and the accuracy of detection. The selected gestures were chosen randomly. The experiment was conducted on four devices equipped with Core-i3 (4 cores) and 4 GB RAM, Core-i7 (8 cores) and 16 GB RAM, Core-i7 (8 cores) and 12 GB RAM, and Core-i7 (4 cores) and 8 GB RAM. The average time for classification was 0.28 seconds. The average accuracy for the 100 trials was 92% . Table 1 shows the results for multi-core performance for different samples of testing.

## 5 Results and Discussions

The results shows an increase in the performance regarding the classification speed. The proposed Multi-core DTW system shows dramatically increasing in performance and could be used for online classifications of gesture templates even if their size increased. However, we know that one of the limitations of our algorithm is the presence of at least a number of templates equally divided on the number of cores so we prevent biased voting. Although the number of templates for a single gesture is not equal to other ones, the dispatcher presented in our system overview guaranteed equally division of template gestures over the N cores.

## 6 Conclusions and Future Work

In this paper, we presented a multi-core Dynamic time warping algorithm in a way where it runs parallel and enhances performance in means of time for gesture classifications. We believe that many gesture classification algorithms can benefit from this system as they can run online with acceptable response time even if they run on a simple multi-core primitive servers. Applications for activity recognition, sports interaction and driving behaviours are a good candidate for such system.

## References


[1] Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. Usable gestures for blind people: Understanding preference and performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 413–422, New York, NY, USA, 2011. ACM.

[2] Amr Alanwar, Moustafa Alzantot, Bo-Jhang Ho, Paul Martin, and Mani Srivastava. Selecon: Scalable iot device selection and control using hand gestures. In *Proceedings of the Second International Conference on Internet-of-Things Design and Implementation*, IoTDI '17, pages 47–58, New York, NY, USA, 2017. ACM.

[3] Lei Jing, Zixue Cheng, Yinghui Zhou, Junbo Wang, and Tongjun Huang. Magic ring: A self-contained gesture input device on finger. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*, MUM '13, pages 39:1–39:4, New York, NY, USA, 2013. ACM.

[4] Ahmad Sufril Azlan Mohamed, Effha Binti Soter, Anand Singh, and Nur Intan Raihana Ruhaiyem. Gesture based help identification for hospital & elderlycare using dynamic time warping: A systematic study. In *Proceedings of the International Conference on Video and Image Processing*, ICVIP 2017, pages 94–98, New York, NY, USA, 2017. ACM.

[5] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49, Feb 1978.

[6] Pat Jangyodsuk, Christopher Conly, and Vassilis Athitsos. Sign language recognition using dynamic time warping and hand shape distance based on histogram of oriented gradient features. In *Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '14, pages 50:1–50:6, New York, NY, USA, 2014. ACM.

[7] A. Atia, S. Takahashi, and J. Tanaka. Smart gesture sticker: Smart hand gestures profiles for daily objects interaction. In *2010 IEEE/ACIS 9th International Conference on Computer and Information Science*, pages 482–487, Aug 2010.

[8] Aya Hamdy Ali, Ayman Atia, and Mostafa-Sami M. Mostafa. Recognizing driving behavior and road anomaly using smartphone sensors. *Int. J. Ambient Comput. Intell.*, 8(3):22–37, July 2017.

[9] Krzysztof Pietroszek, Phuc Pham, and Christian Eckhardt. Cs-dtw: Real-time matching of multivariate spatial input against thousands of templates using compute shader dtw. In *Proceedings of the 5th Symposium on Spatial User Interaction*, SUI '17, pages 159–159, New York, NY, USA, 2017. ACM.

[10] Salwa O. Slim, Ayman Atia, and Mostafa-Sami M. Mostafa. An experimental comparison between seven classification algorithms for activity recognition. In Tarek Gaber, Aboul Ella Hassanien, Nashwa El-Bendary, and Nilanjan Dey, editors, *The 1st International Conference on Advanced Intelligent System and Informatics (AISI2015), November 28-30, 2015, Beni Suef, Egypt*, pages 37–46, Cham, 2016. Springer International Publishing.

[11] Maher Khemakhem, Abdelfettah Belghith, and Labidi Mohamed. The dtw data distribution over a grid computing architecture. *International Journal of Computer Sciences and Engineering Systems (IJCSES)*, pages 241–247, 12 2007.

[12] Stan Salvador and Philip Chan. Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11(5):561–580, 2007.

[13] Diego F. Silva, Rafael Giusti, Eamonn Keogh, and Gustavo E. Batista. Speeding up similarity search under dynamic time warping by pruning unpromising alignments. *Data Min. Knowl. Discov.*, 32(4):988–1016, July 2018.

[14] Nguyen Cong Thuong and Duong Tuan Anh. Comparing three lower bounding methods for dtw in time series classification. In *Proceedings of the Third Symposium on Information and Communication Technology*, SoICT '12, pages 200–206, New York, NY, USA, 2012. ACM.

[15] Kyeongeun Seo and Hyeonjoong Cho. Airpincher: A handheld device for recognizing delicate mid-air hand gestures. In *Proceedings of the Adjunct Publication of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST'14 Adjunct, pages 83–84, New York, NY, USA, 2014. ACM.

[16] Florian Grützmacher, Johann-Peter Wolff, and Christian Haubelt. Exploiting thread-level parallelism in template-based gesture recognition with dynamic time warping. In *Proceedings of the 2Nd International Workshop on Sensor-based Activity Recognition and Interaction*, iWOAR '15, pages 6:1–6:6, New York, NY, USA, 2015. ACM.

[17] Gautham Krishna Gudur, Karthik Subramanian Nathan, Yogesh Kumar B, Ankith A. Prabhu, Ajay Kannan, and Vineeth Vijayaraghavan. A generic multi-modal dynamic gesture recogni-

tion system using machine learning. *CoRR*, abs/1809.05839, 2018.

[18] Josep Maria Carmona and Joan Climent. A performance evaluation of hmm and dtw for gesture recognition. In Luis Alvarez, Marta Mejail, Luis Gomez, and Julio Jacobo, editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pages 236–243, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[19] Khaled Assaleh, Tamer Shanableh, and Mohammed Zourob. Low complexity classification system for glove-based arabic sign language recognition. In Tingwen Huang, Zhigang Zeng, Chuandong Li, and Chi Sing Leung, editors, *Neural Information Processing*, pages 262–268, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[20] G. Costante, L. Porzi, O. Lanz, P. Valigi, and E. Ricci. Personalizing a smartwatch-based gesture interface with transfer learning. In *2014 22nd European Signal Processing Conference (EUSIPCO)*, pages 2530–2534, Sept 2014.

[21] Aya Hamdy Ali, Ayman Atia, and Mostafa Sami. A comparative study of user dependent and independent accelerometer-based gesture recognition algorithms. In Norbert Streitz and Panos Markopoulos, editors, *Distributed, Ambient, and Pervasive Interactions*, pages 119–129, Cham, 2014. Springer International Publishing.

[22] Sagar Hursale, Hiren Makwana, Ankit Panchal, Chintan Bhakkad, and Shubhangi Vairagar. Smartphone's based gesture recognition in air. In *Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015)*, ICARCSET '15, pages 49:1–49:5, New York, NY, USA, 2015. ACM.

[23] Ayman Atia, Shin Takahashi, and Jiro Tanaka. Smart gesture sticker: Smart hand gestures profiles for daily objects interaction. In *Computer and Information Science, ACIS International Conference*, pages 482–487. IEEE Computer Society, 2010.

[24] Seiichi Nakagawa and Hirobumi Nakanishi. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49, February 1978.

[25] Seiichi Nakagawa and Hirobumi Nakanishi. Speaker independent english consonant and japanese word recognition by a stochastic dynamic time warping method. *IETE Journal of Research*, 34(1):87–95, 1988.

[26] Ayman Atia, Shin Takahashi, Kazuo Misue, and Jiro Tanaka. Ubigesture: Customizing and profiling hand gestures in ubiquitous environment. In *Human-Computer Interaction. Novel Interaction Methods and Techniques,13th International Conference, HCI International 2009*, pages 141–151, 2009.

**Ayman Atia** Assistant professor at Faculty of Computer Science, Misr International University. Dr Ayman Atia is member of HCI-LAB , Department of computer Science, Faculty of computers and information systems, Helwan University, Egypt. He received his PhD 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 was senior member of the IEEE computer society till 2016 and current member of ACM SIGHCI.

**Nada Shorim** Graduated from the Faculty of Computer Science, Misr International University (2016) with Honors. Doing pre-masters in Software Engineering at Faculty of Computer Science, Cairo University. Now a teaching assistant in Faculty of Computer Science at Misr International University.