RGC Ref. No.:
UGC/FDS11/E03/15
(please insert ref. above)

RESEARCH GRANTS COUNCIL
COMPETITIVE RESEARCH FUNDING SCHEMES FOR
THE LOCAL SELF-FINANCING DEGREE SECTOR

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report
(for completed projects only)

Submission Deadlines:	1. *Auditor's report with unspent balance, if any: within **six** months of the approved project completion date.*
	2. *Completion report: within **12** months of the approved project completion date.*

Part A: The Project and Investigator(s)

1.   Project Title

Vision-based Two-hand Gesture Recognition and Evaluation System for Healthcare

Training

2.   Investigator(s) And Academic Department(s) / Unit(s) Involved

Research Team	Name / Post	Unit / Department / Institution
Principal Investigator	PANG Wai-man / Associate Professor	School of Computing and Information Sciences, Caritas Institute of Higher Education
Co-Investigator(s)	CHOI Kup-sze, Thomas / Associate Professor	School of Nursing, The Polytechnic University of Hong Kong
Others		

3.   Project Duration

	Original	Revised	Date of RGC / Institution Approval *(must be quoted)*
Project Start Date	15/10/2015		
Project Completion Date	14/10/2018	31/08/2018	Institution Approval granted on 2/8/2018
Duration *(in month)*	36	34.5	Institution Approval granted on 2/8/2018

Deadline for Submission of Completion Report	14/10/2019	31/08/2019	

Part B: The Final Report

5.   Project Objectives

5.1   Objectives as per original application

1.   Collection of two hand gestures from vision-based depth sensor array. A database of a large number of finger and palm configurations will be created from human subjects. The gesture base will be used later in artificial intelligence (A.I.) model training, testing and evaluating the proposed algorithm.

2.   Inventing novel vision-based gesture recognition method specific for two hands and multiple depth sensors. The algorithm will involve registration of multiple depth data of two hands, and capable of detecting finger and palm posture with significant self-occlusion. Proper steps for automatic calibration and synchronization between the sensors will also be included.

3.   A similarity metric on the hand gesture is defined for skill level evaluation. As a result, trainee's performance can be scored and properly guided. Further optimization and acceleration on the recognition algorithm will also be considered.

4.   A prototyping serious game for healthcare training is developed based on the hand gesture technique proposed. The game will involve training tasks using both hands, and the hand motions will be examined and evaluated for correctness.

5.   A pilot testing of the training serious game will be carried out at the end of the project. It will help in collecting related data and comments for an in depth user study to ensure the effectiveness of the proposed training approach.

5.2   Revised objectives

Date of approval from the RGC: N/A

Reasons for the change: N/A

5.3   Realisation of the objectives
(Maximum 1 page; please state how and to what extent the project objectives have been achieved; give reasons for under-achievements and outline attempts to overcome problems, if any)

All the objectives proposed are realized with the extent of each explained below:

1.   Work was done to develop a data acquisition setup prototype with multiple cameras to capture hand gestures data from human subjects. Two small to medium sized dataset with over 50K images and around 30 classes are constructed for testing different machine learning approaches in the gesture recognition. **(objective no. 1)**

2.   Three supervised learning approaches are proposed and tested to recognize two hands gestures. Two of them are multi-view solutions, while the rest requires only a single view to achieve satisfactory results. **(objective no. 2)**

3. When comparing with many existing hand gesture recognition solutions like leap motion and Kinect, which are based on a reconstruction based approach, our proposed methods, referred as a kind of holistic features approach, takes the whole image and all available features into consideration and identify a posture label as its outcome. Experiments show satisfactory performance when applying to many multimedia applications. **(objective no. 3 &4)**

4. The multi-view solutions are effective to avoid occlusion in two hand gestures, and we constructed a three camera setup which require no calibration before use. First, we employed classical image features combined with Support Vector Machine for classification. While, as multiple views are involved, an optimized fusion scheme is proposed to obtain a satisfactory recognition result. The second approach is to propose novel deep convolutional neural network model capable in fast and accurate two hand gestures recognition, which is especially suitable for similarity matching under occlusion. **(objective no. 2 & 3)**

5. In our single view solution, we target to tackle the serious confusion between single hand and double hand postures. We propose a novel bi-level learning to mitigate the mentioned misclassification in the bimanual posture recognition. Specifically, a branched network is designed with a novel marginal likelihood layer tailoring for fusing results from branches. **(objective no. 2 & 3)**

6. Other than vision based approaches, works were done in applying mobile sensor technologies for healthcare purposes. We had explored the use of inertial measurement unit (IMU), Brain Computer Interface (BCI) or related technologies in measuring motions, health status of users and skill levels of health professionals. **(objective no. 4)**

7. To bring the proposed approaches into practice, we integrate the techniques investigated into a number of applications, including gesture games, sign language learner, hand washing training system, child health monitor, as well as food consumption advisor app. With these prototype systems and invited users, we carried out preliminary user studies to better understand the user satisfaction, as well as the effectiveness of the provided functionalities. **(objective no. 5)**

The above works resulted in 7 journal paper published, 2 journal papers currently under review, and 1 in preparation. Also, there are 3 conference papers published. The progress of the research is on schedule as expected. However, due to the change of status of the PI and research assistants, some of the research works have a prolonged revision cycle and cannot be published on time.

5.4   Summary of objectives addressed to date

Objectives *(as per 5.1/5.2 above)*	Addressed *(please tick)*	Percentage Achieved *(please estimate)*
1. Construction of two hand gestures database	✓	100%
2. Inventing novel vision-based gesture recognition method for two hands	✓	100%
3. Proposing novel similarity metric for recognizing hand gesture	✓	100%
4. Prototyping serious game for healthcare training	✓	100%
5. A pilot testing of the training serious game	✓	100%

6. Research Outcome

6.1 Major findings and research outcome
(Maximum 1 page; please make reference to Part C where necessary)

We developed three methods which facilitate the recognition of two hands gestures. First, a multi-view approach is proposed with an optimized fusion scheme [J1,C1] of multiple classifiers to enable the use of both strong and weak features from diversified views. With a comprehensive set of experiment comparing to many previous approaches for hand gesture recognition, a better performance can be achieved with the proposed method.

Later, we attempted to revolute the multi-view scheme with deep neural network as the classification base [J2], thus no hand crafted feature is necessary, all feature extractions are done with an improved layer architecture. Furthermore, to avoid retraining when a different combination of view is used, our network are not ended with a classical Fully Connected (FC) layer, but an element-wise multiplicative layer. As a result, any view and its corresponding network branch can be freely attached and detached from the system.

The final approach is inspired by our observation from the captured datasets, in which single hand is usually observed before the subject performs two hand gestures. In this case, it can easily result in ambiguities when large portion of the dual hand pose resembles a particular single hand counterpart. To mitigate these effects, a bilevel branched network (namely Deep2Hands) is invented, which has an extra neural network branch to classify whether it is a mono and duo hand pose [J3]. By combining this extra network with a novel marginal likelihood layer, we surprisingly found that the original classification result is improved. From an extensive number of experiments, we found that the top-level branch behave like a guidance for a better training to the bottom-level branch which estimate for a particular posture.

It is worth to notice that unlike many of the existing gesture recognition solutions which based on depth information, we focus on using off-the-shelf RGB cameras alone, this allows higher flexibility and at the same time incurs lower costs.

Furthermore, we investigate many existing sensing technologies which have potential application to our problem or healthcare related problems. For example, we exploit Brain Compute Interface (BCI) and virtual environment as visual guidance on the modulation of sensorimotor EEG rhythms generated by hand motor imagery (MI) [J4]. In this work, we further improve the classification accuracy on MI, by proposing an algorithm to automatically extract subject-specific optimal frequency and time bands for the discrimination of ERD patterns produced by left and right hand.

The use of mobile sensing units, like Inertial Measurement Unit (IMU) and camera, is another aspect we would like to explore on their application to healthcare area. For example, we track the water consumption based on activity of users (kids) [J5,C2] and by accompanying with mobile application to facilitate a reminder for the parents. Similarly, a food consumption consolation application is also developed with features like vision based food recognition as well as food label reading capability [J6,C3]

Finally, we combined the methodologies to develop a low-cost and accessible training system for hand washing gesture learning [J7]. Our hand washing system contains three different modes according to different learning levels, namely learning, quiz, and free. In all of these modes, the trainee can interact with the system by performing their pose, and receive response instantly on whether the gesture is correct or not.

6.2   Potential for further development of the research and the proposed course of action
(Maximum half a page)

This research project has proposed a number of enabling techniques for bimanual gesture recognition, as well as the prototype application for hand washing learner. Thus, we can have two directions of the potential development. First, our multi-branch deep neural network is general enough to be extended to more generic image recognition tasks, this is also suggested by many of the reviewers of our paper. By exploiting the bi-level labeled data, it is promising to achieve a higher accuracy of recognition even with unbalanced dataset. On the other hand, our prototype hand washing learning application can be immediately used to collect larger number of data by inviting more healthcare professional and general public. Thus, we can have more in depth analysis on the effectiveness of the learning approach as well as to find out ways in improving the efficiency in learning these healthcare related skills via technology.

7. Layman's Summary
(Describe <u>in layman's language</u> the nature, significance and value of the research project, in no more than 200 words)

The research project concerns the development of advanced solutions for two hands gesture recognition for the purpose of hand washing skill training. From the technical aspect, we tackled the difficulties of self-occlusion commonly seen in two hand poses and robust recognition is not found in any existing commercial solution. With the proposed novel multiple view and deep learning architecture, the recognition performance is improved significantly. Extensive experiments are also carried out to prove our methods champion many other previous research works.

For the sake of developing a machine learning solution, we constructed a multi-view setup and employs it to capture a medium sized bimanual hand gesture dataset useful for training and contributing it as a benchmark for the community.
Finally, with the integration of these robust techniques, we developed a comprehensive hand washing training prototype with many functionalities to facilitate healthcare trainees or general public in learning proper hand washing gestures. We evaluated with invited subjects to ensure the proposed methods are useful and bring effective teaching and learning with better interactivity and innovative technology.
(175 words)

Part C: Research Output

8. Peer-Reviewed Journal Publication(s) Arising **Directly** From This Research Project

(Please attach a copy of the publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The Latest Status of Publications				Author(s) *(denote the corresponding author with an asterisk*)*	Title and Journal / Book *(with the volume, pages and other necessary publishing details specified)*	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this Report *(Yes or No)*	Acknowledged the Support of RGC *(Yes or No)*	Accessible from the institutional repository *(Yes or No)*
Year of Publication	Year of Acceptance *(For paper accepted but not yet published)*	Under Review	Under Preparation *(optional)*						
2019				Geoffrey Poon, Kin Chung Kwan, Wai-Man Pang*	[J1] Occlusion-robust bimanual gesture recognition by fusing multi-views

Multimedia Tools and Applications (2019) 78:Pages 23469–23488 | | **Yes** | **Yes** | **Yes** |
| | | ✓ | | Geoffrey Poon and Wai-Man Pang* | [J2]Deep2Hands: A Deep Bimanual Posture Network with Bi-level Learning

Computer Vision and Image Understanding | | **Yes** | **Yes** | **Yes** |
| | | ✓ | | Geoffrey Poon, Kin Chung Kwan, Yoshihiro Yamada, Masakazu Iwamura, Koichi Kise, Wai-Man Pang* | [J3] Deep Bimanual Static Gesture Recognition with Multi-Cameras

ACM Transaction on Multimedia Computing, Communications, and Applications | | **Yes** | **Yes** | **Yes** |
| 2016 | | | | ShuangLiang, Kup-Sze Choi, Jing Qin, Wai-Man Pang*, Qiong Wang, Pheng-Ann Heng | [J4]Improving the discrimination of hand motor imagery via virtual reality based visual guidance

Computer Methods and Programs in Biomedicine Volume 132, August 2016, Pages 63-74 | | **Yes** | **Yes** | **Yes** |

2016				Geoffrey Poon, Kin Chung Kwan, Wai-Man Pang*, Kup-Sze Choi	[J5]Towards Using Tiny Sensor with Heat Balancing Criteria for Child Care Reminders in International Journal of Semantic Computing, September 2016, Vol. 10, No. 03 : pp. 365-378	**Yes**	**Yes**	**Yes**
2018				Wai-Man Pang*, Vane-Ing Tian, Geo ﬀ Poon	[J6] Food Consumption Tracker with Health Advices by Food Photos and Labels in EAI Endorsed Transactions on Ambient Systems 2018, Vol. 5, No. 17, e3.	**Yes**	**Yes**	**Yes**
			✓		[J7] Wash Your Hands Properly: A Hand Washing Training System with Deep Learned Gesture Recognition	**Yes**	**Yes**	**Yes**

9. Recognized International Conference(s) In Which Paper(s) Related To This Research Project Was / Were Delivered
(Please attach a copy of each conference abstract)

Month / Year / Place	Title	Conference Name	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this Report *(Yes or No)*	Acknowledged the Support of RGC *(Yes or No)*	Accessible from the institutional repository *(Yes or No)*
Jul 2018 Shenzhen, China	Real-time Multi-view Bimanual Gesture Recognition	[C1] IEEE International Conference on Signal and Image Processing 2018		**Yes**	**Yes**	**Yes**

Mar 2016 Taipei, Taiwan	Towards using Tiny Multi-Sensors Unit for Child Care Reminders	[C2]IEEE International Conference on Multimedia Big Data 2016		**Yes**	**Yes**	**Yes**
Nov 2016 Milan, Italy	A Mobile Adviser of Healthy Eating by Reading Ingredient Labels	[C3] 6th EAI International Conference on Wireless Mobile Communication and Healthcare.		**Yes**	**Yes**	**Yes**

10. Whether Research Experience And New Knowledge Has Been Transferred / Has Contributed To Teaching And Learning
(Please elaborate)

The major research output in this project are the novel techniques to recognize bimanual

gestures in real-time, accompanied gesture benchmark dataset as well as the protoype of

hand washing training application. Meanwhile, These new knowledge, dataset and prototypes

are delivered in several computer vision or multimedia related courses within our school's

digital entertainment degree programme to enrich student's exposure and knowledge.

11. Student(s) Trained
(Please attach a copy of the title page of the thesis)

Name	Degree Registered for	Date of Registration	Date of Thesis Submission / Graduation
Not applicable			

12. Other Impact
(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, teaching enhancement, etc.)

The paper [C1] titled "Real-time Multi-view Bimanual Gesture Recognition" received the best paper presentation award in IEEE International Conference on Signal and Image Processing 2018.
Moreover, one of the works [J3] "Deep Bimanual Static Gesture Recognition with Multi-Cameras" is collaborated with research team leading by Dr. Masakazu Iwamura and Prof. Koichi Kise from Osaka Perfecture University.

13. Public Access Of Completion Report
(Please specify the information, if any, that cannot be provided for public access and give the reasons.)

Information that Cannot Be Provided for Public Access	Reasons
Nil	

RESEARCH GRANTS COUNCIL
COMPETITIVE RESEARCH FUNDING SCHEMES FOR
THE LOCAL SELF-FINANCING DEGREE SECTOR

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report - Attachment
(for completed projects only)

RGC Ref. No.:	UGC/FDS11/E03/15
Principal Investigator:	Dr. PANG Wai-Man
Project Title:	Vision-based Two-hand Gesture Recognition and Evaluation System for Healthcare Training

Statistics on Research Outputs

	Peer-reviewed Journal Publications	Conference Papers	Scholarly Books, Monographs and Chapters	Patents Awarded	Other Research Outputs (Please specify)
No. of outputs arising directly from this research project [or conference]	6	3	0	0	1 (under preparation)