DEPLOYING A SMART LIGHTING CONTROL SYSTEM
WITH DYNAMIC HAND GESTURE RECOGNITION

HỆ THỐNG ĐIỀU KHIẾN NHÀ THÔNG MINH
SỬ DỤNG NHẬN DẠNG CỬ CHỈ ĐỘNG CỦA BÀN TAY

Huong Giang Doan[1], Duy Thuan Vu[1]

[1]Control and Automation faculty, Electric Power University

Abstract:

This paper introduces a new approach of dynamic hand gestures controlling method. Different from existing method, the proposed gestures controlling method using a cyclical pattern of hand shape as well as the meaning of hand gestures through hand movements. In one hand, the gestures meet naturalness of user requirements. On the other hand, they are supportive for deploying robust recognition schemes. For gesture recognition, we proposed a novel hand representation using temporal-spatial features and syschronize phase between gestures. This scheme is very compact and efficient that obtains the best accuracy rate of 93.33%. Thanks to specific characteristics of the defined gestures, the technical issues when deploying the application are also addressed. Consequently, the feasibility of the proposed method is demonstrated through a smart lighting control application. The system has been evaluated in existing datasets, both lab-based environment and real exhibitions.

Keywords:

Human computer interaction, dynamic hand gesture recognition, spatial and temporal Features, home appliances.

Tóm tắt:

Bài báo đưa ra một phương pháp tiếp cận mới sử dụng cử chỉ động của bàn tay để điều khiển thiết bị điện tử gia dụng. Điểm mới và nổi bật của bài báo là đưa ra một cách thức điều khiển thiết bị điện gia dụng mới sử dụng các cử chỉ động có tính chất chu kỳ trong cả hình trạng và hành trình chuyển động của của bàn tay. Giải pháp đề xuất nhằm hướng tới đảm bảo tính tự nhiên của cử chỉ và giúp hệ thống dễ dàng phát hiện và nhận dạng. Phương pháp biểu diễn chuỗi cử chỉ động sử dụng kết hợp đặc trưng không gian, đặc trưng thời gian và giải pháp đồng bộ pha giữa các cử chỉ. Kết quả thử nghiệm đạt được với độ chính xác lên tới 93,33%. Hơn nữa giải pháp nhận dạng được thử nghiệm trên bộ cơ sở dữ liệu đề xuất và trên cả các bộ cơ sở dữ liệu của cộng đồng nghiên cứu.

Từ khóa:

Tương tác người máy, nhận dạng cử chỉ động của bàn tay, các đặc trưng không gian và thời gian, thiết bị điện tử gia dụng.

1. INTRODUCTION

Home-automation products have been widely used in smart homes (or smart spaces) thanks to recent advances in intelligent computing, smart devices, and new communication protocols. Their most functionality is to maximize the automating ability for controlling items around the house. The smart home appliances can be a range of products from a simple doorbell or window blind to more complex indoor equipments such as lights, doors, air conditioners, speakers, televisions, and so on. In this paper, we intend deploying a human-computer interaction method, which allows users to use their hand gestures to perform conventional operations controlling home appliances. This easy-to-use system allows user interact naturally without any contact with mechanical devices or GUI interfaces. The proposed system not only maximizes user usability via a gesture recognition module but also provides real-time performance.

Although much successful research works in the dynamic hand gesture recognitions [4,5,7,19], deploying such techniques in real practical applications faces many technical issues. On one hand, a hand gesture recognition system must resolve the real-time issue of hand detection, hand tracking, and gesture recognition. On the other hand, a hand gesture is a complex movement of hands, arms, face, and body. Thanks to the periodicity of the gestures, technical issues such as gestures spotting and recognition from video stream become more feasible. The proposed

gestures in [25] also ensure the naturalness to end-users. To avoid limitations of conventional RGB cameras (shadow, lighting conditions), the proposed system uses a RGB-D camera (e.g., Microsoft Kinect sensor [1]). By using both depth and RGB data, we can extract hand regions from background more accurately. We then analyze spatial features of hand shapes and temporal ones with the hand's movements. A dynamic hand gesture therefore is represented not only by hand shapes but also dominant trajectories which connect keypoints tracked by an optical flow technique. We match a probe gesture with gallery one using Dynamic Time Wrapping (DTW) algorithm. The matching cost is utilized in a conventional classifier (e.g., K-Neighnest Neighbour (K-NN)) for labeling a gesture.

We deploy the proposed technique for a **smart lighting control system** such as turn the lamps on/off or change their intensity. Although a number of lighting control products have been designed to automatically turn on/off bulbs when users enter into or leave out of a room. Most of these devices are focusing on saving energy, or facilitating the control via an user-interface (e.g., remote controllers [10], mobile phones [2,17,16], tablets [8,11], voice recognition [3,23]). Comparing with these product, the proposed system deployed in this study is the first one without requirements of the interacting with a home appliance. Considering about user-ability, the proposed system serves well for common

people, and feasibly support to well-being of elderly, or physical impaired/disabled people. A prototype of the proposed system is shown in Fig. 1. The system has been deployed and evaluated in both lab-based environment and real exhibitions. The assessments of user's feelings are analyzed with promising results.
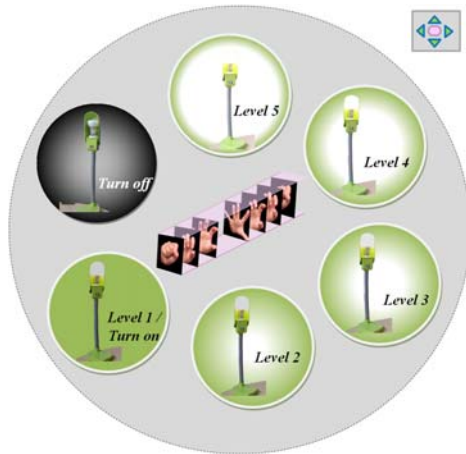
2. PROPOSED METHOD FOR HAND GESTURE RECOGNITION

In this section, we present how the specific characteristics of the proposed hand gesture set will be utilized for solving the critical issues of an HCI application (e.g., in this study, it is a lighting control system). It is noticed that to deploy a real application not only recognition scheme but also some technical issues (e.g., spotting a gesture from video stream) which should be overcome. Fig. 2 shows the proposed framework. There are four main blocks: two first blocks compose steps for extracting and spotting a hand region from image sequence; two next blocks present our proposed recognition scheme which consists of two phases: training and recognition. Once dynamic hand gesture is recognized, lighting control is a straightforward implementation.
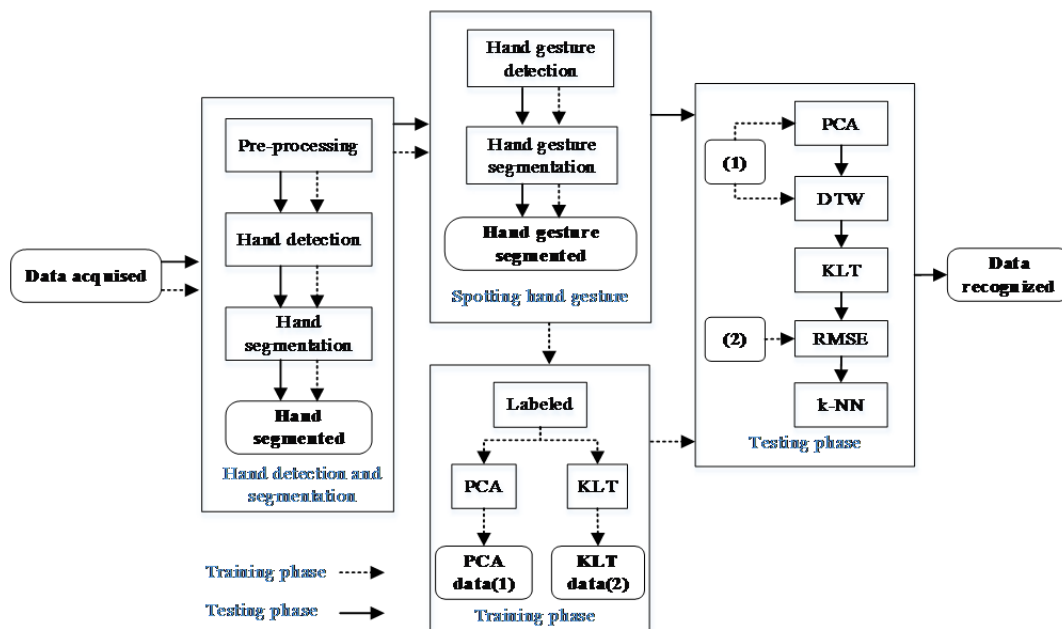

Figure 1. An illustration of the lighting control system. Intensity of a bulb is adjustable in different levels using the proposed hand gestures

Figure 2. The proposed frame-work for the dynamic hand gesture recognition

2.1. Hand detection and segmentation

Pre-processing: Depth and RGB data captured from the Kinect sensor [1] are not measured from the same coordinate system. In the literature, the problem of calibrating depth and RGB data has been mentioned in several works for instance [18]. In our work, we utilize the calibration method of Microsoft due to its availability and ease to use. The result of calibration is showed in Fig. 3 (a)-(b). It is noticed that after the calibration, each pixel in RGB image has corresponding depth value, some boundary pixels of depth image is unavailable.

Figure 3. Hand detection and segmentation procedures.(a) RGB image; (b) Depth image; (c) Extracted human body; (d) Hand candidates

Hand detection: As sensor and environment are fixed, we firstly segment human body using background subtraction (BGS) technique. In general, both depth and RGB images can be used for BGS. However, depth data is insensitive to illumination change. Therefore, in our work we use depth images. Among numerous techniques of BGS, we adopt Gaussian Mixture Model (GMM) [21] because this technique has been shown to be the best suitable for our system [9]. Fig. 3(c) shows human body extraction result.

Hand segmentation: From extracted human body, we continuously extract hand candidates based on distribution of depth features. Fig. 3(d) shows hand candidates obtained at this step. In this example, there are one true positive and

one false positive. The true positive could contain background or miss hand fingers. To remove background and grow hand region to cover all fingers, we apply a step of skin color pruning. Detail of this technique was presented in our previous work [6]. Fig. 4 show intermediate results of hand region segmentation from a hand candidate.
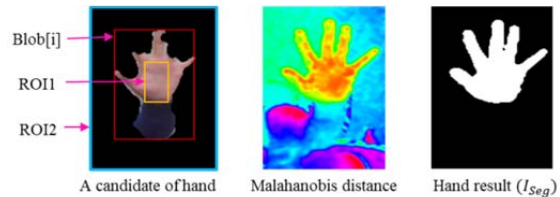

Figure 4. Hand segmentation procedures

2.2. Gesture spotting

In a real application, frames come continuously from video stream. A dynamic hand gesture is a sequence of consecutive hand postures varying in time. Therefore, it is necessary to determine the starting and ending times of a hand gesture before recognizing it. In this study, all pre-defined gesture commands have the same hand shape at starting and ending times. Moreover, hand shapes of a gesture follow a cyclical pattern. We then rely on these properties for gesture spotting as presented in [24].

2.3. Dynamic hand gesture representation

Given a sequence of consecutive frames of a hand gesture, we will extract features for gesture representation. We consider two types of features: spatial features characterize hand shape while temporal features represent hand movement. Both types of these features are important cues for the gesture characterization.

Spatial features: Many types of features could be extracted from hand regions. In this research, we use PCA technique which is now very popular for dimension reduction of feature space. This technique reduces data correlation and computational workload while still keeping enough information to distinguish hand shapes. After segmenting hand region, the image of hand region will be converted to a gray image and resized to the same size X(64x64pixels) and normalized by a standard deviation into X*. Then X* is reshaped into one row matrix Y as (1):

$$Y = [y^1 y^2 .... y^{4096}] \qquad (1)$$

At training phase, we take M hand postures samples from each gesture category G_i with $i = \overline{1, N}$ as (2):

$$S_{Gi} = [Y_{i1} \, Y_{i2} \, ... Y_{iM}] = \begin{bmatrix} y_{i1}^1 & y_{i1}^2 \cdots & y_{i1}^{4096} \\ y_{i2}^1 & y_{i2}^2 \cdots & y_{i2}^{4096} \\ y_{iM}^1 & y_{iM}^1 \cdots & y_{iM}^{14096} \end{bmatrix} (2)$$

A training hand gesture set S = [S_{G1}, S_{G2},..., S_{GN}]^T is input into the PCA algorithm. All parameters and matrices generated from PCA algorithm are stored into a PCA.XML file for further processing. In our work, we keep the first twenty principal components (the most important components) to create a 20-D spatial feature vector for each hand image. Fig. 5 illustrates a sequence of frames of a gesture G_3 (Back) and its projection in the constructed PCA space.
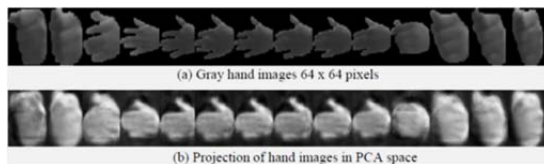

Figure 5. An illustration of the Go_left gesture before and after projecting in the PCA space

Temporal features: In the literature, many methods have been proposed for extracting temporal features of human actions. In our work, we extract hand movement trajectory using KLT (Kanade-Lucas-Tomasi) technique. This technique combines the optical flow method of Lucas-Kanade [14] and the good feature points segmentation method of Shi-Tomasi [20]. This technique was widely utilized in the literature for object tracking or motion representation. The KLT tracker allows describing the trajectory of feature points of hand between two consecutive postures as shown in Fig. 5. This is done through following steps. First, we detect feature points for every frame of the sequence. Then we track these points in the next frame. This is repeated until the end of a gesture. Connecting tracked points in the consecutive frames creates a trajectory. Among generated trajectories, we select the twenty most significant ones to represent a gesture. Fig. 6 illustrates points tracking from several frames and the twenty most significant trajectories.

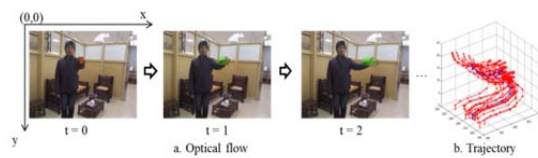**Figure 6. Points tracked using the KLT technique in an image sequence of the gesture G_2 (Next)**

Each trajectory is composed by L= {p_1, p_2, ..., p_L}. Each point p_i has coordinates (x_i, y_i). Taking average of all points gives an average trajectory $G = [\overline{p_1}, \overline{p_2}, ..., \overline{p_L}]$. This average trajectory represents hand directions of a gesture. Fig. 10(b)

illustrates trajectories of 20 feature points and the average trajectory of the Next command in spatial-temporal coordinate. Red circles represent feature points coordinates p_i at the ith frame $i \in [1, L]$. Blue squares represent the average $\overline{p_i}$. These trajectories of training dataset will be saved to "KLT.yml". Using these parameters will be presented in detail in Sec. 4.4.

Phase synchronization:

Given two gestures $\{X_1^T, X_2^T, ..., X_{LT}^T\}$, $\{X_1^P, X_2^P, ..., X_{LP}^P\}$, LT and LP are their length respectively. Let where Z is the projection of corresponding image X in PCA space. The DTW algorithm starts by computing the local cost matrix $C \in R^{LTxLP}$C2 to align T and P. Each element c_{ij} of the matrix C is computed as Euclidean distance between Z_{Ti} and Z_{Pj}.

To determine the minimal cost of the optimal warping path p, we have to compute all possible warping paths between T and P. DTW employs the Dynamic Programming - to evaluate the following recurrence. Here, our DTW algorithm uses the distance function as defined in (3):

DTW(T,P) = min{c_P(T,P), p ∈ $p^{(L_{T\_}L_P)}$} (3)
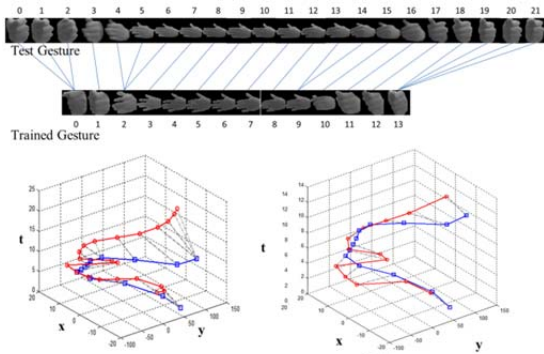

Figure 7. An illustration of the DTW results

K-NN for gesture recognition: To recognize a gesture, we utilize the conventional K-NN technique. Which the most important thing is to define the distance function and the value K. In our work, K is chosen by experiment.

Given two dynamic hand gestures T and P, we apply the step presented in Sec. 4.4.1 and obtain two average trajectories Tr_a^T, Tr_a^P with the same length L. Because end-users do not stand at the same position; or the height of end-users is not the same, the interaction regions of dynamic hand gestures are different in the image coordinate. Therefore, the coordinates of keypoints (x, y) on images of the two sequences could be different. To deal with this problem, we normalize Tr_a^{*T} as (4), (6) and Tr_b^{*T} as (5), (7).

$$Tr_a^T = [\overline{p_1^T}, \overline{p_2^T}, ..., \overline{p_L^T}] \tag{4}$$

$$Tr_a^P = [\overline{p_1^P}, \overline{p_2^P}, ..., \overline{p_L^P}] \tag{5}$$

$$Tr_a^{*T} = [\overline{p_1^T} - (\overline{x^T}, \overline{y^T}), \overline{p_2^T} - (\overline{x^T}, \overline{y^T}) ..., \overline{p_L^T} - (\overline{x^T}, \overline{y^T})] = [P_1^{*T}, P_2^{*T}, ..., P_L^{*T}] \tag{6}$$

$$Tr_a^{*P} = [\overline{p_1^P} - (\overline{x^P}, \overline{y^P}), \overline{p_2^P} - (\overline{x^P}, \overline{y^P}) ..., \overline{p_L^P} - (\overline{x^P}, \overline{y^P})] = [P_1^{*P}, P_2^{*P}, ..., P_L^{*P}] \tag{7}$$

Where $(\overline{x^T}, \overline{y^T})$ and $(\overline{x^P}, \overline{y^P})$ are the average values of all points in the sequence T and P respectively. The distance between Tr_a^T and Tr_a^P is determined by Root Mean Square Error (RMSE) in (8):

$$RMSE(Tr_a^T, Tr_a^P) = \sqrt{\frac{\sum_{k=1}^L (p_k^{*T} - Tr_a^P)^2}{L}} \tag{8}$$

The smaller RMSE value is, the more similar two gestures (T,P) are. Based on RMSE distance, a K-NN classifer is utilized to vote K nearest distances from

template gestures. A label is assigned to a testing gesture based on a maximal number of a label from K. The experimental results in Sec. 4 show that using RMSE is simple but obtains a high accuracy of recognition.

3. DEPLOYING A SMART LIGHTING CONTROL SYSTEM

Based on the designed gestures and the proposed recognition technique, we deploy a solution to control an indoor lighting system, as shown in Fig. 8. The system consists of four components: a Kinect sensor, a PC, a transceiver equipment and a lamp. To test the system, we used a halogen lamp manufactured by Philip company with power ranging from 0W to 200W corresponding to 0-100% brightness. We divided into six levels of brightness (0%, 20%, 40%, 60%, 80%,

100% which is illustrated in the Fig. 9.

We use five pre-defined hand gesture commands to control five levels of brightness corresponding to five states of the lamp. Then state translation scheme according to the incoming command is presented in Fig. 8. Following this scheme, **Next/Back** commands are used to increase or decrease one level of brightness while **Increase/Decrease** used to increase or decrease two levels of brightness. At every state, if the user performs a **Turn_on** command, the lamp will be turned on at the highest level of brightness 5th level. If the user performs a **Turn_off** command, the lamp will be **Turned_off** 0th level. Sec. 4 reports performances of the system tested in a lab-based environment real exhibition with assessments of various end-users.
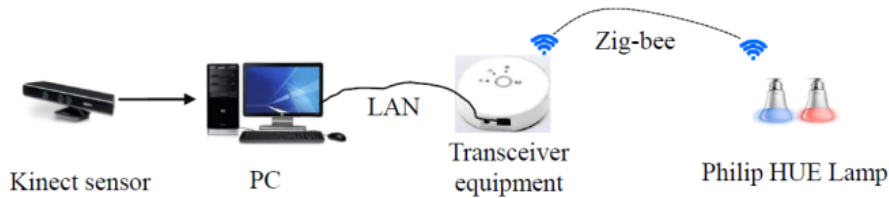
Figure 8. basic components in hand gesture-based lighting control system
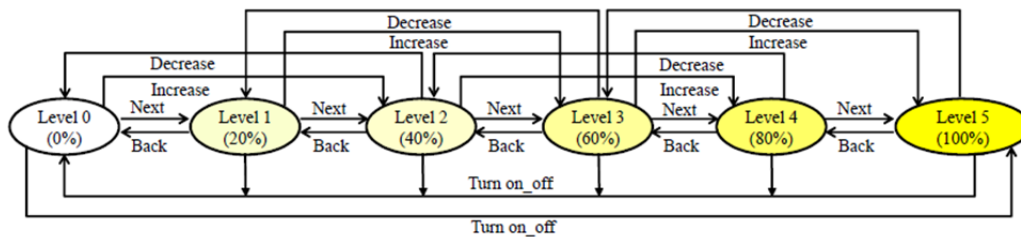

Figure 9. The state diagram of the proposed lighting control system

4. EXPRIMENTAL RESULTS

The proposed framework is warped by a C++ program on a PC Core i5 3.10 GHz CPU, 4 GB RAM. We evaluate the

proposed recognition scheme on four datasets. The first dataset, named MICA1, is acquired in a lab-based environment that is the showroom of our Institution.

The second dataset, named MICA2, is collected in a public exhibition, where is the much noisy environment. Detailed constructions of two datasets MICA1 and MICA2 are presented in detail in [25]. Two other published datasets MSRGesture3D[12] and Cambridge[13] also are utilized for comparing the performance of the proposed recognition technique. We conduct following evaluations: i) gesture spotting; ii) gesture discriminate; iii) gesture recognition and iv) real application of using hand gestures for the lighting control system.

4.1. Evaluation of inter-class and intra-class correlation of designed gestures

Intuitively, the designed gesture vocabulary is quite easy for users to memorize. In this section, we evaluate how they are discriminant for recognition. To aim this end, we take N samples from each gesture class. Then we compute the similarity of every pair of gestures and take the average over all samples. The similarity of two gestures is defined as RMSE computed from two feature vectors representing these gestures.

Table 2 shows the average RMSE computed from interclass and intraclass gestures. We see that the values of RMSE of intraclass gestures (on the diagonal of the matrix) belong to a small range [12.8, 21.3] while the values of RMSE of interclass gestures inside a bigger range [35.4, 49.7]. This means that the interclass gestures are well discriminant and the intraclass gestures are well similar together.

Table 1. RMSE of interclass and intraclass gestures

Gesture	G_1	G_2	G_3	G_4	G_5
G_1	**12.8**	36.5	42.3	36.7	33.4
G_2	35.4	**14.5**	44.2	37.2	38.3
G_3	37.3	41.6	**19.8**	41.4	49.2
G_4	37.2	45.4	49.7	**21.3**	48.0
G_5	39.3	36.8	45.2	41.7	**18.2**

4.2. Evaluation of hand gesture recognition

We evaluate the performance of our hand gesture recognition algorithm on three datasets: MICA1, MSRGesture3D and Cambridge. The dataset MSRGesture3D consists of twelve gestures performed by one hand or two hands. Our current method was designed to recognize one hand gestures. Therefore, we will evaluate our method on a subset of ten one hand gestures. The Cambridge dataset contains five dynamic hand gestures. For all datasets, we perform *Leave-p-out-cross-validation*, with p equals 5.

The recognition result on the dataset MICA1 is given in Tab. 3. In average, the recognition rate is 93.33 ± 6.94 % and the computational time for recognizing one gesture is 167 ± 15 ms. The confusion matrix shows that our algorithm is best with G_0 gesture (recognition accuracy of 100%), good at G_1 and G_5 (recognition accuracy of 97.22%). Some confuses the remaining gestures with Turn on_off gesture: $1G_2$ and $3G_4$. The reason is that in those cases, forearm of hand was not removed that leads to small movement of the hand region. Therefore, our algorithm considers them as G_1 gesture (the hand

shape changes but itself does not move). Moreover, some subjects implemented G_2 and G_4 with small deviation of hand direction that got confuse in Tab. 2.

Table 2. The gesture recognition result of the MICA2 dataset

Pre / Re	G_1	G_2	G_3	G_4	G_5	Regconition rate
G_1	**36**	0	0	0	0	100
G_2	1	**35**	0	0	0	97.2
G_3	0	0	**33**	0	3	91.7
G_4	3	4	0	**29**	0	80.6
G_5	0	0	1	0	**35**	97.2
Avr						**93.3±6.9**

The recognition rate on MSRGesture3D dataset is of $89.19\pm1.1\%$. The recognition rate on Cambridge dataset is of $91.47\pm6.1\%$. Comparing to state of the art methods, our method obtains competitive performance. Currently, our method obtains higher recognition rates on these datasets because we deploy K-NN with K = 9, this method is still good enough on our dataset with well discriminant designed gestures.

Table 3. Competitive performance of our method compared to existing methods

MSRGesture3D			Cambridge		
[13]	[22]	Our method	[12]	[15]	Our method
87.70	88.50	89.19	82.00	91.70	91.47

4.3. Evaluation performance and user-ability in a real show-case

We deploy the proposed method for lighting controls in a real environment of the exhibition. This environment is very complex: background is cluttered by many static/moving surrounding objects and visitors; lighting condition changes frequently. To evaluate the system performance, we follow also Leave-p-out-cross-validation method, with p equals 5. The recognition rate obtains $90.63\pm6.88\%$ that is shown in detail in the Tab. MICA2. Despite the fact that the environment is more complex and noisy than in the lab-case of dataset MICA1, we still obtain good results of recognition.

Pre / Gr	G_1	G_2	G_3	G_4	G_5	Reg rate
G_1	**94**	2	0	0	0	97.9
G_2	10	**83**	0	3	0	86.46
G_3	0	0	**81**	2	13	84.38
G_4	12	3	0	**81**	0	84.38
G_5	0	0	0	0	**96**	100
Avr						**90.6±6.9**

5. DISCUSSION AND CONCLUSION

Discussion: Although a real-case evaluation with a large number of end-users is implemented, as described in Sec. 4. There are existing/open questions which relate to the user's experience or expertise. To achieve a correct recognition system, it is very important that the user replicates the training gestures as close as possible. Moreover, the user's experience also reflect how is easily when an end-user implements the hand gestures. The fact that for a new end-user, without movement of the hand-forearm and only implementations of open-closed gestures of hand palm are quickly adapted. However, the gestures require both open-closed hand palms during hand-forearm's movements could raise difficulties for them.

Conclusion: This paper described a vision-based hand gesture recognition system. Our work was motivated by deploying a feasible technique into the real application that is the lighting control in a smart home. We designed a new set of dynamic hand gestures that map to common commands for a lighting control. The proposed gestures are easy for users to perform and memorize. Besides, they are convenient for detecting and spotting user's command from a video stream. Regarding the recognition issue, we attempted both spatial-temporal characteristics of a gesture. The experimental results confirmed that accuracy of recognition rate approximates 93.33% with the indoor environment as the MICA1 dataset with realtime cost only 176ms/gesture. Besides, 90.63% with the much noise environment as the MICA2 dataset. Therefore, it is feasible to implement the proposed system to control other home appliances.

REFERENCES

[1] http://www.microsoft.com/en-us/kinectforwindows. 2018.

[2] M.T. Ahammed and P. P. Banik, Home appliances control using mobile phone, in International Conference on Advances in Electrical Engineering, Dec 2015, pp. 251-254.

[3] F. Baig, S. Beg, and M. Fahad Khan, Controlling Home Appliances Remotely through Voice Command, International Journal of Computer Applications, vol. 48, no. 17, pp. 1-4, 2012.

[4] I. Bayer and T. Silbermann, A multi modal approach to gesture recognition from audio and video data, in Proceedings of the 15th ACM on ICMI, NY, USA, 2013, pp. 461-466.

[5] X. Chen and M. Koskela, Online rgb-d gesture recognition with extreme learning machines, in Proceedings of the 15th ACM on ICMI, NY, USA, 2013, pp. 467-474.

[6] H.G. Doan, H. Vu, T.H. Tran, and E. Castelli, Improvements of RGBD hand posture recognition using an user-guide scheme,in 2015 IEEE 7th International Conference on CIS and RAM, 2015, pp. 24-29.

[7] A. El-Sawah, C. Joslin, and N. Georganas, A dynamic gesture interface for virtual environments based on hidden markov models, in IREE International Workshops on Haptic Audio Visual Environments and their Applications, 2005, pp. 109-114.

[8] S.M.A. Haque, S.M. Kamruzzaman, and M.A. Islam, A system for smart home control of appliances based on timer and speech interaction, CoRR, vol. abs/1009.4992, pp. 128-131, 2010.

[9] C.A. Hussain, K.V. Lakshmi, K.G. Kumar, K.S.G. Reddy, F. Year, F. Year, and F. Year, Home Appliances Controlling Using Windows Phone 7, vol. 2, no. 2, pp. 817-826, 2013.

[10] N.J., B.A. Myers, M. Higgins, J. Hughes, T.K. Harris, R. Rosenfeld, and M. Pignol, Generating remote control interfaces for complex appliances, in Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology, 2002, pp. 161-170.

[11] R. Kango, P. Moore, and J. Pu, Networked smart home appliances enabling real ubiquitous culture, in Proceedings 3rd IEEE International Workshop on System-on-Chip for Real-Time Applications, 2002, pp. 76-80.

[12] T.K. Kim and R. Cipolla, Canonical correlation analysis of video volume tensors for action categorization and detection, IEEE TPAMI, vol. 31, no. 10, pp. 1415-1428, 2009.

[13] A. Kurakin, Z. Zhang, and Z. Liu, A real time system for dynamic hand gesture recognition with a depth, in 20th EUSIPCO, August 2012, pp. 27-31.

[14] B.D. Lucas and T. Kanade, An iterative image registration technique with an application to stereo vision, in Proceedings of the 7th International Joint Conference on Arti_cial Intelligence Volume 2, San Francisco, CA, USA, 1981, pp. 674-679.

[15] Y.M. Lui, Human Gesture Recognition on Product Manifolds, Journal of Machine Learning Research 13, vol. 13, pp. 3297-3321, 2012.

[16] R. Murali, J.R.R, and M.R.R.R, Controlling Home Appliances Using Cell Phone, International Journal of Cientific and Technology Research, vol. 2, pp. 138-139, 2013.

[17] J. Nichols and B. Myers, Controlling Home and Office Appliances with Smart Phones, IEEE Pervasive Computing, vol. 5, no. 3, pp. 60-67, 2006.

[18] Rautaray, S.S., and A. Agrawal, Vision based hand gesture recognition for human computer interaction: A survey, Artif. Intell. Rev., vol. 43, no. 1, pp. 1-54, Jan. 2015.

[19] S.Escalera, J.Gonzàlez, X.Baró, M.Reyes, and L.A.C.R.S.S.I.Guyon V. Athitsos, H. Escalante, Chalearn multi-modal gesture recognition 2013: Grand challenge and workshop summary, in Proceedings of the 15th ACM on ICMI, USA, 2013, pp. 365-368.

[20] J. Shi and C. Tomasi, Good features to track, in IEEE Conference on Computer Vision and Pattern Recognition - CVPR'94, Ithaca, USA, 1994, pp. 593-600.

[21] C. Staufier and W. Grimson, Adaptive background mixture models for real-time tracking, in Proceedings of Computer Vision and Pattern Recognition. IEEE Computer Society, 1999, pp. 2246 -2252.

[22] J. Wang, Z. Liu, J. Chorowski, Z. Chen, and Y. Wu, Robust 3d action recognition with random occupancy patterns, in Proceedings of the 12th European Conference on Computer Vision - Volume Part II - ECCV'12, 2012, pp. 872-885.

[23] B. Yuksekkaya, A. Kayalar, M. Tosun, M. Ozcan, and A. Alkar, A GSM, internet and speech controlled wireless interactive home automation system, IEEE Transactions on Consumer Electronics, vol. 52, no. 3, pp. 837-843, 2006.

[24] Huong-Giang Doan, Hai Vu, and Thanh-Hai Tran. Recognition of hand gestures from cyclic hand movements using spatial-temporal features, in the proceeding of SoICT 2015, Vietnam, pp. 260-267.

[25] Huong-Giang Doan, Hai Vu, and Thanh-Hai Tran. Phase Synchronization in a Manifold Space for Recognizing Dynamic Hand Gestures from Periodic Image Sequence, in the proceeding of the 12th IEEE-RIVF 2016, pp. 163 - 168, Vietnam.

Biography:

Huong Giang Doan, received B.E. degree in Instrumentation and Industrial Informatics in 2003, M.E. in Instrumentation and Automatic Control System in 2006 and Ph.D. in Control engineering and Automation in 2017, all from Hanoi University of Science and Technology, Vietnam. She is a lecturer at Control and Automation Faculty, Electric Power University, Ha Noi, Viet Nam.

His research interest includes human-machine interaction using image information, action recognition, manifold space representation for human action, computer vision.

Duy Thuan Vu, received B.E. degree in Instrumentation and Industrial Informatics in 2004, M.E. in Instrumentation and Automatic Control System in 2008 from Hanoi University of Science and Technology, Vietnam. Ph.D. in Control engineering and Automation in 2017 in Vietnam Academy of Science and Technology. He is Dean of Control and Automation Faculty, Electric Power University, Ha Noi, Viet Nam.

His research interest includes human-machine interaction, robotic, optimal algorithm in control and automation.