

# Hand number gesture recognition using recognized hand parts in depth images

Dong-Luong Dinh · Sungyoung Lee · Tae-Seong Kim

Received: 8 April 2014 / Revised: 25 September 2014 / Accepted: 7 November 2014 /  
Published online: 2 December 2014  
© Springer Science+Business Media New York 2014

**Abstract** In this paper, we present a novel approach of recognizing hand number gestures using the recognized hand parts in a depth image. Our proposed approach is divided into two stages: (i) hand parts recognition by random forests (RFs) and (ii) rule-based hand number gestures recognition. In the first stage, we create a database (DB) of synthetic hand depth silhouettes and their corresponding hand parts-labeled maps and then train RFs with the DB. Via the trained RFs, we recognize or label the hand parts in a depth silhouette. In the second stage, based on the information of the recognized or labeled hand parts, hand number gestures are recognized according to our derived rules. In our experiments, we quantitatively and qualitatively evaluated our hand parts recognition system with synthetic and real data. Then, we tested our hand number gesture recognition system with real data. Our results show the average recognition rate of 97.80 % over the ten hand number gestures from five different subjects.

**Keywords** Hand parts recognition · Hand number gesture recognition · Hand depth silhouette · Random forests

## 1 Introduction

Among various gesture recognition methodologies using bodies, faces, or hands, a hand-based gesture or sign recognition methodology is one of efficient and natural ways for

---

D.-L. Dinh · S. Lee (✉)  
Department of Computer Engineering, Kyung Hee University,  
1 Seocheon-dong, Giheung-gu, Yongin-si, Gyeonggi-do, Republic of Korea  
e-mail: sylee@oslab.khu.ac.kr

D.-L. Dinh  
e-mail: luongdd@oslab.khu.ac.kr

T.-S. Kim (✉)  
Department of Biomedical Engineering, Kyung Hee University,  
1 Seocheon-dong, Giheung-gu, Yongin-si, Gyeonggi-do, Republic of Korea  
e-mail: tskim@khu.ac.kr

human machine interfaces. Potential applications of the hand gesture or sign recognition include human computer interfaces, automatic sign language interpretations [8], remote controls of electronic appliances [20, 26], and smart robotics [14]. One of key functions in these applications is to recognize hand poses, movements, or both which convey meaningful information of users' intention to give instructions to machines. Among hand gesture recognition of signs such as letters, numbers, or symbols, hand number gesture recognition is an important task to interface between humans and machines. Although various approaches have been proposed for hand number gesture recognition, they can be divided into two commonly used approaches: one is inertial sensor glove-based and the other vision-based approaches [14].

The sensor glove-based approach uses an array of sensors which are attached to a hand glove that transduces fingers and hand actions into electrical signals to recognize the hand number gestures. In [5], a sensor glove was used to process and convey the degree of flexion in each finger of the hand and a tri-axis accelerometer was placed on the back of the hand to provide its orientation information. A neural network method was employed to recognize hand number gestures. In [4, 24], six accelerometers on a hand glove were used to get relative angles between fingers and the palm of the hand. Each number gesture was recognized by a fuzzy rule-based classification system. Another method proposed in [15] introduced a 3D input device with inertial sensors to recognize number drawing gestures in 3D space. The signals of angular velocity and acceleration were considered as motion features. Fisher discriminant analysis was used for hand number gesture recognition. Although, these studies reported some success in hand number gesture recognition, the methods require of wearing gloves and positioning sensor on the glove or hand. Therefore, motion sensor- or glove-based recognition approaches of hand gestures, in general, are not considered as a natural and convenient way in spite of high sensitivity [6].

The vision-based approach uses imaging sensors or cameras to get hand gesture features in colors, shapes, orientations, texture, contours, moments or skeleton [9, 14]. In [2], the authors proposed a method to recognize hand number gestures of Persian sign language from 1 to 10 using the skeleton information of hand silhouettes. The endpoints of the skeleton were extracted as fingertips in recognizing ten gestures. Similar to [25], the authors developed a hand gesture recognition system using the hand orientation and contour features and Hidden Markov Models. Another method proposed in [11, 23] used a color glove in which different colors were imprinted on five fingers of the hand glove. Then finger recognition was performed based on color information to classify hand number gestures. However, these studies exhibited limited success in hand number gesture recognition using color images since the use of color image and its features is generally sensitive to noise, lighting conditions, cluttered backgrounds, and especially occlusions.

Recently, with the introduction of new depth imaging sensors, some studies have been focusing on using depth features for hand gesture recognition [22]. In [13], the method was presented for recognizing hand number gestures by template matching using Chamfer Distance, measuring shape similarity based on the orientation, size, shape, and position features from depth images. In [7], the hand gesture recognition combined two types of hand pose features, namely, distance and curvature features by computing the hand contours. Finally, a support vector machine classifier was employed to recognize hand sign gestures. In these approaches, the same or similar features of color images are extracted from depth images, thereby resulting in marginally improved recognition rate. In another proposed method [19], the finger were detected by hand shape decomposition and hand number gesture recognition was done via template matching by minimizing dissimilarity distance using their proposed Finger-Earth Mover's distance metric. However, this approach suffered from the ambiguity

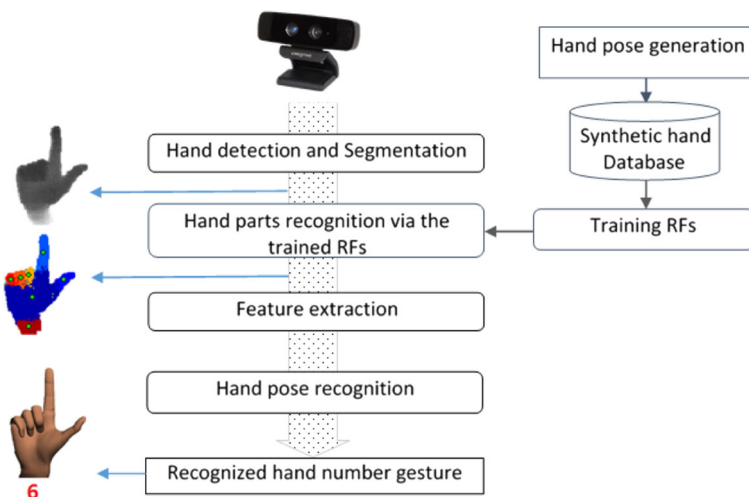
in the height of the thumb finger which is often shorter than other fingers in hand poses when using a height threshold for finger detection, resulting in reduced recognition rate. In general, these previous depth image-based approaches suffer from no information of finger parts, in spite of advantages of depth images such as low sensitivity to lighting conditions.

In this study, we have developed a novel approach of recognizing hand number gestures by recognizing or labeling hand parts in depth images. Our proposed approach consists of two main processes: hand parts recognition by random forests (RFs) classifier and rule-based hand number gestures recognition. The main advantage of our proposed approach is that the state of each finger is directly identified through the recognize hand parts and then number gestures are recognized based on the state of each finger. We have tested and validated our approach on synthetic and real data. Our experimental tests achieved 97.8 % in recognition of ten hand number gestures with five different subjects. Our hand number gesture recognition system should be useful in human and machine interaction applications.

The rest of the paper is organized as follows. In Section 2, we describe the processes of our proposed methodology including hand parts and hand number gesture recognition. In hand parts recognition, we present techniques of synthetic hand database (DB) creation, training and classification per pixel via RFs. In hand number gesture recognition, we describe steps of hand number gesture recognition. Section 3 presents parameter optimizations and obtained results. Section 4 shows our conclusion and discussion.

## 2 The proposed methodology

Overall process of our proposed system for hand number gesture recognition, shown in Fig. 1, consists of two main parts: in the first part, a synthetic hand DB, which contains pairs of depth maps and their corresponding hand parts-labeled maps, was generated. Then, the DB was used to training RFs. In the recognition stage, a depth image was first captured from a depth camera and then a hand depth silhouette was extracted by removing the background. Next, the hand parts of a depth silhouette were recognized using the



**Fig. 1** The framework of our hand number gesture recognition system

trained RFs. Next, a set of features was extracted from the labeled hand parts. Finally, based on the extracted features, hand number gestures were recognized by our rule-based approach.

The ten hand number gestures representing 10 digits from 0 to 9 are shown in Fig. 2. Our final goal is to recognize these gestures with the information of each finger parts recognized by our proposed methodology.

## 2.1 Hand parts recognition

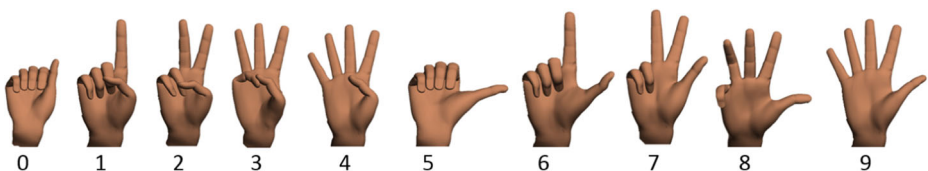
In this part, we describe a methodology of hand parts recognition based on pixel classification of a hand depth silhouette via RFs. The details of hand parts recognition are presented in the following subsections.

### 2.1.1 Synthetic hand DB generation

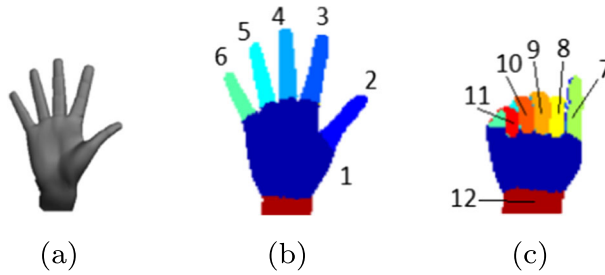
To recognize hand parts from a hand depth silhouette via RFs, the synthetic hand DB, which contains pairs of depth images and their corresponding hand part-labeled maps, is needed to train RFs. We created such the DB with a synthetic hand model using 3Ds-max, a commercial 3D graphic package [1] as shown in Fig. 3a. To identify hand parts, twelve labels were assigned to each hand model as shown in Fig. 3b, c. The five fingers including the thumb, index, middle, ring, and pinkie fingers, were represented by ten corresponding labels including the five front and five back sides of the fingers. The front parts were coded with the index numbers of 2, 3, 4, 5 and 6. Likewise, the five back sides were coded with the index numbers of 7, 8, 9, 10 and 11, respectively. This is critical in our recognition system to identify the open and close state of each finger. In addition, the palm and wrist parts were given the index number of 1 and 12. Our DB comprised 5,000 pairs covering the ten different hand number gestures as shown in Fig. 2. Among them, a set of 3,000 pairs was used in training and a set of the remaining pairs was used in testing. A set of 500 pairs in the DB represents each hand number gesture captured under many different view angles and their orientation in ranging  $[-45...+45]$  degrees compared to vertical original orientation. A set of samples of 3-D hand model and the map of the corresponding labeled hand parts are shown in Fig. 3. The images in the DB had a size of  $240 \times 320$  with 16-bit depth values.

### 2.1.2 Depth feature extraction

To extract depth features  $f$  from pixel  $p$  of depth silhouette  $I$  as described in [12, 21], we computed a set of depth features of the pixel  $p$  based on difference between the depth values of a neighborhood pixel pair in depth silhouette  $I$ . The positions of the pixel pairs were randomly selected on the depth silhouette and they had a relation with the position of



**Fig. 2** Our target hand number gestures, representing numbers from 0 to 9



**Fig. 3** Hand model: **a** 3-D hand model in 3Ds-max and **b, c** Hand parts with twelve labels

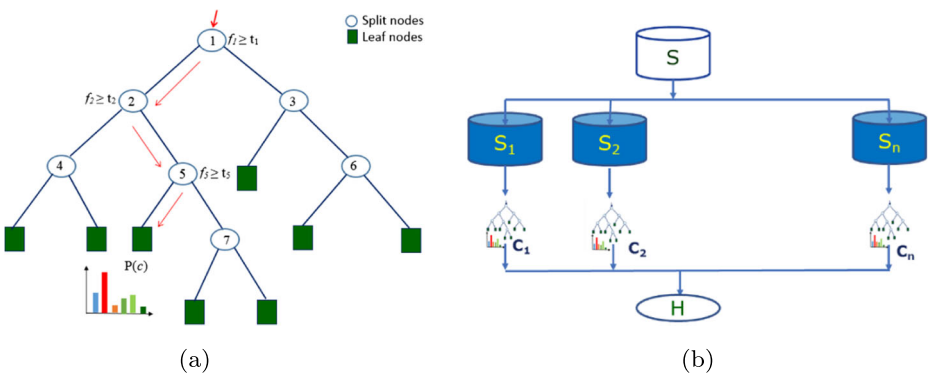
the considered pixel  $p$  by two terms  $o_1$  and  $o_2$  of the coordinates of  $x$  and  $y$ , respectively. Depth features,  $f$ , are computed as follows:

$$f(I, p) = I_d \left( x_p + \frac{o_{1x}}{I_d(x_p, y_p)}, y_p + \frac{o_{1y}}{I_d(x_p, y_p)} \right) - I_d \left( x_p + \frac{o_{2x}}{I_d(x_p, y_p)}, y_p + \frac{o_{2y}}{I_d(x_p, y_p)} \right) \tag{1}$$

where  $I_d(x_p, y_p)$  is the depth value at the coordinates of  $(x_p, y_p)$  and  $(o_1, o_2)$  is an offset pair. The maximum offset value of  $(o_1, o_2)$  pairs was 30 pixels that corresponds to 0.6 meters which was the distance of the subject to the camera. The normalization of the offset by  $\frac{1}{d_I(x_p, y_p)}$  ensured that the features are distance invariant.

2.1.3 RFs for pixel classification

In our work, we utilized RFs for hand parts recognition. RFs are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same sample distribution for all trees in the forest [3, 10, 12, 21]. These concepts are illustrated in Fig. 4. Figure 4a presents a single decision tree learning process as a tree predictor. The use of a multitude of decision trees for training and testing RFs on the same DB of S is described in Fig. 4b. The sample sets  $\{S_i\}_{i=1}^n$  are drawn randomly from the training data S by bootstrap algorithm [10].



**Fig. 4** RFs for pixel-based classification: **a** a binary decision tree and **b** ensemble of  $n$  decision trees for training RFs

In training, we used an ensemble of 21 decision trees. The maximum depth of trees was 20. Each tree in RFs was trained with different pixels sampled randomly from the DB. A subset of 500 training sample pixels was drawn randomly from each synthetic depth silhouette. A sample pixel was extracted as in (1), to obtain 800 candidate features. At each splitting node in the tree, a subset of 28 candidate features was considered. For pixel classification, each pixel  $p$  of a tested depth silhouette was extracted to obtain the candidate features. For each tree, starting from the root node, if the value of the splitting function was less than a threshold of the node,  $p$  went to left and otherwise  $p$  went to right. The optimal threshold for splitting the node was determined by maximizing the information gain in the training process. The probability distribution over 12 hand parts was computed at the leaf nodes in each tree. Final decision to label each depth pixel for a specific hand part was based on the voting result of all trees in the RFs.

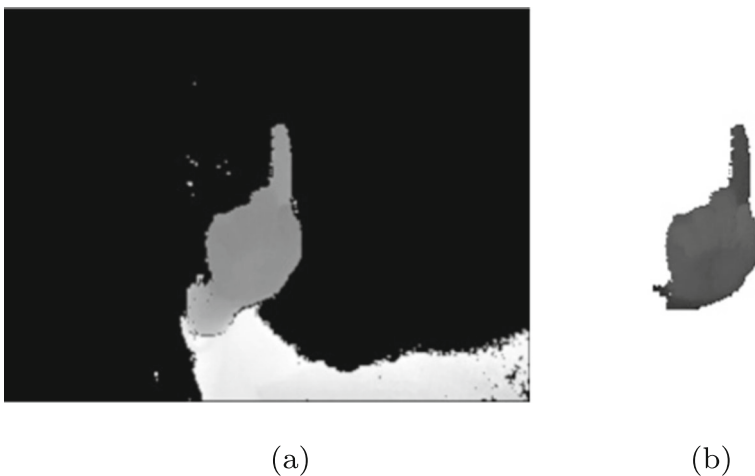
## 2.2 Hand number gesture recognition

In this part, we describe how to recognize hand number gestures from an incoming hand depth silhouette. Based on the recognized hand parts of the hand depth silhouette, a set of features were extracted for recognizing hand number gestures.

### 2.2.1 Hand detection

As shown in Fig. 5, we used the creative interactive gesture camera ([www.intel.com/software/perceptua](http://www.intel.com/software/perceptua)). This device is capable of close-range depth data acquisition. The depth imaging parameters were set to be an image size of  $240 \times 320$ , and frame rate of 30 fps. The distance from hand parts to a camera was in a range of 0.15 m and 1.0 m. The hand parts could be captured in the field of view of 70 degrees.

To detect the hand area and remove background, we used an adaptive depth threshold. The value of threshold was determined based on a specific distance from a depth camera



**Fig. 5** Hand detection and segmentation: **a** the captured hand depth image and **b** its corresponding hand depth silhouette

to the hand parts. To extract hand depth silhouettes, we used the library of Intel Perceptual Computing SDK 2013 which supports the Creative Labs Senz3D Depth and Gesture Camera ([www.intel.com/software/perceptua](http://www.intel.com/software/perceptua)). The detected and segmented hand is shown in Fig. 5. Based on the detected hand depth silhouette, hand orientation is determined by computing its eigenvector to rotate it about vertical original orientation before each hand depth silhouette used for recognizing hand parts.

### 2.2.2 Hand parts recognition

To recognize hand parts of each hand depth silhouette, all pixels of each hand depth silhouette were classified by the trained RFs to assign a corresponding label out of the 12 indices. A centroid point was withdrawn from each recognized hand part and represents each hand part as illustrated in Fig. 6. Figure 6a shows a hand depth silhouette and Fig. 6b presents the result of the labeled hand parts with 1, 2, 3, 9, 10, 11, and 12 indices. Based on Table 1, we know the recognized hand parts including the palm, thumb (front), index (front), middle (back), ring (back), pinkie (back), and wrist parts and their centroid points which are presented with the big dots in green.

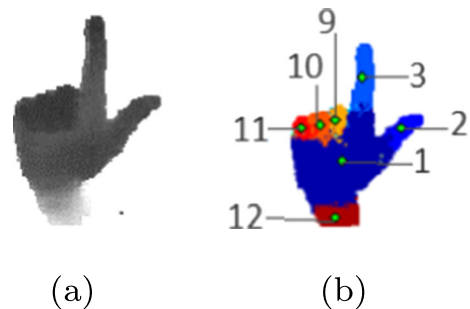
### 2.2.3 Feature extraction and derived rules for hand number gesture recognition

From the recognized hand parts, we extracted a set of features. In our labeling, each finger was represented by two different labels: one label for its front side corresponding to the open state of the finger and another for its back side corresponding to the close state of the finger. From the information of the recognized hand parts, we identify the open or close states of each finger. The states of the five labeled fingers were identified and saved as features, namely  $f_{Thumb}$ ,  $f_{Index}$ ,  $f_{Middle}$ ,  $f_{Ring}$ , and  $f_{Pinkie}$ , respectively.

$$f_{Finger}(i) = \begin{cases} 1 : Open & \text{for Label} = 2, 3, 4, 5, \text{ or } 6 \\ 0 : Close & \text{for Label} = 7, 8, 9, 10, \text{ or } 11 \end{cases} \quad (2)$$

For example, as shown in Fig. 6b,  $f_{Thumb}$  and  $f_{Index}$  become 1 corresponding to the open state of the fingers. In contrast,  $f_{Middle}$ ,  $f_{Ring}$ , and  $f_{Pinkie}$  become 0 corresponding to the close state of the fingers (2). To recognize hand number gestures, we derived a set of recognition rules. The set of five features from the states of all fingers was used to encode the meaning of the ten hand number poses. The derived recognition rules are given in Table 2. One problem in the finger state identification is having both the front and back labels after recognizing the finger parts depending on the configuration of each finger. However in the recognition of the binary state of each finger, we used a maximum likelihood estimation

**Fig. 6** An example of labeled hand parts recognition: **a** a hand depth silhouette and **b** the labeled hand parts



**Table 1** A list of the named and labeled parts in hand model

Label no.	Hand parts	Label no.	Hand parts
1	Palm	7	Thumb (back)
2	Thumb (front)	8	Index (back)
3	Index (front)	9	Middle (back)
4	Middle (front)	10	Ring (back)
5	Ring (front)	11	Pinkie (back)
6	Pinkie (front)	12	Wrist

criterion to select one of the two labels based on estimating the mean probability distribution of all labelled pixels in each set.

### 3 Experimental results

In this part, we have first optimized some main parameters on our own hand synthetic DB and then evaluated hand parts recognition through quantitative and qualitative assessments using synthetic and real data. Final, we have tested to evaluate our proposed hand number gesture recognition methodology on real data.

#### 3.1 Parameter optimization

We performed the experiments on our own DB to optimize the major parameters including the number of trees used in RFs, the number of extracted features, and the maximum value of  $(o_1, o_2)$  offset pairs chosen for feature extraction.

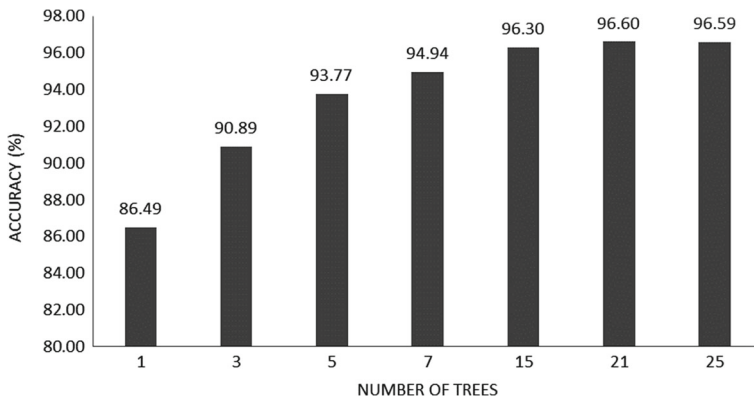
Figure 7 shows the average recognition rates of pixels corresponding to the number of trees used in RFs. The recognition accuracy of pixels increases slowly when the number of trees increases from 1 to around 15 trees. The recognition accuracy is maximum at 21 trees used in RFs and decreases for around 25 trees used in RFs.

**Table 2** Recognition rules of the number gestures based on the states of the five fingers

$f_{Thumb}$	$f_{Index}$	$f_{Middle}$	$f_{Ring}$	$f_{Pinkie}$	Represented numbers
0	0	0	0	0	0
0	1	0	0	0	1
0	1	1	0	0	2
0	1	1	1	0	3
0	1	1	1	1	4
1	0	0	0	0	5
1	1	0	0	0	6
1	1	1	0	0	7
1	1	1	1	0	8
1	1	1	1	1	9

0=Close, 1=Open





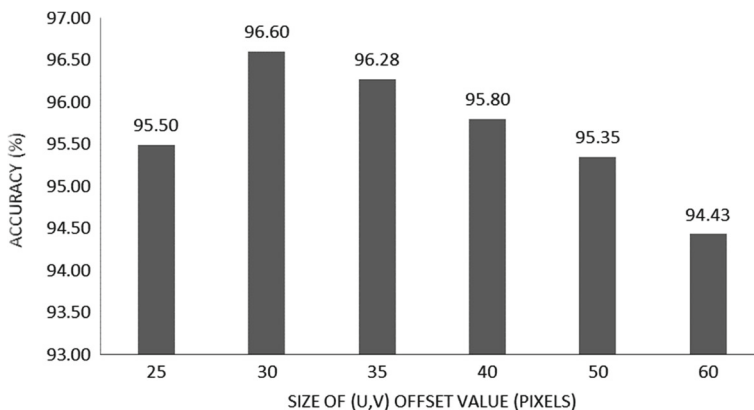
**Fig. 7** The effects of trees in RFs on recognition accuracy

As can be seen in Fig. 8, the average recognition rates of pixels correspond to the selected maximum value changes of  $(o_1, o_2)$  offset pairs. In this experiment, the distance of the subject to the camera is set 0.6 meters. The recognition accuracy is the highest when the maximum value of  $(o_1, o_2)$  offset pairs was chosen 30. This means that ranging of  $(o_1, o_2)$  offset pairs is from 0 to 30. It decreases slowly when the  $(o_1, o_2)$  pairs get the maximum values is around 25, 35, 40, 50, and 60.

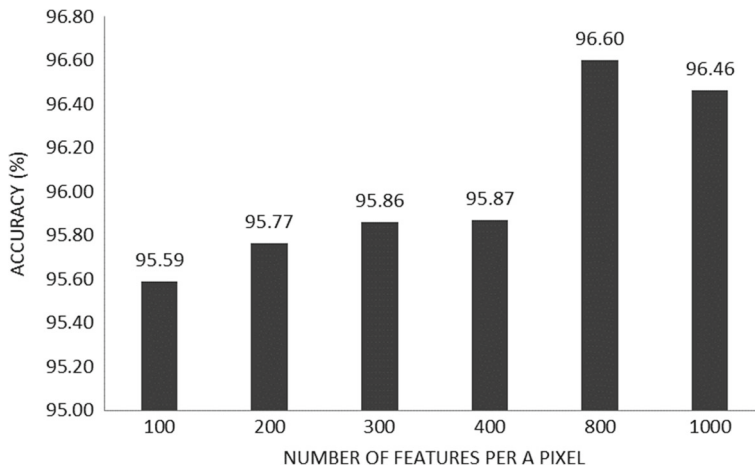
As can be seen in Fig. 9, the feature numbers extracted for each pixels in RFs obtained maximum recognition accuracy at around 800.

### 3.2 Results of hand parts recognition

To evaluate our hand parts recognition via the trained RFs on own synthetic database with optimized parameters qualitative and quantitatively, we tested on a series of 2000 hand depth silhouettes containing various number poses over the ten hand number poses. The qualitative assessment of some hand pose samples on both synthetic data are presented in Fig. 10. As can be seen in Fig. 10, the parts in hand are labeled by different color codes and their centroid points are shown as big dots in green superimposed on the labeled hand parts.



**Fig. 8** The effects of the  $(o_1, o_2)$  offset value pair on recognition accuracy



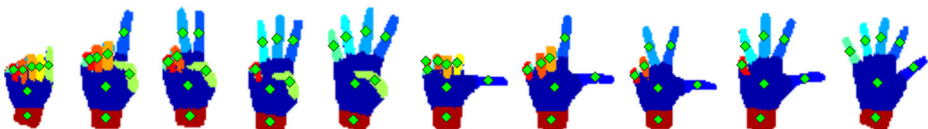
**Fig. 9** The effects of pixel features on recognition accuracy

The quantitative assessment results on synthetic data are provided by a confusion matrix with 12 recognized hand parts as described in Table 3. As revealed by Table 3, the average recognition rate of the palm part is the highest whereas this recognition rate of the pinkie part is the lowest. For finger parts, the recognition results are higher on their front sides than back sides. The average recognition rate of all hand parts was 96.60 %.

### 3.3 Results of hand number gesture recognition

To evaluate our proposed hand number gesture recognition system with real data, we collected a hand pose dataset using the Creative Labs Senz3D Depth and Gesture Camera. A set of 500 samples of the hand poses was captured from five different subjects. Each subject was asked to make 100 hand poses of 10 different number gestures in front of the depth camera. The detected centroid points in the labeled hand regions as illustrated in Fig. 11. Since the ground truth labels are not available with real data, only qualitative visual inspection is possible. The top and bottom rows of Fig. 11 present the hand depth silhouettes and their pixel-wise labeled hand parts results, respectively. The recognized hand parts are presented by different color codes and their detected centroid points are shown as the green dots indicate the centroids. Figure 12 shows the sample results at each step of four hand number gestures 0, 3, 6, and 9 in our hand number gesture recognition system under real scenario.

Based on a set of our derived rules as provided in Table 2 and the state of detected centroid points in the pixel-wise labeled hand regions, the hand number gestures are recognized. Table 4 presents the confusion matrix of the hand number gesture recognition and its average recognition rate was achieved 97.80.



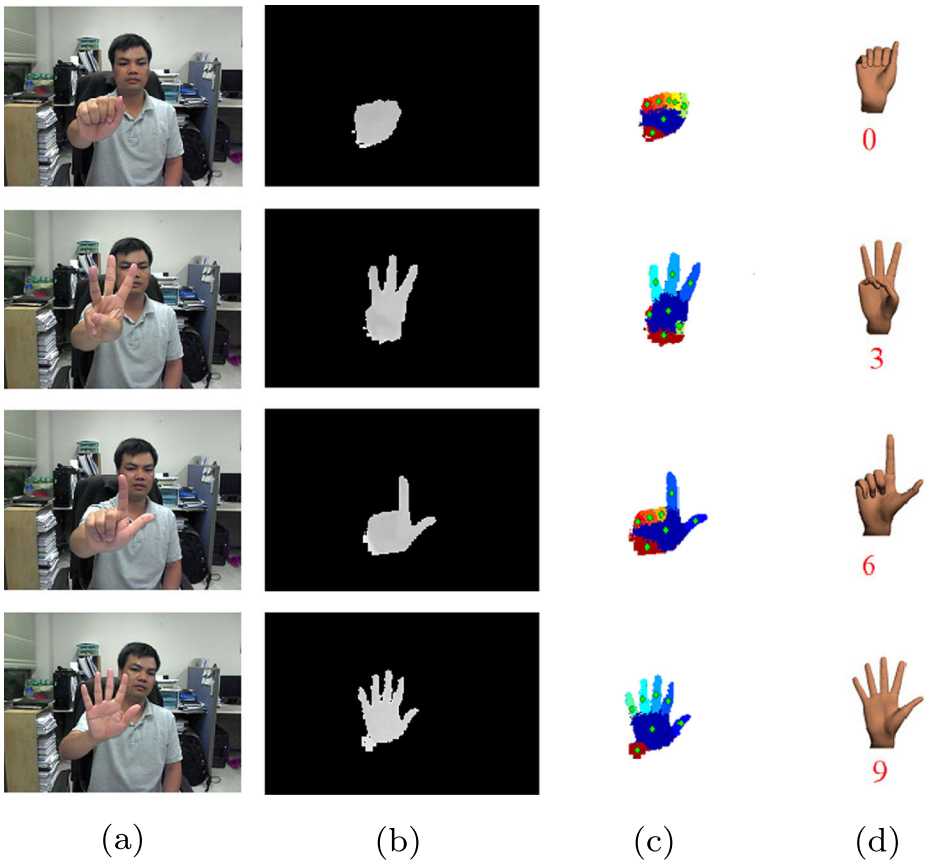
**Fig. 10** Sample results of labeled hand parts via RFs on synthetic data

**Table 3** The accuracy of pixel-based recognition in hand parts via RFs on synthetic data (%)

Hand parts	1	2	3	4	5	6	7	8	9	10	11	12
1	99.24	0.20	0.12	0.04	0.06	0.14	0.01	0.00	0.01	0.03	0.02	0.14
2	2.40	96.15	0.16	0.01	0.00	0.00	1.27	0.01	0.00	0.00	0.00	0.00
3	0.35	0.02	98.72	0.35	0.00	0.00	0.04	0.53	0.00	0.00	0.00	0.00
4	0.25	0.00	0.31	98.25	0.30	0.01	0.10	0.01	0.76	0.01	0.00	0.00
5	0.42	0.00	0.00	0.70	97.52	0.47	0.06	0.00	0.03	0.79	0.02	0.00
6	0.62	0.00	0.00	0.02	0.95	97.41	0.01	0.00	0.00	0.01	0.97	0.00
7	0.23	2.56	0.48	0.25	0.36	0.11	95.45	0.08	0.18	0.16	0.15	0.00
8	0.24	0.07	2.76	0.09	0.00	0.00	0.11	96.67	0.05	0.00	0.00	0.00
9	0.30	0.02	0.03	3.16	0.09	0.01	0.20	0.22	95.72	0.22	0.02	0.00
10	0.97	0.00	0.00	0.21	3.57	0.13	0.23	0.00	0.45	94.10	0.33	0.00
11	1.33	0.01	0.01	0.06	0.30	6.26	0.83	0.00	0.07	0.72	90.43	0.00
12	0.52	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	99.47

*Comparison with another method* In this part, we have evaluated our proposed hand number gesture recognition methodology against the edge moments-based methodology. This approach has actively utilized for static hand gesture recognition [16–18]. In this methodology, first, hand detection and segmentation are applied on color images. From the segmented hand silhouettes, the hand boundaries are extracted by the canny edge detection algorithm. Next, the Hu's moments are computed on the hand boundaries, yielding the moment's features. Finally, these features are used as an input to the SVM classifier. In our experiments, we utilized color and depth images captured at the same time. A set of 1000 color images of the hand number gestures for training and a set of 500 color images of the hand number gestures from five different subjects as collected in our proposed hand number gesture recognition experiment with real data was used. Table 5 shows the results of hand number gesture recognition. As can be shown in Table 5, many gestures are severely confused in the moment-based methodology such as between Gestures 5 and 6, Gestures 7 and 8, and Gestures 8 and 9. This might be due to the fact that their gesture shapes are very similar and the moment features could not differentiate them. The average hand number recognition rate of the color image-based approach archived 88.20 % whereas the proposed 97.80 %.

**Fig. 11** Sample results of labeled hand parts via RFs on real data



**Fig. 12** Some sample results of our hand number gesture recognition system under real scenario: **a** Input scenario, **b** Hand depth silhouette, **c** Labeled hand parts, and **d** Recognized hand number gesture

**Table 4** The confusion matrix of hand number gesture recognition using our proposed system (%)

Numbers	0	1	2	3	4	5	6	7	8	9
0	<b>100</b>									
1	2	<b>98</b>								
2			<b>100</b>							
3				<b>90</b>	10					
4					<b>100</b>					
5				2		<b>98</b>				
6							<b>98</b>	2		
7								<b>98</b>	2	
8									<b>100</b>	
9					2			2		<b>96</b>

**Table 5** The confusion matrix of hand number gesture recognition based the edge moments-based approach (%)

Numbers	0	1	2	3	4	5	6	7	8	9
0	<b>92</b>					8				
1		<b>98</b>				2				
2			<b>100</b>							
3			4	<b>88</b>	8					
4				6	<b>92</b>	2				
5	2					<b>78</b>	20			
6	2					2	<b>92</b>	4		
7				4	2		2	<b>68</b>	18	6
8								2	<b>76</b>	22
9									2	<b>98</b>

#### 4 Conclusion and discussion

In this paper, we have fully presented a novel hand number gesture recognition system using the labeled hand parts via the trained RFs from a hand depth silhouette. We created our own synthetic hand DB and then train RFs with the DB. The recognized hand parts via the trained RFs were implemented on the optimized parameters of our hand DB. As a result, we have achieved the mean recognition rate of 97.80 % over the ten hand number gestures from five different subjects.

In comparison to the conventional color image-based approaches in [2, 11, 23, 25], the main contributions of our proposed approach is that our proposed method identified each finger by recognizing the hand parts and also understood the state of each finger in hand by recognizing its front and back side to identify the open or close states of each finger. Then finally the hand number gestures are recognized according to our derived recognition rules of the states.

Our presented work should be applicable for recognizing sign language gestures as well as humans and machines interaction applications such as household appliances control, health-care supporting systems which help to improve quality of life for everybody, especially for people with learning disabilities, disabled and elderly people.

**Acknowledgments** This research was supported by the MSIP (Ministry of Science, ICT & Future Planning), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency (NIPA-2013-(H0301-13-2001))). This work was also supported by the Industrial Strategic Technology Development Program (10035348, Development of a Cognitive Planning and Learning Model for Mobile Platforms) funded by the Ministry of Knowledge Economy(MKE, Korea) and the Industrial Core Technology Development Program (10049079, Development of Mining core technology exploiting personal big data) funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea)”

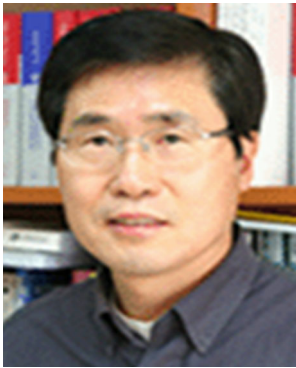
#### References

1. Autodesk 3Ds MAX (2012)
2. Barkoky A, Charkari NM (2011) Static hand gesture recognition of Persian sign numbers using thinning method. In: Proceedings of international conference on multimedia technology (ICMT), pp 6548–6551

3. Breiman L (2001) Random forests. *ML J* 45(1):5–32
4. Bui TD, Nguyen LT (2007) Recognizing postures in Vietnamese sign language with MEMS accelerometers. *IEEE Sens J* 7(5):707–712
5. Cabrera ME, Bogado JM, Fermin L, Acuna R, Ralev D (2012) Glove-based gesture recognition system. In: *Proceedings of international conference on climbing and walking robots and the support technologies for mobile machines*, pp 747–753
6. Dipietro L, Sabatini AM, Dario P (2008) A survey of glove-based systems and their applications. *IEEE Trans Syst Man Cybern-Part C Appl Rev* 38(4):461–482
7. Dominio F, Donadeo M, Marin G, Zanuttigh P, Cortelazzo GM (2013) Hand gesture recognition with depth data. In: *Proceedings of ACM/IEEE international workshop on analysis and retrieval of tracked events and motion in imagery stream*, pp 9–16
8. Elmezain M, Al-Hamadi A, Pathan SS, Michaelis B (2009) Spatio-temporal feature extraction-based hand gesture recognition for isolated American Sign Language and Arabic numbers. In: *Proceedings of international symposium on image and signal processing and analysis*, pp 254–259
9. Hasan H, Abdul-Kareem S (2013) Human-computer interaction using vision-based hand gesture recognition system: a survey. *Neural Comput Appl* 25(2):251–261
10. Hastie T, Tibshirani R, Friedman J (2008) *The elements of statistical learning*, 2nd edn. Springer, New York
11. Lamberti L, Camastra F (2012) Handy: a real-time three color glove-based gesture recognizer with learning vector quantization. *Expert Syst Appl* 39(12):10489–10494
12. Lepetit V, Lagger P, Fua P (2005) Randomized trees for real-time keypoint recognition. *IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2:775–781
13. Liu X, Fujimura K (2004) Hand gesture recognition using depth data. In: *Proceedings of IEEE international conference on automatic face and gesture recognition*, pp 529–534
14. Murthy GRS, Jadon RS (2009) A review of vision based hand gestures recognition. *Int J Inf Technol Knowl Manag* 2(2):405–410
15. Oh JK, Cho SJ, Bang WC, Chang W, Choi ES, Yang J, Cho JK, Kim DY (2004) Inertial sensor based recognition of 3-D character gestures with an ensemble classifiers. In: *Proceedings of the 9th international workshop on frontiers in handwriting recognition*, pp 112–117
16. Otiniano-Rodríguez KC, Cámara-Chávez G, Menotti D (2012) Hu and Zernike moments for sign language recognition. In: *Proceedings of international conference on image processing, computer vision, and pattern recognition*, pp 1–5
17. Priyal SP, Bora PK (2010) A study on static hand gesture recognition using moments. In: *Proceedings of international conference on signal processing and communications (SPCOM)*, pp 1–5
18. Priyal SP, Bora PK (2013) A robust static hand gesture recognition system using geometry based normalizations and Krawtchouk moments. *Pattern Recognit* 46(8):2202–2219
19. Ren Z, Yuan J, Meng J, Zhang Z (2013) Robust part-based hand gesture recognition using Kinect sensor. *IEEE Trans. Trans Multimed* 15(5):1110–1120
20. Shimada A, Yamashita T, Taniguchi RI (2013) Hand gesture based TV control system-towards both user- and machine-friendly gesture applications. In: *Proceedings of 19th Korea-Japan joint workshop on frontiers of computer vision (FCV)*, pp 121–126
21. Shotton J, Sharp T, Kipman A, Fitzgibbon A, Finocchio M, Blake A, Cook M, Moore R (2013) Real-time human pose recognition in parts from single depth images. *Commun ACM* 56(1):116–124
22. Suarez J, Murphy RR (2012) Hand gesture recognition with depth images: a review. In: *Proceedings of the 21st IEEE international symposium on robot and human interactive communication*, pp 411–417
23. Swapna B, Pravin F, Rajiv VD (2011) Hand gesture recognition system for numbers using thresholding. *Commun Comput Inf Sci* 250:782–786
24. Tabata Y, Kuroda T, Okamoto K (2012) Development of a glove-type input device with the minimum number of sensors for Japanese finger spelling. In: *Proceedings of international conference disability, virtual reality and associated technologies*, pp 305–310
25. Vieriu RL, Goras B, Goras L (2011) On HMM static hand gesture recognition. In: *Proceedings of 10th international symposium on signals, circuits and systems (ISSCS)*, pp 1–4
26. Wu CH, Lin CH (2013) Depth-based hand gesture recognition for home appliance control. In: *Proceedings of IEEE 17th international symposium on consumer electronics*, pp 279–280



**Dong-Luong Dinh** received his B.S. and M.S. degree in Information and Communication Technology from Ha Noi University of Science and Technology, Vietnam, in 2001 and 2009 respectively. In 2001, he was a Lecturer with the Department of Information Technology, Nha Trang University, Vietnam. He is currently working toward his Ph.D. degree in the Department of Computer Engineering, Kyung Hee University, South Korea. His research interests include computer vision, machine learning, human computer interaction, human motion analysis and estimation.



**Sungyoung Lee** received his B.S. from Korea University, Seoul, South Korea in 1978. He got his M.S. and Ph.D. degrees in Computer Science from Illinois Institute of Technology (IIT), Chicago, Illinois, USA in 1987 and 1991 respectively. He has been a professor in the Department of Computer Engineering, Kyung Hee University, Korea since 1993. He is a founding director of the Ubiquitous Computing Laboratory, and has been affiliated with a director of Neo Medicinal ubiquitous-Life Care Information Technology Research Center, Kyung Hee University since 2006. Before joining Kyung Hee University, he was an assistant professor in the Department of Computer Science, Governors State University, Illinois, USA from 1992 to 1993. His current research focuses on Ubiquitous Computing, Cloud Computing, Intelligent Computing, Context-Aware Computing, WSN, Embedded Realtime and Cyber-Physical Systems, and eHealth. He has authored/coauthored more than 405 technical articles (130 of which are published in archival journals). He is a member of the ACM and IEEE.



**Tae-Seong Kim** received the B.S. degree in Biomedical Engineering from the University of Southern California (USC) in 1991, M.S. degrees in Biomedical and Electrical Engineering from USC in 1993 and 1998 respectively, and Ph.D. in Biomedical Engineering from USC in 1999. After his postdoctoral work in cognitive sciences at the University of California, Irvine in 2000, he joined the Alfred E. Mann Institute for Biomedical Engineering and Dept. of Biomedical Engineering at USC as a Research Scientist and Research Assistant Professor. In 2004, he moved to Kyung Hee University in Korea where he is currently Professor in the Biomedical Engineering Department. His research interests have spanned various areas of biomedical imaging including Magnetic Resonance Imaging (MRI), functional MRI, E/MEG imaging, DT-MRI, transmission ultrasonic CT, and Magnetic Resonance Electrical Impedance Imaging. Lately, he has started research work in proactive computing at the u-Lifecare Research Center where he serves as Vice Director. Dr. Kim has published more than 90 peer reviewed papers and 150 proceedings, 6 international book chapters, and holds 10 international and domestic patents. He is a member of IEEE from 2005–2013, KOSOMBE, and Tau Beta Pi, and listed in Who's Who in the World '09–'13 and Who's Who in Science and Engineering '11–'12.