International Journal of Fuzzy Logic and Intelligent Systems, vol. 11, no. 4, December 2011, pp. 293-298
http://dx.doi.org/10.5391/IJFIS.2011.11.4.293

Rotation Invariant Histogram of Oriented Gradients

Minkyu Cheon*, Wonju Lee*, Chang-Ho Hyun+ and Mignon Park*+**

*** School of Electrical and Electronic Engineering, Yonsei University**
**** Division of Electrical Electronic and Control Engineering, Kongju National University**

Abstract

In this paper, we propose a new image descriptor, that is, a rotation invariant histogram of oriented gradients (RIHOG). RIHOG overcomes a disadvantage of the histogram of oriented gradients (HOG), which is very sensitive to image rotation. The HOG only uses magnitude values of a pixel without considering neighboring pixels. The RIHOG uses the accumulated relative magnitude values of corresponding relative orientation calculated with neighboring pixels, which has an effect on reducing the sensitivity to image rotation. The performance of RIHOG is verified via the index of classification and classification of Brodatz texture data.

Keywords: Rotation invariant histogram of oriented gradients, Relative orientation, Relative magnitude, Index of classification, Texture classification

1. Introduction

Recently, various algorithms for extracting local image feature have been studied for object matching, target detection, and texture classification. One of the most popular image descriptor is the scale invariant feature transform (SIFT), which is introduced by Lowe [1]. SIFT has illumination, scale, rotation, and affine invariant properties. It has applied to various vision based objects recognition and matching systems [2-6]. In addition, various modified SIFT algorithms including PCA-SIFT [7] have been studied for different applications [8-10]. However, matching systems using SIFT needs relatively long running time compared with systems using other general image features, so SIFT is actually not suitable for real-time object detection or classification.

The SIFT is a sort of sparse descriptor which first detects keypoints (interest points) of a given image and then generates descriptors at keypoint locations. The other descriptor which is called dense descriptor uses the method that extracts features from all over the given image without detecting keypoints. The histogram of oriented gradients (HOG) which is proposed by Dalal and Triggs [11] is one of the most popular dense descriptor. The aim of the HOG is to describe an image by a set of local oriented gradients histogram. These histograms represent occurrences of specific gradient orientation in a local part of images. HOG is usually used to detect a specific object from images. Especially, HOG shows good performance for human detection [11-13]. However, it has a disadvantage that is very sensitive to image rotation. Therefore, HOG is not good choice for classification of textures or objects which can often be detected as rotated image.

In this paper, we propose a new image descriptor which is a rotation invariant histogram of oriented gradients (RIHOG). RIHOG is a kind of dense image descriptors and it fundamentally follows characteristics of HOG that uses oriented gradients. We develop the RIHOG using the fact that relative orientation between a specific pixel and its neighboring pixels cannot be changed according to image rotation. By constructing histogram using those relative orientation and magnitude values, RIHOG can be robust against image rotation.

In this paper, we verify the performance of our proposed descriptor, RIHOG through texture classifications. Brodatz textures [15] which are rotated with specific angles are used for performance verification, and we compare the performance of RIHOG with that of HOG and SIFT.

In Section 2, we briefly introduce HOG and introduce RIHOG in Section 3. In Section 4, we verify the performance of RIHOG through Brodatz textures classification and compare it with HOG and SIFT.

2. Histogram of Oriented Gradients

The aim of HOG is to describe an image with a local oriented gradient histogram. These histograms represent occurrences of specific gradient orientation in a local part of images. The HOG can be calculated by three step-sequence: gradient computation, orientation binning, and histogram generation [11].

In gradient computation step, the gradient of an image is obtained by applying two one-dimensional filters, which are (-1 0 1) for horizontal direction and $(-1\ 0\ 1)^T$ for vertical direction. Either a signed or unsigned gradient can be used; in our case, we use signed gradient whose values range from $-\pi$ to π.

The next step is orientation binning, which is used to

Manuscript received Jul. 20, 2011; revised Oct. 20, 2011; accepted Nov. 2;
+Corresponding author
"This work was supported by the Grant of the Korean Ministry of Education, Science and Technology" (The Regional Core Research Program/Zero Energy Green Village Technology Center)

compute the histogram of orientation. For orientation binning, the image should be divided into blocks of predefined size. Further, for histogram generation, the range of each bin is determined. For example, if the signed gradient is divided into 6 bins with same range, then the range of each bin is 60 degrees ($\pi/3$ radians).

In the histogram generation step, we impose values on histogram of each block. According to the orientation of the gradient, the magnitude of the gradient is accumulated to the bin of the histogram.

Because the HOG shows occurrences of specific gradient orientation, the histogram can be considerably changed by image rotation as shown in Figure 1. Therefore, HOG is not suitable as feature vectors for classification of objects which are often detected as rotated images or texture images.

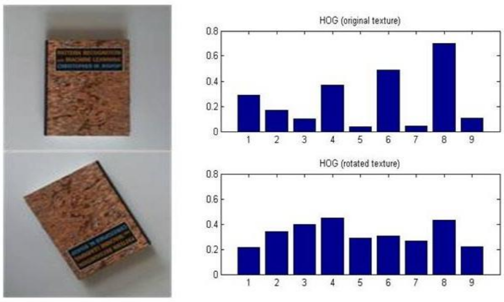

Fig. 1. Histogram of oriented gradients of a book image and the rotated image of the same book

3. Rotation Invariant Histogram of Oriented Gradients

In this section, we introduce our proposed image descriptor, RIHOG. RIHOG uses a method that accumulates relative magnitude values of corresponding relative orientation between a pixel and its neighboring pixels.

3.1. Gradient computation

This step is same with the first step of constructing HOG, gradient computation. In other words, the gradient of an image is obtained by filtering it with two one-dimensional filters, which are (-1 0 1) for horizontal direction and $(-1\ 0\ 1)^T$ for vertical direction. Next, magnitude and orientation values are obtained from those horizontal and vertical gradient values. These magnitude and orientation values which are generated from each pixel are stored for next step.

3.2. Histogram construction

RIHOG is constructed using relative orientation and magnitude values between a pixel and its neighboring pixels. In the case of HOG, the histogram is constructed by accumulating magnitude value of each pixel according to orientation value of the pixel after generating orientation and magnitude of every pixel. However; in the case of RIHOG which we propose in this paper, the histogram is constructed at each pixel using relative orientation and magnitude, and the final histogram is constructed by accumulating these all histograms.

The relative orientation of a pixel is calculated by difference between the orientation of the pixel and orientation of neighboring pixels. In other words, assuming that orientation of a pixel is θ_o , and orientation of n neighboring pixels of corresponding pixel is θ_i (e.g. $\theta_1, \theta_2, \cdots, \theta_8$ in Figure 3.), we define relative orientation as $\theta_o - \theta_i$. In this paper, because we define range of the orientation as $-\pi < \theta \le \pi$, the relative orientation which is outside the pre-defined range is defined as follows:

$$\theta' = \begin{cases} \theta_o - \theta_i + 2\pi & \text{if } \theta_o - \theta_i < -\pi \\ \theta_o - \theta_i - 2\pi & \text{if } \theta_o - \theta_i > \pi \\ \theta_o - \theta_i & \text{otherwise} \end{cases} \quad (1)$$

The relative magnitude is also regenerated by the relative values between the magnitude of the pixel and magnitude of neighboring pixels. If accumulating relative magnitude value is negative, the essential characteristic of accumulated value can be lost by reduction of histogram value. Therefore, relative magnitude should be positive.

In this paper, we propose three ways of calculating relative magnitude as follows:

1) RIHOG1 : Relative magnitude is the absolute value of the difference between magnitude of a pixel and magnitude of neighboring pixels.

$$M_i' = |M_o - M_i| \quad (2)$$

This value shows not ratio but absolute difference of M_o and M_i , and has disadvantage that cannot separate the case of $M_o > M_i$ and $M_i > M_o$.

2) RIHOG2 : Basically, the second method uses the absolute value of the difference between magnitude of a pixel and magnitude of neighboring pixels which is same with RIHOG1. However, two histograms are generated for the case of $M_o > M_i$ and $M_i > M_o$.

$$M_i' = |M_o - M_i| \quad (3)$$

This value shows not ratio but absolute difference of M_o and M_i , and has disadvantage that dimension of the feature vector should be double size of the other RIHOG methods as expanding histogram.

3) RIHOG3 : Relative magnitude of RIHOG3 uses $\tan^{-1}$ value as follows:

$$M_i' = \tan^{-1}\left(\frac{M_o}{M_i}\right) \tag{4}$$

This value shows not absolute difference but ratio of M_o and M_i, so it has different values in the case of $M_o > M_i$ and $M_i > M_o$.

where magnitude of a pixel is M_o, and magnitude of n neighboring pixels of corresponding pixel is θ_i and M_i. (e.g. M_1, $M_2, \cdots, M_8$ in Figure 3.)

After obtaining the relative orientation and magnitude values of neighboring pixels, a histogram is constructed by those values at each pixel. An orientation bin is determined by a relative orientation. Next, the relative magnitude value is accumulated at the orientation bin in histogram and those accumulated values are normalized. Figure 2 shows brief illustration of RIHOG generation, and Figure 3 shows process of generating RIHOG using 8 neighboring pixels.

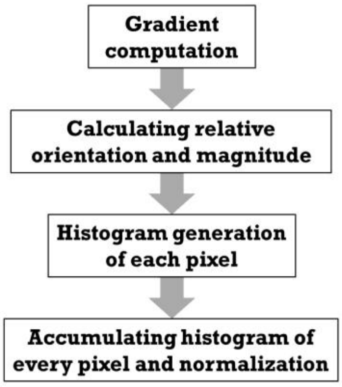

Fig. 2. Brief illustration of rotation invariant histogram of oriented gradients

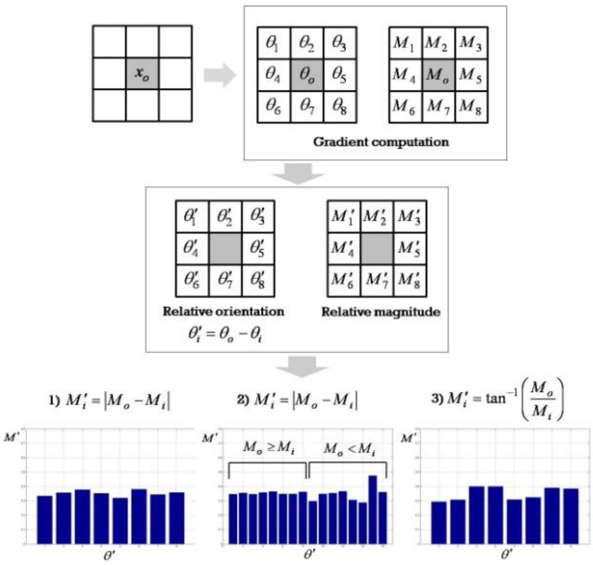

Fig. 3. Brief illustration of rotation invariant histogram of oriented gradients and three kinds of relative magnitude

After obtaining histograms which are generated by relative orientation and magnitude of neighboring pixels at each pixel, all histograms are accumulated, and the accumulated values of each orientation bin are normalized. Therefore, accumulated magnitude values of each orientation bin are always larger than 0, and smaller than 1.

3.3. Neighboring step

In this paper, we basically define neighboring pixels as 8 pixels which are next to a specific pixel [19].

However, we can redefine the neighboring pixels as pixels which are located *n* pixels away from the corresponding pixel. In this paper, concatenations of histograms which use different neighboring steps are also used as feature vectors as shown in Figure 4.

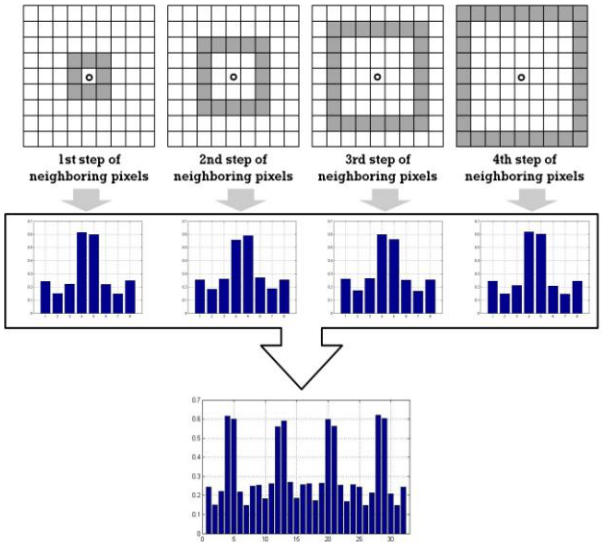

Fig. 4. Neighboring steps and new histogram generation using concatenation of RIHOGs of different neighboring steps

4. Experiment

4.1. Database

We use 111 Brodatz texture data for experiment [15]. (see Table 1) Figure 5 shows several kinds of Brodatz texture data. For performance verification experiment, we generate 12 new data for each texture which are rotated with specific angle and cut by 128×128 size as shown in Figure 6.

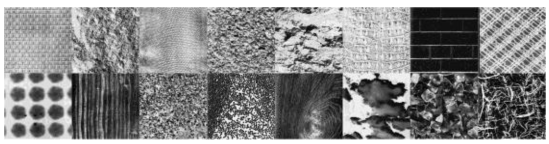

Fig. 5. Example images of Brodatz texture: (from upper-left image) D1, D2, D3, D5, D7, D10, D25, D46, D48, D50, D54, D67, D69, D91, D99, D107

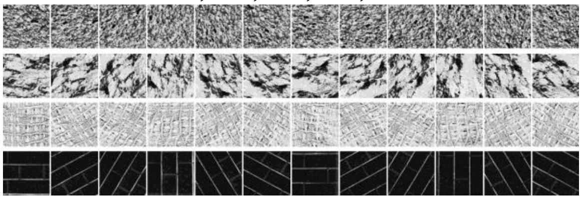

Fig. 6. Rotated Brodatz texture images: each texture is rotated with certain angles (from 0^o to 330^o with 30^o intervals) and cut by 128×128 size at the center of the rotated texture image

Table I. Database for performance verification

	Number of texture	Number of rotation	Total number of data
Brodatz texture	111	12	1332

4.2. Index of classification

In this section, we introduce an index of classification and show the index of classification of RIHOGs using data that we mentioned in Section 4.1. An index of classification is generated by distance between data which are included different classes and distance between data included in the same class. The index of classification can be used to expect that how well these data can be classified. The index of classification of class i can be represented as follows:

$$\begin{array}{c}\text{Index of}\\ \text{classification}\\ \text{for class } i\end{array} = \frac{\begin{array}{c}\text{Average distance between data}\\ \text{in class } i \text{ and all data in different classes}\end{array}}{\begin{array}{c}\text{Average distance between data}\\ \text{in class } i\end{array}}$$

According to above definition of the index of classification, the smaller distance between data included in the same class is and the larger distance between data included in different classes is, the larger the index of classification is. In other words, we can expect more correct classification result when the index of classification is relatively large.

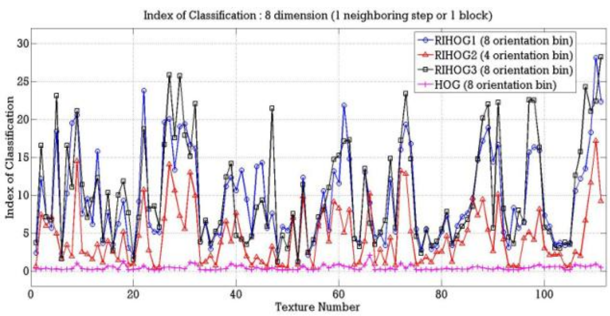

Fig. 7. Index of classification of each class : 8-dimensional HOG (8 orientation bins, 1 block), 8-dimensional RIHOG1 (8 orientation bins, 1 neighboring step), 8-dimensional RIHOG2 (4 orientation bins, 1 neighboring step), 8-dimensional RIHOG3 (8 orientation bins, 1 neighboring step)

We compare the index of classification of HOG and three kinds of RIHOGs using rotated Brodatz texture images which we mentioned in Section 4.1. We use 8-dimensional HOG and RIHOG. Figure 7 shows the index of classification of 8-dimensional HOG and three kinds of 8-dimensional RIHOGs for each class of Brodatz textures, and Table 2 shows average, minimum and maximum index of classification of 8-dimensional HOG and three kinds of 8-dimensional RIHOGs. On average, RIHOG 3 shows the largest index of classification.

Table II. Minimum, maximum, and average Index of classification : 8-dimensional HOG (8 orientation bins, 1 block), 8-dimensional RIHOG1 (8 orientation bins, 1 neighboring step), 8-dimensional RIHOG2 (4 orientation bins, 1 neighboring step), 8-dimensional RIHOG3 (8 orientation bins, 1 neighboring step)

Histogram	Index of classification		
	Minimum	Maximum	Average
HOG	0.10	2.03	0.37
RIHOG1	1.23	28.10	9.57
RIHOG2	0.35	17.12	4.39
RIHOG3	1.14	28.21	10.16

4.3. Texture classification using the nearest neighbor method

In this section, we show texture classification result using the nearest neighbor method. We randomly select 7 texture data as training data and 5 data as test sample data among 12 rotated texture images for each Brodatz texture. In other words, 555 test sample texture data and 777 training texture data are used for classification. We compare result of 8, 16, and 32-dimensional HOG and three kinds of RIHOGs as shown in Table III (a)-(d).

In addition, we also classifies textures using SIFT. We use the most general 128-dimensional descriptors, and regard class label of training data which has the largest number of matching pairs with the test sample data as classification result.

Average number of SIFT descriptors for one 128×128 texture data is 1371, and matching one test sample texture data with all training textures using SIFT descriptors, on average, takes 97.54 sec. On the other hand, in the case of three kinds of RIHOGs, only one RIHOG vector is generated from one texture data, and classification by matching all 32-dimensional RIHOGs test sample data with all RIHOGs training data using the nearest neighbor method, on average, takes only 0.1216 sec. (The hardware spec. of system is 2.66 GHz CPU, 4G RAM and the development software is MATLAB 7.9.)

Table III (a)-(e) and Figure 6 show classification result of HOG, three kinds of RIHOGs, and SIFT. On average, RIHOG3 shows the best classification performance among three kinds of RIHOGs, and several cases of RIHOG show similar or better classification result than SIFT, although feature dimensions and classification time of RIHOG are much less than those of SIFT.

The best classification results of three kinds of RIHOGs are 95.5 percent accuracy for 32-dimensional RIHOG1 (8 orientation bins, and 4 neighboring steps), 90.81 percent accuracy for 32-dimensional RIHOG2 (8 orientation bins, and 2 neighboring steps), and 95.31 percent accuracy for 32-dimensional RIHOG3 (16 orientation bins, and 2 neighboring steps). The classification result using 128-dimensional SIFT descriptors is 94.96 percent accuracy.

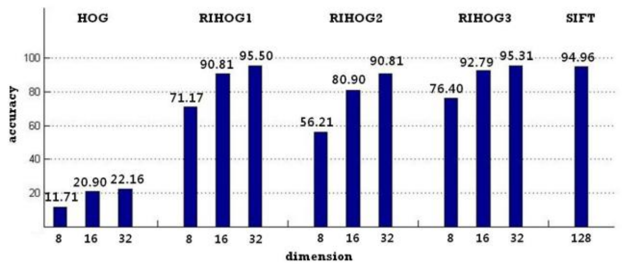

Fig. 8. Classification accuracy of HOG, three kinds of RIHOGs, and SIFT (The best case of 8, 16, and 32-dimensional HOG and RIHOG)

Table III-(a). Classification results of HOG

HOG			
Dimension	Number of orientation bin	Neighboring step	Classification accuracy
8	8	1	11.71
16	8	1, 2	11.17
	16	1	20.90
32	8	1, 2, 3, 4	11.53
	16	1, 2	20.00
	32	1	**22.16**
Average			16.25

Table III-(b). Classification results of RIHOG1

RIHOG 1			
Dimension	Number of orientation bin	Neighboring step	Classification accuracy
8	8	1	71.17
16	8	1, 2	90.81
	16	1	81.98
32	8	1, 2, 3, 4	**95.50**
	16	1, 2	92.97
	32	1	79.28
Average			85.29

Table III-(c). Classification results of RIHOG2

RIHOG 2			
Dimension	Number of orientation bin	Neighboring step	Classification accuracy
8	4	1	56.21
16	4	1, 2	80.90
	8	1	78.56
32	4	1, 2, 3, 4	89.01
	8	1, 2	**90.81**
	16	1	85.59
Average			85.29

Table III-(d). Classification results of RIHOG3

RIHOG 3			
Dimension	Number of orientation bin	Neighboring step	Classification accuracy
8	8	1	7.40
16	8	1, 2	92.79
	16	1	86.67
32	8	1, 2, 3, 4	94.23
	16	1, 2	**95.13**
	32	1	87.03
Average			88.74

Table III-(e). Classification results of SIFT

SIFT		
Dimension	Average number of descriptors for one 128x128 texture image	Classification accuracy
128	1371	**94.96**

5. Conclusions

In this paper, we propose a new image descriptor, the rotation invariant histogram of oriented gradients. In order to overcome a disadvantage of HOG which is very sensitive to image rotation, the rotation invariant histogram of oriented gradients uses method that accumulates relative magnitude values of corresponding relative orientation calculated with neighboring pixels. We show performance of RIHOG through the index of classification and Brodatz texture classification.

Our proposed RIHOG shows relatively large index of classification and much better texture classification result than HOG. In addition, some of classification results of RIHOG are similar or better than that of SIFT, although dimension and classification time of RIHOG are much less than those of SIFT.

Because RIHOG has a remarkable characteristic which is robust against image rotation, we can expect that RIHOG can be applied to many different applications.

References

[1] David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp.91-110, 2004.

[2] Liang Zhang, Xiaomei He, “Fake Shadow Detection Based on SIFT Features Matching,” *2010 WASE International Conference on Information Engineering*, vol. 1,pp.216 – 220, 2010.

[3] Kus, M.C., Gokmen, M., Etaner-Uyar, S., “Traffic sign recognition using Scale Invariant Feature Transform and color classification,” *Computer and Information Sciences, 2008. ISCIS '08. 23rd International Symposium on* pp.1-6, 2008.

[4] Jun Luo, Ma, Y., Takikawa, E., Lao, S., Kawade, M., Bao-

Liang Lu, "Person-Specific SIFT Features for Face Recognition," *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference*, vol. 2, pp.593-596, 2007.

[5] Changhan Park, Sunghun Jung, "SIFT-based object recognition for tracking in infrared imaging system," *Infrared, Millimeter, and Terahertz Waves, 2009. IRMMW-THz 2009. 34th International Conferenc*, pp.1-2, 2009.

[6] Krizaj, J, Struc, V., Pavesic, N, "Adaptation of SIFT features for face recognition under varying illumination," *MIPRO, 2010 Proceedings of the 33rd International Convention*, pp.691-694, 2010.

[7] Y. Ke, R. Sukthankar, "PCA-SIFT : A More Distinctive Representation for Local Image Descriptors," *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, 2004.

[8] Lazebnik, S. Schmid, C., Ponce, J., "A sparse texture representation using local affine regions," *Pattern Analysis and Machine Intelligence, IEEE Transactions*, vol. 27, no. 8, pp.1265-1278, 2005.

[9] Wang Yu-quan, Xia Gui-hua, Zhu Qi-dan, Zhao Guo-liang, "Monte Carlo Localization of Mobile Robot with Modified SIFT," *Measuring Technology and Mechatronics Automation, 2009. ICMTMA '09. International Conference on*, vol. 3, pp.400-403, 2009.

[10]Cheng-Yuan Tang, Yi-Leh Wu, Maw-Kae Hor, Wen-Hung Wang, "Modified sift descriptor for image matching under interference," *Machine Learning and Cybernetics, 2008 International Conference*, vol. 6, pp.3294-3300, 2008.

[11]Dalal, N., Triggs, B., "Histograms of oriented gradients for human detection," *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference*, pp.886-893, 2005.

[12]Bertozzi, M., Broggi, A., Del Rose, M., Felisa, M., Rakotomamonjy, A., Suard, F., "A Pedestrian Detector Using Histograms of Oriented Gradients and a Support Vector Machine Classifier," *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*, pp.143-148, 2007.

[13]Qiang Zhu, Mei-Chen Yeh, Kwang-Ting Cheng, Avidan, S., "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients," *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference*, pp. 1491-1498, 2006.

[14]Jie Chen, Shiguang Shan, Chu He, Guoying Zhao, Pietikäinen, M., Xilin Chen, Wen Gao, "WLD: A Robust Local Image Descriptor," *Pattern Analysis and Machine Intelligence, IEEE Transactions*, vol. 32, no. 9, pp.1705-1720, 2010.

[15]http://www.ux.uis.no/~tranden/brodatz.html

Minkyu Cheon was born in Busan, Korea, in 1981. He received a B.S. in Electrical and Electronic Engineering from Yonsei University, Seoul, Korea, in 2006. He is currently a Combined Master's and Doctoral Program candidate in Electrical and Electronic Engineering at Yonsei University. His main research interests include machine learning, pattern recognition, and computer vision.

E-mail : 1000minkyu@yonsei.ac.kr

Chang-Ho Hyun received the B.S. degrees in control and instrumentation engineering from Kwangwoon University, Seoul, Korea, , and the M.S. and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 1999. 2002, 2008. From 2008 to 2009, he was a senior engineering in Samsung Electronics. Since 2009, he has joined the faculties of the School of Electrical, Electronic and Control Engineering at Kongju National University, where he is currently an assistant professor. His current research interests include intelligent control and application, nonlinear control, robotics, mobile robots.

E-mail : hyunch@kongju.ac.kr

Wonju Lee was born in Paju, Korea, in 1982. He received a B.S. from the Department of Computer Engineering of Hyupsung University in 2005 and an M.S. from the Electronic Engineering of Yonsei University in 2007. He is currently a Doctoral Program candidate in Electrical and Electronic Engineering at Yonsei University. His main research interests include machine learning, pattern recognition, and computer vision.

E-mail : delicado@yonsei.ac.kr

Mignon Park received a B.S. and an M.S. in electronics from Yonsei University, Seoul, Korea, in 1973 and 1977, and a Ph.D. from the University of Tokyo, Tokyo, Japan, in 1982. He was a researcher in the Institute of Biomedical Engineering, University of Tokyo, from 1972 to 1982, as well as at the Massachusetts Institute of Technology, Cambridge, and the University of California Berkeley, in 1982. He was a visiting researcher in the Robotics Division, Mechanical Engineering Laboratory Ministry of International Trade and Industry, Tsukuba, Japan, from 1986 to 1987. He has been a Professor in the Department of Electrical and Electronic Engineering, Yonsei University, since 1982. His research interests include fuzzy control and applications, robotics, and fuzzy biomedical systems.

E-mail : mignpark@yonsei.ac.kr