ISSN 2586-1816 / e-ISSN 2586-4440
융합정보논문지 제7권 제6호 pp. 143-151, 2017 DOI : https://doi.org/10.22156/CS4SMB.2017.7.6.143

Performance Evaluation of the Generalized Hough Transform

Ji-Young Chang*
Department of Computer Engineering, Gwangju University

일반화된 허프변환의 성능평가

장지영*
광주대학교 컴퓨터공학과

Abstract The generalized Hough transform(GHough) can be used effectively for detecting and extracting an arbitrary-shaped 2-D model in an input image. However, the main drawbacks of the GHough are both heavy computation and an excessive storage requirement. Thus, most of the researches so far have focused on reducing both the time and space requirement of the GHough. But it is still not clear how well their improved algorithms will perform under various noise in an input image. Thus, this paper proposes a new framework that can measure the performance of the GHough quantitatively. For this purpose, we view the GHough as a detector in signal detection theory and the ROC curve will be used to specify the performance of the GHough. Finally, we show that we can evaluate the GHough under various noise conditions in an input image.

Key Words : Hough Transform, Generalized Hough Transform, Shape Extraction, Object Recognition, Receiver Operating Characteristic

요 약 일반화된 허프변환은 임의의 형태의 2차원 모델을 입력영상에서 탐지 및 추출하는데 사용되어지는 효과적인 방법이다. 그러나 일반화된 허프변환의 단점으로 실행시간이 오래 걸린다는 것과 과도한 메모리 사용을 들 수 있다. 그래서 현재까지의 대부분의 연구는 일반화된 허프변환의 실행시간과 메모리 사용량을 줄이는데 집중되어왔다. 그러나 실행시간과 메모리 사용을 줄여서 개선된 알고리즘이 입력 영상에 존재하는 노이즈를 고려할 경우 어떤 성능을 제공하는가는 여전히 불분명하다. 그러므로 본 논문은 일반화된 허프변환의 성능 평가를 위한 새로운 프레임워크를 제안한다. 이를 위해 일반화된 허프변환을 신호탐지 이론의 탐지기로 간주하며 ROC 커브를 사용해서 일반화된 허프변환의 성능을 정의한다. 마지막으로 입력 영상에서의 노이즈를 고려한 정량적인 성능 평가가 가능함을 보인다.

키워드 : 허프변환, 일반화된 허프변환, 모양 추출, 물체 인식, 반응자 작용 특성

1. Introduction

With increased computer speed and advanced technology in artificial intelligence, the image recognition is gaining momentum again. The image recognition technology has evolved over time, and many studies are being done in such area as motion detection, biometric recognition, and object tracking [1-5]. For those high level technology to be successful, the Hough transform(HT) is a low level technique that can be used at the preprocessing stage of those algorithms. The HT is mainly used to detect such parametric curves as straight lines, circles, and

Received 2017-11-01 Revised 2017-11-17 Accepted 2017-12-20 Published 2017-12-31
*Corresponding author : Ji-Young Chang (jchang@gwangju.ac.kr)

ellipses[6]. Ballard proposed the generalized Hough transform(GHough) that can detect an arbitrarily scaled and rotated model[7]. The classical survey by Illingworth and Kittler and the recent survey by Mukhopadhyay and Chaudhuri discuss the variants and extensions of both HT and the GHough[8,9]. The GHough is an efficient method when the rotation and scale parameter of a given model is known in advance. However, the GHough requires both heavy computation and an excessive storage of 4-dimensional accumulator array due to the enumeration scheme used in the algorithm. Thus, most of the researches so far have been focused on improving the time and storage requirement of the GHough. A randomized generalized Hough transform was proposed to reduce the execution time of the GHough, Multiple points with their geometric constraints were used to reduce the parameter space from 4-D to 2-D[10-12]. Even though those approaches have been successful in improving both the time and storage requirement of the GHough, it is not still clear how well their proposed algorithms will perform under various noise that can exist in an input image. Thus, this paper propose a framework that allows for measuring the performance of the GHough. This can be done by viewing the GHough as a detector in signal detection theory[13]. This paper consists of the following : (1) We first introduce the theory of hypothesis testing (2) Two important conditional probability functions of the cell counts in the parameter space of the GHough will be derived (3) The ROC(Receiver Operating Characteristic) curve will be used to specify the performance of the GHough under various noise conditions.

2. Generalized Hough Transform(GHough)

The GHough works in 2 phases : model encoding and detection. For the encoding of a given shape, the GHough selects an arbitrary reference point (x_r, y_r) inside the model and stores the displacement vector $\boldsymbol{d_i}$ into the R-table using a gradient angle ϕ_i at each $\boldsymbol{p}_i$ as an index. Fig. 1 shows the model encoding scheme of the GHough. In the second phase of the GHough, voting is done into a 2-dimensional accumulator array A. That is, for each edge pixel $\boldsymbol{P}_i$ in an input image, the edge gradient Φ_i is computed and used to look up the R-table. For each (r, α) value stored there, the location of the potential reference point (x_r, y_r) is computed using the following equation :

$$x_r = x + r\cos\alpha,\ y_r = y + r\sin\alpha \qquad (1)$$

and the $A(x_r, y_r)$ is incremented. Once all the edge pixels in an input image are processed, the accumulator array A is searched to find local maxima which is considered to be the model found. In order to adapt to variations in scale s and rotation θ of the model, the above equation (1) can be modified as

$$x_r = x + rs\cos(\alpha+\theta),\ y_r = y + rs\sin(\alpha+\theta) \qquad (2)$$

and the accumulator array A needs to become 4-dimensional.

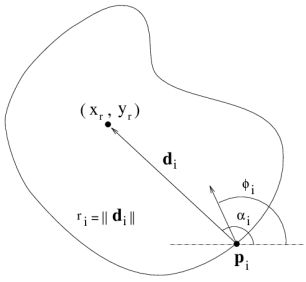


Fig. 1. Model Encoding Scheme of the GHough

3. Hypothesis Testing

To develop an evaluation method for the GHough, we briefly introduce the hypothesis testing. Let us imagine that we measure a single quantity x and choose a classification, based on the outcome of the

measurement of x, of one of two possible hypotheses H_0 and H_1 . Then, the outcome of the measurement x is a random variable with density functions $p_0(x)$ and $p_1(x)$, respectively. Then, x can fall into two regions $R_0(x \leq x_0)$ and $R_1(x > x_0)$, such that the hypothesis H_0 is chosen when x lies in region R_0 and the hypothesis H_1 is chosen otherwise. Fig. 2 shows the two regions.

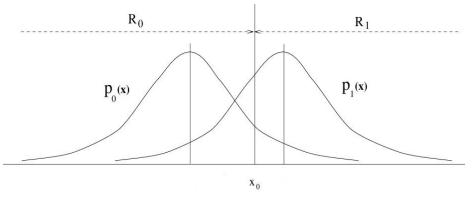


Fig. 2. Probability density functions p_0 and p_1 conditioned under hypothesis H_0 and H_1, respectively

Regardless of what regions we select for R_0 and R_1 with a threshold x_0, we can always consider the following quantities :

$$P\{D_0|H_0\} = \int_{R_0} p_0(x)dx \tag{3}$$

$$P_F = P\{D_1|H_0\} = \int_{R_1} p_0(x)dx \tag{4}$$

$$P\{D_0|H_1\} = \int_{R_0} p_1(x)dx \tag{5}$$

$$P_D = P\{D_1|H_1\} = \int_{R_1} p_1(x)dx \tag{6}$$

The quantities (4) and (6) are called the "probability of false alarm"(P_F), and the "probability of detection" (P_D), respectively. A specific threshold x_0 determines both P_F and P_D of a detector.

4. Receiver Operating Characteristic(ROC)

The Receiver Operating Characteristic(ROC) curve is a plot of P_D versus P_F for varying threshold selected[14]. Fig. 3 shows a typical Receiver Operating Characteristic (ROC) curve for the distribution shown in Fig. 2. What the curve tells us is that, if we set the threshold such that the P_F is equal to 0.1, the detection rate(P_D) we can achieve is 0.7. So, for varying threshold selected, which is implicit in the curve, we can measure the performance of a detector by plotting the probability of detection for varying levels of false alarm probability.

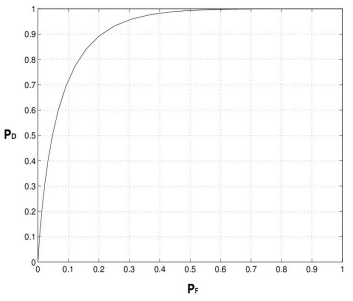


Fig. 3. A sample ROC curve that shows the relationship between P_D and P_F

5. Conditional Probability Functions

In order to obtain the probabilities(P_F and P_D) for the evaluation of the GHough, we need two conditional probability functions : a probability function of a cell count when an instance of an object exists in an image ($\boldsymbol{p_1}$) and a probability function of a cell count when no instance of an object exists($\boldsymbol{p_0}$). These conditional probability functions were derived and used to set the scientific threshold for the GHough considering both Type I and Type II error[15]. Here, we recap the important steps for the derivation of $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$ and extend the work to obtain P_F and P_D for a complete evaluation of the GHough using the ROC curves.

5.1 Spreading Effect in Parameter Space

In this section, we are interested in finding the effects due to the uncertainties in measuring image features in a given image. Suppose an image contains

an instance of the model which is rotated by θ and scaled by s. If the image were ideal with no noise, then it would have a point P_I with gradient angle Φ_I, which is a correct instance of the i^{th} model point p_i with a gradient angle ϕ_i. Thus, the GHough would vote into a point in 2-dimensional translation space $P_{\theta,s}$ corresponding to a rotation θ and a scale s. This point t in $P_{\theta,s}$ is given as

$$t = P_I + sR_\theta d_i \tag{7}$$

Thus, the R-table is indexed using the angle $\Phi_I - \theta$ and the displacement vector $\boldsymbol{d_i}$'s are retrieved and each $\boldsymbol{d_i}$'s needs to be rotated by θ and scaled by s to handle the rotation and scale appropriately.

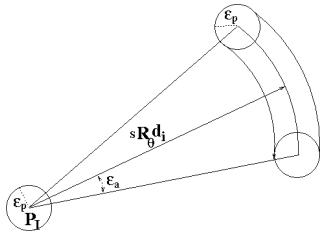


Fig. 4. A set of translations due to measurement errors

However, since an image cannot be ideal, the observed gradient angle Φ_I^{ob} and the observed location P_I will have deviation by ϵ_a and ϵ_p, respectively. That is,

$$|\Phi_I^{ob} - \Phi_I| \le \epsilon_a, \quad \text{and} \tag{8}$$

$$\| P_I^{ob} - P_I \| \le \epsilon_p \tag{9}$$

Considering both positional and angular error, the GHough should vote into a region in the parameter space $P_{\theta,s}$. This region is shown in Fig. 4, and is given by

$$V_{(d_p s)} = \left\{ P_I^{ob} + sR_\theta d_i \mid \left(|P_I^{ob} - P_I| \le \epsilon_p, |\theta_{true} - \theta| \le \epsilon_a \right) \right\} \tag{10}$$

As we can see, this region is related to the length of the i^{th} displacement vector $\boldsymbol{d_i}$ and the amount of positional and angular uncertainty(ϵ_p and ϵ_a). The region corresponding to an arbitrarily shaped model is difficult to obtain. However, we can get the worst case area of spread from the circle whose radius is r in Fig. 5. Note that the distance between A and B is equal to r.

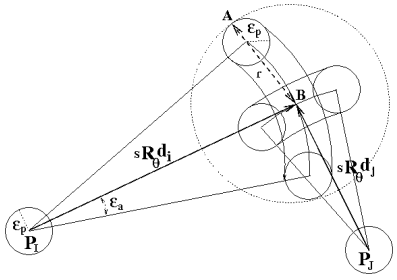


Fig. 5. The region of spread for an arbitrary model

Since the r is given as

$$r = 2 \parallel sd_{\max} \parallel \sin\left(\frac{\epsilon_a}{2}\right) + \epsilon_p$$

the worst case area of spread is

$$G_{spread} = \pi \left[2 \parallel sd_{\max} \parallel \sin\left(\frac{\epsilon_a}{2}\right) + \epsilon_p \right]^2 \tag{11}$$

Where $d_{\max}$ is the length of the longest displacement vector in the model.

Since the votes into the parameter tend to spread into a region instead of a point in the parameter space, we must sum all the votes cast into a moving circle with radius r. Thus, at the peak detection stage, we assume that the GHough will find a maximum sum of cells inside the moving circle of radius r to get the local maxima.

Even though we use the moving circle of radius r at the peak detection stage of the GHough, P_I can vote into the moving circle if the gradient angles in the model and an image match. That is, the I^{th} image point can vote only if the angle $\Phi_I^{ob} - \theta$ matches with

ϕ_i in the R-table. However, since

$$\left|\Phi_I^{ob}-\Phi_I\right| \le \epsilon_a \tag{12}$$

the indexing into the R-table will not always occur. In order to compute the probability that I^{th} image point casts a vote into G_{spread}, let us assume that $2\epsilon_a \le q$. That is, the quantization size q of a gradient angle is sufficiently large to handle the possible angular error. Then, there are 3 cases(Events) to consider as shown in Fig. 6.

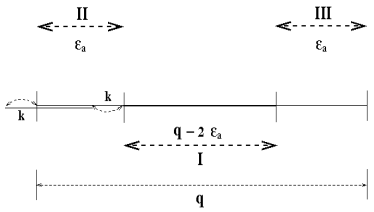


Fig. 6. Non-intersecting sub-intervals I, II, and III of the R-Table bin

Thus, we define the following events :

$A1$: Φ_I lies in the interval I,

$A2$: Φ_I lies in the interval II, and

$A3$: Φ_I lies in the interval III.

Also, if we let A be the event "Indexing occurs." Then, the probability that the I^{th} image point casts a vote is given by

$$\begin{aligned} P_{vote} = {} & P(A0|A1)P(A1)+P(A0|A2)P(A2) \\ & +P(A0|A3)P(A3) \end{aligned} \tag{13}$$

Note that $P(E0|E2)$ depends on exactly where Φ_I lies in the interval II Depending on the value of k in Fig. 6, there are two cases;

1. Φ_I lies close to the end point of the interval I in which case the probability that an indexing occurs is close to 1
2. Φ_I lies close to the end point of quantization interval q in which case the probability that an indexing occurs is close 1/2. Thus, by averaging,

the probability P_{vote} that an I^{th} image point casts a vote into the area G_{spread} is given as :

$$\begin{aligned} P_{vote} &\approx \frac{q-2\epsilon_a}{q}+2\left(\frac{3}{4}\right)\frac{\epsilon_a}{q} \\ &= 1-\frac{\epsilon_a}{2q}, (q \ge 2\epsilon_a) \end{aligned} \tag{14}$$

Thus, the probability that $k1$ model points votes into the moving circle is

$$P(k_1)=\binom{\lfloor f\times n \rfloor}{k_1} P_{vote}^{k_1}\left(1-P_{vote}\right)^{\lfloor f\times n \rfloor - k_1} \tag{15}$$

Where n, N are the number of model points, a number of image points(Edges) and f is the degree of occlusion($0 \le f \le 1$). These are the votes cast by real instances of $\lfloor f\times n \rfloor$ model points in a given image.

Also, since there are n displacement vectors stored in the R-Table, whose size is $\lfloor 2\pi/q \rfloor$, a number of votes cast into the moving circle on the average is given by

$$\beta(n,N,f,q)=\frac{n}{\lfloor 2\pi/q \rfloor}(N-\lfloor f\times n \rfloor) \tag{16}$$

Thus, the probability that $k2$ non-model points cast a vote into the moving circle is

$$P(k_2)=\binom{\beta(n,N,f,q)}{k_2} P_{spread}^{k_2}(1-P_{spread})^{\beta(n,N,f,q)-k_2} \tag{17}$$

Where $P_{spread}=\dfrac{G_{spread}}{1\times 1}$

Here, we assume an image size is 1×1 unit-square and $k1$ and $k2$ are also assumed to be independent, $\boldsymbol{p_1}$ can be obtained as :

$$\boldsymbol{p_1}(k)=P(k_1)\otimes P(k_2) \tag{18}$$

Where $\otimes$ is the convolution operator[16]. For even

moderate values of n and N, the computation of both $P(k1)$ and $P(k2)$ becomes cumbersome, so we approximate them using the Poisson approximation :

$$P(k_1) \approx \frac{\mu^{k_1}}{k_1!}e^{-\mu}, \text{ and } \quad P(k_2) \approx \frac{\nu^{k_2}}{k_2!}e^{-\nu}$$

Where $\mu = \lfloor f \times n \rfloor P_{vote}$ and
$\nu = \beta(n, N, f, q)P_{spread}$.

Thus,

$$\boldsymbol{p_1}(k) \approx \frac{\mu^{k_1}}{k_1!}e^{-\mu} \otimes \frac{\nu^{k_2}}{k_2!}e^{-\nu} \qquad (19)$$

The $\boldsymbol{p_0}$ can also be obtained by setting $f = 0$ in $P(k_2)$ (i.e., no object exists in an image),
and is given as

$$\boldsymbol{p_0}(k) = \binom{\beta(n, N, 0, q)}{k} P_{spread}^{k}(1 - P_{spread})^{\beta(n,N,0,q)-k}$$

which we approximate as

$$\boldsymbol{p_0}(k) \approx \frac{[\beta(n, N, 0, q)P_{spread}]^k}{k!}e^{-\beta(n,N,0,q)P_{spread}} \qquad (20)$$

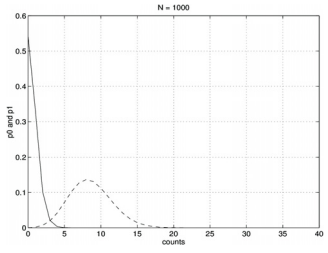
Given $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$, we can evaluate the performance of the GHough under varying conditions using the ROC curves.

6. Performance Evaluation of the GHough

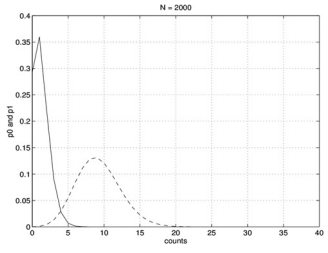
We plot the ROC curves for the GHough to see how it performs by varying the number of edge pixels (N), the amount of uncertainty in measuring edge position (ϵ_p), edge gradient direction(ϵ_a), and the degree of occlusion of the model (f). For the convenience of this evaluation, we set the quantization size q of the R-table to 5 degrees and assumed the model consists of $n = 20$ points and the maximum length of the displacement vector in the model was set to 0.25.

(1) Varying Number of Image Features : N

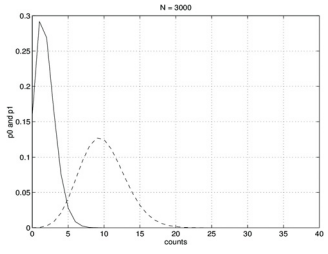
We evaluated the GHough by varying the number of image features(Edges) from 1,000 to 4,000 points in steps of 1,000. Fig. 7 (a) through (d) show both $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$ when N is 1000, 2000, 3000, and 4000, respectively. When N is 1000, $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$ are well separated. However, as N increases, we can see that the overlapping interval of both $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$ increases so that the false alarm probability also increases. Fig. 8(a) is a plot that shows the probability of false alarm (P_F) versus various thresholds for varying number of edge points(N). As we can see, for a given threshold, the P_F increases as N increases. When N is 1,000, setting a threshold such that the P_F is equal to 0.1 allows a P_D of 0.99. However, when N is equal to 4,000, the same threshold allows only a detection rate of 0.95. Fig. 8(b) shows the ROC curves for varying N.



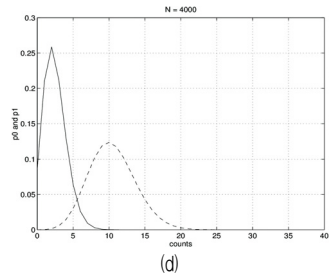
(a)



(b)

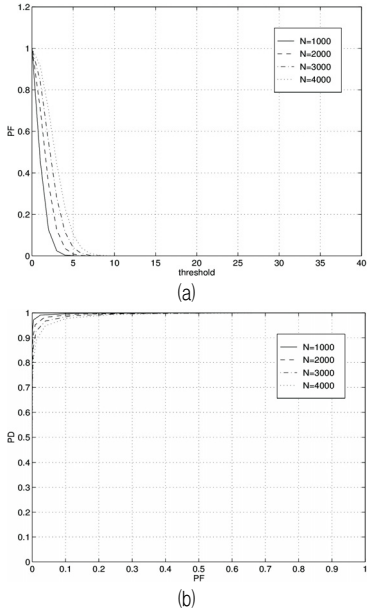


(c)



(d)

Fig. 7. (a), (b), (c) and (d) show p_0 and p_1 when N is 1,000, 2,000, 3,000, 4,000 points, respectively



(a) (b)

Fig. 8. (a) shows the false alarm probability for various threshold when N varies (b) shows the ROC curves when N is 1,000, 2,000, 3,000, 4,000 points, respectively

(2) Varying Degree of Positional Uncertainty :

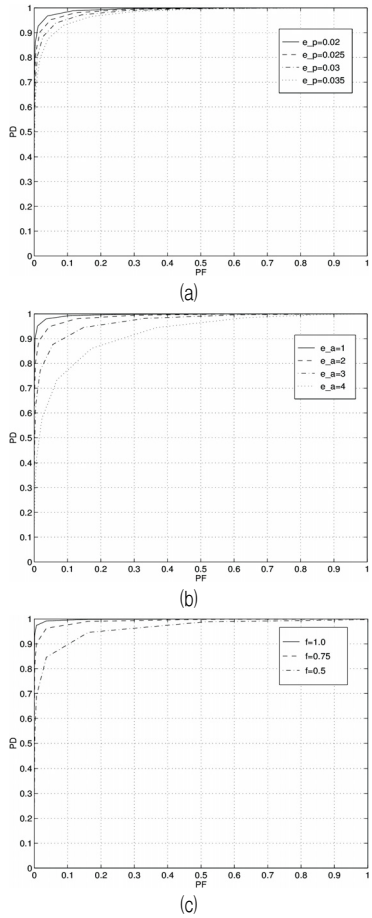
Fig. 9(a) shows the ROC curves as a function of varying positional uncertainty ϵ_p. With the false alarm probability of 0.1, the GHough can only maintain the detection probability of 0.98, but as the positional uncertainty increases, the detection probability falls down up to 0.93.

(3) Varying Degree of Angular uncertainty :

Fig. 9(b) shows the ROC curves as a function of varying degree of angular uncertainty(ϵ_a). As we can see, the performance of the GHough degrades more rapidly compare to the case of positional uncertainty ϵ_p in Fig. 9(a). With the false alarm probability of 0.1, the GHough can only maintain the detection probability of 0.77 when there is angular uncertainty of 4 degrees.

(4) Varying Degree of Occlusion :

Fig. 9(c) shows the ROC curves as a function of varying degree of occlusion(f). In this case, the GHough performs better compare to the case of positional uncertainty ϵ_p , but the detection probability decreases below 0.9, which is very poor if the half of the model is occluded in a given image(f = 0.5).



(a) (b) (c)

Fig. 9. (a), (b) and (c) show the ROC curves for varying positional uncertainty ϵ_p, varying angular uncertainty ϵ_a and varying degree of occlusion f, respectively

7. Conclusion

In this paper, we proposed a framework for the evaluation of the GHough by viewing it as a detector in signal detection theory. We derived the two conditional probability functions(p_0 and p_1) of the peak(Cell counts) in the accumulator array of the GHough. Those functions were expressed by incorporating the uncertainties in measuring the position and the gradient angle of an edge, the number of edges, and the degree of occlusion of the model. Finally, both the false alarm probability(P_F), and the detection probability(P_D) were obtained to specify the Receiver Operating Characteristic(ROC) of the GHough under various noise level. Our proposed evaluation framework can also be used by other GHough variants to specify the performance *quantitatively*.

ACKNOWLEDGMENTS

This study was conducted by research funds from Gwangju University in 2017.

REFERENCES

[1] M. J. Lee. (2014). A Study on Convergence Development Direction of Gesture Recognition Game. *Journal of the Korea Convergence Society, 5(4),* 1-7.
DOI : 10.15207/JKCS.2014.5.4.001

[2] B. SODGEREL, Y. K. Kim & M. H. Kim. (2015). 8-Straight Line Directions Recognition Algorithm for Hand Gestures Using Coordinate Information. *Journal of digital Convergence, 13(9),* 259-267.
DOI : 10.14400/JDC.2015.13.9.259

[3] H. J. Moon, M. H. Lee & K. H. Jeong. (2015). Authentication Performance Optimization for Smart-phone based Multimodal Biometrics. *Journal of digital Convergence, 13(6),* 151-156.
DOI : 10.14400/JDC.2015.13.6.151

[4] S. K. Kang & S. H. Chun. (2016). Real-Time Object Tracking Algorithm based on Pattern Classification in Surveillance Networks. *Journal of digital Convergence, 14(2),* 183-190.
DOI : 10.14400/JDC.2016.14.2.183

[5] Y. K. Kim, J. G. Lim & M. H. Kim. (2016). Lip Reading Method Using CNN for Utterance Period Detection. *Journal of digital Convergence, 14(8),* 233-243.
DOI : 10.14400/JDC.2016.14.8.233

[6] H. P. VC. (1962). *US 3069654.* Washington, DC : U.S. Patent and Trademark Office.

[7] D. H. Ballard. (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition, 13(2),* 111-122.
DOI : 10.1016/0031-3203(81)90009-1

[8] J. Illingworth & J. Kittler. (1988). A survey of the Hough transform. Computer Vision. *Graphics and Image Processing, 44,* 87 - 116.
DOI : 10.1016/S0734-189X(88)80033-1

[9] P. Mukhopadhyay & B. Chaudhuri. (2015) A survey of Hough Transform. *Pattern Recognition, 48,* 993-1010.
DOI : 10.1016/j.patcog.2014.08.027

[10] S. Chiu, C. Wen, J. Lee, K. Lin & H. Chen. (2012). A Fast Randomized Generalized Hough Transform for Arbitrary Shape Detection. *International Journal of Innovative Computing, Information Control, 8(2),* 1103-1116.

[11] C. P. Chau & W. C. Siu. (2004). Adaptive Dual-Point Hough Transform for Object Recognition. *Computer Vision and Image Understanding, 96(1),* 1-16.
DOI : 10.1016/j.cviu.2004.04.005.

[12] P. Tipwai & S. Madarasmi. (2007). A modified generalized Hough transform for image search. *IEICE TRANSACTIONS on Information and Systems, 90(1),* 165-172.
DOI : 10.1093/ietisy/e90-1.1.165

[13] T. W. Anderson. (1984). *An Introduction to Multivariate Statistical Analysis.* USA : John Wiley & Sons.

[14] H. L. V. Trees. (2002). *Detection, Estimation, and Modulation Theory : Detection, Estimation, and Linear Modulation Theory.* USA : John Wiley & Sons.

[15] J. Y. Chang. (2014). A Selection of Threshold for the Generalized Hough Tranform : A Probabilistic Approach. *Journal of Electronics and Information Engineers of Korea, 51(1),* 161-171.
DOI : 10.5573/ieie.2014.51.1.161

[16] S. M. Ross. (2014). *Introduction to probability models.* USA : Academic press.
DOI : 10.1016/B978-0-12-407948-9.00012-8

저 자 소 개

장 지 영(Ji-Young Chang) [정회원]

- 1992년 2월 : Indiana University, 컴퓨터과학 석사
- 1995년 7월 : Indiana University, 컴퓨터과학 박사
- 1995년 9월 ~ 2007년 10월 : 삼성SDS(주), 수석 연구원
- 2008년 3월 ~ 2011년 2월 : 우송대학교 초빙 교수
- 2012년 4월 ~ 현재 : 광주대학교, 컴퓨터공학과 교수

<관심분야> : 컴퓨터 비전, Statistical Pattern Recognition, Medical Imaging