論文 97-22-8-10

Measure of Similarity by Toll Theory and Matching Using Fuzzy Relation Matrix －Focused on 3-Dimensional Images－

Dong Uk Cho*, Kil Sung Han*, Yong Hwan Cho** *Regular Members*

톨이론에 의한 유사도 계산과 퍼지 관계 행렬을 이용한 정합과정의 수행 －3차원 영상을 중심으로－

正會員 조 동 욱*, 한 길 성*, 조 용 환**

ABSTRACT

In this paper, we envisioned a multimedia object recognition system processing and combining information from all available sources, such as 2-D, 3-D, color and sound data. Out of the overall system, we proposed 3-D information extraction and object recognition methods. Firstly, surfaces are classified by z-gradient from the range data, surface features are extracted using the intersection of normal vectors. Also feature relationship such as intersection angle and distance is established between the surfaces. Secondly, recognition is accomplished by matching process which is important step in the image understanding systems. Matching process is very important procedures because of more general and more efficient method is needed in the field of multimedia system. Therefore, we focused the proposal of matching process and in this article, first of all, we deal with the matching process of the 3-D object. Similarity measures are calculated.

요 약

본 논문은 2차원 정보, 3차원 정보 그리고 그 밖의 각종 유용한 정보를 취합하여 이를 상호 보완하여 인식하는 멀티미디어 시스템 중의 한 부분으로 이중 3차원 정보를 추출하고 정합 하는 방법을 제안하고자 한다. 우선 거리 영상으로부터 z축 기울기를 이용하여 표면 분류를 행하고 법선 벡터들의 교점을 통해 각 표면들에 대한 특징을 추출한다. 또한 각 표면들로 부터 이루는 각이나 거리등과 같은 특징 관계를 설정한다. 이후 정합 과정을 통해 인식을 수행하게 되는데 정합 과정은 영상인식의 최종 단계로 대단히 중요한 과정중의 하나가 된다. 왜냐하면 멀티미디어 시스템은 각종 정보를 취합하여 정합 과정을 수행해야 하기 때문에 취합한 모든 정보를 보다 보편적이고

*서원대학교 정보통신공학과
**충북대학교 컴퓨터공학과

論文番號:97193-0609
接受日字:1997年 6月 9日

효율적으로 정합 하는 방법론의 제시가 중요하기 때문이다. 따라서 본 논문에서는 정합 과정 수행에 필요한 보편적 방법론의 제안에 초점을 맞추고자 하며 이중 우선 3차원 물체의 정합 과정에 대해 다루고자 한다. 이를 위해 톨이론을 적용하여 유사도를 측정하며 이를 퍼지 관계 행렬을 구성하여 인식을 수행한다. 최종적으로 실험에 의해 본 논문의 유용성을 입증하고자 한다.

Ⅰ. Introduction

The 2-dimensional image processing, which are based on the gray levels of an image, have found a wide real-life application area such as recognition of characters, car number plates, finger prints and human faces [1]~[4]. However, the images with less contrast such as curved or occluded surface of 3-D objects reveal little or sometimes confusing results after 2-D image processing. This limitation can easily be overcome by 3-D image processing which is based on the range data of an object. Likewise, if 3-D image processing does not differentiate the object of the similar shape, a 2-D image information such as characters on the surface can be a unique determinant for the recognition of the objects. Therefore, for a better object recognition, it is necessary to take a combined approach of complementing 2-D and 3-D image processing. Meantime, to get optimal result, as long as any other information, such as color or sound is available, we should not restrict ourselves to the traditional 2-D and 3-D processing. By constructing a multimedia object recognition system, we can readily distinguish the objects such as the berries of a similar shape by color and the species of same animal family like canine family by their sound. Out of this unified multimedia object recognition system this article describes 3-D image processing and recognition. Especially, for the multimedia system, developing a more general and effective matching algorithm is particularly important. In this paper, in this context, it is focused to propose a matching algorithm, which will be verified in the experimentation applying 3-D images.

Usually, the similarity is evaluated to perform the matching process. Existing methods [5], [6] normalize the feature vectors for an input pattern, and evaluate their distance to estimate the simarity. However, in theses methods, feature vectors must be normalized, which leads to an increase in processing complexity and processing time. For this, in this paper, we propose to apply the toll theory, as a distance function, to evaluate the similarity: the similarity is evaluated as a fractional value between 0 and 1, which is performed without applying a normalization process. Also, as the final matching process, the fuzzy relation matrix is formed using the evaluated similarity.

This paper consists of 5 chapters: Feature extraction for 3-D images is presented in Chapter Ⅱ. In Chapter Ⅲ, after a survey for existing methods for similarity evaluation, a similarity evaluation method is proposed. Also how to form the fuzzy relation matrix is described. In Chapter Ⅳ, the experimentation for the proposed methods is performed. Finally in Chapter Ⅴ, the observation for the experiments is presented as the conclusion.

Ⅱ. Extraction of 3-Dimensional Information

3-D information needed for the next phase processing consists of shape information, surface geometric features and relational vectors[7]. To extract shape information, we define four surface types, i.e., sphere, cylinder, cone and plane, which was based on the fact that about 85% of human made objects consist of these primitives. We humans feel most comfortable with the objects that have a circular cross section, it is to be noted that three types, i.e., sphere, cylinder and cone, among the four have a circular cross section. The change of range value is defined as z-gradient, which is a function of two parameters, i.e., the magnitude and the direction of the change on the Z axis within 2*2 mask. This Z-gradient becomes the basis for surface classification in this study. First, the direction information on the boundary between the background and the object in an input range is acquired, which is illustrated as in Fig 1. The cylinder, as an example in the figure, has a value ⟨4, 4, 0⟩, which means 4 directions(represented as arrows) have ⟨4⟩ cases of 180 degree relation, ⟨4⟩ cases of 90 degree relation, and ⟨0⟩ case of 45 degree relation. The same rule is applied to the case of the inner direction information as in Fig. 2. Also the magnitude of Z-gradient forms the equi-gradient contour, which is used to perform surface classification according to the classification characteristics. For example, uniform distribution is formed in case of the cylinder, triangular distribution for the cone, or exponential distribution for the case of the sphere. Surface classification is finally performed upon all the classification results such as from the boundary direction component, the inner direction comport and the magnitude distribution characteristics.

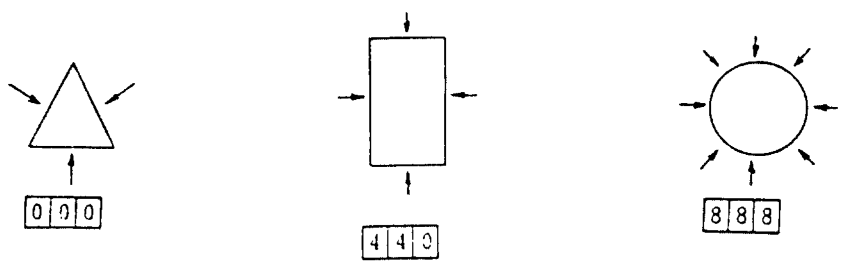
The geometric features about surfaces are obtained from the intersection points out of the normal vector for each surface patch, *i.e.* if we define a mask patch M_P for any chosen point $P(x_0, y_0, z_0)$, the normal vector Vp for P is obtained by

$$V_P=(x_0+At,\ y_0+Bt,\ z_0+Ct) \tag{1}$$

Likewise the normal vector for another point $Q(x_1, y_1, z_1)$ is obtained by

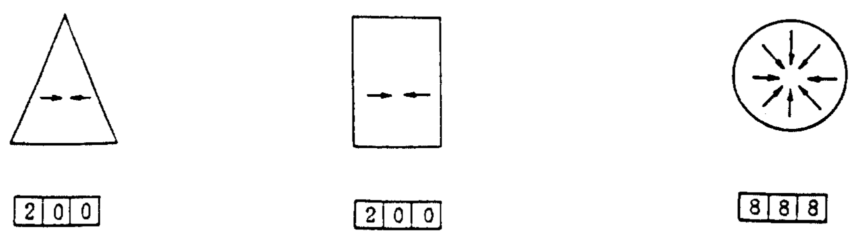
$$V_Q=(x_1+A\,S,\ y_1+B\,S,\ z_1+C\,S) \tag{2}$$

Where t of expression (1) and S of expression (2) are parameter values.



(a) cone (b) cylinder (c) sphere

Fig. 2. Examples of Inner Direction Component



(a) cone (b) cylinder (c) sphere

Fig. 1 Examples of Boundary Direction Component

Now, in case of a sphere , the intersection formed by V_P and V_Q is the center point of a sphere. For a cone and cylinder, the intersection becomes an axis point. The above process is performed on all the mask patches and the results are accumulated, which yields the geometric features. Meantime the relation vectors between surface regions are shown in Table 1.

Table 1. Relation vectors between surface regions

Combinations of features	Extracted features
point , point	distance
line , line	intersection angle
point , line	distance
point , plane	distance
plane , plane	intersection angle
plane , line	intersection angle

Ⅲ. 3-D Object Recognition Using Fuzzy Relation and Similarity Measure Applying to Toll Theory

Utilizing 3-D information extracted from the previous section, the next phase is to perform the normalization operation and undertake the matching process with arising ambiguity in consideration. For this, we apply the fuzzy relation matrix and toll theories.

3.1. Brief Survey of Existing Similarity Measuring Methods [5], [6]

Firstly, let us look into existing similarity measuring methods. Usually, for the similarity measurement, metric-based methods using the distance between feature vectors have been being applied. Since image are represented by a n-dimensional feature vector, and as such represented as points in an n-dimensional space. 'Distance' between a pair of such points represents the dissimilarity between those images. The farther the points are from each other, more dissimilar the images are and vice versa. The distance function defined below can be mapped into the range [0, 1] to represent dissimilarity

$$d_r(x, y) = \left[\sum_{i=1}^{n} |x_i - y_i|^r \right]^{\frac{1}{r}}, \ r \geq 1 \tag{3}$$

This measure is based on a class of distance function known as Minkowski r-metric, which where x and y are two points in an n-dimensional feature space with components, (x_i, y_i), $i = 1, 2, \cdots, n$. For $r = 2$, it is the Euclidean metric, for $r = 1$, it is the city-block metric, and for $r = \infty$, it is the dominance metric. However, to estimate these similarity measures, feature vectors of x and y must be normalized, which leads to such problems as an increase in processing complexity and processing time. Also, from the geometrical view point, the similarity measure for a distance of 0 between two vectors should be 1, whereas the one for a distance of 1 should be 0. In this paper, hence, the toll theory is applied to the similarity measurement, and the fuzzy relation matrix is formed using similarity measures to perform the matching process.

3.2. Measurement of similarity by Fuzzy Relation Matrix

If we define the membership degree of a fuzzy set A for an element x as $\mu_A(x)$, then the fuzzy relation R is to represent the relationship between set A and B and $\mu_A(x, y)$ for $x \in A$ and $y \in B$ is expressed as

$$\mu_R : A \times B \rightarrow [0, 1] \tag{4}$$

Here, $\mu_R(x, y)$ may by more appropriately interpreted as the strength of a relationship than a membership degree. If $\mu_R(x, y) \geq \mu_R(x', y)$ implies(x, y) has stronger relation than (x′, y′), in the discrete case the fuzzy relationship may be represented as a fuzzy relation matrix. To establish a fuzzy relation R for our object matching process, two sets M and F are define as follows:

$$\begin{aligned} M &= \{M_1, M_2, \cdots, M_n\} \\ F &= \{F_1, F_2, \cdots, F_n\} \end{aligned} \tag{5}$$

Here, the set M consists of n model objects and the set F is of g feature vectors. As an example, the set F may contain such feature as area, shape, intersection angle and distance between surface regions. Now, the similarity measures between the feature vectors of the input images and those of the model objects prestored are to be computed and the fuzzy relation matrix R is to be obtained.

An element $r_R(F_i, M_j)$ in the matrix R represents the strength of relationship between the i-th feature vector of the input object and the i-th feature vector of j-th model object, which is in fact the similarity measure between the extracted input feature vector F_i and the i-th feature of the model object M_j. Therefore the relation matrix R expresses the relationship between all input feature vectors and all model objects as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \vdots & & & \\ \vdots & & & \\ a_{j1} & a_{j2} & \cdots\cdots & a_{jn} \end{bmatrix} \quad (6)$$

In the above matrix the columns represent model object set M and the rows depict the feature vector set F of the expression(4). Each element $0 \le a_{ij} \le 1$ means the similarity between the i-th element of the feature vector for the input object, or the i-th evidence vector and the j-th feature vector of the j-th model object.

3.3. Computation of Similarity Measures

To form a fuzzy relation matrix R the required similarity measure S(x) has to be computed.

3.3.1 Similarity Measures from Area and Shape

The normalized value [0, 1] of area or the area ratio is combined with shape information to yield S(x) as

$$S(x) = 1 - \sum_{i=1}^{n} \sum_{j=1}^{m} |A_i - A_j| \quad (7)$$

In the above expression, n and m represent the number of input and model object surface regions respectively, and A_i and A_j are the normalized area ratio values of the surface regions. If two surface regions have completely different shape information, then $A_j = 0$.

3.2 Similarity Measures from Intersection Angle and Distance

We define the similarity measure S(x) for the intersection angle and distance information as

$$S(x) = \frac{1}{1 + \lambda d(x)} \quad (8)$$

where d(x) represent the degree of difference or distance between input feature vector and model feature vector. If d(x) = 0, then there is no similarity between input and model objects and if d(x) = 0, then two objects are the same. To compute d(x) we utilize the toll theory proposed by H. Prade in 1991 IFSA[8]. A toll set is to represent its set membership as a cost which is between 0~∞. The cost of zero means free membership and the infinity means that the membership is forbidden. Now, d(x) is defined as

$$\begin{aligned} d(x) &= -\log\left[-\frac{1}{10}(x-10)\right], \ 0 \le x \le 10 \\ &= \infty \qquad\qquad\qquad\qquad\quad , x > 10 \end{aligned} \quad (9)$$

The above d(x) is a toll membership function and we define the normalized similarity measure for

$$S'(X) = \frac{\sum_{i=1}^{G} S(x_i)}{G} \quad (10)$$

where G is the number of information about distance and intersection angle.

3.4. Matching and Recognition

In the final recognition phase, we perform matching process utilizing the fuzzy relation matrix R. At this time we may give some weight to each feature vector according to their contribution toward recognition, i.e, if we define the weight set, then

$$W = \{ W_1, W_2, \cdots\cdots, W_g \} \tag{11}$$

where $0 \leq W_i \leq 1$. Now, the confidence level in the normalized form is computed as

$$V_i = \frac{W \cdot R}{\sum_{i=1}^{g} W_i} \tag{12}$$

where·is the product-sum operation. The final recognition is done by computing the following expression.

$$V' = \max [V_1, V_2, \cdots\cdots, V_n] \tag{13}$$

Which means the choice of the best match among n model objects.

Ⅳ. Experimental Results and Observations

All the experiments are done on an IBM-PC using the C programming language. Figures 3 and 4 contain the model range images for a cup, a dumbbell, a toy and tumbling doll as shown. The range images of the input objects to be recognized include a tumbling doll and a dumbbell given in Fig. 5. Tables 2 through 5 contain the surface region feature values for the model objects of a cup, a dumbbell, a tumbling doll and a toy respectively, and tables 6 and 7 are those for the input objects of a tumbling doll and a dumbbell respectively. Table 8 and 9 contain all the feature vectors, respectively, for the model objects and for the input objects to be recognized. In table 10 the values [0, 0.79429, 0.03, 0.2] represent the confidence level V computed from the fuzzy relation matrix formed for the first input object and the model objects and [0, 0.43, 0.165, 0.219] is the confidence level value for the second input object and the model objects. The final object recognition is accomplished by taking the maximum value item from the confidence level values. The results of the process show that the first input object is a dumbbell and the second input object is a tumbling doll. At this time we applied the same feature vector weight Wi of 1 for both. As explained, the proposed matching process has been performed utilizing the fuzzy relation matrix formed with the computed similarity between the feature vectors of input and model objects.

Utilization of fuzzy and toll theories in the computation of the similarity measure, S(x) and the required distance d(x) has given much needed flexibility and efficiency in the object recognition phase. Currently we are working on the other parts of the envisioned multimedia object recognition system and the phase of the 2-D images data preprocessing, edge detection from 2-D images are implemented and finished experiments. We expect the progress soon on other phases such as in feature extraction from other multimedia information and the unifying all the processed multimedia information to yield meaningful object recognition, in which some of our preliminary results are very encouraging.

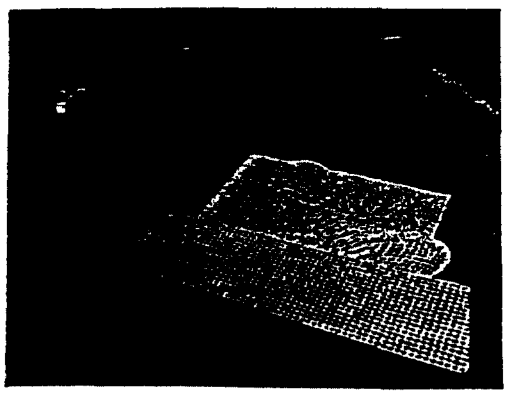


Fig. 3. Range Image (for model)

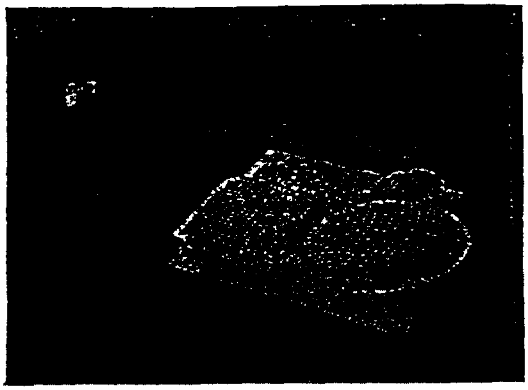


Fig. 4. Range Image (for model)

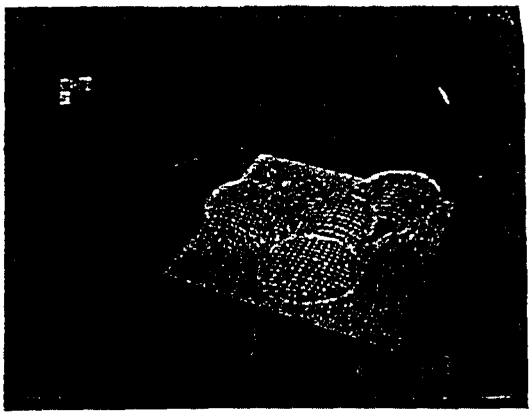


Fig. 5. Range Image (for input)

Table 2. Feature Value of the Surface Region(for cup)

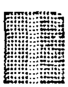
Surface classification	Feature Values of the Surface Region
	Direction Information Boundary = 220 Internal = 200 Axis Equation $A(s)=P_1(67,146,529)+s*(0.0000,0.9998,$ $-0.0208)*\|P_2-P_1\|$ end point $P_2(67,194,528)$ Area Ratio 100

Table 3. Feature Values of the Surface Region (for dumb-bell)

Surface Classification		Feature Values of the Surface Region
S_1 Surface		Direction Information Boundary 445 Internal 888 Shape Equation $(x-107.6)^2+(y-227.6)^2+(z-530.0)^2=(39.1)^2$ Area Ratio 40
S_2 Surface		Direction Information Boundary 440 Internal 200 Axis Equation $A(s)=P_1(183,188,560)+S*(0.5215,$ $0.4212,0.7421)*\|P_2-P_1\|$ end point $P_2(157,209,526)$ Area Ratio 12
S_3 Surface		Direction Information Boundary=441 Internal=888 Sphere Equation $(x-262.8)^2+(y-135.6)^2+(z-541.2)^2=(41.9)^2$ Area Ratio 48

Table 4. Feature Value of the Surface Region (for tumbling doll)

Surface Classification		Feature Values of the Surface Region
S_1 Surface		Direction Information Boundary=888 Internal 888 Shape Equation $(x-214.7)^2+(y-234.8)^2+(z-398.9)^2=(39.9)^2$ Area Ratio 25
S_2 Surface		Direction Information Boundary=888 Internal 888 Sphere Equation $(x-215.2)^2+(y-121.2)^2+(z-397.9)^2=(69.9)^2$ Area Ratio 75

Table 5. Feature Value of the Surface Region (for toy)

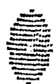
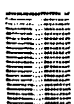
Surface Classification		Feature Values of the Surface Region
S_1 Surface		Direction Information Boundary 888 Internal=888 Shape Equation $(x-59.8)^2+(y-188.2)^2+(z-399.6)^2=(34.7)^2$ Area Ratio 33
S_2 Surface		Direction Information Boundary=440 Internal=200 Axis Equation $A(s)=P_1(59,48,398)+S*(-0.1047,$ $0.0995,0.0000)*\|P_2-P_1\|$ end point $P_2(49,143,398)$ Area Ratio 67

Table 6. feature Value of the Surface Region (for input)

Surface Classification		Feature Values of the Surface Region
S_1 Surface		Direction Information Boundary=888 Internal=888 Shape Equation $(x-159.8)^2+(y-164.8)^2+(z-374.9)^2=(35.1)^2$ Area Ratio 40
S_2 Surface		Direction Information Boundary=888 Internal=888 Sphere Equation $(x-160.1)^2+(y-79.8)^2+(z-374.9)^2=(43.2)^2$ Area Ratio 60

Table 7. Feature Value of the surface Region (for input)

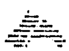
Surface Classification		Feature Values of the Surface Region
S_1 Surface		Direction Information Boundary=666 Internal=888 Shape Equation $(x-209.2)^2+(y-230.1)^2+(z-399.6)^2=(49.9)^2$ Area Ratio 40
S_2 Surface		Direction Information Boundary=000 Internal=200 Axis Equation $A(s)=P_1(124,159,399)+S*(0.6689, 0.7423, 0.0106)*\|P_2-P_1\|$ end point $P_2(187,229,398)$ Area Ratio 20
S_3 Surface		Direction Information Boundary=455 Internal=888 Sphere Equation $(x-59.9)^2+(y-138.9)^2+(z-398.5)^2=(49.8)^2$ Area Ratio 40

Table 8. Feature Extraction (for model)

object name	shape & its area ratio	features between surface regions
cup	1.0	
dumbbell	sp (0.4) cv (0.12) sp (0.48)	40.327 45.231 174.734
toy	sp (0.33) cv (0.67)	49.393
tumbling doll	sp (0.25) sp (0.75)	113.605

Table 9. Feature Extraction (for input)

object name	shape & its area ratio	features between surface regions
input object 1	sp (0.4) cv (0.2) sp (0.4)	38.857 44.352 170.932
input object 2	sp (0.4) sp (0.6)	07.043

Table 10. Results of Recognition

[0, 0.79429, 0.03, 0.2]
input object 1 is a dumbbell

[0, 0.43, 0.165, 0.5219]
input object 2 is a tumbling doll

Ⅴ. Conclusion

In this paper, we propose a method for 3-D information extraction and matching, which is applied to a sub-part of a multi-media system. Specifically, we propose an algorithm for the similarity measurement and matching process which are considered to be some of the most important processes for the multimedia system as it needs to integrate various kinds of information. The efficiency and flexibility of the proposed method are verified by the experiment applying 3-D objects. The distance function which is needed for existing normalization methods is solved by applying the toll theory. Also the matching process is performed by using the fuzzy relation matrix. We expect more experiments on the proposed methods so that they can be fine tuned to the level of real life utility. We also expect much progress soon in all areas of the entire multimedia object recognition system ever closer to the commerce level. Finally, we are thankful Ho Chul Jeon for his fine data processing job.

Reference

1. D.U.Cho et al, "Recognition of Automobile Type

and Extraction of Car Number Plate by Image Processing," 1st Korea-Japan Joint Conference on Computer Vision, 1991.
2. R.Chellappa et al, "Human and Machine Recognition of Faces: A Survey," Proceedings of IEEE, Vol.83, No. 5, 1995.
3. Byung Uk Choi, "Algorithm for Recognizing Handwritten Hanguel Using the Improved Neocognitron," Journal of the KITE, Vol. 32-B, No.2, 1995.
4. Andrew et al, "Automated Fingerprint Recognition Using Structural Matching," Pattern Reconition, Vol.23, No.8, 1990.
5. S. Shafer and D. Rogers, "Similarity and Distance Measures for Celluar Manufacturing: Part i. A Survey," International Journal of Production Research, Vol. 31, No. 5, 1993.
6. S. Shafer and D. Rogers, "Similarity and Distance Measure for Cellular Manufacturing: Part ii. An extension and Comparison," International Journal of Production Research, Vol. 31, No. 6, 1993.
7. D.U.Cho, "Fuzzy-Set Based Feature Extraction for Object of Various Shapes and Appearances," IEEE Int'l Conference on Image Processing, 1996.
8. D.Dubois and H.Prade, "Toll Sets," Proc. IFSA'91, Brussels, Artificial Intelliqence, 1991.

조 동 욱(Dong Uk Cho) 정회원
1983년 2월 : 한양대학교 공대 전자공학과(공학사)
1985년 8월 : 한양대학교 대학원 전자공학과(공학석사)
1985년 9월~1989년 2월 : 한양대학교 대학원 전자통신공학과(공학 박사)
1982년~1983년 : (주)신도리코 기술연구소 연구원
1986년~1993년 : 한양대 공대, 충북대 대학원 강사
1995년~현재 : 한국통신학회 충북지부 임원
1996년 10월 : 동양일보 선정 "충청을 이끌 50인"에 선정
1996년 11월 : 한국통신학회, 한국통신학회 충북지부 학술상 수상
1997년 1월 : 중부매일 선정 "21세기를 이끌 주역"에 선정
1991년 3월~현재 : 서원대학교 정보통신공학과 교수
1997년 3월 : 서원대학교 개교 기념 공로상 수상
1997년 7월 : 한국통신학회 공로상 수상
※주관심 분야 : 영상 처리 및 인식, 퍼지 이론 및 톨 이론

한 길 성(Kil Sung Han) 정회원
1972년 2월 : 공군사관학교(이학사)
1977년 2월 : 서울대학교 전자공학과(공학사)
1981년 7월 : 서울대학교 대학원 전자공학과(공학석사)
1984년 7월~1987년 6월 : Polytechnic Institute of New-York(박사수료)
1996년 7월 : 충북대 대학원 컴퓨터공학과(공학박사)
1977년 3월~1997년 2월 : 공군사관학교 전자공학과 교수
1997년 3월~현재 : 서원대학교 정보통신공학과 교수
※주관심분야 : 디지탈통신, 정보통신, ATM(영상통신)

조 용 환(Yong Hwan Cho) 정회원
1989년 2월 : 고려대학교 대학원 통계학과 (이학박사)
1978년 3월~1981년 10월 : 한국전자통신연구소 통신망 계획실 선임연구원
1982년 6월~현재 : 충북대학교 컴퓨터공학과 교수
1991년 6월~1992년 8월 : 미국 Nevada주립대학교 (UNLV), 전기 및 컴퓨터공학과 교환교수
1993년 11월~현재 : 한국통신학회 이사, 충북지부장
1996년 3월~1996년 6월 : APII Test-bed Forum운영위원회 홍보/출판분과 위원장
1997년 1월~1997년 7월 : 정보통신부 정보통신 번호체계 개선 전담반 위원
※주관심분야 : 트래픽 엔지니어링, ATM망 트래픽 제어, 번호 계획, 영상 통신 및 이해