Regular paper

Journal of the KIECS. pp. 923-932, vol. 12, no. 5, Oct. 31 2017, t. 85, pISSN 1975-8170 | eISSN 2288-2189
http://dx.doi.org/10.13067/JKIECS.2017.12.5.923

의료정보 보호를 위해 얼굴인식에 필요한 효과적인 시선 검출

김숙일* · 석경휴**

Effective Eye Detection for Face Recognition to Protect Medical Information

Suk-Il Kim* · Gyeong-Hyu Seok**

요 약

본 논문에서는 기존의 문제점인 얼굴 움직임이 있을 시 시선 식별이 어려운 점과 사용자에 따른 교정작업이 필요하다는 점을 해결하고자 새로운 시선 식별 시스템과 얼굴인식에 필요한 GRNN(: Generalized Regression Neural Network) 알고리즘을 제안한다. Kalman필터를 사용하여 현재 머리의 위치정보를 이용하여 미래위치를 추정하였고 얼굴의 진위 여부를 판단하기 위해서 얼굴의 특징요소를 구조적 정보와 비교적 처리시간이 빠른 수평, 수직 히스토그램 분석법을 이용하여 얼굴의 요소를 검출한다. 그리고 적외선 조명기를 구성하여 밝은 동공효과를 얻어 동공을 실시간으로 검출, 추적하였고 동공-글린트 벡터를 추출하여 의료정보 보호에 도움을 주고자 한다.

ABSTRACT

In this paper, we propose a GRNN(: Generalized Regression Neural Network) algorithms for new eyes and face recognition identification system to solve the points that need corrective action in accordance with the existing problems of facial movements gaze upon it difficult to identify the user and . Using a Kalman filter structural information elements of a face feature to determine the authenticity of the face was estimated future location using the location information of the current head and the treatment time is relatively fast horizontal and vertical elements of the face using a histogram analysis the detected. And the light obtained by configuring the infrared illuminator pupil effects in real-time detection of the pupil, the pupil tracking was to extract the text print vector. The abstract is to be in fully-justified italicized text as it is here, below the author information.

키워드

Bayesian, Neural Networks, Kalman Filter, Real-Time
베지안, 신경망, 칼만 필터, 실시간

Ⅰ. Introduction

Existing systems use the identification of the line of sight of the face of the movement and eye movement in order to identify the eye. When using only the motion of the face, there is determined the position of the eye based on the location of the face has the drawback can not detect the subtle

* 동강대학교 노인복지과(kim5725@hanmail.net)
** 교신저자 : 청암대학교 병원의료정보과
· 접 수 일 : 2017. 06. 07
· 수정완료일 : 2017. 08. 13
· 게재확정일 : 2017. 10. 18
· Received : Jun 07, 2017, Revised : Aug 13, 2017, Accepted : Oct 18, 2017
· Corresponding Author : Gyeong-Hyu Seok
Dept. of Hospital Medical Information, Cheungam University,
Email : dol27@naver.com

changes in gaze. The largest common gaze estimation method based on the research to date is that the pupil based on the relative position between the pupil and the cornea article lint [1-3]. Another problem with eyes and gaze identification systems is that the corrective action necessary with respect to each user. In order to overcome this limitation, the model of the eye [4] proposes a new gaze tracking technology, and two low viscosity, only correction but simplifies the gaze calibration procedure to that required, smooth operation when relatively movement is less by was, it proposes a precise geometric eye model for each user. Also. Using the two cameras and the geometric characteristics of the eye and image [5] it is possible to eliminate completely the radial calibration procedure. In this paper, the pupils were tracked in consecutive frames by the Kalman filter algorithm, an object tracking algorithm After the detection of the pupil. 1 to verify that you have correctly identified the primary areas adjacent to the identification result has only enhanced the real-time recognition and identification gaze to re-identify[12-13].

II. Features Using Bayesian statistical methods Network

2.1 Bayesian statistical methods Network Theory

Bayesian statistical methods networks to learn from one set of data, each node representing each feature of the set of data and each of (arc) has features and represent the dependencies between the probability of the target node based on the Bayesian network in this study typically predictable[5-7]. The Bayesian statistical methods given network structure consisting of the connection line between each node and the node of the conditional probability, the probability of each node, there is the learning from data can be expressed as a formula the Bayesian statistical methods network as in equation (1).

$$p(x) = \prod_{i=1}^{n} p(x_i|\pi_i) \quad (1)$$

Variables A, B, C, D, E, F, and the respective parameters are to have a high value of yes and no, conditional probability Each column of Table 1-1 shows the possible values assigned to the two variables A and B each line represents the probability to have a condition that the variable D yes or no.

표 1. 조건 확률표

Table 1. Conditional probability Table

	A(y), B(y)	A(y),~B(n)	~A(n), B(y)	~A(n), ~B(n)
D(y)	0.4	0.1	0.8	0.2
~D(y)	0.6	0.9	0.2	0.8

Bayesian statistical methods uses a Bayesian network theory, such as the calculation formula (2) probability conditions.

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \quad (2)$$

When given as the training data set D, P | and (h D) is the posterior probability of h, in order to obtain this P | a (D h), P (h), P (D) should be sought first. Additionally, the maximum posterior probability is the same as equation (3).

$$h_{MAP} \equiv \arg\max_{h \in H} P(D|h)P(h) \quad (3)$$

Assume that in the absence of both the probability of the maximum likelihood and prior probability of hypotheses and then omitted Ph) value is expressed as equation (4).

$$h_{ML} \equiv \arg\max_{h \in H} P(D|h) \quad (4)$$

2.2 Face, eyes with eye identification

Because of gradation between the pupil and the iris the difference is very small, it is difficult to have a low accuracy accurate extraction disadvantages[8-9]. Therefore, the movement of the pupil is hard to overcome this disadvantage by extraction tracking using special light such as infrared. Research to obtain the amount of three-dimensional movement of the face is used a method for estimating a three-dimensional rotation amount, and the amount of movement of the face from the motion of the face feature point projected on the two-dimensional camera image as shown in Figure 1[9-11].

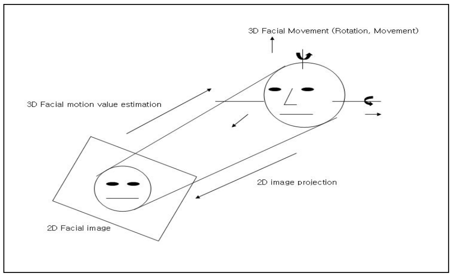

그림 1. 얼굴의 3차원 움직임량 추정

Fig. 1 Three-dimensional movement of the face amount of the estimated

2.3 Face tracking feature

Sirovich and Kirby [10] introduced the method applied by the KL transformation for the beardless face image expressing the face. Hallinan [11-12] was used as the template matching to detect the eye image in Fig. In this paper, the detection head area, it is determined whether or not the proper conditions to extract a face candidate region if it is determined. In addition, after setting the range for the present position of the mouth by using the structural information of the face is obtained a vertical and horizontal histograms for this zone. Due to the nature of the input vertical histogram is then available in only a portion. The sudden changes in the threshold area is more than the amount of the input end position. Hair region has a characteristic that changes in brightness between neighboring pixels ingredient than the face region and can be expressed as a variance, such as formula (5). If the value is more than specified threshold indicates a hair region.

$$V(x,y) = \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1}|Y(x,y) - Y(x+i,y+i)| \tag{5}$$

Image showing the hair region is a hair area 255, and the other is represented by zero. It is possible to obtain a substantial facial area image that is not affected by the hair and the hair image obtained by conventional methods face the logical relationship of the equation (6).

$$f_{RF}(x,y) = f_F(x,y) \cap \overline{f_H(x,y)} \tag{6}$$

2.4 Artificial neural networks for face recognition operation

This research studies the facial recognition algorithm to map GRNN learning. GRNN (: Generalized Regression Neural Network) may leave the base [13] in the observation probability density function of the data to be first used by Specht. Consists of the GRNN is composed of four layers. Type layer (input layer) serves to distribute the input pattern to the neurons in each hidden layer (Hidden Layer) is connected both to the second layer. The most widely used of the neural network BP (: Back Propagation) because it is connected to the neurons (Neuron) for the axonal synchronization (Axon). The input signal input to the neurons becomes combined after multiplied by the weight associated with each input signal value, the sum is represented as an output value put in a non-linear activation function. In this case BP should use the sigmoid (Sigmoid) function such as equation (7).

$$f(x) = \frac{1}{(1-e^{-x)}} \tag{7}$$

The calculated output value is used as input data for the next layer or the final output value according to the position of the neurons. Artificial neural networks are thereby varying the connection weights from the square of the error between the measured value and calculated value that is actually required in the artificial neural network sikineunde learning using learning data. Subtracting the square of the input layer elements from the training data has shown that each hidden layer neuron or an absolute value. Input function of the j-th hidden layer is equal to the equation (8).

$$I_j = \sum_{i=1}^{n} |W_{ij} - X_{ij}| \text{ 또는 } I_j = \sum_{i=1}^{n} (W_{ij} - X_{ij})^2 \quad (8)$$

xi is the weight between the input signal and wij is beonjae i type layer and the j-th hidden layer neurons, and n represents the number of neurons in the input layer. This calculated input value is given over to the non-linear activation function having the form of an exponential function such as equation (9).

$$f(I_i) = \exp\left(\frac{-I_j}{2\sigma^2}\right) \quad (9)$$

Hidden layer the output value of the function calculated on the active layer is passed to the summation, the combined layer by calculating the results by executing the integration of equation (10) and sends to the output layer.

$$Y(X) = \frac{\sum_{i=1}^{n} Y' f(I_i)}{\sum_{i=1}^{n} f(I_i)} \quad (10)$$

Y 'corresponds to the result of the learning data by the output value having the respective hidden layer neurons. Summarizing the above procedure as follows. GRNN is input to the network in each of the learning phase are used to calculate the response to the new input value. And it places the Gaussian kernel function on each learning stage. The result value calculated by the input value is calculated by placing a weight in the resulting average value of the learning phase, the weights are related to the learning phase with the calculated value. Great advantage compared to the back propagation GRNN is having the learning data is relatively less that required is the point. If wave reversal learning data of about 1% in order to obtain similar accuracy have to [13], and the training data is large, it can be used to populate a similar data.

2.5 Probabilistic Graphical Models

When modeling a biological system and an entity associated with the gene expression level system is treated as a random variable of the probabilistic model (4). Random variables include the expression level of the gene, such as shown in the model, such as the hidden characteristics and gene array experiments.

Figures 2 and probability using a hierarchical configuration properties of the network graph to the probability recognition in molecular biology, such as graphs and network discovery proves graph to stochastic signal path that demonstrate the biological signaling pathways and the number of incoming edges of each node as because it has an outgoing edge appears about five times the proven navigation in the actual network.

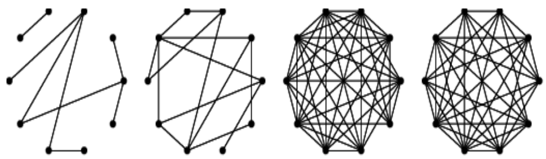

그림 2. 네트워크 탐색 그래프
Fig. 2 Network search graph

Probabilistic graph in Figure 2 is capable of complex objects is so extended in a multi-tiered network as evidenced by the probability search object representation because it can deal with all of the stochastic surface and the surface of the structural object. The lower layer, it is possible to demonstrate the expression of the underlying information, and the upper layer using the learned model searching goes up to the upper layer by the probabilistic inference information from the network which can be obtained by a combination of lower layers of the object.

The so achieved by the characteristics of the stochastic search proved graphs and information is stochastic method such as expression navigation recognition so made the result output from the stochastic graph to model a biological system with a set of derived [1.6] edge network node probabilistic network graph is defined as follows.

Definitions 1. An network graph over $A_V \cup A_E$is a 4-tuple $G=(V(B),E(B),\phi,\epsilon)$ such that

- □ $V(B)$is a finite, nonempty set of vertices;
- □ $E(B) \subset V(B) \times V(B)$is a set of ordered pairs of distinct elements in $V(B)$,called edges;
- □ $A_V$is a finite, nonempty set of vertex labels (primitive descriptions);
- □ $A_E$is a set of edge labels (relation descriptions);
- □ $\phi : V(B) \rightarrow A_V$ is a function, called a vertex interpreter;
- □ $\epsilon : E(B) \rightarrow A_E$ is a function, called an edge interpreter.

The network graph using the above definition consists of the vertices and edges in terms probability parameter a probabilistic graph condition variable in a certain way that is used to represent the characteristics of the recognition target network itself, and connect the stochastic component values at each vertex and edge thereby to generate a probabilistic network graph.

Figure 3 is a probabilistic cross-path models because two basic probabilities to 32, using the coupling between the distribution because it induced a network node of the cross-path probabilistically cross paths with a probability distribution combined with gene combination into a circle probability distribution by configuring the parameters of the edge node and there exists a probability distribution for the biological cross-path direction of the edge in the stochastic model. Node of the two edge portions is connected, there is a probability distribution network chain between cells in the coupling between the two edges.

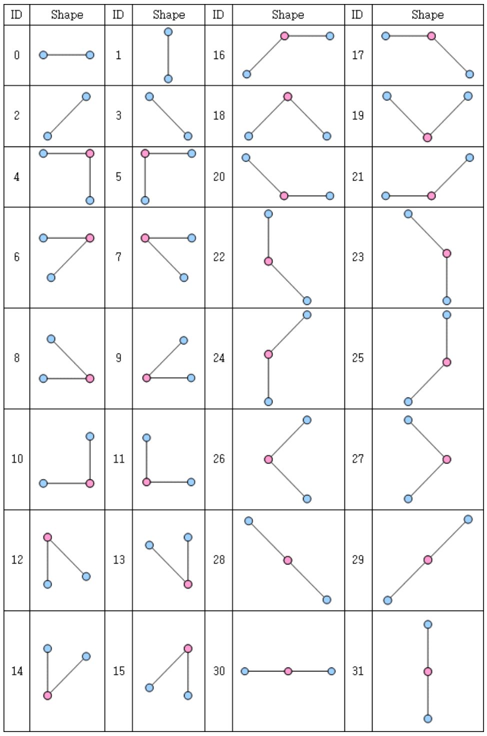

그림 3. 유전자로 결합된 확률분포 모델
Fig. 3 The probability distribution model which is combined with the gene

Matching of the combined probability distribution model is a chain of points of contact and kkeok stochastic graph of each cell and the cells in a matching process to be done first in the recognition between cells, cell recognition.

Extracting an edge representing a biological cross-path connections among the chains to the vertex so that a probability function application to obtain a probabilistic graph can have two probabilities in the probabilistic model graph cell portion can be derived as follows.

In Figure 3, since this distribution to learn and as close as possible to the model of the observed basal an inference model to study parameter estimation with the model selection is a parameter of the conditional probabilities for the model structure. In the model selection, so that best reflect the dependency of the area so are selected from a variety of model structures maximized so that the discrete optimization problem well matched to the data [6] was, to predict the random variable values to calculate a matching distribution [7] was selected.

Ⅲ. Kalman filter algorithm applied

3.1 Kalman filter and pupil tracking using the moving average algorithm

Using an eye tracking apparatus using a bright pupil based on a Kalman filter to determine the location of (c_t, r_t) the pupil in the initial frame and track bright pupil by the Kalman filter. The movement of the pupil in each frame is characterized by a rate of change in the pupil. if the position of (u_t, v_t) the pupil center of the pixel at the time t and referred to at the time t speed change of direction to the c and r, can be expressed at a time point t to the state vector $x_t = (c_t, r_t, u_t, v_t)t$. And the system can be modeled as follows:

$$X_{t+1} = \Phi X_t + W_t \quad (11)$$

wt represents the imperfection of the system. When considered from the point in time t for the pupil position in $Z_t = (\hat{c_t}, \hat{r_t})$, the measurement model can be obtained by the Kalman filter.

$$z_t = Hx_t + v_t \quad (12)$$

In formula 4.2 vt is an indeterminate measurement. Obtaining the estimated position to be a bright pupil effect is estimated by using a threshold value in adjacent pixels. Obtaining an initial value of the state vector x_{t+1} and covariance matrix Σ_{t+1}, is an updated system model and the measurement model for the prediction.

If the snow during a cold pupil tracking Kalman filter or disappear into a bright pupil by an obstruction, so the disadvantage becomes impossible to track by the Kalman filter, using the moving average algorithm increased the accuracy of eye tracking.

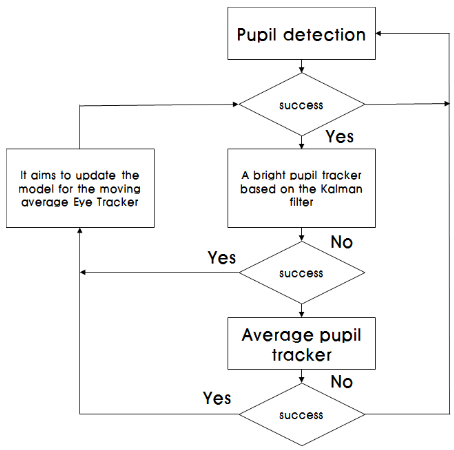

그림 4. 동공 추적 알고리즘
Fig. 4 Pupil tracking algorithm

Average tracking algorithm is repeated moving average by measuring the similarity of two intensity distributions with a bar according to the similarity factor Tara chiya to find the most similar model and candidate regions. Bata's perked estimated coefficients for the target density q in the Y branch and calculate the density of the target candidate.

$$\hat{p}(y) \equiv p[\hat{p}(y), \hat{q}] - \sum_{u=1}^{m} \sqrt{\hat{p}_u \hat{q}_u} \qquad (13)$$

The distance for the two distributions are:

$$d(y) = \sqrt{1 - p[\hat{p}(y), \hat{q}} \qquad (14)$$

Characteristics of the intensity distribution in a region other than the snow and are built quite characteristic by two images even and odd fields. Obtaining a probability distribution of two different characteristics corresponding to a dark pupil image and a bright pupil may lead to a two-dimensional combined histogram.

표 2. 동공 추적 비교

Table 2. Compare pupil tracking

300 Picture frame		Kalman Filter Tracker		The proposed synthetic tracker	
		Recognized frames	Recognition	Recognized frames	Recognition
Left Eye	2260 frame (Open eye)	2012	89%	2255	99%
	330 frame (Close eye)	0	0%	317	95%
	410 frame (Hidden eye)	0	0%	384	94%
Right eye	2125 frame (Open eye)	1945	92%	2117	99%
	330 frame (Close eye)	0	0%	307	93%
	545 frame (Hidden eye)	0	0%	499	92%
Recognition		3957	66%	5879	98%

Ⅳ. Results and Discussion

4.1 facial motion classification analysis

The detected facial feature was used to analyze the facial expression and to determine the movement of each element. Total five awake patient intended for sitting on in the state each room or carelessness with the same set-up as yawning, and estimates the face eye region according to the amount of change of the face and eyelids like jamdeum. Experimental Experiments environment by using the infrared camera system, a total of approximately 1 minute record the face image 5 times with 320 * 240 resolution, 20 frames per second. At this time, the set prior probabilities for all classifications in the bottom-up steps careless yawn, jamdeum in the same operation. Eye detection area for the experiment, head direction estimating head motion estimation, assuming an operation in a combination of face tracking, such as a, b, c, d, e, f, and for carelessness, yawning, jamdeum all steps in each operation a pre-probability measures. Ia, Yw, Fa denote carelessness, yawn, sleep deumeul.

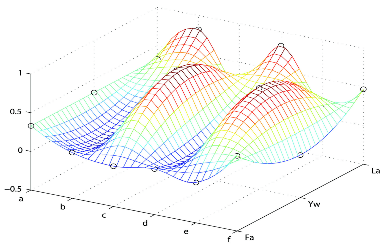

그림 5. 얼굴 동작 분류 결과알고리즘

Fig. 5 Face motion classification

In Figure 5 a was measured Ia, Yw, Fa values are equal to a value smaller than 0.33 to 0.5. This means the state without a change in the operating face when the wakefulness or arousal. b exhibited a high prior probability of the state inadvertently. c This measurement was 0.83 Yw is highly applicable to yawn and d and e are each measured by careless yawn and classification results. That is, a, b, c, d, e is easy to gaze detected in the face region according to various operations of the subject. But f 0.47 was measured in the same manner as the prior probability and a careless jamdeum.

To help identify the gaze position on a monitor by configuring the abduction ever fixture in real-time detection of the pupil obtained a bright pupil effect was traced.

4.2 Re-identification recognition results for adjacent areas

Experimental results on user not involved in the study were identified as a whole indicate the rate of approximately 84%. 3,6 and 4,5 times the area adjacent area is represented in many cases the misrecognized adjacent areas. It is understood that the recognition rate than participating in the study are significantly degraded and can not see that you have identified adjacent to correctly identify the actual gaze area. 23% of 18% of eyes looking at the area 1 is misrecognized eyes looking at the three areas appear as two areas have been identified as areas 4.

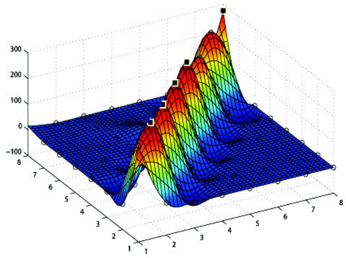

그림 6. 재식별기를 이용한 시선 식별 결과
Fig. 6 Line identification result using a material identifier

As a result, the misrecognized mainly occurs in areas adjacent gaze. Please check the neighboring area for each eye region and perform learning by using only learning data of the adjacent region, so as to re-identify an adjacent area with only verifies the identification result. If it is not the case of the material is improved identification accuracy than 9% was achieved by 93% accuracy. In particular, in error rate was 8% at 18% of the area between the adjacent 1 and 2, the error rate in the area between the eyes 3 and 4 was reduced from the previous 24% to 5%. Also improved error misrecognized by the eye area or other exhibited the same performance. By the generalized corrective action through the eyes GRNN requires no calibration for the individual, even if the movement of the face will be a great advantage because it operates reliably.

Ⅴ. Conclusion

In this paper, by using the Kalman filter using the location information of the current head it was estimated future location is determined whether or not the proper state when the judgment after the hair area detection candidate regions are extracted face. Using both eyes, both nostrils of characteristic elements of the face, by using the ends of the structural information and the mouth relatively fast horizontal processing time, a vertical histogram analysis in order to determine whether or not the authenticity of the face was detected in the face elements.

Experimental results quantizes the monitor coordinates in 4 * 2, to suggest a gaze identification system operating correctly if the face moves, without calibration, a user not involved in the study of the 2400 frames when there is movement of the face It was to identify the right eye with respect to the 2075 frames scored accuracy of about 84%, as a result of the analysis that is not recognized correctly result was found with respect to the adjacent areas that are misrecognized. Therefore, by the re-identification for reducing the misrecognized character in the identification results were ever validate the eye region. When it re-identifies the line of sight of the adjacent regions in the study in error rate was 8% at 18% of between 1 and 2, between 3 and 4 the error rate was decreased from 24% to 5% is increased by about 9% degree of accuracy than the conventional 84% exhibited a 93% accuracy. Proposed Line identifier was identified to improve the performance of the real-time line of sight, or if the calibration according to the user with no natural facial movements. In the future, studies should overcome the limitations of the user's requirements Glasses Come spatial resolution and high attention to further refine your eye area.

References

[1] J. Lee, K. Jeong, G. Park, and K. Sohn, "A Quantitative Reliability Analysis of FPGA-based Controller for applying to Nuclear Instrumentation and Control System," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 9, no. 10, 2014, pp. 1117-1123.

[2] K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20 issue 1, Jan. 2008, pp. 39-51.

[3] T. Yamaguchi, M.Tominage, K, Mrakami, and H. Koshimizu, "Regeneration of facial image eye-contacting with partner on TV Conference environment," *IEEE Int. Conf. on Systems, Man, and Cybernetics*, Chicago, USA, Sept. 2006, pp. 1169-1174.

[4] M. Faruqe and M. Hasan, "Face Recognition Using PCA and SVM," *Anti-counterfeiting, Security, and Identification in Communication*, 2009. *ASID* 2009. *3rd Int. Conf. on*, Hong Kong, Aug. 2009, pp. 97-101.

[5] J. Yang and J. Yang, "Why can LDA be performed in PCA transformed space?," *Pattern Recognition*, vol. 36, no. 2, Feb. 2003, pp. 563-566.

[6] V. Napnik, *The Natue of Statistical Learning Theory*. New York: Springer-verlag, 1995.

[7] P. Liao, J. Liu, M. Wang, J. Ma, and W. Zhang, "Ensemble local fractional LDA for Face Recognition," *Computer Science and Automation Engineering(CSAE)*, 2012 *IEEE Int. Conf. on*, Zhangjiajie, China, vol. 3, May 2012, pp. 586-590.

[8] C. Liu and H. Wechsler, "Independent component analysis of Gabor feature for face recognition," *IEEE Trans. Neural Networks*, vol. 14, no. 4, July 2003, pp. 919-928.

[9] S. El-Khamy, O. Alim, and M. Saii, "Neural Network Face Recognition Using Statistical Feature Extraction," *Radio Science Conf.*, 2000. *17th NRSC '2000. Seventeenth National*, Brisbane, Australia, Feb. 2000, pp. C31/1-C31/8.

[10] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: An application ot face detection," *Proc. IEEE. Computer Society Conf. on. Computer Vision and Pattern Recognition(CVPR)*, San Juan, Puerto Reco. June 1997, pp. 130-136.

[11] M. Yang, "Kernel Eigenfaces vs, Kernal Fisherfaces: Face Recognition Using kernal Methods, Automatrix Face and Gesture Recognition," 202, *Proc. Fourth IEEE Int. Conf.*, Washinton D.C., USA, May 2002.

[12] H. Kim, "Vocal Separation in Music Using SVM and Selective Frequency Subtraction," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 10, no. 1, 2015, pp. 1-6.

[13] J. Jo, "A Car License Plate Recognition Using Colors Information, Morphological Characteristic and Neural Network," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 5, no. 3, 2010, pp. 304-308.

[14] H. Park, "A User Adaptation Method for Hand Shape Recognition Using Wrist-Mounted Camera," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 8, no. 6, 2013, pp. 805-814.

저자 소개

김숙일(Suk-Il Kim)

1996년 중앙대학교 대학원 신문학과 졸업(언론학석사)
2006년 중앙대학교 대학원 신문학과 졸업(언론학박사)
2009년 조선대대학교 사회복지학과 졸업(석사)
1996년 3월 ~ 현재 동강대학교 보육복지과 교수
※ 관심분야 : u-헬스케어, 보육복지, 데이터베이스

석경휴(Gyeong-Hyu Seok)

1995년 호남대학교 전자공학과 졸업 (공학사)
1997년 조선대학교 대학원 컴퓨터학과 졸업(공학석사)
2005년 조선대학교 대학원 컴퓨터학과 졸업(공학박사)
2004년 3월 ~ 현재 청암대학교 병원의료정보과
한국직업능력개발원 통신분야 평가위원
한국의료정보협회 이사
※ 관심분야 : 데이터통신, 신경망, 데이터베이스