

집적 영상의 복원과 통계적 패턴분석을 이용한 왜곡에 강인한 3차원 물체 인식

정회원 엄 석 원*, 이 동 수*, 손 정 영*, 종신회원 김 신 환*

Three-dimensional Distortion-tolerant Object Recognition using Computational Integral Imaging and Statistical Pattern Analysis

Seokwon Yeom*, Dong-Su Lee*, Jung-Young Son* *Regular Members*,
Shin-Hwan Kim* *Lifelong Member*

요 약

본 논문에서는 집적 영상의 획득과 복원을 이용하여 왜곡에 강인한 물체를 인식하는 방법을 연구한다. 해당 화소들의 확률적 특성인 평균과 표준편차를 이용하여 3차원 공간에서 물체를 복원하고 거리를 추정한다. 표적인식은 Fisher 선형판별법(linear discriminant analysis, LDA)과 주성분 분석법(principal component analysis, PCA) 기술을 결합한 통계적 분류기(statistical classifier)로 수행한다. Fisher 선형판별법은 클래스 간의 판별력을 최대로 하고 주성분 분석법은 Fisher 선형판별법을 수행하기 위한 차원축소를 실행한다. 주성분 분석법은 차원축소 후 복원된 벡터와 원 벡터의 오차를 최소화하는 기술로 알려져 있다. 실험 및 시뮬레이션을 통하여 면외(out-of-plane) 회전된 표적을 본 논문에서 제안한 방법으로 분류한다.

Key Words : Integral Imaging, Three-Dimensional Image Processing, Image Reconstruction, Statistical Pattern Recognition, Feature Extraction

ABSTRACT

In this paper, we discuss distortion-tolerant pattern recognition using computational integral imaging reconstruction. Three-dimensional object information is captured by the integral imaging pick-up process. The captured information is numerically reconstructed at arbitrary depth-levels by averaging the corresponding pixels. We apply Fisher linear discriminant analysis combined with principal component analysis to computationally reconstructed images for the distortion-tolerant recognition. Fisher linear discriminant analysis maximizes the discrimination capability between classes and principal component analysis reduces the dimensionality with the minimum mean squared errors between the original and the restored images. The presented methods provide the promising results for the classification of out-of-plane rotated objects.

I. 서 론

집적 영상(integral imaging)은 다수의 마이크로 렌즈로 구성된 렌즈릿 어레이(lenslet array)를 이용하여

3차원의 물체 정보를 저장하고 복원(reconstruction)하는 기술이다^{[1][4]}. 각 렌즈릿은 개별적으로 기본영상(elemental image)을 생성하고 물체정보를 저장한다. 복원은 저장과정의 역으로 행하여지며 컴퓨터를

※ 이 논문은 2008학년도 대구대학교 학술 연구비 지원에 의한 논문임.

* 대구대학교 정보통신공학부(yeom@daegu.ac.kr)

논문번호 : KICS2009-08-358, 접수일자 : 2009년 7월 31일, 최종논문접수일자 : 2009년 10월 8일

이용한 복원과 응용의 연구가 증가하고 있다^{[5],[6]}. 컴퓨터를 이용한 집적 영상의 복원은 가상 핀홀 어레이(pinhole array)를 통하여 임의의 3차원 공간 평면에 물체를 투영한다. 이러한 컴퓨터 복원 기술을 이용하여 물체의 거리를 추정하고 자동으로 표적을 인식하는 연구가 수행되고 있다^{[7],[13]}. 또한 광자 계수 검출기(photon counting detector)와 집적 영상을 이용한 패턴 인식이 제안되었다^{[11],[12]}. 집적영상의 3차원 정보를 이용한 표적의 인식은 3차원 정보를 비간섭(non-coherent) 조명(illumination)을 통하여 단일 노출(single exposure)로 획득할 수 있어 다중 센서 또는 센서의 이동이 필요 없는 간결한(compact)한 시스템을 구성할 수 있고 복원 평면의 거리와 방향을 임의로 설정할 수 있다는 장점이 있다^{[5],[6]}. 이러한 장점을 이용하여 [13]에서는 부분적으로 가려진 표적의 복원과 인식이 수행되었다.

본 논문에서는 집적 영상의 획득과 복원, 그리고 통계적 패턴 인식기술을 이용하여 왜곡(distortion)에 강인한 표적 인식 기술을 연구한다. 해당 화소들의 확률적 특성인 평균과 표준편차를 이용하여 3차원 공간에서 물체를 복원하고 거리를 추정한다. Fisher 선형판별법(linear discriminant analysis, LDA)과 주성분 분석법(principal component analysis, PCA)을 결합하여 면외(out-of-plane) 회전된 표적을 인식한다. Fisher 선형판별법은 벡터를 클래스 간 스캐터(scatter)와 클래스 내 스캐터 행렬의 비율이 최대가 되는 공간에 투영하는 방법이다^[14]. 그러나 학습하는 영상의 수가 벡터의 차원, 즉 영상의 화소의 수보다 적은 경우에 Fisher 선형판별법을 수행하기 위하여 차원의 축소가 필요하다. 주성분 분석법은 벡터를 학습영상의 스캐터 행렬의 고유벡터로 이루어지는 부분공간에 투영하는 방법으로 축소 후 복원된 영상과 원 영상의 mean squared-error를 최소화한다^[14]. 집적 영상의 영상 정보와 PCA-LDA 기법을 이용한 패턴 인식의 연구는 [10]에 보고되었지만 패턴 인식 기술이 집적 영상 복원 과정 없이 적용되었으므로 기본영상의 화소 수가 적은 경우에 물체의 크기 또는 획득 거리에 제한이 따른다. 그러나 본 논문에서는 집적 영상 복원을 선행하여 저해상도 문제를 극복한다. 실험을 통하여 면외 회전된 물체의 거리와 회전에 따른 복원영상의 특성을 보여준다. Fisher 선형판별법과 주성분 분석법으로 투영된 벡터와 판정 경계는 본 논문에서 제안한 방법으로 왜곡된 표적이 인식됨을 보인다.

본 논문의 구성은 다음과 같다. II장에서는 집적

영상의 복원과 거리추정 방법을 설명하고, III장에서는 주성분 분석법과 선형판별법을 이용한 패턴 분류기를 설명한다. IV장에서는 표적의 복원과 인식 실험결과를 V장에서는 결론을 제시한다.

II. 집적 영상 복원과 거리 추정

본 절에서는 집적 영상의 복원과 거리 추정에 관하여 논의한다. 그림 1(a)에서와 같이 집적 영상의 수치적 복원은 가상 핀홀 어레이를 통해 기본 영상을 임의의 거리에 위치한 복원 평면에 투영함으로써 이루어진다.

그림 1(a)의 기호 o_n 는 n 번째의 핀홀이며 g_A 는 기본 영상 평면과 핀홀 어레이 사이의 거리이고 z_A 는 핀홀 어레이와 복원 지점 A 를 포함하는 복원 평면까지의 거리이다. 3차원 공간에서 표적의 정보는 다음과 같이 복원된다^[12].

$$\hat{x}_A = \frac{1}{N_A} \sum_{n=1}^{N_A} x_n \quad (1)$$

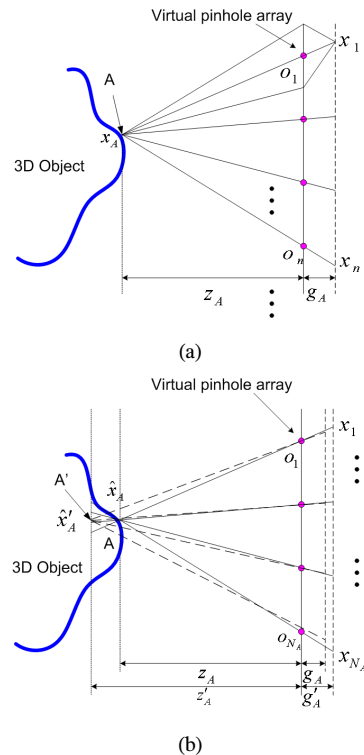


그림 1. (a) 집적 영상의 복원 모델, (b) 물체의 경계 검출 모델

여기서 x_n 는 n 번째 렌즈릿의 기본 영상에서 A 지점에 대응하는 화소의 세기이고 N_A 는 시야각 (viewing angle)이 A 지점을 포함하는 렌즈릿의 수이다. 식 (1)에서 x_n 는 복원 지점 A 에 대응되는 화소와 더불어 A 에 가장 가깝게 투영되는 화소를 포함한다. 이는 컴퓨터를 이용하는 복원은 이산적 (discrete)인 복셀(voxel)에서만 가능하기 때문이다. 그림 1(b)는 거리 추정과 경계 검출 모델이다. 실제로 존재하지 않는 A 지점의 복원에 사용되는 기본 영상의 화소는 A 와 관련 없는 화소에 의하여 형성된다. 그러므로 렌즈릿으로부터 A 지점까지 실제거리 z_A 에 해당하는 화소 세기의 표준편차는 z_A 에서의 표준편차보다 적다. 그러므로 다음과 같이 복원 평면까지의 거리를 추정한다^[13].

$$\hat{z}_A = \operatorname{argmin} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sigma_A(i, j, z) \quad (2)$$

$$\sigma_A^2 = \frac{1}{N_A} \sum_{n=1}^{N_A} (x_n - \hat{x}_A)^2 \quad (3)$$

식(2)에서 N_x 와 N_y 는 복원된 영상의 x 축과 y 축의 크기를 나타낸다. 즉, 표준편차의 합이 최소가 되는 복원평면까지의 거리를 표적의 위치라고 가정한다. 또한 σ_A 는 올바른 거리와 경계에서 선택해지는데지영상을 구성한다. 집적 영상에서 각 기본 영상의 해상도는 일반 카메라의 영상보다 일반적으로 낮지만 수십 또는 수백 개의 기본 영상을 획득한다. 그러므로 복원 지점에 해당하는 화소들 세기의 표준편차를 통하여 경계 부분을 검출하고 물체의 깊이 정보를 추출할 수 있다.

III. 통계적 패턴인식의 적용

본 절에서는 주성분 분석법과 선형판별법을 연계하여 복원영상에 적용한다. 주성분 분석법은 높은 차원(d)의 정보를 표적의 특징이 집중된 낮은 차원(l)으로 투영하는 방법이다. 본 논문에서는 복원된 고차원 영상 벡터 $\mathbf{x} = [x_1 \dots x_d]^t$ 로 구한 스캐터 행렬의 고유 벡터 방향으로 영상 벡터를 선형 투영시켜서 차원을 축소한다. 그러므로 차원 d 는 복원평면에서 화소의 수($N_x \times N_y$)와 동일하고 t 는 전치행렬을 의미한다. 만약 n_c 개의 클래스 중 j 번째 클래스에서 n_j 개의 학습영상이 유용하면 학습영상에

대한 스캐터 행렬을 다음과 같이 구할 수 있다.

$$S = \sum_{j=1}^{n_c} \sum_{n=1}^{n_j} (\mathbf{x}_j(n) - \mathbf{m}_x)(\mathbf{x}_j(n) - \mathbf{m}_x)^t \quad (4)$$

위의 식에서 \mathbf{m}_x 은 모든 학습영상의 표본평균 (sample mean)벡터이다. 식 (4)의 행렬을 이용하여 고유값(eigenvalue) 행렬(A)과 고유벡터(eigenvector) 행렬 $E = [\mathbf{e}_1 \dots \mathbf{e}_d]$ 를 계산한다. 고유값을 크기에 따라 정렬한 후 상위에서 l 번째까지의 고유값에 대응하는 고유벡터로 주성분 분석 행렬 $W_P = [\mathbf{e}_1 \dots \mathbf{e}_l]$ 를 구성한다. 그러므로 투영된 벡터는 \mathbf{y} 는 l 차원의 벡터로 $\mathbf{y} = W_P^t(\mathbf{x} - \mathbf{m}_x)$ 와 같이 구할 수 있다. 투영된 벡터 \mathbf{y} 로부터 복원된 벡터 $\tilde{\mathbf{x}} = W_P \mathbf{y} + \mathbf{m}_x$ 는 \mathbf{x} 과의 mean squared-error를 최소화 한다.

선형판별법은 클래스 내 분산을 나타내는 행렬 (within-class scatter matrix)과 클래스 간 분산을 나타내는 행렬(between-class scatter matrix)의 비를 최대로 하는 선형 투영법이다^[14]. Fisher 선형판별법의 변환 행렬 W_F 는 다음과 같은 조건을 만족한다.

$$W_F = \max_{W_F \in R^{l \times (n_c - 1)}} \frac{|W_F^T S_B W_F|}{|W_F^T S_W W_F|} \quad (5)$$

주성분 분석법을 적용한 후의 클래스 내 스캐터 행렬과 클래스 간 스캐터 행렬은 각각 다음과 같이 나타난다.

$$S_B = \sum_{j=1}^{n_c} n_j (\mathbf{m}_{y_j} - \mathbf{m}_y)(\mathbf{m}_{y_j} - \mathbf{m}_y)^t \quad (6)$$

$$S_W = \sum_{j=1}^{n_c} \sum_{n=1}^{n_j} (\mathbf{y}_j(n) - \mathbf{m}_{y_j})(\mathbf{y}_j(n) - \mathbf{m}_{y_j})^t \quad (7)$$

위의 식에서 \mathbf{m}_y 은 학습영상의 표본 평균벡터 그리고 \mathbf{m}_{y_j} 는 클래스 j 에 속한 학습영상의 조건평균 (class-conditional mean)벡터이다. 그러므로 S_B 는 다른 클래스 간의 분포를 측정하고 S_W 는 동일 클래스의 구성원의 밀집도를 측정한다. W_F 의 역은 $S_W^{-1} S_B$ 의 고유벡터와 동일하다고 알려져 있다^[14]. 일반적으로 차원축소 과정이 없이 영상벡터 \mathbf{x} 에 대하여 선형판별법이 적용될 경우 학습행렬의 수가 영상의 화소의 수보다 일반적으로 적으므로 S_W 는

특이행렬이 되어 투영행렬을 구할 수 없다. 그러므로 본 논문에서 높은 차원의 정보를 낮은 차원으로 투영하는 방법인 주성분 분석법을 선행한다¹⁵⁾.

표적 분류에 대한 판정 방법으로 다음과 같이 유클리드 거리(Euclidean distance)를 사용한다.

$$\hat{j} = \operatorname{argmin} \| W_P^T W_P^d(\mathbf{x}_{test} - \mathbf{m}_{x|j}) \| \quad (8)$$

여기서 \mathbf{x}_{test} 는 미지의 복원 영상 벡터이고 $\mathbf{m}_{x|j}$ 는 학습영상으로부터 얻어지는 클래스 j 의 조건 평균벡터이다. $\| \cdot \|$ 는 유클리드 놈(Euclidean norm)이다.

IV. 실험결과

집적영상 획득 시스템은 렌즈렛 어레이와 픽업 카메라로 구성된다. 각 렌즈렛의 피치(pitch)는 1.09 mm이고 초점 거리는 3.3 mm 정도이다. 세 개의 자동차가 집적영상 픽업 과정을 통하여 획득되었다. 자동차 모형의 크기는 약 4.5 cm x 2.5 cm x 2.5 cm이다. 렌즈렛 어레이와 장애물 사이의 거리는 약 4~5 cm이다. 그리고 렌즈렛 어레이와 표적 사이의 거리는 약 9 cm이다. 각 자동차를 면의 0°, 3°, 6°, 9°, 12°, 15°로 회전하면서 기본영상을 획득하였다. 각 기본 영상 어레이의 크기는 1419 x 1161 화소이며 22 x 18 개의 기본영상으로 구성되어 있다. 그림 2(a)는 모형 자동차의 기본영상 중 하나이다. 그림 2(b)와 2(c)는 복원영상과 에지영상을 보여준

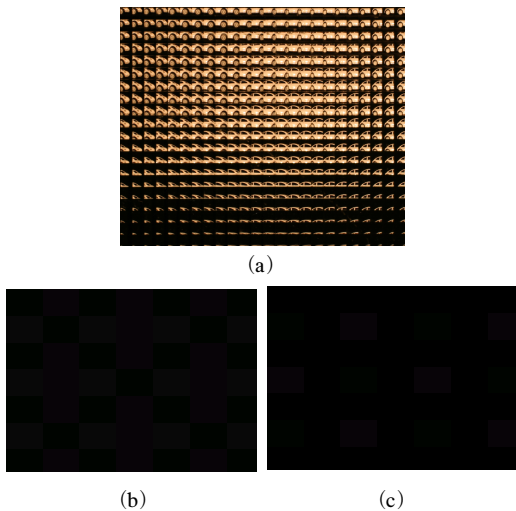


그림 2. (a) 기본 영상, (b) 복원 영상 ($z=49$ mm), (c) 경계 영상 ($z=49$ mm).

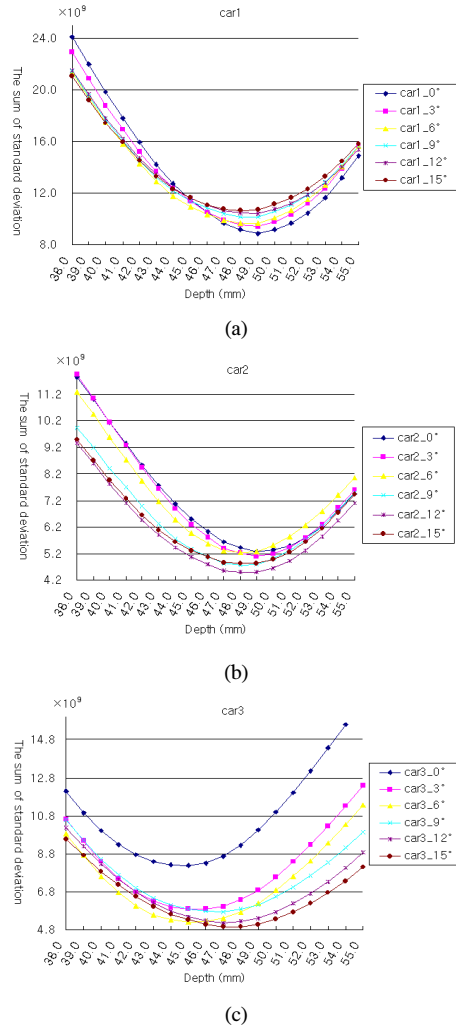


그림 3. 거리에 따른 표준편차의 합, (a) car1, (b) car2, (c) car3.

다. 에지영상은 식 (3)의 σ_A 으로 나타낸다. 그림 3은 복원거리가 38 mm 부터 55 mm로 달리 할 경우 식 (2)의 표준편차의 합을 나타낸다.

학습 과정에서 각 클래스(자동차) 당 두 개의 복원영상이 학습되었다. 학습 영상의 면의 회전각은 9°로 설정되었다. 인식 실험은 9°를 제외한 5개의 각도로 회전한 표적으로부터 얻은 기본 영상에서 복원된 영상 중 5개로 실험하였다. 그러므로 총 테스트 영상은 25개이다. 같은 방법으로 학습영상 각도가 0°일 때 실험을 하였다. 그림 4는 주성분-선형판별법을 적용한 후 학습영상의 평균과 테스트 영상의 위치를 2차원 평면에서 보여준다. 그리고 각 클래스의 평균으로부터 얻어지는 판정경계를 도식하였

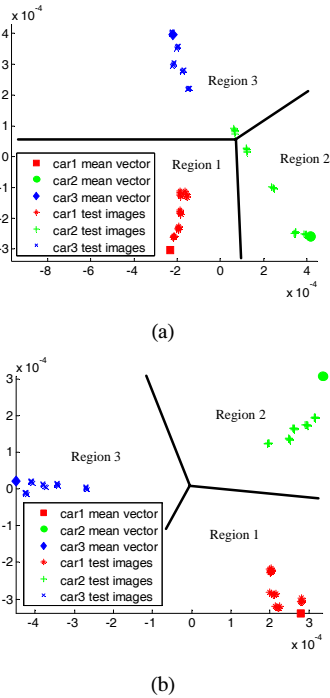


그림 4. 학습영상의 평균과 테스트 영상의 분포, (a) 학습된 표적의 회전각: 0°, (b) 학습된 표적의 회전각: 9°.

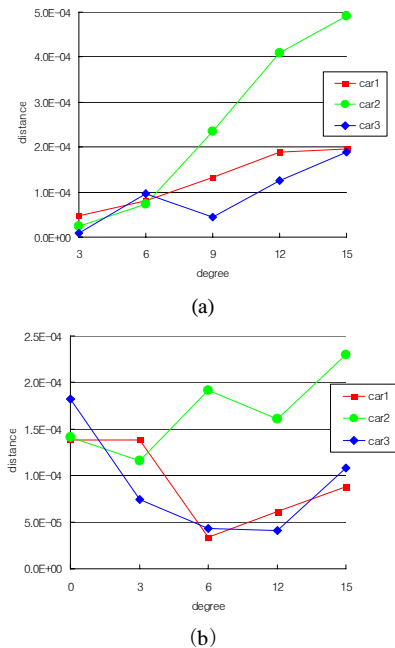


그림 5. 학습영상의 평균과 테스트 영상의 유클리드 거리의 평균, (a) 학습된 표적의 회전각: 0°, (b) 학습된 표적의 회전각: 9°.

다. 학습표적의 회전각도가 0°일 때 클래스 2의 테스트 영상의 일부분이 클래스 1로 판정이 되었다.

그러나 학습표적의 회전각도가 9°일 경우 모든 테스트 영상이 올바른 클래스로 인식되었다. 그림 5는 각 클래스 학습영상의 평균과 동일 클래스의 테스트 영상 간의 유클리드 거리의 평균을 나타낸다. 유클리드 거리는 표적의 회전이 증가할수록 커짐을 보인다. 그림 4에서 보이듯이 선형판별법은 클래스 간 식별력이 최대가 되는 벡터공간으로 투영한다. 이와 같이 선형판별법과 주성분 분석법은 집적영상으로 복원된 영상에 성공적으로 적용되었다.

V. 결론

본 논문에서는 집적 영상으로 복원된 3차원 정보를 이용한 왜곡에 강인한 패턴인식 방법이 연구되었다. 복원과 거리 추정을 위하여 영상 세기의 평균과 표준편차를 이용하였다. 그리고 복원영상을 통계적 패턴 인식기술인 주성분 분석법과 선형판별법을 통하여 학습하고 유클리드 거리를 이용한 판정법으로 분류하였다. 3차원 정보는 수없이 많은 기본 영상으로 얻어지므로 적은 화소 수로 인하여 해상도가 낮을 수 있지만 본 논문에서 수행한 통계적 복원과 패턴 인식 기술로 이를 극복하였다.

참고 문헌

- [1] G. Lippmann, "La photographie integrale," *C. R Acad. Sci.*, 146, pp. 446-451, 1908.
- [2] F. Okano, H. Hoshino, J. Arai, I. Yuyama, "Real-time pickup method for a three-dimensional image based on integral photography," *Applied Optics*, 36(7), pp. 1598-1603, March, 1997.
- [3] J. Son, V. V. Saveljev, Y. Choi, J. Bahn, S. -K Kim, H. Choi, "Parameters for designing autostereoscopic imaging systems based on lenticular, parallax barrier, and integral photography plates," *Optics Engineering*, 42(11), pp. 3326-3333, November, 2003.
- [4] J. Jang, B. Javidi, "Time-multiplexed integral imaging for 3D sensing and display," *Optics and Photonics News*, 15(4), pp. 36-43, April, 2004.
- [5] S. Hong, J. Jang, B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," *Optics Express*, 12(3), pp. 483-491, February 2004.

[6] Y. Hwang, S. Hong, and B. Javidi, "Free view 3-D visualization of occluded objects by using computational synthetic aperture integral imaging," *Journal of Display Technology*, 3(1), pp. 64-70, March 2007.

[7] D. Hwang, D. Shin, S. Kim, E. Kim, "Depth extraction of three-dimensional objects in space by the computational integral imaging reconstruction technique," *Applied Optics*, 47(19), pp. D128-D135, July 2008.

[8] O. Matoba, E. Tajahuerce, B. Javidi, "Real-time three-dimensional object recognition with multiple perspectives imaging," *Applied Optics*, 40(20), pp. 3318-3325, July 2001.

[9] J. Park, J. Kim, B. Lee, "Three-dimensional optical correlator using a sub-image array," *Optics Express*, 13(13), pp. 5116-5126, June 2005.

[10] S. Yeom, B. Javidi, "Three-dimensional distortion-tolerant object recognition using integral imaging," *Optics Express*, 12(23), pp. 5795-5809, November 2004.

[11] S. Yeom, B. Javidi, and E. Watson, "Photon counting passive 3D image sensing for automatic target recognition," *Optics Express*, 13(23), pp. 9310-9330, November 2005.

[12] S. Yeom, B. Javidi, C. Lee, E. Watson, "Photon-counting passive 3D image sensing for reconstruction and recognition of partially occluded objects," *Optics Express*, 15(24), pp. 16189-16195, November 2007.

[13] 이동수, 염석원, 김신환, 손정영, "집적 영상을 이용한 가려진 표적의 복원과 인식", *한국광학회지*, 19(4), pp. 270-275, August 2008.

[14] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern classification 2nd*, Wiley Interscience, New York, pp. 114-124, 2001.

[15] P. N. Belhumer, J. P. Hespanha, D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. on PAMI.*, 19(7), pp. 711-720, July 1997.

염 석 원 (Seokwon Yeom)

정회원



1995년 인하대학교 물리학과 졸업
1997년 고려대학교 전자공학과 졸업
1999년 고려대학교 전자공학과 석사
2006년 코네티컷 대학교 전자공학과 박사

2007년~현재 대구대학교 정보통신공학부 조교수
<관심분야> 신호 및 영상처리, 광정보처리, 패턴인식, 검출 및 추정

이 동 수 (Dong-Su Lee)

정회원



2007년 대구대학교 멀티미디어 공학과 졸업
2009년 대구대학교 정보통신공학과 석사
2009년~현재 대구대학교 정보통신공학과 박사과정
<관심분야> 신호처리, 통신공학

손 정 영 (Jung-Young Son)

정회원



1973년 한국항공대학교 전자공학과 졸업
1982년 테네시대학교 전자공학과 석사
1985년 테네시대학교 응용광학과 박사
1985년~1989년 테네시대학교 박사 후 Space Institute 선임연구원

1989년~2002년 한국과학기술연구소 책임 연구원
2002년~2006년 한양대학교 연구교수
2004년~현재 SPIE Fellow
2006년~현재 대구대학교 정보통신공학부 석좌교수
<관심분야> 3차원 영상, 전자 홀로그래피, 레이저 기반 광학 측정

김 신 환 (Shin-Hwan Kim)

종신회원



1975년 한국항공대학교 통신공학과 졸업
1983년 전북대학교 전기공학과 석사
1988년 동경대학교 전자공학과 박사
1983년~현재 대구대학교 정보통신공학부 교수

<관심분야> 영상 및 비디오 코딩, 스테레오 영상 처리