
두 장의 비교정된 영상으로부터 에피폴라 기하학을 이용한 강건한 대응점 추출

윤용인* · 오인환* · 두경수* · 최종수* · 김진태** · 송호근**

A Robust Correspondence Using the Epipolar Geometry from Two Un-calibrated Images

Yong-In Yoon* · In-Whan Oh* · Kyoung-Soo Doo* · Jong-Soo Choi* · Jin-Tae Kim** · Ho-Keun Song**

이 논문은 교육인적자원부, 산업자원부, 노동부의 최우수실험실 지원사업의 연구비를 지원받았음

요 약

본 논문은 비 교정된 두 장의 영상간에 움직임 정보를 이용하여 먼저 초기 대응되는 후보점들 가운데 잘못된 후보점을 미리 대부분 찾아내서 제거한 후에 에피폴라 평면의 기하학 정보를 이용하여 최소화된 오차를 갖는 기본 행렬을 찾아내는 보다 강건한 알고리즘을 제안한다. 제안된 알고리즘은 초기에 대응되는 후보점을 찾을 때 고전적인 기법인 상관성 대응을 기반으로 대응점을 추출하고 잘못된 대응 후보점들을 정확히 제거함으로써 제안된 알고리즘을 강건하게 하였다. 잘못된 대응 후보점들을 정확히 제거하기 위해 본 논문에서는 두 영상간의 특정점에 대한 이동 벡터의 정보를 이용한다. 다양한 영상을 실험함으로써 제안된 알고리즘이 강건하다는 것을 검증하였고, 그에 따른 실험 결과는 기존 기법의 매칭 알고리즘보다 향상되었음을 보여준다.

ABSTRACT

This paper proposes a robust method to find corresponding points for un-calibrated stereo images by using a classical method based on the epipolar constraints and motion flows. If we detect matching for the only epipolar geometry, the problem is very high. Therefore, in order to find an initial set of matches, we use the correlation technique and then exploit motion vectors to remove mismatches among matching candidates. Then, the epipolar geometry can be accurately estimated using a well adapted criterion and computed the fundamental matrix. The proposed algorithm has been widely tested and works remarkably well in various scenes, evenly, with many repetitive patterns. The results show that the proposed algorithm is better than the conventional.

키워드

비교정된 영상, 에피폴라, 기하학 정보, 기본 행렬, 이동 벡터

* 중앙대학교 첨단영상대학원 영상공학과

접수일자 : 2005. 6. 1

** 한서대학교 컴퓨터정보학과

I. 서 론

자연 영상에서 대응점을 찾는 것은 다중영상 생성(multi-viewpoint image generation)[1], 카메라 교정(camera calibration)[2], 스테레오로부터 3차원 재구성[3]과 같은 다양한 분야에 응용하는데 있어 중요한 역할을 한다. 현재 두 장의 영상 간에 대응점을 정확히 찾는 문제는 컴퓨터 시각(computer vision) 분야에 있어서 중요한 하나의 과제로 남아 있다. 제한 조건은 영상에 대한 일체 기지의 정보 없이 두 영상간의 기하학적 정보만이 있다는 것이다. 영상 매칭의 방법은 크게 두 가지로 분류할 수 있다. 하나는 형판(template) 매칭으로 영상 간에 밝기 값에서 상관관계를 사용하는 알고리즘을 말하며, 유사한 화소 값을 찾아내는 것이다[4]. 다른 하나는 특징 매칭으로써 두 영상 간에 특징이 있는 정보를 추출하는 알고리즘을 말하는 것이다[5].

대응점 찾기 어려운 문제점을 해결하기 위하여, M. Fischer 등이 제안한 RANSAC과 같은 방법[6]은 초기 대응되는 후보점들을 기반으로 직접 에피폴라 평면 기하학을 구성하면서 잘못된 대응점들을 하나씩 제거해 나가는 방식을 취하고 있다. 그렇지만, 이 방법의 경우는 알고리즘에 포함시키는 몇 가지 임계 값을 결정하는데 있어서 초기에 얻어진 특징점들의 분포와 대응 후보점들의 위치에 따라서 실험적으로 그 값을 결정해야 해야 하고, 계산된 값을 중에 임계값들에 대한 결과는 상황에 따라서 큰 차이를 보이기 때문에 적용에 있어서 많은 어려움이 따른다. Z. Zhang의 경우에도 주변 대응점들과의 상호 거리 정보를 이용하여 초기 대응되는 후보점들 가운데 잘못된 후보점들을 제거하는 알고리즘이 포함되어 있지만, 상호 거리 정보를 이용하기 때문에 큰 오차를 가지는 하나의 후보점 쌍에 대해서 주변의 대응 후보점들이 모두 제거된다는 안정적이지 못한 결과를 나타내는 것을 확인하였다[7].

본 논문에서는 두 가지 분류에 있는 방법들 중에 밝기 값의 변화를 기반으로 하는 형판 매칭의 방법을 선택하여 사용하였다. 이를 기반으로 에피폴라 평면상에서 기하학적 정보를 이용하여 강건한 알고리즘을 구성하도록 하였다. 에피폴라 평면의 기하학 조건에서 문제점은 두 영상 사이에 투영된 영상이 존재함으로써 두 영상의 대응점을 찾기가 어려운 문제들이 존재한다. 또한 본 논문의 제안하는 방법은 비교정된 두 장의 영상에 대해서 움

직임 정보를 얻어내어 영상으로부터 먼저 초기 대응되는 후보점들을 가운데 잘못된 후보점들을 미리 대부분 찾아내서 제거한 후에, 에피폴라 평면의 기하학 오차를 감소시키는 보다 강건한 알고리즘을 제안한다. 에피폴라 평면의 기하학을 이용하기 위해, 8쌍의 대응점으로 기본 행렬을 찾아내는 8점 알고리즘을 사용하고[8], 기본 행렬을 찾는 반복수행 과정에서도 RANSAC에서의 중복성을 감소시키기 위해 영상을 분할하여 8쌍의 대응점을 영상 전체에서 골고루 선택될 수 있도록 하였다.

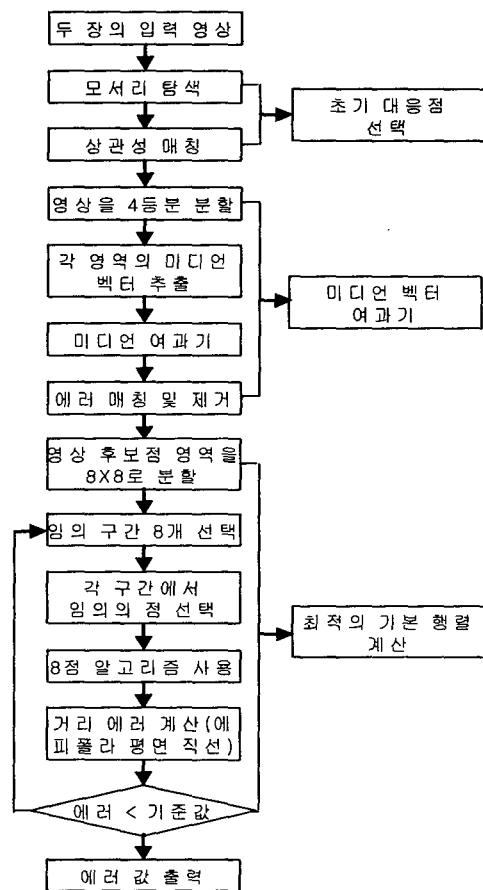


그림 1. 제안된 알고리즘의 전체 블록도
Fig. 1. Block diagram of the proposed algorithm

II. 제안된 알고리즘의 블록도

그림 1은 제안된 알고리즘의 전체 블록도를 나타낸 것

이다. 먼저 입력된 영상으로부터 초기 대응점을 찾기 위하여 고전적인 방법인 모서리 탐색과 상관성 매칭 방법을 사용하여 초기 대응되는 후보점들을 설정한다. 이를 기반으로 잘못된 대응점을 제거하기 위한 방법으로 미디언 벡터 여과기를 사용하였다. 대부분의 잘못된 대응 후보점들을 제거한 후에 에피폴라 평면 기하학적 정보를 사용하여 최적의 기본 행렬을 계산하여 기하학적 에러 값이 최소가 되는 값을 찾는 것이다. 이 방법을 수행하여 실험해보았더니 여러 가지 영상에 대해서 보다 빠르고 정확한 대응관계를 찾아낼 수 있다.[9]

III. 제안된 알고리즘

3.1 특징점 추출

먼저, 입력된 영상으로부터 고전적인 대응점 추출 알고리즘을 사용하여 초기의 대응되는 후보점들을 추출해내고, 추출된 점 중에 포함되어 있는 잘못된 대응점들을 찾아내서 제거하기 위한 방법을 적용하기 위하여 정확한 대응점들의 추출을 수행한다. 그 전에 각 영상에서의 특징점을 추출하는 과정이 선행되어야 한다.

밝기 정보만을 포함하는 좌우 영상 I_L, I_R 에 대해서 밝기의 변화량을 측정함으로써 특징점을 찾아내는 방법으로 좌우 영상에 대한 특징점 정보를 얻어낸다. Harris 모서리 탐색기를 사용하여 다음과 같은 연산자를 사용하였다.

$$R(x, y) = \det[\hat{C}] - k \text{trace}^2[\hat{C}] \quad (1)$$

여기서 \hat{C} 은 다음과 같은 행렬이다.

$$\hat{C} = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (2)$$

단, 밝기 영상 $I(x, y)$ 에서 I_x 와 I_y 는 x 와 y 방향으로의 미분을 나타내는 것이다.

3.2 초기 대응되는 후보점의 추출

두 영상에서 찾아진 특징점을 기반으로 상호연관성을 알아볼 수 있는 상관 함수를 사용함으로써 초기의 대응되는 후보점을 찾아낸다. 상관 함수를 적용하는 상관 윈도

우의 크기는 해당 중심점에 대해서 $(2n+1) \times (2m+1)$ 을 사용하였으며, 두 번째 영상에서 중심점 주위로 사각형의 검색 영역의 크기를 $(2d_u + 1) \times (2d_v + 1)$ 로 선택하고, 첫 번째 영상에서의 점 m_1 과 두 번째 영상에서의 검색 영역 안에 놓여 있는 점 m_2 사이에 상관성을 검사한다.[7]

$$\text{Score}(m_1, m_2) = \frac{1}{(2n+1)(2m+1) \sqrt{\sigma^2(I_1) \times \sigma^2(I_2)}} \times \sum_{i=-n}^n \sum_{j=-m}^m [I_1(u_1+i, v_1+j) - \bar{I}_1(u_1, v_1)] \times [I_2(u_2+i, v_2+j) - \bar{I}_2(u_2, v_2)] \quad (3)$$

여기서,

$$\bar{I}_k(u, v) = \sum_{i=-n}^n \sum_{j=-m}^m \frac{I_k(u+i, v+j)}{(2n+1)(2m+1)} \quad (4)$$

이고, $I_k (k=1, 2)$ 는 점 (u, v) 에서 평균을 나타내며, $\sigma(I_k)$ 는 (u, v) 의 이웃하는 $(2n+1) \times (2m+1)$ 에서 영상 I_k 의 표준 편차를 나타낸다.

$$\sigma(I_k) = \sqrt{\frac{\sum_{i=-n}^n \sum_{j=-m}^m I_k^2(u, v)}{(2n+1)(2m+1)} - \bar{I}_k(u, v)} \quad (5)$$

스코어의 범위는 정규화 과정을 거쳤기 때문에 -1에서 1사이의 값을 가지며 두 영상의 상관도가 정확히 일치할 경우에 이론적으로 1의 값을 갖게 된다. 이때 연산의 대상은 영상 내의 모든 화소 값을 대해서가 아닌 초기 대응되는 후보점에 대해서만 연산을 수행하기 때문에 처리 속도를 향상시킬 수 있다. $I_1(u, v), I_2(u, v)$ 는 각각 좌우 측 영상의 해당 화소 값을 나타내고 n, m 은 상관관계를 계산하는 영역의 크기를 나타내며 일반적으로 $n=7, m=7$ 을 사용한다. 그 결과는 상관관계의 강도를 나타내게 된다. 결과적으로 -1과 1 사이의 값을 중에서 이상적으로 정확한 대응관계에 있는 두 점을 수식에 적용했을 경우 1에 근접한 값을 갖는다. 여기에 일정한 임계값을 주어 대응 후보점으로 추출하게 되고, 본 논문에서는 임계값으로 0.8을 사용하였다.

3.3 잘못된 대응 후보점의 제거

위의 과정에서 초기 대응점에 대한 후보점을 계산한

후에 좌우측의 영상을 각각 4등분으로 분할하여 각 영역마다의 대응점 이동 벡터(motion vector)에 대한 정보를 기반으로 잘못된 후보 점들을 제거한다. 그림 2와 같이 먼저, 영상 전체를 4등분으로 분할하여 각 영역 내에서의 대표되는 이동 벡터를 찾아낸다. 이때 이동 벡터의 대표 값은 빈도가 가장 높은 값에 해당해야 하므로 평균값이 아닌 중간 값을 취하도록 한다. 각각의 영역의 대표 값으로 찾아진 이동 벡터를 기반으로 길이와 방향이 일정한 임계값을 벗어나는 대응 후보 점들을 제거한다. 이때 기준이 되는 임계값은 영상마다 실험적으로 찾아낸다.[9]

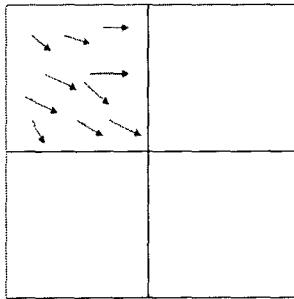


그림 2. 각 구간에 대한 이동 벡터
Fig. 2. Motion vectors of each region

각 구간에서 대표 값으로 찾아진 미디언 벡터(median vector)를 v 라 하고, 해당 영역에서 각 후보 점들의 이동 벡터를 v_i 라 하면 v_i 의 크기에 해당하는 $|v_i - v|$ 성분과 방향에 해당하는 $\tan^{-1}(v_i) - \tan^{-1}(v)$ 성분에 대해서 아래와 같이 두 가지의 경우에 대해서는 잘못된 대응 후보 점으로 제거하게 된다.

- ① $|v_i - v| >$ 크기에 해당하는 임계값
- ② $|\tan^{-1}(v_i) - \tan^{-1}(v)| >$ 방향에 해당하는 임계값

3.4 에피폴라 평면의 기하학적 구성

그림 3에 에피폴라 평면의 기하학적 구성을 나타내었다. 그림 3에서 3차원 물체 M 이 두 대의 카메라 C_1 , C_2 에 투영되어 각각의 영상에 맺힌 좌표를 m_1 , m_2 라고 하면 m_1 , m_2 는 대응점 관계를 갖게 되고, 이때 기준 카메라 C_1 에 대한 C_2 의 기본행렬 F 를 구하면 아래의 식이 성립하게 된다.

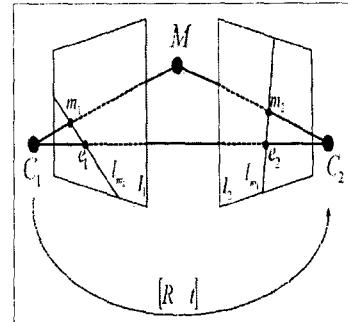


그림 3. 에피폴라 기하학적 표현
Fig. 3. Epipolar geometry representation

$$l_{m_1} = F \tilde{m}_1 \quad (6)$$

$$\tilde{m}_2 F \tilde{m}_1 = 0 \quad (7)$$

$$F \tilde{e}_1 = F^T \tilde{e}_2 = 0 \quad (8)$$

여기서, l_{m_1} 은 좌측 영상에서의 m_1 에 대한 우측 영상으로의 에피폴라 직선이 되고 e_1 , e_2 는 두 카메라의 중심을 연결한 직선이 영상을 지나는 교점인 에피폴라가 된다.

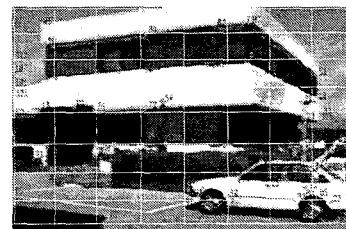


그림 4. 대응점들을 포함하는 8X8 영역 분할
Fig. 4. 8X8 region segmentation with correspondence

RANSAC에서 적용하는 알고리즘과 다른 점은 그림 4에서 보는 바와 같이 좌측 영상에 대해서 대응점들을 포함하는 사각형 테두리를 써우고 그 테두리를 8x8의 영역으로 나누어 8개의 영역을 임의로 뽑아내고 또 다시 각각의 영역 내에서 하나의 대응점을 임의로 뽑는 방식으로 영상 전체에서 골고루 8쌍의 대응점을 얻어내어 8점 알고리즘을 통해 기본 행렬을 구해낸다.[11] 이 과정을 반복 수행하면, 기본 행렬을 기반으로 에피폴라 평면의 기하학을 형성하여 에피폴라 평면 선분과 해당 대응점 간의 거리 오차를 계산한 아래의 값이 작아지는 기본 행렬

을 최종 값으로 선택한다.

$$S = \min \sum_i (d^2(m_{2i}, Fm_{1i}) + d^2(m_{1i}, F^T m_{2i})) \quad (9)$$

IV. 실험 결과

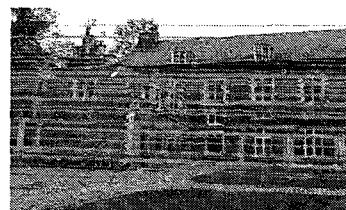
본 논문에서는 IBM-PC 펜티엄 IV 1.4GHz, 256M DDR의 사양으로 컴퓨터 모의실험을 하였으며, 두 종류의 영상을 대상으로 실험하였다. 최적의 기본 행렬에 대해서 에피폴라 평면의 선분을 도시하는 것과 그에 해당하는 거리오차들의 평균을 수치적으로 제시하였다. 본 논문과의 성능비교를 위해 일반적으로 많이 사용되는 RANSAC 방법[6]을 적용한 결과를 제시하였다. 실험 결과의 영상을 그림 5,6에 보이고, 그 결과를 표 1,2에 각각 정리하였다. 결과적으로 RANSAC 방법보다 잘못된 대응 후보점들이 모두 제거되었음에도 불구하고 최종적으로 구해진 대응점의 수가 제안된 방법을 적용했을 때 보다 더 많고, 거리 오차도 작아진 것을 볼 수 있다.



(a) 제안된 방법



(b) RANSAC 방법

그림 5. '건물' 영상
Fig. 5. 'Building' image

(a) 제안된 방법



(b) RANSAC 방법

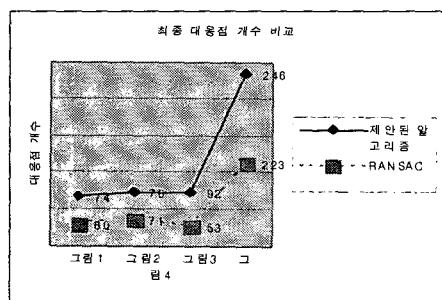
그림 6. '캐슬' 영상
Fig. 6. 'Castle' image표 1. '건물' 영상의 처리 결과
Table 1. Simulation results of 'Building' image

영상의 크기 (512 X 256)	제안된 방법		RANSAC 방법
	처리 시간	대응 점	대응 점
Correlation Matching	0.170 sec	77	X
Median Vector Filter	0.020 sec	74	X
Fundamental 행렬	0.091 sec	74	60
최종 거리 오차 평균 (pixel)	0.438		0.496

표 2. '캐슬' 영상의 처리 결과
Table 2. Simulation results of 'Castle' image

영상의 크기 (384 X 288)	제안된 방법		RANSAC 방법
	처리 시간	대응 점	대응 점
Correlation Matching	0.591 sec	92	X
Median Vector Filter	0.020 sec	76	X
Fundamental 행렬	0.093 sec	76	71
최종 거리 오차 평균 (pixel)	0.236		0.508

각각의 실험 영상에 대한 수치적인 비교는 그림 7에 보인다. 제안된 알고리즘의 경우에 최종적으로 구해진 대응점의 개수가 많고 기본 행렬에 대한 오차 값이 작아졌음을 알 수 있다.



(a) 최종 대응점 수

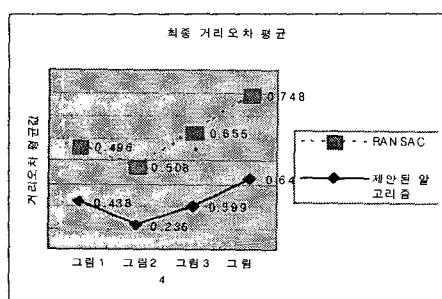


그림 7. 실험 결과
Fig. 7. Experiment results

V. 결 론

본 논문은 비 교정된 두 장의 영상에 대해서 움직임 정보를 이용하여 영상으로부터 초기에 대응되는 후보 점들 가운데 잘못된 후보 점들을 미리 대부분 찾아내서 제거한 후에, 에피폴라 평면의 기하학 정보를 이용하여 기본 행렬에 대한 오차를 감소시키는 보다 강건한 알고리즘을 제안하였다. 일반적인 두 장의 스테레오 영상으로부터 명확한 대응점의 정보를 얻어내기란 쉽지가 않다. 또한 계산된 대응점을 기반으로 구해지는 기본 행렬의 경우는 그 오차에 매우 민감하게 변화하기 때문에 최소의 오차를 포함하는 기본 행렬을 구하는 것은 굉장히 중요한 과

제이다. 강건한 대응점 정보를 얻어내기 위해서 잘못된 대응관계를 가지는 후보 점들을 사전에 제거하고, 에피폴라 평면 기하학을 구성하도록 하였다. 또한 기본 행렬을 구하는 과정에서도 영상에 고루 펼쳐진 대응점 8쌍을 선택하여 8점 알고리즘에 적용해야만 신뢰할 수 있는 기본 행렬을 구할 수 있다. 본 논문에서는 일반적인 영상에서도 보다 강건하다는 것을 실험 결과를 통하여 증명하였다.

본 논문의 내용은 다양한 스테레오 영상의 특성에 대해서, 일반적으로 신뢰할 수 있는 결과를 얻기 위한 과정이었고, 보다 정확한 에피폴라 평면 기하학을 구성하고자 한다면, 영상마다의 특성을 고려하여 임계값이나 영상 분할 방법 등을 수정함으로써 어느 정도의 개선할 수 있다.

참고문헌

- [1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.
- [2] Z. Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1-20, 1998.
- [3] R. Cipolla and E. Boyer, "3D Model Acquisition from Uncalibrated Images," *IAPR Workshop on Machine Vision Applications*, pp. 559-568, 1998.
- [4] P. Fua, "A Parallel Stereo Algorithm that Produces Dense Depth Maps and Preserves Image Features," *Machine Vision and Applications*, vol. 6, no. 1, 1993.
- [5] R. Horaud and T. Skordas, "Stereo Correspondence Through Feature Grouping and Maximal Cliques," *IEEE Trans. PAM*, vol. 11, no 12, pp. 1310-1180, 1989.
- [6] M. Fischler and R. Bolles, "Random Sampling and Algebra of Point and Line Correspondences Between N Images," *Proc. International Conference on Computer Vision*, pp. 951-962, 1995.
- [7] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong, "A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry," *Artificial Intelligence Journal*, vol. 78, pp. 87-119, 1995.
- [8] Q. Luong and O. Faugeras, "The Fundamental Matrix: Theory, Algorithms, and Stability Analysis," *The*

International Journal of Computer Vision, vol. 17, no. 1, pp. 43-76, 1996.

[9] B. Triggs, "Matching Constraints and the Joint Image," IEEE Int. Conf. Computer Vision, 1995.

저자소개



윤 용 인(Yong-In Yoon)

1986년 중앙대학교 전자공학과 졸업
(공학사)
1988년 중앙대학교 대학원 전자공학
과 졸업(공학석사)
2003년 중앙대학교 첨단영상대학원 영상공학과 졸업(공
학박사)
1991년 ~ 1993년 국제전자(주) 연구원
1994년 ~ 1999년 대우고동기술연구원
2003년 ~ 2005년 중앙대학교 첨단영상대학원 박사후과정
2006년 ~ 중앙대학교 문화콘텐츠연구원 영상콘텐츠연구원
※관심분야: 증강현실, 신호처리, 컴퓨터 비전, 영상처리 등



오 인 환(In-Whan Oh)

2002년 중앙대학교 전자공학과 졸업
(공학사)
2004년 중앙대학교 대학원 전자공학
과 졸업(공학석사)

2004년 ~ 현재 신도리코 연구소
※관심분야: 신호처리, 영상처리 등



두 경 수(Kyoung-Soo Doo)

1989년 중앙대학교 전자공학과 졸업
(공학사)
1991년 중앙대학교 대학원 전자공학
과 졸업(공학석사)
2001년 중앙대학교 첨단영상대학원 영상공학과 졸업(공
학박사)
2002년 ~ 2004년 일본 소프트재팬 주임연구원
2004년 ~ 2005년 중앙대학교 첨단영상대학원 계약교수
2005년 ~ 현재 중앙대학교 첨단영상대학원 전임강사
※관심분야: 영상처리, 전방향카메라시스템, 증강현실 등



최 종 수(Jong-Soo Choi)

1975년 인하대학교 전기공학과 졸업
(공학사)
1977년 서울대학교 대학원 전자공학
과 졸업(공학석사)

1981년 일본 게이오대학 대학원 전기공학과 졸업(공학
박사)
1981년 ~ 1999년 중앙대학교 전자공학과 교수
1995년 ~ 현재 중앙대학교 첨단영상대학원 교수

※관심분야: 증강현실, 컴퓨터 비전, 영상정보기술 등



김 진 태(Jin-Tae Kim)

1987년 중앙대학교 전자공학과 졸
업(공학사)
1989년 중앙대학교 대학원 전자공
학과 졸업(공학석사)

1993년 중앙대학교 대학원 전자공학과 졸업(공학박사)
1993년 ~ 1994년 중앙대학교 기술과학연구소 선임연구원
1994년 ~ 1995년 서울대학교 자동제어특화연구센터 선
임연구원

1995년 ~ 현재 한서대학교 컴퓨터정보학과 부교수
※관심분야: 영상통신, 얼굴인식, 디지털 워터마킹 등



송 호 근(Ho-Keun Song)

1991년 중앙대학교 전자공학과 졸
업(공학사)
1993년 중앙대학교 대학원 전자공
학과 졸업(공학석사)

1997년 중앙대학교 대학원 전자공학과 졸업(공학박사)
1996년 ~ 현재 한서대학교 컴퓨터정보학과 부교수
※관심분야: 내용기반 영상/비디오 검색, 3차원 영상 처
리/인식, 영상통신/회의 등