

논문 2013-50-8-26

효율적인 호모그래피 추정을 통한 파노라마 영상 생성

(Efficient Homography Estimation for Panoramic Image Generation)

서 상 원*, 정 수 웅*, 한 윤 상*, 최 종 수**, 이 상 근**

(Sangwon Seo, Soowoong Joeng, Yunsang Han, Jongsoo Choi, and Sangkeun Lee[©])

요 약

본 논문에서는 고해상도 영상에서 호모그래피를 효율적으로 추정하는 방법을 제안하였다. 호모그래피를 정확하게 추정하는 것은 영상 정합 기술에서 가장 중요한 부분이다. 하드웨어의 급속한 발전으로 고해상도 영상을 쉽게 취득 할 수 있게 되었지만, 데이터의 크기가 증가함에 따라 정확한 일치관계를 추정하는데 많은 계산량이 요구되었다. 또한, 고해상도의 영상에서 확률적으로 부정확한 호모그래피가 추정되는 문제가 발생하였다. 따라서 우리는 원 영상을 각 스케일별로 다운 샘플링한 다수의 영상을 생성하여 각 다운 샘플링된 영상에 해당하는 호모그래피를 추정한 후 투영 오차가 가장 작은 것을 선택하여 원 영상에 적용할 수 있게 보정할 수 있는 방법을 제안하였다. 또한, 영상을 정합하는 과정에서 기준 영상과 대상 영상의 색상 톤의 차이가 큰 경우에는 중첩영역의 지역 정보만을 이용하여 기준 영상과 대상 영상의 색상 톤을 일치시키는 방법을 제안하였다. 실험 결과 기존의 방법보다 3.2M 픽셀의 해상도 영상에서 약 3배, 8M 픽셀의 해상도 영상에서 약 5배 이상 빠른 결과를 얻을 수 있음을 확인하였다. 이는 입력 영상의 해상도가 커질수록 제안한 방법의 효과가 더 커진다는 것을 보여준다.

Abstract

An efficient homography estimation method for large sized images is proposed. Estimating an accurate homography is one of the most important parts in image stitching processes. Since hardwares have been advanced, it has been possible to take higher resolution images. However, computational cost for estimating homography has been also increased. Specifically, when too many features exist in the images, it requires lots of computations to estimate a correct homography. Furthermore, there is a high probability of obtaining an incorrect homography. Therefore, we propose a numerical method to extract the appropriate correspondences from several down-scaled images to estimate and compensate the homography numerically for restoring an original homography. Also, if there is an unbalance in color tone between the reference and the target images, we make them balanced by using local information of the overlapped regions. Experimental results show that proposed method is three times faster in 3.2 mega pixel images, five times faster in 8mega pixel images than the conventional approach. Therefore, we believe that the proposed method can be a useful tool to efficiently estimate a homography.

Keywords : Homography, DLT, Stitching, SIFT, Color Balancing

I. 서 론

* 학생회원, ** 평생회원, 중앙대학교 첨단영상대학원 영상학과

(Graduatee School of Advanced Imaging Science, Multimedia, and Film Chung-Ang University)

[©] Corresponding Author(E-mail: sangkny@cau.ac.kr)

※ 이 논문은 2013년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2013-029824)

접수일자: 2013년3월21일, 수정완료일: 2013년7월25일

영상 정합 기술은 중첩된 영역이 있는 입력 영상을 정합하는 영상처리 알고리즘으로 다수의 영상을 이용하여 하나의 고해상도의 영상을 생성하는 방법이다. 우리가 사용하는 디지털 장비는 최대해상도의 한계를 가지고 있으나 이 기술을 이용하면 한계를 극복한 넓은 시

야의 초고해상도의 영상을 얻을 수 있다. 일반적으로 파노라마 영상을 생성할 때와 HD(High-Definition) 영상을 이용하여 UHD(Ultra-High-Definition) 영상을 생성할 때 이용되고 있으며^[1] 스마트폰, 태블릿 등 모바일 장비에서도 사용되고 있다. 대표적인 애플리케이션으로는 구글과 네이버의 지도 서비스인 스트리트 뷰의 항공 영상 서비스와 위성 영상 서비스 등이 있다.

영상 정합 기법에서 사용자의 특별한 입력 없이 특징점 기반으로 정합하는 기법이 많이 이용되고 있다. 이때 가장 중요한 것은 입력 영상들 사이의 대응 관계를 나타내는 호모그래피를 정확하게 추정하는 것이며 기본적으로 3단계의 과정을 통하여 추정한다. 먼저 중첩된 영역이 있는 기준 영상과 대상 영상에서 SIFT(Scale Invariant Feature Transform)^[2] 알고리즘과 같은 특징점 추출 알고리즘을 이용하여 특징점을 추출하고 기준 영상과 대상 영상들에서 추출된 특징점들 사이에서 일치되는 특징점을 찾는다. 그리고 RANSAC(RANdom SAmple Consensus)^[3-6] 알고리즘을 이용하여 일치된 특징점들 사이에서 부정확하다고 판단되는 점들을 제거하고 정확하게 일치되었다고 판단되는 임의의 4점을 추출한다. 마지막으로 임의의 4점을 DLT(Direct Linear Transform)^[7]식에 적용하여 호모그래피를 추정하며, 그 중 최소의 에러를 갖는 최적의 호모그래피가 추출될 때까지 이 과정을 반복한다.

SIFT 알고리즘과 RANSAC 알고리즘을 이용하여 파노라마 영상을 생성하는 기존의 방법으로^[8-11] 영상을 정합할 때 주로 HD급 이하의 저해상도 영상들을 이용하였기 때문에 과도한 특징점의 추출로 인하여 계산량이 급격하게 증가하고 부정확한 호모그래피가 추정되는 문제는 크게 부각되지 않았다. 그러나 최근 HD 영상 4장을 정합하여 UHD 영상을 제작하는 방식이 제안되고^[12] 기술의 발전으로 고해상도 영상을 정합하는 경우가 늘어나면서 저해상도 영상들을 정합할 때에 비해 고해상도 영상을 정합할 때 여러 가지 문제가 부각되고 있다. 그 원인으로는 영상의 해상도가 커질수록 SIFT 알고리즘으로 추출되는 특징점의 수가 급격하게 증가하고, 특징점을 일치시키는 과정에서 계산량이 기하급수적으로 증가하기 때문이다. 또한, 부정확한 특징점이 일치되는 것을 방지하기 위해서 RANSAC 알고리즘을 적용할 때 오히려 너무 많은 데이터양 때문에 부정확한 호모그래피가 추정될 확률이 높아진다^[13]. 따라서 우리

는 고해상도 영상을 정합할 때 각 스케일별로 다운 샘플링한 다수의 영상에서 추정된 최적의 호모그래피를 선택하고 보정하는 방법으로 호모그래피를 효율적으로 추정하는 방법을 제안한다. 또한 영상 정합을 위한 중첩된 영역이 있는 연속적인 영상을 촬영 하더라도 카메라의 시점이 변하면서 색상 톤이 달라 질 수 있다. 따라서 우리는 입력 영상들의 중첩 영역의 지역 정보만을 이용하여 색상 보정 행렬을 생성하고 이것을 대상 영상에 적용함으로써 색상 톤을 기준 영상에 일치시키는 방법을 제안하였으며, 결과적으로 기준 영상과 대상 영상의 색상 톤이 보다 자연스러워지는 결과를 얻을 수 있었다.

본 논문의 I장은 기존의 영상 정합 방법과 최근에 부각된 문제점, II장은 기존 방법의 문제점을 해결한 효율적인 호모그래피 추정 방법, III장은 기존 방법과 비교한 실험 결과, IV장은 결론으로 구성되었다.

II. 본 론

본 논문에서는 고해상도의 입력 영상들로부터 호모그래피를 효율적으로 추정하기 위해서 입력으로 들어온 기준 영상과 대상 영상을 $1/\alpha$ 로 다운샘플링 한다. 그리고 SIFT알고리즘을 이용하여 특징점을 추출한 후 특징점을 일치시킨다. 부정확하게 일치된 특징점들(outlier)을 제거하기 위해서 RANSAC알고리즘을 이용하고 정확하게 일치된 특징점(inlier)을 추정한다. 그리고 기준 영상과 대상 영상에서 최적의 4쌍의 특징점을 결정하고 이것을 DLT식에 적용하여 호모그래피를 추정한다. 이 과정을 SIFT를 통해 일치된 특징점의 개수가 β 이하가 될 때 까지 반복하여 수행한다. 이렇게 추정된 다수의 호모그래피의 투영 오차(projection error)를 계산하고 투영 오차가 가장 적은 것을 최적의 호모그래피로 결정한다. 이 때 최적이라 판단된 호모그래피를 원 영상에 그대로 적용하면 부정확한 결과가 나오기 때문에 보정 행렬식을 이용하여 최적의 호모그래피를 보정한다. 제안된 방법을 이용하면 기존의 방법보다 효율적으로 호모그래피를 추정할 수 있다.

영상을 촬영할 때 같은 카메라로 같은 시간에 중첩 영역이 있는 영상을 연속적으로 촬영해도 영상들의 색상 톤이 다를 수 있다. 그 이유는 영상을 정합하기 위하여 중첩 영역이 있는 영상을 촬영할 때 카메라의 시점

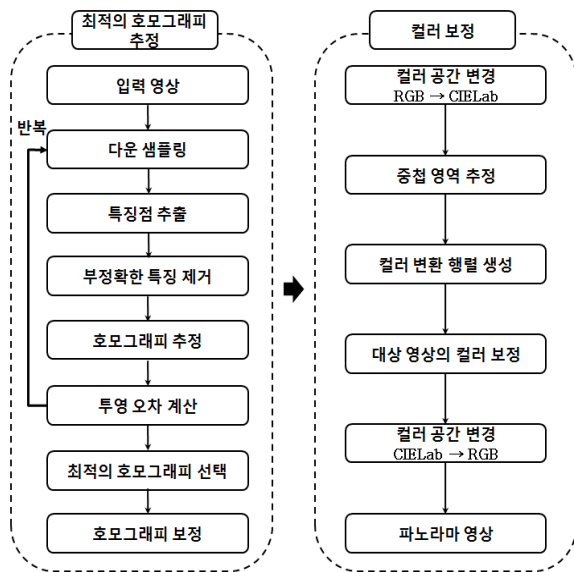


그림 1. 제안된 방법의 흐름도
Fig. 1. Flowchart of the proposed method.

이 바뀌면서 카메라가 받아들이는 빛의 양이 달라지기 때문이다. 따라서 정합된 영상에서 더 자연스러운 결과를 얻기 위해서는 색상을 보정하는 과정이 필요하다. 따라서 본 논문에서는 중첩된 영역에서의 지역 정보를 바탕으로 SVD(Singular Value Decomposition)를^[14] 이용하여 색상 변환 행렬을 만들었다. 그리고 이것을 대상 영상에 적용하여 기준 영상에 색상을 보정해 주었다. 제안된 알고리즘을 이용하면 기존 알고리즘보다 효율적으로 호모그래피를 추정할 수 있으며 기준 영상과 대상 영상의 색상 톤이 다른 입력 영상을 정합하는 경우에도 더 자연스러운 결과를 얻을 수 있었다.

1. 특징점 추출

특징점을 기반으로 파노라마 영상을 생성할 때에는 일반적으로 크기, 회전, 조명성분 등의 변화에 강건하게 특징을 추출하는 SIFT알고리즘을 이용하여 특징점을 추출하고 특징점들의 대응 관계를 찾는다.

SIFT알고리즘에서 특징점을 찾는 방법은 크게 4단계로 이루어져 있다. 첫 번째 단계에서는 특징점의 후보가 될 수 있는 점들을 찾는다. 입력 영상이 들어오면 가우시안 필터링을 이용하여 가우시안 피라미드를 만들고 피라미드의 전 영상과 후 영상의 연속된 2장의 영상을 이용하여 DOG(Difference of Gaussian)영상을 만든다. 그리고 각 픽셀에서 현재 영상의 인접한 8개의 픽셀과 전 영상과 후 영상에서 각각 인접한 9개씩의

픽셀씩, 총 26개의 픽셀 중에서 극값을 찾아서 특징점이 될 수 있는 후보를 만든다. 두 번째 단계에서는 특징점 후보 중에서 안정적인지 못한 특징점들을 제거하고 정확한 위치에 없는 특징점들은 테일러 급수를 이용하여 정확한 위치로 옮겨준다. 세 번째 단계에서는 회전 변환에 강건한 특성을 만들어 주기 위해서 방향과 크기를 결정해 준다. 가우시안 블러 시킨 영상에서 특징점 주변 16×16 픽셀에서 방향과 크기의 값을 얻는다. 마지막 단계에서는 특징점 기술자를 만들기 위해서 특징점을 중심으로 주변의 기울기(gradient)값들의 방향을 구하고 가우시안 윈도우를 사용하여 크기의 1.5 배에 해당하는 값을 가중치로 준다. 그리고 16×16 블록의 방향과 크기로 이루어진 4×4 크기의 특징점 기술자를 생성한다. 특징점 기술자가 생성된 후에는 kd-tree^[15]를 이용하여 기준 영상과 대상 영상의 대응되는 점들을 일치시킨다. 이 때 일치된 특징점들이 모두 정확하다고 판단할 수 없으므로 RANSAC알고리즘을 이용하여 잘 못 일치된 점들을 제거하고 호모그래피 추정을 위한 4쌍의 점을 추출한다. RANSAC알고리즘은 에러를 포함하고 있는 데이터 집합에서 반복적인 작업을 통해서 확률적으로 최적의 값을 도출하는 알고리즘이다. SIFT알고리즘에서 추출된 특징점 중에서 RANSAC알고리즘을 통해서 임의의 4점을 추출하고 확률적으로 최적이라고 판단되는 호모그래피가 도출될 때까지 반복적으로 추정한다.

2. 호모그래피 추정

파노라마 영상을 생성하기 위해서 가장 중요한 요소는 정확한 호모그래피를 추정하는 것이다. 파노라마 영상을 얻을 때 기본 전제 조건은 3차원 실세계에 있는 영상을 카메라의 2차원 이미지 평면에 투영시키는데 두 영상은 동일평면상에 있다고 가정을 한다. 여기서 호모그래피란 3×3 행렬로 기준 영상의 점과 그에 대응되는 대상 영상의 점들 사이의 관계를 나타낸 것이다. 기준 영상과 대상 영상의 대응 관계는 식 (1)로 정의할 수 있다.

$$X = HX' \tag{1}$$

여기서 X 와 X' 은 기준 영상과 대상 영상의 점이고 H 는 X 와 X' 의 관계를 나타낸 호모그래피 행렬이다. 호모그래피 행렬은 식 (2)와 같다.

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad (2)$$

호모그래피를 추정하기 위해서는 기본적으로 기준 영상의 4점과 그에 대응되는 대상 영상의 4점이 필요하다. 일반적으로 SIFT 알고리즘과 RANSAC 알고리즘을 통해서 나온 기준 영상과 대상 영상 사이에 대응되는 특징점 4쌍을 DLT식 (3)에 적용하여 호모그래피를 추정한다.

$$\begin{bmatrix} 0^T & -w'_i X_i^T & y'_i X_i^T \\ w'_i X_i^T & 0^T & -x'_i X_i^T \end{bmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0 \quad (3)$$

식 (3)에서 두 번째 행렬의 h^j 는 식 (2)의 호모그래피 행렬의 첫 번째, 두 번째, 세 번째 열을 나타낸다. X_i^T 는 기준 영상에서 i 번째 특징점의 좌표(x_i, y_i, w_i)를 나타내고 X_i^T 는 i 번째 특징점에 대응되는 대상 영상의 특징점 좌표(x'_i, y'_i, w'_i)를 나타낸다. i 는 호모그래피를 추정하기 위한 기준 영상과 대상 영상에 일치하는 특징쌍에서 첫 번째 대응 관계에 있는 특징점부터 네 번째 대응 관계에 있는 특징점을 의미한다. 식 (3)에서 첫 번째 행렬은 A_i , 두 번째 행렬을 h 로 표현하면 식 (4)와 같다.

$$A_i h = 0 \quad (4)$$

여기서 A_i 는 기준 영상과 대상 영상 사이에 대응되는 i 번째 특징점을 나타낸다. 호모그래피 행렬은 A_i 를 SVD 식 (5)에 적용시켜서 추정할 수 있다^[7].

$$A = U \Sigma V^* \quad (5)$$

SVD 식 (5)는 특이값 분해를 통해서 선형방정식의 해를 구할 때 이용한다. A 행렬은 U, Σ, V 로 분해할 수 있으며 A 행렬이 $m \times n$ 행렬이면 Σ 는 특이값을 나타내는 $m \times n$ 대각 행렬이고 U 와 V 는 특이값에 대응되는 $m \times m, n \times n$ 특이 벡터들을 나타내는 직교 행렬이다.

3. DLT 특성

호모그래피를 구하기 위해서 기준 영상과 대상 영상의 일치되는 4점을 기반으로 DLT방법을 적용한다. 제

안된 방법에서는 효율적으로 입력 영상들의 호모그래피를 추정하기 위해서 다수의 다운샘플링을 수행한다. 다운샘플링을 수행하면 영상 해상도의 변화로 추출되는 특징점의 위치가 달라지기 때문에 스케일마다 같은 값의 호모그래피 행렬이 나오지 않는다. 따라서 영상을 다운샘플링하여 구해진 호모그래피는 원 영상에 바로 적용할 수 없다. 이러한 문제를 해결하기 위해서 스케일 변화에도 적용할 수 있도록 DLT식의 특성을 분석하여 호모그래피를 보정하는 방법을 제안한다. DLT식에 입력으로 들어가는 8개의 특징점 좌표를 이용하여 추정된 호모그래피와 8개의 특징점 좌표의 공약수를 이용하여 호모그래피를 추정하고 식 (7)을 통하여 보정된 결과가 정확히 일치하는 것을 확인하였다. 8개 좌표 공약수들의 집합 C 는 식 (6)과 같다.

$$C = \{n_1, n_2, n_3, \dots, n_k\} \quad \text{for } 1 < k < N \quad (6)$$

n_k 는 8개 좌표의 공약수이고 N 은 약수의 개수이다. 집합 C 에 속한 약수 n_k 를 이용하여 8개의 좌표를 나눠서 호모그래피를 추정하고 식 (7)을 이용하여 보정하면 원 영상의 호모그래피를 정확하게 복원할 수 있다.

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/n \end{bmatrix} \widehat{H}^d \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & n \end{bmatrix} \quad (7)$$

H 는 원 영상에서 나온 특징점으로 추정된 호모그래피이고 \widehat{H}^d 은 약수로 나눠서 추정된 호모그래피이다.

4. 호모그래피 선택

제안된 방법에서는 최적의 호모그래피를 선택하기 위해서 고해상도의 입력 영상에서 특징점이 β 개 이하가 될 때까지 반복하여 해상도를 $1/\alpha$ 로 다운샘플링하고, 각 다운샘플링된 해상도별로 호모그래피를 추정한다. 이렇게 반복하여 구해진 n 개의 호모그래피 중 투영 오차가 최소가 되는 호모그래피를 최적이라고 판단한다. 이러한 방법을 사용할 수 있는 이유는 식 (8)과 표 1에서 보이듯이 원 영상(scale 1)에서 호모그래피를 추정하는데 걸리는 시간이 영상의 해상도를 줄여서 해상도 별로(scale 2~9) 다수의 호모그래피를 추정하는 시간보다 크기 때문이다.

$$t_1 > \sum_{i=\alpha}^n t_i \quad (8)$$

표 1. 영상의 스케일별 호모그래피 추정 시간 (초)

Table 1. Computational time at each scale of images. (sec.)

영상의 스케일	1	1/2	1/3	1/4	1/5	1/6	1/7	1/8	1/9	제안한 방법의 수행 시간
호모그래피 추정 시간	89.46	14.64	5.69	3.12	1.93	1.31	0.96	0.71	0.55	28.94

여기서 좌측 항은 다운샘플링을 하지 않은 방법의 수행 시간이며 우측 항은 n 번 다운샘플링하여 호모그래피를 추정한 수행 시간의 합을 의미한다. 해상도가 큰 영상일 수록 우측 항이 좌측 항보다 작으므로 반복수행하더라도 효율적으로 호모그래피를 추정할 수 있다. 최적의 호모그래피를 결정하기 위한 투영 오차는 식 (9)를 이용해서 구할 수 있다.

$$\varepsilon = \sum_{i=1}^4 |X_i^d - \widehat{X}_i^d| \quad \text{where} \quad \widehat{X}_i^d = \widehat{H}^d X_i^d \quad (9)$$

여기서 X_i 는 RASNAC을 통해서 나온 기준 영상의 4개의 특징점 좌표이다. \widehat{X}_i 은 기준 영상에 대응되는 대상 영상의 4개의 특징점 좌표가 호모그래피 행렬을 통해서 기준 영상의 어떤 점에 대응되는지를 나타내는 추정된 좌표이다.

5. 호모그래피 보정

다운샘플링된 영상에서 추정된 호모그래피를 보정 없이 원 영상에 적용하면 그림 2(c)와 같이 부정확하게 정합된 결과가 발생한다. 따라서 보정 행렬에 적용하여 호모그래피 행렬의 값들을 보정해 주어야 한다. 호모그래피 보정식은 DLT식을 이용하여 구하는데 앞 단계에서 설명한 DLT 특성을 적용하여 증명할 수 있다. DLT 식 (3)을 풀어서 쓰면 식 (10)과 같다.

$$\begin{bmatrix} 0 & 0 & 0 & -x_1 & -y_1 & -1 & x_1 y_1 & y_1 y_1 \\ x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1 x_1 & -y_1 x_1 \\ 0 & 0 & 0 & -x_2 & -y_2 & -1 & x_2 y_2 & y_2 y_2 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2 x_2 & -y_2 x_2 \\ 0 & 0 & 0 & -x_3 & -y_3 & -1 & x_3 y_3 & y_3 y_3 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3 x_3 & -y_3 x_3 \\ 0 & 0 & 0 & -x_4 & -y_4 & -1 & x_4 y_4 & y_4 y_4 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4 x_4 & -y_4 x_4 \end{bmatrix} \begin{bmatrix} h_1 \\ h_4 \\ h_7 \\ h_2 \\ h_5 \\ h_8 \\ h_3 \\ h_6 \end{bmatrix} = \begin{bmatrix} -y_1 \\ x_1 \\ -y_2 \\ x_2 \\ -y_3 \\ x_3 \\ -y_4 \\ x_4 \end{bmatrix} \quad (10)$$

이때 해상도를 줄인 영상에서 기준이 되는 점(x_i, y_i)과 대상이 되는 점(x'_i, y'_i)을 n 배씩 곱해주면 식 (11)과 같이 나타낼 수 있다.

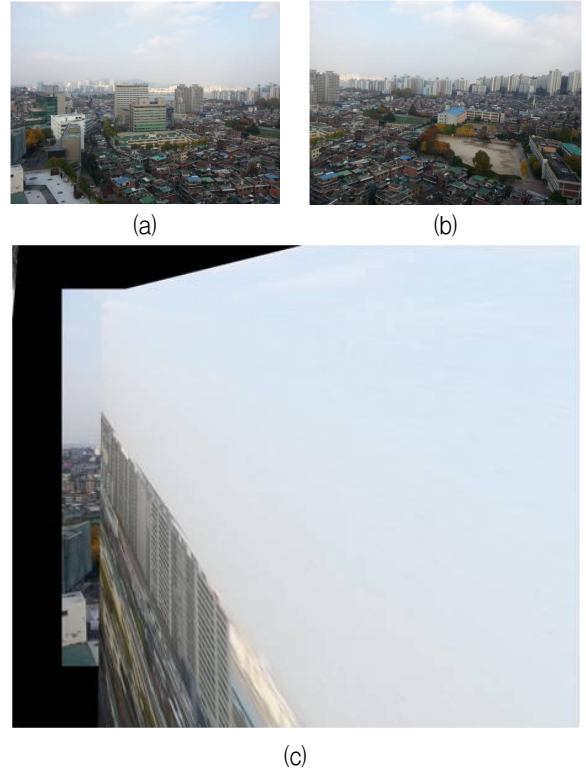


그림 2. 작은 영상에서 추정된 호모그래피를 보정 과정 없이 원 영상에 적용한 결과 (a) 기준 영상, (b) 대상 영상, (c) 원 영상에 맞지 않는 호모그래피를 적용하여 나온 결과

Fig. 2. An example without applying the homography compensation method: (a) Reference image, (b) Target image, and (c) Incorrect output without applying the proposed method.

$$\begin{bmatrix} 0 & 0 & 0 & -x_1 n & -y_1 n & -1 & x_1 y_1 n^2 & y_1 y_1 n^2 \\ x_1 n & y_1 n & 1 & 0 & 0 & 0 & -x_1 x_1 n^2 & -y_1 x_1 n^2 \\ 0 & 0 & 0 & -x_2 n & -y_2 n & -1 & x_2 y_2 n^2 & y_2 y_2 n^2 \\ x_2 n & y_2 n & 1 & 0 & 0 & 0 & -x_2 x_2 n^2 & -y_2 x_2 n^2 \\ 0 & 0 & 0 & -x_3 n & -y_3 n & -1 & x_3 y_3 n^2 & y_3 y_3 n^2 \\ x_3 n & y_3 n & 1 & 0 & 0 & 0 & -x_3 x_3 n^2 & -y_3 x_3 n^2 \\ 0 & 0 & 0 & -x_4 n & -y_4 n & -1 & x_4 y_4 n^2 & y_4 y_4 n^2 \\ x_4 n & y_4 n & 1 & 0 & 0 & 0 & -x_4 x_4 n^2 & -y_4 x_4 n^2 \end{bmatrix} \begin{bmatrix} \widehat{h}_1^d \\ \widehat{h}_4^d \\ \widehat{h}_7^d \\ \widehat{h}_2^d \\ \widehat{h}_5^d \\ \widehat{h}_8^d \\ \widehat{h}_3^d \\ \widehat{h}_6^d \end{bmatrix} = \begin{bmatrix} -y_1 n \\ x_1 n \\ -y_2 n \\ x_2 n \\ -y_3 n \\ x_3 n \\ -y_4 n \\ x_4 n \end{bmatrix} \quad (11)$$

여기서 \widehat{H}^d 는 영상의 해상도를 줄인 영상들 사이의

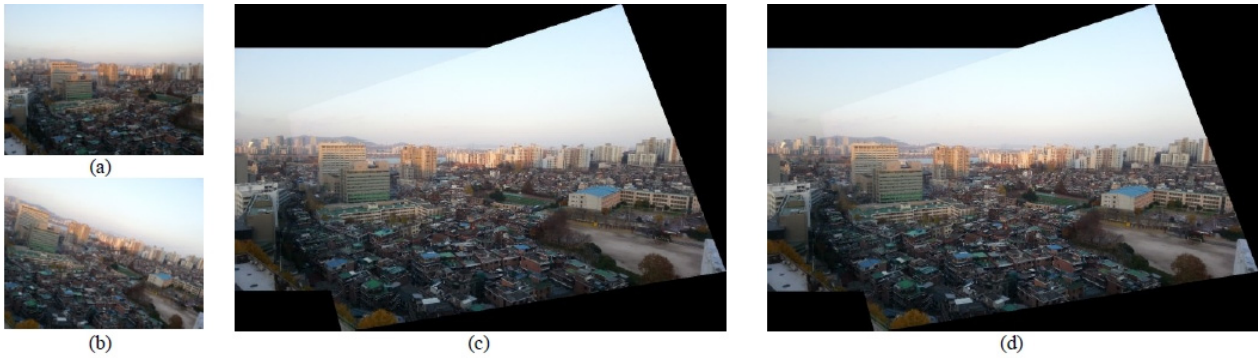


그림 3. 기존 방법과 제안한 방법 비교 (a) 기준 영상, (b) 대상 영상, (c) 기존 방법의 결과, (d) 제안한 방법의 결과
 Fig. 3. Comparative results with complicated scene contents: (a) Reference image, (b) Target image, (c) The general output from a conventional method, and (d) The proposed output.

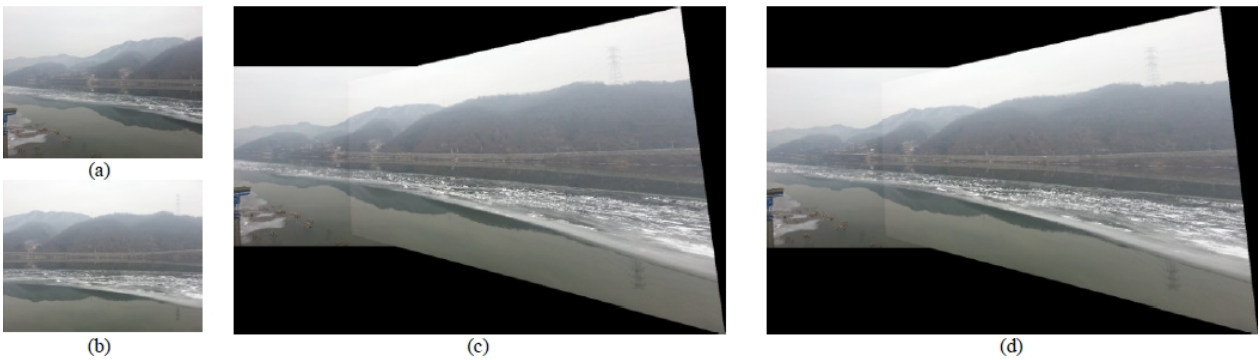


그림 4. 기존 방법과 제안한 방법 비교 (a) 기준 영상, (b) 대상 영상, (c) 기존 방법의 결과, (d) 제안한 방법의 결과
 Fig. 4. Comparative results with simple scene contents: (a) Reference image, (b) Target image, (c) The general output from a conventional method, and (d) The proposed output.

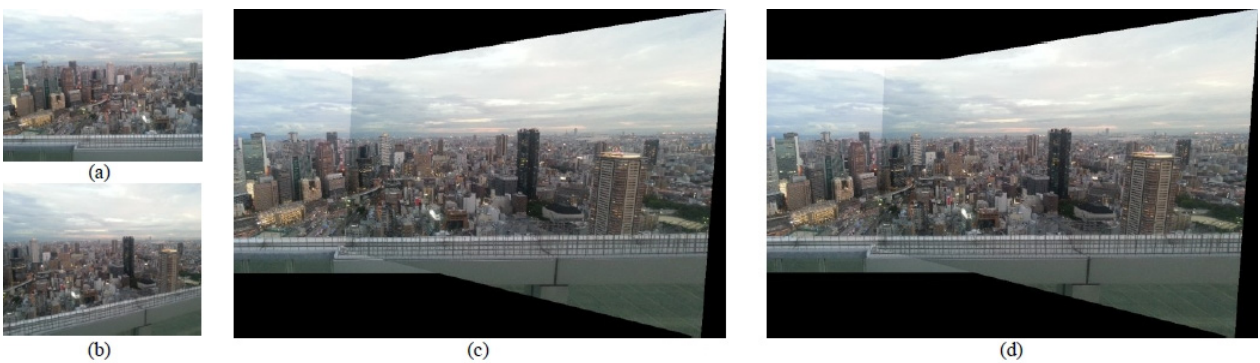


그림 5. 기존 방법과 제안한 방법 비교 (a) 기준 영상, (b) 대상 영상, (c) 기존 방법의 결과, (d) 제안한 방법의 결과
 Fig. 5. Comparative results with complicated contents and large size images: (a) Reference image, (b) Target image, (c) The general output from a conventional method, and (d) The proposed output.

관계에서 추정된 호모그래피이다. 3×3 행렬 \widehat{H}^d 에서 $(\widehat{h}_1^d, \widehat{h}_4^d, \widehat{h}_7^d)$ 는 첫 번째 열, $(\widehat{h}_2^d, \widehat{h}_5^d, \widehat{h}_8^d)$ 는 두 번째 열, $(\widehat{h}_3^d, \widehat{h}_6^d)$ 는 세 번째 열을 나타낸다. 이 때 A 행렬은 n 을

공통으로 갖게 된다. 따라서 A 행렬을 n 으로 나누어 주면 A 행렬의 3열과 6열에는 $1/n$ 이 남고 7열과 8열에는 n 이 남게 된다. 따라서 n 과 $1/n$ 을 \widehat{H}^d 행렬로 넘겨주면

\hat{h}_3^d 와 \hat{h}_6^d 는 n 배가 되고 \hat{h}_7^d 과 \hat{h}_8^d 은 $1/n$ 배가 된다. 이것을 식 (2)와 같이 정리하면 식 (12)와 같다.

$$H = \begin{bmatrix} \hat{h}_1^d & \hat{h}_2^d & n \cdot \hat{h}_3^d \\ \hat{h}_4^d & \hat{h}_5^d & n \cdot \hat{h}_6^d \\ 1/n \cdot \hat{h}_7^d & 1/n \cdot \hat{h}_8^d & \hat{h}_9^d \end{bmatrix} \quad (12)$$

식 (12)는 다시 호모그래피 보정식 (7)과 같이 나타낼 수 있다. 이런 과정을 통해서 기존의 방법보다 효율적으로 호모그래피를 추정할 수 있다. 호모그래피 보정 방법으로 그림 2와 같은 문제점을 해결할 수 있었으며 3.2M 픽셀의 고해상도의 영상을 정합한 경우, 8M 픽셀의 초고해상도의 영상에서 안개등의 영향으로 특징이 적은 경우와 일반적인 환경에서 특징이 많은 경우에 그림 3-5을 통해서 정확하게 정합된 것을 확인할 수 있다.

6. 색상 보정

파노라마 영상 생성을 위하여 중첩된 영역이 있는 영상을 촬영할 때 영상마다 색상 톤이 다를 수 있다. 그 이유는 영상을 촬영할 때 카메라의 시점에 변화를 줄 때와 촬영하는 사람의 기술에 따라서 카메라가 받아들이는 빛의 양이 변하기 때문이다. 색상 톤이 다른 영상들을 색상 보정 과정 없이 전 단계에서 구한 호모그래피를 이용하여 단순히 정합하면 그림 6(c)와 같이 색상 톤이 부자연스러운 결과를 얻게 된다. 따라서 영상을 정합한 다음 후처리 과정으로 색상 톤의 차이를 최소화할 수 있는 색상 보정 과정이 필요하다.

색상의 보정을 하기 위해서 식 (13-16)을 통해서 색상 공간을 RGB에서 사람의 시각에 더 가까운 CIELab으로 바꿔 준다^[16-17].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (13)$$

RGB 색상 공간을 CIELab으로 바꾸기 위해서는 2 단계의 과정을 거치는데 일단 RGB 색상 공간을 XYZ 색상 공간으로 바꾼 다음 다시 $L^* a^* b^*$ 로 바꿔준다.

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16 & \text{if } \frac{Y}{Y_n} > 0.008856 \\ 903.3 \left(\frac{Y}{Y_n} \right) & \text{if } \frac{Y}{Y_n} \leq 0.008856 \end{cases} \quad (14)$$

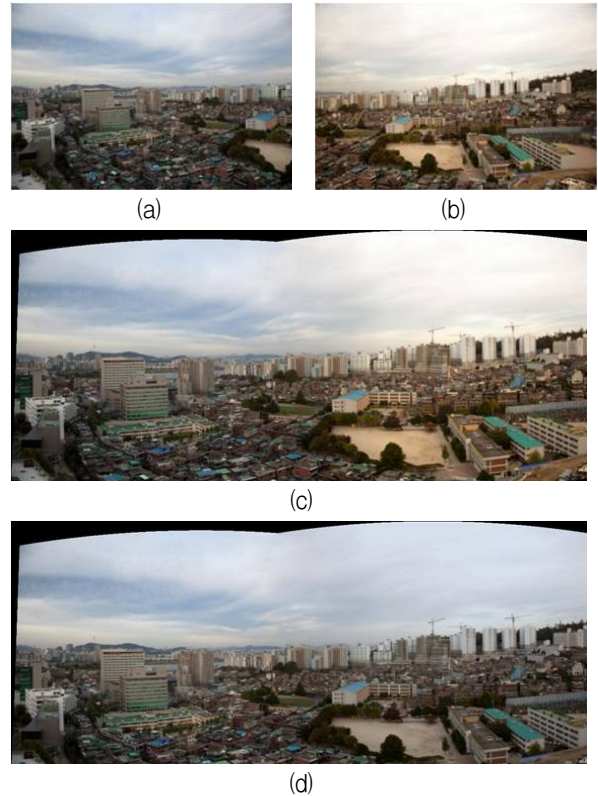


그림 6. 색상 톤의 차이가 있는 영상에서 색상 보정 알고리즘을 적용한 결과와 적용하지 않은 결과 비교 (a) 기준 영상, (b) 대상 영상, (c) 색상 보정을 하지 않은 결과, (d) 대상 영상의 색상을 기준 영상에 맞게 보정한 결과

Fig. 6. Comparison of color correction results: (a) Reference image, (b) Target image, (c) The result without applying a color correction method, and (d) The result with applying a color correction method.

$$a^* = 500 \left(f \left(\frac{X}{X_n} \right) - f \left(\frac{Y}{Y_n} \right) \right) \quad (15)$$

$$b^* = 200 \left(f \left(\frac{Y}{Y_n} \right) - f \left(\frac{Z}{Z_n} \right) \right) \quad (16)$$

색상 보정을 할 때 빠르게 하기 위해서 중첩 영역의 지역 정보만을 이용하여 색상 변환 행렬을 생성한다. 기본적으로 중첩 영역은 기준 영상과 대상 영상에 호모그래피를 적용한 영상을 “and” 연산을 통해서 추정한다. 중첩 영역을 찾으려면 기준 영상과 대상 영상의 중첩 영역의 Lab 채널 정보들을 이용하여 색상 변환 행렬을 생성한다. 이 때 중첩 영역의 정보를 이용하여 각각 $3 \times N$ 행렬을 생성한다. 그리고 식 (17)에 넣어서 색상

변환 행렬을 생성한다.

$$M_{ref} = AM_{tar} = \begin{bmatrix} L_1 L_2 & L_n \\ a_1 a_2 \cdots a_n \\ b_1 b_2 & b_n \end{bmatrix} = \begin{bmatrix} a_{11} a_{12} a_{13} \\ a_{21} a_{22} a_{23} \\ a_{31} a_{32} a_{33} \end{bmatrix} \begin{bmatrix} L'_1 L'_2 & L'_n \\ a'_1 a'_2 \cdots a'_n \\ b'_1 b'_2 & b'_n \end{bmatrix} \quad (17)$$

여기서 M_{ref} 와 M_{tar} 행렬은 기준 영상과 대상 영상의 중첩 영역의 정보를 가지고 있는 행렬이다. 그리고 A 행렬은 기준 영상과 대상 영상의 지역 정보를 이용하여 생성된 색상 변환 행렬이다. 이 때 색상이 선형으로 변한다고 가정하고 식 (5)를 이용하여 선형 방정식의 해를 구하면 색상 보정 행렬 A 를 생성할 수 있다. A 행렬을 통해서 그림 6(c)와 같이 색상 톤의 차이가 있을 때에 색상 톤을 보정하여 그림 6(d)처럼 자연스러운 결과를 얻을 수 있다.

III. 실험 결과

제안된 방법의 성능을 알아보기 위해서 3.2M 픽셀 및 8M 픽셀 등의 고해상도 영상을 1세대 i5-750 CPU, 8G RAM 환경에서 실험하였다. 다운샘플링을 위한 α 의 값은 2로 설정하였으며, 반복횟수를 결정하기 위한 β 의 값은 500으로 설정 하였다. 이때 영상은 픽셀들로 표현하기 때문에 α 는 2이상의 정수로 나타난다. 그림 3은 일반적인 환경에서 3.2M 해상도의 영상으로 실험한 결과이고 그림 4는 8M 해상도에서 안개의 영향으로 특징이 적게 있는 영상, 그림 5는 8M 해상도에서 특징이 많이 있는 영상에서 실험한 결과이다. 그림 3 ~ 그림 5에서 볼 수 있듯이 기존에 가장 좋은 성능을 보이는 SIFT와 RANSAC 알고리즘을 이용한 방법의 결과와 제안된 방법에 따라 생성된 파노라마 영상이 크게 다르지 않은 것을 확인할 수 있다. 표 2에서 확인할 수 있

듯이 제안한 방법으로 3.2M 영상에서 기존의 방법 대비 호모그래피를 3배 이상 빠르게 추정하고 8M 영상에서 5배 이상 빠르게 추정하였다. 영상의 복잡도가 커질수록 기존의 방법 보다 제안한 방법에서 더 효율적으로 호모그래피를 추정하는 것을 확인할 수 있었다. 그림 6에서는 색상 톤의 차이가 있는 영상을 입력영상으로 하는 경우 정합 후 기준 영상과 대상 영상의 색상 톤이 달라 부자연스러운 것을 볼 수 있지만 제안된 방법으로 색상을 보정 한 결과 더 자연스러운 결과를 만들 수 있음을 볼 수 있다.

IV. 결론

다수의 다운 샘플링 영상을 이용하여 다수의 호모그래피를 추정하고 최적의 호모그래피를 선택하여 보정하는 방법으로 호모그래피를 효율적으로 추정하는 방법과 중첩 영역의 지역적 정보를 이용하여 최소자승법에 기반을 둔 최적화 기법을 통하여 영상의 전반적인 색상을 효과적으로 보정하는 방법을 제안하였다. 제안한 방법으로 고해상도 영상에서 과도한 특징점의 추출로 인하여 계산량이 급격히 증가하는 문제를 해결하였으며 과도한 특징점의 추출로 인하여 RANSAC 알고리즘에서 에러가 포함된 데이터를 추출하여 부정확한 호모그래피가 추정되는 문제도 효율적으로 해결하였다. 그 결과 3.2M 픽셀 해상도의 영상에서 호모그래피를 3배 이상 빠르게 추정할 수 있었으며 8M 픽셀 해상도의 영상에서 5배 이상 빠르게 추정할 수 있었다. 색상 톤의 차이가 있는 영상이 입력으로 들어온 경우에도 중첩 영역의 지역 정보를 이용하여 색상을 빠르게 보정할 수 있었다. 정합 결과 면에서도 기존에 가장 좋은 성능을 보인다고 알려진 SIFT와 RANSAC 알고리즘을 이용한 결과와 같은 결과를 얻을 수 있었다.

표 2. 기존 방법과 제안된 방법의 호모그래피 계산 시간 비교 (초)

Table 2. Computational time comparisons for homography estimation. (sec.)

사용된 영상의 특징	2048×1536 해상도 (많은 개수의 특징점)	3264×2448 해상도 (적은 개수의 특징점)	3264×2448 해상도 (많은 개수의 특징점)
SIFT + RANSAC	64.02	133.09	422.42
PROPOSED	19.97	24.13	76.74

REFERENCES

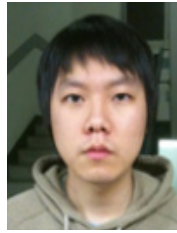
- [1] R. Szeliski, *Image Alignment and Stitching: A Tutorial, Handbook of Mathematical Models in Computer Vision*, Springer, 2005.
- [2] D. Lowe, "Distinctive image features from scales-invariant keypoints," *International Journal of Computer Vision*, vol.60, no.2, pp. 91-110, November 2004.
- [3] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography," *Communication of the ACM*, Vol. 24, No. 6, pp. 381-395, June 1981.
- [4] O. Chum and J. Matas, 'Randomized RANSAC with Td,d test,' *Proc. of the British Machine Vision Conference*, pp. 448 - 457, London, UK, September 2002.
- [5] O. Chum and J. Matas, "Matching with PROSAC - progressive sample consensus," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 220 - 226, San Diego, CA, June 2005.
- [6] S. Y. Ye, A. Y. Jeon, G. R. Jeon, and K. G. Nam, "EMSAC 알고리즘을 이용한 대응점 추출에 관한 연구," *전자공학회 논문지-SP*, 제44권, 제4호, 44-50쪽, 2007.
- [7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge university Press 2nd edition, 2004.
- [8] M. Brown and D. Lowe, "Recognising Panorama," *International Conference on Computer Vision*, Vol. 2, pp. 1218-1225, October 2003.
- [9] M. Brown and D. Lowe, "Invariant Features from Interest Point Groups," *British Machine Vision Conference*, pp. 656-665, September 2002.
- [10] M. Brown and D. Lowe, "Automatic panorama image stitching using invariant features," *International Journal of Computer Vision*, Vol. 74, No. 1, pp. 59-73, August 2007.
- [11] S. H. Lee and M. Y. Kim, "다중 카메라 기반 대 영역 고해상도 영상 획득 시스템과 실시간 영상 정합 알고리즘," *전자공학회 논문지-SC*, 제49권, 제4호, 10-16쪽, 2012.
- [12] Y. H. Kim and S. K. Lee, "A Simple and Effective Image Color Balancing for HD-to-UHD Conversion," *International Conference on Electronics Information and Communication*, January 2013.
- [13] E. Vincent and R. Laganier, "Detecting planar homographies in an image pair," *Image and Signal Processing and Analysis*, pp. 182-187, June 2001.
- [14] G. H. Golub and C. F. Van Loan, *Matrix Computation*. JHU Press, 1996.
- [15] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, *Introduction to Algorithms*, Prentice Hall India, 2002.
- [16] Y. S. Han, S. H. Lee, J. S. Choi, and S. K. Lee, "A simple and efficient color recovering system for content sharing website," *IEEE Transaction on Consumer Electronics*, Vol. 56, No. 2, pp. 863-869, May 2010.
- [17] S. H. Lee and J. S. Choi, "Design and Implementation of Color Correction System for Images Captured by Digital Camera," *IEEE Transactions on Consumer Electronics*, Vol. 54, No. 2, pp. 268-276, May 2008.

저 자 소 개



서 상 원(학생회원)
 2011년 강남대학교 전자공학과
 학사 졸업.
 2011년~현재 중앙대학교 첨단
 영상대학원 영상학과
 석사 과정.

<주관심분야 : 컴퓨터 비전, 컬러 보정>



정 수 응(학생회원)
 2010년 남서울대학교 멀티미디어
 학과 학사 졸업.
 2012년 중앙대학교 첨단영상대학
 원 영상학과 석사 졸업.
 2012년~현재 중앙대학교 첨단영
 상대학원 영상학과
 박사 과정.

<주관심분야 : 영상 개선, 영상 검색, 패턴 인식>



한 윤 상(학생회원)
 2006년 한서대학교 컴퓨터정보
 공학과 학사 졸업.
 2009년 중앙대학교 첨단영상대학
 원 영상학과 석사 졸업.
 2009년~현재 중앙대학교 첨단
 영상대학원 영상학과
 박사 과정.

<주관심분야 : super-resolution, 컴퓨터 비전, 컬러 복원, 영상처리>



최 중 수(평생회원)
 1975년 인하대학교 전기공학과
 학사 졸업.
 1977년 서울대학교 전자공학과
 석사 졸업.
 1981년 게이오대학교 전기공학과
 박사 졸업.

1981년~1981년 Aloka Co. Ltd 연구소 연구원.
 1981년~1999년 중앙대학교 전자공학과 교수
 1999년~현재 중앙대학교 첨단영상대학원
 영상학과 교수.

<주관심분야 : 영상신호처리, 컴퓨터 비전, 증강 현실, 영상합성, 영상통신>



이 상 근(평생회원) -교신저자
 1996년 중앙대학교 전자공학과
 학사 졸업.
 1999년 중앙대학교 전자공학과
 석사 졸업.
 2003년 조지아공과대학교 전기 및
 컴퓨터 공학과 박사 졸업.

2003년~2008년 Samsung Information and
 System America, Irvine, CA.
 Staff Research Engineer.

2008년~현재 중앙대학교 첨단영상대학원
영상학과 교수

<주관심분야 : 멀티미디어 신호 및 통신, 영상 압축, 영상 개선 및 복구, 영상 검색, 컬러 보정>