

실린더 파노라마 영상의 재구성을 위한 단순화된 사영 변환

이강호*

Simplified projective transform for reconstruction of cylindrical panorama

Kang-Ho Lee *

요약

본 논문은 고정된 카메라를 수평 이동(panning)시켜 얻어진 영상으로부터 단순화된 변환식을 이용 실린더 파노라마 영상(cylindrical panorama)을 재구성하는 방법을 제안한다. 일반적인 영상간의 사영 변환식(projective transform)은 3차원 공간의 X, Y, Z축에 대한 카메라의 회전 성분을 모두 고려하지만, 실제 파노라마 영상을 구성하기 위해서는 Y축 중심의 카메라 회전만을 고려한다. 이러한 제약은 기존의 8자유도(degree of freedom)를 갖는 사영 변환식을 4자유도를 갖는 사영 변환식으로 단순화 시키고, 그 결과 변환식을 유도하기 위해 필요한 대응점의 개수가 절반으로 줄어들어 전체적인 계산량을 크게 감소시킬 수 있다. 모의실험 결과 제안된 알고리즘은 기존의 알고리즘과 계산량을 감소시키면서 유사한 성능을 보이고 있고, 단순화된 사영 변환식을 이용한 영상을 구성하는 것을 보여주었다.

Abstract

In this paper we propose a method of reconstruction of cylindrical panorama using simplified projective transform from the panning image on the fixed camera. For the practical construction of cylindrical panorama we consider the rotation of the camera on the Y-axis only, even though considering the rotation components on all of the X,Y,Z axis on three-dimensional space for projective transform between general panoramas. The restriction mentioned above simplifies projective transform with existing 8 degrees of freedom into the one with 4 degrees of freedom. In the results, overall computation for projective transform can be decreased to the great extents in quantity, because the number of corresponding points required for inducing the transforming formula is gone down by half. Proposed algorithm from the simulation carried out in this paper shows similar performance and decreased computational quantity compared with existing algorithm. Also, it shows the construction of cylindrical panorama using simplified projective transform.

▶ Keyword : 실린더 파노라마 영상(cylindrical panorama), 사영 변환식(projective transform) 8자유도 (degree of freedom)

• 제1저자 : 이강호
• 접수일 : 2006.03.3, 심사완료일 : 2006.05.18
* 한국재활복지대학 정보보안과 교수

1. 서론

가상현실 시스템(virtual reality system)에서 가상환경은 3차원 그래픽 기반으로 만들어져 있어 복잡한 모델링(modeling) 작업과 실시간으로 가상환경을 만들어내기 위한 특별한 용도의 값비싼 하드웨어가 필요하며, 영상의 복잡도(complexity)가 제한되어 현실감있는 가상환경 표현이 어려웠다[1][2][3][4][5]. 이에 비해 실세계의 환경으로부터 표현된 파노라마(panorama) 영상을 이용하여 가상환경을 표현하게 되면 특별한 용도의 값비싼 하드웨어의 필요 없이 실시간 표현이 가능하게 되었다.

실세계의 환경으로부터 파노라마 영상을 생성하는 방법들은 하드웨어적인 방법과 소프트웨어적인 방법이 있다. 하드웨어적인 방법은 어안(fish eye) 렌즈와 같은 넓은 시각 영역(field of view)을 가지는 렌즈를 이용하거나 파노라마 카메라를 사용하여 영상을 직접 표현하는 방법이다. 소프트웨어적인 방법은 이미지 모자이크(image mosaics) 방법이 있다. 이미지 모자이크 방법은 카메라나 비디오로부터 얻어진 영상들을 정렬시킨 다음, 정렬된 영상을 하나의 커다란 시각 영역을 가지는 파노라마 영상으로 재구성(reconstruction)하는 작업이 필요하다[5][6].

실린더 파노라마(cylindrical panorama) 영상은 운동시차(motion parallax)가 영상들 사이에 나타나지 않게 하기 위해 카메라를 삼각대에 고정시켜 놓은 다음, 수평(panning) 방향으로 회전시켜 얻어진 영상을 이용하여 표현한다[5]. 카메라의 회전각(panning angle)과 초점거리를 정확히 알고 있는 경우는 각 영상을 실린더에 투영하여 파노라마 영상을 생성할 수 있다. 그러나 이들을 알지 못하는 경우 카메라의 내부 요소인 초점거리(focal length)를 먼저 추정하고, 영상을 실린더에 와핑(warping)시킨 후, 많은 반복 계산에 의해 영상들 간의 움직임 요소를 복원하여 파노라마 영상을 생성할 수 있다.

일반적인 영상간의 사영 변환식(projective transform)은 8자유도(degree of freedom)를 갖는다. 따라서 변환식을 유도하기 위해서는 4쌍의 대응점이 필요하다[1][2][3][5]. 그러나 카메라의 수평 이동(panning)만을 고려하는 파노라마 영상의 구성의 경우 Y축 중심의 카메라 회전만을 고려하

기 때문에 변환식의 자유도를 절반으로 줄일 수 있다. 따라서 본 논문에서는 단순화된 사영 변환식을 이용해서 실린더 파노라마 영상을 재구성하는 방법을 제안한다. 영상의 특징점 및 대응점 검출은 유사 특징점(quasi feature point)과 변형된 블록 정합 알고리즘(modified block matching algorithm)을 이용한다. 제안된 알고리즘은 다양한 영상에 적용하여 그 성능을 평가하였고, 그 결과 계산량은 크게 감소하였으며 기존의 8자유도를 갖는 사영 변환식을 이용하는 경우와 유사한 성능을 보임을 확인하였다.

II. 중첩 영역 및 특징점 추출

2.1 위상 상관을 이용한 중첩 영역 추출

연속된 두 영상의 중첩된 영역을 추출하기 위해서는 영상을 주파수 공간(frequency domain)으로 푸리에 변환(Fourier transform)한 후 식(1)처럼 cross power spectrum을 계산하여 두 영상의 변위(displacement)를 구한다. 이 변위 정보로부터 두 영상의 중첩 영역을 추출할 수 있다[3][7][10].

$$\frac{F_1(u, v)F_2^*(u, v)}{|F_1(u, v)F_2^*(u, v)|} = e^{-j2\pi(ux_0+vy_0)} \dots\dots\dots (1)$$

2.2 유사 특징점 추출 및 대응점 검출

영상에서 텍스처(texture) 정보가 풍부한 영역은 그렇지 않은 다른 영역과 구별되는 특징을 갖는다. 유사 특징점은 2.1에서 추출된 기준 영상의 중첩 영역을 4개의 부영역(subarea)으로 나눈 뒤, 각 영역에서 특징점을 추출했다. 그리고 그 대응점은 주변 조영 조건의 변화와 카메라의 움직임으로 인해 발생하는 영상의 왜곡(distortion)을 고려한 블록 정합 알고리즘을 이용해서 검출했다. 본 논문에서는 유사 특징점과 블록 정합 알고리즘을 이용하고 4자유도의 사영 변환식을 이용하기 때문에 특징점은 4개의 부영역에서 임의로 선택한 2개의 특징점을 이용한다. 식(2), (3), (4)는 특징점 및 대응점 검출에 대한 식을 나타내고 있다.

$$q_i = \arg \max[\sigma^2_{k,i}], 0 \leq i \leq 1$$

$$\sigma^2_{k,i} = \sum_k \left[\sum_g^{G_{MAX}} g \cdot h_g - M_k \right] \dots\dots\dots (2)$$

q_i 는 i 번째 부 영역에서 추출된 특징점이고, $\sigma^2_{k,i}$ 는 i 번째 부 영역에서 k 번째 블록의 국부 분산(local variance)을 나타낸다.

$$E(d_x, d_y) = \sum_{(x,y) \in B} |I(x,y,m) - I(x,y,m+1)| \cdot w_{i,j} \dots\dots\dots (3)$$

$$[\hat{d}_x, \hat{d}_y] = \arg \min_{(d_x, d_y)} [E(d_x, d_y)] \dots\dots\dots (4)$$

$E(\cdot)$ 는 대응점 정합에 사용된 평가함수이고, $[\hat{d}_x, \hat{d}_y]$ 는 최소 오차를 갖는 정합점의 변위값을 나타낸다.

III. 평면 사영 변환식

그림 1은 3차원 공간상의 한 점 P 가 시선각 (view point)이 다른 두 장의 영상으로부터 투영되는 평면 사영 변환(planar projective transform)의 관계를 나타내고 있다.

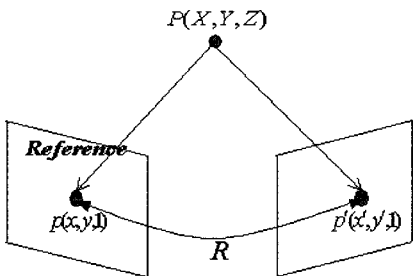


그림 1. 사영 변환의 원리
Fig 1. Fundamental of projective transform

3.1 일반적인 사영 변환식

두 점 $p(x,y)$, $p'(x,y)$ 의 변환 관계식은 식(5)~(7)로부터 얻어지며, V_0, V_1 은 각 카메라의 내부 파라미터 (intrinsic parameter)를 나타낸다.

$$p = V_0 \cdot P \quad \Leftrightarrow \quad P = V_0^{-1} \cdot p \dots\dots\dots (5)$$

$$p' = V_1 \cdot R \cdot P = V_1 \cdot R \cdot V_0^{-1} \cdot p = M_{projective} \cdot p \dots\dots\dots (6)$$

$$M = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \dots\dots\dots (7)$$

M 은 일반적인 8자유도(degree of freedom)를 갖는 사영 변환 행렬이다.

3.2 단순화된 사영 변환식 유도

식(6)에서 회전 행렬은 식(8)과 같이 3차원 공간상에서 X, Y, Z축 대한 모든 회전을 고려한다.

$$R = R_X \cdot R_Y \cdot R_Z \dots\dots\dots (8)$$

그러나 실린더 파노라마(cylindrical panorama)를 구성하기 위한 영상의 취득 단계에서는 영상의 운동 시차(motion parallax) 성분을 감쇄시키기 위해 카메라를 삼각대위에 고정시키고, 수평 방향으로 이동(panning)만을 고려한다. 즉, Y축을 중심으로 하는 회전만을 고려하기 때문에 식(8)에서 R_X, R_Y 는 단위 행렬이 된다. 따라서 식(6)은 식(9)와 같이 변형된다.

$$M = V_1 \cdot R_1 \cdot V_0^{-1}$$

$$= \begin{bmatrix} f_1 & 0 & 0 \\ 0 & f_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} 1/f_0 & 0 & 0 \\ 0 & 1/f_0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} f_1/f_0 \cos\beta & 0 & f_1 \sin\beta \\ 0 & f_1/f_0 & 0 \\ -\sin\beta/f_0 & 0 & \cos\beta \end{bmatrix} \dots\dots\dots (9)$$

β 는 Y축을 중심으로 카메라의 회전 각도(panning angle)를 나타낸다. $\beta \neq n\pi/2, (n=1,2,3,\Lambda)$ 이면 식(9)는 식(10)처럼 정규화된 형태(normalized form)로 표현할 수 있다.

$$M' \sim \begin{bmatrix} f_1/f_0 & 0 & f_1 \tan \beta \\ 0 & f_1/f_0 \cos \beta & 0 \\ -\tan \beta/f_0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} m_0 & 0 & m_1 \\ 0 & m_2 & 0 \\ m_3 & 0 & 1 \end{bmatrix} \dots\dots\dots (10)$$

식(10)의 M' 과 식(7)의 M 을 비교하면 필요한 파라미터의 개수가 1/2로 줄어든 것을 알 수 있다. 즉, 4자유도의 사영 변환식으로 변형되었고, 결론적으로 파라미터를 계산하기 위해 필요한 특징점의 개수 역시 1/2로 감소하였다. 대응되는 한 쌍의 점을 $(x_k, y_k), (x'_k, y'_k)$ 라 하면, 두 점 사이의 관계는 식(11)과 같이 정의된다.

$$\begin{aligned} m_0 x_k + m_1 - m_3 x_k x'_k &= x'_k \\ m_2 y_k - m_3 x_k y'_k &= y'_k, 0 \leq k \leq 1 \end{aligned} \dots\dots\dots (11)$$

2장에서 추출한 2쌍의 대응점을 식(11)에 대입하면 vector matrix 형태의 선형 방정식이 주어진다. 이 방정식의 해를 구하면 식(10)의 4개 파라미터를 얻을 수 있다.

IV. 실린더 파노라마

실린더 파노라마(cylindrical panorama)는 구성의 용이함 때문에 일반적으로 많이 이용된다. 그림 2는 실린더 파노라마의 구성 원리와 구성 과정에 대한 Flow Chart를 보여준다. 실린더 파노라마는 운동시차(motion parallax)가 영상들 사이에 나타나지 않기 위해 카메라를 삼각대 위에 고정시켜 놓은 다음, 수평(panning)방향으로 회전시켜 얻은 영상을 취득한다. 취득된 영상에서 카메라의 내부 요소인 초점거리를 먼저 추정하고, 영상을 실린더 좌표계로 변환시킨 후, 반복 계산에 의해 영상들 간의 움직임 요소를 복원하여 파노라마 영상을 구성한다.

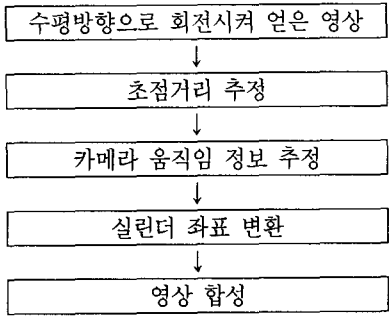
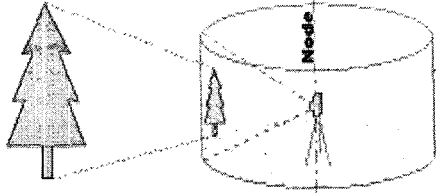


그림 2. 실린더 파노라마의 구성 원리와 Flow Chart
Fig 2. Fundamental of cylindrical panorama and Flow Chart

4.1 초점거리 추정

회전 행렬의 처음 두 행(또는 열)은 서로 같은 크기의 절대값(norm)을 갖고, 서로 직교(orthogonal)하는 성질을 갖는다[4]. 따라서 식(7)에 이러한 성질을 적용하면 카메라의 초점거리를 추정할 수 있다. 식(12)는 이러한 관계를 설명한다.

$$f_0 = \sqrt{\frac{m_5^2 - m_2^2}{m_0^2 + m_1^2 - m_3^2 - m_4^2}} \quad \text{or} \quad f_0 = \sqrt{\frac{m_2 m_5}{m_0 m_3 + m_1 m_4}} \dots\dots\dots (12)$$

마찬가지 방법으로 f_1 을 계산할 수 있고, 만약 고정된 초점거리를 가정한다면 f_0, f_1 의 기하 평균(geometrical mean)을 초점거리로 취한다.

$$f = \sqrt{f_0 \cdot f_1} \dots\dots\dots (13)$$

그러나 식(13)은 일반적으로 8자유도를 갖는 사영 변환식으로부터 추정된 초점거리다. 본 논문에서 정의한 4자유도에 대한 사영 변환식에 대한 초점거리 역시 유사한 방법을 이용해서 추정할 수 있다. 식(10) 역시 처음 두 개의 행

(또는 열)은 식(14)처럼 같은 절대값을 갖고, 서로 직교하는 성질을 갖는다.

$$m_0^2 + m_1^2 / f_0^2 = m_2^2, m_0^2 + m_3^2 / f_1^2 = m_2^2 \dots\dots\dots (14)$$

따라서 식(10)에서 추정된 초점거리 f_0, f_1 은 식(15)와 같이 정의되고, 식(13)과 같이 기하 평균을 추정된 초점거리로 사용한다.

$$f_0 = \sqrt{\frac{m_1^2}{m_2^2 - m_0^2}}, f_1 = \sqrt{\frac{m_3^2 - m_0^2}{m_2^2}} \dots\dots\dots (15)$$

4.2 실린더 좌표 변환

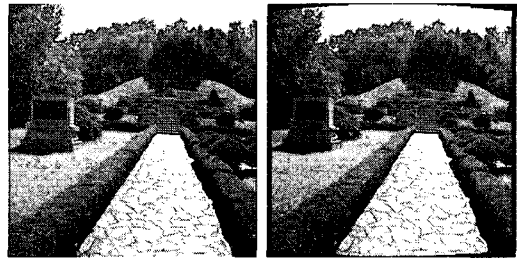
실린더 파노라마 영상을 구성하기 위해서는 세계 좌표 $P(X, Y, Z)$ 는 2차원 실린더 좌표계로 변환되어야 한다. 이러한 관계는 식(16)과 같다.

$$\theta = \tan^{-1}(X/Z), v = Y/\sqrt{X^2 + Z^2} \dots\dots\dots (16)$$

θ 는 카메라의 회전 각도, v 는 스캔 라인(scan line)을 나타낸다[5]. 그러나 우리가 알고 있는 것은 영상 좌표 $p(x, y)$ 이기 때문에 식(16)은 영상 좌표계의 변환식으로 변환되어야 한다. 카메라 방정식 $x = f \cdot X/Z, y = f \cdot Y/Z$ 를 식(16)에 대입하고 정리하면 식(17)과 같이 직교 좌표계에 대한 변환식이 정의된다.

$$\theta = \tan^{-1}(x/f), v = y/\sqrt{x^2 + f^2} \dots\dots\dots (17)$$

식(17)을 이용하면 직교 좌표계의 영상을 실린더 좌표계의 영상으로 변환시킬 수 있다. 그림 3은 변환된 결과를 보여주고 있다.



(a) (b)
 그림 3. 실린더 좌표 변환의 예 :
 (a) 원 영상, (b) 실린더 좌표 변환 영상
 Fig 3. Example of cylindrical transform
 (a) Original image (b) Cylindrical transform image

4.3 카메라 움직임 정보 추정

그림 3과 같이 식(17)을 이용해서 입력 영상이 실린더 좌표계로 변환되면 영상간에는 수평 또는 수직 방향으로의 변위 성분만이 존재한다. 따라서 실린더 파노라마로 구성하기 위해서는 변환된 영상들간의 이동 변위를 계산해야 한다. 본 논문에서는 2장에서 입력 영상간의 중첩 영역을 추출하기 위해 사용된 영상의 위상 상관(phase correlation)을 이용한다. 위상 상관은 주파수 공간에서 최적의 정합을 탐색할 뿐만 아니라 계산량이 적고, 다른 정합 방법에 비해서 잡음에 강건한 성질을 갖는다[1][2][7].

4.4 영상 합성

영상간의 사영 변환식과 추정된 카메라 초점거리를 이용하여 실린더 좌표계로 변환된 영상에 대해서 우리는 두 영상을 하나의 영상으로 합성할 수 있다. 즉, 유도된 사영 변환식을 이용해서 대상 영상을 기준 영상으로 와핑(warping)시킬 수 있고, 와핑된 영상은 기준 영상과 합성될 수 있다. 대상 영상을 사영 변환식을 이용해서 와핑시키는 경우, 영상의 와핑 맵(warping map)은 여러 가지 방법으로 계산할 수 있다. 본 논문에서는 역 매핑(inverse mapping) 방법을 이용함으로써 와핑된 영상에 대해서 보간(interpolation)의 필요성을 제거한다[9].

합성된 영상에서 시각적인 결함(visual artifacts)을 없애기 위해서 그림 4에서와 같은 과도기 함수(transient function)를 적용한 선형 가중치 함수를 정의한다. 이 가중치 함수는 영상간의 경계선 제거를 위해 많이 사용되는 방법 중에 하나이다. 식(18)에서 이 함수를 정의한다.

$$I_{blend} = W_A \cdot I_{Reference} + W_B \cdot I_{Target}$$

$$W_A + W_B = 1.0 \quad \dots\dots\dots (18)$$

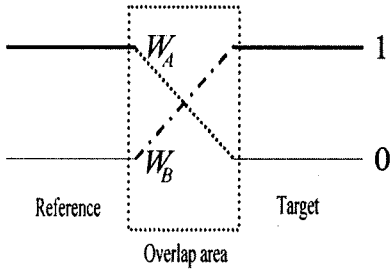


그림 4. 과도기 함수를 이용한 가중치 함수
Fig 4. Bilinear weighting function using the transient function

V. 모의실험 결과

3.2와 4.1에서 유도한 4자유도의 사영 변환식을 적용하여 실린더 파노라마 영상을 재구성하는 모의 실험을 하였다. 그림 5는 실험에 사용된 영상은 5프레임으로 구성되었으며 영상 취득은 디지털 카메라를 이용하여 취득한 320×256의 RGB 칼라 영상을 사용하였다.



그림 5. 실험 영상
Fig 5. Experiment images

제안된 알고리즘의 성능을 비교하기 위해 기존의 논문에서 최적화를 위해 사용되었던 평가함수를 이용한다[1][2][3][4]

(5). 평가함수는 식(19)에 정의되어 있다.

$$E = \frac{1}{N} \sum_R [I'_{ref}(x, y) - I'_{tar}(x, y)]^2 \quad \dots\dots\dots (19)$$

N은 중첩 영역 R에서 화소의 개수를 나타내고, I_{ref}와 I_{tar}은 각각 실린더 좌표계로 변환된 기준 영상과 대상 영상을 나타낸다. 표1은 위 식을 이용해서 주어진 영상의 각 프레임간의 성능을 평가한 결과를 나타낸다.

표 1. 단순화된 사영 변환식을 이용한 실린더 파노라마 영상의 성능 평가

Table 1. Performance comparision of cylindrical panorama image using the simplified projective

구분	E	
	8자유도 사영 변환식	4자유도 사영 변환식
1 2 frame	42	46
2 3 frame	32	34
3 4 frame	43	42
4 5 frame	20	23

실험 결과를 살펴보면 제안된 4자유도를 갖는 사영 변환식을 이용하여 실린더 파노라마 영상을 구성한 결과는 기존의 8자유도를 갖는 사영 변환식을 이용하는 경우와 유사한 성능을 보이고 있음을 알 수 있다. 또한 기존의 8자유도를 이용하는 경우 한 쌍의 영상에 대해서 전체 처리 시간이 평균 5초가 걸렸지만 제안된 방법은 3초의 시간이 걸렸다. 결과적으로 제안된 방법은 기존의 방법과 유사한 성능을 보이면서 보다 빠른 연산을 가능하게 한다. 그림 6은 제안한 방법을 이용해서 구성한 실린더 파노라마 영상이며, 시각적으로도 좋은 결과를 나타내고 있다.

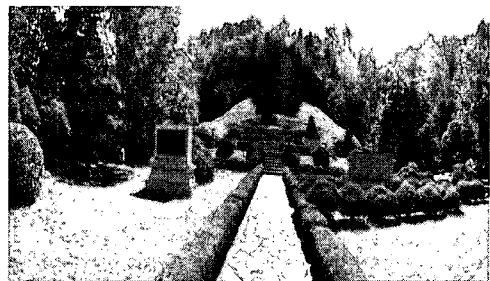


그림 6. 실린더 파노라마 영상으로 재구성된 결과
Fig 6. Reconstructed result of cylindrical panorama

VI. 결 론

본 논문에서는 4자유도(degree of freedom)의 사영 변환식을 이용하여 실린더 파노라마 영상을 구성하는 방법을 제시하였다. 기존의 8자유도의 사영 변환식을 유도하기 위해서는 4쌍의 대응점이 필요했지만, 제안한 방법은 2쌍의 대응점만을 이용하기 때문에 전체적인 계산시간을 줄일 수 있었으며, 8자유도의 변환식을 이용하는 경우와 유사한 성능을 나타내고 있다. 실험 결과를 통해서도 제안한 방법이 시각적으로도 만족할 만한 결과를 나타내고 있음을 확인할 수 있었다.

참고문헌

[1] 김대현, 최중수, "유사 특징점을 이용한 모자이크 영상의 구성", 전자공학회 논문지 제38권 4호 PP. 381-391, 2001.

[2] 김대현, 윤용인, 최중수, "변형된 블록 정합을 이용한 이미지 모자이크", 제13회 신호처리합동학술대회 Vol.13, No.1, pp.393-396, 2000.

[3] R. Szeliski, "Image Mosaicing for TeleReality applications", IEEE Workshop on Applications of Computer Vision (WACV'94), pp. 44-53, Florida, Dec. 1994.

[4] Z. Imad, O. Faugeras, R. Deriche, "Using geometric corners to build a 2D mosaic from a set of images", Proc. 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.420-425, Puerto Rico, June, 1997.

[5] R. Szeliski, H. Y. Shum, "Creating full view panoramic image mosaics and environment maps", Computer Graphics (SIGGRAPH'97), pp.251-258, August 1997.

[6] S. E. Chen, "QuickTime VR an Image based Approach to Virtual Environment Navigation", Computer Graphics (SIGGRAPH'95), pp. 29-38, Aug. 1995.

[7] B. S. Reddy, B. N. Chatterji, "An FFT Based Technique for Translation, Rotation, and Scale Invariant Image Registration", IEEE Trans. On Image Processing, Vol.5 No.8 pp.1266-1271, Aug. 1996.

[8] E. Trucco, A. Verri, Introductory Techniques for 3D Computer Vision, Prentice Hall, Upper Saddle River, 1998.

[9] J. Gemes, L. Velho, Image Processing for Computer Graphics, S Springer, New York, 1997.

[10] 차정희, "DEM과 산영상을 이용한 비전기반 카메라 위치인식" 한국컴퓨터정보학회 논문지 vol.10, No.6, pp.177-185, 2005.

저 자 소 개



이 강 호

1986. 2. 중앙대학교 전자공학과 졸업(공학석사)
 1991. 8. 중앙대학교 전자공학과 졸업(공학박사)
 1990. 3. 대덕대학 인터넷정보과 교수
 2000. 3. 송호대학 멀티미디어과 교수
 2003. 3. ~ 국립한국재활복지대학 정보보안과 교수
 2006. 3. ~ 한국컴퓨터정보학회 회장
 <관심분야> 정보보안, 디지털 영상 처리

