

특집논문-06-11-4-02

변수화된 영상 보정을 통한 전방향 영상 생성 방법

김 유 나^{a)}, 심 동 규^{a)†}

Omni-directional Image Generation Algorithm with Parametric Image Compensation

Yu-na Kim^{a)}, and Dong-gyu Sim^{a)†}

요 약

본 논문에서는 변수화된 영상 보정을 통한 전방향 영상 (Omni-directional image) 생성 방법을 제안한다. 제안한 방법은 모든 방향의 영상을 포함하고 영상 합성 시에 발생하는 왜곡과 에러를 최소화하는 전방향 영상을 생성하기 위해, 원형 좌표계를 기반으로 평면 영상을 원형 영상으로 변환한다. 카메라로 획득한 각각의 영상은 카메라의 시스템과 외부 환경 조건에 따라 발생하는 비그네이팅 효과와 조명 변화를 포함하고 있으므로, 최적의 비그네이팅 및 조명 변수를 추출하여 자연스러운 전방향 영상 생성을 위한 영상 보정 과정을 수행한다. 실험 결과에서는 제안한 방법에 따라 생성된 전방향 영상을 확인하고 비그네이팅 효과 및 조명 변화를 고려하지 않은 방법과 제안한 방법의 성능을 비교, 분석하여 제안한 방법이 1~4dB 정도 더 효과적임을 보인다.

Abstract

This paper proposes an omni-directional image generation algorithm with parametric image compensation. The algorithm generates an omni-directional image by transforming each planar image to the spherical image on spherical coordinate. Parametric image compensation method is presented in order to compensate vignetting and illumination distortions caused by properties of a camera system and lighting condition. The proposed algorithm can generate realistic and seamless omni-directional video and synthesize any point of view from the stitched omni-directional image on the spherical image. Experimental results show that the proposed omni-directional system with vignetting and illumination compensation is approximately 1~4dB better than that which does not consider the said effects.

Key Words : Omni-directional image, panoramic, vignetting, illumination

I. 서 론

최근 방송 및 멀티미디어 서비스 기술은 아날로그 시대의 흑백/컬러 영상 시대로부터 디지털 시대의 고해상도 TV

와 대화형 멀티미디어 방송으로 발전해가고 있다. 이로 인해 기존의 방송 서비스에 비해 영상의 해상도가 높아지고 다채널의 고품질 음향이 제공되기 때문에 시청자가 느끼는 리얼리티는 더욱 높아지고 있다. 또한, 인터넷을 통한 데이터 전달 기술 및 속도의 증가로 다양한 멀티미디어 콘텐츠의 이용이 가능하게 됨에 따라 사용자들의 요구가 단순 소비 형태의 콘텐츠 이용에서 대화형 콘텐츠, 실감 콘텐츠의 이용을 요구하는 수준에 이르렀다. 이러한 환경 속에서 차

a) 광운대학교 영상처리연구소

Image Processing Systems Laboratory Dept. of Computer Engineering, Kwangwoon University

† 교신저자 : 심동규(dgsim@kw.ac.kr)

* 본 연구는 “서울시 산학연 협력사업”을 통하여 이루어졌음

세대 기술로 떠오르고 있는 3차원 영상 디스플레이 기술은 시각 정보를 보다 실감나고 사실적으로 표현하여 사용자들의 다양한 욕구를 충족시킬 수 있는 대안이다. 3차원 영상 디스플레이 기술은 자유로운 시점과 넓은 화면을 제공하는 다시점 비디오 (Multiview video) 기술로 표현될 수 있는데, 현재 다시점 비디오 기술은 전방향 비디오 (Omni-directional video), 자유 시점 비디오 (Free viewpoint video), 양방향 스테리오 비디오 등 크게 3가지의 응용 분야를 갖는 3DAV의 이름으로 연구가 진행 중이다^[1]. 이 중 전방향 비디오는 사용자에게 단 방향이 아닌 360도의 전 방향 또는 넓은 시야 화면을 제공하는 비디오 기술로, 사용자가 전방향 비디오 시스템 환경에 있는 것만으로 실제 환경에 있는 것과 같이 현장감과 몰입감을 느끼게 해준다는 특징을 가진다. 이 기술은 방송 및 3차원 멀티미디어 콘텐츠 산업뿐만 아니라 감시 카메라 시스템의 비디오 압축, 위성 영상 처리나 항공 영상 처리 및 시뮬레이션 등에서 유용하게 사용될 수 있고^{[2]-[4]}, 현재 방송에서 사용되고 있는 가상현실 (Virtual reality)에 적용하여 컴퓨터 그래픽이 아닌 실제 영상으로 사실감을 더해줄 수 있다.

전방향 영상을 생성하기 위한 방법으로는 넓은 시야각을 가지는 특수한 카메라를 이용하여 영상을 획득하는 방법과, 다중 카메라를 이용하여 영상을 획득한 후 3D 메쉬 (mesh) 객체에 비디오 데이터를 텍스처 (texture)로 맵핑하는 방법, 카메라의 내부 및 외부 변수를 이용하여 영상을 생성하는 방법 등이 있다^{[4]-[7]}. 하지만, 어안렌즈 (fish-eye), 볼록 반사경 (convex mirror) 등의 특수한 카메라로 촬영된 영상은 일반 카메라로 촬영한 영상에 비해 넓은 시야의 영상을 획득할 수 있지만 영상의 왜곡이 심해 고해상도의 영상 생성에는 적절치 않다는 한계를 지니고 있다. 텍스처 맵핑과 카메라 변수를 이용한 방법 역시 계산량과 복잡도의 증가로 비디오 코딩을 하기에 어렵다는 단점이 있다. 뿐만 아니라 위의 방법들은 카메라의 시스템에 의해 발생하는 현상과 영상들 사이의 조명 변화를 고려하지 않기 때문에 다수의 영상을 이용하여 전방향 영상을 생성할 때 부자연스러운 전방향 영상이 생성될 수 있다. 비그네이팅 효과와 조명 변화를 고려하고 이를 보정하는 파노라믹 영상 생성 방법^[8]은 원형 형태의 전방향 영상을

생성하여 모든 방향의 영상을 포함할 수 없다는 한계가 있다.

본 연구에서는 효과적인 전방향 영상 생성을 위한 방법을 제안한다. 제안한 방법은, 평면 영상으로 전방향 영상을 생성한 경우 등방 왜곡이 발생하고 원통형 영상으로 생성할 경우에는 전체 영상을 포함할 수 없다는 점을 고려하여, 등방 왜곡을 최소화하면서 전체 영상을 포함할 수 있도록 구의 한 부분을 나타내는 원형 영상을 이용하여 전방향 영상을 생성한다. 이를 위해 카메라로 획득한 평면 영상을 원형 좌표계를 이용하여 원형 영상으로 변환하여, 모든 방향의 영상을 포함하고 영상을 합성할 때 발생하는 왜곡을 최소화한다. 그리고 빛의 양과 카메라의 회전 등의 외부 환경 조건에 의해 발생하는 조명 변화와 영상의 주변으로 갈수록 영상이 어두워지는 비그네이팅 효과 (vignetting effect)를 보정하여 보다 사실적이고 자연스러운 전방향 영상을 생성한다.

본 논문의 구성은 다음과 같다. 2장에서는 제안한 전방향 영상 생성 방법에 대해 설명하고, 3장에서는 제안한 방법을 통해 생성된 전방향 영상의 결과를 비교 분석한다. 마지막으로 4장에서는 제안한 방법에 대한 결론을 기술한다.

II. 변수화된 영상 보정을 통한 전방향 영상 생성 방법

본 장에서는 제안한 변수화된 영상 보정을 통한 전방향 영상 생성 방법을 소개한다. 그림 1은 제안한 방법의 블록도를 나타낸다. 앞 장에서 논의된 바와 같이 모든 방향의 영상을 포함하고 합성시에 발생하는 왜곡과 에러를 최소화하는 전방향 영상을 생성하기 위해서는 평면 영상을 변환하는 과정이 필요하다. 이를 위해 회전 카메라로부터 획득된 원본 프레임인 N 번째 프레임과 $N+1$ 번째 프레임을 원형 좌표계를 기반으로 원형 영상으로 변환한다. 그리고 두 영상 사이에 공통으로 존재하는 영역 내에서 합성을 위한 가장 적합한 영역을 찾기 위해, N 번째 프레임에서 특징점을, $N+1$ 번째 프레임에서 특징점과 가장 일치하는 대응점을 찾

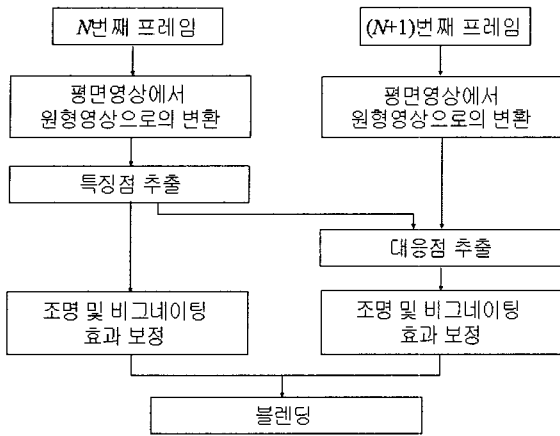


그림 1. 전방향 영상 생성을 위한 방법의 블록도
Fig. 1 Block diagram for omnidirectional image generation

는다. 또한 각각의 영상은 외부 환경 조건과 카메라 시스템으로 인해 발생하는 비그네이팅 효과와 영상들 사이에 조명 변화를 가지고 있으므로, 이러한 현상을 보정하기 위하여 가장 최적의 비그네이팅과 조명 변수를 구하여 영상이 가지고 있는 조명 및 비그네이팅 효과를 보정한다. 마지막으로 특징점과 대응점을 기준으로 N번째 프레임과 N+1번째 프레임을 합성하고, 정합 부분에서 발생하는 경계면 처리와 함께 영상합성 과정을 수행한다. 그림 2는 N장의 평면 영상을 이용하여 전방향 영상을 생성하는 과정을 나타낸다. N장의 평면 영상을 이용한 최종의 전방향 영상은 그림 1의 과정을 N-1번 수행함으로써 생성된다. 다음의 절에서는 그림 1의 각 과정에 대해 소개한다.

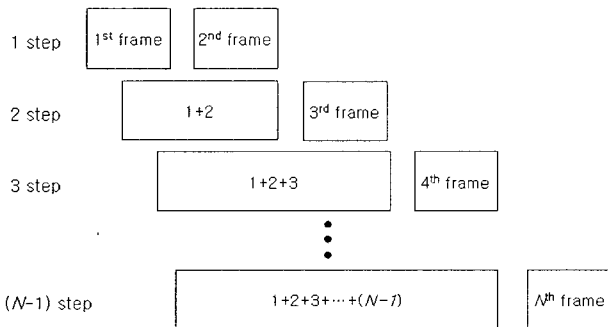


그림 2. N장의 영상을 이용한 전방향 영상 생성 과정
Fig. 2 Process of omnidirectional image generation using N frames

1. 평면 영상에서 원형 영상으로의 변환

전방향 영상은 회전 카메라에 의해 획득된 평면 영상들을 서로 합성함으로써 생성된다. 각각의 평면 영상들은 서로 독립적인 좌표를 가지고 있으므로 평면 영상 사이에 공통으로 존재하는 영역을 기준으로 합성을 하더라도 영상이 매끄럽게 연결되지 않고 서로 어긋나는 현상이 발생한다. 이들을 합성하면서 발생하는 왜곡과 에러를 줄이기 위해서는 공통의 좌표계가 요구된다. 본 연구에서는 공통의 좌표계로 원형 좌표계를 이용하여 2D 평면 영상을 3D 원형 영상으로 변환한 후, 전방향 영상을 생성한다. 그림 3은 평면 영상에서 원형 영상으로의 변환을 위해 평면 영상 내의 한 픽셀을 원형 영상 면으로 투영하는 과정을 나타낸 것이다. I_p 와 I_s 는 각각 평면 영상 면과 원형 영상 면을 나타내고, 픽셀 A와 픽셀 B는 임의의 좌표로써, 평면 영상면 내의 픽셀 A의 좌표 (P_x, P_y, P_z) 는 직각 좌표, 평면 영상면 내의 픽셀 B의 좌표 (S_x, S_y, S_z) 는 원형 좌표를 나타낸다. 이 때 S_θ 는 픽셀 B를 x-y 평면으로 투영했을 때의 점과 x축 사이의 각, S_ϕ 는 x-y 평면과 픽셀 B와의 각을 나타내고, 하나의 원형 영상이 가지는 S_θ 와 S_ϕ 의 최대값은 카메라의 화각과 같다. 본 연구에서 사용된 카메라의 화각은 가로축 $46.4^\circ (T_\theta)$, 세로축의 $35.64^\circ (T_\phi)$ 이고, 이 화각을 이용하여 카메라의 초점거리를 나타내는 변수 f 를 예측할 수 있다. 한 원형 영상이 가질 수 있는 S_θ 와 S_ϕ 의 범위와, f 를 계산하는 수식은 다음과 같다.

$$\begin{aligned}
 &-\frac{T_\theta}{2} < S_\theta < +\frac{T_\theta}{2} \\
 &-\frac{T_\phi}{2} < S_\phi < +\frac{T_\phi}{2} \\
 &f = \text{width} / 2 \times \tan \frac{T_\theta}{2}
 \end{aligned}
 \tag{1}$$

원형 영상은 그림 3과 같이 평면 영상의 중심을 직각 좌표계의 x축에 위치시키고, 원형 영상의 픽셀에 B에 대응되는 평면 영상의 픽셀 A를 픽셀 B에 투영함으로써 생성된다. 이 때 픽셀 A는 픽셀 B와 원점이 일치선이 되는 곳과 평면 영상이 만나는 곳에 위치한 점으로, 픽셀 B로 투영을

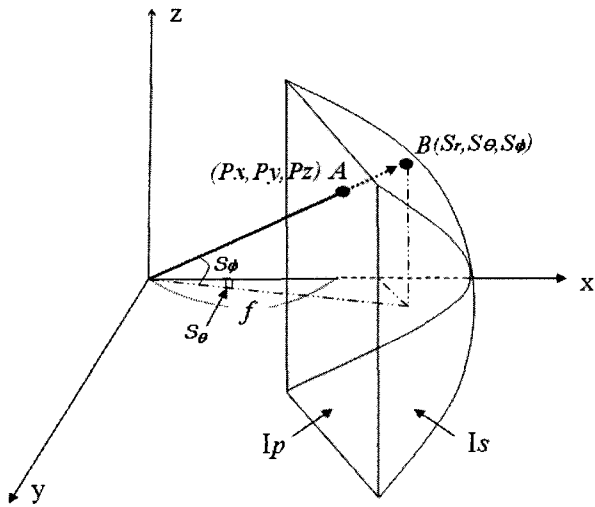


그림 3. 평면 영상 내의 한 픽셀을 원형 영상면으로 투영
Fig. 3 Projection of a pixel in the planar image to the spherical image

위해서는 픽셀 A의 좌표 값이 필요하다. 카메라의 화각과 f를 이용하여 원형 영상 내에 점 B의 원형 좌표 (Sr, Sθ, Sφ)가 주어질 때, 그림 4의 과정을 통해 픽셀 A의 좌표를 구하고 이를 원형 영상 면에 맵핑할 수 있다. 픽셀 A의 값은 식 (2)를 이용하여 픽셀 B의 원형 좌표 (Sr, Sθ, Sφ)로

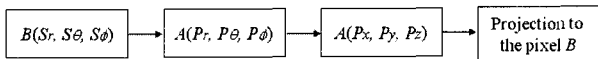


그림 4. 평면 영상 내의 값을 대응되는 원형 영상의 좌표에 맵핑하는 과정
Fig. 4 Process of mapping the pixel value in planar image to the corresponding coordinate in spherical image

부터 픽셀 A의 원형 좌표 (Pr, Pθ, Pφ)를 구하고, 식 (3)을 이용하여 픽셀 A의 원형 좌표를 직각 좌표로 변환함으로써 얻어진다.

$$\begin{aligned} P_r &= f / \sin S_\theta \cos S_\phi \\ P_\theta &= S_\theta \\ P_\phi &= S_\phi \end{aligned} \tag{2}$$

$$\begin{aligned} P_x &= f \\ P_y &= P_r \sin P_\theta \sin P_\phi \\ P_z &= P_r \cos P_\phi \end{aligned} \tag{3}$$

위의 과정을 통해 구한 픽셀 A의 좌표는 소수점의 위치 (fractional position)를 갖는다, 그러므로 선형 보간법을 통해 픽셀 A의 주변의 네 값들을 이용하여 픽셀 A의 값을 계산한 후, 이 값을 원형 영상의 픽셀 B로 투영한다. 위의 과정을 통해 평면 영상을 원형 영상으로 변환한 영상이 그림 5에 나타나 있다.

2. 특징점 및 대응점 추출

회전 카메라에 의해 취득된 연속적인 프레임 중 전방향 영상 생성을 위해 사용되는 프레임은 프레임 간에 20% 정도의 공통의 영역을 가지고 있는 프레임으로 선택된다. 이 공통의 영역을 기준으로 두 영상이 합성되는데 이를 위해

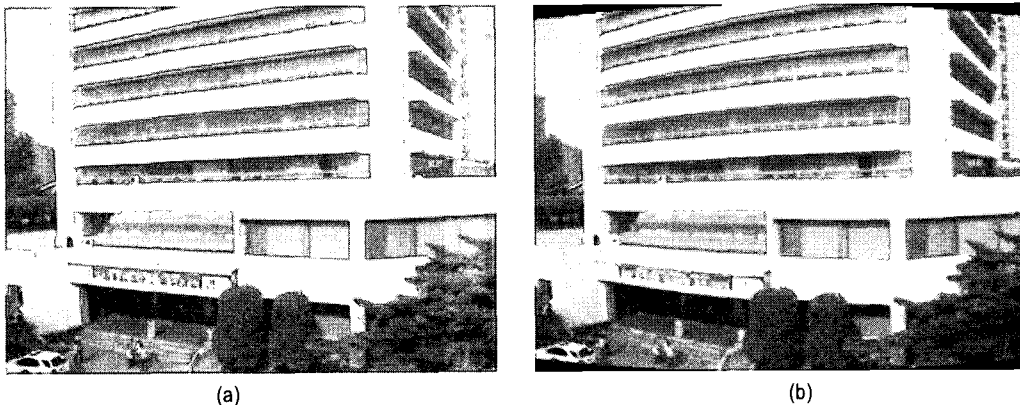


그림 5. 평면 영상을 원형 영상으로 변환 (a) 평면 영상 (b) 원형 영상
Fig. 5 Transform of a planar image to the spherical image (a) Acquired planar image (b) Transformed spherical image

서는 공통으로 존재하는 영역 내에서 합성을 위한 최적의 영역을 추출하는 과정이 필요하다. N 번째 프레임과 $N+1$ 번째 프레임을 이용하여 전방향 영상을 생성한다고 할 때, N 번째 프레임에서 합성을 위한 가장 최적의 특징 영역을 특징점 블록, $N+1$ 번째 프레임에서 이 특징점 블록과 대응되는 영역을 대응점 블록이라고 하자. N 번째 프레임과 $N+1$ 번째 프레임 사이에 가장 최적이 되는 특징점 블록과 대응점 블록 추출 과정은 그림 6에 나타나있다. 그림 6의 N 번째 프레임에서 회색 영역은 $N+1$ 번째 프레임과 공통된 영역으로, 특징점 블록 추출을 위한 탐색 범위이다. 최적의 특징점 블록 F 는 탐색 범위 내에서 32×32 블록 단위로 모든 블록의 분산을 계산한 후, 가장 큰 분산 값을 가지는 블록으로 선택된다. 특징점 블록과 가장 일치하는 대응점 블록 추출은 $N+1$ 번째 프레임의 회색 영역 내에서 수행된다. 대응점 블록 추출의 탐색 범위는 특징점 블록 추출과는 달리, 대응점 블록이 N 번째 프레임에서 추출된 특징점 블록과 비슷한 높이를 가질 것이라고 가정하여, 그림 6의 $N+1$ 번째 프레임의 회색 영역과 같이 설정된다. 특징점 블록과 가장 일치하는 대응점 블

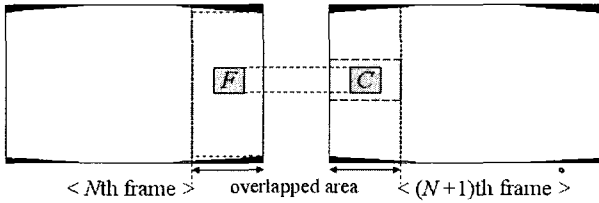


그림. 6 N 번째 프레임과 $N+1$ 번째 프레임 사이의 특징점 및 대응점 추출을 위한 탐색 범위

Fig. 6 Search range of feature and correspondence detection between the N th frame and the $(N+1)$ th frame

$$E = \sum_{x,y}^{\text{overlapped area}} \{I_{Nth}(x,y) - I_{(N+1)th}(x,y)\}^2$$

$$I_{Nth}(x,y) - I_{(N+1)th}(x,y) = I'_{Nth}(x,y) \left\{ V \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{x^2 + y^2}}{f} \right) \right) + 1 - V \right\}$$

$$- I'_{(N+1)th}(x,y) \left\{ \frac{V \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{x^2 + y^2}}{f} \right) \right) + 1 - V}{L_0} \right\} + \frac{L_1}{L_0}$$

(5)

록 (C)은, 탐색 범위인 $N+1$ 번째 프레임의 회색 영역 내에서 32×32 블록 단위로 특징점 블록과의 상호 상관 계수를 계산하여, 가장 큰 계수 값을 가지는 블록이 선택된다.

3. 조명 및 비그네이팅 효과 보정

카메라로부터 획득된 영상은 카메라 시스템이나 외부 조명 환경에 의한 비그네이팅 효과와 조명 변화를 가진다. 비그네이팅 효과는 렌즈 시스템에 의하여 영상의 주변에 입사 빛이 차단되는 현상으로, 영상의 주변으로 갈수록 영상이 어두워 보이는 현상을 말한다. 또한, 조명 변화는 영상이 획득될 때 주변 환경의 밝기 조건의 변화에 따라 같은 장소에서 짧은 시간 간격을 두고 촬영된 영상일지라도 영상들 사이에 서로 다른 밝기 값을 가지고 있는 현상을 말한다. 이러한 현상이 포함된 영상들을 이용하여 전방향 영상을 생성할 경우 같은 영역이라도 부분적으로 서로 다른 밝기 값을 가진 영상이 생성될 수 있으므로, 영상을 합성하기 전에 비그네이팅 효과와 조명 변화를 보정하는 과정을 통해 전체적으로 동일한 밝기 값을 가지는 전방향 영상을 생성하는 과정이 필요하다. 비그네이팅 효과와 조명 변화의 보정을 위해서는 다음의 수식을 이용한다^{[9][10]}.

$$I'(x,y) = \frac{L_0 I(x,y) + L_1}{V \cos^4 \left(\arctan \left(\frac{\sqrt{x^2 + y^2}}{f} \right) \right) + 1 - V} \quad (4)$$

$I(x,y)$ 는 조명 및 비그네이팅 효과를 가지고 있는 원형 영상이고, $I'(x,y)$ 는 위의 효과가 보정된 후의 영상으로 예측

$$I_{Nth}(x, y) = I'_{Nth}(x, y) \left\{ V \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{x^2 + y^2}}{f} \right) \right) + 1 - V \right\}$$

$$I_{(N+1)th}(x, y) = \frac{I'_{(N+1)th}(x, y) \left\{ V \cos^4 \left(\arctan \left(\frac{\Delta \sqrt{x^2 + y^2}}{f} \right) \right) + 1 - V \right\} - L_1}{L_0} \quad (6)$$

된 영상이다. L_0 과 L_1 는 조명 변화 변수, V 는 비그네이팅 변수, f 는 카메라의 초점거리를 나타낸다. 여기서 미지수인 조명 변수 L_0 , L_1 와 비그네이팅 변수 V 를 구하기 위해서, N 번째 프레임과 $N+1$ 번째 프레임 사이의 겹쳐진 영역에서 식 (5)의 비용 함수를 최소화하는 최적의 변수 L_0 , L_1 , V 를 구한다.

식 (5)를 통해 최적의 변수 L_0 , L_1 , V 를 구한 후, 이 변수를 식 (6)에 대입하여 조명 및 비그네이팅 효과가 보정된 영상인 I_{Nth} 와 $I_{(N+1)th}$ 를 생성한다.

4. 영상 합성

조명 및 비그네이팅 효과 보정을 통해 영상 사이의 전체적인 밝기를 보정한 후, 전방향 영상을 생성하기 위해 각 원형 영상을 합성하는 과정을 수행한다. 영상 합성은 두 영상 사이에서 추출된 특징점과 대응점 블록을 기준으로 수행되는데, 두 영상 사이의 픽셀 값은 약간의 차이가 있기 때문에 별도의 처리 없이 영상을 합성할 경우, 정합 부분에서 경계 면이 발생할 수 있다. 이를 위해 본 연구에서는 두 영상 사이의 픽셀 값이 가중치를 부여하여 영상을 합성하였다. 이는 N 번째 프레임과 $N+1$ 번째 프레임 중 각 프레임에 더 가깝게 위치한 픽셀에 가중치를 부여하여 전방향 영상의 픽셀 값을 생성하는 방법으로, 이를 통해 발생할 수 있는 에러가 최소화된 전방향 영상을 생성할 수 있다.

III. 실험 결과

1. 전방향 영상 생성

본 논문에서는 전방향 영상을 생성하기 위하여 4개의 시

표 1. 실험에 사용된 비디오 시퀀스

Table 1. Video sequences used in our evaluation

시퀀스	영상 크기	촬영 장소	프레임 수
시퀀스 1	320×240	실외	105
시퀀스 2	320×240	실내	117
시퀀스 3	320×240	실외	140
시퀀스 4	320×240	실외	121

퀀스를 취득하였다. 표 1은 실험을 위해 취득한 비디오 시퀀스들을 나타낸다. 본 실험에 사용된 시퀀스는 실내 및 실외에서 취득하였고, 모든 시퀀스는 팬틸트 (Pan-tilt) 드라이버와 캠코더 (GZ-MC200)를 사용하여 자동적으로 회전을 하며 취득되었다. 그림 7은 실험에 사용된 비디오 시퀀스 획득 시스템을 보여준다.



그림 7. 실험에 사용된 비디오 시퀀스 획득 시스템

Fig. 7 Video sequence acquisition system used in our experiment

제한한 방법을 통한 전방향 영상 생성은 각 시퀀스의 프레임 중 20%정도 겹치는 영역을 가지는 네 장의 프레임을 선택하여 생성한다. 그림 8의 (a)~(d)는 전방향 영상 생성

을 위해 시퀀스 1에서 선택된 프레임들을 나타낸다. 이 중 그림 8(d)는 영상이 취득될 때 받은 순간적인 빛의 변화로 인해 그림 8(a)~(c)와 다른 밝기를 가지고 있다. 그림 8(e)는 그림 8(a)~(d)의 4장의 원본 프레임을 이용하여 조명 및 비그네이팅 보정을 수행하지 않고 생성한 전방향 영상을 나타낸다. 조명 변화를 가지는 그림 8(d)로 인해 그림 8(e)의 표시된 부분에서 보듯이 오른쪽으로 갈수록 나머지 영역과는 다른 밝기 값을 가지는 부자연스러운 전방향 영상이 생성됨을 알 수 있다. 그러나 제안한 방법에 따라 조명 및 비그네이팅 효과 보정 후 생성된 전방향 영상인 그림 8(f)는 부분적으로 다른 밝기 값을 가지는 영역이 발생하지

않고, 영상 내의 모든 영역이 서로 비슷한 조명을 가지는 것을 확인할 수 있다.

2. 성능 평가

제안한 방법의 효율성을 평가하기 위해, 두 가지의 관점에서 성능을 비교하였다. 먼저 카메라로부터 획득된 원본 프레임인 평면 영상을 이용하여 영상을 합성한 결과와, 제안한 방법인 원형 좌표계를 기반으로 평면 영상을 원형 영상으로 변환한 후 영상을 합성한 결과를 비교하였다. 그림 9는 두 결과 영상을 나타낸 것으로 그림 9(a)는 평면 영상

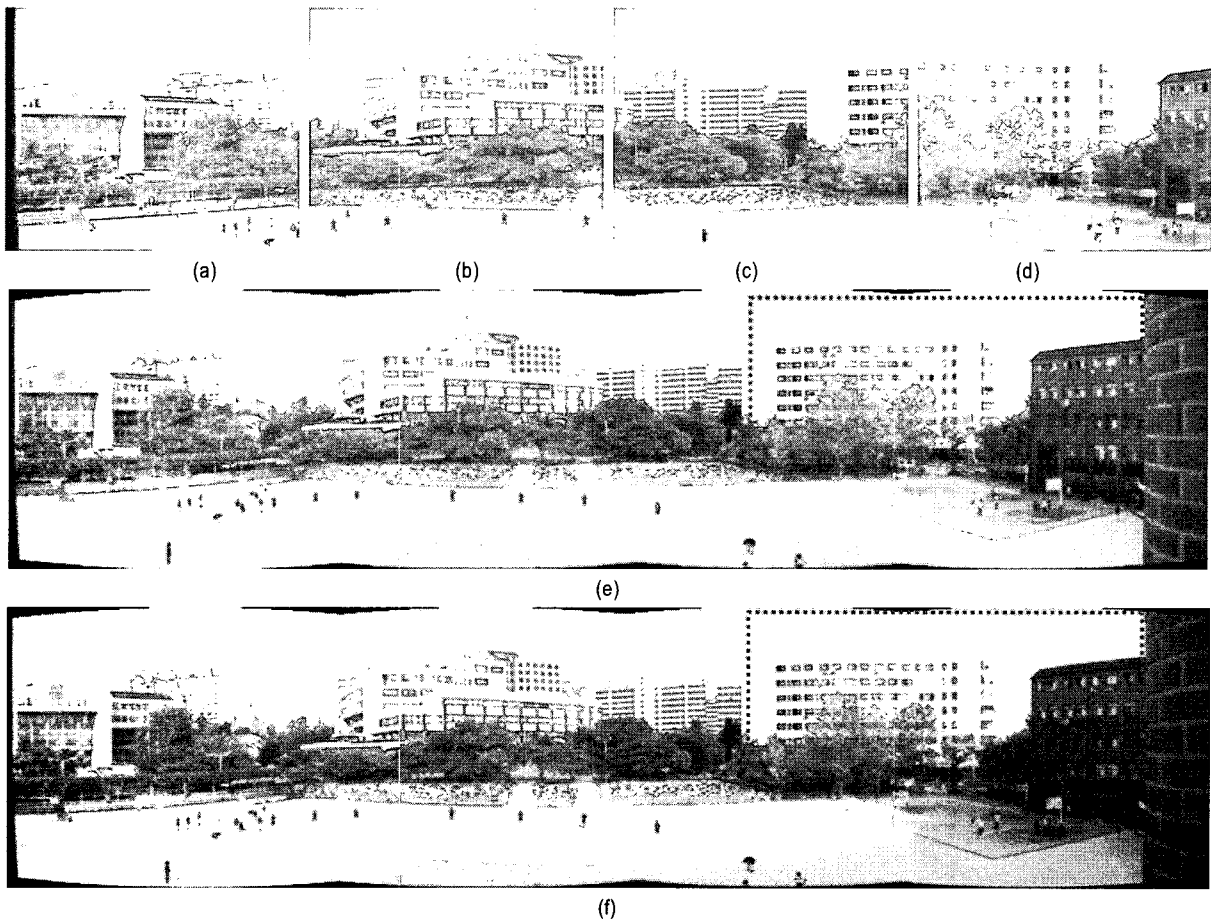


그림 8. 4장의 프레임을 이용한 전방향 영상 생성 (a)~(d) 시퀀스 1 (1, 35, 70, 105번째 프레임) (e) 조명 및 비그네이팅 보정 없이 생성된 전방향 영상 (f) 조명 및 비그네이팅 보정 후 생성된 전방향 영상

Fig. 8 Omni-directional image generation using 4 frames (a)~(d) Sequence 1 (1st, 35th, 70th, 105th frame) (e) Generated omni-directional image without vignetting and illumination compensation (f) Generated omni-directional image with vignetting and illumination compensation

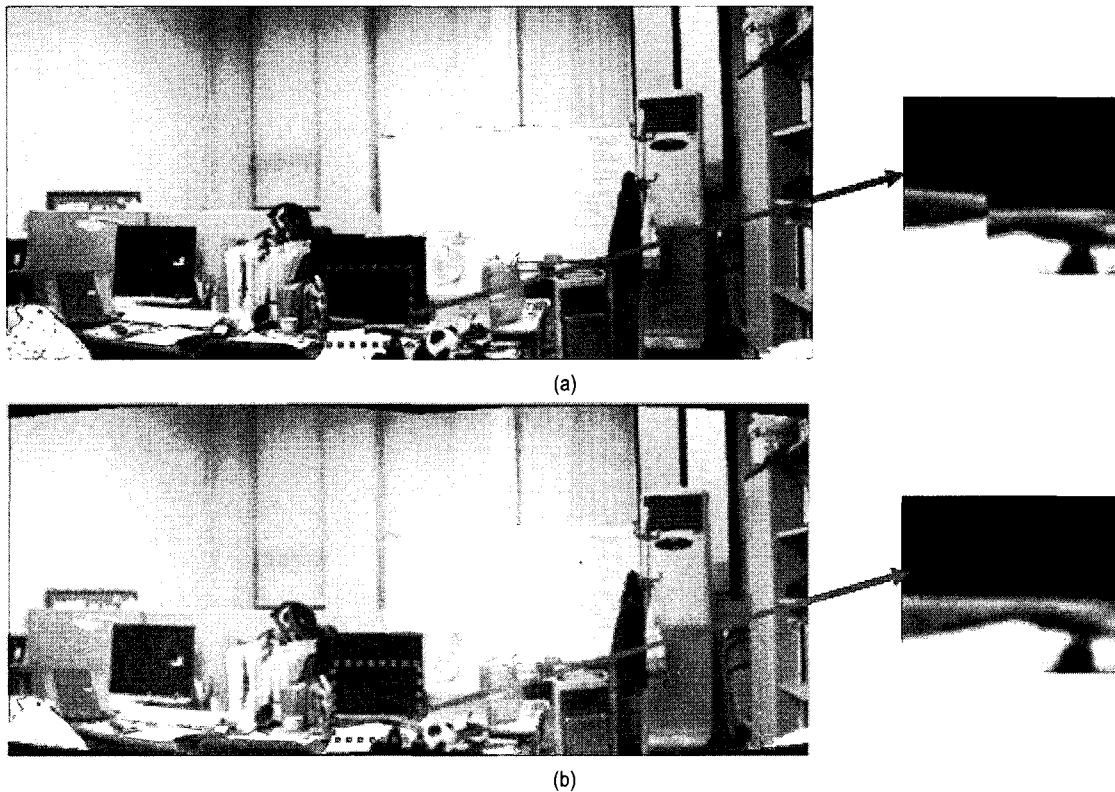


그림 9. 평면 영상을 이용한 합성과 원형 영상을 이용한 합성 비교 (a) 평면 영상을 이용한 영상 합성 (b) 원형 영상을 이용한 영상 합성
 Fig. 9 Comparison of blended images using planar mapping and spherical mapping (a) Blending using planar mapping (b) Blending using spherical mapping

을 이용하여 합성한 영상, 그림 9(b)는 원형 영상으로 변환한 후 합성한 영상을 나타낸다. 그림 9(a)와 그림 9(b)내에서 공통된 영역을 확대한 영상에서 보듯이, 원형 영상으로 변환하지 않고 영상을 합성한 경우 영상 사이에 등방 왜곡이 발생하는 반면, 원형 영상을 합성한 경우 합성 시에 발생하는 왜곡과 에러가 최소화됨을 확인할 수 있다.

두 번째로 조명 및 비그네이팅 효과 보정의 효율성을 보이기 위해, 제안한 방법과 조명 및 비그네이팅 효과를 보정하지 않은 방법의 성능을 평가하였다. 이를 위해 두 방법을 통해 생성된 전방향 영상으로부터, 원본 프레임에 대응하는 위치에서 평면 영상으로 역 변환을 수행한 후, 원본 프레임과의 PSNR을 계산하였다. 이 때 원본 프레임에 대응되는 전방향 영상 내의 원형 영상은 조명 및 비그네이팅 효과가 보정된 영상이므로, 식 (6)을 역으로 수행하여 조명 및 비그네이팅 효과를 가지는 영상을 생성하고 이를 역 변

환한다. 역 변환은 평면 영상을 원형 영상으로 변환하는 2.1절의 과정을 역으로 수행하는 과정으로, 식 (2)와 식 (3)을 역으로 계산함으로써 원형 영상으로부터 평면 영상으로 변환될 수 있다. 그림 10은 원본 프레임과 두 방법을 통해 생성된 전방향 영상으로부터 역 변환 된 영상 사이의 PSNR을 나타낸 것이다. 여기서 원본 프레임인 그림 10(a)는 시퀀스 1의 75번째 프레임을 나타내고, 그림 10(b)와 그림 10(c)는 시퀀스 1을 이용하여 두 방법을 통해 생성된 전방향 영상으로부터, 75번째 프레임에 대응하는 위치에서 역 변환 된 영상이다. 그림 10을 통해 조명 및 비그네이팅 효과를 보정한 방법의 PSNR이 더 높은 것을 확인할 수 있다.

그림 11은 각 실험 시퀀스의 전체 프레임에 대해 두 방법을 적용한 결과에 따른 PSNR을 그래프로 나타낸 것이다. 그림 11을 통해 4개의 시퀀스 모두 조명 및 비그네이팅



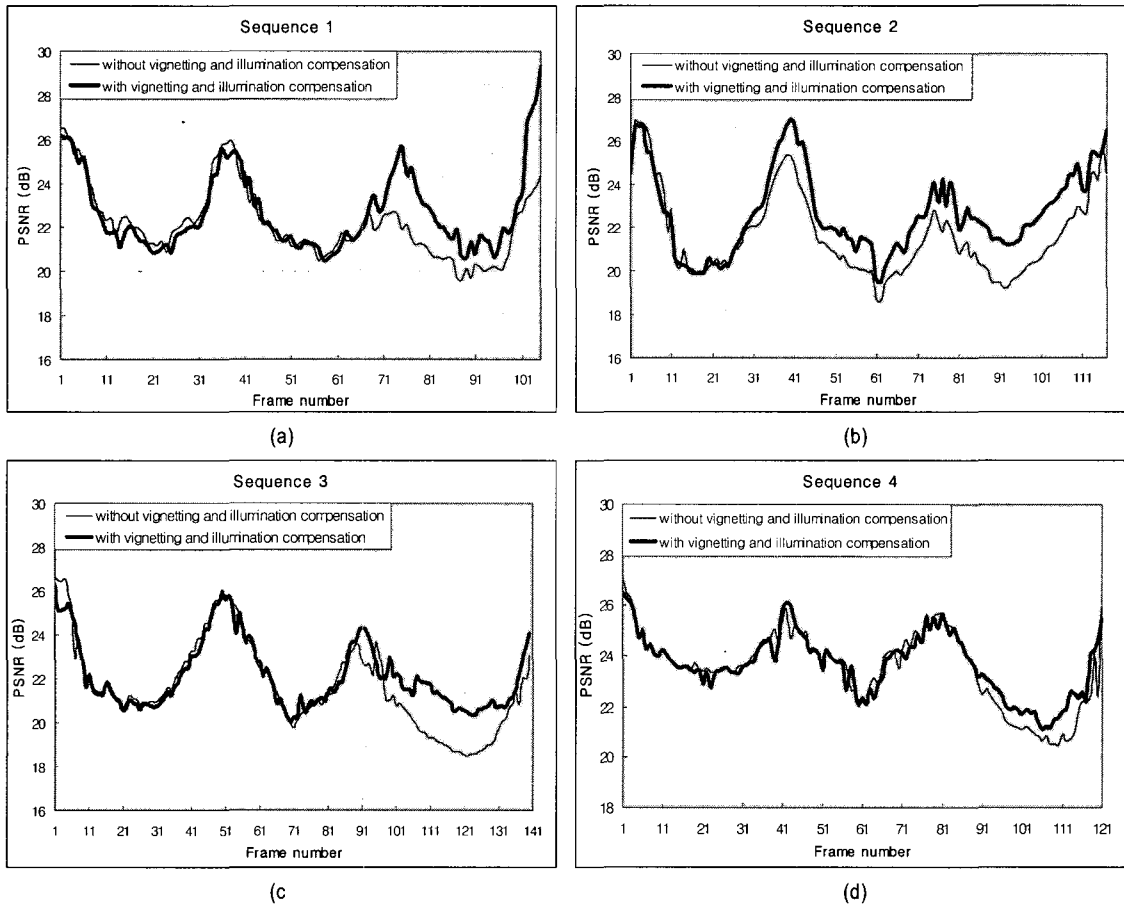
(a)

(b)

(c)

그림 10. 원본 영상과 역변환 된 영상 사이의 PSNR (a) 원본 영상 (시퀀스 1의 75번째 프레임) (b) 조명 및 비그네이팅 보정 후 역변환 된 영상과 원본 영상과의 PSNR (c) 조명 및 비그네이팅 보정 없이 역변환 된 영상과 원본 영상과의 PSNR

Fig. 10 PSNRs between original and inverse transformed images (a) Original image (75th frame in sequence 1) (b) Inverse transformed image with vignetting and illumination compensation (25.7dB) (c) Inverse transformed image without vignetting and illumination compensation (22.6dB)



(a)

(b)

(c)

(d)

그림 11. 4개의 시퀀스에 대해 조명 및 비그네이팅 효과 보정을 수행한 영상과 수행하지 않은 영상의 PSNR (a) 시퀀스 1 (b) 시퀀스 2 (c) 시퀀스 3 (d) 시퀀스 4
 Fig. 11 The PSNRs with and without the compensation of illumination and vignetting effects for 4 test sequences (a) Sequence 1 (b) Sequence 2 (c) Sequence 3 (d) Sequence 4

효과를 보정한 방법이 보정하지 않은 방법보다 1~4dB 더 높은 성능을 보이는 것을 확인할 수 있다. 이 때 실외 영상의 경우 효과를 보정한 방법이 부분적으로 성능을 보이는 반면, 그림 11(b)의 실내 영상은 효과를 보정한 방법이 전체 프레임에 걸쳐 성능을 보인다. 이는 실내와 실외의 조명 환경이 다르기 때문에 발생하는 현상이다. 실외에서 획득된 시퀀스의 경우 대부분의 빛이 일정하고 카메라의 회전에 따른 빛의 변화가 크지 않으므로 영상이 획득될 때 순간적으로 빛의 변화를 받은 영상에 대해서만 좋은 성능을 보인다. 하지만, 실내 영상의 경우 조명이 한정적으로 한 곳에 집중되어 있고, 이로 인해 카메라의 회전 방향에 따른 조명의 변화가 크기 때문에 대부분의 프레임이 비그네이팅 효과와 조명 변화를 가지고 있다. 그러므로 조명 및 비그네이팅 효과를 보정했을 때, 대부분의 영상에서 보정하지 않은 방법보다 높은 성능을 보인다. 또한 평면 영상으로 역 변환된 영상이 전방향 영상을 생성할 때 사용되었던 4장의 원본 프레임과 비슷한 위치에 있을수록 더 높은 PSNR을 가지는 것을 확인할 수 있다.

IV. 결론

본 논문에서는 비그네이팅 효과 및 조명 변화를 고려한 원형 좌표계 기반의 효과적인 전방향 영상 생성 방법을 제안하였다. 제안한 방법을 통한 전방향 영상 생성은 회전 카메라로 획득한 평면 영상을 원형 좌표계를 기반으로 원형 영상으로 변환한 후 영상을 합성함으로써, 모든 방향의 영상을 포함하고 합성 시에 발생할 수 있는 왜곡과 에러를 최소화하였다. 그리고 카메라 시스템과 외부 환경 조건에 의해 발생하는 비그네이팅 효과와 조명 변화를 보정하여 보다 정확한 전방향 영상을 생성하였다. 실험 결과에서는 제안한 방법에 따라 생성된 전방향 영상을 확인하였고, 원본 프레임인 평면 영상을 이용하여 합성한 영상과 원형 영상으로 변환한 후에 전방향 영상을 생성하는 것이 보다 효과적임을 보였다. 그리고 제안한 방법과 조명 및 비그네이팅 효과를 보정하지 않은 방법을 비교하여 제안한 방법이

1~4dB 더 높은 성능을 보이는 것을 확인하였다.

현재 제안한 방법에 따라 생성된 전방향 영상은 카메라를 회전하여 획득한 영상을 이용하여 생성한 파노라마 형태의 영상이므로, 원형 좌표계의 특성을 바탕으로 모든 방향의 영상을 획득하여 반구 또는 구 형태의 전방향 영상을 생성할 예정이다. 이를 실제 전방향 시스템 환경에 적용하여 재생한다면 사용자의 시청 실감을 한층 더 높일 수 있을 것으로 예상된다.

참고 문헌

- [1] A. Smolic and D. MaCutchen, "3DAV exploration of video-based rendering technology in MPEG," *IEEE Trans. on Circuit and Systems for Video Technology*, vol. 14, no. 3, pp. 348-356, March, 2004
- [2] I. Pavlidis, V. Morellas, P. Tsiamyrtzis, S. Harp, "Urban surveillance systems: From the laboratory to the commercial world," *Proceedings of IEEE*, vol. 89, no. 10, pp. 1478-1497, October, 2001
- [3] N. Atzpadin, P. Kauff, O. Schreer, "Stereo analysis by hybrid recursive matching for real-time immersive video conferencing," *IEEE Trans. on CSVT*, vol. 14, no. 3, pp. 321-334, March, 2004
- [4] W. Tang, T. Wong, P. Heng, "A system for real-time panorama generation and display in tele-immersive application," *IEEE Trans. on Multimedia*, vol. 7, no. 2, pp. 280-292, April, 2005
- [5] S. Akeda, T. Sato, N. Yokoya, "High-resolution panoramic movie generation from video stream acquired by an omni-directional multi-camera system," *IEEE Conf. on Multisensor Fusion and Integration for Intelligent System*, vol. 7, pp. 155-160, August, 2003
- [6] S. Nayar, A. Karmarkar, "360x360 mosaics," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Hilton Head Island, SC, USA, vol. 2, pp. 388-395, 2000
- [7] R. Y. Tsai, "An efficient and accurate camera calibration technique for 3D machine vision," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, pp. 364-374, 1986
- [8] D. G. Sim, "New panoramic image generation based on modeling of vignetting and illumination effects," *Pacific Rim Conference on Multimedia*, Jeju, Korea, part 1, pp. 1-12, 2005
- [9] Y. Altunbasak, R. Merswewau, A. Patti, "A fast parametric motion estimation algorithm with illumination and lens distortion correction," *IEEE Trans. Image Processing*, vol. 12, no. 4, pp. 395-408, April, 2003
- [10] N. Asada, A. Amano, M. Baba, "Photometric calibration of zoom lens systems," *Proceedings of the 13th International Conference on Pattern Recognition*, Vienna, Austria, vol. 1, pp. 186-190, 1996

 저 자 소 개



김 유 나

- 2006년 2월 : 광운대학교 컴퓨터공학부 컴퓨터공학 전공 (학사)
- 2006년 2월 ~ 현재 : 광운대학교 컴퓨터공학과 석사과정
- 주관심분야 : 3D Video; H.264 Video Coding, Multiview Video Coding



심 동 규

- 1993년 2월 : 서강대학교 전자공학과 (학사)
- 1995년 2월 : 서강대학교 전자공학과 (석사)
- 1999년 2월 : 서강대학교 전자공학과 (박사)
- 1999년 3월 ~ 2000년 8월 : (주)현대전자 선임연구원
- 2000년 9월 ~ 2002년 3월 : (주)바로비전 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 조교수
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전