

일반논문-09-14-2-07

깊이맵의 전처리와 깊이 정보의 기하학적 분석을 통한 최적의 스테레오스코픽 영상 자동 생성 기법

이재호^{a)}, 김창익^{b)†}

Stereoscopic Image Generation with Optimal Disparity using Depth Map Preprocessing and Depth Information Analysis

Jaeho Lee^{a)} and Changick Kim^{b)†}

요 약

DIBR(depth image-based rendering) 기법은 한 장의 칼라 영상과 그에 대응하는 깊이맵을 이용하여 수신측에서 가상의 좌우 영상을 생성하여 시청자에게 깊이감을 제공해주는 방법이다. 이 때 생성되는 좌우 영상의 화질과 깊이감은 좌우 영상의 시점에 해당하는 가상 카메라의 간격에 영향을 받게 된다. 본 논문에서는 가상 카메라의 간격 조절을 통해 양질의 깊이감을 제공하는 기법을 제안한다. 영상의 기하 분석에 따르면 스테레오스코픽 영상의 변이가 증가할수록 영상의 왜곡이 줄어들어 깊이감이 좋아지게 되나, 이 경우에는 그에 따른 영상의 열화가 고려되지 않았다. 실험결과에 따르면 영상열화나 시역형성의 제한 등으로 인해 변이의 증가폭에 제한이 있음을 알 수 있다. 변이 증가폭을 늘리기 위해서는 영상의 열화 현상을 줄여야 하므로 깊이맵의 전처리 과정을 추가하였다. 제안하는 방법은 시청자의 요구에 따라 카메라의 위치와 시점을 추가적으로 조절할 수 있게 하는 양방향 서비스를 추가로 제공하므로 3차원 TV 뿐만 아니라, DMB와 같은 이동형 방송 시스템에도 적용이 가능하다.

Abstract

The DIBR(depth image-based rendering) method gives the sense of depth to viewers by using one color image and corresponding depth image. At this time, the qualities of the generated left- and right-image depend on the baseline distance of the virtual cameras corresponding to the view of the generated left- and right-image. In this paper, we present a novel method for enhancing the sense of depth by adjusting baseline distance of virtual cameras. Geometric analysis shows that the sense of depth is better in accordance with the increasing disparity due to the reduction of the image distortion. However, the entailed image degradation is not considered. Experimental results show that there is maximum bound in the disparity increasement due to image degradation and the visual field. Since the image degradation is reduced for increasing that bound, we add a depth map preprocessing. Since the interactive service where the disparity and view position are controlled by viewers can also be provided, the proposed method can be applied to the mobile broadcasting system such as DMB as well as 3DTV system.

Keywords : 3DTV, Stereoscopic, Depth map, DIBR

a) 한국과학기술원 정보통신공학과

Dept. of Information and Communications Engineering, Korea Advanced Institute of Science and Technology (KAIST)KAIST

b) 한국과학기술원 전기 및 전자공학과

Dept. of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST)

† 교신저자 : 김창익 (cikim@ee.kaist.ac.kr)

※ 본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업(IITA-2009-(C1090-0902-0017))과 교육인적자원부 두뇌한국21 (BK21) 정보 기술사업단 지원의 연구결과로 수행되었습니다.

· 접수일(2008년7월7일), 수정일(1차:2008년9월11일, 2차:11월 3일), 게재확정일(2009년3월2일)

I. 서론

디지털 방송 기술이 비약적으로 발전함에 따라, 단순히 고해상도의 화면을 제공하는 것만으로는 시청자의 다양한 요구를 만족할 수 없는 상황에 이르렀다. 기존의 영상에 깊이 정보(depth information)을 더하여 시청자에게 현장에 있는 듯한 몰입감을 제공할 수 있는 3차원 비디오는 이러한 추세에 발맞추어 미래의 TV 시스템을 위한 좋은 해결책의 하나로 인식되고 있다.

3차원 TV는 시청자의 수에 따라 단일시청자(single viewer)를 대상으로 하는 경우와 다중시청자(multiple viewer)를 대상으로 하는 경우의 2가지로 분류할 수 있다^[1]. 일반적으로 TV라 하면 시청자의 수에 구애받지 않고 화면을 제공하는 것으로 여겨지나, 3차원에서는 입체감을 주기 위해 시청자의 양안에 각기 다른 영상을 제공하여야 하므로 그로 인한 시역의 한계 등으로 인해 동시에 여러 명에게 동일한 영상을 제공하기가 어렵다. 물론 특수 안경을 이용하여 시역의 제한이 없이 다수의 시청자에게 입체감을 줄 수 있으나, 안경의 휴대와 착용의 불편과 같은 문제로 인해 현재는 무안경식 방식이 더 선호되고 있다.

무안경식 3차원 TV는 깊이감을 이용하여 시청자들에게 보다 나은 시청환경을 제공할 수 있지만, 깊이감의 표현에서 나타나는 잦은 시역의 변화와 불일정한 깊이감 등으로 인해 어지러움을 유발할 수 있는 문제점이 있다. 뿐만 아니라 2차원에서는 MSE(Mean Squared Error)나 PSNR(Peak Signal-to-Noise Ratio)처럼 널리 사용되는 객관적인 품질 측정 방식이 있으나, 3차원의 깊이감은 주관적인 요소여서 객관적인 측정 표준을 정하기 어렵다는 문제가 있다. 따라서 보다 논리적으로 깊이감을 제어할 수 있는 방안에 대한 연구가 필요하다.

본 논문에서는 단일 사용자를 대상으로 한 DIBR 기법의 3차원 TV 서비스에서 좌우 가상 카메라의 간격을 이용하여 자동으로 최적의 깊이감을 제공하는 알고리즘을 제안한다. 한 장의 칼라영상과 그에 해당하는 깊이맵으로부터 스테레오스코픽 영상을 생성할 경우 일반적으로 좌우영상에 해당하는 가상 카메라의 간격은 사람의 양안 간격 정도로 설정하여지나, 이렇게 생성된 스테레오스코픽 영상의 깊이

감은 카메라 변수나 촬영 환경에 따라 다르게 나타나기 때문에 올바른 방식이 아니다. 스테레오스코픽 영상의 경우 촬영 환경 및 디스플레이 환경에서의 여러 변수들을 바탕으로 깊이감의 왜곡 현상을 예측할 수 있는데^[2], 이러한 예측 결과를 토대로 스테레오스코픽 영상을 만들 경우에 변수들을 조절하여 깊이감의 왜곡 현상을 줄일 수 있다. 이렇게 개선된 스테레오스코픽 영상을 통해 시청자는 더욱 몰입감이 높고 피로감이 적은 시청환경을 제공받게 된다. 그러나, 가상 카메라의 간격이 늘어남에 따라 기존 카메라에서는 물체에 가려져서 보이지 않았던 영역이 새로 생성된 좌우 영상에 나타나게 된다. 일반적으로 주변 영역의 정보를 바탕으로 이러한 영역을 메우게 되는데, 이는 화질 열화의 원인이 되어 시청자가 깊이감을 느끼는데 있어서 방해 요소가 된다. 따라서 이러한 가려짐 영역을 줄이고, 동시에 효과적으로 메우기 위한 알고리즘이 함께 요구된다.

본 논문의 구성은 다음과 같다. 2장에서 관련된 배경 지식에 대해 소개하며, 3장에서는 영상 열화를 줄이기 위한 깊이맵의 전처리 방법 및 자동으로 최적의 카메라 간격을 추출하기 위한 방법을 제시한다. 4장에서는 제안된 방법의 실험 결과와 타당성을 검증하기 위한 설문조사의 결과를 보여주며 5장에서 결론 및 추후과제에 관해서 논의한다.

II. 배경

일반적으로 스테레오스코픽 영상을 얻기 위해서는 일정 간격 떨어진 두 대의 카메라로부터 얻어진 영상이 필요하다. 이 때 영상을 얻는 카메라의 배열 방식에 따라 그림 1-(a),(b)에 나타난 바와 같이 수렴타입과 평행타입의 두 가지 방식으로 나눌 수 있다^[3]. 평행방식은 두 카메라의 광축(camera axis)이 평행하게 배열된 형태로 무한대의 깊이를 바라보고 있는 배열이다. 반면 수렴방식은 영상을 촬영하는 두 대의 카메라가 공간상의 동일한 수렴점(convergence point)에 초점을 가지고 있는 형태이다. 이는 마치 인간이 장면 상에서 특정 물체에 초점을 가지고 바라보는 것과 같은 이치이다. 영상 획득 과정에서 이러한 배열의 차이로 인해 3차원 영상을 재생할 때 차이가 나타나게 된다. 스테레오스코픽 디스플레이

이에서는 카메라와 디스플레이에 따른 여러 가지 변수에 의해 왜곡 현상이 나타나게 되는데, 대표적인 현상으로 광축을 따라 깊이감이 일정하지 않게 나타나는 **depth non-linearity**와 촬영 지점에서부터 동일한 거리의 면이 휘어져서 나타나는 **depth plane curvature**가 있다^[3,4]. **depth plane curvature** 현상은 수렴방식의 카메라 배열에서 카메라의 CCD 면이 기울어져 있으므로 영상의 가장자리가 중앙보다 더 멀게 느껴지는 현상이다. 이러한 현상은 영상 획득 과정에서 카메라 배열 방식에 따라 다르게 나타난다. 예를 들어, 평행방식에

서는 **depth plane curvature** 현상은 없으나 영상의 수렴점이 무한대에 존재하므로 영상 내의 모든 물체가 주변이를 가져서 튀어나와 보인다는 단점이 있다.

이러한 현상들은 그림 1-(c)에 나타난 것처럼 카메라의 CCD를 각각 좌우로 이동시켜 준 변형된 평행 타입 방식으로 해결할 수 있다^[5]. 변형된 평행 방식은 수렴점을 가지고 있으므로 영상 내의 물체들이 위치에 따라 주변이와 정변이를 함께 가지게 된다. 따라서 화면상에서 튀어나와 보이는 물체와 들어가 보이는 물체가 공존하게 되어 시청자에

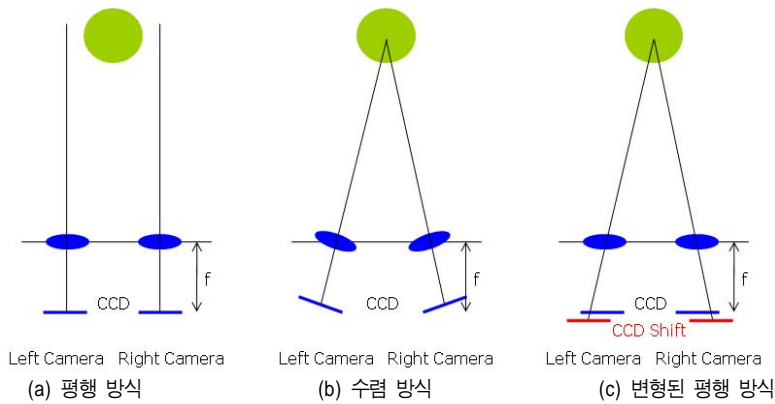


그림 1. 카메라의 배열 방식
Fig 1. Different camera array

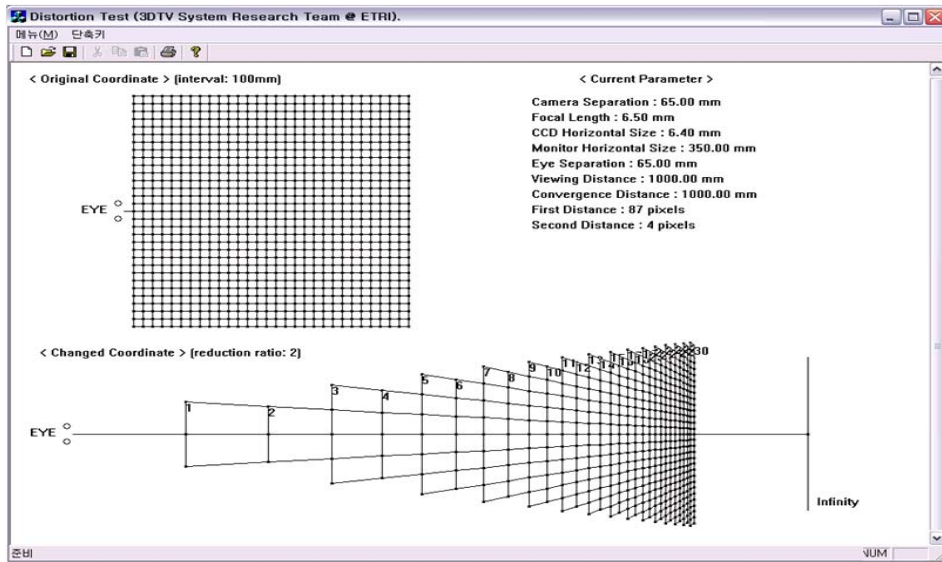


그림 2. 변형된 평행 방식에서의 영상 왜곡 현상
Fig. 2. Image distortion at modified parallel array

게 더욱 현실적인 깊이감을 제공해 줄 수 있다. 뿐만 아니라 언급한 바와 같이 CCD를 강제로 이동시켜서 CCD면이 기울지 않고 평행한 상태로 수렴점을 만들어 주므로, 그림 2처럼 변형된 공간에서 동일한 거리의 면이 휘지 않고 평평하게 되어 depth plane curvature 현상을 제거할 수 있는 장점이 있다. 그러나 변형된 평행 방식에서도 광축을 따라 깊이감이 일정하지 않게 되는 문제(depth non-linearity)는 여전히 존재한다. 본 논문에서는 DIBR 기법으로 생성된 좌우 영상을 변형된 평행방식으로 획득된 것으로 가정하고, depth non-linearity 문제를 추가적으로 해결할 수 있는 방법에 대해 논의하였다.

III. 제안하는 알고리즘

본 논문에서 제안하는 알고리즘의 전체 흐름은 그림 3에 나타난 바와 같다. 본 논문에서는 주로 점선으로 표시된 사각형 안의 내용에 대해 언급할 것이다. DIBR 기법으로 가상 시점의 좌우 영상을 생성할 때 발생하는 가려짐 영역을

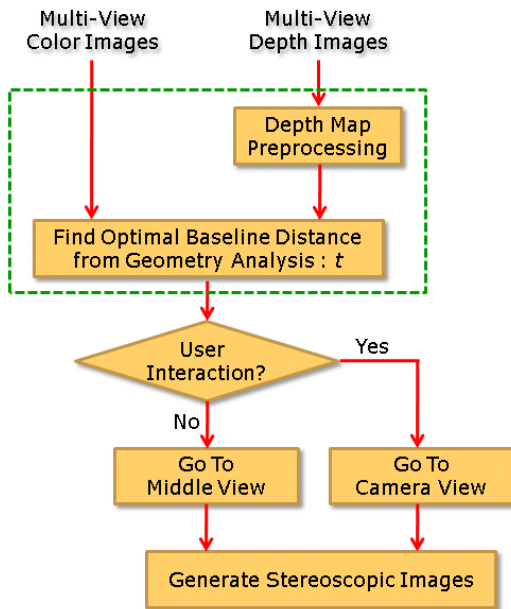


그림 3. 제안하는 방법의 전체 흐름도
Fig. 3. Overall flow of proposed method

해결하기 위해 깊이맵은 전처리 과정을 거친다. 그리고 촬영 환경 변수와 디스플레이 환경 변수들을 이용하여 영상을 기하학적으로 분석하여 깊이감을 극대화할 수 있는 최적의 카메라 간격(t)을 찾아낸다. 다시점 카메라 환경에서는 제공되는 카메라의 시점이 다수 존재하므로 시청자의 선택에 따라 카메라의 시점을 조절할 수 있다. 만약 시청자로부터 어떠한 정보도 받지 않는다면, 중앙에 위치한 카메라(middle view)의 영상을 이용하여 스테레오스코픽 영상을 생성하게 된다.

1. 깊이맵의 전처리

일반적으로 스테레오스코픽 영상을 생성할 때, 화소 단위의 깊이맵이 정교할수록 영상의 품질은 좋아진다. 그러나 Z-Cam 등을 이용해 얻어지는 깊이맵의 경우, 그 품질이 좋지 않아서 그대로 이용할 경우 DIBR 결과로 좋은 영상을 기대하기 어렵다. 따라서 사용하고자 하는 목적에 맞게 깊이맵을 적절히 전처리 해주는 과정이 필요하다. 본 논문에서 제안하는 전처리 과정은 그림 4와 같다.



그림 4. 깊이맵의 전처리 과정
Fig. 4. Preprocessing of depth map

1.1 중간값 처리 (Median Filtering)

DIBR을 이용하여 스테레오스코픽 영상을 생성할 때 주로 문제가 되는 부분은 물체의 경계 부분이다. 물체에 가려 보이지 않던 배경 영역이 가상 시점에서 보이는 현상이 발생하기 때문이다. Z-Cam이나 스테레오 정합 등을 이용해 얻어진 깊이맵의 경우 그림 5-(a)에서 보는 바와 같이 경계 부분에서 깊이값은 대체로 매끄럽지 않게 나타난다. 이를 위해 경계 부분에서의 적절한 마스크 처리 과정이 필요하며, 깊이맵의 경우 중간값 처리 과정이 가장 효과적으로 이러한 역할을 수행할 수 있다^[6]. 그림 5-(b),(c)는 마스크의 크기에 따른 깊이맵의 중간값 처리 결과를 보여준다.

그러나 중간값 처리 과정의 경우, 마스크의 크기가 커질

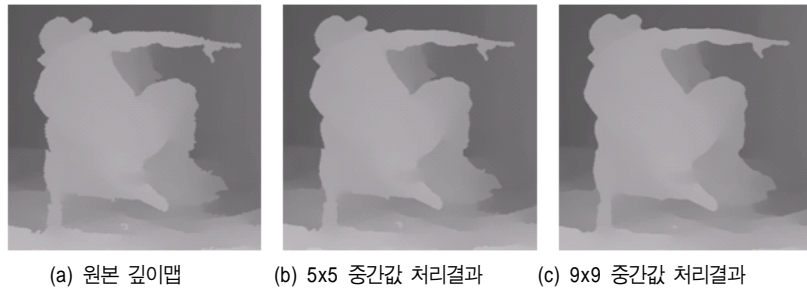


그림 5. 마스크의 크기에 따른 깊이맵의 중간값 처리결과
Fig. 5. Median filtered results of depth map with mask size

수록 처리 속도가 증가한다는 단점이 있기 때문에 영상에서 MSE 등을 이용하여 마스크의 크기를 적응적으로 결정하는 방법^[7,8]이 연구되어 왔다. 그러나 여기서는 깊이맵의 경계부분에서의 불연속성만 보완하면 되므로 영상 전체에 걸쳐 에너지 최적화 상태를 계산하는 과정은 낭비에 가깝다. 제안하는 알고리즘에서는 캐니 에지 검출기(canny edge detector)를 통해 깊이맵의 경계 부분만을 추출한 뒤, 해당 영역에만 중간값 처리 과정(5x5)을 적용하여 경계가 아닌 부분에서의 계산 과정을 생략하였다. 해당 과정은 식 (1)과 같다.

$$I(x,y) = \begin{cases} median(I(x,y)) & \text{if } c(x,y)=1 \\ I(x,y) & \text{if } c(x,y)=0 \end{cases} \quad (1)$$

$I(x,y)$ 는 (x,y) 에서의 깊이값을, $c(x,y)$ 는 캐니 에지 검출기의 결과값을 의미한다. $c(x,y)=1$ 은 해당 화소가 경계에 해당함을 의미한다.

1.2 경계 부분 방향성 검출

중간값 처리 결과로 바뀌어진 깊이맵의 경계를 소벨 마스크(sobel mask)를 통하여 구한다. 이렇게 처리된 결과 영상은 경계화소들간의 분산값을 최소화시키도록 분류하는 방법인 Ostu method^[9,10]를 통해 4단계의 강도(O_0, O_1, O_2, O_3)로 구분된다. 이전 단계에서와 같이 캐니 에지 검출기를 사용하지 않는 이유는, 경계로 구해진 화소의 강도를 측정하기 위해서이다. 이를 통해 1.3절에서 가우시안(gaussian) 마스크 처리를 할 때, 경계 화소에 따라 적응적으로 가우시

안 마스크의 크기를 조절할 수 있다.

스테레오스코픽 영상을 만들 때 물체에 의해 가려진 영역으로 인해 문제가 되는 것은 주로 수직 방향의 경계 부분이다. 따라서 식 (2)와 같이 소벨 마스크의 연산 과정에서 구해지는 x, y 방향으로의 경사도(gradient) 값을 이용하여 각 경계 화소에서의 기울기를 구한다. 이 기울기 값 $ang(x,y)$ 를 이용하여 이후의 전처리 과정에서 수평 방향의 경계 부분은 연산 과정에서 제외시킨다.

$$ang(x,y) = \arctan\left(\frac{Grad_y(x,y)}{Grad_x(x,y)}\right) \quad (2)$$

$Grad_x(x,y)$ 는 (x,y) 에서 x 축 방향으로의 경사도 값이고, $Grad_y(x,y)$ 는 (x,y) 에서 y 축 방향으로의 경사도 값이다. 그림 6은 강도에 따른 경계 부분 검출 결과와 방향성 검출 후 세로 방향 성분의 경계 부분만 남은 경계맵의 결과이다.



(a) 경계화소 강도에 따른 분류 (b) 방향성 검출 후의 경계화소 분류
그림 6. 경계화소 전처리 후의 결과
Fig. 6. Preprocessing results of boundary pixels

1.3 비대칭형 가우시안 마스크 처리

8비트로 표현된 깊이값의 불연속성 때문에 스테레오스코픽 영상에서는 정보가 없는 화소가 많이 발생하게 된다. 특히 물체의 경계 부분에서는 깊이값의 차이가 크게 나타나므로 문제가 더욱 심각해진다. 따라서 경계 부분에서의 불연속성을 완화하기 위하여 평활화 처리(smoothing filtering)를 해주어야 한다. 일반적으로 사람의 양안이 가로로 배열되어 있다는 특성 때문에 스테레오스코픽 영상 생성 과정에서의 변이는 가로 방향으로 발생한다. 따라서 세로 방향의 성분이 강하도록 비대칭형 가우시안 평활화 처리를 해주면 문제를 크게 개선할 수 있다^[11,12].

본 논문에서는 1.2절에서 구해진 결과를 바탕으로, 경계에서의 깊이값과 강도에 따라 가우시안 평활화 처리 정도를 적응적으로 조절하면서 가우시안 평활화 처리를 하도록 하였다. 따라서 평활화 처리 결과로 나타난 깊이값 $\hat{I}(x,y)$ 는 다음과 같이 나타난다.

$$g_{\sigma_x, \sigma_y}(\hat{I}(x,y)) = \frac{\sum_{\nu=-\omega/2}^{\omega/2} \left\{ \sum_{\mu=-\omega/2}^{\omega/2} (\hat{I}(x-\mu, y-\nu)g(\mu, \sigma_\mu))g(\nu, \sigma_\nu) \right\}}{\sum_{\nu=-\omega/2}^{\omega/2} \left\{ \sum_{\mu=-\omega/2}^{\omega/2} g(\mu, \sigma_\mu)g(\nu, \sigma_\nu) \right\}} \quad (3)$$

$$\alpha(x,y) = \begin{cases} \frac{D(x,y)}{D_{\max}} & \text{if } D(x,y) < D_{\max} \\ 1 & \text{if else} \end{cases} \quad (4)$$

$$\hat{I}(x,y) = \alpha(x,y) \times I(x,y) + (1 - \alpha(x,y)) \times g_{\sigma_x, \sigma_y}(\hat{I}(x,y)) \quad (5)$$

w 는 가우시안 마스크 크기, σ 는 평활화 강도, $g(\mu, \sigma_\mu)$, $g(\nu, \sigma_\nu)$ 는 각각 x , y 방향으로의 1차 가우시안 평활화를 의미

하며, $D(x,y)$ 는 현재 화소에서 가장 가까운 경계 화소까지의 거리를 의미한다. D_{\max} 값은 1.2절에서 구해진 경계 화소에서의 강도(O_0, O_1, O_2, O_3)에 따라 다르게 적용하였다. 그리고 생성되는 영상은 각각 좌우방향 중 한쪽방향으로만 이동하므로 물체의 좌우 경계 중에 한쪽 방향으로만 평활화 과정을 적용하면 된다. 예를 들어 좌측 시점 영상의 경우, 물체의 좌측 경계만 문제가 되므로 좌측 경계에 해당하는 부분만 위의 과정을 적용한다. 이와 같이 적용된 좌우 시점을 위한 깊이맵의 결과는 그림 7과 같다.

2. 스테레오스코픽 영상의 기하학적 분석과 조절

스테레오스코픽 영상을 통해 시청자는 깊이감을 느낄 수 있다. 이 때, 카메라의 촬영 환경 변수에 따라 시청자가 느끼는 깊이감에는 차이가 나타나게 된다^[13,14]. 앞으로의 3차원 디지털 방송에서는 영상 데이터 자체뿐만 아니라, 그에 따른 다양한 부가 정보들이 포함되어 전송되므로^[15], 이러한 촬영 환경 변수 또한 유용하게 이용될 수 있다. 촬영 환경 변수에는 카메라의 간격, 수렴거리, 초점거리, CCD의 가로길이, 카메라 시야각 등이 있다. 이중 변형된 평행 방식에서 영향을 미치는 요인은 그림 2에서 나타난 바와 같이, 카메라 간격, 초점거리, CCD 가로길이 등이다. 그러나 초점거리나 CCD 가로길이는 촬영 단계에서 이미 결정되는 사항이므로 DIBR 단계에서 조절할 수 없다. 따라서 본 장에서는 촬영 환경 변수 중 카메라 간격의 조절을 통해 자동으로 최적의 깊이감을 제공하는 스테레오스코픽 영상을 생성하는 방안을 제시하였다.

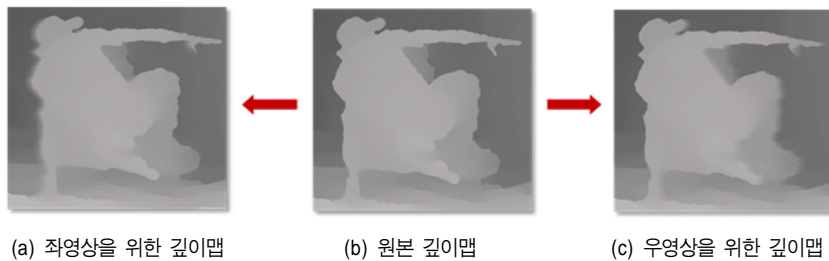


그림 7. 깊이맵의 전처리 결과
Fig. 7. Preprocessing results of depth map

2.1 깊이감의 기하학적 변환

스테레오스코픽 영상을 통해 나타나는 깊이감을 정확하게 측정하기란 대단히 어렵다. 그러나 촬영 환경 변수 및 시청 변수 등을 통해 생성된 스테레오스코픽 영상의 깊이감을 기하학적으로 변환하여 측정해봄으로써 이를 예측할 수 있다. 본 논문에서는 Andrew^[3] 등이 제안한 방법을 기초로 깊이감을 측정하였다. 앞서 소개된 그림 2는 일반적인 관찰조건 아래에서 Visual C++을 이용한 변형된 평행 방식의 시뮬레이션 결과를 보여주고 있다.

원래의 물체공간은 촬영 환경 변수들을 통해 스테레오스코픽 영상 생성시 깊이감이 왜곡되어 나타나게 된다. 변형된 물체공간에서 볼 수 있듯이, 일반적인 촬영 환경 아래에서는 카메라에서 가까운 위치일수록 깊이감이 잘 나타나며, 멀어질수록 깊이감이 줄어들게 된다. 또한 무한대 평면이 존재하게 되어서 일정 거리 이상의 깊이감을 표현하는 것이 불가능하다. 이는 실외 영상의 촬영과 같이 촬영 거리의 제한이 없을 때 문제를 야기시킬 수 있다.

이러한 깊이감의 왜곡 현상은 촬영 환경 변수를 이용하여 조절할 수 있다. 그림 8은 촬영 환경 변수 중 카메라 간격의 변화에 따른 왜곡 현상의 변화를 보여주고 있다. 일반적으로 카메라의 간격과 초점 거리, 그리고 CCD의 크기가 증가할수록 깊이감의 왜곡이 줄어드는 것으로 나타나며

^[3,13], 시청자에게 인식되는 깊이감 또한 왜곡이 적은 경우에 더욱 좋아지는 것으로 조사 되었다^[14]. 이러한 현상을 바탕으로 스테레오스코픽 영상에서 왜곡이 적도록 제어하여 양질의 깊이감을 제공할 수 있다.

2.2 자동적인 깊이감의 조절

디스플레이의 크기나 시청거리와 같은 시청 변수의 경우에는 시청자의 환경에 따라 변하는 조건이므로 스테레오스코픽 생성 단계에서 제어할 수 없다. 따라서 용도에 따라 시청자의 환경을 예상하여 그것의 크기와 시청거리를 포함 시켜야 한다.

그러나 촬영 환경 변수의 경우에는, 카메라 간격의 조절을 통해 깊이감을 제어할 수 있다. 초점거리나 CCD의 크기는 촬영 당시 사용된 카메라에 의해 결정되지만 카메라의 거리는 DIBR 단계에서 결정되므로 이를 조절하여 시청자에게 전달되는 깊이감을 조절하는 것이다. 이의 영향은 앞선 그림 8에 나타나있다. 이상적으로 영상의 왜곡을 없애기 위해서는 변형된 물체 공간에서의 깊이감이 실제 공간에서의 깊이감과 정확히 일치해야 한다. 이를 나타내기 위하여 깊이비례도(r)를 정의하였다. 깊이비례도는 그림 9과 식 (6)에 나타난 것처럼, 가상 카메라의 위치에서 가장 가까운 공간에서의 깊이감과 먼 공간에서의 깊이감의 비례도이다. 실제 공간에서

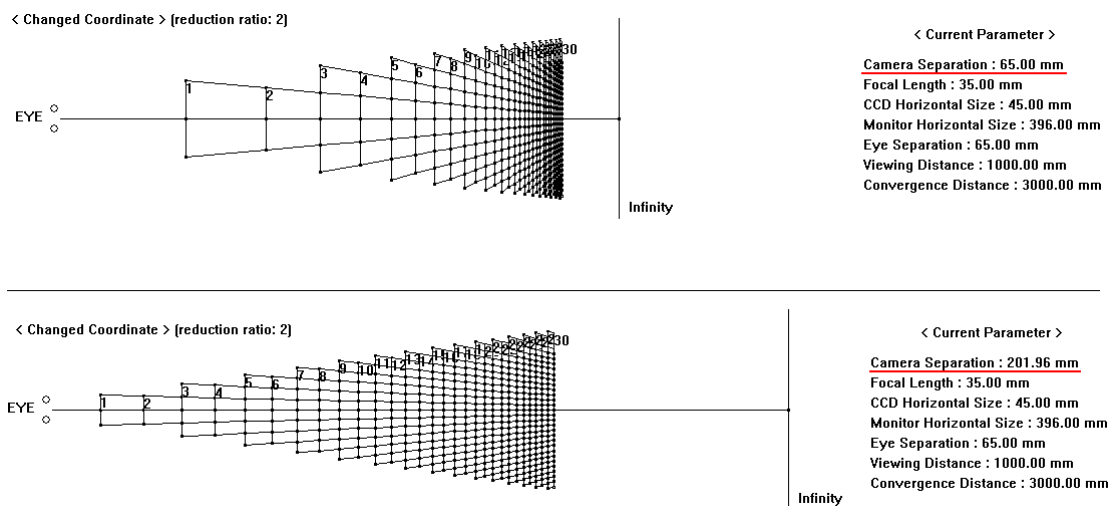


그림 8. 촬영 환경 변수에 따른 영상 왜곡 현상의 변화
 Fig. 8. Variation of image distortion with shooting parameters

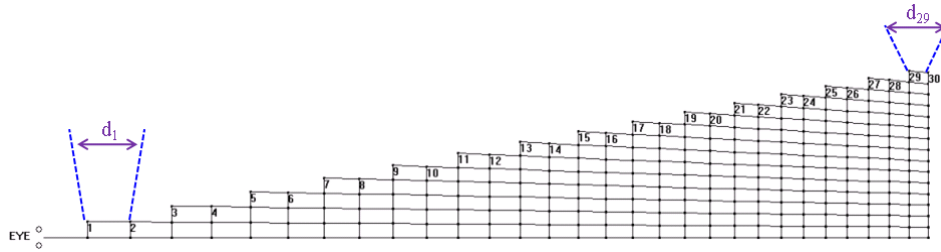


그림 9. 깊이 비례도 표현 방식
Fig. 9. Expression of depth proportion

는 \$d_1\$과 \$d_{29}\$의 길이가 동일하나 왜곡에 의해 변하게 된다. 일반적으로 왜곡에 의해 \$r\$은 1보다 작은 값을 가진다.

$$r = d_{29}/d_1 \tag{6}$$

앞서 설명한 바와 같이 일반적으로 카메라의 간격이 넓을수록 영상의 왜곡은 줄어든다. Andrew^[3]에 따르면 변형된 평행 방식에서의 왜곡 현상을 정의하는 구조적 변환식은 다음과 같다.

$$Z_i = \frac{Ve}{e + Mf \frac{t}{Z_0}} \tag{7}$$

여기서 \$e\$는 양안거리, \$V\$는 시청거리, \$M\$은 CCD와 디스플레이 장치의 크기비, \$f\$는 초점거리, \$Z_0\$는 실제 공간에서의 위치, \$t\$는 카메라 간격이다. 따라서 깊이비례도 \$r\$은 실제 공간에서의 위치 \$Z_0\$에 따라 식 (8)과 같이 정의된다.

$$r = \frac{Z_{i30} - Z_{i29}}{Z_{i2} - Z_{i1}} \tag{8}$$

식 (7)의 분모에서 2번째 요소가 1번째 요소인 양안거리 \$e\$에 비해 충분히 커서 양안거리가 무시가능하다면 식 (8)에서 깊이비례도 \$r\$이 1이 됨을 알 수 있다. 앞서 설명한 바와 같이, CCD나 디스플레이의 크기, 초점거리 등은 사전에 결정되는 값이므로 카메라 간격 \$t\$를 늘려주어서 깊이비례도를 개선시킬 수 있다. 하지만 카메라의 간격이 충분히 확보

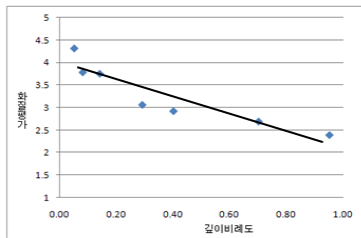
되면, 늘어나는 카메라의 간격에 비례하여 개선되는 영상 왜곡의 효과는 미비해진다. 또한 카메라 간격이 커짐에 따라 생성되는 스테레오스코픽 영상의 변이가 식 (9)과 같이 커지게 된다.

$$x_l = x_c + \left(\frac{t f}{2 Z}\right), \quad x_r = x_c - \left(\frac{t f}{2 Z}\right) \tag{9}$$

기준 영상에서의 \$x_c\$에 위치한 화소는 각각 좌우 시점 영상에서 가상 카메라 간격 \$t\$에 비례하여 위치하게 된다. \$f\$는 카메라의 초점 거리, \$Z\$는 \$x_c\$에서의 깊이값이다. 따라서 이 가상 카메라의 간격이 커질수록 그림 8의 아래 그림에서처럼 깊이감은 좋아지게 되나, 기준 카메라에서는 보이지 않았던 가려짐 영역이 많이 나타나게 되어 화질은 나빠지는 현상이 발생한다.

깊이비례도가 높아질수록 화질이 떨어지는 trade-off 현상을 해결하기 위해 시청자들이 느끼는 화질 열화 정도를 주관적인 평가를 통해 알아보고자 하였다. 설문을 통해 시청자들이 화질 열화 정도를 둔감하게 인식하는 범위 내의 최대 깊이비례도 \$r\$을 추정하도록 한다. 설문 방식은 DSCQS 방식^[16]으로 진행하였고 실험 영상으로는 break-dancing과 ballet^[17]을 이용하였다. 기준 카메라의 칼라 영상과 깊이맵을 이용하여 좌우 시점 영상을 생성하였으며, 기준 카메라에서 보이지 않았던 영역은 3장에 소개된 방식으로 메웠다. 19명을 대상으로 평가하였으며, 평가방식은 그림 10-(a)와 같이 진행되었다. 설문 결과는 그림 10-(b)와 같으며, 결과의 상관도는 -0.9163이 나왔다. 따라서 깊이비례도가 커질수록 화질열화 현상은 심해진다고 판단할 수

Excellent	5
Good	4
Fair	3
Poor	2
bad	1



(a) 평가방식

(b) 화질열화에 대한 평가 결과

그림 10. 화질열화에 대한 설문조사 방식과 결과

Fig. 10. Question method of image degradation and result

있다. 결과에서 볼 수 있듯이, 시청자들은 대체로 약 0.5의 비례도까지는 화질열화에 대해 둔감한 편이므로 비례도 r 은 0.5로 설정하였다.

3. 스테레오스코픽 영상 생성

화질 열화의 감소와 깊이감의 증가를 고려한 카메라 간격이 결정되면 이를 바탕으로 최종적으로 스테레오스코픽 영상을 생성하게 된다. 앞서 설명한 바와 같이 다시점 카메라에서 기준이 되는 카메라가 시청자에 의해 선택되면, 해당 카메라를 기준으로 앞서 설정된 변이만큼 떨어진 좌우 영상을 DIBR 방식으로 생성한다.

DIBR은 현재 시점의 장면을 표현하는 칼라 영상과 그에 해당하는 깊이맵을 전송하고, 수신단에서 이 정보들을 이용하여 양안 시점에 해당하는 스테레오스코픽 영상을 재생해내는 방식이다^[11,13]. 깊이맵은 주로 8비트로 표현되고 영상 내에서 이러한 깊이값의 변화가 심하지 않으므로 고주파 성분이 적다는 특징이 있다. 따라서 압축 효율이 일반 칼라 영상을 보내는 것보다 좋으므로 대역폭을 줄일 수 있는 장점이 있다^[13]. 또한 DIBR 기술을 이용하여 임의 시점에서의 장면을 생성할 수 있다는 장점도 있으므로 최근에는 두 개의 칼라 영상을 보내는 것보다는 DIBR을 이용한 스테레오스코픽 재생방법이 많이 이용되고 있다^[11].

그림 11은 DIBR 방식을 이용한 스테레오스코픽 영상 생성의 개요를 나타내고 있다. 기준 시점에서의 칼라 영상과 그에 해당하는 깊이맵을 입력으로 하여 임의 시점에서의 두 영상을 그림과 같이 만들어 낼 수 있다. 이 두 영상은 양안에 각기 다른 입력으로 들어가게 된다. 이렇게 만들어진 두 영상은 그림 3에서의 아래 그림과 그림 11에서 볼 수 있듯이 물체 뒤의 가려짐 영역이 정보 부족으로 인하여 나타나지 않는 문제점이 있다(생성된 좌우 영상에서 검은 색으로 나타난 부분). DIBR에서는 이렇게 가려진 부분을 효과적으로 채우기 위한 알고리즘이 요구된다.

변이가 작으면 영상 생성이 발생하는 가려진 영역이 작기 때문에 기존의 정보를 이용하여 해당 부분을 메워도 화

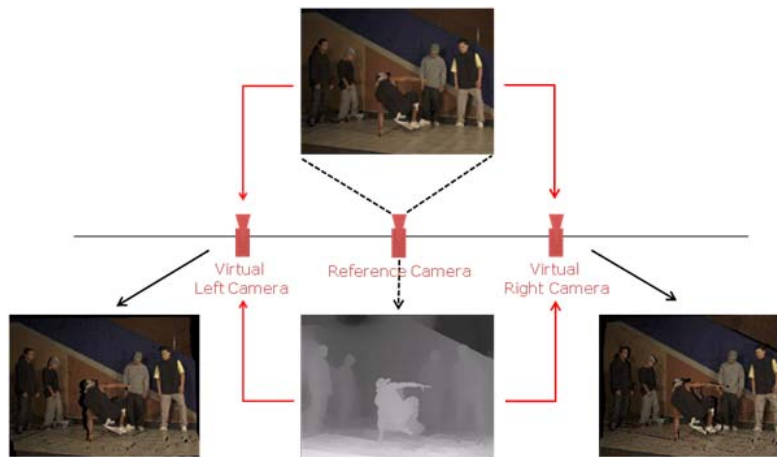


그림 11. DIBR의 개요와 occlusion 문제

Fig 11. Outline of DIBR and occlusion problem

질 열화 현상이 두드러지게 나타나지는 않는다. 반면 변이가 큰 경우에는 화질 열화 현상이 심하게 나타날 수 있으나, 앞서 카메라 간격을 설정할 때 이러한 부분을 고려하였으므로 최대 변이는 깊이비례도가 0.5를 넘지 않는 선에서 조절되게 된다.

2장에서 DIBR을 위한 깊이맵의 전처리 과정이 설명되었다. 해당 과정을 통해 스테레오스코픽 영상 생성시 발생하는 가려진 영역을 줄일 수 있다. 남아있는 부분을 채우기 위해서는 주변 화소들의 색상을 참조하여야 한다. 이러한 부분은 대체로 물체의 경계 부분에서 발생하기 때문에 주변 화소들의 색상에는 다소 큰 차이가 나타나게 된다. 따라서 메워질 화소를 중심으로 상하좌우와 대각선의 4개의 방향을 비교한 후, 가장 작은 색상 차이를 가지는 2개의 방향을 이용하여 메우도록 한다. 그림 12에서 DIBR의 결과를 확인할 수 있다.

IV. 실험 결과

1. 깊이맵 전처리에 따른 화질 평가

깊이맵의 전처리 과정에 따른 화질 개선 결과를 평가하기 위해 Zhang의 실험 결과^[11]와 비교하였다. Zhang의 방법에서는 비대칭 가우시안 마스크를 통하여 깊이맵을 부드럽게 만들어주고, 이를 이용하여 스테레오스코픽 영상을 생성한다. 본 논문에서는 물체 경계부분에서의 부드러운 윤곽선 도출을 위해 중간값 처리를 선행하였고, 경계 부분에만 비대칭 가우시안 마스크를 사용하여 깊이맵의 왜곡을 줄였다. 또한 물체의 경계 부분도 좌우영상 생성시에는 한 쪽 방향으로만 문제가 발생하므로 좌우 경계를 판단하여 해당 부분만 마스크 처리하도록 하였다.

실험 영상은 그림 13과 같이 MPEG FTV Group에서 제공되는 테스트 영상들을 사용하였다. 깊이맵은 Nagoya 대



그림 12. DIBR을 이용한 임의 시점 영상 생성
Fig. 12. Arbitrary view generation using DIBR

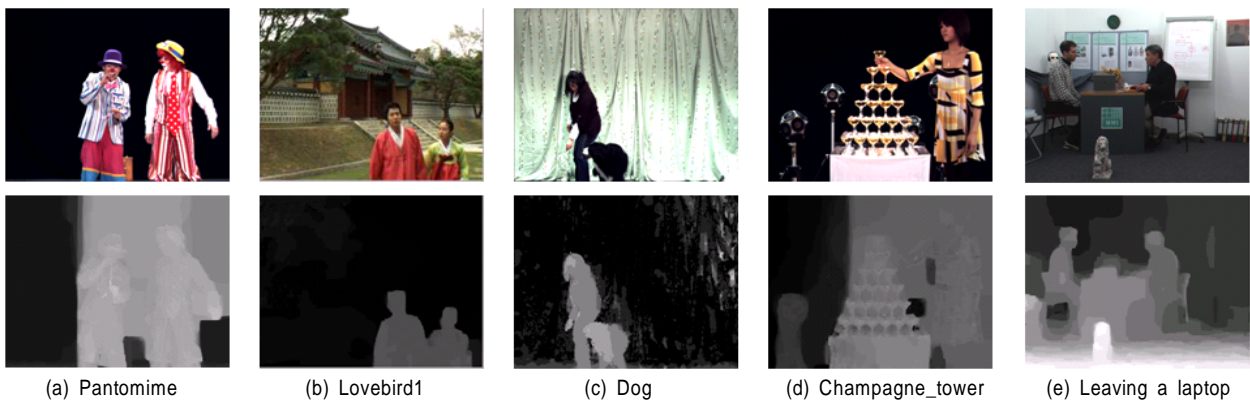


그림 13. 실험 영상들
Fig. 13. Test images

학에서 제공하는 BP 기반의 스테레오 매칭 기법을 이용하여 생성되었다. 전처리 과정을 거친 깊이맵을 이용하여 기준 카메라에서 가장 가까운 곳에 위치한 좌우 카메라에 해당하는 영상을 DIBR 방식으로 생성하고, 이를 본래 카메라에서 촬영된 영상과 비교하여 PSNR을 계산하였다. 실험결과는 표 1과 같다.

표 1. 깊이맵 전처리의 화질 테스트 결과

Table 1. The results of the image quality of depth map preprocessing

(dB)	Pantomime	Lovebird1	Dog	Champagne tower	Leaving a laptop
Zhang ^[11]	28.17	27.31	27.88	25.42	26.68
제안한 방법	29.44	27.96	28.69	26.67	28.07

제안한 방법으로 깊이맵을 전처리하여 영상을 생성할 경우, PSNR이 평균적으로 약 1dB 정도 상승함을 확인할 수 있다. 특히 “Champagne_tower”나 “Leaving a laptop”처럼 영상에 고주파 성분이 많은 경우에 더욱 좋은 개선효과가 나타났다. 이러한 화질의 향상으로 인해 시청자들은 스테레오스코픽 영상을 보는데 있어서 더욱 깊은 몰입감을 느끼게 된다.

전처리 과정이 깊이도 측정에서 미치는 영향을 알아보기 위해 위의 결과를 이용하여 평가해 보았다. 총 14명을 대상으로 평가를 진행하였으며, 위 실험에서 사용한 5종류의 영상에 대해 Zhang의 전처리 방법을 이용하여 만든 영상과 제안한 방법을 이용하여 만든 영상을 보여주면서 깊이비례도의 변화에 따라 인식되는 깊이감에 대해 평가하도록 진행하였다. 실험결과는 그림 14와 같으며 결과는 5종류 영상에 대한 평균 수치이다.

Zhang의 방법으로 전처리 과정을 거친 후 생성한 스테레오스코픽 영상에서는 깊이비례도 0.3~0.4의 구간에서 만족도가 좋게 나타나지만, 제안한 방법으로 전처리 과정을 거친 후 생성한 스테레오스코픽 영상에서는 깊이비례도 0.4~0.5의 구간에서 만족도가 좋게 나타난다. 따라서 전처리 과정을 통해 화질을 개선할 경우 이에 따라 적용할 수 있는 깊이비례도의 값이 커짐을 알 수 있다. 만약 DIBR에

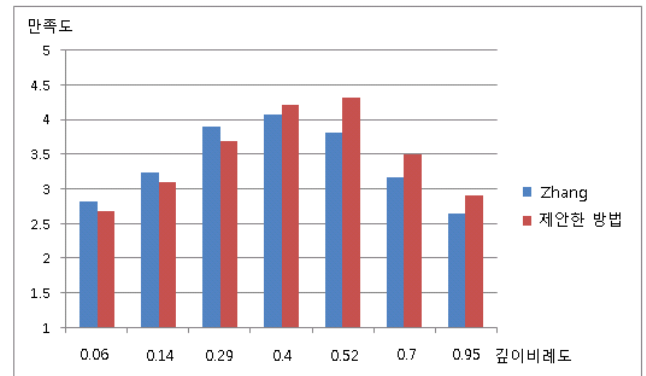


그림 14. 전처리 과정에 따른 깊이비례도 차이

Fig. 14. Depth ratio difference between previous and proposed method

서 생성된 가상 시점의 좌우 영상에서 나타나는 가려짐 영역을 완벽하게 복원할 수 있다면, 깊이비례도의 값은 1에 수렴할 것이다.

2. 깊이비례도에 따른 평가

깊이비례도에 따른 스테레오스코픽 영상 생성 결과를 평가하기 위해 3차원 모니터(MIRACUBE C190S)를 이용하였다. 총 14명을 대상으로 하였으며, 피실험자들은 각각의 영상에 대해서 인식되는 깊이감에 대해서 평가하였다. 카메라 간격이 증가하여 깊이비례도가 증가할수록 깊이감은 좋아지나, 지나치게 증가하면 시역형성이 어려워져서 어지러움을 느끼거나 깊이감을 느끼기 어렵게 된다.

그림 15는 실험에 사용된 영상을 보여주고, 그림 16은 각 영상에 대한 조사 결과이다. 실험 결과, 시청자들은 대체로 깊이비례도가 증가함에 따라 깊이감을 더욱 잘 느끼게 되나 0.5-0.6 이상의 깊이비례도에서는 앞선 그림 10의 결과와 같이 화질열화 현상이 발생해 깊이감을 저해하기 시작하였다. 그리고 0.8 이상의 깊이비례도에서는 시역형성이 되지 않아서 깊이감을 거의 느낄 수 없었다.

위의 조사를 통해, 앞선 실험에서 설정한 0.5의 깊이비례도에서는 대체로 시청자들이 만족스러운 깊이감을 느낄 수 있다. 따라서 스테레오스코픽 영상의 초기 생성시에 적합한 카메라의 간격은 깊이비례도 0.5 정도의 값이 적당



(a) Breakdancing 영상 (b) Ballet 영상

그림 15. 실험에 사용된 interlaced 영상
Fig. 15. Experimental interlaced images

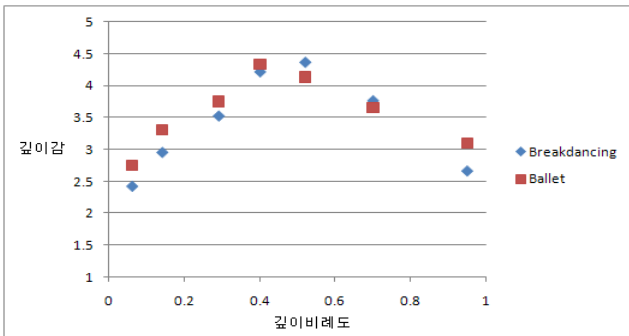


그림 16. 깊이감 측정 결과
Fig. 16. Experimental results of sense of depth

하다.

결정된 깊이비례도에 의한 스테레오스코픽 영상 생성의 성능을 확인하기 위해, 4.1절에서 사용된 다시점 영상 콘텐츠들을 이용하여 비교 실험해 보았다. 식 (7)과 (8)을

이용하여 깊이비례도가 0.5가 될 때까지 카메라 간격 t 를 늘려가면서 0.5가 되는 순간 멈추면 원하는 카메라 간격이 된다. 계산된 카메라 간격과 식 (9)를 이용해 만들어진 좌우 영상으로 스테레오스코픽 영상을 생성하고, 기준 카메라에서 가장 가까운 곳에 위치한 좌우 카메라의 원영상으로 스테레오스코픽 영상을 생성하여 두 영상에 대한 깊이감을 주관적으로 평가하였다. 그림 17-(a)는 실험에 사용된 영상들 중 “Leaving a laptop” 영상에서 6번 카메라를 기준으로 5번과 7번 카메라에서 촬영된 영상으로 스테레오스코픽 영상을 만든 것이며, 그림 17-(b)는 6번 카메라의 칼라 영상과 깊이맵을 이용하여 깊이비례도가 0.5가 되도록 스테레오스코픽 영상을 생성한 결과이다. 해당 영상을 3차원 모니터를 통해 전시한 후, 피실험자들에게 깊이감을 평가하게 하였다. 실험에 사용된 영상은 4.1절에서 사용된 실험 영상과 동일하며, 14명을 대상으로 실험을 진행하였다.

실험 결과 그림 18에 나타난 것처럼 깊이비례도에 따라 생성된 스테레오스코픽 영상의 경우가 다시점 카메라에서 주어진 영상을 그대로 활용하는 것보다 깊이감이 더 잘 느껴지는 것으로 나타났다. “Dog”, “Pantomime”, “Champagne_tower” 영상처럼 배경이 단순하여 상대적인 깊이감을 느끼기 어려운 환경에서는 느껴지는 깊이감의 차이가 작지만, 배경이 다소 복잡하게 존재하는 경우에는 상대적인 깊이감을 느낄 물체가 많아지므로 깊이감의 차이가 많이 나는 것



(a) 원영상을 이용한 스테레오스코픽 영상



(b) 깊이비례도에 따른 스테레오스코픽 영상

그림 17. 스테레오스코픽 영상 생성
Fig. 17. Stereoscopic image generation

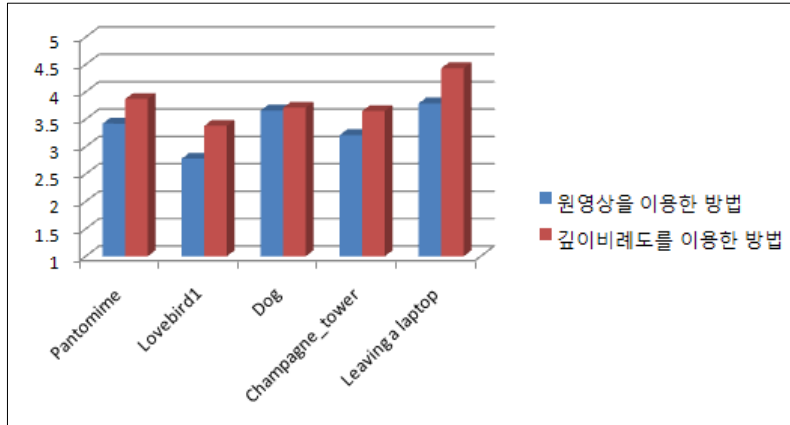


그림 18. 제안된 방법을 이용한 깊이감 평가
 Fig. 18. Test of the depth of sense by using proposed method

을 확인할 수 있다. 실험에 사용된 영상이 제한된 조건에서의 촬영 영상이며, 카메라의 대수가 많지 않아서 카메라 간의 기본 간격이 크다는 한계점은 있지만, 제안된 방식으로 인해 스테레오스코픽 영상을 생성하는 경우 더욱 효과적인 깊이감을 준다는 사실을 알 수 있다.

V. 결론 및 향후 과제

본 논문에서는 DIBR 방식에 의해 스테레오스코픽 영상이 생성될 경우, 기준 카메라에서 물체에 의해 가려져서 보이지 않았던 부분을 효과적으로 채우기 위한 깊이맵의 전처리 과정을 제안하였다. 또한 카메라 변수와 촬영 변수를 이용하여 스테레오스코픽 영상의 기하 분석을 통하여 시청자에게 최적의 깊이감을 주기 위한 카메라의 간격을 도출하는 방법도 함께 제안하였다. 실험 결과를 통해 제안된 방법으로 약 1dB 정도 화질이 개선된 영상을 얻을 수 있으며, 개선된 영상을 바탕으로 카메라 간격과 영상 왜곡 사이의 관계를 이용하여 최적의 카메라 간격을 도출하고자 하였다. 카메라 간격에 비례하여 영상의 왜곡 현상은 줄어들어 상대적인 깊이감이 더 잘 느껴지지만 왜곡을 줄이기 위해 무작정 깊이비례도를 높게 설정하면, 깊이맵의 전처리 과정에도 불구하고 스테레오스코픽 영상 생성 과정에서 영상의 열화 현상이 심해지므로 적절한 깊이비례도값 0.5를 실험을 통해 도출하였다.

제안된 방법은 영상의 종류에 무관하게 카메라 변수와 촬영 변수만을 이용하여 영상의 기하학적 분석을 수행하고 있다. 따라서 앞으로 영상의 콘텐츠에 따라서 인식되는 깊이감의 차이나, 기하구조 외에도 깊이감을 측정하기 위한 다른 방법에 대한 연구가 필요하다. 제안된 방법은 시청자의 요구에 따라 카메라의 위치와 시점을 추가적으로 조절할 수 있게 하는 양방향 서비스를 추가로 제공하므로 3차원 TV 뿐만 아니라 DMB나 mobile phone과 같은 개인형 방송에도 이용될 수 있다.

참 고 문 헌

- [1] C. Fehn, P. Kauff, M. Op de Beek, F. Ernst, W. A. IJsselstein, M. Pollefeys, L. Van Gool, E. Ofek, and I. Sexton, "An Evolutionary and Optimised Approach on 3D-TV," in Proc. of International Broadcast Conference '02, pp. 357-365, Sept. 2002.
- [2] Larry F. Hodges, "Basic Principles of Stereographic Software Development," in Proc. of SPIE, Stereoscopic Displays and Applications II, vol. 1457, pp. 9-17, 1991.
- [3] A. Woods, T. Docherty, and R. Koch, "Image Distortions in Stereoscopic Video Systems," in Proc. of SPIE, Stereoscopic Displays and Applications IV, vol. 1915, pp. 36-48, 1993.
- [4] H. Yamanoue, "The Relation between Size Distortion and Shooting Conditions for Stereoscopic Images," SMPTE journal, vol. 106, no. 4, pp. 225-232, 1997.
- [5] C. Smith, "3-D or not 3-D?," New Scientist, vol. 102#1407, pp. 40-44, 1984.
- [6] Y. Ho, S. Na, K. Oh, and C. Lee, "Depth Coding and Virtual View

Synthesis for FTV”, ISO/IEC JTC1/SC29/WG11 M14996, October 2007.

[7] H. Huang and T. Lee, “Data Adaptive Median Filters for Signal and Image Denoising Using a Generalized SURE Criterion,” IEEE Signal Processing Letters, vol. 13, no. 9, pp. 561-564, 2006.

[8] S. Tsekeridou, C. Kotropoulos, and I. Pitas, “Morphological Signal Adaptive Filter for Noise Removal,” IEEE Conf. Electronics, Circuits, and Systems, vol. 1, pp. 191-194, 1996.

[9] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” IEEE Trans. on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, 1979.

[10] P. S. Liao, T. S. Chen, and P. C. Chung, “A Fast Algorithm for Multilevel Thresholding,” Journal of Information Science and Engineering 17, pp. 713-727, 2001.

[11] L. Zhang and W. J. Tam, “Stereoscopic Image Generation Based on Depth Images for 3D TV,” IEEE Trans. on Broadcasting, vol. 51, no. 2, pp. 191-199, June 2005.

[12] G.M. Um, F. Speranza, L. Zhang, W. J. Tam, R. Renaud, L. B. Stelmach and C. H. Ahn, “Investigation of the Effect of Disparity-Based Asymmetrical Filtering on Stereoscopic Video,” in Proc. of Visual Communications and Image Processing, vol. 5150, no. 3, pp. 110-118, July 2003.

[13] C. Fehn, “Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV,” in Proc. of SPIE, Stereoscopic Displays and Virtual Reality Systems XI, vol. 5291, pp. 93-104, 2004.

[14] H. Tamanoue, M. Okui and F. Okano, “Geometrical Analysis of Puppet-Theater and Cardboard Effects in Stereoscopic HDTV Images”, IEEE Trans. on CSVT, vol. 16, no. 6, pp. 744-752, June 2006.

[15] I. Park, M. Kim, S. Lee, Y. Lee, J. Kwon, W. Hwang, H. Kim, J. Bae, J. Yoo, and H. Byun, “Interactive Multi-view Video and View-dependent Audio under MPEG-21 DIA,” 3DTV Conference 2007, Kos Island, Greece, pp. 7-9, May 2007.

[16] T. Alpert and J. Evain, “Subjective Quality Evaluation - The SSCQE and DSCQE Methodologies,” EBU Technical Review, no. 271, 1997.

[17] <http://research.microsoft.com/vision/InteractiveVisualMediaGroup/3DVideoDownload/>, MSR.

저 자 소 개



이 재 호

- 2006년 : 부산대학교 전자전기통신공학부 학사
- 2007년 ~ 현재 : 한국전자통신연구원 실감방송시스템연구팀 위촉연구원
- 2008년 : 한국정보통신대학교 (ICU) 공학부 석사
- 2008년 ~ 현재 : 한국과학기술원 (KAIST) 정보통신공학과 박사과정
- 주관심분야 : 3D 영상처리, 의료영상처리, 패턴인식



김 창 익

- 1989년 : 연세대학교 전기공학과 학사
- 1991년 : 포항공과대학교 (POSTECH) 전자전기공학과 석사
- 1991년 ~ 1997년 : SKC Ltd. R&D 센터 선임연구원
- 2000년 : 워싱턴주립대학교 전기공학과 박사
- 2000년 ~ 2005년 : Senior member of Technical Staff, Epson Palo Alto Laboratory, Epson R&D Inc.
- 2005년 ~ 2009년 : 한국정보통신대학교 (ICU) 공학부 부교수
- 2009년 ~ 현재 : 한국과학기술원 (KAIST) 전기 및 전자공학과 부교수
- 주관심분야 : 3-D video processing, image/video understanding, intelligent media processing, video coding for IPTV