임의의 가상시점 홀로그램 서비스를 위한 중간시점 영상 및 디지털 홀로그램 생성

서영호* · 이윤혁** · 구자명** · 김동욱***

Intermediate View Image and its Digital Hologram Generation for an Virtual Arbitrary View-Point Hologram Service

Young-Ho Seo* · Yoon-Hyuk Lee** · Ja-Myung Koo** · Dong-Wook Kim***

이 논문은 2013년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2010-0026245).

요 약

본 논문은 최근 관심이 고조되고 있는 디지털 홀로그램의 시야각을 확보하기 위하여 시청자의 시점을 추적하여 그 시점에 해당하는 데이터를 생성하고, 이를 디지털 홀로그램으로 변환하는 방법을 제안한다. 이 방법은 제어하는 시야각의 맨 좌측과 맨 우측 시점에 대한 정보(깊이정보와 컬러 또는 명도정보)가 주어졌다고 가정한다. 이 방법은 주어진 좌, 우측의 깊이영상을 대상으로 스테레오 정합에 의해 단위 깊이 당 의사변위증분을 구하여 사용한다. 이를 이용하여 주어진 좌, 우측시점으로부터 원하는 가상시점의 정보를 생성하고, 그 결과의 두 영상을 결합하여 해당 시점의 정보를 획득한다. 이 경우 발생하는 비폐색 영역을 정의하고 이를 채우는 방법을 제안한다. 이 방법을 구현하여 실험한 결과 생성한 중간시점의 깊이영상과 RGB 영상의 평균 화질은 각각 33.83[dB]과 29.5[dB]이었으며, 평균 수행속도는 프레임 당 250[ms]이었다. 또한 이 방법을 이용하여 시청자와 인터랙티브하게 디지털 홀로그램을 서비스하는 시스템의 프로토타입을 제안한다. 이 시스템에는 좌, 우 시점의 영상정보를 획득, 카메라 캘리브레이션과 영상보정, 중간시점 영상 생성, 컴퓨터-생성 홀로그램(computer-generated hologram, CGH) 생성 및 홀로그램 영상 복원 기능을 포함한다. 이 시스템은 LabView® 환경에서 구현되며, CGH 생성과 홀로그램 영상 복원은 GPGPU로, 나머지는 소프트웨어로 구현한다. 구현 결과 평균 수행 속도는 초당 약 5 프레임을 처리할 수 있는 속도이었다.

ABSTRACT

This paper proposes an intermediate image generation method for the viewer's view point by tracking the viewer's face, which is converted to a digital hologram. Its purpose is to increase the viewing angle of a digital hologram, which is gathering higher and higher interest these days. The method assumes that the image information for the leftmost and the rightmost view points within the viewing angle to be controlled are given. It uses a stereo-matching method between the leftmost and the rightmost depth images to obtain the pseudo-disparity increment per depth value. With this increment, the positional informations from both the leftmost view point and the rightmost view point are generated, which are blended to get the information at the wanted intermediate view point. The occurable dis-occlusion region in this case is defined and a inpainting method is proposed. The results from implementing and experimenting this method showed that the average image qualities of the generated depth and RGB image were 33.83[dB] and 29.5[dB], respectively, and the average execution time was 250[ms] per frame. Also, we propose a prototype system to service digital hologram interactively to the viewer by using the proposed intermediate view generation method. It includes the operations of data acquisition for the leftmost and the rightmost view points, camera calibration and image rectification, intermediate view image generation, computer-generated hologram (CGH) generation, and reconstruction of the hologram image. This system is implemented in the LabView® environments, in which CGH generation and hologram image reconstruction are implemented with GPGPUs, while others are implemented in software. The implemented system showed the execution speed to process about 5 frames per second.

키워드

디지털 홀로그램, 시청자 인터랙티브, 중간시점 영상 생성, 컴퓨터-생성 홀로그램, 디지털 홀로그램 서비스 시스템 프로토타입

Key word

digital hologram, viewer-interactive, intermediate view image generation, computer-generated hologram (CGH), digital hologram service system prototype

* 종신회원 : 광운대학교
** 준회원 : 광운대학교
*** 종신회원 : 광운대학교 (교신저자, dwkim@kw.ac.kr)

접수일자 : 2012. 07. 27
심사완료일자 : 2012. 08. 20

Open Access <http://dx.doi.org/10.6109/jkiice.2013.17.1.15>

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

영화 ‘아바타’가 촉발시킨 3D 영상의 열풍은 과거 몇 차례 출시되었던 3D 영화와는 완전히 다른 양상으로 전개되고 있다. 이것은 아바타가 최고의 관심을 끌어 모음에 따라 세계 유수의 3D 관련 기관들이 그간 준비해놓은 많은 장비, 장치, 도구들을 출시하기 시작하였고, 여기에는 3D 영상을 획득하는 장치로부터 3D 영상 콘텐츠를 편집하고 저장하는 도구까지 모든 분야를 아우르고 있다. 따라서 3D와 관련된 열풍은 과거와는 달리 이제 돌이킬 수 없는 상황에 이르렀다고 볼 수 있다.

현재의 3D 영상/비디오는 대부분 좌우의 두 시점에 해당하는 영상을 좌안과 우안에 각각 강제로 주입시키는 스테레오스코피(stereoscopy) 방식이며, 능동식(셔터 글래스 방식) 또는 수동식(편광방식)의 안경식 모니터를 사용하고 있다. 이러한 현재의 방식에 대해 시청자들은 세 가지 정도의 불편함/불편감을 호소하고 있다[1][2]. 그 첫 번째가 안경을 착용하여야 하는 불편함이다. 현재 무안경 방식의 기술이 상당한 수준에 이르고 있으나, 뒤에 설명할 불편감이 무안경식의 경우 더욱 두드러지기 때문에 아직은 상용화에 적극적인 행보를 보이지 못하고 있다. 두 번째는 시청자의 움직임에 좌우 영상이 서로 교차하거나 각 눈의 초점에 맞지 않아 입체감을 제대로 느끼지 못하는 현상이다.

이 문제를 해결하기 위하여 다시점(multi-view) 또는 극초다시점(ultra multi-view), 궁극적으로는 오토스테레오스코피(auto-stereoscopy) 기술이 개발되고 있는데, 이들 또한 대부분 무안경 방식이라 동일한 문제점을 안고 있으며, 과다한 전송 데이터양의 문제까지 갖고 있다. 마지막으로 두 시점의 영상을 양안에 강제 주입하는 방식으로 인한 시청자의 불편감이다. 이것은 일반적으로 사람이 사물을 볼 때와 조금 다른 메카니즘 때문에 발생하는 것으로 알려져 있으며, 심한 경우 현기증이나 혼절에 이르기까지 한다. 이에 최근 시청자의 안전한 시청을 위해서 국제 표준화기구[3] 및 일본[4]에서 시청안정성을 위한 가이드라인을 발표한 바 있으며, 우리나라에서도 시청안전성을 위한 시청 및 콘텐츠 제작 가이드라인을 발표한 바 있다[5].

이에 최근 주목받고 있는 기술이 홀로그래프 기술이다. 홀로그래프는 1946년 Dennis Gabor가 발명한 이래 조용하지만 꾸준히, 주로 유럽, 미국, 일본에서 연구되어 왔다[6]. 홀로그래프 기술은 공간상에 3D영상을 맺히는 기술이므로, 일반 사물을 보듯이 홀로그래프 영상을 안경을 착용하지 않고 볼 수 있기 때문에 위에서 언급한 문제를 모두 해결할 수 있다.

따라서 이 기술은 디지털 영상분야의 궁극적인 목표로 여겨지고 있다[6][7]. 최근 우리나라도 디지털 홀로그래프에 대한 연구에 박차를 가하고 있으며, 2009년부터 정부의 연구개발과제도 시작되었다. 그러나 아직 홀로그래프 영상의 크기와 해상도에 있어서 일반적인 디스플레이 수준에 이르지 못하고 있으며, 현재의 기술수준에 의한 영상의 시야각이 매우 좁다는 한계 때문에 홀로그래프를 서비스하는 시스템에 대한 연구는 거의 진행되지 않고 있다. 그러나 [7]에서는 시청자의 눈을 추적하여 초점 근처의 작은 영역만 홀로그래프 영상으로 하고 나머지는 2D영상을 보여주는 디스플레이 시스템을 개발하기도 하였다.

본 논문에서는 디지털 홀로그래프를 서비스하는 시스템을 타겟으로 하고, 홀로그래프의 시야각이 좁은 단점을 보완하기 위해서 시청자의 시점을 추적하여 그 시점에 맞는 홀로그래프 영상을 디스플레이하는 시스템을 제안하고자 한다. 이를 위해서 시청자의 시점에 해당하는 시점의 영상정보를 실시간으로 추출하는 알고리즘을 개발한다. 또한 개발한 중간시점 영상 생성 방법을 적용하여 홀로그래프를 서비스하는 시스템의 프로토타입(prototype)을 구현하고자 한다. 이 시스템에는 데이터 획득을 위한 카메라 시스템, 카메라 보정 및 영상보정, 중간시점 영상정보 생성, 컴퓨터-생성 홀로그래프(computer-generated hologram, CGH) 생성, 홀로그래프 영상 복원 기능을 포함한다.

본 논문의 구성의 구성은 다음과 같다. 2장에서는 본 논문이 타겟으로 하는 홀로그래프 서비스 시스템의 프로토타입에 대해 설명하고, 3장에서는 이 시스템에 사용될 중간시점 영상생성 방법을 제안한다. 4장에서는 제안한 중간시점 영상생성 방법과 홀로그래프 서비스 시스템 프로토타입을 구현하여 그 성능을 보이고, 이를 토대로 5장에서 본 논문의 결론을 맺는다.

II. 홀로그램 서비스 시스템 프로토타입

본 논문이 홀로그램의 서비스 시스템을 타겟으로 하고, 그 프로토타입을 구현하는 것이 목적이기 때문에 이 프로토타입 시스템부터 설명한다. 본 논문에서는 기존의 2D 또는 3D 영상/비디오를 서비스하는 시스템의 프레임워크를 그대로 유지하는 것으로 가정하였다. 즉, 송신측에서 영상데이터를 획득하고 가공, 압축, 정보보호 및 보안 솔루션을 적용한 다음 송신하면, 수신측에서 이를 수신하여 정보보안과 압축 디코딩을 수행하여 디스플레이하는 시스템이다.

이 서비스 프레임워크를 유지한 홀로그램 서비스 시스템 프로토타입을 그림 1에 흐름도의 형태로 보이고 있다. 이 시스템에서 기존의 시스템 과정 중 압축과 정보보호 및 보안 솔루션은 포함되지 않았는데, 이들은 본 연구팀이 이미 연구를 진행한 바 있으나[8][9][10], 프로토타입이 목적이기 때문에 이 과정들을 제외시켰다.

그림 1은 위쪽과 아래쪽의 두 개의 흐름도로 구성되어 있다. 위쪽 흐름도는 송신측의 연산능력이 더 좋은 경우로, 중간시점 영상생성 및 CGH 생성을 송신측에서 수행하여 그 정보를 수신측으로 보내는 경우이다.

그림 1. 구현할 프로토타입 홀로그램 서비스 시스템
Fig. 1 The prototype hologram service system to be implemented

반면, 아래쪽 흐름도는 수신측의 연산능력이 더 좋아 수신측에서 이 과정들을 수행하는 흐름도이다. 서론에서 밝힌 바와 같이 본 논문의 목적 중 하나는 홀로그램

영상의 시야각을 넓히기 위하여 시청자의 시점을 추적하여 그 시점에 해당하는 홀로그램 영상을 디스플레이하는 것이다. 따라서 얼굴추적에 의한 시점제어(View Point Control by Face Tracking) 기능이 포함되어 있으며, 이 데이터에 따른 중간시점 영상 생성(Intermediate view Rendering) 기능을 포함하고 있다. 중간시점 영상생성 기능에 대해서는 다음 장에서 상세히 다루고, 이 장에서는 나머지 기능블록들에 대해서만 설명한다.

2.1. CGH 생성 (CGH Generation)

이 기능블록은 송신측 맨 끝에 있으나, 이 기능블록을 먼저 설명하는 것이 앞으로의 설명을 더욱 잘 이해할 것으로 판단하여 먼저 설명한다. 홀로그램은 두 개의 광파(물체파와 기준파)를 간섭시켜 만든 프린지 패턴(fringe pattern)이다. 그러나 간섭현상을 일으키는 장치나 도구가 복잡하고 정교하며, 또한 휴대하기 용이치 않기 때문에 CGH를 생성하여 디지털 홀로그램으로 사용한다[6][7]. CGH는 간섭현상을 수학적으로 모델링한 것인데, 식 (1)에 그 식을 나타내었으며, 식 (2)는 홀로그램이나 물체평면의 크기가 홀로그램 평면과 물체와의 거리(z_j)보다 매우 작다고 가정하고 식 (1)을 간략화한 것이다.

$$I_{\alpha j} = \sum_j A_j \cos \left[\frac{2\pi}{\lambda} \sqrt{(p_{\alpha} x_{\alpha} - p_j x_j)^2 + (p_{\alpha} y_{\alpha} - p_j y_j)^2 + (p_j z_j)^2} \right] \quad (1)$$

$$I_{\alpha j} \approx \sum_j A_j \cos \left[2\pi \left(\frac{p_j z_j}{\lambda} + \frac{1}{2\lambda p_j z_j} [(p_{\alpha} x_{\alpha} - p_j x_j)^2 + (p_{\alpha} y_{\alpha} - p_j y_j)^2] \right) \right] \quad (2)$$

여기서, $I_{\alpha j}$ 는 물체평면의 좌표 (x_j, y_j, z_j) 의 광원(빛의 세기 A_j)이 홀로그램 평면의 좌표 (x_{α}, y_{α}) 에 미치는 간섭영향을 나타내고, p_j 와 p_{α} 는 물체평면과 홀로그램평면의 화소크기를 각각 나타낸다. 간섭현상을 일으키는 두 광파는 동일한 것으로 간주하고(분파하여 사용함), 그 파장은 λ 이다.

식 (2)에서 코사인 함수의 각도항(argument)을 분리하여 계산할 수 있고, 또 홀로그램 평면의 한 열에서 맨 좌측 또는 한 행에서 맨 위의 화소값을 계산할 때 생성되는 중간결과들을 사용하면 반복연산 방법으로 나머지 화소값들을 계산할 수 있다. 이때 하나의 광원이 모든 홀로그램 화소에 미치는 영향을 계산하면 물체 화소의 z_j 의 변화를 최소로 할 수 있다[11]. 그러나 이 방법

의 단점은 각 광원으로 홀로그램 전체에 미치는 영향을 계산한 결과를 다음 광원에 대한 영향을 계산하기 전에 모두 메모리에 저장하였다가 다음 광원의 영향에 누적 덧셈을 하기 위해 다시 메모리에서 꺼내 와야 하기 때문에 메모리 접근횟수와 시간 때문에 병목현상이 심하게 발생한다. 이에 본 연구팀은 한 홀로그램 화소에 모든 광원이 미치는 영향을 계산하고 다음 홀로그램 화소로 진행하는 방식을 제안하여, 다른 모든 성능은 변함이 없고 메모리 접근 횟수를 약 1/20,000로 줄인 계산방식을 제안한 바 있으며[12], 이 논문에서도 이 방법을 사용한다.

2.2. 영상 데이터 획득 시스템(Image Acquisition System)

본 논문의 목적이 디지털 홀로그램 서비스이고 식 (1) 또는 (2)의 CGH를 생성하여 사용하므로, 이 식에 나타난 물체의 각 화소에 대한 (x_j, y_j, z_j) 좌표와 광강도 A_j 가 필요하다. 이들 데이터를 획득하는 시스템은 아래와 같은 방법이 가능하다.

- ① 두 대의 2D 카메라로 두 시점의 영상을 획득, 시차(disparity)를 구하여 z_j 를 얻는 방법
- ② 한 대의 깊이카메라와 한 대의 2D카메라를 사용하는 방법

이 두 가지 방법 중 ①은 스테레오 정합(sterео matching)을 수행하여야 하는데, 다양한 정합방법이 제안되었으나, 정확도와 연산시간 간에 상보적인 관계를 갖는 것이 보통이다. 따라서 실시간을 목표로 하는 본 논문의 시스템에 사용하기에는 적합지 않다. 스테레오 정합으로 획득한 시차(d)로 깊이정보를 계산하는 식은 식 (3)과 같다.

$$z = \frac{Bfs}{d} \quad (3)$$

여기서, z 는 깊이값, B 는 두 카메라간 거리, f 는 카메라의 초점거리를 각각 나타낸다.

(a)

(b)

그림 2. 수직리그를 이용한 깊이카메라와 RGB카메라의 배열; (a) 수직리그에 배치된 두 카메라, (b) 사용한 콜드미러의 특성

Fig. 2 Arrangement of a depth camera and an RGB camera: (a) arranged two cameras, (b) characteristics of the used cold mirror

②의 방법이 최근 많이 사용하는 방법인데, 이 방법에서는 두 카메라의 시점을 정확히 일치시키고, 광각과 해상도를 조정하여야 한다. 시점을 일치시킨다는 관점에서는 최근 수직 리그(rig)를 사용하는 경향이 두드러지고 있는데, 본 논문에서도 수직리그를 사용하여 카메라를 배치하였으며, 그림 1 (a)에 그 사진을 보이고 있다. 여기에 사용된 깊이카메라는 Mesa Image사의 SR4000[13]이고 해상도는 176×144 이었으며, RGB카메라는 회사는 Allied vision technologies사의 Pike F210C IRF24[14]이고, 해상도는 $1,920 \times 1,080$ 이었다. 따라서 두 영상간의 해상도 및 크기를 조절하여 동일한 영상의 동일한 해상도로 만들어 사용하였다. 사용한 깊이카메라는 적외선을 사용하므로, 그림 2 (a)에서 보는 바와 같이 콜드미러(cold mirror)를 제작하여 사용하였는데, 그 파장에 대한 반사율 특성을 그림 2 (b)에 보이고 있다.

SR4000이 사용하는 적외선의 파장이 850nm이므로 이 파장 대에서는 반사율이 5% 정도이고, RGB카메라가 사용하는 가시광선 대에서는 거의 100%의 반사율을 보이고 있다.

(a)

(b)

(c)

(d)

(f)

그림 3. 수평리그를 이용한 두 SR4000 카메라의 배치 및 획득한 영상의 예 : (a) 카메라 시스템; 좌측 카메라로 획득한 (b) 깊이영상, (c) 명도영상; 우측카메라로 획득한 (d) 깊이영상, (e) 명도영상
Fig. 3 Two SR4000 camera arrangement using a horizontal rig and the acquired image examples: (a) the camera system; acquired (b) depth image, (c) luminance image from left camera; acquired (d) depth image, (e) luminance image from right camera

최근 Microsoft사에서 저가의 RGB+깊이 카메라 (Kinect)[15]를 출시하여 큰 관심을 끌고 있는데, RGB와 깊이영상 모두 640×480 의 꽤 높은 해상도를 제공한다. 이 카메라는 적외선의 구조광을 조사하여 그 패턴과 물체에서 반사되어온 빛의 패턴을 스테레오 정합하여 깊이정보를 획득한다. 따라서 정합의 기준이 되는 영상에 따라 좌 또는 우측에 비폐색(disocclusion)영역이 발생하여 정확성을 요하는 곳에는 사용하기 곤란하다. 한편 SR4000 또한 깊이영상 뿐만 아니라 빛의 밝기에 해당하는 명도영상을 제공한다. 비록 해상도가 낮기는 하지만, 깊이영상의 정확도가 상대적으로 높다. 따라서 본 논문에서는 컬러영상이 필요한 경우는 그림 2의 시스템을 사용하고, 컬러가 아닌 광강도(luminance, 명도)만을 필요로 할 때에는 SR4000을 사용한다.

본 논문의 목적 중 하나가 시청자의 시점을 추적하여 시청자의 시점에 해당하는 홀로그램을 서비스하는 것이다. 따라서 최소 두 개의 시점정보가 필요한데, 그림 3(a)에 두 개의 SR4000을 수평리그에 배치한 획득시스템을 보이고 있고, (b)(c)(d)(e)는 좌우 카메라에서 획득한 깊이영상과 명도영상을 각각 보이고 있다.

2.3. 카메라 캘리브레이션 및 영상 보정 (Camera Calibration and Image Rectification)

본 논문의 목적 중 하나가 시점을 제어하는 것이고, 이를 위해서는 중간시점의 영상정보를 생성하여야 한다. 이것은 내, 외부의 카메라 파라미터들[16]을 알고 있어야 가능한데, 이 파라미터들이 주어지는 경우도 있으나 그렇지 않은 경우도 있다. 따라서 본 논문에서는 획득된 영상으로부터 카메라 파라미터들을 획득하여 사용한다. 또한 그림 3의 두 카메라의 수평을 최대한 맞추었다고 하더라도 두 카메라로 획득한 영상의 수평이 맞지 않을 수 있으므로 두 획득영상의 수평과 수직을 맞추는 영상보정을 수행한다.

그림 4는 카메라 캘리브레이션과 영상보정을 위해 본 연구팀이 만든 GUI를 보이고 있다. 이 그림은 SR4000으로부터 획득한 명도영상으로 캘리브레이션과 보정을 수행하는 예인데, 여기서는 체스(chess) 보드를 사용한다. 총 30쌍의 영상을 획득하여 사용할 수 있으며, 캘리브레이션이 끝나면 체스보드의 각 경계점들의 (x, y) 좌표를 이용하여 영상보정을 수행한다.

그림 4. 카메라 캘리브레이션과 영상보정을 위한 GUI

Fig. 4 GUI for camera calibration and image rectification

카메라 캘리브레이션은 처음 영상을 획득할 때 한 번 수행하면 되므로, 이 기능은 시스템 내에 넣지 않고 외부에서 수행하여 그 결과를 받아 사용한다.

2.4. 얼굴추적에 의한 시점제어 (View Point Control by Face Tracking)

시청자의 시점에 해당하는 디지털 홀로그램을 서비스하기 위해서는 시청자의 얼굴을 추적하여야 하는데, 본 연구팀은 실시간 동작이 가능한 얼굴탐색 및 추적 알고리즘을 이미 제안한 바 있다[17]. 여기서는 이 방법을 간략히 소개한다. 이 방법은 얼굴검출과 얼굴추적의 두 과정으로 이루어져 있으며, 얼굴검출은 장면이 바뀌거나 시청자가 사라지는 경우 등에 있어서 시청자의 얼굴을 찾는 과정이다. 따라서 이 과정은 초기에 한 번만 수행하게 되고, 일단 얼굴을 검출하면 추적과정을 반복수행하면서 지속적으로 얼굴을 추적한다.

얼굴검출 과정은 깊이영상과 RGB영상을 사용하는데, 깊이영상은 그전 프레임과의 차영상을 구하여 움직임이 있는 영역을 검출하고, RGB영상으로는 피부색깔 영역을 검출한다. 검출된 두 영상의 공집합에 해당하는 부분을 [18]의 Adaboost 알고리즘을 사용하여 얼굴을 검출한다. 추적결과는 얼굴에 해당하는 좌표, 검출된 얼굴의 RGB영상 및 깊이영상이며, 이들을 템플릿(template)으로 하여 얼굴추적 과정이 수행된다.

얼굴추적 과정은 템플릿 정합방법을 사용하는데, 추적을 위한 탐색영역을 얼굴의 깊이값에 따라 가변적으로 정하며, 그 값을 실험적으로 구하여 사용하였다. 템플릿 정합에 사용한 비용함수는 SAD(sum of absolute differences)이며, 기본적으로는 탐색영역 중 가장 작은 비용함수 값을 갖는 위치를 추적결과로 간주한다. 그러나 정합 중 일정한 값(문턱치 α) 이하의 비용함수 값이 발생하면, 그 위치를 추적결과로 간주하고 추적을 조기 종료한다. 또한 탐색영역의 비용함수 값 중 가장 작은 값이 주어진 값(문턱치 β)을 넘으면 추적에 실패한 것으로 간주하고 얼굴검출 과정으로 돌아가서 얼굴을 재검출한다. 여기에 사용한 문턱치 α 와 β 는 실험적으로 구하여 사용하였다.

이 방법의 구현 결과는 얼굴검출의 평균시간은 full-HD(1,920×1,080) 영상의 프레임 당 157.31ms, 얼굴추적은 평균 8.42m가 소요되었다. 오검출율은 얼굴탐색이 1.39%, 얼굴추적이 1.3%이었다. 얼굴추적 역시 시스템의 기능과 직접적인 관계가 없으므로 외부에서 수행하고 그 결과를 받아 사용한다.

2.5. 홀로그램 영상의 복원 (Holographic Display)

수신측의 최종단계는 디지털 홀로그램을 영상으로 복원하는 것이다. 실제의 경우는 광학장치를 이용하여 공간상에 영상을 복원하여야 하지만, 연구단계에서는 광학장치보다 시뮬레이션에 의한 복원을 많이 사용한다. 본 연구에서는 이 두 가지 방법을 모두 고려한다.

2.5.1. 광학장치에 의한 복원(Optical Holographic Display)

그림 5에 홀로그램 영상을 복원하는 광학장치를 홀로그램을 생성하는 장치와 같이 나타내었다. 그림에서 보는 바와 같이 홀로그램을 생성할 때 단일 레이저광을 분파하여 하나는 물체에 조사하고(물체파) 다른 하나는 직접 홀로그램 평면에 조사하여 두 광파의 간섭현상을 디지털 장치로 획득한다. 영상복원 장치는 획득장치와 매우 유사하며, 분파된 두 광파 중 하나만을 사용하는 것이 다르다.

그림 5. 디지털 홀로그램 획득 및 영상복원;
(a) 획득, (b) 영상복원

Fig. 5 Digital hologram acquisition and its image reconstruction; (a) acquisition, (b) image reconstruction

디지털 홀로그램 영상을 복원할 때 획득과 동일한 광원을 사용하는 것이 일반적이며, 획득과 동일한 조건의 장치를 사용한다. 디지털 홀로그램은 광파의 회절현상을 적절히 일으키는 장치(공간광변조기, *spatial light modulator*, SLM)에 띄우고, 여기에 획득할 때의 기준파와 동일한 광파를 조사하면 원래의 물체 위치에 그 물체의 영상이 맺히게 된다.

2.5.2. 시뮬레이션에 의한 복원 (Holographic Display Simulation)

그림 5의 광학장치에 의하지 않고 시뮬레이션에 의해 영상을 복원할 수도 있다. 물론 이 경우는 복원될 거리를 지정하여야 하기 때문에 그 거리의 평면에 맺히는 2차원 영상을 복원하게 된다. 이 방법은 Fresnel 변환을 이용하며, 식 (4)와 같다[6].

$$F(x, y) = \frac{e^{ikz}}{i\lambda z} e^{\frac{i\pi}{\lambda z}(x^2 + y^2)} \iint f(\xi, \eta) e^{\frac{i\pi}{\lambda z}(\xi^2 + \eta^2)} e^{-\frac{i2\pi}{\lambda z}(\xi x + \eta y)} d\xi d\eta \quad (4)$$

여기서 $f(\xi, \eta)$ 와 $F(x, y)$ 는 홀로그램 영상과 복원된 영상의 화소값을 각각 나타내며, $i = \sqrt{-1}$ 이다. z 는 복원영상의 깊이값, k 와 λ 는 사용한 광파의 파수(wave number)와 파장을 각각 나타낸다.

III. 의사변위 값을 이용한 중간시점 영상 생성

앞에서 언급한 바와 같이 시청자의 시점에 해당하는 홀로그램 영상을 제공하기 위해서는 시청자의 시점에 해당하는 영상을 생성하고 그 영상에 대한 홀로그램을 생성하여야 한다. 일반적으로 중간시점 영상을 생성하는 방법을 도식적으로 나타내면 그림 6과 같다. 즉, 보정된 좌영상과 우영상을 입력받아 주어진 방법에 의해 좌영상과 우영상 각각으로부터 원하는 시점의 영상을 생성한다. 그 결과의 두 영상을 결합(blending)하고, 결합된 영상에서 채워지지 않은 부분을 후처리과정에서 채워 넣는다.

주어진 좌영상과 우영상으로부터 원하는 시점의 영상을 생성하는 가장 일반적인 방법은 3D워핑(warping) 방법이다[19-22]. 특히 [21]에서는 MPEG-4에서 3D워핑을 이용하여 중간시점 영상을 생성하는 참조 소프트웨어(reference software)를 제공하고 있다. 그러나 3D워핑 방법은 카메라의 내부 파라미터들을 모두 알고 있어야 가능하며, 매 프레임 워핑에 의해 중간시점 영상을 생성하여야 하기 때문에 상당히 많은 시간이 소요된다.

그림 6. 중간시점 영상 생성
Fig. 6 Intermediate view image rendering

따라서 본 논문에서는 소요시간을 최소화 할 수 있는 중간시점 깊이영상 및 RGB영상을 생성하는 방법을 제

안한다. 이 방법을 그림 7에 나타내었는데, 깊이값의 증분에 대한 변위의 증분을 계산하여 사용하며, 변위는 주어진 두 깊이영상에 대해 스테레오 정합으로 구한다. 단위 깊이값에 대한 변위증분의 연산은 한 장면(scene)의 첫 번째 프레임만을 대상으로 수행하므로 이 과정은 전처리(pre-processing) 과정에 해당한다. 일단 변위증분이 계산되면 나머지 프레임들은 이 변위증분을 사용만 하면 된다.

3.1. 스테레오 정합에 의한 의사변위증분 계산

단위 깊이값에 대한 변위증분을 계산하기 위해서는 최소 두 깊이값에 대한 변위값을 알아야 하는데, 본 논문에서는 스테레오 정합을 사용한다.

그림 7. 변위증분을 사용한 중간시점 영상 생성
Fig. 7 Intermediate view Image rendering by disparity increment

3.1.1. 두 깊이값에 대한 스테레오 정합

일반적으로 스테레오 정합 방법은 상당한 연산시간을 소요하는 방법이 아니면 신뢰도가 많이 떨어진다. 그러나 본 논문에서는 연산시간을 많이 소요하는 전역적(global) 방법을 사용하지 않고 국부적(local) 방법을 사용한다. 대신 신뢰도를 높이기 위하여 동일 깊이값에 대해 한 화소가 아니라 여러 화소에 대해 스테레오 정합을 사용하여 변위값의 빈도수가 가장 높은 것을 사용한다.

이를 위해서 먼저 깊이영상을 스캔하여 상위 10%와 하위 10% 내에 있는 깊이값들 중 빈도수가 가장 많은 깊이값을 각각 선정한다(high-frequent high & low depth values extraction). 이 두 깊이값을 갖는 화소들에 대해 주어진 좌우영상을 대상으로 스테레오 정합을 수행한다. 이 때 사용한 정합방법은 단일 정합창(matching window)을 사용한 국부적 방법이며, 정합창은 5×5 이었다. 상위 10%와 하위 10%내의 각 깊이값에 대한 정합결과 가장 빈도수가 높은 각 값을 해당 깊이값에 대한 변위값으로 결정한다.

3.1.2. 변위증분 계산

변위와 깊이와의 관계는 식 (3)과 같다. 이 식에서 보는 바와 같이 변위와 깊이와의 관계에서는 카메라 파라미터인 초점거리(f)와 두 카메라간의 거리(B)가 필요하다. 그러나 본 논문에서는 이런 파라미터들이 알려져 있지 않다고 가정하고, 이 파라미터들은 사용하지 않는다. 이 식에 나타난 바와 같이 깊이와 변위는 반비례의 관계를 갖는다. 식 (3)을,

$$d_j = \frac{fB}{z_j} = fBd_{p,j} \quad (5)$$

와 같이 변위와 깊이의 비례식으로 변형할 수 있다. 여기서 $j \in \{H, J\}$ 이며, z_j 는 스테레오 정합을 수행한 상위 10%와 하위 10%내의 깊이값을 각각 지칭한다. 또한 d_j 는 해당 깊이값의 측정된 변위값이며, $d_{p,j}$ 는 깊이값을 변환한 변위값(여기서는 의사변위값(pseudo-disparity value)라 칭한다)이다. 여기서 의사변위증분을 다음과 같이 정의한다.

$$d_{p,icmt} = \frac{d_H - d_L}{d_{p,H} - d_{p,L}} = fB \quad (6)$$

한편, 그림 8과 같이 좌영상과 우영상 사이의 임의의 가상시점을 생각하자. 이 시점의 카메라와 좌측 카메라간의 거리를 B' , 우측 카메라와의 거리를 B'' 이라고 하자. 좌영상, 우영상 그리고 가상시점에서 동일 물체에 대한 깊이값은 같으므로, 좌영상과 우영상에 대비한 이 가상시점에서의 의사변위증분 값을 각각 $d'_{p,icmt}$ 과

$d''_{p,icmt}$ 라 하면, 이들은 카메라간의 거리에 비례하므로,

$$d'_{p,icmt} = fB' = d_{p,icmt} \frac{B'}{B} \quad (7)$$

$$d''_{p,icmt} = fB'' = f(B - B') = d_{p,icmt} \frac{B''}{B} = d_{p,icmt} - d'_{p,icmt} \quad (8)$$

로 각각 구할 수 있다.

그림 8. 가상시점
Fig. 8 Virtual view point

3.2. 초기 중간시점 영상 생성

식 (6)은 또한 좌영상 또는 우영상 내의 임의의 변위 값 d_i 와 해당 의사변위 값 $d_{p,i}$ 에 동일하게 적용된다.

$$d_{p,icmt} = \frac{d_i - d_L}{d_{p,i} - d_{p,L}} = \frac{d_H - d_i}{d_{p,H} - d_{p,i}} \quad (9)$$

따라서 좌영상 또는 우영상의 임의의 깊이 값 z_i 에 대한 실제의 변위 d_i 는,

$$d_i = d_L + d_{p,icmt}(d_{p,i} - d_{p,L}) = d_H - d_{p,icmt}(d_{p,H} - d_{p,i}) \quad (10)$$

로 구할 수 있다. 식 (10)은 좌영상 또는 우영상에서 측정된 두 위치의 변위 중 하나를 기준으로 임의의 깊이 값을 갖는 화소의 위치를 정의하고 있다. 그림 8의 임의의 가상시점에서는 식 (8)에 의해 해당 의사변위 증분값을 구하고, 좌영상과 우영상의 d_L 또는 d_H 에 해당하는 화소를 기준으로 임의의 깊이 값 z'_i 와 z''_i 를 갖는 화소의 위

치를 다음과 같이 계산할 수 있다.

$$d'_i = d_L + d'_{p,icmt}(d_{p,i} - d_{p,L}) = d_H - d'_{p,icmt}(d_{p,H} - d_{p,i}) \quad (11)$$

$$d''_i = d_L + d''_{p,icmt}(d_{p,i} - d_{p,L}) = d_H - d''_{p,icmt}(d_{p,H} - d_{p,i}) \quad (12)$$

또한 식 (10)이 깊이값에 관여되지만, 해당 깊이값에 대한 변위 또는 위치가 결정되면 그 화소에 해당하는 RGB 또는 명도영상의 화소의 위치도 동일하다. 따라서 식 (10)을 이용하면 중간시점의 깊이영상 뿐만 아니라 해당 시점의 RGB 또는 명도영상을 구할 수 있다. 이렇게 생성한 중간시점 영상을 그림 7에서는 초기 중간시점 영상(initial IVPI(intermediate view point image))라고 명명하였다. 그림 7에서 보는 바와 같이 본 논문에서는 좌영상과 우영상으로부터 각각 초기 깊이 IVPI와 RGB 또는 명도 IVPI를 각각 생성한다.

3.3. 후처리 과정

3.3.1. 초기 중간시점 영상 합성

좌영상과 우영상으로부터 생성된 각각의 IVP 깊이영상 또는 RGB(명도)영상은 동일한 가상시점의 영상들이다. 후처리의 첫 번째 과정은 이 두 영상을 결합하는 것이며, 중간시점 영상의 좌표 (i, j)에 대해 다음과 같이 수행한다.

① 해당 화소의 좌영상으로부터 생성한 중간시점 깊이영상(left_depth_initial_IVPI(i,j))과 우영상으로부터 생성한 중간시점 깊이영상(right_depth_initial_IVPI(i,j))이 모두 값을 가지지 않을 때는 결합한 중간시점 깊이영상(depth_blended_IVPI(i,j)) 및 RGB 영상(RGB_blended_IVPI(i,j))의 해당 화소를 비특색영역으로 정의한다.

②left_depth_initial_IVPI(i,j)과 right_depth_initial_IVPI(i,j) 중 하나만 값을 가질 때는 그 값을 depth_blended_IVPI(i,j)의 값으로 하고, 해당 RGB 화소를 RGB_blended_IVPI(i,j)의 값으로 결정한다.

③left_depth_initial_IVPI(i,j)과 right_depth_initial_IVPI(i,j) 모두 값을 가질 때는 두 값의 평균값을 depth_blended_IVPI(i,j)의 값으로 하고, 마찬가지로 두 깊이화소에 해당하는 RGB화소의 평균값을 RGB_blended_IVPI(i,j)의 값으로 한다.

3.3.2. 홀필링과 인페인팅

생성한 두 중간시점 영상을 결합한 결과영상에 홀(hole) 또는 비페색영역이 생기는 이유는 다음과 같다. 첫 번째 이유는 좌영상과 우영상 원본 깊이영상의 깊이 값이 정확치 않는 경우이며[20], 특히 특정 위치에 해당하는 두 깊이값이 모두 정확치 않을 때이다. 이런 경우는 영역으로 나타나기 보다는 주로 한두 화소가 값을 갖지 못하는 홀이 생긴다. 두 번째 이유는 좌영상과 우영상 모두에서 가려진 영역이 가상의 중간시점에서 보이는 경우이다. 이 경우가 전형적인 비페색영역이며, 이 경우는 한 영역이 값을 갖지 못하는 것으로 나타난다. 그러나 본 논문에서는 이 두 경우를 구분하지 않고 다음과 같이 처리한다.

3.3.2.1. 중간값 필터링

먼저, 결합된 깊이영상(depth_blended_IVPI)의 각 홀에 대해 그 홀을 중심으로 3×3 창을 씌운다. 홀 화소를 제외한 8개의 화소 중 5개 이상이 값을 가질 경우 그 값들 중 중간값을 그 홀의 깊이값으로 결정한다(median filtering). 이 때 해당 RGB영상의 홀은 깊이영상의 홀을 채운 화소에 해당하는 RGB화소의 값으로 채운다. 8개 중 5개 미만의 화소만 값을 갖는 경우에는 그 홀을 채우지 않고 다음 단계에서 처리한다.

3.3.2.2. 인페인팅

중간값 필터링을 거친 결합영상에는 어느 정도의 영역이 값을 갖지 않은 비페색영역만 존재한다. 위에서 언급한 바와 같이 이 영역들의 정보는 좌영상이나 우영상에 없을 가능성이 높으므로, 이 영역에 대한 정확한 값을 찾는 것은 거의 불가능하다. 지금까지 비페색영역의 인페인팅 기법들은 한 시점이 주어졌다고 가정하기 때문에 비페색영역을 주로 배경영역의 값으로 채워 넣는다. 예를 들어, [22]에서는 비페색영역의 대상 화소를 기준으로 8 방향으로 탐색하여 가장 가까운 비페색영역이 아닌 화소들을 찾아서 그 중 배경에 해당하는 화소들만 골라 그 평균값을 그 화소의 값으로 결정한다.

(a)

(b)

그림 9. 인페인팅 방법
(a) 비페색영역의 발생 (b) 인페인팅.
Fig. 9 Inpainting scheme
(a) appearing disocclusion region (b) inpainting

그러나 본 논문에서는 좌, 우의 두 시점이 주어졌다고 가정하였기 때문에 비페색영역은 그두 물체사이의 영역에서만 발생한다. 이 상황을 그림 9 (a)에 나타내었는데, 이 경우는 가상시점이 좌영상에 가까운 경우이다.

그림에서 나타낸 것과 같이 굵은 점선으로 나타낸 부분이 비페색영역에 해당하며, 이 부분이 굵은 실선으로 나타낸 크기로 영상에 나타나게 된다. 그림에서 보는 바와 같이 이 비페색영역의 실체는 가상시점의 위치에 크게 좌우된다. 또한 시점의 변화가 좌우로만 이

루어진다 고 가정하면 비페색영역에 채워질 내용은 상하의 값보다 좌우의 값에 거의 의존한다.

이런 점을 이용하여 본 논문에서는 다음의 비페색영역 인페인팅 방법을 제안하며, 이 방법을 그림 9 (b)에 도시하였다. 이 방법은 비페색영역의 특정 화소에 대하여 6 방향 탐색을 실시하여, 각 방향의 가장 근접한 비페색영역이 아닌 화소들을 추출한다.

그림에서 각 화소의 이름을 p_{ij} , 화소로부터 탐색된 화소까지의 거리를 d_{ij} 로 나타내었는데, $i \in \{l, r\}$ 는 해당 화소의 좌측방향(l)인지 우측방향(r)인지를 나타내고, $j \in \{u, c, d\}$ 는 탐색방향이 위쪽(u)인지, 아래쪽(d)인지, 수평방향(c)인지를 나타낸다. 그림 9 (a)에 나타난 카메라간 간격을 사용하여 비페색영역 내의 화소 p_d 의 값 $v(p_d)$ 는 탐색된 화소값 $v(p_{ij})$ 과 거리 d_{ij} 로 식 (11)과 같이 계산한다. 이 식은 깊이영상과 RGB 영상 모두에 적용한다.

$$v(p_d) = \frac{\sum_j \left[\frac{B'}{B} v(p_{rj}) d_{rj} + \frac{B'}{B} v(p_{lj}) d_{lj} \right]}{3 \sum_i \sum_j d_{ij}} \quad (13)$$

IV. 구현 및 실험 결과

이 장에서는 3장에서 제안한 중간시점 영상생성 방법과 2장에서 설명한 홀로그램 서비스 시스템 프로토타입을 구현하고 실험한 결과를 보인다.

4.1. 중간시점 영상생성

3장에서 제안한 중간시점 영상생성 방법을 Microsoft Window7 환경에서 Microsoft Visual Studio 2008을 사용하여 구현하였다. 이 때 Intel Core 2 Quad 2.4 GHz CPU와 4GB RAM을 가지고 있는 PC를 사용하였다. 사용한 테스트 시퀀스는 표 1에 보인 바와 같이 4 가지이며, 모두 MPEG에서 사용하는 것이다[21]. 모든 시퀀스는 동일한 길이를 갖도록 200 프레임 씩 잘라서 사용하였다. 이들 시퀀스들은 획득한 깊이정보를 그대로 나타내지 않고 약간의 후처리를 수행한 것으로 판단된다. 즉, 동일시간의 좌영상과 우영상, 그리고 중간시점 영상의 배경이나

물체의 깊이값과 RGB값이 차이를 보이는 경우가 많다. 따라서 제공되는 모든 시점의 깊이값들을 100% 신뢰할 수는 없지만, 객관적인 데이터를 도출하기 위하여 이 시퀀스들을 사용하였다.

그림 10에 카페의 한 영상에 대해 본 논문에서 제안한 방법을 적용한 과정의 결과들을 보이고 있다. (e)와 (f), (g)와 (h)는 각각 좌영상과 우영상으로부터 생성한 초기 깊이 IVPI와 RGB IVPI를 보이고 있는데, 사용한 원 영상에 따라 비페색영역이 방향성을 가지는 것을 볼 수 있다. 이들 영상들을 결합한 영상(i)와 (k)을 거쳐 최종 중간시점 영상(j)와 (l)이 생성되었다. 이 두 영상과 중간시점의 원영상(m)과 (n)을 비교하면, 거의 차이가 나지 않는 것을 확인할 수 있다.

표 1. 테스트 시퀀스
Table. 1 Test sequences

Name	Made by	Resolution	Length
Cafe	GIST	1,920×1,080	200 frames
Lovebird 1	ETRI	1,024×768	200 frames
Poznan street	Poznan Univ.	1,920×1,080	200 frames
Mobile	Philips	720×540	200 frames

표 2. 생성한 중간시점 영상의 평균 화질
Table. 2 Average image qualities of the generated intermediate view images

Name	Depth image [dB]	RGB image[dB]	
	Proposed	Proposed	VSRS 3.5a
Cafe	34.55	29.63	30.53
Lovebird 1	34.18	24.87	25.77
Poznan street	33.50	31.48	32.92
Mobile	33.14	32.05	36.38
Average	33.83	29.50	31.40

표 2에서는 본 논문에서 제안한 방법을 표 1의 시퀀스에 적용하여 구한 평균 화질을 MPEG의 가상시점 영상생성 참조 소프트웨어(reference software) VSRS 3.5a [21]의 결과와 비교하여 보이고 있다.

표에서 보이는 VSRS 3.5a의 결과는 정수화소(integer-pel) 단위로 수행한 것으로, 경계부분의 잡음을

그림 10. 카페영상의 중간시점 영상생성 예; 원본의 (a) 깊이 좌영상, (b) 깊이 우영상, (c) RGB 좌영상, (d) RGB 우영상; 초기 IVPI (e) left_depth, (f) right_depth, (g) left_RGB, (h) right_RGB; (i) depth_blended_IVPI, (j) final_depth_IVPI, (k) RGB_blended_IVPI, (l) final_RGB_IVPI; 원본 IVPI (m) 깊이, (n) RGB

Fig. 10 Example of IVPI generation with Cafe image; original image of (a) left depth, (b) right depth, (c) left RGB, (d) right RGB; initial_IVPI of (e) left_depth, (f) right_depth, (g) left_RGB, (h) right_RGB; (i) depth_blended_IVPI, (j) final_depth_IVPI, (k) RGB_blended_IVPI, (l) final_RGB_IVPI; original IVPI of (m) depth, (n) RGB

제거하지 않고 제안한 방법과 같이 좌우측 영상을 결합하였으며, 홀-필링은 GIST에서 제안한 방법을 추가로 사용하는 방식으로 수행하였다. VSRS는 그 목적이 컬러영상을 생성하는 것이어서 깊이영상을 따로 생성하지 않기 때문에 깊이영상에 대한 정보를 따로 확보하지는 못하였다.

표 2에서 보는 바와 같이 제안한 방법은 컬러 또는 명도영상보다 깊이영상에 대해 평균 4[dB] 이상의 좋은 결과를 보였다.

RGB영상에 대해 VSRS 3.5a와 비교하면 대부분은 제안한 방법이 약 1[dB] 정도 낮은 값을 보였는데, Mobile에서는 4[dB] 이상의 차이를 보였다. 이것은 대상 영상의 특성에 따른 것으로, Mobile과 같이 단순한 영상에서는 VSRS가 더욱 좋은 결과를 보이는 것으로 사료된다. 전체적으로 대상 시퀀스들은 좌영상과 우영상의 깊이영상 및 RGB영상이 일관되지 않은 양상을 많이 보이고 있는데, 이런 경우 VSRS가 더욱 우수한 화질을 생성하는 것으로 판단된다.

표 3. 중간시점 영상 생성 평균 수행시간
Table. 3 Average execution time of intermediate view image generation

Name	Proposed[sec]	VSRS 3.5a[sec]
Cafe	0.589	4.242
Lovebird 1	0.170	1.640
Poznan street	0.278	2.064
Mobile	0.055	0.344
Average	0.250	2.000

표 3에서는 제안한 방법과 VSRS 3.5a의 수행시간을 비교하였다. 여기서 제안한 방법은 깊이영상과 RGB영상을 모두 생성하는 시간이며, 첫 번째 프레임에서 수행하는 스테레오 정합시간까지 모두 포함한 시간이다. 참고로, 첫 번째 프레임의 수행시간은 다른 프레임의 시간보다 4에서 10배 정도의 시간이 소요된다. 표에서 보는 바와 같이 제안한 방법이 VSRS 3.5a보다 평균 8배 정도 빠른 수행속도를 보여, 표 2에서의 화질저하에 대한 보상으로 간주할 수 있으리라 생각된다.

4.2. 홀로그램 서비스 시스템 프로토타입

그림 1의 프로토타입 시스템은 National Instrument사의 LabView® 환경에서 구축하였으며, 각 기능블록은 이 환경에 맞는 dll(dynamic linking library) 파일을 만들어 사용하였다. 총 6개의 CPU 쓰레드(thread)와 3개의 GPGPU를 사용하여 구현하였는데, 각 재원을 어디에 사용하였는지를 그림 11에 나타내었다.

CPU는 카메라로부터 데이터를 획득하고 카메라 캘리브레이션과 영상보정, 중간시점 영상 생성에 각각 한씩 사용하였다. 또 하나의 CPU는 그림 11에 전처리(Pre-processing)에 사용하는 것으로 되어 있는데, 이것은 그림 2에서 보인 바와 같이 수직리그를 이용하여 깊이영상+RGB영상 획득할 때 두 영상의 광각과 해상도를 맞추는 작업이다. 6개의 CPU중 나머지 세 개는 세 개의 GPGPU 구동에 각각 사용하였으며, 세 GPGPU는 CGH 생성에 두 개, 시뮬레이션에 의한 영상복원에 1개를 각각 사용하였다. GPGPU는 nVidia사의 GPU를 사용하였으며, CUDA로 프로그래밍하였고, GPGPU를 사용하지 않은 기능블록들은 C/C++로 프로그래밍하였다.

그림 11. 프로토타입 시스템의 자원 할당
Fig. 11 Resource assignment to the prototype system

앞에서 설명한 바와 같이 데이터 획득에는 MESA Image사의 SR4000®을 사용하여 깊이영상+명도영상을 획득하거나 수직리그를 사용하여 깊이영상+RGB영상을 획득하여 사용하였는데, 본 연구에서 컬러 홀로그램이 아닌 단색 홀로그램을 대상으로 하기 때문에(그림 5) 주로 SR4000®을 사용하였다. 이 카메라의 해상도는 깊이영상과 명도영상 모두 176×144이다. 두 대의 카메라를 교차신호잡음(crosstalk) 없이 동시에 사용하기 위하여 두 카메라의 변조주파수로 29MHz와 30MHz를 각각 사용하였고, CGH는 1,280×1,024의 해상도(SLM의 해상도)로 생성하였다. 그림 12에 구현한 결과의 최상위 계층 결과를 보이고 있다. 내부 기능블록은 그림 11의 구성도와 동일한 위치에서 1:1 대응된다.

그림 12. 최상위 계층의 구현 결과
Fig. 12 Top-level implementation result

그림 13에 이 시스템을 동작시키기 위해 설계한 GUI를 보이고 있다. 그림의 맨 상단 가운데에 있는 횡으로 표시된 척도는 생성한 중간시점의 시점을 나타낸다. 좌측카메라와 우측 카메라가 -100과 +100에 해당하며,

생성할 중간시점은 -250에서 +250까지 가능하도록 되어 있다. 아래 가운데 제어버튼은 중간시점 생성 방법을 3D 위핑과 변위증분방법 중 선택할 수 있는 버튼이다. 그 아래에 있는 6개의 사진은, 위쪽 3개는 RGB 또는 명도영상, 아래쪽 3개는 깊이영상을 나타내며, 좌우측 각각의 두 영상은 좌우측의 각 카메라로부터 획득한 영상이며, 가운데 두 영상은 생성된 중간시점 영상들이다. 좌측 아래의 회전침은 좌측 카메라로부터 획득된 영상, 중간시점 영상, 우측 카메라로부터 획득된 영상을 각각 CGH 생성에 사용할 수 있도록 하는 제어창이다.

그림 13. 시스템 동작을 위한 GUI
Fig. 13 GUI for system operation

6개의 사진 아래에 작은 두 개의 버튼이 있는데, CGH 생성여부와 복원방법을 선택하는 버튼이다. 그 아래의 두 영상 중 좌측은 생성된 CGH, 우측 시뮬레이션에 의해 복원된 영상이며, 광학장치에 의해 복원된 영상을 촬영한 영상은 맨 오른쪽 아래에 있다. 그림의 맨 우상단에 있는 두 개의 숫자 쌍 표시창은 좌우측 카메라의 변위와 사용 변조주파수를 각각 나타낸다. 그 아래의 min과 max 표시창은 중간시점의 깊이값의 최소와 최대값을 8-비트 스케일로 나타낸다. 우측 버튼은 좌측 카메라와 우측 카메라의 영상을 서로 바꿔주는 기능을 가지고 있다. 아래 작은 표시창들은 홀로그램 생성에 사용된 파라미터들을 나타내고 있는데, CGH의 높이와 너비, 광원의 확대 배율, 광원과 CGH의 화소크기, 사용한 광파의 파장, 물체평면과 홀로그램 평면간의 거리, 복원영상의 밝기 배율, 카메라에 대한 CGH 및 복원영상의 좌우 회전각도를 각각 나타내고 있다.

광학장치에 의해 복원된 영상을 촬영한 영상 위에 4개의 표시창은 복원영상 중 표시되는 부분의 좌우상하 경계를 나타낸다.

표 4. 기능별 프레임당 평균 수행시간
Table. 4 Average execution time per frame for each function

Function	Execution time [ms]
Image acquisition	21
Camera calibration	347
Image rectification	17
Face tracking	3.78
Intermediate view rendering	5.35
CGH generation	219
Image reconstruction by simulation	221

이 GUI 환경에서 각 기능별 평균 수행시간을 표 4에 보였다. 이 표에서 카메라 캘리브레이션은 첫 번째 프레임에서 한 번만 수행하므로 초기 처리시간으로 간주할 수 있고, 따라서 실제 수행시간에는 영향을 미치지 않는다. 표 4에서 중간시점 렌더링 시간이 5.35[ms]로 되어 있고 이것은 표 3의 시간에 비해 매우 낮은 값이다. 이것은 앞에서 언급한 바와 같이 이 실험에 사용된 영상의 해상도가 표 3에 사용된 해상도보다 매우 낮기 때문이다. 또한 표 4의 항목 중 영상획득, 영상보정과 얼굴추적 및 중간시점 렌더링, CGH 생성은 각각 병렬로 수행된다. 또한 시뮬레이션에 의한 영상복원은 수신측에서 수행하는 것으로 본다면 CGH 생성과는 병렬로 수행된다고 간주할 수 있다. 따라서 이 네 동작의 수행시간 중 가장 긴 것에 전체적인 동작속도가 결정된다. 표 4에서 보는 바와 같이 시뮬레이션에 의한 영상복원이 가장 긴 시간을 소요하므로 이 기능을 포함할 경우 프레임당 221[ms]가 소요된다. 그러나 그림 5에서 설명한 바와 같이 광학장치에 의해 영상을 복원할 경우 복원에 소요되는 시간은 디지털 홀로그램을 SLM에 업로드만 하면 되므로, 수행시간은 CGH 생성시간에 좌우된다. 어느 경우이건 현재의 시스템은 초당 약 5프레임 정도를 처리할 수 있는 것으로 나타났다.

V. 결 론

본 논문에서는 시청자의 시점을 추적하여 그 시점에 해당하는 디지털 홀로그램을 서비스하는 인터랙티브 서비스 시스템을 타겟으로 하여, 그 요소 기술인 중간시점 영상 생성 방법과 타겟 시스템의 프로토타입을 제안하고 구현하였다. 제안한 중간시점 영상생성 방법은 깊이영상을 대상으로 두 개의 깊이값에 대한 스테레오 정합을 통해 단위 깊이 당 의사변위증분(pseudo-disparity increment)을 구하고, 이를 사용하여 원하는 가상시점의 화소위치를 결정하는 방법이다. 또한 본 논문에서는 맨 좌측과 맨우측 영상이 주어지고 그 사이의 중간시점 영상을 생성하는 것을 가정하였는데, 이때 발생하는 비폐색영역을 채우는 방법을 제안하였다. 제안한 방법의 성능을 MPEG의 VSRS 3.5a와 비교하였을 때 생성된 영상의 화질은 평균 1dB 정도 낮았으나 수행속도는 8배 정도 빨랐다.

또한 본 논문에서는 제안한 중간시점 영상 생성 방법을 포함한 시청자-인터랙티브 디지털 홀로그램 서비스 시스템의 프로토타입을 구현하였다. 이 시스템은 좌, 우의 깊이영상과 RGB영상 정보 획득, 카메라 캘리브레이션 및 영상보정, 중간시점 영상 생성, CGH 생성 및 홀로그램 영상 복원 기능을 포함하고 있다. 이 시스템은 LabView 환경에서 구현되었으며, CGH 생성과 시뮬레이션에 의한 영상 복원은 GPGPU로, 나머지는 소프트웨어로 구현하였다. 모든 기능은 dll 파일로 만들어 LabView에서 불러 쓸 수 있도록 하였다. 이 프로토타입 시스템의 속도는 가장 많은 시간이 소요되는 CGH 생성 또는 시뮬레이션에 의한 영상 복원에 중속되며, 수행속도는 초당 5 프레임 정도를 처리할 수 있었다.

제안한 중간시점 영상생성 방법은 화질 면에서는 기존의 방법에 뒤지지만, 속도 면에서는 매우 우수한 성능을 보였다. 따라서 이 방법은 고속동작이 필요한 경우에 유용하게 사용될 수 있으리라 생각된다. 또한, 현재 대부분의 가상시점 영상 렌더링 연구에서 그렇듯이, 비폐색영역을 채우는 빠르고 정확한 방법을 개발하는 것이 절실히 필요하다고 판단된다. 본 논문에서 구현한 디지털 홀로그램 서비스 시스템의 프로토타입은, 이런 시스템을 구현하는 것을 아직 연구하고 있지 않기 때문에, 기존

연구와의 비교를 할 수가 없었다. 그러나 디지털 홀로그램 역시 서비스가 궁극적인 목적이므로 향후 이런 연구가 구체적으로 진행될 것으로 예상할 수 있다. 현재 구현된 시스템은 프로토타입이고 또 성능도 실시간에 미치지 못하지만, 향후의 연구에 좋은 참고가 될 것으로 기대된다. 이 시스템에 대해서는 GPGPU로 구현된 CGH 생성 및 영상복원 기능을 전용하드웨어로 구현하거나 고속 알고리즘을 개발하여 시스템 속도를 실시간 동작 속도로 높이는 것이 필요하며, 이것이 본 연구팀의 향후 연구분야이기도 하다.

감사의 글

이 논문은 2013년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2010-0026245).

참고문헌

- [1] T. Motoki, H. Isono, and I. Yuyama, "Present Status of Three-Dimensional Television Research," Proc. IEEE 83(7): 1009-1021(July 1995).
- [2] Edited by O. Schreer, et al., 3D Video Communications, John Wiley & Sons Ltd., Atrium, England, 2005.
- [3] ISO IWA3, Image safety reducing the incidence of undesirable biomedical effects caused by visual image sequences, 2005.
- [4] 3D Consortium, "3DC Safety Guideline for Popularization of Human-friendly 3D," 2006.
- [5] 한국정보통신기술협회 (TTA), "3DTV 방송 안전 가이드라인," TTAK.KO-07.0086, 2010. 12. 23.
- [6] S. A. Benton and V. M. Bove, Jr., Holographic Imaging, John Wiley and Sons Inc., Hoboken NJ, 2008.
- [7] Reichelt, et al., "Holographic 3-D Displays - Electro-

- holography within the Grasp of Commercialization”, A part of the Book named as Advances in Lasers and Electro Optics, INTECH, pp. 683-710, April 2010.
- [8] Y. H. Seo, et al., “Digital hologram compression technique by eliminating spatial correlations based on MCTF,” Optics Communications, Vol. 283, pp. 4261-4270, July 2010.
- [9] H. J. Choi, et al., “Digital watermarking technique for holography interference patterns in a transform domain,” OPTICS and LASERS in ENGINEERING, Vol. 46, No. 4, pp. 343-348, Jan. 2008.
- [10] H. J. Choi, et al., “A Hybrid Encryption Technique for Digital Holography using DCT and DWT,” Intl. J. of KIMICS, Vol. 9, No. 3, pp. 271-275, Aug. 2011.
- [11] T. Shimobaba, T. Ito, “An efficient computational method suitable for hardware of computer-generated hologram with phase computation by addition”, Computer Physics Communications, vol. 138, pp. 44-52, 2001.
- [12] Y.-H. Seo, H.-J. Choi, J.-S. Yoo, and D.-W. Kim, “Cell-based hardware architecture for full-parallel generation algorithm of digital holograms”, Optics Express, Vol. 19, Issue 9, pp. 8750-8761, 2011.
- [13] <http://www.mesa-imaging.ch/prodview-4k.php>
- [14] <http://www.alliedvisiontec.com/emea/home.html>
- [15] Kinect data sheet, http://www.microsoftstore.com/store/msstore/en_US/pd/Kinect-for-Xbox-360/productID.216507400.
- [16] Oliver Schreier, et al., 3D Video Communications, John Wiley & Sons Ltd., Atrium, England, 2005.
- [17] Yoon-Jin Hae, et al., “A Fast Face Detection and Tracking Algorithm by using Depth and Color Information,” KICS Winter Conference, pp.--, Feb. 8th, 2012.
- [18] P. Viola and M. J. Jones, “Robust Real-Time Face Detection,” Computer Vision, Vol. 52, No. 2, pp. 137-154, 2004.
- [19] S. Zinger, et al., “View Interpolation for Medical Images on Autostereoscopic Displays,” IEEE Trans. on Circuits and Systems for Video Technology, Vol. 22, No. 1, pp. 128-137, Jan. 2012.

- [20] Lu Yang, et al., “Artifact Reduction using Reliability Reasoning for Image Generation of FTV,” Elsevier J. Vis. Commun. Images R. 21, pp. 542-560, 2010.
- [21] http://wg11.sc29.org/svn/repos/MPEG-4/test/trunk/3Dview_synthesis/VSRS
- [22] S. Zinger, et al., “Free-Viewpoint Depth Image based Rendering,” Elsevier J. Vis. Commun. Images R. 21, pp. 533-541, 2010.

저자소개

서영호(Young-Ho Seo)

1999년 2월 : 광운대학교 전자재료
공학과 졸업(공학사)
2001년 2월 : 광운대학교
일반대학원 졸업(공학석사)

2004년 8월 : 광운대학교 일반대학원 졸업(공학박사)
2003년 9월 ~ 2004년 6월 : 한국전기연구원 연구원
2005년 9월 ~ 2008년 2월 : 한성대학교 조교수
2008년 3월 ~ 현재 : 광운대학교 교양학부 부교수
※관심분야: 2D/3D 영상 및 비디오 처리, 디지털
홀로그램, SoC 설계, 워터마킹/암호화

이윤혁(Yoon-Hyuk Lee)

2012년 2월 : 광운대학교 전자재료
공학과 졸업(공학사)
2012년 3월 ~ 현재 : 광운대학교
일반대학원 석사과정

※관심분야: SoC 설계, 디지털 홀로그램

구자명(Ja-Myung Koo)

2012년 2월 : 안양대학교 정보통신
공학과 졸업(공학사)
2012년 3월 ~ 현재 : 광운대학교
일반대학원 석사과정

※관심분야: 3D 영상처리, 디지털 홀로그램

김동욱(Dong-Wook Kim)

1983년 2월 : 한양대학교
전자공학과 졸업(공학사)
1985년 2월 : 한양대학교
전자공학과 졸업(공학석사)

1991년 9월 : Georgia공과대학 전기공학과 졸업
(공학박사)

1992년 3월~현재 : 광운대학교 전자재료공학과 정교수

2000년 3월~2001년 12월 : 인티스닷컴(주) 연구원

2006년 3월~현재 : (사)실감미디어산업협회 이사

2009년 3월~현재 : 광운대학교 실감미디어 연구소
소장

※ 관심분야 : 3D 영상처리, 디지털 홀로그램, 디지털
VLSI Testability, VLSI CAD, DSP 하드웨어 설계