

논문 2006-43IE-4-15

웹 기반의 입체 동영상 통신을 위한 3차원 입출력 시스템의 최적화 구현

(Implementation of Optimized 3D Input & Output Systems for
Web-based Real-time 3D Video Communication)

고 정 환*, 이 정 석**, 안 영 환**

(Jung-Hwan Ko, Jung-Suk Lee, and Young-Hwan An)

요 약

본 논문에서는 IEEE 1394 카메라, Intel Xeon 서버 시스템, Microsoft DirectShow 기술을 이용하는 웹 기반의 실시간 입체 동영상 통신을 위한 3차원 입출력 시스템을 제안하고 그 구성요소인 스테레오 카메라 시스템, 3D 디스플레이 시스템, 신호처리 시스템을 최적화하기 위한 조건들을 분석하였다. 즉, 웹 기반의 다시점 3차원 디스플레이 응용에 적합하도록 각 구성요소의 필수 조건에 맞추어 입출력 장치를 선택하고 이들을 조합하여 하나의 3차원 입체영상통신 시스템을 구성하였다. 그리고 제안된 시스템은 Visual C++.Net 환경에서 DirectX 9.1 SDK를 사용하여 최적화 구현하였다. 실험결과 서버 시스템에서 자연스럽게 입체감 있는 다시점의 실시간 3차원 입체 동영상을 관찰할 수 있었고 또한, 클라이언트 시스템에서도 전송된 영상들에 대하여 디스플레이를 통해 3차원 입체 영상을 관찰함으로써 웹 기반의 실시간 원격 3차원 입체 동영상 통신 시스템의 실질적인 응용 가능성을 제시하였다.

Abstract

In this paper, 3D input and output systems for a web-based real-time 3D video communication system using IEEE 1394 digital cameras, Intel Xeon Server system and Microsoft Directshow library is proposed. And some conditions for optimizing the operations of the stereo camera, 3D display and signal processing system are analyzed. Input & output systems are carefully selected, which can satisfy the required optimization conditions and the final 3D video communication system is implemented by using three optimized devices. The overall control system is developed with Microsoft Visual C++.Net and Microsoft DirectX 9.1 SDK. Some experimental results show that the observer can feel the natural presence from multi-view(4-view) 3D video of server system in real-time and also can feel the natural presence from 3D video of client system and finally suggest an application possibility of the proposed web-based real-time 3D video communication in real fields.

Keywords : 3D communication, disparity estimation, IVS, reconstruction

I. 서 론

최근, 보다 자연스럽게 현실감 있는 비전 시스템의 요구가 증가함에 따라 3차원 영상 디스플레이 및 전송

기술에 대한 많은 연구들이 진행되고 있다^[1-3]. 특히, 심해저, 우주 공간 및 고 방사선 구역 등의 극한 작업 환경이나 원격 화상회의 시스템, 원격 의료서비스, 무인감시 시스템, 공장 자동화, 지능형 교통 시스템 등 다양한 산업 분야에서 기존의 2차원 영상시스템의 한계를 넘어 설 수 있는 3차원 영상 통신시스템에 대한 필요성이 날로 증가되고 있으며, 일부는 이미 실용화가 추진되고 있는 실정이다.

특히 원격지 상호간의 동일 시간, 동일 장소에서 회

* 정희원, (주) 인코시스 기술연구소

(R&D Department, INCOSYS Co., Ltd.)

** 정희원, 인하공업전문대학 메카트로닉스과

(Department of Mechatronics, Inha Technical College)

접수일자: 2006년6월9일, 수정완료일: 2006년12월4일

의하는 것과 같은 효과를 주는 원격 입체 화상회의 시스템의 경우, 보다 현실감 있고 임장감 있는 회의 환경을 조성할 수 있다는 측면에서 그 관심이 더욱 고조되고 있으며, 이에 대한 활발한 연구가 진행되고 있다. 이러한 원격 3차원 화상 통신시스템을 구축하기 위해서는 다시점의 3차원 영상 획득 및 정확한 실시간 전송, 그리고 이를 현장감 있게 표현해 줄 수 있는 디스플레이에 이르기까지 고 난이도의 방법과 기술들이 요구된다. 그러나 현재 보급 중인 원격 화상 통신 시스템의 경우, 대형 디스플레이, 영상/음성 처리 장치, 미팅룸 구성, 멀티캐스트 지원 및 네트워크와 같은 확장성 있는 하드웨어 시스템의 요구와 이를 뒷받침하기 위해 상호 호환될 소프트웨어 개발의 한계가 있으며, 영상 전송에 있어서도 Vic에 근거한 구형 압축방식에 의존하여 300kbps 수준의 2차원 영상만을 전송할 수 있는 문제점을 안고 있다. 또한, 기존의 스테레오 영상의 화상신호에 기반을 두어 개발된 대부분의 3차원 화상 전송기술은 화질과 속도, 전용장비 및 회선관리에 따른 네트워크 운영에 대한 추가적인 비용 등에 능동적으로 대처하는데 많은 어려움이 있다. 더욱이, 현재의 전송기술로서 필요로 하는 대역폭을 3차원 화상전송의 가능성에 대해 고려해 볼 때, 전송매체 및 설비 등의 문제점이 나타나게 되며, 지상에 있어서 무선전송 설비에 대한 주파수 할당이 진행되고 있고 3차원 화상전송을 위한 전용 통신로의 확보가 어려운 문제점이 나타나게 된다^[2-3]. 또한, 기존의 3D 입체 화상 시스템은 다안식 입체 카메라를 통해 다시점의 영상획득과 디스플레이가 이루어지기 때문에 시점의 개수가 증가함에 따라 시역이 확대되고 보다 자연스러운 3D 디스플레이가 가능하지만^[4-9], 시점의 개수가 증가됨에 따라 영상 데이터도 동시에 증가하기 때문에 실시간적 영상처리 프로세서 및 초고속, 광대역의 전송 채널이 또한 요구되게 된다^[2-3].

따라서 본 논문에서는 웹 기반의 실시간 3차원 디스플레이 응용을 위한 카메라 및 프로세싱 시스템의 최적화에 관한 연구를 통하여 현재의 웹 기반 시스템에 3차원 시각 시스템을 적용하기 위한 각 부분별 시스템의 필수 조건들을 분석하고 이런 조건들을 만족할 수 있는 시스템을 구성하여 제안하고자 한다. 이 시스템은 웹 기반의 3차원 응용에 적합하도록 각 구성요소의 필수 조건에 맞추어 장치를 선택하고 이들을 조합하여 하나의 3차원 디스플레이 시스템을 구성한다. 또 이 다시점 3차원 영상 시스템을 최적의 상태로 제어하고 응용할 수 있도록 Visual C++.Net을 기반으로 DirectX 9.1

SDK를 사용하여 직접 제어프로그램을 구현한다. 제안된 시스템에서 서버 시스템을 통하여 자연스럽게 실재감 있는 실시간 3차원 영상을 관찰할 수 있었고 클라이언트 시스템에서도 전송된 영상들에 대하여 3차원 영상을 관찰할 수 있었다. 이를 통하여 실시간 웹 기반의 원격 3차원 통신 시스템의 실질적인 응용 가능성을 제시하고자 한다.

II. 실시간 3차원 영상 획득 및 디스플레이 시스템의 구현

본 논문에서 제안한 웹 기반의 실시간 3차원 디스플레이 시스템은 그림 1과 같이 크게 3 부분으로 이루어진다. 즉, 최적화된 시차 검출을 위한 스테레오 카메라 시스템, 스테레오 영상의 데이터를 빠른 속도로 연산하고 웹 전송을 지원하기 위한 서버 시스템 그리고 웹을 통하여 좌영상과 시차 영상을 전송받아 IVS (intermediate view synthesis) 기반의 다시점 3차원 영상 복원 및 실시간적인 포맷에 맞추어 디스플레이하기 위한 클라이언트 시스템으로 나누어지며^[10-11], 서버 시스템과 클라이언트 시스템에서는 획득된 다시점 영상을 3차원 디스플레이 장치에 실시간으로 디스플레이가 가능하다.

즉, 스테레오 카메라로 부터 스테레오 영상이 서버 컴퓨터 시스템으로 획득된다. 획득된 스테레오 영상은 시차 추정기법을 통하여 최적화된 시차를 검출한 후 좌영상과 함께 웹을 통하여 클라이언트 컴퓨터 시스템에 전송되며, 전송된 좌 영상과 시차 데이터는 IVS 기법을 이용하여 다시점 3차원 영상으로 복원하여 실시간으로 3차원 디스플레이 장치에 자연스러운 입체화상을 디스플레이 한다.

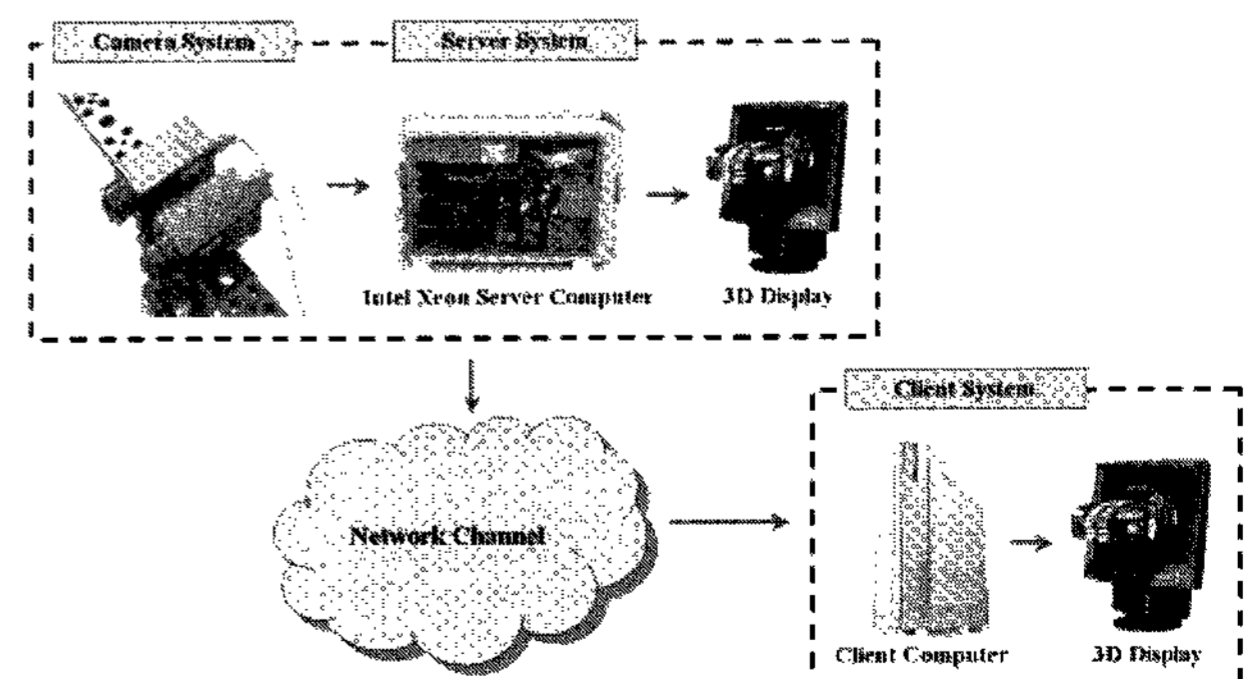


그림 1. 웹기반 실시간 3차원 디스플레이 시스템의 흐름도

Fig. 1. Flow diagram of real-time web-based 3D display system.

1. 스테레오 카메라 시스템

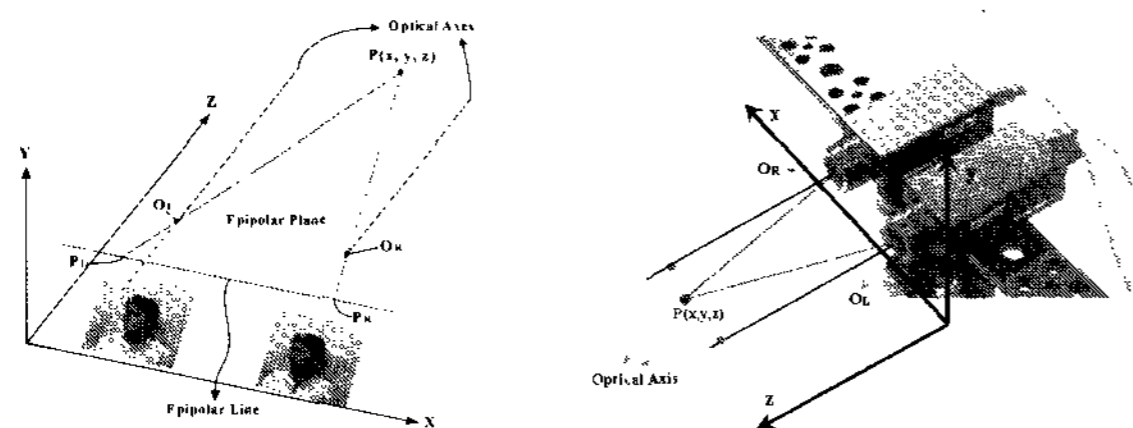
본 논문에서 제안된 스테레오 카메라 시스템은 최적의 3차원 영상 획득을 위한 중요한 전처리 과정으로 실감 3차원 디스플레이를 위해 세밀하고 정교한 시차 검출이 요구된다. 3차원 영상 시스템에서 좌영상과 우영상을 정합하는 대응점 일치 문제는 스테레오 영상 시스템에서 가장 근본적인 문제로, 본 논문에서는 에피폴라 기하학(Epipolar Geometry)^[7]을 기반으로 구현하였다.

그림 2는 대응되는 한 점을 기준으로 투영된 스테레오 영상간의 에피폴라 기하학과 이를 뒷받침하기 위해 렌즈의 축이 서로 병렬로 놓인 상태로 구성된 다시점 카메라의 구조적인 관계를 각각 나타낸 것이다.

그림 2(a)에서 공간상의 임의의 점 P는 좌, 우 카메라 렌즈의 중심 O_L, O_R 을 통해 좌, 우 영상 평면인 P_L, P_R 로 투영된다. 따라서 그림 2(a)와 같은 평행식 구조의 스테레오 영상에서는 에피폴라 선이 두 영상에 대해 평행함으로 변이벡터가 항상 수평 값만 갖게 되기 때문에 P_L 에 해당하는 대응점 P_R 을 찾기 위해서는 에피폴라 선 상에서 1차원적으로 탐색하게 된다^[4]. 스테레오 3차원 디스플레이 시스템에서 관찰자는 스테레오 영상의 시차 정보로부터 3차원 입체를 인식하기 때문에 카메라 시스템의 구조는 매우 중요하다.

스테레오 카메라 시스템에서 획득된 스테레오 영상은 IEEE 1394a PCI 인터페이스 장치를 통하여 서버 시스템으로 전송된다. 그림 3은 스테레오 영상이 서버 컴퓨터 시스템으로 전송되는 과정을 보여준 것이며, 1394 어댑터는 3개의 1394a 포트와 1개의 PCI 1.1 슬롯을 가진 인터페이스 장치로 IEEE 1394와 PCI 1.1간의 데이터 통신을 위한 인터페이스 장치이다.

2개의 1394 어댑터를 사용하였으며 1개당 카메라 1대가 연결되어 있다. 또 두개의 1394 어댑터는 서버 컴퓨터의 PCI 1.1 버스에 설치하였으며, 스테레오 카메라 입



(a) 에피폴라 기하학 (b) 스테레오 카메라 시스템
(a) Stereo camera system (b) Stereo camera system

그림 2. 에피폴라 기하학과 다시점 카메라 시스템

Fig. 2. Multi-view camera system based on Epipolar geometry.

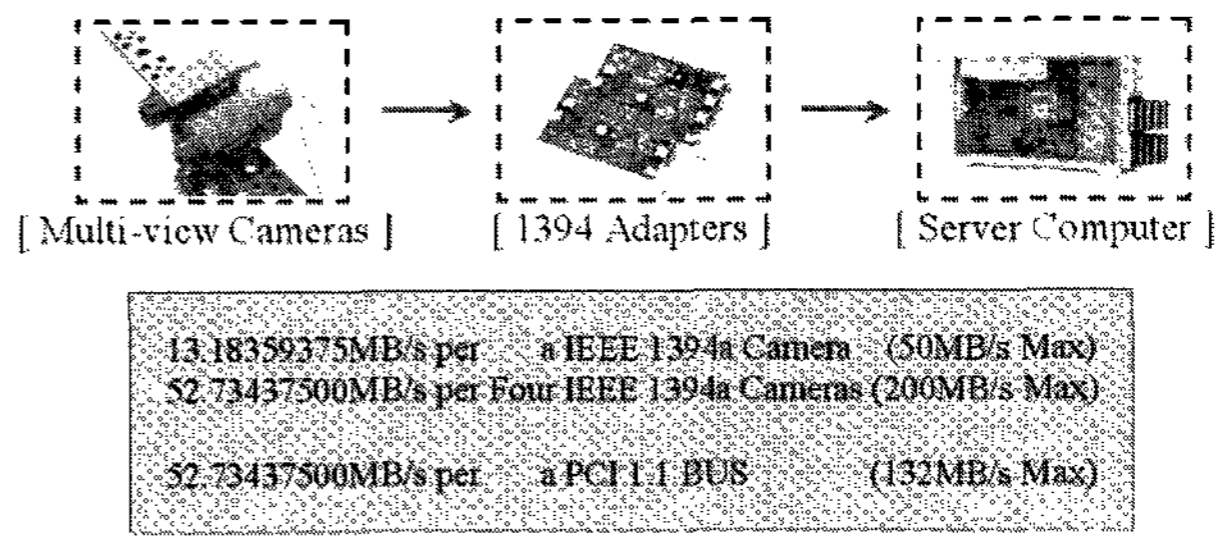


그림 3. 다시점 영상의 서버시스템 전송

Fig. 3. Transmission to server system of multi-view image data.

력의 최대 대역폭은 132MB/s 이다. 제안된 카메라 시스템은 VGA급의 RGB24 영상을 각각의 카메라에서 초당 15프레임 획득한다. 한대의 카메라가 발생시키는 데이터 량은 $13.18359375MB/s (=640(\text{Width}) \times 480(\text{Height}) \times 24(\text{bits}) \times 15(\text{fps}))$ 으로 IEEE 1394a 규격이 50MB/s (=400Mbps)를 지원하므로 전송 대역폭이 충분하다.

또한 스테레오 카메라 시스템 전체가 구동될 경우 $26.3671875MB/s (=13.18359375MB \times 2 \text{ Cameras})$ 의 대역폭이 발생하는데 이 데이터 량도 2개의 IEEE 1394a 포트를 통해 서버 시스템의 PCI 1.1 버스로 이동하기에 충분한 대역폭이다. 이론상의 대역폭이 26.3671875MB/s으로 PCI 1.1(132MB/s)에서 전송하기에 충분하지만 이것은 서버시스템으로의 전송 자체에만 쓰이는 데이터 량이므로 매우 방대하다.

2. 실시간 다시점 영상 처리를 위한 서버 시스템

스테레오 카메라 시스템에서 실시간 영상획득과 알고리즘 프로세싱을 위해서는 방대한 데이터 량이 발생한다. 이러한 데이터 량을 실시간 구동하기 위해서는 컴퓨터 시스템의 성능이 매우 중요하며 여기에서 CPU와 메모리, 그리고 데이터 버스의 전송 대역폭이 성능 결정의 주된 요인으로 작용하게 된다.

그림 4는 방대한 데이터 량을 실시간적으로 전송하

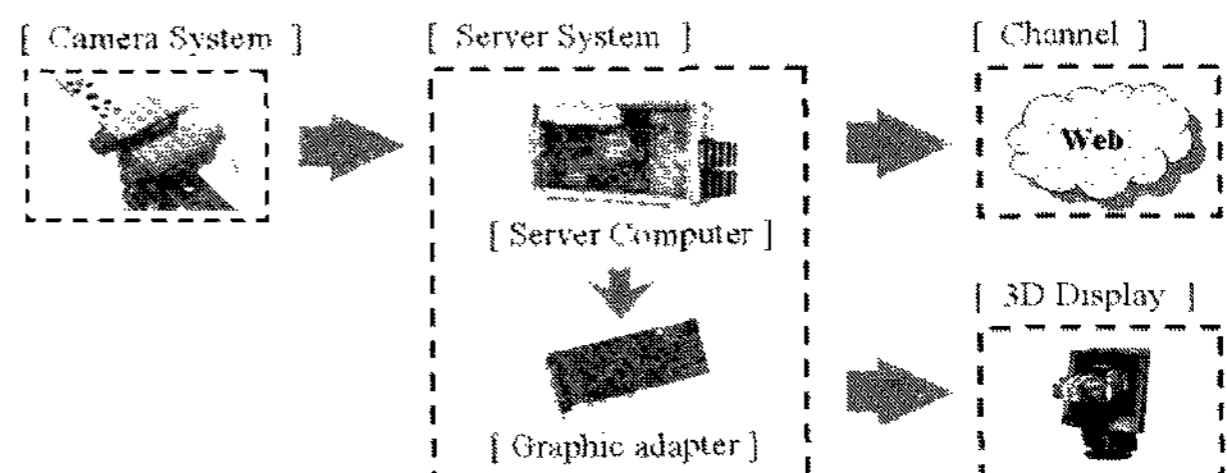


그림 4. 고속 스테레오 영상처리를 위해 구현된 서버시스템

Fig. 4. Server system for high speed stereo image processing.

기 위해 사용된 서버 시스템으로 인텔사의 제온(Xeon) 서버시스템을 이용한 고속 스테레오 영상처리를 위해 구현된 서버시스템을 나타낸 것이다.

즉, 방대한 데이터의 고속처리를 위해 Intel사에서 개발된 제온 프로세서 2.4GHz를 듀얼(dual) 모드, 두개의 독립적인 512M DDR 메모리와 SCSI Ultra320 HDD를 서버보드 SE7501HGZ와 구성하였으며, 특히 Dual 모드 2.4GHz의 Xeon CPU는 일반적인 사용자용의 CPU가 아닌 듀얼 모드 전용의 CPU로 방대한 양의 데이터를 고속으로 처리할 수 있도록 특별히 설계된 CPU로 다시점 3차원 영상처리에 매우 적합하다. 두개의 512MB의 DDR 메모리는 각각의 CPU에 일대일로 연결되어 메모리가 서로 종속적이지 않도록 구성되어 시스템의 안정적으로 동작할 수 있도록 해준다. 서버보드로 사용된 SE7501HGZ의 경우, 32 bit/33 MHz PCI(132MB/s), 2개의 64 bit/100 MHz PCI-X(800MB/s), 1개의 64 bit/133 MHz PCI-X(1056MB/s)와 같이 3개의 독립적인 PCI 버스를 가진다. 이 3가지의 버스는 서로 독립적이므로 서로의 데이터 전송에 영향을 주지 않는다. 이 점은 방대한 양의 데이터를 처리하는 다시점 3차원 영상처리의 특성에 매우 적합하다. 또한 PCI-X 버스의 경우 현재 일반적으로 사용되는 PCI 1.1(32 bit/33 MHz, 132MB/s) 규격의 버스에 비교하여 매우 큰 용량의 데이터를 전송할 수 있다. 그러나 현재 일반적으로 사용되는 PCI 장치가 PCI 1.1의 속도만을 지원하고 PCI-X 버스에 대하여 하위호환을 지원하는 수준이므로 PCI-X의 전송용량을 현재로서는 사용할 수 없으나 서버보드 SE7501HGZ의 독립적이고 큰 대역폭의 버스 특성은 향후 3차원 영상처리에 매우 중요한 자원이다.

한편, 본 논문에서는 데이터양의 축소와 함께 보다 효율적인 실시간 영상 전송을 위해 평행식 스테레오 카메라를 통해 검출된 시차지도와 좌 영상만을 전송하게 된다.

식 (1)은 스테레오 영상의 시차를 검출하기 위해 사용된 MSE 함수를 나타낸 것이며, 식 (1)에서 N_x 와 N_y 는 블록의 크기를, I_L 과 I_R 은 좌 영상 블록과 이에 대응하는 우 영상 블록을 각각 나타낸 것이다.

$$MSE = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |I_L(i, j) - I_R(i + d, j + l)|^2 \quad (1)$$

일반적으로, 인간시각 시스템에서는 수직시차는 수평시차에 비해 그 영향이 적기 때문에^[2] 이러한 인간 시

각 시스템의 특성을 고려하여 2.1에서 소개된 바와 같이 스테레오 카메라를 병렬로 구성하게 되면 수평시차만이 존재하게 된다. 이 경우 우영상에 대한 좌영상의 시차는 식 (2)와 같이 표현된다. 한 블록과 정합(matching)되는 가장 작은 MSE값을 가지는 영역이 선택되면 좌, 우 영상의 최적화된 시차벡터가 식 (2)에 의해 결정된다.

$$d_{ij} = \arg \min \frac{1}{N_x N_y} \sum_{(i, j) \in \beta_i} |I_L(i, j) - I_R(i + d, j)|^2 \quad (2)$$

식 (2)에서 d_{ij} 는 변이벡터의 검색영역을 의미하며, β_i 는 $N_x \times N_y$ 크기의 I_L 의 i - 번째 블록이다. 식 (2)에서 검출된 시차벡터는 블록의 대표 값이 되고 이 결과는 IVS 기반의 다시점 영상 생성을 위한 시차지도를 표현하는데 사용된다.

3. 실시간 네트워크 전송 및 클라이언트 시스템

그림 5는 네트워크 전송과 지연 요인을 나타내는 개념도로써, 이전단계에서 처리된 좌 영상 및 시차는 네트워크를 통하여 매회 압축하여 네트워크를 통하여 클라이언트 시스템으로 전송한다.

클라이언트 시스템은 네트워크로부터 좌 영상 및 시차정보를 전송받아 IVS 기반의 다시점의 3차원 영상을 복원한 후, 3차원 디스플레이 장치를 통해 디스플레이 한다. 기본적으로 컴퓨터 기반 시스템이기 때문에 서버 시스템과 같은 디스플레이 방식과 같은 인터페이스와 확장성을 지닌다. 그러나 네트워크로부터 영상을 전송받아 디스플레이 해야 하기 때문에 클라이언트 시스템

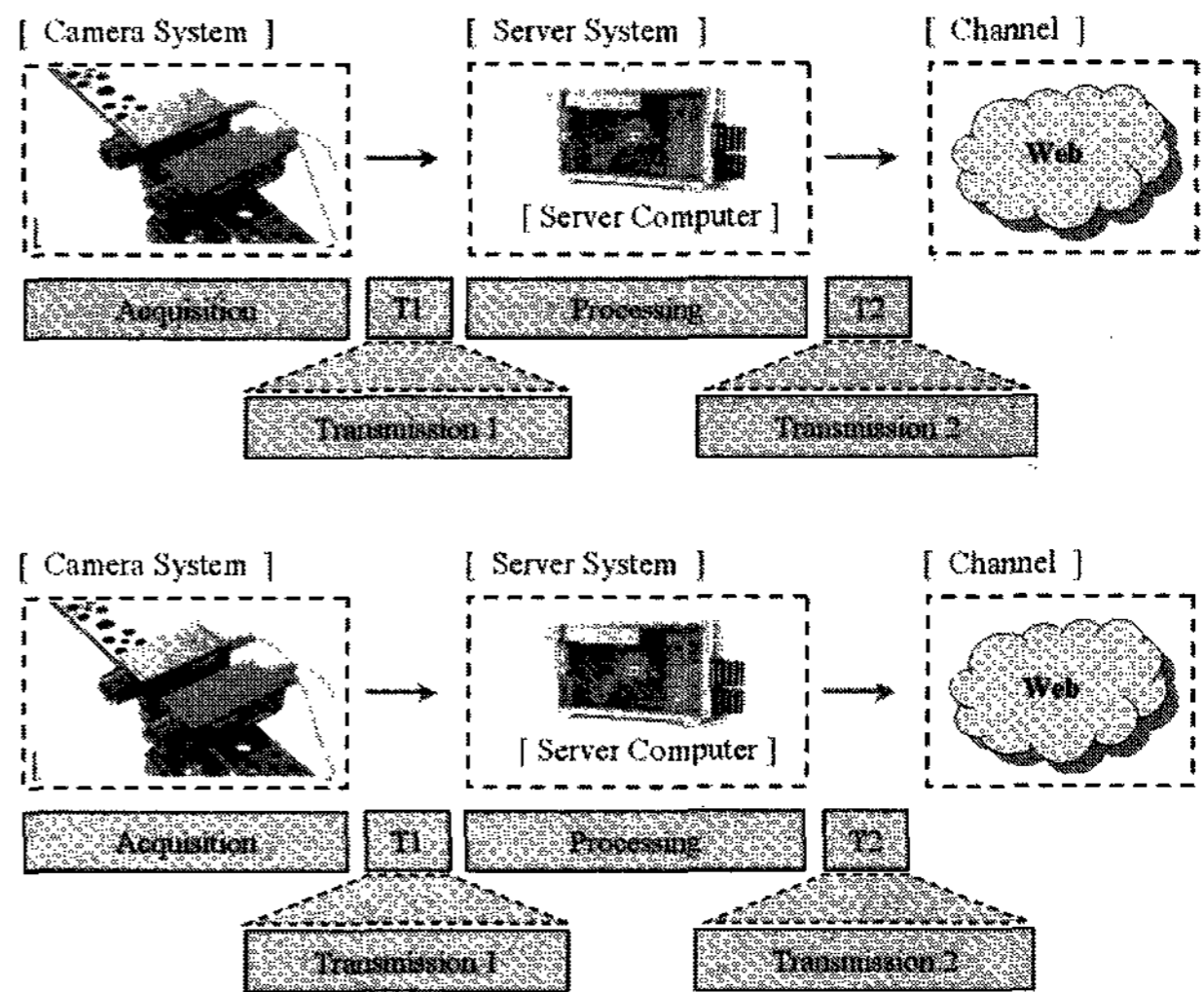


그림 5. 서버 시스템의 전송지연 요인
Fig. 5. Time delay of server system.

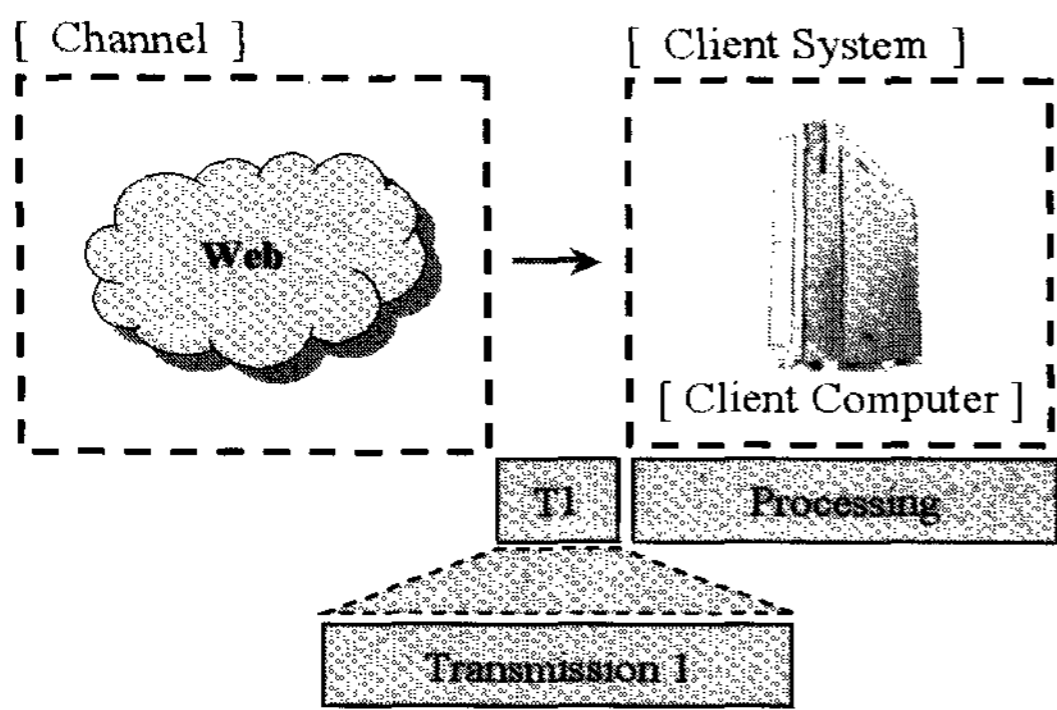


그림 6. 클라이언트 시스템의 디스플레이 지연 요인
Fig. 6. Time delay of client system.

의 3차원 디스플레이 성능은 네트워크의 대역폭에 종속적이다.

그림 6은 클라이언트 시스템의 디스플레이 지연 요인에 대한 개념을 보여준 것으로, 서버 시스템으로부터 전송된 좌 영상과 시차정보는 클라이언트 시스템에 도착하기 까지 전송 채널 특성의 영향을 받게 된다.

클라이언트 시스템에서 한 번에 받는 패킷의 크기는 서버 시스템이 보낸 패킷의 크기와 같다고 보장할 수 없다. 서버는 특정 크기의 패킷을 한 번에 보내지만 클라이언트는 서버가 보낸 패킷의 크기를 부분으로 잘라서 받을 수도 있고 한 번에 받을 수도 있다. 일반적으로 채팅 프로그램과 같은 데이터 크기가 작은 단위의 패킷은 한번의 송수신으로 전송이 완료되지만 영상 데이터 같은 대용량의 데이터는 네트워크의 특성과 전송 당시의 채널상태에 따라서 수신되는 패킷 크기가 변동되므로 수신시 패킷의 전송완료 여부를 확인하여야 한다. 또 실시간 응용을 기반으로 하고 있으므로 수신시 일정 시간을 기다린 후 완료되지 않았다면 수신한 영역만을 적용해야 한다. 또한, 클라이언트 시스템에서 영상 수신시 버퍼를 두어 늦게 도착하는 데이터에 대하여 버퍼링을 하여 데이터의 손실을 줄이는 방법도 있다. 이러한 버퍼링 방법은 실시간 3차원 입체 방송 시스템에 효과적인 성능을 기대할 수 있다.

그림 7은 버퍼링이 적용된 클라이언트 시스템 3차원 디스플레이의 개념도를 보여준 것으로, 버퍼링 사이즈가 크다면 그만큼의 손실 가능성을 줄이고 대역폭이 지원되는 최대의 성능에 접근할 수 있다.

그러나 버퍼링의 사이즈가 크다는 것은 그만큼 디스플레이를 위한 지연시간이 크다는 것과 동일하므로 실시간 화상회의와 같은 응용에는 크기가 큰 버퍼링의 적용은 적합하지 않다. 제안된 실시간 3차원 디스플레이 시스템에서 전체 지연 시간은 서버 시스템에서 카메라

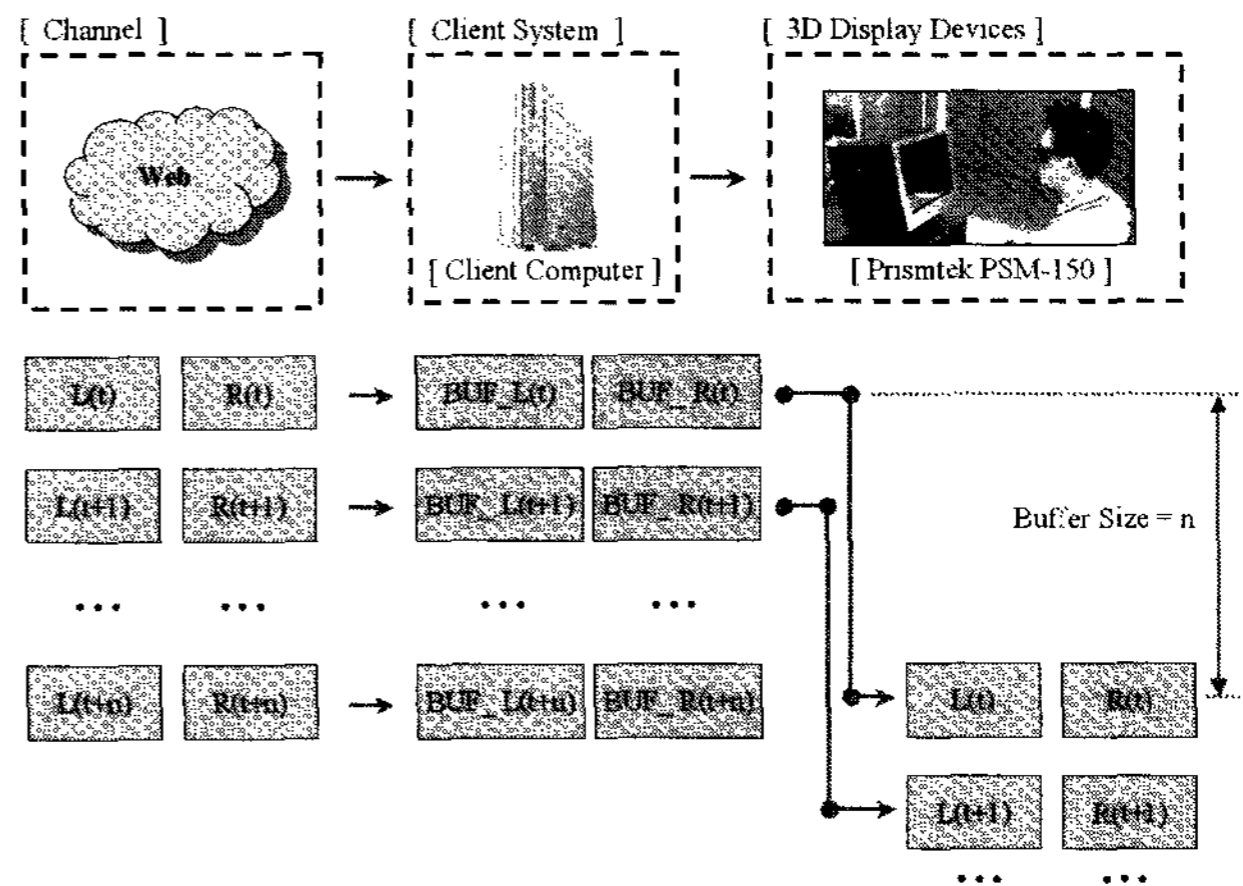


그림 7. 버퍼링이 적용된 클라이언트 시스템 3차원 디스플레이
Fig. 7. Buffering-based client system for 3D display.

에서 영상을 획득하여 압축한 다음 전송하여 클라이언트에서 재 정돈 되어 화면상에 디스플레이 되기까지의 시간이다.

4. IVS 기법을 이용한 실시간 다시점 3차원 입체 디스플레이

서버 시스템으로 부터 전송되어온 좌 영상과 시차정보는 접속된 클라이언트 시스템의 IVS 기법에 따라 식 (3)과 같이 수평방향의 시차와 좌우 영상간의 관계를 통해 우 영상을 복원된 뒤, 연쇄적인 중간시점의 영상을 생성함으로써 보다 자연스러운 다시점 3D 영상 통신이 이루어지게 된다^[10-11].

$$I_R = \begin{bmatrix} i_R \\ j_R \end{bmatrix} = \begin{bmatrix} i_L + \hat{d}(i_L, j_L) \\ j_L \end{bmatrix} = I_L + \begin{bmatrix} \hat{d}(i_L, j_L) \\ 0 \end{bmatrix} \quad (3)$$

식 (3)에서 I_L 과 \hat{d}_{ij} 는 전송된 좌 영상과 탐색범위 내의 시차값을 각각 나타낸 것이며, I_R 은 복원된 우 영상을 나타낸 것이다.

그림 8은 본 논문에서 제안한 다시점 입체 화상 통신 시스템을 구현하기 위해 이용된 IVS 기법의 개념도를 나타낸 것이다.

즉, 식 (4)와 같이 시차 정보의 시점의 위치 α 에 따라 가중 평균값으로 내삽하는 경우를 나타낸 것으로, I_P 는 좌, 우 영상의 시차가 교차되는 지점에서 생성된 중간시점 영상의 위치를 각각 나타낸 것이다^[10-11].

$$I_P(i, j) = (1 - \alpha) \cdot I_r(i - \hat{d}_{ij}(i, j), j) + \alpha \cdot I_l(i - \hat{d}_{ij}(i, j), j) \quad (4)$$

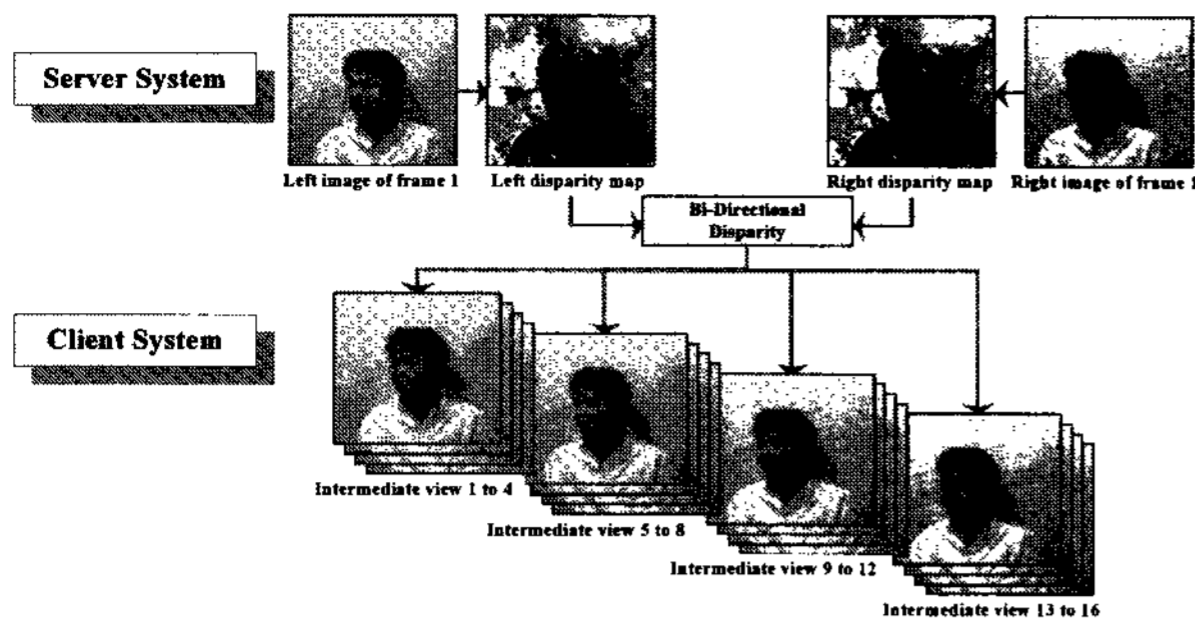


그림 8. 서버 시스템의 실시간 다시점 영상 복원의 개념도
 Fig. 8. Flow diagram of server system for real-time multi-view image reconstruction.

일반적으로, 중간시점의 영상을 생성할 때 스테레오 영상 쌍의 한쪽 영상에서는 서로 가려진 영역(occluded region)이 존재하게 된다. 이 때, 이러한 가려진 영역에는 변이가 할당되지 못한 영역이 존재할 수 있기 때문에 이러한 영역에 대해서는 시차 평활화 과정을 통하여 주변 변이 값들의 평균값을 구하여 그 변이 값으로 대체하게 된다^[11].

한편, IVS 기법을 통해 생성된 다시점 영상들은 3차원 디스플레이 장치를 통하여 관측자에게 입체영상으로 보이게 된다. 3차원 영상 디스플레이에서는 스테레오 영상의 동기설정이 매우 중요하며, 좌 영상과 우 영상의 동기가 일치되지 않은 채 관측자에게 보일 경우 왜곡된 시점의 영상이 관측자에게 보여 지게 된다. 따라서 좌 영상과 우영상이 모두 획득된 후 두 영상이 동시에 디스플레이 되어야 하며, 서버 시스템을 통해 디스플레이 장치의 출력 포맷에 따라 영상의 출력 형식을 바꾸어 주어야 한다. 제안된 시스템에서 사용된 디스플레이 어댑터는 Enseco 사의 MVP Pro ZX로 4개의 RGB 15핀 D-SUB 포트를 지원한다. 이 장치를 통하여 생성된 다시점 영상을 4대의 별개의 디스플레이 장치로 전송할 수 있다. 이 디스플레이 어댑터는 현대의 컴퓨터 시스템에 최대 4개의 장치를 동시에 설치할 수 있으며 이는 현대의 컴퓨터 시스템에서 최대 16대의 디스플레이 장치를 지원할 수 있다는 것을 의미한다^[13-14]. 제안된 시스템은 컴퓨터 기반의 시스템으로 일반적인 모니터 디스플레이에 쓰이는 아날로그 RGB 15핀 D-SUB 단자를 사용하였다. 사용된 디스플레이 어댑터가 4포트의 출력을 지원하고 영상의 형식은 디스플레이 장치에 따라 유연하게 변형이 가능하므로 다시점 3차원 디스플레이는 물론 스테레오 안경식 또는 무안경식 디스플레이를 지원할 수 있다.

III. 실험 및 결과 분석

본 논문에서 제안한 웹 기반의 실시간 3차원 디스플레이 응용을 위한 카메라 및 프로세싱 시스템은 그림 9와 같이 사용하였으며, 카메라 시스템은 VGA급의 24비트 컬러영상을 지원하는 Unibrain 사의 Firei-400을 사용하였으며, 전체 실험과정에서 Microsoft DirectX 9.0 SDK와 Microsoft Visual C++.Net을 사용하여 직접 구현한 응용 프로그램 사용하여 다시점 디지털 카메라 시스템의 제어, 획득된 다시점 영상의 압축 및 전송 프로세싱, 다시점 영상의 디스플레이에 이르는 모든 과정을 제어하였다.

Microsoft DirectX SDK는 윈도우즈 시스템에서 멀티미디어를 지원하기 위한 개발 키트로 멀티미디어 장치의 구동, 데이터 획득, 압축, 해제, 디스플레이 등을 지원한다. Microsoft Visual C++.Net은 카메라 기반의 멀티미디어 프로그램을 개발하기 위하여 Microsoft DirectX 9.1 SDK와 함께 사용되었다. 전송방식으로 IEEE 1394a를 사용하였으며, 실험에 사용될 수 있는 컬러형식과 데이터양을 표 1에 나타내었다.

표 2는 다시점 카메라 시스템의 환경을 일치 시키기 위하여 사용된 카메라의 변수이며, 다시점 카메라 시스

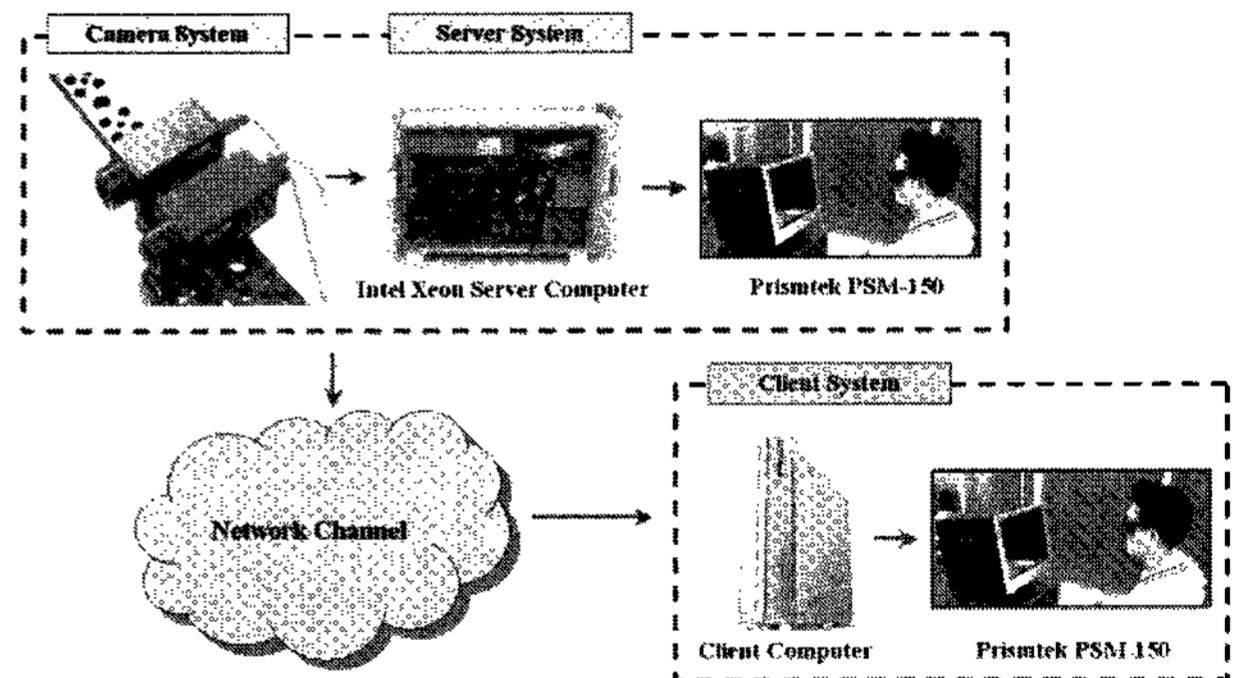


그림 9. 웹 기반 실시간 3차원 디스플레이 시스템의 흐름도
 Fig. 9. Data flow of the real-time web-based 3D display system.

표 1. 시스템에 적용 가능한 컬러형식과 데이터량
 Table 1. Amount of image data by multi-view camera system.

640(W) × 480(H) × 24(RGB24) × 15(fps) × 4(Cameras)	= 52.734375MB/s
640(W) × 480(H) × 16(YUYV) × 15(fps) × 4(Cameras)	= 31.5625MB/s
640(W) × 480(H) × 12(YUV411) × 30(fps) × 4(Cameras)	= 52.734375MB/s

표 2. 실험에 사용된 스테레오 카메라의 변수값
Table 2. Parameters of multi-view camera system.

대비 (contrast)	색조 (hue)	선명도 (sharpness)	감마 (gamma)
142	66	20	1
후광보정 (backlight compensation)	화이트 밸런스 (white balance)	밝기 (brightness)	노출 (exposure)
0	96	auto	auto



(a) 좌 영상 (b) 우 영상 (c) 시차영상
(a) Left image (b) Right image (c) Disparity image

그림 10. 획득된 1st 프레임의 2시점 원본 영상과 시차 영상

Fig. 10. Captured 2-view original image and disparity image of 1st frame.

템에서 획득된 영상을 보다 이상적으로 획득하기 위한 기본 조건이다.

본 실험에서는 에피폴라 기하학의 구속조건에 따라 설정된 스테레오 카메라를 이용하여 실제 촬영된 640 X 480 해상도의 스테레오 영상을 사용하였다. 즉, 그림 10 과 같이 배경에 인물의 그림자(silhouette)와 함께 측면으로 다른 물체가 존재하는 영상과 이를 이용한 검출된 시차 영상을 각각 나타낸 것이다.

그림 10과 같이 획득된 좌 영상과 시차영상은 JPEG 형식으로 압축되어 클라이언트 시스템으로 전송하였다. JPEG 압축 기법은 MPEG4나 H.264와 같은 알고리즘보다 큰 대역폭이 필요하지만 본 실험의 목적인 웹기반의 3차원 응용을 위한 분석과 그 가능성을 제안하는 것에 충분하다. 실험에서 쓰인 JPEG 압축 기법에 의하여 320 x 240급의 RGB영상은 평균 13kB, 640 x 480급의 RGB영상은 평균 34kB의 압축 크기를 나타내었다.

표 3은 시스템 설정에 따른 전송에 필요한 대역폭을 나타낸 것이다.

현재 근거리 Ethernet 기반의 근거리 네트워크 시스템(LAN)은 100Mbps로 12.5MB/s의 대역폭을 가진다. 전송 대역폭 만으로 볼 때 640x480급의 고해상도 다시점 영상도 근거리 네트워크 내에서 다시점 전송/ 디스플레이가 가능하다. 그러나 서버 시스템에서 다시점 카

표 3. 시스템 설정에 따른 전송에 필요한 채널 대역폭

Table 3. Required bandwidths of channels for experiments

해상도	평균압축	평균 FPS	시점	전송용량
320 x 240	13kB	15	2	390kB
			4	780kB
		30	2	780kB
			4	1560kB
640 x 480	34kB	15	2	1020kB
			4	2040kB
		30	2	2040kB
			4	4080kB

표 4. 표준 압축 기법의 시점수에 따른 전송 대역폭
Table 4. Required bandwidths of channels for applications.

압축 표준	단위 대역폭	시점수	필요 대역폭
MPEG2 SD	4~8 Mbps	2	8 ~ 16 Mbps
		4	16 ~ 32 Mbps
MPEG2 HD	25 Mbps	2	50 Mbps
		4	100 Mbps
MPEG4	64 kbps~2 Mbps	2	128 kbps ~ 4 Mbps
		4	256 kbps ~ 8 Mbps
		6	384 kbps ~ 12 Mbps
		8	512 kbps ~ 16 Mbps

메라 시스템으로부터 다시점 영상을 받아 압축하여 전송하여 클라이언트 시스템에서 재분배, 디스플레이하기까지의 여러 가지의 지연시간이 존재한다.

표 4는 표준 압축 기법의 시점수에 따른 전송 대역폭을 나타낸 것으로, 본 논문에서는 상대적으로 구현이 간단한 JPEG에 의한 영상 압축을 사용하여 실험하였으며 H.264나 MPEG4 기법에 의하여 3차원 다시점 영상을 전송할 경우 보다 고성능을 기대할 수 있다.

그림 11과 12는 전체 30 frame의 전송된 입력 영상 중 5th 프레임과 10th 프레임의 영상에 대해 원본 좌 영상과 식 (3), (4)에 의해 생성된 중간시점의 영상들 중 5번째 시점(N=5)과 10번째 시점(N=10)의 중간영상과 복원된 우 영상을 각각 나타낸 것이다. 그림 11과 12에서 영상속의 사람이 보이는 시점이 각 영상마다 조금씩 다르며 시각적으로 보기에 별다른 오정합 없이 성공적으로 합성되었음을 알 수 있다.

표 5는 본 논문에서 제안한 3D 화상 전송 시스템의 시점별 중간영상 합성 소요시간을 나타낸 것이다. 표 5에서, 중간영상 획득시 발생하는 지연시간이 4시점의 결과에서 초당 15프레임에 해당하는 66ms에 근접하기

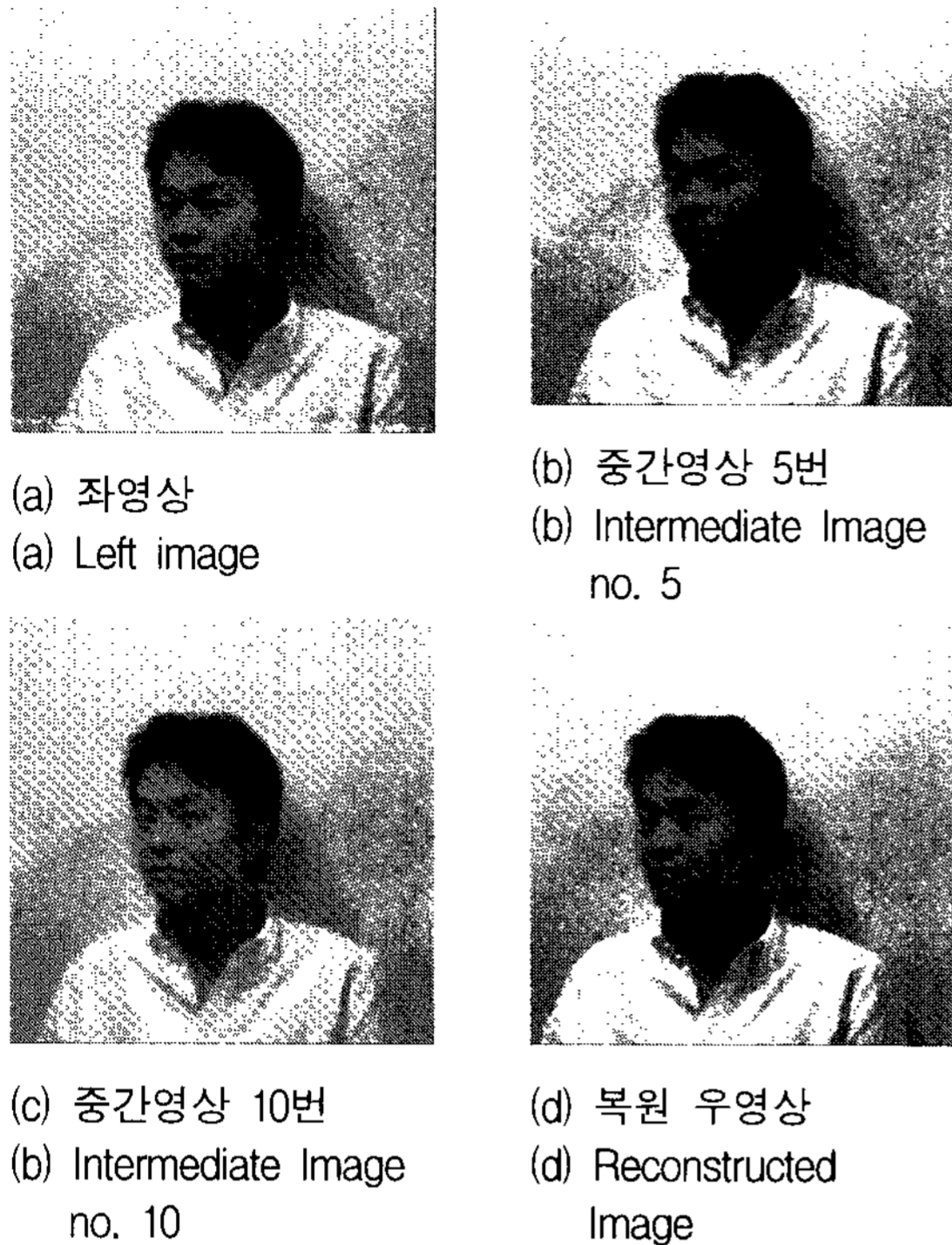


그림 11. 전송된 영상으로 복원된 프레임 5번의 4시점 영상
 Fig. 11. The reconstructed intermediate views (4 views) of frame 5.

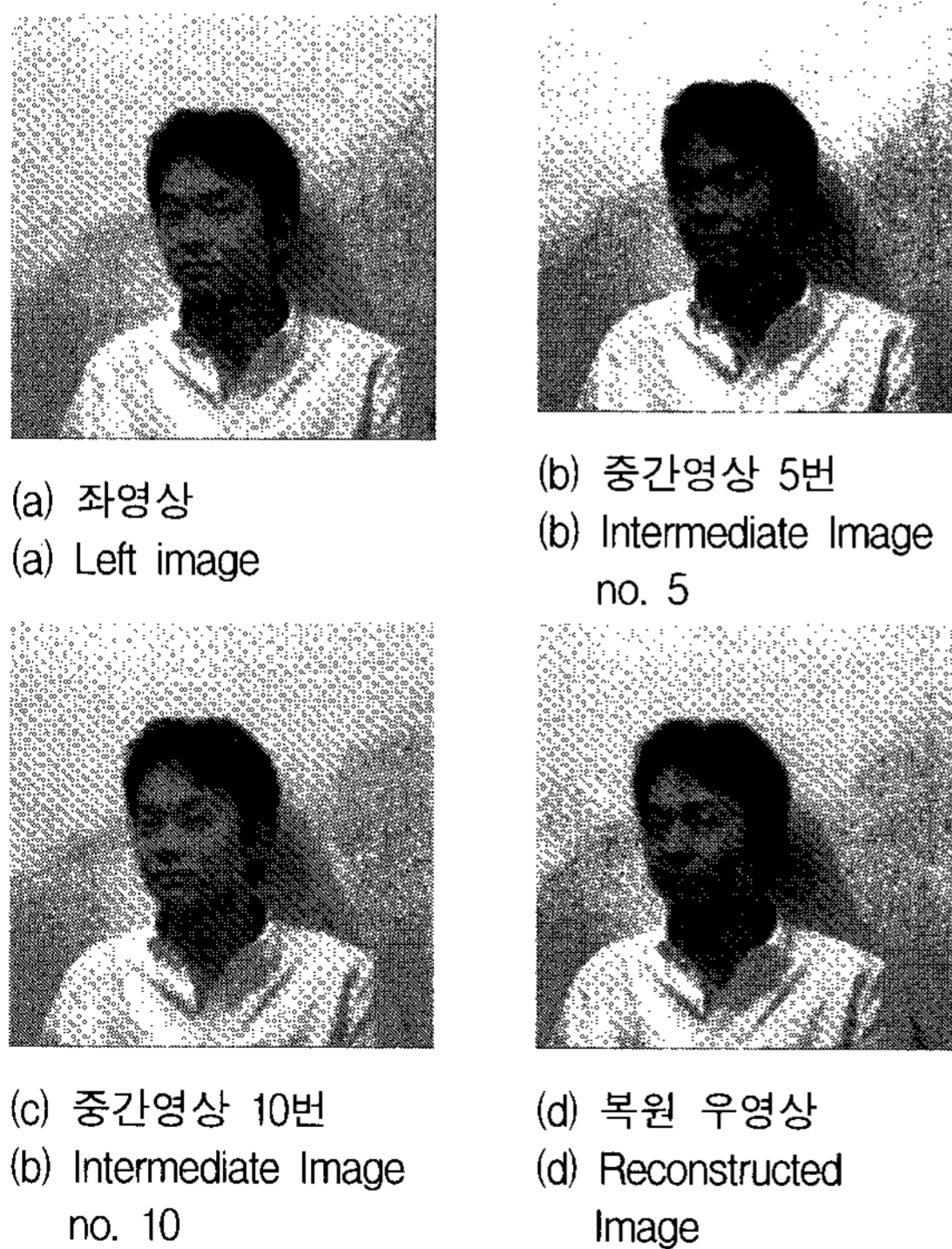


그림 12. 전송된 영상으로 복원된 프레임 10번의 4시점 영상
 Fig. 12. The reconstructed intermediate views (4 views) of frame 10.

표 5. 시점별 중간영상 합성시 소요시간

Table 5. Synthesis processing time of intermediate image as view.

실험횟수	소요시간		
	1 시점	4 시점	16 시점
1	30 ms	70 ms	281 ms
2	31 ms	67 ms	276 ms
3	30 ms	63 ms	280 ms
4	30 ms	66 ms	280 ms
5	30 ms	70 ms	277 ms
평균	30.2 ms	67.2 ms	278.8 ms

때문에 실시간 15프레임의 전송에는 4시점 합성이 현재 제안된 시스템에서의 한계이다. 또한 1시점에서는 지연시간이 30ms로 초당 30프레임의 처리가 가능하며 16시점에서는 초당 3프레임으로 제한된다. 전송되는 영상은 좌 영상과 변이지도이기 때문에 1시점의 중간영상 합성에는 $\alpha = 0.5$ 의 중간영상과 우 영상 복원에 해당하는 2회의 영상합성이 발행하게 된다.

따라서 좌 영상, 중간영상, 우 영상의 3시점을 얻게 되며, 4시점에서는 총 6시점의 영상을 얻게 되고 16시점에서는 총 18시점의 영상을 얻을 수 있다는 분석을 내릴 수 있다.

클라이언트 시스템에 수신되어 재 정렬된 다시점 영상은 서버시스템과 마찬가지로 Prismtek PSM-150에 전송되어 3차원 입체 디스플레이 하게 된다. 각각의 과정에서 발생하는 지연시간은 서버/ 클라이언트 컴퓨터 시스템의 자원 상태나 채널의 상태에 따라 달라진다. 특히 영상을 초당 15프레임 전송할 경우 모든 지연시간의 합이 1/15초(66ms) 이내로 완료되어야 하므로 각 파트의 지연시간에 매우 종속적이다. 한편, 서버보드로는 SE7501HGZ 그리고 주 저장 장치로는 SCSI Ultra320 HDD가 사용되었다. 각각 최소 132MB/s의 대역폭을 가지는 PCI 버스를 3개 가지며 다시점 카메라시스템으로부터 획득한 4시점의 영상을 실시간으로 전송할 수 있다. 디스플레이 어댑터는 Enseio MVP Pro ZX로 1개의 PCI 슬롯과 4개의 비디오 포트를 입출력 장치로 가진다. 이 디스플레이 어댑터는 4개의 디스플레이 장치를 지원한다. 따라서 획득한 4개의 다시점 영상은 4대의 각각의 디스플레이 장치에 출력될 수 있다. 다시점 디스플레이 장치를 통하여 다시점 3차원 디스플레이를 할 수도 있으나 본 실험에서는 그림 13과 같이 스테레오 3차원 디스플레이 장치인 Prismtek PSM-150을 사용하여 각각의 시점에 대하여 분석하였으며, 이 방식을 통하여 각각의 시점에 대하여 보다 명확하게 인지할 수

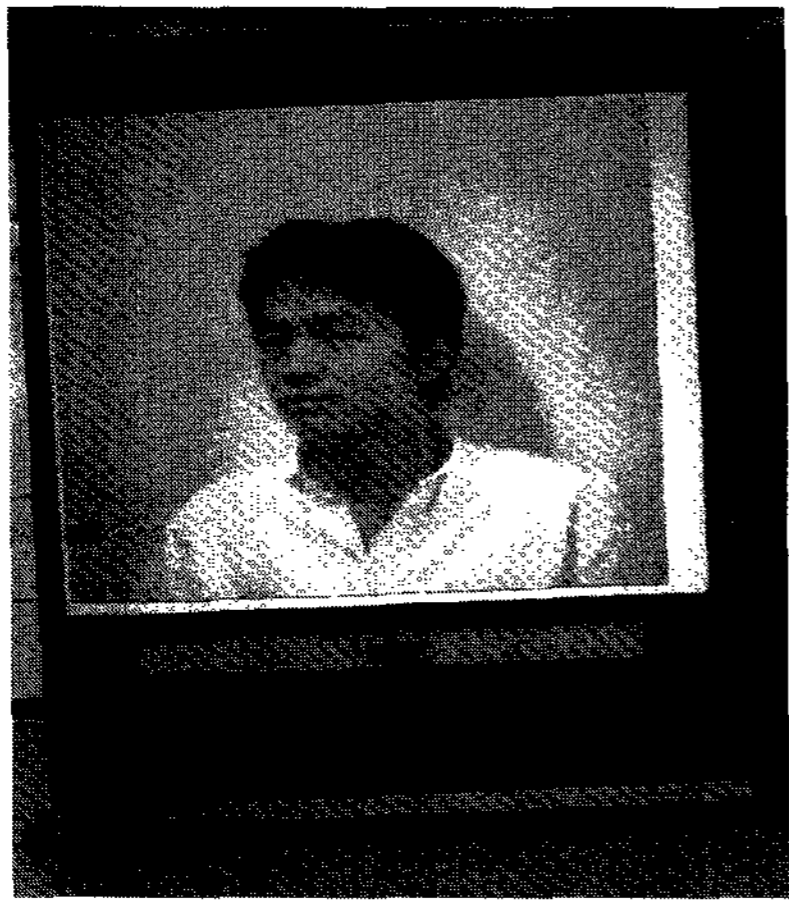


그림 13. Prismtek PSM-150을 통한 4시점 디스플레이
Fig. 13. -view display using Prismtek PSM-150.

있었다.

Prismtek PSM-150은 편광방식의 3차원 디스플레이 장치로 입체효과를 느끼기 위해서는 디스플레이 장치와 편광이 동일한 편광안경이 필요하다. 안경을 통하여 영상을 관한 결과 매우 자연스러운 3차원 영상을 볼 수 있었으나, 실험에서 오브젝트가 카메라와 너무 가까운 거리에 있을 경우에는 눈이 피로함을 느꼈으며 실험 결과 평균 1미터 이상의 거리를 유지할 때 자연스럽게 편광안 3차원 영상을 느낄 수 있었다. 제안된 시스템은 디지털 영상 기반으로 구성되어 있으므로 디스플레이 형식은 호환성이 매우 크다. 실험에 쓰인 3차원 영상 디스플레이 장치 외에도 아날로그 RGB 15핀 D-SUB 단자를 지원하는 모든 시스템에 바로 응용 가능하며 출력 단자가 다를 경우에도 컴퓨터 기반의 시스템이므로 호환성이 매우 큼을 알 수 있었다.

또한, 제안된 시스템에서 실험적인 결과로 서버 컴퓨터 시스템에서는 카메라에서 15fps로 획득한 VGA급의 4시점 영상에 대하여 실시간 디스플레이 하는데 성공하였다. 4시점 디스플레이에서는 VGA급으로 2시점 디스플레이에서는 SXVGA급으로 크기를 실시간 조절하여 영상을 디스플레이 할 수 있었다. 디스플레이 어댑터는 RGB 15핀 D-SUB 포트를 통하여 출력되므로 RGB 15핀 D-SUB를 지원하는 모든 3차원 디스플레이 장치에 바로 응용 가능하며 다른 출력 포트에 대해서도 매우 큰 호환성을 가진다. 웹 기반의 3차원 영상 디스플레이 실험은 서버 컴퓨터에서 다시점 영상을 획득, 압축하여 네트워크를 통하여 전송한 다음 클라이언트에서 전송받은 압축 데이터를 재구성하여 3차원 영상 디스플레이 장치에 전송하였다. 클라이언트 시스템에서 수신한 데이터를 재구성하여 320×240급의 스테레오 영상에 대하

여 7.5fps를 근거리 네트워크에서 실시간 디스플레이 하였다. 이 스테레오 영상을 편광방식의 3차원 디스플레이 장치인 Prismtek PSM-150을 통하여 관찰한 결과 3차원 영상을 관찰할 수 있었다.

IV. 결 론

본 논문에서는 IEEE 1394 카메라, Intel Xeon 서버 시스템, 그래픽 어댑터, Microsoft DirectShow 기술을 이용하는 웹 기반의 실시간 응용에 적합한 3차원 영상 디스플레이 시스템을 제안하고 그 구성요소인 카메라 및 프로세싱 시스템이 최적화되기 위한 조건들을 분석하였다. 제안된 시스템에서는 보다 효율적인 IVS 기법을 이용하여 웹 기반의 원격 3D 화상 통신 시스템의 구현하였으며, 스테레오 카메라를 사용하여 촬영한 영상을 웹 기반의 실시간 전송 및 중간영상 합성 실험결과, 전송에서 4시점과 16시점의 중간영상을 획득하는데 소요되는 지연시간 역시, 각각 67.2ms, 278.8ms로 나타남으로써 이를 통한 새로운 실시간 웹 기반의 원격 3D 화상 통신 시스템의 실질적인 응용 가능성을 제시하였다.

참 고 문 헌

- [1] V. S. Grinberg, G. Podner, and M. W. Siegel, "Geometry of binocular imaging", Proc. of SPIE, vol. 2177, pp. 56 - 65, 1994.
- [2] N. A. Dodgson, J. R. Moore, and S. R. Lang, "Multi-view autostereoscopic 3D display", IBC '99 (International Broadcasting Convention), 10th-14th, pp. 497-502, 1999.
- [3] C. H. Yang, "Geometric models in stereoscopic video", Rapport technique de l'INRS-T'el'ecomunications, no. 95~12, 1999.
- [4] J. Harrold, A. Jacobs, G. J. Woodgate, David Ezra, "Performance of a Convertible 2D and 3D Parallax Barrier Autostereoscopic Display", Proceedings of the SID, 20th International Display Research Conference, Palm Beach Florida USA, 2000.
- [5] I. Sexton, "Parallax Barrier Display Systems", Stereoscopic Television, IEE Colloquium on, pp.5/1-5/5, 1992.
- [6] S. Pastoor, M. Wopkinga, "3-D displays: A review of current technologies", Display, vol. 17, no. 2, pp. 100~110, 1997.
- [7] A. Schmidt, A. Grasnich, "Multi-viewpoint Autostereoscopic Displays from 4D-Vision" Proc

- of SPIE , vol. 4660, 2002.
- [8] C. V. Berkel, D. W. Parker and A. R. Franklin, "Multiview 3D-LCD", Proc of SPIE, vol. 2653, pp. 32~39, 1996.
- [9] C. V. Berkela and J. A. Clarke, "Characterisation and optimisation of 3d-lcd module design", Proc of SPIE, vol. 3012, pp. 179~187, 1997.
- [10] K. H. Bae, J. J Kim, and E. S. Kim, "New disparity estimation scheme based on adaptive matching window for intermediate view reconstruction", Optical Engineering, vol.42, no.6, pp.1778-1786, 2003.
- [11] K. H. Bae, J. H. Ko, E. S. Kim, "Stereo Image Reconstruction using Regularized Adaptive Disparity Estimation", JJAP, vol. 45, no. 5a, pp.4107-4114, 2006.
- [12] <http://www.dti3d.com/>
- [13] <http://www.4d-vision.de/>

 저 자 소 개



고 정 환(정회원)
 1999년 광운대학교 제어계측
 공학과 학사졸업
 2001년 광운대학교 전자공학과
 석사졸업
 2005년 광운대학교 전자공학과
 박사졸업

2005년~2006년 광운대학교 산학협력단
 책임연구원
 2006년~현재 (주) 인코시스 기술연구소
 책임연구원
 <주관심분야 : 로봇비전, 3D 영상통신>



안 영 환(정회원)
 1983년 인하대학교 기계공학과
 학사 졸업
 1990년 매릴랜드대학 기계공학과
 석사졸업
 1998년 웨스트버지니아대학
 기계공학과 박사졸업

현재 인하공업전문대학 메카트로닉스과 부교수
 <주관심분야 : 자동제어, 신경회로망>



이 정 석(정회원)
 1985년 광운대학교 전기공학과
 학사졸업
 1990년 광운대학교 전기공학과
 석사졸업
 2001년 광운대학교 제어계측
 공학과 박사졸업

1990년~1997년 국방과학연구소 선임연구원
 현재 인하공업전문대학 메카트로닉스과 조교수
 <주관심분야 : 제어계측, 마이크로프로세서, 머신
 비전>