

방송 및 모바일 실감형 2D/3D 콘텐츠 변환 방법 및 플랫폼

2D/3D conversion algorithm on broadcast and mobile environment and the platform

송혁, 배진우, 유지상, 최병호*
(Hyok Song, Jin-woo Bae, Ji-sang Yoo, Byeoung ho Choi)

Abstract : TV technology started from black and white TV. Color TV invented and users request more realistic TV technology. The next technology is 3DTV. For 3DTV, 3D display technology, 3D coding technology, digital mux/demux technology in broadcast and 3D video acquisition are needed. Moreover, Almost every contents now exist are 2D contents. It causes necessity to convert from 2D to 3D. This article describes 2D/3D conversion algorithm and H/W platform on FPGA board. Time difference makes 3D effect and convolution filter increased the effect. Distorted image and original image give 3D effect. The algorithm is shown on 3D display. The display device shows 3D effect by parallax barrier method and has FPGA board.

Keywords: 3DTV, 2D/3D conversion, 3D display

I. 서론

컴퓨터 그래픽, 컴퓨터 비전 및 멀티미디어 통신 기술들은 새로운 응용 기술인 Multi-view 비디오 코딩과 같은 신기술을 탄생시켰으며 발전시키고 있다. 이 기술들은 흑백TV, 컬러TV 그리고 HDTV로부터 발전하여 오늘에 이르고 있다. 다시점 TV또는 3DTV기술은 사용자가 임의의 시야각에서 영상을 감상할 수 있도록 해 준다. 다시점 영상 영상의 부가 데이터인 Depth data 또는 Disparity data는 자연영상을 이용한 처리나 컴퓨터 그래픽에 사용하게 된다.

Depth data 또는 Disparity 데이터를 획득하는데 가장 부분은 대응점 예측이다. 이는 컴퓨터 비전의 스테레오 영상 처리에서 대응점 예측은 매우 연구가치가 높은 분야중의 하나이다. 스테레오 매칭의 목적은 스테레오 영상에서 대응점을 찾아서 깊이 데이터를 획득하는 것이다. 이 분야는 지난 수십년간 지속적으로 연구되어 왔으나 현재 컴퓨터 영상처리 분야에서 계속적인 관심을 끄는 분야 중에 하나이다.

디스플레이 분야에서 다양한 3D 영상 디스플레이 장치들이 생산되고 있으며 3D 영상의 효과는 계속 개선되고 있다. 3D 영상처리의 과정은 스테레오 카메라 또는 Multi-view 카메라를 통하여 획득된 영상이 Depth data 또는 Disparity 데이터를 이용하여 3D 영상으로 변환되고 다양한 디스플레

이 장치에 보여지게 된다.

본 논문의 구성은 다음과 같다. 2장에서 3D 시스템을 소개하며 3장에서는 2D/3D 변환알고리즘을 보이며 4장에서는 실험결과를 보인다.

II. 3DTV 시스템과 깊이감

3D 시스템은 크게 세 파트로 나눌 수 있다. 첫 번째로 획득부이다. 스테레오 카메라, 평행 Multi-view 카메라, 사각 카메라 및 Depth 카메라 등이 다양한 Multi-view 카메라들이다. 획득된 이미지는 코딩/전송의 단계를 거쳐 디코딩된다. 디코딩된 데이터는 3D 디스플레이 장치에 사용된다.

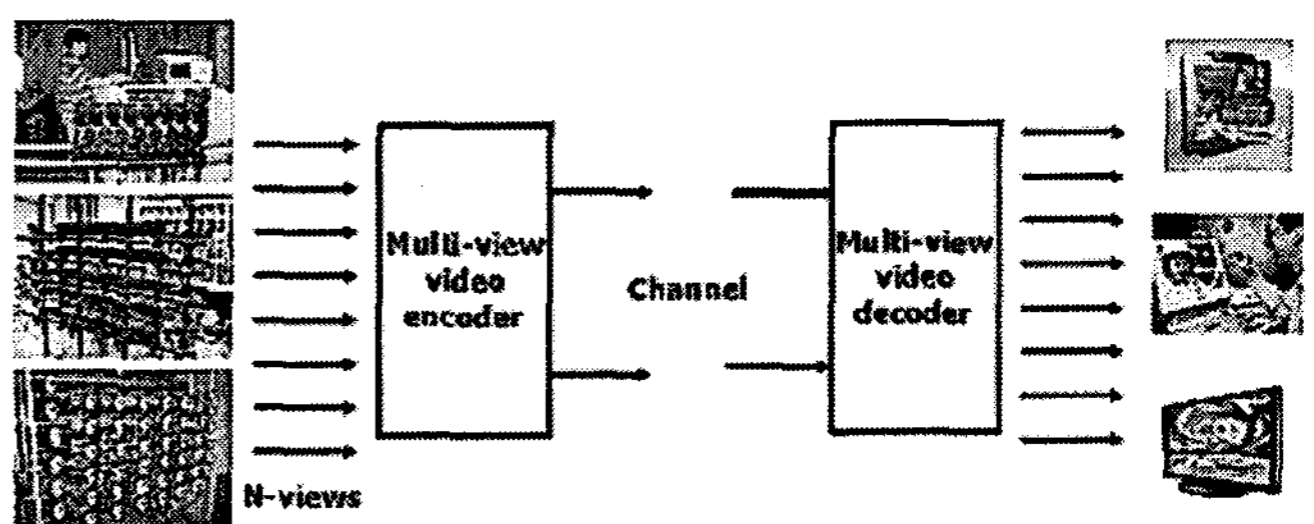


그림 1. 3DTV 시스템의 구조

그림 1은 3DTV 시스템의 블록 다이어그램이다. 전자부품 연구원에서는 1차원 평행 다시점 카메라를 제작하여 영상을 획득하였고 이를 H.264 표준을 이용하여 압축 및 전송하였다[1]. CMU는 48개의 카메라를 2차원 배열하여 영상을 획득하였다[2]. 나고야 대학에서는 100개의 카메라를 1차원 그리고 2차원으로 배열하여 영상을 획득하였다. 다시점 카메라는 카메라의 조정이 필요하며 영상의 정렬이 요구된다. 이러한 과정은 영상 스트림을 이용하여 영상을 압축하는데 있어서 이득이 된다. 또한 이 영상은 효율적으로 3D 효과를 줄 수 있다. 압축된 영상은 압축되지 않은 영상보다 채널을 이용한 전송시에 채널의 Bandwidth를 이용하는데 효율적이다. 다시점 카메라로부터 획득된 영상과 부가 데이터를 이

* 책임저자(Corresponding Author): 최병호
논문접수 : 2007. 08. 06., 채택확정 : 2007. 08. 06.
송혁, 최병호 : Korea Electronics Technology Institute
hsong@keti.re.kr, bhchoi@keti.re.kr
배진우 : Korea Invention Promotion Association
bjw8751@kipa.org
유지상 : Kwangwoon University
jsyoo@daisy.kw.a.ckr

※ 본 연구는 정보통신부 및 정보통신연구진흥원의 IT 신성장동력핵심기술개발사업의 일환으로 수행하였음. [2006-S-050-01, 지능형 로봇 시청각 신호처리 SoC]

용하여 디코딩된 영상은 다양한 3D 디스플레이 장치에 표시된다.

III. 인터플레이션

영상의 Geometric 처리에 있어서 영상의 크기 조절, 영상의 회전, 그리고 영상의 Translation 이 있다[3]. 영상을 조작할 시에 Hole과 Overlap일 발생하는데 이를 처리하기 위한 방법으로 인터플레이션이 있다. 가장 간단한 방법으로는 Nearest neighbor 방법이며 가장 많이 사용하는 방법으로 Bilinear interpolation 이다. Nearest neighbor 방법은 원 영상 픽셀을 이용하므로 새로운 픽셀을 위한 연산이 필요 없으며 Bilinear interpolation 방법은 가중치가 주어진 주변 픽셀들의 연산으로 이루어진다. 고차 인터플레이션 방법으로는 Cubic convolution 방법과 B-spline 필터를 이용한 방법을 많이 사용한다. 인터플레이션 함수는 아래와 같이 정의한다.

$$\hat{f}(x) = \sum_k C_k \beta(x - x_k) \quad (1)$$

여기에서 $\hat{f}(x)$ 는 인터플레이션 함수이고, $\beta(x)$ 는 인터플레이션 커널이다. 그리고 x 와 x_k 는 각각 연속함수 값과 Discrete 값을 의미한다.

본 논문에서는 Cubic convolution filter 를 사용하였다. Cubic convolution filter는 filter의 파형이 양 끝에서 고정된 값을 가지므로 연산이 용이하다. 아래는 일반적인 Cubic convolution 함수이다.

$$\begin{aligned} \hat{f}(x) = & \alpha_k (f(x_{k-1}) - f(x_{k+1}))A(s) \\ & + \alpha_k (f(x_k) - f(x_{k+2}))B(s) \\ & + (f(x+k) - f(x_{k+1}))C(s) \\ & + f(x_k) \end{aligned} \quad (2)$$

여기에서

$$\begin{aligned} A(s) &= (S^3 - 2S^2 + s) \\ B(s) &= (S^3 - S^2) \\ C(s) &= (2S^3 - 3S^2) \end{aligned} \quad (3)$$

이다.

III. 2D/3D 변환

대부분의 콘텐츠는 2D 카메라로 획득/저장/압축된다. 입체로 표현되는 스테레오스코픽 영상은 두 개의 다른 각도의 영상으로 표현되나, 현재는 스테레오스코픽 카메라가 개발되어 3D 데이터로도 저장되기도 하나 소수이므로 2D/3D 변환이 필요하다.

Garcia의 한 시각적 특성을 이용하여 시간차이를 둔

연속영상을 이용하여 좌우 영상을 구성하면 #효과를 만들어낼 수 있다[5]. 이 방법은 매우 간단하면서 좋은 효과를 만들어 낼 수 있어 쉽게 응용 가능하다. 연속된 영상의 특성을 파악하여 시간차이를 조절하여 적절한 좌우 영상을 선택하여 깊이감을 조절하여 시청자에게 피로감을 주지 않는 시차를 만들어 낸다.

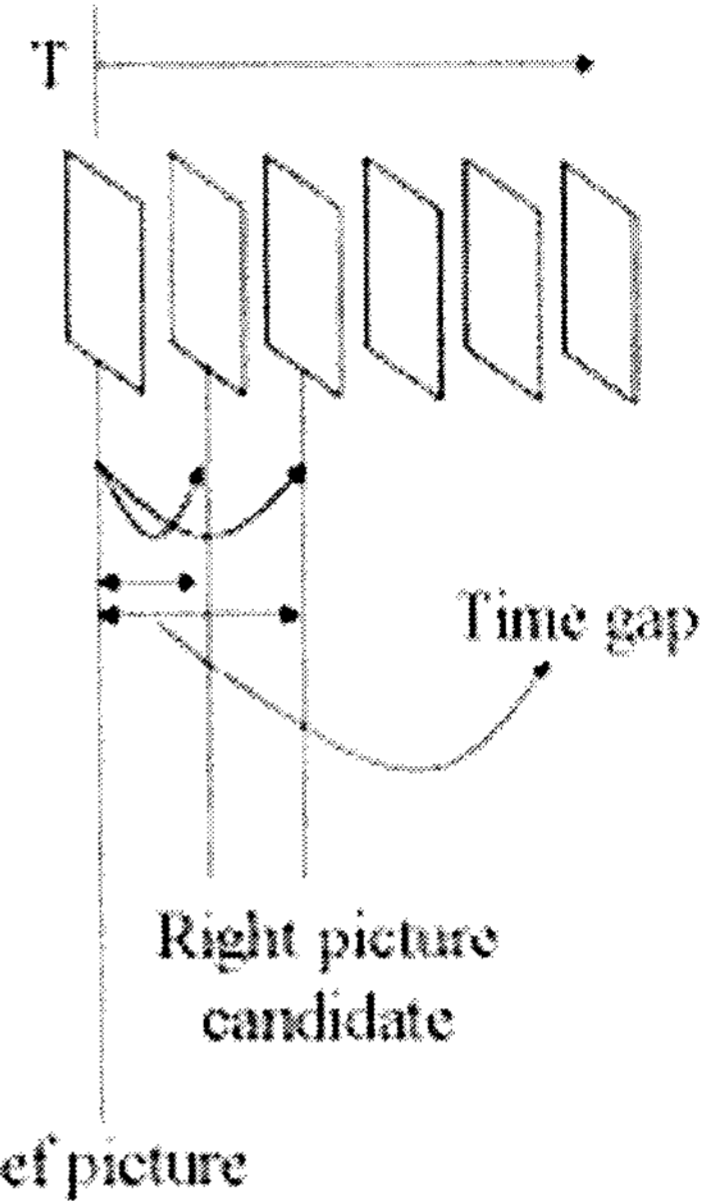


그림 2. 시간차를 이용한 3D 변환 방법

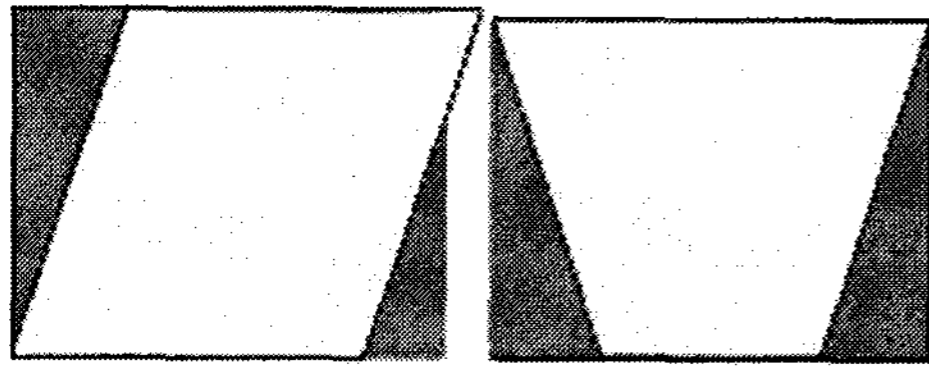
영상의 일반적인 특성을 이용하여 3D 효과를 만들어내는 방식으로 영상의 상하에 배치된 객체와 배경에 시차를 달리 부여하는 방법이 있다. 일반적으로 영상의 상단은 영상의 배경으로 객체에 비해 먼 위치에 있고 객체는 아래측에 가까운 쪽에 배치되는 특성을 이용하여 시차를 달리 만들어 낼 수 있다. 이러한 효과는 필터를 이용하여 영상을 왜곡하여 만들어 낸다. S 값에 의하여 필터의 계수가 주어지며 다양한 영상의 왜곡이 가능하다. S를 구하는 식은 아래와 같다.

$$\begin{aligned} f_{temp} &= B_{hor} \times i - \frac{J}{H} \times D - UD \\ J_{int} &= (\text{int}) f_{temp} \\ S &= f_{temp} \times J_{int} \end{aligned} \quad (4)$$

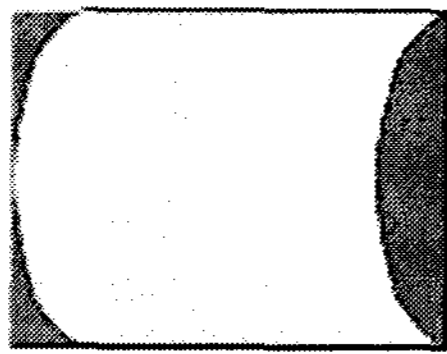
식(1)과 식(2)에서 s를 구하기 위한 식은 위와 같다. B_{hor} 는 입력사이즈와 출력사이즈의 비율이며 는 픽셀의 수평위치이다. J는 픽셀의 수직위치이며 H는 영상의 높이를 나타낸다. D는 왜곡하는 영상의 중심 기준을 의미하며 UD는 왜곡되는 전체 영상의 시차를 의미한다. 본 식을 이용하여 다양한 영상의 왜곡이 가능해진다. 영상의 왜곡의 예는 그림 3과 같다.

식(2)와 식(3)을 이용하여 영상의 특성을 이용하여 상하 시차를 달리하여 그림 4와 같은 결과를 얻었다.

그림 4는 PC 시스템에서 소프트웨어를 이용하여 알고리즘을 실행하여 얻은 결과이다. 그림 2의 (a)는 D값은 0이 되며 UD를 조절한 결과이며 (b)는 D를 0으로 하여 UD값을 점진적으로 변화하여 나온 결과이다.



(a) (b)



(c)

그림 3. 영상왜곡의 예

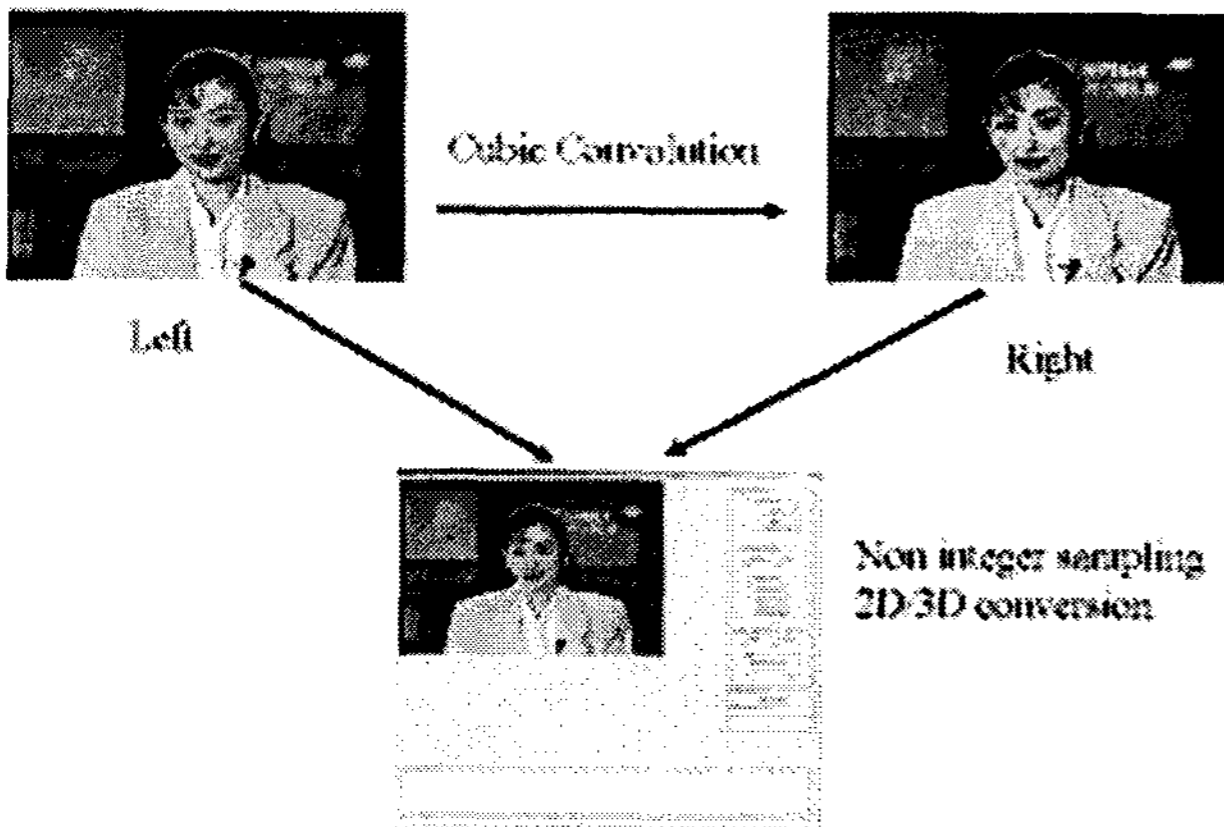


그림 4. 좌우 시차를 달리한 영상

3D 디스플레이 장치는 안경 유무에 따라서 나안식과 안경식으로 나뉘며 시점에 따라서 스테레오 및 다시점 디스플레이로 나뉜다. 3D 효과를 생성하는 방식에 의하여 Parallax barrier 방식 Reticular 방식 이 주로 사용된다. 이런 3D 디스플레이 장치는 Subpixel 방식으로 화소의 RGB 요소가 뒤섞여 표현하게 된다. 따라서 영상의 Formatter가 필요하다. 본 논문의 구현에 사용한 디스플레이 장치인 나안식 Parallax barrier 방식으로 구현할 경우 Subpixel muxing이 필요하다. 그림 5는 3D 디스플레이 장치에서의 Subpixel muxing을 나타낸다.

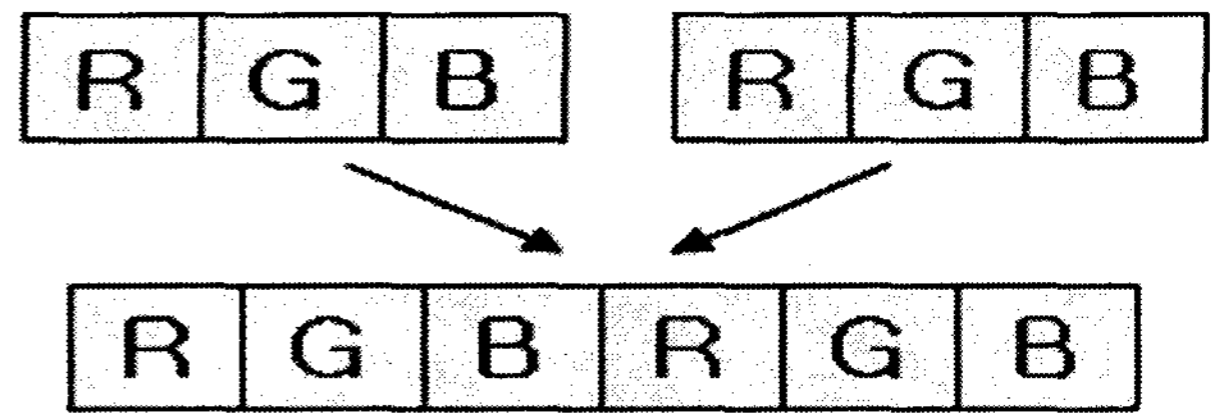


그림 5. Subpixel muxing

IV. 실험결과

본 논문의 구현은 Cubic convolution filter를 FPGA로 구현하여 이를 3D LCD에 표현하였다. 실험에는 600,00개의 게이트를 가진 Actel APA600-BG456을 이용하여 하드웨어를 구성하였다. 제작된 하드웨어의 블록도는 아래와 같다.

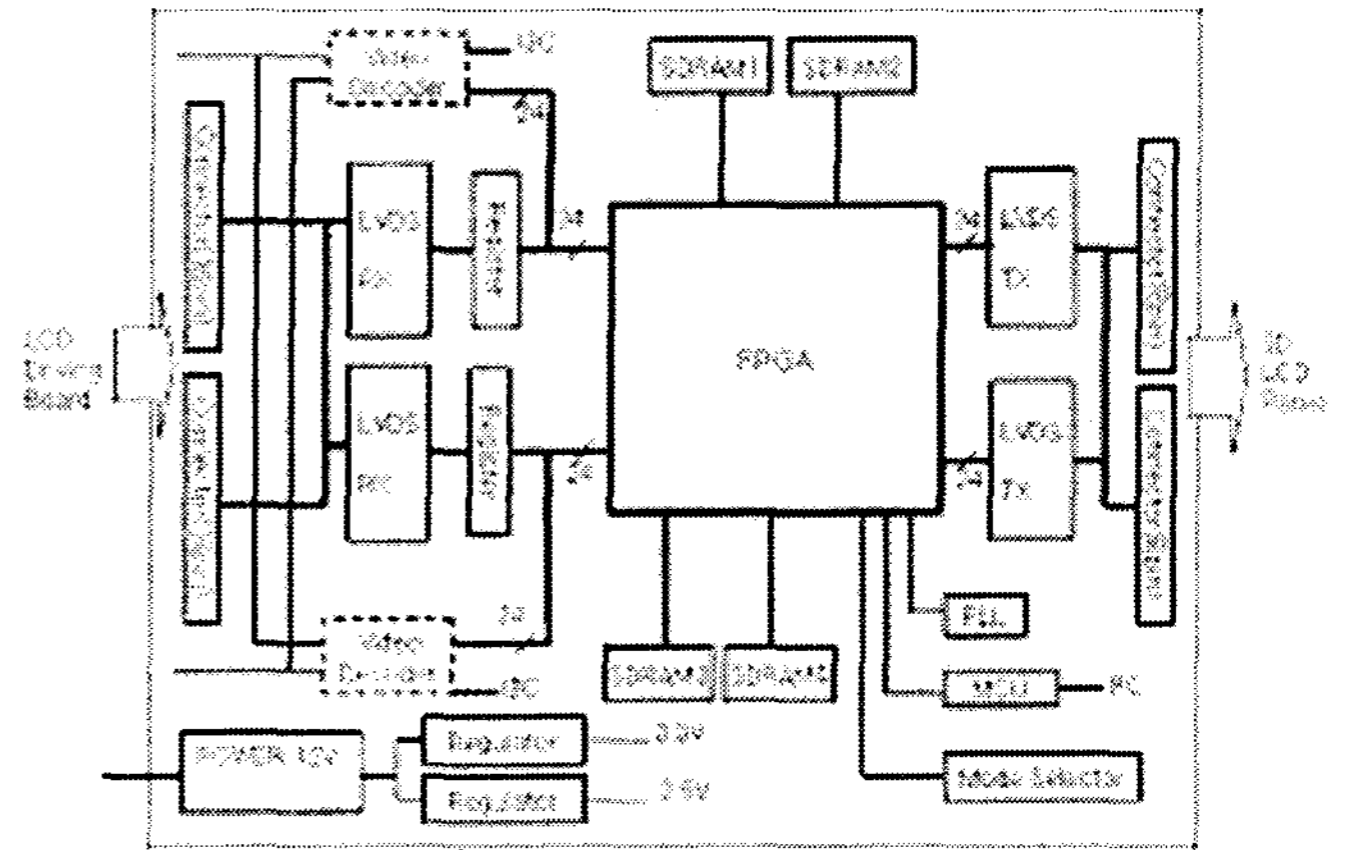


그림 6. 제작된 FPGA 보드의 블록도

그림 6에서 입력 영상은 PC의 VGA 출력을 통하여 FPGA 보드로 전달되며 FPGA 보드에서는 LVDS 신호를 FPGA 신호 레벨에 맞게 변환하여 메모리로 저장하였다. 이 때 LVDS 신호 변환 하드웨어를 사용하였다. 입력되는 영상 데이터는 최소 2프레임이 메모리에 저장되어 원영상과 왜곡된 입체 영상으로 두 영상을 이용하여 입체 효과를 나타내었다. 메모리에 저장된 영상은 원영상이며 원영상 데이터를 왜곡시켜 왜곡된 영상과 muxing 한다. 제작된 플랫폼 보드는 아래와 같다.

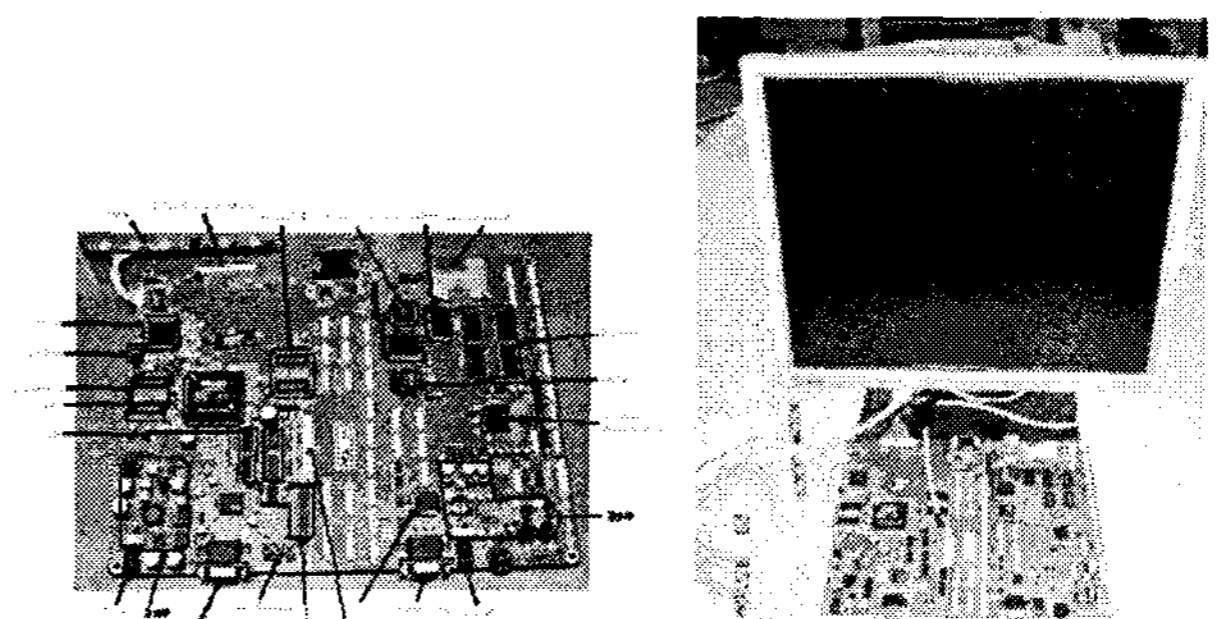


그림 7. 제작된 플랫폼 보드

영상의 시차는 디스플레이 장치 관점에서 Pixel 단위로 표현한다. 영상의 왜곡은 수식에 의하여 조절되나 3D 디스플레이 장치에 표현된 영상을 사용자가 감상할 경우 사용자의 위치에 따라 입체감이 매우 다르게 표현된다. 사용된 디스플레이 장치의 경우 1280x1024 의 해상도를 가지며 1미터 거리에서 사용자가 감상할 경우 눈의 피로를 느끼지 않는 한도내에서 입체감을 형성할 경우 시차의 최대치는 약 7pixel이라는 실험적 데이터를 얻었다. 이는 지속적인 시차를 가질 때와 순간적 시차를 가질 경우의 데이터가 다르며 사용자의 위치가 디스플레이 장치로부터 멀어질 경우 그 시차는 커지게 된다.

VI. 결론

대부분의 영상 콘텐츠가 2D 데이터이므로 2D/3D 변환은 매우 중요한 분야이다. 홈씨어터 분야에서 2D 영상 콘텐츠를 3D 프로젝터를 이용하여 감상할 수 있으며 모바일 환경에서 3D 모바일 디스플레이 장치를 이용하여 입체감을 느낄 수 있다. 본 논문은 영상의 압축 표준과 상관없이 모든 동영상에서 3D 영상으로 감상할 수 있는 알고리즘으로 다양한 분야에 사용될 수 있다. 그러나 이 방법은 단순한 영상의 Translation만을 고려하여 왜곡된 영상이므로 더 나은 영상의 왜곡을 위하여 영상 생성시에 카메라 파라미터를 고려하여 임의의 카메라로부터 Intrinsic 파라미터와 Extrinsic 파라미터를 구하여 이를 이용하여 영상을 구하여야 한다.

참고문헌

- [1] C. K. Jung, H. Song, J. H. Park, B. H. Choi, "MVC multi-view video test sequence from KETP", *ISO/IEC JTC 1/SC 29/WG 11 and ITU-T SG 16, JVT-X023*, Geneva, Swiss, June 2007
- [2] Ted Square, Tsuhan Chen, "CMU ultimate eye multi-camera data", *ISO/IEC JTC 1/SC 29/WG 11, Document M12003*, Busan, Korea, April 2005
- [3] Randy Crane, "A simplified approach to image processing", Prentice hall 1997
- [4] Jong-ki Han, Seung-ung Baek, "Parametric cubic convolution scaler for enlargement and reduction of image", *IEE transactions on Consumer Electronics*, Vol. 46, 2000
- [5] T. Okino, "New television with 2D/3D image conversion techniques", SPIE, Vol. 2653, Photonic west, 1995



송 혁

1999년 광운대학교 제어계측공학과(공학사). 2001년 광운대학교 전자공학과(공학석사). 2000년~현재 전자부품연구원 연구원. 관심분야는 3DTV, 영상압축, 로봇 영상 등임.



배 진 우

1998년 순천대학교 전자공학과(공학사). 2003년 광운대학교 전자공학과(공학석사). 2006년 광운대학교 전자공학과 전자공학과(공학박사) 2007~현재 발명진흥회 전문위원. 관심분야는 영상신호처리, 3DTV, 영상압축 등임



유 지 상

1985년 서울대학교 전자공학과(공학사). 1987년 서울대학교 전자공학과(공학석사). 1993년 Purdue university Electrical engineering(공학박사). 1994년~1997년 한림대학교 조교수. 1997년~현재 광운대학교 정교수. 관심분야는 웨이블릿 기반 영상처리, 영상압축, 영상인식, 비선형신호처리 등임.



최 병 호

1993년 한양대학교 전자공학과(공학사). 1997년 한양대학교 전자공학과(공학석사). 1997~현재 전자부품연구원 책임연구원. 관심분야는 디지털 영상처리, 3DTV 및 스테레오 비전 등임