

실시간 입체 영상 변환기 구현

고윤호*, 김경원*, 권병헌**, 최명렬*
*한양대학교 전자전기제어계측공학과
**유한대학 정보통신과
e-mail:ciel0611@asic.hanyang.ac.kr

Implement of a Real-Time Stereoscopic Image Converter

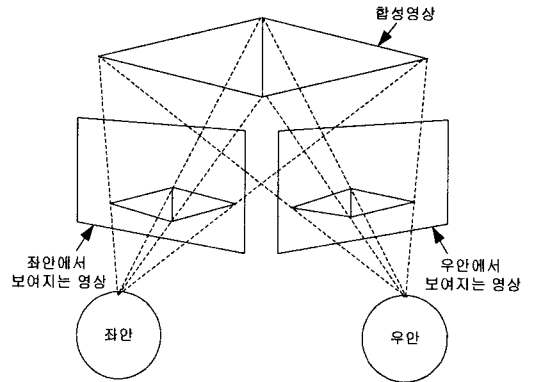
Yoon-Ho Ko*, Kyong-Won Kim*,
Byong-Heon Kwon**, Myung-Ryul Choi*
*Dept of EECL, Han-Yang University
**Dept of Information and Telecommunication, Yuhan College

요 약

실시간으로 2차원 영상을 입체 영상으로 변환하는 알고리즘을 제안하였으며 입체 영상 변환기를 구현하여 검증하였다. 입체 영상 변환기는 motion-parallax를 이용하여 물체의 움직임 방향과 속도에 관계 없이 실감나는 실시간 입체 영상을 재생 할 수 있다. 본 논문에서는 입체 영상의 원리에 대해 논하였고 기존의 입체 영상 변환기에 기법에 대해 서술하였으며 실시간 알고리즘을 제안하였고 입체 영상 변환기의 구현 방법에 대해 서술하였다.

1. 서론

3차원 영상은 그림 1과 같이 양안을 통해 보이는 서로 다른 영상을 인식함으로써 인간의 뇌에서 3차원 영상이 합성된다. 하지만, 2차원 영상의 경우에는 양안이 동일한 영상을 보게 되므로 입체적인 사물을 볼 때와 달리 불편하지만 사람들은 지금까지 반복된 경험에 의해 자연스럽게 평면으로 받아들인다. 따라서 현실과 비슷한 입체감을 얻기 위해서는 처음부터 입체 카메라를 사용하여 영상을 만들거나 또는 수작업을 통하여 입체로 변환하여 주거나, 컴퓨터 그래픽인 경우에는 양안을 위하여 두 번씩 다르게 렌더링 하여야한다. 하지만, 이 작업은 많은 비용과 시간이 소요되며, 기존의 2차원 기반으로 제작된 방대한 양의 영상 자료를 3차원으로 변환할 수 없는 단점이 있다.



(그림 1) 입체 영상의 재생 원리

2. Depth 지각의 요인

2.1 단안의 Depth의 요인

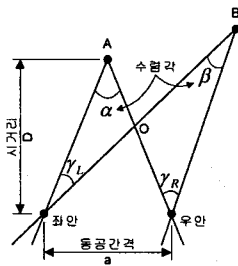
먼 곳에 있는 물체일수록 흐리거나 희미해져서

콘트라스트가 약하게 보인다는 공기투기(Aerial Perspective)의 효과나 선과 선의 간격이 좁아지는 것 같은 도형을 볼 때에 깊이감이 생기는 선 원근법(Linear Perspective), 멀어질수록 방막에 투영되는 상의 텍스처 치밀하며, 또 많은 상이 조밀하게 투영되게 되는 텍스처 그라디언트(Texture Gradient)의 효과도 있다. 또 물체에 생기는 명암의 효과도 중요

한 실마리가 되며, 뒤쪽의 것은 앞 쪽의 것에 의해 그 일부가 덮이는 중첩(Interposition)의 효과도 모두 단안에 의한 요인이다.

2.2 단안의 Depth의 요인

인간의 눈이 가로방향으로 약 6.5cm 떨어져서 2개 존재한다는 것에 따른 양안으로의 실마리는 오행판단에 특히 중요하며, 이것에는 수렴과 양안시차로 불리는 것이 있다.

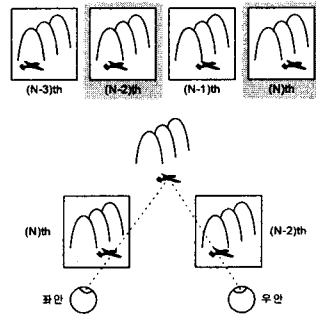


(그림 2) 수렴과 양안시차

그림 2와 같이 어떤 대상 A를 바라볼 때, 양안은 안쪽으로 회전하여, 그 대상 위에서 만난다. 이와 같은 양안의 작용을 수렴이라 부른다. 이 때 대상 A와 시선이 이루는 각(그림 2의 α)을 수렴 각이라 부른다. 수렴에 의한 오행의 감도는 근거리에서 큰 효과를 나타내는데 20cm정도까지 유효하다고 한다. 그러나 거리가 멀어지면 수렴 각이 작아지므로 이 효과는 급격히 작아진다. 또, 양안이 일정 간격 떨어져있기 때문에, 어떤 물체를 바라보았을 때 양안의 망막상은 같아지지 않고, 주시점으로부터 떨어진 위치에서는 대체로 간격이 생긴다. 양안에서의 이와 같은 간격 량의 차이가 양안시차이며, 그림 2에서 A를 응시하고 있을 때 이것과 다른 방향에서 다른 오행을 갖는 B와의 사이에서 생기는 차, 즉 각도로 표현하면 $\gamma_L - \gamma_R$ 또는 $\beta - \alpha$ 로 주어진다. 양안 상에 간격이 있을 때, 일반적으로는 2중상으로 보이게 되지만, 양안 시차가 어느 정도 이하일 때는 상이 하나로 융합하여 간격 량의 크기 및 방향에 따라 주시하고 있는 점 앞에 또는 뒤로 명확한 오행이 느껴진다. 일반적인 입체 디스플레이에서는 이 효과를 가장 많이 활용하고 있다.

3. 기존의 입체 변환 방법

스테레오 변환 기술의 최초 등장은 1993년 일본 산요 전기에서 2차원-3차원 변환 TV를 상업용 목적으로 개발한 이후로, 종종 기술 소개 및 제품이 등장하고 있다. Okino 그룹은 세계 최초로 MTD(Modified Time Difference)방식을 이용하여 2차원-3차원 동영상 변환 TV를 개발하였다. 이 변환 기술은 현재 영상과 카메라 및 물체 운동이 존재할 때에 시간적으로 지연된 영상을 각각 좌영상과 우영상으로 보여 입체 효과를 얻는 방식이다. 이 방식은 움직임이 크지 않을 때에는 효과가 크지만 운동량이 커지면 고스트가(ghost)가 생기는 단점이 있다.



(그림 3) MTD 방식의 원리

4. 제안한 입체 영상 변환 방법

제안한 입체 영상 변환 방법은 2차원 영상에서 운동 시차를 이용하여 영상 내 운동 물체의 운동 방향 및 속도와는 상관없이 서로 다른 원근 오행을 갖는 입체 영상을 실시간으로 제공한다.

4.1 샘플 영상의 취득

먼저 운동 시차의 효율적인 계산과 실시간 처리를 위하여 현재 및 이전 영상의 해상도 보다 작은 해상도의 샘플 영상을 PD1, PD2의 등간 격으로 그림 5와 같이 취득한다. 등간 격으로 샘플링된 샘플 영상은 원영상의 형상 정보 및 휘도 분포 특성이 동일하다. 즉, 샘플 영상에 대한 히스토그램의 평균 및 표준편차는 원영상에 대한 히스토그램의 평균 및 표준편차가 동일하기 때문에 운동 시차를 실시간으로 계산하기 위해서 샘플 영상을 이용하는데 큰 무리가 없다. 또한, 취득된 샘플 영상의 형상 정보를 추출하기 위해서 컬러 신호를 휘도 신호를 변환한다.

4.2 움직임 검출 및 영역 분할

현재 및 이전 영상에 대한 샘플 영상의 휘도 신호에서 움직임이 있는 픽셀을 검출한다. 움직임 검출 방법은 현재 ($P_{(M)^*}$)와 이전 ($P_{(N-1)^*}$) 영상의 픽셀 차의 절대값 (D_{Pixel})을 구하고, 임계값 (D_{th}) 비교를 통하여 정지 및 움직임 픽셀로 구분하였다. 여기서는 정지 픽셀과 움직임 픽셀 이렇게 두 가지 상태로만 검출을 하였는데, 정지 픽셀은 일반적으로 배경을 이루는 픽셀로 상대적으로 먼 거리에 위치해 있는 것으로 가정하고, 반대로 움직임 픽셀은 상대적으로 가까운 위치에 있는 것으로 가정한다.

상기한 방법으로 검출된 픽셀은 배경 또는 운동 물체의 영역을 구성하는 픽셀로, 이 픽셀 값을 이용하여 픽셀을 구성하는 영역으로 확장시키는 단계가 필요하다. 샘플 영상을 수직 방향으로 8등분하여, 각 영역에서 정지 또는 움직이는 픽셀 값들에 대한 대푯값을 계산한다. 샘플 영상을 수직 방향으로 8등분하는 이유는 움직이는 픽셀 값이 영상 전역에 걸쳐 동일한 계조 값이 아닌 서로 다른 계조값으로 이루어져 있을 경우 검출 에러를 줄이기 위함이다.

4.3 오행 지도(Depth Map) 생성

샘플 영상에서 각 영역당 계산된 8개의 움직임 픽셀의 대푯값 (P_{th})을 이용하여 원영상의 해상도 크기의 오행 지도를 작성한다. 즉, 움직임 픽셀의 대푯값 ($P_{(M)^*}$)을 중심으로 실험 결과에 의하여 상하위 25%의 오차를 갖는 픽셀 값을 움직이는 물체를 구성하는 픽셀 군으로 정하고, 상대적으로 가까운 위치에 있는 영역이므로 오행 값을 작게 설정해 주며, 그 외의 픽셀군은 오행 값을 크게 설정한다.

4.5 양시차(Positive Parallax) 처리

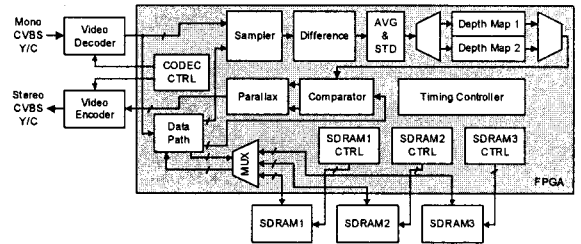
본 논문에서는 배경 및 운동 물체의 시차 처리는 모두 양시차 처리를 행하여 움직이는 물체가 화면 안쪽에 위치하고 배경 또한 움직이는 물체보다 더 뒤쪽으로 화면 안쪽에 위치하도록 하였다. 배경과 운동 물체의 오행 차이를 4픽셀로 한정했으며, 오행

차이에 의한 교합 문제는 ZOI(Zero Order Interpolation) 및 FOI(First Order Interpolation) 과 같은 보간 알고리즘을 사용하여 해결하였다.

5. 하드웨어 구조

제안된 변환기는 비디오 코덱과 SDRAM과 FPGA로 구성되어 있다.

비디오 디코더는 TV 또는 VTR, DVD에서 들어오는 아날로그 입력 CVBS(Composite Video Broadcast Signal) 또는 S-video(Y/C)를 입력으로 받아 CCIR-601 색채 요소로 변환해 준다. 비디오 인코더는 디지털 YUV 색채 신호를 NTSC, PAL의 CVBS 또는 S-video로 변환해 준다. SDRAM은 비디오 디코더에 의해 생성된 ITU 656 YUV 4:2:2 포맷의 데이터를 저장한다.

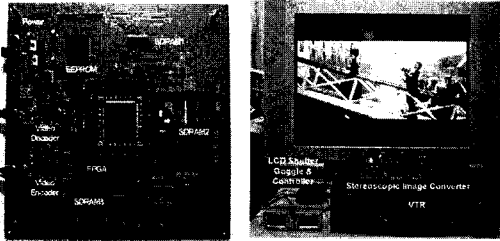


(그림 4) 제안한 변환기의 블록다이어그램

그림 5는 제안한 변환기의 블록다이어그램이다. Data path부와 MUX부는 비디오 디코더에서 나오는 데이터를 SDRAM과 sampler부에 보내거나 SDRAM 가운데 이전영상을 선택하는데 사용된다. Sampler부에서는 현재와 이전영상에서 원영상보다 작은 샘플영상을 만든다. 이때 만들어진 영상은 Y신호만으로 구성된다. Difference부에서는 이전 영상과 현재 영상 사이에 움직임 픽셀을 찾아낸다. AVG & STD부에서는 샘플영상을 8단계로 나누어 각각의 움직임 픽셀의 평균값을 구한다. Depth Map과 Comparator부에서는 이전 블록에의 평균값과 입력되는 영상과 비교를 하여 입력되는 픽셀이 움직임 픽셀인지 배경 픽셀인지를 구분한다. Parallax부에서는 제안된 보간 알고리즘을 이용하여 3차원 영상을 만든다. Timing Controller와 SDRAM CTRL은 전 블록을 제어하여 전체 동작의 순서를 제어한다.

5. 구현

VHDL을 이용하여 제안한 입체 영상 변환 기법을 설계하고 그림 5의 (a)와 같이 구현하였다. 우리는 300만 Gate FPGA와 비디오 코덱과 외부 SDRAM을 사용하였다. 전체 평균 Gate 수는 31,925 이고 최대 주파수는 75MHz이다.



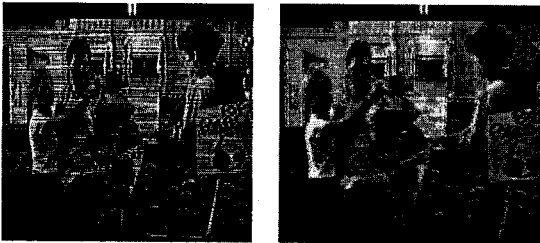
(a) 입체영상 보드 (b) 입체영상 시스템

(그림 5) 입체영상 변환기 보드와 입체영상 시스템

5. 검증

그림 5의 (b)와 같이 제안한 변환기를 검증하기 위해 SD급 TV와 LCD 셔터 안경과 콘트롤러와 VTR을 사용하였다.

그림 6의 (a)는 기존 TV의 2차원 영상이고 그림 6의 (b)는 입체 변환기를 사용하여 제작된 입체영상이다. 그림에서 보는 것과 같이 일반 2차원 영상으로부터 제작된 입체영상을 원근정보를 가지고 있는 입체 영상으로 변환되었다.



(a) 2차원 영상 (b) 입체영상

(그림 6) 2차원 영상과 입체영상 비교

5. 결론

이 논문에서는 실시간 입체영상 변환기에 대해서 서술하였다. 이 입체 영상 변환기는 물체의 방향과 속도에 상관없이 실감나는 입체 영상을 만들 수 있다.

입체 영상 변환기는 실시간 입체영상 변환에 적합

하며 또한 VTR, TV, cable TV, CD, DVD 에 적용 가능하며 다양한 해상도를 지원 할 수 있다.

참고문헌

- [1] Lipton, et al., Stereoscopic Television System with Field Storage for Sequential Display of Right and Left Images, U.S. Patent No. 4,562,463, 1985.
- [2] T. Okino, et al., "New Television with 2D/3D Image Conversion Technologies," SPIE Photonic West, vol. 2653, pp. 96-103, 1995.
- [3] H. Murata, et al., "Conversion of Two-Dimensional Image to Three Dimensions," SID '95 DIGEST, pp. 859-862, 1995.
- [4] Y. Matsumoto, et al., "Conversion System of Monocular Image Sequence to Stereo using Motion Parallax," SPIE Photonic West, vol. 3012, pp. 108-115, 1997.
- [5] B. J. Garcia, "Approaches to Stereoscopic Video Based on Spatio-Temporal Interpolation," SPIE Photonic West, vol. 2635, pp. 85-95, San Jose, 1990.