

# 운동 객체 검출 방식의 실시간 입체 영상 변환 장치 구현

정재성\*, 조화현\*, 윤종호\*, 최명렬\*  
\*한양대학교 전자전기 제어계측공학과

e-mail: starsung@asic.hanyang.ac.kr

## Implementation of a Real-Time Stereoscopic Image Converter Using Detection Method of Moving Object

Jae-Sung Jung\*, Hwa-Hyun Cho\*,  
Jong-Ho Yoon\*, Myung-Ryul Choi\*  
\*Dept of EECI, Hanyang University

### 요 약

본 논문에서는 영상 내 객체들의 운동시차를 이용하여 실시간으로 2차원 영상을 입체 영상으로 변환하는 알고리즘을 제안하였다. 제안한 입체 영상 변환 알고리즘은 영상 내 객체 분할, 운동시차를 이용한 운동 객체 검출, 시차 처리를 이용하여 입체 영상으로 변환한다. 입체 영상 변환 알고리즘의 성능 평가를 위해 좌안과 우안 영상의 절대 차이 영상을 이용하여 기존의 MTD 방식과 시물레이션 결과를 비교, 분석하였다. 제안한 알고리즘은 영상의 운동 방향 및 속도에 상관없이 다양한 영상원에 대해 실시간 입체 영상 변환이 가능하다. 입체 영상 변환 장치는 VHDL로 설계하였으며, FPGA 테스트 보드 구현을 통해 성능 및 기능을 검증하였다.

### 1. 서론

3차원 입체 영상 기술은 인간이 마치 영상이 제작되고 있는 장소에 있는 것 같은 생동감 및 현실감을 느낄 수 있게 만드는 기술이다. 현대 사회가 점차 발전해 감에 따라 세계 각국은 ‘보고 듣는’ 멀티미디어형 서비스로부터 3차원 정보중심의 ‘보다 자연스럽고 실감있게 보고 즐길 수 있는’ 실감형 3차원 입체 멀티미디어 서비스로 전환을 준비하고 있다. 입체 영상 기술은 정보통신·방송·의료·교육·항공·군사·게임·애니메이션·가상현실 등 응용분야가 매우 다양한 차세대 입체 멀티미디어 핵심 기반기술이다.

3차원 영상은 양안을 통해 보이는 서로 다른 영상을 인식함으로써 인간의 뇌에서 입체 영상으로 합성된다. 하지만, 2차원 영상의 경우에는 양안이 동일한 영상을 보게 되므로 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. 본 논문에서는 이러한 사전 작업 없이 기존의 2차원 영상을 실시간으로 3차원 영상으로 변환하는 입체 영상 변환 방법을 제안, 구현 및 검증하였다. 본 논문의 2장에서는 입체 영상의 원리, 3장에서는 기존의 입체 영상 변환 방법에 대하여 설명하였다. 4장에서는 제안한 알고리즘을 논하였으며 5장에서는 성능 평가, 6장에서는 구현한 입체 영상 변환 장치, 7장에서는 결론에 대하여 논하였다.

### 2. 입체 영상의 원리

#### 2.1 입체시의 원리

인간의 두 눈은 일반적으로 약 6cm 정도 서로 떨어져 있다. 그래서 그림 1에서와 같이 사물을 볼 때 양쪽 눈은 각각 약간 다르게 사물을 인식하게 되며, 각각의 사물들이 어느 정도 떨어져 있는지에 대해 파악할 수 있게 된다. 이를 양안시차라고 하는데, 이 차이의 느낌이 뇌에서 해석되어 입체감을 가지는 상으로 지각되어진다[1].

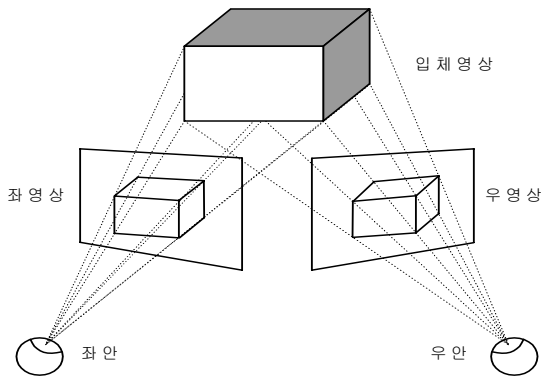


그림 1. 입체시의 원리

## 2.2 Parallax

Parallax는 그림 2와 같이 negative parallax, zero parallax, positive parallax의 3 종류가 있다. Negative parallax는 영상이 디스플레이 되는 평면보다 앞에서 양안의 초점이 모아져서 깊이가 평면보다 앞에 있게 되어 물체가 가깝다고 느껴진다. Zero parallax는 양안의 초점이 영상이 디스플레이 되는 평면에 모인 상태를 나타내며 이때의 깊이는 눈과 평면간의 거리이다. Positive parallax는 영상이 디스플레이 되는 평면보다 먼 뒤쪽에서 양안의 초점이 모아진 상태를 말하고, 깊이는 눈과 평면간의 거리 보다 멀게 느끼게 된다[2].

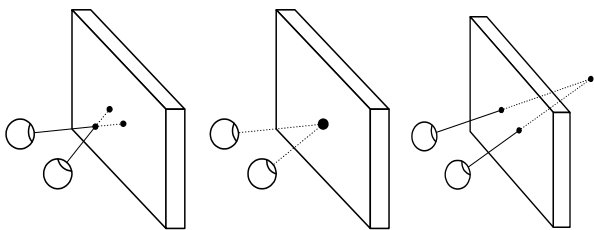


그림 2. Parallax

## 3. 기존의 입체 변환 방법

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. 그림 3과 같이 영상에서 물체가 우측으로 운동하고 있고, 카메라는 정지하고 있는 상태일 때, 현재 (N)번째 영상을 좌영상으로 하고, 지연 영상 중에서 (N-2)번째 영상을 우영상으로 구성하여 입체 영상을 만든 후 양안에 디스플레이하면 물체는 모니터 앞쪽으로 튀어나오듯이 보여지고, 배경은 모니터에 디스플레이되어

3차원 입체감을 느끼게 해준다[3]. 이 방식은 그림3과 같이 움직이는 물체가 비교적 저속의 정수평 운동일 때는 입체 효과가 좋지만, 만일 움직이는 물체의 방향이 수평이 아닌 비수평 운동일 경우 움직이는 물체는 하나의 상으로 융합하지 못하고 이중상으로 보여, 입체 효과를 얻을 수 없다. 또한, 움직이는 물체의 속도에 따라서 지연 영상 중 어떤 영상을 선택하는 문제가 발생한다.

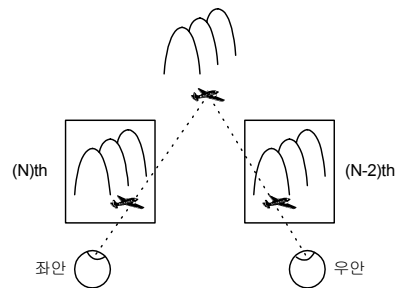
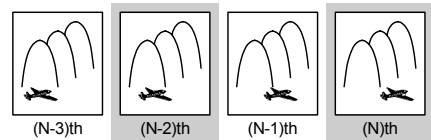


그림 3. MTD 방식의 원리

## 4. 제안한 입체 영상 변환 방법

본 논문에서 제안한 입체 영상 변환 방법은 2차원 영상에서 객체를 분할하고, 운동 시차를 이용하여 영상 내 운동 물체의 운동 방향 및 속도와는 상관없이 서로 다른 깊이감을 갖는 입체 영상을 실시간으로 제공한다.

### 4.1 휘도 영상 변환 및 객체 분할

2차원 입력 영상의 형태 정보를 추출하기 위해 컬러 신호를 휘도(Y) 신호로 변환한다. 휘도는 양안에 의해 인지되는 빛의 양에 비례하기 때문에, 객체의 형태 정보를 모두 포함하고 있다. 효과적인 입체 영상 변환을 위해 객체를 분할한다. 동일한 객체는 유사한 형태 정보를 가지고 있으므로, 유사한 범위의 휘도 계조로 이루어져있다고 가정할 수 있으며, 영상을 다단계의 휘도 계조로 분할함으로써 객체를 분할할 수 있다. 또한 동일 객체내에서 발생할 수 있는 오류를 방지하기 위해 마스킹 처리를 수행한다. 주변 픽셀들과의 상관성을 고려하여 객체 분할에서 발생하는 노이즈를 제거한다.

## 4.2 객체 분할 및 마스크

실시간 처리를 위하여 현재 및 이전 영상의 해상도보다 작은 해상도의 표본 영상을 등간격으로 취득한다. 등간격으로 샘플링된 표본 영상은 원영상의 형상 정보 및 휘도 분포 특성이 동일하다. 즉, 표본 영상에 대한 히스토그램의 평균 및 표준편차는 원영상에 대한 히스토그램의 평균 및 표준편차가 동일하기 때문에 운동 시차를 실시간으로 계산하기 위해서 표본 영상을 이용하는데 큰 무리가 없다.

## 4.3 운동 객체 검출

분할된 객체 기반의 현재 및 이전 표본 영상에서 움직임이 있는 객체를 검출한다. 운동 객체의 검출 방법은 현재 ( $Y_{(N)}^{\#}$ )와 이전 ( $Y_{(N-1)}^{\#}$ ) 영상의 픽셀 차의 절대값 ( $D_{Pel}$ )을 구하여 객체별 ( $D_{Pel}$ )의 대푯값을 결정하고, 전체 영상에서 각 객체들의 움직임 정도를 판별하여 정지 및 움직임 객체를 구분한다.

## 4.4 깊이 지도(Depth Map) 생성

표본 영상에서 결정된 정지 및 운동 객체에 대한 정보를 바탕으로 원영상 해상도에 대한 깊이 지도를 생성한다. 표본 영상에서의 한 픽셀에 대해 등간격 샘플링 범위내에서 동일한 휘도 계조의 픽셀에 대해 동일한 깊이값을 설정한다. 정지 픽셀은 일반적으로 배경을 이루는 픽셀로 상대적으로 먼 거리에 위치해 있는 것으로 가정하여 깊이값을 크게하고, 반대로 움직임 픽셀은 상대적으로 가까운 위치에 있는 것으로 가정하여 깊이값을 작게 설정한다.

## 4.5 시차 처리

배경 및 운동 객체의 시차 처리는 모두 양시차 처리를 행하여 움직이는 물체가 화면 안쪽에 위치하고 배경 또한 움직이는 물체보다 더 뒤쪽으로 화면 안쪽에 위치하도록 하였다. 배경과 운동 물체의 시차 차이를 영상 내에서 3픽셀로 한정했으며, 시차 처리에 의한 교합 문제는 ZOI(Zero Order Interpolation) 보간 알고리즘을 사용하여 해결하였다.

## 5. 성능 평가

제안한 알고리즘의 성능 평가를 위해 정원 영상과 탁구 영상에 대해 기존의 대표적인 MTD 방식과 시뮬레이션 결과를 비교, 분석하였다. 성능 평가의 효율성을 위해 입체 변환 영상의 좌/우 영상의 절대 차이값을 구하여, 변환 영상의 깊이감이 적절히 적용되었는지 평가하였다. 그림 4에서 검게 나타나는 부분이 실제 변환 영상에서 깊이감이 나타나는 부분이다.

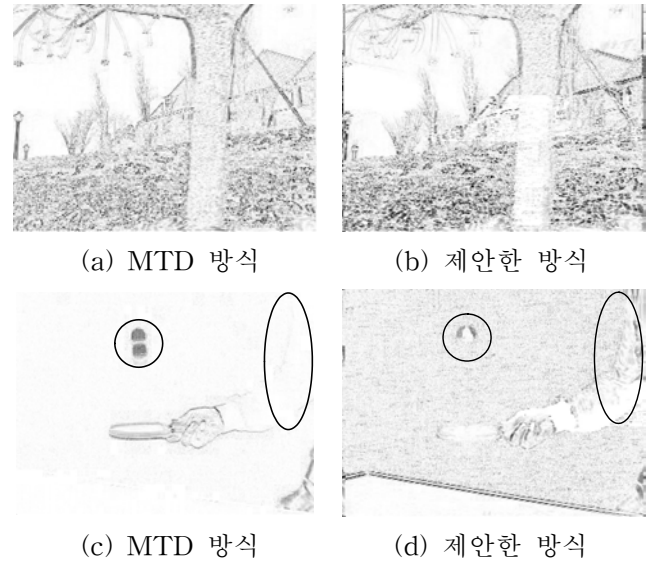


그림 4. 시뮬레이션 결과

정수평 운동인 정원 영상의 경우 그림 4의 (a),(b)에서 큰 차이를 느낄 수 없지만, 수직 방향으로 속도가 큰 탁구 영상의 경우 (c),(d)에서 알 수 있듯이, MTD 방식은 탁구공에서 이중상이 발생되며 팔 부분의 깊이감이 제대로 적용되지 못한 반면, 제안한 알고리즘은 영상의 운동 속도와 방향에 상관없이 적절한 깊이감이 적용됨을 알 수 있다.

## 6. 입체 영상 변환 장치 구현

입체 영상 변환 장치는 그림 5와 같이 비디오 디코더, SDRAM, FPGA와 비디오 인코더로 구성하였다. 비디오 디코더는 TV 또는 VTR, DVD에서 들어오는 아날로그 입력 CVBS(Composite Video Broadcast Signal) 또는 S-video(Y/C)를 입력으로 받아 컬러 신호로 변환해 준다. SDRAM은 비디오 디코더에 의해 생성된 컬러 신호의 영상 데이터를 저장한다. FPGA는 제안한 알고리즘을 처리한다. 비디오 인코더는 변환된 입체 영상 데이터를 NTSC, PAL의 CVBS 또는 S-video로 변환해 준다.

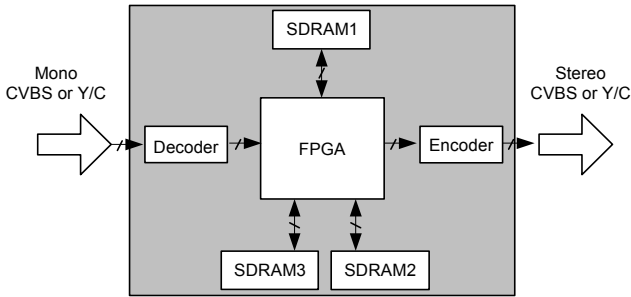


그림 5. 입체 영상 변환 장치 구성도

그림 6은 입체 영상 변환 장치의 타이밍도이다. 각 SDRAM은 비디오 디코더로부터의 Even과 Odd 필드 영상 데이터를 저장하고, 다음 4 필드 주기 동안 저장한 데이터를 내보낸다. FPGA는 제안한 알고리즘 처리를 수행하여 각 필드의 영상 데이터를 객체 분할한 후 이전 영상 데이터와의 연산을 통해 깊이 지도를 생성하고, 다음 필드 주기에 SDRAM으로부터 출력되는 이전 영상 데이터를 parallax 처리하여 내보낸다. 비디오 인코더는 최종 CVBS 또는 Y/C 신호로 변환하여 영상 데이터를 출력한다.

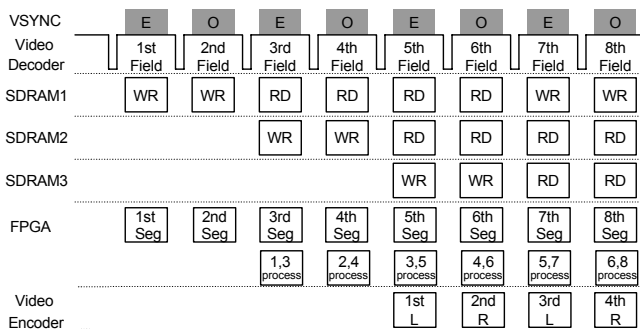


그림 6. 입체 영상 변환 장치 타이밍도

그림 7은 입체 영상 변환 장치의 블록도이다. 각 블록은 제안한 알고리즘의 각 단계를 수행한다.

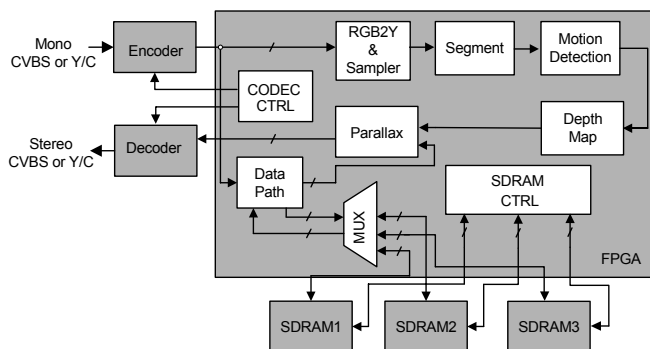


그림 7. 입체 영상 변환 장치 블록도

입체 영상 변환 장치는 VHDL을 이용하여 제안한 입체 영상 변환 기법을 설계하고, 그림 8과 같이 구현하였다. Xilinx FPGA XCV300E를 타겟으로 구현한 결과 전체 Gate 수는 약 35,000 게이트이고, 최대 동작 주파수는 75MHz이다.



그림 8. 입체영상 변환 장치와 데모 시스템

### 7. 결론

본 논문에서는 영상의 운동 방향이나 속도에 상관없이 실시간으로 입체 영상 변환을 수행하는 알고리즘을 제안하고, 설계 및 구현하였다. 입체 영상 변환 장치는 실시간 입체영상 변환에 적합하며 TV, VTR, DVD 등에 적용 가능하고, 다양한 해상도를 지원 할 수 있다. 추후 개선된 객체 분할 및 시차 처리 알고리즘과 다양한 형식의 입출력, 해상도를 지원할 수 있는 하드웨어 구현 방법을 연구할 예정이다.

### 참고문헌

- [1] T. Okino, et al., "New Television with 2D/3D Image Conversion Technologies," SPIE Photonic West, vol. 2653, pp. 96-103, 1995.
- [2] S. Nagada, How to Reinforce Perception of Depth in Single Two-Dimensional Pictures, Tyler & Francis, 1991.
- [3] H. Murata, et al., "Conversion of Two-Dimensional Image to Three Dimensions," SID '95 DIGEST, pp. 859-862, 1995.
- [4] B. J. Garcia, "Approaches to Stereoscopic Video Based on Spatio-Temporal Interpolation," SPIE Photonic West, vol. 2635, pp. 85-95, San Jose, 1990.
- [5] 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.
- [6] A Murat Tekalp, Digital Video Processing, Prentice Hall, 1995.