

HD 해상도 스테레오 영상 카메라 구현과 거리 인식 응용

한병완^{1*}, 임성준²

¹동원대학교 컴퓨터애니메이션과, ²동원대학교 인터넷정보과

Development of HD Resolution Stereoscopic Camera and Apparatus for Recognizing Depth of Object

Byung-Wan Han^{1*} and Sung-Jun Lim²

¹Division of Computer Animation, Tongwon University

²Division of Internet Information, Tongwon University

요 약 3D 입체 영상을 만들어 내기 위해서는 사람의 오른쪽 눈과 왼쪽 눈의 역할을 하는 양안 카메라가 필요하다. 즉, 왼쪽과 오른쪽 두 대의 카메라로부터 입력되는 영상을 합치는 3차원 영상 처리를 통하여 입체 영상으로 만들어내는 과정을 거치게 된다. 본 논문에서는 HD 해상도를 가지는 스테레오 영상 카메라 구현을 위하여 두 대의 고해상도 줌 카메라로부터 입력되는 영상데이터를 스테레오 영상 포맷으로 변환하는 알고리즘을 FPGA를 통하여 실시간으로 동작하도록 구현하였다. 또한 생성된 스테레오 영상으로부터 좌우 영상간의 물체의 위치 차이를 계산하여 구하는 방법을 제시하였다.

Abstract Two cameras which function like human eyes, are needed to make 3D stereoscopic image. That is, stereoscopic image is made via 3 dimensional image processing for combining two images from left and right camera. In this paper two high resolution zoom cameras are used to make HD resolution stereoscopic camera. And the algorithm which convert to stereoscopic image from HD resolution zoom camera image, is implemented using FPGA for real-time operation. The algorithm which measure the depth of object between left and right image is proposed.

Key Words : HD resolution, Stereoscopic Camera, Zoom Lens

1. 서론

스테레오 영상 기술은 사용자에게 영상 현실감을 극대화 시킬 수 있는 미래 영상 기기의 핵심 기술로서 입체 영상 실험 방송 및 영화 제작과 같은 상용화를 위한 시도가 다양하게 진행되고 있다. 스테레오 영상 기기 개발 동향으로 스테레오 영상촬영, 영상처리, 영상전송, 영상 디스플레이 기기 등의 분야에서 활발한 연구를 진행하고 있으며 새로운 기술 개발 및 표준화를 진행하고 있다.[1]

입체감을 느낄 수 있는 생리적 요인으로 양안 시차의 원리를 이용한다. 좌우 눈에 보이는 상은 눈에 가까울수록 차이가 많고 멀어질수록 차이가 적어진다. 두 눈은 모

양체의 근육을 긴장 혹은 이완시켜 초점을 조절하고 있다. 물체를 볼 때 가까우면 가까울수록 안구를 안쪽으로 회전하여야 한다. 이때 근육의 긴장이 거리감을 주게 된다. 사람은 좌우 양안에 투영되는 상의 차이에 의해 3차원 공간 속의 장면 및 사물에 대해 입체감을 인식하게 된다.

스테레오 영상 기술은 서로 다른 위치에서 영상 카메라로 촬영한 2개의 영상을 이용하여 영상에 입체감을 부여하게 된다. 입체 영상을 이용하면 영상에 있는 물체의 거리를 측정할 수 있다. 스테레오 영상 카메라에서 물체의 거리를 측정하는 방법은 두 개의 카메라간의 거리와 렌즈 화각을 이용하여 좌우 영상간의 물체의 위치 차이를 계산하여 구할 수 있다. 스테레오 영상 카메라에서 피

*Corresponding Author : Byung-Wan Han (Tongwon University)

Tel: +82-10-2217-0985 email: bwahan@tw.ac.kr

Received November 27, 2012 Revised (1st December 17, 2012, 2nd December 31, 2012) Accepted January 10, 2013

사체의 거리를 측정할 수 있다면 피사체의 이동경로 추적, 최적 영상크기 설정, 카메라간의 연계 동작 등 다양한 응용 분야에 이용할 수 있다.

사람의 좌우 눈은 서로 다른 2차원 영상을 보게 되며 각각의 영상이 망막을 통하여 뇌로 전달되면 이를 서로 융합하여 영상의 깊이감과 실제감을 만들게 된다. 이러한 원리를 3D 영상 카메라 설계에 적용하면 두 대의 카메라를 두 눈의 간격과 유사하게 약65mm 간격을 두고 위치시키며 각각 영상을 촬영하게 된다. 두 개의 카메라를 위치시키는 방법에 따라 평행식, 복합식 등의 다양한 형태로 배치하게 된다.[2]

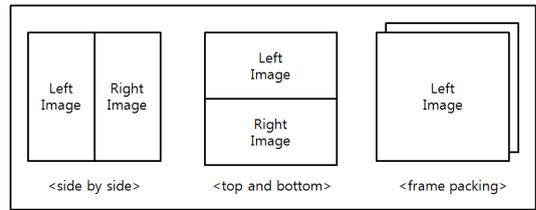
본 연구에서는 줌렌즈의 설치와 기구 설계의 편의를 위하여 평행식 설치 방안을 채택하였으며 HD 해상도와 고배율 줌 렌즈를 채용하여 스테레오 영상 카메라를 구현하고, 이를 움직이는 물체에 대하여 실시간 거리 측정에 응용 방안에 대하여 연구하였다.

2. 본론

2.1 스테레오 영상 데이터 포맷

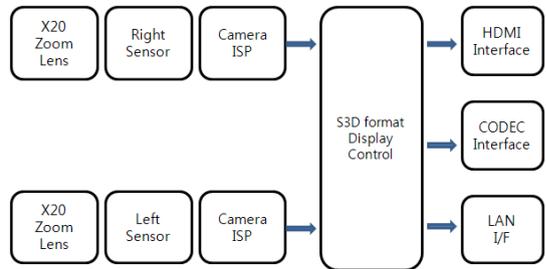
스테레오 영상을 만들어 내기 위해서 사용하는 영상 포맷은 side by side 방식, top and bottom 방식과 frame packing 방식이 주로 사용된다. side by side 방식은 영상을 전송할 때 전송 화면의 한 라인을 구성할 때, 반은 왼쪽 카메라의 영상을 나머지 반은 오른쪽 카메라의 영상을 보내는 방식이다. top and bottom 방식은 영상을 전송할 때 전송 화면의 위쪽부터 반은 왼쪽 카메라의 영상을 나머지 반은 오른쪽 카메라의 영상을 전송하는 방식이다. 이에 반하여 frame packing 방식은 한 프레임은 왼쪽 카메라의 영상을 다음 한 프레임은 오른쪽 카메라의 영상을 보내는 방식이다.[3]

side by side 방식은 한 라인에 왼쪽 카메라의 한 라인과 오른쪽 카메라의 한 라인 데이터를 보내야 하기 때문에 일반적으로 샘플링 방식을 사용하여 데이터를 반으로 줄여서 전송한다. 따라서 수평 해상도가 낮아지는 단점을 가지고 있다. top and bottom 방식은 수평 해상도의 저하는 사람의 눈에 민감하게 반응하기 때문에 상대적으로 덜 민감한 수직 해상도를 반으로 줄여서 전송하는 방식이다. 이에 반하여 frame packing 방식은 매 프레임마다 왼쪽과 오른쪽 카메라로부터 받은 하나의 완전한 영상을 번갈아 보내는 방식으로 해상도의 저하가 없어서 최근에 많이 사용하는 방식이다. Fig. 1에 앞에서 설명한 세 가지 전송 방식의 영상 포맷을 나타내었다.



[Fig. 1] Stereo Video Data Forat

2.2 스테레오 영상 카메라의 구성

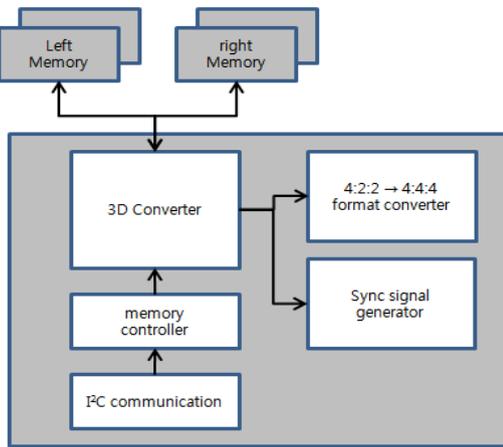


[Fig. 2] Block Diagram of Stereo Camera

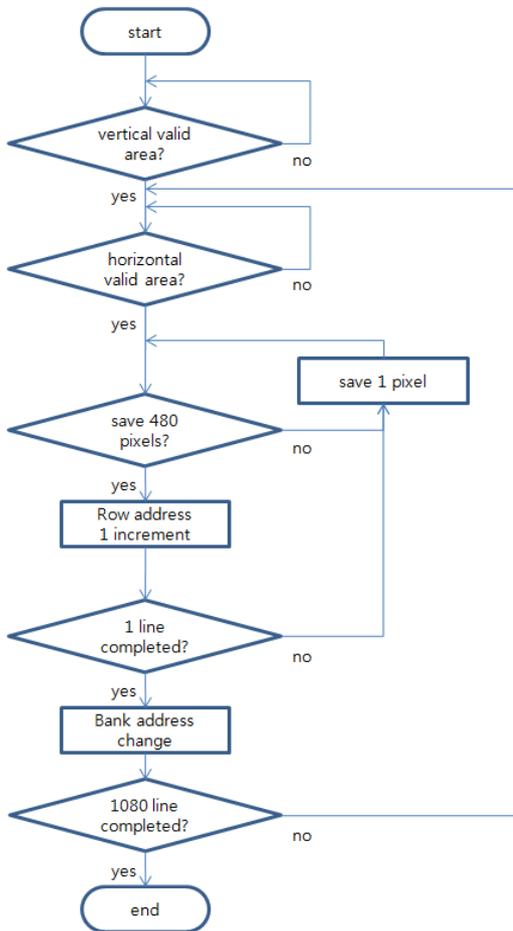
두 개의 줌 영상 카메라로부터 입력된 영상을 3D 영상 규격을 만들게 된다. 3D 영상 규격을 지원하는 영상을 만들기 위하여 Xilinx사의 Spartan-6 FPGA를 사용하였으며 동작 주파수는 74.25MHz를 사용하였다. 본 연구에서 개발된 FPGA에서 지원하는 3D 영상규격은 Side-by-Side, Top-Bottom 이며 Full HD 해상도의 3D 영상에 대하여 초당 30 프레임의 실시간 동작이 가능하도록 설계하였다. [Fig. 2]는 본 논문에서 구현한 스테레오 카메라의 구성도이다.

2.3 스테레오 영상 데이터 변환

입력된 두 개의 카메라 영상으로부터 스테레오 영상 데이터를 만들기 위해 구현한 FPGA의 블록도는 Fig. 3과 같다. I2C 통신부는 3D 변환 모드를 설정하기 위하여 프로세서와 I2C 통신을 하여 side by side, top and bottom, left only, right only 중 하나의 모드를 설정한다. 메모리 제어부와 3D 변환기에서는 I2C 통신부를 통하여 입력된 3D 포맷에 따라서 right 용 메모리 2개와 left용 메모리 2개를 제어하여 원하는 데이터 포맷으로 변환한다. 이때 사용한 메모리는 128M SDRAM으로 16비트 메모리를 사용하였다. 4:2:2 => 4:4:4 포맷 변환기에서는 4:2:2 형태로 입력된 영상 데이터를 4:4:4 형태의 영상 데이터로 변환해주는 역할을 해주며, 동기신호 발생기에서 HSync, VSync와 DE 신호를 만들어 준다.



[Fig. 3] Block Diagram of Stereo Format Converter



[Fig. 4] Flowchart of memory writing operation

Left Memory와 Right Memory를 두 개씩 사용한 이유

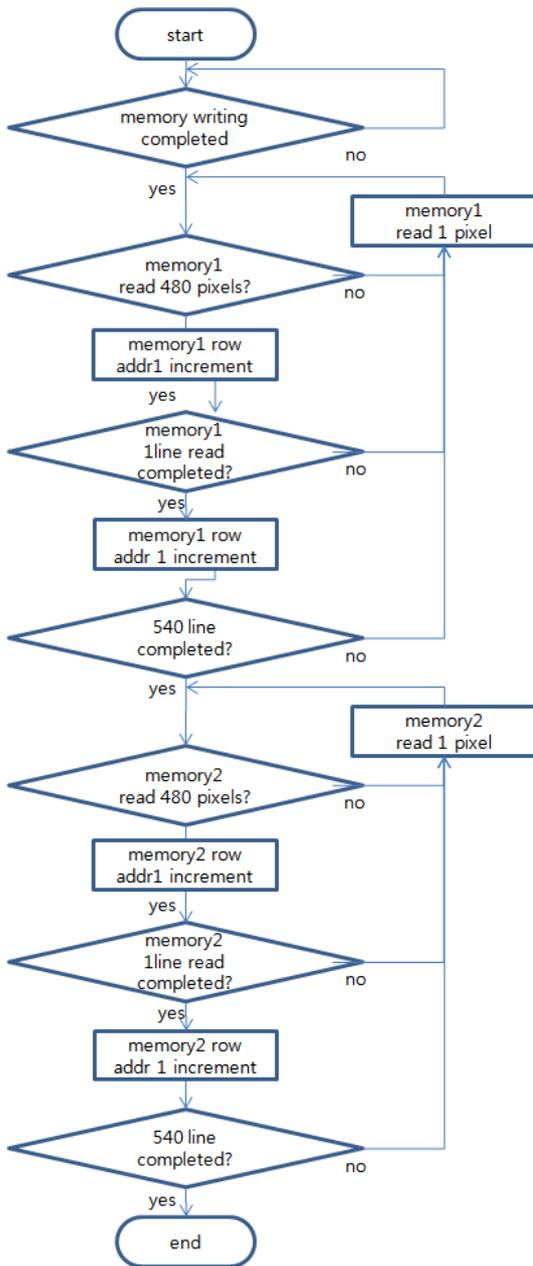
는 계속해서 실시간으로 입력되는 영상 데이터를 실시간으로 출력해야 하기 때문에 하나의 메모리를 이용하여 제어하는 경우에는 입력률 시간의 부족으로 SDRAM 같이 입력과 출력을 공유하는 메모리를 사용하면 문제를 발생시키게 된다. 따라서 본 논문에서는 하나의 메모리에 쓰기 동작을 할 때에는 다른 메모리에서 데이터를 읽고, 반대의 경우에는 반대편 메모리에 쓰고, 나머지 메모리에서 데이터를 읽는 작업을 한다.

사용한 메모리의 용량은 8Mb x 16로 한 row 당 512개의 데이터를 저장할 수 있다. 그러나 본 연구에서는 한 row 당 480 픽셀 데이터만 저장하였다. 그 이유는 Full HD 카메라의 경우 하나의 라인이 1920 픽셀로 구성되어 있기 때문에 어드레스 제어를 간단히 하기 위해서이다. 즉 메모리 데이터 저장 시에 유효화소 구간이 시작되면 480 픽셀 데이터를 하나의 row에 저장한 후, 다음 row에 저장하는 방식을 사용하였기 때문에 카메라 1라인을 저장하기 위해서는 4번의 row 주소 증가가 필요하였다. 그리고 top and bottom 모드에서는 두 개의 라인 중 하나의 라인 데이터는 버리고 나머지 하나의 라인 데이터만 사용하기 때문에 저장 시에 매 카메라 라인마다 뱅크 주소를 0과 1을 번갈아 사용하였다. Fig. 4에 메모리 저장 동작에 관련된 순서도를 나타내었다.

top and bottom 모드가 선택된 경우에는 우선 Left Memory에서 1920 픽셀의 데이터를 540 라인 읽어 들인 후, 다시 Right Memory에서 1920 픽셀의 데이터를 540 라인 읽어 들여 사용한다.

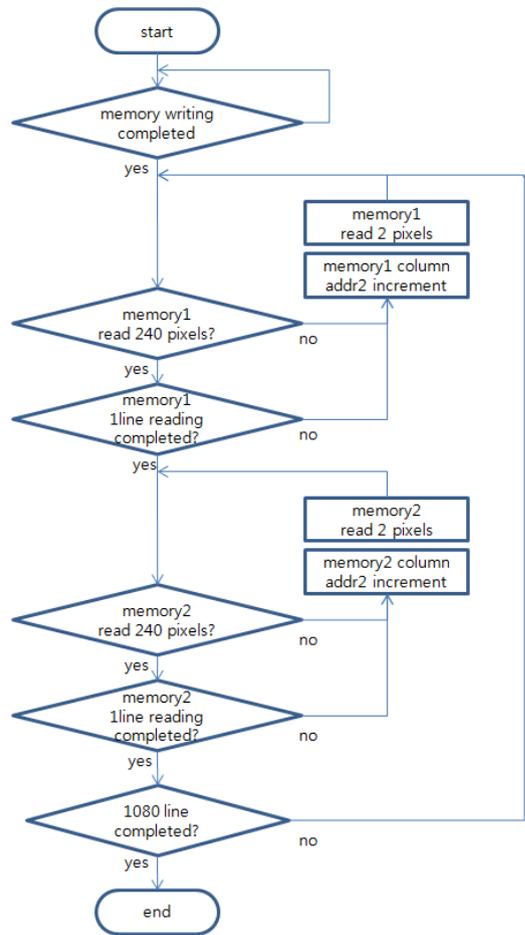
이 경우 카메라 영상의 왜곡을 막기 위하여 메모리1과 메모리2의 픽셀 데이터를 읽을 때 한 라인 데이터를 읽고 다음 라인의 데이터는 읽지 않는 방식을 사용한다, 메모리 데이터의 저장 시에 매 라인마다 뱅크 주소를 0과 1을 반복하여 사용하였기 때문에 top and bottom 모드의 읽는 순간에는 하나의 뱅크만 선택하여 읽으면 자동으로 540 라인만 선택하여 읽을 수가 있다. [Fig. 5]에 top and bottom 모드의 읽기 동작에 대한 순서도를 나타내었다.

side by side가 선택된 경우에 메모리 읽기 동작은 다음과 같다. 우선 메모리1에서 카메라의 1 HD가 시작되면 한 라인의 데이터를 읽은 후 나머지 시간에 메모리2의 한 라인 데이터를 읽어 들이는 방식으로 메모리 데이터를 읽는다. 이때 메모리1과 메모리2의 데이터는 각각 저장되어 있는 1920개의 픽셀 데이터를 읽는 것이 아니라 각각 960개의 픽셀데이터만 읽어 들여야 한다. 따라서 메모리를 읽을 때 4개의 이미지 데이터 중 세 번째와 네 번째 데이터만 읽고 나머지 데이터는 읽지 않으면서 메모리를 읽는 동작을 수행해야 한다. Fig. 6에 side by side 모드에서의 메모리 읽기 동작에 대한 순서도를 나타내었다.



[Fig. 5] Flowchart of memory reading operation at top and bottom mode

left only와 right only 모드는 왼쪽이나 오른쪽 하나의 카메라 영상 데이터 만 전송하는 방식이므로 Left Memory나 Right Memory 하나의 메모리만 사용하게 된다.



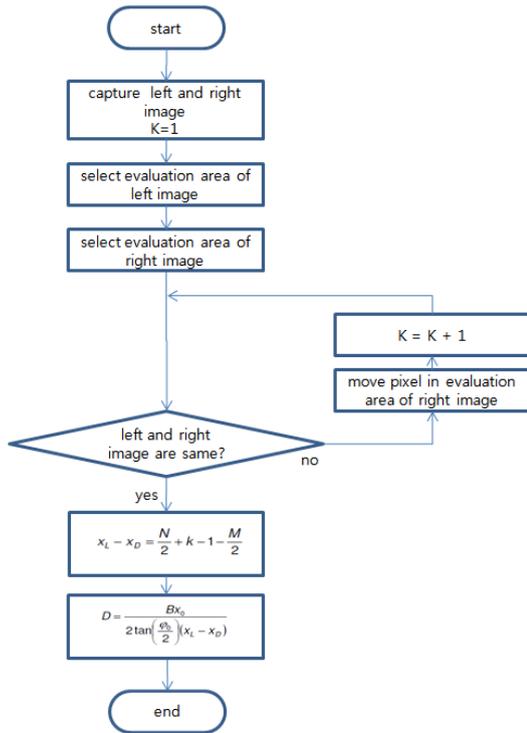
[Fig. 6] Flowchart of memory reading operation at side by side mode

2.4 스테레오 카메라를 이용한 거리계측

피사체의 거리를 측정하는 방법으로 수동적 방법과 능동적 방법이 있을 수 있다. 본 논문에서 채용한 방법은 수동적 방법이며 스테레오 카메라를 이용하는 방법을 채택하였다. 스테레오 카메라를 이용한 피사체의 거리 측정 은 측정 대상물과 양 카메라간의 상대적인 거리를 바탕으로 계산할 수 있다.[4-6]

피사체의 거리를 감지할 수 있는 스테레오 시각시스템 은 투시변환에 의해 3차원 공간을 2차원 공간으로 매핑 하는 스테레오 카메라 2대를 이용하여 스테레오 카메라 와 피사체간의 기하학적인 배치로부터 3차원 정보를 복원하는 원리이다. 피사체의 거리를 측정하는 방법은 2개의 카메라를 이용하여 좌측 영상과 우측 영상을 촬영하고 각 영상에 대하여 윤곽선을 검출하여 상호 매칭되는 영상 부분에 대한 분석 방법을 사용한다.

스테레오 카메라를 이용한 거리 측정에서 우선 고려하여야 할 것은 양 카메라의 기구적으로 직 및 수평적으로 정확하게 배치되어야 영상의 분석 오차를 줄일 수 있다. 또한, 움직이는 물체에 대한 거리 측정에서 동일 시점에 동기 되어 영상을 촬영하여야만 피사체의 이동에 의하여 발생하는 오차를 줄일 수 있다.



[Fig. 7] Flowchart of measuring depth of object

스테레오 카메라를 이용한 정확한 거리측정을 위하여 거리 측정 범위에 적합한 렌즈의 사양을 선정하는 것이 중요하다. 거리 측정에 주로 이용되는 렌즈 사양은 화각과 최대 해상도와 관련되어 있다.

스테레오 영상 카메라를 이용하여 피사체의 거리를 측정하기 위하여 FPGA 에서 구현된 실시간 흐름도를 [Fig. 7]에 나타내었다. 여기서, D는 계산된 거리값을 나타내며, N과 M 값은 좌우 카메라에서 영상 창의 크기를 나타내는 것으로 $M > N$ 이다. 또한, 좌측 카메라에서 N(수직) x N(수평) 크기의 영상 영역을 나타내며, 우측 카메라에서는 N(수직) x M(수평) 크기의 영상 영역으로 설정되었다. 좌우 카메라에서 촬영된 동일 피사체에 대한 영상에 대하여 좌측 영상 영역 N x N 영상을 우측 영상 영역 N x M 내에서 왼쪽에서 오른쪽으로 한 화소씩 차례로 이동하면서 좌우 카메라 영상의 일치여부를 판단하게

된다. 좌우 카메라 영상에서 가장 일치도가 높은 지점 k 는 가장 왼쪽으로부터 이동 화소 수를 나타내며, B 는 좌우 카메라가 설치된 간격, X_0 는 좌우 각 카메라의 최대 화소수를 나타낸다.^[7]

본 연구에서 사용된 줌렌즈와 화소수에 대한 정보를 바탕으로 스테레오 영상 카메라의 좌우 카메라에서 동일 물체에 대한 상대적인 차이에 대한 위치 값을 바탕으로 기하학적인 공식을 이용하여 피사체의 거리를 측정하는 방법을 흐름도 Fig. 7에 나타내었다.

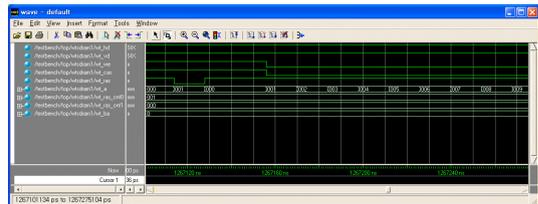
2.5 구현 및 검증

본 논문에서 3D 변환기를 설계하기 위해서 Xilinx FPGA인 Spartan 6를 사용하였으며, Verilog를 이용하여 코딩하였다. 또한 동작 주파수는 입력되는 데이터의 주파수와 일치하는 74.25MHz를 사용하였다. 메인 클럭은 외부에 오실레이터를 달아 사용하지 않고, LVDS receiver 를 통해 추출해 놓은 클럭을 메인 클럭으로 사용하였다.



[Fig. 8] Stereo Camera Board

논문에서 제시한 설계를 검증하기 위하여 ModelSim 을 이용하여 시뮬레이션을 하였다. Fig. 8은 본 논문을 위하여 구현된 스테레오 카메라 모습이다.

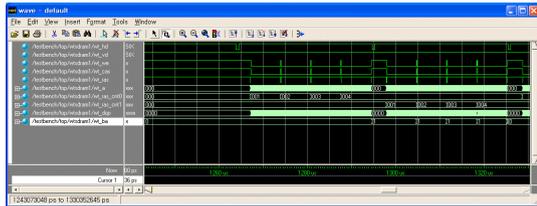


[Fig. 9] Simulation result of memory writing operation - column address

Fig. 9는 메모리 쓰기 동작에서의 column 주소의 변화를 보이고 있다. 메모리 쓰기 동작 시에는 side by side 모

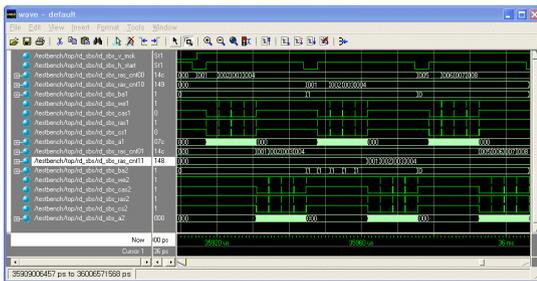
드나 top and bottom 모드에 상관없이 매 클릭마다 1씩 column 주소가 증가되는 것을 볼 수 있다.

Fig. 10은 메모리 쓰기 동작에 대한 시뮬레이션 결과를 나타내고 있다. 앞서 언급한 것과 같이 top and bottom 모드시의 메모리 읽기 동작을 간단히 하기 위하여 카메라 한 라인 마다 뱅크 주소가 0과 1로 바뀌는 것을 볼 수 있다.



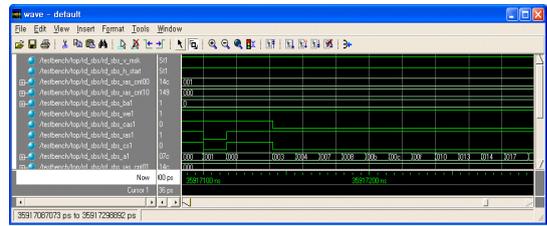
[Fig. 10] Simulation result1 of memory writing operation- row address and bank address

Fig. 11은 side by side 모드에서의 메모리 읽기 동작에 대한 시뮬레이션 결과를 나타내고 있다. 앞서 언급한 것과 같이 매 라인 마다 메모리에서 960 픽셀 데이터를 읽고 이어서 메모리2에서 960 픽셀 데이터를 읽어 들이는 것을 볼 수 있다.



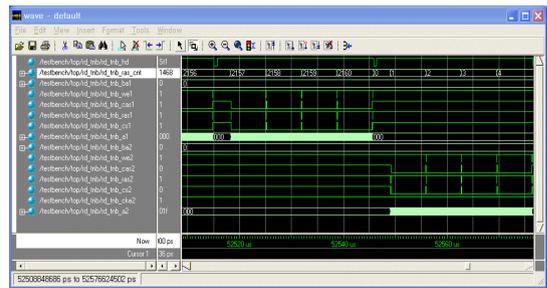
[Fig. 11] Simulation result of memory reading operation at side by side mode - row address

Fig. 12는 side by side 모드에서의 메모리 읽기 동작 시 column 주소의 변화를 보여주고 있다. 각 메모리에 저장된 1 라인의 데이터를 읽을 때 전체 라인을 읽는 것이 아니라 1/2 라인 데이터를 읽기 위하여 column 주소가 연속적으로 증가하는 것이 아니라 두 개의 픽셀데이터를 읽고 나머지 두 개의 픽셀데이터는 버리기 위하여 불연속적으로 증가하고 있는 것을 알 수 있다.



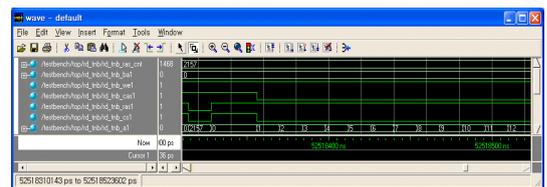
[Fig. 12] Simulation result of memory reading operation at side by side mode - column address

Fig. 13은 top and bottom 모드에서의 메모리 읽기 동작에 대한 시뮬레이션 결과를 나타내고 있다. 앞서 언급한 것과 같이 메모리에서 540 라인 데이터를 읽은 후 메모리2에서 나머지 540 라인 데이터를 읽고 있는 것을 볼 수 있다. 시뮬레이션 상에서 row 주소가 2160으로 보이는 것은 한 라인마다 4개의 row를 사용하였기 때문이다. 또한 앞서 설명한 것과 같이 뱅크 주소는 변하지 않고 있는 것을 볼 수 있다.



[Fig. 13] Simulation result of memory reading operation at top and bottom mode - row address and bank address

Fig. 14는 top and bottom 모드에서의 읽기 동작 시의 column 주소의 변화를 보여주고 있다. 이 경우 side by side 모드와는 달리 주소가 연속적으로 변하는 것을 볼 수 있다.



[Fig. 14] Simulation result of memory reading operation at top and bottom mode - column address

3. 결론

본 논문에서 스테레오 영상 포맷을 구현하기 위한 알고리즘과 이를 이용하여 거리 측정 응용에 대한 최적의 구현 방법에 대하여 논하였다. Xilinx FPGA Spartan6를 사용하여 실시간으로 동작 테스트를 하였으며, 동작주파수 74.25MHz와 같이 고속의 동작 조건에서 설계한 FIFO가 완전하게 동작하는 것을 확인하였다. 다양한 데이터 포맷을 지원하도록 설계하였으나 본 논문에서 일반적인 3D 영상 포맷을 지원하도록 설계하고 실시간 동작을 확인하였으나 FPGA 구현 시 내부 PLL을 이용하여 주파수 168.5MHz 까지 시스템이 동작하도록 설계하여 시뮬레이션 상에서 검증하였으므로 추가 연구에서는 Frame Packing 방식도 지원할 수 있을 것으로 예상된다. 또한 스테레오 영상을 만들기 위해서 사용하는 카메라의 줌 배율에 따라 화각이 달라지므로 다양한 줌 배율에 따른 거리 측정 알고리즘의 검증이 필요할 것이다.

References

[1] Y. S. Yu, "Technical Trend and Market Prospects of 3D Display", ETRI

[2] S. D. Park, "Photographing and Editing of 3D Stereoscopic Image", www.kocca.kr, 2010년

[3] T. G. Moon "Technical Trend and Outlook of 3D Display Technology", KEIT, PD ISSUE VOL 11-7, 2011.

[4] H. Walcher, "Position sensing - Angle and distance measurement for engineers, 2nd edition", Butterworth Heinemann, 1994.

[5] "Navigation 2, Radio Navigation, Revised edition", Oxford Aviation Training, 2004.

[6] J. Carnicelli, "Stereo vision: measuring object distance using pixel offset", <http://www.alexandria.nu>

[7] Jernej Mrovlje and Damir Vrancic, "Distance measuring based on stereoscopic pictures", 9th International PhD workshop and System and Control, 2008.

한 병 완(Byung-Wan Han)

[정회원]



- 1985년 2월 : 경북대학교 전자공학과 (공학사)
- 1987년 2월 : 연세대학교 전자공학과 (공학석사)
- 2006년 2월 : 연세대학교 정보대학원 (정보학박사)
- 1987년 1월 ~ 1998년 2월 : LG 전자 멀티미디어연구소
- 1998년 3월 ~ 현재 동원대학 컴퓨터애니메이션과 부교수

<관심분야>

멀티미디어네트워크, 3D 애니메이션

임 성 준(Sung-Jun Lim)

[정회원]



- 1989년 2월 : 고려대학교 전기공학과 (공학사)
- 1991년 2월 : 고려대학교 전기공학과 (공학석사)
- 2007년 2월 : 연세대학교 정보대학원 (정보학박사)
- 1991년 1월 ~ 1999년 2월 : LG 전자 멀티미디어연구소 선임
- 1999년 3월 ~ 현재 동원대학교 인터넷정보과 부교수

<관심분야>

멀티미디어 시스템, 웹 프로그램