

특집논문 (Special Paper)

방송공학회논문지 제19권 제5호, 2014년 9월 (JBE Vol. 19, No. 5, September 2014)

<http://dx.doi.org/10.5909/JBE.2014.19.5.616>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

FPGA와 GPU를 이용한 스테레오/다시점 변환 시스템

신 홍 창^{a)†}, 이 진 환^{a)}, 이 광 순^{a)}, 허 남 호^{a)}

Stereo-To-Multiview Conversion System Using FPGA and GPU Device

Hong-Chang Shin^{a)†}, Jinwhan Lee^{a)}, Gwangsoon Lee^{a)}, and Namho Hur^{a)}

요 약

본 논문에서는 FPGA와 GPU를 이용한 실시간 스테레오 다시점 변환 시스템을 소개한다. 해당 시스템은 이중의 연산장치를 이용하며 그에 따라 크게 두 부분으로 나뉜다. 첫 번째 부분은 변이 추출 부분으로서 실시간 계산을 위해 FPGA 기반으로 구현되었다. 기본적으로 DP(Dynamic programming) 기반의 스테레오 정합 방법을 통해 초기 변이 영상이 계산되며, 후처리를 통해 개선된다. 개선된 변이 영상은 USB3.0과 PCI-express를 통해 GPU 장치로 전송된다. 스테레오 입력 영상이 GPU 장치로도 전송되면, 변이 영상의 변이 값을 이용하여 중간 시점에서의 영상을 합성한다. 생성된 시점 영상들은 무안경 다시점 3차원 디스플레이의 특성에 맞게 하나의 영상으로 화소 또는 부분화소 단위로 재배치되는 시점 다중화 과정을 거쳐 최종적으로 4K 무안경 다시점 디스플레이에 실시간으로 재생된다. 스테레오 정합을 제외한 나머지 연산은 모두 GPU에서 병렬처리된다.

Abstract

In this paper, we introduce a real-time stereo-to-multiview conversion system using FPGA and GPU. The system is based on two different devices so that it consists of two major blocks. The first block is a disparity estimation block that is implemented on FPGA. In this block, each disparity map of stereoscopic video is estimated by DP(dynamic programming)-based stereo matching. And then the estimated disparity maps are refined by post-processing. The refined disparity map is transferred to the GPU device through USB 3.0 and PCI-express interfaces. Stereoscopic video is also transferred to the GPU device. These data are used to render arbitrary number of virtual views in next block. In the second block, disparity-based view interpolation is performed to generate virtual multi-view video. As a final step, all generated views have to be re-arranged into a single image at full resolution for presenting on the target autostereoscopic 3D display. All these steps of the second block are performed in parallel on the GPU device.

Keyword : 3D, autostereoscopic, stereo-to-multiview, FPGA, GPU

a) 한국전자통신연구원(Electronics and Telecommunications Research Institute)

† Corresponding Author : 신홍창 (Hong-Chang Shin)

E-mail: hcshin@etri.re.kr

Tel: +82-42-860-1725

※ 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신-방송 연구개발 사업의 일환으로 수행되었습니다. [14-000-02-001, 초고품질 콘텐츠 지원 UHD 실감방송/디지털시네마/사이버 융합서비스 기술 개발 사업]. 본 연구에 사용된 FPGA 기반 스테레오 영상 변이 추출 하드웨어 구현에 수고하신 ㈜비전메트리 연구원 여러분께 감사드립니다.

※ 이 논문의 연구결과 중 일부는 “3DSA 2014”에서 발표한 바 있음.

· Manuscript received July 21, 2014 Revised September 5, 2014 Accepted September 25, 2014

1. 서론

오토스테레오스코픽 디스플레이(autostereoscopic display)는 두 시점 이상의 영상을 한 화면에 동시에 재현함으로써 안경을 쓰지 않아도 3차원 영상을 볼 수 있게 한다. 최근의 콘텐츠 수요 동향에 따르면, 현 방송시스템에서 무안경 다시점 3차원 디스플레이에 맞는 콘텐츠를 제공하기 위해서는 스테레오 영상을 다시점 영상으로 자동으로 변환해주는 스테레오/다시점 변환 시스템이 필요하다.

스테레오 영상을 다시점 영상으로 자동 변환하기 위해서는 크게 두 가지 문제가 해결되어야 한다. 첫 번째로 스테레오 영상 각각에 대한 3차원 정보를 추출해야 하고, 두 번째는 추출된 3차원 정보를 이용하는 DIBR(depth image based rendering)과 같은 영상 합성 방법을 통해 중간 시점에서의 영상을 생성할 수 있어야 한다.

3차원 정보 추출은 입력된 영상을 토대로 장면 속의 객체와 배경의 깊이(depth) 정도를 추정하는 것으로, 일반적으로 영상의 모든 화소에 대해서 두 영상 간에 서로 대응하는 대응점(correspondences)들을 찾고 이를 대응점과의 거리(disparity)로 표현하는 변이 영상(disparity map)을 추출하는 스테레오 정합(stereo matching) 방법을 이용한다. 스테레오 정합은 컴퓨터 비전 분야에서 십수 년 동안 꾸준히 연구가 되어온 분야이며 다양한 방법들이 제시되어 왔지만 상충관계에 있는 다양한 환경에서의 강인성(robustness)과 실시간에 준하는 처리속도를 모두 일정 수준 이상으로 만족시키는 방법은 여전히 난제로 남아있다.

3차원 정보를 알 수 있는 깊이 영상(혹은 변이 영상)이 주어지면 그 정보를 토대로 영상 워핑(warping) 방법을 통해 가상 시점에서의 영상을 생성할 수 있다. 널리 쓰이는 기법인 DIBR 방법을 통해 좌우 영상에서 각각의 깊이 정보를 참조하며 가상 시점으로 화소들을 투영시킨 후에 이를 혼합(blending)하는 과정을 거친다. 이때, 참조 시점에서 보이지 않는 영역인 차폐영역(disocclusion)과 같이 정보가 없는 홀(hole) 영역이 생기게 되는데 이는 주변정보를 이용하여 보간하는 인페인팅(inpainting) 과정을 통해 채워지게 된다.

이러한 스테레오 정합 기반 영상 합성 기법과 관련된 오

토스테레오스코픽 디스플레이용 콘텐츠를 생성하기 위해 다시점(multiview) 영상을 고속으로 생성하는 방법에 대한 연구사례가 있었다. F. Zilly는 카메라 사이의 거리인 베이스라인(baseline)이 좁은 경우와 넓은 경우를 모두 촬영할 수 있도록 4개의 카메라가 장착된 다시점 영상 획득 카메라 시스템으로 MVD(multi-view plus depth)를 실시간 생성이 가능하게 하는 시스템을 개발하였다^[1]. 깊이 영상(depth map) 추출을 위해 라인 단위의 혼합형 재귀 정합 방법(L-HRM)을 사용하여 초기 깊이 영상을 구하였고, 교차형 양방향 중앙값 필터링(cross-bilateral median filter)을 이용하여 깊이 영상을 개선하였다^[2]. 교차형 양방향 중앙값 필터링 방법은 경계부분을 보존하는 특성을 지닌 양방향 필터링(bilateral filtering)을 기반으로 한 방법으로서 좌우 일관성 검사(left right consistency check)와 주변 화소와의 제로 평균 상호 상관성(zero-normalized cross-correlation)을 통해 계산된 가중치를 토대로 중앙값을 구하는 가중치 중앙값 필터링 방법(weighed median filtering)을 통해 깊이 정보를 개선한다.

여기서 한 가지 주목할 점은 차이 제곱값의 합(sum of squared differences)과 같이 지역적 스테레오 정합(local stereo matching) 방법 중에 비교적 간단한 정합 방법으로 초기 변이 영상을 먼저 구한 후에 경계 부분을 보존하는 유도영상 필터(guided image filtering)를 사용하여 변이 영상을 개선하는데 품질과 속도 면에서 상당히 좋은 결과를 보였다는 것이다.

이에 유사하게 본 논문에서도 스테레오 정합을 두 과정으로 분리하여 실시간성에 중점을 두는 지역적 스테레오 정합 방법을 이용하여 고속으로 초기 변이값을 구한 후 수직 방향 투표 과정과 양방향 필터링을 이용하여 변이정보를 개선한다. 초기 변이 영상을 고속으로 추출하기 위해 동적 프로그래밍(Dynamic programming) 기반으로 한 스테레오 정합 방법을 FPGA로 구현한 모듈을 사용한다.

한편, 깊이 정보를 이용하는 응용 애플리케이션이 점차 늘어남에 따라 FPGA를 기반으로 한 스테레오 정합 시스템을 개발하는 연구도 활발히 진행되어 왔다. S. Jin은 Census transform 기반의 스테레오 정합 시스템을 개발하였다. 영상 정렬, 좌우 일관성 검사, 유일성 검사 등의 전후처리까지

모두 포함하여 하나의 FPGA로 구현하였다. VGA(640x480) 해상도 영상에서 64 화소의 변이 범위 내로 230 fps까지 깊이 정보를 생성할 수 있다^[3]. L. Zhang은 FPGA 기반의 실시간 HD 스테레오 정합 시스템을 개발하였다. 하드웨어 구현이 가능한 Mini-Census transform과 Variable-Cross 스테레오 정합 알고리즘을 사용하여 FPGA로 구현하였고, 1024×768 해상도인 영상을 60fps까지 깊이 정보를 생성할 수 있다^[4].

FPGA와 GPU를 모두 사용하여 스테레오 영상으로부터 다시점 영상을 생성하는 연구 사례가 있었다^[5]. HD 크기의 스테레오스코픽 영상을 입력받아서 최종적으로 qHD(quarter-HD)크기의 9개 중간 영상을 20fps의 속도로 생성하였다. 라인 단위로 처리하는 알고리즘 특성상 변이 결과영상에서 수평 방향으로 띠 모양으로 변이가 어긋나는 잡음(streaking artifacts)이 발생되었으며, 품질과 실시간 성능 면에서 부족한 면이 있었다.

본 논문에서는 스테레오/다시점 변환 기법을 스테레오 정합 모듈과 다시점 영상 합성 모듈의 두 부분으로 나누고 두 모듈을 이중의 장치에서 독립적으로 구현을 하였다. 스테레오 정합과 다시점 영상 생성은 모두 계산 집약적인 기법이며 실시간 동작을 위해서는 고속화 기법이 필요하며, 실제로 고속화를 위해 FPGA 기반 하드웨어 구현 방법과 GPU 기반의 병렬처리 소프트웨어 구현 방법이 널리 쓰이고 있다. 각각은 모두 장단점을 가지는데 이와 관련된 연구 사례가 있었다^[6]. 실시간 기반의 광류 계산(optical flow), 스테레오 정합과 지역적 특징점 추출 기법 측면에서 FPGA와 GPU의 성능을 비교하였는데, GPU 기반의 구현 결과가 FPGA 기반 구현 결과 대비 대체적으로 우수한 결과를 보였다. 생산비용과 소비전력 측면에서의 고려를 제외하면 FPGA는 메모리 사용량이 적고 많은 수의 비트 연산이 필요한 알고리즘에 강점을 보이며, GPU는 영상 워핑(image warping) 혹은 렌더링과 같이 메모리 사용량이 많은 경우와 임의의 메모리 접근이 필요한 텍스처(texture) 처리와 같은 영상 처리 알고리즘에 강점을 보인다.

이에 따라 본 논문에서는 비트 연산이 많은 스테레오 정합 모듈은 FPGA 기반으로, 4K 해상도 이상의 메모리 접근이 필요한 다시점 영상 합성 모듈은 GPU 기반으로 하여,

각각을 이중의 장치(hybrid system)에서 구현하였으며, 스테레오 정합 모듈에서는 실시간성에 주로 초점을 두고 고속으로 초기 변이값을 구한 후에 수직방향의 투표 과정과 양방향 필터링 기법을 이용하여 변이 영상을 개선한다. 개선된 변이 영상은 영상 합성에 사용되며, 이후의 다시점 영상 합성 과정은 모두 GPU기반으로 구현되었다. 최종 결과 영상은 4K 해상도의 다시점 영상이 된다.

본 논문의 구성은 다음과 같다. 서론에 이어 2장에서는 제안된 시스템의 전체 구조에 대해서 설명을 하고, 3장에서는 FPGA 기반의 변이 추출과 블록 단위의 영상 합성 방법에 대해 설명한다. 4장에서는 실험 결과를 보이며 5장에서는 맺음말로 결론을 맺는다.

II. 시스템 구조

이번 장에서는 그림 1에 나타난 제안 시스템의 전체 구조를 설명한다. 기본적으로 영상 생성은 DIBR(Depth-image-based rendering) 방식을 사용하였으며, 입력되는 스테레오 영상으로부터 좌우 각각에 대해 변이를 추출한 뒤 변이를 이용하여 중간 시점에서의 영상을 합성한다.

본 시스템은 크게 두 블록을 구성되는데, 첫 번째는 FPGA 기반의 변이 추출 블록이며, 두 번째는 GPU 기반의 중간 영상 합성 블록이다.

3DTV 수신기가 HDMI 1.4를 통해 HD 해상도의 양안 영상을 수신한다. 입력된 영상의 규격은 UYVY422이며, 수신된 좌우 영상은 각각 크기 및 색상 변환 모듈에서 HD의 1/4 크기인 가로 960, 세로 540 해상도의 RGB 규격으로 변환되어 USB 3.0 인터페이스를 통해 변이 추출 모듈의 입력 영상으로 사용된다. 또한, 영상 합성 모듈로의 입력 크기에 맞게 변환된 RGB 영상이 중간 영상 생성 모듈의 입력 영상으로 사용된다. 이처럼 GPU로 양안 영상이 전달된 이후에는 모든 전후처리가 GPU에서 처리된다.

quarter-HD 크기인 좌우 RGB 영상이 FPGA 모듈에 전송되면 좌우 각각에 대한 변이 영상이 추출된 후에 다음 프레임이 입력되는 시점에서 변이 결과 영상이 출력되므로 1 프레임의 지연 시간(latency)이 발생한다. 즉, 현재 프레임

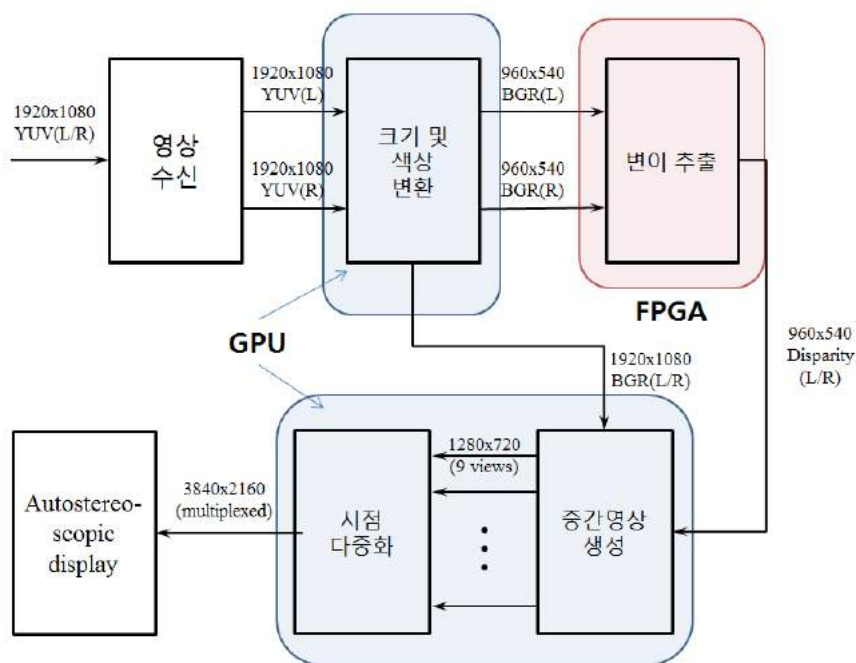


그림 1. 하드웨어 구조도

Fig. 1. The overall hardware architecture of the proposed system

의 변이 추출 결과가 다음 프레임의 입력 영상이 입력될 때 동시에 받아지므로 전체적으로 1 프레임 지연된 상태로 실시간 처리된다. FPGA 모듈에서 영상 합성 시 출력된 변이 영상은 그대로 쓰이지 않고 영상 합성의 입력 크기에

맞게 변이 영상을 업샘플링(upsampling)하는 과정을 거친다.

그림 2는 변이 추출 모듈의 시제품을 보여준다. 입력과 출력을 위해 USB 3.0 포트 두 개를 사용하며, USB 3.0 인

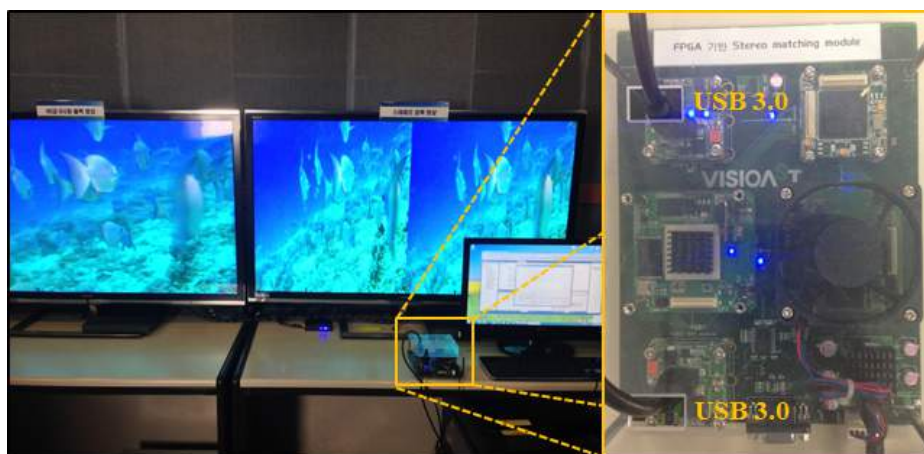


그림 2. 변이 추출 모듈 및 실시간 재생 시스템

Fig. 2. A prototype of the FPGA-based stereo matching module

터페이스를 통해 입출력되는 영상의 대역폭은 각각 600 Mbps, 300 Mbps이다. 영상 변환은 Spatran6 FPGA를 사용하였으며, 스테레오 정합과 변이 영상 개선은 둘 다 Kintex-7 FPGA를 사용하였다.

III. 알고리즘

1. FPGA 기반 변이 추출 모듈

실시간 스테레오 정합 방법으로서 Trellis 동적 프로그래밍(Dynamic programming) 기법을 사용하였다. 먼저 좌우 영상의 중앙에 위치한 중앙 변이맵(centered disparity map)을 구한 후, 좌우 영상으로 변환을 한다^[7]. 이 방법은 병렬화처리가 가능한 연산이며, 병렬처리를 할 경우에 $O(MN^2)$ 의 복잡도가 $O(MN)$ 으로 줄어든다. 라인 단위로 처리하는 알고리즘 특성상 변이 결과영상에서 수평 방향으로 띠 모양으로 변이가 어긋나는 잡음(streaking artifacts)이 발생되

는 문제가 있다. 따라서 후처리를 통해 개선되어야 한다. 변이 개선 후처리는 두 단계로 이루어진다^[8].

먼저 수직방향으로 투표 과정(vertical voting)을 거친다. 수평 단위로 발생하는 잡음을 수직방향의 정보를 더 포함하여 잡음을 완화하는 방법으로써, $d_{\max} + 1(d_{\max}$ 는 최대 변이 레벨)만큼의 히스토그램을 구성하고 기준 화소의 상하 수직 방향에 위치한 N개의 변이에 대해서 화소의 색상 차이가 일정 문턱치(threshold)값보다 작은지를 확인하여 유사도가 가장 높은 변이를 투표를 통해 선정한다. 이런 과정을 거치면 넓은 영역에서 발생하는 띠 모양의 변이 잡음이 효과적으로 제거된다.

그 다음에는 인접한 4 방향에서 양방향 필터링(bilateral filtering)이 수행된다. 4 방향으로 연결된 이웃한 점에서 정확한 정합 결과를 참고하여 라인 사이의 일관성을 보완한다. 개선된 변이 영상은 그림 4에서 확인할 수 있다. 이렇게 후처리 개선까지 완료된 최종 변이 영상의 해상도는 HD의 1/4 크기인 960×540 해상도를 가지며 다음 모듈인 중간 영상 합성 모듈에서 입력 데이터로 사용된다.

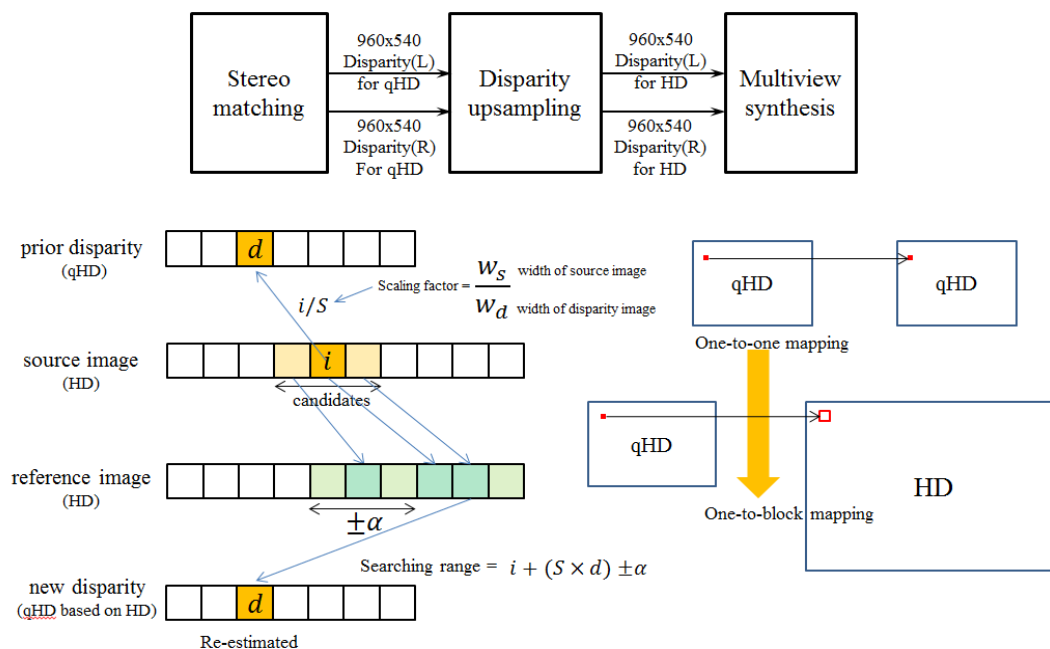


그림 3. 변이 업샘플링

Fig 3. linear disparity upsampling

2. GPU기반 중간 영상 합성 모듈

변이 추출이 완료되어 좌우 각각에 대한 변이영상이 출력되면 바로 GPU로 다시 전송한다. 입력된 변이 영상은 해상도가 원본 영상의 1/4 크기로 생성되었으므로 이를 다시 HD 해상도에 맞게 변환해 주는 과정이 필요하다. 여기서는 간단한 변이 확대(disparity upsampling) 방법을 이용하였다.

그림 3을 보면 추출된 이전 변이 정보(prior disparity)를 참고하여 검색 범위 내의 후보군을 지정하고, 해당 후보군의 변이로 검색 범위를 지정한 후 다시 정합을 수행하여

가장 적합한 변이를 선택한다. 변이 확대 과정이 끝나면 변이 영상의 모든 변이가 HD 기준으로 바뀌게 되지만, 하나의 화소는 원본 영상에서 4개 화소의 변이를 표현하므로 영상 합성 시에 블록 단위로 영상 합성을 한다.

블록 단위의 영상 합성 알고리즘은 화소 단위의 알고리즘에 비해 연산량이 현저히 줄어들지만 화질 저하가 거의 발생하지 않는다^[9]. 중간 영상 합성은 좌우 영상에서 각각 투영된 두 영상의 색상 차이를 고려한 가중치-평균 방법을 이용하는 식 (1)의 혼합(blending) 과정을 거친다.

$$V_{\alpha}(p) = \lambda(p, \alpha) \cdot V_{\alpha,l}(p) + [1 - \lambda(p, \alpha) \cdot V_{1-\alpha,r}(p)] \quad (1)$$

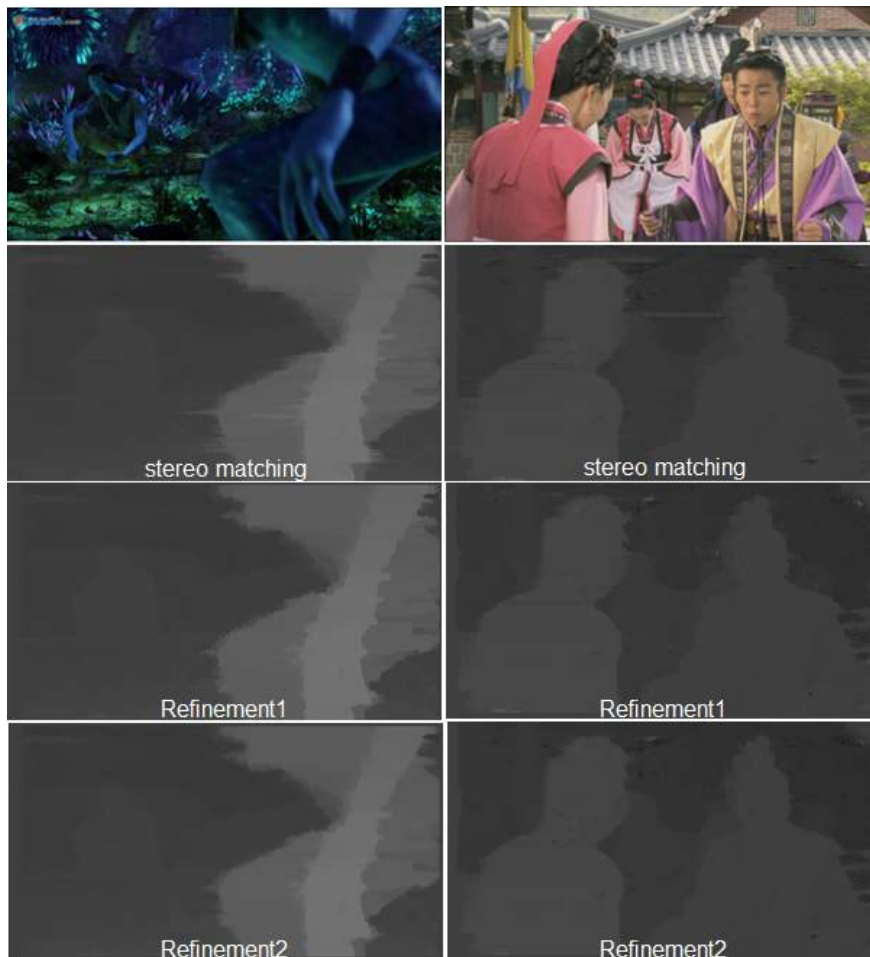


그림 4. 변이 추출 및 변이 개선 결과 영상

Fig. 4. The result images of the disparity estimation and refinement

식 (1)에서 $V_{\alpha,l}(p)$ 과 $V_{1-\alpha,r}(p)$ 는 각각 좌우 영상으로부터 가상 시점의 위치인 α 로 투영된 영상을 나타낸다. α 는 0과 1사이의 값을 가지며, $V_{\alpha}(p)$ 는 이 두 영상이 혼합된 영상이다. 이때, λ 는 가중치이며 아래의 식 (2)에 의해서 계산된다.

$$\lambda(p, \alpha) = \frac{(1-\alpha) \cdot \lfloor e_{1-\alpha,r}(p) \rfloor}{\alpha \cdot e_{\alpha,l}(p) + (1-\alpha) \cdot e_{1-\alpha,r}(p)} \quad (2)$$

가중치 계산에 활용되는 오차값 e 는 변이값을 토대로 기준 영상과 참조 영상에서의 가져온 화소값 차이를 식 (3)과 같이 계산하여 반영한다.

$$e_{\alpha,l} = \left\{ \left| V_{\alpha,l}(p) - I_r[q+d_l(q)] \right|, \left| V_{\alpha,l}(p) - I_l[q+d_r(q)] \right| \right\} \quad (3)$$

혼합 과정까지 거치면 좌우 시점 영상에서 보이지 않는 차폐 영역(disocclusion area)은 홀(hole) 영역으로 남아 있게 되는데 이 부분은 주변의 인접 화소 정보를 이용하여 채워지게 된다^[10].

IV. 실험 및 결과분석

그림 4는 변이 추출 결과를 보여준다. 초기 변이 영상은 FPGA기반의 Trellis 동적 프로그래밍 기법을 이용하여 생성된 영상이며, 이후에 두 차례의 후처리 과정을 거쳐 최종 변이 영상이 만들어진다. 앞서 설명한 것처럼 첫 번째는 수직 투표 방식을 통해 개선된 영상이며, 두 번째는 4개의 이웃한 점을 이용한 양방향 필터링을 통해 개선된 영상이다. 경계 부분에서 보였던 라인 단위의 띠 모양의 잡음을 포함한 잡음들이 단계적으로 제거되었음을 확인할 수 있다.

표 1은 각 단계별 연산 시간을 나타낸다. 변이 추출이 독립적인 장치에서 별도로 연산을 수행하기 때문에 GPU가 장착되어 있는 호스트와 FPGA 모듈 간의 데이터 전송 과정이 있는데, 이 때 한 프레임 지연(latency)이 발생한다. 첫 번째 프레임이 FPGA에 입력되어 추출된 결과 변이 영상이

두 번째 프레임이 입력될 때 출력되므로 1 프레임 지연이 발생한다.

표 1. 각 단계별 연산 시간

Table 1. Computation time of each step

| 단계 | 연산 시간(s) |
|-------------------------|----------|
| FPGA -> GPU 데이터 전송 (2K) | 0.002 |
| 크기 및 색상 변환 | 0.008 |
| FPGA 기반 변이 추출 | 0.033 |
| GPU 기반 변이 업샘플링 | 0.012 |
| GPU 기반 중간영상 합성 | 0.018 |
| 다시점 다중화 (4K) | 0.024 |
| GPU -> Host 데이터 전송 (4K) | 0.015 |

입력 영상은 가로 1920, 세로 1080 해상도를 가진 side-by-side 규격의 스테레오 영상이며, 크기 및 색상 변환 과정에서 가로 1280, 세로 720 해상도로 변환된다. 변환된 영상은 중간 영상 합성 단계의 입력 영상으로 쓰이며, FPGA 변이 추출 모듈에서 전송된 영상을 이용하여 가로 1280, 세로 720 해상도를 가진 9개의 중간 시점 영상을 생성한 후에 목표 디스플레이의 규격에 맞게 부분 화소 단위로 재배치하는 시점 다중화(view multiplexing) 과정을 거쳐 최종적으로 4K 해상도의 영상이 4K 디스플레이에서 재생된다. 실험에 사용된 무안경 다시점 디스플레이는 그림 2에 보이는 Toshiba의 55인치 4K 무안경 다시점 디스플레이이다.

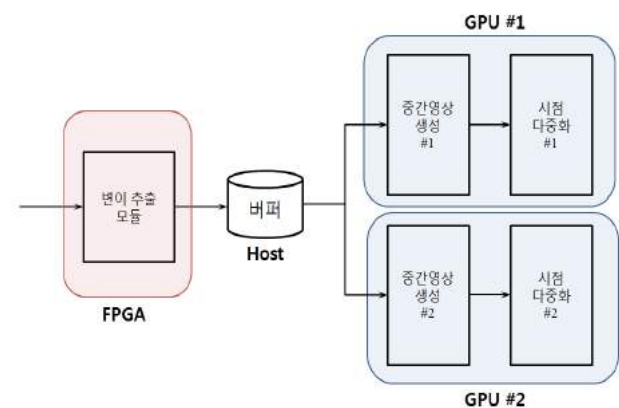


그림 5. 다중 GPU 시스템 구성도

Fig. 5. multi-GPU system

연산 시간을 보면 실시간 재생기준이 되는 30 fps에 못 미치는데, 이는 그림 5처럼 다중 GPU 시스템을 구성하여 속도를 30fps까지 높일 수 있었다. FPGA에서 출력되는 비디오 영상을 호스트의 버퍼를 거쳐 두 GPU 장치가 동시에 병렬로 수행할 수 있게 프레임이 분배되도록 구성하였다. 실험에 사용된 GPU 디바이스는 NVIDIA사의 Geforce

GTX 690이며 CUDA를 기반으로 구현하였다.

합성된 영상의 품질을 객관적 수치로 측정하기 위해 다시점 카메라 영상이 있는 MPEG GIST Cafe (그림 6(h))의 카메라 시점 영상의 위치에 중간 시점을 생성하여 카메라 시점과 중간 합성 영상의 크기를 맞추어 두 영상 간의 PSNR을 측정한 결과 31dB 정도가 나옴을 확인하였다.



(a)



(b)



(c)



(d)



(e)

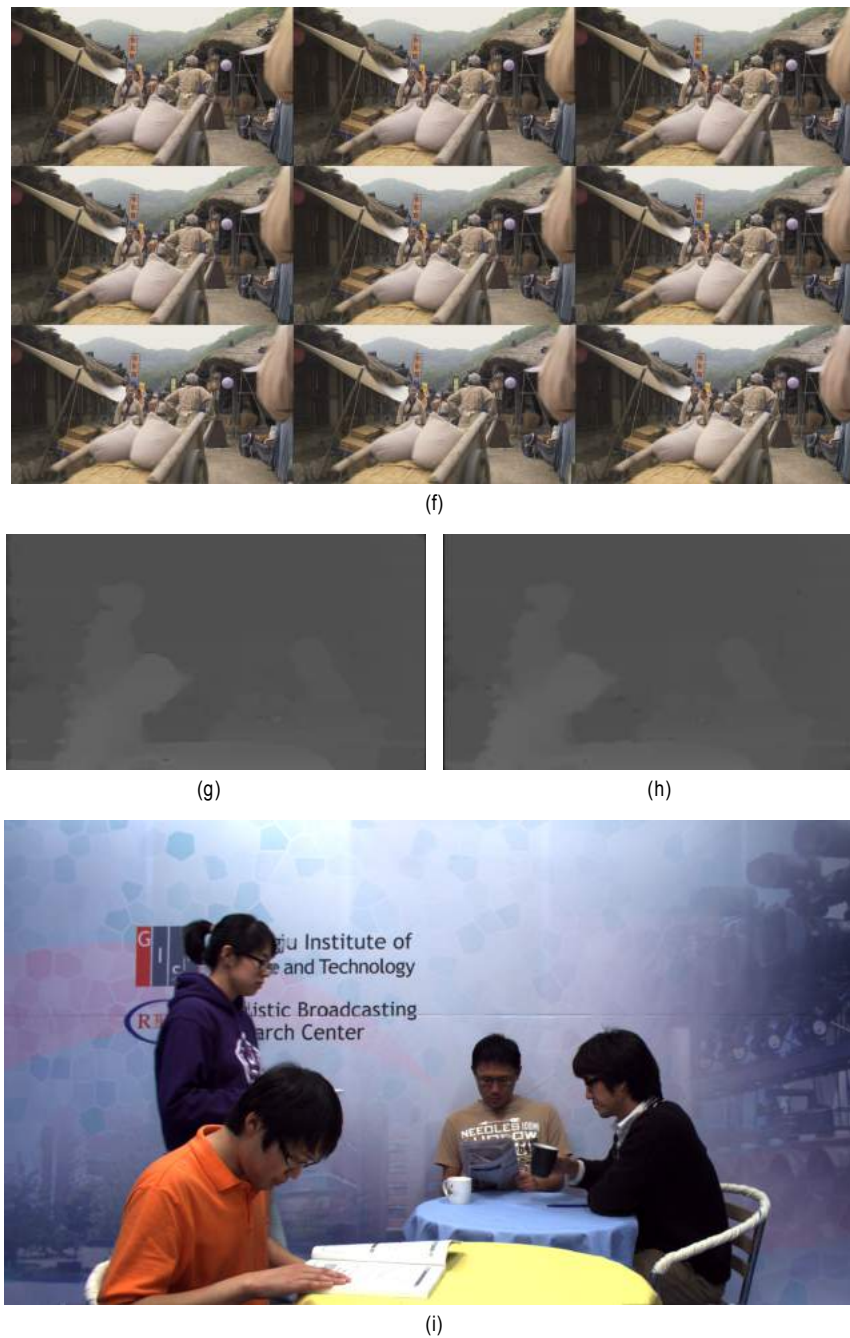


그림 6. 변이 추출 결과 영상과 중간 시점 합성 영상. (a) 좌영상, (b) 우영상, (c) 좌측 변이영상, (d) 우측 변이영상, (e) 합성된 중간시점 영상($\alpha=0.5$), (f) 합성된 9시점 영상, (g) 좌측 변이영상, (h) 우측 변이영상, (i) 합성된 중간시점 영상
 Fig. 6. The result images of disparity estimation and intermediate view using "Gaebeack" and MPEG "GIST Cafe" sequences, (a) left source image, (b) right source image, (c) left disparity, (d) right disparity, (e) synthesized intermediate view($\alpha=0.5$), (f) synthesized 9-intermediate views, (g) left disparity, (h) right disparity, (i) synthesized intermediate view

V. 결 론

본 논문에서는 FPGA와 GPU를 이용한 실시간 무안경 3차원 비디오 영상을 생성하는 시스템을 소개하였다. 메모리 사용량이 비교적 적고 많은 수의 비트 연산이 필요한 스테레오 정합을 FPGA 기반으로 구현하였으며, 메모리 사용량이 많고 임의의 메모리 접근이 빈번한 중간 영상 합성 모듈은 GPU 기반으로 구현한 후 이중 간의 모듈을 결합하여 하나의 통합된 시스템으로 구성하였다. 또한, 영상 합성 모듈의 연산 속도를 실시간까지 올리기 위해 다중 GPU 시스템을 구성하여 실시간까지 속도를 높일 수 있었다. 결과로서 제안된 시스템은 HD 스테레오스코픽 영상을 받아 9개의 HD 중간시점 영상을 30fps 속도로 생성한 후 4K 해상도인 하나의 최종 출력 영상으로 다중화하여 다시점 디스플레이에 재생한다.

향후에는 변이 영상의 품질을 더욱 향상시키기 위해 원본 영상을 유도영상(guidance image)으로 이용하여 가중치 기반 중앙값 필터링(weighted median filtering)으로 변이 영상을 실시간으로 개선하는 방법을 구현, 적용할 예정이다.

참 고 문 헌 (References)

- [1] F. Zilly, C. Riechert, M. Muller, P. Eisert, T. Sikora, and Peter Kauff, "Real-time generation of multi-view video plus depth content using

- mixed narrow and wide baseline," *Journal of Visual Communication and Image Representation*, 2013.
- [2] M. Muller, F. Zilly, P. Kauff, "Adaptive cross-trilateral depth map filtering," in *3DTV-Conference*, pp. 1-4, June 2010.
- [3] S. Jin, J. Cho, X. D. Pham, K. M. Lee, S.-K. Park, M. Kim, and J.W. Jeon, "FPGA design and implementation of a real-time stereo vision system," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 1, pp. 15 - 26, 2010.
- [4] L. Zhang, K. Zhang, T. S. Chang, G. Lafruit, G. K. Kuzmanov, and D. Verkest, "Real-time high-definition stereo matching on FPGA," *Proceedings of the 19th ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, pp. 55 - 64, 2011.
- [5] H.-C. Shin, G.-M. Um, C. Kim, W.-S. Cheong, and N. Hur, "Autostereoscopic 3D video generation from stereoscopic videos using FPGA and GPU," *Proceedings of IEEE International Conference on 3D Imaging(IC3D)*, 2012.
- [6] K. Pauwels, M. Tomasi, J. Dias, E. Ros, and M. M. Van Hulle, "A Comparison of FPGA and GPU for Real-Time Phase-Based Optical Flow, Stereo, and Local Image Features," *IEEE Transaction on Computers*, vol. 61, no. 7, pp. 999-1012, 2012.
- [7] Y. Oh and H. Jeong, "Trellis-based parallel stereo matching," *Proceedings of IEEE conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2143-2146, 2000.
- [8] X. Sun, X. Mei, S. Jiao, M. Zhou, and H. Wang, "Stereo matching with reliable disparity propagation," *Proceedings of IEEE International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT)*, 2011.
- [9] L. Zhang, "Fast stereo matching algorithm for intermediate view reconstruction of stereoscopic television images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 10, pp. 1259-1270, 2006.
- [10] L. Do, G. Bravo, S. Zinger, P. H. de With, "GPU-accelerated Real-time Free-viewpoint DIBR for 3DTV," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, May 2012.

저 자 소 개



신 홍 창

- 2005년 2월 : 세종대학교 컴퓨터공학과 학사 졸업
- 2008년 8월 : 한양대학교 전자통신컴퓨터공학과 석사 졸업
- 2009년 3월 ~ 현재 : 한국전자통신연구원 방송통신미디어연구소 입체방송연구실 연구원
- 주관심분야 : 스테레오 정합, 영상합성, 차세대 방송기술, 병렬처리

저 자 소 개



이 진 환

- 1987년 2월 : 한국항공대학교 통신공학과 학사
- 2000년 2월 : 한국과학기술원 통신공학과 석사
- 1989년 2월 ~ 현재 : 한국전자통신연구원 방송통신미디어연구소 입체방송연구실 책임연구원
- 주관심분야 : 입체방송, 디지털방송시스템



이 광 순

- 1993년 : 경북대학교 전자공학과 학사
- 1995년 : 경북대학교 전자공학과 석사
- 2004년 : 경북대학교 전자공학과 박사
- 2001년 ~ 현재 : 한국전자통신연구원 입체방송연구실 실장
- 주관심분야 : 3DTV 방송시스템, 무안경 3D, 영상처리 등



허 남 호

- 1992년 2월 : 포항공과대학교 전기전자공학과 공학사
- 1994년 2월 : 포항공과대학교 대학원 전자전기공학과 공학석사
- 2000년 2월 : 포항공과대학교 전자전기공학과 공학박사
- 2000년 5월 ~ 현재 : 한국전자통신연구원 방송통신미디어연구소 방송시스템연구부 부장
- 주관심분야 : 차세대 방송기술, 제어 및 전력전자