

가상 변이맵 탐색과 수렴 조건 판단을 이용한 실시간 가상시점 생성 방법

준회원 신인용*, 종신회원 호요성**

Real-time Virtual View Synthesis using Virtual Viewpoint Disparity Estimation and Convergence Check

In-Yong Shin* *Regular Member*, Yo-Sung Ho** *Lifelong Member*

요약

본 논문에서는 양안식 3차원 디스플레이 장치에 사용가능한 실시간 가상시점 영상을 생성방법을 제안한다. 제안하는 방법은 변이 탐색 (disparity estimation) 과정에서 중복되거나 불필요한 계산을 제거하기 위해, 변이맵 탐색 위치를 가상시점과 동일한 위치로 설정한다. 또한, 전역 방법을 사용한 변이맵 탐색 과정은 정확도는 높지만 복잡한 계산량이 문제이기 때문에 최적화 과정에서 변이값이 수렴하는 위치를 탐색하고 정보 갱신을 멈추는 방법을 사용하였다. 이렇게 얻어진 가상시점의 변이맵은 실제 참조가 가능한 영역인지 판단하는 수렴 조건 판단 (convergence check) 절차를 거쳐 영상을 합성한다. 영상 합성을 위한 모든 과정은 실시간 처리를 위하여 단일 명령 복수 쓰레드 (Single Instruction Multiple Threads) 구조인 그래픽 처리장치(GPU)를 활용한 프로그래밍을 통해 고속화를 이루었다. 실험 결과를 통해 제안하는 방법이 합성 영상의 화질을 유지하면서도 실시간으로 가상시점 영상을 합성한 것을 확인할 수 있었다.

Key Words : Stereo matching, belief propagation, GPU programming, CUDA, view synthesis

ABSTRACT

In this paper, we propose a real-time view interpolation method using virtual viewpoint disparity estimation and convergence check. For the real-time process, we estimate a disparity map at the virtual viewpoint from stereo images using the belief propagation method. This method needs only one disparity map, compared to the conventional methods that need two disparity maps. In the view synthesis part, we warp pixels from the reference images to the virtual viewpoint image using the disparity map at the virtual viewpoint. For real-time acceleration, we utilize a high speed GPU parallel programming, called CUDA. As a result, we can interpolate virtual viewpoint images in real-time.

I. 서론

최근 들어 3차원 영상서비스에 대한 소비자들의 관심과 그것을 위한 연구가 활발히 진행되고 있다. 3차원 영상 서비스는 인간의 양안 깊이 단서를 제공하여 시청자에게 기존의 방송과는 다르게 입체감

과 몰입감을 선사한다. 이러한 3차원 방송 서비스의 활용범위는 영상회의 시스템, 영화, 의료영상, 입체 TV와 같이 다양하다. 그에 따라 동영상관련 표준화 기구인 MPEG에서는 3DAV(3D audio-viual) 표준화에 노력을 기울이고 있다^[1].

현재 활발히 연구되는 3차원 비디오 시스템이 많

* 광주과학기술원 실감방송연구센터(hoyo@gist.ac.kr), (° : 교신 저자)

논문번호 KICS2011-07-278, 접수일자 : 2011년 7월 1일, 최종논문접수일자 : 2012년 1월 10일

많은 발전을 이루었음에도 불구하고 수용치를 벗어난 양안시차나 수렴거리 불일치 등의 이유로 시각적 불편과 피로의 문제가 발생된다. 따라서 인간 시각 시스템의 기능장애와 관련한 시각피로도 유발요인의 경우 그것을 억제하려는 노력과 시도가 필요하다.

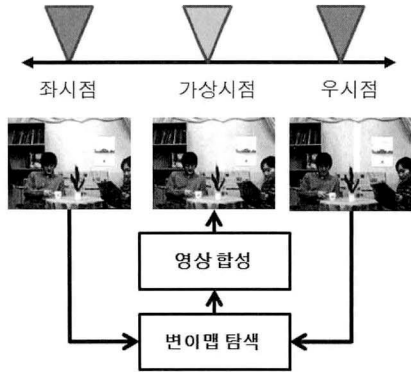


그림 1. 가상시점 합성과정

MPEG (moving picture experts group) 3DVC 그룹에서는 이러한 관점에서 스테레오스코픽을 이용하여 두 시점의 사이에 시청자에게 적절한 양안시차를 제공하는 임의시점의 영상을 합성한다. 합성된 영상은 스테레오스코픽 원영상중 하나의 영상과 함께 새로운 스테레오스코픽 영상 조합을 이루어 시청자에게 제공되는데 이것을 고급-스테레오스코픽 (advanced stereoscopic) 디스플레이 시스템이라고 한다^[2].

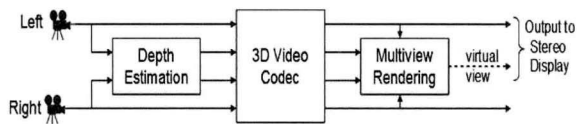


그림 2. 고급-스테레오스코픽 디스플레이 시스템

이렇게 시점 거리가 조절된 양안식 영상은 사용자가 느끼는 피로의 정도에 따라 조절이 가능한 장점이 있다. 하지만 새로운 스테레오스코픽 영상을 만들어 내는 과정인 가상시점 합성단계에 필요한 변이맵을 탐색과정은 실시간 처리가 힘들기 때문에 그림 2와 같이 전송과정 이전 단계에서 변이맵을 탐색하고 추가적으로 스테레오 영상과 함께 변이맵을 전송해야하는 대역폭 부담이 생기는 단점이 있다.

본 논문에서는 고급-스테레오스코픽 디스플레이 시스템의 단점인 수신단에서 변이맵을 사전에 탐색하여 압축하고 스테레오스코픽 영상과 별개로 추가

전송해야 하는 비용적인 문제를 해결하기 위해 영상 송신단이 아닌 수신단에서 양안식 영상을 이용하여 실시간으로 변이맵을 합성하고 이를 이용해 가상시점을 합성하는 방법을 제안한다.

논문의 구성은 다음과 같다. 2장에서는 고급-스테레오스코픽 디스플레이 시스템에서 사용되는 변이맵 탐색과 영상합성에 대해 설명한다. 3장에서는 제안하는 실시간 가상 시점 생성 방법을 설명한다. 4장에서는 구현 결과를 보이고 기존 연구와 비교한다. 마지막으로 5장에서는 결론을 맺는다.

II. 고급-스테레오스코픽 디스플레이 시스템

제안하는 방법에 앞서 고급-스테레오스코픽 디스플레이 시스템에 대하여 소개하고자 한다. 3차원 영상의 시각적 피로감은 다양한 요인에 의해서 발생되고 그것들 중 일부분은 어느 정도 해결이 가능한 수준에 이르렀다.

고급-스테레오 스코픽 디스플레이 시스템이 이러한 인간 시각 시스템의 기능장애와 관련한 시각피로도를 줄이기 위해 제안되었다. 영상합성을 위해서는 기본적으로 좌시점과 우시점의 영상정보와 각 시점에서의 변이정보가 필요하다.

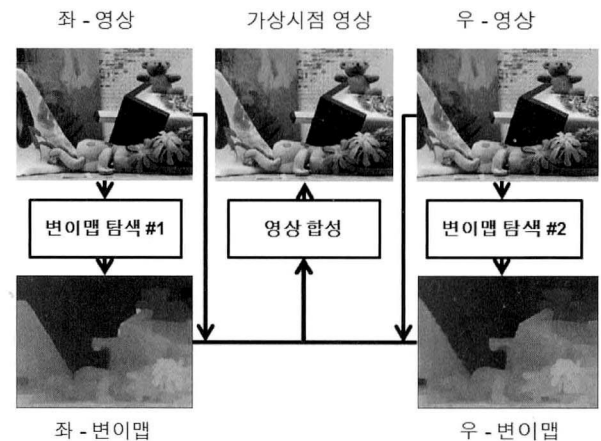


그림 3. 고급-스테레오스코픽 디스플레이 시스템의 영상합성 방법

변이맵의 경우 좌시점과 우시점에 대하여 각각 탐색이 수행되어 모두 2개의 변이맵을 얻는다^{[3][4]}. 이렇게 얻어진 변이맵은 사용자가 원하는 가상 시점으로 참조 시점의 영상정보를 이동시키는데 사용된다. 일반적으로 변이맵을 가상시점에 먼저 전방 위핑 시키고 미디언 필터를 거친 후 위핑된 가상시점의 변이정보를 사용하여 후방 위핑 시키는 방법

을 사용한다. 마지막으로 좌시점과 우시점에 모두 존재하지 않는 화소의 경우 홀영역으로 지정하고 주변의 영상정보를 활용하여 홀을 채우게 된다.

III. 제안하는 실시간 가상시점 영상 합성

본 장에서는 고급-스테레오스코픽 디스플레이 시스템의 단점을 보완하는 실시간 가상시점 합성 방법을 설명한다. 양안식 영상과 각각의 시점에 해당하는 변이맵이 존재하는 경우 추가적인 변이맵 탐색과정이 없이 다양한 위치의 가상시점 생성이 용이하다. 하지만 합성하고자 하는 시점이 고급-스테레오스코픽 디스플레이 시스템과 같이 단지 하나의 시점으로 고정된 경우 두 번의 변이탐색 과정은 영상합성에는 전혀 사용되지 않는 변이값들을 계산하고 과정이 포함된다. 그림 4는 고급-스테레오스코픽 디스플레이 시스템에서 입력된 스테레오 영상과 각 시점에 해당하는 변이맵을 탐색한 결과이다. 그림 4의 변이맵에서 빨간색으로 칠해진 영역의 경우 가상시점을 가운데에 생성하였을 경우 전경에 의해 가려지게 되어 영상합성에 전혀 쓰이지 않는 부분이다. 따라서 이러한 영역은 변이탐색과정에서 탐색할 필요가 없는 부분들이다. 그리고 파란색으로 칠해진 영역은 같은 변이값을 갖는 영역으로 이중으로 변이탐색을 하기 때문에 중복된 계산으로 복잡도만 증가하게 된다. 따라서 불필요한 영역은 변이탐색에서 제외하고 중복된 영역은 한번만 탐색하는 방법이 효율적이다.

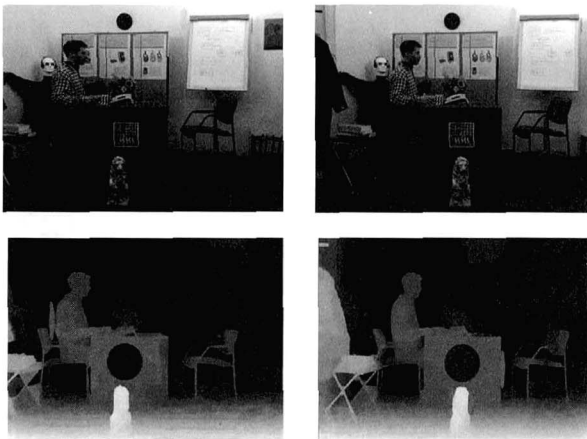


그림 4. 좌시점과 우시점 변이맵의 중복되거나 불필요한 변이 정보

가상시점 합성 과정에서 앞서 말한 단점들을 극복하기 위해 가려지는 영역들을 제외하고 변이값이

중복으로 탐색되는 부분을 피하기 위해서 변이맵을 가상시점에서 직접 탐색하는 방법을 사용하였다. 기존의 참조영상과 같은 시점에 해당하는 변이맵을 탐색하는 것이 아니고 가상시점에 위치하는 변이맵을 탐색하여 앞서 말한 단점들을 보완한다. 제안하는 방법은 다음 과정을 통해 수행된다.

1단계 : 제안하는 새로운 에너지 함수를 사용하여 가상시점의 변이맵을 직접 탐색. 최적화를 위해 계층적 신뢰진과 알고리즘을 사용함.

2단계 : 가상시점 변이맵을 이용하여 참조시점의 영상정보를 합성을 위해 가져올 것인지 결정하는 가려진 영역(occlusion region)을 결정한다.

3단계 : 양쪽의 참조시점 모두에 대하여 참조가 불가능한 홀영역에 대해서 주변에 위치하는 영상정보를 활용하여 홀을 채운다.

3.1 가상시점 변이맵 탐색

변이맵을 탐색하기 위해서는 에너지 함수 선언하고 에너지 함수가 최소값이 되는 변이맵을 구해야 한다⁵⁾. 일반적으로 전역적 방법과 지역적 방법이 사용되고 있으며 지역적 방법은 복잡도가 낮은 반면에 정확도가 떨어지고 전역적 방법은 복잡도는 높지만 정확도가 높은 장점이 있다⁶⁾⁷⁾. 본 논문에서는 정확도를 위해 전역적 방법에 사용되는 에너지 함수를 적용하였다. 에너지 함수는 일반적으로 식 1과 같이 이루어져 있다. 전역적 방법의 경우 데이터 부분과 주변노드의 변이값 연속성을 위한 에너지를 모두 고려한다.

$$E(x, y, d) = E_{data}(x, y, d) + E_{smooth}(x, y, d) \quad (1)$$

$$\sum_{x,y} |I_L(x + d_{V_L}(x, y), y) - I_R(x - d_{V_R}(x, y), y)| \quad (2)$$

$$d_{V_R}(x, y) = \text{Alpha} \times d_{V_L}(x, y) \quad (3)$$

가상시점의 변이맵을 얻기 위해서는 데이터 부분을 식 2와 같이 새롭게 사용하였다. d_{V_L} 과 d_{V_R} 은 가상시점의 각 화소위치에 대응하여 가상시점과 참조시점간의 변이값이다. 두 변수의 관계는 식 3과 같이 나타낼 수 있으며 Alpha값은 그림 5와 같이

가상시점과 우시점의 상대적 거리를 나타낸다.

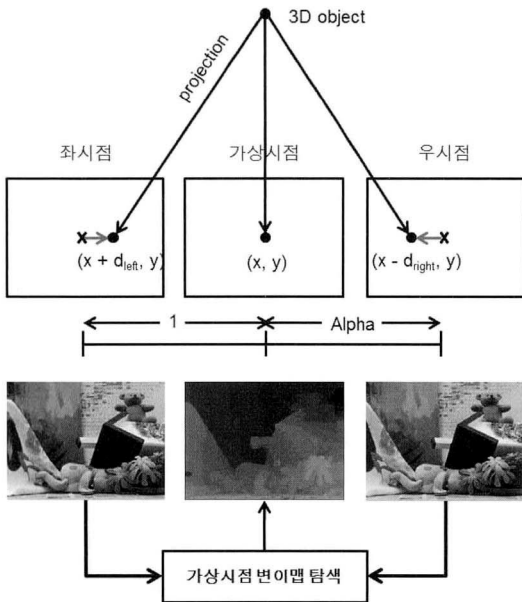


그림 5. 가상시점 변이맵 탐색

전역적인 변이맵 탐색 방법중 속도와 정확도를 위해 계층적 신뢰전과 알고리즘을 사용하였다. 계

3.2 수렴판단을 통한 계산량 축소

계층적 신뢰전과 알고리즘은 획기적으로 메시지 업데이트량을 줄였으나 여전히 모든 노드에 대하여 수많은 반복이 필요하다. 그림 6은 계층별 변이맵 결과로 평탄한 변이정보를 갖는 영역에 대해서는 변이값의 변화가 적은 것을 알 수 있다. 이러한 특징을 이용하여 이미 수렴한 변이정보를 갖는 노드를 검출하고 이러한 영역에 대하여 메시지 계산을 제외시키는 방법을 적용하였다.

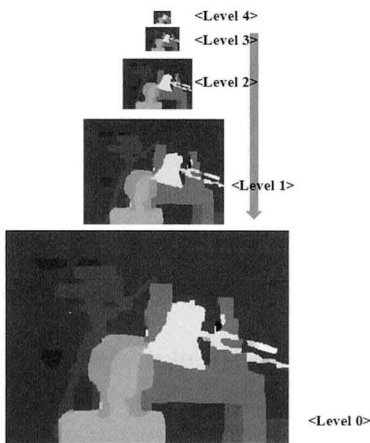


그림 6. 계층별 변이맵

그림 7은 계층별 변이맵을 구하고 현재의 변이맵과 이전에 얻은 변이맵과의 차이를 이용하여 수렴 여부를 판단한 결과이다. 그림 7의 아래에 위치한 영상들이 수렴영역을 판단한 영상으로 검정색 부분이 수렴하는 위치이다.

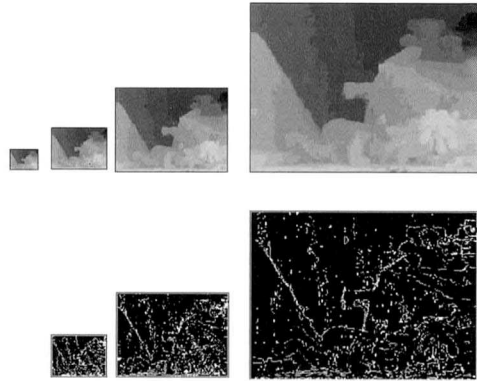


그림 7. 계층별 변이맵과 수렴영역 판단

3.3 가려진 영역 결정 및 후방 워핑

앞선 과정에서 가상시점 변이맵을 얻게 되면 후방 워핑을 통하여 참조시점의 색상정보를 가져올 수 있다. 하지만 가상시점의 화소가 참조시점에서는 가려진 경우 색상정보는 전혀 다른값을 갖고 있기 때문에 참조해서는 안된다. 이러한 경우를 피하기 위해 가상시점의 각 화소에 대하여 참조시점에서의 가려짐 여부를 판별해야 한다. 가려짐 여부를 판별하기 위해서는 획득한 가상시점 변이맵을 좌우 참조시점으로 워핑시키면서 참조시점의 각 위치에 대하여 중복되어 워핑되는 부분에 대해서 변이값이 가장 작은 것이 원래 존재했던 가상시점에서의 위치정보를 따로 저장하여 가상시점 영상합성을 위한 후방워핑시 참조하지 못하게 하는데 사용한다.

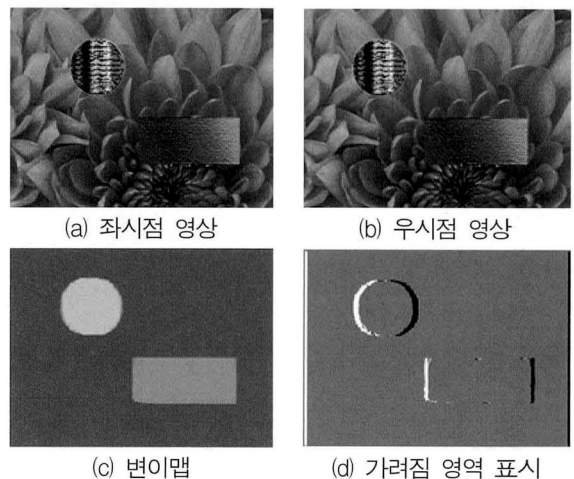


그림 8. 가려짐 영역 결정

그림 8은 인공적으로 생성한 양안식 영상과 그것으로 만든 가상시점 변이맵 그리고 영상합성을 위한 가려짐 영역을 표시한 내용이다. 그림 8(d)에서 흰색은 좌시점에 대한 가려짐을 나타내고 검정색은 우시점에 대한 가려짐을 나타낸다. 그리고 회색은 좌시점과 우시점 모두에 대하여 참조 가능함을 나타낸다.

3.4 홀채움

양쪽의 참조시점이 존재하는 중간시점 영상합성의 경우 참조가 가능한 시점이 하나도 존재하지 않는 홀영역은 거의 나타나지 않는다. 하지만 홀영역이 존재하는 경우 참조가능한 영상정보가 없더라도 주변 화소값과의 유사성을 이용하여 홀채움이 가능하디^{[8][9]}. 이러한 홀영역에 대해서는 속도를 위해 주변의 영상정보를 그대로 복사하는 방법을 사용하였다.

3.5 GPU 병렬 구현

실시간 처리를 위해서 병렬처리 방법중 하나인 GPU를 이용한 병렬처리 프로그래밍을 사용하였다. GPU의 경우 CPU와는 구조적인 차이로 수치계산에 관련된 코어의 수가 월등히 많아 병렬처리에 유리하다. GPU 병렬처리의 경우 단일 명령 복수 데이터 구조를 지원하는데 이것은 여러 화소에 대하여 동일한 명령을 줄 수 있어 영상처리에 적용이 용이한 편이다 [10].

IV. 실험 결과 및 분석

제안한 방법의 특징은 종전의 두 참조시점에 대한 변이맵을 탐색 및 생성하여 가상시점 영상을 합성하지 않고 직접 가상시점에 해당하는 변이맵을 생성하여 그것을 이용해 가상시점을 합성하는 것이다. 따라서 제안한 방법의 성능평가를 위해 방법 1인 두 참조시점의 변이맵을 탐색하여 영상을 합성하는 방법과 방법 2인 두 참조시점중 하나의 참조시점에 대한 변이맵을 탐색하여 영상을 합성하는 방법 두가지 방법을 사용 및 비교하였다. 가상시점의 품질평가를 위해 3시점의 영상이 존재하는 Middlebury에서 제공하는 영상과 MPEG의 MVD 테스트 영상을 사용하였다.

표 1. 실험 영상 정보

| 실험영상 | Teddy | Poster | Cones | News paper | Book arrivals |
|------|----------|----------|----------|------------|---------------|
| 해상도 | 640 x480 | 480 x416 | 480 x416 | 640 x480 | 640 x480 |
| 변이범위 | 30 | 20 | 20 | 50 | 50 |

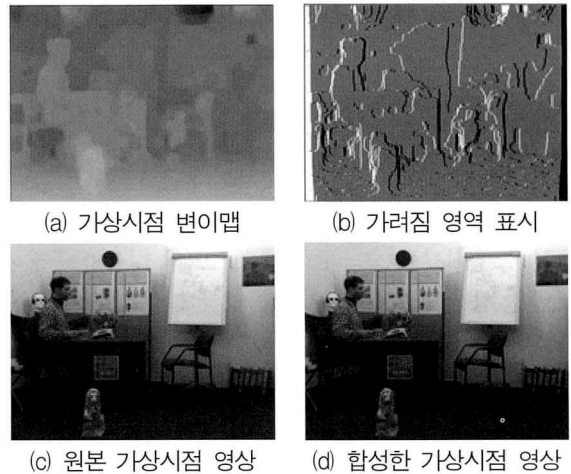


그림 9. 변이맵 및 가상시점 영상 합성 결과

그림 9는 시점이 3개 존재하는 영상에서 좌우 시점 영상만을 입력으로 중간시점에 대하여 가상시점을 생성하고 가려짐 영역을 판단하여 가상시점 영상을 합성한 결과를 보여준다. 그림 9(c)는 합성한 가상시점의 품질을 확인하기 위한 동일한 시점의 원본과 PSNR값을 측정하였다. 그림 10은 방법 1,2와 제안하는 방법의 가상시점 영상 PSNR값 비교 결과이다. 그림 11은 수행 시간 측정치 비교를 나타내는데 CPU 상에서 방법 1,2와 제안하는 방법의 속도측정을 하고 추가적으로 제안하는 방법을 GPU 구현시 속도를 측정하여 표시하였다. 그 결과 실시간 처리가 가능한 속도증가를 나타내었다.

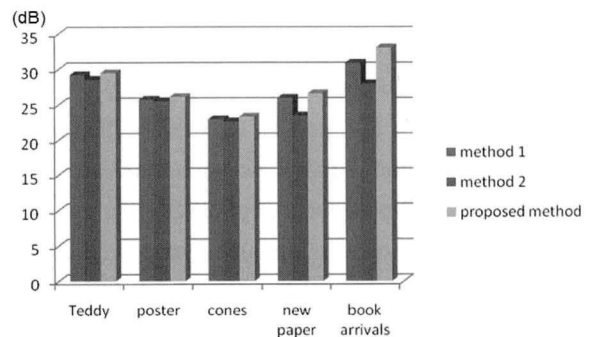


그림 10. 합성한 가상시점 영상의 품질 비교

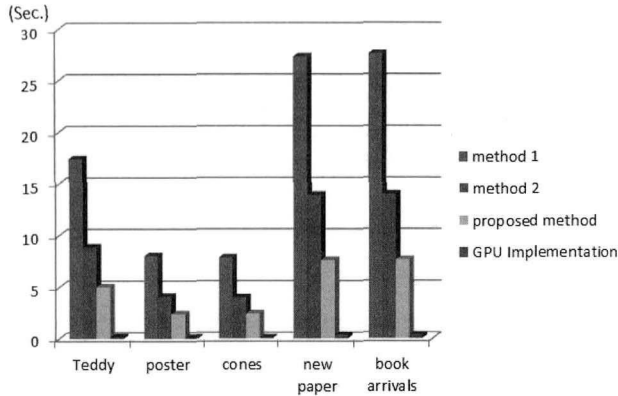


그림 11. 수행 시간 측정치 비교

V. 결 론

본 논문에서는 고급-스테레오스코픽 디스플레이 시스템을 위한 실시간 가상시점 영상 합성 방법을 제안하였다. 기존 방법의 단점을 보완하기 위하여 불필요하거나 중복된 계산을 줄이는 가상시점에서 변이맵을 직접 탐색하고 계층적 신뢰전파 과정에서 수렴부분을 검출하여 계산량을 줄이며 변이맵 탐색과정을 수신단에서 실시간으로 가능하게 하는 방법을 제안하였다. 모든 과정은 GPU로 구현하여 처리속도를 더욱 높였다. 그 결과 고품질의 영상품질을 유지하면서 실시간으로 가상시점 영상을 생성하였다.

참 고 문 헌

[1] ISO/IEC JTC1/SC29/WG11, "Survey of algorithms used for multi-view video coding," N6909, January 2005.

[2] Video and Requirements, "Applications & requirements on 3D video coding," ISO/IEC JTC1/SC29/WG11 Doc. N11061, Xian, China, Oct. 2009.

[3] ISO/IEC JTC1/SC29/WG11, "Contribution for 3D Video Test Material of Outdoor Scene," M1537, April 2008.

[4] J. D. Oh, S. Ma, and C. C. J. Kuo, "Disparity estimation and virtual view synthesis from stereo video," in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), pp. 993 - 996, New Orleans, La, USA, May 2007.

[5] M. Tappen and W. Freeman, "Comparison of Graph Cuts with Belief Propagation for Stereo," Proc. IEEE Int'l Conf. Computer Vision, vol. 1,

pp. 508-515, 2003.

[6] P. F. Felzenszwalb and D. P. Huttenlocher, "efficient Belief Propagation for Early Vision," CVPR, vol. 1, pp.261-268, 2004.

[7] Q. Yang, L. Wang, R. Yang, H. Stewenius, and D. Nister, "Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling," CVPR, June 2006.

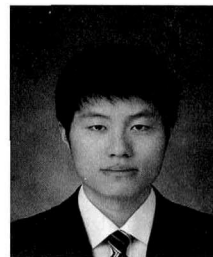
[8] K. Oh, S. Yea, and Y. Ho, "Hole Filling Method using Depth Based In-painting for View Synthesis in Free Viewpoint Television and 3-D Video," Picture Coding Symposium, pp. 39 (1-4), 2009.

[9] M. Oliveira, B. Bowen, R. McKenna, Yu-Sung Chang, "Fast Digital Image Inpainting", Proceedings of the International Conference on Visualization, Imaging and Image Processing (VIIP 2001), Marbella, Spain, September 3-5, 2001

[10] nVIDIA Corporation, "CUDA 3.2 Programming Guide," in http://www.nvidia.com/cuda_develop.html, 2010

신 인 용 (In-Yong Shin)

준회원



2009년 인하대학교 전자공학과 졸업(학사)
 2009년~현재 광주과학기술원 정보통신공학과 석사과정
 <관심분야> 디지털 신호처리, H.264/AVC, GPU 병렬처리

호요성 (Yo-Sung Ho)

중신회원



1981년 서울대학교 공과대학전
지공학과 졸업(학사)

1983년 서울대학교 대학원
전자공학과 졸업(석사)

1989년 Univ. of California,
Santa Barbara, Dept. of
Electrical and Computer
Engineering.(박사)

1983년~1995년 한국전자통신연구소 선임연구원

1990년~1993년 미국 Philips 연구소, Senior
Research Member

1995년~현재 광주과학기술원 정보통신공학과 교수
<관심분야> 디지털 신호처리, 영상 신호 처리 및
압축, 디지털 TV와 고선명 TV, 멀티미디어 시스
템, MPEG 표준, 3차원 TV, 실감방송