

사용자 환경에 적응적인 3DTV 시스템

준회원 백 윤 기*, 최 미 남*, 박 세 환*, 정회원 유 지 상*

3DTV System Adaptive to User's Environment

Yun-ki Baek*, Mi-nam Choi*, Se-whan Park* Associate Members,
Ji-sang Yoo* Regular Member

요 약

본 논문에서는 사용자의 시점과 환경을 고려한 3DTV 시스템을 제안한다. 제안된 시스템은 다시점 영상의 부·복호화기, 얼굴 추적, 2D/3D 동영상 입체 변환으로 구성된다. 다시점 영상에 입체감을 부여하기 위한 방법 중 실시간 처리가 가능한 2차원 영상을 3차원 영상으로 변환하는 2D/3D 동영상 입체 변환을 이용하여 입체 영상을 생성하였다. 기존의 2D/3D 동영상 입체 변환은 다시점 영상을 고려한 변환 방법이 아니기 때문에 다시점 영상의 시점 변화 시 다른 시점의 영상을 사용하여 불안정한 입체 영상을 생성한다. 본 논문에서는 이러한 문제점을 해결하기 위하여 다시점 영상에 적합하도록 시점 변화 시를 고려한 2D/3D 변환방법을 사용함으로써 시점 변화 시 안정된 입체영상을 생성할 수 있었다.

Key Words : 다시점 영상, 2D/3D 동영상 입체 변환, 얼굴 추적

ABSTRACT

In this paper, we propose a 3DTV system that considers user's view point and display environment. The proposed system consists of 3 parts - multi-view encoder/decoder, face-tracker, and 2D/3D converter. The proposed system try to encode multi-view sequence and decode it in accordance with the user's view point and it also gives a stereopsis to the multi-view image by using of 2D/3D conversion which converts decoded two-dimensional(2D) image to three-dimensional(3D) image. Experimental results shows that we are able to correctly reconstruct a stereoscopic view that is exactly corresponding to user's view point.

1. 서 론

수요 변화와 사용자들의 욕구 증대로 인해 CRT를 이용한 TV, 평면 브라운관, LCD, PDP 등의 소자들과 여러 영상 압축 기술들이 개발되었다. 그로 인해 현실감이 증대되었고 보다 나은 영상 표시가 가능하게 되었다. 이러한 2차원 영상의 발전에도 불

구하고 끊임없는 사용자의 욕구는 단방향의 2차원 고화질 영상에서 사용자의 의견을 반영할 수 있는 양방향 3차원 영상으로 변화 할 것이라는 것은 자명하다.

최근 3차원 기술 중 가장 활발한 활동을 하는 연구 분야는 다시점 영상의 부호화이다. MPEG에서는 다시점 영상과 관련된 표준화 작업을 58차 회의부

* 본 연구는 2007년도 교내 연구비 지원과 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업 (IITA-2006-C1090-0603-0017)의 연구결과로 수행되었습니다.

* 광운대학교 전자공학과 디지털 미디어 연구실 (byk100@image.gwu.ac.kr, minam@image.gwu.ac.kr, sehwanpark@image.gwu.ac.kr, jsyoo@daisy.kw.ac.kr)

논문번호 : KICS2006-12-540, 접수일자 : 2006년 12월 21일, 최종논문접수일자 : 2007년 9월 14일

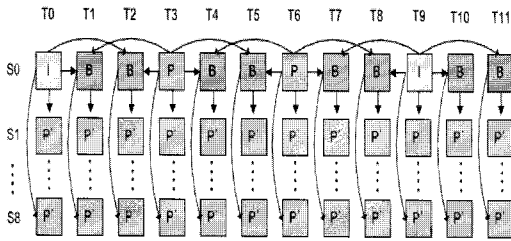


그림 2. 예측구조
Fig. 2 Estimation structure

가능해야 한다. 또한 복호화기 자원의 효율적인 관리, 부·복호화기 구현에 용이하도록 병렬처리가 지원되어야 한다.

본 논문의 시스템을 구성하기 위해서는 위 다시점 영상의 부호화와 관련된 사항을 만족해야 한다. 이러한 다시점 영상의 부호화 표준화가 진행 중에 있으며 JMVM(Joint Multi-view Video Model)이 참조 소프트웨어로 나왔다. 하지만 현 참조 소프트웨어는 실시간성과 자원의 관리 면에서 적절하지 못하기 때문에 본 논문에서는 시스템의 검증에 위하여 MPEG-2에 기반하여 그림 2와 같은 예측구조를 가진 부호화기를 사용하였다⁴⁾.

3.2 2D/3D동영상 입체 변환

본 논문에서는 이미 실시간 변환이 검증된 2차원 영상의 3차원 영상 변환 기법을 다시점 영상에 적합하게 구현하였다⁵⁾.

2D/3D 동영상 입체 변환의 기본 원리는 Ross가 발견한 정신물리학 이론인 Ross 현상에 기반하고 있다. Ross는 양안의 시간적 지연이 입체 비전의 단서(cue)라는 사실을 발견하였다⁶⁾.

예를 들어 Ross 현상을 설명하면 그림 3에서 영

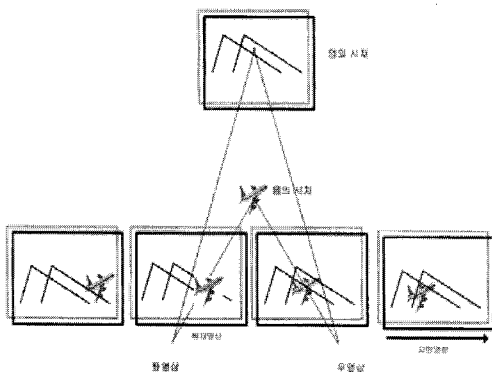


그림 3. Ross 현상 기반의 2D/3D 변환
Fig. 3 Ross phenomenon

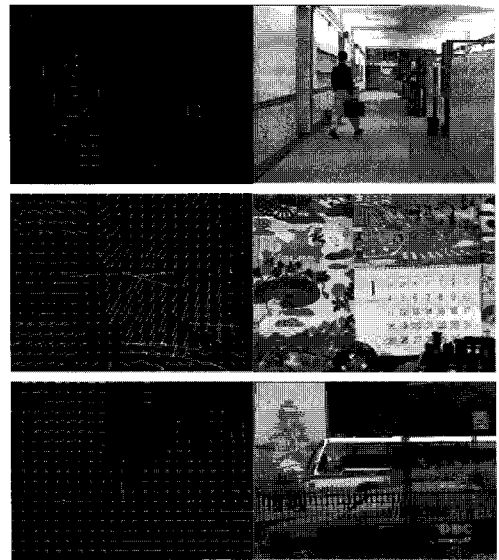


그림 4. 움직임 벡터
Fig. 4 Motion vector

상내의 비행기가 좌에서 우로 이동하고 배경이 되는 산은 고정되어 있다고 하면 좌영상으로 현재 영상을 사용하고 우영상으로 지연영상을 사용한다. 이때 좌영상과 우영상에 시차가 발생한다. 이로 인해 비행기는 화면의 앞쪽에 수렴점이 생기는 음의 시차가 발생하고 배경이 되는 산은 화면의 뒤쪽에 수렴점이 생기는 양의 시차가 발생하여 입체감을 느낄 수 있다. 이런 현상은 움직임이 수평운동의 경우에 적용되며 그렇기 때문에 우선적으로 각 프레임의 운동 유형을 파악하는 것이 필요하다. 운동 유형의 분

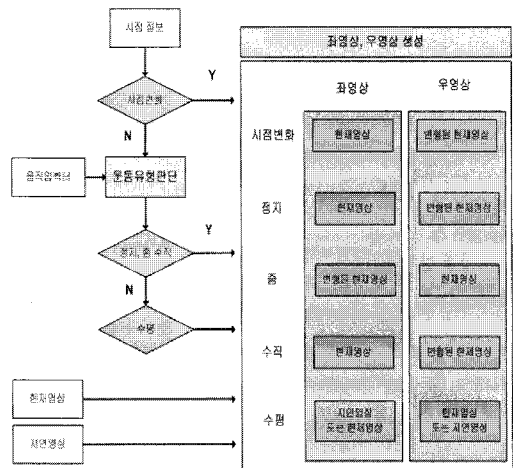


그림 5. 2D/3D동영상 입체 변환
Fig. 5 2D/3D stereoscopic conversion

류에 따라 수평운동을 제외한 운동 유형에 경우 지연영상이 아닌 그에 적합한 변환 방법이 필요하다.

운동유형의 판단을 위해 움직임 벡터를 분석하며 움직임 벡터는 그림 4와 같이 각 프레임의 운동 정보를 분석 할 수 있는 중요한 정보이다. 본 논문에서는 움직임 벡터 분석의 분석을 통해 운동유형을 정지, 줌, 수직, 수평 운동으로 구분하였다.

그림 5는 입체 변환의 전체적인 순서도로 현재 영상, 지연 영상, 움직임 벡터, 시점 정보를 이용하여 운동유형, 시점 변화를 판단하고 좌영상과 우영상을 생성하는 과정이다.

다시점 영상에서 시점의 변화 시 지연 영상을 가지고 좌영상 또는 우영상으로 사용하면 다른 시점의 이전 영상을 가지고 변환을 하게 된다. 스테레오 영상은 두 대의 카메라를 사용해서 획득한다. 이때 두 카메라간 거리는 일반적인 사람의 눈 사이의 거리 65mm와 비슷한 거리를 유지하게 된다. 하지만 다시점 영상은 보통 카메라간 거리가 65mm 이상의 간격으로 획득되기 때문에 다른 시점의 영상을 사용하여 변환된 영상을 사용하게 되면 입체감을 느낄 수 없게 된다. 본 논문에서는 이러한 문제점을 해결하기 위해 시점의 변화 시 그림 5와 같이 모든 운동유형에 우선하게 하여 타시점의 지연 영상의 사용을 방지한다. 변환은 현재 영상과 변환된 현재 영상을 사용한다. 이때 대부분의 영상이 위쪽보다 아래쪽이 더 가까이 있는 일반적인 영상의 특성을 고려하여 위쪽은 화면보다 들어가 보이게 하는 양의 시차를 주고 아래쪽은 화면보다 나와 보이는 음의 시차를 줌으로써 변환한다. 또한 영상의 중심에서부터 멀어질수록 이동값을 크게 하여 시차를 많이 주도록 한다.

카메라나 물체가 움직이지 않으면 정지 영상으로 판단한다. 정지 영상 형태의 입체 변환은 현재 영상과 변환된 현재 영상을 사용한다. 시점 변화 시와 같이 위쪽은 화면보다 들어가 보이게 하는 양의 시차를 주고 아래쪽은 화면보다 나와 보이는 음의 시차를 줌으로써 변환한다.

줌 영상은 영상이 점점 확대되거나 축소되는 장면을 연출한다. 따라서 줌 영상의 판단 방법은 영상 내에서 중심으로 향하는 움직임 벡터의 비율이 입계값보다 크게 되면 줌 영상으로 결정한다. 여기서 움직임 벡터의 방향은 식 (1)에서 dx와 dy가 이루는 각으로 결정한다. dx는 수평방향으로의 움직임 벡터, dy는 수직방향으로의 움직임 벡터를 나타낸다. 변환은 영상의 외곽에서 중심으로 가까워질수록

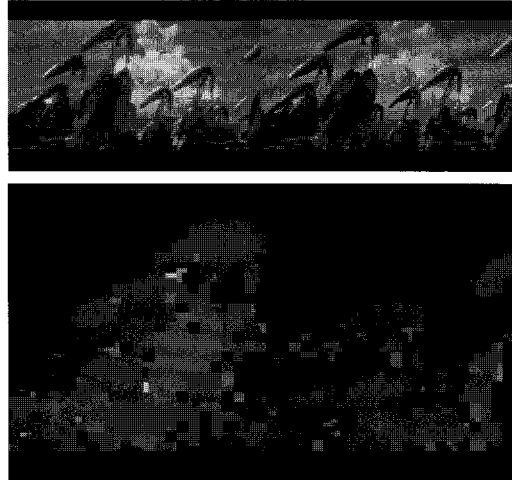


그림 6. 움직임 벡터의 세기값
Fig. 6 Motion vector intensity

큰 시차를 준다. 따라서 영상의 중심으로 가까워질수록 모니터 밖으로 나와 보이는 음의 시차가 크게 나타난다.

$$\theta = \tan^{-1} \frac{dy}{dx} \quad (1)$$

수직 영상의 결정은 수평 방향 움직임 벡터와 수직 방향 움직임 벡터가 이루는 각으로 결정한다. 수직 영상의 각의 결정 역시 식 (1)을 이용하여 방향을 결정한다. 영상의 중심 부분에서 물체들이 운동하고 있다고 가정할 때 중심에서 멀어질수록 양의 시차를 크게 두어 입체감을 느끼게 해준다.

정지 영상, 줌 영상, 수직 영상 형태가 아니면 수평 영상으로 판단한다. 수평 영상은 지연 영상과 현재 영상을 사용해 입체 변환한다^[7]. 좀 더 입체감을 높이기 위해 정지 영상에서 사용한 방법과 비슷하게 중심에서 위쪽으로 갈수록 지연 영상에 시차를 준다. 그림 6은 N번째 프레임과 N+1번째 프레임 그리고 두 영상 사이의 움직임 벡터의 세기값이다. 아래 그림에서 움직임 벡터의 세기값을 통해 볼 때 지연 영상과 현재 영상을 사용하였을 경우 적절한 깊이값을 반영하여 앞에 있는 물체의 시차가 크고 뒤쪽으로 갈수록 적은 시차가 발생한다는 것을 확인할 수 있다.

3.3 얼굴 추적

얼굴 추적의 첫 단계인 피부색 검출을 위해 여러 색상 공간에서 피부색을 검출하여 그 성분들을 재

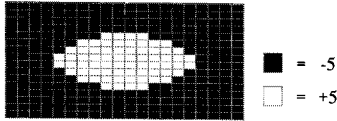


그림 7. 눈 형판
Fig. 7. Eye template

조합하여 사용한다[8][9]. 선택된 색상성분과 선택 범위는 식 (2)와 같으며, $f(x, y)$ 는 이진화 된 영상의 화소값이다.

$$f(x, y) = \begin{cases} 0 & \text{if } (127 < Cr < 170) \cap (110 < Q < 235) \cap (I < 70) \cap (H < 30) \\ 1 & \text{Otherwise} \end{cases} \quad (2)$$

검출된 피부색 영역 중에서 블록별 면적과 블록의 가로 세로 비율, 최외각 화소를 기준으로 얼굴과 비 얼굴을 1차적으로 가려낸다. 선택 되어진 얼굴 영역에서 얼굴을 검출하기 위해 본 논문에서는 특징점으로 눈과 입을 사용하였다. 눈 검출을 위해 그림 7과 같은 눈 형판을 이용하여 최대의 유사도를 갖는 눈 영역을 검출한다^[10].

2D/3D 동영상 입체 변환 시 안경식 모니터의 경우 입체안경을 착용하기 때문에 다른 특징점이 필요하게 된다. 입은 얼굴에서 상대적으로 적색 성분을 많이 포함하고 있다. 그러므로 Cr 성분에서 상대적으로 높은 화소 값을 갖고 Cb 에서 낮은 화소 값을 갖는다. 그러한 성질을 이용하여 Cr 성분을 부각하여 입 영역을 찾는다^[11].

IV. 실험 및 결과

그림 8은 캠(명준정보통신 알파캠 M2), 3D 모니터(DIMEN G170A 안경 식), 3D 안경, 사용자 인터페이스로 이루어진 전체적인 시스템 구성이다. 실험은 캠을 통해 입력 받은 사용자의 얼굴을 추적하여 사용자의 시점이 정확하게 해당하는 시점을 반영하는지를 확인한다. 이후 얼굴 추적을 통해 얻은 사용자의 시점정보를 이용하여 원하는 시점을 복호하고 시점 변화 시 2D/3D 동영상 입체 변환의 안정성에 관해서 확인한다.

제한한 시스템의 실험을 위해 사용한 영상은 그림 9, 그림 10의 Crossing People, Chatting 영상이다. 영상은 카메라 거리 20cm, 사이즈 352 × 480 인 9시점의 영상이다^[12].

그림 11은 사용자가 시점을 이동하면 그에 따라서 해당시점이 변화하는 결과를 보여준다. 그림

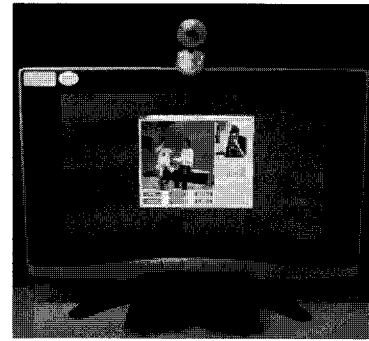


그림 8. 시스템 구성
Fig. 8 System structure

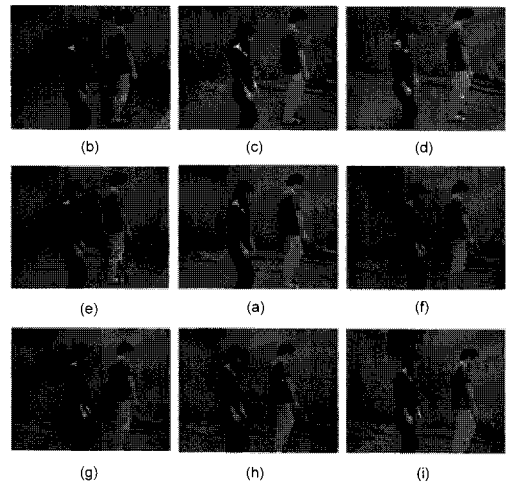


그림 9. Crossing People : (a)~(i) 1번 시점~9번 시점
Fig. 9 Crossing People : (a)~(i) 1st view~9th view

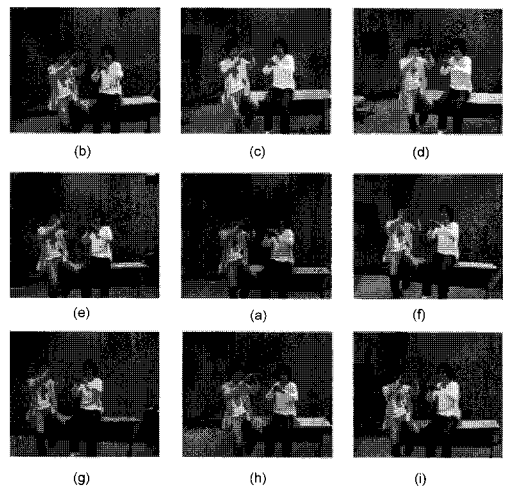


그림 10. Chatting : (a)~(i) 1번 시점~9번 시점
Fig. 10 Chatting(a)~(i) 1st view~9th view

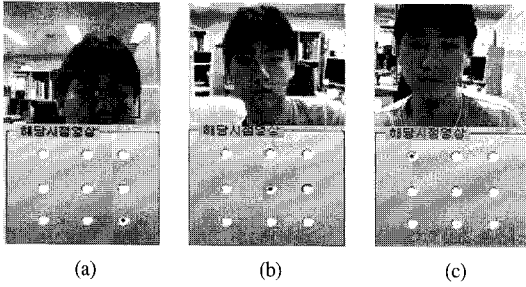


그림 11. 사용자의 시점에 따른 해당시점 변화: (a) 9번 시점 (b) 1번 시점 (c) 2번 시점
 Fig. 11 View point : (a) 9th view (b) 1st view (c) 2nd view

표 1. 얼굴추적 성능 검증

영상 개수/크기	113/200*200	시간(초)	0.609
검출률(%)	86%	검출률(%) 3D 안경착용	70%

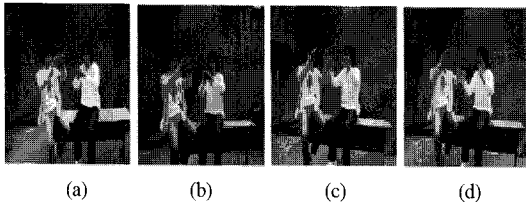


그림 12. 시점 변화에 따른 복호 된 영상 : (a) 1번 시점의 영상 (b) 5번 시점의 영상 (c) 6번 시점의 영상 (d) 4번 시점의 영상
 Fig. 12 Reconstructed frame : (a) 1st view image (b) 5th view image (c) 6th view image (d) 4th view image

11(b)에서 사용자의 얼굴의 중심이 중앙에 있기 때문에 해당시점 영상 역시 중심에 위치하고 있다. 그림 11(a)는 사용자가 좌측 하단에 위치하고 있고 해당시점도 좌측 하단에 있는 그림이고 그림 11(c)는 사용자가 우측 상단에 위치하고 있고 해당시점도 우측 상단에 위치하고 있다.

또한 얼굴 추적의 성능을 검증하기 위하여 그림 11과 같은 사용자의 얼굴이 포함된 영상 113 프레임을 테스트 하였으며 그 결과는 표 1과 같다.

검출시간은 평균 0.609초로 처리 시간이 짧기 때문에 사용자의 시점의 변화를 추적하기에는 적절하였다. 또한 검출률은 86%, 3D 안경 착용 시는 70%로 3D 안경 착용 시에도 1초 이내로 사용자의 시점을 확인 할 수 있다.

그림 12은 디스플레이 중 사용자가 시점을 변경하였을 때 시점에 해당하는 영상이 디스플레이 되는지를 확인한 실험 결과이다. 디스플레이 순서로 1

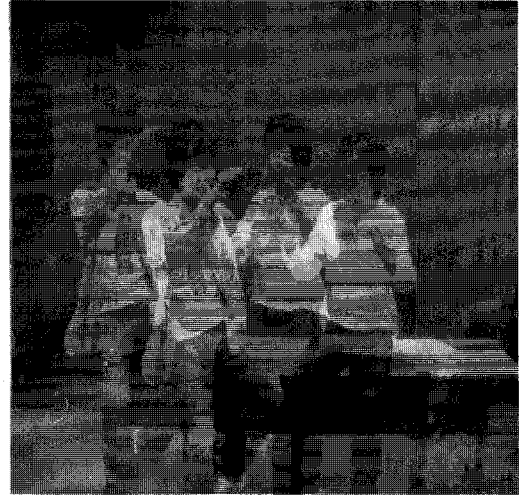


그림 13. 기존의 2D/3D변환의 시점 변경 시 입체 영상
 Fig. 13 3 dimensional image with conventional 2D/3D conversion



그림 14. 제안한 시스템의 시점 변화 시 2D/3D동영상 입체 변환
 Fig. 14 3 dimensional image with proposed 2D/3D conversion

번 프레임에서 1번 시점이 선택되어서 디스플레이 되어 지고 14번 프레임에서 5번, 40번 프레임에서 6번, 75번 프레임에서 4번 시점이 선택되어 디스플레이 되는 결과 영상으로 얼굴 추적을 통해 얻은 시점의 정보를 통해서 원하는 시점이 복호하는 것을 확인하였다.

2D/3D동영상 입체 변환 시 기존의 2D/3D변환 방법을 사용하게 되면 다른 시점의 지연영상을 사용하게 되며 그림 13과 같은 입체영상을 생성하게 된다. 다른 시점의 영상의 간격은 일반적인 스테레오 영상과 비교해 멀기 때문에 입체감을 느낄 수 없고 수직시차가 생겨서 눈에 피로감이 증가하게 된다.

그림 14는 다시점 영상에서 시점 변화 시 2D/3D 동영상 입체 변환 결과이다. 시점 변화 시 현재 영상과 변형된 현재영상을 사용함으로써 실험 결과와 같이 타시점과 겹치는 현상이 없이 안정적으로 입체 변환이 된 것을 확인 할 수 있다.

V. 결론 및 향후 연구 방향

본 논문에서는 9시점의 다시점 영상을 부호화하고 사용자의 시점에 따라 그 시점만을 복호하였다. 그리고 복호화 과정에서 사용되는 움직임 벡터의 분석을 통해서 2차원 영상을 3차원 영상으로 변환하는 3DTV 시스템을 제안하였다. 영상의 부호화를 위하여 기준이 되는 시점을 정하고 기준이 되는 시점을 참조하여 타시점을 부호화하였고 시점 추적을 통해 사용자의 시점 정보를 얻었다. 또한 스테레오 영상을 생성하기 위하여 실시간 처리가 가능한 2D/3D 동영상 입체 변환을 사용하였으며 기존의 2D/3D 입체 변환으로는 시점의 변화가 있을 경우 서로 다른 시점의 영상을 사용하여 안정적이지 못한 입체영상을 생성하는 문제점을 해결하기 위해 시점 변화 시 현재 영상과 변형된 현재 영상을 사용하여 다른 시점이 겹치는 것을 방지하여 안정된 입체 영상을 생성 할 수 있었다.

본 논문에서 제안한 시스템은 입체 영상을 생성하기 위하여 변위예측과 같은 복잡도가 높은 기법을 사용하지 않고 복호되는 움직임 벡터의 분석을 통해서 입체 영상을 생성하기 때문에 실시간의 시스템의 구성이 가능하다. 또한 얼굴추적과 같은 사용자와의 상호작용을 통해서 보다 실감있는 3DTV 시스템을 구성하였다.

본 논문의 시스템을 구성하기 위한 기술 요소들 각각은 많은 연구 기관들에서 현재에도 연구 중에

있고 이들 기술 요소가 향후 더욱 발전 된다면 제안한 3DTV 시스템의 성능도 더불어 향상 될 것으로 기대된다.

참고 문헌

- [1] MPEG/ISO/IEC JTC1/SC29/WG11, "Subjective Test Results for The CfP on Multi-View Video Coding," M13009, Bangkok, Jan. 2006.
- [2] Matsumoto, Y., Terasaki, H., Sugimoto, K., Arakawa, T., "Conversion System of Monocular Image Sequence to Stereo Using Motion Parallax," SPIE Photonic West, Vol.3012, pp.108-115, 1997.
- [3] 김만배, 박상훈, "키 프레임을 이용하는 2차원 동영상의 3차원 입체 변환," HCI 학회, 2002년 2월.
- [4] MPEG/ISO/IEC JTC1/SC29/WG11, "MVC Reference Software Implementation Plan", JVT-U061, Hangzhou, China, 20 - 27 October, 2006
- [5] 홍호기, 백윤기, 이승현, 김동욱, 유지상. "2D h.264 동영상의 3D 변환" 한국통신학회논문지, 제 31권 제 12C호, pp.1208-1215, 2006.
- [6] Ross, J., "Stereopsis by Binocular Delay," In Nature, Vol.248, pp.354-364, 1974.
- [7] Ross, J., Hogben, J. H., "Short-term Memory in Stereopsis," In Vision Research, Vol.14, pp.1195-1201, 1974.
- [8] Sobottka, K., Pitas, I., "Extraction of Facial Regions and Features Using Color and Shape Information," IEEE International Conference on Pattern Recognition, Vol.3, Aug. 1996.
- [9] Dai, Y., Nakano, Y., "Extraction of Facial Images from Complex Background Using Color Information and SGLD Matrices," Proc. of Int. Workshop on Automatic Face-and Gesture-Recognition. pp.238-242, 1995.
- [10] 한충신, 장세훈, 배진우, 유지상, "다시점 영상 시스템을 위한 얼굴 추적," 한국통신학회 논문지, 제 30권 제 2C호, pp.16-24, 2005년 2월.
- [11] Rein, L. H., Mohamed, A. M., Anil K. J., "Face Detection in Color Images," IEEE Transactions on Pattern Analysis and

Machine Intelligence, Vol.24, pp.696-706, May 2002.

[12] Keio Multi-View Video Coding(MVC) Testing Data, <http://www.ozawa.ics.keio.ac.jp/~moriyama/mpeg-3dvideo/index.htm>

백 윤 기 (Yun-ki Baek)

준회원



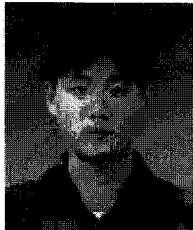
2006년 2월 광운대학교 전자 공학과 졸업

2006년 3월~현재 광운대학교 전자공학과 석사과정

<관심분야> 3차원 영상처리, 동영상 코덱

최 미 남 (Mi-nam Choi)

준회원



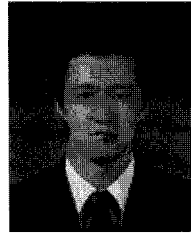
2006년 2월 광운대학교 전자 공학과 졸업

2006년 3월~현재 광운대학교 전자공학과 석사과정

<관심분야> 3차원 영상처리, 동영상 코덱

박 세 환 (Se-whan Park)

준회원



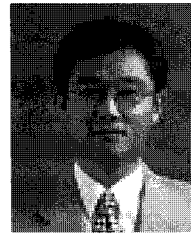
2007년 2월 광운대학교 전자 공학과 졸업

2007년 3월~현재 광운대학교 전자공학과 석사과정

<관심분야> 3차원 영상처리, 동영상 코덱, IVR

유 지 상 (Ji-sang Yoo)

정회원



1985년 2월 서울대학교 전자 공학과 졸업(공학사)

1987년 2월 서울대학교 대학원 전자공학과 졸업(공학석사)

1993년 5월 Purdue 대학교 전기 공학과 졸업(Ph.D.)

1993년 9월~1994년 8월 현대전

자산업(주) 산전연구소 선임연구원

1994년 9월~1997년 8월 한림대학교 전자공학과 조교수

1997년 9월~2001년 8월 광운대학교 전자공학과 조교수

2001년 9월~2006년 8월 광운대학교 전자공학과 부교수

2006년 9월~현재 광운대학교 전자공학과 교수

<관심분야> 웨이블릿 기반 영상처리, 영상압축, 영상인식, 비선형 신호처리