

3D 내시경 영상시퀀스의 MPEG-2 코딩 효율에 관한 연구

송철규, 이영목, 이상민, 김원기, *이제호, **이명호
삼성생명과학연구소 임상의공학센터,
*삼성의료원 산부인과, **연세대학교 전기공학과

A Study on the MPEG-2 Coding Performance of 3D Endoscopic Image Sequence

C. G. Song, Y. M. Lee, S. M. Lee, W. K. Kim, *J. H. Lee, and **M. H. Lee
Biomedical Engineering Center, Samsung Biomedical Research Institute
*Dept. of OBGY, Samsung Medical Center
**Dept. of Electrical Engineering, Yonsei University

ABSTRACT

In this study, for the coding of stereoscopic video sequence, two approaches are presented based on simulcast method and sidefield image format. The field sequential method for stereoscopic visualization have been specified, Also, camera parameter and shooting conditions for each test sequence are studied. Coding method based on sidefield format structure revealed better performance over simulcast in PSNR.

서론

미래의 화상정보 전달은 보다 더 자연스럽게 친밀감을 줄 수 있는 3차원 영상기술의 개발을 요구한다. 이러한 3차원 영상기술의 활용분야로는 위성방송, 외과적 수술 및 시뮬레이션, 오락, 스포츠 레저 등 범위가 다양하며, 이를 구현하기 위한 3차원 영상매체 기술은 액정서터 방식, 렌티큘러 방식, 홀로그래픽 방식 등이 있다. 홀로그래픽 방식은 지금은 필름으로 부터 입체영상을 얻는 것이 가능할 뿐 인터랙티브 영상을 실시간으로 제어하기 위해 고가의 장비를 필요로 한다. 또한 2개의 초소형 LCD를 이용한 HMD(head mounted display)를 이용하는 경우, 1280x1024 해상도와 120도의 시야각의 제품이 등장하고 있으나 가격이 수천만원대에 이르고, 착용무게로 인한 작업상의 제약이 따른다. 현재의 기술로는 해상도와 시야각, 정보처리 속도 및 광학적 셋업문제를 고려할 때 액정서터 방식이 가장 유리하다.

3차원 영상의 정보량은 상황에 따라 다르지만 물체의 한 시점이 가지는 정보는 관측 대상, 콘트라스트, 색, 시간정보 등을 추가하면 기존의 전송 매체를 통해 처리 할 수 없는 만큼 방대한 양이 된다.

CCIR 601 포맷인 경우, 전체적인 비트율이 약 250Mbps(720[HT]x480[VT]x8[Bits]x3[R,G,B]x30[frame/sec])가 되며, 양안 입체 영상 데이터로 처리할 경우에는 500 Mbps가 된다. 비압축시 스펙트럼 효율성이 2 bit/Hz인 경우라면 요구되는 기저대역폭은 250 MHz가 된다. 따라서 2차원 압축 및 코딩방법과는 달리 해상도에 따라 방대한 정보량이 생성되므로 효율적인 중복성 제거와 신호대역폭 압축, 영상처리 및 코딩기술이 요구된다.

본 연구에서는 3D 내시경 영상시퀀스에 대한 부호화 효율을 분석하기 위하여 입체비디오 시스템을 구성하고, 얻어진 입체 내시경 비디오를 필드시퀀셜 포맷으로 처리한 후, MPEG-2 대역폭 제한별 코딩효율을 분석하였다. 비교를 위한 표준 입체영상으로는 프랑스 CCETT사[1]의 MPEG-2 영상 시퀀스를 사용하여 Simulcast 방식코딩을 수행하였고, 입체 내시경 영상에 대해서는 sidefield 포맷 변환 후, MPEG-2 코딩을 실행한 후 신호대 잡음비와 x, y축 방향 상관관계 등의 분석을 통해 코딩 효율을 실험하였다.

3D 내시경 영상 시스템

1. 시스템의 특징

복강경을 이용한 현재의 외과수술은 체내에 삽입하는 내시경의 2차원 수술영상을 모니터로 관찰하면서 수술이 진행되므로[2] 의사 자신등이 갖고 있는 해부학적 지식 및 경험을 바탕으로, 경우에 따라서는 탐촉자에 의해 화면의 깊이를 판단하면서 수술을 실행하고 있다. 따라서 비교적 간단한 수술작업은 2차원 복강경의 관찰하에서도 수행이 가능하나 보다 복잡한 수술조작은 곤란하고, 장시간을 요하거나 수술의 안전성 및 확실성에 대한 점에서는 문제가 되고 있다. 이러한 점들을 해결하는 3D 내시경은 체내에 삽입되는 telescope가 2채널 렌즈계로 구성이 되

고, 이를 3D 카메라와 결합시킴으로써 수술부위를 3차원으로 볼 수 있는 특수 내시경으로 시술의사에 실제의 입체영상을 제공하여 편리한 수술조작을 수행할 수 있도록 하는 장치이다. 본 연구에서는 3D 카메라와 3D telescope, 그리고 LCD 셔터형의 시분할 방식 제어기로 구성되는 시스템으로 내시경 수술 후 얻어진 3D 비디오 시컨스를 이용하였다.

2. 시스템의 구성과 입체 파라미터

본 연구에서 적용한 3D 내시경의 특징은 8°의 시차각을 가진 2중 렌즈계 telescope와 3D 카메라(독일, Wolf사)로서 수술부위 좌,우 영상을 얻는다. 그리고 좌, 우 영상은 모니터 상에 좌, 우의 영상출력을 디스플레이 하기 위해 Micom제어에 의해 1/120 Hz의 주기로 절환을 하게된다. 이때, 고글(goggle) 구동 회로를 절환주기에 동기시켜 주면 LCD 안경을 통해서 각각 좌, 우의 영상이 눈에 전달되며, 두 눈 사이의 불일치(binocular disparity)때문에 3D 내시경 영상을 지각하게 된다.

필드시컨셜 방식의 3D 영상의 경우 하나의 이미지를 구성하는 주사선의 수가 525개에서 절반으로 감소함에 따라 해상도가 그만큼 떨어지나, 이를 LCD 셔터 안경을 통해 사람의 눈으로 볼 때는 거의 구별할 수 없다. 본 연구에서는 피로감을 느끼지 않도록 하기 위해 최대 양의 수평시차(vertical parallax)와 음의 수평시차가 2 cm를 넘지 않아야 됨을 확인하였다. 그림 2는 필드 시컨셜 디스플레이를 위해 입체 영상저장 포맷을 sidefield 방식으로 구성하는 것을 나타내고 있다. 그림 1(a)와 (b)는 3D 카메라와 입체 디스플레이와 관련한 파라미터를 각각 도시하고 있다[3]. 그림 1(a)는 3D 카메라 시스템을 나타내고 있으며, 구성요소로는 카메라간의 거리(2.2mm), 렌즈로부터 카메라간 광학 축이 교차하는 지점까지의 거리인 convergence 거리, 그리고 렌즈의 초점거리와 CCD의 크기로 결정되는 시야각으로 구성이 된다.

그림 1(b)는 디스플레이 시스템을 결정하는 3가지 파라미터가 관측거리, 디스플레이 시스템의 크기, 그리고 관측자 두눈의 거리로 구성이 된다.

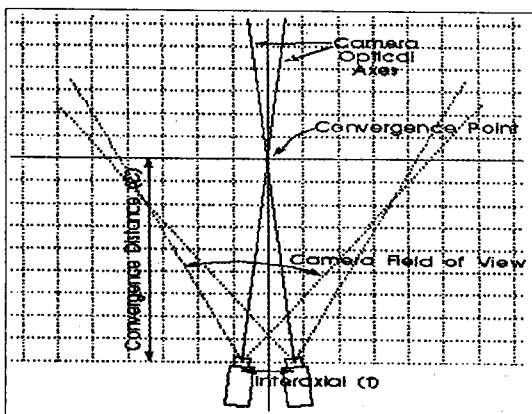


그림 1(a) 3D 카메라 시스템의 parameter

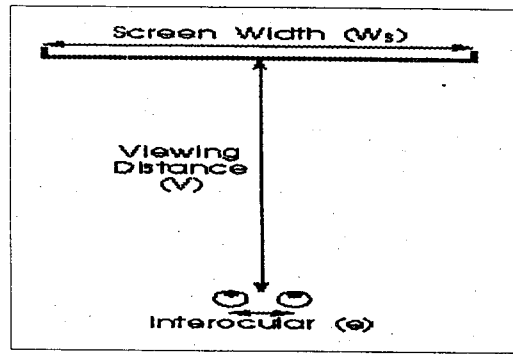


그림 1(b) 3D 디스플레이 파라미터

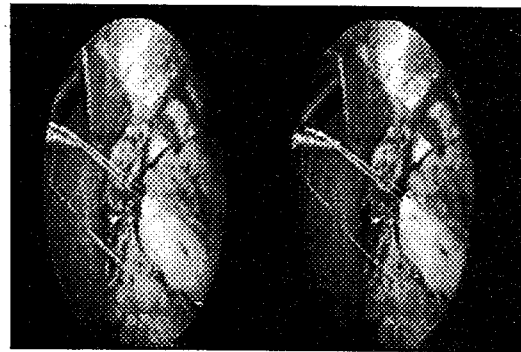


그림 2 입체 내시경 수술영상의 sidefield 포맷

입체내시경 카메라 및 디스플레이 시스템의 중요한 3D 파라미터는 다음과 같다.

- baseline 거리(t = 두 카메라 렌즈 중심간의 거리)
- C = Convergence 거리, $t/(2\tan[\beta+\arctan(h/f)])$
- f = 두 카메라 렌즈간의 초점거리
- Wc = CCD width(3D 내시경의 경우는 6.4 mm)
- β = camera convergence 각도
- h = CCD센서의 광학축 offset
- α = 시야각 = $\arctan[(Wc/2+h)/f+\arctan[Wc/2-h)/f]$
- V = 관측거리
- e = 양안거리(65 mm)
- Ws = 디스플레이 스크린의 수평 사이즈

Sidefield 영상의 필드 시컨셜 처리

3D 비디오 영상을 표준 NTSC영상으로 변환하는 방법은 여러 가지가 있으나 필드 시컨셜 방법과 사이드 필드 방법이 가장 효율적이다. sidefield 방법은 한 프레임내에 좌, 우영상을 수평축으로 각각 1/2로 압축하여 한 프레임 내에 좌, 우 영상을 나란히 저장하는 방법이다. 또한 120 Hz 필드시컨셜 방법은 우측 눈에 해당하는 영상을 odd필드에, 좌측 눈에 해당하는 영상을 even필드에 각각 비디오 동기를 시키는 방법이다. 본 연구에서는 VCR에 녹화시에는 sidefield 방법으로 저장을 하고, 모니터에 재생할 때

에는 그림 3과 같이 subfield구조의 필드 시퀀셜 영상으로 변환하였다[4].

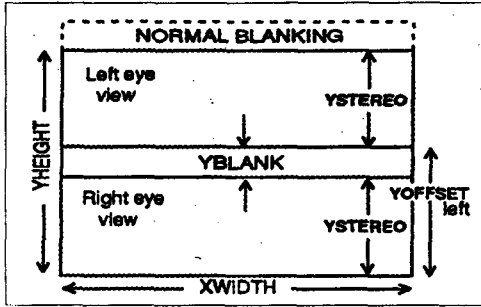


그림 3 필드 시퀀셜 디스플레이

그림 3에서 좌,우에 해당하는 상, 하의 subfield영상은 YBLANK에 의해 분리되며, 두 눈에 보이게 되는 영상의 크기는 XWIDTH(x축), YSTEREO(y축)가 된다.

$$YSTEREO = \frac{YHEIGHT - YBLANK}{2} \quad (1)$$

우측 눈에 보이는 영상의 YOFFSET_{right}는 0 이고, 좌측 눈에 보이는 영상의 YOFFSET_{left}는 식 (2)가 된다.

$$YOFFSET_{left} = YSTEREO + YBLANK \quad (2)$$

결국, 우측 눈에 보이는 영상의 좌표는 식(3)과 같고

$$\left(XWIDTH \frac{x - x_{min}}{x_{max} - x_{min}}, YSTEREO \frac{y - y_{min}}{y_{max} - y_{min}} \right) \quad (3)$$

좌측 눈에 보이는 영상의 좌표는 아래의 식이 된다.

$$\left(XWIDTH \frac{x - x_{min}}{x_{max} - x_{min}}, YSTEREO \frac{y - y_{min}}{y_{max} - y_{min}} + YOFFSET_{left} \right)$$

실험결과 및 고찰

대역폭 제한에 따른 코딩효율을 실험하기 위해서 참고로 사용된 표준 입체영상 시퀀스와 입체 내시경 시퀀스의 사양은 표 1과 같다.

먼저 표준 영상시퀀스 Discussion, Manege, Train의 포맷은, CCIR 601포맷으로 720x480의 해상도와 초당 30 frame으로 구성되며, 각 프레임은 2개의 필드와 4:2:2의 크로마 포맷을 갖는다.

본 연구에서는 입체영상 시퀀스 코딩을 위해 기존의 simulcast방식을 적용하였다. 이 방식은 압축 및 전송을 좌, 우 영상 채널별로 각각 처리하는 방법으로 좌측영상에 대한 예측은 좌측영상 자체만을 가지고 하며, 우측 영상도 우측만을 이용한다. 따라서 좌, 우 영상간의 높은 상관관계를 갖는 공간적 redundancy[5]를 전혀 이용할 수 없는 단점이 있다. 그리고 입체 내시경 영상을 sidefield 포맷으로

표 1 CCETT 입체 영상 시퀀스와 3D 내시경 영상시퀀스의 카메라 셋업 파라미터

Sequence	Discussion	Manege	Train	3D 내시경 영상
시퀀스 길이 (frames)	30	30	50	450
축적거리	20 mm	30 mm	40 mm	3 - 20mm
Baseline거리	30 cm	30 cm	8.75 mm	22 mm
Converging Distance	3 m	9.2 m	2.8 m	0.1 m
Object Range	6 -7 m	12m~∞	unkown	0.1~0.5 m

하나의 프레임 내에 좌, 우영상이 있는 비디오 시퀀스를 만들고, 이를 이용하여 MPEG-2 전송 대역폭을 변화시키면서 가변적 부호화를 실행하였다.

그림 4는 전송 대역폭을 변화시키면서 표준 입체 영상에 대해서 얻은 simulcast방식의 PSNR과, sidefield 방식을 적용한 입체 내시경 영상 시퀀스에 대해 휘도 PSNR을 비교 하였다. 입체 내시경 수술 영상 시퀀스는 사이드필드 포맷으로 MPEG-2 부호화시, 단일 GOP내의 프레임 개수 N=15로 하고 P 픽처간 프레임 수는 M=3으로 하였다. 표준 영상시퀀스에서 얻어진 PSNR 값은 좌, 우 영상의 평균치이며, 우측 영상이 좌측 영상 보다 5 % 이상의 PSNR비가 높았다.

표준 입체영상 중에 전송 대역의 증가에 따른 신호대 잡음비 성능이 가장 높은 시퀀스는 Train이고 가장 낮은 성능을 나타내는 시퀀스는 Manege임을 알 수 있다. 전송 대역폭이 커질수록 PSNR의 값은 급격히 증가하지만, 실제 실험에서 대역폭이 10 Mbps 이상 증가할수록 처리속도의 급격한 감소를 보였다. 디코딩된 입체 내시경 비디오 시퀀스를 재생하였을 때, 부호화에 따른 화질의 점차적인 차이를 확인 하였다. 그림 5는 입체 내시경 영상에서 (a)는 좌,우 영상이 120 Hz의 수직 주파수로 필드 시

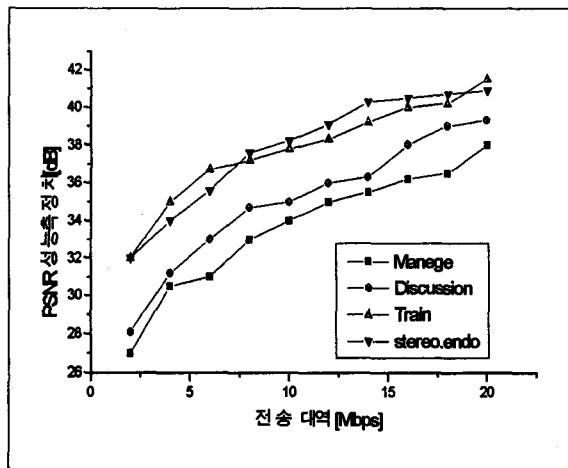
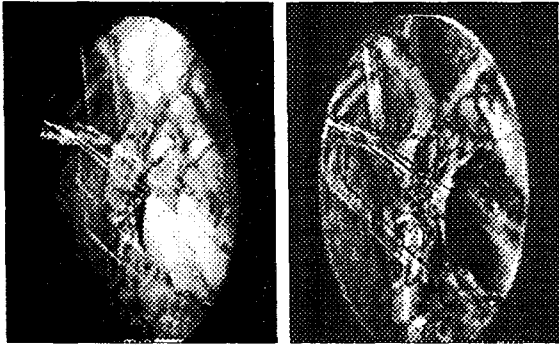


그림 4 표준 입체영상과 3D 내시경 영상의 전송대역 변화에 따른 부호화성능

컨실하게 오버레이 될 때의 영상이고, (b)는 좌, 우 영상의 차에 의한 disparity영상이다.



(a) 오버레이 된 영상 (b) Disprity 영상
그림 5 입체 내시경 영상의 특성

[5] S. Sethuraman, M.W. Siegel, and A.G. Jorgan, "A Multiresolution Framework for Stereoscopic Image Sequence Compression," Proc. ICIP'94, pp. 361-365, Oct. 1994

결론

본 연구에서는 3D 내시경 영상시퀀스에 대한 MPEG-2 코딩 효율을 실험하였으며 얻어진 결과는 다음과 같다.

첫째, 표준 입체영상 중에 대역폭 증가에 따른 신호대 잡음비 효율이 가장 높은 시퀀스는 Train이고 가장 낮은 시퀀스는 Manege임을 알 수 있었다.

둘째, 입체 내시경 영상을 sidefield 포맷으로 변환한 후 부호화 시킨 경우의 PSNR은 표준영상 시퀀스에 대해 simulcast방식으로 코딩한 경우보다 모든 경우의 전송대역에 대해 전체적으로 더 높게 나타났다. 셋째, 좌, 우 내시경 영상으로 부터 얻어진 disparity 영상에서 x축, y축 방향의 correlation 값을 구한 결과, x축 방향으로의 상관관계가 더 높은 특징을 보였다.

참고문헌

[1] CCETT, "Description of Stereoscopic Test Sequences: Train and Manege," Oct., 1994
 [2] C. G. Song, Y. M. Lee, S.M. Lee, W. K. Kim, J.H. Lee, and M. H. Lee, "Implementation of real-time digital endoscopic image processing system, Proc. SPIE, Vol. 3164, 1997
 [3] A. Woods et. al," Image Distortion in Stereoscopic Video Syetms,"Proc. SPIE, Vol. 1915, 1993
 [4] L. F. Hodges, "Basic Principles of stereoscopic software development," Proc. SPIE, Vol. 1457, 1991