

고화질 스테레오스코픽 콘텐츠의 모바일 콘텐츠로의 변환에 관한 연구

이봉호, *박영수, 정원식, 허남호, 남승필**

한국전자통신연구원, *과학기술연합대학원대학교, 두두원**

leebh@etri.re.kr

A Study on the Resizing of HD Stereoscopic Contents to Mobile Contents

Bongho Lee, Youngsoo Park, Won-sik Cheong, Namho Hur, Seung-pil Nam

요 약

DMB 와 같은 모바일 기기에서 3 차원 입체 영상 서비스를 제공하기 위해서는 스테레오스코픽 콘텐츠를 생성하는 단계에서 모바일 기기의 특성, 즉 3D 디스플레이의 특성을 고려하여 촬영하거나 또는 가공하여야 한다. 본 논문에서는 일반적으로 HD 급 고해상도로 촬영된 스테레오스코픽 3 차원 콘텐츠를 모바일 환경에 적합하도록 크기를 변환하기 위한 알고리즘 및 그에 따라 변환된 모바일 3 차원 영상의 분석 결과를 제시하고자 한다.

1. 서론

2009 년 영화 아바타(Avatar)의 세계적인 흥행 성공 이후, 3 차원 영상 서비스에 대한 기대감이 높아졌다. 영화뿐만 아니라 방송 및 기타 영역에서도 3 차원 서비스를 제공할 수 있는 기술적인 가능성이 확인된 바 있으며, 이미 디지털 방송에서는 프레임 호환 방식의 3D 방송이 상용 서비스로 제공되고 있다. 최근에는 보다 더 좋은 고화질의 3D 서비스를 제공하기 위한 움직임으로 서비스 호환 방식에 대한 기술 개발 및 표준화 작업이 전 세계적으로 진행 되고 있다.

모바일 방송의 경우, 국내에서는 DMB (Digital Multimedia Broadcasting) 방송이 이미 활성화 되었고, 3D 방송 서비스로의 확장도 용이한 상태이다. 또한 스마트폰의 출현으로 DMB 와 같은 방송망이 아닌, 이동통신이나 무선 인터넷 망을 통해 3D 콘텐츠를 보다 쉽게 전달할 수 있는 환경이 조성되었다. 단말 측면에서 보면, 3D 디스플레이를 지원하는 DMB 폰이 이미 출시 된 바 있고, 3D 영상 디스플레이와 콘텐츠 기술의 성숙과 더불어, 향후 다수의 3D 카메라 및 3D 디스플레이를 탑재한 3D 스마트폰이 출시될 예정이다.

이러한 시장의 추세 및 상용 서비스를 활성화하는 측면을 고려해보면, 무엇보다도 중요한 것은 3D 콘텐츠의 다양화인데, 아직은 즐기고 감상할 수 있는 3D 콘텐츠가 절대적으로 부족한 실정이다. 이는 3D 콘텐츠 제작의 어려움에 기인하고 있으며, 그나마도 대부분의 3D 콘텐츠가 고해상도의 디스플레이를 통한 감상을 목적으로 촬영 또는 생성되고 있기 때문이다. 따라서, 모바일 단말을 통한 3D 콘텐츠의 감상을 위해서는 무엇보다 3D 콘텐츠의 다양화가 이루어져야 하고, 가장 쉬운 해결방법은 고화질 3D 콘텐츠를 모바일 3D 단말에 볼 수 있도록 변환하는 것이고 이를 통해 OSMU(One Source Multi Purpose)에 의한 3D 영상 콘텐츠의 다양화를 이룰 수 있다.

3D 영상의 경우는 2D 영상과는 달리 단순하게 해상도의 변경만으로는 전반적인 품질을 만족시킬 수 없다. 이는 3D 영상이 갖고 있는 깊이감 인지에 따른 특징으로 고해상도 특히 HD (1920x1080, High Definition)급 3D 영상을 모바일 단말 화면의 크기로 단순하게 크기 변환할 경우, 3D 모바일

단말의 특성에 따라 HD 영상의 그 것과는 다른 입체감이 느껴지고, 눈의 피로감을 유발할 수 도 있다. 따라서 모바일 단말을 통한 3D 콘텐츠 감상을 위해서는 이러한 점을 고려한 크기 변환 방법이 마련되어야 한다[1].

본 논문에서는 DMB 단말을 목표 디스플레이 및 시청 환경으로 가정하고 HD 급 고해상도 3D 영상을 DMB 단말을 통해 편안하게 볼 수 있도록 하는 크기 변환 방법의 개념 및 기술에 대해서 소개하고자 한다. 2 장에서는 3D 크기 변환 방법의 개요를 설명하고 3 장에서는 DMB 기반 3D 크기 변환 방법을, 마지막으로 4 장에서는 3 장의 기술을 이용하여 실제로 크기 변환이 된 결과 및 영상에 대한 분석을 다루고자 한다.

2. 3D 크기 변환 개요

3D 크기 변환은 3 차원 입체 영상을 목표하는 해상도 및 원하는 영상 품질에 적합하도록 바꾸는 것으로 정의할 수 있다. 3D 디스플레이는 시청자에게 비교적 안정적으로 입체감을 제공할 수 있는 변위 범위(disparity range)를 가지고 있다. 일반적으로 이 변위 범위를 넘어서 입체감이 표현되면, 시청자는 눈의 피로감이나 불편함을 느낄 수 있다. 만약, 이러한 3 차원 영상의 특성이 고려되지 않고, 2D 영상과 동일한 방법의 크기 변환이 이루어지면, 3D 영상을 감상할 때 크로스톡(crosstalk) 및 고스팅(ghosting) 현상이 나타날 수 있고, 카드보드 효과와 같은 입체감 왜곡이 발생하기도 한다 [2].

본 논문에서는 이러한 특징을 고려하여 모바일 단말에서 입체감을 보다 더 편안하게 감상할 수 있는 DMB 기반의 3 차원 입체 영상 크기 변환 기법을 다음과 같이 소개하고자 한다.

3. DMB 기반 3D 크기 변환

본 장에서는 DMB 시청 환경에 적합한 3D 콘텐츠 크기 변환 방법을 제안한다. 일반적으로 크기 변환을 위해서는 목표

단말의 물리적 특성, 입체감에 영향을 미치는 disparity 정보

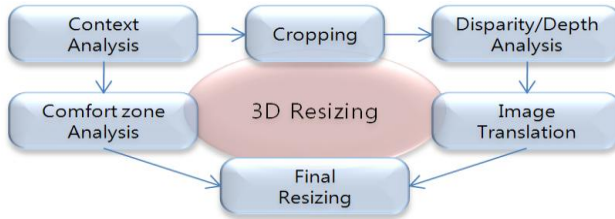


그림 1. 3D 리사이징 시 고려해야 할 기능 구성도

획득 및 안정 영역 (comfort zone)에 대한 분석이 선행되어야 한다. 구체적으로는 그림 1 과 같이 도식화 될 수 있으며 크게 이미지 분석, 영상 자르기, 안정 영역 및 변위 계산, 이미지 변형 및 영상 처리를 포함한 크기 변환으로 나누어 볼 수 있다. 영상 자르기 및 안정 영역 분석 단계에서 분석된 매개변수를 이용하여 실제로 영상의 크기를 변환하는 데에는 다양한 방법이 존재하지만 본 논문에서는 2D 영상 기반의 크기 변환, HIT(Horizontal Image Translation)기반 크기 변환 및 깊이 지도 (depth map)기반 변환 방법에 대해서 기본 구조와 특징 및 장단점에 대해서 분석하고자 한다.

3.1 영상 자르기 (Cropping)

고화질의 3D 콘텐츠를 모바일 단말에서 보기 위해서는 모바일 단말의 화면에 맞도록 영상의 크기를 변환시켜줘야 한다. 영상의 크기를 변환하는 방법에는 영상 자르기 (cropping), 선형적 영상 변환, 비선형적 영상 변환 등이 있다. 본 논문에서는 영상 자르기 방법과 선형적 영상 크기 변환 방법에 대해서만 고려해 보기로 한다. 영상의 크기 변환을 위해서 가장 많이 쓰는 방법은 선형적인 영상 크기 변환 방법으로 이 방법을 사용할 경우, 고화질 3D 콘텐츠의 내용 모두를 모바일 영상의 화면에 담을 수 있는 장점이 있다. 하지만 고화질 콘텐츠와 모바일 단말의 가로/세로 비율 (aspect ratio)이 서로 다를 경우, 이에 따른 영상의 왜곡이 발생하게 된다. 영상 자르기 방법의 경우는 모바일 단말의 화면 크기에 맞게 영상을 잘라내기 때문에 가로/세로 비율에 따른 영상의 왜곡은 없지만, 잘려나간 부분으로 인한 영상 정보의 손실이 일어난다. 따라서 본 논문에서는 이 두 방법의 장점을 살리고자, 모바일 단말의 가로/세로 비율에 맞추어서 영상 자르기 방법을 사용한 후, 이를 다시 선형적 크기 변환 방법을 통해 모바일 단말의 크기에 맞도록 크기를 줄이는 방법을 사용하였다. 단순한 2D 영상의 크기 변환을 위해서는 이 방법들을 통해서도 만족할 만한 결과를 얻을 수 있지만, 3D 스테레오스코픽 영상의 경우에는 고려해야 할 점이 더 많다. 특히 입체감을 편안하게 볼 수 있는 안정 영역을 벗어난 부분이 많이 생기지 않도록 미리 영상의 크기 변환 과정에서 그 부분을 잘라내는 방법도 필요하고, 입체감을 느끼는데 방해를 하는 스테레오스코픽 윈도우 위반 (Stereoscopic Window Violation)을 피할 수 있도록 영상을 잘라내는 방법에 대한 고려도 필요하다.

3.2 안정 영역 분석 (Comfort Zone Analysis)

시청자들이 3D 콘텐츠를 감상할 때, 가장 문제가 되는 부분은 눈의 피로감을 쉽게 느낄 수 있다는 점이다.

스테레오스코픽 입체 영상을 감상할 때 눈의 피로감을 느끼는 원인에는 여러 가지가 보고되고 있지만, 가장 큰 원인은 일반적인 자연환경과는 달리 눈의 주시와 폭주가 서로 일치하지 않는데 따른 것이다. 우리가 보고자 하는 물체가 화면 상에 위치한 것처럼 보이는 상태를 ZPS (Zero parallax Setting)라 하는데 이 경우에는 왼쪽과 오른쪽 영상 간의 변위 (disparity)가 0 이 되어 일반적인 자연 환경처럼 주시와 폭주가 일치하게 된다. 하지만 입체감을 표현하기 위해서 왼쪽과 오른쪽 영상 사이에 변위가 생기면 우리가 보고자 하는 물체는 화면보다 앞, 또는 뒤에 위치하고 있는 것으로 인지된다. 이때 너무 과도하게 화면 앞쪽으로 나와 보이거나, 뒤쪽으로 들어가 보일 경우 눈의 피로감을 유발하게 된다. 따라서 눈의 피로감이 적은 3D 콘텐츠를 위해서는 물체가 앞으로 나와 보이거나 뒤로 들어가 보이는 정도에 따른 변위 값을 일정한 한도로 제한하여야 하는데 이 한도 안의 변위 값을 갖는 영역을 스테레오스코픽 영상 콘텐츠의 안정 영역 (comfort zone)이라고 한다. 안정 영역에 대한 문제는 3D 콘텐츠의 제작단계에서부터 고려되어 제작되고 있지만, 고화질 3D 콘텐츠를 모바일 단말에 맞도록 변환하는 과정에서 다시 발생하게 된다. 안정 영역이라는 것이 시청자들의 입체감을 인지하는 능력뿐만 아니라 화면의 크기, 시청 거리와 같은 시청 환경에 따라서도 달라지기 때문이다. 따라서 콘텐츠의 변환 과정에서 달라지는 시청환경에 따른 안정 영역의 분석도 필요하다.

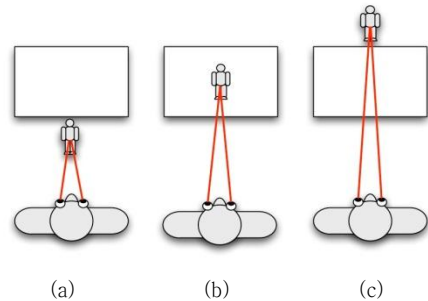


그림 2. 변위에 따라 다르게 인지되는 물체의 깊이감: (a) negative parallax; (b) ZPS; (c) positive parallax.

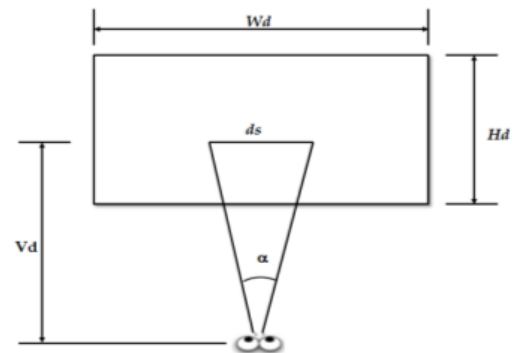


그림 3. 화면 크기와 시청 거리에 따른 변이 값.

그림 3 은 화면의 크기와 시청거리에 따른 입체 영상의 변이 값에 대해 나타내고 있다. 최적 시청 거리는 화면 높이의 3 배에 해당하는 지점이라고 할 때 [3], 화면 크기와 시청 거리에 따른 변이 값 d_s 은 다음과 같이 구할 수 있다.

$$d_s = 6 \cdot L_x \left(\frac{H_d}{W_d} \right) \cdot \tan\left(\frac{\alpha}{2}\right) \quad (1)$$

여기에서 L_x 는 영상의 가로 방향의 화소 수이고, H_d, W_d 는 각각 화면의 길이와 높이를 나타낸다. 또 각 α 는 눈의 위치로부터 화면의 왼쪽과 오른쪽 영상 사이의 각을 표시하며, HVA (Horizontal Visual Angle)이라고 부른다 [4]. 이 HVA 를 이용하여 눈의 피로감을 주지 않고, 입체 영상을 감상할 수 있는 안정 영역에 대한 범위를 구할 수 있는데 이는 시청자의 입체감 인지 능력이나 3D 디스플레이의 특성에 따라 조금씩 달라진다. Valyus 는 HVA 의 값 α 을 $\pm 1.6^\circ$ 이내로 할 것을 제안하였고 [5], Yeh 와 Silverstain 은 negative parallax 일 경우, 빠른 물체가 보여지는 일부에서 눈의 피로감을 유발하는 반면, positive parallax 에서는 보다 큰 값을 가져도 피로감을 느끼지 않는다는 점에서 착안하여 negative parallax 에서는 1.57° 이하, positive parallax 에서는 -4.93° 이상으로 할 것을 제안하였다 [6]. 또 Lambooji et al.은 안정영역이 DOF (Depth Of Field)와 깊은 연관이 있다고 밝혔으며 [7], Yano et al.은 실험을 통하여 $\pm 0.2D$ (diopter) 또는 컨버전스 점과 양 눈 사이의 각이 $\pm 0.82^\circ$ 이상이면 눈의 피로를 느낄 수 있다고 발표하였다 [8].

3.3 2D 기반 크기 변환

2D 기반의 크기 변환 방법은 좌우 스테레오 영상에 대해 다음 그림과 같이 1920x1080 크기의 영상을 320x240 크기로 크기 변환을 수행한다. HD 영상과 DMB 영상은 서로 다른 가로/세로 비율을 가지고 있으므로 가로/세로 비율 변환에 따른 왜곡을 줄이기 위해서는 영상 자르기(cropping)가 필요하다. 본 변환 방법에서는 먼저 320x240 의 가로/세로 비율을 유지하기 위해 1440x1080 크기로 영상 자르기 방법을 사용한다.

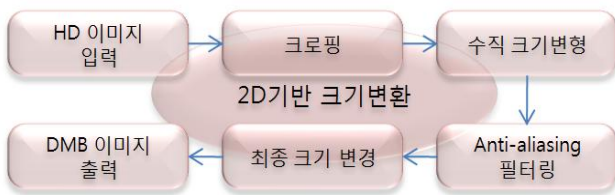


그림 4. 2D 기반 3D 이미지 크기변환 구성도

이때 인터레이스 영상의 경우는 필드 별 왜곡을 최소화 하기 위해 수직 방향에 대해 하나의 필드만을 선택하는 크기변환을 수행한다. 이 과정에서 1440x540 크기로 변환된다. 크기 변환은 앨리어싱(aliasing) 왜곡이 발생하므로 이를 제거하기 위한 가우시안 필터링을 통해 발생할 수 있는 왜곡을 최소화 하고 마지막으로 영상을 320x240 크기로 변환하는 것을 특징으로 한다.

3.4 HIT 기반 크기 변환

2D 기반 크기 변환은 복잡도 측면에서는 큰 장점을 가지지만 입체감을 조절 할 수 없기 때문에 이에 대한 개선책이 필요하다. 이를 위해 본 논문에서는 HIT 를 적용하였다.

HIT 는 가로방향으로 영상의 화소들을 이동 시켜, 3D 영상에서 물체가 전체적으로 화면 앞쪽으로 나와 보이도록 하거나, 뒤쪽으로 들어가 보이도록 하여 입체감을 조절할 수

있다.



그림 5. HIT 기반 3D 이미지 크기변환 구성도

그림 5 는 HIT 기반의 3D 영상 크기 변환에 대한 구성도를 보여주고 있다. 입력된 영상은 크기 변환을 위한 처리와 목표로 하고 있는 디스플레이의 환경 및 안정 영역 분석을 위한 부분으로 나누어져 진행 된다. 1 차 크기변환은 목표로 하는 해상도와 화면 가로/세로 비율을 고려하여, 크기 변환 계수는 240 화소 크기를 기준으로 4.5 로 정하고 이에 비례하여 크기를 변환하며, 결과적으로 426x240 크기의 영상을 생성한다. 이 단계에서 발생할 수 있는 앨리어싱 왜곡을 줄이기 위해 가우시안 필터링을 적용한다. 본 논문에서는 5x5 커널을 적용하였다.

다음으로 안정영역 분석은 3.2 절에서 설명한 방법을 적용하여 교차 변위(crossed disparity)를 위한 최대 각도 및 비교차 변위(uncrossed disparity)를 위한 최대 각도를 구하여 안정영역의 변위 범위를 구한다.

다음으로는 변환된 영상의 깊이감을 안정 영역을 고려하여 조정하기 위해 각 영상이 가지는 변위를 분석해야 하는데, 스테레오 영상 정합 방법을 통해 원하는 변위를 구한다. 본 논문에서는 계산 량 및 용도를 고려하여 sparse 한 스테레오 영상 정합을 전제로 하였으며 블록 별로 처리하였다.

HIT 처리 및 필터링 단계에서는 안정영역 분석 단계 및 변위 계산 단계에서 얻어진 변위 정보를 이용하여 HIT 를 수행한다. HIT 는 여러 가지 방법으로 적용할 수 있겠지만 본 논문에서는 최대 변위 값을 갖는 교차 영역을 안정 영역 안으로 이동 시키는 처리를 하였다. 이 경우 비교차 변위를 갖는 영역 중 일부는 안정 영역을 벗어날 수 있는데, 이에 따른 시각적 불편함을 줄이기 위하여, 해당 영역에 가우시안 필터를 적용할 수 있다.

마지막 단계로는 DMB 호환 영상을 만드는 작업으로 320 보다 큰 영역, 즉 경계 화소(border pixel)을 제거하여 최종적으로 320x240 의 3D 이미지를 생성한다.

3.5 깊이 지도 기반 크기 변환

깊이 지도 기반의 크기 변환 방법은 좌우 영상으로부터 스테레오 정합을 수행하여 중간 시점 영상을 만들고, 중간시점 영상과 좌우 영상간의 깊이 지도(disparity map)을 이용하여 크기 변환된 좌우 영상을 생성해 내는 것이다. 본 방식은 그림 6 과 같고, 정확한 스테레오 정합이 요구되며, 크기 변환 이후 영상의 생성시 폐색 영역(Occlusion region)의 처리가 필요하지만, HIT 방식에 비해 영상의 3D 기하학적 구조에 대한 왜곡을 줄일 수 있다.

이 방법에서는 보다 정확한 스테레오 정합 방법을 통해 생성된 깊이 지도 영상을 이용하여 중간 시점 영상(center

view) 생성하고 이 영상을 1차적으로 크기를 변환한다.

다음 단계로 HIT 와 동일한 과정을 통해 얻은 안정 영역을 고려하여 입체감을 조정하기 위하여 1차적으로 크기 변환된 중간 시점 영상에 대한 변위 지도를 갱신한다. 즉, 최대 최소 변위 내에 위치하도록 변위 값을 변경하는 작업에 해당한다. 다음으로는 전 단계에서 생성된 중간 시점 영상 및 변위 지도를 이용하여 좌우 영상을 합성하고 경계 화소들을 제거하는 영상 자르기 방법을 적용함으로써 최종적으로 모바일 환경에 적합한 3차원 영상을 생성하게 된다.



그림 6. 깊이 지도 기반 3D 크기변환 구성도

4. 실험 결과 및 결론

본 논문에서는 3장에서 제시한 방법을 검증하기 위해 데스크톱 PC 환경에서 구현을 하였다. 2D 기반, HIT 기반 및 깊이 지도 기반 크기 변환 기능을 분리하여 구현을 하였으며 시뮬레이션은 visual studio 6.0 을 사용하여 수행하였다. 각 방법의 기능을 구현하기 위해 윈도우 플랫폼 SDK 및 OpenCV SDK 를 사용하였다.

실험 영상의 경우, 움직임이 큰 스포츠, 빠른 영상, 느린 영상, 어두운 영상, 단순 영상 및 복잡한 3D 영상이 포함된 HD 급의 여러 영상들을 사용하였고, 실험 영상 각각에 대하여 각 방식에 따른 크기 변환 방법을 적용하여 DMB 호환 3D 영상을 생성하였다.



그림 7. 크기변환된 영상

2D 기반의 경우는 단순히 2D 에 적용되는 방법을 적용하여 입체감을 조정하는데 한계가 있었으며 이에 반해 HIT 의 경우 측정된 안정 영역에 HIT 를 적용하여 깊이감이 조절되는 효과를 확인하였다. 하지만 이 경우 2D 와 달라진 점은 입체감을 일정하게 뒤로 이동 시키는 효과를 가져와 비교차 영역 부분에서 깊이감 왜곡을 확인할 수 있었다. 이를 위해서 가우시안 필터를 적용한 블러링 처리를 하였으나 이에 대한 전반적인 효과는 주/객관적인 화질 평가를 통해서 분석이 필요함을 확인하였다.

깊이 지도 기반 크기 변환 방법의 경우에는 HIT 기반의 방법에 비해 입체감 조정이 더 정밀하게 이루어 짐을 확인

하였으나 폐색 영역(occlusion region)이 발생하였고, 이를 보완하는 과정에서 왜곡이 발생함을 확인할 수 있었다.

보완해야 할 점으로는 3.1 절에서 설명한 바와 같이 내용 적응형 영상 크기 변환 방법을 이용해 시청자에게 정보 전달을 극대화 하면서도 피로감 및 불쾌감을 최소화 하여 3 차원 입체감을 최적화 할 수 있는 영역을 분리 해내는 기술 개발이 필요하다. HIT 및 깊이 지도 기반 크기 변환 방법의 경우에는 안정 영역의 정확한 측정 및 계산이 필수적이므로 이에 대한 다양한 방법이 개발 되어야 하며 또한 깊이 지도 기반의 경우에는 깊이 지도 정보의 정확도에 따라 크기 변환된 3D 영상의 품질이 달라지므로, 이에 대한 기술 역시 추가적으로 연구 및 개발 되어야 한다.

감사의 글

This research was supported by the KCC(Korea Communications Commission), Korea, under the development and standardization of Terrestrial Stereoscopic 3DTV Broadcasting System technology by the KCA(Korea Communications Agency) (KCA-2011-10912-02001)

참고문헌

- [1] 김옥중 외 “스테레오 3D 영상 스케일링에 대한 기하학적 분석 및 anti-aliasing 필터”, 방송공학회 논문지, 2009년 제 14권 제5호.
- [2] Atanas Boev, Danilo Hollosi, Atanas Gotchev and Karen Eqiazarian, “Classification and simulation of stereoscopic artifacts in mobile 3DTV content”, SPIE Stereoscopic Displays and Applications, 20 Feb 2009.
- [3] H. Yamanoue, M. Tadenuma and I. Yuyama, “3D Hi-vision,” International Broadcasting Convention (IBC), 1992, pp 533-537.
- [4] L. F. Hodges, “Tutorial:Time-multiplexed stereoscopic computer graphics,” IEEE Computer Graphics and Applications, Vol. 12, 1992, pp 20-30.
- [5] N. A. Valyus, Stereoscopy, The Focal Press, London and New York, 1966.
- [6] Y. Y. Yeh and L. D. Silverstein, “Limits of Fusion and Depth Judgements in Stereoscopic Color Displays,” Humna Factors, Vol. 32, No. 1, pp 45-60, 1990.
- [7] M. T. M. Lambooji, W. A. Ijsselsteijn and I. Heynderickx, “Visual discomfort in stereoscopic displays: a review,” SPIE Stereoscopic Displays and Virtual Reality Systems, Vol 6490, 28 Jan 2007.
- [8] S. Yano, M. Emoto and T. Mitsuhashi, “Two factors in visual fatigue caused by stereoscopic HDTV images,” Displays, Vol 25, Issue 4, Nov 2004, pp 141-150.