Journal of Korea Multimedia Society Vol. 19, No. 4, April 2016(pp. 706-716)
http://dx.doi.org/10.9717/kmms.2016.19.4.706

시각적 편안함을 위한 입체적 삼차원 영상 편집 방법

김정언†, 강행봉††

Stereoscopic 3D Video Editing Method for Visual Comfort

Jung-Un Kim†, Hang-Bong Kang††

ABSTRACT

Each year, significant amounts of Stereoscopic 3D(S3D) contents have been introduced. However, viewers who enjoy the contents readily experience a sense of fatigue on account of various factors. Consequently, many improvement studies have been conducted with the domain of disparity by, for example, simply controlling the disparity or optimizing the reaction speed of viewers' eyes to vergence. However, such studies are limited to the disparity domain and therefore are restricted to a limited number of applications. In this study, we attempted to transcend this limitation and analyzed how a reconstruction in color and brightness, as well as disparity and other important features, affects eyes in terms of vergence adaptation. As a result, we found that, the higher the color similarity is, the better it positively affects vergence adaptation during viewing. Based on this analysis, we propose in this paper a similar color extraction method between takes that are applicable to real-life situations. In an evaluation, the algorithm was applied to publicly available S3D contents and produced a converted color optimized image. The vergence adaptation time of this applied contents was significantly decreased. Also it was minimized through color reconstruction, thereby, being resulted in enhancing viewer concentration.

Key words: Disparity, Accommodation, Vergence, S3D, Color Editing

1. 서 론

2000년대 들어 커지기 시작한 S3D 콘텐츠 시장은 하드웨어와 기술의 발전에 힘입어 급격하게 성장하였다. 영화, 영상, 게임, 산업 현장, 의료 분석 등 다양한 분야에서 보다 현실적인 결과물을 보여주는 도구로 사용되고 있다. 이러한 S3D 콘텐츠는 양안시차에 따라 발생하는 디스패리티(Disparity)를 통해 깊이를 느낀다는 점에 착안하여 촬영 단계에서 2개의 카메라를 통해 사람이 두 눈을 통해 바라보는 영상을 각기 촬영하고 이를 양안에 보여줌으로써 실제 현실 세계를 볼 때와 유사한 경험을 제공한다. 이처럼 오늘날의 S3D 콘텐츠는 HVS(Human Visual System) 관점에서 현상을 재현하고 있으나 실제 사람은 이외에도 보다 다양한 시각 정보들을 통해 상호 보완을 하며 현실 세계를 파악하고 구성하게 된다[7,20]. 이러한 디스패리티를 통해 만들어낸 콘텐츠는 시청자에게 깊이감(Depth)을 제공하여 큰 성공을 거두었지만 시청자들은 쉽게 피로에 노출되었고 이를 해결하는 것이 현재 가장 중요한 이슈가 되었다. 직관적으로 시청자에게 깊이감을 느끼게 만드는 요인으로서 피로감 유발의 가장 큰 원인이 되는 것은 테이크

※ Corresponding Author : Hang-Bong Kang, Address: (420-743) Dept. of Digital Media., Catholic Univ. of Korea, 43-1 Yeoggok 2-dong, Wonmmi-gu, Bucheon-si, Gyeonggi-do, Korea, TEL : +82-2-2164-4598, E-mail : hbkang@catholic.ac.kr
Receipt date : 2015-12-22, Revision date : 2016-01-28
Approval date : 2016-02-17

† Dept. of media Engineering, Catholic University of Korea (E-mail : amysh@catholic.ac.kr)
†† Dept. of media Engineering, Catholic University of Korea
※ This work was supported by the ICT R&D program MSIP/IITP, Korea (10041937, Development of Personalized Stereoscopic 3D Editing Tool and Rendering Process).

(Take)간의 디스패리티 차이이다. 어두운 영화관에서 영상의 테이크가 전환되며 발생되는 급격한 디스패리티 변화는 안구의 수렴 및 발산 운동(Vergence, 이하 안구 운동)을 일으켜 양안의 초점 변환을 유도하게 되고 이 과정에서 눈은 자연히 순응(Accommodation) 시간이 필요하게 된다. 하지만 잦은 테이크의 변화로 이러한 순응이 끝나기 전에 연속적으로 변화하는 디스패리티는 사람으로 하여금 어지러움을 느끼게 하고 피로감을 축적시키게 된다. 따라서 이에 대한 가장 직접적인 해결책은 이러한 디스패리티 변화량을 조절하는 것이며 이와 관련하여 많은 연구가 진행되어 왔다. 여러 연구 끝에 디스패리티 변화와 안구 운동 반응과의 관계를 도식화하고, 연결할 두 테이크간의 프레임(Frame) 중 최적의 안구 운동 반응을 이끌어낼 수 있는 프레임을 선택하여 연결하는 방법이 제안되었다[14]. 이 방법은 디스패리티 관점에서 최적의 결과를 이끌어내지만 이는 하나의 디스패리티라는 관점에서만 이루어진 개선방법이기 때문에 이 이상의 개선을 위해서는 영역(Domain)을 확장할 필요가 있다. 따라서 안구 운동의 반응속도 개선을 통해 시각적 편안함(Visual comfort)을 증진시킨다는 방향성은 유지하고 디스패리티 외의 안구 운동에 영향을 미치는 다른 특징을 고려할 필요가 있다.

이를 위해 S3D 영상을 구성하는 많은 특징들을 조사하였고 그 중 테이크간의 색상 변화 및 밝기 유사도에 주목했다. 색상과 밝기 정보는 영상을 구성하는 가장 기본적인 특징(Low-Level Feature)이며 동공의 움직임과 직접적인 연관성을 지니기 때문에 빠르게 변화하는 테이크 사이에서 순응 과정에 도움을 줄 것이라는 가설을 세웠다. 즉, 디스패리티 변화에 따라 안구 운동이 발생할 때 두 테이크 간의 색상 및 밝기 유사도를 증가시킴으로써 동공의 움직임을 최소화하고 이것이 안구 운동의 반응 속도에 영향을 준다면 결과적으로 눈의 피로도를 감소시키게 될 것이다. 따라서 동일한 디스패리티 변화를 갖는 두 영상(원본과 색상 변화를 준 영상)을 시청할 때의 안구 운동을 반복 측정하여 이 특징의 유효성을 검증한다. 하지만 이 과정에서 원영상에 부여된 콘텐츠 제작자의 의도를 훼손하게 되는 문제가 발생하게 된다. 따라서 색상 변화를 적용할 영역을 시선(Gaze)에 기반을 둔 중요도 영상(Saliency Map)을 이용하여 시청자의 인지에 영향을 최소화 할 수 있는 영역을 선택하여 변화를 주는 방법을 함께 제시한다.

정리하면 본 연구의 목적은 크게 3가지로 나타낼 수 있다.

1) 디스패리티의 변화 과정에서 색상 및 밝기의 유사도가 안구 운동에 미치는 영향을 분석
2) 연결된 두 테이크간의 색상 및 밝기 유사도를 높이기 위한 색상 변환 기법 제안
3) 변환할 색상을 적용할 비중요 영역 검출 방법 제안

위의 목적을 달성하기 위해 본 연구는 아래와 같은 과정을 통해 수행되었다.

우선 관련 연구 조사를 통해 S3D 콘텐츠의 구성과 시청 환경, 시청 시의 눈의 움직임의 연관성에 대해 알아보고 이 때 발생하는 피로도의 원인과 이를 감소시키기 위해 어떠한 연구가 진행되었는지를 알아본 뒤 색상 변형 모델을 통한 연구 확장 가능성을 검증한다. 이를 기반으로 색상 유사도 모델이 실제로 안구 반응에 영향을 미치는지 사전 실험을 통해 확인한다. 다음은 색상 유사도를 증가시키기 위해 연결된 테이크 간의 색상 유사도를 평가하고 이를 기반으로 유사도를 높여 줄 색상 변형 모델을 제안한다. 또한 사용자의 집중도가 낮은 비중요 영역에 대한 설정 방법을 함께 제시한다. 이를 실제 영상에 적용하여 원영상과의 비교 평가를 통해 실제 환경에서 변환 모델이 시각적 피로도 감소에 영향을 주는지를 확인하고 제안된 기법이 적용될 수 있는 영역에 대한 고찰을 통해 실제로 쓰일 수 있는 기술로의 가능성을 검증한다.

2. 관련연구

실험에 앞서 S3D 영상과 안구 운동간의 메커니즘을 살펴보았다. 디스패리티에 의한 깊이 인지 과정에 대한 많은 연구가 있어 왔고 2004년 Meesters[18]가 정리한 조사 연구에서는 S3D와 안구 운동의 관계에 대해 정리하고 있다. 주로 S3D 환경에서 나타나는 디스패리티 변화와 함께 발생하는 안구 움직임 변화를 정리한 것으로 디스패리티가 수렴하고 발산하는 과정에서 영상이 맺히는 영역과 변화 형태를 안구 운동과 순응, 두 관점에서 특징을 분석하였다.

깊이 지각(Depth Perception) 이러한 일련의 안구 움직임을 통한 인간의 시각정보 인지 구조를 HVS(Human Visual System)라 하는데 사람은 깊이 인식에 있어 여러가지 정보(cue)에 기반하여 정보를 획득하고 분류하게 된다(가려짐, 원근감, 사물간의 상대적 크기, 표면의 무늬나 밝기 변화량, 그림자 등). 다시 말해 이러한 정보들은 사물과의 거리에 매우 밀접한 연관을 가지고 있는데[4] 이 중 입체시(Stereopsis)는 매우 강한 깊이 정보로서 HVS에서는 물체를 바라보는 양안의 디스패리티에 의해 거리감을 인지하게 된다[20]. 이에 따라 본 연구에서는 모든 데이터 및 문장의 디스패리티 표기는 픽셀을 기본 단위로 사용하고 좌우 영상의 특성상 수평성분만을 고려하였다(z축의 깊이 변화를 중심으로 실험).

디스패리티의 감도(Sensitivity of Disparity)- 이러한 디스패리티는 깊이 정보를 표현한다는 관점에서 보면 많은 부분에서 밝기(Brightness)의 속성과 유사하다. 디스패리티 검출 시의 검출 경계(Threshold)는 0.3-0.5 CPD주변의 주파수에서 결정되는데 이는 휘도 영역(Luminance Domain)에서의 대비 민감도 함수(Contrast Sensitivity Function)과 유사하게 표현된다[2,3,17]

쌍안 융상(Binocular Fusion) 깊이 지각에 직접적으로 영향을 주는 것은 디스패리티지만 이미지는 망막에 맺히는 것으로 인지하게 된다. 이 이미지는 Panum's Fusional Area라 불리는 호롭터(Horopter) 주변에 융합하여 맺히게 되고 그렇지 못하면 복시가 발생한다. 이러한 융합(Fusion)은 개인차나 조명, 텍스쳐(Texture) 등 여러 요인에 영향을 받게 되는데 이러한 요인들로 인해 일시적인 혼란이 오거나 집중력 저하가 올 수 있다. 이는 일반적으로 어두운 실내에서 감상하게 되는 S3D 콘텐츠의 특성 상 영상의 색상과 밝기 값을 통해 텍스쳐가 구성되며 동시에 조명(광원)으로 동작하게 되므로 이를 개선하는 것이 피로도 감소에 유효할 수 있다는 가설의 기반이 된다.

안구 운동 측정(Vergence Measurements) 모든 실험 단계에서 안구 운동의 측정은 비교하고자 하는 속성외에는 제한을 두어야 한다. Erkelens[5]와 Hung[8] 등의 연구자들은 이를 위해 실험의 영상을 외부 환경적 요소를 배제한 물리적 사물로 제한하고 수동적으로 제어되는 스크린을 통해 실험을 진행하였다. 나아가 안구 운동에 영향을 줄 수 있는 다른 요인을 제거하기 위해 수직선 영상(Vertical Line Image)같은 간단한 자극원(Stimuli)을 통해 실험을 진행하기도 하였으며 핀홀(Pinhole) 기법을 사용하여 순응 과정에서의 변화를 억제하기도 했다[11]. 3D 모니터에서 출력되는 3D 깊이와 달리 실제 사람이 지각하는 깊이는 모니터에서 출력되는 이미지가 투영되는 하나의 가상 평면에 나타나기 때문에 왜곡으로 인한 시각적인 불편함이 발생하므로 모니터에 출력되는 이미지와 일치하거나 유사한 이미지를 출력하는 3D 디스플레이를 개발하여 왜곡과 시각적인 불편함을 줄이기도 하였다[6]

시각적 편안함(Visual Comfort) 초점심도 범위 내에 혹은 바깥에 이미지가 존재하더라도 이미지가 움직일 때 Step Pulse Function에 따라 시각적인 불편함이 조성된다. 정적인 영상은 디스패리티의 범위가 시각적인 편안함에 크게 영향을 미치고, 동적인 영상은 수평적인 모션이 큰 영향을 미쳤다[15]. 또한 공통적으로 장면전환 후에 나타나는 시간적 디스패리티 연속성이 시각적인 불편함을 유도한다. 그래서 Mendiburu[19]등은 디스패리티를 조절하는 방법으로 수평 이미지 변환 기법을 제안하였다.

디스패리티와 안구 운동의 관계(Relation between Disparity and Vergence) 디스패리티의 속도는 디스패리티의 조절에서 가장 중요한 요인중 하나이다. Lang[16] 등은 영상 내의 디스패리티 불연속성을 감소시키기 위해 테이크의 돌출된 디스패리티를 주변과 유사한 디스패리티의 범위로 대체하는 방법을 제안했다. 또한 piotr[21]은 휘도의 대비와 디스패리티의 상호작용을 설명하고 휘도값를 이용하여 Joint Luminance Contrast, Disparity Manipulation, Auto-stereoscopic Contents 등의 최적화와 같은 새로운 가능성을 제시하였다. 이어 Krzysztof[14]은 테이크들이 연결되는 분할(cut) 지점에서의 디스패리티 변화에 따른 안구 운동 반응을 분석하여 디스패리티 구성과 안구 운동의 반응 속도와의 관계를 체계화시켰다. 이를 응용하여 기존에 구성된 테이크내의 프레임을 분석하고 최적의 디스패리티 연결성을 지니는 분할 지점을 연결하여 안구 운동 반응을 최적화시키는 결과를 얻었다.

이상 기존의 연구들을 살펴볼 때 대부분 콘텐츠의 깊이 변화, 즉 디스패리티에 중점을 두고 연구를 진

행하였다. 하지만 디스패리티라는 한 가지 측면에서의 개선방법에는 한계가 있기 때문에 시각적 피로도를 감소시킬 수 있는 다른 도메인을 찾아 연구를 확장해야 한다. 이를 위해 본 연구에서는 연속적으로 변화하는 테이크에서의 급변하는 색상과 밝기는 동공 움직임에 대한 부담감을 주게 되고 이는 곧 안구의 피로도를 누적시켜 안구 운동의 반응에도 영향을 줄것이라는 가설을 세웠다. 그리고 이를 확인하기 위해 테이크간의 색상과 밝기를 변형한 실험 영상을 만들어 원본을 시청할 때와의 안구 움직임을 비교하여 분석하였다.

3. 제안 기법

3.1 사전 실험

앞서 Krzysztof 등은 디스패리티가 급격히 변화하는 분할 영역에서의 안구 운동 변화 속도를 통해 눈의 편안함 정도를 평가하였는데 먼저 이를 재현해 보기 위해 Di =0, ∓90px 에 해당하는 500×500픽셀의 흑백 색상의 잡음 패치(Noise Patch)를 통해 동일한 방법으로 피실험자 16명에게 30회씩 반복 실험을 하였다. 하지만 이와 같은 반복실험만으로는 Krzysztof가 제시한 Gompertz Curve($v=ae^{be^{ct}}+d$)[22]에 정확하게 수렴하지는 않았다. 이는 피실험자의 집중도나 안구 건조에 의한 눈떨림 등에 영향을 받아 오차값들이 발생하기 때문으로 추정된다. 따라서 저역필터(Low-Pass Filter)를 통해 오차값을 보정하였고 이를 각각 30회 반복하여 얻어진 측정값들을 모아 샘플링(Sampling)을 통해 평균 대표 커브(Curve)를 얻어내었다. 추가적으로 색상 유사도가 안구 움직임에 미치는 영향을 비교하기 위해 패치에 RGBCMYK 7가지 색상을 입혀 디스패리티 변화와 함께 색변화를 함께 발생시켰다. 마찬가지로 이를 각각 30회씩 반복 재생을 통해 안구 움직임의 변화여부를 측정하여 안정화에 대한 일반적인 임계값인 95%에 해당하는 안구 운동 반응(Convergence)의 순응 지점(95%-Point) p를 찾아 비교하고자 하였다. Gompertz Curve의 수식에서 a는 점근선에 대한 배율, b와 c는 Gompertz Curve의 양의 값을 갖는 상수들로 b는 x축에 대한 변위를 나타내고 c는 곡률(Growth Rate)을 의미하는데 우리가 알고자 하는 95% 순응 지점 p는 곧 (v-d)/a의 크기가 0.95가 되는 t시점을 의미하므로 아래 식 (1)과 같이 정리할 수 있다.

$$p_{95} = \ln\left(\frac{\ln(0.95)}{b}\right)/c \tag{1}$$

그 결과 연결된 두 패치 영상의 색상이 동일할 때 안구 운동 시작 시점으로부터 순응 시점까지 걸리는 시간이 감소하는 것을 확인하였다.

3.2. 색상 변환 모델의 구조

위의 실험 결과를 바탕으로 실제 S3D 콘텐츠의 색상을 변형하여 시각적 피로도를 최소화 할 수 있는 변환 모델을 제안한다. 본 모델은 연결된 두 테이크의 색상 및 밝기 유사도를 높여주어 분할 지점에서의 급격한 시각적 변화를 줄여 눈의 적응력을 높이고 안구 운동을 안정시켜 반응성을 향상시키기 위한 방법이다. 이를 위해 연결되는 두 테이크의 디스패리티를 분석하여 두 영상의 유사도 색상을 결정한다. 다음으로 영상의 비중요영역을 산출하여 이를 통해 시청자의 집중도가 낮은 영역에 대해 색상을 변환시키게 되는데, 이는 어떤 형태로든 색상 변환이 일어나게 되면 원작자의 의도와 달라지게 되며 시청자의 시각에서도 위화감을 줄 수 있기 때문이다. 또한 색상 변환을 하기 전에 디스패리티로 인한 안구 운동 반응을 최적화 하기 위해 디스패리티가 발산하는 구조의 연결 지점에 대해서는 수렴하는 형태로 변환 후 색상 변환을 적용한다. 아래 Fig. 1은 본 과정의 전체적인 흐름을 나타내고 있다.

디스패리티 이동(Disparity Shift) 연결된 2개의 테이크로 구성된 시퀀스가 들어오면 각각 테이크의 디스패리티를 산출하고 변화되는 형태를 분류한다. 이 결과가 ±30px 로부터 발산(Divergence)하는 형태를 보인다면 두 번째 테이크의 디스패리티를 0으로 이동시켜 수렴(Convergence)하는 형태의 디스패리티로 변형한다. 이는 기존 Krzysztof의 연구 결과 가운데 안구 운동 반응이 ±30px에서 발산할 때 가장 느리고 0으로 수렴할 때 가장 빠른 것을 이용, 해당 발산 구간을 이동시켜 수렴하게 만들어 안구 운동의 반응 속도를 높여 피로도를 감소시킨다.

색상 분석(Color Analysis) 테이크가 바뀌며 컷이 발생하는 시점에서의 급격한 색상 변화를 보다 조화롭게 변화시키기 위해 색상 변환 모델을 적용한

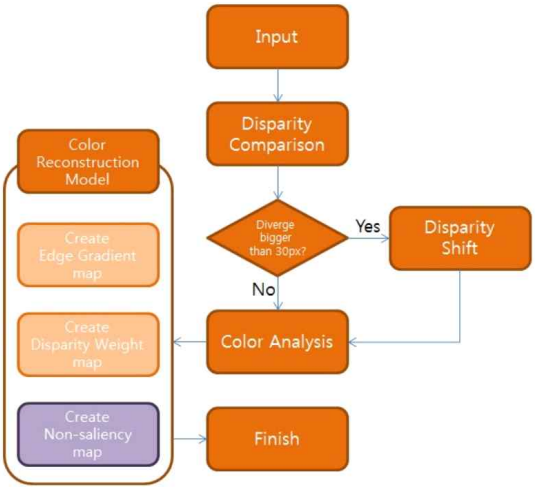

Fig. 1. Process of the proposed model. The figure presents a process of analyzing and transforming an input sequence that is composed of two detailed takes and produces a result with less eye fatigue.

다. 적용을 위한 조건은 가우시안 피라미드(Gaussian Pyramid)를 이용하여 두 테이크의 대표색상을 추출한 뒤 이를 HSL Color Space로 변환하여 두 테이크의 H와 L값을 각각 비교한다. Hue는 색상을 0~360도의 각도로 표현한 값으로 ±로 이동함에 따라 색상이 점진적으로 변화하는 특성을 지닌다. 즉 인근의 색상일수록 색상 유사도가 높게 되는데 색상을 인지할 때 통상 20%인 72도 내외의 색상을 유사 색상이라고 보고 두 테이크의 Hue 거리가 72도를 초과하는 경우 테이크 1의 Hue값을 테이크 2의 Hue값으로 대체시킨다. 또한 밝기를 의미하는 L(Luminance) 영역 역시 20% 범위를 벗어나는 경우 두 번째 테이크의 L값을 가져와 테이크 1의 L을 대체한다. 이렇게 결정된 테이크 1의 색상이 최초 테이크 1의 색상과 변화가 없는 경우(대체가 일어나지 않은 경우)에는 이러한 색상 변환 모델을 적용하지 않고, 대체가 일어난 경우에는 테이크 1의 색상을 변환시킨다.

$$RH_1 = 360 * Threshold_h > |H_1 - H_2| ? H_1 : H_2 \quad (2)$$

$$RL_1 = 360 * Threshold_l > |L_1 - L_2| ? L_1 : L_2 \quad (3)$$

위 식 (2)는 테이크 1의 Hue값(H_1)과 테이크 2의 Hue값(H_2)의 크기를 비교하여 Hue값에 대한 임계값($Threshold_h$)을 넘는 경우 테이크 1에 적용할 변환 색상(RH_1)을 얻어내게 된다. 또한 식 (3)은 Hue와 마찬가지로 Luminance에 대해 변환 값을 구할 수 있게 된다.

색상 변환(Color Reconstruction) 색상 분석 단계에서 색상 변형 여부가 결정되면 첫 테이크의 색상을 변형한다. 다만 영상 전역에 색상 변환을 적용하는 경우 시청자들로 하여금 혼란을 줄 수 있기 때문에 위화감을 최소화 하기 위해 시청자의 주목도가 낮은 영역(Less Saliency Area)을 찾아 해당 영역에 대해서만 가중치를 두어 변형을 진행한다. 이를 위해 영상의 색상과 에지(Edge) 정보, 디스패리티 맵을 통한 전방 물체 추정(Foreground Object Detection) 등의 특징 정보들을 통합해 Less Saliency Area을 생성하고 이 맵의 밝기 정보를 가중치로 사용하여 색상 분석 단계에서 결정한 변환 색상값으로 테이크 1의 색상 변환을 실행한다.

3.2.1 색상 변환 모델의 세부 과정

위 선행실험에서 얻어낸 색상과 안구 운동의 관계를 바탕으로 실제 업무 환경에서 사용할 수 있는 색상 변환 모델을 제안한다. 일련의 색상 변환은 디스패리티가 급격히 변화하는 연속된 두 테이크 사이에서 수행하며 두 번째 테이크의 색상을 기준으로 첫 테이크의 색상을 조정한다. 하지만 완성된 영상에 대한 추가적인 변환 작업이기 때문에 몇 가지 주의점이 있다. 첫 번째로 콘텐츠 제작자의 의도를 보존해야 한다. 본 모델은 시청자의 원활한 영상 감상을 돕기 위한 추가적인 장치이기 때문에 이 변환으로 인해 원작자의 메시지 전달을 방해하는 것은 좋지 않다. 또한 시청자의 입장에서도 변환된 콘텐츠의 색상 변화로 인한 위화감을 느낄 수 있기 때문에 이를 최소화할 필요가 있다. 따라서 영상 전체에 대한 색상 변환이 아닌 시청자의 집중도가 떨어지는 영역에 대해 색상 변환을 적용하여 시청자의 몰입도 저하를 최소화하며 색상 유사도를 높이는 작업이 중요하다.

비 중요영역(Less Saliency Area) Less Saliency Area는 영상을 구성하는 영역 중 중요도가 떨어지는 영역으로 일반적으로 배경에 해당되는 영역이며, 카메라의 초점 바깥쪽의 영역으로 흐릿하게 나타나는 경우가 많다. 이는 디스패리티를 통해 생성되는 깊이 영상(Depth Map)의 배경 영역(Low Intensity Area)과 경계선 검출(Edge Detection)을 통해 얻어진 비

경계영역 영상을 가우시안 피라미드를 통해 합산하여 원 영상에 대한 비 중요영역 영상을 생성한다. 일반적인 Saliency Area 검출 알고리즘에서는 경계선 정보로는 파악하기 힘든 중요 객체(Saliency Object, 주로 Foreground Object)의 검출을 위해 색상의 유클리디안 거리나 텍스처, 경계선 정보 등 여러 가지 정보를 통해 객체를 추정하지만, S3D 영상의 경우 영상 내의 디스패리티 차이를 통해 전경 영역을 특정할 수 있기 때문에 가우시안 피라미드(Gaussian Pyramid)를 통해 생성한 전역 경계선 영상(Blurred Edge Map)과 디스패리티 맵을 더해 반전시키면 우리가 원하는 비 중요영역(Less Saliency Area)을 생성할 수 있다. 본 실험에서 사용한 경계선 영상은 가버 필터(Gabor Filter)를 통해 가 0, 45, 90, 135도에 해당하는 4 방향에 대한 경계선 영상을 생성한 뒤 식 (4)[23] 가우시안 피라미드를 통해 각기 5단계의 영상을 만들어 식 (5)[24]를 통해 원 영상의 크기로 동일하게 확대한 뒤 동일한 가중치로 하나의 영상을 만들어 사용한다.

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(\frac{-x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right)$$
$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta \quad (4)$$

$$G_0 = I(x,y), for\ level, l = 0$$
$$G_l(x,y) = \sum_{m=-2}^{2}\sum_{n=-2}^{2}\omega(m,n)G_{l-1}(2x+m,2y+n), otherwise \quad (5)$$

Fig. 2는 가버 필터에 의한 경계선 영상(좌상단)과 디스패리티 영상(좌하단)을 더해 하나의 비 중요영상(우)을 만드는 과정이다. 좌측의 두 영상은 영상 내에 중요 정보에 대한 가중치를 내포하고 있으므로 이를 반전하여 비 중요영상(Less Saliency Map)을 생성할 수 있다. 이를 이용하여 첫 번째 테이크의 비 중요영역에 대해 색상 변환을 적용하게 된다. 이때 대표 색상은 색상 분석(Color Analysis) 단계에서 결정된 HSL 색상의 RGB 값을 사용한다.

4. 실험 및 평가

4.1 실험군

이공대 학생들 20명으로 구성된 피실험자를 통해 실험을 진행하였으며(10M, 10F), 21살부터 35살까지 연령대의 스테레오 적합성 테스트를 통과한 인원으로 구성되어 있다. 이 중 7명이 안경을 착용하고 있고 이들은 클립형 스테레오 글라스를 이용하였다. 이들은 원본 영상 및 변환 영상을 각각 30차례씩 연속으로 감상하며 눈의 움직임을 측정 하였다. 또한 해당 변환으로 인한 영상 변형 및 체감하는 이질감 정도를 5단계로 평가하였다(이질감이 높을수록 낮은 점수로 평가).

4.2 실험 장비

실험은 2580*1440 해상도의 Samsung 27 Inch 3D Monitor와 HP WorkStation 장비를 이용했고 눈의 움직임을 추적하기 위해 EyeLink 1000 Plus 장비를 사용했다. EyeLink 1000 모델은 초당 1000 Sample

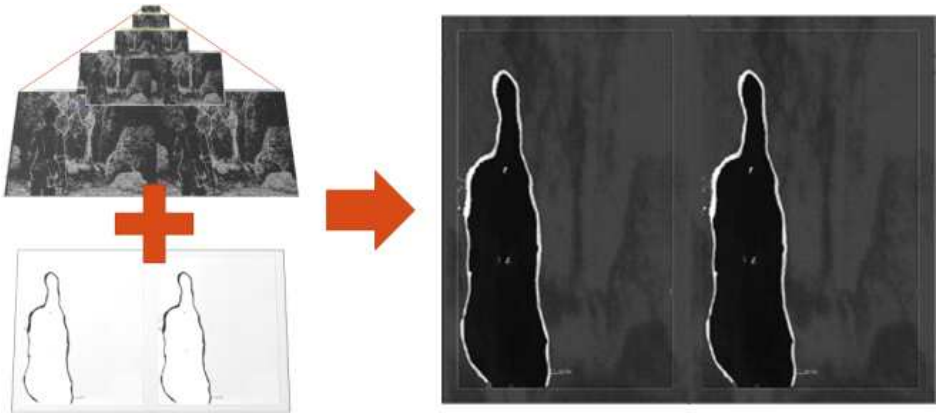

Fig. 2. Process of generating a non-saliency map. A saliency map generated with the edge and color information of the original image was used to generate a final non-saliency map by combining the Gaussian pyramids product with a foreground object map generated with the disparity difference>.

(Binocular Tracking은 각 500)을 추적하여 기록할 수 있으며 눈이 바라보고 있는 모니터의 x,y 좌표를 실시간으로 기록한다. 이 때 피실험자의 머리를 책상에 고정하여 움직이지 않게 하였고 눈과 모니터간의 거리는 55cm로 실험을 진행하였다.

4.3 자극원(Stimuli)

애니메이션 라푼젤에서 일부 테이크들의 분할 지점 전후 프레임에 위 색상 변환 모델을 적용하여 색상 유사도를 높인 영상과 원영상을 비교하여 실험하였다. 이때 모든 프레임이 아닌 위 색상 분석 단계에서 검출 경계 이상의 값을 갖는 프레임에 대해서만 실험을 진행하였다. 이 때 정확한 비교를 위해 H,L 값을 제외한 영상의 디스패리티 등 다른 안구 운동에 영향을 줄 수 있는 요인들에 대해서는 제약을 두었다. Fig. 3은 이 과정에서 발생하는 두 프레임간의 색상 변화 과정을 나타낸다.

4.4 실험 과정

모든 실험은 Eye Tracker 매뉴얼에 있는 대로 시선 추적에 앞서 Calibration과 Validation 과정을 거친 후 추적을 진행하였다. 각 테이크 당 30회의 반복 측정을 수행하였고 실험을 수행하는 과정에서 측정 장비가 눈을 놓치는 경우 다시 Calibration 과정을 진행 후 실험을 재개하였다.

4.5 결과 분석

모든 과정이 끝나면 Eye Tracker가 추적한 데이터를 C# API를 통해 데이터를 읽어 분석하였다. Tracker에서 생성된 결과에는 시간 축에서의 시선의 변화 값(x,y)만이 기록되기 때문에 눈 깜빡임(Eye Blinking), 단속성운동(Saccade Movement) 등이 함께 출력되고 동시에 눈의 떨림 등 오류 값 역시 함께 출력되기 때문에 이를 구분하여 보정해주어야 한다. 눈 깜빡임은 Tracker에서 x,y좌표를 0으로 인식하기 때문에 0을 제거하고, 단속성 운동은 좌우 눈의 좌표 변화가 동시에 같은 방향으로 일어나기 때문에 이를 기준으로 구분이 가능하다. 이 두 가지를 구분해 제거한 뒤 저역필터를 통해 오차를 보정한 결과 값을 사용하였다.

4.6 본 실험

위 색상 변형 모델을 통해 변형된 영상을 시청할 때와 원영상을 시청할 때의 눈의 반응속도를 비교 측정하였다. 실험의 핵심은 테이크 분할 영역에서의 색상 유사도를 증가시켰을 때의 안구 움직임을 측정하고 그 결과를 비교, 분석하여 시간 축(z축)의 결과 값을 토대로 색상 유사도와 안구 반응에 따른 피로도의 상관관계를 증명하는 것이다. 이 과정에서 원영상에 대한 왜곡된 변형이 발생하여 시청자로 하여금 급작스런 영상 변화에 따른 이질감이 발생하기 때문

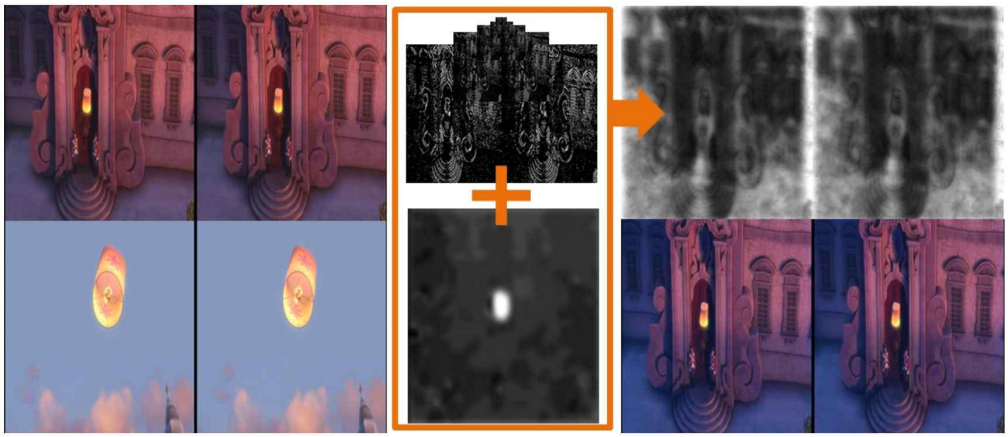

Fig. 3. Left : Original Input Frame. Center : saliency area creation. Right top : inverted saliency map(less saliency map). Right bottom : Result image. The color features of the second take image were made similar to those of the original image at the cut time, when the degree of fatigue occurred, to accomplish a smooth scene change.

에 피실험자들로 하여금 실험 후의 설문을 통해 이질감에 대한 평가를 진행하였다. 테이크가 전환되는 과정에서 발생하는 색상 변화가 눈에 이질감을 주지 않을수록 높은 점수를 주도록 하였다.

그 결과 색상의 변화가 눈에 보이기는 하지만 대체적으로 큰 이질감을 느끼지 못했다는 결론을 얻어낼 수 있었다.

4.6.1 실험 및 결과

실험 영상은 애니메이션 라푼젤에서 추출한 8초짜리 테이크를 사용하였으며 각 중요영역(Saliency Area)의 평균 디스패리티를 -90px와 90px로 고정하였다(중요영역의 디스패리티를 기준으로 기준값이 맞도록 영상을 이동하였다).

Fig. 4의 좌측 그래프는 본 실험의 두 가지 실험 결과를 나타낸다. 해당 그래프는 피실험군 16명에 대한 평균 반응시간으로 오렌지색 선은 원영상의 반응을 나타내고 파란 선은 제안한 모델을 통해 색상 유사도를 높인 모델에 대한 반응이다. 두 그래프를 통해 비교해 보면 cut 시작지점으로부터 안구 운동 반응이 일어나는 시점까지의 시간이 원영상에서는 122 ms, 변환영상에서는 84ms로 나타나는 것을 확인 할 수 있다. 이는 테이크 분할 영역에서의 색상 유사도가 시청자의 안구 반응속도 개선에 긍정적인 영향을 미친다는 점을 나타낸다. 또한 그래프 우측에 있는 95% Point는 피실험이군 전체에 대한 평균 안구 운동 안정화 시간을 나타낸다(안구 운동 변화 그래프가 극대값 도달까지 95%가 되는 지점). 위 그래프를 통해 색상 유사도를 높였을 때 안구 운동 반응의 안정화가 빠르게 나타난다는 점을 확인할 수 있다. 즉, 테이크 간의 색상 유사도의 증가는 안구의 피로도를 감소시키는데 효과적이며, 보다 편안한 영상 감상을 도울 수 있다는 결론을 도출했다. 이와 함께 눈이 인지하는 영상 내 사물의 깊이 역시 원영상의 경우보다 더 작아지는 것을 확인할 수 있는데 이는 기존의 색상 대비와 깊이의 관계에 대한 연구에서 xy평면에 나타나는 주파수 차이에 따라 깊이가 다르게 느껴지는 효과를 보인 바 있다. 이로 미루어 볼 때 S3D 영상에서의 깊이는 영상의 x,y축 외에 시간 축의 z축간 주파수 차에 따라서도 체감 깊이가 달라지며 이로 인해 안구 운동 반응의 안정화까지의 시간이 단축된다는 결론을 얻을 수 있다.

5. 개발 어플리케이션

데이터 분석 툴(Data Analysis Tool) Eye Tracker를 통해 트래킹된 눈의 움직임 데이터를 분석하여 각 영상이 시청자에게 주는 부담감을 수치화하여 보여주는 도구를 개발하였다. 이 분석 도구는 변화하는 테이크에 따른 눈의 움직임을 추적하고 그 데이터를 통해 자극에 따른 안구 반응속도, 안구 운동 반응의 변화시간 등을 비교하고 상대적으로 시각적 피로감을 유발할 수 있는 구성인지의 여부를 알려주는 역할을 한다. 콘텐츠 제작자는 이를 통해 전체 콘텐츠 내의 상대적 피로유발 구간을 파악할 수 있으며 아래에서 제안하는 색상 변환 툴을 이용하여 보완

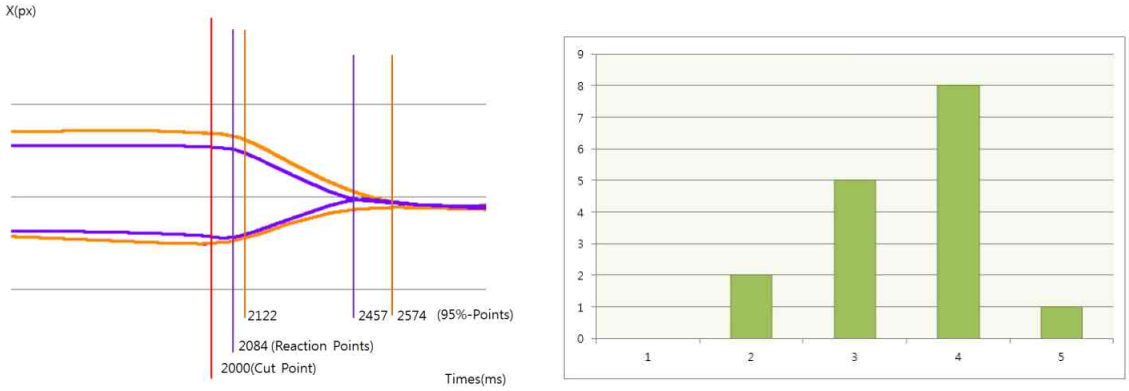

Fig. 4. Left : A comparison graph between original takes and reconstruction takes. Right : Evaluation of the sense of difference perceived during image conversion. The greater the sense of difference, the lower the point became.

Fig. 5. The results of Color Reconstruction model by increasing the color similarity between two images that has different color and brightness. First and 3rd rows are two figures that connect to configure the original scene. Second rows show non-saliency region of first row. Bottoms are the results of this model. It more reduces the fatigue of the eyes through natural connections than the original in vergence adaptation.

할 수 있다.

색상 변환 툴(Color Reconstruction Tool) 본문에서 제안한 색상 변환 과정을 적용하여 연속된 두 테이크의 색상 및 디스패리티를 편집하도록 하는 툴을 제작하였다. 입력된 영상을 급작스런 디스패리티 변화를 기준으로 분리하고 각각의 대표색상을 산출, 색상 변환을 하거나 디스패리티가 발산하는 경우 디스패리티를 이동시켜 수렴형태로 바꾸는 등의 기능을 제공한다. 이러한 기술을 적용하여 현업에서 콘텐츠를 제작하고 편집함에 있어 하나의 가이드 라인으로 활용할 수 있다.

6. 결론 및 한계

본 실험은 빠르게 진화하는 3D 콘텐츠 제작에 있어 매우 중요한 요소인 시청자 피로도를 감소시키는데 목적이 있고 결과를 통해 알 수 있듯 테이크의 분할 영역에서의 색상 및 밝기를 변형하고 디스패리티를 이동시켜 영상을 감상하는 시청자의 안구 부담을 개선하고 안구 반응 속도를 향상시켜 시청자의 피로도를 감소시킬 수 있다는 결론을 얻었다. 본 실험을 통해 얻어낸 결과를 토대로 기존에 출시된 S3D 영상을 재구성할 수 있고, 애초 콘텐츠 제작 단계에서 영상을 구성할 때 색상 변화에 대한 영향을 고려하여 테이크들을 구성한다면 시청자를 더욱 고려한 쾌적한 영상을 제공할 수 있을 것이다. 다만 본 실험은 -90px~90px의 디스패리티 변화가 발생하는 테이크 간의 변화를 기준으로 진행하였기 때문에 디스패리티 변화가 적은 영역의 변형에 대해서는 고려되지 않았고, 또한 색상 변화가 극단적인 경우 생성되는 비 중요영역에 대한 가중치를 적용하더라도 개인에 따라 거부감을 느낄만한 이질감을 발생시키기도 한다. 또한 기준치를 초과하는 색상 차이를 가진 테이크가 매우 짧게 반복되는 경우 색상의 변화가 빠르게 나타나며 시청자에게 혼란을 야기할 수 있다. 이와 관련하여 반복주기에 대한 가중치 경감 등 추가적인 연구가 필요하다.

REFERENCE

[1] Y. Kim and H. Kang, “An Analysis of Recovery Rate and a Change of Depth Recognition After Watching 3D Videos,” *Journal of Korea*

Multimedia Society, Vol. 18.1, pp. 88-96, 2015.

[2] M.F. Bradshaw and B.J. Rogers, *Sensitivity to Horizontal and Vertical Corrugations Defined by Binocular Disparity*, Vision Res. 39.18.304956, 1999.

[3] A. Brookes and K. Stevens, "The Analogy Between Stereo Depth and Brightness", *Perception.* 18.5, pp. 601-614. 1989.

[4] J. CUTTING and P. VISHTON, *Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In Perception of Space and Motion (Handbook of Perception and Cognition).* Academic Press, 1995.

[5] Erkelens, C. J., R. M. Steinman, and H. Collewijn., "Ocular vergence under natural conditions II. Gaze- shifts between real targets differing in distance and direction." *Proceedings of the Royal Society of London B: Biological Sciences,* 236.1285, pp. 441-465, 1989.

[6] Hoffman, David M., et al. "Vergence - accommodation conflicts hinder visual performance and cause visual fatigue." *Journal of vision* 8.3 pp. 33-33, 2008.

[7] Palmer, Stephen E. *Vision science: Photons to phenomenology.* Vol. 1. Cambridge, MA: MIT press, 1999.

[8] Hung, George K., et al. "Vergence eye movements under natural viewing conditions." *Investigative ophthalmology & visual science* 35.9 pp. 3486-3492, 1994.

[9] Hung, George K. "Adaptation model of accommodation and vergence." *Ophthalmic and Physiological Optics* 12.3 pp. 319-326, 1992.

[10] Hung, George K. "Dynamic model of the vergence eye movement system: simulations using MATLAB/SIMULINK." *Computer methods and programs in biomedicine* 55.1, pp. 59-68, 1998.

[11] Hung, George K. *Models of oculomotor control.* Hackensack: World Scientific, 2001.

[12] Krishnan, V. V., Faramarz Farazian, and Lawrence Stark. "AN ANALYSIS OF LATENCIES AND PREDICTION IN THE FUSIONAL VERGENCE SYSTEM*." *Optometry & Vision Science* 50.12, pp. 933-939, 1973.

[13] Krishnan, V. V., DOUGLAS SHIRACHI, and LAWRENCE STARK. "Dynamic measures of vergence accommodation." Optometry & Vision Science 54.7, pp. 470-473, 1977.

[14] Templin, Krzysztof, et al. "Modeling and optimizing eye vergence response to stereoscopic cuts." *ACM Transactions on Graphics (TOG)* Vol. 33.4, 2014.

[15] Lambooij, Marc, Wijnand A. IJsselsteijn, and Ingrid Heynderickx. "Visual discomfort of 3D TV: Assessment methods and modeling." *Displays* 32.4, pp. 209-218, 2011.

[16] Lang, Manuel, et al. "Nonlinear disparity mapping for stereoscopic 3D." *ACM Transactions on Graphics (TOG).* Vol. 29. No. 4. ACM, 2010.

[17] Lunn, Peter D., and Michael J. Morgan. ""The analogy between stereo depth and brightness": A reexamination." *Perception* 24.8, pp. 901-904, 1995.

[18] Meesters, Lydia MJ, Wijnand A. IJsselsteijn, and Pieter JH Seuntiëns. "A survey of perceptual evaluations and requirements of three-dimensional TV." *Circuits and Systems for Video Technology, IEEE Transactions on* 14.3, pp. 381-391, 2004.

[19] Mendiburu, Bernard. *3D movie making: stereoscopic digital cinema from script to screen.* CRC Press, 2012.

[20] Palmer, Stephen E. *Vision science: Photons to phenomenology.* Vol. 1. Cambridge, MA: MIT press, 1999.

[21] Didyk, Piotr, et al. "A luminance-contrast-aware disparity model and applications." ACM Transactions on Graphics (TOG) 31.6, pp. 184,

2012.
[22] Zwietering, M. H., et al. "Modeling of the bacterial growth curve." Applied and environmental microbiology 56.6, pp. 1875-1881, 1990.
[23] Marĉelja, S. "Mathematical description of the responses of simple cortical cells*." JOSA 70.11, pp. 1297-1300, 1980.
[24] Burt, Peter J. "Fast filter transform for image processing." Computer graphics and image processing 16.1, pp. 20-51, 1981.

김 정 언

2007년 가톨릭대학교 컴퓨터공학과(학사)
2009년 가톨릭대학교 컴퓨터공학과(석사)
2010년~현재 가톨릭대학교 미디어공학과 박사과정

관심분야 : 기계학습, 영상처리, 자율주행

강 행 봉

1980년 한양대학교 전자공학과(학사)
1986년 한양대학교 전자공학과(석사)
1989년 Ohio State Univ. 컴퓨터공학(석사)
1993년 Rensselaer Polytechnic Institute 컴퓨터공학(박사)
1993년~1997년 삼성종합기술원 수석연구원
1997년~현재 가톨릭대학교 디지털미디어학과 교수
2005년UC Santa Barbara, Visiting Professor
관심분야 : 컴퓨터비전, HCI, 컴퓨터그래픽스, 인공지능