

# 객체 기반 3D 입체 영상 변환 기법

준회원 이 왕 로\*, 강 근 호\*, 종신회원 유 지 상\*

## Object-based Conversion of 2D Image to 3D

Wang-Ro Lee\*, Keun-Ho Kang\* Associate Members, Jisang Yoo\* Lifelong Member

### 요 약

본 논문에서는 움직임 추정(motion estimation, ME), 컬러 라벨링(labeling) 그리고 Non-local mean 필터를 이용하여 2D 영상을 3D 입체 영상으로 변환하는 기법을 제안한다. 제안하는 기법에서는 먼저 프레임 간의 움직임을 추정하여 객체의 움직임 벡터를 추출하고 주어진 영상에 대해 컬러 라벨링 작업을 수행하여 영상을 분리한다. 움직임 추정 결과와 컬러 라벨링 결과를 비교 분석하여 영상내의 객체를 추출하고 추출된 객체를 이동하여 우 영상을 생성하게 되는데 이때 우 영상을 생성하는 과정에서 채워지지 않은 가려짐 영역이 발생하며 전체 화소간의 상관도를 고려하는 Non-local mean 필터를 사용하여 보상한다. 이후 원본 영상인 좌 영상과 생성된 우 영상으로 비월 주사하여 최종 3D 입체 영상을 재현한다. 실험 결과를 통해 제안된 기법으로 생성된 3D 입체 영상에서 객체위주의 안정된 입체 변환이 수행되는 것을 확인할 수 있었다.

**Key Words** : 3D conversion of 2D image, Motion estimation, Non-local mean filter, Color labeling  
Common hole filling

### ABSTRACT

In this paper, we propose an object based 2D image to 3D conversion algorithm by using motion estimation, color labeling and non-local mean filtering methods. In the proposed algorithm, we first extract the motion vector of each object by estimating the motion between frames and then segment a given image frame with color labeling method. Then, combining the results of motion estimation and color labeling, we extract object regions and assign an exact depth value to each object to generate the right image. While generating the right image, occlusion regions occur but they are effectively recovered by using non-local mean filter. Through the experimental results, it is shown that the proposed algorithm performs much better than conventional conversion scheme by removing the eye fatigue effectively.

### I. 서 론

3D 아바타 영화 이후 3D 미디어에 관한 관심이 고조되고 있고, TV 제조사들도 3D 기능이 탑재된 다양한 3DTV를 출시하고, 노트북, 디지털 카메라, 휴대전화 등의 개인용 기기도 3D 제품이 상용화되어 출시 보급되고 있으며, 전 세계적으로 다양한 3D 입체 방송 서비스도 시행되고 있다. 따라서 3D 콘텐츠에 대

한 수요는 앞으로 계속 증가할 것으로 예상되나 이에 비해 3D 콘텐츠의 공급은 아직 그 수요를 따라가지 못하고 있는 실정이다.

일반적으로 입체감을 느낄 수 있는 스테레오 영상을 제작하기 위해서는 양안에 해당하는 두 대의 카메라, 즉 스테레오스코픽(stereoscopic) 카메라 시스템을 사용하는 것이 일반적이다. 하지만 스테레오 카메라 시스템을 사용하여 3D 콘텐츠를 제작하는 경우 아직

※ 본 연구는 2011년도 광운대학교 교내 학술연구비 지원에 의해 연구되었음.

\* 광운대학교 전자공학과 Digital Media Lab(wangcarrot@kw.ac.kr, keunhodivine@kw.ac.kr, jsyoo@kw.ac.kr)

논문번호 : KICS2011-06-238, 접수일자 : 2011년 6월 6일, 최종논문접수일자 : 2011년 9월 15일

도 많은 비용과 노력을 필요로 하고 있기 때문에 기존에 2D로 제작된 영상을 3D 입체 영상으로 변환하여 3D 콘텐츠를 제작하는 방법이 대안으로 사용되고 있다.

2D 영상을 3D 입체 영상으로 변환하는 방법은 여러 가지가 있다. 실시간 서비스를 위해서는 자동 변환이 필요한데, 이 경우 변환된 3D 입체 영상의 품질은 두 대의 카메라를 직접 이용하여 제작한 경우에 비해 현저히 떨어지는 것을 알 수 있다. 영화 등의 변환에는 품질을 높이기 위하여 수동 방법을 이용하고 있으나 여전히 시간과 노력에 비해 만족할 만한 품질을 얻기가 쉽지 않다. 자동 기법의 경우는 품질은 떨어지나, 입체감을 줄임으로써 자연감을 더하는 방법으로 연구가 진행되고 있다.

가장 기본적인 3D 입체 변환 기법으로는 MTD(modified time difference)를 이용한 방법과 Garcia의 공간 시간 보간(spatial-temporal interpolation)을 이용한 방법, Matsumoto의 영상의 깊이 정보를 활용한 방법 등이 있다<sup>[1-3]</sup>. 하지만 기존 대부분의 방법들은 영상 변형을 통해 입체 영상의 화질이 떨어지고, 움직이는 물체의 속도와 방향을 결정해야 하며, 수직, 수평 운동 방향을 구분하는 과정을 거쳐야 한다<sup>[4]</sup>.

일반적으로 사람은 사물을 바라볼 때 움직임이 없는 배경보다는 움직임이 있는 객체에 대해 입체감을 더 느낀다. 따라서 본 논문은 객체의 움직임 방향에 영향을 받지 않고 객체를 구분하여 깊이감을 부여함으로써 입체 변환 시 부자연스러움을 최대한으로 줄여주는 변환 기법을 제안한다. 이를 위해 먼저 주어진 영상에서 객체를 분리하여 각각의 객체에 깊이감을 따로 부여하게 된다. 객체를 분리하는 과정에서 앞 뒤 프레임들로부터 효과적인 움직임 정보를 추출하고 컬러 라벨링 전처리 과정을 통하여 효율적으로 객체를 검출한다. 또한 주어진 2D 영상을 좌 영상으로 하고 우 영상 생성 시 발생하는 가려짐 영역을 Non-local mean 필터를 이용하여 효과적으로 보상한다. 최종적으로 생성된 우 영상과 원본 영상인 좌 영상을 비월 주사하여 3D 영상을 재현한다. 실험결과 재현된 3D 입체 영상은 기존의 프레임 차이를 이용한 객체기반 입체 변환 방법보다 정교한 객체 추출 결과를 보였으며 가려짐 영역의 보상도 기존의 방법들보다 Non-local mean 필터를 이용한 방법이 뛰어난을 실험을 통해 확인 할 수 있었다.

본 논문의 구성은 다음과 같다. 2장에서는 일반적인 3D 입체 변환 원리에 대해 소개하고, 3장에서는 제안하는 기법의 전체적인 구조와 세부적인 과정을 설명한다. 4장에서는 제안한 기법의 성능 비교를 위해

다양한 실험 결과를 보여주고, 마지막으로 5장에서 결론을 맺는다.

## II. 3D 입체 영상 변환 원리

3D 입체 영상은 양의 시차, 음의 시차, 영의 시차 등 크게 세 종류의 양안시차(binocular disparity)를 제공함으로써 사용자에게 영상에서의 깊이감을 느낄 수 있게 한다. 시차의 종류는 인간이 스크린을 통해 입체 영상을 보았을 때 수렴점이 어느 위치에 생기느냐에 따라서 결정된다<sup>[5]</sup>. 그림 1에 보이는 것처럼 영의 시차는 스크린 상에, 양의 시차는 스크린의 뒤쪽에, 음의 시차는 스크린의 앞쪽에 수렴점이 존재한다.

그림 2는 Ross 현상을 이용한 입체 영상의 원리를 보여주고 있다. 좌 영상으로는 현재영상을 사용하고 우 영상은 지연영상을 사용한다. 영상내의 비행기는 좌에서 우로 이동하고 있으며 배경이 되는 산은 정지하고 있다. 좌안과 우안은 모두 배경이 되는 산과 비행기를 바라보게 되며 그때 좌 영상과 우 영상의 차이인 시차가 발생하게 된다. 따라서 비행기는 스크린보다 앞쪽에 수렴점이 생기는 음의 시차를 갖게 되고,

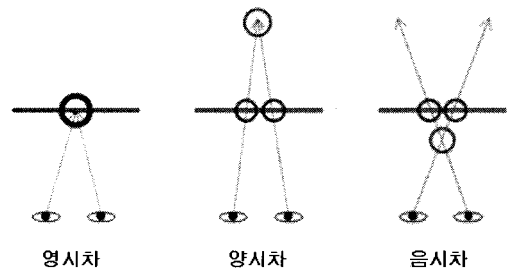


그림 1. 양안 시차의 종류  
Fig. 1. Types of Binocular Disparity

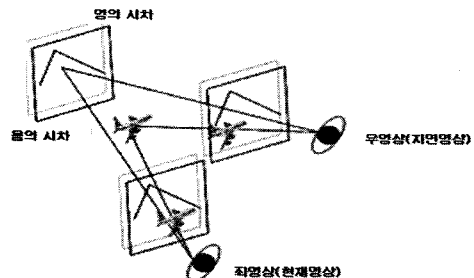


그림 2. Ross 현상 기반의 3D 변환 원리  
Fig. 2. Basic principle of 3D conversion based Ross phenomenon

산은 스크린에 수렴점이 생기게 되어 영의 시차를 가진다<sup>6)</sup>. 이와 같은 Ross 현상에 기반을 둔 입체 영상에서 시청자는 음의 시차와 양의 시차에 따라 영상 내 객체가 화면 바깥쪽으로 나와 보이거나 화면 안쪽으로 들어가 보이는 현상 등으로 입체감을 느낄 수 있다<sup>7)</sup>.

일반적인 디스플레이 장치에서 재현되는 영상은 2D 영상이다. 이러한 2D 영상을 가지고 좌 영상과 우 영상으로 구성되는 3D 입체 영상을 생성하는 것이 3D 입체 변환의 기본 원리이다. 구현 방법은 크게 세 가지로 구분된다. 첫째 방법은 현재 영상이 좌, 우 영상 중 하나가 되면, 다른 하나는 지연 영상들 중에서 1개를 선택하는 것이다. 둘째, 현재 영상은 좌, 우 영상 중 하나를 선택하고 다른 영상은 합성된 영상으로 구현하는 것이다. 마지막으로 좌 영상과 우 영상을 모두 합성 영상으로 구성하는 경우이다. 두 번째와 세 번째 방법에서 합성된 영상의 품질을 향상시키기 위해서는 가려짐 영역을 보충하는 다양한 후처리 기법이 필요하다<sup>8)</sup>.

변환 시 전경에 객체가 존재할 때, 일반적으로 인간은 배경보다 전경 객체의 입체감에 몰입한다. 따라서 객체별로 정확한 입체감을 부여할 수 있다면 고품질의 3D 입체 영상을 얻을 수 있다. 따라서 영상에서 정확하게 전경의 객체와 배경을 분리하는 작업이 필요하다. 하지만 자동으로 객체를 정확하게 분할하는 기법은 매우 어려운 일이며 많은 방법이 제안되었으나 아직도 연구가 필요한 분야이다<sup>8)</sup>. 따라서 고품질의 3D 입체변환 작업에는 대부분 수작업으로 객체를 분리하고 있으며 실시간 처리보다는 비실시간 처리의 응용에 사용되고 있다<sup>8)</sup>.

### III. 제안하는 3D 입체 영상 변환 기법

본 논문에서 제안하는 변환 기법의 블록도를 그림 3에 보였다. 먼저 블록 기반의 움직임 추정(motion estimation, ME) 방법을 이용하여 프레임 간의 움직임 정보를 추출한다. 다음은 영상의 색상(hue) 성분을 이용한 라벨링(labeling) 작업을 두 차례 수행하여 일정 크기 이하의 영역은 잡음으로 판단하여 제거하며 영상을 분리하게 된다<sup>9),10)</sup>.

앞에서 획득한 움직임 추정 결과와 컬러 라벨링 결과를 동시에 이용하여 영상내의 객체를 정밀하게 추출한다. 주어진 영상을 좌 영상으로 하고 추출된 객체 중심으로 좌 영상의 해당 영역을 이동하여 우 영상을 생성한다. 우 영상을 생성하는 과정에서 빈 공간이 발생하게 되는데 본 논문에서는 Non-local mean 필터<sup>11)</sup>

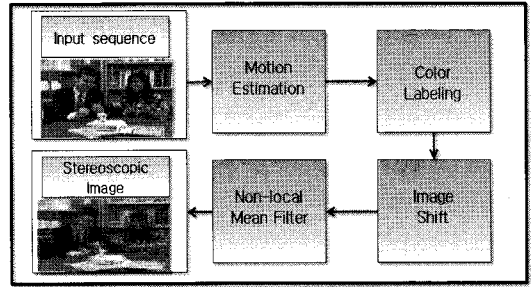


그림 3. 제안하는 3D 입체 영상 변환 기법의 블록도  
Fig. 3. Block diagram of the proposed 3D conversion algorithm

를 이용하여 보충한다. 최종적으로는 원 영상인 좌 영상과 생성된 우 영상의 행을 교차시켜 3D 입체 영상을 재현하게 된다.

#### 3.1 영상의 움직임 정보 추출

영상의 움직임 정보는 블록 기반의 움직임 추정 방법을 이용하여 추출한다. 움직임 정보의 유사성을 평가하는 기준으로는 뵈셀과 덧셀 연산만을 사용하기 때문에 곱셀 연산에 비해 계산량이 적고 연산 속도가 빠른 식 (1)의 SAD (sum of absolute difference)를 사용한다.

$$SAD = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} |f_N(i, j) - f_{N-1}(i+x, j+y)| \quad (1)$$

$$\min\{SAD(i, j)\}, \quad -D \leq (i, j) \leq D$$

여기서  $f_N$ 과  $f_{N-1}$ 은 현재 프레임과 이전 프레임을 의미하고,  $i, j, x, y$ 는 화소의 위치,  $M$ 은 블록의 가로 및 세로의 크기를,  $D$ 는 검색 지역의 가로 및 세로의 크기를 나타낸다.

SAD는 블록 범위 안에 존재하는 화소들과 이전 프레임의 화소들 간 차이의 절대 값의 합이며 가장 작은 SAD 값을 가지는 블록이 가장 정합이 잘된 블록으로 선택된다<sup>9)</sup>.

본 논문에서는 다양한 크기의 블록을 사용하며 그림 4에 블록 크기별 움직임 추정 수행 결과를 보였다. 2x2 크기의 블록으로 움직임을 추정을 하면 크기가 큰 블록의 경우보다 움직임 정보에 잡음이 많아지지만 더 세밀한 움직임 정보를 추출할 수 있다는 장점이 있다. 또한 본 논문에서는 움직임 추정 시 앞 뒤 각각 9개 프레임, 총 18 프레임을 참조하여 움직임 정보를 획득한다. 앞 뒤 각각 9개씩 총 18 프레임에 대하여 움직임 추정을 하면 잡음의 분포는 분산적이기 때문

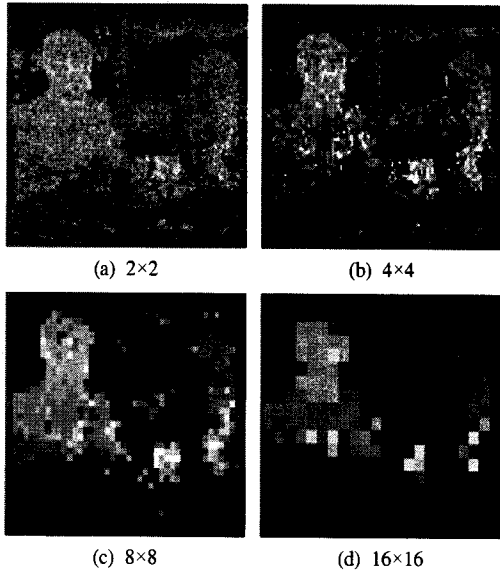


그림 4. 블록 크기에 따른 움직임 추정치  
Fig. 4. Motion estimation based on block size

에 비교적 객체부분을 잘 추출할 수 있다.

총 18프레임의 움직임 추정 결과에 가중치를 다르게 적용하여 움직임 정보를 확률적으로 찾으려 하고, 현재 프레임과 가까운 프레임일수록 가중치 값을 크게 부여한다. 가중치를 다르게 적용한 18개의 결과 프레임들의 각 화소 값에서 움직임 정보가 검출되는지 비교한다. 움직임 정보가 검출되면 확률적으로 중복되는 것만 움직임으로 선택하고 그렇지 않은 경우는 움직임이 없다고 판정하며 그 결과 움직임의 정보를 갖지만 확률적으로는 움직임이 아닐 가능성이 큰 경우를 제거하는 효과가 있다.

즉 그림 5에서 보면 움직임 예측에 사용된 프레임의 수가 많아질수록 객체 부분이 강조되는 것을 알 수 있다. 또한 그림 5(c)와 그림 5(d) 영상처럼 객체임도 불구하고 움직임이 적은 부분은 더 많은 프레임을 참조하여 움직임을 추정함으로써 객체로 판별하는데 도움이 된다. 또한 너무 과도한 움직임을 갖는 객체를 판별하기 위해서 현재 프레임에 가까운 참조 프레임의 움직임 추정치에 더 큰 가중치를 두고 움직임 정보를 구하게 된다. 그림 5(f)는 가중치를 적용하여 구한 움직임 벡터를 모두 더하여 평균을 구한 결과 영상이다.

그림 5(f)의 앞 뒤 프레임 참조로 평균을 구한 최종 움직임 추정 결과에서 남아 있는 잡음과 객체 안에 채워지지 않은 홀(hole)들은 형태학적(morphological) 연산인 팽창(dilation)과 침식(erosion) 연산을 이용하여 채우게 된다<sup>[12]</sup>.

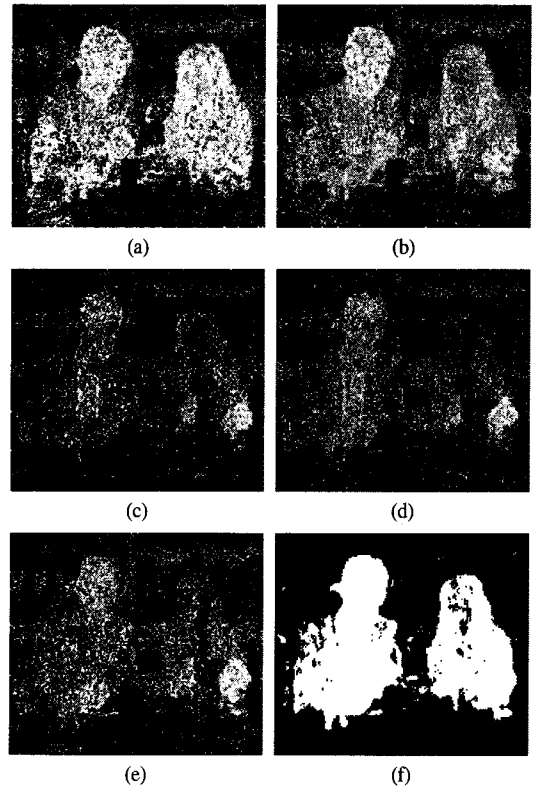


그림 5. 앞 뒤 프레임 참조에 따른 확률기반 움직임 추정 현재와 과거 각 (a) 아홉 프레임 (b) 네 프레임 (c) 한 프레임, 현재와 미래 각 (d) 두 프레임 (e) 네 프레임 (f) 최종 결과  
Fig. 5. Motion estimation based on probability of reference of front and rear frames Current and past (a) 9 frames (b) 4 frames, (c) 1 frames, and current and future (d) 2 frames, (e) 4 frames, and (f) final result

그림 6은 원 영상의 현재 프레임과 침식 연산 및 팽창 연산을 거친 최종 움직임 추정 결과 영상이다.

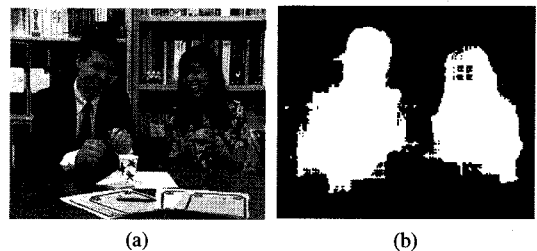


그림 6. (a)원 영상의 현재 프레임과 (b) 침식 팽창 결과 얻은 움직임 추정치  
Fig. 6. Motion estimation based on (a) current frame of the original image and (b) the results of dilation and erosion

### 3.2 컬러 라벨링(Color Labeling)

본 논문에서는 컬러 라벨링의 방법으로 널리 알려진 Rosenfeld의 방법을 사용한다<sup>[10,13]</sup>. 객체 추출을 위

한 효율적인 컬러 라벨링 전처리 과정으로 RGB 색상계를 사람이 사물을 바라보고 해석 할 때 가장 유사하다는 HSI 색상계로 변환 후 mean-shift 필터를 적용한 후 컬러 라벨링 작업을 수행한다.

3.2.1 HSI 색상계 변환

컬러 영상에서의 객체 추출의 경우, 다양한 종류의 색상계가 적용될 수 있다. 일반적으로 RGB 색상계는 디지털 영상에서 공통적으로 사용되는 컬러 모델이다. 이는 기본적인 삼원색 빨강색(red), 초록색(green), 파랑색(blue)으로 구성된다. 그러나 RGB 색상계는 단순히 색상만을 이용하므로 채도와 명도에 대한 분석이 어렵다는 문제점을 가지고 있다<sup>[14]</sup>. 따라서 본 논문에서는 RGB 색상계가 가지고 있는 한계를 보완하기 위하여 HSI 색상계를 이용한다.

HSI 색상계는 색상(hue), 채도(saturation), 명도(intensity) 등의 요소로 색을 표현하는 방식이다. 색상은 빨강, 초록, 파랑 등의 색의 종류를 의미하고, 채도는 색의 탁하고 선명한 정도를 표현한다. 진한 빨강색을 분홍색으로 바꾸기 위해서는 단순히 채도를 조절하면 된다. 명도는 빛의 밝기를 나타낸다<sup>[14]</sup>.

HSI 색상계는 그림 7과 같이 하나의 원뿔이 뒤집어져 있는 이중 원뿔 모양의 원통 좌표계(cylindrical coordinate system)로 정의된다. 이때 색상은 원형으로 정의되며 0°에서 360°의 범위를 갖는 각  $\theta$ 로 표현하며 명도는 원뿔의 아래쪽 꼭지점에서 최소값 0을 갖고, 위쪽 꼭지점에서 최대값 1을 갖는다. 채도는 명도를 고정하였을 때 나타나는 0부터 1의 범위를 가지는 수평축의 길이가 된다.

색상을 이용한 영상 유사도 검사를 위하여 RGB 색상계를 HSI 색상계로 변환하는 공식은 식(2)와 같다. H에 대한 조건이 다음과 같을 때, H,S,I는 각각 다음과 같이 결정된다<sup>[15]</sup>.

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}$$

$$\theta = \cos^{-1} \left[ \frac{\frac{1}{2}[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right] \quad (2)$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

$$I = \frac{1}{3}(R+G+B)$$

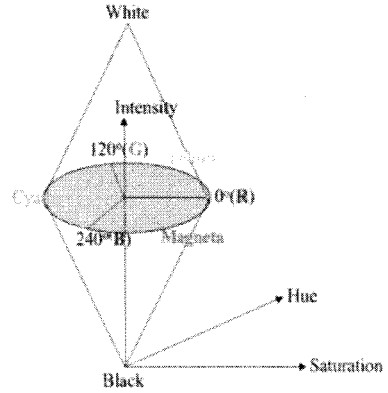


그림 7. HSI 색상계  
Fig. 7. HSI color model

HSI 색상계로의 변환 이외에도 YUV 색상계의 휘도값 Y를 이용해서 라벨링을 할 수도 있지만 휘도 값 Y의 범위는 0부터 255까지이고, HSI의 색상(hue)값의 범위가 0부터 360°이기 때문에 색상(hue) 라벨링 작업을 통해 휘도가 같지만 색상이 다른 부분들을 다른 라벨로 분리할 수 있다는 장점이 있다<sup>[14,15]</sup>.

3.2.2 Mean-shift 필터링

Mean-shift 필터의 적용 범위는 가능한 한 비슷한 색상을 가지는 영역 안에서만 이루어지고 다른 색상을 가지는 영역의 경계를 넘지 않는다. 따라서 탐색 범위 안에 같은 색상의 화소를 많이 포함하는 방향으로 탐색 범위를 이동하여 필터를 적용하게 된다. 이와 같은 방법으로 탐색 범위가 정해지면 그 범위에서의 화소들의 평균값을 구하여 초기 탐색창의 중심점 값이 되도록 한다<sup>[13]</sup>. 이렇게 함으로써 단순히 blurring 되는 문제를 해결하여 경계를 보존할 수 있다. 처리에 적합한 조건은 탐색 영역을 원의 형태를 띤 범위로 설정하는 것이다. 탐색창의 반지름을 크게 할수록 참조하는 화소 수가 많아져 연산 속도가 느려지기 때문에 본 논문에서는 실험 결과 적절한 값을 선택한다.

식 (3)은 색상이 다른 영역의 경계 부근에서 탐색 영역 안에 포함되는 화소를 결정하는 기준이다.

$$dR^2 + dG^2 + dB^2 < threshold \quad (3)$$

여기서  $dR$ ,  $dG$ ,  $dB$ 는 탐색창의 중심화소와 현재 화소와의 R, G, B 성분의 차이이다. 임계값을 작게 할수록 영상의 경계는 유지되어 mean-shift 필터의 결과로 생기는 blurring 효과를 줄일 수 있다. 본 논문에서는 임계값을 적절하게 선택하여 컬러 라벨링의 전 처



그림 8. Mean-shift 필터링 결과  
Fig. 8. Result of mean-shift filtering

리 과정으로 사용한다. 그림 8은 mean-shift 필터를 적용한 결과 영상이다.

### 3.2.3 컬러 라벨링(Color Labeling)

휘도 성분 Y의 범위인 256 보다 범위가 크기 때문에 HSI 색상계의 색상(hue) 성분을 이용하면 색상에 따른 라벨링을 더 정확하게 할 수 있다는 장점이 있다. 색상 분을 36 단계로 나뉘어서 1 차적으로 라벨링한 후에 같은 색상의 영역을 결합하여 2차 라벨링을 한다. 그림 9는 mean-shift 필터를 적용한 후 컬러 라벨링을 한 결과이다.



그림 9. 컬러 라벨링 결과  
Fig. 9. Result of color labeling

### 3.3 객체 추출 및 우 영상 생성

앞 절에서 구한 움직임 추정 결과와 컬러 라벨링 결과를 비교 분석하여 객체를 추출한다. 컬러 라벨링 결과 각 라벨 영역안의 화소들이 일정 비율 이상 움직임 정보를 갖지 않으면 비 관심 객체로 판단하여 제거한다. 최종적으로 남은 영역에서 움직임 정보가 있는 관심 객체를 추출할 수 있다. 그림 10(a)에 최종적으로 추출된 영상내의 객체를 보여주고 있다. 추출한 객



그림 10. (a) 추출된 객체 영상 (b) 생성된 우 영상  
Fig. 10. (a) Extracted objects and (b) generated right image

체를 입체 영상으로 보기 위해서는 동일점이 좌 영상과 우 영상에 투시되어야 한다. 두 눈으로 영상을 바라 볼 때 투시되는 영상에 미묘한 차이가 있어 입체감을 느낄 수 있는 것과 같은 원리이다<sup>[4]</sup>. 앞서 입체 영상을 생성할 때 언급한 세 가지 구현 방법 중에서 제안하는 기법은 현재 영상은 좌, 우 영상 중 하나를 선택하고 나머지 영상은 합성 영상으로 구현한다. 제안하는 방법은 현재 영상을 좌 영상으로 선택하고 추출된 객체에 실험결과 일정한 값을 부여하고 이 값에 따라 객체를 오른쪽으로 이동시켜 우 영상을 생성한다.

이때 시각적으로 객체만 튀어나와 보이는 부자연스러운 부분을 없애기 위해서 배경도 같은 방법으로 이동시킨다. 그림 10(b)는 생성된 우 영상이다. 생성된 우 영상은 홀 채움(hole filling) 과정을 거친 후 좌 영상과 비월주사(interlace) 방식으로 합성되어 입체 영상으로 재현된다.

### 3.4 홀 채움(Hole Filling)

추출한 객체를 이용하여 우 영상 생성 시 값이 채워지지 않은 빈 공간(hole)들이 발생한다. 이것은 해당 화소가 값을 가지지 않기 때문에 검은색으로 표현되어 잡음처럼 보이기도 하며, 이 영역이 커지면 영상에서 객체의 형태가 일그러져 부자연스럽게 보이게 된다. 따라서 우 영상을 생성한 후 이러한 채워지지 않은 홀을 채우기 위한 후 처리 과정이 반드시 필요하다. Non-local mean 필터는 잡음 제거에 많이 사용되는 기법 중 하나이다. 잡음을 제거하고자 하는 화소 주변과 유사한 지역 구조 정보를 지닌 지역이면 어디든지 참조를 하기 때문에 기존의 영상 잡음 제거 필터들 보다 높은 잡음 제거 성능을 보여준다<sup>[11]</sup>.

Gaussian filtering과 anisotropic filtering과 같은 기존의 대표적인 잡음 제거 방법은 영상의 세밀한 부분과 구조를 제거하나 Non-local means는 각각의 화소에 대해서 지역적인 부분에 국한되지 않기 때문에 다른 방법보다 가장자리 보존 등에 있어 매우 뛰어난

결과를 얻을 수 있다<sup>[6]</sup>. 본 논문에서는 채워지지 않은 화소들을 잡음으로 가정하고 이 기법을 적용한다. 식 (4)는 Non-local mean 필터의 동작 원리를 보여준다<sup>[17]</sup>.

$$v(i) = u(i) + n(i)$$

$$v = \{v(i) | i \in I\}$$
(4)

여기서  $v(i)$ 는 관측된 영상,  $u(i)$ 는 원 영상,  $n(i)$ 는 잡음성분이다. 이 때  $i$ 는 화소의 2차원 좌표  $(x,y)$ 를 의미한다. 이 경우 잡음이 포함된 영상  $v(i)$ 는 가려짐 영역을 처리 하지 않은 우 영상에 해당한다. 잡음이 있는 우 영상  $v(i)$ 에서 잡음이 없는  $u(i)$ 를 복원한다는 의미이다. Non-local mean 필터의 장점은 잡음의 값을 복원하기 위해 주위의 화소 값만 참조하지 않고 영상 전체 화소 참조한다는 것이다<sup>[18]</sup>. 식 (5)는 Non-local mean 필터의 구체적인 연산 과정을 보여준다.

$$NL[v](i) = \sum_{j \in I} w(i,j)v(j),$$

$$w(i,j) = \frac{1}{Z(i)} e^{-\frac{\|v(N_i) - v(N_j)\|_{2,a}^2}{h^2}}, 0 \leq w(i,j) \leq 1$$
(5)

$$\sum_j w(i,j) = 1$$

$$Z(i) = \sum_j e^{-\frac{\|v(N_i) - v(N_j)\|_{2,a}^2}{h^2}}$$

여기서  $NL[v](i)$ 는 화소  $i$  위치에서의 추정된 값이며,  $w(i,j)$ 는  $(i,j)$  위치에서의 화소 값과 유사한 정도에 따라 가중치를 준 것이다.  $h$ 는 필터의 차수를 나타내며,  $\|v(N_i) - v(N_j)\|_{2,a}^2$ 은 유클리디안 거리를,  $N_k$ 은 화소  $k$  위치를 중심으로 이웃 블록을 나타



그림 11. Non-Local mean 필터 처리 결과  
Fig. 11. Result of Non-Local mean filtering

낸다.

Non-local mean 필터는 각 화소에 대해 영상 전체의 화소를 참조하기 때문에 잡음 제거에 좋은 성능을 갖는데 반해 많은 계산 시간을 요구하는 단점이 있다<sup>[19]</sup>. 따라서 본 논문에서는 계산 시간을 단축하기 위하여 우 영상 생성 시 생기는 홀 영역의 화소에서만 필터링을 수행한다. 그림 11은 그림 10(b)의 영상에 Non-local mean 필터를 적용한 결과 영상이다.

#### IV. 실험 결과

본 실험에서는 ASU(Arizona State University) 홈페이지에서 제공되는 그림 12(a)의 CIF format (352×288)의 “Paris”와 “Silent” 영상과 그림 13의 QCIF format (176×144)의 “Akiyo”, “Claire” 영상을 실험 영상으로 사용하였다.

본 논문에서 제안한 기법을 2D 영상에 적용하여 생성한 입체 영상을 그림 12에 보였다. 그림 12(a)의 원 영상으로부터 객체 추정 과정을 통해 우 영상을 생성 후 가려짐 영역을 보상한 영상은 그림 12(b)에, 마지막으로 이 두 영상을 비월주사 하여 만들어낸 3D 입체 영상을 그림 12(c)에 보였다.

제안한 기법의 성능을 비교 검증하기 위한 실험도 수행하였다. 앞서 언급했듯이 ASU 홈페이지에서 제공 되는 QCIF format (352×288)의 “Akiyo”, “Claire” 영상을 실험 영상으로 사용하였으며 좌 영상을 원 영상으로 하고 객체 추출을 통해 생성된 영상을 우 영상 생성하였다.

제안된 변환 기법의 성능을 객관적으로 비교하는

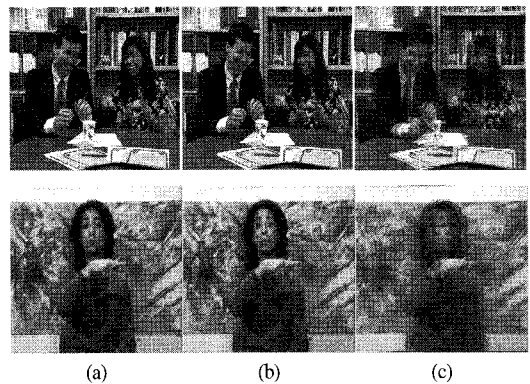


그림 12. Paris(첫 행)와 Silent(두 번째 행)  
(a) 원 영상 (b) 가려짐 영역 보상 영상 (c) 3D 입체 변환 결과 영상  
Fig. 12. Paris(first row) and Silent(second row) (a) Original image, (b) occlusion compensation image and (c) result of converted 3D stereoscopic image

것은 매우 어렵다. 따라서 본 논문에서는 ITU-R 권고 이중자극 연속품질척도법인 DSCQS(double-stimulus continuous quality-scale) 방법을 가지고 주관적 화질 평가를 수행하였다. 표 1은 10명의 평가자가 직접 영상을 보고 평가하는 방법으로 평가 기준은 입체감, 화면 떨림 현상, 눈의 피로도 3 가지로 구분되어 나타내었다. 비교 검증은 기존에 제안된 방법으로 프레임간의 차이를 이용한 객체 기반의 입체변환 기법<sup>[4]</sup>을 이용하였다.

그림 13(a)는 제안한 변환 기법을 이용하여 생성된 3D 입체 영상이며 그림 13(b)는 기존의 기법으로 만들어진 입체 영상이다. 전반적으로 기존의 기법보다 가려짐 영역에 대한 보상이 자연스럽게 처리 되었으며 객체만 이동하여 우 영상을 생성한 기존의 방법보다 객체위주의 이동과 함께 배경도 함께 이동시켜 우 영상을 생성한 제안한 기법의 입체 영상이 보다 자연스러운 입체감을 느끼게 한다. 제안한 기법은 그림 13과 같이 배경과 객체가 잘 구분되어 뛰어난 입체 효과를 느낄 수 있고 특히 뉴스화면, 일기예보와 같이 배경은 멈추어 있고 객체인 사람만 움직이는 영상에서는 더욱 효과적인 성능을 보여 주었다.

표 1의 결과에서도 보듯이 기존 기법에 비해 제안한 기법의 성능이 입체감, 화면 떨림 현상, 눈의 피로도 등의 3 가지 항목에 대하여 전반적으로 우수하다는 것을 알 수 있다.

표 1. 주관적 평가 결과(5점 만점)  
Table 1. Result of DSCQS

평가기준 \ 기준	프레임간의 차이를 이용한 객체기반 기법	제안하는 기법
입체감	2.5	3.0
화면 떨림 현상	2.3	3.1
눈의 피로도	1.8	2.4
합 계	2.2	2.83

V. 결 론

본 논문에서 제안한 3D 입체 영상 변환 기법은 움직임 정보와 컬러 라벨링 작업을 이용하여 영상내의 객체를 먼저 추출하고 추출된 객체를 이동시킴으로써 양안시차를 객체별로 분리하여 부여함으로써 자연스러운 입체감을 구현할 수 있다는 장점이 있다. 양쪽 눈에 맞는 영상은 서로 다른 시차를 가져야 입체감을 느낄 수 있다. 따라서 2D 영상에서의 정확한 객체 추출 과정이 매우 중요하다.

본 논문에서는 움직임 정보와 색상정보를 이용하여 효과적으로 객체를 분리함으로써 정확한 객체 추출이 가능하기 때문에 입체적인 효과 면에서 자연스러움을 잘 구현할 수 있었다. 또한 Non-local mean 필터를 사용하여 우 영상 생성 시 발생하는 홀 영역을 보상함으로써 매우 성능이 우수한 변환 기법을 구현할 수 있었다.

참 고 문 헌

- [1] T. Okino, G. Murata, K. Taima, T. Iinuma and K. Oketani, "New television with 2D/3D image conversion techniques," In *Proc. SPIE*, Vol. 2653, pp.96-103, Jan. 1996.
- [2] B. J. Garcia. "Approaches to stereoscopic video based on spatial-temporal interpolation," In *Proc. SPIE*, Vol.2653, pp.85-95, Photonic West, 1990.
- [3] Y. Matsumoto, Hajime Terasaki, Kazuhide Sugimoto, and Tsutomu Arakawa, "Conversion system of monocular image sequence to stereo using motion parallax," In *Proc. SPIE*, Vol. 3012, pp.108-115, Photonic West, Feb. 1997.
- [4] 한효정, 변혜란, "2차원 동영상으로부터 객체 기반의 3차원 입체 변환 기법," 한국컴퓨터 종합학술대회 논문집, 제33권 제1(B)호, pp.361-363,

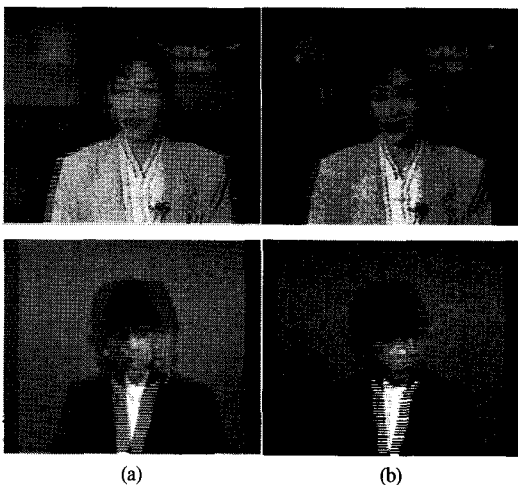


그림 13. Akiyo(첫행)와 Claire(두번째 행) (a) 제안한 기법의 결과 (b) 기존 기법의 결과  
Fig. 13. Akiyo(first row) and Claire(second row) (a) Result of the proposed algorithm and (b) result of conventional conversion algorithm



2006년 6월.

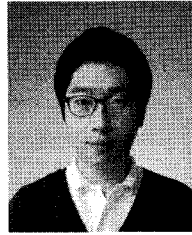
- [5] D. F. McAllister, *Stereo computer graphics and other true 3D technologies*, Princeton, NJ: Princeton University Press, 1993.
- [6] 홍호기, 백윤기, 이승현, 김동욱, 유지상, "2D H.264 동영상의 3D 입체 변환," 한국통신학회 논문지, 제 31권 제12C호, pp.1208-1215, 2006년 12월.
- [7] 김수동, 유지상, "Depth layer partition을 이용한 2D 동영상의 3D 변환 기법," 방송공학회 논문지, 제16권 제1호 pp.44-53, 2011년 1월.
- [8] 김민배, "2D/3D 입체 변환 기술," 한국통신학회지, 제27권 제3호, pp.23-28, 2010년 2월.
- [9] Philip H. S. Torr and Andrew Zisserman "Feature Based Methods for Structure and Motion Estimation," *ICCV Workshop on Vision Algorithms*, pp.278-294, 1999.
- [10] Azriel Rosenfeld and John L. Pfaltz, "Sequential operations in digital processing," *JACM*, Vol.13, pp.471-494, 1966.
- [11] A. Buades, B. Coll and J. M. Morel. "A non-local algorithm for image denoising," *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Vol.2, pp.60-65, Jun, 2005.
- [12] I. Pitas, *Digital Image Processing Algorithms And Applications*, WILEY, 2000.
- [13] D. Comaniciu and P. Meer "Means Shift Analysis and Applications," *IEEE Int'l Conf. Computer Vision*, Vol.2, pp.1197-1203, Sept. 1999.
- [14] 우희숙, 노명중, 박준규, 조우석, 김병국, "HSI 모델을 이용한 모자이크 영상의 품질 평가," 한국측량학회지, 제27권 4호, pp.429-435 2009년 4월.
- [15] R. C. Gonzalez, and R. E. Woods, *Digital image processing*, Second Edition, Prentice-Hall, 2002
- [16] J. Salmon, "On two parameters for denoising with non-local means," *IEEE Signal Process. Lett*, Vol.17, 3, pp.369-372, Mar. 2010.
- [17] 강근호, 이왕로, 유지상, "객체 추출과 Non-Local 필터를 이용한 2D 영상의 3D 변환," 한국방송공학회 추계학술대회, pp.184-187, 2010년 11월.
- [18] A. Buades, B. Coll, J. M Morel. "A review of image denoising algorithms, with a new one," *SIAM J. Multiscale Modeling and Simulation*,

Vol.4 No.2, pp.490-530. 2005.

- [19] 정혜진, 강행봉, "Non-local means 기반 빠른 잡음제거 기법," 대한전자공학회 하계종합학술대회, pp.1088-1089, 2009년 7월.

이 왕 로 (Wang-Ro Lee)

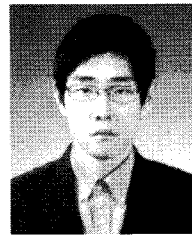
준회원



2011년 2월 광운대학교 전자공학과 졸업  
 2011년 3월~현재 광운대학교 전자공학과 석사과정  
 <관심분야> 3D 영상 신호 처리, 다시점 영상 생성, 영상 신호 처리

강 근 호 (Keun-Ho Kang)

준회원



2011년 7월 광운대학교 전자공학과 졸업  
 2011년 8월~현재 광운대학교 전자공학과 석사과정  
 <관심분야> 3D 영상 신호 처리, 동영상 코덱, 영상신호 처리

유 지 상 (Jisang Yoo)

종신회원



1985년 서울대학교 전자공학과 학사  
 1987년 서울대학교 전자공학과 석사  
 1993년 Purdue Unic. EE, Ph.D.  
 1997년~현재 광운대학교 전자공학과 교수

<관심분야> 3D 입체 영상처리 / 압축, 웨이블릿 기반 영상처리, 비선형 디지털 신호처리