

# 2차원 동영상으로부터 다시점 동영상 생성 기법

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

## Multi-View Video Generation from 2 Dimensional Video

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

### 요 약

본 논문에서는 2차원 동영상에서 다시점 동영상을 생성하는 방법을 제안한다. 다시점 동영상의 생성을 위해 공간적으로 색상 정보, 시간적으로 움직임 정보를 이용한다. 색상 정보는 객체의 정확한 외곽을 추출하기 위하여 사용된다. 색상의 동질영역을 구분하기 위하여 휘도와 색차 정보를 이용하며, 움직임 정보를 구하기 위하여 정합창을 이용한 화소기반의 움직임 예측을 수행한다. 다음 단계는 색상 정보와 움직임 예측을 통해 얻어진 결과를 결합한다. 움직임 예측을 통해 얻은 움직임의 세기 값을 색상정보를 이용해 얻은 각 영역에 할당하고 이를 깊이 값으로 변환한다. 2차원의 입력 영상과 변환되어진 깊이 정보를 통해 회전변환 과정을 거쳐 최종적으로 다시점 동영상을 생성한다. 기존의 2차원 동영상의 3차원 동영상 변환과의 주관적 평가를 통해 효과적으로 다시점 영상을 생성함을 확인 할 수 있었다.

**Key Words** : 영역 분할, 깊이 정보, 다시점

### ABSTRACT

In this paper, we propose an algorithm for generation of multi-view video from conventional 2 dimensional video. Color and motion information of an object are used for segmentation and from the segmented objects, multi-view video is generated. Especially, color information is used to extract the boundary of an object that is barely extracted by using motion information. To classify the homogeneous regions with color, luminance and chrominance components are used. A pixel-based motion estimation with a measurement window is also performed to obtain motion information. Then, we combine the results from motion estimation and color segmentation and consequently we obtain a depth information by assigning motion intensity value to each segmented region. Finally, we generate multi-view video by applying rotation transformation method to 2 dimensional input images and the obtained depth information in each object. The experimental results show that the proposed algorithm outperforms comparing with conventional conversion methods.

### 1. 서 론

HDTV, DMB 등의 기술의 발달로 인해 현실감이 증대되고 보다 나은 영상 표시가 가능하게 되었다. 이러한 기술의 발달은 고화질의 영상, 어디서나 볼

수 있는 영상을 제공한다는 점에서 사용자들에게 2차원 영상의 제한된 조건 속에서 최상의 서비스를 제공한다. 하지만 인간의 끊임없는 욕구는 사용자가 화면 속 장면에 있는 것과 같은 보다 사실적이고 실감 있는 영상을 추구하게 되었다. 이러한 욕구는 자연스럽

※ 본 논문은 교육인적자원부, 산업자원부, 노동부의 출연금 및 보조금으로 수행한 최우수실험실지원사업의 연구결과입니다.

\* 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 : KICS2007-05-232, 접수일자 : 2007년 5월 26일, 최종논문접수일자 : 2007년 12월 20일

계 3차원 영상에 대한 관심의 증대로 이어졌으며 3차원 영상에 대한 연구도 활발히 이루어졌다. 유럽에서는 3DTV에 대한 연구로 COST230(1992~1994)프로젝트, PANORAMA(1995~1998), ATTEST(2002~2004)를 수행하였으며 3DTV(2004~2009) 프로젝트가 수행되고 있다<sup>[1][2][3]</sup>. 국내에서는 한국전자통신연구원이 2002 FIFA 한일 월드컵을 스테레오 동영상 방송 시스템을 구축하여 실험 방송을 하였다. 또한 일본에서는 NHK의 3차원 HDTV 프로젝트 등 3차원 관련 분야에 대한 다양한 연구가 진행 중이며 MPEG의 3DAV 그룹에서도 다시점 동영상 부호화 표준화 작업이 진행 중이다<sup>[4]</sup>.

다시점 영상은 여러 시점에서 영상을 획득하기 때문에 시점이 한 곳에 고정되는 스테레오 영상의 단점을 극복할 수 있다. 다시점 영상의 시점 제약을 극복할 수 있는 특징은 3차원 영상의 디스플레이에서 가장 중요한 기능이다. 다시점 영상을 얻기 위한 방법으로는 다시점 카메라를 통해 직접 영상을 획득하는 방법과 Depth 카메라를 이용하는 방법이 있다.

그 중에서도 다시점 카메라로 영상을 직접 획득하는 방법이 가장 일반적인 방법이다. 하지만 시점의 증가에 따라 데이터양이 방대해지는 단점이 있다. 이러한 문제점을 해결하면서 효율적으로 스테레오 영상이나 다시점 영상을 얻기 위한 방법으로 Depth 카메라를 이용한 방법이 있다. Depth 카메라를 이용하면 RGB 텍스처 데이터와 8 비트의 깊이 정보를 얻는다. 스테레오 영상이나 다시점 영상을 얻기 위해서는 Depth 카메라에서 획득된 깊이 정보를 시차 정보로 변환하고 주어진 시점에 따른 적절한 스테레오 영상을 합성할 수 있는 알고리즘이 필요하다. 이런 알고리즘의 개발과 관련하여 깊이 정보의 처리 및 깊이 정보를 이용한 스테레오 영상 합성 기법에 대한 연구가 진행되었다<sup>[5][6][7]</sup>. 하지만 다시점 카메라와 Depth 카메라의 보급은 아직 미흡한 편이고 콘텐츠도 부족하다.

2D/3D 변환은 2차원 영상에서 스테레오 영상을 생성하는 변환 기법으로 그 변환 방법은 Ross가 발견한 Ross 현상에 기반을 두고 있다. Ross 현상은 양안의 시간적인 지연이 입체감을 느끼게 할 수 있다는 것이다<sup>[8]</sup>. 2차원 영상을 변환하여 3차원 영상을 생성하기 때문에 기존의 카메라나 장비, 존재하고 있는 모든 콘텐츠를 활용 할 수 있기 때문에 저비용, 고효율의 기법이다. 하지만 기존의 기법은 좌영상과 우영상만을 생성하여 시역을 제한하는 단점과 양안의 시간적인 지연을 통해 입체감을 형성

하기 때문에 입체감이 효과적이지가 못한 단점을 가지고 있다.

본 논문에서는 2차원 동영상의 영역 분할을 통해 다시점 동영상으로 변환하는 방법을 제안함으로써 기존의 2D/3D 변환이 가지고 있는 단점을 극복하고 입체감을 향상시키고자 한다.

본 논문의 2장에서는 기존의 2D/3D 변환 기법을 소개하며 3장에서는 다시점 동영상을 생성하기 위해 본 논문에서 제안한 방법을 설명하고, 4장에서는 실험결과를 확인, 분석하며 5장에서 결론을 내린다.

## II. 기존의 기법

기존의 2차원 영상으로부터 좌영상과 우영상의 입체영상을 생성하기 위한 방법으로는 Ross현상을 이용한 MTD(modified time difference)방법이 사용되어진다<sup>[9][10][11]</sup>. 그림 1은 Ross현상을 설명하기 위한 예를 보여준다. 영상에서 배경이 되는 산은 고정이 되어있으며 비행기는 좌에서 우로 이동한다. 이때 좌영상은 현재영상을 사용하고 우영상은 지연영상을 사용하면 배경이 되는 산은 스크린에 수렴점이 생기는 영의 시차를 열고 비행기는 스크린보다 앞쪽에 수렴점이 생성되는 음의 시차를 얻는다.

MTD 방법은 위에서 설명한 Ross현상을 이용하여 현재영상과 지연영상을 사용하여 입체영상을 생성한다. 그렇기 때문에 기존의 2D/3D 변환은 효과적인 지연영상을 선택하는 기법 및 장면전환점과 수직움직임과 같이 MTD기법을 적용 할 수 없을 경우 입체영상을 생성하기 위한 방법을 위주로 연구 하였으며 기본적인 변환의 과정은 그림 2와 같다.

그림 2와 같이 움직임 벡터로부터 운동방향과 운동유형을 판단하며 이로부터 수평운동의 경우 현재

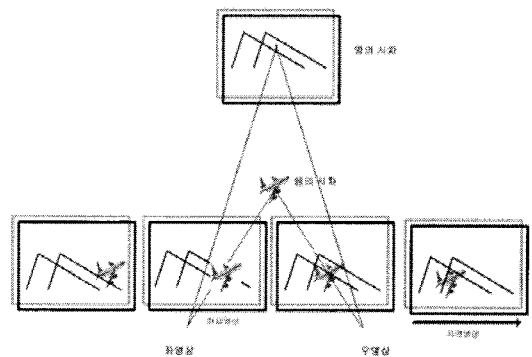


그림 1. Ross 현상  
Fig. 1 Ross phenomenon

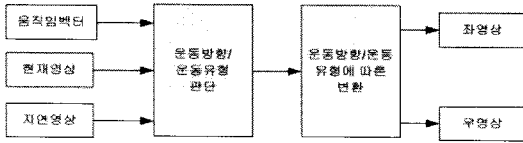


그림 2. 입체 변환 과정  
Fig. 2 3D conversion of 2D video

영상과 지연영상을 사용하고 그 외의 경우는 운동 유형과 운동방향에 따른 적절한 변환을 통해 좌영상과 우영상을 생성한다.

### III. 제안하는 기법

그림 3에서 앞에 있는 물체(비행기)와 뒤에 있는 물체(비행기)가 같은 속도로 움직인다면 움직임 탐색의 결과는 앞의 물체가 뒤의 물체에 비해 크게 나올 것이다. 본 논문에서는 움직임의 정보가 깊이 정보를 반영하는 하나의 요소로서 작용한다는 점을 이용한다. 또한 영상에서 영역 분할을 통해 각 영역에 하나의 깊이 값을 부여함으로써 움직임의 세기를 분할된 영상의 깊이 값으로 변환하는 과정을 거쳐 깊이 정보를 생성하게 된다.

제안하는 기법은 그림 4와 같이 영상의 영역 분할 및 움직임 예측, 깊이 정보 생성, 다시점 영상 생성의 3 단계로 나누어진다. 1 단계에서는 색상 정보를 이용해 입력 영상을 영역 분할한다. 또한 현재 영상과 이전 영상 간의 움직임 예측을 통해 움직임의 세기를 구하게 된다. 2 단계에서는 영역 분할된 영상과 움직임 정보를 이용하여 깊이정보를 생성한다. 마지막으로 3 단계에서는 깊이정보와 입력된 2차원 영상을 회전변환을 통해 다시점 영상을 생성한다. 세부과정은 다음 절에서 설명한다.

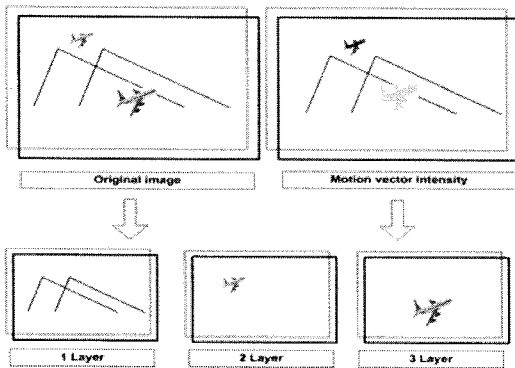


그림 3. 움직임 정보의 깊이 정보 변환  
Fig. 3 Transform of depth with motion vector

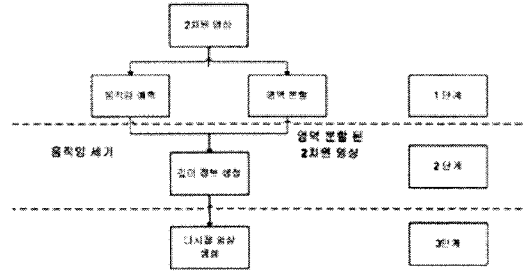


그림 4. 제안하는 기법의 블록도  
Fig. 4 Block diagram of proposed method

#### 3.1 영역 분할 및 움직임 예측

본 논문에서는 영역을 분할하기 위하여 색상 정보를 이용하였다. 색상 정보는 동질의 영역에 대한 정보를 포함하고 있는 정보로서 색상의 유사도는 영상을 분할하기 위한 척도로 사용되어진다<sup>[12]</sup>.

색상 정보를 이용한 영역 분할의 목적은 객체의 외곽을 정확하게 추출하기 위함이다. 정합창(measurement window)을 사용한 움직임 예측 시에 객체의 외곽보다 큰 움직임 영역이 구해지기 때문에 색상정보를 이용한 영역 분할 결과가 이를 보정해 주는 역할을 한다.

영역 분할은 영역 초기화 단계, 영역 융합(region merging), 최소 화소 이하의 영역처리 단계로 나누어진다. 초기화 단계에서는 2차원 영상의 화소 각각이 하나의 영역으로 초기화된다. 영역 융합 단계는 유사한 색상을 가진 영역을 하나의 영역으로 융합하는 과정이다. 유사한 색상의 구분을 위한 척도로 본 논문에서는 휘도 값과 채도 값을 사용하였으며 유사도를 비용함수(cost function)로 사용한다.  $Y$ ,  $U$ ,  $V$  각각의 비용값은 식 (1)을 통해 구하며 이 값이 작을수록 두 영역이 비슷한 하나의 영역임을 의미한다.

$$\begin{aligned}
 C_Y(R_i, R_j) &= \{Y(R_i) - Y(R_j)\} \\
 C_U(R_i, R_j) &= \{U(R_i) - U(R_j)\} \\
 C_V(R_i, R_j) &= \{V(R_i) - V(R_j)\}
 \end{aligned} \tag{1}$$

$R_i$ 는 영상에서  $i$ 번째 영역을 의미한다. 그리고 여기서  $Y(R_i)$ 는 영역  $R_i$ 의  $Y$  성분의 평균값,  $U(R_i)$ 는 영역  $R_i$ 의  $U$  성분의 평균값,  $V(R_i)$ 는 영역  $R_i$ 의  $V$  성분의 평균값을 의미하며  $C_Y(R_i, R_j)$ 는 영역  $R_i$ 와  $R_j$  사이의  $Y$  성분의 비용값(cost value),  $C_U(R_i, R_j)$ ,

$C_r(R_i, R_j)$ 는 영역  $R_i$ 와  $R_j$  사이의 U, V 성분의 비용 값을 의미한다.

식 (1)과 같이 계산된 비용값이 오차보다 작다면 두 영역을 융합하는 과정이 수행된다. 식 (2)는 영역 융합 조건이다. 여기서 허용 오차값은 융합단계에서 각 영역의 화소수를 결정하는 값으로 이 값이 커질수록 각 영역의 크기도 커지며 잘못된 융합이 발생할 수 있다. 본 논문에서는 이 값을 임의의 작은 값(0~5)으로 설정하여 영역 융합 단계에서 잘못된 융합이 발생하지 않도록 하였다.

$$\begin{aligned} C_Y(R_i, R_j) &< \text{허용오차} \cap \\ C_U(R_i, R_j) &< \text{허용오차} \cap \\ C_V(R_i, R_j) &< \text{허용오차} \end{aligned} \quad (2)$$

영역의 융합과정에서 영역 정보인 영역의 Y, U, V의 값이 재설정되어진다. 식 (3)은 재설정에 관한 수식을 나타낸다.

$$\begin{aligned} Y(R_i) &= \frac{1}{(PC_i + PC_j)} \left( \sum_{k=0}^{PC_i-1} Y_{(i,k)} + \sum_{k=0}^{PC_j-1} Y_{(j,k)} \right) \\ U(R_i) &= \frac{1}{(PC_i + PC_j)} \left( \sum_{k=0}^{PC_i-1} U_{(i,k)} + \sum_{k=0}^{PC_j-1} U_{(j,k)} \right) \\ V(R_i) &= \frac{1}{(PC_i + PC_j)} \left( \sum_{k=0}^{PC_i-1} V_{(i,k)} + \sum_{k=0}^{PC_j-1} V_{(j,k)} \right) \end{aligned} \quad (3)$$

식(3)에서  $Y(R_i)$ ,  $U(R_i)$ ,  $V(R_i)$ 는  $i$ 번째 영역의 Y, U, V 성분의 값을 의미한다. 또한  $PC_i$ 는  $i$ 번째 영역의 화소의 수,  $Y_{(i,k)}$ ,  $U_{(i,k)}$ ,  $V_{(i,k)}$ 는  $i$ 번째 영역의  $k$ 번째 화소의 Y, U, V 값을 의미한다.

영역 융합을 거치고 나면 영역이 과도하게 분할된 결과를 얻을 수 있다. 하지만 과분할(oversegmentation) 영역의 결과는 차후 과정인 영상에 깊이값을 부여하는 과정에서 오차를 발생하게 되는 요인으로 작용한다. 마지막 3 단계는 영역의 화소수가 설정되어진 최소 화소보다 적은 수의 화소를 보유하고 있다면 수행되어진다. 먼저 영역에 이웃하고 있는 모든 영역과의 유사도를 비교하기 위하여 식 (4)와 같이 비용 함수를 계산한다.

$$C(R_i, R_j) = C_Y(R_i, R_j) + C_U(R_i, R_j) + C_V(R_i, R_j) \quad (4)$$

그리고 각각의 영역은 이웃한 영역들과의 색상의

유사도를 비용값을 가지고 비교하여 가장 작은 비용값을 가진 영역과 융합하는 과정을 정해진 최소의 화소 이상이 될 때까지 반복한다.

그림 5는 지금까지 영역 분할 과정의 예이다. 그림 5에서 굵은 선은 영역의 경계이며 각 영역은 R1에서 R23으로 명칭한다. 또한 그림에서 각 영역의 표시된 수는 각 영역의 유사도의 척도이며 본 논문에서는 Y, U, V성분을 모두 사용하였으나 그림 5에서는 하나의 성분만을 예를 들어 설명한다. 초기화 단계에서는 그림 5에서와 같이 각 화소가 하나의 영역으로 초기화 되며 초기 유사도의 척도는 각 화소의 화소값이 된다. 그리고 영역 융합과정에서 식 (2)를 이용하여 영역의 유사도를 판단하며 허용 오차가 4로 설정 되어있으면 R1,R2,R3,R7,R8은 하나의 영역으로 융합되며 유사도의 척도는 식 (3)를 이용하여 재설정된다. 마찬가지로 R4,R8,R9,R10,R14,R15,R16도 하나의 영역으로 융합되며 영역 융합 후에 그림 5의 영역 융합결과를 얻을 수 있다. 영역 융합 후 화소 값이 임의로 설정하여 둔 최소 화소 이하의 영역이 존재 할 수 있다. 이런 과분할 결과를 제거하기 위하여 최소 화소이하의 영역 처리가 필요하다. 만약 최소 화소가 4로 설정되어 있다면 미소영역은 R12, R20가 된다. R12 영역 주변의 영역인 R1, R20 중 식 (4)에 따라서 R12는 비용값이 작은 R1과 융합되며 R20은 R5와 융합된다.

본 논문에서는 영상에서 움직임 예측을 위하여 화소 기반의 움직임 예측을 한다. 화소 기반의 움직임 예측은 모든 화소가 변이를 갖는 방식으로 블록 기반의 움직임 예측보다 신뢰도 있는 세밀한 움직임

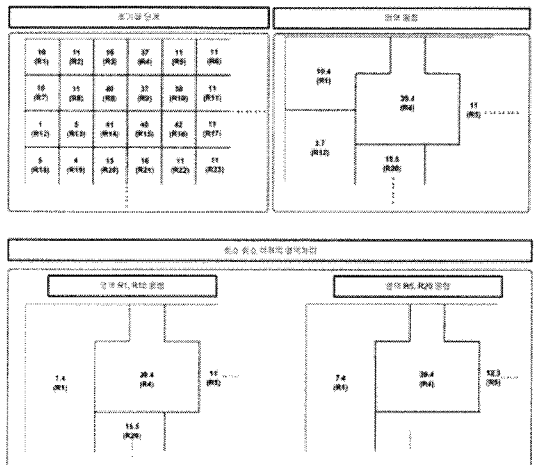


그림 5. 영역 분할 과정  
Fig. 5 Example of segmentation

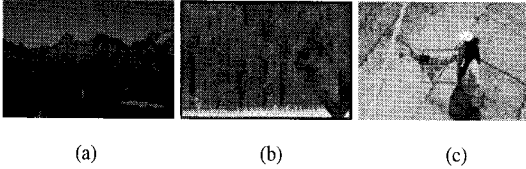


그림 6. 정합창 설정: (a) 수평 움직임, (b) 수직 움직임, (c) 일반적 움직임  
 Fig. 6 Measurement window: (a) Vertical motion, (b) Horizontal motion, (c) General motion

예측을 할 수 있다<sup>13)</sup>. 화소 기반의 움직임 예측을 하기 위해 찾고자 하는 화소의 주변 영역의 값을 정합창으로 사용한다. 움직임 예측 과정은 탐색 범위 내에서 정합창 내의 화소의 비용값을 비교함으로써 이루어진다. 탐색 범위내의 가장 작은 비용값을 가지는 화소를 시간상의 정합되는 점으로 간주한다. 본 논문에서는 비용값을 위해 화소가 가지는 Y값을 사용하였다. 화소기반의 움직임 예측을 정확히 하기 위해서는 정합창의 크기와 탐색 범위의 설정이 중요하다. 정합창의 크기는 정합되는 점의 신뢰도를 높일 수 있으나 정합창이 커지게 되면 계산량도 많아지고, 영상내의 객체의 외곽을 정확히 따라가기가 어렵게 된다<sup>14)</sup>. 또한 탐색 범위의 설정도 결과에 중요한 영향을 미친다. 너무 작거나 큰 탐색 범위의 설정은 잘못된 정합점을 찾을 가능성을 크게 만든다. 또한 수평의 움직임이 있는 영상에서는 수평방향으로의 탐색 범위가 정합점을 찾을 가능성을 높여주고 수직 움직임이 있는 영상에서는 수직 방향으로의 탐색 범위가 정합점을 찾을 가능성을 높여준다. 본 논문에서는 영상의 움직임 유형에 따라 그림 6과 같이 수평방향과 수직방향, 정방향의 탐색범위를 사용하였으며 이 설정은 영상의 전체적인 움직임을 얻기 위하여 16×16 블록으로 초기움직임을 구하였다. 또한 이 초기 움직임에 기반하여 탐색 범위를 설정하였고 정합창은 움직임 유형에 따라 10×5, 5×10, 8×8로 설정하였다. 그림 6에서 사각형 영역은 탐색범위의 한 예로 (a)는 수평방향 움직임으로 수평방향 탐색을 중요시하며, (b),(c)는 각각 수직방향, 일반적인 움직임의 탐색을 나타낸 것이다.

### 3.2 깊이 정보 생성

지금까지의 과정은 영역 분할과 움직임 예측을 위한 과정이었다. 이제 움직임 값을 깊이 정보로 변환하기 위하여 식 (5)의 과정을 거친다.

$$D(x, y) = Motion_x + Motion_y \quad (5)$$

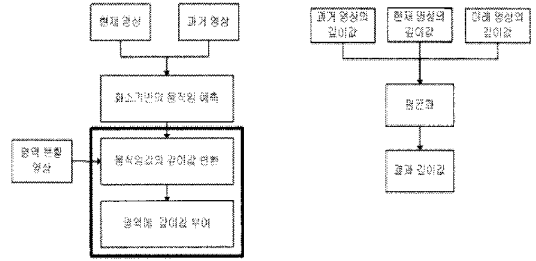


그림 7. 움직임 추정 및 깊이 정보 생성  
 Fig. 7 Motion estimation and depth map generation

여기서  $D(x, y)$ 는 화소  $x, y$ 의 깊이값이다. 그리고  $Motion_x$ 는  $x$  방향으로의 움직임 크기(intensity)이고,  $Motion_y$ 는  $y$  방향으로의 움직임 크기이다.

$D(x, y)$ 는 많은 오차값을 가지고 있고 한 영역 내에서 화소들 간에 편차가 크기 때문에 다음 과정의 처리가 필요하다. 본 논문에서는 구한  $D(x, y)$ 에 대해 영역내의 화소들의 깊이값 평균을 구해서 한 영역 내에서는 같은 값을 가지도록 한다. 이 과정을 식 (6)에 나타내었다.

$$D(R_i) = \frac{1}{m \times n} \sum_{k=0}^{m \times n - 1} D(k) \quad (6)$$

식 (6)에서  $D(R_i)$ 는  $m \times n$  화소를 가진  $i$  번째 영역에서의 깊이값을 의미하며  $D(k)$ 는 영역의  $k$  번째 화소의 깊이값이다.

영상에서 물체나 카메라의 움직임은 일정하지 않다. 그렇기 때문에 위 과정의 결과를 깊이 정보로 사용하게 되면 다른 시점의 영상을 생성하였을 경우 영상이 끊기는 느낌을 받게 되거나 영상이 깨지는 것 같이 느끼게 된다. 위와 같은 현상을 방지하기 위해 본 논문에서는 과거 영상과 미래 영상 그리고 현재 영상의 깊이값의 평균값을 취하게 된다. 그림 7은 지금까지의 움직임 추정 및 깊이 정보를 생성하기 위한 과정을 나타낸 블록도이다.

### 3.3 다시점 영상 생성

그림 8은 디스플레이 장치를 바라보는 관찰자의 위치를 나타낸다.

관찰자의 위치 이동에 따른 각을  $\theta$ 라 하면 새로운 위치의 카메라에 투영되는 영상은  $\theta$ 만큼 회전된다.  $\theta$ 만큼 회전된 영상 화소의 위치는 다음 식 (7)에 의해서 결정된다<sup>15)</sup>.

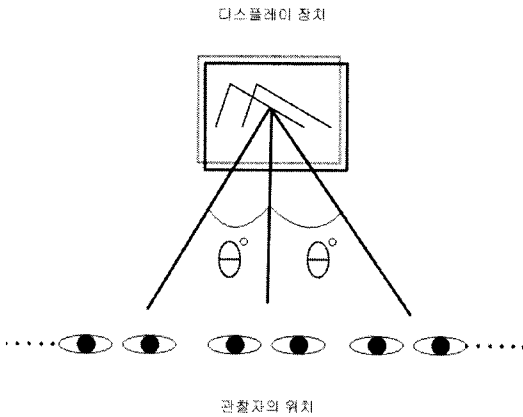


그림 8. 디스플레이 장치를 바라보는 관찰자의 위치  
Fig. 8 Multiple view points

$$\begin{bmatrix} x_\theta \\ y_\theta \\ z_\theta \end{bmatrix} = R(\theta) \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (7)$$

여기서  $(x_\theta, y_\theta, z_\theta)$ 는  $\theta$ 도 회전된 영상의 좌표,  $(x, y, z)$ 는 원 영상의 좌표를 의미한다.  $R(\theta)$ 는 식 (8)의 회전 변환 행렬이다.

$$R(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \quad (8)$$

깊이 영상 기반의 영상생성 시 한 시점의 텍스처 영상을 사용하여 가려진 영역을 처리하기 때문에 가려진 영역을 처리하기에 어려움이 있다. 일반적으로 이러한 가려진 영역을 처리하는 방법으로는 배경과 전경사이의 선형보간법(Linear interpolation of foreground and background)과 배경색의 외삽법(background color extrapolation)의 방법이 사용되며 본 논문에서는 가려진 영역의 처리를 위하여 선형보간법을 사용하였다.

#### IV. 실험 결과 및 분석

실험으로 영역분할, 움직임 예측 및 깊이 정보 생성의 각 과정과 생성되어진 깊이 정보를 이용한 다시점 영상 생성 결과를 분석한다. 그림 9는 영역

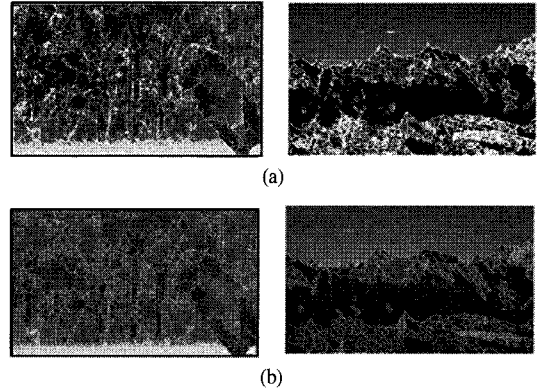


그림 9. 수평, 수직 움직임 영상의 영역분할 결과 영상:  
Fig. 9 Segmentation result of vertical and horizontal motion image : (a) Before small region process, (b) After small region process

분할 결과 영상이다. 그림 9(a)는 최소 화소 적용 이전의 수평 움직임과 수직 움직임의 결과 영상으로 비용함수만을 계산한 영역 융합의 결과이다. 그림 9(a) 영상에서 위쪽 하늘 영역은 거의 하나의 영역으로 융합되었고 아래쪽의 복잡도가 높은 영역은 많은 미소 영역이 존재하는 것을 알 수 있다. 또한 그림 9(b)의 경우에도 복잡한 나무의 잎과 같은 부분은 많은 미소 영역이 존재 하고 있는 것을 확인할 수 있다. 그림 9(a)와 같은 결과는 미소 영역이 존재하기 때문에 영역 분할 결과로는 부적합하다. 그림 9(b)는 최소 화소의 설정을 통해서 미소 영역을 일정수준 이상으로 융합한 결과이다. 그림 9를 통해서 영상에서 유사한 색상의 영역이 하나의 영역으로 융합 된 것을 확인할 수 있다.

다음 과정은 현재 영상과 과거 영상의 화소 기반 움직임 예측이다. 그림 10은 움직임 예측 결과이다. 본 논문에서 수평 움직임이 있는 영상에서는  $10 \times 10$ 의 정합창과  $10 \times 2$ 의 탐색범위를 설정하였고 수직 움직임이 있는 영상에서는  $10 \times 10$ 의 정합창과  $3 \times 10$ 의 탐색범위를 설정하였다.



그림 10. 움직임 예측 결과: (a) 수평 움직임 영상, (b) 수직 움직임 영상  
Fig. 10 Motion estimation results :  
(a) Horizontal motion image, (b) Vertical motion image

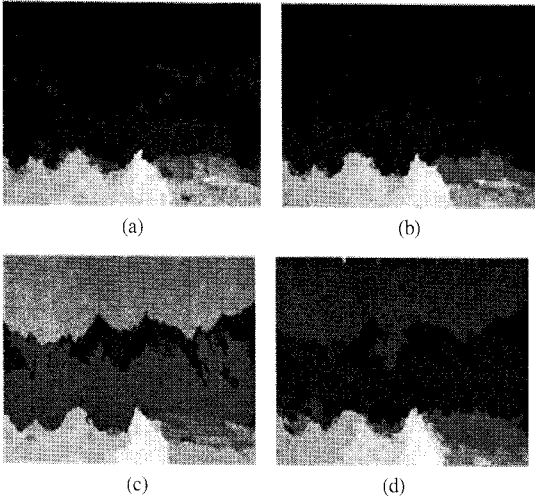


그림 11. 깊이값 결과: (a) 1번 프레임, (b) 2번 프레임, (c) 3번 프레임, (d) 4번 프레임  
 Fig. 11 Depth map images : (a) 1st frame, (b) 2nd frame, (c) 3rd frame, (d) 4th frame

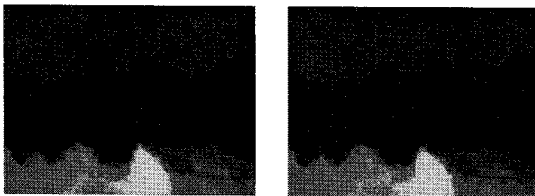


그림 12. 평균값이 부여된 결과  
 Fig. 12 Final depth map images

움직임 예측 후 결과를 깊이값으로 변환한다. 그림 9(b)와 그림 10의 결과로부터 그림 11의 결과영상을 얻는다. 그림 11은 그림 7의 영역 분할된 영역에 그림 9의 움직임 예측의 결과값의 평균값을 계산한 결과이다.

그림 11은 잡음이 있고, 특히 그림 11(c)와 같이 카메라나 물체의 움직임 크기의 변화, 움직임 예측의 오차로 인해 불연속적인 결과를 얻게 된다.

본 논문에서는 그림 11과 같이 연속된 영상에서 변화가 큰 프레임의 문제를 처리하기 위해 현재 영상과 과거, 미래 영상과의 평균값을 사용한다. 그림 12은 그림 10의 3, 4번 프레임의 결과 영상이다.

그림 13은 그림 12의 깊이영상과 입력영상을 회전 변환하여 얻은 9 시점 영상이다.

그림 14는 9 시점 영상 중 -4도 영상과 +4도 영상의 차영상으로 앞으로 나와 있는 부분에 시차가 크게 적용된 것을 확인할 수 있다.

기존의 기법과의 비교를 위해서 기존의 2D/3D

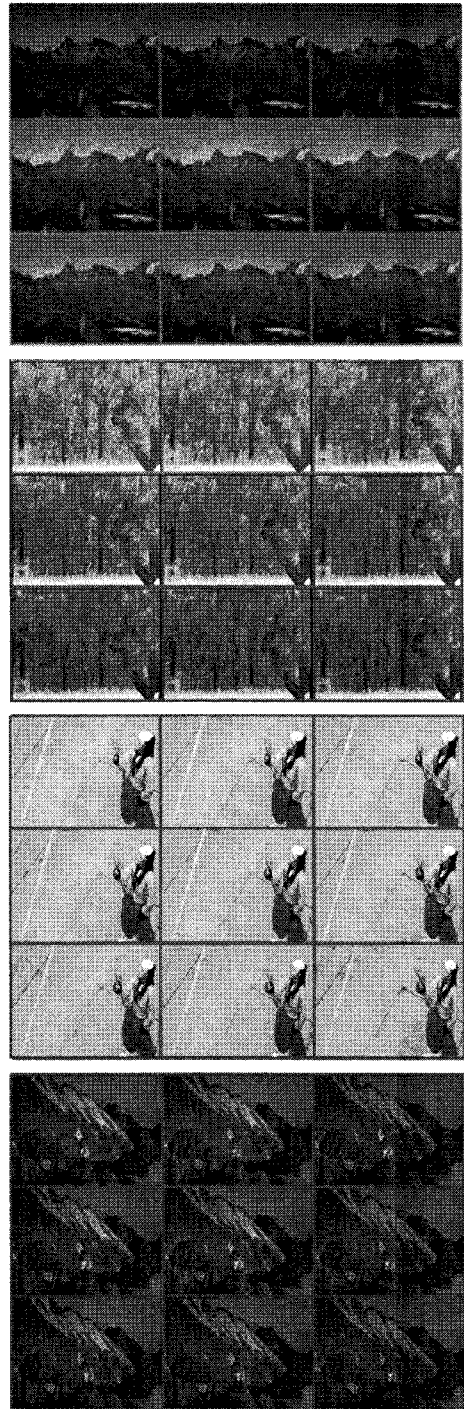


그림 13. 깊이영상과 입력영상을 회전 변환하여 얻은 9 시점 영상(-4도~+4도 영상)  
 Fig. 13 Generation of nine view images(-4 degree ~+4 degree)

변환의 결과와 주관적 평가를 진행하였다.

주관적 평가를 위하여 DSCQS 방법을 이용하여 입체감에 대한 평가를 했다[16]. 기존의 기법으로

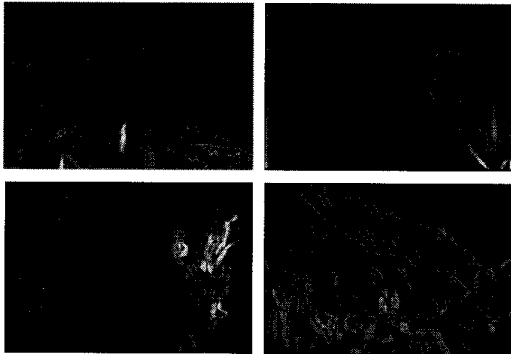


그림 14. -4도 +4도 영상의 차영상  
Fig. 14. Difference between -4 degree and +4 degree image

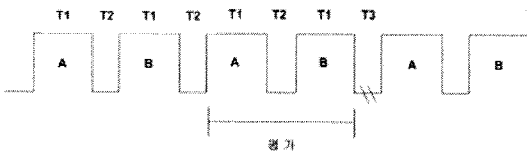


그림 15. 주관적 평가 방법  
Fig. 15 Subjective assessment of the quality

표 1. 주관적 평가 결과  
Table 1. Result of subjective assessment

	기존의 기법	제안한 기법
평가	Pair(58)	Good(70)

생성된 입체영상과 제안한 방법으로 생성된 다시점 영상 중 0도 영상과 +1도 영상을 사용하여 생성된 입체 영상을 비교 하였다. 평가 방법으로는 그림 15와 같이 평가자에게 기존의 기법과 제안한 기법으로 생성된 영상을 순차적으로 두 번 보여 준다. 그리고 두 번째로 영상을 보면서 평가자는 평가를 하게 된다. 15명의 평가자가 평가를 했으며 테스트 영상으로는 위 결과영상을 사용하여 각각 120frame 총 480frame을 사용하였다. 기존의 기법과 제안한 기법의 순서는 무작위로 설정하였으며 순서를 바꿔가며 총 3회의 평가를 하였다.

각 구간별 시간은 T1, T2, T3 각각 6초, 3초, 5초이다.

평가 결과는 Excellent, Good, Fair, Poor, Bad의 5등급으로 이루어 졌으며 이를 0~100의 점수로 환산하여 평가 하였다. 평가 결과는 아래 표1과 같으며 기존의 기법에 비해 제안한 기법의 입체효과가 우수함을 알 수 있다. 이러한 결과는 기존의 기법이 수평의 움직임에서는 높은 입체감이 있지만 그 이

외의 움직임에서는 입체효과가 낮기 때문이라고 분석된다.

## V. 결론 및 향후 연구 방향

3차원 비디오의 획득 및 처리기술은 아직 보편화 되지 않은 기술로써 통신, 방송, 의료, 교육, 군사 훈련, 게임 등의 다양한 응용 가능성을 가지고 있는 분야로 관심을 받고 있다. 본 논문에서 제안한 영역 분할을 통한 2차원 동영상의 다시점 동영상 생성 기법은 이런 관심을 반영한 결과라고 할 수 있다.

단일 시점의 영상으로부터 깊이 정보를 얻고 이로부터 가상 시점의 영상을 생성하는 것은 매우 어려운 연구이다. 하지만 다시점 카메라나 Depth 카메라가 보급되지 않은 상황과 현재 2차원 영상의 콘텐츠를 활용하기 위해서는 꼭 필요한 연구이다.

본 논문에서는 다시점 동영상의 생성을 위해 영역 분할과 움직임 예측을 하였으며 움직임 정보가 깊이 정보를 반영하는 정보로 사용될 수 있다는 점을 이용하여 움직임 예측의 결과를 깊이 정보로 변환하였다. 또한 단일 시점의 움직임 정보를 깊이 정보로 변환하는 과정에서 움직임을 사용함으로써 생길 수 있는 깊이 정보의 변화가 심한 문제점을 현재 영상의 시간적 주변 영상의 깊이 값을 사용하여 해결할 수 있었다.

하지만 단일시점의 영상에서 깊이 정보를 획득하기 위해서는 본 논문의 움직임 정보 이외의 다른 단서를 활용하는 기법의 연구가 필요하다. 또한 본 논문의 기법 중 영역분할 기법은 영상 색상의 동질성(homogeneity)과 강직성(rigidity)에 민감하기 때문에 이에 강인한 기법과 단일시점으로부터 가상의 시점 생성 시 발생하는 가려진 영역의 처리에 관한 연구는 향후 더욱 연구 되어져야 하는 연구 분야이다.

## 참 고 문 헌

- [1] R. Franich, R. Lagendijk and R. Horst, Reference model for hardware demonstrator implementation, RACE DISTIMA deliverable 45/TUD/IT/DS/B/003/bl, October, 1992
- [2] A. Redert, et al., "ATTEST : Advanced Three-Dimensional Television System Technologies", 3D Data Processing Visualization and Transmission, 2002. First International Symposium, pp.313-319, Jun, 2002



[3] 3DTV NoE, <http://www.3dtv-research.org/>

[4] MPEG/ISO/IEC JTC1/SC29/WG11 "Subjective test results for the CfP on multi-view video coding," N7779, Bangkok, Thailand, Jan. 2006

[5] C. Fehn, "Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV", In Proceedings of SPIE Stereoscopic Displays and Virtual Reality Systems XI, pages 93-104, San Jose, CA, USA, January 2004

[6] 김성열, 이상범, 김유경, 호요성, "깊이맵의 정보 분해와 경계 평탄 필터링을 이용한 다시점 영상 생성 방법", 한국방송공학회 논문지, 제 11권 제 4호, pp.471-482, 2006

[7] 서강욱, 한충신, 유지상, "다시점 영상 합성을 위한 깊이 정보의 전처리", 한국방송학회논문지, 제11권 제1호, pp.91-99, 2006

[8] J. Ross, "Stereopsis by binocular delay", In Nature, Vol.248, pp.354-364, 1974

[9] 김만배, 박상훈, "키 프레임을 이용하는 2차원 동영상의 3차원 입체 변환", HCI 학회, 2002.2.

[10] 홍호기, 백윤기, 이승현, 김동욱, 유지상, "2D H.264 동영상의 3D 입체 변환", 한국통신학회논문지, 제31권 제12C호, pp.1208-1215, 2006

[11] Y. Matsumoto, H. Terasaki, K. Sugimoto, T. Arakawa, "Conversion System of Monocular Image Sequence to Stereo Using Motion Parallax," SPIE Photonic West, Vol.3012, pp.108-115, 1997

[12] S. Cooray, N. O'Connor, S. Marlow, N. Murphy, T. Curran, "Semi-automatic video object segmentation using recursive shortest spanning tree and binary partition tree" WIAMIS 2001 - Workshop on Image Analysis for Multimedia Interactive Services, Tampere, Finland, 16-17 May 2001

[13] Y. Wang, J. Ostermann, Y. Zhang, "Video processing and communication," Prentice Hall, 2002

[14] 박종일, 이노우에세이키, "여러 대의 카메라를 이용한 계층적 깊이정보 추출 알고리즘," 한국방송공학회 논문지 제2권 제1호, pp. 45-56, 1997

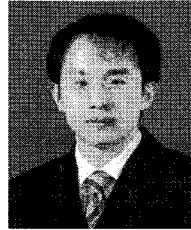
[15] 장세훈, 한충신, 배진우, 유지상, "깊이 정보를 이용한 실시간 다시점 스테레오 영상 합성", 한국통신학회 논문지 제 30권 제 4C호, pp.

239-246, 2005

[16] "Methodology for subjective assessment of the quality of television picture", ITU-R Recommendation BT.500-11

백 윤 기 (Yun-ki Baek)

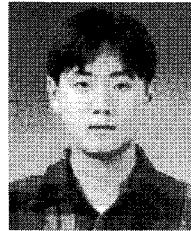
준회원



2006년 2월 광운대학교 전자 공  
학과 졸업  
2006년 3월~현재 광운대학교 전  
자공학과 석사과정  
<관심분야> 3차원 영상처리, 동  
영상 코덱

최 미 남 (Mi-nam Choi)

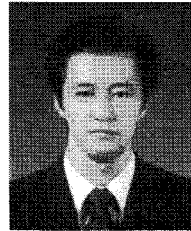
준회원



2006년 2월 광운대학교 전자 공  
학과 졸업  
2006년 3월~현재 광운대학교 전  
자공학과 석사과정  
<관심분야> 3차원 영상처리, 동  
영상 코덱

박 세 환 (Se-whan Park)

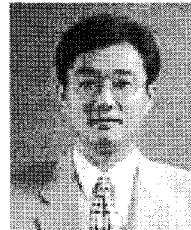
준회원



2007년 2월 광운대학교 전자 공  
학과 졸업  
2007년 3월~현재 광운대학교 전  
자공학과 석사과정  
<관심분야> 3차원 영상처리, 동  
영상 코덱, IVR

유 지 상 (Ji-sang Yoo)

정회원



1985년 2월 서울대학교 전자 공  
학과 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, 영상인  
식, 비선형 신호처리