

특집논문-02-07-3-02

키 프레임 기반 스테레오스코픽 변환 방법

김만배*, 박상훈*

Stereoscopic Conversion based on Key Frames

Man Bae Kim* and Sang Hoon Park*

요 약

본 논문에서는 2D 동영상을 스테레오스코픽(Stereoscopic) 동영상으로 변환하는 새로운 방법을 제안한다. 일반적으로 영상의 운동 정보를 이용하여 스테레오스코픽 영상을 생성한다. 그러나 얻어진 운동 정보는 신뢰도가 낮을 수가 있고, 이것은 스테레오스코픽 영상 생성 단계에서 오류 발생의 원인이 된다. 상기 문제점의 해결 방안으로 운동 벡터 정보의 신뢰도가 높고 추정된 운동 유형이 정확한 키 프레임(Key Frame)을 활용하는 스테레오스코픽 변환 방법을 제안한다. 또한 각 운동 유형에 따른 스테레오스코픽 변환 방법을 소개한다. 제안한 스테레오스코픽 변환 방법의 성능 검증을 위해 다양한 실험 영상에 적용하고 키 프레임의 운동유형의 정확도를 측정하였다. 실험에서는 평균 90 % 이상의 키 프레임과 그 운동 유형을 정확히 찾는 성능을 얻었다.

Abstract

In this paper, we propose a new method of converting 2D video into 3D stereoscopic video, called stereoscopic conversion. In general, stereoscopic images are produced using the motion informations. However unreliable motion informations obtained especially from block-based motion estimation cause the wrong generation of stereoscopic images. To solve for this problem, we propose a stereoscopic conversion method based upon the utilization of key frame that has the better accuracy of estimated motion informations. As well, the generation scheme of stereoscopic images associated with the motion type of each key frame is proposed. For the performance evaluation of our proposed method, we apply it to five test images and measure the accuracy of key frame-based stereoscopic conversion. Experimental results show that our proposed method has the accuracy more than about 90 percent in terms of the detection ratio of key frames.

I. 서 론

기존의 스테레오스코픽 동영상 획득 방법은 스테레오스코픽 카메라 등의 스테레오스코픽 영상 입력 기기를 이용하는 것이었다. 하지만 이 방법은 제작상의 어려움으로 인해 많은 콘텐츠를 제작하기가 어려운 단점이 있다. 이에 반해 본 논문에서 제안하는 스테레오스코픽 변환(Stereoscopic

Conversion) 기술을 이용하면 기존의 2D 동영상을 스테레오스코픽으로 변환할 수 있는 장점이 있다. 그림 1은 스테레

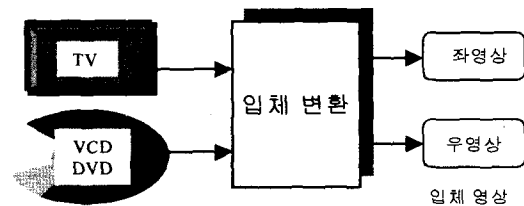


그림 1. 스테레오스코픽 변환의 구조도
Fig. 1. Block Diagram of Stereoscopic Conversion

* 강원대학교 컴퓨터정보통신공학과
Department of Computer, Informations and Telecommunication,
Kangwon National University

오스코픽 변환의 간단한 구조를 보여준다. TV, VCD (Video Compact Disk), DVD(Digital Versatile Disk) 등의 2D 동영상이 모두 변환 대상으로 가능하다. 이러한 2D 동영상은 본 논문에서 제안하는 스테레오스코픽 변환을 이용하여 좌영상과 우영상으로 구성되는 스테레오스코픽 영상으로 생성되게 된다.

지금까지 스테레오스코픽 변환과 관련된 많은 연구가 있었지만 주요 연구를 요약하면 다음과 같다. Garcia는 인간 시각 특성인 공간 시간 보간(spatial-temporal interpolation) 기능을 변환 방법에 적용하였다^[1]. Okino 그룹은 세계 최초로 MTD(Modified Time Difference) 방법을 이용하여 상업용 스테레오스코픽 동영상 변환 TV를 개발하였다^[2]. Kim은 수직 시차의 수평 시차 변환을 이용하여 스테레오스코픽 변환 방법을 제안하였다^{[3][4][5][6]}. Matsumoto는 영상의 깊이 정보를 이용하여 2-D 동영상에서 스테레오스코픽 동영상을 구현하였다^[7]. 본 논문에서는 TV 및 MPEG(Moving Picture Experts Group) 방식의 동영상을 변환 대상으로 하는 새로운 스테레오스코픽 변환 알고리즘을 제안한다. 또한 제안 기술은 동영상 포맷에 관계없이 적용되는 장점이 있다.

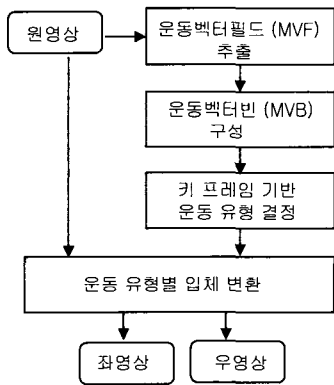


그림 2. 스테레오스코픽 변환 순서도
Fig. 2. Flowchart of Stereoscopic Conversion

그림 2는 본 논문에서 제안하는 스테레오스코픽 변환의 순서도이다. 먼저, 각 (8×8) 블록(Block) 또는 (16×16) 매크로블록(Macroblock) 단위로 운동 벡터(Motion Vector: MV)를 추출한다. 각 영상마다 운동 벡터 필드(Motion Vector Field: MVF)을 얻은 후에 운동 벡터 빈(Motion Vector Bin: MVB)을 구성한다. MVB는 MV의 운동 크기 및 운동 방향을 계산한 후에 해당 셀(Cell)에 저장된다. MVB를 이용하여 현재 프레임이 어떠한 운동 유형에 속하는지를 결정한다. 운동 유형 결정에는 안정된 스테레오스코

픽 동영상을 얻기 위하여 키 프레임(Key Frame: KF) 개념을 이용하였다. KF는 운동 정보의 정확도가 높은 영상을 말한다. 따라서, 신뢰도가 떨어지는 운동 유형을 그대로 이용할 경우 인접한 영상간의 급격한 스테레오스코픽 시차 변화로 인해 눈의 피로감을 발생시키는 문제가 보완된다. 운동 정보의 정확도를 구분하여 높을 경우에는 현재 영상을 KF의 운동 유형으로 결정하고, 반대의 경우에는 가장 최근의 KF의 운동 유형을 이용한다. 마지막으로 결정된 운동 유형에 따라 스테레오스코픽 영상을 생성한다.

본 논문의 구성은 다음과 같다. II장에서는 MVB의 구성을 설명하고, MVB를 이용한 운동 유형의 결정은 III장에서 설명한다. IV장에서는 운동 유형과 원영상에서 스테레오스코픽 영상을 생성하는 방법을 소개한다. V장은 실험 영상들에 적용하여 얻은 결과를 분석하고, 마지막으로 VI장에서는 결론을 맺는다.

II. 운동 벡터 빈 (MVB)

운동 정보를 분석하기 위해서는 먼저 영상의 운동 벡터 필드(MVF)를 추출해야 한다. MVF는 영상에서 보통 블록 또는 매크로블록 당 하나씩 존재하게 되는 운동 벡터(MV)의 집합을 의미한다. MV는 다음 두 가지 방법으로 얻을 수 있다. TV 방식의 동영상에서는 다양한 운동 예측 기법을 이용하여 구할 수 있고^[9], HDTV, VCD, DVD에서 사용하고 있는 MPEG 방식의 동영상에서는 비트스트림(Bitstream)에서 직접 추출할 수 있다^[10].

MVF를 추출한 후에는 운동 유형을 결정하기 위해 사용되는 데이터인 MVB를 구성한다. 그림 3은 MVB의 셀 구성을 나타낸다. MV의 수평 및 수직 성분을 각각 MV_x 및 MV_y 라고 하면 MV의 운동 방향 θ 와 크기 $|MV|$ 는 식

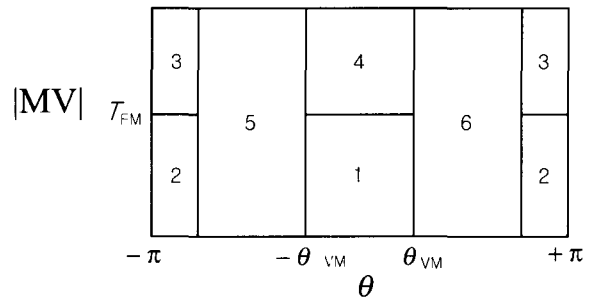


그림 3. Motion Vector Bin(MVB)의 구조
Fig. 3. Structure of MVB

(1)과 같이 구할 수 있다.

$$\theta = \tan^{-1}(MV_y / MV_x)$$

$$|MV| = \sqrt{MV_x^2 + MV_y^2} \tag{1}$$

여기서 θ 와 $|MV|$ 는 각각 수평, 수직 축의 파라메타이다. MVB는 7개의 셀로 구성되며 각 셀은 자신의 운동크기 및 방향 특성에 부합되는 MV들을 포함한다. 그림 3에는 도시되지 않지만 셀 0은 크기가 0인 MV를 포함하게 된다.

고속 운동 임계값, T_{FM} 과 비수평 운동 임계값, θ_v 은 정신물리학에서의 실험 결과에 근거한다^{[6][8]}. 이론에 따르면 동영상 내에서 물체가 수평 방향으로 움직이게 되면 양안 시차가 발생하기 때문에 인간은 깊이감을 느끼게 된다. 이때 물체가 수평 방향으로 움직이되 수평 퓨전 각도 임계값(Horizontal Fusion Threshold), θ_H 이상의 고속 운동이면 인간이 느낄 수 있는 양안 시차의 범위를 넘어가게 된다. 또한 물체가 비수평으로 움직일 경우에는 수직 시차가 발생하게 되는데 수직 퓨전 임계값(Vertical Fusion Threshold), θ_V 이상의 수직 운동이면 인간의 눈에 피로감을 발생하게 한다. 정신물리학에서의 실험 결과를 보면 θ_H 는 27' (minutes of arc)로, 비수평 운동에서의 임계값 θ_V 는 음의 시차에서는 4.93°(degree)로, 양의 시차에서는 1.57°(degree)로 각각 측정되고 있다. 이 값에 의해 θ_{VM} 과 T_{FM} 은 식 (2)와 같이 구할 수 있다.

$$\theta_{VM} = \frac{\theta_V}{60}$$

$$T_{FM} = 2D \tan\left(\frac{\theta_H}{2}\right) \left(-\frac{N_x}{W_x}\right) \tag{2}$$

여기서 θ_{VM} 은 θ_V 를 각도(degree)로 변환한 것을 의미한다. T_{FM} 은 θ_H 를 수평거리로 변환한 값으로 여기서 D 는 인간의 눈과 모니터간의 거리를, N_x 는 영상의 수평 해상도(pixel)를, W_x 는 모니터의 수평 길이(cm)를 각각 의미한다. 그러므로 θ_{VM} 과 T_{FM} 이 결정되면 MV의 크기와 방향에 따라 MVB를 구성할 수가 있다. MVB에서 각 셀의 운동 특성은 다음과 같다. 셀 0은 크기가 0인 운동 벡터들의 집합으로서 정지 영상의 특성을 갖고 있다. 셀 1, 2는 $|\theta|$ 와 $|MV|$ 가 모두 작으므로 수평 저속 운동의 특성을 갖고 있고, 운동 방향은 서로 반대 방향이다. 또한, 셀 3, 4는 수평 고속 운동, 셀 5, 6은 비수평 운동 특성을 가지게 된다.

III. 키프레임 기반 운동 유형 결정

이 장에서는 MVB를 이용한 키 프레임 기반 운동 유형 결정 방법에 대해 설명한다. 운동 유형은 크게 정지(Static), 줌(Zoom), 수평 운동(Horizontal Motion), 비수평 운동 (Vertical Motion) 및 고속 운동(Fast Motion)으로 분류된다. 그러나 MV를 이용하여 구해진 MVB는 영상의 실제 운동 유형과 비교해서 정확도가 떨어진다. 그 이유는 MV를 구하는 방법으로는 블록 매칭(Block Matching) 방식을 많이 사용하는데, 이 경우 MV가 블록의 실제 운동 방향과 비교해서 정확도가 떨어지기 때문이다^[9]. 그림 4는 이렇게 운동 정보의 신뢰도가 떨어지는 MVB의 예를 보여준다. 이 영상은 정지 배경에 인물이 이동하는 영상이다. 그러나 MVB를 보면 인물과 관계없는 MV가 많이 존재함을 관찰할 수 있다. 따라서 이 MVB를 이용하여 얻어진 운동 유형은 정확하지 않게 된다. 따라서 신뢰성 있는 스테레오스코픽 동영상을 생성하는데 어려움이 있다. 본 논문에서는 이를 보완하기 위해서 키 프레임(Key Frame: KF)에 기반한 운동 유형 결정 방법을 제안한다.

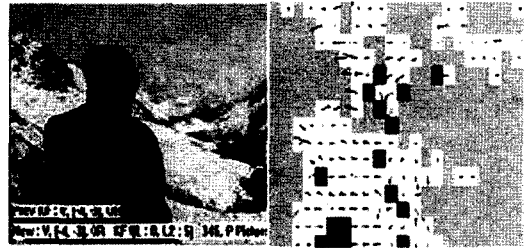


그림 4. 신뢰도가 떨어지는 MVB의 예
Fig. 4. Example of Unreliable MVF

KF는 정확도가 높은 운동 유형을 가지는 영상이다. 따라서 운동 유형 결정에서 운동 정보의 정확도가 높은 경우에는 현재 영상은 KF의 운동 유형을 가지고, 아니면 가장 최근의 키 프레임의 운동 유형과 동일하게 결정한다. 그림 5는 KF로

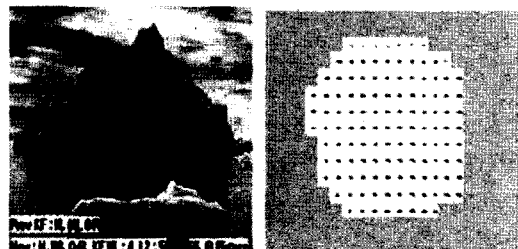


그림 5 키 프레임과 MVB의 예
Fig. 5. Example of Key Frame and its MVF

결정될 수 있는 영상의 예이다. 이 영상은 정지 배경에 물체가 우측 운동을 하는 영상과 MVF이다. MVF를 관찰하면 운동 벡터가 실제 운동과 거의 일치함을 알 수 있다.

그림 6은 본 논문에서 제안하는 KF기반 운동 유형 결정 흐름도이다. 먼저 영상이 줌 영상인지를 판단하게 된다. 줌 영상의 판단에서는 KF 개념에 의한 생략(skip) 조건은 없다. 이 시점부터는 줌 영상이 현재 KF(current KF)로 사용이 된다. 줌 영상이 아니라면 운동 유형 결정의 다음 단계로 넘어가게 된다. 다음은 정지 영상을 결정하는 단계로서 여기에는 KF 개념에 의한 생략 조건이 있는데 이 조건을 만족한다면 운동 유형은 현재 KF의 운동 유형과 같은 것으로 결정된다. 정지 영상으로 결정되었다면 새로운 KF(new KF)는 정지 KF로 설정된다. 아니면 non-KF이므로 다음 단계에서 처리된다. 지금까지의 운동 유형 결정 조건을 하나도 만족하지 않는다면 수평 운동(HM) KF 또는 비수평 운동(VM) KF 중의 하나로 결정하게 된다. 두 운동 모두가 아니면 non-KF이므로 현재 KF의 운동 유형으로 처리되게 된다.

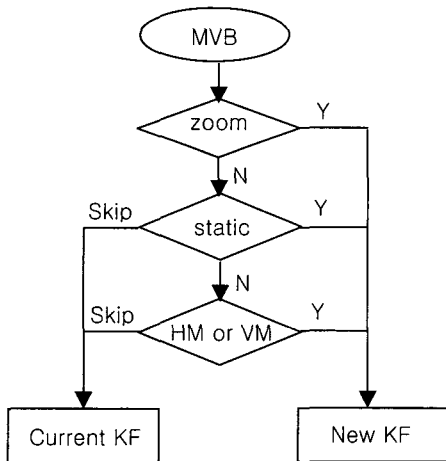


그림 6. KF 기반 운동 유형 결정 흐름도
Fig. 6. Flowchart of Motion Type Decision

줌 영상은 MV의 방향이 중심으로 향하는 줌 인(Zoom In)과 밖으로 향하는 줌 아웃(Zoom Out)으로 구성된다. 이 특성을 이용하기 위해 영상을 4개의 4분면(Quadrant), Q_i 로 분할한다. 각 Q_i 에 존재하는 MV의 개수, N_i 를 구하는데 MV의 방향은 식 (3)을 만족하여야 한다.

$$\frac{\pi}{2}(i-1) \leq \theta_i < \frac{\pi}{2}i, \quad i=1,2,3,4 \quad (3)$$

또한 각 좌 영역과 우 영역에서 최대 MV 개수를 식 (4)와 같이 구한다.

$$N_L = \text{MAX}(N_2, N_3), \quad N_R = \text{MAX}(N_1, N_4) \quad (4)$$

줌 KF를 결정하는 알고리즘은 다음과 같다.

Step 1) N_L 과 N_R 이 T_N 보다 작다면 줌 영상이 아닌 것으로 판단한다. T_N 은 식 (5)와 같이 구한다.

$$T_N = \frac{N_x \cdot N_y}{4} \cdot \alpha_N \quad (5)$$

여기서 N_x 와 N_y 는 가로, 세로 블록 개수이다. α_N 은 최소 MV 개수 임계값이다.

Step 2) N_L 와 N_R 에 속하는 MV를 각각 +, - 방향으로 나누어 식 (6)의 두 조건을 비교한다.

$$T_Z^{-1} \leq \frac{N_{L+}}{N_{L-}} \leq T_Z, \quad T_Z^{-1} \leq \frac{N_{R+}}{N_{R-}} \leq T_Z \quad (6)$$

여기서 T_Z 는 임계값이다. 식 (6)의 두 조건을 만족하지 못한다면 줌 영상이 아닌 것으로 판단한다.

정지 KF 영상은 MVB의 셀 0에 포함되어 있는 MV의 비율, R_S 를 이용하여 결정한다. 즉 R_S 는 MVB에서 $(MV_x, MV_y) = (0,0)$ 인 MV의 비율이다. 식 (7)을 만족하면 정지 KF로 결정한다.

$$R_S \geq T_S \quad (7)$$

여기서 임계값 T_S 는 임의로 정한다. 정지 KF 다음의 연속되는 영상이 식 (7)을 만족하지 않지만 $R_S = [T_S - \alpha_S, T_S]$ 범위에 있다면 현재 KF의 운동유형과 동일하게 결정한다. 임계값 α_S 역시 임의로 정할 수 있다.

수평 운동 KF는 R_H , R_V 를 이용한다. MVB의 셀 i 에 있는 MV 개수를 영상 블록의 전체 개수로 나눈 비율을 $R(i)$ 라고 정의하면 R_H , R_V 는 식 (8)과 같이 구할 수 있다.

$$\begin{aligned} R_H &= R(1) + R(2) + R(3) + R(4) \\ R_V &= R(5) + R(6) \end{aligned} \quad (8)$$

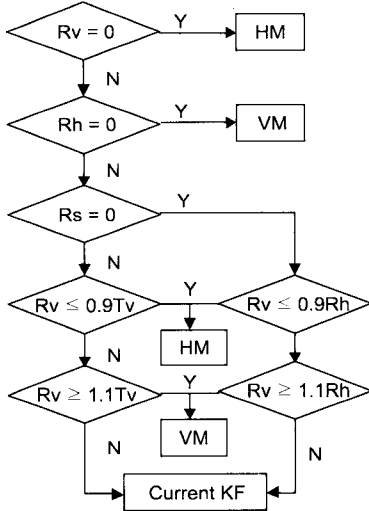


그림 7. 수평 운동(HM) 및 비수평 운동(TM)의 결정 순서도
Fig. 7. Flowchart of Horizontal and Non-Horizontal Motion Decision

그림 7은 수평 운동과 비수평 운동의 결정 흐름도이다. 그림을 보면 수평 운동으로 판단되는 경우는 세 가지가 있다: 1) $R_v=0$, 2) $R_v \leq 0.9$, 3) $R_v \leq 0.9R_H$. 그림의 하단 부분을 보면 R_v 와 T_v 를 비교하고 R_v 와 R_H 를 비교하는 부분이 있다. 각각 0.9와 1.1 사이의 생략 범위가 있는데 여기에 포함될 경우는 non-KF이므로 KF의 운동 유형과 같다.

현재 영상이 수평 운동으로 결정되면 영상 내에 존재하는 카메라와 물체 운동 유형 및 운동 방향 등을 결정해야 한다. 카메라 및 물체의 운동 유형을 결정하는 방법과 같다. R_s 가 임계값 α_s 보다 크거나 같으면 물체 운동으로 정의한다. 반대로 작으면 카메라 운동으로 정의한다. 운동방향은 식 (9)와 같이 구할 수 있는 γ 값을 이용한다.

$$\gamma = \frac{R(2) + R(3)}{R(1) + R(4)} \quad (9)$$

여기서 $R(1)$ 과 $R(4)$ 는 셀 1, 4의 MV의 비율이며 우측 운동 특성을 가진다. $R(2)$ 와 $R(3)$ 는 셀 2, 3의 MV의 비율이며 좌측 운동 특성을 가진다. 물체 운동이고, γ 가 임계값 T_H 보다 크거나 같으면 물체의 좌측 운동, 카메라 운동이면 카메라 우측 운동으로 판단한다. γ 가 T_H 보다 작으면 물체 운동이면 물체의 우측운동, 카메라 운동이면 카메라 좌측 운동으로 판단한다.

비수평 운동 KF는 수평 운동과 같이 그림 7에서 결정된

다. 운동 유형이 비수평 운동으로 결정되면 수평 운동처럼 스테레오스코픽 변환에 필요한 카메라 및 물체 운동 유형과 운동 방향이 필요하다. 비수평 MV의 수평 성분들의 평균인 $AVG(MV_x)$ 가 0보다 작다면, 물체 운동인 경우는 물체 좌측 운동으로, 카메라 운동인 경우는 카메라 우측 운동으로 판단한다. 아니면, 물체 운동은 우측 운동, 카메라 운동은 좌측 운동이다. $AVG(MV_x)$ 는 식 (10)과 같이 계산된다.

$$AVG(MV_x) = \frac{1}{N} \sum_{i=1}^N MV_{xi} \quad (10)$$

여기서 N 은 MVB에서 셀 5와 6의 크기의 합이다.

고속 운동은 수평 운동 중에서 운동 속도가 빠른 것으로 정의한다. 고속 운동은 MVB의 셀 3, 4에 있는 MV 개수로 판단한다. 만약 개수가 T_N 보다 크면 고속 운동으로 결정한다. T_N 은 식 (11)과 같이 정의된다.

$$T_N = N_x \times N_y \times T_F \quad (11)$$

여기서 T_F 는 고속 운동에 대한 임계값이다.

생략 조건으로 정지 KF에서의 α_s 와 수평 및 수직 운동 KF에서 0.9와 1.1 값을 사용하였다. 이 값들은 일반 동영상에서 정지 KF로 판단되는 영상들의 MV의 분석 결과이다. 즉 (0, 0)으로 나오는 MV의 비율을 계산하여 결정한 경험적 수치이다. α_s 의 경우 80~90%의 MV가 (0,0)인 경우에는 정지 KF로 판단하기에 모호하기 때문에 KF 결정을 생략한다. 0.9와 1.1은 수평 운동과 수직 운동에 포함되는 MV들간의 비율이 이 범위에 포함된다면 해당 영상은 복합적인 운동일 가능성이 높기 때문에 안정성을 위하여 KF 결정을 하지 않는다.

지금까지 줌, 정지, 수평, 비수평, 및 고속 KF의 운동 유형을 판단하는 방법에 대해 설명하였다. 다음 장에서는 운동 유형에 따른 스테레오스코픽 영상 생성 방법을 설명한다.

IV. 스테레오스코픽 영상 생성

이 장에서는 각 운동 유형에 따른 스테레오스코픽 영상 생성 방법을 제안한다. 줌 영상의 스테레오스코픽 변환에서는 좌영상은 원영상을 그대로 사용하고 우영상은 원영상을 화소(pixel) 단위의 수평 이동을 하여 구한다. 줌은 물체의

확대 및 축소를 이용하는데, 원점을 중심 좌표로 하여 적절한 위치 이동 값을 설정한다. 이 방법은 양의 시차로 중심에 가까울수록 시차 값을 크게 하여 상대적으로 주변보다 깊이가 있게 한다.

Step 1) 화소의 좌표 (x, y) 에서 원의 중심까지의 거리를 식 (12)와 같이 구한다.

$$r = \sqrt{x^2 + y^2} \quad (12)$$

Step 2) r 을 이용하여 화소의 시차(parallax), p 값을 식 (13)과 같이 구한다.

$$p = \text{MAX}_p \times \frac{r}{\text{MAX}(r)} \quad (13)$$

여기서 $\text{MAX}(r)$ 은 중심에서 모서리까지의 거리이다. MAX_p 는 최대 시차 값으로 사용자가 정의한다.

Step 3) 각 화소를 p 만큼 수평축으로 이동하여 우영상을 만든다.

정지 영상의 스테레오스코픽 변환에서는 좌영상은 원영상을 그대로 사용하고 우영상은 원영상을 $N \times N$ 블록당 수평 이동을 하여 구한다. 블록, B 의 수평 이동 값은 식 (14)와 같이 구한다.

$$P_i = \text{MAX}_p \cdot \frac{\frac{1}{N^2} \sum_{k \in B} x_k}{255} \quad (14)$$

P_i 는 i 번째 블록의 수평 이동 값을 의미하며 MAX_p 는 최대 시차 값이다. x_k 는 화소의 밝기이며 255는 최대 밝기 값이다.

수평 운동 영상의 스테레오스코픽 변환에서는 좌영상과 우영상을 생성하기 위해 현재 영상과 지연 영상(delayed image)을 사용한다. 표 1은 수평 운동의 운동 유형에 따른 현재 영상과 지연 영상을 이용한 좌영상 및 우영상 선택 방법을 보여준다^[6].

표 1. 수평 운동에서 좌우 영상 선택
Table 1. Selection of Left and Right Images for HM

운동유형	좌영상	우영상
물체 좌측	지연 영상	현재 영상
물체 우측	현재 영상	지연 영상
카메라 좌측	현재 영상	지연 영상
카메라 우측	지연 영상	현재 영상

고속 운동 영상은 수평 운동과 같이 원영상과 지연 영상을 사용하여 스테레오스코픽 영상을 생성한다. 그러나 고속 운동은 영상간의 움직임이 크기 때문에 수평 운동과 동일하게 처리할 경우에는 큰 수평 시차 때문에 양안의 퓨전이 어려워진다. 제안 방법은 지연 영상을 사용하여 시차를 줄이고 운동 방향에 따라 수평 이동을 하여 시차를 조절하는 것이다. 지연 영상의 수평 이동 값은 MVB에서 셀 3, 4에 속하는 MV의 수평 성분의 평균이다. 카메라 좌측 운동이거나 물체 우측 운동일 경우는 지연 영상을 수평 이동 값만큼 이동하고, 카메라 우측 운동 또는 물체 좌측 운동일 경우는 반대 방향으로 이동한다.

비수평 운동 영상의 스테레오스코픽 변환은 고속 운동과 유사하다. 비수평 운동도 좌영상과 우영상을 생성하기 위해 현재 영상과 지연 영상을 사용한다. 고속 운동에서는 수평 이동만 수행하였지만, 비수평 운동에서는 수직 시차가 발생하기 때문에 수직 방향 이동도 필요하다. 수평 및 수직 이동 양은 각각 MVB의 셀 5, 6에 존재하는 MV의 수평 성분 및 수직 성분의 평균으로 한다.

V. 실험 및 분석

이 장에서는 제안한 스테레오스코픽 변환 방법을 실험 영상에 적용하여 실험 결과를 보여준다. 블록 매칭 운동 예측 기법을 적용하여 얻은 MV를 이용하거나, 또는 MPEG 데이터와 같이 부호화기(Encoder)에서 구한 MV를 이용할 수 있다. 실험에서는 전자와 후자의 차이가 없기 때문에 MPEG 데이터를 이용하였다. 즉 MPEG 비트스트림에서 직접 블록의 MV를 추출하여 키 프레임(KF)의 운동 유형을 구한 후, 대응하는 스테레오스코픽 변환 기법을 적용하여 스테레오스코픽 동영상을 생성하였다.

실험 동영상으로는 Forest(줌인), Conversation(정지), Cube1(육면체의 수평 운동), Cube2(육면체의 수직 운동), Mountain(카메라 고속 운동)를 사용하였다. Conversation을 제외하고는 각각의 운동 유형별 영상 시퀀스를 인위적으로 제작한 후 MPEG 부호화기로 압축하여 생성하였다. Conversation은 영상 내에서 약간의 움직임 벡터가 존재하는 동영상으로 실험하기 위하여 일반 영화의 MPEG 동영상에서 임의로 선택하였다. 영상 해상도는 352×240 이고 전체 프레임 수는 각각 300장이다. 정지 KF에서의 임계값인 T_s 와 α_s 는 0.9와 0.1로, 고속 운동 KF에서의 임계값인 T_v 는 52로 각각 정하였다.

표 2. 그림 8의 MVB 데이터

Table 2. MVB Data of Images in Fig. 8

실험영상	N_{MV}	N_1	N_2	N_3	N_4	N_5	N_6	N_7	R_S	R_H	R_V
Cubel's KF	260	133	124	0	0	0	3	0	0.51	0.48	0.01
Mountain's KF	260	0	19	0	0	241	0	0	0	1	0

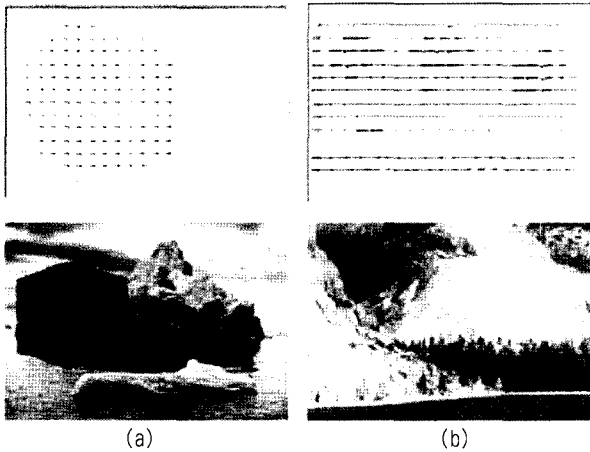


그림 8. 실험 영상에서 KF 영상의 예
Fig. 8. Examples of KF for Test Images

그림 8은 KF로 결정된 영상의 MVF와 비월영상의 예를 보여준다. (a)는 Cubel1에서 KF로 결정된 영상의 예로서, MV가 대부분 수평 방향을 향하는 것을 관찰할 수 있다. (b)는 Mountain에서 KF 영상으로, MV는 대부분 수평 방향이지만 크기가 큰 것을 관찰할 수 있다. 표 2는 그림 8의 영상의 MVB 데이터이다. N_{MV} 는 MV의 전체 개수를 나타내고, N_i 는 MVB에서 각 셀의 크기를 나타낸다. 먼저 (a)의 영상 R_H 는 0.48, R_V 는 0.01로서, 수평 운동 KF 결정 조건에서 $R_V \leq 0.9$ 조건을 만족하여 수평 운동 KF로 결정된 것을 관찰할 수 있다. (b)는 N_4 가 상대적으로 큼을 볼 수 있는데, $N_3 + N_4$ 가 T_S 보다 크므로 고속 운동 KF로 결정된 것을 관찰할 수 있다.

그림 9는 실험 영상들을 제안 알고리즘에 모두 적용한 후, 연속된 5 장의 영상, I_1, I_2, I_3, I_4, I_5 를 임의로 추출한 결과를 보여준다. 각 결과에서 위부분은 해당 영상의 MVF를 나타내고, 아래부분은 생성된 좌우영상의 비월영상을 나타낸다. 각 영상의 우측상단 부분에 KF 표시된 영상은 KF를 의미한다. 먼저 (a)와 (b)는 Forest와 Conversation의 실험 결과로서, Forest는 대부분의 MV 방향이 바깥쪽을 향하므로 모두 줌 KF로 결정되었고 Conversation은 대부분 영상의

표 3. 실험 영상의 성능 측정

Table 3. Performance Measure of Test Images

실험영상	운동 유형	N_T	N_{KF}	N_{CKF}	$R_{KF} (%)$	$R_{CKF} (%)$
Forest	줌인	300	300	276	100	92
Conversation	정지	300	265	216	88.3	91.5
Cube1	수평운동, 물체우측	300	259	252	86.3	97.3
Cube2	비수평운동	300	202	174	67.3	86.2
Mountain	고속운동, 물체운동	300	280	250	93.3	89.3
평균		300	261	234	87.1	89.4

MV가 몇 개 존재하지 않기 때문에 모두 정지 KF로 결정되었다. (c)와 (d)는 Cube1과 Cube2의 실험 결과이다. Cube1은 I_3 를 제외하고는 영상의 대부분 MV 방향이 물체가 움직이는 영역을 중심으로 수평 방향을 갖고 있기 때문에 수평 운동 KF로 결정되었다. 그러나 I_3 는 움직임이 없는 배경 부분에 MV가 많이 존재하므로 상대적으로 운동 정보가 부정확하다. 그러므로 알고리즘에 따라 KF로 결정되지 않고 이전 KF(I_2)의 운동 유형과 동일하게 결정되었다. Cube2는 대부분의 MV가 수직 성분을 갖고 있기 때문에 모두 비수평 운동 KF로 결정되었다. 마지막 (e)는 Mountain의 실험 결과로서, I_1, I_3, I_5 는 MV가 대부분 크기 때문에 고속 운동 KF로 결정되었으나 I_2, I_4 는 운동 정보가 부정확하므로 이전 KF의 운동 유형으로 결정되었다.

표 3은 실험 영상들의 성능 측정값을 보여준다. N_T 는 전체 프레임 개수로 300프레임이다. N_{KF} 는 KF로 결정된 영상의 수를 의미하고, R_{KF} 는 이것의 비율을 나타낸다. N_{CKF} 는 KF 영상 중에서 실제 운동 유형과 같은 영상의 수를 의미하고, R_{CKF} 는 이것의 비율을 나타낸다. R_{KF} 는 동영상에 대한 운동벡터의 정확성을 나타내는데, 이것은 부호화에서 어떠한 움직임 예측 방법을 사용하느냐에 따라 차이가 날 수 있다. 성능을 높이기 위해서는 성능이 우수한 움직임 예측 방법을 사용하는 것이 필요하다. MPEG 실험 영상은 블록 매칭 방법을 사용했다. R_{KF} 는 67%에서 100%까지의 분포를 갖고 있고 평균은 89.4%이다. 블록 매칭 방법을 사용하더라도 전체적으로 높은 비율을 가지게 되는 것을 알 수 있다. R_{CKF} 는 운동 유형 결정 알고리즘의 정확도

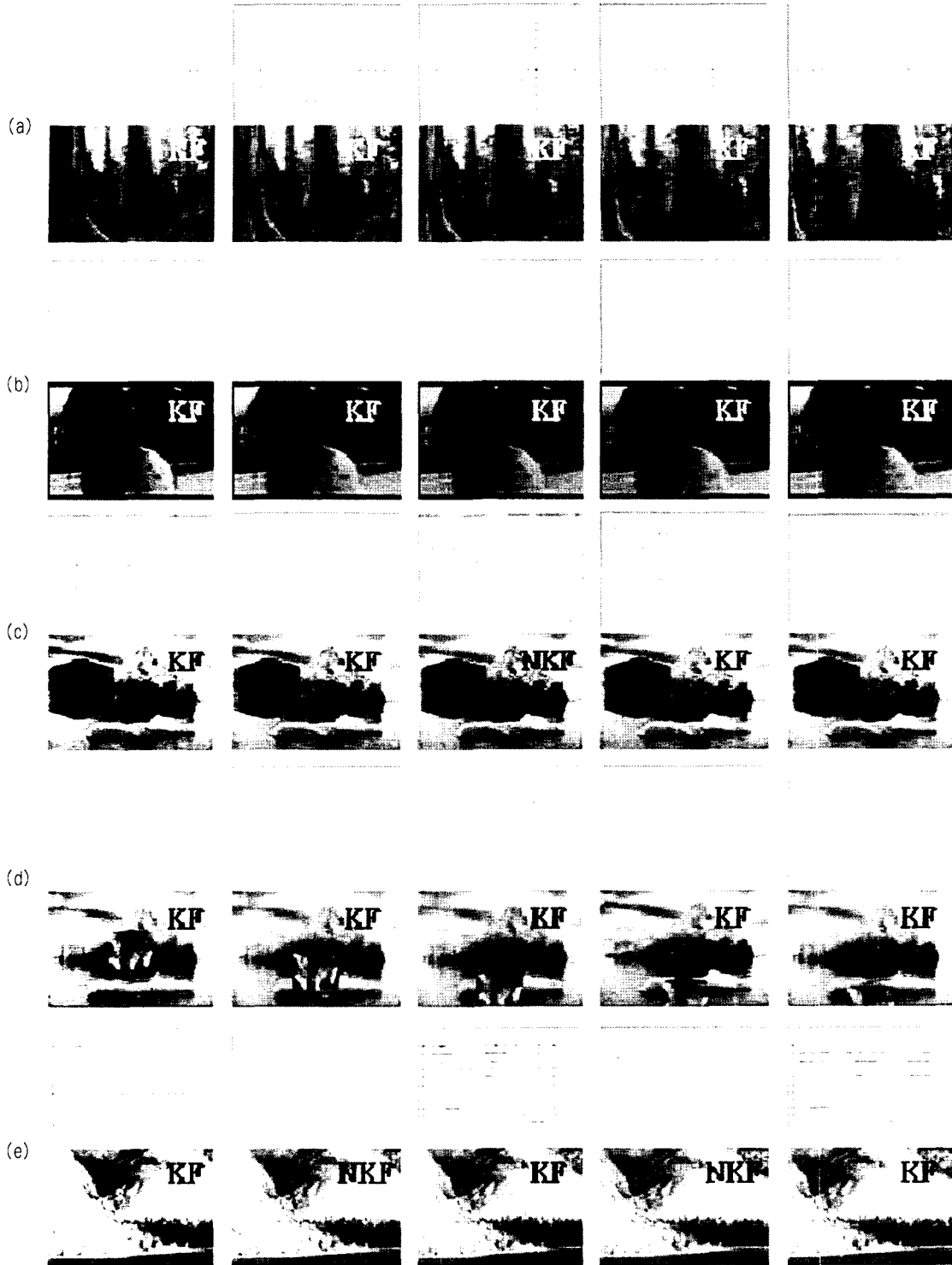


그림 9. 실험 영상의 MVF와 비월영상 (a) Forest (b) Conversation (c) Cube1 (d) Cube2 (e) Mountain
 Fig. 9. MVFs and interlaced Images (a) Forest (b) Conversation (c) Cube1 (d) Cube2 (e) Mountain

를 나타낸다. 81%에서 97%까지 전체적으로 높은 정확도를 가지는 것을 알 수 있는데, 이것은 본 논문에서 제안하는 운동 벡터를 이용한 운동 정보 분석 및 운동 유형 결정 방법이 신뢰도가 있음을 보여준다.

상대적으로 KF를 사용하지 않고, 매 프레임마다 얻어지는 운동 유형을 이용하면 그림 9에서 보는 바와 같이 부정확한 운동 벡터로 인해 바람직하지 않은 스테레오스코픽 변환을 수행하게 된다. 결과적으로 스테레오스코픽 영상의 왜곡이 발생하여 부적절한 스테레오스코픽감을 주게 된다.

VI. 결 론

본 논문에서는 2D 동영상을 스테레오스코픽 동영상으로 변환하는 스테레오스코픽 변환 방법을 제안하였다. 스테레오스코픽 변환 기술의 핵심은 정확한 운동 유형 결정과 이에 따른 스테레오스코픽 변환 방법이라고 할 수 있다. 다양한 운동 유형 결정 및 키 프레임을 이용하여 급격한 스테레오스코픽 시차 변화를 완화하는 두 제안 방법이 본 논문의 주요 사항이다. 각 운동 유형의 키 프레임을 결정하는 알고리즘을 제안하였고, 대응하는 스테레오스코픽 변환 방법을 소개하였다. 줌, 정지, 비수평 운동, 고속 운동 영상의 스테레오스코픽 변환은 본 논문에서 새로 소개하였다.

실험에서는 각 운동 유형 및 만들어진 스테레오스코픽 영상들을 보여주었다. 스테레오스코픽 변환된 동영상을 실시간으로 시청한 결과 안정된 스테레오스코픽을 주는 것으

로 판단되었다. 향후, 스테레오스코픽 변환 기술은 TV, HDTV, VCD, DVD 등에 응용될 예정이다. 영상에서 각 영역마다 다른 스테레오스코픽 변환을 하기 위해서는 프레임에서 객체를 분류하는 영역 분할 과정이 필요하다. 객체 기반 압축 기법인 MPEG-4 데이터의 스테레오스코픽 변환은 제안 기술의 새로운 응용 영역이 될 것이다.

참 고 문 헌

- [1] B. J. Garcia, "Approaches to stereoscopic video based on spatial-temporal interpolation," *SPIE*, Vol. 2635, Photonic West, 1990.
- [2] T. Okino and et. al, "New television with 2D/3D image conversion techniques," *SPIE*, Vol. 2653, Photonic West, 1995.
- [3] 김만배, 최광철, 김도균, 송문섭, "2차원 동영상의 3차원스테레오스코픽 동영상 변환", 제 11회 신호처리합동학술대회, 서울대학교, 1998.
- [4] 김만배, "3차원 스테레오스코픽 변환을 위한 MPEG 압축 데이터에서의 영상 처리 기법", 한국방송공학회 학술대회, 중앙대학교, pp. 3-8 1998.
- [5] M. Kim and et al., "Stereoscopic conversion of monoscopic video by the transformation of vertical-to-horizontal disparity," *SPIE*, Vol. 3295, Photonic West, Jan, 1998.
- [6] Man Bae Kim and Sang Hun Lee, "A new method for the conversion of MPEG encoded data into stereoscopic video," *J. of the Society for 3D Broadcasting and Imaging*, Vol. 1, No. 1, pp. 48-59, Jun, 2000.
- [7] Y. Matsumoto and et al., "Conversion system of monocular image conversion technologies," *SPIE*, Vol. 3012, Photonic West, 1997.
- [8] D. F. McAllister, *Stereo computer graphics and other true 3D technologies*, Princeton, NJ: Princeton University Press, 1993.
- [9] A. M. Tekalp, "Digital Video Processing," Prentice-Hall, 1995.
- [10] K. R. Rao and J. J. Hwang, "Techniques and Standards for Image, Video, and Audio Coding," Prentice-Hall, 1996.

 저 자 소 개

**김 만 배**

- 1983년 : 한양대학교 전자공학과 학사
- 1986년 : University of Washington (Seattle) 전기공학과 석사
- 1992년 : University of Washington (Seattle) 전기공학과 박사
- 1993년 : Georgetown University, Visiting Scientist
- 1996년 : University of Rochester, Visiting Scientist
- 1992년~1998년 : 삼성종합기술원 신호처리/소프트웨어 Lab 수석연구원
- 1998년~현재 : 강원대학교 컴퓨터·정보통신학과 부교수
- 주관심분야 : MPEG 입체 변환, MPEG-21 DIA, Image/Video 신호 처리

**박 상 훈**

- 2001년 2월 : 강원대학교 정보통신공학과 학사
- 2003년 2월 : 강원대학교 컴퓨터·정보통신공학과 석사 (예정)
- 주관심분야 : Image/Video 신호 처리, 멀티미디어 통신