

# 실사 양안식 정지영상 및 동영상 콘텐츠 지원을 위한 합성 방법 연구

이인재, 정세윤, 김규현  
한국전자통신연구원 방송미디어연구부  
[ninja@etri.re.kr](mailto:ninja@etri.re.kr)

## Synthesis Method for Stereoscopic Still Pictures and Moving Pictures

Injae Lee, Seyoon Jeong and Kyuheon Kim  
Broadcasting Media Technology Department,  
Electronics and Telecommunications Research Institute

### 요약

As there is a growing tendency to represent the 3D content instead of the 2D content, researches for the stereoscopic image and video are under way in a variety of fields such as acquisition, compression, transmission, authoring and display. The authoring technique for stereoscopic contents has given emphasis to virtual stereoscopic contents. Thus the authoring technique for stereoscopic pictures is insufficient. When we compose a stereo scene with stereoscopic pictures, stereoscopic contents may not match the stereo scene because each stereoscopic picture may have different camera condition. To solve this problem, stereoscopic pictures have been modified manually. It is a laborious work and will be spent much time. Also it is difficult for a user who does not have an elementary knowledge of stereopsis. In this paper, we propose the synthesis method to compose a natural stereo scene with stereoscopic still pictures and moving pictures. Experimental results show that the proposed method in this paper allows a user to synthesize stereoscopic contents easily and compose a stereo scene conveniently.

### 1. 서론

최근 인간의 감성에 호소하는 영상시스템으로서 3차원 입체 영상에 대한 기대가 고조되고 있다. 이는 일상적으로 보고 있는 자연계의 정경에 보다 가깝고 자연스러운 영상 시스템을 추구하고자 하는 지속적인 인간의 욕망에 의해 보다 많은 정보가 요구됨으로써 발생하는 자연스러운 현상이다. 3차원 입체 영상 기술은 차세대 방송, 통신 서비스의 핵심기술로서 정보통신, 방송, 의료, 교육, 게임, 애니메이션, 가상현실 등 그 응용분야가 매우 다양하다. 또한 3차원 입체 영상 기술은 현재 입체영상생성기술, 편집처리기술, 디스플레이기술, 입체 시각기술 등 다양한 방면에서 기술 개발이 진행되고 있는 단계이다[1]. 편집처리 기술은 영상재구성 기술, 부호화기술, 영상 처리기술 등으로 분류될 수 있다. 이중 영상재구성 기술은 가상 객체를 이용한 VRML 기반의 표현 기술 위주로 발전되었으며, Trimension system Company의 ReACTor[2], MultiGen Company의 SmartScenel[3], 한국전자통신연구

원의 immersive authoring system [4] 등의 VRML 기반 저작도구들이 개발되어 있는 상태이다. 상기 저작도구들은 주로 가상 콘텐츠를 목표로 개발되었으므로 이를 통해 실사 콘텐츠를 이용하여 장면을 구성하면 각각의 실사 콘텐츠들은 각기 촬영조건이 다를 경우 어색한 결과를 낼 수 있다. 또한 가상 객체를 목표로 개발된 VRML 기반의 양안식 콘텐츠 저작도구는 임의형상을 가진 실사 콘텐츠를 지원하지 않으므로 실사 콘텐츠를 다루기에 부적합하다.

따라서 본 논문에서는 실사 정지영상 및 동영상 콘텐츠를 이용하여 보다 자연스러운 장면을 구성하기 위한 합성 방법에 대해 제안하였다. 본 논문의 2장에서는 실사 양안식 콘텐츠를 위한 합성 방법에 대해 기술하였고, 3장에서는 제안된 방법을 이용한 실험결과를 분석하였다. 4장에서는 본 논문의 결론을 맺는다.

### 2. 양안식 콘텐츠 정보 변환 장치를 이용한 합성 방법

양안식 콘텐츠를 이용해 특정 장면을 합성할 때, 장면을 구성할 양안식 콘텐츠들 간의 카메라 정보가 다를 경우 이를 그대로 이용하면 어색한 결과를 낼 것이다. 또한 사용자의 만족도에 따라 양안식 콘텐츠의 깊이감을 조절하고자 할 경우 해당 콘텐츠에 대한 수정이 필요할 것이다. 따라서 본 논문에서는 상기에 기술한 기준 카메라에 기반하여 양안식 콘텐츠 조율 및 깊이감 조절 기능을 수행하는 양안식 콘텐츠 정보 변환 장치를 통해 양안식 콘텐츠 변환에 영향을 미치는 정보들의 값을 구하고, 이 정보들을 합성에 이용함으로써 비교적 적은 노력으로 짧은 시간에 해당 양안식 콘텐츠들을 합성하는데 그 목적이 있다. 양안식 콘텐츠 정보 변환 장치를 통해 획득된 콘텐츠의 위치정보, 깊이 정보, 양안시차, 크기 정보 등은 MPEG-4의 객체 기반 장면 기술을 위해 사용되어질 수 있다.

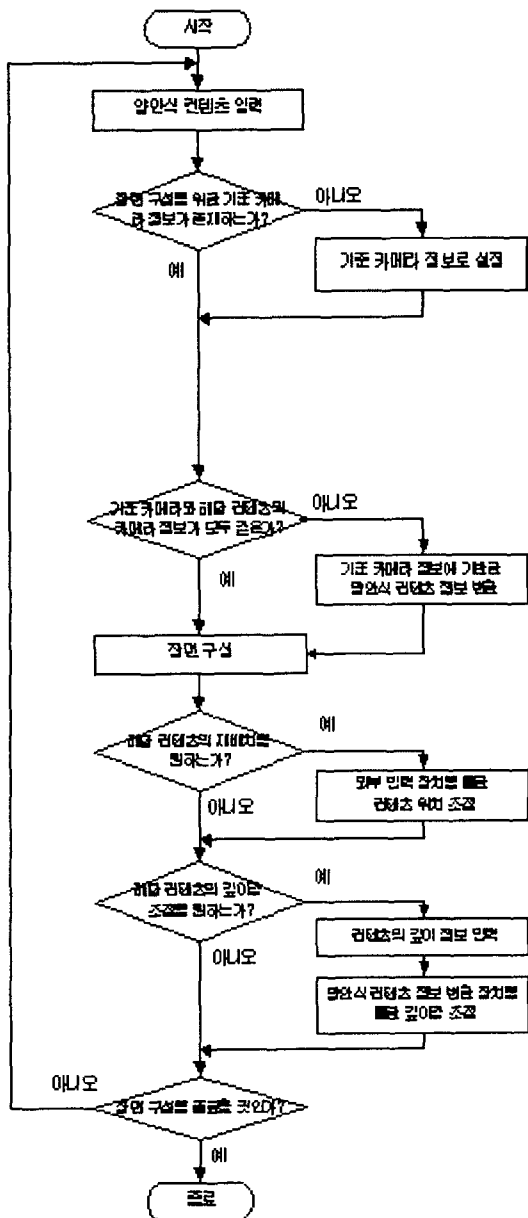


그림 1. 양안식 콘텐츠의 합성 방법에 대한 흐름도

## 2.1. 실사 양안식 콘텐츠를 이용한 장면 합성 방법

그림 1은 양안식 콘텐츠 합성 방법에 대한 처리 흐름도를 나타낸다.

먼저 합성하고자 하는 양안식 콘텐츠가 입력으로 들어오면, 장면 구성을 위한 기준 카메라 정보가 존재하는지 판단한다. 각각의 양안식 콘텐츠마다 촬영 당시 카메라 정보가 다를 수 있으며, 각기 다른 카메라 정보를 가지는 양안식 콘텐츠들을 그대로 이용하여 장면을 구성할 경우 어색한 결과를 낼 수 있으므로 구성하고자 하는 장면의 기준 카메라 정보를 초기에 설정해 주고, 이에 맞게 해당 양안식 콘텐츠들을 변환시킴으로써 보다 자연스럽게 합성시키고자 한다. 만약 기준 카메라 정보가 설정되어 있지 않을 경우, 사용자가 카메라 위치, 방향, 초점거리(focal length) 등의 값을 직접 입력시킬 수 있다. 그러나 컴퓨터 비전에 관한 지식이 없는 일반 사용자들이 카메라 정보를 의미있는 값으로 설정해 주는 일은 쉽지 않고, 경험이 풍부하지 않다면 원하는 장면 구성을 위한 기준 카메라 정보를 예측하는 일 또한 어려운 일이 될 것이다. 따라서 장면 구성 초기에 입력하고자 하는 양안식 콘텐츠에 대한 촬영당시 카메라 정보를 장면 구성을 위한 기준 카메라 정보로 설정해 줄 수 있다.

장면 구성을 위한 기준 카메라 정보가 설정되어 있으면, 이에 맞추어 양안식 콘텐츠를 변환하여 화면에 출력한다. 이를 위해서 먼저 해당 콘텐츠의 촬영당시 카메라 정보와 전체 장면 구성을 위한 기준 카메라 정보를 비교한다. 이때 양안식 콘텐츠 영상 변환에 영향을 미치는 중요한 카메라 정보는 카메라 위치 정보(카메라의 3차원 좌표  $(x,y,z)$ ), 카메라 방향 정보, 초점거리(focal length) 정보 등이 있다. 두 카메라 정보를 비교하여 두 카메라의 위치 정보 중 깊이 관련 정보(카메라의 3차원 좌표 중  $z$  좌표)가 다르거나 혹은 초점거리 정보 등이 다르다면 양안식 콘텐츠 정보 변환 장치를 통하여 해당 양안식 콘텐츠를 기준 카메라 정보에 맞게 변환시킨다. 이렇게 기준 카메라 정보에 맞게 변환된 해당 양안식 콘텐츠를 사용하여 장면을 구성한다.

또한 사용자의 만족도에 따라 해당 양안식 콘텐츠의 재배치를 원한다면 외부 입력 장치를 통해 콘텐츠의 위치를 바꿔 줄 수 있다. 그리고 해당 양안식 콘텐츠의 깊이감을 조절해 주고자 한다면 사용자가 원하는 깊이 정보를 입력 시킨 후, 해당 입력 정보를 기반으로 양안식 콘텐츠 정보 변환 장치를 통해 양안식 콘텐츠를 변환시킨다.

제안된 합성 방법을 통해 보다 편리하게 양안식 콘텐츠를 합성시켜 줄 수 있다.

## 2.2. 양안식 콘텐츠 정보 변환 장치

그림 2는 장면의 합성에 사용되는 양안식 콘텐츠 정보

변환 장치의 입출력 블록도를 나타낸다. 양안식 콘텐츠 정보 변환 장치는 장면 구성시 필요한 기준 카메라 정보와 해당 양안식 콘텐츠의 촬영당시 카메라 정보를 비교하여 카메라 위치 정보 중 깊이 관련 정보, 즉 카메라 3차원 좌표 중 z값만이 다르거나 초점거리 등이 서로 다를 경우, 양안식 콘텐츠 정보 변환 장치를 통해 해당 콘텐츠를 기준 카메라 정보에 맞게 변형시켜 준다. 이는 카메라의 위치 정보 중 깊이 관련 정보(z 좌표) 혹은 초점거리 등의 차이에 따른 상대적인 크기정보, 양안시차 정보를 구해줌으로써 해당 양안식 콘텐츠를 변형시킨다. 또한 사용자의 만족도에 따라 깊이감을 조절해 줄 수 있다. 이를 위해 깊이감 표현에 있어 중요한 요소인 해당 양안식 콘텐츠의 양안 시차 정보를 사용자가 직접 변경시켜 줄 수 있다. 또한 사용자에게는 더 익숙하고 다루기 쉬운 콘텐츠의 깊이 정보를 입력시킬 경우, 깊이 정보에 대응되는 양안시차 정보로 변환해 준다. 또한 보다 자연스러운 장면 구성을 위해 양안시차 정보와 더불어 상대적인 크기 정보도 계산해 준다. 해당 콘텐츠의 깊이 정보가 커질수록 즉, 카메라와 콘텐츠 간의 거리가 멀어질수록 상대적으로 양안시차 정보와 크기 정보는 작아진다[5]. 이러한 특성을 기반으로 깊이정보, 양안시차 정보, 크기 정보, 초점거리 간의 상관관계를 통해 크기정보와 양안시차 정보를 얻어내어 합성시 이용한다.

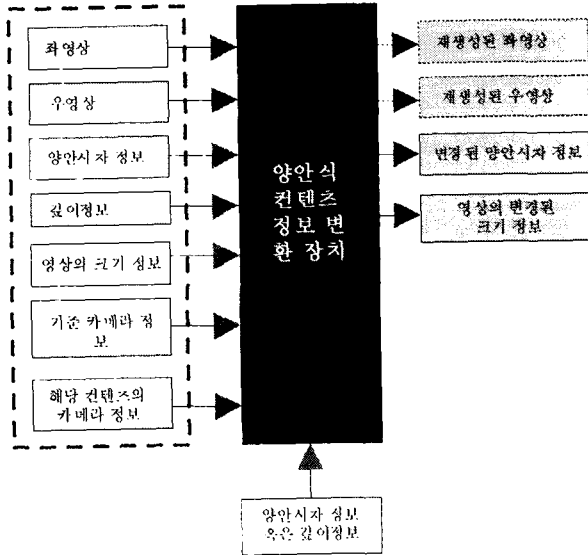


그림 2 양안식 콘텐츠 정보 변환 장치의 입출력 블록도

실사 양안식 콘텐츠의 양안시차 정보, 깊이 정보, 크기 정보는 사용자 입력에 의해 주어지는 양안시차 혹은 깊이 정보를 기반한 상대적인 값으로 변환된다. 만약 사용자에게 의해 깊이정보가 주어질 경우, 양안시차 정보와 크기정보는 식 (1), (2)를 기반으로 계산 되어진다.

$$disparity_{new} = disparity_{source} + \left( \frac{depth_{source}}{depth_{new}} \right) \times \alpha - \left( \frac{focallength_{source}}{focallength_{new}} \right) \times \beta \quad (1)$$

$$size\_x_{new} = size\_x_{source} \times \left( \frac{depth_{source}}{depth_{new}} \right) \times \gamma,$$

$$size\_y_{new} = size\_y_{source} \times \left( \frac{depth_{source}}{depth_{new}} \right) \times \gamma \quad (2)$$

만약 사용자에게 의해 양안시차 정보가 입력되면, 크기정보는 식 (2)에 의해, 깊이 정보는 식 (3)에 의해 변환되어진다.

$$size\_x_{new} = size\_x_{source} \times \left( \frac{depth_{source}}{depth_{new}} \right) \times \gamma,$$

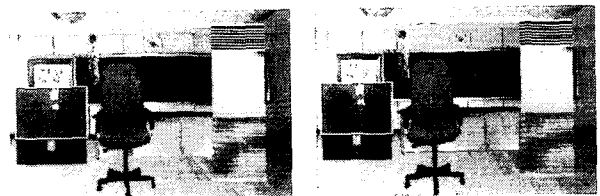
$$size\_y_{new} = size\_y_{source} \times \left( \frac{depth_{source}}{depth_{new}} \right) \times \gamma \quad (3)$$

$$depth_{new} = \frac{depth_{source} \times \phi}{Temp + \phi}$$

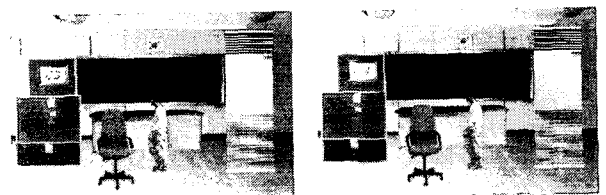
### 3. 실험결과 및 고찰

제안한 합성 방법을 검증하기 위해 JPEG, BMP 포맷의 실사 양안식 정지영상과 AVI 포맷의 실사 양안식 동영상을 사용하여 실험하였다.

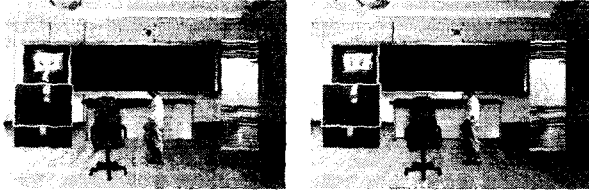
그림 3은 실사 양안식 콘텐츠를 이용한 장면 합성 과정을 보여준다. 그림3에서 좌측의 영상은 해당 양안식 영상의 좌영상, 우측의 영상은 해당 양안식 영상의 우영상을 나타낸다. 그림 3(a)에서 보여지는 것처럼 각기 다른 카메라 정보를 가지는 실사 양안식 콘텐츠를 이용하여 장면을 구성할 경우 어색한 결과를 낼 수 있다. 그림 3(b)는 장면 구성을 위한 기준 카메라 정보를 기반으로 각각의 양안식 콘텐츠들을 변환 시켜준 초기화 결과이다. 그림 3(c)는 사용자의 만족도에 따라 재배치되고 깊이감이 조절된 결과이다.



(a)



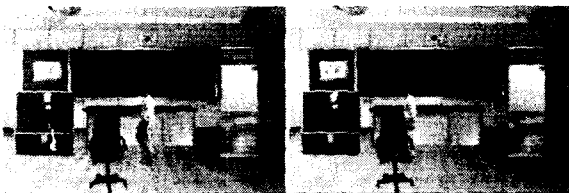
(b)



(c)

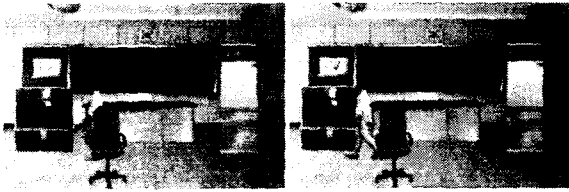
그림 3. 양안식 정보변환 장치를 이용한 장면 합성 결과

그림 4는 합성된 장면의 연속적인 결과이며, 그림에서 보여지듯이 실사 양안식 정지영상 뿐 아니라 실사 양안식 동영상도 장면에 자연스럽게 합성되었다. 그림 3, 4의 합성된 장면 결과에서 보여지듯이 제안한 방법을 통해 실사 양안식 콘텐츠를 이용하여 보다 편리하게 장면을 구성해 줄 수 있었으며, 만족스러운 결과를 얻을 수 있었다.



(a)

(b)



(c)

(d)

그림 4. 실사 양안식 동영상의 연속적인 장면 합성 결과

#### 4. 결론

본 논문에서는 실사 양안식 콘텐츠를 이용하여 보다 자연스러운 양안식 장면을 구성하기 위한 양안식 콘텐츠 합성 방법을 제안하였다. 실사 양안식 콘텐츠들은 각기 다른 카메라 조건 아래 획득될 수 있으므로 이러한 콘텐츠를 이용하여 장면을 구성할 경우 사용자에게 있어 만족스럽지 못한 결과를 낼 수 있다. 따라서 양안식 콘텐츠를 구성하고자 하는 장면에 맞게 합성하는 기술이 필요하다. 제안된 방법은 먼저 서로 다른 카메라 정보를 가지는 양안식 콘텐츠를 기준 카메라 정보에 맞게 변환하여 장면을 구성하고, 양안식 콘텐츠 정보 변환 장치를 통하여 양안식 콘텐츠를 재배치하거나 양안식 콘텐츠의 깊이감을 조절한 후 장면을 재구성하는 두가지 과정으로

이루어진다. 실험결과를 통해 제안된 합성 방법을 이용함으로써 보다 편리하게 만족스러운 결과를 얻을 수 있었으며, 시간 및 노력을 절감할 수 있었다. 또한 제안한 합성 방법을 이용할 경우 양안식 영상에 대한 기반 지식이 없는 사용자들도 쉽게 양안식 콘텐츠를 합성할 수 있다.

제안한 양안식 장면 합성 방법은 3DTV와 같이 양안식 콘텐츠를 다루는 시스템을 위한 저작 도구에 사용될 수 있으며, 양안식 콘텐츠를 위한 렌더러에도 활용될 수 있다.

#### 참고문헌

[1] Bernd Jahne, Horst Haubecker, Computer vision and applications (ACADEMIC Press, 2000).  
 [2] <http://www.trimension-inc.com>, Trimension system.  
 [3] Multigen Inc, SmartScene™ GUI toolkit programmer's guide, 1997.  
 [4] Jeongdan Choi, Chansu Lee, Kwangman Oh, Chanjong Park, A Design and Implementation of Immersive Virtual Space Authoring System, Proceedings of International Conference on Virtual Systems and MultiMedia, Dundee, U.K, 1999, 316-312.  
 [5] Ramesh Jain, Machine vision (MIT Press, 1995).