

영역분할 및 예측을 통한 객체기반 스테레오 동영상 부호화

정회원 권순규*, 배태면*, 한규필*, 정의윤*, 하영호*

Object-based Stereoscopic Video Coding Using Image Segmentation and Prediction

Sun-Kyu Kwon*, Tae-Min Bae*, Kyu-Phil Han*, Eui-Yoon Jung*,
Yeong-Ho Ha* *Regular Members*

요 약

본 논문에서는 스테레오 동영상의 객체기반 부호화 기법을 제안하였다. 스테레오 영상열의 정보량을 줄이면서 블럭화 현상이나 모스quito 현상과 같은 블럭정합 기반 방법의 예측오차를 줄이기 위해서는 객체기반 부호화 기법이 필요하다. 객체기반 방법에서는 부호화에 적절한 객체를 추출하기가 어렵고, 추출된 객체에 대해서 프레임이 지남에 따라 갱신해 주어야 하는 문제점이 발생한다. 이를 개선하기 위해서 제안된 방법에서는 전처리, 객체추출, 객체갱신 과정을 사용하였다. 전처리 과정에서는 움직임과 변이 예측의 신뢰성이 낮은 영역을 비객체영역으로 분할하여 부정확한 객체의 추출을 방지하였다. 객체추출 과정에서는 좌영상의 예측을 향상시키기 위해 기존의 움직임과 밝기값을 고려하는 영상분할법에 변이를 추가함으로써 객체기반 부호화에 적합한 객체를 추출하였다. 그리고 객체갱신 과정에서는 새로운 객체의 추출, 객체분할, 객체 병합을 적용하여 누적 오차를 줄였다.

ABSTRACT

Object-based stereoscopic video coding scheme is presented in this paper. In conventional BMA based stereoscopic video coding for low bit rate transmission, image prediction errors such as block artifacts and mosquito phenomena are occurred. In order to reduce these errors, object-based coding scheme is adopted. The proposed scheme consists of preprocessing, object extraction, and object update procedures. The preprocessing procedure extracts non-object regions having low reliability for motion and disparity estimation. This procedure prohibits extracting inaccurate objects. For the better prediction of left channel image, the disparity information is added to the object extraction. And the proposed algorithm can reduce the accumulated error through the object update procedure that detects newly emerging objects, merges objects that have the same object-disparity and object-motion, and splits object which has large image prediction error. The experimental results show that the proposed algorithms improve the quality of the prediction without block artifacts and mosquito phenomena.

I. 서 론

가상 현실에 대한 관심이 증가함에 따라 3차원 영상을 경험할 수 있는 입체영상 시스템에 대한 필

요성이 증가하고 있다. 특히 3차원 효과를 나타내는 여러 매체 중 좌·우 시각차의 원리를 이용하는 스테레오 영상을 이용한 3차원 영상압축기술^[1,2] 및 영상표현 시스템^[3-7]에 대한 연구가 최근 활발히 이루어지고 있다.

* 경북대학교 전자전기공학부(ksk@palgong.kyungpook.ac.kr)

논문번호 : 99160-0424

접수일자 : 1999년 4월 24일

※본 연구는 정보통신부에서 지원하는 대학기초 연구지원 사업으로 수행되었습니다.

스테레오 동영상열은 좌우 두개의 단일 동영상열로 이루어져 있으므로 정보량이 기존의 동영상들보다 두 배 이상 크므로 효율적인 저장이나 전송을 위해서는 시·공간적 중복성을 효과적으로 제거해야만 한다. 초기에 연구된 스테레오 동영상 부호화기는 기존의 MPEG 방식을 스테레오 영상에 적용하여 블럭정합 방식의 움직임(motion)과 변이(disparity)를 각각 예측함으로써 압축하였다.^[8] 그러나, MPEG 방식의 부호화는 스테레오 영상의 시·공간적 상관성을 고려하기가 어렵기 때문에 영상 예측과 압축에 있어서 한계가 있다. 특히, 예측 영상에 있어서 블럭화 현상 및 모스키토 현상과 같은 블럭정합 기법의 모순들이 발생하고 좌·우 예측영상의 화질차이가 커진다. 이를 극복하기 위해서 점차 스테레오 영상의 특성을 고려할 수 있는 객체(object) 기반의 부호화 기법에 대한 연구가 증가하고 있다.^[9-10]

객체기반 부호화 기법은 움직임과 변이가 동일한 부분을 영상에서 객체로 추출하여 객체 단위로 부호화 함으로써 영상을 압축하는 방식이다. 이는 단일 객체에 대해서 오차가 최소가 되는 움직임과 변이를 가졌기 때문에 영상을 정확하게 예측할 수 있다. 또한 지금까지 연구되어진 스테레오 정합에 대한 종래의 기법을 쉽게 적용할 수 있으므로 정확한 변이 예측과 추출된 변이 정보의 압축에서도 효율적인 부호화를 할 수 있다. 그러나, 기존의 단일 동영상열에 대한 객체기반 부호화기법에서 사용하는 움직임과 밝기값의 유사성을 고려하는 객체분할법을 스테레오 동영상열 부호화에 이용하면 좌·우영상의 공간적인 상관성을 나타내는 변이를 고려하지 않으므로 좌·우 예측영상의 화질은 차이가 크게 된다. 그리고 프레임이 지남에 따라 객체의 움직임으로 인한 예측오차가 발생하는데 이를 개선하지 못하면 동영상열에 대한 효과적인 부호화를 하지 못한다.

본 논문에서는 움직임이나 변이의 예측에 있어서 신뢰성이 낮은 영역을 전처리 과정을 통하여 추출하여, 객체추출에서 이 영역을 제외시킴으로써 잘못된 객체추출에 의한 영상예측 오차를 줄일 수 있었다. 그리고 움직임, 밝기값, 변이의 유사성을 고려함으로써 우영상은 물론 좌영상 예측에도 적합한 객체를 추출하였으며, 새롭게 발생하는 영역, 폐색영역(occlusion region), 예측오차가 큰 객체에 대해서는 유사성에 기반한 영상분할과 병합을 적용하여 객체를 갱신하였다.

II. 기존의 스테레오 영상 부호화

1. 블럭정합 기반의 스테레오 동영상 부호화

기존의 스테레오 동영상 부호화기는 MPEG-2와 같은 단일 동영상열 부호화 기법을 기반으로 스테레오 영상에 확장한 형태로 일반적으로 3가지 방법으로 나누어진다.^[11] 좌·우영상을 독립적으로 부호화하는 동시부호화(simulcast stereo coding), 좌·우영상 중 한쪽 영상을 먼저 부호화하고 다른 쪽 영상은 먼저 부호화한 영상을 이용하여 부호화 하는 겸용부호화(compatible stereo coding), 그리고 좌·우영상을 동시에 부호화 하는 접합부호화(joint stereo coding)가 있다. 동시부호화 방법은 좌·우영상 사이의 상관성을 나타내는 변이 정보를 이용하지 않기 때문에 공간적인 중복성을 제거할 수 없으므로 부호화 효율은 떨어지지만 구현시 병렬화가 가능하고 설계가 간단하다. 이에 비해 겸용부호화 방법은 좌영상열을 변이 보상에 의해 예측하는 방법으로 변이가 움직임에 비해 값이 크고 두 카메라의 기하학적인 특성이나 광학적인 특성이 완전히 동일하지 않을 경우 움직임에 의한 예측보다 정확한 결과를 얻지 못한다.^[12] 접합부호화 방법은 기준영상열에 관계없이 좌우영상열 모두 변이와 움직임 보상에 의해서 예측되는데 겸용부호화에 비해 연산량이 증가하는 단점을 가진다. 그림 1은 일반적인 블럭정합 기반의 스테레오 동영상 부호화기의 구조이다. 이때 dv 는 좌·우영상간의 변이 정보이며, R_{in} 과 L_{in} 은 각각 부호화기의 좌·우 입력영상을 나타내고, R_{out} 과 L_{out} 은 복호기의 복호된 영상이다. 블럭정합 기반의 스테레오 부호화는 블럭 단위로 움직임과 변이를 예측하여 이 정보들로부터 영상을 예측하는 기법으로 전송정보량을 줄일 경우 단일 영상열의 블럭정합에서와 마찬가지로 블럭화 현상과 모스키토 현상 등이 발생하는 단점이 있다.

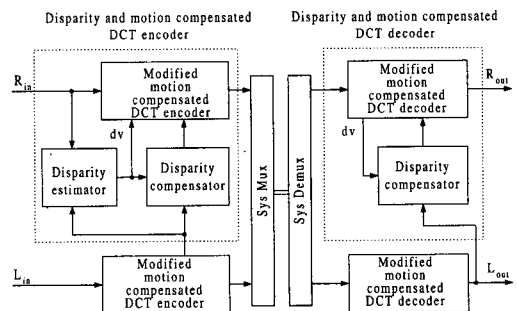


그림 1. 블럭 정합기반 스테레오 부호화기

2. 기존의 객체기반 스테레오 동영상 부호화

단일 영상열 부호화에서의 객체기반 부호화를 스테레오 동영상열에 적용할 경우, 변이가 추가되어야 좌·우 영상을 예측하는데 적합한 객체를 분할할 수 있다. 기존의 객체기반 스테레오 동영상 부호화에서는 객체를 움직임과 변이가 모두 동일한 영역으로 정의하여 영상을 예측하였다.^[13] 이러한 방법은 객체를 움직임과 변이가 모두 동일한 영역으로 정의하였기 때문에 객체 추출의 정확성은 화소 단위의 움직임과 변이 예측에 의존적이고, 객체의 수가 증가하여 전송해 주어야 할 정보량이 증가한다. 특히 화상회의 영상과 배경영역이 균일한 영상의 경우, 배경 영역에서의 움직임과 변이가 실제로는 동일하지만 예측된 값은 동일하지 못하여 정확한 객체 추출이 어렵다. 그림 2는 기존의 객체기반 스테레오 동영상 부호화기의 구조를 나타낸다.

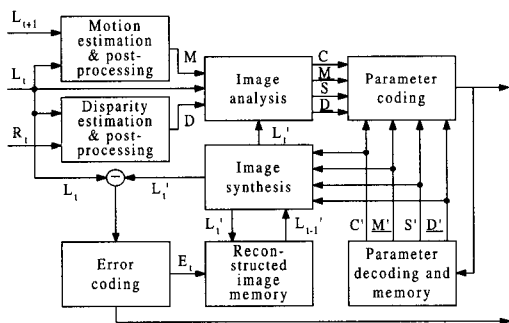


그림 2. 기존의 객체기반 스테레오 부호화기

III. 제안한 스테레오 부호화기

1. 제안한 스테레오 부호화기의 구조

제안한 부호화기는 움직임 및 변이 예측이 정확하지 않는 영역을 비객체영역으로 분할하고 움직임, 변이, 그리고 밝기값이 유사한 영역을 객체로 분할하며, 분할된 객체단위로 움직임과 변이를 찾아 좌·우영상을 예측한다. 그림 3은 제안한 스테레오 영상 부호화기의 블록도이다. 전처리(preprocessing) 블록에서는 현재 프레임의 좌·우영상을 입력으로 사용하여 우영상(기준영상)을 객체영역과 비객체영역으로 분할한다. 추출된 객체영역에 대해서 객체추출(object extraction) 블록에서는 우영상열의 t 프레임과 $t+1$ 프레임, 좌영상열의 t 프레임 영상을 사용하

여 객체를 추출한다. 추출된 객체로부터 우영상은 다음과 같이 예측한다. 객체 움직임 예측(object motion estimation) 블록에서 객체단위의 움직임을 예측하고 움직임 보상법(motion compensation)에 의해서 우영상을 예측하게 된다. 객체갱신(object update) 블록에서는 예측된 우영상과 원영상 사이의 오차 정보로부터 새로운 객체의 추출, 객체분할, 객체병합을 통해서 객체정보를 갱신한다. 갱신된 객체정보로부터 좌영상은 다음과 같이 예측한다. 우영상 예측 전에 예측된 객체 변이(object disparity estimation), 객체갱신 이후에 예측된 객체 변이 그리고 우영상에 대한 객체 움직임 정보를 사용하여 좌영상에 대한 움직임 정보를 예측함으로써 우영상의 경우와 마찬가지로 움직임 보상법에 의해서 좌영상을 예측하게 된다. 예측된 각각의 좌·우 영상에 대해서 PSNR(peak signal to noise ratio)계산 블록에서는 기준객체(reference object) 갱신의 유무를 결정한다. 이 때 부호화기에서 복호화기로 전송되는 정보는 인트라 프레임 정보(intra right, intra left), 객체영역에 대한 정보(object region), 분할된 객체 정보(reference object), 객체의 움직임(object motion) 및 변이정보(object disparity), 객체갱신 정보(object update), 좌·우 영상에 대한 예측오차 정보(right error, left error)이다.

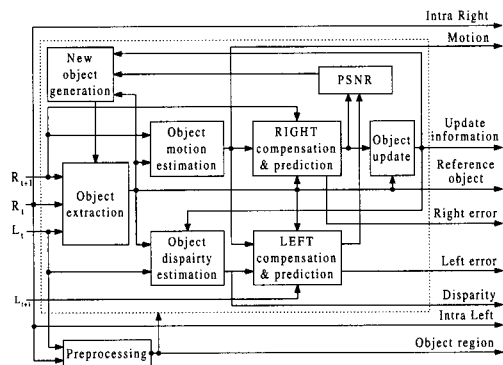


그림 3. 제안한 스테레오 동영상 부호화기의 블록도

2. 객체영역 추출을 위한 전처리 과정

객체기반 부호화에서는 객체를 얼마나 정확하게 추출하느냐에 따라 그 성능이 좌우된다. 잘못 추출된 객체가 많을수록 예측오차에 의해 전송해 주어야 할 정보량이 커지기 때문에 객체의 수를 줄이면서 정확한 추출이 필요하다. 특히 화상회의 영상의 경우 밝기값이 균일한 배경(background)과 사람(foreground)은 반드시 분리되어야 한다. 이러한 배

경영역은 평탄하고 경계선과 같은 특징점이 없기 때문에 객체가 추출되어서는 안된다. 그러므로 객체 추출 이전에 전처리 과정을 통해서 기준영상을 객체영역과 비객체영역으로 분할함으로써 이후 정확한 객체를 추출할 수 있다.

일반적으로 밝기값의 분포가 유사한 영역에서는 움직임 또는 변이 정보를 정확하게 예측할 수 없다. 즉, 블럭밝기값의 차이에 의해서 정보를 예측할 경우 예측된 값에서의 블럭밝기값의 차이가 다른 값에서의 것과 비슷하다면 예측된 그 값은 신뢰성이 낮아지게 된다. 블럭밝기값 차이에 대한 분포를 조사함으로써 영상을 객체영역과 비객체영역으로 분할할 수 있다.

그림 4는 전처리과정을 나타내는 블록도이다. 좌우영상으로부터 변이를 예측할 경우 우영상의 기준블럭에 대해서 좌영상을 탐색하면서 탐색영역 내에서의 블럭밝기값 차이의 분산을 구할 수 있다. 분산값이 작을수록 예측된 값의 신뢰성은 작아지기 때문에 분산값에 대한 문턱치를 적용함으로써 비객체영역(non-object region)에 대한 마커(marker) 후보영역을 추출할 수 있다. 마커 후보영역 추출에 대한 수식적 표현은 다음과 같다.

$$Var[\sum_i \sum_j R(i, j) - L(i+k, j+l)] < T_{var} \text{ for } (k, l) \in A(1)$$

여기서, T_{var} 는 블럭밝기값 차이의 분산에 대한 문턱치값으로 발견적(heuristic)으로 구할 수 있고, A 는 탐색영역을 나타낸다.

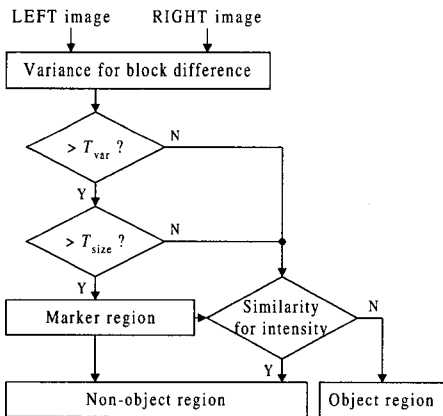


그림 4. 전처리과정의 순서도

마커 후보영역 중에서 크기가 작은 영역은 단순

히 밝기값이 유사한 객체일 확률이 크기 때문에 영역의 크기에 대한 문턱치(T_{size})를 적용함으로써 마커영역을 추출할 수 있다. 문턱치 값은 200화소 정도가 적당하다.

추출된 마커영역과 인접한 블럭들 중에서 밝기값의 분포가 마커영역과 유사한 블럭은 그 블럭밝기값 차이의 분산이 크다고 하더라도 비객체영역으로 볼 수 있다. 예를 들면, 화상회의 영상의 경우 배경영역 중에서 사람의 경계선 부분과 인접한 배경영역은 블럭밝기값 차이의 분산은 크지만 실제로 배경영역에 속한다. 이러한 영역을 객체영역으로 추출할 경우 배경영역에서의 불필요한 객체를 추출하게 되는 오류를 범하게 된다. 마커영역 인접 블럭의 평균 블럭밝기값과 마커영역의 평균밝기값의 차이에 대한 문턱치를 적용함으로써 비객체영역에 대한 마커영역을 성장시킬 수 있다. 마커성장에 대한 수식적 표현은 다음과 같다.

$$\left| \frac{1}{N} \sum_{(i,j)} R(i, j) - \frac{1}{N} \sum_{(k,l) \in \Omega} \sum_{(i,j)} R(i+k, j+l) \right| < T_{avg} \quad (2)$$

여기서, N 은 블럭 내의 화소의 수이고, Ω 는 마커영역의 인접블럭을 의미하며, T_{avg} 는 평균밝기값 차이에 대한 문턱치 값은 실험적으로 2.5이다.

3. 객체추출

객체기반 부호화에서 정확한 객체를 추출하기 위해서는 객체에 대한 적절한 정의가 필요하다. 기존의 단일 영상열에 대한 객체기반 부호화에서는 밝기값과 움직임 정보를 고려한 객체를 정의하였다. 그러나 스테레오 동영상열에 대한 객체기반 부호화 시에는 밝기값과 움직임 정보만을 고려할 경우 좌영상의 예측이 어렵다. 좌영상을 정확하게 예측하기 위해서는 좌우 영상 사이의 공간적인 상관성을 나타내는 변이를 객체추출시 추가해야 한다. 본 논문에서는 밝기값, 움직임, 변이가 유사한 영역을 객체로 정의하였다. 밝기값의 유사성을 고려한 것은 밝기값의 경계부분이 움직임의 경계부분일 확률이 크고 이러한 경계가 객체의 경계일 확률 또한 높기 때문이다. 움직임의 유사성은 움직임 보상(motion compensation)에 의한 우영상 예측시 객체 단위의 움직임을 예측하기 때문에 객체 내의 각 픽셀에서의 움직임과 객체의 움직임에 대한 오차를 줄이기 위해서 필요하다. 이러한 오차를 줄이면 우영상 예측의 오차가 감소한다. 변이의 유사성은 좌영상 예측에 있어서 매우 중요하다. 객체정보는 우영상에

대해서만 존재하기 때문에 좌영상 예측을 위해서는 좌·우 영상의 상관관계를 나타내는 변이정보가 필요하며 객체 단위의 변이 예측에 대한 오차를 줄이기 위해서 객체 추출시 변이의 유사성은 필요하다. 객체추출의 순서는 화소 단위로 변이와 움직임을 예측하고 다중큐를 이용한 변이, 움직임, 밝기값의 유사성에 기반한 영상분할법으로 이루어진다.

3.1. 변이 예측

스테레오 영상은 극상선 조건(epipolar line constraint)을 가정할 경우, 좌·우 영상 사이의 변이는 일차원 탐색만으로도 예측이 가능하다.^[14] 정확한 객체추출을 위해서는 객체의 경계부분이 명확히 구분되어야 하기 때문에 변이를 화소단위로 예측해야 한다. 또한 탐색 범위의 제한 없이 변이를 예측하게 되면 연산량이 증가하고 정확한 예측이 어려워진다. 그러므로 적절한 탐색 범위를 구하기 위해서 계층적 방법을 사용한다.^[13,15] 먼저 큰 블록 단위의 블록정합 방법에 의해서 전 영역에 대해서 변이를 예측한 후, 중간값 필터(median filter)를 사용하여 인접 블록의 변이와 연관시킴으로서 변이도를 평탄하게 만든다. 그리고, 변이의 분포를 조사하여 탐색 범위를 결정한다.

비객체영역은 변이가 동일하다고 보아도 오차가 크지 않기 때문에 하나의 값으로 대체할 수 있다. 그 값은 비객체영역에 대해서 변이분포를 조사하여 가장 많이 존재하는 값으로 정하면 된다. 객체영역 내의 화소들에 대한 변이 예측은 이전 라인의 값과 이전 화소의 값을 고려하는 평탄성 제약(smoothness constraint)을 적용한다. 평탄성 제약을 적용함으로써 객체 내부에서 변이의 불연속성을 방지할 수 있고, 또한 객체경계 부분에서의 불연속성을 유도하여 이후 객체분할에 대한 정확한 정보를 제공할 수 있다. 평탄성 제약을 적용한 비용함수는 아래와 같다.

$$J = \sum_i \sum_j |R(i, j) - L(i, j + d)| + \alpha |d_p - d| + \beta |d_t - d| \quad (3)$$

여기서 d_p 는 이전 화소의 변이, d_t 는 이전 라인의 변이, 그리고 α 와 β 는 각각 이전 화소, 이전 라인에 대한 가중치를 나타내고 d 의 범위는 $[d_{min}, d_{max}]$ 이다.

3.2. 움직임 예측

물체의 움직임이 없고 카메라가 움직이는 경우, 블럭정합 방법으로 움직임을 예측하면 전 영역에서

움직임이 동일하다. 이런 경우 예측된 움직임은 경계 부분에서의 불연속성과 같은 객체추출을 위한 정보를 얻을 수 없다. 그러나 광류(optical flow)를 사용하면 밝기값의 변화에 대한 움직임이 예측되기 때문에 전체 영상에서의 움직임은 다르다.^[16-17] 즉 광류는 영상에 관계없이 객체추출을 위한 움직임 예측이 가능하다.

3.3. 다중큐를 이용한 객체추출

예측된 변이와 움직임에 밝기값을 추가하여 유사성 조건에 기반한 영상분할로부터 객체를 추출한다. 그림 5는 영상분할법을 이용한 객체추출을 나타내는 순서도이다. 먼저 변이, 움직임, 밝기값의 유사성과 화소의 위치에 대해서 우선 순위를 적용하는 단일큐(queue)를 사용하여 영역을 분할하고, 이 영역들

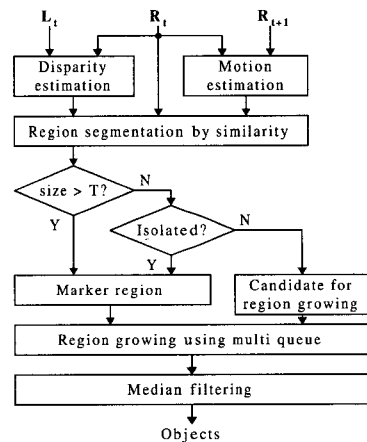


그림 5. 객체추출 순서도

의 크기에 대해서 문턱치를 적용한다. 크기가 작은 영역은 변이와 움직임의 부정확한 예측에 의해서 발생한 것이기 때문에 이런 영역을 객체로 분할할 경우 객체의 수는 많아지고 또한 전송해 주어야 할 정보량도 많아진다. 즉, 모든 영역이 객체가 될 수 있는 것은 아니라, 크기가 문턱치보다 큰 영역이 객체추출을 위한 마커가 된다. 그러나 크기가 문턱치보다 작더라도 그 영역 내에서의 변이와 움직임이 주위의 값들과 차이가 크다면 이 영역은 독립된 영역으로서 마커가 된다. 추출된 마커를 성장시키기 위해서 다중큐(multi-queue)^[19]를 사용하는데, 영역 성장시 성장적도는 변이, 움직임, 밝기값이 된다. 그러나 영역 분할 단계에서는 유사성을 조사할 때 움직임과 변이에 대한 가중치를 많이 할당하지만, 여기서는 밝기

값의 유사성에 대한 기준치를 많이 할당한다. 그 이유는 밝기 차이가 큰 경계부분은 서로 다른 객체의 경계가 될 가능성이 크기 때문이다. 즉, 영역분할 단계에서는 객체의 위치를 찾는 반면 영역성장 단계에서는 할당된 영역에 대한 정확한 경계를 찾게 된다. 추출된 객체들에 대해서 수평·수직 방향의 좁은 선 영역은 움직임에 의해 사라질 확률이 높기 때문에 중간값 필터(median filter)로 제거시킨다.

4. 스테레오 영상특성을 이용한 좌영상 예측

우영상에 대해서 영역 분할하여 객체를 추출하기 때문에 우영상 예측은 움직임 보상(motion compensation)에 의해 가능하다. 그러나 좌영상에 대해서는 객체가 존재하지 않기 때문에 움직임 자체를 얻을 수 없으므로 움직임 보상에 의한 예측은 불가능하다. 좌영상에 대해서 객체를 추출할 경우 객체추출에 필요한 연산량과 전송해 주어야 할 정보량이 증가하므로 바람직하지 않다. 좌영상의 예측은 보통 변이가 움직임보다 범위가 크기 때문에 변이 보상(disparity compensation)보다는 움직임에 의한 예측이 바람직하다.^[20] 스테레오 영상은 움직임과 변이 사이의 상관관계를 가지므로 이를 이용하여 좌영상에 대한 움직임을 예측할 수 있다.^[12] 그림 6은 스테레오 영상열 간의 관계를 나타내는데 움직임 정보에 의해 시간적 상관성을 예측할 수 있고 변이 정보를 통해서 공간적 상관성 또한 예측할 수 있다.

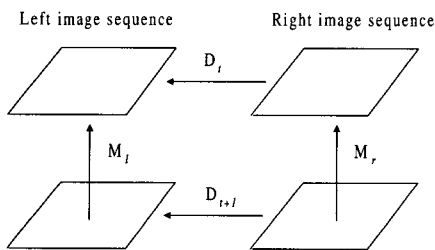


그림 6. 스테레오 영상에서 움직임과 변이의 관계

움직임과 변이는 벡터량이므로 이들의 시·공간적인 합은 0이라는 벡터의 정의로부터 아래 식과 같은 변이와 움직임의 관계 즉, 시·공간 상관관계를 알 수 있다.

$$M_l - D_l - M_r + D_{l+1} = 0 \tag{4}$$

식 (4)로부터 좌영상 움직임을 구하면 다음과 같다.

$$M_l = D_l - D_{l+1} + M_r, \tag{5}$$

식 (5)로부터 좌영상의 움직임은 우영상의 움직임과 현재 프레임의 변이, 다음 프레임의 변이로부터 구할 수 있음을 알 수 있다. 그러나 식 (5)로부터 좌영상 전체의 움직임 정보를 얻을 수 있는 것은 아니다. 우영상과 정합되지 않는 좌영상의 폐색영역에 대해서는 식 (5)가 성립되지 않는다. 폐색영역은 스테레오 카메라의 특성에 의해 우영상에서는 보이지 않는 부분에 의해서 발생하는데 이러한 영역은 움직임 정보가 존재하는 인접영역과 동일한 객체라고 가정할 수 있기 때문에 움직임 또한 동일하다고 볼 수 있다.

5. 객체갱신

본 논문에서 정의한 객체는 시간에 따라 모양이 변하지 않고 그것의 움직임이 2차원으로 제한되기 때문에 프레임이 지남에 따라 오차가 누적된다. 누적 오차를 줄이기 위해서는 프레임마다 객체갱신이 필요하다. 그러나 객체갱신에도 불구하고 누적오차가 커지면 부호화 측에서 좌·우 예측영상의 PSNR을 구해서 문턱치보다 작을 경우 새롭게 객체분할을 수행하여 전송함으로써 해결된다.

5.1. 새로운 객체추출

객체의 움직임에 의해 새로 출현하는 영역인 폐색영역에서는 객체가 존재하지 않고 또한 영상내 부호화(intra-picture coding)해야 하기 때문에 전송해야 할 정보량은 급격하게 증가한다. 폐색영역 내에 인접 객체들과 움직임, 변이, 밝기값이 다른 영역이 존재한다면 새로운 객체로 인식해야 한다. 객체추출에서 사용한 영상분할법을 폐색영역에 적용함으로써 새로운 객체를 인식하고 새로운 객체가 아닌 폐색영역에 대해서는 정보의 유사성을 고려하여 인접 객체에 귀속시킨다. 그러나, 위의 두 경우를 만족하지 않는 영역에 대해서는 영상내 부호화로 처리해야 한다.

5.2. 객체분할

프레임이 지남에 따라 객체내부 화소들의 움직임, 변이, 밝기값은 유사성을 유지하지 못한다. 이러한 객체에 대해서 하나의 움직임을 적용한다면 예측오차는 증가하므로 객체분할이 필요하다. 식 (6)에서와 같이 움직임 보상에 의한 우영상 예측시 각 객체의 예측오차를 계산하여 오차가 허용치보다 큰

객체는 움직임, 변이, 밝기값의 유사성에 기반한 영
상분할법을 사용하여 각기 다른 객체로 분할한다.

$$\frac{1}{N} \sum_{(i,j) \in \text{Object}} |\text{Origin}(i,j) - \text{Predict}(i,j)| > T_{\text{error}} \quad (6)$$

여기서 N 은 객체 내에 존재하는 화소의 수를,
 $\text{Origin}(i,j)$ 는 원영상을, $\text{Predict}(i,j)$ 는 예측된 영
상을, 그리고 T_{error} 는 예측오차에 대한 문턱치값을 나
타낸다.

5.3. 객체병합

프레임이 지남에 따라 단일 객체 내에서 정보가
달라지는 경우도 있지만 인접한 서로 다른 객체의
정보가 동일해지는 경우도 있다. 객체 움직임과 객
체 변이가 동일한 인접 객체를 병합하여 객체의 수
를 줄인다. 객체병합을 통해서 객체의 수를 줄이지
않는다면 다음 프레임의 우영상 예측시 객체의 움
직임에 의한 폐색영역이 증가하여 전송정보량은 증
가한다. 이를 줄이기 위해서 객체병합은 반드시 필
요하다.

IV. 실험 및 고찰

실험에서는 그림 7에 나타나는 384×384 크기를
가지는 25프레임의 'man'영상과 그림 12에 나타
는 640×512 크기의 'aqua'영상이 실험에 사용되었
다. 실험에 사용된 스테레오 영상은 극상선 조건
(epipolar line constraint)을 만족시키므로 일차원 탐
색만을 수행하여 변이를 예측하였다. 'Man'영상은
움직임의 범위가 블럭정합 방법을 사용할 경우 최
대 25화소 정도이고, 변이의 범위는 $-40 \sim -100$ 화소
정도이다. 움직임의 범위가 변이의 범위보다 작기
때문에 좌영상은 변이 보상보다는 움직임 보상에
의해서 정확하게 예측하였다. 그림 8의 (a)는 전처
리과정의 결과를 나타내는데 배경과 사람영역이 뚜
렷이 구분되었음을 볼 수 있다. 그리고 그림 8의
(d)는 추출된 객체도로서 객체의 크기에 대한 문턱
치를 200화소로 하였을 때 초기 객체의 수는 17개
이다. 그림 9는 'man'영상에 대한 제안한 방법을
적용한 결과로서 블럭정합 기반 방법에서 일어나는
경계 부분에서의 블럭화 현상이나 모스키토 현상이
나타나지 않음을 볼 수 있다. 그림 10은 기존의 블
럭정합 기반 방법과 제안한 방법을 비교하기 위해
서 네번째 프레임 좌 예측영상에 대한 경계선 부분
을 확대한 영상이다. 그림 10의 (a)는 경계부분과

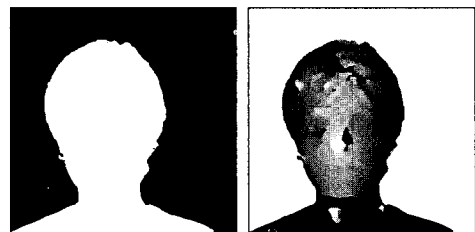
배경부분 등 여러 영역에서 블럭화 현상이 나타나
지만, 그림 10의 (b)에서는 경계부분에서의 블럭화
현상이 나타나지 않고 명확한 경계가 예측됨을 알
수 있다. 이는 제안한 방법이 객체를 분할할 때 밝
기값을 함께 고려하여 밝기값의 경계영역을 객체로
분할하기 때문이다. 특히, 밝기값의 경계영역이 사
람의 화질인식에 더 민감하기 때문에 제안한 방법
에 의한 결과가 주관적인 측면에서도 화질이 더 우
수함을 알 수 있다. 그림 11에서 보면 'man'영상은
우영상의 PSNR이 좌영상보다 더 높는데 그 이유는
우영상열을 기준영상열로 사용하여 우영상보다 좌영
상의 움직임 예측 오차가 크기 때문이다. 'Aqua'영
상은 움직임의 범위가 최대 6화소 정도이고 변이의
범위가 최대 30화소 정도인 영상으로 그림 13에서
볼 수 있듯이 경계부분을 화질의 열화 없이 예측할
수 있었다.

실험 결과에서는 'man'영상과 같은 배경이 균일
한 화상회의 영상의 경우, 배경과 객체를 명확히 구
분하여 경계에서의 열화가 발생하지 않았고, 불필요
한 객체추출에 의한 예측오차를 줄일 수 있었다.
'Aqua'영상과 같은 여러 객체가 존재하는 영상의
경우, 객체 추출에 변이를 함께 사용함으로써 객체
의 경계를 명확히 구분하여 오차를 줄일 수 있었다.



(a) (b)

그림 7. 'Man' 스테레오 영상 (a) 첫번째 프레임 좌영상 (b) 첫번째 프레임 우영상.



(a) (b)

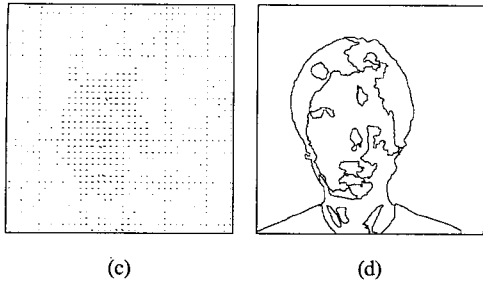


그림 8. 'Man' 스테레오 영상 (a) 전처리과정 결과영상 (b) 변이도 (c) 광류 (d) 객체도

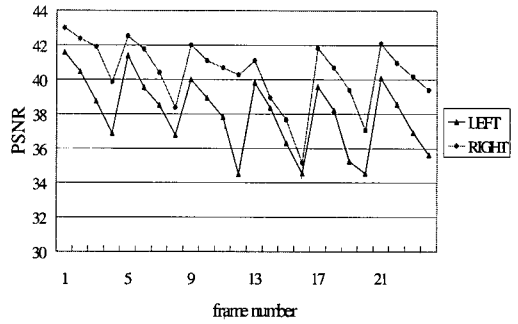


그림 11. 'Man' 영상열에 대한 PSNR 그래프

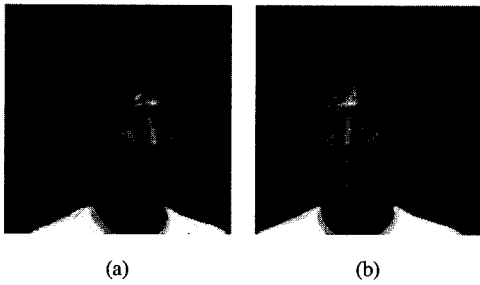


그림 9. 제안한 방법에 의한 결과영상 (a) 네번째 프레임 좌 영상 (b) 네번째 프레임 우영상 (c) 좌 예측오차영상 (d) 우 예측오차영상

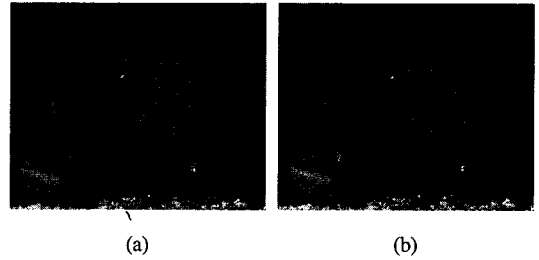


그림 12. 'Aqua' 스테레오 영상 (a) 첫번째 프레임 좌영상 (b) 첫번째 프레임 우영상

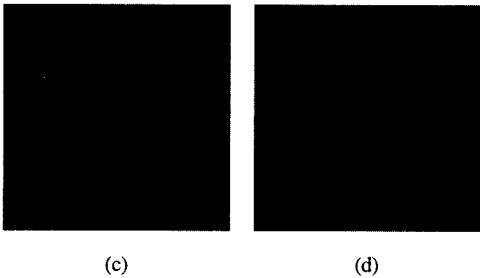


그림 10. 기존의 방법과의 비교 (a) BMA기반의 방법에 의한 결과 (b) 제안한 방법에 의한 결과

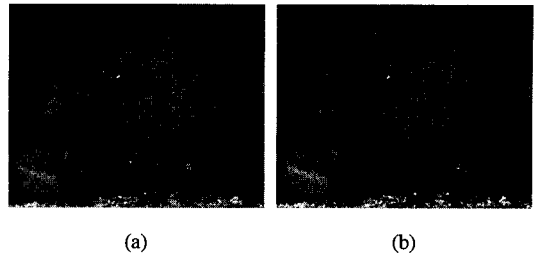


그림 13. 제안한 방법에 의한 예측영상 (a) 네번째 프레임 좌영상 (b) 네번째 프레임 우영상 (c) 좌 예측오차영상 (d) 우 예측오차영상

V. 결론

본 논문에서는 객체기반 스테레오 동영상 부호화기를 제안하였다. 기존의 블럭정합 기반의 스테레오 동영상 부호화기가 전송정보량을 줄일 경우 경계부분에서의 화질저하를 가져오는데 반해 제안된 부호

화기에서는 움직임 및 변이 예측이 어려운 영역을 비객체영역으로 정의하여 잘못된 객체추출을 방지하였고, 기존의 움직임과 밝기값의 유사성을 고려한 객체추출 방법에 변이를 추가하여 좌영상 예측을 개선하였으며, 객체갱신을 통해 새롭게 발생하는 영역과 기존 영역의 소멸 및 변형에 의한 부호화 효율저하를 줄일 수 있었다. 실험 결과에서는 제안된 방법이 기존의 블럭정합 기반의 부호화에서 나타나는 블럭화 및 모스키토 현상을 제거할 수 있고 예측영상 화질을 개선시켰다는 것을 보여주었다.

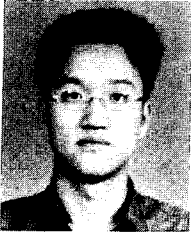
제안된 방법에서는 객체가 단단하고 그 움직임이 2차원이라고 가정을 하였기 때문에 몇 가지 문제점이 발생한다. 이는 객체의 3차원 모델링과 더불어 3차원 공간상의 움직임을 예측함으로써 개선할 수 있을 것이다.

참 고 문 헌

- [1] F. Chassaing, B. Choquet, and D. Pele, "A stereoscopic television system (3D-TV) and compatible transmission on a MAC channel (3D-MAC)", *Signal Processing: Image Communication*, vol. 4, pp. 33-43, 1991
- [2] R. Horst, "A digital codec for 3D-TV transmission", *Signal Processing of HDTV*, pp. 489-495, 1993.
- [3] E. M. Izquierdo, "Stereo image analysis for multi-viewpoint telepresence applications," *Signal Processing: Image Communication*, vol. 11, pp. 231-254, Mar. 1998.
- [4] S. Pastoor and K. Schenke, "New auto-stereoscopic display system," *Proc. SPIE*, pp. 31-40, Feb. 1995.
- [5] M. R. Jewell, G. R. Chamberlain, D. E. Sheat, P. Cochran, and D. McCartney, "3D imaging systems for video communication applications," *Proc. SPIE*, pp. 4-10, Feb. 1995.
- [6] E. Izquierdo, "Stereo matching for enhanced telepresence in 3D-video communications," *IEEE Trans. Circuit and System for Video Technology*, vol. 7, pp. 629-643, 1997.
- [7] A. Puri, R. V. Kollarits, and B. G. Haskell, "Compression of stereoscopic video using MPEG-2," *Critical Reviews Conference on Standards and Common Interfaces for Video Information Systems, published by SPIE*, Oct. 1995.
- [8] D. Tzovaras, M. G. Strinzi, and H. Sahinoglou, "Evaluation of multiresolution techniques for motion and disparity estimation," *Signal Processing: Image Communication*, vol. 6, no. 1, pp. 59-67, Mar. 1994.
- [9] D. Tzovaras, N. Grammalidis, and M. G. Strinzi, "Object-based coding of stereo image sequences using joint 3-D motion/disparity compensation," *IEEE Trans. on Circuit and System for Video Technology*, vol. 7, no. 2, pp. 312-327, Apr. 1997.
- [10] G. Martinez, "Shape estimation of articulated 3D objects for object-based analysis-synthesis coding(OBASC)," *Signal Processing: Image Communication*, vol. 9, pp. 175-199, 1997.
- [11] A. Puri, R. V. Kollarits, and B. G. Haskell, "Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4," *Signal Processing: Image Communication*, pp. 201-234, Oct. 1997.
- [12] 배태면, 한규필, 김용석, 이호영, 하영호, "다해상도 블럭정합 기법을 이용한 스테레오 동영상의 부호화," *한국통신학회 하계종합 학술 발표대회 논문집*, pp. 420-424, 1997.
- [13] S. Panis, M. Ziegler, and J. P. Cosmas, "The use of stereo and motion in a genetic object-based coder," *Signal Processing: Image Communication*, vol. 9, pp. 221-238, 1997.
- [14] D. Geiger, B. Landendorf, and A. Yuille, "Occlusions and binocular stereo," *International Journal of Computer Vision*, 1993.
- [15] D. Tzovaras, N. Grammlidis, and M. G. Strinzi, "Disparity field and depth map coding for multiview 3D image generation," *Signal Processing: Image Communication*, vol. 11, pp. 205-230, 1998.

권 순 규(Sun-Kyu Kwon)

정회원



1997년 2월 : 경북대학교 전자

공학과 졸업

1999년 2월 : 경북대학교 전자

공학과석사

1999년 3월~현재 : 경북대학교

전자공학과 박사과정

<관심분야> 영상신호처리, 컴퓨터 비전 등

배 태 면(Tae-Min Bae)

정회원

한국통신학회 제 23권 9호 참조

현재 : 경북대학교 대학원 전자공학과 박사과정

한 규 필(Kyu-Phil Han)

정회원

한국통신학회 제 23권 9호 참조

현재 : 경북대학교 대학원 전자공학과 박사 졸업

정 의 윤(Eui-Yoon Jung)

정회원

한국통신학회 제 23권 9호 참조

현재 : 경북대학교 대학원 전자공학과 박사과정

하 영 호(Yeong-Ho Ha)

정회원

한국통신학회 제 24권 6호 참조

현재 : 경북대학교 전자전기공학부 교수