

MPEG-4 동영상 부호화를 위한 객체 합성 방법

이 정 우, 호 요 성

광주과학기술원 정보통신공학과

전화 : 062-970-2263 / 핸드폰 : 011-9610-2943

Object Composition for MPEG-4 Video Coding

Jeong-Woo Lee and Yo-Sung Ho

Kwangju Institute of Science and Technology (K-JIST)

E-mail : {jeongwoo, hoyo}@kjist.ac.kr

요 약

본 논문은 MPEG-4 비디오 부호화 알고리즘에서 객체 기반 부호화를 위한 비트율 제어를 수행하는 경우에, 복호기에서 객체들을 합성하는 과정에서 발생할 수 있는 객체 합성 문제(composition problem)에 대해 소개하고, 이를 해결하기 위한 알고리즘을 제안한다. 제안된 알고리즘은 복호기에서의 합성문제 최소화 알고리즘과 복호기에서의 합성문제 제거 알고리즘으로 구성된다. 제안된 알고리즘은 인간의 시각에 영향을 주지 않으면서 객체기반 부호화를 수행하여 부호화 성능을 향상시킬 수 있다.

I. 서 론

MPEG-4 비디오 부호화 알고리즘[1]은 기존의 다른 표준과는 달리, 객체기반 부호화를 지원하기 때문에 각 객체는 서로 다른 시간 해상도(temporal resolution)를 가질 수 있다. 그래서 비트율 제어를 수행하는 데 있어서 MPEG-4 알고리즘은 H.263이나 MPEG-2와 같은 다른 비디오 부호화 방식에 비해 프레임율과 양자화 변수를 결정하는데 있어서 큰 유연성을 가진다. 또한 비트율 제어가 객체 단위에서 수행될 수 있기 때문에 프레임 기반 부호화 알고리즘에 비해 부호화 성능을 극대화 할 수 있는 장점을 가진다.

MPEG-4 부호기를 통해 부호화된 각 객체 비트열은 복호기에서 복원된 후에 사용자에게 의도에 따라 한 화면으로 합성되어 재생되는 과정을 거치는데, 각 객체가 서로 다른 해상도를 가질 수 있기 때문에 하나로 합성된 화면에서 사용자가 원하지 않는 부분에서 각

객체에 의해 정의되지 않는 영역이 생기는 객체 합성 문제(composition problem)가 발생할 수 있다.

현재까지 제안된 MPEG-4 비트율 제어 알고리즘은 부호화되는 모든 객체가 같은 시간 해상도를 가지는 경우에 대해 적용되었다 [2, 3]. 즉, 현재 부호화 시간에 부호화하고자 하는 객체들은 모두 부호화되든지 아니면 모두 부호화되지 않는다. 객체기반 비트율 제어를 수행하는 경우[4, 5]에도 부호화 성능은 최대한 높일 수 있지만, 객체 단위로 부호화했기 때문에 합성 문제를 해결하는 알고리즘의 개발이 필요하다.

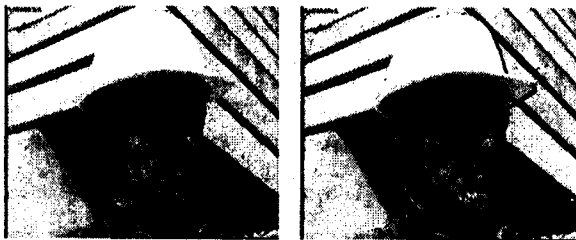
본 논문에서 제안된 알고리즘은 복호기에서 모양 힌트 정보를 추출하여 합성 문제를 최소화하는 알고리즘과 복호기에서 합성 문제를 일으키는 영역을 검출하고 이를 복원하는 알고리즘으로 구성된다. 즉, 복호기에서 합성 문제가 최소화되도록 비디오 객체들을 부호화하고, 복호기에서 합성 문제가 일어나지 않도록 합성 문제가 일어난 영역을 복원한다.

II. 객체 합성 문제

그림 1은 FOREMAN 시퀀스에서 발생한 합성 문제의 예를 보인 것이다. FOREMAN 시퀀스는 전경을 나타내는 객체와 배경을 나타내는 객체로 구성되어 있다. 그림 1(a)는 두 객체를 모두 30Hz로 부호화하고 이를 복호화한 후에 객체를 합성한 결과이고, 그림 1(b)는 전경을 나타내는 객체를 30Hz로 부호화하고 배경을 나타내는 객체를 15Hz로 부호화하여 이를 복원한 후에 객체를 합성한 결과이다.

그림 1(a)에서 볼 수 있듯이, 두 객체가 같은 시간 해상도로 부호화되는 경우에는 복호기에서 영상을 복원하는 과정에서 합성 문제가 발생하지 않으므로 인간

의 시각에 별 영향을 주지 않는다. 즉, 합성 화면내에 있는 모든 위치의 화소가 완전히 정의된다. 하지만, 그림 1(b)에서 볼 수 있듯이, 각 객체가 다른 시간 해상도로 부호화되는 경우에는 인간의 시각에 영향을 주는 문제가 발생한다. 즉, 시퀀스 내에서 객체의 모양이 움직임을 가지고, 해당 객체들이 다른 시간 해상도로 부호화된 경우에 재구성된 화면에서 복원된 객체의 모양 정보로 정의되지 않는 부분이 발생한다. 이러한 영역들은 인접한 객체 혹은 겹쳐진 (overlapping) 객체들은 갱신되지 않고, 갱신되는 객체들은 움직임을 가지기 때문에 발생한다.



(a) 같은 해상도인 경우 (b) 다른 해상도인 경우

그림 1. 합성 문제의 예

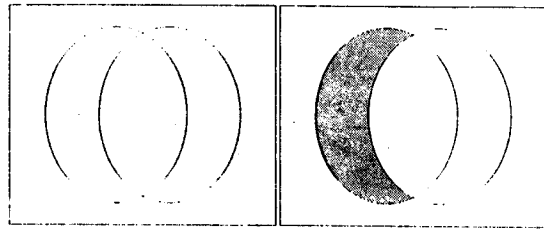
이와 같이 한 영상내에 있는 서로 다른 객체가 다른 시간 해상도를 가지게 되어 합성 영상 내에 객체들의 모양 정보로 정의되지 않는 영역이 발생하는 문제를 합성 문제라고 한다. 이러한 합성 문제가 발생한 영역들은 모든 객체들이 같은 시간에 부호화되어 재동기화되는 경우에 사라진다.

III. 모양정보 추출

움직임 객체의 모양 정보를 추출하는 가장 잘 알려진 방법은, 그림 2(a)에 보이는 바와 같이, 이전 화면과 현재 화면 사이의 모양 정보를 이용하여 달라진 화소의 개수를 계산하는 Hamming Distance를 이용하는 방법이다. 하지만, 그림 2(b)에 보이는 바와 같이, 합성 문제 영역은 건너뛰어진 객체들이 정의하지 못하는 화소들로 구성되기 때문에 시간 i 에서 j 번째 객체의 모양 정보 $d_{j,i}$ 는 다음과 같이 정의된다.

$$d_{j,i} = |A_{j,p} - A_{j,c}| \quad (1)$$

여기서 $A_{j,p}$ 와 $A_{j,c}$ 는 각각 j 번째 객체의 이전 화면과 현재 화면에서 모양 정보의 화소 위치를 포함하는 집합이고, "-"은 집합의 차(difference), "|·|"은 집합 원소의 개수(cardinality)를 계산하는 집합 연산을 나타낸다.



(a) Hamming Distance (b) Composition Distance

그림 2. 모양 정보 추출 방법

IV. 객체 합성 방법

1. 부호기에서 객체 합성 방법

먼저 부호기는 현재 시퀀스를 구성하는 객체들이 합성 문제를 발생시킬 가능성이 있는지를 조사한다. 합성 문제가 발생하기 위해서는 적어도 두 개의 움직이는 객체가 필요하므로 객체의 수가 2보다 적으면 합성 문제가 발생하지 않는다.

이 조건을 만족하지 않는 경우에도 객체들의 움직임이 복원된 화면의 PSNR 값에 주는 영향이 미미한지 여부를 조사하는 것이 필요하다. 이를 위해 각 객체의 모양 정보를 이용하여 객체의 움직임 정도를 계산한 후에 실험적으로 구해진 임계값과 비교하여, 모양 움직임 정보가 임계값보다 큰 경우에는 복원된 화면의 PSNR 값에 큰 영향을 끼친다고 결론을 내릴 수 있다.

하지만 움직임을 가지는 객체의 모양 정보가 임계값을 초과하는 경우라도, 각 객체의 움직임이 같은 방향으로 발생한다면 객체의 움직임은 복원 화면의 PSNR 값에 주는 영향이 적을 것이다.

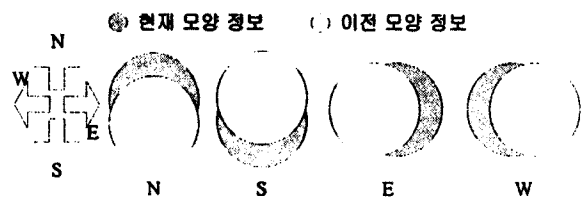


그림 3. 움직임 객체의 방향 측정

그림 3에 보인 것처럼, 움직임 객체의 방향을 구하기 위해 해당 객체의 이전 화면의 모양 구성 정보와 현재 화면의 모양 구성 정보의 차를 이용한다. 비록 본 논문에서는 4 방향을 고려한 간단한 알고리즘을 이용하였지만, 객체의 움직임 방향을 측정할 수 있는 다른 알고리즘을 사용할 수 있다.

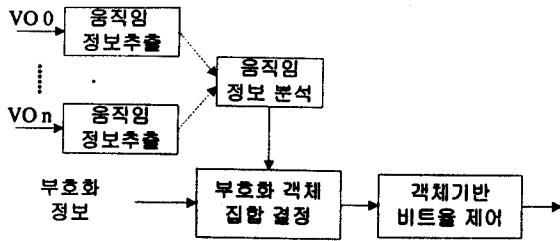


그림 4. 부호화 과정

그림 4에 보이는 바와 같이, 부호기는 먼저 각 객체의 모양 정보를 구한 후에 모양 정보 분석기를 통해 반드시 함께 부호화 되어야 하는 객체들을 검사하고, 이를 고려하여 객체기반 부호화를 수행한다.

실험을 통해 제안된 알고리즘은 비트율 측면에서 25% 정도 감소 효과를 가져왔으며, 각 객체들간의 모양 경계 사이가 매우 작기 때문에, 재구성 화면에서 합성 문제를 일으키는 영역은 PSNR 값에도 영향을 거의 주지 않는다. 하지만 인간의 시각은 PSNR 값에 영향을 주지 않는 구멍에 대해서도 민감하다는 것을 주목해야 한다. 이러한 현상을 없애기 위해 복호기에서도 합성문제 제거 알고리즘을 고려한다.

2. 복호기에서 객체 합성 방법

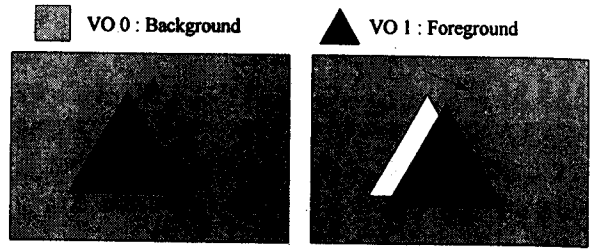
복원 화면에서 발생하는 합성 문제를 완전히 제거하기 위해 제안된 알고리즘은 먼저 복원 화면에서 문제가 발생하는 영역을 검출한 후에, 검출된 영역을 복원한다. 이를 위해 영역을 검출하는 알고리즘과 영역을 복원하는 알고리즘이 요구된다.

화면을 구성하는 객체들 중에 하나는 부호화되고, 나머지 객체들은 건너뛰어졌다고 가정하자. 복원되어야 하는 영역이 검출되었다면, 부호화 객체는 이미 완전한 모양 정보를 가지고 있기 때문에 검출된 영역의 화소들은 건너 뛰어졌던 객체들의 영역에서 복원되어야 한다. 합성 화면에서 문제 영역을 검출한 후에, 해당 영역 내에 있는 각 화소는 건너 뛰어졌던 인접한 객체들의 모양 경계와 최소의 거리를 가지는 위치의 값으로 복원된다.

그림 5는 배경에 해당하는 객체가 화면 전체 크기와 같은 경우에 문제 영역을 어떻게 검출하는지 보여준다. 시간 i 에서 합성 화면 내에 존재하는 문제 영역을 나타내는 집합 H_i 는 다음의 식에 의해 계산된다.

$$A_{j,d} = A_{j,b} - A_{j,c} \quad (2)$$

$$H_i = \bigcup_{j=0}^{M-1} A_{j,d} - \bigcup_{j=0}^{M-1} A_{j,c} \quad (3)$$



(a) 이전 프레임

(b) 현재 프레임

그림 5. 화소 복원 (배경이 있는 경우)

여기서 M 은 비디오 객체의 수이고, H_i 는 합성 문제가 발생하는 영역에 해당하는 화소의 위치 정보를 저장한다.

합성 화면이 배경 객체를 포함하지 않는 임의의 모양을 가지는 객체들로 구성되는 경우에 제안된 영역 검출 알고리즘은 실제 복원되기 원하는 영역을 정확하게 검출하기 위해 약간의 제한을 둔다. 집합 H_i 에 속하는 각 원소에 대해, 탐색 범위 내에 건너 뛰어졌던 인접한 객체의 영역이 존재하지 않으면 해당 원소를 제거한다. 합성 문제에 대한 영향을 이미 부호기에서 최소화하였기 때문에, 탐색 영역의 범위를 크게 할 필요는 없다. 본 논문에서는 실험적으로 구해진 8을 탐색 영역으로 정하였다.

V. 실험 결과 및 분석

본 논문에서 제안된 알고리즘의 성능을 검사하기 위해 배경과 전경 2개의 객체로 화면을 구성하는 CIF 형식의 FOREMAN 시퀀스와 배경을 고려하지 않고 전경에 해당하는 2개의 객체들로 화면을 이루는 COASTGUARD 시퀀스를 이용하였다.

그림 6과 그림 7은 배경 객체를 포함하는 화면 합성과 포함하지 않는 화면 합성의 경우에 제안된 알고리즘을 이용하여 합성 문제를 해결한 결과를 보여준다. 그림 6(a)와 그림 7(a)는 실제 합성 문제가 일어난 현상을 보여주고, 그림 6(b)와 그림 7(b)는 식 (2)와 식 (3)에 의해 검출된 영역을 나타낸 것이다. 배경을 포함하는 경우에 원하는 영역을 정확하게 검출하지만, 객체를 포함하지 않는 경우에는 에러가 발생하는 것을 알 수 있다.

그림 7(c)는 탐색 영역을 고려하여 합성 문제가 일어난 영역을 검출한 영상이다. 그림 7(b)에 비해 합성 문제 영역을 올바르게 검출한다. 그림 6(c)와 그림 7(d)는 배경 영역에 해당하는 부분에서 가장 최소의 거리를 가지는 위치의 화소값으로 복원한 후에 최종적으로 화면에 보여지는 영상을 보여준다.

VI. 결 론

본 논문에서는 다중 객체의 부호화를 지원하는 MPEG-4 비디오 알고리즘에서 각 객체의 부호화 성능을 높이기 위해 객체기반 부호화 과정을 수행하는 경우, 각 객체가 서로 다른 시간 해상도를 가짐으로써 발생할 수 있는 합성 문제에 대해 소개하고 이 문제를 해결하기 위한 방법을 제안하였다. 제안된 알고리즘은 부호기에서의 합성문제 최소화 알고리즘과 복호기에서의 합성문제 제거 알고리즘으로 구성된다. 합성문제 최소화 알고리즘은 부호기에서 객체의 모양 정보를 이용하여 부호화되는 객체의 선택에 영향을 주고, 합성문제 제거 알고리즘은 복호기에서 인간의 시각에 민감할 수 있는 합성문제 영역들을 제거한다. 프레임기반 부호화 과정에 비해 객체기반 부호화 과정을 수행함으로써 1-1.5 dB 정도의 부호화 성능이 향상되고 제안된 합성문제 해결 방법에 의해 복호기에서 복원된 영상에서 인간의 시각에 거슬리는 부분이 보이지 않는다.

감사의 글

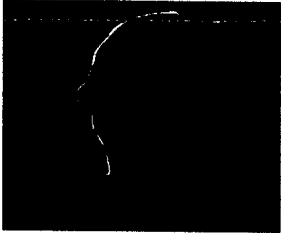
본 연구는 광주과학기술원 (K-JIST) 초고속광네트워크 연구센터 (UFON)를 통한 한국과학재단 우수 연구센터 (ERC)와 교육부 두뇌한국21 (BK21) 정보기술사업단의 지원에 의한 것입니다.

참 고 문 헌

- [1] ISO/IEC 14496-2, "Information technology - Coding of audio/visual objects," Part 2 : Visual, 1998.
- [2] A. Vetro, H. Sun and Y. Wang, "MPEG-4 rate control for multiple video objects," *IEEE Trans. Circuit and Systems for Video Technology*, vol. 9, pp. 186-199, Feb. 1999.
- [3] J.W. Lee, A.Vetro, Y.Wang and Y.S. Ho, "Object-based Rate Allocation with Spatio - Temporal Trade-offs," *Proceedings of SPIE VCIP*, pp. 374-384, Jan. 2002.
- [4] 이정우, 호요성, "시공간 왜곡을 이용한 MPEG-4 객체기반 비트율 제어 방법," *영상처리 및 이해에 관한 워크샵*, pp 57-62. Jan. 2002.
- [5] J.W. Lee, A. Vetro, Y. Wang and Y.S. Ho, "MPEG-4 video object-based rate allocation with variable temporal rates," *IEE Electronics Letters*, vol. 38, no. 19, pp. 1088-1090, Sept. 2002.



(a) 합성 문제가 일어난 영상



(b) 검출된 영역

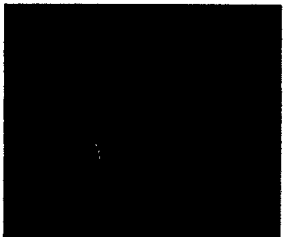


(c) 복원된 영상

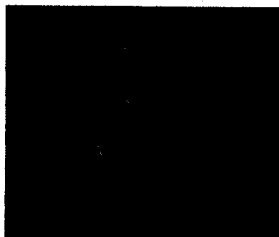
그림 6. 실험 결과 (배경이 있는 경우)



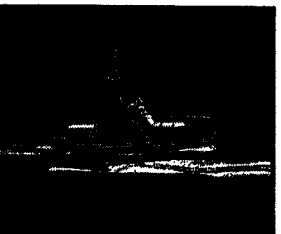
(a) 합성 문제가 일어난 영상



(b) 비제한 검출 영역



(c) 제한 검출 영역



(d) 복원된 영상

그림 7. 실험 결과 (배경이 없는 경우)