

크기, 움직임 및 왜곡 정보에 의한 목표비트 분배

한 학 수, 황 희 련, 황재정
군산대학교 전자정보공학부

Distribution of Target Bits based on Size, Motion and Distortion

Hak-Su Han, Hee-Ryung Hwang, Jae-Jeong Hwang
Dept. of Radiocommunication Eng., Kunsan National Univ.
{haksu, hhr, hwang}@ace.kunsan.ac.kr

Abstract

An efficient bit rate distribution technique that distributes available bits for multiple objects based on motion vector magnitude, size of object shape, and coding distortion is presented. This coding concept using the three parameters was exploited in MPEG-4 multiple object coding. But the scheme is likely to produce poor results such as allocating more bits to less important objects and degrading picture quality, due to the lack of analysis and research in view of human visual aspect. In this paper importance of each object is represented by the three parameters and visually analyzed. Target bits are distributed according to coding distortion using the pre-assigned shape and motion information.

I. 서론

지금까지 많은 논문에서 비트율 분배에 대한 문제가 다루어졌다. 세 변수(크기, 움직임, 왜곡률)를 이용하여 부호화하는 개념은 MPEG-4에서 다중객체를 부호화하는 데에 이용되었다. 그러나 변수들을 시각적인 관점에서 분석하여 이용하는 연구가 미흡하여 시각적으로 중요하지 않은 객체에 많은 비트가 할당되어 화질이 저하

되는 결과를 초래하였다. 본 논문은 MPEG-4 MVO 율제어를 기준으로 다중객체 부호기에 관련된 율제어 기법을 설명한다. 또한 다중객체의 객체별 중요성을 크기, 움직임, 왜곡으로 표현하여 전체 목표비트를 객체별로 분배하는 기법을 제한한다. 이를 토대로 다양한 테스트를 조건하에서 실행한 후 이전 복원 결과에 따라 목표치를 다중객체 전체에 대해 설정하여 적절히 분배한 결과를 보인다. 끝으로 실험을 통해 얻어진 결과를 비교 분석하여 향후 나아갈 방향을 제시한다.

II. 다중객체 부호기

그림 1은 다중객체 율제어 기법을 위한 블록도이다. 실제 율제어 기법은 사후부호화 단계에서 얻어진 QP를 부호화 엔진에 제공하는 역할을 한다. 목표치를 예측하여 n 개의 객체에 적절히 분배해주어야 하며 이때 스킵될 프레임이 얼마인지를 고려한다(사전프레임 스킵 제어). 여기서 중요한 문제는 단일객체와는 달리 영상정보가 부호화에서 고려되어야 한다. 임의의 형상과 텍스처 간의 비트율 분배 문제에서 비트율이 낮으면 형상을 위한 비트가 상대적으로 크게 되며 반대의 경우 낮게 된다. 따라서 다중객체 부호화에서 가장 중요한 요소로 대두된다.

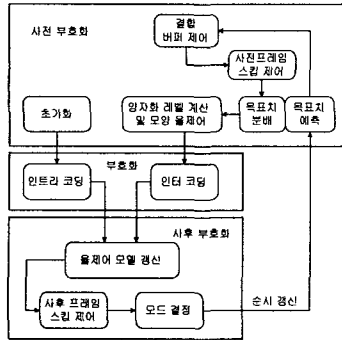


그림 1. 다중객체 부호화를 위한 블록도

다중의 객체는 의미적으로 상호 관련되어 있고 화질은 VO(video object)들의 종합된 화질로써 복잡한 함수가 된다. 객체들은 채널을 공유하므로 성능의 최적화는 전역 제어를 요구한다. 입력되는 n 개의 VOL(video object layer)은 각각의 특성과 중요도를 제시한다. 전체 시퀀스의 목표 비트율은 물리적인 채널용량에 따르며 주어진 비트율이 각 VOL에 배분된다.

1. 버퍼 초기화

첫 프레임을 부호화하는데 사용된 비트수 T_f 를 제외하고 시퀀스의 나머지 프레임들을 위해 가용한 비트수 b_r 을 결정한다.

$$b_r = t_s \cdot R_s - T_f \quad (1)$$

(t_s : 시퀀스의 주기를 초로 나타냄, R_s : 시퀀스의 원하는 비트율)

또한 하나의 VOP(video object plane)가 부호화된 후 버퍼로부터 제거되는 비트수는 다음과 같다.

$$R_{drain} = \frac{b_r}{n_r} \quad (2)$$

(n_r : I 프레임 후 부호화 되는 P 프레임 수)

이들 변수는 각각 MVO(multiple video object)마다 주어진 비트율에 따라 다르게 정의된다.

2. 초기 목표 비트율 예측

이전 프레임의 i -번째 객체에 사용된 비트율 $R_{p,i}$ 와 비례하여 i -번째 객체를 위한 비트율을 배정하는 방식으로 목표치가 설정된다.

$$T_i = \max \left\{ \frac{R_s}{M \cdot F_s}, (1 - w_p) \frac{T_r}{M \cdot N_r} + (w_p) R_{p,i} \right\} \quad (3)$$

(R_s : 시퀀스 비트율, F_s : 소스의 프레임율, T_r : 가용 비트수,

N_r : 부호화될 잔여 P 프레임 수, M : VO(video object) 수)

w_p 값의 증가는 개별 목표치를 $R_{p,i}$ 에 더욱 비례하도록 만든다. 실험에서는 $w_p = 0.25$ 가 사용되었다.

3. 결합 버퍼 제어

초기 목표치가 결정되면 전체비트 T_s 는 VO별 목표비트 T_i 의 합이다.

$$T_{joint} = T_s \cdot \frac{B_c + 2 \cdot (B_s - B_c)}{2 \cdot B_c + (B_s - B_c)} \quad (4)$$

(B_c : 버퍼레벨, B_s : 버퍼크기)

오버플로우와 언더플로우를 피하기 위해 최종 목표치는 다음과 같이 조정한다.

$$T = \begin{cases} (1 - \delta)B_s - B_c, & \text{if } B_c + T_{joint} > (1 - \delta)B_s \\ R_{drain} - B_c + \delta B_s, & \text{if } B_c - R_{drain} + T_{joint} < \delta B_s \\ T_{joint}, & \text{otherwise.} \end{cases} \quad (5)$$

대표적으로 안전마진 δ 값은 0.1이나 낮은 비트율에서 영상 정보를 부호화하기 위해 0.2~0.25로 증가될 수 있다.

4. 목표치 분배

목표 비트수 T_i , $i \in M$ 을 만족하도록 각각의 임의 영상 VO에 결합 버퍼제어의 출력목표치가 분배된다. 주어진 목표치에서 객체 i 의 목표치는 다음과 같다.

$$T_i = T \cdot (w_s SIZE_i + w_m MOT_i + w_v VAR_i) \quad (6)$$

정규화 이전의 i -번째 VO에 대한 변수는 각각 다음과 같이 정의된다.

$$\begin{aligned} MOTION[i] &= (abs(MV_x) + abs(MV_y)) / N_{MB}[i] \\ SIZE[i] &= N_{MB}[i] \\ MAD^2[i] &= (error_{sq}[i]) / N_{MB}[i] \end{aligned} \quad (7)$$

하중치 $\{w_s, w_m, w_v\} \in [0, 1]$ 은 $w_s + w_m + w_v = 1$ 을 만족한다. 그러나 영상의 크기는 배경 VO에서 가장 크게 되며 가장 많은 비트가 할당된다. 움직임에서도 만일 특정 객체가 계속 큰 움직임을 갖는 경우 계속 많은 비트가 할당되게 된다. 그러나 지속적으로 움직이는 물체의 부호화는 갑자기 움직임이 시작되거나 끝난 객체에 비해 시간적으로 덜 중요하다는 점을 간과하고 있다. 이 문제의 해결책을 다음 장에 논의한다.

III. 시각적 비트율 분배

인간의 공간주파수적 감도 특성은 그림 2와 같은 특성을 갖는다[9]. 즉 중간주파수(약 5내지 10cycles/degree)에서 감도가 가장 높게 된다. 형상에 대한 하중치는 식(8)에 대입하여 적응적으로 얻어진다. 기준 주파수(f)가 정해지면 그에 따라 관측거리를 선정하고 객체별 공간주파수가 계산된다.

$$H(f) = 2.6(0.0192 + 0.114f) \exp(-0.114f)^{1.1} \quad (8)$$

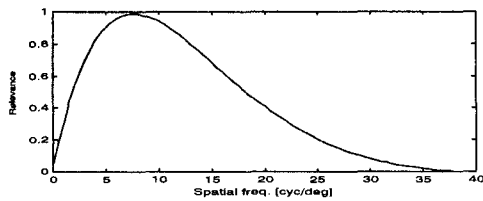


그림 2. 인간 시각시스템의 공간주파수 특성

움직임의 변화에 의한 감도특성 또한 공간주파수 특성과 마찬가지로 중간 정도의 움직임에서 높은 감도를 가진다. 왜곡량은 절대크기 자체를 이용한다. 마지막 단계로 과거 프레임과의 차이값(DM, DS, DV)을 얻어 이 차이값에 의한 목표비트의 재분배를 다음과 같이 행한다.

$$T_i = T_i \frac{DM_i + DS_i + DV_i}{\frac{1}{numVOs} \sum_j (DM_j + DS_j + DV_j)} \quad (9)$$

각 객체에 대한 전체 비트수가 결정되면 텍스처를 위한 가용 비트수는 움직임, 형상, 다른 부가정보를 제외하고 얻어진다. 양자화 변수는 앞에 언급한 바와 같이 각 객체에 대한 올-양자화 모델을 이용하여 얻는다.

IV. 실험 및 결과

올-왜곡 함수를 이차 율제어 모델로 근사화하여 다중 객체 부호화에 적용하였다. 실험에 사용된 시퀀스는 Akiyo(2: 정적인 배경, 앵커), News(4: 배경, 춤추는 댄서의 삽입된 비디오, 두 명의 앵커, 텍스트), Coastguard(4: 정적인 바닷물 배경, 왼쪽에서 진입하는 배, 화면 중간의 조그만 보트, 상단에 해안 배경)이다. "Coastguard" 영상에 대해 그림 3에 보인다.

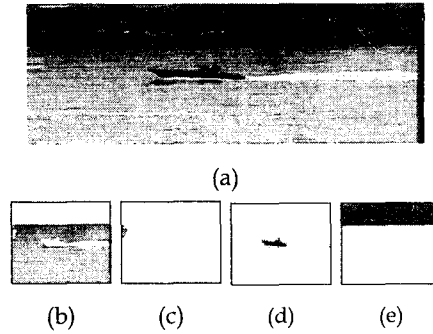


그림 3. 실험 Coastguard 영상과 네 개의 객체

다중객체를 효과적으로 부호화하기 위해 객체별 중요도를 평가하여 중요도에 따른 목표 비트를 설정하고 비트량을 분배하는 기법을 실험하였다. 실험에 사용된 시스템은 MPEG-4 VM8을 기초로 하여 객체별 중요도를 고려하였다. 네 개의 객체를 가진 Coastguard 영상의 객체별 크기, 움직임, 분산량은 그림 4, 그림 5, 그림 6과 같이 얻어진다.

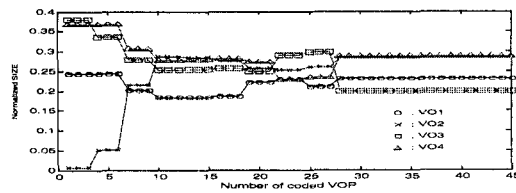


그림 4. Coastguard 영상의 객체별 크기 변화

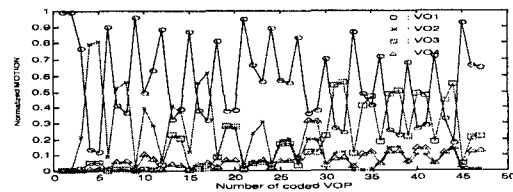


그림 5. 네 개의 객체를 가진 Coastguard 영상의 움직임 벡터크기(128kbps)

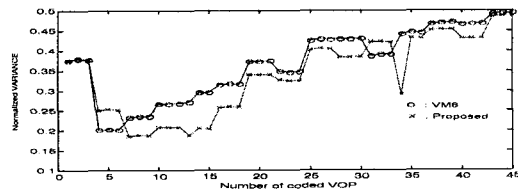


그림 6. 자승 MAD로 표현한 Coastguard 영상의 VO2 분산량

그림 7 a)에서 보면 MPEG-4 VM의 결과는 VO1에

지나치게 많은 비트를 할당하는 경우를 볼 수 있다. 왜냐하면 움직임이 존재하고 크기가 작지 않기 때문에 발생하는 현상이다. 그림 7 b)에서 제시하는 결과는 고르게 비트를 분배함을 보인다. 이처럼 객체의 중요성에 비추어 적절한 비트를 분배하는 데에는 크기, 움직임, 왜곡량이 프레임간 부호화 과정에서 절대량과 변화량을 이용해야 한다.

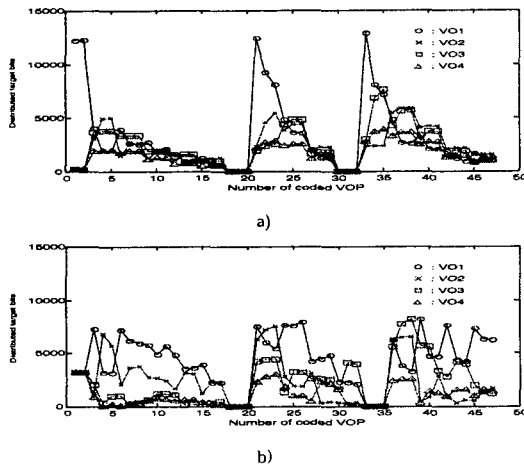


그림 7. 목표비트의 객체별 분배(128kbps의 Coastguard 영상): a) VM10, b) 제안된 기법

V. 결론

다양한 종류의 비트율과 공간 및 시간해상도에서 다중 객체를 효과적으로 부호화하는 데에 필요한 비트량 분배기법을 제안하였다. 이 분배 기법은 SVO 기법에 첨가하여 MVO로 확장하는 시스템에 이용될 수 있다. 사진 부호화 단에서 움직임 및 영상부호화 후에 가용 비트를 분석하지만 버퍼상태와 시공간해상도가 버퍼 안정을 위해 조정된다. 또한 시간해상도에 근거하여 동작모드를 선택함으로써 울제어를 적용시킨다. 서로 다른 부호화 조건에 적용토록 하며 공간 및 시간해상도간, 영상과 텍스처 왜곡간 트레이드오프를 얻도록 비트율을 분배하는 기법을 제안하였다. 객체간 중요성을 분석하여 VOP로 나뉘어지고 객체기반 모델링에 의해 비트가 할당된다. 객체간 목표치 분배는 움직임벡터 크기, 객체의 크기, 이전 프레임의 왜곡 등을 반영하는데 단순한 비교에 의한 분배보다 변수의 변화량을 시각적으로 중시하여 목표치를 분배하였다. 제안한 기법을 MPEG-4 VM10 부호기에 사용된 기법과 비교하였다.

현재 사용된 울-양자화기 모델이 텍스처 부호화에 꽤 적절하지만 영상에 효과적인 R-D 모델 개발이 아직도 요구된다. 그러한 모델은 영상의 기하학적 상태에 의존할 것이다. 개별 R-D 모델외에 영상 및 텍스처 왜곡에

대한 감각적 왜곡의 이해가 영상 및 텍스처 결합 울제어를 위해 요구된다. 또한 서로 다른 객체가 다른 프레임 울로 부호화되도록 하여 합성문제를 해결할 도구가 요구된다. 그러면 더 복잡한 버퍼링 기법을 필요로 하지만 가능성은 실로 크다.

참고문헌

- [1] M. R. Pickering and J. F. Arnold, "A perceptually efficient VBR rate control algorithm," *IEEE Trans. Image Processing*, vol. 3, pp. 527-532, Sept. 1994.
- [2] P. Fleury, J. Reichel, and T. Ebrahimi, "Image quality prediction for bit rate allocation," in *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, pp. 339-342, Oct. 1996.
- [3] X. Marichal, T. Delmot, C. DeVleeschouwer, V. Warscotte, and B. Macq, "Automatic detection of interest areas of an image or of a sequence of images," in *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, pp. 371-374, Oct. 1996.
- [4] J. B. Lee and A. Eleftheriadis, "Spatio-temporal model-assisted compatible coding for low and very low bit rate video-telephony," in *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, pp. 429-432, Oct. 1996.
- [5] L. Wang and A. Vincent, "Joint rate control for multi-program video coding," *IEEE Trans. Consumer Electron.*, vol. 42, pp. 300-305, Aug. 1996.
- [6] MPEG video group, "Text of ISO/IEC FDIS 14496-2," Doc. ISO/IEC JTC1/SC29/WG11 N2502, Atlantic City, NJ, Oct. 1998.
- [7] T. Sikora, "The MPEG-4 video standard verification model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 19-31, Feb. 1997.
- [8] A. Vetro and H. Sun, "Joint rate control for coding multiple video objects," in *Proc. IEEE Workshop Multimedia Signal Processing*, Princeton, NJ, pp. 181-186, June 1997.
- [9] J. L. Mannon and D. J. Sakrison, "The effects of a fidelity criterion on the encoding of images," *IEEE Trans. Inform. Theory*, vol. 20, pp. 525-536, July. 1974.