

정규논문 (Regular Paper)

방송공학회논문지 제17권 제4호, 2012년 7월 (JBE Vol. 17, No. 4, July 2012)

<http://dx.doi.org/10.5909/JBE.2012.17.4.670>

## 깊이정보를 이용한 관심영역의 화질 제어 방법

권순각<sup>a)†</sup>, 박유현<sup>a)</sup>

### Picture Quality Control Method for Region of Interest by Using Depth Information

Soon-kak Kwon<sup>a)†</sup> and Yoohyun Park<sup>a)</sup>

#### 요 약

정지영상 및 동영상내의 화면에서 관심영역이 설정되고, 비관심영역에 비하여 관심영역에 높은 화질이 제공되면 전체적으로 주관적 화질이 증대하게 된다. 관심영역은 색상 카메라만을 이용하여 추출하게 되면 계산이 복잡해질 뿐만아니라 추출의 정확도도 떨어지게 된다. 따라서 본 논문에서는 관심영역을 설정하기 위하여 깊이 카메라를 이용하며, 카메라와 관심영역 객체별 거리를 산출하고, 거리에 따라 객체의 화질을 차별적으로 제어하는 방법을 제안한다. 즉, 거리가 먼 객체에는 높은 양자화 계단크기를 적용하지만, 상대적으로 거리가 가까운 객체에는 낮은 양자화 계단크기를 적용하여 화질을 좋게 한다. 본 논문에서 제안된 방법에 의해 객체의 거리별로 양자화 계단의 크기를 차등적으로 적용함으로써 주관적 화질의 개선이 이루어졌으며 모의실험을 바탕으로 확인할 수 있다.

#### Abstract

If the region of interest (ROI) is set within the picture of image and video and the high quality is provided in ROI compared to Non ROI, then overall subjective picture quality can be increased. ROI extracted by the color camera only increases the calculation complexity and reduces the extraction accuracy. In this paper, we use depth camera to set the ROI and calculate the object distance from camera, then propose a method that the different picture quality is controlled by depending on the distance of an object. That is, we apply a high quantization step size to the far object, but relatively a low quantization step size to the close object, so better picture quality can be provided. Simulation results show that applying the differential quantization step size to the distance of objects by the proposed method can improve the subjective picture quality.

Keyword : Quantization step size, Region of interest, Depth information, Picture quality control

## 1. 서론

최근 스마트폰, 스마트TV 등의 다양한 컴퓨터 및 정보단말기의 보급으로 말미암아 정지영상 및 동영상 등의 멀티미디어 서비스가 활발히 진행되고 있다. 영상 신호는 다른 멀티미디어 신호에 비하여 데이터가 방대하기 때문에 전송

a) 동의대학교 컴퓨터소프트웨어공학과 (Dept. of Computer Software Engineering, Dongeui University)

† 교신저자 : 권순각 (Soon-kak Kwon)

E-mail: [skkwon@deu.ac.kr](mailto:skkwon@deu.ac.kr)

Tel: +82-51-890-1727 Fax: +82-51-890-2629

· 접수일(2012년5월14일), 수정일(2012년6월13일), 게재확정일(2012년6월13일)

및 기록을 위해서 압축 부호화가 필요하다. 압축 부호화는 제한된 부호화율 상황에서 최적의 화질을 제공하기 위해 다양한 부호화 기법이 적용되고 있다. 정지영상을 위해서는 ISO/IEC에서 JPEG, JPEG2000 부호화 표준을 제정하였으며, 동영상 압축표준으로는 ISO/IEC와 ITU-T 두 기관에서 MPEG-1, 2, 4와 H.261, H.262, H.263, H.264 표준을 제정하였다<sup>[1]</sup>. 또한 최근에는 차세대 비디오 부호화 표준으로 HEVC(High Efficiency Video Coding)에 대한 표준화 작업이 진행중에 있다<sup>[2-4]</sup>.

또한, 부호화 성능을 개선시키기 위하여, 2차원 공간상에서 관심영역을 설정하고, 비관심영역에 비하여 관심영역에 더 높은 화질을 제공함으로써, 관심영역의 주관적 화질을 증대시키는 방법에 대한 연구가 진행되고 있다.

관심영역을 설정하는 방법은 현재까지 다양한 방법들이 제시되었으며, 연속하는 화면간의 차이값으로부터 움직임 부분을 찾아 관심 영역으로 설정하는 방법<sup>[5]</sup>, 사람의 얼굴 부분과 배경 부분을 검출하여 얼굴부분을 관심 영역으로 설정하는 방법<sup>[6]</sup>, 화면간 차 신호와 피부색깔을 동시에 고려하여 관심영역으로 설정하는 방법<sup>[7]</sup> 등이 제안되어져 왔다. 또한, 화면내 객체추출이 어렵기 때문에 객체추출없이 영상의 움직임을 판단하여 중앙부분, 외곽부분, 상위 또는 하위부분 등으로 관심영역을 설정하는 방법<sup>[8,9]</sup>이 제안되었다. 그러나, 영상 화소값을 획득하는 색상 카메라만으로는 다수의 객체가 존재할 경우에 각각 객체의 윤곽선을 정확하게 검출하고, 각 객체에 관심도의 우선순위를 다르게 부여하여 화질을 제어하기는 어려울 수 밖에 없다. 이를 해결하기 위하여 깊이 카메라와 색상 카메라를 동시에 이용하여 관심영역을 설정하는 방법<sup>[10]</sup> 등이 제안되어져 왔다. 깊이 카메라로부터 얻은 거리정보와 색상 카메라로부터 얻은 영상의 텍스처 및 움직임 정보 값을 이용하여 객체를 추출하고 관심영역으로 설정하기 때문에 관심영역 설정이 더 정확해질 수 있다. 그러나, 거리정보를 관심영역 추출에만 적용하고, 화질제어에는 적용하지 않기 때문에 다수의 객체가 존재할 경우, 객체간 화질제어에 어려움이 발생하게 된다.

따라서, 본 논문은 깊이(Depth) 및 색상(Color) 카메라를 동시에 이용하여 관심영역을 추출하고, 카메라와 목표 객체 사이의 거리와 화질제어를 위한 양자화 계단크기의 관

계식을 제시하여 관심영역에 효율적으로 화질 제어를 하기 위한 방법을 제안한다. 즉, 카메라로부터 멀리 있는 객체에는 상대적으로 높은 양자화 계단크기를 적용하지만, 가까이 있는 객체에는 상대적으로 작은 양자화 계단크기를 적용하고 높은 화질을 제공하게 해준다.

본 논문의 구성은 2장에서는 깊이 카메라로부터 카메라와 객체 간의 거리를 측정하는 방법에 대해 설명하고, 3장에서는 객체간 거리 정보로부터 양자화 계단크기를 달리 적용하는 제안한 방법에 대해 살펴본다. 4장에서는 모의실험 및 결과를 분석하며 5장에서 결론에 대해 기술한다.

## II. 깊이 카메라를 이용한 거리 측정 및 객체 라벨링

깊이 카메라는 비행시간(TOF: Time-of-Flight) 방식으로 작동하는 적외선 펄스 레이저로 구현되고 있다. 적외선 펄스 레이저는 적외선 레이저 다이오드를 포함하며, 다이오드에서 방출된 펄스광은 공간으로 조사되고 피사체의 표면에서 산란된 후에 적외선 펄스 레이저의 광검출기에서 검출할 때까지의 시간이 측정된다. 이 시간을 측정함으로써 객체와의 거리를 측정할 수 있다. 측정된 피사체의 거리 정보는 화면으로 구성되며, 화면은 각 화소(pixel) 별로 구성된다.

화면내 같은 객체는 비슷한 거리에 존재하기 때문에, 수십cm 이내 거리에 있는 화소들을 그룹화하고 라벨링하는 과정을 수행한다. 라벨링된 화면은 배경영역에는 '0', 첫 번째 객체영역에는 '1', 두 번째 객체영역에는 '2' 등으로 객체 수만큼의 크기 값으로 화소값이 저장된다.

본 논문에서는 깊이 카메라로 Microsoft사에서 개발한 Kinect<sup>[11]</sup>를 사용한다. Microsoft사는 PrimeSense사의 움직임 센서 도입, GestureTek사의 특허 라이선스 사용, 3DV Systems사와 Canesta사의 기업 인수 등을 통하여 게임분야에 적용하기 위한 Kinect를 개발하였다. Kinect 기술은 레이저를 이용해 사용자에 적외선을 투사하고 나서 적외선 점(dot)들의 반사파 강도를 측정함으로써 이미지들을 나눈다. 반사 강도가 약한 점들은 멀리서부터 반사된 것으로 간

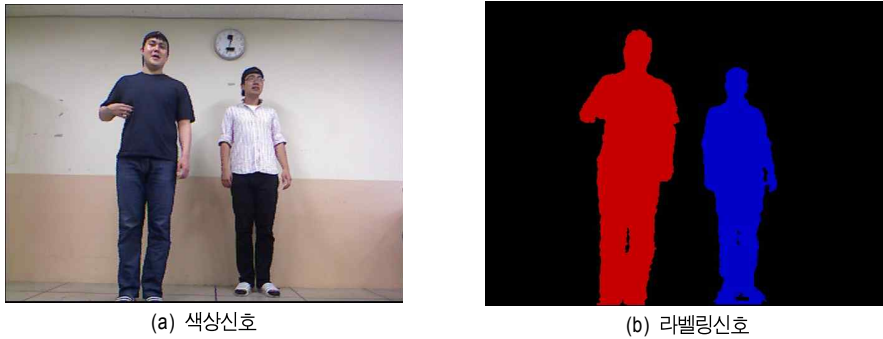


그림 1. Kinect로부터 얻어진 색상신호와 라벨링신호  
 Fig. 1. Color and labelling signals obtained from Kinect

주되며 보다 강도가 센 점들은 전면에 있는 사용자로부터 온 것으로 추정한다. 그림 1은 PrimeSense사의 오픈 라이브러리인 OpenNI<sup>[12]</sup>를 이용하여 색상 영상과 거리정보로부터 객체정보를 얻어 라벨링해준 화면을 나타낸다.

### III. 거리정보를 이용한 화질제어 방법

동영상에서 한 화면 또는 정지영상의 화면에는 다수의 객체가 존재한다. 객체들은 카메라로부터 서로 다른 거리상에 존재한다. 앞에서 살펴본 것처럼 깊이 카메라는 객체마다 거리 정보를 산출해준다. 일반적으로 배경부분은 가장 먼 거리에 존재하고, 관심의 대상이 되는 객체는 비교적 가까운 거리에 존재한다. 따라서, 카메라로부터 객체간의 거리정보를 이용하여 화질을 차등적으로 제어해줄 필요가 있다.

객체의 거리 정보로부터 화질을 차별적으로 제어하는 제안방법에 대한 순서도는 그림 2에 나타낸다.



그림 2. 제안된 방법의 순서도  
 Fig. 2. Flowchart of proposed method

거리 영상 산출부에서는 카메라에서부터 각각의 객체까지의 거리 정보를 산출하여 특정 해상도를 갖는 거리 영상을 얻는다. 거리 영상의 해상도는 깊이 카메라에서 제공되는 수평 및 수직의 공간적 해상도, 즉 적외선 센서의 정확도에 의존한다. 객체 추출부에서는 거리 영상에 근거하여 비슷한 거리에 있는 화소들을 그룹화하여 동일 객체로 라벨링하는 과정이 수행된다. 일반적으로 기준거리에서 수십 cm 또는 수m 이내에 존재하는 화소들은 같은 거리에 존재하는 화소들로 근사화된다. 화질제어부에서는 객체별로 근사화된 거리 정보를 이용하여 거리에 따라 객체별로 화질을 차이나게 제어한다. 화질제어는 다양한 방법으로 사용될 수 있으며, 본 논문에서는 양자화 계단크기 방법을 적용한다.

영상부호화 기법에서 화질이 양자화 계단크기를 통하여 제어될 경우에는 정지영상 및 동영상내 화면이 수평방향x 수직방향으로 8x8 또는 4x4 블록단위로 처리되며, 하나의 블록에 대하여 동일한 값의 양자화 계단크기가 적용된다. 양자화 계단크기가 크면 손실이 발생하여 화질이 떨어지고, 양자화 계단크기가 작아지면 손실이 줄어들어 화질이 좋아지게 된다. 이러한 양자화 계단크기는 제한된 전체 부호화율 조건을 만족시키는 범위안에서 최적의 화질이 제공되도록 제어되어야 한다.

그림 3은 객체내의 각 블록의 거리에 따라 양자화 계단크기를 다르게 적용하는 경우를 나타낸다. 양자화 계단크기 값은 거리에 따른 관계식으로 정해진다. 거리에 따라 양자화 계단크기가 선형으로 증가되게 할 수 있다. 선형은 거리

에 따라 정비례로 증가하게 제어된다.

#### IV. 모의실험 및 결과

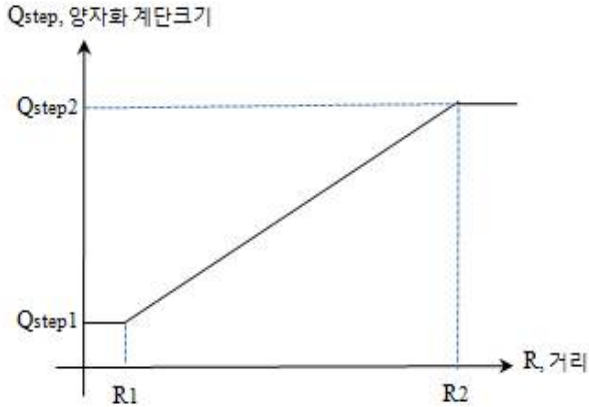


그림 3. 거리에 따른 양자화 계단크기의 설정  
 Fig. 3. Assigning quantization step-size according to distance

$$Q_{step}(R) = \alpha R + \beta, \quad R_1 \leq R \leq R_2 \quad (1)$$

여기서  $Q_{step}$ 은 양자화 계단크기 값이며,  $R$ 은 카메라로부터 객체간 거리 값이다. 또한,  $\alpha, \beta$ 는 상수 값이다. 특정 거리  $R_1$ 보다 작고,  $R_2$ 보다 큰 경우에는 거리에 따라 양자화 계단크기를 변화하지 않고 고정시킨다. 즉,  $Q_{step1}$ 은 부호화기에서 제공되는 양자화 계단크기의 최소값 이상이 되어야 하고,  $Q_{step2}$ 는 양자화 계단크기의 최대값 이하가 되어야 한다.

모의실험을 위해서 깊이카메라와 라벨링은 2장에서 설명한 Kinect와 PrimeSense사의 OpenNI가 사용되었고, 영상부호화기는 H.264 JM10.2<sup>[13]</sup>이 사용되었다. 영상내 객체는 그림1과 같이 2개이며, 카메라로부터 객체1, 객체2, 배경은 서로 다른 거리에 존재한다. 객체1은 객체2에 비하여 약1m, 객체2는 배경에 비하여 약1m 더 가깝게 위치한다. 영상의 각 화면은 수평방향 및 수직방향으로 각각 640, 480 화소로 구성되었고, 초당 30화면으로 총 90 화면이 사용되었다.

그림 4는 본 논문에서 사용하고 있는 색상영상을 고정 길이 부호화율로 부호화하고, 각 경우에 대한 원영상과 복원 영상의 PSNR을 표시한 것이다. 부호화 조건으로는 H.264 Main Profile이며, GOP내에 15개 화면이 있고, 이웃한 P-화면사이에는 2개의 B-화면이 포함된다.

제안된 방법은 객체마다 양자화 계단크기를 차별적으로 제어하는 것으로서, 식(1)의  $\alpha$ 에 의해서 조절이 가능하다. 매크로블록마다 양자화 계단크기가 조절되며,  $\alpha = 4$ 로 설정되면, 1m거리 마다 H.264 양자화 변수가 4씩 차이나도록 설정되는데, 객체1은 객체2에 비하여 양자화 변수가 4 작게, 객체2는 배경에 비하여 4 작도록 제어하는 것을 의미한다. 즉, 부호화율 제어를 단순화하기 위하여 H.264 JM에 근거하여 목표 양자화 변수가 결정될 때, 객체2는 그 값을 그대로 적용하고, 객체1은 그 값보다 4 작게 할당하며 배경

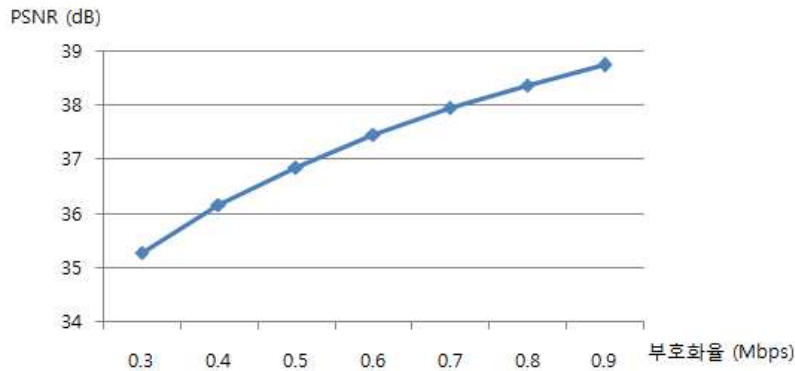


그림 4. 색상영상의 부호화율과 PSNR 관계  
 Fig. 4. Relationship of bitrate and PSNR for color video

은 4 크게 할당하게 한다.

제안된 방법의 성능은 객관적 비교척도인 PSNR보다는 객체마다 주관적 화질의 차이를 측정하는 것이 바람직하므로, 주관적 화질 측정의 비교척도를 사용하였다. 기존방법 및 제안방법은 모두 0.4Mbps로 부호화율이 설정되었고, 제안방법은 다양한  $\alpha$  값에 따라 주관적 화질을 측정하였다. 주관적 화질의 측정 방법은 20명의 평가자에게 부호화 방법에 대한 사전예고 없이 복원된 화면을 무작위로 배치하고 주관적 화질이 우수한 순서로 점수를 부여하게 하였다. 가장 우수한 화면은 5점, 그다음으로 우수한 화면은 4점, 이런 순서로 하여 가장 화질이 좋지 않는 화면은 1점을 부여하도록 하였다. 표 1은 20명의 평가자가 부여한 점수를 평균한 것을 나타낸다.

표 1. 주관적 화질 측정 결과  
Table 1. Investigation result of subjective picture quality

기존방법	제안방법, $\alpha$					
	1	2	3	4	5	6
3.50	3.55	3.60	3.70	3.75	3.85	3.80

제안방법은 기존방법에 비하여 주관적 화질이 더 좋음을 알 수 있다.  $\alpha$ 가 작은 경우, 즉, 객체간 양자화 계단크기 차이가 크지 않는 경우에는 주관적 화질의 개선이 많지 않지만, 객체간 양자화 계단크기가 4이상 증가하면, 제안방법에 의한 주관적 화질이 기존방법에 비하여 상당히 개선됨을 알 수 있다.

## V. 결론

본 논문은 깊이카메라와 색상카메라를 동시에 이용하여 관심 객체의 주관적 화질을 개선시키기 위한 방법을 제안하였다. 카메라에 가까이 있는 객체는 배경 또는 멀리 떨어진 객체보다 상대적으로 관심이 높기 때문에, 깊이카메라로부터 카메라와 객체간 거리정보를 얻고, 가까이 있는 객체는 배경 또는 멀리 떨어진 객체보다 양자화 계단크기를 상대적으로 작게 할당하도록 해줌으로서 주관적 화질을 좋

게 해준다.

깊이카메라로 Kinect를 사용하고, 영상부호화 기법으로 H.264를 사용하여 거리정보에 따라 관심 객체의 화질을 차별적으로 제어하는 방법을 구현하였으며, 다양한 깊이카메라 및 영상부호화 기법에 연계되어 확장 응용가능할 것으로 기대된다.

## 참 고 문 헌

- [1] S.-k. Kwon, A. Tamhankar, K. R. Rao, "Overview of H.264/MPEG-4 Part 10", *Journal of Visual Communications and Image Representation*, Vol. 17, No. 2, pp. 186-216, 2006.
- [2] K. Ugur, K. Andersson, A. Fuldseth, G. Bjøntegaard, L. P. Endresen, J. Lainema, A. Hallapuro, J. Ridge, D. Rusanovskyy, C. Zhang, A. Norkin, C. Priddle, T. Rusert, J. Samuelsson, R. Sjöberg, Z. Wu, "High Performance, Low Complexity Video Coding and the Emerging HEVC Standard", *IEEE Trans. Circuit Syst. Video Technology*, Vol. 20, No. 12, pp.1688-1697, 2010.
- [3] W.-J. Han, J. Min, I.-K. Kim, E. Alshina, A. Alshin, T. Lee, J. Chen, V. Seregin, S. Lee, Y. M. Hong, M.-S. Cheon, N. Shlyakhov, K. McCann, T. Davies, J.-H. Park, "Improved Video Compression Efficiency through Flexible Unit Representation and Corresponding Extension of Coding Tools", *IEEE Trans. Circuit Syst. Video Technology*, Vol. 20, No. 12, pp.1709-1720, 2010.
- [4] F. Bossen, V. Drugeon, E. Francois, J. Jung, S. Kanumuri, M. Narroschke, H. Sasai, J. Sole, Y. Suzuki, T. K. Tan, T. Wedi, S. Wittmann, P. Yin, Y. Zheng, "Video Coding using a Simplified Block Structure and Advanced Coding Techniques", *IEEE Trans. Circuit Syst. Video Technology*, Vol. 20, No. 12, pp.1667-1675, 2010.
- [5] H. Song, C.-C. J. Kuo, "A Region-Based H.263+ Codec and its Rate Control for Low VBR Video", *IEEE Trans. Multimedia*, Vol. 6, No. 3, pp.489-500, 2004.
- [6] L. Tong, K. R. Rao, "Region of Interest Based H.263 Compatible Codec and its Rate Control for Low Bit Rate Video Conferencing", *International Symposium on Intelligent Signal Processing and Communication Systems*, pp.249-252, Dec. 2005.
- [7] Y. Liu, Z. G. Li, and Y. C. Soh, "Region-of-Interest Based Resource Allocation for Conversational Video Communication of H.264/AVC", *IEEE Trans. Circuit Syst. Video Technology*, Vol. 18, No. 1, pp. 134-139, 2008.
- [8] M.-C. Chi, C.-H. Yeh, and M.-J. Chen, "Robust Region-of-Interest Determination Based on User Attention Model through Visual Rhythm Analysis", *IEEE Trans. Circuit Syst. Video Technology*, Vol. 19, No. 7, pp. 1025-1038, 2009.
- [9] S.-k. Kwon, A. Punchihewa, D. G. Bailey, S.-W. Kim, J. Lee, "Adaptive Simplification of Prediction Modes for H.264 Intra-Picture Coding", *IEEE Trans. Broadcasting*, Vol. 58, No. 1, pp. 125-129, 2012.

[10] Y. Zhang, G. Jiang, M. Yu, Y. Yang, Z. Peng, and K. Chen, "Depth Perceptual Region-of-Interest Based Multiview Video Coding", Journal of Visual Communication and Image Representation, Vol. 21, No. 5-6, pp. 498-512, 2010.

[11] <http://www.microsoft.com/en-us/default.aspx>

[12] <http://openni.org>

[13] JM Reference Software Version 10.2, <http://iphome.hhi.de/suehring/tml/download>.

---

## 저 자 소 개

---



### 권 순 각

- 1990년 : 경북대학교 전자공학과 공학사
- 1992년 : KAIST 전기및전자공학과 공학석사
- 1998년 : KAIST 전기및전자공학과 공학박사
- 1997년 ~ 1998년 : 한국전자통신연구원 연구원
- 1998년 ~ 2001년 : 기술보증기금 팀장
- 2001년 ~ 현재 : 동의대학교 컴퓨터소프트웨어공학과 교수
- 2003년 ~ 2004년 : University of Texas at Arlington 교환교수
- 2010년 ~ 2011년 : Massey University 교환교수
- 주관심분야 : 멀티미디어신호처리, 영상통신



### 박 유 현

- 1996년 : 부산대학교 전자계산학과 이학사
- 1998년 : 부산대학교 전자계산학과 이학석사
- 2008년 : 부산대학교 전자계산학과 이학박사
- 2000년 : 한국국방연구원(KIDA) 연구원
- 2001년 ~ 2009년 : 한국전자통신연구원(ETRI) 선임연구원
- 2009년 ~ 현재 : 동의대학교 컴퓨터소프트웨어공학과 조교수
- 주관심분야 : 클라우드 컴퓨팅, 모바일 시스템, IT융합 시스템