DBpia

논문 99-24-11B-18 한국통신학회논문지 '99-11 Vol.24 No.11B

애니메이션 동영상을 위한 새로운 부호화 기법

정회원 민 병 석*, 정 제 창*, 최 병 욱*

A novel codec for animation moving picture

Byongseok Min*, Jechang Jeong*, Byunguk Choi* ***Regular Members***

요 약

본 논문에서는 애니메이션(animation) 동영상을 효율적으로 압축 부호화할 수 있는 부호화기를 제안한다. 현재 동영상을 압축부호화 하기 위한 소프트웨어 부호화기로는 Cinepak, MPEG-1, Indeo 5.0이 사용되고 있으나 자연 영상(natural image)을 대상으로 개발된 것들이다. 따라서, 애니메이션 동영상에 적용해서 고능률 압축시 색상 및 윤곽선의 열화가 매우 크게 발생해서 주관적 화질이 매우 좋지 않다는 단점을 갖고 있다. 일반적으로 애니메이션 동영상은 화소당 RGB 24비트 색상을 사용하는 자연 영상과 달리 8비트 즉 256색상 이하의 색상으로 칼라 양자화를 수행하더라도 원 영상과 거의 구별할 수 없는 특징을 갖고 있다. 장면내에서 객체들의 영역들은 단일 색상으로 이루어져 있어 공간내에서의 이웃한 화소의 용장성이 매우 높다. 또한, 객체의 윤곽선이 뚜렷해서 윤곽선이 크게 열화되지 않아야 주관적 화질이 우수해 진다. 그리고, 특정 장면에서 배경은 고정되어 있고 객체가 움직이는 경우가 많은데 움직임이 매우 커서 MPEG-1과 같이 움직임 정보를 부호화하기 위해서 블록단위로 처리할 경우 배경까지도 열화가 크게 발생한다. 국제 표준은 아니지만 애니메이션 동영상을 대상으로 하는 Autodesk사의 FLC 방법은 런길이(run-length) 방식을 기반으로 하기 때문에, MPEG-1 보다는 주관적 화질이 우수하지만 전역 팔레트를 사용하고 있어 색상의 변화로 인한 화질 열화가 발생한다.

본 논문에서는 이러한 애니메이션 동영상의 특징을 반영해서 화면내 부호화와 화면간 부호화로 구성된 새로운 부호화기를 제안한다. 화면내 부호화 방법으로써 칼라 양자화(color quantization), 적응 차분 부호화(ADPCM), 전방분류화(forward classification)와 Golomb-Rice 부호화를 제안하고, 화면간 부호화 방법으로써 움직임 특성에 따른 움직임 블록 정보 부호화 방식에 기반한 부호화기를 제안한다. 그리고, 기존에 널리 사용하는 방식인 MPEG-1, FLC, Intel Indeo 5.0등과 객관적 및 주관적 화질을 비교함으로써 성능을 평가한다. 평가 결과 100:1 정도의 압축률에서 주관적 화질 및 객관적 화질이 MPEG-1, Autodesk사의 FLC, Intel사의 Indeo 5.0 보다 우수한 결과를 얻었다.

ABSTRACT

In this paper, we propose a new efficient codec to compress animation video. Conventional software codecs to compress moving pictures are Cinepak, MPEG-1 and Indeo 5.0 but these are for the natural image. Therefore, when we compress animation moving pictures at low bit rate by those codec, the quantization artifacts, e.g., false color, contours and blocking artifacts, are occured so that the quality of images are very poor. Generally, the natural image requires 24 bits per pixel, which are red, green and blue components with 8 bits per pixel respectively. Though animation are quantized in color below the rate of 8 bpp, having 256 colors, the subjective quality is approximately equal to original images. In a scene, the region of objects are made of one color so that the correlation between neighbor pixels is very high.

* 한양대학교 전자통신공학과

논문번호 : 99010-0111, 접수일자 : 1999년 1월 11일

※본 연구는 '99 한양대학교 교내 연구비의 지원에 의해 수행되었습니다.

www.dbpia.co.kr

Also, because the contours of object are thick, the quality is high unless those are too degraded. It is often that the background is fixed and the objects are moving in a scene. Its quality will be degraded when we adpot block-processing like MPEG-1 because the motion is very big. With FLC of Autodesk, not being International Standards, which uses run-length code, the subjective quality is good but color is degraded for using the global palette.

In this paper, considering the characteristics of animation we design a new codec which has intra frame coder and inter frame coder. The intra mode coding method consists of color quantization, ADPCM, forward classification and Golomb-Rice coder. The inter coding method is proposed, which is the block-based processing dependent on the characteristics of motion. Finally, we appreciate the proposed codec by the comparision of objective and subjective qualities with conventional codecs like MPEG-1, FLC and Intel Indeo 5.0.

Ⅰ. 서 론

현재, 개인용 컴퓨터(personal computer)에서 동영상을 압축 부호화 하기 위한 소프트웨어 코덱(codec)으로는 Cinepak, MPEG-1, Indeo 5.0 등이 사용되고 있으나, 자연 영상(natural scene)을 대상으로 개발된 것들이다[1,3,5,6]. 이러한 코덱은 애니메이션 동영상에 적용해서 고능률 압축시 색상 및 화질의 열화가 매우 크게 발생해서 주관적 화질이 매우 좋지 않다는 단점을 갖고 있다. 국제 표준은 아니지만 상용 소프트웨어 코덱으로 Autodesk사의 FLC는 애니메이션 저작 도구의 화일 형태로 사용되고 있다[5,6]. 애니메이션 저작시에는 매 화면 단위로 지역 팔레트를 수작업으로 지정하지만, 기존의 RGB 24비트 동영상 화일을 FLC로 변환할 때는 일반적으로 자동처리를 통해 256 색상의 전역 팔레트를 설정한다. 따라서, 동영상에서 장면(scene)이 많이 변화하는 경우, 색상의 열화가 크게 발생하고 움직임이 작은 프레임(frame)을 생략하기 때문에 움직임이 자연스럽지 못하다는 단점을 갖고 있다.

일반적으로 애니메이션 동영상은 화소당 RGB 24 비트 색상을 사용하는 자연 영상과 달리 8비트 즉 256색상 이하의 색상으로 칼라 양자화를 수행하더라도 원 영상과 거의 구별할 수 없는 특징을 갖고 있다. 화면내에서 객체들의 영역들은 단일 색상으로 이루어져 있어 공간내에서의 이웃한 화소의 용장성이 매우 높고, 객체의 윤곽선이 뚜렷해서 윤곽선이 크게 열화되지 않아야 주관적 화질이 우수해 진다. 움직임이 발생하면 많은 경우 지역적으로 발생하는데 특정 장면에서 배경은 고정되어 있고 객체의 일부만 움직인다. 예를 들면 주인공이 대화하는 경우 입 주변에서만 움직임이 발생하므로 앞뒤 프레임간의 시간적인 용장성이 매우 높다. 이러한 경우, 움직임이 발생한 특정 영역만을 부호화하면 압축 효율을 높일 수 있다. 반면, 움직임이 화면 전체에서 크게 발생하는 경우 MPEG-1과 같이 움직임 정보를 부호화하기 위해서 블록단위로 처리할 경우 배경까지도 열화가 크게 발생한다. 이러한 경우, 앞뒤 프레임간의 시간적인 용장성이 거의 없기 때문에, I 프레임 형태로 부호화하는 것이 증가하는 정보량에 비해 주관적 화질을 크게 개선할 수 있다.

본 논문에서는 이러한 애니메이션 동영상의 특징을 반영해서 새로운 방식의 부호화기를 제안하고자 한다. 제안하는 부호화 방식을 그림 1에 나타내었다. 제안하는 방법의 중요한 부분들은 칼라 양자화, 화면내 부호화 및 화면간 부호화, 엔트로피 부호화이다. 우선 RGB 24 비트 영상을 입력으로 색상 정보를 감축하기 위한 8진 트리에 기반한 칼라 양자화(color quantization)가 수행된다[10,11]. 칼라 양자화된 영상을 대상으로 앞 뒤 영상간에 움직임이 발생하는 정도에 따라 화면내 부호화 및 화면간 부호화 모드를 선택한다. 위에서 설명한 바와 같이 움직임이 크면 정보량에 비해 화질의 열화가 크므로 화면내 부호화를 수행하고, 움직임이 작으면 화면간 부호화를 수행한다. 화면내 부호화(intra frame coding)는 이웃한 화소간의 정보를 감축하기 위해서 MPEG-4의 적응 차분 부호화(ADPCM)[7]와 엔트로피 부호화의 효율을 높이기 위한 전방 분류화(forward classification) 그리고 최종 엔트로피 부호화로서 JPEG-LS에서 채택된 Golomb-Rice 부호화[8]로 구성된다. 화면간 부호화(inter frame coding)는 화면내 부호화에서 사용한 칼라 양자화된 영상을 대상으로, 움직임이 발생하는 특성에 따라 I 모드(intra mode), P1 모드(prediction 1 mode)와 P2 모드(prediction 2 mode)로 구분된다. I 모드에서는 움직임이 화면 전체에서 발생하는 경우로, 움직임을

www.dbpia.co.kr

추적해서 부호화하는 것보다 화면내 부호화하는 것이 압축효율과 처리시간 및 화질의 상보성에 비추어 볼 때 화면내 부호화가 효율적이다. P1 모드는 앞 뒤 영상이 동일해서 부호화하지 않고 동일 화면이라는 정보만 부호화하는 모드이고, P2 모드는 움직임이 지역적으로 발생하므로 부호화 효율을 높이기 위해 움직임이 발생한 영역만을 추출해서 부호화하는 모드이다.

본 논문의 구성으로, 2장에서 화면내 부호화를 제안하고 3장에서 화면간 부호화를 제안하고 4장에서 실험 결과에 대해 기존의 방법과 비교 및 고찰을 하고 5장에서 결론을 맺는다.

Ⅱ. 화면내(Intra Frame) 부호화

화면내 부호화는 한 화면내에서 이웃한 화소간의 용장성을 최소화시켜 정보량을 감축하는 방법이다. 기존의 방법으로 이산 여현 변환 또는 웨이블릿 변환을 기반으로 하는 주파수 영역에서의 부호화 방식이 높은 정보량 감축 결과를 나타내고 있으나, 공간 영역에서의 방식 보다 시스템의 구조가 복잡하다는 단점이 있다[1,2,9].

본 논문에서는 애니메이션 동영상을 부호화하기 위해서 공간 영역에서의 처리를 기반으로 하는 부호화 방식을 제안하고자 한다. 애니메이션 입력 영상은 화소당 24비트 색상정보와 320x240 화소의 크기를 갖는다. 화면내 부호화 측면에서 보면 애니메이션의 특징은 다음과 같다. 첫번째, 화소당 24비트(대략 1600만) 색상을 8비트(256) 색상 내지 5비트(32) 색상으로 감축해도 주관적 화질의 저하가 크지 않다. 두번째, 애니메이션 제작 과정상 배경과 객체가 분리되어 제작되는데, 객체들을 몇개의 영역들로 구성되며, 특정 영역내의 색상은 단일 색상으로 이루어져 있어 용장성이 매우 높다. 그리고, 객체들의 윤곽선이 뚜렷해서 윤곽선에서 작은 열화가 발생하더라도 전체적인 주관적 화질이 크게 떨어진다.

제안하는 방식은 칼라 양자화, 분류화, 엔트로피 부호화의 3단계를 갖는다. 그림 2는 제안하는 화면내 부호화기의 구조를 나타내고 있다. 그림 2에 나타나 있는 것처럼, 24비트 입력 영상에 대해 색상을 감소시키기 위해서 8진 트리를 이용한 칼라 양자화를 수행하고, 이웃한 화소들간의 용장성을 줄이기 위해 MPEG-4의 적응 차분 부호화를 수행한다. 이때 발생하는 차분 신호들의 통계적 분포의 특성을 엔트로피 부호화기에 적절한 분포로 변환하는 매핑을 거친 후, 전방 분류화를 통해 상위 분류(Significant class)와 하위 분류(Nonsignificant class), 그리고 분류의 위치를 나타내는 2진 정보인 맵 정보(Map)를 생성한다. 이 단계에서 상위 분류와 하위 분류는 그림 2에서 처럼 Golomb-Rice 부호화를 통해 엔트로피 부호화된다. 그리고, 맵 정보는 런 길이 부호화 후 Golomb-Rice 부호화된다.

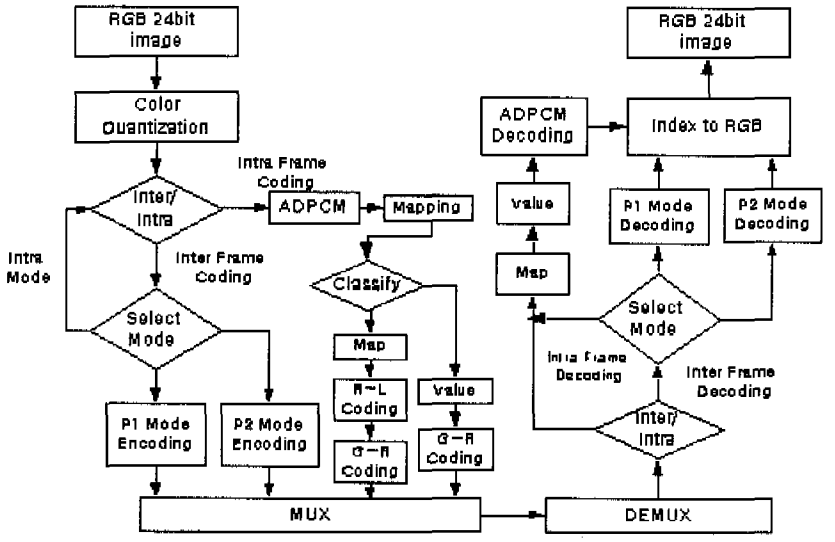


그림 1. 제안하는 부호화의 구조

www.dbpia.co.kr

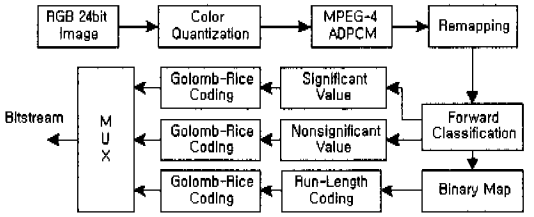


그림 2. 화면내 부호화

1. 8진 트리를 이용한 칼라 양자화

칼라 양자화는 컴퓨터 그래픽 및 영상처리에서 정보를 감소시키는 방법으로서 색공간에서 양자화하는 방법과 주파수 영역에서 양자화하는 방법으로 크게 구분된다. 주파수 영역에서 양자화하는 방법은 YUV 또는 YIQ 색공간에서 이산여현변환(Discrete Cosine Transform)을 기반으로 하는 MPEG-1, H.263, MPEG-2등이 있고 Wavelet 변환을 기반으로 하는 방법 등이 있다[4,8,9,13]. 이러한 방법에서는 주로 자연영상을 처리대상으로 하고 있으며 인간시각의 주파수 특성을 이용하고 있으므로, 색공간 처리에 비해 주파수 영역으로의 변환과 역변환을 위한 처리시간이 많이 필요하다.

색공간에서의 처리는 RGB, YUV, YIQ과 Lu^*v^* 등의 색좌표계 공간에서 색상의 통계적 성질에 따라, 장면 내에서 주요한 색상만으로 영상을 표현하는 방법이다. 이러한 방식으로는 메디안 컷(median cut), 시퀀셜 스칼라 양자화(sequential scalar quantization), Kohonen 신경망을 이용하는 방법과 8진 트리(octree)등이 있다[10,11]. 본 논문에서는 화질이 비교적 우수하며 처리속도가 빠른 8진 트리를 이용하였다.

(a)

(b)

(c)

(d)

그림 3. 색상수에 따른 영상
(a) 24비트(약 1600만색상) (b) 8비트(256색상)
(c) 6비트(64색상) (d) 5비트(32색상)

2. 적응 차분 부호화(ADPCM)

애니메이션 영상의 특성상 많은 경우 배경을 제외한 객체들은 수 개의 색상으로 구성되어 있고 단순한 영역들이 단일 색상으로 구성되어 있다. 또한, 객체와 영역을 나타내는 윤곽선이 수개의 화소 크기인 에지 형태로 뚜렷하다. 이러한 특성의 영상에

www.dbpia.co.kr

대해 이산 여현 변환과 같은 주파수 영역에서 고주파 성분을 크게 양자화하면 윤곽선이 손상되어 주관적 화질이 떨어진다. 오히려, 영상내에서 공간상의 상관관계가 매우 높기 때문에 차분 부호화를 적용하는 것이 부호화의 효율을 높일 수 있다.

본 논문에서는 차분 부호화 방법으로 애니메이션 영상에 적합하면서 복잡도가 낮은 MPEG-4에서 제안된 I-VOP의 직류성분(DC) 영상에 대한 예측 부호화를 적용하였다[7]. 여기에서 현재 화소의 값에 대한 예측을 수직성분과 수평성분의 신호로부터 현재의 화소값을 예측하므로 수평 또는 수직 화소만을 이용해서 1차원으로 처리하는 것보다는 복잡도가 증가하지만, 방향성분을 고려하고 있기 때문에 동일한 부호화율에서는 화질을 향상시킬 수 있다. 그리고, 예측 차수를 높이고 수평과 수직 화소에 가중치를 부가하는 예측 부호화도 가능하지만 애니메이션 특성상 1차 차수의 예측이 적절하였다. 식(1)에서 MPEG-4 적응 차분 부호화 과정을 나타내고 있다. 그림 4에서 현재 화소의 위치 X와 예측에 사용되는 화소 A, B와 C를 나타내었다.

$$\begin{aligned} &\text{if } (|A-B|<|A-C|) \\ &\quad \hat{X}=C \\ &\text{else} \\ &\quad \hat{X}=A \\ &er = X-\hat{X} \end{aligned} \tag{1}$$

X의 예측 값 $\hat{X}$은 A, B와 C에서의 값들의 관계로부터 얻어진다. A에서의 값으로부터 B와 C에서의 값을 각각 빼어 그 차이가 큰 쪽으로 $\hat{X}$의 값을 예측한다. 예를 들면, |A-C| 가 |A-B| 보다 크면 $\hat{X}$는 C가 되고 부호화 대상인 오차 er는 X-C가 된다.

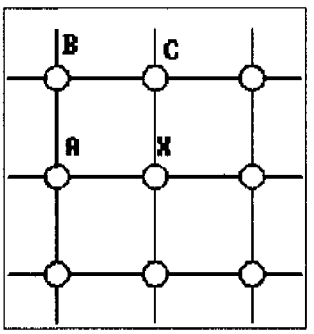


그림 4. MPEG-4 적응 예측

그림 5 (a)에서는 실험에 사용된 256 칼라 양자화된 애니메이션 영상 중에서 한 줄을 보이고 있으며 (b)에서는 MPEG-4 ADPCM을 수행한 결과를 보이고 있다. 가로축은 화소의 위치를 나타내며 세로축은 색인 값을 나타내고 있다. 애니메이션 영상의 특성상, 256 양자화를 수행했어도 유사한 색상이 이웃함을 알 수 있다. 색인의 차분 신호의 특성은 전체적으로 낮은 값을 갖고 색상의 경계에서만 큰 값을 갖는다. 따라서, 부호화 대상인 색인의 차분 신호는 라플라시안 분포를 나타낸다.

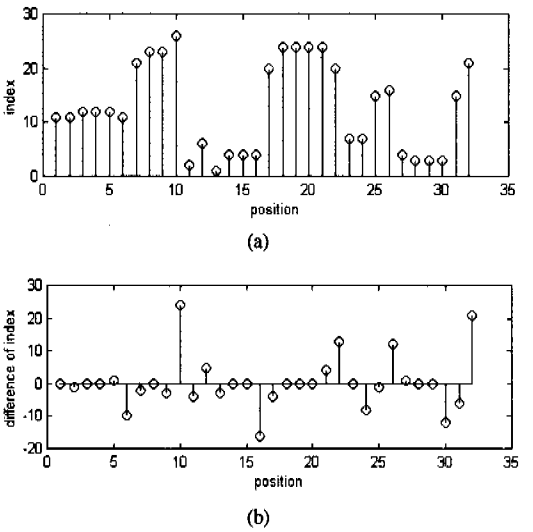


그림 5. MPEG-4 적응 차분 부호화
(a) 색인 영상 한 줄의 색인값
(b) (a)의 ADPCM

3. 차분신호의 매핑

색인 영상의 ADPCM한 신호는 음수와 양수 서포트(support)를 갖는 라플라시안 분포를 나타내므로, 양수 서포트를 갖는 지수 분포에 적합한 Golomb-Rice 부호화기를 적용하기 위해서 이 신호를 지수 분포에 근사하도록 매핑을 수행한다. 변환하는 방법을 식(2)에 나타내었다.

$$\begin{aligned} &\text{if } (index>0) \\ &\quad new_index = 2\times index+1 \\ &\text{else} \\ &\quad new_index = -2\times index \end{aligned} \tag{2}$$

식 (2)에서 색인의 차분신호인 index가 갖는 범위는 -255~255이고, 양의 index를 홀수로 매핑하고 음의 index를 짝수로 매핑하며 0의 index는 0으로 해서, new_index는 0~511의 범위를 갖으므로 index와 new_index의 엔트로피는 변함이 없으면서

www.dbpia.co.kr

분포의 특성만 변화한다.

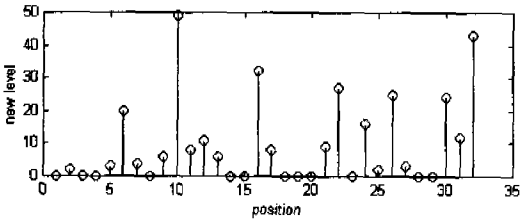


그림 6. 차분신호의 매핑

4. 전방 분류(Forward Classification)

압축 부호화에서 효율을 높이기 위한 방법으로, 엔트로피 부호화에 앞서 신호를 통계적 성질 또는 문맥에 따라 분류하는 전처리 과정이 적용되고 있다. 분류과정은 크게 전방 분류와 후방 분류(backward classification)로 구분된다. 후방 분류 방식은 과거 신호들로 부터 현재 신호의 분류화가 이루어지는데 분류의 정확도를 높이기 위해서는 부호화와 복호화 단계가 복잡하다. 반면에, 전방 분류는 신호의 통계적 특성에 따라 문턱치(threshold)를 적절하게 선택해서 신호들을 분류화하고 분류된 신호와 분류정보를 전송하는 방식이다.

본 논문에서는, 부호화 효율을 높이기 위하여 다음 절에서 설명하는 엔트로피 부호화기의 전처리 과정으로 단일 문턱치를 사용한 2원 분류화를 적용하였다. 그림 7에서 제안하고 있는 전방 2원 분류화 과정을 나타내었다. 입력 신호들에 대한 통계적 특성에 따라 적절한 문턱치를 구한다. 이것을 그림 7 (a)에서 나타내었다. 문턱치를 입력 신호에 적용함으로써 입력신호는 상위 분류와 하위 분류, 그리고 이진 맵정보(binary map image)를 발생한다.

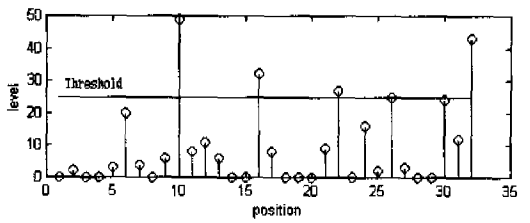


그림 7. 전방 분류화

입력 신호가 이산 무기억 표본(discrete memoryless sample)이면 분류화 과정에서 발생하는 맵정보는 신호의 엔트로피에 영향을 주지 않는다. 따라서, 입력 신호들의 정보가 2개의 분류 신호들과 맵정보의 정보량을 모두 합한 것과 같으므로 전체 정보량의 증가는 발생하지 않는다[12].

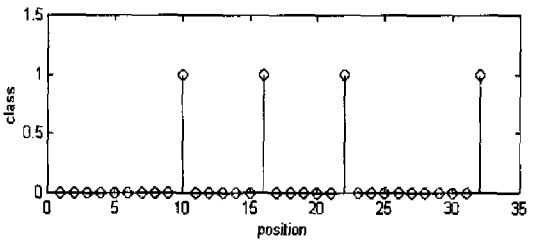


그림 8. 2원 분류 결과

5. Golomb-Rice 부호화

색인 영상에 대해 적응 차분 부호화, 매핑과 전방 분류화를 수행하고, 2원 분류 신호와 맵정보를 부호화하기 위해서 엔트로피 부호화 방법으로 Golomb-Rice 부호를 적용하였다. Golomb은 식(3)과 같은 기하적(geometreic)인 확률분포를 가지는 정보원에 대해 효율적인 부호화 방법을 제안하였다[8].

$$P(n) = (1-p_0)p_0^n \quad \text{where } n \geq 0,\ 0 < p_0 < 1 \qquad (3)$$

여기서 p_0는 상수이다. 양의 정수 m이 주어질 경우 Golomb 부호 G_m은 음이 아닌 정수 n을 다음과 같이 부호화 하게 된다. 첫 번째 단계로 n을 m으로 나눈 후, 몫을 q라고 하고 나머지를 r 로 표현한다(즉, n = mq + r). 다음 단계로 q를 1진(unary) 수로, r을 2진수로 바꾸어 나타내면 n에 대한 부호 G_m이 된다. 그리고 식(3)의 기하적인 확률분포에 대해 m이 식(4)과 같이 주어질 경우 부호 G_m은 최적이 된다.

$$m = \left\lceil -\frac{\log(1+p_0)}{\log(p_0)} \right\rceil \qquad (4)$$

Rice는 $m = 2^k$일 경우에 Golomb 부호화 방법으로 부터 매우 간단한 부호화와 복호화 과정을 제안하였다. 이 특별한 부호화 방법이 Golomb-Rice 부호화이다. 예를들어, k=2, n=22 일때, n값에 대한 Golomb-Rice 부호화 방법의 과정은 다음과 같다. 첫 번째 단계로 22(=n)를 2진부호로 표현하면 10110이 되고, 다음 단계로 10110을 2(=k) bit 만큼 자리이동을 하면 101=5가 된다. 그리고, 이것을

www.dbpia.co.kr

1진(unary) 부호로 다시 표현하면 000001이 되고, 자리이동한 2(=k)의 2진수 $(10)_2$ 과 다음과 같이 (000001)(10)로 연결하게 되면 22(=n)에 대한 Golomb-Rice 부호인 00000110를 얻을 수 있다. 여기에서 압축 효율에 영향을 주는 k값은 입력 신호의 특성에 따라 실험적으로 결정한다.

Golomb-Rice 부호의 장점은 부호화와 복호화 과정에서 사용되는 모든 연산이 계수기(counter)와 자리이동 연산에 의해 이루어지기 때문에 Huffman 부호화 방법이나 산술연산 부호화 방법에 비해 매우 낮은 복잡도를 갖는다. 그러므로, 실시간 복호를 필요로 하는 응용 분야에 적합하다.

Ⅲ. 화면간 부호화(Interframe Coding)

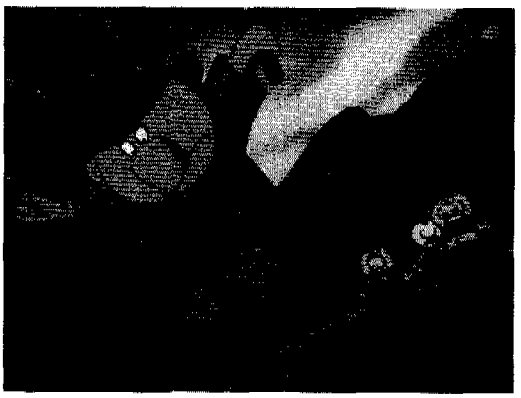
일반적으로 동영상은 이전 장면과 현재 장면간에 많은 용장성을 갖고 있으며, 이러한 정보를 줄이기 위한 방법으로는 MPEG-1, MPEG-2 등의 표준에서 채택하고 있는 움직임 추정 및 움직임 보상 등이 있다. 움직임 추정 및 보상에서는 화소 및 블록 단위로 이전 장면으로 부터 현재 화면과 가장 가까운 부분을 탐색해서 차이 성분만 전송하는 방식을 취하고 있다[1,2]. 이러한 방식은 시스템 구조가 복잡하고, 우수한 화질을 얻기 위해서 탐색시간이 많이 소비되고, 움직임이 크게 발생하는 경우 정보량이 증가해서 부호화 단계에서의 단점이 되고 있다[4,8].

본 논문에서는 애니메이션 동영상의 특징을 분석해서 새로운 방식의 화면간 부호화기를 제안하고자 한다. 화면간 부호화 관점에서 일반적인 영상과 다른 애니메이션 동영상의 특징은 다음과 같다. 첫번째, 움직임이 발생하는 장면간에서는 움직임이 매우 크다. MPEG-1과 MPEG-2에서는 초당 30화면을 재생하므로 움직임이 크게 발생하지 않지만, 애니메이션 동영상은 초당 12~15 화면을 재생하므로 움직임이 매우 크다. 두번째, 이전 화면과 현재화면이 동일한 경우가 많다. 즉, 움직임 정보가 발생하지 않는 경우가 많다. 세번째, 움직임이 발생하는 영역이 제한되어 있는 경우가 많다. 특정 주제 장면(scene)에서 배경은 고정되어 있고 인물이라던가 특정 객체만 비교적 움직임이 크게 발생한다. 그림 9에서 한 인물의 움직임이 특정영역에서 크게 발생하고 배경과 다른 인물의 움직임은 없다. 그림 10은 배경과 인물들의 움직임이 없는 경우이다.

(b)

그림 9. 움직임이 큰 경우
(a) 이전 화면 (b) 현재 화면

본 논문에서는, 이와 같이 자연영상과 다른 애니메이션 동영상 특징을 이용해서, 효율적인 장면간 부호화 방법을 제안한다. 장면간 부호화에서는 화면의 특성상 크게 3가지 모드로 분할해서 처리한다. 첫번째, 이전 화면과 현재 화면이 상관성이 거의 없는 경우는 I-프레임 모드로 부호화하고, 두번째, 이전



(a)

(a)

www.dbpia.co.kr

(b)

그림 10. 움직임이 없는 경우
(a) 이전 화면 (b) 현재 화면

화면과 현재 화면이 동일한 경우 P1-프레임 모드로 부호화하고, 세번째, 이전 화면과 현재 화면에서 움직임이 발생한 경우 P2-프레임 모드로 부호화함으로써 화면의 특성에 맞는 부호화 알고리듬을 적용한다.

I 프레임 모드

이 모드는 이전 화면과 현재 화면간의 시간적 용장성이 거의 없으므로 움직임 정보를 부호화하는 것보다 인트라 모드로 부호화하는 것이 효율적인 경우이다. 일반적으로 화면이 새로운 내용으로 전환되거나 움직임이 전체 화면에서 크게 발생하는 경우에 해당된다. 앞에서 기술한 인트라 모드에서의 부호화 방법을 적용한다.

P1 프레임 모드

애니메이션 동영상에서는, 현재 화면이 이전 화면과 동일한 경우가 많다. 이러한 경우, 이전 화면의 헤더만을 부호화함으로써 정보량을 크게 감축시킬 수 있다. 본 논문에서 제안한 부호화 방법에서는 이러한 화면이 많을수록 압축률은 향상된다.

P2 프레임 모드

화면간 부호화 모드에서 움직임이 발생하는 경우에 적용하는 모드로서, 이전 화면과 현재 화면간 차이 화면의 특성이 지역적으로 발생하는 경우에 적용한다. 애니메이션 동영상은 배경의 고정되어 있고 인물과 같은 객체들만 움직이는 경우가 많다. 이러한 경우, 움직임은 전체적으로 발생하는 것이 아니라 지역적으로 발생한다. 따라서, 움직임이 있는 영역의 정보만을 부호화해서 정보량을 감축시킬 수 있다. 움직임 정보의 정확도를 높이기 위해 채도, 색도와 명암의 색공간(IHS 색공간)으로 변환하고, 움직임이 발생한 영역을 효율적으로 탐색하기 위해서 8x8 화소 크기의 블록 단위로 처리한다. 블록의 크기를 크게 할 경우, 인물의 입 또는 눈 등에서 지역적으로 발생하는 작은 움직임으로 인한 정보량이 증가한다. 이러한 3개값에 의한 차이가 문턱치 보다 낮으면 움직임이 없음을 0으로 나타내고, 문턱치 보다 높으면 1로 나타내어서 움직임 발생 블록 맵을 구성한다. 움직임이 발생한 블록이 90%이상 되면 I 프레임 모드로 부호화를 수행한다. 그렇지 않은 경우, 블록의 색인 정보는 인트라 모드에서와 같이 ADPCM하고 매핑과 분류화를 거쳐, Golomb-Rice 부호화하고, 분류를 나타내는 분류 맵은 런길이와 Golomb-Rice 부호화를 적용한다. 다만, 움직임 발생 블록의 위치를 나타내는 맵은 정보량이 많지 않으므로 8비트씩 묶어 저장한다.

화면간 부호화의 구성을 다음 그림 11에 나타내었다.

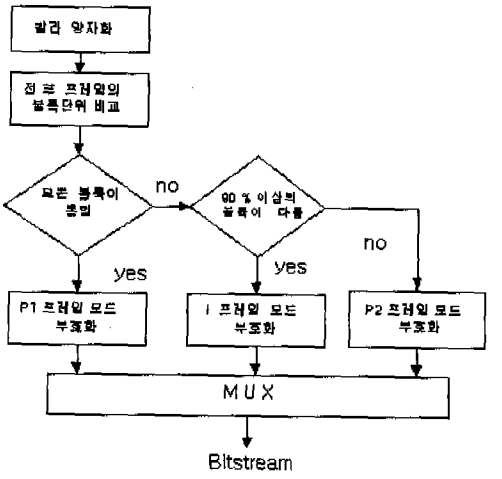


그림 11. 화면간 부호화

① 색인 영상을 8 × 8 화소 크기의 블록으로 전체 영상를 분할한다.

② 현재 영상에 대해 블록별로 이전 영상의 같은 위치에 있는 블록과 채도, 색도 및 명암도를 비교해서 이전 영상과 동일하면 0의 값을 지정하고 그렇지 않으면 1의 값을 부여해서 블록 맵을 구성한다. 이것을 식(5)에 나타내었다.

$$diff = \sum_{i=0}^{i=7}\sum_{j=0}^{j=7}(\Delta Hue_{m,n}(i,j)^2 + \Delta Saturation_{m,n}(i,j)^2 +$$

www.dbpia.co.kr

$$\Delta Graylevel_{m,n}(i,j)^2)$$

$$\begin{aligned} &\text{if } (\ diff \geq Threshold) \\ &\quad BlockMap(m,n) = 0 \\ &else \\ &\quad BlockMap(m,n) = 1 \end{aligned} \tag{5}$$

여기에서 각 항은 다음과 같다.

$$\Delta Hue_{m,n}(i,j) = Hue_{m,n}(i,j)_{current\ frame} - Hue_{m,n}(i,j)_{previos\ frame}$$

$$\Delta Saturation_{m,n}(i,j) = Saturation_{m,n}(i,j)_{current\ frame} - Saturation_{m,n}(i,j)_{previos\ frame}$$

$$\Delta Graylevel_{m,n}(i,j) = Graylevel_{m,n}(i,j)_{current\ frame} - Graylevel_{m,n}(i,j)_{previos\ frame}$$

③ 이전 화면과 비교하여 현재 화면이 모든 블록에서 전부 동일하면, P1 프레임 모드로 부호화를 수행한다.

④ 현재 화면에서 90%이상의 블록들이 이전 화면의 블록과 같지 않을 때는 I 프레임 모드로 부호화한다.

⑤ 현재 화면에서 이전 화면의 모든 블록이 동일하지 않고, 90%이하의 블록이 동일하지 않을 경우는 P2 프레임 모드로 부호화한다.

Ⅳ. 실험 결과 및 고찰

본 논문에서 애니메이션 동영상을 부호화의 성능을 평가하기 위해 전체적인 성능을 기존의 부호화 방법인 MPEG-1, FLC와 Indeo 5.0 등과 압축률, PSNR과 주관적 화질을 비교하고 각각의 특성을 고찰한다. 실험 대상은 애니메이션 영화 "콩딱콩 이야기"에서 추출한 동영상을 대상으로 하였다. 한장의 영상 크기는 320×240 화소이고 화면수는 164장이고 초당 15장으로 상영된다. 그리고, 영상은 R, G, B 각각의 색상이 8비트인 RGB 24비트 색상으로 구성되어 있다. 압축률 R은 식(6)와 같이 입력 동영상의 전체 비트로 부호화된 전체 비트를 나눈 것으로 정의한다.

$$R = \frac{the\ total\ number\ of\ encoded\ bits}{the\ total\ number\ of\ image\ sequence\ bits} \tag{6}$$

또한, 객관적 화질을 나타내는 기준으로 $PSNR$과 평균 $PSNR$은 식(7)와 같이 정의해서 평가하였다.

$$PSNR = 10\log_{10}\left(\frac{255\times255\times3}{\sigma_n^2}\right) \quad [dB]$$

$$\sigma_n^2 = \frac{\sum_i\sum_j (R(i,j)-\hat{R}(i,j))^2+(G(i,j)-\hat{G}(i,j))^2+(B(i,j)-\hat{B}(i,j))^2}{the\ number\ of\ pixels} \tag{7}$$

$$PSNR_{ave} = 1/N\sum_{i=1}^{N} PSNR_{i-th\ frame}$$

여기에서 각 항은 다음과 같다.

$R(i,j), G(i,j), B(i,j)$는 입력 영상의 (i,j)번째 화소의 R, G, B성분이다.

$\hat{R}(i,j)\ \hat{G}(i,j), \hat{B}(i,j)$는 복호된 영상의 (i,j)번째 화소의 R, G, B성분이다.

*the number of pixels*은 한 화면의 화소수이다.

N은 입력 영상의 화면수이다.

본 논문에서 제안한 방법으로 "콩딱콩 이야기" 동영상을 압축 부호화시 압축률은 다음과 아래 표 1과 같다.

표 1. 압축률과 평균 PSNR

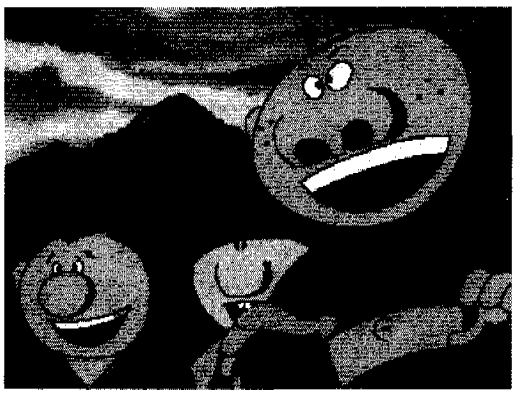
색상수	압축률	$PSNR_{ave}$
16색	111.14 : 1	25.21
32색	110.47 : 1	29.22
64색	94.60 : 1	32.12
128색	85.21 : 1	34.86
256색	69.50 : 1	37.12

표1에서 나타낸 것과 같이 색상수가 감소할 수록 전체적으로 압축률이 향상되고 있지만, 색상수와 압축률이 정비례하고 있지는 않은데 그 이유는 다음과 같다. 첫번째, 화면내 부호화에서 사용하는 Golomb-Rice 부호화 단계에서 입력 신호의 분포가 변화하기 때문이다. 둘째로, 화면간 부호화의 P2 모드에서 움직임이 발생하는 블록이 증가하면서 블록맵 정보가 증가하기 때문이다. 또한, I 모드와 P2의 Golomb-Rice 부호화 단계에서 입력 신호의 분포가 변화하기 때문이다.

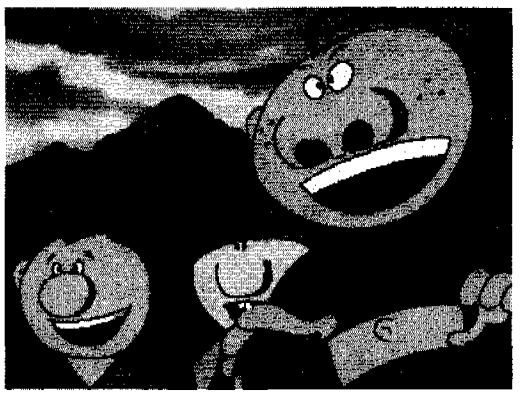
본 논문에서 애니메이션 동영상을 대상으로 제안하고 있는 방법과 기존의 소프트웨어로 구현된 부호화 방법을 비교하였다. 애니메이션 동영상을 MPEG-1으로 부호화했을 때, 발생하는 화질의 특성은 다음과 같다. MPEG-1의 I 프레임과 제안하는 방법에서 I 프레임 모드로 부호화한 결과를 그림 12에 나타낸 것과 같이 동일한 압축률에서 제안하는

www.dbpia.co.kr

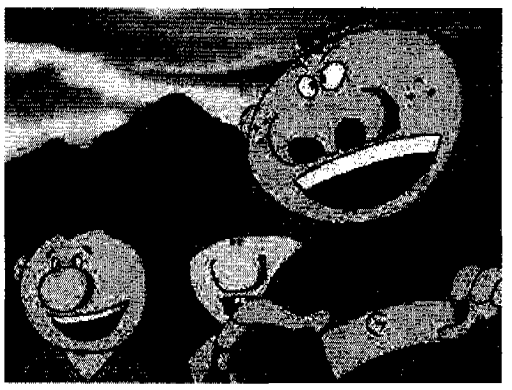
방법에 비해 MPEG-1의 경우 주관적 화질이 매우 떨어진다. 특히, 블록화 현상이 화면 전체에서 발생하고 있고, 애니메이션 동영상에서 주관적 화질에 크게 영향을 끼치는 윤곽선 및 에지부분이 많이 손상되고 있다. MPEG-1의 경우 이산 여현 변환을 기반으로 하는 주파수 영역에서의 양자화와 특히 색상 정보의 간축 기법을 사용하므로, 고능률로 압축 부호화시 색상이 원 영상과 많은 차이가 발생함을 알 수 있다.



(a)



(b)



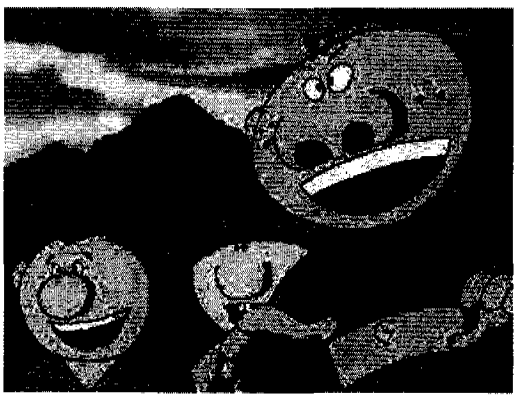
(c)

(d)

그림 12. 제안한 방법과 MPEG-1의 I 프레임 영상 비교
(a) 입력 영상
(b) 제안하는 방법(32색상,압축률 110:1)
(c) MPEG-1(압축률 110:1)
(d) MPEG-1(압축률 50:1)

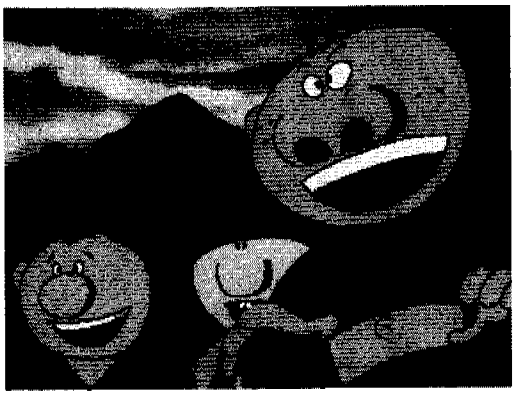
본 논문에서 제안하고 있는 방법에서는 MPEG-1의 P 프레임에 대응하는 것이 P2 모드로 부호화한 것에 해당된다. 그림 13에 MPEG-1의 P 프레임과 제안하는 방법의 P2모드 프레임을 비교하였다. MPEG-1 P 프레임의 경우, 주관적 화질에 영향을 주는 윤곽선과 에지가 크게 열화되어 있고, 움직임이 거의 발생하지 않았음에도 불구하고 움직임이 발생해서 나타나는 블록화 현상 등이 화면 전반적으로 나타나고 있다. 또한, 색상의 열화 뿐만 아니라 링 현상 등이 주관적 화질을 매우 저하시키고 있다. MPEG-1은 부호화과정에서 복호된 I 프레임으로 부터 P 프레임을 예측하기 때문에, I 프레임의 열화로 인한 오차 전파가 발생한다. 따라서, 고압축률로 부호화할 때, I 프레임의 화질이 좋지 않을 경우 P프레임의 화질은 더욱 저하된다.

국제 표준은 아니지만 상용으로 제작된 애니메이



(a)

www.dbpia.co.kr



(b)

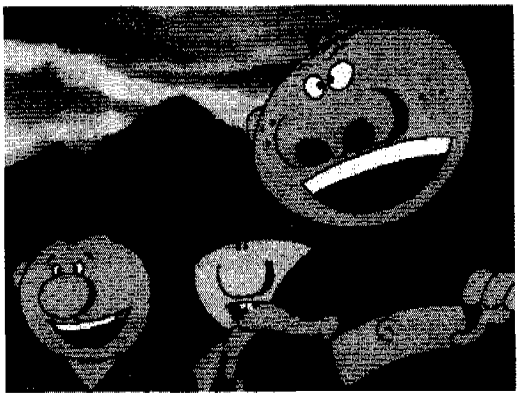
그림 13. MPEG-1의 P 프레임과 제안하는 방법의 P2 모드 프레임의 비교
(a) MPEG-1 P 프레임 (압축률 110:1)
(b) 제안하는 방법의 P2 모드 프레임 (32색상, 압축률 110:1)

션 형식으로 Autodesk사에서 개발한 FLC로 부호화한 결과와 비교한다. 이 방식의 특징은 영상 전체에 대한 256색상이하의 전역 팔레트를 사용하고, I 프레임 모드와 Skip 프레임 모드로 구성된다. I 프레임 모드에서 런길이의 변형된 word line coding방식으로 압축부호화를 한다. Skip 프레임은 전후영상을 비교해서 움직임이 작은 경우 해당 프레임을 삭제하고 앞화면을 반복 재생하도록 하고 있다. 본 논문에서 제안하고 있는 P1 모드와 유사하다. 그림 14 (a)에 FLC로 부호화한 결과를 나타내고 있다. 동영상 전체에서 256색상의 전역 팔레트만을 사용하고 있기 때문에 원래의 영상과 색상에서 많은 차이가 발생하였다. 본 논문의 방법과 비교했을 때, 움직임이 지역적으로 발생하거나 전체적으로 발생하거나 모든 프레임을 I 모드로 부호화하기 때문에 압축률이 저하된다.

"콩딱콩이야기"를 FLC으로 압축하였을 경우, Skip 프레임이 존재하므로 평균 PSNR을 구할 수 없어 PSNR만을 산출하였다. 전체 압축률은 55.07 : 1 이었으며 해당 프레임 영상의 PSNR은 25.08 dB가 산출되었다.

현재 개인용 컴퓨터에서 널리 사용되고 있는 동영상 소프트웨어 코덱인 Intel Indeo 5.0과 결과를 비교하였다. Intel Indeo 5.0의 구체적인 부호화 방식이 공개되어 있지 않아 전체 압축률과 PSNR만을 비교하였다. 35.24 : 1의 압축률에서 해당 프레임의 PSNR이 29.26으로 산출되었다. 본 논문에서 제안하고 있는 방법에서 32색상과 압축률 110:1 정도로 부호화한 것과 유사한 화질의 결과를 얻었다.

(a)



(b)

그림 14. FLC와 Indeo 5.0으로 부호화된 영상
(a) FLC로 부호화한 경우(압축률 55.07:1)
(b) Intel Indeo 5.0으로 부호화한 경우 (압축률 35.24:1)

V. 결 론

본 논문에서는 애니메이션 동영상의 특징을 반영해서 화면내 부호화와 화면간 부호화의 I 모드, P1 모드 및 P2 모드를 갖는 고능률의 새로운 압축방식을 제안하였다. 화면내 부호화에서는 색상수를 줄이기 위한 칼라 양자화와 화면내 용장성을 줄이기 위한 적응 차분 부호화 및 엔트로피 부호화를 수행해서 정보를 압축하고, 화면간 부호화에서는 애니메이션 동영상에서 발생하는 움직임의 특성에 따라 적절한 모드를 선택해서 영역별 적응 차분 부호화 및 엔트로피 부호화를 수행해서 정보를 압축 부호화하였다.

현재 널리 사용되고 있는 소프트웨어 코덱인 MPEG-1에서는 블록화 현상 및 영역별 색상의 열화가 현저했고 FLC 방식에서는 전역 팔레트를 사

www.dbpia.co.kr

용하므로 색상의 열화가 나타났고 작은 움직임을 갖는 프레임이 삭제되어 부자연스러운 움직임이 나타났다. 그리고, Intel Indeo 5.0은 MPEG-1과 FLC 방식 보다는 화질이 우수했지만, 실험결과를 비교해 보았을 때, 제안하고 있는 방법이 객관적 화질 및 주관적 화질이 우수함을 알 수 있었다. 향후, 요구되는 전송율 또는 압축률에 따른 비트율 제어(bitrate control)와 다해상도를 지원하는 스케일러빌리티(scalablity) 제어 등이 연구되어야 할 것으로 사료된다.

감사의 글

실험에 사용한 "콩딱콩 이야기" 동영상을 제공해 주신 (주)애니하우스에 감사드립니다.

참 고 문 헌

[1] 정제창 번역, 최신 MPEG, 교보문고, 1995
[2] 정제창 번역, 응용 MPEG, 교보문고, 1997
[3] Chwan-Hwa John Wu and J. David Irwin, Emerging Multimedia Computer Communication Technologies, Pentice Hall, 1998
[4] K. R. Rao and J. J. Hwang, Techniques and Standards for Image, Video, and Audio Coding, Prentice Hall, 1996
[5] James D. Murray and William Vanryper, Graphics File Formats, OReilly & Associates, INC., April, 1996.
[6] File format from http://www.wotsit.org/
[7] Touradj Ebrahimi, MPEG-4 Video Verification Model Version 11.0, ISO/IEC JTC1/SC29 WG11 MPEG98/N2172, Tokyo, March 1998
[8] Vasudev Bhaskaran and Konstantinos Konstatinides, IMAGE AND VIDEO COMPRESSION STANDARDS Algorithms and Architectures second edition, Kluwer Academic Publishers, 1997
[9] Arun N. Netravali and Barry G. Haskell, Digital Pictures 2nd Ed., PLENUM, 1995
[10] Dean Clark, "Color octree routines" , Dr. Dobbs Journal, June, 1995
[11] Ian Ashdown, A Programmer's Perspective, John Willey & Sons Inc., 1995
[12] 유훈, 이창호, 민병석, 정제창, "영상 압축을 위한 저 복잡도를 갖는 고효율 엔트로피 부호화기," 제11회 신호처리 합동 학술대회, 제11권, 제1호, pp. 789-792, 10월 1998년
[13] Jain, Fundamentals of Digital Image Processing, Prentice Hall, 1989

민 병 석(Byongseok Min) 정회원
현재 : 한양대학교 전자통신공학과 박사과정
충청대학 전자통신과 교수
참조 : 한국통신학회 논문지 제 23권 4호

정 제 창(Jechang Jeong) 정회원
현재 : 한양대학교 전자통신공학과 교수
참조 : 한국통신학회 논문지 제 23권 4호

최 병 욱(Byunguk Choi) 정회원
현재 : 한양대학교 전자통신공학과 교수
참조 : 한국통신학회 논문지 제 23권 4호

www.dbpia.co.kr