

초저속 동영상 전송

류영진 김남철
경북대학교 전자공학과

Transmission of Motion Video at Very Low Data Rates

Young Jin Ryoo, Nam Chul Kim

Dept. of Electronics, Kyungpook National Univ.

<Abstract>

A new transmission scheme is presented for transmitting motion video at very low data rates. In this scheme, two-level cartoons are extracted from gray-level images, and their interframe differences are coded using 2D run-length coding. Experimental results show that the proposed scheme yields compression ratio as high as 216:1.

I. 서론

4.8 - 19.2 Kbps의 초저속으로 동영상 전송하는 기술은 영상전화 시스템 및 청각장애자 통신에 응용될 수 있다. ^(1,2) 그런데 이렇게 낮은 전송률로 동영상을 전송하기 위해서는 원래의 동영상을 경제적인 이진 영상으로 나타내는 것과 추출된 이진 영상을 효율적으로 부호화하는 것이 중요하다. 특히 이러한 분야에서는 사람의 시각으로 영상을 인식하고 해석하기 때문에, 원영상으로부터 사람의 시각적 특성이 고려된 이진 영상을 추출하여야 한다. 일반적으로 사람의 시각을 고려했을 때, 영상의 특징들만을 나타내는 스케치 영상보다는 카본 영상이 ⁽¹⁾ 사람의 시각에 좋다.

본 논문에서는 초저속으로 동영상을 효율적으로 전송하기 위하여, 비교적 사람의 시각적 특성을 잘 고려하는 DIP(difference of inverse probabilities)

^(3,4) 연산자와 이진화 기법을 이용하여 카본 영상을 추출한 후 프레임간에 카본 영상의 차영상을 2차원 조건부 run length 부호화하는 전송 시스템을 제안하였다.

II. DIP 연산자

DIP 연산자의 알고리즘을 간단히 표현하면 다음과 같다.

$$DIP = \begin{cases} \frac{\bar{f}(i,j)}{f(i,j)} - \frac{\bar{f}(i,j)}{f_m(i,j)}, & f_m(i,j) - f(i,j) > T_1 \\ 0, & f_m(i,j) - f(i,j) \leq T_1 \end{cases} \quad (1)$$

여기서 $f(i,j)$ 는 화소의 밝기를 나타내며 $\bar{f}(i,j)$ 는 3x3 창내의 합이고 $f_m(i,j)$ 는 3x3 창내의 최대치를 나타낸다. 따라서 이 연산자는 3x3 창내에서 최대치와 중심 화소의 차가 주어진 역치보다 큰 화소에 대해 중심 화소의 극부 확률의 역에서 최대값을 갖는 화소의 극부 확률의 역을 뺀 값(DIP)을 결과값으로한다.

이 연산자는 최대치와 중심 화소의 차를 고려하여 벨리 및 에지에서 큰 값을 갖고 밝기의 변화가 작은 영역에서는 작은 값을 갖는다. 또한 $f(i,j)$ 로 극부적 밝기를 고려하여 같은 밝기의 변화에서도 밝은 부분에서보다 어두운 부분에서 더 큰 값을 갖는다.

따라서 DIP 연산자는 밝기의 변화 뿐만 아니라 극부적 밝기를 고려하여 벨리와 에지를 추출함으로써

스케치 특성을 잘 나타낸다. 또한, DIP 연산자는 수행이 매우 간단하여 복잡도에 있어서 기존의 미분 연산자와 큰 차이가 없다.

III. 전송을 위한 부호화 과정

본 연구에서 제안한 전송 시스템의 블록도는 그림 1과 같다.

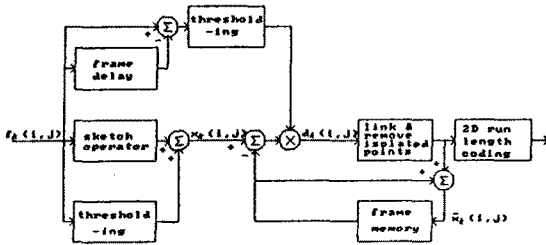


그림 1. 제안된 전송 시스템의 블록도

1. 카본화

원영상을 이진 영상으로 나타내는데 있어서 사람의 시각적 특성을 고려하면, 물체의 윤곽만을 나타내는 스케치 영상보다는 매우 어두운 부분은 검은색 1로 채워서 카본으로 나타내는 것이 더욱 좋다. 따라서 다른 특징 추출 연산자에 비해 시각적 특성을 잘 고려하는 DIP 연산자와 간단한 이진화 기법을 이용하여 다음과 같이 카본을 추출하였다.

$$x(i,j) = \begin{cases} 1, & f(i,j) < T_2 \text{ or } DIP > T_3 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

여기서 DIP는 DIP 연산자의 결과값이다. 카본화하는 과정을 간단히 설명하면, 화소의 밝기가 임계치 T_2 보다 작은 화소는 1로 하고, 임계치 T_2 보다 큰 화소에는 DIP 연산자를 적용하여 이진화한다. 이렇게 추출된 카본 영상은 스케치 영상보다 시각에 좋을 뿐만 아니라 부호화하였을때 감축율을 더욱 높게한다.

2. 차영상 취득

동영상에서는 프레임간에 중복성(redundancy)이 매우 많기 때문에 매 프레임마다 카본 영상을 그대로 전송하는 것은 효율적이지 못하다. 본 연구에서는 높은 감축율을 얻기 위해 현재 프레임의 카본 영상과 이전 프레임의 재생된 카본 영상과의 차영상을 부호화함으로써 프레임간의 중복성을 제거하였다. 다음 식은 차영상을 구하는 과정을 간단히 표현한 것이다.

$$d_k(i,j) = \begin{cases} x_k(i,j) \oplus \bar{x}_{k-1}(i,j) \\ \quad, & |f_k(i,j) - f_{k-1}(i,j)| > T_k \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

여기서 $x_k(i,j)$ 는 현재 프레임의 카본 영상이고, $\bar{x}_{k-1}(i,j)$ 는 이전 프레임의 재생 영상이다. 그런데 프레임간의 차영상을 구할 때, 화소 전체에 대해 구하지 않고 (3)식과 같이 원영상간의 밝기차가 어떤 임계치보다 큰 화소에 대해서만 카본 영상의 차를 구한다. 이렇게 하는 이유는 모든 화소에 대해 카본 영상의 차를 구할 경우, 물체의 움직임이 배경의 밝기 및 조명에 영향을 미쳐 프레임간의 카본 영상의 차가 상당히 많이 나타나 감축율이 크게 떨어지기 때문이다.

실제적으로 차영상을 구해보면 고립점이 많고 연결성이 좋지 못하다. 그런데 보다 높은 감축율을 얻기 위해서는 차영상에 고립점이 적어야 하고 연결성이 좋아야한다. 따라서, 차영상에서 5x5 창내에 중심 화소 자신만 차가 있으면 고립점으로 판단하여 제거하였다. 또한, 5x5 창내에 중심 화소가 프레임간의 차가 없는 것으로 나타나 있으면, 수직, 수평, 45 대각선 및 135 대각선 방향으로 차가 있는 것으로 나타난 화소의 수가 어떤 임계치보다 크면 연결하였다.

3. 2차원 run length 부호화

프레임간의 차영상을 부호화함에 있어서, 본

연구에서는 그래픽 전송에 사용되는 조건부 run length (CRC) 부호화를 사용하였다.^[5]

IV. 실험 결과 및 검토

실험에 사용한 영상으로는 크기가 192x128 화소의 256개 명암도를 갖는 ELLEN 영상을 사용하였다. 그림 2에서 보는 바와 같이 ELLEN 영상은 다소 복잡한 배경에서 한 여성이 전화하고 있는 장면이다.

그림 3은 ELLEN 영상의 첫번째 프레임과 세번째 프레임에 대해 DIP 연산자를 적용한 후 적절한 역치로 이진화 시킨 결과 영상이다. DIP 연산자가 극부적 밝기의 영향을 받아 어두운 부분에서 특징점을 잘 추출하고, 손의 윤곽을 뚜렷하게 추출하므로 벨리 또한 잘 추출함을 알 수 있다. 그림 4는 DIP 연산자를 이용하여 ELLEN 영상으로부터 추출한 카본 영상인데, 밝기가 매우 어두운 머리 부분이 검은색으로 채워져 그림 3보다 전체적으로 보기에 좋다.

그림 5(a)는 첫번째 프레임과 세번째 프레임간의 모든 화소에 대해 카본 영상의 차를 구한것이다. 본문에서 언급한 바와같이 배경 부분에서도 프레임간의 차가 상당히 많이 나타남을 볼 수 있다. 그림 5(b)는 프레임간에 원영상의 차가 어떤 임계치보다 큰 화소에 대해서만 카본 영상의 차를 구한 결과이다. 배경 부분에서 차는 적게 나타나지만 고립점이 다소 존재하고 연결성이 좋지 않다. 그림 5(c)는 그림 5(b)에서 고립점을 제거하고 연결성을 개선한 결과이다. 전체적으로 고립점이 제거되고 연결성이 다소 개선됨을 알 수 있다. 그림 5(d)는 재생된 영상인데 전체적으로 그림 4(b)와 동일하다.

최종적으로 그림 5(c)의 차영상을 2차원 CRC 부호화한 결과, 216배의 감축율을 얻었으며 초당 30 프레임을 전송할 때 27.3 Kbps의 전송율을 얻었다.

V. 결론

초저속으로 동영상상을 효율적으로 전송하기 위하여,

본 논문에서는 사람의 시각적 특성에 부합하는 카본 영상을 추출한 후, 프레임간에 카본 영상의 차영상을 구하여 부호화하는 전송 시스템을 제안하였다. 제안된 방법을 동영상에 적용한 결과, 추출된 카본 영상이 사람의 시각에 좋을 수 있었고 216배의 높은 감축율을 얻을 수 있었다. 앞으로 전처리 과정과 이동 보상 방식(motion compensation)을 추가하면 9.6 Kbps까지 낮추어지리라 기대된다.

VI. 참고문헌

1. D. E. Pearson and J. A. Robinson, "Visual communication at very low data rates," Proc. IEEE, vol.73, pp. 795-812, Apr. 1985.
2. J.F. Abramatic, Ph. Letellier and M. Nadler, "A narrow-band video communication system for the transmission of sign language over ordinary telephone lines," Image Sequence Processing and Dynamic Scene Analysis in NATO ASI series, Springer-Verlag, pp. 314-336, 1983.
3. Y. J. Ryoo and N. C. Kim, "A valley operator for extracting sketch features : DIP," accepted for publication in Electronics Letters, Apr, 1988.
4. 김남철, 류영진, "스케치 특징의 추출을 위한 벨리 연산자," 통신학회 영상처리 워크샵 논문집, pp. 16-19, 1987년 10월.
5. R. B. Arps, "An introduction and digital facsimile compression review," proc. ICC, pp. 7.1-3, 1975.

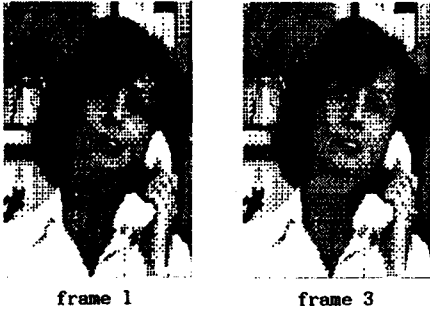


그림 2. ELLEN 영상

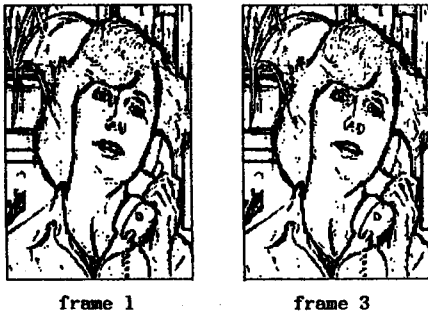


그림 3. ELLEN 영상에 대한 DIP 연산자의 결과



그림 4. 카본화한 결과 영상

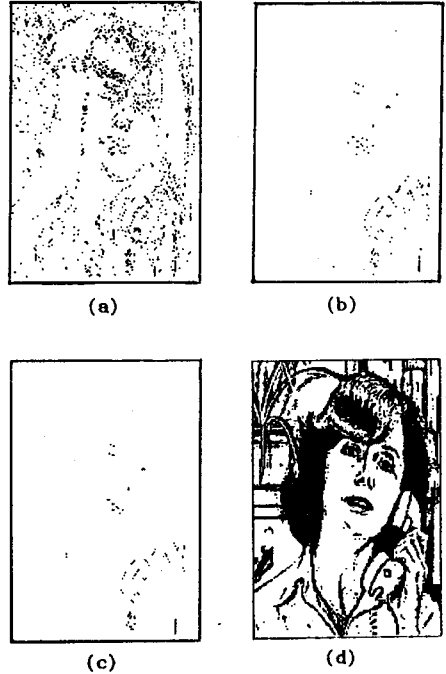


그림 5. (a) 첫번째 프레임과 세번째 프레임간에 카본 영상의 차영상
 (b) (a)에 현영상의 차를 고려한 결과
 (c) (b)에서 노이즈 제거와 연결성능 개선시킨 결과
 (d) 재생된 카본 영상