

저 전송율 비디오 부호화를 위한 효율적인 고속 움직임추정 알고리즘과 영상 분할기법

이 병석, 한수영, 이동규, 이두수
한양대학교 전자통신전파공학과 신호처리 연구실
전화 : 02-2290-0358 / 핸드폰 : 019-480-2943

Efficient Fast Motion Estimation algoritmm and Image Segmentation For Low-bit-rate Video Codina

Byung-Seok Lee, Su-Young Han, Dong-Gyu Lee, Doo-Soo Lee
Dept. of Electronics Engineering, Hanyang University
E-mail : spee@ihanyang.ac.kr

Abstract

This paper presents an efficient fast motion estimation algorithm and image segmentation method for low bit-rate coding. First, with region split information, the algorithm splits the image having homogeneous and semantic regions like face and arm regions in image. Then, in these regions, We find the motion vector using adaptive search window adjustment. Additionally, with this new segment based fast motion estimation, we reduce blocking artifacts by intensively coding our interesting region(face or arm) in input image. The simulation results show the improvement in coding performance and image quality.

I. 서론

우리는 현재 멀티미디어의 시대에 살고 있다. 멀티미디어에 있어서 가장 중요한 정보는 영상과 음성이며, 이러한 영상과 음성정보는 그 특성상 엄청난 양의 데이터를 가지고 있다. 이를 전송하거나 저장하기에는 현재의 제한된 대역폭으로는 큰 어려움이 있으므로 높은 압축률을 통하여 낮은 비트율에서 동작하는 부호화기가 필수적이며 특히 실시간으로 전송이 요구되는 화상전화나 영상회의에 있어서 압축률은 재생 시 화질

에 큰 영향을 미치게 된다[1].

동영상은 그 특성상 시간 영역과 공간 영역, 비트 열들 사이에서 많은 상관관계를 가지고 있으므로, 이를 제거하면 압축률을 높일 수 있다. 현행 ITU-T의 H.26x 계열과 MPEG등 대부분의 압축 표준안들은 시간적 상관성을 제거하는데 움직임 추정 기법(motion estimation)을, 공간적 상관성을 제거하는데 DCT변환을, 비트 열들 사이의 확률적 상관 관계를 제거하는데 가변길이 부호화(variable length coding)를 사용하고 있다. 하지만 전 영역의 움직임을 탐색하는 완전탐색 알고리듬(full search algorithm)은 많은 계산량과 수행 시간을 필요로 하며, DCT등의 블록 기반(block based)의 변환을 행하면 블록 잡음(blocking effect)과 같은 화질의 열화가 수반된다. 따라서 보다 나은 부호화 효율을 얻기 위한 고속 움직임 추정(fast motion estimation), 웨이블릿 기반 부호화(wavelet based coding), 의미론 기반 부호화(semantic based coding), 영역 분할 기반 부호화(segment based coding), 객체 기반 부호화(object based coding)등의 연구가 한창이다[1][2].

본 논문에서는 효율적인 저 전송률의 부호화를 구현하기 위해서, 영역 분할 정보를 이용하여 영상을 동일한 특성을 가지는 영역들(homogeneous regions)로 구성한 뒤 관심 영역을 추출하여 각 영역별로 탐색영역(search window)을 가변적으로 적용하여 계산량과 수

행시간을 줄이게 된다. 이러한 영역 분할 기반의 효율적인 고속 움직임 추정기법과 함께 입력 영상 중 관심 영역을 보다 충실히 부호화 함으로써 화질의 저하를 보완하는 기법을 제안한다.

본 논문의 구성은 다음과 같다. II, III장에서는 제안된 영역분할 및 움직임 추정 알고리듬에 관하여 고찰하고 IV장에서는 모의실험에 대한 결과를 보이며 V장에서는 앞으로의 개선점으로 결론을 맺는다.

II. 영역 분할

비디오 코딩과 같은 응용분야에 있어서는 입력 영상의 복잡한 배경(background region)으로부터 움직임이 보다 많은 전경(foreground region)을 분리하여 영역기반 부호화를 행하는 것은 현재의 제한된 채널에서 보다 개선된 화질을 보장하게 한다[3][4].

본 논문에서는 화상전화나 화상회의와 같은 영상 시퀀스의 특성이 배경과 사람의 얼굴부분, 가슴부분으로 구성된다는 점에 착안하여 입력영상 중 좀 더 의미 있는 얼굴영역과 팔 영역을 추출하였으며 영역 분할 단계의 전체 알고리듬 구성은 아래 그림과 같다.

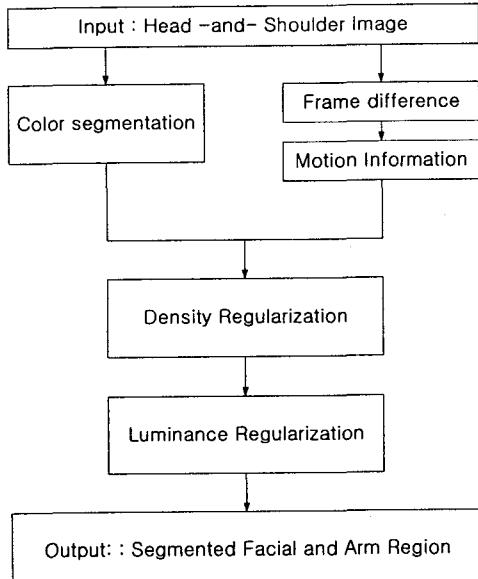


그림 1. 영역 분할 알고리듬

실시간 처리를 요구하는 응용분야에서는 고속으로 관심영역을 추출해 내는 것이 중요하므로 간단하고 시간이 적게 걸리는 차 영상에 의한 방법과 함께 움직임 정보, 얼굴의 색정보를 이용하였다.

2.1 칼라를 이용한 분할(Color Segmentation)

첫번째 단계로 입력 영상의 색성분 신호인 Cr, Cb 값을 이용하여 색상을 포함하는 픽셀을 추출하게된다[3].

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

$$O_1(x, y) = 1, \text{ if } [Cr(x, y) \in R_{Cr}] \cap [Cb(x, y) \in R_{Cb}] \\ 0, \text{ otherwise}$$

즉, Cr, Cb 값이 동시에 그 상대적인 범위(respective range) 안에 존재하면 색 영역으로 판정한다.

2.2 차 영상을 이용한 분할(Frame Difference)

배경 영역에 포함된 색성분을 가지는 영역은 움직임이 거의 없으므로 두 프레임간의 차 영상을 이용하여 배경으로부터 사람의 얼굴과 팔을 포함하는 영역을 쉽게 분리해 낼 수 있다.

2.3 움직임 추정기법을 이용한 분할

차 영상으로부터 구해진 영역에 대하여 좀 더 움직임이 많은 영역을 구하기 위하여 움직임 추정기법을 이용한다.

2.4 분리된 영역의 정규화(Regularization)

검출된 영역에서의 noise 제거와 세밀한 분할을 위하여 모폴로지 필터를 이용하여 침식연산과 팽창연산을 행한다[7]. 사용된 모폴로지 필터는 다음의 식과 같이 프레임 차에 따라서 가변크기를 갖도록 하였다.

$$r = R_1, \text{ if } |FD(x, y)| \leq Th$$

$$r = R_2, \text{ otherwise}, \text{ where } R_1 \geq R_2$$

프레임 차가 큰 영역은 상대적으로 움직임이 많고 세밀한 영역화가 필요하므로 모폴로지 필터의 작용범위를 줄여준다. 반면에, 프레임 차가 작은 영역은 움직임이 거의 없는 영역이므로 이때는 모폴로지 필터의 작용범위를 크게하여 가능한 한 같은 영역들로 분리 되도록 하였다.

2.5 결과 영상 획득

제안된 알고리즘에 의하여 부호화의 대상이 되는 얼굴 영역과 손 영역을 분리해 낸 영상을 얻어 부호화하고 전송한다.

III. 적응적인 탐색 위치를 가지는 고속 움직임 추정 알고리듬

본 논문에서는 영역 분할과 함께, 각 영역을 부호화 시 전 영역 탐색 블록매칭 알고리듬(FSBMA)의 복잡한 계산량을 줄이기 위하여 블록의 종류에 따라서 탐색영역을 가변적으로 적용하여 탐색영역을 감소시키는 알고리듬(ASBMA)을 이용한다[5].

- 움직임 추정 단계는 크게 다음의 3 단계로 구성된다.
1. 블록 분류(Classifying Block)
 2. 탐색 영역의 설정 및 BMA 수행
 3. 임계값의 재 추정(Update the Threshold)

3.1 블록 분류

우선, 분할된 결과 영상을 이용하여 다음과 같이 4종류의 블록으로 분할한다.

1. 분할된 영역(얼굴과 손 영역)에 존재하는 블록(Static Block)
2. 분할되지 않은 영역에 존재하는 블록(Active Block)
3. SB에서 AB영역(CB1)으로 또는 AB영역에서 SB영역(CB2)으로 변경되는 블록(Change Block)

3.2 탐색 영역의 설정 및 BMA 수행

SB와 CB2 영역에 대하여는 탐색 영역을 적게 설정 하며(SB:w/4, CB2:w/2, w는 움직임 벡터의 최대 범위 초기 값), AB와 CB1 영역에 대하여는 탐색 영역을 크게 설정(AB, CB1:w) 하여 움직임 추정 상관함수인 MAD(Mean Absolute Difference)를 이용해 블록 매칭 움직임 추정을 수행한다[6].

$$MAD(i,j) = \frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |I_t(x+m, y+n) - I_{t+1}(x+m+i, y+n+j)|$$

그림2에 설정된 탐색영역에 대해서 모션 벡터를 찾는 방법을 나타내었다.

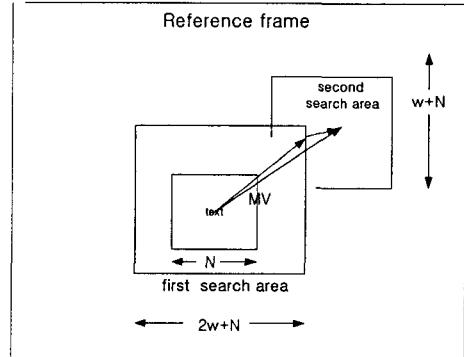


그림 2. 설정된 탐색영역 안에서의 움직임 추정

MV가 첫 번째 탐색영역의 경계를 가리키면 두 번째 탐색영역으로 이동하여 움직임 추정을 수행한다. 두 번째 MV가 두 번째 탐색영역 내에 결정이 되면 움직임 추정과정을 멈추고 두 MV간의 벡터 합을 구하여 최종 MV를 얻는다.

3.3 임계값의 재 추정

첫 번째 단계에서 분류되었던 블록들이 변화가 있는지를 검사하여 임계값을 새로 수정한다.

IV. 모의 실험

본 절에서는 제안된 영역분할 기법과 움직임 추정기법 알고리듬을 이용하여 실제 영상에 적용한 모의 실험결과를 제시한다. 모의 실험에 사용된 영상은 360×288 의 "salesman" 시퀀스 30프레임과 "clare" 시퀀스 50프레임을 사용하였다.

그림3에는 영역분할의 각 단계별 결과 영상을 나타내었다. (a)와 (b)는 각각 5번째, 7번째 프레임의 "salesman" 영상이며 (c)는 두 프레임간의 차 영상이다. (d)는 차 영상에서 좀 더 움직임이 많은 부분을 찾기 위하여 움직임 추정을 한 결과이며 (e)는 칼라정보를 함께 이용하여 관심영역을 추출한 최종 결과 영상이다. 그리고 (f)에는 20번째 frame "clare" 영상의 복호 영상을 나타내었다.

차 영상(c)에서는 움직임이 거의 없는 배경영역에서 전경영역을 추출한 결과를 볼 수가 있으며 추출된 전경영역에서도 우리의 관심영역인 얼굴과 손 영역에서 움직임이 많이 있음을 (d)영상에서 관찰할 수가 있다. (e)영상은 차 영상과 움직임 정보, 사람의 살색정보를

함께 이용하여 보다 정확한 영역 추출을 한 최종 결과 영상을 보여주며 (f)에는 관심영역만을 부호화 하였을 때 수신 단에서 얻어진 복호화 결과 영상을 나타내었다.



그림 3. 각 단계별 결과 영상((a) 5th 프레임, (b) 7th 프레임, (c) 두 프레임간의 차 영상, (d) 움직임 추정 영상, (e) 분할 영상, (f) 20th 복호화 영상)

표 1. ASBMA 수행결과

Test Sequence	평균 PSNR(dB)	
	FSBMA	ASBMA
Clare	45.64	45.17
Salesman	42.06	41.32

표1에는 제안된 알고리듬을 이용하여 부호화를 수행하였을 때의 결과를 FSBMA와 비교하였다. 제안된 알고리듬이 전 영역을 탐색하였을 때보다 "clare" 시퀀스의 경우 PSNR이 0.4 [dB], "salesman" 시퀀스의 경우 0.7 [dB] 감소하였다.

V. 결론 및 개선점

본 논문에서는 저 전송률 비디오 부호화를 위하여 관심영역을 추출하여 가변적인 탐색영역을 설정한 후 움직임 추정을 수행하는 영역기반 부호화 기법을 제안하였다. 관심영역을 추출할 때에는 속도 향상을 위하여 차 영상과 움직임 정보를 이용하였고, 배경에서의 움직임이 있는 경우를 위하여 사람의 살색 정보를 함께 이용하였다. 탐색영역의 설정은 추출된 영역의 종류에 따라 4가지 블록으로 분류하여 각각 적절한 범위를 설정하여 주었다.

제안된 알고리듬이 부호화의 측면에서 성능의 향상이 있음을 모의 실험결과를 통해 알 수 있었으며 앞으로의 개선점은 움직임이 적은 화상전화나 화상회의 영상뿐 아니라 보다 움직임이 많은 영상에도 확대가 되도록 많은 연구가 이루어져야 할 것이다.

Reference

- [1] Ralf Shafer, Thomas Sikora' "Digital Video Coding Standards and Their Role in Video Coding" Proceedings of The IEEE, vol. 83, No.6, June 1995
- [2] A. Murat Tekalp " Digital Video Processing", Prentice Hall, 1995
- [3] Douglas Chai, King N.Ngan, "Face Segmentation Using Skin-Color Map in Videophone Applications" IEEE Transactions on circuits and systems for video technology, June. 1999.
- [4] Hualu Wang, Shih-Fu Chang "A Highly Efficient System for Automatic Face Region Detection in Mpeg Video" IEEE Transactions on Circuits and System for Video Technology, August. 1997
- [5] Hwang-Seok Oh, Choong-Hoon Lee "A New Block Matching Algorithm Based on An Adaptive Search Area Adjustment using Spatio-Temporal Correlation" IEEE Transactions on Consumer Electronics, June. 1999
- [6] Borko Fuhr, Joshua Greenberg, "Motion Estimation Algorithms for Video Compression", Kluwer Academic Press, 1997
- [7] Anil K. Jain, "Fundamentals of Digital Image Processing", Prentice Hall, 1989