

움직임 분할 기반의 프레임간 보간 기법

오근영*, 나인호*, 고남영*, 김동욱**, 김진태***

*군산대학교 정보통신공학과

**전주대학교 전기전자정보통신공학부

***한서대학교 컴퓨터학과

Interframe interpolation technique based on motion segmentation

*Keun-Young Oh, *In-Ho Ra, Nam-Young Ko

Dong-Wook Kim, *Jin-Tae Kim

*Kunsan National University

Jeonju University, *Hanseo University

E-mail : dwkim@jeonju.ac.kr, jtkim@gaya.hanseo.ac.kr

요 약

본 논문에서는 움직임 필드를 분할하고, 이를 바탕으로 가변 보간율을 이용한 새로운 비디오 보간 기법을 제시한다. 움직임 필드의 분할은 수리형태학적 필터링 및 영역별 움직임 추정을 이용하여 구성된다. 구성된 움직임 필드는 각각의 움직임 특성을 바탕으로 배경 또는 정지 영역, 움직임 영역, 드러난 영역, 가려진 영역 등의 분류에 바탕을 둔다. 제시된 보간기법은 움직임이 복잡하거나 비선형적일 때, 보간율을 적게 하며, 움직임이 적거나 선형적일 때 보간율을 높게 함으로써 일정한 화질을 유지할 수 있도록 하였다.

1. 서 론

초저비트율 부호화를 통한 동영상의 전송은 통신채로의 전송능력의 발달에도 불구하고 여전히 그 필요성이 존재한다. 초저비트율 부호화를 위한 많은 기법들이 존재한다. 이러한 기법 중 프레임간 보간 기법은 디스플레이 장치의 시간적 확장성에 대한 해결 방안이 될 수 있는데, 즉 다양한 멀티미디어를 표현하는 단말기 시스템에서 프레임율을 높여야 할 필요가 있을 때 프레임 보간 기법이 유용한 수단이 된다. 또한 프레임간 보간 기법은 기존의 동영상 압축 알고리즘과 결합할 경우 매우 유용한 압축 수단이 될 수 있다. 즉, 기존 프레임으로 설정된 프레임을 움직임 보상 및 DCT 등 기존의 동영상 부호화 기법으로 전송하고, 스킵된 프레임에 대해서 수신단에서 재현해 내는 기법이다. 프레임간 보간기법은 codec의 구현에 있어서 다소 복잡함이 따르지만 전송 비트율을 줄이기 위해서 유용한 수단이 될 수 있다.

여러 가지 움직임 보상형태의 영상 보간기법이 제시되었는데, 이를 위해 여러 가지 움직임 추정 기법이 적용되었다 [1][2][5]. [2]에서는 순방향 및 역방향 움직임 트래킹을 위해서 recursive 알고리즘이 사용되었다. R. Thoma [4]등은 계층적 블록

매칭기법을 이용한 움직임 벡터 필드의 추출기법이 제시되었다. 이러한 기법들은 대부분 화소단위 또는 블록 기반의 움직임 추정 기법을 적용한다. 그러나, 화소기반은 여러 가지 잡음의 영향 등으로 국부최소치 문제에 빠지기 쉽고, 블록 정합기법은 물체의 형태를 고려하지 않기 때문에 오정합의 문제가 발생되기 쉽다. 또한, 이러한 기법들은 일정한 보간율을 갖는다. 즉, 기존의 알고리즘의 공통적인 점은 움직임의 양이 크지 않고, 움직임의 복잡도가 낮다는 가정을 전제로 움직임의 복잡도에 상관없이 일정한 보간율을 갖는다.

일반적으로 프레임간 보간은 적절한 움직임 추정을 행하여 얻어진 움직임 벡터를 바탕으로 움직임 필드를 구성하고 기존 프레임간 일치되는 블록, 영역 또는 화소를 찾아 이를 바탕으로 보간될 프레임을 구성한다. 이 때, 보간된 프레임의 화질은 기존되는 프레임간의 움직임 정도에 의해 많이 좌우된다. 따라서, 움직임이 많은 경우 보간되는 프레임의 화질은 저하된다. 특히, 움직임이 비선형적일 경우 화질 저하 정도는 더욱 악화된다. 따라서, 보간되는 프레임 간격이 커질수록 움직임의 정도에 더 큰 영향을 받는 것이 일반적이다. 따라서, 기존의 기법들은 아래와 같은 문제점을 야기할 수 있다.

움직임이 비선형적인 경우 화소간에 비정합점이 발생하기 쉽다.

움직임에 의존하여 보간된 영상간의 화질차가 변화가 심하다.

본 논문에서는 이러한 문제점을 해결하기 위하여 영역기반의 가변 간격 프레임 보간 기법을 제시한다. 제안된 보간 기법은 기준프레임(anchor frame) 간의 간격을 움직임의 정도에 적용적으로 변화시킴으로써 프레임간 화질 변화를 적게 유지할 수 있도록 하고, 또한 영역기반의 움직임 추정 및 벡터 교정을 통해 실제 움직임에 근접된 움직임 필드를 구성하고, 이를 프레임 보간에 적용한다.

II. 가변간격의 프레임 보간기법

II-1. 움직임 영역 추출

일반적으로 프레임 보간 기법이 적용되는 영상은 움직임의 범위 및 움직임이 크지 않은 영상을 대상으로 하는 것이 일반적이다. 먼저 카메라 잡음 등 여러 가지 요인으로 인한 잡음의 영향을 제거하기 위해 배경을 제외한 움직임 영역을 추출한다. 움직임 추출은 배경영역의 제거와 움직임 에지의 추출, 두 단계로 구성된다.

II-1-1. 움직임 영역 추출

배경영역은 카메라가 정지되어 있는 경우 전후 프레임간 공간적 변화가 없는 영역이며, 이는 프레임간 차를 바탕으로 이루어진다. 즉,

$$C_{k,i}(x_1, x_2) = \begin{cases} 1 & \text{if } (|FD_{k,i}(x_1, x_2)| > T_1) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

여기서, $FD_{k,i}(x_1, x_2)$ 는 프레임 k, i 간의 (x_1, x_2) 에서의 차를 의미하며, $C_{k,i}$ 는 문턱치 적용 후 얻어지는 값으로 변화 영역에 속하는 화소는 1, 배경에 속하는 화소는 0이다.

II-1-2. 수리형태학적 필터링 (morphological filtering)을 이용한 영역 스무딩

본 단계는 위에서 얻어진 결과를 바탕으로 배경과 움직임 영역의 경계를 스무딩시키는 단계이다. 이를 위해 본 논문에서는 수리형태학적 필터링[6]을 적용한다. 본 논문에서는 수리형태학에서 정의된 팽창(dilation)과 침식(erosion)처리를 바탕으로 개방(opening)과 폐쇄(closing)을 순서대로 행한다.

$$\begin{aligned} open(A, B) &\equiv dil(er(A, B), B) \\ &= A(A(-)B)(++)B \\ &= (A[-](-B))[+]B \end{aligned} \quad (2)$$

$$\begin{aligned} close(A, B) &\equiv er(dil(A, -B), -B) \\ &= A(A(++)(-B))(-)(-B) \\ &= (A[+](+B))[-]B \end{aligned} \quad (3)$$

여기서, 식(2)는 개방을 식(3)은 폐쇄를 나타내는데, 개방은 침식, $er(A, B)$ 을 행한 후 팽창, $dil(er(A, B), B)$ 을 행하고, 폐쇄는 팽창, $dil(A, -B)$ 을 행한 후 침식, $er(dil(A, -B), -B)$ 을 행한다. 이 때, A는 움직임 영역이며, B는 이 때 적용된 구성소(structuring element)이다.

II-1-3. 스킵 프레임수 결정

프레임간 보간은 영역 내부보다는 영역별 경계인 움직임 에지를 중심으로 한 부분에서 더욱 더 적절한 보간기법이 요구된다. 따라서, 본 논문에서는 위에서 얻어진 움직임 영역에 대해 움직임 에지를 포함하는 소영역을 구한다. 이를 움직임 에지 영역(Motion Edge Region: MER)이라 칭하기로 한다. MER은 일반적으로 프레임간 차를 구하여 차가 큰 경우에 발생하므로 식 (1)의 문턱치를 높게하여 구하는데, $T_1 \ll T_2$ 이다.

$$ME_{k,i}(x_1, x_2) = \begin{cases} 1 & \text{if } (|FD_{k,i}(x_1, x_2)| > T_2) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

즉, $ME_{k,i}$ 가 1일 때 이 화소는 움직임 에지 영역으로 판정한다. 이것을 움직임 영역에 대해 제한적으로 검사한다. 'salesman' 영상(그림 1)을 대상으로 적용했을 때 그림 2는 움직임 영역을, 또 그림 3은 MER을 구한 결과이다.

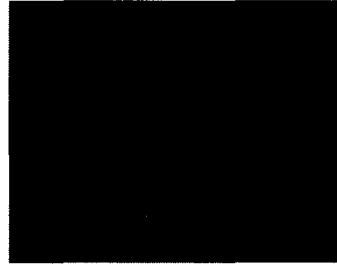


그림 1 'salesman' 영상

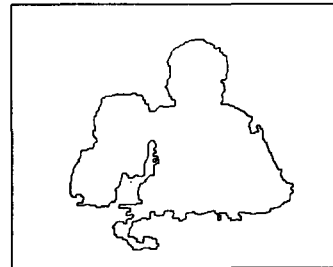


그림 2 움직임 영역



그림 3 MER

언어진 MER은 스킵 프레임 수 결정 및 보간에 이용된다. 먼저 스킵 프레임은 MER의 화소수를 바탕으로 결정된다. 즉,

$$MER_{NUM}(k, l) = \sum_S MER(x_1, x_2) \quad (5)$$

여기서, $MER(k, l, x_1, x_2)$ 는 프레임 k , l 간의 공간 S 의 MER을 나타낸다. 언어진 $MER_{NUM}(k, l)$ 을 바탕으로 해당 프레임으로 스킵할 것인가 아닌가를 결정한다. 만일 움직임 블록의 개수가 사전에 설정된 임계치 이하이면 스킵하고 그렇지 않을 경우 그 프레임을 기준 프레임으로 설정한다. 만일 스킵되면, 기준 프레임과 다음 프레임간에 같은 평가 계산을 행한다. 이를 정리하면 아래와 같다.

두 기준 프레임을 각각 I_A, I_B 라고 할 때,

단계 1: 시작되는 기준 프레임 I_A 을 정하고, 바로 인접한 k 번째 I_k 를 기준 프레임 I_B 로 설정.

단계 2: I_k 프레임의 스킵 여부를 결정한다. 즉, 기준 프레임 I_A 와 인접한 기준 프레임 I_k 간의 움직임 정도를 평가하여 스킵프레임인지 아닌지를 결정.

단계 3: 만일 이 스킵프레임이면 기준 프레임을 I_{k+1} 로 설정하고, 단계 2로 감.

단계 4: 만일 단계 3의 결과가 스킵 프레임이 아니면 그 때의 기준 프레임을 보간을 위한 기준 프레임 I_B 로 결정.

II-2. 프레임간 보간기법

II-2-1. 움직임 벡터 필드의 구성

프레임간 보간기법의 성능을 좌우하는 것 중의 하나는 정확한 움직임 벡터의 추출이다. 먼저 정확한 움직임 벡터의 추출을 위해서 여러 가지 기법들이 적용될 수 있다. 움직임 추출을 위해서 대표적으로 화소 기반, 블록 기반 및 영역 기반의 3가지 형태로 고려될 수 있다. 먼저 화소 기반의 경우 pel-recursive 기법이 대표적으로 적용될 수 있으며, 블록 기반은 블록 매칭 알고리즘(block matching algorithm)이 대표적이다. 마지막으로

영역 기반은 움직임 추정에 앞서 정확한 영역 분할이 선행되어야 하는데, 일반적으로 분할된 결과는 영상의 구성에 의존하여 매우 민감하다. 본 논문에서는 이러한 문제점을 해결하기 위해 영역 기반을 기반으로 한 움직임 추정 기법을 적용하고, 이를 바탕으로 벡터 교정기법을 적용한다.

II-2-2. 움직임 추정

동영상 압축에 널리 사용되고 있는 블록기반 기법은 물체의 실질적인 움직임을 추정하기에는 물체의 경계 등을 고려하지 않고 이로 인하여 오정합이 발생되기 쉽기 때문에 비적합한 방법이다. 따라서 본 논문에서는 수신된 2장의 기준 프레임에 대해 영역 기반의 움직임 추정기법을 적용한다. 움직임 모델은 여러 가지 모델이 고려될 수 있는데 아래와 같이 6개의 파라미터로 구성되는 affine 움직임 모델 [3]로 표현될 수 있다.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_1 + a_2x + a_3y \\ a_4 + a_5x + a_6y \end{pmatrix} \quad (6)$$

$$= \begin{pmatrix} s_x \cos \theta_x & -s_y \sin \theta_y \\ s_x \cos \theta_x & s_y \cos \theta_y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \end{pmatrix}$$

본 논문에서는 영역 기반의 움직임 추정에 계산적 복잡도 등을 고려하여 다음과 같이 병진이동만 고려를 한다.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \end{pmatrix} \quad (7)$$

II-2-3. 벡터의 교정

영역별로 언어진 움직임 벡터는 몇 가지 측면에서 교정될 필요가 있다. 일반적으로 영역 분할은 알고리즘에 적용되는 문턱치의 설정에 따라 영향을 크게 받으며, 영역간 밝기 값 또는 칼라 값의 차가 크지 않은 경우 정확한 영역 분할을 장담할 수 없다. 따라서, 본 논문에서는 영역별 움직임 추정 결과를 바탕으로 화소별 벡터를 교정한다. 화소별 벡터 교정은 크게 두단계를 가지는데, 첫째는 DFD를 바탕으로 한 벡터 교정이다. DFD를 바탕으로 한 화소벡터 교정은 교정될 임의의 화소에 대해 거리 $r(x, y)$ 를 정의하고, 이 범위내의 영역별 벡터를 얻고 이를 v_R 로 정의한다. 즉, v_R 은

$$v_R = \{v_{R1}, v_{R2}, \dots, v_{Rn}\} \quad (8)$$

이다. 이 벡터들을 이용하여 DFD를 구하고 최소 DFD를 갖는 벡터를 (x, y) 의 화소 벡터 $v_p(x, y)$ 로 한다. 즉,

$$v_p(x, y) = \min[DFD\{v_{Rn}\}] \quad (9)$$

여기서 $DFD(v_R)$ 는 $[I_n(x, y) - I_{n+dir}(x_i, y_i)]^2$ 이다.

화소별 DFD 기반의 교정이 끝나면, 프레임내의 그래디언트 기반의 벡터 병합이 행해진다. 즉, 인접한 화소끼리 유사한 밝기 값을 가지는 경우 동일한 물체로 볼 수 있으며, 이러한 점을 고려하여 화소 벡터의 재교정을 행한다. 이 때, 인접한 화소간 gradient를 구하기 위하여 4근방(4-neighbor) 화소점을 고려한다. 일반적으로 Laplace 오퍼레이터는 다음과 같은 filter mask를 갖는다. 즉,

$$L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (10)$$

화소 (x, y) 에 대해 Laplace 오퍼레이터를 취한 후 얻어지는 필터값을 $L(x, y)$ 라고 할 때,

$$|L(x, y)| < T$$

인 조건을 만족하면 동일 영역으로 취급한다. 여기서 T 는 문턱 값으로서 작은 값을 가진다. 만일 이 경우 벡터는 4근방 화소 점의 벡터와 동일한 값을 부여한다.

III-3. 보간 알고리즘

스킵된 프레임은 재구성하기 위해서는 기준 프레임간에 구해진 양방향 움직임 벡터 필드를 바탕으로 이루어진다. 움직임에 따라 발생하는 화소의 성격은 크게 다음과 같이 분류할 수 있다. 즉, 정지된 배경, 움직이는 부분(moving part), 가려지는 부분(covered part), 드러난 부분(uncovered part) 등이다. 먼저, 움직이는 부분은 배경과 함께 해당 화소의 움직임이 일치하는 경우이다.

기준 프레임을 각각 I_A, I_B 일 때, N 을 기준 프레임간 간격, k 를 I_A 로부터 간격이라 할 때, 만일 보간되는 프레임 I_k 의 임의의 화소 (x_k, y_k) 가 움직이는 부분일 경우, 양 기준 프레임의 벡터가 일치한다면, 해당되는 기준 프레임의 화소값의 평균으로 보간된다. 드러난 배경의 경우 I_B 의 일부분이 I_A 에 존재하지 않을 때 발생된다. 이 경우 I_k 에는 I_B 로부터 보간된다. 이 때 움직이는 부분 또는 드러난 배경 영역 등에는 채워지지 않는 고립점들이 존재할 수 있다. 이 경우, 인근 8근방점들의 화소 밝기값을 이용하여 보간된다.

III. 컴퓨터 모의실험

제시된 알고리즘의 성능 평가를 위해서 컴퓨터 모의실험을 행하였다. 대상 영상으로는 'sales-

man' 영상을 사용하였으며, 352*288 크기를 가지며, 200 프레임에 대해 실험을 행하였다. 그림 4에서 아랫쪽 숫자는 스킵되는 프레임 수를 나타내는 것이며, MER을 바탕으로 스킵된 프레임 수를 나타낸다. 그림 4에서 위의 두 개의 그래프는 제시된 가변 간격기법을 적용하여 보간된 결과와 기존의 고정 스킵수 적용한 결과를 제시한 것이다. 고정 보간율의 경우 2 프레임을 스킵한 경우로 고정시키고, BMA, 블록 정합법에 의해 얻어진 벡터를 기반으로 얻어진 결과이다. 그림에서 가변 간격의 기법의 경우 고정 간격에 비해 화질의 변화가 적음을 알 수 있다. 따라서, 가변 간격을 3 프레임 이상 스킵하는 경우로 확대할 경우 화질 변화를 더욱 줄이면서 전송에 소요되는 비트량을 줄일 수 있음을 예상할 수 있다.

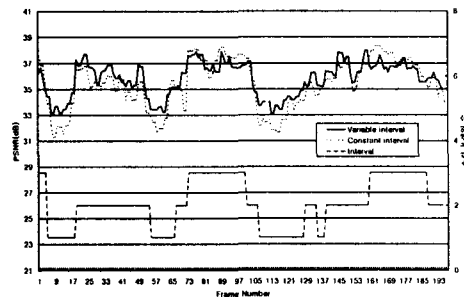


그림 4 보간된 프레임 화질

IV. 결론

본 논문에서는 가변 간격의 프레임간 보간 기법을 제시하였다. 제시된 기법은 초저비트율 전송 시스템에서 스킵되는 프레임 수를 고정할 경우에 비해 프레임별 화질의 변화를 줄일 수 있으며, 시간적 확장성에 효과적으로 대처할 수 있다.

참고문헌

- [1] H. G. Musemann, P. Pirsch, and H. J. Grallert, "Advances in picture coding," *Proc. IEEE* 73, pp. 523-548, 1985.
- [2] Cafforio C. and Rocca F., and Tubaro S., "Motion compensated image interpolation," *IEEE Trans. Commun.* COM-38, 215-222.
- [3] V. Seferidis and M. Ghanbari, "General approach to block matching motion estimation," *Opt. Eng.*, vol. 32, no. 7, pp. 1464-1474, July 1993.
- [4] R. Thoma and M. Biering, "Motion compensating interpolation considering covered

and uncovered background," *Signal Processing: Image Communication I*, pp. 191-212, 1989.

[5] J. Ribas-Corbera and Sklansky, "Interframe interpolation of cinematic sequences," *Journal of Visual Comm. and Image Rep.*, vol. 4, no. 4, pp. 392-406, DEC. 1993.

[6] Robert J. Schalkoff, *Digital image processing and computer vision*, John Wiley & Sons, INC, 1989.