

동영상 압축 방식을 위한 적응 움직임 벡터 예측 방식

김지희, 홍민철

송실대학교 정보통신공학과

Adaptive Motion Vector Prediction for Video Coding

Ji-Hee Kim and Min-Cheol Hong

School of Electronic Engineering, Soongsil University

Email: mhong@e.ssu.ac.kr

요 약

본 논문에서는 최소 계승 선형 예측 방식의 에지 방향성을 이용하여 공간영역에서의 다양한 움직임 벡터 예측기를 적응적으로 설정하는 방식을 제안하고자 한다. 적응 움직임 예측 방식은 동영상 움직임 벡터가 국부 통계적인 특성의 돌연한 변화로 특징지어진다는 것을 바탕으로 예측기를 움직임 벡터의 통계적인 특성에 따라 전환하는 방식이다. 본 논문에서 사용된 최소 계승 예측 방식은 움직임 벡터의 다양한 통계적 특성을 이용하여 국부적으로 움직임 벡터 예측 계수를 최적화하지만 최적화 과정에서 매우 큰 계산량을 갖게 됨으로 실제적으로 적용하기가 어려웠다. 그러므로 본 논문에서는 최소 계승 예측 방식을 에지 방향성의 관점에서 재해석하여 적응적으로 움직임 벡터 예측기를 개선하므로 계산량을 줄이면서 일정한 성능을 유지함을 확인 할 수 있었다.

1. 서 론

움직임 벡터 예측 방식은 인접한 움직임 벡터의 상관관계를 이용하여 비대칭(skewed)된 분포를 갖는 예측 오류(prediction error)를 생성하므로 압축 효율을 증가시킨다. “비대칭 (skewed)”은 어떤 예측 오류가 다른 예측 오류보다 더 높은 확률을 갖는다는 것을 의미한다[1]. 차등 펄스 부호화 변조(Differential Pulse Code Modulation(DPCM))는 예측 오류의 확률분포가 비대칭되는 대표적인 방식으로 사용되는 예측기에 의해 움직임 벡터 압축 방식의 성능이 좌우된다.

ITU 의 H.264 와 같은 차세대 동영상 압축 방식은 움직임 벡터를 부호화하는 과정에서 미디언 예측기 (median predictor)를 사용한다. 미디언 예측 방식은 인접한 블록의 움직임 벡터를 이용하는 문맥 기반의 (context-based) 비선형 예측 방식으로 에지를 보전하는 성질로 인해 평균 예측 방식보다 성능이 우수하고, 전송된 예측 오류와 이미 처리된 인접 블록의 움직임 벡터만으로 움직임 벡터를 재구성하는 이점을 갖는다. 그러나, 미디언 예측 방식은 인접한 세 블록의 움직임 벡터 크기를 비교하여 예측기를 설정하기 때문에 평

균 예측 방식보다 계산량이 크게 나타난다. 또한, 미디언 예측방식은 빠르고 복잡한 움직임을 갖는 동영상에서 선형 예측방식을 이용하여 얻어진 예측 오류의 분산 값보다 큰 값을 가지므로 움직임 벡터를 나타내는 비트 수를 최소화하기에 부적합하다[2-4]. 이러한 문제점을 해결하기 위해서 본 논문에서는 최소 계승 방식을 기반으로 하는 적응 움직임 벡터 예측 방식에 대해 제안한다.

적응 움직임 벡터 예측 방식은 움직임 벡터가 국부 통계적인 특성의 돌연한 변화(에지)로 특징지어진다는 것을 바탕으로 예측기를 움직임 벡터의 통계적인 특성에 따라 전환하는 방식이다. 일반적인 전환 방식 (Switching function)을 이용하여 간단하고 고정된 예측기들 중 하나를 적응적으로 결정하는 방식과 각 움직임 벡터의 예측 계수만을 적응적으로 설정하는 방식으로 크게 구분된다. 전자의 대표적인 방식은 문맥 기반의 무 손실 적응 영상 압축 방식(Context-based Adaptive Lossless Image Coding)[5]의 기울기 보정 예측 방식 (Gradient Adjusted Prediction)과 적은 계산량을 갖는 문맥기반의 무 손실 영상 압축 방식(Low Complexity, context-based, lossless image coding)[6]의 미디언 에지 검출 방식(Median edge Detection)이고, 후자의 대표적인 방식은 최소 계승 적응 방식(Least Square Adaptation)이다.

최소 계승 예측 방식(Least Square prediction)은 움직임 벡터의 다양한 통계적 특성을 이용하여 국부적으로 움직임 벡터 예측 계수를 최적화하지만 최적화 과정에서 매우 큰 계산량을 갖게 됨으로 동영상 표준화 방식에 실제적으로 적용하기 어려웠다. 그러므로 본 논문에서는 최소 계승 예측 방식을 에지 방향성의 관점에서 재해석하여 적응적으로 움직임 벡터 예측기를 개선하므로 계산량을 줄이는 방식을 제안한다.

본 논문은 다음과 같이 구성되어 있다. 2 장에서 제안된 방식의 이론적 배경인 최소 계승 방식을 살펴보고, 3 장에서 최소 계승 방식의 에지 방향성을 설명한다. 제안된 방식인 적응 움직임 벡터 예측 방식은 4 장에서 제시하고 5 장에서 제안된 방식을 뒷받침하는 실험결과를 포함한다. 최종적으로 6 장에서 제안된 방식에 대한 결론을 맺는다.

2. 이론적 배경

영상내의 움직임 벡터가 스택(stack) 순서로 저장되었다고 가정하여 동영상내의 한 프레임의 공간좌표를 하나의 인덱스(index) n 으로 명시한다.

최소 계승 방식을 기반으로 하는 움직임 벡터 예측 방식은 움직임 벡터의 N 차 마르코비안(Markovian) 성질에 의해 N 개의 이미 처리된 인접한 움직임 벡터만을 고려한다.

$mv(n-11)$	$mv(n-8)$	$mv(n-6)$	$mv(n-9)$	$mv(n-12)$
$mv(n-7)$	$mv(n-3)$	$mv(n-2)$	$mv(n-4)$	$mv(n-10)$
$mv(n-5)$	$mv(n-1)$			

그림 1. 이미 처리된 인접한 움직임 벡터의 순서 (N=12)

그림 1 과 같이 움직임 벡터 예측 방식에 사용되는 이미 처리된 인접한 움직임 벡터의 순서에 따라 예측된 움직임 벡터는 다음과 같다

$$pred_mv(n) = \sum_{k=1}^N a(k)mv(n-k) \quad (1)$$

예측기를 설정하기 위해 명확하게 움직임 에지를 검출하고 그의 방향성을 추정하는 것은 방향성 정보를 이용하여 움직임 벡터를 예측하는 방식은 것은 매우 어려운 문제이다. 예를 들어, 기존의 문맥 기반의 적응 예측 방식인 기울기 보정 예측 방식(Gradient Adjusted Prediction)과 미디언 에지 검출 방식(Median edge Detection)은 움직임 배경(smooth) 영역에서 움직임 에지 영역까지 타당한 적응성을 얻기 위해 움직임 에지를 검색하여 전환하는 방식을 사용하였다. 그러나, 이러한 방식은 임의의 방향성을 갖는 움직임 에지 주위의 영역에서 넓은 예측 오류를 생성한다.

움직임 에지의 검색이나 방향의 명백한 추정 대신에, 최소 계승 방식을 기반으로 하는 움직임 예측 방식은 아래와 같은 학습창(training window) 안에 예측 계수를 지역적으로 최적화하여 움직임 에지 방향에 적응적이고 최적의 근사한 예측기를 설정한다.

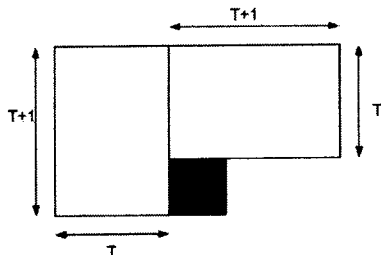


그림 2. 움직임 벡터 예측 계수를 최적화하기 위해 설정된 학습창 (training window)

그림 2 에서와 같이, 움직임 벡터 예측 계수를 최적

화하기 위해 설정된 학습창은 두 개의 직사각형 안에 $M=2T(T+1)$ 개의 이미 처리된 인접한 움직임 벡터들로 구성된다. 학습창을 $M \times 1$ 열 벡터 $\bar{y} = [X(n-1) \cdots X(n-M)]^T$ 로 명시하면 \bar{y} 를 예측하기 위한 인접한 움직임 벡터들은 $M \times N$ 형태인 다음과 같은 행렬로 주어진다.

$$C = \begin{bmatrix} mv(n-1-1) & \cdots & mv(n-1-N) \\ \vdots & \ddots & \vdots \\ mv(n-M-1) & \cdots & mv(n-M-N) \end{bmatrix} \quad (2)$$

식 (2)의 $mv(n-j-k)$ 는 $mv(n-j)$ 를 예측하기 위한 k 번째 이웃하는 움직임 벡터이다. 학습창내의 움직임 벡터 예측 계수는 최소 계승 최적화 과정을 통해 얻어진다

$$\min \|\bar{y} - C\bar{a}\|_2^2 \quad (3)$$

식(3)은 다음과 같은 해를 갖는다

$$\bar{a} = (C^T C)^{-1} (C^T \bar{y}) \quad (4)$$

이때, $\bar{a} = [a(1) \cdots a(N)]^T$ 은 최적화된 움직임 벡터 예측 계수 이다.

위의 결과는 전통적인 선형 예측 이론[8]에 의해서도 유도될 수 있다. 정적인 가우시안 분포(Stationary Gaussian process)에 대한 최소 평균 제곱 오류 예측 방식(MMSE prediction)은 2 차 통계적인 특성인 코베리언스(covariance)로 결정된다.

$$\bar{a} = R_{XX}^{-1} \bar{r}_X \quad (5)$$

$\bar{r}_X = [r_1 \cdots r_k \cdots r_N]$, ($1 \leq k \leq N$), $r_k = Cov\{mv(n)mv(n-k)\}$ 이고, $R_{XX} = [R_{kl}]$, $R_{kl} = Cov\{mv(n-k)mv(n-l)\}$ ($1 \leq k, l \leq N$) 을 나타낸다. 기하학적으로 식 (5)는 $mv(n-1) \cdots mv(n-N)$ 에 따라 펼쳐진 부분공간(subspace)상에 $mv(n)$ 을 정사영한 것이다. 그러나, 일반적으로 움직임 벡터는 정적인 가우시안 분포를 갖는다는 가정에 위반된다. 정적이지만(Nonstationary) 움직임 벡터를 다루기 위한 접근 방식은 움직임 벡터의 예측하는 과정에서 국부 통계적 특성을 추정하여 해결할 수 있으며 위에서 정의된 C 와 \bar{y} 를 이용하여 추정된 2 차 통계적인 특성, $(\hat{R}_{XX}, \hat{r}_X)$ 은 다음과 같이 주어진다.

$$\hat{R}_{XX} = \frac{1}{M} C^T C \quad \hat{r}_X = \frac{1}{M} C^T \bar{y} \quad (6)$$

식(6)을 식(5)에 대입하면, 식(4)를 얻는다.

3. 움직임 에지 방향성

움직임 배경(smooth) 영역에서 움직임 에지 영역에 이르기 까지 움직임 벡터 예측 방식의 적응성은 적응 예측 방식의 성능을 좌우한다. 임의적인 에지 방향은 에지 영역에서의 적응성을 떨어뜨리는 원인이 된다. 최적의 예측방식이 에지 방향에 다를 것이라는 생각은 쉬워 보임에도 불구하고 그에 대한 수행은 매우 어렵다. 앞에서 기술한 바와 같이, 에지를 검출하여 방향을 추정하는 것은 그 자신의 한계(예 robustness)를 가지고 있다. 그와는 반대로, 최소 계승 방식을 기반으로 하는 적응 움직임 예측 방식은 에지 방향성을 이용하여 방향성을 최적에 가깝게 적합화함으로써 움직임

임 벡터를 예측한다.

최소 계승 방식의 최적화 과정에서 설정된 학습 창안에 이미 처리된 인접한 움직임 벡터들은 움직임 에지 주변의 움직임 벡터들과 움직임 에지에서 떨어져 있는 움직임 벡터들로 분류된다. 움직임 에지에서 떨어져 있는 움직임 벡터들에서 행렬 C 는 종종 특이 행렬이 되어(not full-ranked) 최소 계승 방식의 최적화 과정은 유일한 해를 갖을 수 없다. 실제로 움직임 에지에서 떨어져 있는 움직임 벡터들에서의 최적 예측기 집합은 N 차 영역의 다중평면(hyperplane) $\sum_{k=1}^N a(k)=1$ 안에 존재한다. 반면에 움직임 에지 주변의 움직임 벡터들에서는 행렬 C 가 정칙행렬이 되므로(full-ranked) 해가 유일하게 존재한다. 그러므로, 움직임 에지 주변의 움직임 벡터들의 최적 예측기 집합은 다중평면 $\sum_{k=1}^N a(k)=1$ 의 부분 집합이 된다. 분류된 움직임 벡터들의 위와 같은 성질은 최소 계승 방식의 최적화 과정에서 움직임 에지 주변의 움직임 벡터들이 주도적인 역할을 담당하므로 에지 방향성을 가능하게 한다.

학습창을 기반으로 하는 최소 계승 방식의 최적화 과정은 움직임 에지를 검출하는 단계가 없어도 에지 주변의 움직임 벡터들에 대한 최적의 예측 계수를 발견할 수 있는 편리한 방식을 제공한다. 또한, 대부분의 처리하고자 하는 블록의 움직임 벡터가 에지 주변의 움직임 벡터들과 동일한 움직임 에지에 포함되기 때문에 에지 주변의 움직임 벡터들에 대해 최적화된 예측 계수를 처리하고자 하는 블록의 움직임 벡터에 적용시킬 수 있다.

4. 제안된 방식

에지 방향성을 이용하는 예측방식의 주요한 단점은 막대한 계산량이다. 최소 계승 방식의 최적화 과정의 문제점은 식(6)의 코베리언스 행렬(covariance matrix) $C^T C$ 을 연산하는 것이다. 이러한 문제점을 해결하기 위해 본 논문에서는 일정한 성능을 유지하면서 움직임 벡터들 중 일부분만을 최소 계승 방식의 최적화 과정에서 사용함으로써 계산량을 줄이는 방식에 대해 제안한다.

움직임 에지 방향성을 이용하는 예측 방식은 다음의 두 가지 면을 바탕으로 한다. 첫째, 움직임 에지 주변의 움직임 벡터들을 최적화하는 예측 계수는 같은 움직임 에지를 따르는 인접한 움직임 벡터에 대해서도 적용 가능하다. 둘째, 하나의 움직임 에지에서 최적화된 예측기 집합은 배경(smooth)영역의 최적화된 예측기 집합의 부분 집합이다. 움직임 에지 방향성을 이용하는 예측 방식의 이러한 성질은 움직임 에지 주변에 움직임 벡터에서 최적화된 예측 계수를 저장하여 다음 번 움직임 에지를 만날 때까지 반복적으로 사용하는 것을 가능하게 한다. 그러므로, 최소 계승 방식의 최적화 과정은 움직임 벡터 단위가 아닌 움직임 에지 단위로 수행된다.

본 논문에서는 움직임 예측 방식을 움직임 에지 단위로 수행하기 위해 다음과 같은 전환 방식을 제안한

다. 예측 오류 $e(n) = mv(n) - m\hat{v}(n)$ 의 크기가 이미 설정된 임계치 이상이면 예측 계수를 개선하기 위해 최소 계승 방식의 최적화 과정이 수행되고 이하이면 다음 움직임 벡터를 예측하기 위해 저장된 예측 계수를 사용한다.

5. 실험 결과

제안된 적응 움직임 예측 방식을 실험하기 위해 학습창은 $M=4, N=4$ 로 설정하였고, 예측 오류와 비교하여 에지 주변의 움직임 벡터들과 에지에서 떨어져 있는 움직임 벡터들로 나누는 임계치 값은 실험을 통해 91로 설정하였다



그림 4. 실험에 사용된 학습창

제안된 적응 움직임 예측 방식은 JVT JM3.9 의 미디언 예측 방식과 객관적, 주관적인 성능을 비교하기 위해 여러 동영상에 다양한 비트율 및 프레임 율에서 실험하였다. 본 논문에서는 QCIF 크기를 갖고 30 frames/sec 인 300 frames 의 동영상이 입력 되었을 때, 10 frames/sec 로 압축하여 100 frame 을 출력된 Foreman 과 News 영상에 대해 기술한다

제안된 방식의 객관적인 성능을 측정하기 위해 PSNR(Peak to Signal to Noise Ratio)이 사용되었으며, 이는 화소당 8bit 로 구성된 $M \times N$ 영상 크기에 대해 다음과 같이 정의된다.

$$PSNR = 10 \log \frac{MN \times 255^2}{\|f - \hat{f}\|^2} \quad (7)$$

식 (7)의 $\|\cdot\|$ 은 유클리디안 노름(Euclidean norm)을 의미하며, f 및 \hat{f} 은 각각 원 영상과 복원 영상을 의미한다.

제안된 방식의 계산량을 측정하기 위해 실험 영상을 각 방식에 적용하여 프레임당 부호화되는 시간(E.T.F: Encoding Time per Frame)을 이용하였다.

$$saving(\%) = \frac{E.T.F(JM3.9) - E.T.F(JM3.9 + LS)}{E.T.F} * 100$$

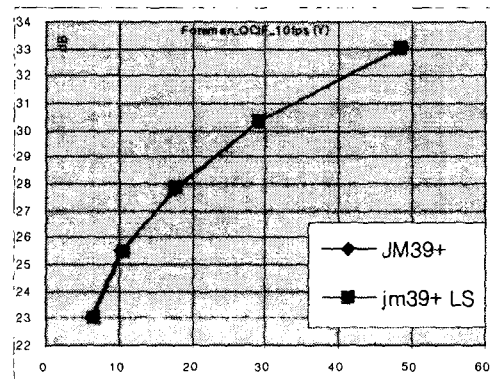


그림 5. Foreman 영상의 비트율에 따른 PSNR 비교

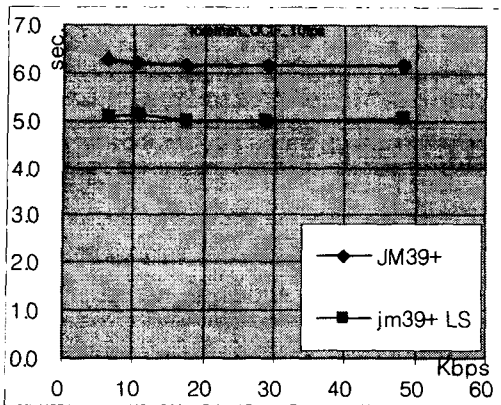


그림 6. Foreman 영상의 비트율에 따른 계산량 비교

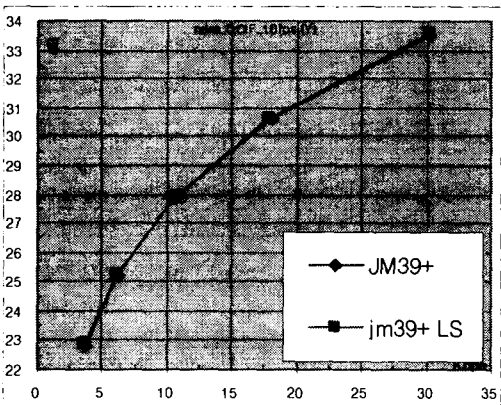


그림 7. News 영상의 비트율에 따른 PSNR 비교

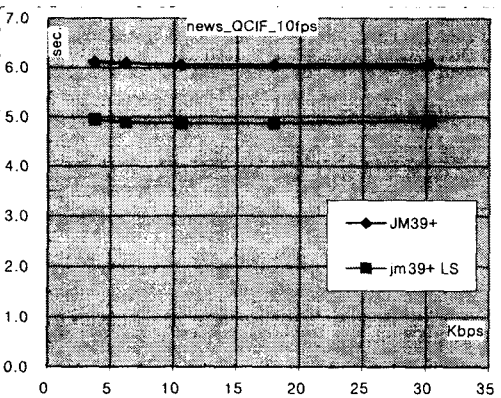


그림 8. News 영상의 비트율에 따른 계산량 비교

Foreman 과 News 영상에서 비트율에 따른 객관적인 성능인 PSNR 을 나타내는 그림 5 와 7 은 제안된 방식이 비트율과 무관하게 JM3.9 의 미디안 예측 방식과 매우 유사하게 나타낸다. 반면에, 계산량을 나타내는 그림 6 과 8 는 제안된 방식이 JM3.9 의 미디안 예측 방식보다 감소한 것을 나타내었다. 이는 움직임 에지 방향성을 기반으로 움직임 벡터를 적응적으로 개선한 결과이다. 이는 움직임 에지 주변의 움직임 벡터들을 최적화하는 예측 계수가 같은 움직임 에지를 따르는 인접한 움직임 벡터에 대해서도 적용 가능하고, 하나의 움직임 에지에서 최적화된 예측기 집합은 배정

(smooth) 영역의 최적화된 예측기 집합의 부분 집합이라는 최소 계승 방식의 에지 방향성에 기인한다

6. 결론

본 논문에서는 예측될 움직임 벡터의 이미 처리된 인접한 블록의 움직임 벡터를 움직임 에지 주변의 움직임 벡터와 움직임 에지에서 떨어져 있는 움직임 벡터로 나누어 에지 주변의 움직임 벡터에서만 최소 계승 예측 방식을 사용하여 예측기를 개선한 적응 움직임 벡터 예측 방식을 제안하였다. 그러므로, 제안된 방식은 에지 주변에 움직임 벡터에서 최적화된 예측 계수를 저장하여 다음 번 에지를 만날 때까지 반복적으로 사용하여 최소 계승을 이용한 예측 방식의 방대한 계산량을 줄일 수 있었다.

제안된 방식의 성능을 향상시키기 위해 학습창을 구성하는 인접한 블록의 움직임 벡터 수를 결정하는 방법과 임의적으로 임계치를 설정한 종전의 실험방식에서 주어진 영상의 정보를 이용하여 임계치를 보정하는 방법을 연구 하고 있다.

참고 문헌

- [1] Khalid Sayood, *Introduction to Data Compression*, Morgan Kaufmann, 2000
- [2] S. Chakravarti, T. P. Jung, S. C. Ahalt, and A.K. Krishnamurthy, "Comparison of Prediction Methods for Differential Image Processing Application," in *Proc. Int. Conf. Systems Engineering*, pp. 210 -213, 1991
- [3] S. D. Kim and J. B. Ra, "An Efficient Motion Vector Coding Scheme Based on Minimum Bitrate Prediction," *IEEE Trans. Image Processing*, vol. 8, pp.1117-1120, Aug.1999
- [4] D. H. Kang, J.H. Choi, Y.H. Lee, and C. Lee, "Application of a DPCM System with Median Predictors for Image Coding," *IEEE Trans. Con. Elec.* Vol. 38, pp. 429-435, Aug.1992
- [5] X. Wu and N. Memon, "Context-based adaptive lossless image coding," *IEEE Trans. Commun.*, vol.45, pp. 437 - 444, Apr. 1997
- [6] M. Weinberger, G. Seroussi and G. Sapiro, "The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS," *IEEE Trans. Image Processing*, vol. 9, pp. 1309-1324, Aug. 2000
- [7] X. Li and M. Orchard, "에지 directed prediction for lossless compression of natural images," in *Proc. Int. Conf. Image Processing*, vol. 4, pp.. 58-62, Oct.1999
- [8] N. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Englewood Cliffs, NJ: Prentice-Hall, 1984
- [9] X. Li and M. Orchard, "에지 directed prediction for lossless compression of natural images," in *Proc. Int. Conf. Image Processing*, vol. 10, pp.. 813-817, Jun.2001
- [10] H. Ye, G. Deng, and J. Devlin, "Least squares approach for lossless image coding," in *Proc. 5th Int. Symp. Signal Processing Applications*, vol. 1, pp. 63-66, 1999