

정합 오차 기준을 확장한 제한된 1비트 변환 알고리즘

*이상구 **윤장혁 ***정제창

한양대학교 전자컴퓨터통신공학과

*esg803@hanyang.ac.kr **yjh2748@hanyang.ac.kr ***jjeong@hanyang.ac.kr

Constrained One-Bit Transform using Extension of Matching error criterion

*Lee, Sang-gu **Yun, Jang-hyeok ***Jeong, Je-chang

Department of Electronics and Computer Engineering, Hanyang University

요약

본 논문은 제한된 1비트 변환 (Constrained One-Bit Transform : C1BT) 알고리즘의 정합 오차 기준의 동적 범위를 확장하는 알고리즘을 제안하였다. C1BT는 정합 오차 기준으로 SAD (Sum of Absolute Differences)를 사용하지 않고 CNNMP (Constrained Number of Non-Matching Points)를 사용하여 하드웨어 구현을 용이하게 하고 속도를 대폭 향상시켰다. 이는 기존의 움직임 예측 방법인 전역 탐색 알고리즘 (Full Search Algorithm: FSA)과 비교하여 연산량을 크게 줄였으나 움직임 예측의 정확도를 현저히 감소시켰다. 이 점을 개선하기 위해 이 논문에서는 C1BT의 정합 오차 기준을 확장하여 움직임 예측의 정확도를 높이는 알고리즘을 제안하였다. 기존의 C1BT와 제안하는 알고리즘을 비교한 결과에서 제안하는 알고리즘이 기존의 C1BT에 비해 움직임 예측의 정확도의 기준인 PSNR 측면에서 더 우수한 성능을 보였다.

1. 서론

유한한 통신용량과 대역폭을 가지고 있는 멀티미디어 환경 내 영상 통신에서는 압축 기술이 필수적이다. 이를 위해 영상 통신에서는 움직임 예측 알고리즘을 사용한다. 이 알고리즘은 영상 내 연속되는 프레임 사이의 시간적 중복성을 제거하여 기존의 송신 데이터양을 획기적으로 감소시킨다. 가장 간단하고 보편적인 움직임 예측 기법은 블록 매치 알고리즘 (Block Matching Algorithm : BMA) 이다. 이는 각 영상 내 프레임을 일정한 크기의 정사각형 블록으로 분할하고 압축하고자 하는 프레임 내 각 블록들과 정합 오차가 가장 작은 정합블록을 이전 프레임에서 지정된 탐색 범위 내에서 찾아내는 방법이다. 이후 탐색된 정합 오차가 가장 작은 정합블록의 위치로부터 움직임 벡터 (motion vector)가 결정된다. 이와 같이 프레임 내 각 정합블록과 정합 오차를 가장 작게 하여 압축 효율을 최대로 높일 수 있는 움직임 벡터를 제공하는 것이 블록 매칭 알고리즘이다[1][3][4].

최적의 움직임 벡터를 찾는 알고리즘은 여러 가지 방법이 있는데 이 중 가장 적합한 방법으로 알려져 있는 알고리즘은 전역 탐색 알고리즘 (Full Search Algorithm : FSA)이다. 이 알고리즘은 현재 프레임 내 정합블록과 가장 정합오차가 적은 정합블록을 찾기 위해 참조 프레임의 탐색 영역 내 모든 정합후보블록들을 탐색하며 정합 오차 기준으로 SAD (Sum of Absolute Difference)를 사용한다. 그러나 전역 탐색 알고리즘은 탐색 영역 내에서 모든 후보블록들을 탐색하면서 각 후보블록마다 SAD를 계산하여 방대한 연산량을 요구한다. 이는 실시간 동영상 부호화기의 구현에 활용하기 어려움을 시사한다[1][2][5].

이 같은 전역 탐색 알고리즘의 문제점을 해결하기 위해서 다양한 고속 블록 정합 알고리즘들이 제안되고 있다. 이 중 하나가 정합 오차 기준을 SAD 대신 NNMP (Number of Non-Matching Points)를 사용하는 이진 블록 매칭 움직임 예측 알고리즘이다. 이 알고리즘은 영상 내 프레임을 이루고 있는 8비트 픽셀 (pixel)을 이진 변환하여 움직임을 예측한다. 이에 따라 이진 변환된 프레임들을 불리언 (Boolean) 연산을 사용하면서 최적의 움직임 벡터를 찾고 하드웨어 구현 시 고속 움직임 예측이 가능하게 하였다. 각 픽셀의 이진 변환을 위한 알고리즘으로는 1비트 변환^[1] (One-Bit Transform : 1BT) 알고리즘, 2비트 변

환^[2] (Two-Bit Transform : 2BT) 알고리즘 그리고 제한된 1비트 변환^[3] (Constrained One-Bit Transform : C1BT) 알고리즘이 있다.

먼저 제안된 1BT의 경우 원 영상 프레임을 이진 변환한 평면 1개만 사용하면서 기존 프레임의 정보를 많이 잃어버려 정확하지 않은 움직임 벡터를 구할 확률이 증가하였다. 이에 따라 움직임 예측의 정확도가 전역 탐색 알고리즘보다 매우 저하되는 단점을 가지게 되었다. 그래서 2BT와 C1BT는 원 영상 프레임의 이진 변환된 평면을 2개 사용하면서 1BT의 움직임 예측의 정확도를 개선하기 위하여 제안되었다[2][3][4].

본 논문의 구성은 다음과 같다. 제 2장에서는 기존의 알고리즘들인 1BT, 2BT, C1BT에 대해 자세히 기술하고 제 3장에서는 본 논문에서 제안하는 정합 오차 기준을 확장하여 움직임 예측의 정확도를 개선하는 알고리즘에 대해 소개한다. 제 4장에서는 실험결과를 통해 기존의 알고리즘과 제안하는 알고리즘의 PSNR을 비교하고 제 5장에서는 결론을 맺는다.

2. 기존의 알고리즘들

2.1 1비트 변환을 사용하는 움직임 예측 알고리즘

1비트 변환을 사용하는 움직임 예측 알고리즘은 원 영상 프레임을 이진 변환하기 전에 식 (1)에서 정의된 17×17 대역 통과 필터 커널 K 를 이용하여 새로운 영상 프레임을 만들어낸다.

$$K(i, j) = \begin{cases} 1/25, & \text{if } i, j \in [0, 4, 8, 12, 16] \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

이 후 필터가 이용된 프레임 I_f 과 원본 영상 프레임 I 의 각 화소를 비교하여 식 (2)와 같이 원본 영상 프레임을 1비트 변환을 사용하여 이진 변환한다.

$$B(i, j) = \begin{cases} 1, & \text{if } I(i, j) \geq I_F(i, j) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

식 (2)에서 $I(i, j)$, $I_F(i, j)$ 는 각각 현재 프레임 내 (i, j) 위치의 화소값, 필터가 적용된 프레임 내 (i, j) 위치의 화소값을 나타낸다. 그 후 기존의 움직임 예측 기법인 전역 탐색 알고리즘과 동일하게 지정된 탐색 범위 내에서 최적의 움직임 벡터를 찾게 된다. 하지만 적합한 움직임 벡터를 결정할 때 기존 알고리즘에서 사용하는 정합 오차 기준인 SAD와 달리 1비트 변환 알고리즘에서는 불리언 연산자를 사용하는 NNMP (Number of Non-Matching Points)를 이용하며 식 (3)과 같이 정의된다.

$$NNMP_{1BT}(m, n) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{B^t(i, j) \oplus B^{t-1}(i+m, j+n)\} \quad -s \leq m, n \leq s-1 \quad (3)$$

식 (3)에서 $B^t(i, j)$, $B^{t-1}(i, j)$ 는 각각 현재 프레임의 1비트 변환된 프레임, 이전 프레임의 1비트 변환된 프레임이다. 또 N 은 정합블록 크기이고 s 는 탐색 범위의 크기를 결정한다. (m, n) 은 탐색 범위 내 후보 정합블록들의 변위를 나타낸다.

2.2 2비트 변환을 사용하는 움직임 예측 알고리즘

2BT를 사용하는 움직임 예측에서는 이전 변환된 평면을 1개만 사용하여 원 영상 프레임의 정보를 많이 잃어버리면서 생기는 1BT의 부정확한 움직임 예측을 개선하기 위하여 제안된 움직임 예측 알고리즘이다. 2BT에서는 식 (4)와 같이 2BT에서는 국부 주변 (local surrounding) 화소들의 평균, 표준편차와 근사치 표준편차 (approximate standard) 값들이 사용된다.

$$\begin{aligned} \mu &= E[I_{tw}] \\ \sigma^2 &= E[I_{tw}^2] - E^2[I_{tw}] \\ \sigma_a &= 15 + 0.0125\sigma^2 \end{aligned} \quad (4)$$

식 (4)에서 I_{tw} 는 2비트 변환되는 8×8 정합블록 주변의 40×40 크기의 국부 문턱 윈도우 (local threshold window) 안의 화소값을 의미한다. 이를 통해 구해진 평균과 근사치 표준편차는 원 영상 프레임을 이전 변환하기 위해 사용된다. 식 (5)와 같이 2BT를 통해 이전 변환되는 두 비트평면은 원 영상 프레임 I 의 화소와 국부 문턱 윈도우 안의 화소들의 평균 및 근사치 표준편차를 비교하여 얻어진다.

$$\begin{aligned} B_1(i, j) &= \begin{cases} 1, & \text{If } I(i, j) \geq \mu \\ 0, & \text{otherwise} \end{cases} \\ B_2(i, j) &= \begin{cases} 1, & \text{if } I(i, j) \geq \mu + \sigma_a \text{ or } I(i, j) \leq \mu - \sigma_a \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (5)$$

식 (5)에서 $B_1(i, j)$, $B_2(i, j)$ 는 각각 2비트 변환된 프레임의 첫 번째 비트 평면, 두 번째 비트 평면을 의미한다. 2BT에서 이용되는 정합 오차 기준인 $NNMP_{2BT}$ 는 식 (6)과 같이 정의된다.

$$NNMP_{2BT}(m, n) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{B_1^t(i, j) \oplus B_1^{t-1}(i+m, j+n)\} \oplus \{B_2^t(i, j) \oplus B_2^{t-1}(i+m, j+n)\} \quad -s \leq m, n \leq s-1 \quad (6)$$

식에서 $B_n^t(i, j)$ 는 2비트 변환된 현재 프레임의 n 번째 비트 평면이고 $B_n^{t-1}(i, j)$ 는 2비트 변환된 이전 프레임의 n 번째 비트 평면이다.

2BT 역시 1BT와 같이 전역 탐색 알고리즘보다 움직임 예측이 다소 부정확하지만 이전 변환된 비트 평면을 한 개 더 이용하면서 정확도가 개선되었다.

2.3 제한된 1비트 변환을 사용하는 움직임 예측 알고리즘

2BT와 마찬가지로 C1BT를 사용하는 움직임 예측 알고리즘도 1BT의 단점인 움직임 예측의 부정확성을 개선하기 위해 제안된 움직임 예측 알고리즘이다. C1BT에서도 2BT와 동일하게 두 개의 비트평면을 사용하나 평균과 근사치 표준편차를 이용하지 않고 식 (7)과 같이 문턱 값 D 를 이용하여 원 영상 프레임을 이전 변환한다.

$$CM(i, j) = \begin{cases} 1, & \text{if } |I(i, j) - I_F(i, j)| \geq D \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

식에서 알 수 있듯이 2BT와 크게 다른 점은 영상 프레임을 이전 변환할 때 특별한 연산과정 없이 1BT에서 사용하는 필터를 적용한 프레임 I_t 와 문턱 값 D 를 이용하여 두 번째 비트 평면인 제한된 마스크 (Constraint Mask : CM)를 얻는다는 점이다. 제한된 마스크를 사용하는 C1BT는 원 영상 프레임 내 두 화소값이 실제로 비슷하나 1비트 변환되면서 두 화소가 서로 다른 값이 되면서 생기는 움직임 예측의 부정확성을 개선한다. C1BT에서 사용되는 정합 오차 기준인 CNNMP (Constrained NNMP)는 식 (8)과 같이 정의된다.

$$CNNMP(m, n) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{[CM^t(i, j) \oplus CM^{t-1}(i+m, j+n)] \cdot [B^t(i, j) \oplus B^{t-1}(i+m, j+n)]\} \quad -s \leq m, n \leq s-1 \quad (8)$$

여기서 $CM^t(i, j)$, $CM^{t-1}(i, j)$ 는 제한된 1비트 변환된 이전 프레임의 제한된 마스크, 제한된 1비트 변환된 현재 프레임의 제한된 마스크를 의미한다. C1BT 역시 전역 탐색 알고리즘보다 움직임 예측의 정확도가 떨어지나 1BT의 문제점을 개선하며 정확도가 향상되었다.

3. 제안하는 알고리즘

이번 장에서는 C1BT의 움직임 예측을 개선하기 위한 알고리즘에 대해 기술하고자 한다. 기존의 C1BT 알고리즘의 정합 오차 기준인 CNNMP의 동적 범위는 정합블록의 크기가 $N \times N$ 일 때 $0 \leq CNNMP \leq N \times N$ 으로서 기존의 정합 오차 기준인 SAD ($0 \leq SAD \leq 256 \times N \times N$)의 동적 범위를 $1/256$ 감소시킨다. 이에 따른 정보 손실로 인해 FSA에 비해 움직임 예측이 다소 부정확해지는데 이 점을 개선하고자 CNNMP의 식 (9)을 다음과 같이 수정하였다.

$$\begin{aligned} NNMP_{C1BT,1} &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{CM^t(i, j) \cdot [B^t(i, j) \oplus B^{t-1}(i+m, j+n)]\} \\ NNMP_{C1BT,2} &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{CM^{t-1}(i, j) \cdot [B^t(i, j) \oplus B^{t-1}(i+m, j+n)]\} \\ NNMP_{C1BT,3} &= 2 * NNMP_{C1BT,1} + NNMP_{C1BT,2} \end{aligned} \quad (9)$$

식 (9)와 같이 기존의 정합 오차 기준인 CNNMP의 식에서 AND 연산을 현재 프레임의 제한된 마스크 값과 이전 프레임의 제한된 마스크 값에 분배하고 OR 연산을 덧셈 연산으로 수정하면서 현재 프레임의 제한된 마스크 값에 가중치를 주었다. 이에 따라 $NNMP_{C1BT}$ 의 동적 범위가 $0 \leq NNMP_{C1BT} \leq 3 \times N \times N$ 으로 기존 CNNMP에 비해 3배 늘어났다.

4. 실험결과

실험은 움직임 특성이 서로 다른 6개의 CIF(352 x 288) 영상을 이용하여 움직임 예측을 측정하기 위한 알고리즘이므로 휘도 성분만을 고려하였다. 실험조건으로 탐색 범위와 움직임 블록 크기를 각각 $\pm 16, 16 \times 16$ 으로 결정하고 정합블록 탐색 순서는 나선형(spiral) 기법을 이용하였다. 또한 제안하는 알고리즘과 기존 알고리즘의 성능 비교를 위해 화질 비교에 자주 사용되는 복원 영상의 PSNR (Peak Signal-to-Noise Ratio)를 사용하였다.

	C1BT	D = 8	D = 9	D = 10	D = 11	D = 12
Sequence	PSNR (dB)	PSNR (dB)	PSNR (dB)	PSNR (dB)	PSNR (dB)	PSNR (dB)
children	28.51	28.51	28.55	28.58	28.59	28.62
flower	26.48	26.52	26.53	26.54	26.55	26.56
sea_football	23.22	23.23	23.28	23.31	23.35	23.38
singer	36.76	36.82	36.86	36.88	36.89	36.89
stefan	23.14	23.16	23.18	23.19	23.2	23.21
table	30.5	30.54	30.57	30.58	30.58	30.58
Average	28.10	28.13	28.16	28.18	28.19	28.21

표 1. 기존 C1BT와 PSNR 비교

	1BT	2BT	C1BT	Proposed	FSA
Sequence	PSNR (dB)	PSNR (dB)	PSNR (dB)	PSNR (dB)	PSNR (dB)
children	28.21	28.18	28.51	28.62	29.18
flower	26.42	26.32	26.48	26.56	26.68
sea_football	22.84	23.16	23.22	23.38	24.08
singer	36.29	36.26	36.76	36.89	37.11
stefan	22.94	23.22	23.14	23.21	23.9
table	29.87	30.24	30.5	30.58	31.46
Average	27.76	27.9	28.10	28.21	28.74

표 2. 기존 이진 변환 움직임 예측 알고리즘과 전역 탐색 알고리즘 PSNR 비교

표 1은 제안하는 알고리즘에서 최적의 문턱 값을 결정하기 위해 문턱 값을 변화시키면서 PSNR을 측정한 결과이다. 실험결과 문턱 값이 12일 때 제안하는 알고리즘이 가장 좋은 성능을 보였고 기존의 알고리즘보다 PSNR이 0.11 dB 상승하였다. 표 2는 기존의 이진 변환 움직임 예측 알고리즘 1BT, 2BT, C1BT 그리고 전역 탐색 알고리즘인 FSA와 PSNR을 비교한 결과이다. 제안하는 알고리즘은 가장 좋은 성능을 보인 문턱 값 12를 사용하였다. 표와 같이 기존의 이진 변환 움직임 예측 알고리즘보다 월등한 성능을 보여준다. 특히 1BT, 2BT, C1BT에 비해 PSNR이 0.44dB, 0.30dB, 0.11dB 상승하였다. 하지만 전역 탐색 알고리즘에 비해 움직임 예측의 정확도가 다소 부족하는데 이는 정합 오차 기준의 동적 범위가 감소하면서 생기는 정보 손실을 확실하게 개선하지 못하였음을 시사한다.

5. 결론

본 논문에서는 정합 오차 기준의 동적 범위를 확장시키면서 기존의 이진 변환 움직임 예측 알고리즘인 C1BT와 비교하여 움직임 예측의 정확도를 개선하는 알고리즘을 제안하였다.

실험결과를 통해 제안하는 알고리즘은 기존의 알고리즘보다 PSNR 측면에서 0.11 dB 상승하였고 다른 이진 변환 알고리즘인 1BT, 2BT에 비해 0.44 dB, 0.30dB 상승하였다. 이는 고화질 디지털 영상의 구현에 적합한 알고리즘이 될 수 있겠다.

6. 감사의 글

이 논문은 2013년도 정부(교육부)의 재원으로 한국연구재단

의 지원을 받아 수행된 기초연구사업임 (NRF-2011-0011312)

참고문헌

[1] B. Natarajan, V. Bhaskaran, and K. Konstantinides, "Low-complexity Block-based Motion Estimation via One-Bit Transforms," IEEE Trans. Circuits Syst. Video Technol., vol.7,no.5,pp.702-706, Aug. 1997

[2] A. Erturk and S. Erturk, "Two-Bit Transform for Binary Block Motion Estimation," IEEE Trans. Circuits Syst. Video Technol., vol.15, no.7, pp.938-946, July 2005.

[3] O. Urhan and S. Erturk, "Constrained one-bit transform for low complexity block motion estimation," IEEE Tans. Circuits and Syst. Video Technol., vol.17, no.4, pp.478-482, Apr. 2007.

[4] O. Urhan, "Constrained one-bit transform based fast block motion estimation using adaptive search range," IEEE Tans. Consumer Electron, vol.56, no.3, pp.1868-1871, Aug. 2010.

[5] C. Choi and J. Jeong, "Enhanced two-bit transform based motion estimation via extension of matching criterion," IEEE Tans. Consumer Electron, vol.56, no.3, pp.1883-1889, Aug. 2010.