

# 움직임탐색에 있어서 DCT를 이용한 탐색점 배치

송지연, 김준한, 김상곤, 윤영우  
영남대학교 컴퓨터공학과

Search point displacement using DCT for Motion Estimation

Ji-Youn Song, Jun-Han Kim, Sang-Gon Kim, Young-Woo Yoon  
Dept. of Computer Engineering, Yeungnam University

## 요 약

블록의 DCT계수를 이용하여 동영상 움직임 탐색(motion estimation)에서 탐색점 수를 결정하는 선택적 탐색 기법을 제안한다.

DCT(Discrete Cosine Transform)연산과 블록정합 기법인 3단계 탐색기법을 기반으로 한다.

DCT연산 결과인 블록의 주파수성분에 가중치레이블을 적용하여 고주파성분을 많이 가진 블록일수록 탐색점의 수를 증가시켜 국부극소에 빠질 확률을 줄여 화질을 개선한다. MP@ML 영상 데이터에 대해 제안된 기법을 사용하여 시뮬레이션하였다.

제안하는 기법은 3단계 탐색기법에 비하여 연산량은 증가하나, 연산량 증가에 비하여 좋은 화질을 제공한다.

## 1. 서 론

디지털 영상 데이터의 압축은 시간적, 공간적인 중복성을 줄이는 방법과 발생부호의 통계적 특성을 이용하여 줄이는 방법으로 크게 3가지 방법을 이용하고 있다. 시간적인 중복성을 줄이는 대표적인 방법중 하나가 바로 움직임 추정방법이다.

동영상에서 움직임추정하는 과정이 전체 부호화 시간의 약 30%로 많은 시간이 소요되기 때문에, 빠르고 효과적인 움직임 추정방법을 찾아내는 것은 중요하다.

이것을 위해 가장 많이 이용되는 방법이 간단하고 하드웨어 구현이 용이한 블록정합 알고리즘(BMA)이다.[1] 블록정합 알고리즘 가운데서 기본적인 전역탐색(Full Search, FS)은 주어진 탐색영역내의 모든 후보점을 탐색점으로 하여 정합에러가 적은 후보점을 이용, 움직임벡터를 결정하는 것으로 성능은 가장 우수하나 계산량이 많아 시간이 많이 소요된다.

따라서, 수행시간 단축이나 성능 개선을 목표로 여러 알고리즘이 제안되고 있다.

그 중 3단계 탐색(Three Step Search, TSS)은 탐색점의 수가 적기 때문에 수행속도는 빠르지만, 정합에러의 국

부극소점(local minimum)에 수렴하는 단점이 있어 전역탐색에 비해 성능이 크게 떨어진다.

블록정합 기법에서 탐색점의 수와 화질은 비례한다.

탐색점이 많으면 연산량은 증가하지만 국부극소에 빠질 확률이 줄어들어 화질 개선을 기대할 수 있다.

본 논문은 공간주파수가 높을수록 단위면적당 국부극소점의 수가 많아지고 따라서 국부극소점에 수렴할 확률이 높아짐에 착안하여 영상의 공간주파수 성분을 분석하여 고주파성분을 많이 가진 블록에 탐색점을 상대적으로 많이 배치함으로써 국부극소에 빠질 확률을 줄여 화질을 개선하는 방법을 제안한다.

영상의 공간주파수는 이전 I-프레임의 DCT계수를 참조하여 산출한다.

## 2. 블록 정합기법

블록 정합기법은 한 블록내의 모든 화소들이 같은 방향으로 평행이동한다는 가정아래 후보 블록을 선택하여 블록내 화소값의 정합에러를 비교하여 가장 에러가 적은 블록을 선택하는 기법이다.

대표적인 블록 정합기법에는 전역탐색기법, 선택적 탐색기법, 화소 간척기법으로 분류된다.

전역탐색기법은 일정한 탐색범위내에서 모든 화소에 대한 블록을 순차적으로 비교하는 방법으로 비교적 정확한 움직임 벡터를 구할 수 있지만 계산량이 많은 단점을 가진다.

이를 개선하기 위해 제안된 선택적 탐색기법중 삼단계 탐색기법은 동일한 탐색거리에 위치한 8방향의 최초 후보점들을 선택하여 정합에러를 계산하고 이 중에서 최소의 정합에러를 가지는 후보점을 하나 선택한 다음, 탐색거리를 다시 반으로 줄여 탐색거리가 1이 될 때까지 동일한 국부탐색 과정을 반복하여 최종 움직임벡터를 찾아낸다. 전역 탐색기법에 비하여 계산량은 줄지만 국부극소에 빠져 움직임 추정을 정확히 할 수 없다.[2]

블록정합의 판단 기준으로 평균절대오차(Mean Absolute Difference: MAD)와 평균자승오차(Mean Squared Difference :MSD)가 많이 사용된다.[3]

MAD가 MSD에 비하여 계산량이 적고 하드웨어 구현이 용이하므로 본 논문에서는 MAD를 사용한다.

$$MAD(x, y) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |X_i(t, j) - X_{i-1}(t+k, j+1)| \quad (식 1)$$

MN : 블록의 화소수, (i,j) : 화소의 좌표,  $X_i$  : 현재 프레임,  $X_{i-1}$  : 이전 프레임, (i+k, j+1) : 탐색영역의 탐색점의 좌표

### 3. DCT(Discrete Cosine Transform)

DCT는 직교변환의 하나로서 영상 정보 특유의 인접화소간의 높은 상관성을 이용하여 영상 신호를 특정 좌표축으로 집중시켜 영상 신호의 중복성을 줄여 줌으로서 정보 압축을 실현한다. 하나의 영상 프레임에서 DCT를 하는 경우, 먼저 그 프레임을 작은 크기의 N×N 화소 블록들로 분할한 후, 각 블록에 대해 2차원 DCT처리를 한다[4]. 8×8 블록의 입력 픽셀 f(x,y)에 대한 2차원 DCT 변환계수는 다음과 같다.

$$F(\mu, \nu) = \frac{C(\mu)}{2} \frac{C(\nu)}{2} \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos((2x+1)\mu\pi/16) \cos((2y+1)\nu\pi/16) \quad (식 2)$$

( $\mu = 0$  일 때  $C(\mu)=1/\sqrt{2}$ ,  $\mu \neq 0$  일 때  $C(\mu)=1$ )

한 블록당 1개의 DC계수와 63개의 AC계수가 얻어진다. DC는 8×8 화소의 평균값을 나타낸다.

AC값은 왼쪽에서 오른쪽으로 진행함에 따라 고주파의 수평 주파수 성분을, 위에서 아래로 나아갈수록 고주파의 수직 주파수 성분을 많이 포함하게 된다.

### 4. 제안 알고리즘

제안하는 기법은 다음 2단계로 구성된다.

- (가) 1단계 : 공간주파수를 나타내는 지수 r의 산출
- (나) 2단계 : 지수 r에 비례하여 탐색점 배치

#### 4.1. 공간주파수를 나타내는 지수 r 산출

##### 4.1.1. 화면내 공간주파수의 특성

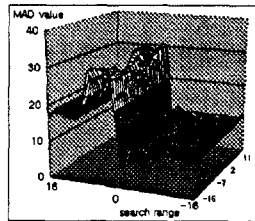
화면 전체가 유사한 밝기의 같은 색이라면 밝기는 다소 달라도 인접한 화소들이 같은 색이어서, 백과 적의 경우처럼 색이 완전히 다를때보다 인접한 값들을 유추하기가 훨씬 쉽다. 이러한 경우를 “그림이 매우 평활한 경우”라고 한다.

그림이 평활하고 천천히 변화할 때 “공간주파수가 낮다”고 말한다. 이와 반대로 체크무늬처럼 그림이 세밀하고 변화가 많을 때 “공간주파수가 높다”고 말한다. [5]

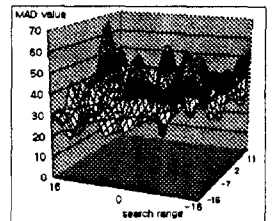
일반적으로 공간주파수가 낮은 화면은 탐색점간의 MAD차이가 작아서 탐색점간의 거리가 멀어도 무방하며 반대로 공간주파수가 높은 화면은 탐색점간의 MAD차이가 크기 때문에 탐색점간의 거리가 멀면 국부극소에 빠지기 쉽다는 직관적인 지식을 실험을 통해 증명한다.

그림1은 MPEG 영상 352×288 크기의 4:2:0 Stefan의 첫번째 프레임에 대하여 모의실험 한 결과이다.

영상이 복잡하여 공간주파수가 높은 관중석부분의 한 픽셀과 비교적 영상이 평활하여 공간주파수가 낮은 ground부분의 한 픽셀을 탐색창의 크기를 -16 ~ +16 으로 하여 모의실험하였다. 결과는 그림1과 같다.



(a) ground 부분



(b) 관중석 부분

그림1. Stefan #1의 MAD 분포도

그림1 에서 공간주파수가 높은 (b)는 공간주파수가 낮은 (a)에 비하여 인접 영역의 MAD차가 크므로 동일한 탐색점 배치시 국부극소에 빠질 확률이 높다.

#### 4.1.2. 지수 r의 산출

영상을 8×8의 블록으로 분할한 다음, 영상의 공간주파수를 나타내는 DCT계수를 구한다. 그 다음 주파수에 비례하는 가중치(weight)를 다음 식에 따라 각 블록의 DCT계수에 부여한다.

표 1. 가중치 테이블 A(i, j)

		1 : x 좌표의 가중치						
		계수화수 → 고주파수						
j · y 좌 표 의 가 중 치	0	1	2	3	4	5	6	7
	1	2	3	4	5	6	7	8
	2	3	4	5	6	7	8	9
	3	4	5	6	7	8	9	10
	4	5	6	7	8	9	10	11
	5	6	7	8	9	10	11	12
	6	7	8	9	10	11	12	13
	7	8	9	10	11	12	13	14

$$r = \frac{\sum_{i=0}^M \sum_{j=0}^N A(i,j)K(i,j)}{\sum_{i=0}^M \sum_{j=0}^N K(i,j)} \quad \dots (식 3)$$

MN : 블록의 화소수  
A(i,j) : 가중치  
K(i,j) : 블록의 DCT계수  
(i,j) : 화소의 좌표

4.2. 지수 r에 비례하여 탐색점 배치

제안된 기법은 3단계 기법을 바탕으로 탐색점간 거리 조절을 통하여 탐색점의 수를 증가시킨다. 3단계 기법에서 탐색점간 거리는 탐색영역/2로 최초 9개의 탐색점이 선택된다. 여기서 탐색점간 거리를 줄여 탐색점을 증가시킨다. threshold(문턱치)1, 2에 따라 세가지 경우를 모의실험한다.

- (가)  $r \leq \text{threshold1}$  : 탐색점간 거리를 2배 (탐색점간 거리=탐색영역/2)로 줄임, 3단계 기법
- (나)  $\text{threshold1} < r \leq \text{threshold2}$  : 탐색점간 거리를 3배로 줄임.
- (다)  $r > \text{threshold2}$  : 탐색점간 거리를 4배로 줄임.

다음 탐색 단계에서도 동일한 방법으로 탐색점간 거리를 조정하여 탐색점의 수를 증가시킨다. 따라서 최대 탐색점의 수는 탐색영역이 -16에서 16일 때, 지수 r 값에 따라서 33개에서 97개, 193개로 증가한다.

각 탐색기법에 대한 최대 탐색점수를 탐색범위 16에 대하여 다음과 같이 표로 작성하였다.

표2. 각 탐색기법의 복잡도(블록당 MAD연산)

탐색기법	탐색점 수(A)	연산량
FSBM	$(2\omega + 1)^2$	1,089
TSS(t=2)	$1 + 8\log_2 \omega$	33
제안방법 (t=2, 3, 4)	$(1 + 8\log_2 \omega) * i + (1 + 24\log_2 \omega) * j + (1 + 48\log_2 \omega) * k$	$33 \leq A \leq 193$

FSBM : Full Search, TSS : Thress Step Search  
 $\omega$ : 탐색영역(=16), t: 상수(탐색점간 거리= $\omega/t$ )  
 i: t=2인 블록수, j: t=3인 블록수, k: t=4인 블록수

5. 모의실험 및 결과분석

제안기법을 MPEG 영상 Flower Garden, Hockey 각각 50프레임, 720x480 크기, 4:2:0 영상에 대하여 모의실험하였다. 블록크기는 8x8이고 탐색범위는 -16 ~ +16으로 설정하였다. 지수 r에 대한 threshold1과 2는 각각 0.3과 1로 주었다.

화질의 척도로는 PSNR(Peak Signal to Noise Ratio)을 사용하여 평가하였다.

$$PSNR = 10 \log_{10} \frac{255^2}{MSD} \quad [dB] \quad (\text{식 4})$$

그 결과는 그림 2, 표 3과 같다.

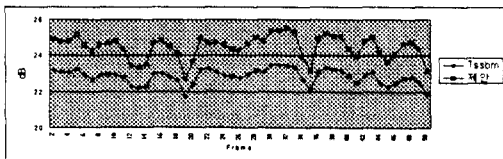


그림 2. Flower Garden(W=16, 50프레임)의 PSNR비교

그림2는 Flower Garden 50프레임에 대하여 3단계 탐색기법과 제안하는 방법에 대한 PSNR값을 나타낸다.

표 3. 각 기법의 50프레임 평균 PSNR[dB]과 블록당 연산량(MAD)

탐색기법	Flower Garden		hockey	
	PSNR	연산량	PSNR	연산량
전역탐색(FSBM)	27.92	1,089.00	29.53	1,089.00
3단계 탐색(t=2)	22.85	32.49	27.59	32.43
3단계 탐색(t=4)	25.57	92.77	28.37	94.55
제안 방법	24.49	51.83	27.76	44.34

$\omega$ : 탐색영역(=16), t: 상수(탐색점간 거리= $\omega/t$ )

제안한 방법은 표3에서 보듯이 3단계 탐색방법보다 연산량은 증가하였지만, 연산량 증가에 비하여 PSNR 값이 뚜렷이 개선되었다.

6. 결론

본 논문에서는 공간주파수 성분을 나타내는 지수 r에 비례하여 탐색점 수를 조정하여 국부극소에 빠지기 쉬운 고주파영역에 탐색점 수를 증가시킨다. 그 결과 연산량은 3단계 탐색기법에 비해 증가하지만 화질은 개선되었다.

Flower Garden 과 Hockey를 각각 50프레임으로 모의실험한 결과 제안한 기법이 3단계 기법에 비하여 연산량은 증가하지만 PSNR은 0.3~1.5[dB]증가하여 좋은 성능을 보였다.

또, 영상 특성에 맞는 지수 r의 threshold를 조정함으로써 원하는 화질을 얻을 수 있는 장점이 있다.

앞으로의 연구과제로 영상의 주파수 성분에 따라 자동적으로 threshold를 적용하여 탐색점 배치를 할 수 있는 연구가 더 이루어져야 한다.

참고 문헌

- [1] H. G. Muamann, P. Pirch, and H. J. Grallert, "Advances in picture coding" Proc. of IEEE, Vol.73, No4, pp. 523-548, April 1985.
- [2] A. K. Jain, "Image data compression: A review" Proc. of IEEE, Vol 69, No3, pp 349-389, Mar. 1981
- [3] Borko Furht, Joshua Greenberg, Raymond Wet-water, "Motion Estimation Algorithms for Video Compression", KAP, 1997
- [4] 정화자, "DCT를 이용한 윤곽선 추출", 정보과학회 논문지 제 3권 제 1호, 1997.2 pp 90-97
- [5] 정제창 譯, 일본 멀티미디어 통신연구회 編, "그림으로 보는 최신 MPEG", 교보문고, 1995