

H.264/AVC에서 효율적인 정화소 움직임 추정

윤효순*, 김미영**

*전남대학교 전산학과

**전남도립대학교 보건의료학과

e-mail:estheryon@hotmail.com, imkimmee@naver.com

Efficient Integer pixel Motion Estimation on H.264/AVC

Hyo-Sun Yoon*, Mi-Young Kim**

*Dept of Computer Science, Chonnam National University

**Dept of Health Medical, Jeonnam Provincial College

요 약

H.264/AVC에서 움직임 추정에 소요되는 시간을 줄이면서 보다 나은 화질을 유지하기 위하여 본 논문에서는 정화소 움직임 추정 기법을 제안하였다. 본 논문에서는 정화소 움직임을 위하여 대칭적인 탐색패턴을 이용한 계층적인 탐색 기법을 사용하였다. 제안한 정화소 움직임 추정 탐색 기법은 십자가 탐색 패턴, 다중 사각형 탐색 패턴, 작은 사각형 탐색 패턴, 다이아몬드 탐색 패턴들로 이루어져 있다. 제안한 움직임 추정 탐색 기법에서 사용한 탐색 패턴들은 블록 움직임이 수직으로 크거나 블록 움직임이 크면서 규칙적인 영상에서 국부적 최소화 문제를 해결하고 움직임 추정에 소요되는 시간을 줄이기 위하여 탐색 영역 내에 탐색 점들을 규칙적, 대칭적으로 배치하였다. 제안한 기법의 성능을 전역 탐색 기법의 성능과 비교하였을 때 움직임 추정에 소요되는 시간에 있어서 약 4~5.5 배의 속도 향상을 가져왔으며, 영상 화질에 있어서 전역 탐색 기법의 화질과 같거나 약간의 화질 저하를 보였다. 반면에, 비대칭 다중 육각형 탐색 기법의 성능과 비교하였을 때 움직임 추정 속도면에 있어서 약간의 성능 향상과 화질에 있어서 비슷하거나 최대 0.05 (dB)정도 향상을 보였다

1. 서론

일반적으로 동영상 압축 표준안들은 정화소 움직임 추정, 부화소 움직임 추정 단계를 거쳐 최적의 움직임 벡터를 찾아내는데, 기존의 제안된 고속 정화소 기법들은 바로 이 정화소 움직임 추정에 소요되는 계산량을 줄이기 위한 알고리즘들이다. 대표적인 정화소 기법인 3단계 탐색(Three Step Search:TSS)[3], 새로운 3단계 탐색(New Three Step Search:NTSS)[4], 2차원 로그형 탐색(2 Dimension LOGarithmic search:2DLOG)[5], 4단계 탐색(Four Step Search:4SS)[6], 다이아몬드 탐색(Diamond Search:DS)[7][8], 2단계 탐색(2 Step Search:2SS)[9][10], 움직임 벡터를 추정하는 움직임 벡터 필드 적응적 탐색 기법(MVFAST)[11], 예측된 움직임 벡터 필드 적응적 탐색 기법(PMVFAST)[12]등은 상대적으로 움직임이 작은 영상이나 영상 크기가 작은 경우에 좋은 결과를 보이지만, 영상의 움직임이 크거나 영상의 크기가 큰 경우에는 국부적 최소화 문제에 빠지게 되어 화질이 저하되는 결과를 볼 수 있다.

H.264/AVC[1][2]는 기존 압축 표준안 H.26x (x : 1,2,3)와 MPEG-y (y : 1,2,4)보다 압축 효율과 영상 화질에 있어서 성능이 향상됨을 보였다. H.264/AVC는 압축 효율과 영상 화질을 향상시키기 위하여 가변 블록 사이즈 움직임 추정, 1/4화소 움직임 추정, 여러 개의 참조 영상

을 사용하는 움직임 보상, 디블러킹 필터 내장등과 같은 기술들을 사용했기 때문이다. 그러나 이러한 기술들은 움직임 추정에 있어서 많은 계산량을 필요로 한다.

H.264/AVC에서는 정화소 움직임 추정 시 영상의 움직임이 크거나 영상 사이즈가 상대적으로 큰 영상에서도 움직임 벡터 추정을 비교적 잘 수행할 수 있는 비대칭 다중 육각형 탐색 기법[13]을 사용하여 화질 저하를 개선하고자 하였으나, 이 기법은 탐색 영역 내에 배치하는데 있어서 비대칭적, 불규칙적으로 배치함으로써 블록 움직임이 수직으로 크거나 블록 움직임이 크면서 규칙적인 영상에서 국부적 최소화에 빠질 수 있는 단점이 있다.

본 논문에서는 블록 움직임이 수직으로 크거나 블록 움직임이 크면서 규칙적인 영상에서 국부적 최소화에 빠지지 않도록 탐색 점들을 탐색 영역 내에 대칭적, 규칙적으로 배치하여 정화소 움직임 벡터를 추정하는 기법을 제안하였다.

본 논문의 구성은 2장에서 H.264/AVC에서 정화소 움직임 추정에 대하여 설명하고, 3장에서 제안한 기법을 기술한다. 그리고 4장에서 기존 기법들과 성능을 비교한 후, 5장 결론으로 맺는다.

2. H.264에서 정화소 움직임 추정

동영상 압축 표준안에서는 영상의 시간적 상관성을 이

용하여 영상을 압축하기 위하여 각각의 영상을 분할해서 이전에 부호화된 영상으로부터 움직임 추정하여 현재의 블록을 예측하는 움직임 보상 기법을 사용하고 있다. H.261/263 그리고 MPEG-1/2/4 압축 표준안에서 움직임 추정은 일반적으로 16×16 블록 단위로 수행된다. 그러나 H.264/AVC에서의 움직임 추정은 16×16 블록 단위부터 4×4 블록 단위까지 7종류의 움직임 보상 블록 크기를 사용하여 움직임 벡터를 추정한다. 움직임 보상의 블록 크기를 보다 작게 하면 정확한 움직임 예측을 할 수 있지만 계산량과 부호량이 증가하게 된다.

H.264/AVC에서는 매크로 블록을 어떤 모드 부호화 할지를 결정하는 것이 H.264 /AVC 부호기의 핵심이다. H.264/AVC는 최적의 모드를 결정하기 위하여 왜곡도와 함께 발생 비트를 고려한다. H.264/AVC는 16×16, 16×8, 8×16, 8×8, 4×8, 8×4 그리고 4×4의 7가지 블록 모드에 대하여 움직임 추정을 수행한다. 움직임 보상과 모드 판정에 대해서 다음과 같은 비용 함수 식 (1), (2)를 정의하고 그 비용 함수를 최소로 하는 최적의 모드를 결정한다[14][15].

$$J_{cost}(s, c) = SAD(s, c) + \lambda motion\ Genbit(\Delta MV) \quad (식\ 1)$$

$$J_{mode}(s, c) = SSD(s, c, Mode|QP) + \lambda modeR(s, c, Mode|QP) \quad (식\ 2)$$

위 식 (1)에서 SAD (Sum of Absolute Difference)은 원 영상 (s)와 참조 영상(c)의 예측 오차로 절대값의 합이고, $\lambda motion$ 은 움직임 탐색에 대한 라그랑제 승수로 양자화 계수에 의존하는 값이다. 그리고 ΔMV 는 실제의 움직임 벡터와 예측된 움직임 벡터의 차이이고 $Genbit(\Delta MV)$ 는 ΔMV 을 비트로 표현한 것으로 헤더부분의 부호량이다. 식(2)에서 $Mode$ 는 예측모드이고 QP 는 양자화계수이다. 그리고 SSD (Sum of Square Difference)는 원영상과 참조영상과의 제곱오차 합이고 $\lambda mode$ 는 모드판정에 대한 라그랑제 승수이며 $R(s, c, Mode|QP)$ 는 헤더부분 뿐만 아니라 직교 변환과 양자화 변환 처리 후의 변환계수 부분의 비트도 포함한다. 식 (2)의 비용 함수를 최소로 하는 모드를 현재 블록의 모드로 선정한다.

비대칭 십자가 탐색 패턴, 불규칙 다중 육각형 탐색 패턴, 육각형 탐색 패턴, 작은 다이아몬드 탐색 패턴들로 이루어진 비대칭 다중 육각형 탐색 기법이 H.264/AVC에서 고속 정 화소 움직임 추정 기법으로 사용된다.

비대칭 다중 육각형 탐색 기법은 아래와 같은 알고리즘을 수행하여 움직임 벡터를 추정한다.

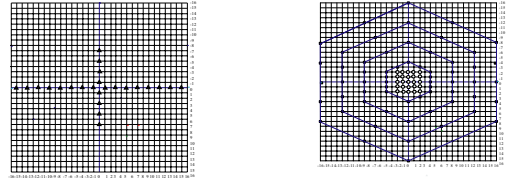
[단계1] : 초기 탐색 시작점을 원점으로 그림 1의 (a)와 같이 탐색 점들을 배치한다. 탐색 점들 중 최소 비용함수 값을 가지는 점이 다음 단계의 탐색 시작점이 된다.

[단계2] : 그림 1의 (b)와 같이 탐색 범위 2의 전역 탐색과 다중 육각형 탐색을 실시한다. 최소 비용 함수 값을 가지는 점이 다음 단계의 탐색 시작점이 된다.

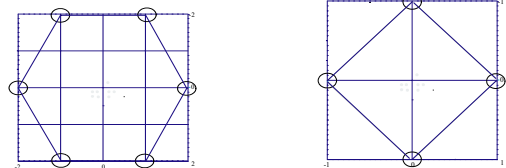
[단계3] : 그림 1의 (c)와 같이 탐색 점들을 배치하고 이 탐색 점들 중 최소 비용함수를 가진 점을 결정한다. 만약

최소 비용함수를 가진 점의 위치가 탐색 패턴의 중심에 있는 점이면 다음 단계를 수행하고, 그렇지 않으면 이 단계를 반복 수행한다.

[단계4] : 그림 1의 (d)와 같이 탐색 점들을 배치하고 이 탐색 점들 중 최소 비용함수를 가진 점을 결정한다. 최소 비용함수를 가진 점을 정화소 움직임 벡터이다.



(a)비대칭 십자가 탐색 패턴 (b)불규칙 다중 육각형 탐색 패턴



(c) 육각형 탐색 패턴 (d) 작은 다이아몬드 탐색 패턴

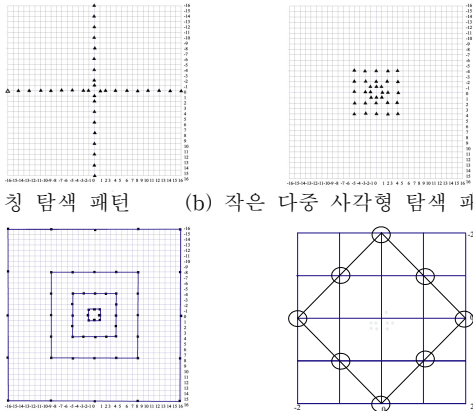
(그림 1) 비대칭 다중 육각형 탐색기법의 탐색 패턴들

3. 제안한 정화소 움직임 추정 기법

비대칭 다중 육각형 탐색 기법 중 1단계에서 사용하는 그림 1(a)의 비대칭 십자가 탐색 패턴은 수직 방향보다 수평 방향에 더 많은 탐색 점들을 배치함으로써 영상의 수직 움직임이 큰 영상이나 대각선 방향에 움직임이 있는 영상에서 국부적 최소화에 빠짐으로서 화질 저하를 야기할 수 있다. 그리고 비대칭 다중 육각형 탐색 기법 중 2단계에서 사용하는 그림 1(b)의 불규칙 다중 육각형 탐색 패턴은 탐색 범위 2의 전역 탐색과 다중 육각형 탐색으로 이루어져 있다. 만약 비대칭 십자가 탐색 패턴을 수행한 후 얻어진 움직임 벡터가 정확하지 못한 움직임 벡터라면 탐색 범위 2의 전역 탐색을 수행하는 것은 시간 낭비가 된다. 그리고 다중 육각형 탐색은 움직임이 크면서 불규칙한 영상에 적합한 탐색 패턴으로 움직임이 크고 규칙적인 영상에서 잘못된 움직임 벡터를 추정할 수도 있다.

제안한 정화소 움직임 추정 기법 중 그림 2(a)의 대칭 탐색 패턴은 정화소 움직임 벡터가 대각선 방향 보다 수평, 수직 방향에 많이 존재한다는 사실을 이용하여 탐색 점들을 수평, 수직 방향으로 규칙적, 대칭적으로 배치함으로써 움직임이 수직 방향으로 큰 영상에서 국부적 최소화에 빠지지 않도록 하였고 또한 다음 단계에서 보다 정확한 탐색의 시작점을 사용할 수 있도록 대각선 방향에도 탐색 점들을 배치하였다. 그리고 그림 2(b)의 작은 다중 육각형 탐색 패턴은 대칭 탐색 패턴의 수행 결과를 이용하여 초기에 정화소 움직임 추정을 종료하기 위하여 사용한다. 만약 대칭 탐색 패턴에서의 최소 비용함수를 가진 점을 중심으로 작은 다중 육각형 탐색 패턴에 있는 점들을 배치한 후 최소 비용함수를 가진 점을 결정하고 이 점의 비용함수의 값에 따라 다음에 사용할 탐색 패턴이 달

라진다. 그림 2(c)의 큰 다중 사각형 탐색 패턴은 움직임이 크고 규칙적인 영상에 국부적 최소화에 빠지지 않도록 탐색 점들을 탐색 영역 내에 대칭적으로 일정한 간격을 두고 배치하였다. 그리고 제안한 정화소 움직임 추정 기법은 국부적인 움직임 추정을 위하여 그림 1(c)의 육각형 탐색 패턴 대신 그림 2(d)의 큰 다이아몬드 탐색 패턴을 이용하였는데, 실험 결과 큰 다이아몬드 탐색 패턴을 사용한 다이아몬드 탐색 기법 (Diamond Search :DS)이 육각형 탐색 패턴을 이용한 육각형 탐색 기법(Hexagon Search :HS)보다 화질과 속도면에 있어서 더 나은 결과를 보였기 때문이다. 실험 결과는 표 1, 표2에 나타나 있다. 표 1의 실험 결과를 보면 영상 화질면에 있어서 다이아몬드 탐색 기법이 육각형 탐색 기법보다 좋은 결과를 보였다. 다이아몬드 탐색 기법이 육각형 탐색 기법보다 더 정확한 움직임 벡터를 추정하므로 표 2의 실험 결과에 다이아몬드 탐색 기법이 육각형 탐색 기법보다 움직임 추정 시간을 더 많이 필요로 하는 경우가 발생한다.



(a) 대칭 탐색 패턴 (b) 작은 다중 사각형 탐색 패턴
(c) 큰 다중 사각형 탐색 패턴 (d) 큰 다이아몬드 탐색 패턴
(그림 2) 제안한 정화소 움직임 추정 탐색 패턴들

제안한 정화소 기법은 비대칭 다중 육각형 탐색 기법처럼 메디안 예측, 상위 모드를 이용한 예측, 이전 화면의 대응 블록을 이용한 예측, 이웃한 참조 화면을 이용한 예측을 이용하여 초기 탐색 시작점을 결정한다. 제안한 정화소 움직임 추정 기법은 다음과 같은 알고리즘을 수행하여 정화소 움직임 벡터를 추정한다.

[단계1] : 초기 탐색 시작점을 원점으로 그림 3의 (a)와 같이 탐색 점들을 배치한다. 이들 탐색 점들 중 최소 비용 함수 값을 가지는 점이 다음 단계의 탐색 시작점이 된다.

[단계2] : 그림 3의 (b)와 같이 작은 다중 사각형 탐색을 실시한다. 최소 비용 함수 값을 가지는 점의 비용 함수 값이 충분히 작으면 4단계를 수행하고 그렇지 않으면 다음 단계를 수행한다.

[단계3] : 그림 3의 (c)와 같이 다중 사각형 탐색을 실시한다. 최소 비용 함수 값을 가지는 점이 다음 단계의 탐색 시작점이 된다.

[단계4] : 그림 3의 (d)와 같이 탐색 점들을 배치하고 이

탐색 점들 중 최소 비용함수를 가진 점을 결정한다. 만약 최소 비용함수를 가진 점의 위치가 탐색 패턴의 중심에 있는 점이면 다음 단계를 수행하고, 그렇지 않으면 이 단계를 반복 수행한다.

[단계5] : 그림 2의 (d)와 같이 탐색 점들을 배치하고 이 탐색 점들 중 최소 비용함수를 가진 점을 결정한다. 최소 비용 함수를 가진 점을 정화소 움직임 벡터로 추정하고 탐색을 마친다.

<표 1> QCIF 영상과 CIF영상의 평균 PSNR

QCIF 영상	DS	HS	CIF영상	DS	HS
Akiyo	36.83	36.81	Akiyo	38.44	38.44
Container	34.75	34.75	Bus	32.71	32.69
Foreman	34.33	34.23	Children	35.55	35.54
M&D	34.93	34.93	Container	34.7	34.69
Mobile	31.64	31.61	Foreman	35.62	35.57
Salesman	34.01	34.02	Mobile	32.52	32.5
Silent	34.36	34.35	M&D	37.53	37.5
Stefan	32.43	32.41	Paris	33.86	33.86
Table	34.23	34.22	Stefan	33.52	33.52

<표 2> QCIF 영상과 CIF영상의 움직임 추정 시간

QCIF영상	DS	HS	CIF영상	DS	HS
Akiyo	51.179	52.946	Akiyo	204.556	210.013
Container	40.163	50.783	Bus	324.977	324.075
Foreman	50.776	48.514	Children	207.939	172.449
M & D	48.393	58.672	Container	188.895	209.161
Mobile	52.998	65.668	Foreman	268.199	206.546
Salesman	43.669	55.979	Mobile	246.418	234.393
Silent	47.129	60.78	M & D	200.395	179.765
Stefan	54.81	67.22	Paris	216.77	176.238
Table	64.455	67.439	Stefan	278.03	259.175

4. 실험결과

제안한 움직임 추정 기법은 H.264의 JM 9.5에서 실험하였으며, QCIF인 Akiyo, Container, Foreman, Mother and Daughter (M&D), News, Salesman, Silent, Stefan, Suzie and Table 영상과 CIF인 Akiyo, Bus, Children, Container, Foreman, Mobile, Mother and Daughter, Paris and Stefan 영상의 각각 150 프레임을 사용하였고 비교 탐색 기법으로는 전역 탐색 기법 (Full Search :FS), 비대칭 다중 육각형 탐색 기법 (UMHexagonS)들을 사용하였다. 성능 비교 함수로는 영상 화질의 품질을 평가하기 위하여 PSNR(Peak Signal-to-noise Ratio)을 사용하였는데, PSNR은 다음과 같이 정의된다.

$$MSE = \left(\frac{1}{MN}\right) \sum_{m=1}^M \sum_{n=1}^N [x(m, n) - \hat{x}(m, n)]^2 \quad (식 4)$$

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (식 5)$$

여기에서 식(4)에서 M, N은 영상의 가로와 세로의 크기이고, $x(m, n)$ 는 원 영상화면을 나타내고, $\hat{x}(m, n)$ 는 움직임 보상 예측된 화면을 나타낸다. 각 영상에 대한 실험 결과를 표 3, 표 4, 표 5, 표 6에 나타냈는데, 표 3, 표 4에 각 실험 영상의 움직임 추정 시간을 나타내었고 표 5, 표 6에 각 실험 영상의 평균 PSNR(dB)의 평균을 나타낸다.

<표 3> QCIF 영상의 움직임 추정 시간

	FS	UMHexagons	제안
akiyo	299.06	65.559	64.922
container	353.569	69.372	66.935
foreman	340.414	93.236	93.183
M&D	304.484	80.481	79.179
news	299.969	70.904	71.035
salesman	298.077	72.832	74.083
silent	302.51	79.78	79.488
stefan	302.541	106.706	105.732

<표 4> CIF 영상의 움직임 추정 시간

	FS	UMHexagons	제안
akiyo	1204.366	266.988	265.731
BUS	1241.977	511.289	506.342
children	1172.804	319.013	313.653
container	1221.643	290.917	286.447
foreman	1249.968	392.383	388.66
mobile	1230.134	416.083	412.227
paris	1180.629	302.616	304.003

5. 결론

본 논문에서는 H.264에서 움직임 추정에 소요되는 계산량을 줄이면서 화질을 유지하기 위한 정화소 움직임 추

<표 5> QCIF 영상의 평균 PSNR

	FS	UMHexagons	제안
akiyo	36.75	36.7	36.71
container	34.63	34.62	34.63
foreman	34.45	34.4	34.4
M&D	34.93	34.88	34.93
news	35.12	35.1	35.15
salesman	34.02	33.99	34.02
silent	34.36	34.33	34.34
stefan	32.5	32.47	32.47

<표 6> CIF 영상의 평균 PSNR

	FS	UMHexagons	제안
akiyo	38.38	38.38	38.38
BUS	33.29	33.27	33.27
children	35.68	35.64	35.64
container	34.6	34.59	34.59
foreman	35.71	35.68	35.68
mobile	32.47	32.46	32.46
paris	33.82	33.77	33.8

정 기법과 부화소 움직임 추정 기법을 제안하였다. 제안한 정화소 움직임 추정 기법은 계층적인 탐색 기법으로 대칭적 십자가-엑스 탐색 패턴, 다중 사각형 탐색 패턴, 큰 다이아몬드 탐색 패턴, 작은 다이아몬드 탐색 패턴들로 이루어져있다. 이 탐색 패턴들은 영상 화질에 영향을 미치는 국부적 최소화 문제를 해결하기 위하여 탐색 영역 내에 탐색 점들을 수평, 수직, 대각선 방향으로 규칙적, 대칭적으로 배치하였다. 그리고 제안한 부화소 움직임 추정은 블록 크기에 따라 탐색 패턴을 적응적으로 결정하여 화질을 더욱 좋게 하였다

실험을 통하여 알 수 있듯이, 제안한 기법을 전역 탐색 기법과 비교하였을 때 움직임 추정 속도면에 있어서 약 4.5배(움직임이 작은 영상) ~ 6배(움직임이 많은 영상) 향상을 보였고 평균 0.01(dB) 화질저하를 보였다. 그리고 제안 기법을 비대칭 다중 육각형 탐색 기법과 비교하였을 때 움직임 추정 속도와 화질에 있어서 각각 약 1.2배와 약 0.02 (dB)정도 향상을 보였다.

Acknowledgment

"이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2010-0024120)

참고문헌

[1] T.Wiegand, G.J.Sullivan, A.Luthra, "Overview of the H.264/AVC Video Coding Standard", IEEE Trans. on Circuits and System for Video Technology Vol. 13 No. 7 pp. 560-576, 2003.

[2] Draft ITU-T recommendation and final draft international standard of joint video specification. (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC in Joint Video Team(JVT) of ISO/ICE/ MPEG and ITU-T VCEG, JVTG050, 2003.

[3] T.Koga, K.Inuma, A.Hirano, Y.Ishiguro, "Motion compensated interframe coding for video conference", Proc. NTC81, pp.G5.3.1-5.3.5, Nov. 1981.

[4] R.Li, B.Zeng and M.L.Liou, "A New Three Step Search Algorithm for Block Motion Estimation", IEEE Trans. on Circuits and System for Video Technology Vol. 4 No. 4 pp. 438-441, Aug. 1994.

[5] J.R.Jain and A.K.Jain, "Displacement measurement and its application in Interframe image Coding", IEEE Trans. on Communications, Vol. 29 No. 12 pp.1779-1808, Dec. 1981.

[6] L.M.Po and W.C.Ma, "A Novel Four Search Algorithm for Block Motion Estimation". IEEE Trans. on Circuit and Systems for Video Technology Vol. 6. pp.313-317, June 1996.

[7] J.Y. Tham, S.Ranganath and A.A.Kassim, "A Novel Unrestricted Center-Biased Diamond Search Algorithm for Block Motion Estimation", IEEE Transactions on Circuits and Systems for Video Technology. Vol. 8 pp.369-377, Aug. 1998.

[8] S.Zhu and K.K.Ma "A New Diamond Search Algorithm for Fast Block Matching Motion", IEEE Transaction on Image Processing Vol. 9 No. 2 pp287-290, Feb. 2000.

[9] Yuk Ying Chung ,Neil W.Bergmann "Fast Search Block Matching Motion Estimation Algorithm using FPGA", Visual Communications and Image Processing 2000. Proc.SPIE Vol. 4067 pp913-921.

[10] Danian Gong, Yun He, "Fast Motion Estimation Algorithm using Horizontal and Multi-grid Search Strategy", Picture Coding Symposium 2001. pp362- 365

[11] Prabhudev Irappa Howur, K. K. Ma, "Report on Performance of Fast Motion Estimation using Motion Vector Field Adaptive Search Technique", ISO/IEC/JCTI/SC29/WG11 M5453, Dec. 1999.

[12] A. M. Tourapis, O. C. Au, M. L. Liou, "Fast Block Matching Motion Estimation using Predictive Motion Vector Field Adaptive Search Technique", ISO/IEC/JCTI/SC29/WG11 MPEG2000/M5866, Mar. 2000.

[13] Z. Chen, P. Zhou, Y. He, " Fast Integer pel and Fractional pel motion estimation in for JVT" in JVT of ISO/ICE/ MPEG and ITU-T VCEG, 6th meeting, 2002.

[14] M. Wein, "Variable Block-Size Transforms for H.264/AVC", IEEE Trans. on Circuits and System for Video Technology Vol. 13 No. 7 pp. 604-613, 2003.

[15] T. Wiegand, B. Girod, "Lagrangian Multiplier Selection in hybrid video coding control", ICIP 2001.