동영상 부호화를 위한 DCT 변환에 의한 트리 구조의 임베이드 부호화
Tree structured embedded coding based on DCT for moving image

조경식, 이연문, 정차근
호서대학교 제어계측공학과
Kyung-Sik Cho, Youn-Moon Lee, Cha-Keon Cheong
Dept. of C&I Eng., College of Eng., Hoseo University
e-mail : 994901@mail.hoseo.ac.kr

요 약 문

본 논문에서는 정확한 비트 제어가 가능한 동영상 통신을 위한 새로운 영상 부호화 방법을 제안한다. 기존의 표준 부호화기와 호환성을 최대한 유지하기 위해 MC/DCT 하이브리드 부호화 구조에 의거한 순차 전송의 임베이드 트리 구조의 부호화 방법을 기술한다. 제안 방법의 유효성을 검증하기 위해 표준 동영상에 대한 컴퓨터 모의 실험 결과 제안방법은 기존의 부호화 방법에 비해 비트 제어가 용이하고 부호화 성능이 개선됨을 확인했다.

1. 서 론

최근 멀티미디어 통신이 급격히 증가하는 추세에 따라서 영상 전달이 더욱 많아지고 있다. 그러나, 영상 정보는 크기가 매우 커서 압축을 하지 않고는 실시간에 전달하기는 거의 불가능하다. 현재 널리 퍼져 있는 가정용 전화선 정도의 회선을 이용하기 위해서는 압축율이 뛰어난 영상 압축 알고리즘이 필요하다. 이를 위하여 64kbps 이하의 통신 환경에서 영상 통신을 할 수 있는 영상 압축 알고리즘이 개발되었으며, 이는 H.263으로 표준화되었다.

일반적으로 I 프레임에서 발생되는 비트 량은 전체 비트 량의 40-70%을 차지할 정도로 비중이 매우 높으며 연속되는 P 프레임이나 B 프레임의 부호화에도 많은 영향을 준다. 따라서 낮은 비트레이트 부호화기의 경우 I 프레임의 효율적인 부호화는 전체적인 부호기의 성능향상에 매우 중요하다.

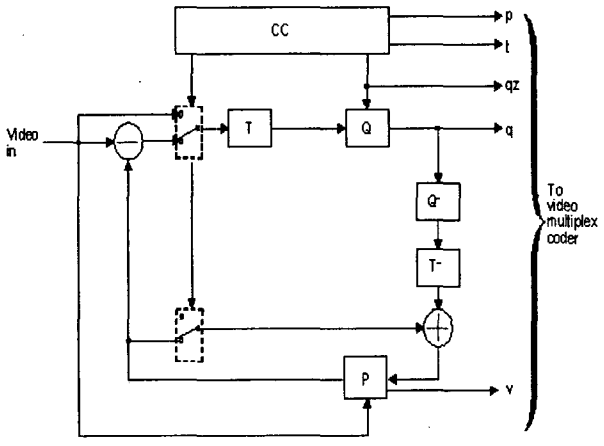
또한 낮은 비트레이트 부호화기에서 전송 대역폭을 효율적으로 사용하기 위해서는 비트 율 제어가 특히 중요한데 H.263 하이브리드 부호화기의 경우 양자화 값을 변화시켜 비트 율을 제어하고 있으나 양자화 레벨에 따라 발생되는 비트 량의 변동폭이 크며 비트 율의 정확한 제어가 어렵다. 따라서, 우수한 성능의 저 전송률 영상 부호화기를 개발하기 위해서는 I 프레임의 효율적인 부호화 방법과 정확한 비트률 제어가 가능한 부호화기가 필요하다.

본 논문에서는 H.263 부호화기의 인트라 프레임의 영상을 DCT변환한 계수와 인터 프레임의 움직임 보상된 영상의 DCT 변환계수를 3 레벨의 웨이브렛 구조로 보고 제로 트리 구조로 재구성 후 임베이드 제로트리 부호화기를 적용하여 기존 MC DCT 하이브리드 부호화기와 호환성을 가지면서 비트제어가 정확하고 성능이 좋은 부호화기를 제안하였다. 제안 부호화기는 인트라 프레임의 부호화의 경우 1bpp이상에서는 기존 부호화기 보다 성능이 우수한 결과를 얻었지만 1bpp 미만에서는 약간 떨어지는 문제가 발생하였다. 하지만 비트율 제어면에서는 임베이드 제로트리 부호화기의 적용으로 기존의 방법보다 정확한 제어가 가능했다.

2장에서 실험에서 사용한 기존 화상회의용 동영상 표준 부호화기인 H.263 부호화기에 대하여 간단히 알아보고 3장에서는 임베이드 영상 부호하기에 대하여 설명한다. 4장에서는 본 논문에서 제안한 방법인 DCT 변환 계수에 임베이드 변환 부호화를 적용하는 방법과 엔트로피 부호화 방법에 대해 설명한다. 6장에서는 기존의 방법과 비교하기 위해 인트라와 인터 프레임의 영상에 대해 H.263과 성능을 비교 분석한다.

2. H.263

H.263은 ITU-T에서 제안된 저 전송률을 가지는 통신 선로(64kbps이하)에서 영상 회의나 비디오 전화 등을 위한 멀티미디어 통신 서비스의 동영상 부분에 대한 압축에 쓰일 수 있는 부호화된 표현에 대한 권고안이다. 기본적인 영상 소스 코딩 알고리즘은 H.261에 기반하고 있다. 그러나 H.261에 비하여 같은 화질의 영상에 대해 거의 반정도의 비트를 생성한다. 그림 1은 H.263 부호화기의 알고리즘을 나타내는 블록도 이다.



T Transform
Q Quantizer
P Picture Memory with motion compensated variable delay

CC Coding control
p Flag for INTRA/INTER
t Flag for transmitted or not
qz Quantizer indication
q Quantizing index for transform coefficients
v Motion vector

그림 1 H.263 Source Coder

입력 영상은 그림 2 처럼 먼저 여러개의 매크로 블록으로 이루어진 GOB(Group of block)으로 나뉘어 지고 매크로 블록은 4개의 8x8 휘도 블록과 2개의 색차 블록으로 구성된다.

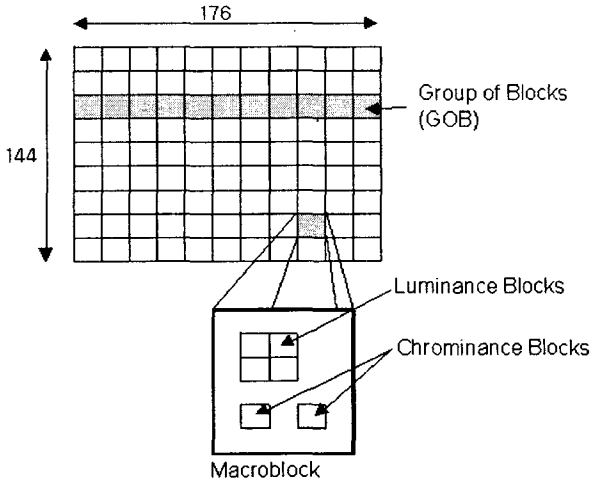


그림 2 H.263 Block Hierarchy

QCIF(176x144)의 경우 16라인이 하나의 GOB을 구성하여 9개의 GOB로 이루어지며 하나의 GOB는 11개의 매크로 블록으로 구성되어 총 99개의 매크로 블록으로 이루어져 있다. 부호화는 각 매크로 블록에 대해 좌에서 우로, 위에서 아래로 스캔 방식으로 이루어지는데 각 매크로 블록의 6개의 8x8블록에 대해 DCT변화(T)을 한 후 양자화(Q)가 이루어진다. 인터 프레임의 부호화는 전 프레임을 Q 인버스와 T인버스에 위해 복호화된 프레임과 현재 프레임의 차분을 구해 움직임 벡터(v)를 발생시키고 움직임 보상 에러에 대하여 부호화를 한다. 각 매크로는 인트라 블록의 첫 번째 계수(DC)값을 제외하고 모든 계수에 대해 동일하게 31개의 양자화 값으로 양자화 된다[9].

비트율을 제어하기 위해서는 이 31개의 양자화 값을 비트량에 따라 조정해야 되는데 양자화 값에 따라 발생되는 비트량이 일정하지 않고 원하는 비트가 되었다고 비트 발생을 중단하면 그림 2 에서 비트 발생이 중단된 이후의 매크로 블록은 부호화가 되지 않아 문제를 발생 하므로 부호화를 중단할 수 없기 때문에 비트율 제어가 매우 어렵다. 그래서 본 논문에서는 DCT변환 계수를 제로 트리 구조로 재구성하여 임베이드 제로 부호화를 하였다.

이 부호화 방식은 비트 발생이 임베이드하게 발생되기 때문에 원하는 비트량에서 부호화을 중단 하더라도 복호측에서 전체 영상을 복원하는 것이 가능하므로 비트 제어면에서 매우 좋은 특성을 가지고 있다.

3. 임베이드 영상 부호화기

Shaprio에 의해 EZW (Embedded Zerotree Wavelet)[1]의 개념이 소개된 후 정지 영상 압축에서 기존의 DCT 기반 영상 부호화기보다 뛰어난 rate-distortion 성능 향상을 갖는 임베이드 제로 트리 영상 부호화 방법에 관한 많은 연구가 진행되어 왔다 [2][3][4]. 이들의 방법은 주로 웨이브렛 변환 등 대역 분할된 계수의 효율적 부호화를 위한 것으로 비교적 간단한 알고리즘으로 높은 부호화 성능과, 계수의 위치 및 크기정보에 따라 부호화 함으로써 출력 비트 스트림의 임베이드화가 가능한 특징을 갖는다.

즉, 비트 스트림 발생의 임베이드화로 인해, 다양한 해상도가 가능한 scalability 특성과 정확한 비트율 제어가 가능하다. 제로 트리구조에 의한 임베이드 영상 압축은 SAQ (Successive Approximation Quantization)알고리즘을 기본으로 사용하고 있다. SAQ는 자기 유사 성질(self-similarity)을 이용해 대역간 중요도가 높은 계수의 위치를 예측하는 significance map coding과 웨이브렛 계수의 크기가 순차적으로 근사화가 되도록 하는 refinement로 이루어진다. 구체적인 SAQ알고리즘은 다음과 같다.

SAQ 알고리즘:

1. Set Threshold T

$T=2^n, n=\log_2(\max_{(i,j)}\{|c_{i,j}|\})$ (1)

$C_{i,j}$는 변환 계수를 의미.

2. Significance map coding

정의된 방식으로 모든 계수를 비교,
임계(Threshold)값 보다 크면 symbol을 출력하고 계수를 ±1.5T값으로 셋팅.

3. Refinement

significance계수를 refine.
계수값±0.25T.

4. $n=n-1, T=2^n$ goto step 2 (2)

n을 1감소시켜 0.5T에 대하여 다시 significance map coding을 함.

Significance map coding을 효율적으로 수행하기 위해 그림 3 과 같은 제로 트리구조를 사용한다. 이는 상위레벨에 있는 웨이브렛 계수가 임의의 주어진 임계값 T 보다 작다면 그것의 하위 레벨의 계수도 역시 작을 확률이 높다는 것을 이용한 것이다.

Shapiro의 EZW부호화기를 향상시킨 임베이드 알고리즘으로 SPECK(Set Partitioned Embedded bloCK coder)[2]과 SPIHT(Set Partitioning in Hierarchical Trees)[3]등이 있다. SPECK 알고리즘은 Quadtree partitioning과 Octave band

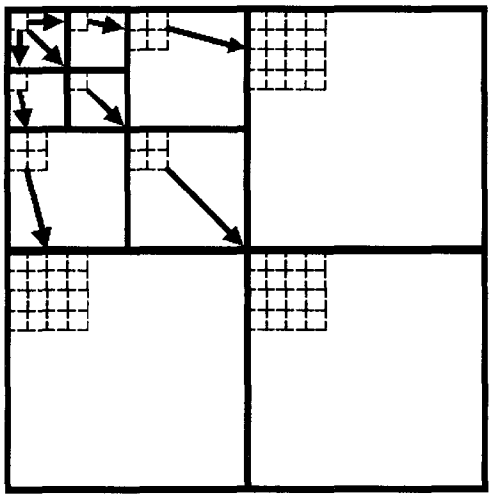


그림3 An Example of 3 Level Zerotree Structure

partitioning을 사용해 SPIHT 알고리즘 보다 구조가 간단하다. 그러나 SPIHT가 SPECK보다 rate-distortion 면에서 약간 우수하다[3].

기존의 제로 트리 부호화 기법들은 웨이브렛 변환을 사용하여 주로 높은 해상도(512x512) 영상을 이용하는 정지 영상 압축 부호화에는 매우 효율적이지만 QCIF(176x144)와 같은 낮은 해상도를 주로 사용하는 저전송율 비디오 압축에서는 필터 특성 및 불충분한 대역 분해, 즉 웨이브렛의 공간 주파수 특성 저하로 부호화 성능이 현저히 저하되는 문제점이 있으며, 실용화가 진행되고 있는 MC/DCT 하이브리드 표준 코덱과의 호환성을 갖지 못한다. 또

한 대역분할에 의한 EZW는 정지 영상 부호화만을 대상으로 하고 움직임 검출에 의한 동영상 부호화에 적용이 어려운 문제가 있다.

4. 임베이드 DCT 변환 부호화

그림 4는 제안 방식의 블록 다이어그램을 나타낸다. 기존 방식의 양자화 방식 대신에 계수들을 제로 트리 구조로 재배치(A)한 후 임베이드 제로 트리 부호화의 양자화 방법인 연속 추정 양자화(SAQ)을 하였다.

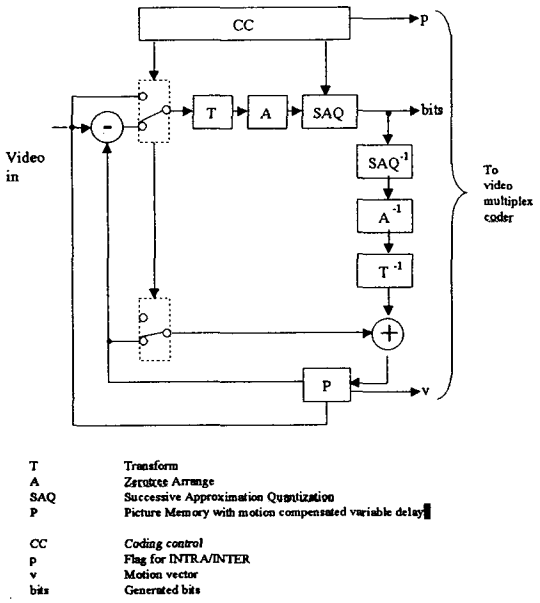


그림 4 제안한 부호기의 블록 다이어그램

일반적으로 전 프레임에 대한 DCT 변환에 의한 화상의 표현은 웨이브렛 변환에 의한 화상의 표현과 동일하다[7]. 이 성질을 이용해서, DCT에 의한 입력 화상의 대역 분할을 피라미드 구조로 재배치함으로써, DCT의 높은 에너지 compaction 특성을 활용하여 부호화 성능을 개선시키는 것이 가능하다. 그림 5는 본 논문에서 제안한 DCT 계수를 피라미드 구조로의 재배치 방법을 간략히 나타낸 것이다. 즉, 한 프레임 전체에 대해 8x8크기의 각 블록에 대한 DCT변환으로부터, DC성분, 저역 주파수성분, 고역 주파수 성분으로의 3레벨 피라미드 구조로 재배치한다. 그림 5 에서 DCT 변환 계수의 parent-children 관계는 i($1 \le i \le 63$)에 대해 parent는 [i/4]이고, j($1 \le j \le 15$)에 대해 children은 [4j, 4j+1, 4j+2, 4j+3]이다.

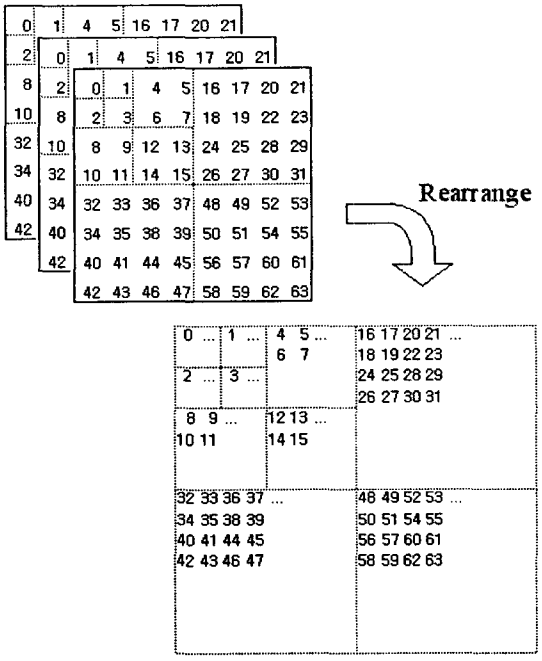


그림 5 DCT의 3 Level 구조로의 재배치

인트라 프레임의 경우 DC계수의 경우 다른 주파수 계수들에 비해 계수 값이 매우 커서 refinement시 불필요한 bit가 발생한다. 이를 방지하기 위해 DC레벨 값을 평균값으로 바이어스 시켜 부호화의 효율을 높였다.

재배열된 제로 트리 구조의 DCT계수를 양자화하기 위해 향상된 임베이드 제로 트리 부호화의 한 방법으로 SPIHT[3]을 이용한다. 이 임베이드 제로 트리 부호화 방법은 significant계수들이 주로 가장 상위 대역에 주로 분포한다는 성질을 이용함으로써 계수간 중복 성을 효율적으로 제거한다. DCT변환된 각 블록의 DC는 주변 블록의 DC값과 공간적 상관성이 매우 크다. 따라서, DCT 계수의 재배열로 가장 상위 대역에는 공간적 상관성이 큰 DC 계수들이 주로 분포하므로 임베이드 제로트리 부호화에 매우 적합하다.

전체 pixel set의 subset T_m에 대해 다음과 같은 significance test를 한다.

$$S_n(T_m) = \begin{cases} 1, & \max_{(i,j)\in T_m} \{|c_{i,j}|\} \ge 2^n \\ 0, & otherwise \end{cases} \quad (3)$$

식(2)에 위해 주어진 임계값 보다 크면 비트 '1'을 발생하고 그렇지 않으면 '0'을 발생한다. S_n이 0이면 다음 pass에서 n-1에 대해 다시 test를 하고 1

이면 T_m을 새로운 subset $T_{m,l}$로 나누어 significance test를 한다. subset이 단일 계수가 될 때까지 모든 계수에 대해 적용하여 중요한 계수부터 부호화을 행한다.

significant 정보는 list of insignificant sets(LIS), list of insignificant pixels(LIP), 그리고 list of significant pixels(LSP)의 세 개의 list를 사용하여 부호화 된다. 이때, 부호화 과정은 significance map coding으로 significance 정보의 위치 및 계수의 부호(sign)를 부호화 하는 sorting pass 와 significance의 계수를 연속적으로 추정하는 refinement pass 두 단계로 이루어진다. Sorting과 refinement pass에 위해 발생된 비트는 이진 적응 산술 부호화(binary adaptive arithmetic coding)을 사용하여 엔트로피 부호화을 하였다[8].

5. 모의 실험 결과

제안한 임베이드 MC DCT 변환 부호화의 성능을 검증하기 위해 QCIF 크기의 carphone 스트림을 사용하여 휘도 성분에 대해서만 모의실험을 수행했다. 본 실험에서는 기존 방법과 제안 방법에 대해 bit-rate에 따른 PNSR을 인트라 프레임에 대하여 비교해 보았다(그림 6). 그리고 인터 프레임에 대한 실험은 첫 번째 인트라 프레임을 양자화 하지 않은 영상을 가지고 움직인 보상된 첫 번째 인터 프레임에 대하여 bit-rate에 따른 PSNR을 비교했다(그림 7).

실험 결과 1bpp 이하에서 부호화의 성능이 약간 떨어지지만 1bpp 이상에서는 성능이 더 우수한 것을 확인 할 수 있었다. 무엇보다도 제안 방법은 임베이드 부호화로 주어진 비트 발생량에서 코딩을 멈추어도 복호시 전체 프레임을 복원 할 수 있었다.

그림 8과 9는 H.263으로 부호화 한 후 복후화 한 영상이며 그림 10과 11은 제안 방법으로 H.263과 같은 bpp로 부호화후 복호화한 영상이다.

6. 결 론

본 논문은 낮은 비트 레이트에서도 고품질의 동영상 통신을 위해 비트율 제어가 용이한 효율적인 DCT기반 임베이드 영상 부호화기를 제안했다.

DCT와 웨이브렛 변환의 특성을 이용하여 DCT 계수를 3-Level 피라미드 구조로 재배치하고 임베이드 제로트리 부호화 기법으로 부호화 하여 성능의 개선과 비트율 제어가 용이함을 확인했으며 현재 1bpp이하에서도 부호화 성능을 높이기 위한 방법으로 전체 프레임에 임베이드 제로 트리 부호화를 적용하지 않고 매크로 블록 단위로 적용하여 에너지가 낮은 블록에 대해서는 부호화를 하지 않는 방법과 새로운 임베이드 제로 트리 부호화 알고리즘에 대해 연구하고 있다.

<참고문헌>

[1] J. M. Shapiro, "Embedded image coding using zerotree of wavelet coefficients," IEEE Trans. on Signal Processing, Vol. 41, No. 12, pp 3445-3456, Dec 1993.

[2] A. Islam and W. Pearlman, " An embedded and efficient low-complexity hierarchical image coder."

[3] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," IEEE Trans. Circuits and Systems for Video Technology, Vol. 6, No. 3, pp. 243-250, June 1996.

[4] Z. Xiong, O. Guleryuz, and M. T. Orchard, "A DCT-based Embedded Imge Coder," IEEE Signal Processing Letters, vol. 3, pp.289-290, Nov. 1996.

[5] W. B. Pennebarker and J. L. Mitchell, JPEG Still Image Data Compression Standard, New York: VanNostrand Reinhold, 1992.

[6] KHalid Sayood, Introduction to Data Compression, Morgan Kaufmann, 1996.

[7] M. Ve Herli and C. Herley "Wavelets and filter banks: Theory and design," IEEE Trans. Signal Processing, Vol. 40, no. 9, pp. 2207-2232

[8] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," Commun. ACM. vol. 30, pp. 520-540, June 1987.

[9] DRAFT ITU-T Recommendation H.263

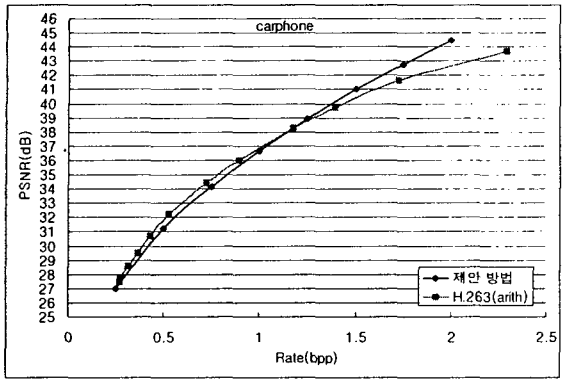


그림 6 bit rate에 따른 PSNR (INTRA)

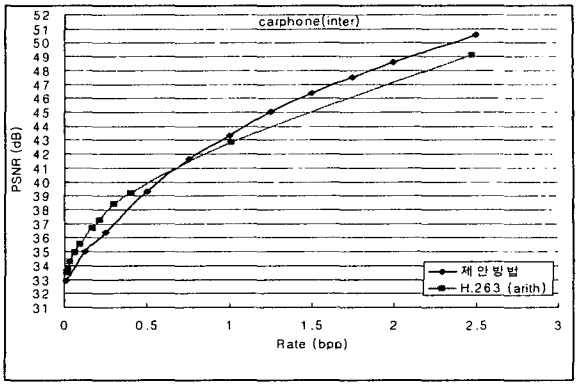


그림 7 bit rate에 따른 PSNR (INTER)

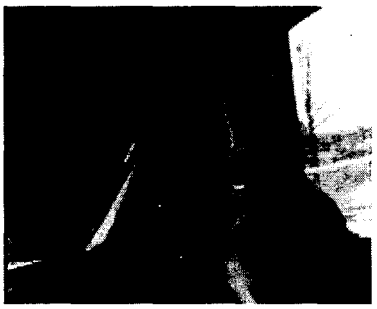


그림 8 H.263 psnr_y=38.25
(Intra I=5, 1.178bpp)

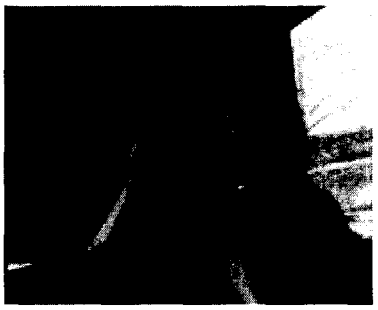


그림 9 H.263 PSNR_Y=35.12
(Inter I=0, q=10 0.067bpp)

그림 11 제안방법(Intra)
psnr_y=38.31, 1.178bpp

그림 10 제안방법(Inter)
(psnr_y=34.75, 0.067bpp)