



ABSTRACT

본 논문에서는 현재 ITU-T와 ISO에 의해서 공동으로 표준화 작업이 진행되고 있는 H.264 | MPEG-4 Part 10 표준의 주요 테스트 모델 사항들에 대해서 특징을 서술하고 그 기능을 설명하고자 한다. 대표적인 테스트 모델 사항으로서는 RD 최적화 및 변환을 이용한 부호 모드 결정 기법, 오류 내성을 가지는 부호 모드 결정 기법 등을 들 수 있다. 현재 대역폭 제어 및 고속 움직임 추정에 관한 사항들이 테스트 모델의 발전을 위한 주요 과제로 연구가 진행되고 있다.

1. 서론

현재 국제 멀티미디어 응용을 위한 동영상 분야의 표준으로는 대표적으로 H.263 과 MPEG-4 등을 들 수 있다. 이중 H.263은 ITU-T SG16 Q.15 (VCEG : 현재는 Q.6임)에 의해서 표준화가 진행

되어 왔으며, 이 단체는 H.263+ 및 H.263++에 이어서 차세대 멀티미디어 동영상 응용을 위한 표준으로서 H.264의 표준화 작업을 1998년부터 시작해 왔다.

ITU-T의 이러한 표준화 작업은 중도에 ISO/IEC JTC1 SC29 WG11 (MPEG)과의 공동작업으로 이어져, 2001년 11월 태국 Pattaya에서 열린 VCEG-O 회의에서 JVT(Joint Video Team)의 결성이라는 결실을 맺게 되었다. JVT는 현재 ITU | ISO 공동 표준안을 마련하기 위해서 표준화 작업을 계속해서 진행하고 있으며, 2002년 8월 현재 위원회 최종안 (FCD : Final Committee Draft)를 마련하기 위한 노력을 계속중이다[1][2].

본 논문에서는 위원회안 (CD : Committee Draft)를 기준으로 하여, 현재까지 JVT에서 진행된 표준안 작업중에서 테스트 모델에 해당하는 사항들에 대해서 설명하고자 한다. 특히 테스트 모델

의 여러 사항들 중에서도 성능에 많은 영향을 끼치며, 연구 가치가 높다고 여겨지는 최적화 모드 결정 기법과 오류가 있는 환경에서의 전송을 위한 부호기 최적화 기법에 집중하여 살펴보기로 한다. 이외에도 주요 테스트 모델 사항들 중의 하나인 고속 움직임 추정(fast motion estimation)과 대역폭 제어(rate control)에 대해서도 현재 JVT에서는 많은 노력을 들여서 성능을 개선시키기 위한 연구를 계속하고 있다.

2. 배경

JVT에서 진행하고 있는 표준화 작업을 위한 테스트 모델은 JM 3.9(Joint Model)까지 공표가 되어있는 상황이다. JM 3.9에서 구현되어 있는 테스트 모델에 관한 사항은 크게 다음과 같이 몇 가지로 나눌 수 있다. 첫번째는 RD 최적화 기법으로서 이는 압축 성능의 최적화를 위해서 레퍼런스 소프트웨어(reference software)에 포함되어 있는 사항이다. 최적화 기법은 저복잡도(low complexity) 모드 결정 기법과 고복잡도(high complexity) 모드 결정 기법으로 크게 나뉘어 질 수 있다. 저복잡도 모드 결정 기법은 RD 최적화 기법을 도입하지 않고 낮은 성능 사양을 가지는 시스템에서도 구현이 용이하도록 제안한 기법이며, 고복잡도 모드 결정 기법은 높은 성능 사양을 가지는 시스템에서 최대의 성능을 보이도록 RD 최적화 기법을 도입하여 제안한 개념이다.

두번째 사항은 오류방지 부호화 기법으로서 오류가 존재하는 채널 환경에서의 영상 압축 비트 스트림을 전송하게될 경우에 오류의 전파 방지를 통한 화질 개선이 그 주요 목적이다. 오류방지 부호화 기

법은 두 가지로 나뉘어지는데, 첫 번째 기법은 McubeWorks와 SKT가 공동으로 제안하여 채택된 드리프트 잡음 모델링 기법(Drift noise modeling)이며(3)(4)(5), 두 번째 기법은 독일의 HHI가 제안하여 채택되어 있는 다중 복호기 기법(Multiple decoder)이다. 드리프트 잡음 모델링 기법은 채널 환경에서 주어지는 오류의 확률을 기반으로 복호기에서 발생할 수 있는 오류의 전파를 수식적으로 전개하여, 이 결과에 기반하여 최적의 부호화 모드를 결정하는 기법이다. 또한 다중 복호기 기법에서는 부호기 측에서 복수개의 복호기를 내장하여 복호기 측에서 발생할 수 있는 오류의 전파를 실험적으로 파악하여 최적의 부호화 모드를 결정하는 기법이다.

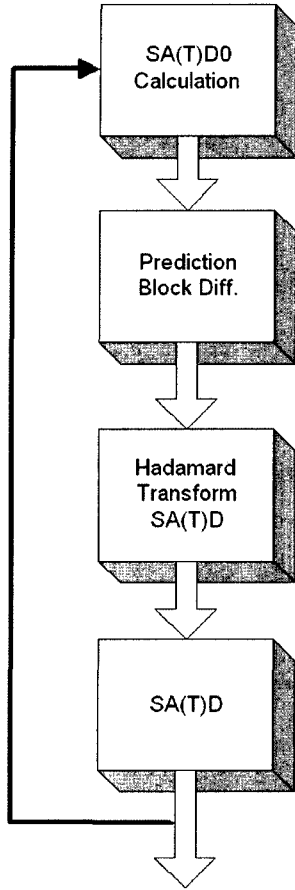
기타 주요 사항으로는 현재 활발하게 연구가 진행되고 있는 고속 움직임 추정 및 대역폭 제어등을 들 수 있다. 이러한 사항들에 대해서는 현재 JVT의 회원들간에 성능 개선을 위한 많은 연구가 진행되고 있다.

3. 최적의 부호 모드 결정 기법

최적의 부호 모드를 결정하기 위해서는 RD 최적화 기법 및 변환 기법 등이 사용되고 있다. 본 장에서는 이러한 요소들에 대해서 각각을 열거하고 설명한 후, 각각의 요소들을 사용하여 각각 저복잡도(low complexity) 및 고복잡도(high complexity) 모드에서의 최적 부호 모드를 결정하는 기법에 대해서 설명한다.

3.1 저복잡도 모드 결정 기법

저 복잡도 모드에서는 SAD(Sum of Absolute



〈그림 1〉 모드 결정의 개략적인 흐름도

Difference) 혹은 SATD(Sum of Absolute Transform Difference)를 사용하여 최적의 부호 모드를 결정하는 것을 제안하고 있다. 이는 RD 최적화를 통하지 않고, 단순히 잡음만을 결정의 요소로 사용하기 때문에 구현에 필요한 연산이 단순하게 줄어들게 된다. 전체적인 모드 결정의 사항을 〈그림 1〉에서 보여지는 바와 같게 된다.

3.1.1 SAD / SATD의 사용

매크로블록의 부호화 모드를 결정하기 위해서 움직임 벡터를 찾거나 Intra 모드(Intra mode)로 결

정하기 위한 기준의 하나로 SAD(Sum of Absolute Difference) 혹은 SATD(Sum of Absolute Transform Difference)를 사용하는 기법이 제안되어 있다.

모드 결정을 위한 과정의 시작에서 각각의 경우에 다음과 같이 SA(T)D0라는 값을 계산함으로써 시작하게 된다. SA(T)D0는 각 모드의 결정시에 모든 모드들이 생성해 내는 경우에 대해서 동일한 가중치를 주는 것을 피하기 위하여서 제안된 요소이다. SA(T)D0에 의해서 Intra 모드일 경우와 움직임 벡터를 찾는 경우가 〈표 1〉에서와 같이 각각 다르게 정의되게 된다.

〈표 1〉 각 모드에 따른 SA(T)D0의 값

SA(T)D0	
Intra	QPo(QP) x PM
Motion Vector Search	QPo(QP) x (BC + 2 x CN)

QPo(QP) : H.263에서의 QP값을 JM에서의 해당값으로 환산한 가중치

PM : 각각의 예측모드에 따른 가중치

BC : 해당 움직임벡터를 부호화하는데 드는 비트의 수

CN : 레퍼런스 프레임을 부호화하는데 사용되는 부호어의 번호

이때 SKIP 모드가 결정되는 것과 너무 많은 숫자의 구획이 Intra 4x4 모드로 부호화되는 것을 막기 위해서, 다음의 두 경우에는 아래에 설명된 바와 같이 예외를 가지게 된다.

예외 1) Intra 모드일 경우
 Intra 16x16 모드가 추천되는 움직임 벡터가 0 벡터일 경우에는 Intra QPo(QP)의 값을 SATD로부터 계산한 값을 Intra 모드에 해당하므로 부호화 하는바 같은 바

트를 절약할 수 있는 경우에 해당하여, 이 경우에 더 높은 기중치를 두기 위함이다.

예외 2) Intra 4x4 모드일 경우

모든 Intra 4x4 모드의 경우에는 16x QPo(QP)에 해당하는 값을 더해준다. 이것은 실험적으로 결정된 값이며, 필요이상으로 많은 개수의 구획이 Intra 4x4 모드로 부호화 되는 것을 막기 위한 조치이다.

- SAD의 계산

움직임 결정 및 Intra 모드 부호화를 위해서는 식 (1) 같이 구해지는 Diff값을 이용해서 SAD 혹은 SATD를 구하여 사용한다. 식 (1)에서 설명된 값은 움직임 추정 혹은 공간 예측(spatial prediction)을 통해서 예측된 값과 원영상의 화소값과의 차이를 나타낸다.

$$Diff(i,j) = Original(i,j) - Prediction(i,j) \quad (1)$$

움직임 벡터를 구하는 경우에 있어서 정수화소(integer pel)의 위치에서는 계산의 복잡도를 고려하여 SATD가 아닌 SAD를 사용할 것을 제안하고 있다. SAD는 식 (2)에서 정의된 바와 같이 구해진다.

$$SAD = \sum_{i,j} |Diff(i,j)| \quad (2)$$

- SATD의 계산

움직임 추정시에 정수 화소가 아닌 분수 화소(fractional pel)의 위치에서의 모드 결정에는 SAD가 아닌 SATD를 사용하게 된다. 이것은 H.264 | MPEG-4 Part 10 또한 기존의 국제 동영상 부호화 표준과 같이 잔류 신호(residual signal)를 변환한 후에, 변환 계수를 부호화하는 기법을 취하고 있기

때문이다. 즉, 식 (2)에 의해서 계산된 SAD를 모드 결정의 기준으로 삼을 경우, 변환된 계수의 특성을 완전하게 반영하지 못하여 최적의 움직임 벡터 혹은 공간 예측 모드를 얻어내는 것이 용이하지 않을 수 있다는 문제점을 해결하기 위해서 제안한 것이다. 따라서 변환에는 H.264 | MPEG-4 Part 10에서 채택한 정수 변환(integer transform)을 채택하여 사용하는 것이 보다 최적의 모드를 결정하는데 더 효과적 이겠으나, SATD를 사용할 경우에 야기될 수 있는 구현상에서의 복잡함을 제거하기 위해서 식 (3)에서 정의된 바와 같은 기저(Kernel)를 가지는 하다마드 변환(Hadamard transform)을 사용할 것을 제안하고 있다. 식 (3)에서 정의된 바와 같은 기저를 가지는 하다마드 변환을 2차원으로 행하여 DiffT를 구한후 최종적으로 SATD를 구하여, 이 값을 SA(T)D0와 더하여 모드 결정을 행할 것이 제안되어 있다.

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad (3)$$

식 (3)에서 정의된 기저를 이용하여, 식 (4)에서와 같이 DiffT를 구할 수 있다.

$$DiffT(x,y) = H [Diff(i,j)] \quad (4)$$

$(x, y = 0 \dots 3 \quad i, j = 0 \dots 3)$

H[] 하다마드 변환 연산자 (Hadamard transform operator)

3.1.2 모드 결정

상기한 식 (1)~식 (4)를 통해서 구해진 SAD 혹은

은 SATD를 이용하여 최종적으로 식 (5)와 같이 SA(T)D와 SA(T)D0의 값의 합을 최소화하는 모드를 최적의 모드로 결정할 것이 제안되어 있다.

$$SA(T)D_{min} = \min(SA(T)D + SA(T)D0) \quad (5)$$

3.2 고복잡도 모드에서의 모드 결정 기법

본 절에서는 고복잡도 모드(high complexity mode)에서의 모드 결정 기법에 대해서 서술하고자 한다. 고복잡도 모드에서는 저복잡도 모드에서 고려하지 않았던 부호화에 필요한 비트수의 개념을 도입하여 RD(Rate-Distortion) 최적화 기법을 사용한다. 고복잡도 모드에서는 복잡도가 별로 문제가 되지 않는 경우, 예를 들면 연산 능력(computational power)이 충분히 주어지는 경우에 보다 나은 압축 및 오류 방지 성능을 얻어내기 위해서 사용하게 된다.

- 고복잡도 움직임 벡터 추정기법

정수 화소단위의 움직임 추정에서는 고복잡도 모드에서도 MC_range라는 변수로 주어진 값을 사용한 추정 범위내에서 H.263에서 제안되었던 기법과 유사한 나선형(spiral) 구조로 추정을 실시한다. 이 때 모든 Inter 모드에서는 16×16 크기의 구획이 가지는 예측 벡터를 나선형 추정의 중심점으로 삼아서 복잡도를 낮출 것을 제안하고 있다. 현재 구현되어 있는 구조는 매크로블록을 이루는 4×4 크기의 구획에 대해서 각각 SAD를 구한후 이를 사용하여 단계적으로 더 큰 크기의 구획에 대해서 추정을 실시하여 구현시에 필요한 연산을 줄이는 구조로 되어있다. 고복잡도 모드에서는 정수 화소뿐만 아니라, 분할 화소에서도 식 (6)에서 정의된 바와 같이 RD 최적화 기법을 사용하여 최적의 움직임 벡

터를 추정한다.

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} \cdot R(m-p)$$

$m = (m_x, m_y)^T$: 움직임 벡터

$p = (p_x, p_y)^T$: 예측 벡터 (6)

λ_{MOTION} : Lagrangian 계수 ($= \sqrt{0.85 \cdot 2^{p \cdot 1.3}}$)

$R(m-p)$: 최종적으로 부호화되어야 할 움직임 정보의 비트수

이 때 식 (6)에서 SAD는 식 (7)과 같이 구해진다.

$$SAD(s, c(m)) = \sum_{x=1, y=1}^{B, B} |s[x, y] - c[x - m_x, y - m_y]| \quad (7)$$

SATD는 3.1.1에서 설명한 바와 같은 과정을 통하여 하다마드 변환을 이용해서 구해지게 된다. 이 때 $J(m, \lambda_{MOTION})$ 을 최소화 하는 값을 구함으로써 RD 최적화를 이용한 최적의 움직임 벡터를 구할 수 있게 된다.

- 최적 레퍼런스 프레임 결정 기법

최적의 움직임 벡터를 결정하는 기법을 연장하여 Inter 모드에서의 최적의 NxM 모드와 최적의 성능을 나타내는 레퍼런스 프레임을 식 (8)에서 구해진 $J(REF | \lambda_{MOTION})$ 값을 최소화하는 과정을 통하여 구현해 낼 수 있다.

$$J(REF | \lambda_{MOTION}) = SATD(s, c(REF, m(REF))) + \lambda_{MOTION} \cdot (R(m(REF) - p(REF)) + R(REF)) \quad (8)$$

- 최적 매크로블록 부호화 모드 결정

상기한 최적 움직임 벡터 탐색 및 최적 레퍼런스 프레임 결정 기법은 매크로블록에서의 최적의 부호

화 모드 결정 과정에 유사한 연장을 통하여 식 (9)에서와 같이 유사하게 적용될 수 있다.

$$J(s, c, MODE | QP, \lambda_{MOTION}) = SSD(s, c, MODE | QP) + \lambda_{MOTION} \cdot R(s, c, MODE | QP) \quad (9)$$

식 (9)에서 λ_{MOTION} 는 모드 결정을 위한 Lagrangian 계수를 지칭하며, MODE는 가능한 부호화 모드들의 집합들의 원소들을 표시하는 변수이며, <표 2>에 설

< 표 2 > 프레임 형식에 따라 MODE가 가질 수 있는 각각의 값들

Intra 프레임	MODE ∈ {INTRA 4x4, INTRA 16x16}
Inter 프레임	MODE ∈ {INTRA 4x4, INTRA 16x16, SKIP, 16x16, 16x8, 8x16, 8x8}

명된 것 같은 값들을 가질 수 있게 된다.

또한 식 (9)에서 정의된 SSD(Sum of Squared Difference)는 식 (10)과 같이 구해진다.

$$SSD(s, c, MODE | QP) = \sum_{x=1, y=1}^{16, 16} (s_r[x, y] - c_r[x, y, MODE | QP])^2 + \sum_{x=1, y=1}^{8, 8} (s_b[x, y] - c_b[x, y, MODE | QP])^2 + \sum_{x=1, y=1}^{8, 8} (s_g[x, y] - c_g[x, y, MODE | QP])^2 \quad (10)$$

또한 식 (9)에서 설명된 $R(s, c, MODE | QP)$ 는 해당 MODE와 QP를 선택하여 매크로블록을 부호화했을 경우에 사용되는 비트의 수이다. 이 값은 매크로블록의 헤더, 움직임 벡터 및 변환계수들을 부호화했을 경우에 사용되는 값들을 포함하고 있다.

$c_r[x, y, MODE | QP]$ 와 $s_r[x, y]$ 는 각각 복원된 영상과 원영상의 휘도성분값을 나타낸다. 마찬가지로 c_b, c_g 와 s_b, s_g 는 각각 해당 색상 성분을 지칭한다.

그리고 Lagrangian 계수는 식 (11)에서 정의된 것과 같은 값을 사용할 것을 제안하고 있다.

$$\lambda_{MODE} = 0.85 \cdot 2^{QP/2} \quad (11)$$

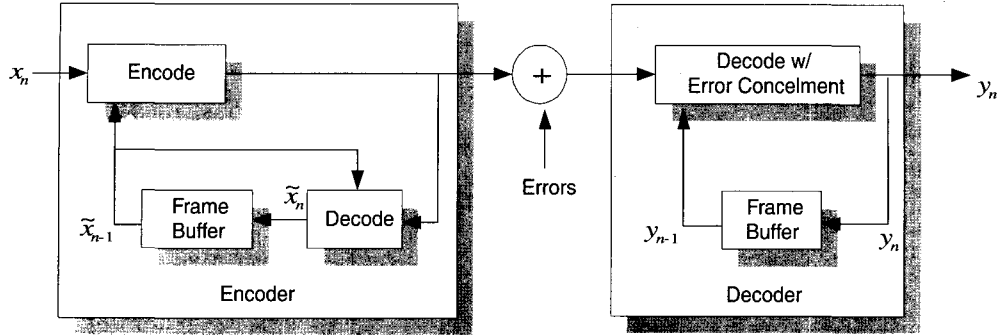
- 8x8 모드 결정

현재 H.264 | MPEG-4 Part 10은 8x8 구획을 하위 구획 (8x4, 4x8, 4x4)로 나누어 움직임 추정하는 것을 허용하고 있다. 따라서 8x8 구획으로 결정된 구획에 대해서 하위 구획으로 나누어 부호화하는 것을 결정하는 것도 이미 소개한 RD 최적화 기법을 연장하여 적용하는 것이 가능하다. 이와 같은 이유로 인하여 8x8 모드 결정은 매크로블록 모드 결정기법과 매우 비슷한 방법으로 진행된다. 식 (12)는 식 (9)에서 매크로블록의 최적 구획을 결정하였듯이, 8x8으로 결정된 각각의 구획에 대해서 이를 최적의 하위 구조의 구획으로 나누어 결정하는 것이 가능하다는 것을 보여주고 있다.

$$J(s, c, MODE | QP, \lambda_{MOTION}) = SSD(s, c, MODE | QP) + \lambda_{MOTION} \cdot R(s, c, MODE | QP) \quad (12)$$

기존의 모드 결정 기법에서와 같이 QP는 매크로블록에 사용되는 QP값이며, λ_{MODE} 는 각 모드에 대해서 사용되는 Lagrangian 계수이다. 또한 매크로블록에서와 같이 MODE는 모드 결정시에 결정이 가능한 모든 모드들의 집합 중에서 선택되게 된다. 이때 기준이 되는 모드는 8x8 크기의 구획이므로 선택이 가능한 모드들의 집합은 8x8 크기의 구획을 포함하여 식 (13)에서 보이는 것과 같은 요소들을 가지게 된다.

$$MODE \in \left\{ \begin{array}{l} INTRA \ 4x4 \\ 8x8, \ 8x4, \ 4x8, \ 4x4 \end{array} \right\} \quad (13)$$



〈그림 2〉 부호기 / 복호기 모델의 개략도

4. 오류방지 부호화 기법

본 절에서 서술하고자 하는 기법은 전송오류가 존재하는 채널에서의 오류의 전파를 최대한 방지하고자 부호기에서 효율적으로 부호화를 하는 기법들이다. 현재 JM에는 두 가지의 기법을 통하여 이를 달성하고자 하는 노력을 하고 있다. 첫번째 기법은 엠큐브웍스와 SK Telecom에서 공동으로 제안한 드리프트 잡음 모델링 기법이며 다른 하나는 독일의 HHI에서 제안한 다중 복호기 기법이다. 첫번째 기법은 오류의 전파를 채널 환경에서의 오류 확률을 기반으로 하여, 잡음의 전파과정을 모델링 하여 이를 최소한으로 줄이는 것을 제안한 기법이다. 반면에 두 번째 기법은 부호기에서 복수개의 복호기를 가정하고 이들 각각이 오류 환경이 있는 채널을 통해서 전송된 비트스트림을 수신하였을 경우에 발생할 수 있는 오류 현상을 통계적으로 분석하여 오류 전파를 최소한으로 줄일 수 있는 기법을 제안한 것이다. 본 장에서는 이들 기법 각각을 개략적으로 서술하고자 한다.

4.1 드리프트 잡음 모델링 기법

전송 오류가 발생하는 채널을 가정하는 경우, 오

류 전파의 가장 큰 원인은 부호기와 복호기에서 존재하고 있는 복원 영상의 차이에 기인하게 된다. 개략적으로 부호기와 복호기 사이에 오류가 존재할 수 있는 전송채널이 존재하는 경우의 블록도를 〈그림 2〉에서와 같이 가정할 수 있다.

〈그림 2〉의 블록도에서 채널 오류에 기인하여 영상 저하의 가장 큰 원인이 되는 부호기와 복호기간의 복원영상의 차이를 드리프트 잡음(drift noise)라고 정의하면, 복원 영상이 가지는 잡음의 양은 식 (13)에서와 같이 양자화 잡음(quantization noise)과 드리프트 잡음(drift noise)의 합으로 나타나게 된다(3).

$$\begin{aligned}
 D_{total}^n &= E(x_n - y_n)^2 \\
 &= E(x_n - \tilde{x}_n + \tilde{x}_n - y_n)^2 \\
 &= E(x_n - \tilde{x}_n)^2 + (\tilde{x}_n - y_n)^2 = D_q^n + D_{drift}^n
 \end{aligned}
 \tag{14}$$

식 (14)에서 보여진 D_q^n 와 D_{drift}^n 은 각각 식 (15)에서와 같이 정의되는 양자화 잡음과 드리프트 잡음을 나타낸다.

$$\begin{aligned}
 D_q^n &= E(x_n - \tilde{x}_n)^2 \\
 D_{drift}^n &= E(\tilde{x}_n - y_n)^2
 \end{aligned}
 \tag{15}$$

4.2 RD-최적화된 부호 모드 결정 방법

최적의 부호 모드는 식 (16)에서 보는 바와 같이 총 잡음과 부호화에 필요한 비트수 간의 상관관계를 고려하여 결정하게 된다.

$$Cost = D_{total}^n + \lambda_m \cdot R \quad (16)$$

여기서 R 은 부호화에 필요한 비트 수이며, Lagrangian 계수 λ_m 은 영상을 실제로 부호화하기에 앞서 그 부호화 파라미터(양자화 계급 등)에 따라 미리 결정되는 상수이다.

4.3 드리프트 잡음의 추정과 갱신

드리프트 잡음은 복원 영상을 통한 움직임 추정 및 보상 과정에 의해서 계속적으로 프레임간에 전파되게 되고, Intra 모드로 부호화 되기 전까지는 계속적으로 존재하게 된다.

이러한 사실에 기초하여 Inter, Inter 및 SKIP 모드일 경우의 각각의 드리프트 잡음은 식 (17)~식 (19)에서와 같이 표현되게 된다.

정상적인 Inter 모드로 부호화된 구역의 경우

$$D_{drift}^n = E(\tilde{x}_n - y_n)^2$$

$$= E([x_n - \tilde{x}_{n-1}(mv) + q_r] + \tilde{x}_{n-1}(mv) - [x_n - \tilde{x}_{n-1}(mv) + q_r] - y_{n-1}(mv))^2 \quad (17)$$

$$= E(\tilde{x}_{n-1}(mv) - y_{n-1}(mv))^2$$

$$= D_{drift}^{n-1}(mv)$$

정상적인 Intra 모드로 부호화된 구역의 경우

$$D_{drift}^n = E(\tilde{x}_n - y_n)^2$$

$$= E(x_n + q_a - x_n - q_a)^2 \quad (18)$$

$$= 0$$

오류가 발생한 구역의 경우:

$$D_{drift}^n = E(\tilde{x}_n - y_n)^2$$

$$= E(\tilde{x}_n - y_{n-1})^2$$

$$= E(\tilde{x}_n - \tilde{x}_{n-1} + \tilde{x}_{n-1} - y_{n-1})^2$$

$$= E(\tilde{x}_n - \tilde{x}_{n-1})^2 + D_{drift}^{n-1} \quad (19)$$

식 (17)~식 (19)에서 정의된 바와 같이 드리프트 잡음이 구해질 때, 각 경우에 드리프트 잡음의 기대치는, 패킷 손실 확률이 일 때 식 (20)~식 (22)와 같이 추정될 수 있다[5].

- Inter 모드의 경우,

$$D_{drift}^n = E(\tilde{x}_n - y_n)^2$$

$$= (1-p) \cdot E(\tilde{x}_{n-1}(mv) - y_{n-1}(mv))^2$$

$$+ p \cdot \{E(\tilde{x}_{n-1} - y_{n-1})^2 + E(\tilde{x}_n - \tilde{x}_{n-1})^2\}$$

$$= \{(1-p) \cdot D_{drift}^{n-1}(mv) + p \cdot D_{drift}^{n-1}\}$$

$$\cdot C_m(p, n - n_0) + p \cdot C_o(QP) \cdot (\tilde{x}_n - \tilde{x}_{n-1})^2 \quad (20)$$

$C_m(p, n - n_0)$: 움직임 영역에서의 확산계수

- Intra 모드의 경우,

$$D_{drift}^n = E(\tilde{x}_n - y_n)^2$$

$$= p \cdot \{E(\tilde{x}_{n-1} - y_{n-1})^2 + E(\tilde{x}_n - \tilde{x}_{n-1})^2\}$$

$$= p \cdot D_{drift}^{n-1} + p \cdot C_o(QP) \cdot (\tilde{x}_n - \tilde{x}_{n-1})^2 \quad (21)$$

-SKIP 모드의 경우

$$D_{drift}^n = E(\tilde{x}_n - y_n)^2$$

$$= (1-p) \cdot E(\tilde{x}_{n-1} - y_{n-1})^2$$

$$+ p \cdot \{E(\tilde{x}_{n-1} - y_{n-1})^2 + E(\tilde{x}_n - \tilde{x}_{n-1})^2\}$$

$$= C_s(p, n - n_0) \cdot D_{drift}^{n-1} \quad (22)$$

$C_s(p, n - n_0)$: 정지 영역에서의 확산 계수

각 영역에서의 확산계수들은 시간적인 진행에 따

라서 식 (23)과 식 (24)에서 정의한 바와 같이 지수 함수적으로 증가하도록 정의하여 사용한 결과 좋은 성능을 얻을 수 있음이 제안된 바 있다[5].

$$C_s(p, n - n_0) = \exp \left\{ p \left[1 + \frac{n - n_0}{M} \right] \right\} \quad (23)$$

$$C_m(p, n - n_0) = \exp \left\{ p \left[1 + 4 \cdot \frac{n - n_0}{M} + 4 \cdot \left(\frac{n - n_0}{M} \right)^2 \right] \right\} \quad (24)$$

여기서 $M = \begin{cases} 9 & QCIF \\ 18 & CIF \end{cases}$

식 (22)에서 주목할 만한 점은 정지영역으로 판단되어 SKIP 모드로 결정되는 구역에도 $C_s(p, n - n_0)$ 이라는 확산 계수를 정의하여, 정지영역이라 할 지라도 드리프트 잡음의 값이 프레임의 진행과 더불어 계속해서 증가해가도록 식 (23)에서와 같이 $C_s(p, n - n_0)$ 를 정의하고 있다. 이러한 이유는 정지영역을 Intra로 부호화할 경우 오류가 발생하게 되면, 이러한 오류의 대처 방법으로서 시간적인 은닉(temporal concealment)을 사용하는 것을 가정하게 될 경우, 은닉을 위한 예측 구획의 화소값들이 복원 영상의 해당 구획의 값들과 일치함으로써, 드리프트 잡음이 존재하지 않는 상황으로 인식하게 되고, 이는 결과적으로 오류가 발생한 구획을 계속해서 Intra로 갱신하지 않고 오류가 발생한 상태로 지속시키는 결과를 가지고 오게 되기 때문이다.

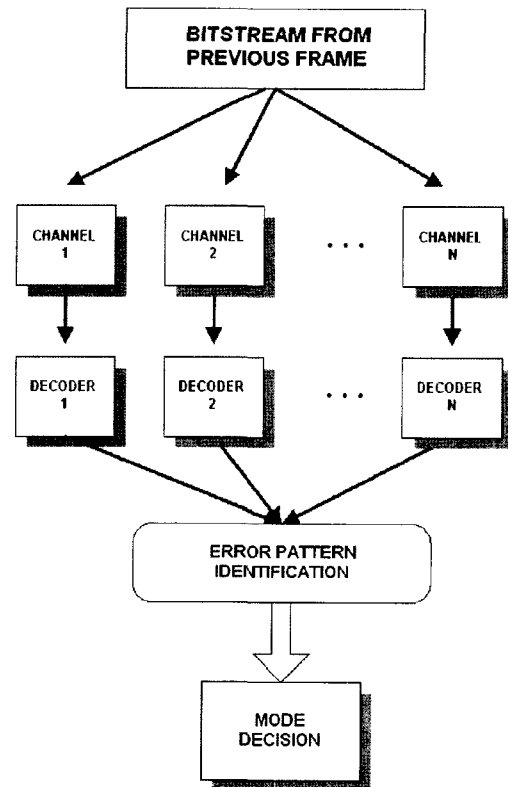
최종적으로 어떠한 구역을 Intra 모드로 결정하게 되면 식 (25)에서와 프레임 인덱스를 초기화 시켜주는 작업을 시행한다.

$$n_0 = n \quad (25)$$

4.2 다중 복호기 기법

4.1 절에서는 드리프트 잡음을 모델링하여 이를 오류 확률 p를 통하여 open loop로 추정하여 각 구역에서의 최적화된 부호화 모드를 결정하는 기법을 설명한 바 있다. 본 절에서는 부호기측에서 복수개의 복호기를 가정하여, 오류가 있는 환경을 가정하여, 복호기에서 수신이 가능한 비트열(bit stream)을 실험적으로 재생해 내고, 이를 통하여 얻어지는 오류의 현상을 바탕으로 최적화된 부호 모드를 결정하게 된다.

이와 같은 일련의 과정을 나타내는 과정은 <그림 3>에서 보는 바와 같다. <그림 3>에서 부호기는 이



<그림 3> 다중 복호기 기법의 개략도

전 프레임을 부호화한 비트 스트림을 복수개의 채널 시뮬레이터(channel simulator)를 통과시킨 후 복수개의 복호기로 복호 과정을 거쳐서 해당 프레임에서의 오류 발생 정보를 실험적으로 예측한다. 이렇게 실험적으로 예측된 오류 정보를 통하여 현재 프레임에서 취할 수 있는 최적의 부호화 조치를 취하게 된다.

5. 테스트 모델 영역에서의 연구 과제

주요한 테스트 모델 작업중의 하나로서 고속 움직임 추정(fast motion estimation)에 관한 연구가 계속되고 있다. 가변 구획 크기(Variable Block Size)를 채택하는 JVT 부호기의 특성상, 기존의 국제 동영상 부호 표준들에 비해서 움직임 추정이 더 복잡한 연산을 필요로 하는 것이 사실이다. 이러한 문제점을 해결하기 위해서 정수 화소 위치에서의 고속 움직임 추정 및 분수 화소 위치에서의 움직임 추정에 관한 제안들이 다수 제안된 바 있으며, 이에 대한 연구도 계속해서 진행되고 있다.

또한 대역폭 제어(rate control)에 관한 연구도 주요한 진행 작업중의 하나이다. 현재 2002년 7월에 오스트리아 Klagenfurt 회의에서 마이크로소프트가 제안한 대역폭 제어 기법을 중심으로 이를 발전시키기 위한 연구가 진행되고 있다(6).

이미 설명한 고속 움직임 추정, 대역폭 제어 그리고 오류 내성을 위한 모드 결정방법등을 포함한 테스트 모델의 성능 향상을 위한 연구를 원활하게 진행하기 위해서 JVT 산하에 기술 분과 위원회(Ad-Hoc Group : chair Chul-Woo Kim, McubeWorks Inc.)가 결성되어서, 이와 관련된 연구를 계속해서 진행해 나가고 있다.

6. 결론

본 논문에서는 현재 ITU와 ISO에서 공동으로 표준화 작업이 진행중인 JVT 부호기의 테스트 모델에 관한 사항들에 관하여 자세히 서술하였다. 현재 JVT에서 테스트 모델 분야에서 주요하게 관심을 가지고 진행시켜 왔던 분야는 상기한 바와 같이 최적의 부호화 기법, 오류 내성을 가지는 모드 결정 기법등이며, 현재에도 고속 움직임 추정 기법 그리고 효율적인 대역폭 제어 기법등에 대한 연구를 활발하게 진행해 오고 있다.

최적화 기법은 크게 나누어서 저복잡도(low complexity) 모드 결정 기법과 고복잡도(high complexity) 모드 결정 기법으로 크게 나뉘어 지며, 고복잡도 모드 결정 기법에서는 최대의 성능을 보이도록 RD 최적화 기법을 도입하여 제안한 개념이 포함되어 있다.

오류방지 부호화 기법 또한 크게 두 가지로 나뉘어 지는데, 드리프트 잡음 모델링 기법(Drift noise modeling)으로서, 채널 환경에서 주어지는 오류의 확률을 기반으로 복호기에서 발생할 수 있는 오류의 전파를 드리프트 잡음의 개념을 도입하여 모델링함으로써 최적의 부호화 모드를 결정하는 기법이다. 다중 복호기 기법에서는 부호기 측에서 복수개의 복호기를 내장하여 복호기측에서 발생할 수 있는 오류의 전파를 실험적으로 파악하여 최적의 부호화 모드를 결정하는 기법임을 본문에서 설명하였다.

또한 현재 표준화의 최종적인 성과를 보기 위하여 박차가 가해지고 있는 이 시점에서, 고속 움직임 추정 기법과 대역폭 제어에 관한 기법은 여러 전문가들의 관심을 받고 있으며, 연구의 대상으로서도 그 가치가 높다고 판단된다.

참고 문헌

- (1) Gisle Bjontegaard, "H.26L Test Model Long Term Number 8 (TML-8) draft 0," June, 2001.
- (2) Thomas Wiegand & Gary Sullivan, "Working Draft Number 2, Revision 8, JVT-G112r3," April, 2002.
- (3) Chul-Woo Kim et. Al., "High complexity mode decision for error prone channel," ITU-T SG16/Q6 Contribution: VCEG-O43, Nov. 2001.
- (4) Chul-Woo Kim et al., "High complexity mode decision for error prone channel," ITU-T SG16/Q6 Contribution: JVT-C101, May 2002
- (5) Chul-Woo Kim et al., "High complexity mode decision for error prone channel (revised)," ITU-T SG16/Q6 Contribution: JVT-D118, Jul. 2002
- (6) Siwei Ma et al., "Rate control on JVT standard," ITU-T SG16/Q6 Contribution: JVT-D030, Jul. 2002

필자 소개



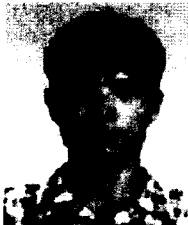
김철우

-1987년 3월~1991년 2월 : 서울대학교 전자공학과 학사
 -1991년 3월~1993년 2월 : 서울대학교 전자공학과 석사
 -1993년 3월~1998년 2월 : 서울대학교 전자공학과 박사
 -1998년 3월~2000년 3월 : LG 전자 DM연구소 선임연구원
 -2000년 3월~2000년 12월 : ACS Technology 책임연구원
 -2001년 1월~현재 : (주)엠큐브웍스 기술이사
 -주관심분야 : H.263 및 H.26L, MPEG-4 비디오 표준화 및 멀티미디어 통신 시스템



김후종

-1988년 2월 : 서강대학교 전자공학과 학사
 -1995년 8월 : 서강대학교 전자공학과 석사
 -2000년 2월 : 국민대학교 전자공학과 박사 재학 중
 -1988년 1월~1995년 8월 : LG정보통신 이동통신연구단 선임연구원
 -1995년 9월~현재 : SK 텔레콤 Terminal 개발 팀장
 -주관심분야 : 이동통신 시스템, Wireless Terminal



황인성

-1993년 : 인하대학교 전자공학과 졸업 학사
 -1995년 : 인하대학교 전자공학과 대학원 졸업 석사
 -1995년 1월~1999년 11월 : LG전자 멀티미디어사업본부 평택연구소 주임연구원
 -1999년 11월~현재 : SK텔레콤 과장
 -주관심분야 : H.263, H.26L, MPEG-4 비디오 표준화 및 3GPP 멀티미디어 표준