

신연재

MPEG-4 비디오 압축 표준

호요성

광주과학기술원 정보통신공학과 교수

1. MPEG-4 표준의 특징

최근 컴퓨터, 통신, 디지털 신호 처리 기술이 발전함에 따라 인간의 시각과 청각을 이용한 디지털 멀티미디어 정보 서비스들이 급속도로 확산되고 있다. 이에 발맞추어 국제 표준화 기구인 ISO/IEC의 JTC1/SC29/WG11에서는 MPEG-1 표준과 MPEG-2 표준을 제정하여 영상을 중심으로 한 비디오 CD, 영상 회의, 원격 강의, 주문형 비디오, 디지털 TV, 고선명 TV 등 여러 응용 분야에 성공적으로 이용하고 있다. 이러한 MPEG-1 표준과 MPEG-2 표준은 디지털 비디오/오디오 신호를 디지털 저장 매체에 효율적으로 저장하거나, 일정한 전송율을 가지는 통신로를 이용하여 디지털 정보를 경제적으로 전송하는데 중점을 두었다.

하지만 이러한 표준들은 지금까지 서로 다른 영역으로 나뉘어져 있던 방송, 통신과 컴퓨터의 융합으로

이루어진 멀티미디어 표준이라고 할 수 없다. 따라서, 이에 대응하는 새로운 표준 방식의 필요성에 따라 MPEG-4라고 불리는 표준화 작업이 진행되었다. MPEG-4의 정식 프로젝트 명칭은 Very Low Bitrate Audio Visual Coding 이다. 그림 1은 각 표준이 지원하는 전송 비트율에 대한 복원 영상의 화질을 도시하였다.

기존의 MPEG-1과 MPEG-2 방식에서는 입력 영상의 내용과는 무관하게 정해진 데이터를 부호화 능력에 따라 압축 전송하였다. 그러나 이러한 방식은 화소값만을 직접 처리하므로, 기능적 측면에서 많은 제약이 따르기 때문에 현재의 많은 요구들을 모두 충족시킬 수 없다. 따라서 MPEG-4에서는 영상의 내용을 바탕으로 하는 내용기반 부호화

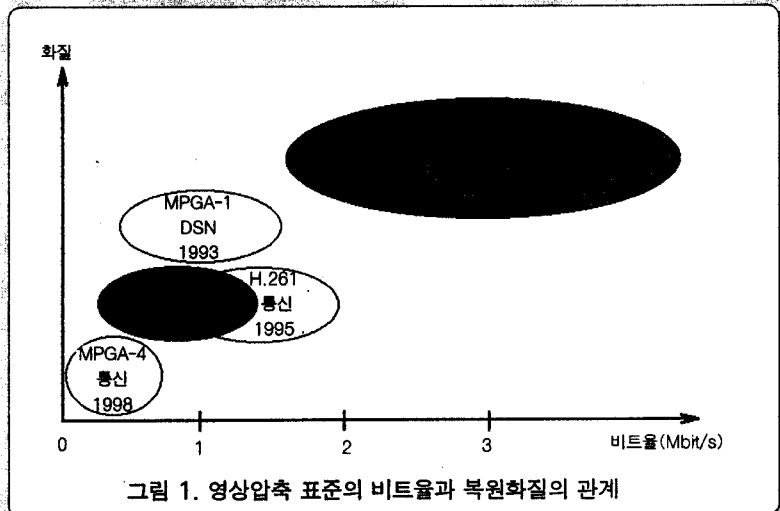


그림 1. 영상압축 표준의 비트율과 복원화질의 관계

(Content-based Coding)에 초점을 맞추어 표준화가 진행되고 있다. MPEG-4는 장면에서 뽑아낸 오디오 비디오(Audio-Visual, AV) 데이터를 부호화, 복호화 과정을 통하여 특정한 부분의 객체를 합성하여 장면을 복원할 수 있다. 이러한 객체 단위 부호화는 다양한 멀티미디어 응용에서 객체 사이의 상호 관계를 정립하여 이들의 접근과 조작을 수월하게 하여 사용자의 요구에 따른 장면의 구성이 가능하게 한다. 이러한 기능을 제공하기 위하여 MPEG-4는 내용기반 대화형 기능(Content-based Interactivity), 초고속 압축 기능, 광범위한 접속기능(Universal Access)의 3가지 방향에 중점을 두고 표준화가 이루어지고 있다.

MPEG-4표준에서는 부호화 대상이 되는 객체도 풍경, 인물, 음성과 같은 자연적인 것뿐만 아니라 미디어 음원, 합성 음원과 같은 인위적으로 합성된 것도 처리할 수 있다. 따라서 노래방 기기처럼 전자음악과 미디어, 영상과 자막 등이 섞인 것들도 MPEG-4를 이용하여 간단하게 실현할 수 있을 것이다.

기존의 MPEG-1, MPEG-2 방식이 주로 하드웨어로 구성하는 것을 전제로 하였으나, MPEG-4는 소프트웨어로 구현함을 기본으로 한다.

또한 MPEG-4는 손상된 정보의 복원 능력을 기존의 표준보다 더욱 향상시키는데 관심을 두고 있다. 이러한 특징이 가능하도록 하기 위하여 비트열의 형태로 조절하는 다중화(Multiplexing), 동기화(Synchronizing) 등의 방식을 결정했다. MPEG-4의 또 다른 특징은 전송율을 가변적으로 변화시킬 수

있다는 것이다. 따라서 고속 전송 망이든 저속 전송 망이든 관계없이 제공자가 압축 비율을 달리 함으로써, 전송이 가능하다. 아울러 MPEG-4 표준화 활동에서는 일반 전화망이나 이동 통신망과 같이 전송 주파수 대역이 아주 작은 채널에서도 디지털 비디오/오디오 통신 서비스를 제공할 수 있도록 초당 64 kbps 이하의 매우 낮은 비트율을 갖는 부호화 방식을 개발하고 있다.

현재까지의 연구에 따르면 8~64 kbps의 낮은 전송속도에서도 전송이 가능하다. 현재 MPEG-4 비디오 그룹은 5~40 kbps 대역까지의 적용을 목적으로 크게 형태 정보 부호화(Shape Information Coding), 움직임 정보 부호화(Motion Information Coding), Texture 정보 부호화(Texture Information Coding)와 관련된 도구들을 개발하고 있다.

이러한 특징을 갖는 MPEG-4 표준은 제작자(Author)와 사용자(User) 그리고 서비스 제공자(Service Provider)들의 요구를 모두 수용하여야 할 것이다.

2. MPEG-4 자연영상 부호화의 요구사항

여기서는 일반적인 MPEG-4의 비디오부의 표준을 위한 요구사항을 알아보기로 하겠다. 요구사항은 자연영상, 합성영상의 두가지 부분으로 나누어져 있다. 본 절에서는 자연영상에 대해 살펴보기로 한다.

*

(a) 객체 기반 표현 : MPEG-4

표준은 임의의 모양으로 나타나는 비디오 객체의 합성으로 이해되는 비디오 장면 표현 방법을 제공할 것이다. MPEG-4는 한 장면내의 개별 객체가 다른 매개변수로서 또는 다른 품질과 다른 부호화 알고리즘으로 부호화 될 수 있는 수단을 제공할 것이다.

(b) 비디오 내용(Video Content) : MPEG-4 비디오은 모든 형태의 화소 기반 영상 내용을 처리하기 위한 수단을 제공할 것이다.

(c) 객체 기반 비트열 편집과 처리 : MPEG-4 표준은 영상 객체의 편집을 위한 수단을 제공할 것이다.

(d) 객체 기반 랜덤 액세스(Object-based Random Access) : MPEG-4 표준은 한 장면내의 몇몇의 객체 또는 모든 객체들에 대해서 지정된 시간 안에 좋은 해상도를 가지고 객체 기반 랜덤 액세스를 하기 위한 효과적인 방법을 제공할 것이다.

(e) 객체 품질과 충실도(Object Quality and Fidelity) : MPEG-4는 비슷한 조건에서 현재 가능한 또는 현재 존재하는 표준보다 더 나은 주관적 영상 품질을 제공할 것이다.

(f) 다중 발생 데이터 비트열의 부호화(Coding of Multiple Concurrent Data Streams) : MPEG-4는 장면의 다중 관점/사운드 트랙을 효과적으로 부호화하고 기본 비트열 사이의 충분한 재동기를 보장하는 수단을 제공할 것이다.

(g) 오류와 손실 환경에 강인함 : MPEG-4표준은 갖가지 종류의 무선, 유선 네트워크와 저장매체 상에서 여러 회복 비디오 비트열을 가능하게 하는 도구들을 제공할 것이다.

이것은 낮은 비트율 응용에서도 지원될 것이다. 오류 복구 방법으로는 은폐, 결합 내성, 완벽한 복구 등을 생각할 수 있다.

(h) 객체 기반 계위 부호화 : MPEG-4는 내용이나, 품질, 모양 정보 항목에서의 계위 부호화를 가능하게 하는 도구와 구문 요소들을 제공할 것이다. 이러한 형태의 계위 부호화는 영상 정보의 유연한 내용 기반 계위 설정을 가능케 할 것이다.

(i) 시간 지연 양식(Delay Modes) : MPEG-4는 낮은 종단간(End-to-End) 시간지연 모드와 낮은 복호화 시간지연 모드를 포함하여 여러 가지 종류의 시간지연 모드를 지원할 것이다.

(j) 비트율 모드(Bitrate Modes) : MPEG-4 비디오는 64 kbit/s (low), 64 384 kbit/s (intermediate), 384 kbit/s 4 Mbit/s (high)의 비트율 범위에서 최적화된다.

(k) 복잡성 모드(Complexity Modes) : MPEG-4 표준은 낮은 복잡성 영상 부호화기와 복호화기를 포함하여, 여러 가지 종류의 복잡성

모드를 제공할 것이다. 복잡성 계위 부호화 가능 영상 조건이 규정 지어질 것이다.

(1) 정지영상 : MPEG-4 비디오는 정지영상의 효과적인 부호화 방법을 지원할 것이다.

3. MPEG-4 비디오 객체의 부호화 기술

비디오 객체 역시 영상이 자연영상과 합성영상으로 나누어지는 것처럼 인위적으로 합성된 것과 자연적인 것으로 나뉘어진다. 본 절에서는 자연영상 객체의 주요 부호화 기술에 대하여 전반적으로 기술하고자 한다.

1) 자연적인 Texture, 영상, 비디오

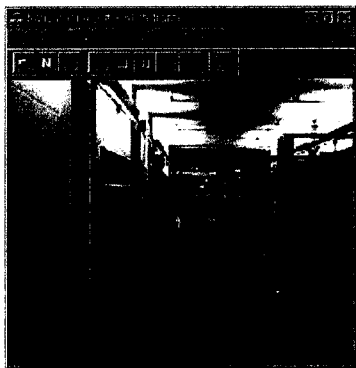
MPEG-4 비디오 표준에서, 자연영상을 표현하기 위한 도구들은 다중 매체 환경에서 효율적인 저장, 전송을 가능하게 하고 Texture, 비디오, 영상 데이터 처리를 가능하게 하는 표준적인 핵심 기술들을 제공하기 위한 것이다. 이러한 도구들은 영상 객체의 극소 단위인 영상 객체

(VO)의 표현과 복호화를 가능하게 한다. 이러한 단위 영상 객체는 그림 2처럼 VOP(Video Object Plane)라는 마스크를 통해 정의될 수 있다. 이러한 VOP는 MPEG-4의 객체 중심 부호화에서 부호화의 기본단위가 된다.

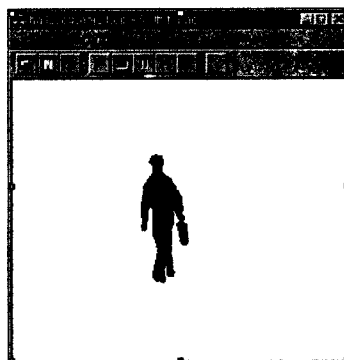
2) 자연 영상을 표현하기 위한 도구들의 구조

MPEG-4 영상이나 비디오 부호화 알고리즘은 내용기반 기능을 지원한다는 목적으로, 비디오 객체나 임의의 모양에 대한 효과적인 표현을 할 수 있게 해 준다. 이것들은 MPEG-1과 MPEG-2에서 이미 지원한 대부분의 기능들을 포함할 뿐만 아니라, 입력형식, 프레임율, 가변 비트율 레벨, 공간적/시간적/질적 계위 부호화의 레벨에서 표준적인 4각 크기의 영상 시퀀스를 효율적으로 압축하는 기능들을 지원한다.

현재 영상과 비디오의 MPEG-4 비디오 표준에서 제공하고 있는 비트율과 기능들을 아래의 그림 3에 나타내었다.



(a) 원영상



(b) Video Object Plane

그림 2. 영상 객체의 표현

그림 3에서 VLBV 핵심(Very Low Bit-rate Video: VLBV)는 낮은 공간 해상도와 낮은 프레임율을 지원하면서, 5~64 kbits/s 사이의 비트율에서 동작하는 응용분야를 위한 알고리즘과 도구들을 제공한다. VLBV 핵심에서 지원하는 기본 기능들을 간단하게 정리해 보면 다음과 같다.

(a) 높은 부호화 효율과 높은 오류 내성/복구를 갖는 4각 크기 영상 시퀀스의 VLBV 부호화, 실시간 멀티미디어 통신 응용을 위한 낮은 복

잡성

(b) VLBV 멀티미디어 데이터베이스 저장과 액세스 응용을 위한 랜덤 액세스(Random Access), 고속 전방향(Fast Forward), 고속 역방향(Fast Reverse) 작동을 위한 준비

*

위에 간략히 설명된 기본적인 기능들은 높은 비트율에서도 지원된다. 계획된 비트율의 범위는 64 kbits/s에서 4 Mbits/s이다. 높은 비트율

의 응용을 위해서, 복잡한 신호의 부호화를 위한 도구들을 MPEG-4 표준에서 명시하고 있다.

내용기반 기능은 장면에서 물리적인 객체(VO)에 따라 분리된 부호화와 복호화를 지원한다. 압축 영역에서, MPEG-4는 영상이나 영상 VO의 내용을 유연하게 표현하고 처리하기 위한 가장 기본적인 메커니즘만을 제공한다.

합성영상 데이터와 자연영상 데이터의 하이브리드(Hybrid) 부호화를 위하여, MPEG-4의 내용기반 부호화 기능은 서로 다른 장면으로부터 추출한 여러 개의 VO를 혼합하는 기능을 지원한다.

3) 내용 기반 기능의 지원

MPEG-4 표준은 임의의 형태 영상이나 비디오의 복호화뿐만 아니라 기존의 4각 영상이나 비디오의 복호화도 지원한다. 이 개념을 아래의 그림 4에 나타내었다.

기존의 4각 영상이나 비디오의 부호화는 MPEG-1/2 부호화와 유사하게 이루어지는 것을 볼 수 있다.

확장된 MPEG-4 내용기반 접근 방법은 임의의 형태에 대한 MPEG-4 VLBV 핵심 또는 높은 비트율 도구의 논리적인 확장인 것으로 생각할 수 있다.

4) MPEG-4 비디오 영상과 부호화 형태

그림 5는 임의의 모양 입력 영상 시퀀스 뿐만 아니라 기존의 4각 모양의 입력 영상 시퀀스를 부호화하기 위한 MPEG-4 알고리즘을 간단하게 나타낸 것이다. 기본 부호화 구

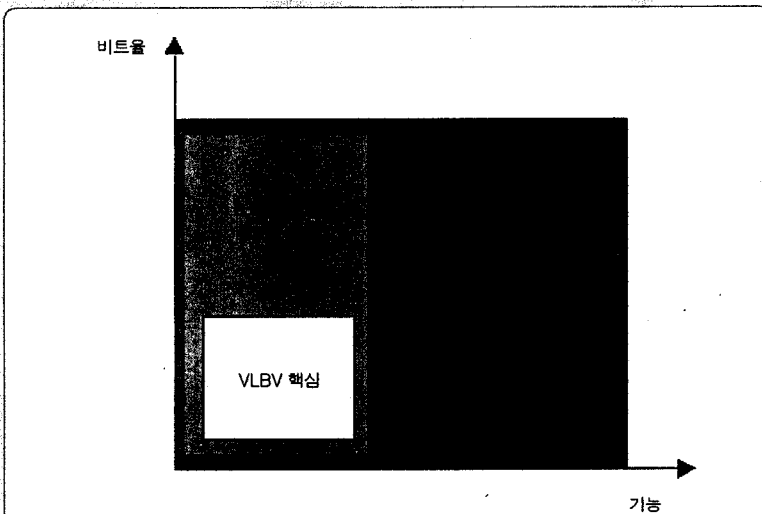


그림 3. MPEG-4 영상, 영상 부호화 알고리즘 그리고 물들의 분류

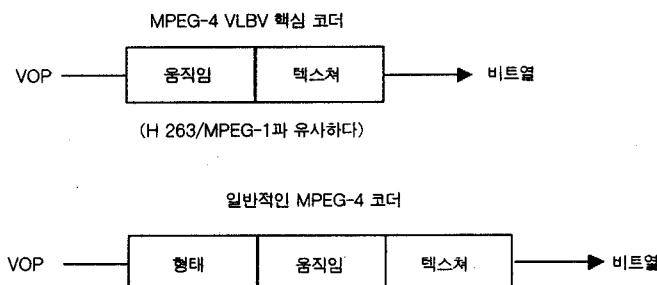


그림 4. VLBV 핵심과 일반적인 MPEG-4 부호화기

조는 DCT 기반 Texture 부호화 뿐만 아니라 임의의 모양의 VO에 대한 형태 부호화와 움직임 보상을 포함한다.

MPEG-4가 채택한 내용기반 부호화 방법의 중요한 이점은 장면의 각 객체마다 알맞은 객체 기반 움직임 예측 도구(Tool)를 사용함으로써, 비디오 시퀀스를 위한 압축효율이 매우 향상되었다는 것이다. 다음과 같은 많은 움직임 예측 기술들이 객체의 효과적인 부호화와 유연한 기술을 위하여 사용될 수 있다.

*

(a) 표준 88 또는 1616 화소 블록 기반 움직임 추정과 보상

(b) Affine Transformation을 기술하는 8개의 움직임 매개변수를 이용한 전역적 움직임 보상

(c) 정적 스프라이트(Sprite)를 기반으로 하는 전체 움직임 보상. 정적 스프라이트는 객관적인 배경을 나타내는 가능한 최대 크기의 정지 영상이다. 연속되는 영상에서, 카메라 움직임을 나타내는 8개의 전역 움직임 매개변수는 객체를 재구성할 때 사용되도록 부호화된다. 이러한 매개변수들은 첫번째 프레임에서 전송된 스프라이트의 적당한 Affine Transformation을 기술한다.

*

동적 스프라이트를 기반으로 하는 전역 움직임 보상. 스프라이트는 첫번째 프레임에 전송되지 않고, 동적으로 장면 위에 생성된다.

5) Texture와 정지 영상의 부호화

Texture와 정지영상의 효과적인

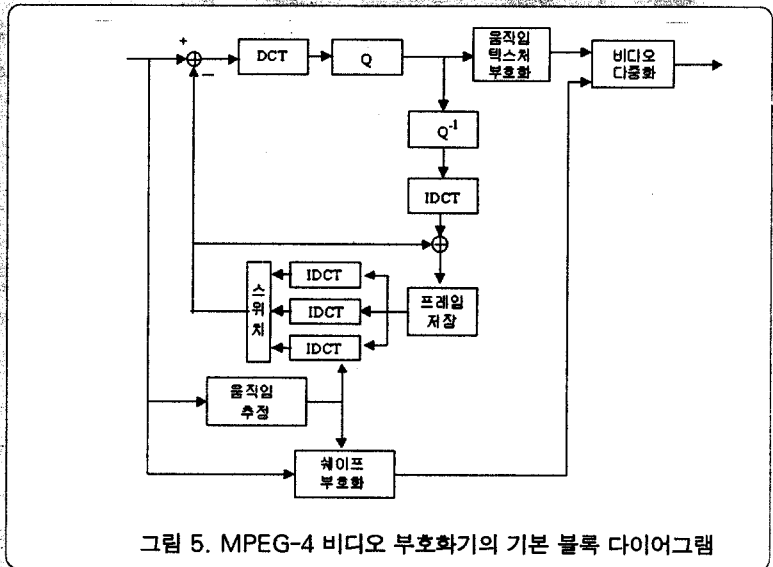


그림 5. MPEG-4 비디오 부호화기의 기본 블록 다이어그램

부호화는 매우 낮은 비트 압축률에서 매우 높은 압축률까지 VLBV 핵심, 높은 비트율, 내용기반 도구들에 의해서 지원된다. 제로 트리 웨이블렛 해석을 기반으로 하는 Texture와 영상 부호화 알고리즘이 지원된다.

6) 비디오 객체의 계위 부호화 (Scalable Coding)

MPEG-4는 공간적, 시간적 계위 부호화 방법을 가지는 영상과 비디오의 부호화를 지원한다. 계위 부호화는 다음과 같은 기능을 수행하며, 비트열의 일부를 복호화하는 능력과 영상이나 영상 시퀀스를 재구성하는 능력을 말한다.

*

- (a) 복호화기의 복잡성을 줄인다. 따라서 질이 떨어진다.
- (b) 공간적 해상도를 낮춘다.
- (c) 시간적 해상도를 낮춘다.
- (d) 같은 시간적, 공간적 해상도에서 질을 낮춘다.

*

이러한 기능은 수신기가 완전한 해상도나 완전한 질을 갖는 영상이나 비디오 시퀀스를 디스플레이 하지 못할 때를 위해서, 또는 이중 네트워크상에서 영상이나 비디오의 순차적 부호화를 위해서 요구되었다. 이러한 문제는 처리능력이나 디스플레이 해상도가 제한되었을 때 발생할 것이다.

7) 오류 발생 환경에서의 오류 강인성

MPEG-4는 많은 종류의 저장, 전송 미디어 상에서 영상이나 비디오 정보를 액세스 하기 위하여 여러 강인성과 복구성(Resilience)을 제공한다. 특히, 이동통신이 급격하게 발전함에 따라, 이것은 오디오나 비디오 정보를 무선 네트워크상에서 액세스 하기 위하여 매우 중요하다. 즉, 오류가 발생하는 환경에서 낮은 비트율(64 kbps보다 낮은)로 오디오

와 비디오 압축을 수행해야 한다는 것이다.

MPEG-4를 위해 개발된 오류 복구 도구들은 크게 재동기, 데이터 복구, 오류 은닉의 세 부분으로 나눌 수 있다.

가) 재동기

재동기 도구들은 잉여(Residual) 오류 또는 오류가 발견된 후에 복호화와 비트열 사이에서 재동기를 할 수 있게 하기 위한 것이다. 일반적으로 오류 이전의 동기점과 동기를 복구한 첫번째 점 사이의 데이터를 버린다. 만약 재동기 접근 방법이 복호화에 의해서 버려진 총 데이터를 작은 부분에 제한하는데 효과적이라면, 데이터를 복구시키거나 오류의 효과를 숨기는 다른 형태의 도구들의 효과가 매우 향상될 것이다.

MPEG-4에 의해서 채택된 재동기 접근방법은 H.261과 H.263의 ITU-T 표준에서 사용된 GOB (Group of Picture) 구조와 유사하다. 이 표준의 GOB는 매크로블록(Macroblock, MB)의 하나 또는 그 이상의 열을 의미한다. 새로운 GOB의 시작에서, GOB 헤더라고 불리는 정보는 비트열 내에 위치한다. 이 헤더 정보는 GOB 시작 코

드를 가지고 있다. 게다가 이 헤더 안에는 복호화 과정을 재시작 하는데 필요한 정보, 패킷내에 포함되어 있는 첫번째 매크로블록의 매크로 숫자와 그 첫번째 매크로블록을 복호화하기 위한 양자화 매개변수들이 포함되어 있다.

재동기를 위한 GOB 접근방법은 공간적 재동기를 기초로 한다. 즉 일단 어떤 특정한 매크로블록 위치가 부호화 과정에 도달하면, 재동기 마커가 비트열 내에 삽입된다. 이러한 접근방법의 잠재적 문제는 만약 부호화 처리가 변화한다면, 재동기 마커가 비트열 전역에 일정하게 위치하지 않는다는 것이다. 그러므로 움직임이 많은 장면 부분은 오류에 더 민감하고 오류를 은폐하기가 더욱 어려워질 것이다.

MPEG-4에서 채택한 비디오 패킷 접근방법은 비트열 전역에 주기적 재동기 마커를 삽입하는 것을 기본으로 한다. 즉, 비디오 패킷의 길이는 매크로블록의 숫자에 기초하는 것이 아니라, 패킷 내에 포함된 비트의 숫자를 기본으로 한다. 현재의 비디오 패킷내에 포함되어 있는 비트의 숫자가 미리 결정된 문턱값보다 크면, 새로운 비디오 패킷이 다음의 매크로블록의 시작에서 만들어진다. VOP 시작 코드와 재동기 마커는 오

직 비트열내의 지정된 고정 영역 위치에만 나타난다. 이 방법은 시작 코드 에뮬레이션(Emulation)과 관련된 문제를 피하는데 도움이 된다. 즉, 오류가 비트열 내에 존재할 때, 이 예러가 VOP 시작 코드와 유사할 수도 있다. 이러한 경우에 복호화기 측에서 사용되는 고정 영역 동기신호는 VOP 시작 코드를 고정 영역의 시작에서 찾는다. 따라서 시작 코드 에뮬레이션과 관련된 문제를 해결할 수 있을 것이다.

나) 데이터 복구

동기가 재 일치된 이후에 데이터 복구 도구들은 일반적으로는 잃어버리는 데이터를 복구시키려 할 것이다. 이러한 툴들은 단순한 오류 정정 코드가 아니라 오류 복구 방법으로 데이터를 부호화하는 기술이다.

비디오 그룹에 의해서 보증된 도구 중에 하나인 RVLC(Reversible Variable Codes)를 예로 들 수 있다.

이 방법에서, 가변장 코드워드는 정방향과 역방향 어느쪽으로 읽어도 동일하도록 설계된다. 111, 101, 010와 같은 것을 예로 들 수 있다. 100과 같은 코드워드는 사용되지 않는다. 이러한 방법은 압축효율을 감소 시키지만, 오류 회복에서의 개선

