

H.263을 기반으로 한 확장 가능한 비디오 코덱

노 경 택*

H.263-Based Scalable Video Codec

Kyung-Taeg Rho*

요 약

계층화 비디오 코딩은 확장성을 갖기 위해 다중 비디오 스트림 안에 비디오 정보가 전송되도록 한다. 계층화 코딩 구조는 두 가지 면에서 이론상의 장점을 갖는다. 첫째, 클라이언트 프로세싱 능력과 네트워크 대역폭 면에서 네트워크와 수신자의 이질성을 허락한다. 둘째, 여러단계의 비디오 화질이 요구될 때 이용가능한 대역폭에 맞출수 있다. 본 논문에서 우리는 움직임 보상을 이용한 확장가능한 비디오 코덱 구조를 제시한다. 이것은 패킷 네트워크 상에서 실시간 오디오, 비디오 통신에 적합하다. 코딩 알고리즘은 ITU-T 추천 H.263+에 맞추면서 복잡도를 감소시키기 위한 기법들을 제시한다. 빠른 움직임 검출은 H.263 베이스 레이어에서 이루어지고 상위계층에서 이용된다. 시각적 성질에 따른 매크로 블록 배제는 움직임 검출에 앞서서 모든 계층에서 이루어진다. 패킷 손실로 인한 에러확산은 주기적으로 각 계층에서 프레임 안에 인트라 블록단위의 행렬을 도입함으로써 차단되어질 수 있다.

Abstract

Layered video coding schemes allow the video information to be transmitted in multiple video bitstreams to achieve scalability. they are attractive in theory for two reasons. First, they naturally allow for heterogeneity in networks and receivers in terms of client processing capability and network bandwidth. Second, they correspond to optimal utilization of available bandwidth when several video quality levels are desired. In this paper, we propose a scalable video codec architectures with motion estimation, which is suitable for real-time audio and video communication over packet networks. The coding algorithm is compatible with ITU-T recommendation H.263+ and includes various techniques to reduce complexity. Fast motion estimation is performed at the H.263-compatible base layer and used at higher layers, and perceptual macroblock skipping is performed at all layers before motion estimation. Error propagation from packet loss is avoided by periodically rebuilding a valid predictor in Intra mode at each layer.

* 서울보건대학 사무자동화과 조교수

알고리즘의 주요 특징을 2절에서 설명하고 끝으로 3절에서 결론을 맺겠다.

I. 서론

소스기반 멀티캐스트 전송에서 실시간 비디오 스트림은 네트워크 상에 모든 수신자들에게 일정한 전송률로 전송된다. 그러나 주어진 전송에 관심을 갖는 모든 수신자들을 만족하는 고정된 전송률을 사용하는 것은 다양한 처리능력의 수신자들에게 적합하지 못하다. 예를 들면 낮은 대역폭을 지닌 사용자들은 상당한 패킷손실을 겪게될 것이다. 한편 높은 대역폭을 소유한 사용자들은 그들의 자원을 충분히 이용하지 못하게 하는 저품질의 소스에 불평할 것이다. 이러한 문제점을 극복하기 위한 해결책이 소스에서 수신자들로 비트 전송률을 적당히 하는 것이다. 따라서 소스는 여러 계층을 이용하여 다양한 화질을 분배하고 수신자들은 그들이 등록할 수 있는 계층을 선택함으로써 수신 비트율을 조절하는 것이다.

인터넷의 성장은 비디오 등과 같은 실시간 서비스 제공을 가능하게 한다. 따라서 이들 서비스는 패킷손실과 지연을 다룰 수 있어야한다. 연속된 서브스트림의 추가로 화질이 개선되는 계층화코딩방법은 이에 적합하다. 네트워크 혼잡상태에서 서브스트림에서 패킷손실이 일어날 수 있지만 약간의 화질이 떨어지더라도 영향을 받지 않은 하위계층에서 패킷들을 디코딩 함으로써 화상을 복원할 수 있다.

McCanne et al.이 설명했듯이 확장가능하고 계층적인 비디오 코딩 알고리즘은 대역폭이나 네트워크 혼잡에 따른 비디오 계층들에 동적인 가입을 허락하는 RLM (Receiver driven Layered Multicast)과 같은 알고리즘과 함께 사용될 때 IP 네트워크와 수신자들의 이질성(heterogeneity)을 수용하게 된다[1]. 본 논문은 조건적 보충과 결합된 움직임 보상 예측을 이용하는 계층화된 알고리즘을 제시한다. 알고리즘은 움직임 보상 예측에 의한 성능을 유지하면서 조건적인 보충에 의한 장점을 보존한다. 또한 손실된 패킷의 영향을 발견하고 감추게 하는 단순하지만 효율적인 에러회복 방안을 포함한다. 여기서 움직임보상 알고리즘은 낮은 복잡도 (complexity)를 베이스계층에서만 이용된다.

II. 계층화된 코덱 알고리즘

1. 코덱개요

인코더는 낮은 비트율의 베이스계층 스트림과 높은 비트율을 지닌 상위계층 스트림으로 구성된 비디오 스트림을 만들어낸다. 그림 1에서 코덱구조를 설명하는데 E_0 는 베이스계층 인코더, E_1 , E_2 는 상위계층 인코더를 나타낸다. E_0 , E_1 , E_2 는 메타 인코더 E 를 구성한다. 베이스계층 디코더 D_0 와 상위계층 디코더 D_1 , D_2 는 메타 디코더 D 를 구성한다[2]. 따라서 H263 디코더에 의해 디코딩될 수 있다. [1]에서 McCanne에 의해 제안된 코덱과 달리, 우리의 베이스계층 인코더는 움직임 보상 예측을 하기 위해 움직임 평가를 이용한다. 상위계층 스트림은 H.263+를 기초로 한다[3]. 상위계층 인코더는 몇 가지 예측을 이용하는데, 즉 그림 2에서 설명하듯이 동기적으로 다시 만들어진 프레임들 (EI 모드)과 베이스계층에서 계산된 움직임 벡터들을 이용한 움직임 보상 예측 (EP 모드)과 같은 예측을 이용한다. 코딩모드 결정은 다음과 같은 예측 에러에 대한 평가를 기초로 한다:

```
if (dfd_mad < fd_mad)
    coding_mode = INTER;
else
    coding_mode = EI;
```

여기서 dfd_mad 는 베이스 계층에서 얻어진 움직임 벡터를 이용한 두 프레임간의 차에 대한 평균 절대값 차를 말하고 fd_mad 는 베이스 계층에서 다시 만들어진 프레임과 원래의 프레임간의 평균 절대값 차를 말한다.

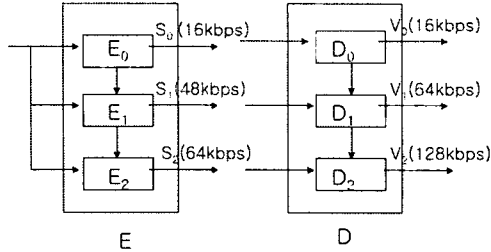


그림 1. Embedded codec. The encoder generates one base layer stream S0 and two enhancement layer streams S1 and S2.

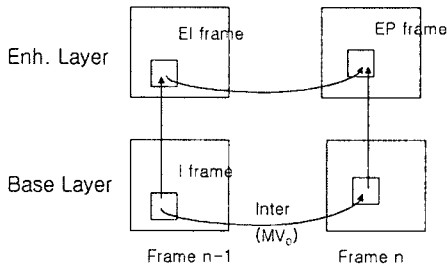


그림 2. Prediction and coding modes at base and enhancement layers. MV0 denotes the motion vector field computed at the base layer

1) 레이트 컨트롤 (rate control)

레이트 컨트롤은 각 계층에서 독립적으로 이루어진다. 각 계층에서 인코더 버퍼의 용량은 해당 계층에서 설정한 비트율의 반으로 한다. 버퍼가 버퍼용량의 80%에 도달되면 들어오는 버퍼 프레임들은 해당계층에서 스킵 되어진다. 이와 같은 단순한 방법이 버퍼를 버퍼용량의 반정도 크기로 유지되도록 한다.

2) 에러 컨트롤

UDP 프로토콜은 패킷전달이 보장되지 않기 때문에 디코더는 패킷손실을 발견하고 그것으로부터 회복할 필요가 있다. 회복을 위한 알고리즘은 서로 다르지만 상호 관련된 두 가지 모듈로 구성된다. 첫째, 디코더는 RTP 패킷헤더의 순서필드로부터 손실된 패킷을 발견한다. 손실된 매크로 블록과 일치하는 이미지 영역을 바로 전 프레임으로부터의 데이터가 단순히 반복되어진다. 둘째, 에러 데이터로 인한 움직임 보상 예측으로부터 야기되는 에러확산을 차단하기 위해 인트라블록들로 구성된 매크로 블록라인을 주기적으로 인코딩 한다. 이 방안은 그림 3에서 설명된다. F fps의 코딩비율을 가정할 때 이 방법은 mb_lines/F 초만에 에러회복이 된다. 여기

서 mb_lines는 프레임당 매크로블록 라인수이다.

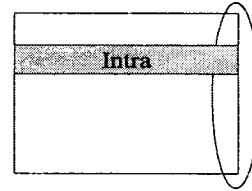


그림 3. Periodic encoding of lines of Intra macroblocks for error propagation suppression

2. 복잡도 감소방안 (Complexity reduction)

시간에 관련된 압축 (temporal compression)을 위해서 인코더는 비디오 화질을 유지하면서 복잡도를 크게 줄이기 위한 두 가지 기법을 이용한다: 빠른 움직임 검출과 시각적 성질에 따른 조건적 보충이 있다.

1) 빠른 움직임 검출

움직임 검출은 베이스계층에서만 이루어진다. H.263+는 예측을 위해 상위계층에서 움직임 벡터를 이용한다. 우리 인코더는 계층마다 베이스 계층 움직임 벡터를 이용한다. 움직임 검출 알고리즘은 16×16 휘도 블록들에 관한 ±15 범위 안에서 logarithmic cross-search를 이용한 블록매칭 알고리즘이다(4, 5).

2) 시각적 성질에 따른 조건적 보충

조건에 따른 보충은 프레임간에 거의 변경되지 않은 이미지 영역을 발견해서 알려주는 코덱의 역할을 말한다 (6). H.263은 블록스킵을 통해 매크로 블록들의 조건적 보충을 허락하지만 여기서는 시각적 성질에 따른 조건적 보충은 복잡도를 감소시키기 위해 움직임 검출에 앞서 이루어진다. C는 8×8 휘도 블록에 대한 변화값을 가리킨다. m은 8×8 입력 이미지 블록 curr(i,j)의 평균 휘도값을 말한다. C는 다음과 같이 계산된다:

```

for (i=0; i<8; i++) {
  for (j=0; j<8; j++) {
    diff = ABS(curr(i,j) - prev(i,j));
    if (diff > T(m)) {
      C += diff/T(m);
    }
  }
}
if ( C > max_above_T) change = True;
else change = FALSE;
    
```

매개변수 max_above_T 는 6으로 선정했다. 만약 어떤 8×8 블록도 큰 변화가 없다면 16×16 매크로 블록은 스킵 된다. 평균 블록 휘도값 m 에 대한 스톱스레시홀드 $T(m)$ 의 의존성은 Chiu와 Berger의 연구에 기초한다 [7].

III. 결론

본 논문에서 우리는 움직임 보상을 이용한 확장 가능한 비디오 코덱 구조를 제시하였다. 에러제어 방안으로 손실된 패킷을 RTP 헤더의 순서 필드로 발견할 수 있다. 손실된 매크로 블록과 일치하는 이미지 영역을 바로 전 프레임으로부터의 데이터를 반복시킨다. 또한 에러 데이터로 인한 움직임 보상 예측으로부터 야기되는 에러확산을 차단하기 위하여 인트라 블록들로 구성된 매크로 블록라인을 주기적으로 인코딩 하였다. 베이스 계층에서 빠른 움직임 검출이 이루어지도록 하고 상위계층에서 이용하도록 하였다. 프레임간에 거의 변경되지 않은 이미지 영역을 발견해서 전송하지 않음으로써 복잡도를 감소시켰다.

참고문헌

- [1] S. McCanne, M. Vetterli, V. Jacobson, Low-complexity Video Coding for Receiver-driven Layered Multicast. Trans. JSAC, vol. 16, no. 6, pp.983-1001, August 1997.
- [2] ITU-T Draft H.263: video coding for low bitrate communication, December, 1995.
- [3] ITU-T Draft H.263+: video coding for low bitrate communication. september, 1997.
- [4] K.R. Rao, P. Yip, Discrete Cosine Transform. Algorithms. Advantages, Applications. Academic Press, 1990.
- [5] H.G. Musmann, P. Pirsch, H.J. Grallert, Advances in Picture Coding. Proc. IEEE, April 1985.
- [6] F.W. Mounts, A video encoding system with conditional picture-element replenishment. Bell Systems Technical Journal, vol. 48, no. 7, pp 2545-2554, Sept. 1969.
- [7] Y.J. Chiu, T. Berger Perceptual rate control of video sequences. Proc. Fourth Data Compression Industry Workshop, March 1997, Snowbird, Utah.

저자소개



노경택

1986년: 중앙대학교 전자계산기 공학과 졸업(공학사)

1986년: 제일제당 전산실

1989년: New Jersey Institute of Technology 컴퓨터학과(이학석사)

1992년: University of South Carolina 컴퓨터학과(박사수료)

1999년: 고려대학교 컴퓨터학과(박사수료)

1993년 3월 ~ 현재 서울보건대학 사무자동화과 조교수

관심분야: image and video compression and coding, video communication