

Frame Dropping Transcoder에서 활동정보 및 중첩영역의 크기를 고려한 모션벡터 합성 기법

김성민*, 김현희, 탁광옥, 이승원, 정기동
부산대학교 컴퓨터공학과
e-mail:morethannow@melon.cs.pusan.ac.kr

Motion Vector Composition Scheme using activity information and overlapped extent on the Frame Dropping Transcoder

Sung-Min Kim*, Hyun-Hee Kim, Kwang-ok Tak, Seung-Won Lee,
Ki-Dong Chung
Dept of Computer Engineering, Pusan National University

요 약

여러 응용 서비스를 유·무선을 포함한 다양한 네트워크를 통해 제공하기 위해서는 네트워크에 적응할 수 있는 서비스 형태가 요구된다. 그 가운데 멀티미디어 서비스의 경우 네트워크 서로 다른 환경에 적응할 수 있는 해결책으로 트랜스코딩 기술이 제시되었다. 하지만, 트랜스코딩을 위해 필요한 복호·부호의 처리 과정은 실시간으로 제공되는 멀티미디어 스트리밍의 경우에 제약조건으로 작용하고, 이에 따른 처리 과정을 대폭 줄이는 일부 기술들은 사용자 측의 서비스 품질에 문제점을 안고 있다.

본 논문에서는 트랜스코딩을 통한 처리 과정과 사용자 측 서비스 품질의 두 가지 측면을 고려하는 frame dropping 시의 모션 벡터 합성 기법에 대해서 언급한다. 또한, 본 논문에서는 기존의 기법과는 달리 양방향 예측 프레임이 포함된 경우에도 적용할 수 있는 확장성을 제공한다.

1. 서론

최근 들어 시작된 멀티미디어 서비스에 대한 수요의 증가는 현재 유선에서의 서비스 형태를 벗어나 무선에 이르기까지 다양한 형태의 서비스로 사용자에게 제공되고 있다. 현재 화상 전화, 원격 교육 시스템 및 주문형 비디오등 네트워크를 통한 멀티미디어 서비스는 사용자들에게 많은 각광을 받고 있으며, 앞으로도 많은 수요가 예상되는 분야이다.

인터넷이 사용되던 초기에는 주로 유선에 국한된 서비스 형태가 대부분이었고, 멀티미디어 서비스도 근래 들어 이용 가능하게 된 서비스이지만, 현재 사용자들의 요구는 무선에서조차도 자유로운 미디어 서비스를 요구하게 되었다. 따라서 현재의 멀티미디어

서비스를 위한 시스템은 유선뿐만이 아니라 무선을 포함한 다양한 네트워크의 상황에 맞는 적응형 서비스를 제공해야 한다.

미디어는 보통 대용량의 특성을 가지고 있으므로, 압축된 스트림의 형태로 사용자에게 제공된다. 하지만, 규격화된 압축 스트림을 이용해서 서비스를 제공할 경우에는 여러 가지 특성을 가지는 네트워크에 적응할 수 없게 된다. 이때 트랜스코딩 기술은 서비스를 제공하는 측과 서비스를 제공받는 측의 중간에서 양측의 상황에 적응할 수 있는 해결책으로 사용할 수 있다[1,2,3]. 멀티미디어를 제공하는 서버에서 높은 재생율을 가지는 비디오를 사용자에게 제공할 경우, 사용자의 네트워크가 서버에서 제공하는 재생율

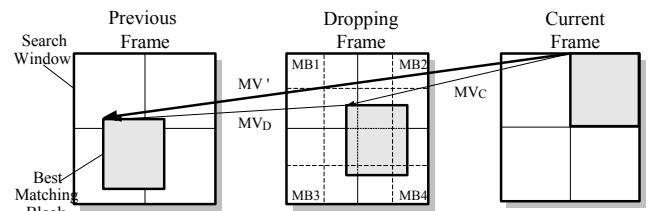
을 처리 못할 경우 정상적인 서비스를 기대하기 어렵다. 이럴 경우 트랜스코더는 사용자측의 네트워크 특성을 고려하여 낮은 재생율을 가지도록 전송 스트림을 조절하게 되고, 이는 사용자 측의 입장에서 보다 나은 서비스를 제공받는 방법이 된다. 하지만, 트랜스코딩의 절차는 서버로부터 전송된 스트림과 사용자 측으로 전송될 스트림 사이에서 전송율을 조절하므로, 복호화와 부호화의 처리과정을 거치게 되어 실시간 서비스를 위해서는 심각한 문제점으로 작용한다. 전송율 조절을 위해 프레임을 drop시키는 트랜스코더에서는 복호화와 부호화의 처리과정중 가장 많은 처리가 필요한 움직임 벡터 예측을 drop시키는 프레임의 모션벡터를 재사용하여 선택하는 기법들이 제안되었다[4,5,6,7]. 이러한 대표적인 기법들에는 FDVS(Forward Dominant Vector Selection) 기법과 ADVS(Activity Dominant Vector Selection) 기법이 있다. 하지만, 이런 기법들은 사용자 측 서비스 화질이 떨어지는 단점을 가지고 있으며, 서비스의 질을 향상시키기 위한 추가 기법들은 처리 과정의 시간을 늘리는 단점으로 작용되었다.

본 논문에서 제안하는 모션 벡터의 합성 기법은 기존의 선택 기법이 가지는 모션벡터의 선택 및 그에 따른 추가적인 기법 적용이 아니라, 모션 벡터와 활동성 정보 및 중첩 영역의 가중치를 고려한 합성 기법이다. 따라서 사용자 측의 화질향상을 위해 추가적인 적용기법이 필요 없으므로 비슷하거나 보다 적은 처리 과정 속에서 사용자 측의 서비스 품질을 향상시킬 수 있다. 프레임을 drop시킬 경우 drop되는 프레임에 대한 다음 프레임의 모션벡터들은 drop된 프레임의 특정위치를 가르키게 된다. 이때 겹쳐지는 drop 프레임 매크로 블록의 모션벡터들과 중첩되는 매크로 블록과의 영역을 함께 고려하여 새로운 모션벡터를 합성하게 된다. 합성된 모션벡터는 drop된 프레임의 특성을 충분히 반영하여 drop 프레임의 이전 프레임을 참고할 수 있게 되므로, 기존 기법들이 가지는 단순 선택에 비해서 보다 나은 성능을 보일 수 있다. 그리고 제안하는 기법은 양방향 예측 프레임을 포함하는 비디오 시퀀스에 대해서도 적용할 수 있는 확장성을 지닌다.

본 논문은 다음과 같이 구성된다. 2장에서는 관련 연구에 대해서 소개하고, 3장에서는 본 논문에서 제안하는 활동정보 및 중첩영역의 크기를 이용한 모션 벡터 합성 기법 및 예를 보인다. 마지막으로 4장에서는 결론 및 향후과제를 제시한다.

2. 관련연구

전송율 조절을 위해 프레임을 drop시키는 경우에 움직임 벡터를 새로이 추출해 내는 것은 트랜스코더의 복잡도를 증가시키므로, drop되는 프레임의 움직임 벡터를 재사용하는 기법들이 연구되었다. [그림 1]은 프레임이 drop될 경우의 움직임 벡터가 수정되어야 하는 방법에 대해서 설명하고 있다. 하지만, 정확한 움직임 벡터를 얻기 위해서는 움직임 벡터를 새로이 추출해 내야 하므로, FDVS나 ADVS와 같은 기법들은 drop되는 프레임의 움직임 벡터를 재사용하는 방법을 이용한다. 정확한 움직임 벡터의 값은 $MV_C + MV_D$ 의 값인 MV' 이지만, 위 두 기법은 MB1~MB4가 가지는 모션벡터 중에서 MV_D 의 값과 가장 유사하리라고 판정된 모션벡터를 이용한다.

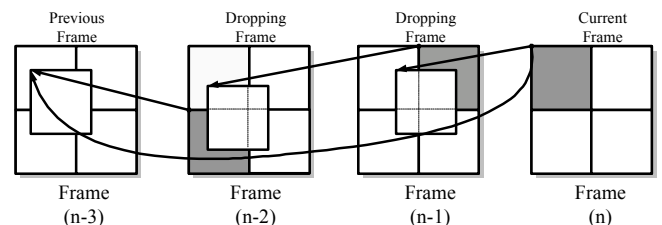


[그림 1] 모션 벡터의 재구성

2.1. FDVS(Forward Dominant Vector Selection)

FDVS 기법은 [그림 1]에서 drop되는 프레임의 MB1~MB4의 모션벡터들 중 가장 큰 영역이 중첩되는 매크로 블록인 MB4의 모션벡터 값을 MV_D 로 예측해서 사용한다[6]. MB_i 의 모션벡터를 MV_i 로 표현할 때, drop되는 프레임의 모션벡터를 재사용해서 만들어진 모션벡터 $MV' = MV_4 + MV_C$ 와 같이 구성된다.

[그림 2]는 중간에 여러 프레임이 drop될 경우의 모션 벡터를 재구성하는 모습을 보여준다.



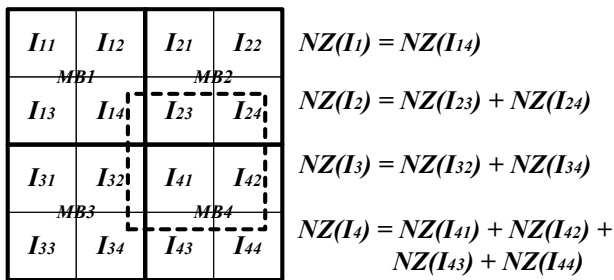
[그림 2] FDVS에서의 모션 벡터 재구성

2.2. ADVS(Activity Dominant Vector Selection)

ADVS 기법은 [그림 1]에서 drop되는 프레임의

MB1~MB4가 가지는 각각의 블록들과 중첩되는 영역을 고려하여 모션벡터를 선택한다[7]. 각 매크로 블록의 중첩되는 블록이 가지는 0이 아닌 DCT 계수의 개수 합을 통해 가장 큰 매크로 블록의 모션벡터를 선택한다.

[그림 3]은 모션벡터를 선택하기 위한 그림으로 [그림 1]의 dropping 프레임을 보여준다. ADVS 기법은 $NZ(I_1) \sim NZ(I_4)$ 중에서 가장 큰 값을 갖는 MB의 모션벡터를 선택하게 된다.



[그림 3] ADVS에서 모션 벡터 재구성

3. 활동정보 및 중첩영역을 이용한 모션 벡터 재구성

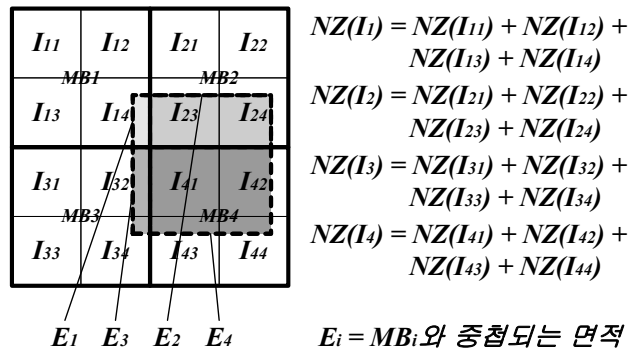
3.1. 기법에 대한 개요

관련연구에서 소개된 두 기법은 drop되는 프레임의 특정 모션벡터를 선택함으로써 움직임 예측의 복잡한 처리과정은 제거했으나, 모션벡터의 선택이 가지는 정밀함의 한계는 사용자 측의 화질 저하를 초래했다. 따라서 활동성 정보를 지니고 있는 0이 아닌 DCT 계수의 개수 합과 중첩되는 매크로 블록의 영역비율을 모두 고려하여, 새로운 모션벡터를 만들어 내는 기법을 소개한다.

기존의 기법들이 선택의 한계를 드러내는 반면, 본 논문에서 제안하는 가중치 기반의 모션벡터 합성 기법은 drop되는 프레임에 의해 손실되는 모션벡터 정보의 복원력이 뛰어나다. 모션 벡터의 합성을 위한 단계는 다음과 같다.

step 1. 모션벡터 합성을 위한 매크로 블록 별 가중치 계산 : 가중치는 중첩되는 영역의 매크로 블록의 0이 아닌 DCT 계수의 개수 합과 중첩되는 영역의 좌표크기 면적을 통해서 구해진다. [그림 4]를 통해서 얻어지는 각 매크로 블록의 가중치 W_i 는 다음과 같은 식으로 구해진다.

$$\bar{W} = (E_i * NZ(I_i)) / \sum_{i=1}^4 E_i * NZ(I_i)$$



[그림 4] 활동정보 및 중첩영역을 통한 모션벡터 재구성

step 2. 각 매크로 블록 별 weight를 통해 새로운 모션 벡터의 합성 : step 1을 통해서 얻어진 각 매크로 블록의 가중치와 해당 매크로 블록의 모션벡터의 곱을 통해서 새로운 모션 벡터를 합성한다. 합성식은 다음과 같다.

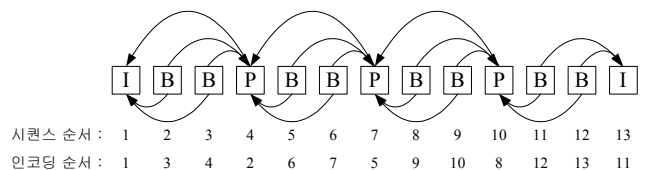
합성된 모션 벡터 :

$$\sum_{i=1}^4 MV_i * \bar{W}_i$$

3.2. 양방향 예측 프레임이 있는 시퀀스에 대한 예

본 논문에서 제안하는 기법은 I와 P프레임의 시퀀스에 대해 초점을 맞추고 있는 기존의 기법들과 달리 I,B,P프레임의 시퀀스에서도 적용 가능한 확장성을 지닌다. I,B,P를 모두 가지는 시퀀스의 재생을 조절을 위해 프레임을 drop할 경우, 상대적으로 용량이 작은 B프레임보다는 P프레임을 drop시키는 것이 훨씬 효율적이다. 하지만, B프레임의 양방향성 예측 특성에 따라 drop되는 P 프레임은 주변의 여러 B프레임에 영향을 미치게 된다. 따라서 B프레임이 drop된 P프레임을 참고할 경우 직전의 P프레임 또는 I프레임을 통해서 복구함에 따라 I,B,P프레임이 모두 존재하는 시퀀스에서도 확장하여 사용가능하다.

[그림 5]는 GOP가 12개의 프레임으로 이루어진 MPEG 계열의 시퀀스와 참조되는 상태를 나타낸다.

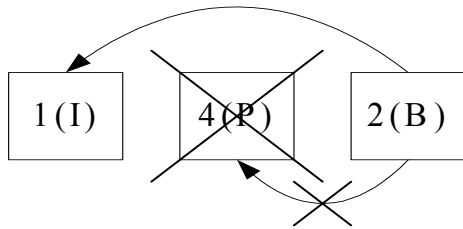


[그림 5] I,B,P 시퀀스의 프레임 참조

B프레임은 부호화하는 매크로 블록의 특성에 따

라 forward, backward, bidirectional의 모션 벡터를 가진다. 이때 bidirectional일 경우에는 두 개의 모션 벡터를 가지게 된다.

[그림 5]에서 시퀀스 순서의 4번째인 P프레임을 삭제하게 되면 2,3,5,6의 B프레임이 영향을 받게 된다. [그림 6]은 시퀀스 순서 2번 프레임인 B 프레임의 예측상태를 나타낸다.



[그림 6] P프레임이 drop될 경우 B프레임의 참조 변화

이 때, 시퀀스 순서 4번 프레임인 P프레임이 비트를 조절을 통해 drop될 경우 B프레임이 가지는 세 가지 예측에 따라 다음과 같이 처리할 수 있다. forward 예측인 경우는 I프레임을 참고하므로 추가적인 처리가 필요하지 않고, bidirectional 예측인 경우는 forward 예측만 사용하면 된다. 그리고, backward 예측일 경우는 [그림 6]의 경우와 같이, 앞서 제안한 기법을 통해서 모션벡터를 합성할 수 있다. 따라서 앞서 제안했던 활동정보 및 중첩영역 기반의 모션벡터 합성 기법은 I,P프레임만을 가진 시퀀스 뿐만 아니라 I,B,P를 모두 가지는 시퀀스에 대해서도 적용이 가능하다.

4. 결론 및 향후과제

유·무선을 포함하는 멀티미디어 서비스 시스템에서는 서로 다른 환경을 가지는 네트워크 특성에 적응성을 가지는 서비스가 중요하다. 본 논문에서는 서버에서 제공되는 전송율과 사용자 측의 네트워크 또는 단말기 능력에 따라 적응성을 제공할 수 있는 트랜스코더 내부의 모션벡터 합성 기법에 대해서 언급했다. 기존의 트랜스코딩 기술들이 가지던 계산 복잡도나 사용자 측 화질 저하의 문제점을 극복할 수 있는 기법으로 I,P프레임으로 구성된 시퀀스 뿐만 아니라 I,B,P프레임으로 구성된 시퀀스에도 적용 가능하다. 하지만, I,B,P프레임으로 구성되어 있는 경우에도 연속적인 P프레임 drop의 경우 복잡한 모션벡터의 합성 및 사용자 측의 화질에 심각한 저하를 초래할 수 있다. 따라서 I,B,P프레임으로 구성된 시퀀

스에서 연속적으로 P프레임을 drop시킬 수 있는 기법에 대한 연구가 계속되어야 할 것이고, 실험을 통한 성능의 객관적인 검증이 필요하다.

참고문헌

- [1] A. Vetro and C. Christopoulos and H. Sun, "Video Transcoding Architectures and Techniques: An Overview," IEEE signal Processing Magazine, ISSN: 1053-5888, Vol. 20, Issue 2, pp. 18-29, March 2003.
- [2] N. Bjork and C. Christopoulos, "Video Transcoding for Universal multimedia Access," ACM Workshop on Multimedia Standards, Interoperability and Practice (MM2000 Workshop), Proceedings of ACM Multimedia 2000, pp. 75-79, November 4, 2000.
- [3] Y. Nakajima and M. Sugano, "MPEG bit rate and format conversions for heteroheneous network/storage applications," IEICE Trans. Electron., E85-C, (3), pp. 492-503, Mar. 2002.
- [4] Sun, and C. W. Lin, "Motion Estimation for High Performance Transcoding," IEEE Trans. Consumer Electron., vol. 44, pp. 649-658, Aug. 1998.
- [5] Kai-Tat Fung and Wan-Chi Siu, "DCT-based Video Frame-skipping Transcoder," Proceedings, IEEE International Symposium on Circuits and Systems (ISCAS'03), Vol. II pp. 656-659, Bangkok Thailand, May 2003.
- [6] M. J. Chen, M. C. Chu and C. W. Pan, "Efficient motion estimation algorithm for reduced frame-rate video transcoder," IEEE Trans. Circuits syst. Video Technol., vol. 12, pp. 269-275, Apr. 2002.
- [7] J. Youn, M. T. Sun and C. W. Lin, "Motion vector refinement for high performance transcoding," IEEE Trans. Multimedia, vol. 1, pp. 30-40, Mar. 1999.