

GOP 구조 변환에 있어서의 효율적인 트랜스코딩 기법

EFFICIENT VIDEO TRANSCODING IN THE GOP STRUCTURE CONVERSION

이강준*, 김정준**, 정제창***
Kangjun Lee, Jeongjun Kim, Jechang Jeong

Abstract - Recently, for satisfying many application demands such as coding delay, computing power, transporting channel characteristic, etc, many profiles are supported in video coding standards. Therefore, in transcoding between same standards or between other standards, the functional difference of profiles supported by application occur many problems. In this paper, transcoding MPEG-2 main profile to H.264/AVC baseline profile which has restriction in the number of reference frame is focused. In this case, the bidirectional prediction supported in MPEG-2 main profile is not supported in H.264/AVC baseline profile. Also, in the restriction of reference frame, motion vectors in the MPEG-2 decoder as predictor should be adjusted. In this paper, the proposed algorithm is based on the characteristic of which motion vector is uniform according to the distance from reference frame. The adaptive search techniques through the determination of the uniformity extremely reduce the computational complexity.

Key Words : Transcoding, H.264/AVC, MPEG-2, GOP Conversion

1. 서론

동영상 처리 기술의 발전과 더불어 새로운 비디오 압축 표준들이 등장하고 있다. 따라서 많이 비디오 콘텐츠가 다양한 형태의 압축 표준에 따라 만들어지고 있다. 또한 다양한 기능적 요구사항에 따라 많은 프로파일들이 지원되고 있다. 예를 들어, MPEG-2에 있어서는 5개의 프로파일이 지원되고 있다[1]. 간단한 구현이 가능한 베이스라인 프로파일은 양방향 예측을 지원하지 않고, 많은 기능을 가진 메인 프로파일은 HDTV, DVD와 같은 기기에서 이용되고 있다. SNR, Spatial, High 프로파일은 다양한 스케일러블 부호화 방법을 제공한다. H.264/AVC에 있어서는 베이스라인 프로파일은 영상회의와 모바일 방송을 위한 기능을 제공한다[2]. 엔트로피 부호화에 있어서는 CAVLC(Context-based Adaptive Variable Length Code)를 사용한다. 방송과 스토리지 장치를 위해 개발된 메인 프로파일은 양방향 예측을 위한 B slice와 엔트로피 코딩을 위해 CAVLC(Context-based adaptive binary arithmetic code)를 사용한다. 스트리밍 서비스를 위한 Extended 프로파일에서는 SP/SI slice, data partition, CAVLC 기술이 사용된다.

위에서 언급했던 것과 같은 기능적으로 다른 프로파일을 사용하는 표준간의 트랜스코딩은 많은 문제점을 야기한다[3]. 트랜스코딩 과정에 있어서 핵심은 복호기에서의 정보를 이용

하여 부호기의 계산량을 낮추면서 화질의 열화를 최소화하는 것이다. 움직임 추정과 부호기에서 가장 많은 연산량을 차지하는 부분이다. 따라서, 복호기의 움직임 벡터를 이용하여 부호기의 움직임 추정 과정에서 연산량을 줄이는 작업에 관해 많은 연구가 진행되어 왔다. 하지만, MPEG-2 메인 프로파일에서 H.264/AVC 베이스라인 프로파일로의 트랜스코딩에 있어서는 기존의 방법을 사용하는데 많은 문제점이 있다. MPEG-2 메인 프로파일에서는 양방향 예측을 이용한 움직임 추정을 하지만, H.264/AVC 베이스라인에서는 양방향 예측을 지원하지 않는다. 또한, H.264/AVC 베이스라인에서 구현의 편리성을 위해 참조 프레임의 수를 한 장으로 제한하는 경우에는 레퍼런스 프레임의 불일치로 인한 문제점도 발생한다.

따라서 본 논문에서는 MPEG-2 메인 프로파일에서 레퍼런스 프레임에 제약을 가진 H.264/AVC 베이스라인 프로파일로의 트랜스코딩에 중점을 두고 GOP 구조가 바뀌는 트랜스코딩에 있어서 복호기의 움직임 벡터를 효율적으로 이용하기 위한 방법을 설명한다.

2장에서 움직임 벡터의 선형성에 기반한 적응적 움직임 재추정 알고리즘에 대해 설명하고 3장에서 실험결과를 보여준 뒤, 4장에서 결론을 제시할 것이다.

2. 제안하는 방법

2.1 선택적 재탐색 영역 결정

본 논문에서는 시간축에서의 움직임 벡터의 선형성에 기반하여 선택적인 재탐색 영역을 설정하는 방법을 제시한다. MPEG-2 복호기로부터의 움직임 벡터는 아래의 식 (1)에 의해 재조정된다.

저자 소개

* 이강준: 한양대학교 전자통신컴퓨터공학과 박사과정

** 김정준: 한양대학교 전자통신컴퓨터공학과 석사과정

*** 정제창: 한양대학교 전자통신컴퓨터공학과 정교수

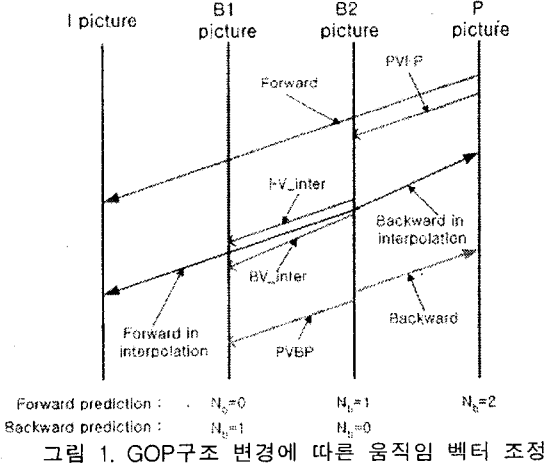
If (forward prediction)

$$MV_{adjust} = \text{Round} (MV_{original} / (N_b + 1))$$

Else if(Backward prediction)

$$MV_{adjust} = - \text{Round} (MV_{original} / (N_b + 1)) \quad (1)$$

위의 식(1)에 의해 조정된 움직임벡터를 아래의 그림1에 나타내었다. 그림1에서 PVFP는 전방향예측에서 유도된 Predicted Motion Vector(PMV)이고 PVBP는 후방향예측에서 유도된 PMV, FV_inter와 BV_inter는 양방향예측에서 각각 유도된 PMV를 나타낸다.



양방향예측에서 유도된 두 개의 PMV의 차를 비교하여 해당되는 매크로블럭의 움직임의 시간축에 따른 선형성을 예측할 수 있다. 아래의 표1은 양방향예측에서 유도된 두 개의 움직임벡터의 차에 따른 PMV와 실제 전역탐색방법에 따른 움직임벡터의 차를 나타낸다. 표1에서 각 PMV의 차가 3이상 일 때 PMV의 정확도가 많이 떨어지는 것을 볼 수 있다.

표 1. 양방향 예측 움직임벡터 차이에 따른 PMV 정확성

시퀀스	3이하			3이상		
	x	y	number	x	y	number
Akiyo	0.11	0.15	290.72	9	5.33	0.16
Bus	0.57	0.24	203.27	8.86	11.98	43.65
Coast.	0.36	0.17	210.07	8.44	8.2	25.2
Football	0.44	0.36	176.84	16.76	13.54	66.88
Foreman	0.57	0.5	213.9	16.76	6.58	30.66
Table.	0.51	0.41	226.25	9.66	6.92	34.06

따라서 본 논문에서는 양방향예측에서 두 개의 PMV값을 비교하여 3이상의 차이를 보이는 경우 재탐색영역을 확장하여 움직임 추정의 정확도를 높인다.

2.2 움직임예측 오차의 보정

식 (1)에 의해 추정된 PMV의 예측 오차에 있어서 주변 블럭과의 연관성을 가지고 있다. 주변 매크로블럭의 PMV와 같은 PMV를 가지고 있는 매크로블럭은 예측 오차에 있어서도 같은 오차를 나타내는 경우가 많다. 따라서, 예측 오차를

줄이기 위해 주변 블럭과 같은 PMV를 가지고 있는 매크로블럭의 PMV는 주변 매크로블럭의 오차를 더해 예측 오차를 감소시킨다. 아래의 의사코드는 위의 과정을 나타내고 있다.

```

If (PV_upper = PV_upper_right = PV_left = PV_current)
{
    If( Error_upper =Error_upper_right = Error_left)
    {
        PV_Current = PV_Current + Error_upper
    }
}
    
```

제안한 알고리즘의 전체적인 순서도는 아래의 그림 2와 같다. 제안한 알고리즘에서 선택적 탐색 영역 설정을 위해 다이아몬드 탐색 기법에 기반한 방법을 사용하였다[4]. 또한, P 픽처와 B 픽처에 다른 탐색 방법을 적용하여 움직임예측의 정확도를 높였다.

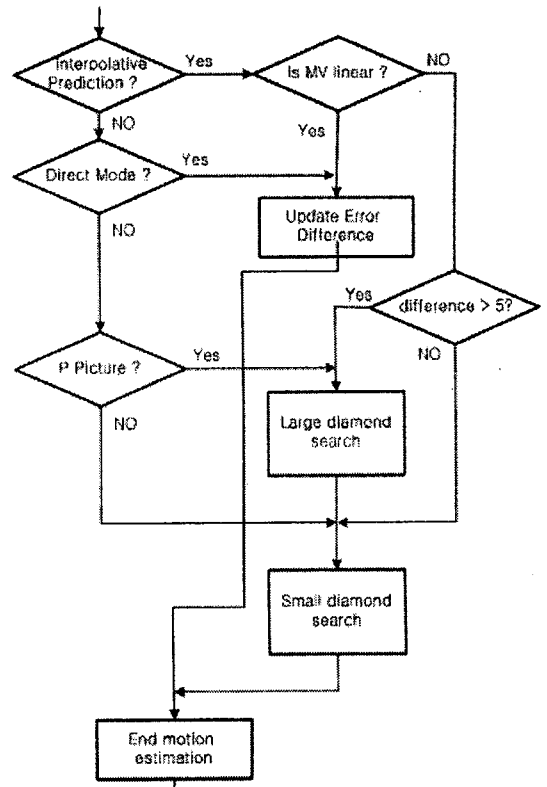


그림 2. 전체적인 순서도

3. 실험 결과

실험에서 CIF(352×288) 해상도를 가지고 4.7Mbps에서 부호화된 MPEG-2 영상을 사용하였고 GOP 구조는 N=12, M=4를 사용하였다. 복호된 영상은 제안한 방법에 의해 H.264/AVC 베이스라인 부호기에서 부호화되었다. 제안한 방법은 GOP 구조를 변경하는 다른 트랜스코딩에서도 사용될 수 있으므로 H.264/AVC의 인터모드는 16×16으로 제한하였다. Rate Control 기능은 사용하지 않았고 RD 최적화 기능을 사용하였다.

표 2. 제안한 알고리즘의 성능비교

sequence		PSNR	Bit-rate	Search-point
Akiyo	Full(4)	39.5	6588	32076
	Full(2)	39.51	6588	9900
	Diamond	39.49	6597	8904
	proposed	39.50	6615	1248
Bus	Full(4)	34.45	56200	32076
	Full(2)	34.45	56974	9900
	Diamond	34.45	56531	8584
	proposed	34.45	56805	3201
Coastguard	Full(4)	34.37	51485	32076
	Full(2)	34.37	51820	9900
	Diamond	34.37	51477	8303
	proposed	34.37	51599	2662
Football	Full(4)	34.32	75787	32076
	Full(2)	34.33	76014	9900
	Diamond	34.32	75786	10748
	proposed	34.32	75980	5519
Foreman	Full(4)	36.47	19813	32076
	Full(2)	36.47	19923	9900
	Diamond	36.45	20006	10391
	proposed	36.45	19934	3765
Table tennis	Full(4)	35.51	30936	32076
	Full(2)	35.51	31177	9900
	Diamond	35.50	31268	9301
	proposed	35.51	31390	3310

위의 표 2는 bit-rate, search-point, PSNR에 있어서의 다양한 트랜스코딩 방법의 성능을 나타내고 있다. 전역탐색방법에서는 식 (1)에 의해 유도된 PMV주변 [-2,2], [-4,4]에서 전역탐색방법에 의한 결과를 나타내고 있다. 다이아몬드탐색에 있어서는 식(1)에 의한 PMV에서 다이아몬드 탐색에 의한 방법의 결과이다. PSNR과 Bit-rate에 있어서 제안한 방법은 전역 탐색 방법과 거의 동일한 성능을 나타낸다. search-point에 있어서는 제안한 방법이 전역탐색방법과 다이아몬드 방법과 비교하여 45%에서 88%까지 연산량을 절감한 것을 볼 수 있다.

그림 3은 Bus 시퀀스의 RD 성능을 보여준다. 제안한 방법이 [-2,2]의 영영을 갖는 전역탐색방법을 능가하는 것을 볼 수 있다. Football 시퀀스에 있어서의 RD 성능은 그림 4에 나타나 있다. PMV 주변[-2,2]에서의 전역탐색방법과 동일한 성능을 보임을 관찰할 수 있다.

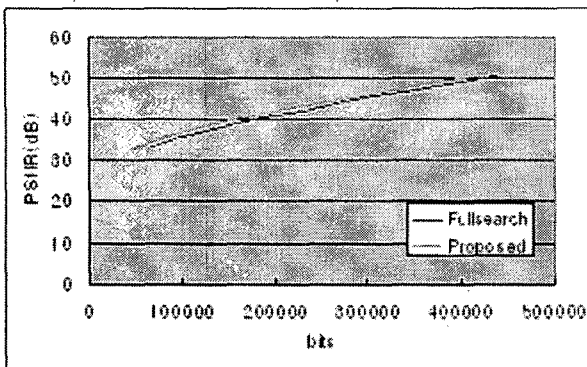


그림 3. Bus 시퀀스의 RD 성능

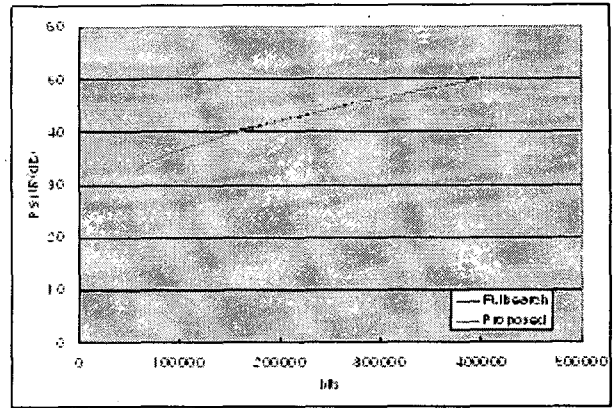


그림 4. Football 시퀀스의 RD 성능

4. 결론

본 논문에서는 GOP 구조 변환을 포함하는 트랜스코딩에 있어서 움직임 벡터의 선형성에 따른 선택적 탐색 패턴을 적용하여 움직임 추정의 복잡도를 줄이고 화질열화를 최소화하는 방법을 제안했다. 제안한 방법은 화질과, bit-rate를 유지하면 계산량을 획기적으로 줄였다.

<Acknowledgement>

본 연구보고서는 정보통신부 출연금으로 ETRI, SoC산업진흥센터에서 수행한 IT SoC 핵심설계인력양성사업의 연구 결과입니다.

참 고 문 헌

- [1] "Generic Coding of Moving Pictures and Associated Audio Information - Part 2: Video," ITU-T and ISO/IEC JTC 1, ITU-T Recommendation H.262 and ISO/IEC 13818-2(MPEG-2), 1994.
- [2] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, ITU-T Rec. H264|ISO/IEC 14496-10 AVC, 2003.
- [3] T. Shanableh and M. Ghanbari, "The importance of the bi-directionally predicted picture in video streaming," IEEE Trans. Circuits Syst. Video Technol., vol. 11, pp. 402-402, March 2001.
- [4] Shan Zhu and Kai-Kuang Ma, "A new diamond search algorithm for fast block-matching motion estimation," IEEE Trans. Image Proc. vol. 9, no. 2, Feb. 2000.