

레터논문 (Letter Paper)
방송공학회논문지 제22권 제3호, 2017년 5월 (JBE Vol. 22, No. 3, May 2017)
<https://doi.org/10.5909/JBE.2017.22.3.397>
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

Homography 참조 픽처를 사용한 화면 간 예측 효율 향상 방법

김 태 현^{a)}, 박 광 훈^{a)‡}

Improvement of Inter prediction by using Homography Reference Picture

Tae Hyun Kim^{a)} and Gwang Hoon Park^{a)‡}

요 약

최근 드론, 액션캠 등과 같은 촬영 장비의 활성화로 다양한 전역 움직임을 내포한 영상들이 많이 생성되고 있다. 이때 회전, 확대, 축소 등의 움직임이 발생한 경우, 2D motion vector를 활용하는 기존의 화면 간 예측 방법은 높은 부호화 효율을 기대하기 어렵다. 본 논문에서는 전역 움직임을 homography 참조 픽처를 통해 반영한 비디오 부호화 방법을 제안한다. 제안방법으로, 1) 현재 픽처와 참조 픽처간 전역 움직임 관계를 homography로 파악하여 새로운 참조 픽처를 생성하는 방법, 2) homography 참조 픽처를 화면 간 예측에 활용하는 방법이 있다. 실험은 HEVC 참조 소프트웨어인 HM 14.0에 적용하였고, 실험결과 RA 기준 6.6% 부호화 효율이 증가했다. 특히, 회전 전역 움직임을 지니는 영상을 이용한 실험 결과에서는 기존대비 최대 32.6%의 부호화 효율이 증가하는 결과를 나타내어, 드론과 같이 복잡한 전역 움직임이 자주 나타나는 비디오에서 높은 효율을 보일 수 있을 것으로 기대된다.

Abstract

Recently, a lot of images containing various global movements have been generated by the activation of the photographic equipment such as the drone and the action cam. In this case, when the motion such as rotation, scaling is generated, it is difficult to expect a high coding efficiency in the conventional inter-picture prediction method using the 2D motion vector. In this paper, we propose a video coding method that reflects global motion through homography reference pictures. As a proposed method, there are 1) a method of generating a new reference picture by grasping a global motion relation between a current picture and a reference picture by homography, and 2) a method of utilizing a homography reference picture for inter-picture prediction. The experiment was applied to the HEVC reference software HM 14.0, and the experimental result showed an increase in encoding efficiency of 6.6% based on RA. Especially, the results using the videos with rotational motion have a maximum coding efficiency of 32.6%, which is expected to show high efficiency in video, which is often represented by complex global motion such as drones.

Keyword : HEVC(High Efficiency Video Coding)/H.265, MPEG(Moving Picture Experts Group), JVET(Joint Video Exploration Team), Homography

a) 경희대학교 컴퓨터공학과(Dept. of Computer Science and Engineering, Kyung Hee University)

‡ Corresponding Author : 박광훈(Gwang Hoon Park)

E-mail: ghpark@khu.ac.kr

Tel: +82-31-201-3680

ORCID: <http://orcid.org/0000-0001-7133-8285>

· Manuscript received December 6, 2016; Revised February 2, 2017; Accepted February 2, 2017.

※ 이 논문은 2017년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (B0117-17-1006, 초고실감 미디어 서비스 실현을 위해 HEVC/3DA 대비 2배 압축을 제공하는 5세대 비디오/오디오 표준 핵심 기술 개발 및 표준화)

I. 서론

최근 드론(Drone) 촬영, 액션캠(Action cam), 자동차 블랙박스, 스마트 폰 카메라와 같은 촬영 장비의 직접적이고 다양한 움직임이 많은 동영상 촬영 시스템들이 도입되고 있다. 이러한 시스템으로 촬영된 영상들은 촬영 장비 자체의 움직임 발생으로 인하여 프레임 간 전역 움직임이 많이 발생한다. 기동성이 높은 항공 촬영이 가능한 드론 촬영이나, 보다 극적인 움직임을 촬영하는 액션캠의 경우 영상의 회전과 같은 복잡한 움직임이 빈번하게 발생한다. 이러한 복잡한 움직임의 발생은 동영상 부호화 효율을 감소시키는 원인이 될 수 있다. 본 논문에서는 HEVC/H.265에서 화면 간 예측에 존재하는 문제점을 거론하고 문제점에 대한 해결 방안을 제시하여 부호화 효율을 증가시키는 방법을 제안한다.

II. 기존 연구 및 문제점

기존 화면 간 예측 기술에 활용되는 블록 단위 2차원 움직임 벡터 방식은 동영상의 각 프레임을 작은 정사각형 또는 직사각형 단위의 블록으로 나누어, 블록 단위의 움직임을 찾아 2차원 움직임 벡터로 표현하는 방식을 말한다. 블

록의 크기가 작을수록 화면 간 예측의 정밀도가 높아지나 부호화해야 하는 2차원 움직임 벡터의 수가 늘어나고, 블록의 크기가 클수록 화면 간 예측의 정밀도가 떨어지나 부호화해야 하는 2차원 움직임 벡터의 수는 줄어든다. 2차원 움직임 벡터 방식은 화면 내 물체의 평행 이동만을 표현할 수 있다. 따라서 각 블록 영역의 움직임이 평행 이동일 경우에는 화면 간 예측 정밀도가 높아 부호화 효율이 높아지지만, 블록 영역의 움직임이 비평행 이동일 경우에는 화면 간 예측 정밀도가 떨어지게 된다.

HEVC/H.265에서는 이러한 문제를 해결하기 위해 다양한 부호화 구조를 적용하고 있다. CTU(Coding Tree Unit) 구조는 쿼드트리(Quad tree) 방식으로 4개의 정사각형 모양으로 단계적으로 분할 되거나 분할되지 않을 수 있다. 이를 통해 영역별로 부호화 효율을 고려한 서로 다른 크기의 블록 크기를 지정할 수 있다. CU는 예측 부호화 기법에 사용하기 위해 PU(Prediction Unit)로 분할 될 수 있으며, PU는 다양한 직사각형들의 조합으로 분할 될 수 있어 보다 다양한 형태의 분할을 지원한다. 서로 다른 움직임이 일어나는 물체 경계를 고려하여 화면 간 예측에서 각 블록이 가급적 동일한 움직임을 가지는 영역에 위치하도록 분할하며, 화면 간 예측 정밀도와 부호화하는 움직임을 벡터를 포함한 부가 정보들을 고려하여 분할 여부를 결정하는 것으로 높은 부호화 효율을

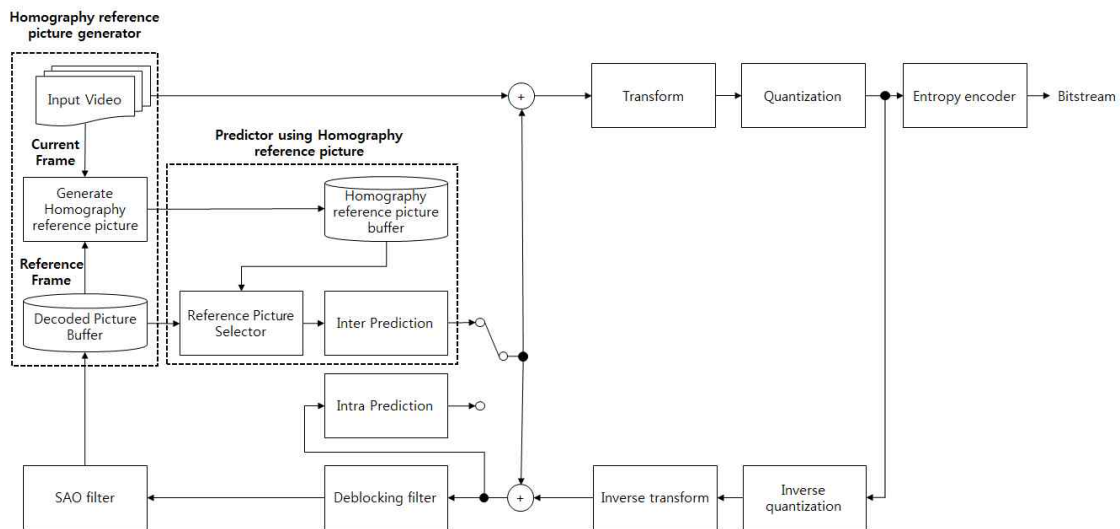


그림 1. 제안하는 알고리즘이 적용된 부호화기 구조
Fig. 1. Encoder structure with proposed algorithm

가질 수 있다. 그러나 분할 구조의 종류가 다양하거나, 분할이 많이 발생하게 될 경우 분할 구조에 대한 Bit-stream 내 정보가 늘어나 부호화 효율을 떨어뜨릴 수 있다^{[1][2]}.

III. 제안된 Homography 참조 픽처를 활용한 화면 간 예측 방법

Homography 참조 픽처를 활용한 동영상 부호화기 구조는 <그림 1>과 같다. 본 논문의 제안 방법은 크게 Homography 참조 픽처 생성부와 Homography 참조 픽처를 활용한 화면 간 예측부로 나뉜다.

Homography를 이용한 참조 픽처는 현재 프레임과 참조 프레임의 전역 움직임 관계를 파악하여 참조 픽처를 현재 픽처와 유사한 상태로 기하 변환(Warp) 하여 생성한다^[3]. 현재 픽처와 참조 픽처는 각각 특징점을 찾고 묘사(Description)하여 비교한 후, 현재 프레임에서 찾아진 특징점들을 참조 프레임에서 찾아진 특징점들 중 가장 유사한 특징점과 매칭시킨다. 매칭된 특징점 그룹을 이용하여 두 영상 간 전역 움직임 정보를 3x3 행렬로 도출해낸다. 이때 매칭된 특징점 그룹 중 outlier들을 배제하고 유효 매칭만을 활용한다. 도출한 행렬은 참조 픽처의 한 지점을 현재 픽처의 대응되는 지점으로 기하 변환하는데 사용한다. 도출한 행렬은 복호화기에서 동일하게 도출할 수 없으므로 참조 픽처마다 행렬 정보를 부호화 할 필요가 있다. 행렬의 값 중 하나를 1로 정규화 하여 매개변수 9개 중 8개만을 부호화할 수 있다. 색차(Chroma) 영상의 행렬은 휘도(Luma) 영상의 행렬로부터 유도된다. 색차 영상과 휘도 영상의 해상도가 동일한 YUV 4:4:4 포맷의 경우 색차 영상은 휘도 영상의 행렬과 동일한 행렬을 사용하여 Homography 참조 영상을 생성한다. 색차 영상의 해상도가 휘도 영상의 1/4인 YUV 4:2:0 포맷의 경우 휘도 영상의 행렬로부터 색차 영상의 행렬을 유도한다. 휘도 영상의 행렬이 식(1)의 H_l 과 같을 때, 색차 영상의 행렬은 식(1)의 H_c 와 같이 유도될 수 있다.

$$H_l = \begin{bmatrix} l_1 & l_2 & l_3 \\ l_4 & l_5 & l_6 \\ l_7 & l_8 & 1 \end{bmatrix}, H_c = \begin{bmatrix} l_1 & l_2 & l_3/2 \\ l_4 & l_5 & l_6/2 \\ 2l_7 & 2l_8 & 1 \end{bmatrix} \quad (1)$$

기하 변환된 픽처는 추가적인 참조 픽처로 화면 간 예측에 활용된다. 기존의 참조 픽처와 기하 변환된 참조 픽처 양쪽을 사용하여 예측하며 따라서 일반 참조 픽처와 기하 변환 참조 픽처 중 어느 참조 픽처를 선택하였는지를 나타내는 신호가 필요하다. 이때 기하 변환 참조 픽처를 선택하였는지를 나타내는 신호는 PU단위로 적용되기 때문에 Bit-Stream내에서 많은 양을 차지하여 부호화 효율을 저하시키게 된다. 이러한 부호화 효율 저하를 줄이기 위해서 Merge 모드에 기하 변환 참조 픽처의 선택 여부 정보를 포함하여 활용한다. 각 Merge 후보는 해당 후보를 복호화할 때 사용된 기하 변환 참조 픽처를 사용 여부 정보를 가지고 있으며, 현재 PU는 선택된 Merge 후보와 동일하게 기하 변환 참조 픽처 사용 여부를 결정한다. 본 제안에서는 화면 간 예측의 정밀성이 상대적으로 높아 Merge 모드의 선택빈도가 높아지므로, Merge 모드를 활용한 기하 변환 참조 픽처 선택 정보의 예측 방법으로 추가 전송 정보를 많이 줄일 수 있다^[4].

IV. 실험 및 결과

본 논문에서 제안한 알고리즘의 성능을 확인하기 위하여 HEVC 참조 소프트웨어인 HM 14.0에 구현하여 HM14.0과의 비교실험을 수행하였다. 실험영상은 1920x1080 해상도 영상을 사용하였으며, 전역 움직임 특성에 따라 평행이동(Panning) 4종, 회전(Rotate) 4종, 확대(Zoom in) 3종, 축소(Zoom out) 3종의 네 그룹 14개 영상을 사용하였다. 실험조건은 HEVC의 공통 실험 조건의 저지연-P(Low-delay-P) 환경과 저지연(Low-delay) 환경, 임의 접근(Random-access) 환경, 그리고 Main 프로파일을 준수하였다. 제안 방법의 부호화 효율을 측정하기 위해 BD-Bitrate 방법을 사용하였다^[5]. <표1>는 제안한 방법을 이용하여 부호화한 BD-Bitrate를 나타낸 것이다. 저지연-P 환경에서 평균 12.8%, 저지연 환경에서 평균 9.3%, 임의 접근 환경에서 평균 6.6%의 기존 대비 부호화 효율 증대를 확인하였다. 각 그룹별 실험 결과는 평행이동 영상에서는 부호화 효율이 감소되는 경향을 보이며, 회전 영상에서는 높은 부호화 효율의 증가가 관측되었다. 기존의 움직임 벡터 방식으로 화면 간

표 1. HEVC/H.265 대비 제안 알고리즘의 BD-rate
 Table 1. BD-rate of proposed algorithm compared to HEVC / H.265

Sequence Type	Low-delay-P Main			Low-delay Main			Random-access Main		
	Y BD-rate	Cb BD-rate	Cr BD-rate	Y BD-rate	Cb BD-rate	Cr BD-rate	Y BD-rate	Cb BD-rate	Cr BD-rate
Panning	3.7%	4.9%	3.9%	9.1%	8.9%	8.9%	5.9%	6.2%	6.2%
Rotate	-32.6%	-26.5%	-27.1%	-28.5%	-21.0%	-21.4%	-20.6%	-16.0%	-16.0%
Zoom in	-7.3%	-4.4%	-3.0%	-2.8%	-0.5%	0.2%	-4.1%	-1.8%	-1.4%
Zoom out	-14.0%	-12.1%	-13.1%	-14.9%	-10.9%	-11.3%	-7.1%	-5.4%	-5.6%
Overall Average	-12.8%	-9.7%	-10.1%	-9.3%	-5.9%	-5.9%	-6.6%	-4.3%	-4.3%

예측의 수행이 잘 이루어지는 평행이동의 경우 제안 방법을 사용할 시 추가 정보 송신이 필요하여 부호화 효율이 감소하는 반면, 기존의 움직임 벡터 방식으로 예측이 어려운 회전이동의 경우 추가 정보를 송신함에도 화면 간 예측 정밀도가 높아져 부호화 효율이 높아진다. 실험을 통해 비평행인 복잡한 전역 움직임을 가지는 영상에서 제안 방법이 적용되면 부호화 효율이 높아지는 것으로 분석된다.

V. 결론

본 논문에서는 드론과 같은 촬영 장비의 다양화로 인한 보다 복잡한 전역 움직임을 가지는 영상생성이 많아지는 추세에 대응하여, 전역 움직임을 반영한 참조 픽처를 생성하고 이를 이용한 부호화 방안에 대하여 다루었다. 실험 결과를 통해 기존 2차원 움직임 벡터 방식으로 화면 간 예측 효율이 저하되는 영상에서 효율을 볼 수 있음을 확인하였다. 본 논문에서 제안하는 알고리즘을 통해 향후 보다 다양성을 가지고 복잡해져 가는 영상 속 움직임에 대한 부호화 효율이 증대될 것으로 기대된다. 본 논문에서 얻어진 결과

를 통해, 향후 제안 알고리즘으로 인해 발생하는 부가정보를 보다 효율적으로 압축하여 더욱 부호화 효율을 높일 수 있는 방안에 대한 연구가 진행되어야 하겠다.

참고 문헌 (References)

- [1] G. J. Sullivan, J. M. Boyce, Y. Chen, J. R. Ohm, C. A. Segall, and A. Vetro, "Standardized extensions of high efficiency video coding (HEVC)". IEEE Journal of selected topics in Signal Processing, Vol.7, No.6, pp.1001-1016, Dec 2013.
- [2] I. K. Kim, J. Min, T. Lee, W. J. Han, and J. H. Park, "Block Partitioning Structure in the HEVC Standard", IEEE Transactions on CSVT, Vol.22, No.12, pp.1697-2012, Dec 2012.
- [3] O. D. Faugeras, and F. Lustman. "Motion and structure from motion in a piecewise planar environment." International Journal of Pattern Recognition and Artificial Intelligence Vol.2, No.3, pp.485-508, June 1988.
- [4] G. J. Sullivan, J. R. Ohm, W. J. Han and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard", IEEE Trans. Circuits Syst. Video Technol., Vol. 22, No. 12, pp.1649-1668, Sep 2012.
- [5] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," ITU-T SG16 Q.6, VCEG-M33, Texas, USA, Apr. 2001.