

특집논문-06-11-4-04

다시점 동영상 부호화를 위한 가변형 다시점GOP 예측 구조

윤재원^{a)}, 서정동^{a)}, 김용태^{a)}, 박창섭^{b)}, 손광훈^{a)‡}

Flexible GOP prediction structure for multi-view video coding

Jaewon Yoon^{a)}, Jungdong Seo^{a)}, Yongtae Kim^{a)}, Changseob Park^{b)}, and Kwanghoon Sohn^{a)‡}

요약

본 논문에서는 다시점 동영상 부호화를 위한 참조 소프트웨어의 부호화 성능을 높이기 위해 가변형 다시점GOP 예측 구조로 부호화 하는 방법을 제안한다. 다시점 동영상 부호화를 위한 참조 소프트웨어에서는 고정된 시공간 예측구조를 사용하여 다시점 동영상을 부호화한다. 그러나 다시점 동영상 부호화의 성능은 영상의 특성에 따라 예측 부호화 구조를 가변적으로 변경하는 것에 영향을 받는다. 따라서 다시점 동영상의 전역 변이를 이용하여 부호화의 기준 시점을 정하고 카메라 간의 간격을 고려하여 B-픽처의 개수를 조절하여 영상의 특성에 따라 다시점 동영상의 부호화 단위인 다시점GOP 예측 구조를 가변적으로 적용하는 방법을 제안한다. 실험 결과에서 제안된 가변형 다시점GOP 예측구조의 부호화 방법이 기존의 참조 소프트웨어보다 우수한 성능을 보여 줌을 확인하였다. 제안 예측 부호화 구조는 기존의 부호화 구조와 비교하여 7.1%의 비트량 감소를 보였다.

Abstract

In this paper, we propose a flexible GOP prediction structure to improve coding efficiency for multi-view video coding. In general, reference software used for MVC uses the fixed GOP prediction structure. However, the performance of MVC depends on the base view and numbers of B-pictures between I-picture(or P-picture) and P-picture. In order to implement the flexible GOP prediction structure, the location of base view is decided according to the global disparities among the adjacent sequences. Numbers of B-pictures between I-picture(or P-picture) and P-picture are decided by camera arrangement such as the baseline distance among the cameras. The proposed method shows better result than the reference software of MVC. The proposed prediction structure shows considerable reduction of coded bits by 7.1%.

Keyword : MVC(Multi-view Video Coding), GOP(Group of GOP), camera arrangement

I. 서론

3차원 영상에 대한 관심이 증대되면서 많은 연구 기관들이 3차원 영상 처리와 디스플레이 시스템에 대한 연구 결과를 발표하였다. 유럽에서는 3DTV에 대한 연구로 디지털

스테레오 영상을 획득하고 부호화하고 전송하는 시스템을 개발하기 위한 목적으로 DISTIMA(Digital STereoscopic IMaging and Applications) 프로젝트가 수행되었다^[1-3]. DISTIMA의 후속 프로젝트로서 PANORAMA(PAKage for New OperRational Autostereoscopic Multiview systems and Applications)가 시작되었는데 이 프로젝트의 목적은 3차원 입체감을 표현하는 통신을 통해 시청자에게 생생한 현장감을 전달하는데 있다^{[4][5]}. IST(Information Society Technologies)의 VIRTUE(VIRtual Team User

a) 연세대학교 전기전자공학부

Dept. of Electrical and Electronic Engineering, Yonsei University

b) KBS방송기술연구팀

KBS Broadcast Technical Research Institute

‡ 교신저자 : 손광훈(khsohn@yonsei.ac.kr)

Environment) 프로젝트에서는 화상 회의 시스템을 통해 몰입감있는 입체감을 전달하고자 하였다^[6]. 이 프로젝트에서는 가상 환경에서 실물 크기의 상반신 비디오 영상을 표현하여 양방향 화상 회의를 구현하고자 하였다. 또한 국내에서는 ETRI(Electronics and Telecommunications Research Institute)가 FIFA(Federation Internationale de Football Association) 2002 한일 월드컵 때 스테레오 동영상 방송 시스템을 구축하여 실험 방송하였다^{[7][8]}. 또한 일본에서는 NHK(Nippon Hoso Kyokai)의 3차원 HDTV(High Definition Television) 프로젝트등 3차원 관련 분야에 대한 여러 연구가 진행 중이다^{[9][10]}. 이러한 3차원 다시점 영상 시스템에 대한 관심 증대로 인해 MPEG(Moving Picture Expert Group) 3DAV(3 Dimensional Audio Video)가 ISO/IEC SC29WG11에 설립 되었고 현재는 ISO/IEC (International Organization for Standardization/International Electrotechnical Commission)와 ITU-T(International Telecommunications Union-Telecommunication)의 공동으로 합작 동영상 팀(Joint Video Team 이하, JVT)에서 SVC(Scalable Video Coding)의 참조 모델(Reference Model) 기반의 다시점 동영상 부호화(Multi-view Video Coding)에 대한 표준화가 진행 중이다. 다시점 동영상 부호화의 표준화 작업을 통해 교육, 훈련, 3차원 영화 산업, 의료 산업, 화상 회의 및 3차원 방송 등 여러 분야에서 더 자연스럽게 현실감 있는 영상을 전달할 수 있을 것으로 기대된다.

다시점 동영상은 여러 대의 카메라를 통해 획득한 영상이므로 영상간의 상관도가 높다. 다시점 동영상을 부호화할 때 시점 수에 비례하여 전송해야 할 영상의 정보량과 계산 복잡도가 증가하게 된다. 그러나 다시점 동영상을 부호화하는데 있어서 영상간의 유사성을 이용하면 압축 효율을 높일 수 있다.

다시점 동영상 부호화의 참조 소프트웨어(Reference Software)는 계층적 B-픽처(Hierarchical B-picture)를 사용하였고 부호화 구조는 고정하여 수행되고 있다^[11]. 다시점 동영상 부호화의 성능은 카메라 구조에 따른 각 영상의 특성과 밀접한 관계가 있다. 다시점 영상들 중에서 가장 왼쪽의 카메라로 획득한 영상과 가장 오른쪽의 카메라로 획득한 영상의 화면은 다른 시점의 화면과 겹쳐지는 부분이 적다. 그러나 중앙에 위치한 카메라로 획득한 영상은 다른

시점의 영상과 겹쳐진 부분이 상대적으로 많게 된다. 따라서 다시점 비디오 압축할 때에 카메라 배열의 중앙에 위치한 카메라로 획득한 영상을 기준 시점으로 정하게 되면 좀 더 높은 압축 효율을 얻을 수 있을 것이다. 그러나 다시점 동영상 부호화의 참조 소프트웨어의 부호화 구조는 영상의 특성에 관계없이 가장 왼쪽의 카메라에서 획득한 영상을 기준 시점으로 고정시켜 놓았다. 실험 영상의 각각의 특성에 맞도록 부호화 구조를 변경한다면 예측 부호화를 수행할 때에 좀 더 신뢰도 높은 변이 벡터를 찾게 되고 이를 통해서 압축 성능을 높일 수 있을 것이다.

또한, 다시점 영상에서 I-픽처(또는 P-픽처)와 P-픽처 사이의 B-픽처의 개수도 부호화 효율에 영향을 미치게 된다. 카메라 간의 간격이 넓을 때는 I-픽처(또는 P-픽처)와 P-픽처 사이의 B-픽처의 개수를 줄여서 중간 시점의 영상을 압축하는데 필요로 하는 참조 영상을 인접영상으로 이용한다. 그러나 카메라 간의 간격이 좁을 때는 I-픽처(또는 P-픽처)와 P-픽처 사이의 B-픽처의 개수를 늘려도 참조하는 영상의 최대 변이 값이 크지 않기 때문에 예측의 신뢰도가 떨어지지 않아 효율적으로 압축 할 수 있다.

본 논문의 전체적인 구성은 다음과 같다. 제2장에서는 다시점 동영상 부호화의 참조 소프트웨어에 대해서 설명하고, 제3장에서는 본 논문의 제안 부호화 방법인 가변형 다시점GOP(Group of GOP) 예측 구조에 대해서 설명하였다. 제4장에서는 제안된 다시점 동영상을 위한 부호화 방식과 기존의 방식을 비교하여 제안 방법의 우수성을 검증하였다. 마지막으로 제5장에서 결론을 맺는다.

II. 다시점 동영상 부호화 참조 소프트웨어

다시점 동영상 부호화 참조 소프트웨어는 실험 영상에 관계없이 시공간 예측 구조를 고정된 다시점GOP(Group of GOP) 단위로 부호화를 수행한다. 그림 1은 8시점 영상에 GOP(Group of GOP)의 크기가 8인 다시점 동영상 부호화 참조 소프트웨어의 시공간 예측 구조를 보여준다^[11]. 그림 1에서 세로 방향은 시점 축을 나타내고 가로 방향은 시간 축을 나타낸다. 다시점 동영상 부호화 참조 소프트웨

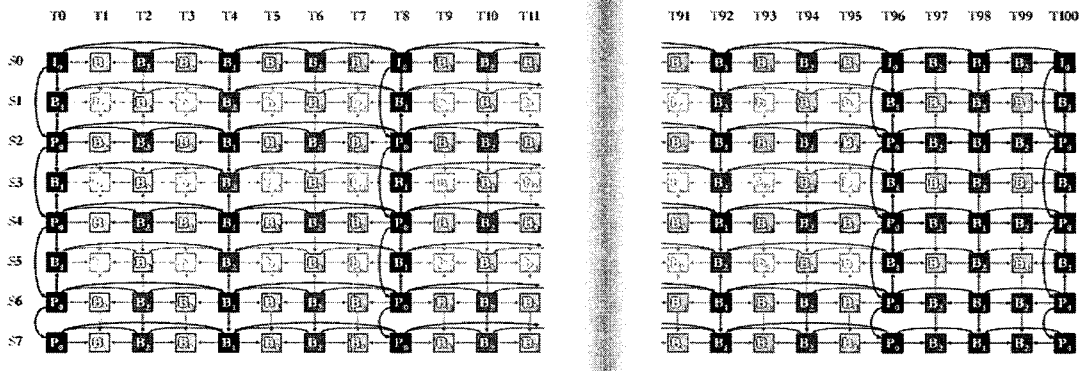


그림 1. 다시점 동영상 부호화의 참조 소프트웨어의 시공간 예측 구조
 Fig. 1. Inter-view temporal prediction structure of Reference software of MVC

어의 기본적인 부호화 구조는 각 시점마다 계층적 B-픽처를 사용한다. 그리고 그림 1의 S1, S3, S5에서는 시점간의 예측(inte-view prediction)을 통해서 부호화를 수행한다. 짝수개의 시점을 가진 영상에서 마지막 시점(그림 1에서 S7)의 예측 구조는 그림 1의 S2, S4, S6과 다르게 시점간의 예측을 이용한다. 부호화의 동기를 위해서 각 GOP(S0/T0, S0/T8)는 I-frame으로 시작한다. 또한 그림 1은 영상의 전체 길이가 GOP의 정수배와 맞지 않을 때 영상의 마지막 부분은 짧은 길이를 가지는 GOP로 부호화 하는 것을 보여준다. 시점간의 예측을 위한 참조 영상의 처리를 위해서 다시점 동영상 부호화 참조 소프트웨어의 Decoded Picture Buffer는 $2 \times GOP_length + number_of_views$ 로 할당한다.

다시점 동영상 부호화의 참조 소프트웨어에서 효과적인

메모리 관리를 위한 프레임 재배열은 그림 2와 같다. 부호화는 모든 시점의 첫 번째 프레임을 순서대로 수행하고 한 개의 다시점GOP의 나머지 프레임들은 시간 축을 따라 지그재그 방향으로 진행된다. 지그재그 패턴은 다음 다시점GOP에도 적용된다. 따라서 다시점 동영상 부호화의 참조 소프트웨어에 사용되는 입력 영상은 부호화하기 전에 하나의 yuv파일로 재배열하는 과정이 필요하게 된다. 그러므로 다시점 동영상 복호화기에서는 하나의 yuv결과 파일을 각 시점의 파일로 나누는 역재배열 과정을 수행해야한다.

다시점 동영상 부호화의 참조 소프트웨어는 부호화 수행을 위해서 참조 영상 리스트 재순서화(Reference Picture List Reordering, 이하 RPLR)와 메모리 관리 조절(Memory Management Control Operation, 이하 MMCO)의 관리를 해야 한다.

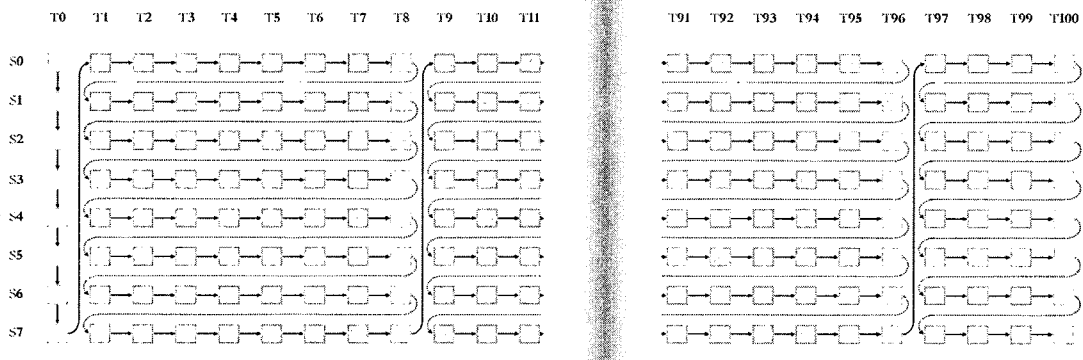


그림 2. 다시점 영상의 프레임 재배열
 Fig. 2. Frame reordering of multi-view sequence

RPLR은 부호화할 프레임의 참조 영상이 버퍼(buffer)상에 어느 위치에 있는지 압축할 프레임과 참조 영상사이의 상대적 거리를 나타낸다. 버퍼의 용량에는 한계가 있기 때문에 참조 영상으로 사용될 프레임을 버퍼상에 저장하고 있다가 더 이상 참조 영상으로 쓰이지 않을 영상을 버퍼에서 지우기 위한 명령어가 MMCO이다. RPLR과 MMCO의 관리가 되어야만 다시점 동영상을 오류 없이 압축을 수행할 수 있게 된다.

다시점 동영상 부호화는 기존의 2차원 압축 표준(H.264/AVC)에 시점간의 예측 구조를 사용하기 때문에 효율적인 부호화를 할 수 있다. 시점간의 예측은 변이 추정을 수행하여 변이 벡터를 사용하는 것이다. 변이 벡터는 좌영상과 우영상이 대응되는 지점의 상대적인 벡터로 정의된다. 변이 벡터의 예측은 참조 영상으로부터 현재영상을 복원하고 객체의 정보를 표현하기 때문에 매우 중요하다. 움직임 벡터와 비교했을 때, 변이 벡터는 두 가지 큰 특성을 가지고 있다^[12-14]. 첫째, 동영상에서 움직이는 물체만이 움직임 벡터를 가지고 있고 배경과 정지된 객체는 움직임 벡터를 가지고 있지 않다. 이런 성질은 움직이는 객체만을 움직임 추정 및 보상 기법으로 부호화하여 높은 부호화 효율을 얻을 수 있게 해준다. 그러나 다시점 영상에서 모든 객체와 배경은 변이 벡터를 가지고 있다. 그러므로 변이 추정은 카메라 사이의 간격이 좁을수록 부호화 효율이 높아지게 되고, 움직임 추정은 작은 수의 객체가 적게 움직일 때 적합하다. 두 번째, 동영상의 가려진 부분은 움직이는 객체에 의해서 발생한다. 그리고 다시점의 가려진 영역은 변이 벡터의 불연속과 보이는 영역의 제한으로 발생하게 된다. 가려진 영역은 움직임과 변이 추정의 성능을 감소시킨다. 많은 연구가 가려진 영역에 대해 진행되어왔으며 다양한 알고리즘이 개발되었지만 가려진 영역을 완벽하게 감지하고 부호화하는 것은 불가능하여 추가적인 정보를 전송해야한다^{[15][16]}.

다시점 영상의 부호화 효율은 시점간의 변이 추정 성능의 영향을 받는다. 변이 추정의 신뢰도를 높이기 위해서는 영상간의 상관도를 고려하여 기준 시점을 중앙에 위치하도록 하는 것이 유리하다. 그림 1에서 다시점 동영상 부호화의 참조 소프트웨어의 부호화 기준 시점은 영상의 특성에 관계없이 가장 왼쪽 시점으로 잡아 압축하기 때문에 예측 효율이 떨어질 수 있다.

다시점 동영상 부호화에서 기준 시점과 압축 효율의 관계를 알아보기 위한 실험을 해보았다. 동일한 영상에 대해

서 기준 시점을 1번째 시점, 4번째 시점, 8번째 시점으로 다르게 하여 부호화를 수행하였다. 그림 3은 "Ballroom" 영상의 기준 시점에 따른 부호화 결과를 보여준다. 기준 시점이 중앙에 위치할 때 가장 좋은 결과를 보였다. 그림 3의 결과로부터 다시점 동영상 부호화에서는 기준 시점이 부호화 성능에 영향을 미치고 여러 시점들 중에서 중앙에 위치할 경우에 부호화 효율이 높아지는 것을 알 수 있다.

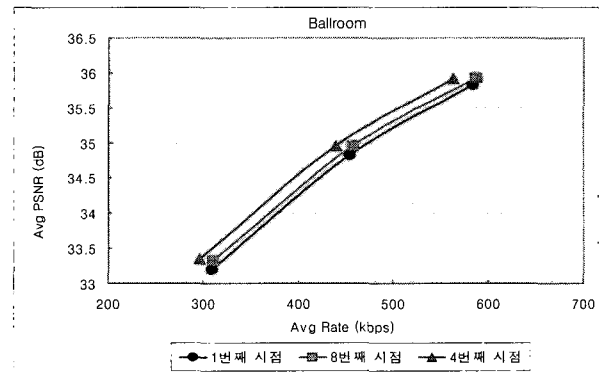


그림 3. "Ballroom"영상에서 기준 시점에 따른 부호화 결과
Fig 3. Results for "Ballroom" according to the base view

III. 가변형 다시점GOP 예측 구조

그림 4는 가변형 다시점GOP 예측 구조 결정과정의 블록도이다. 제안한 다시점 동영상 부호화에 다시점 영상이 입

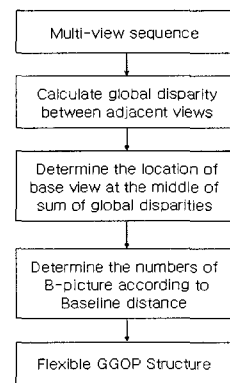


그림 4. 가변형 다시점GOP 예측 구조의 블록도
Fig. 4. Blockdiagram of the flexible GGOP prediction structure

력으로 들어오면 첫 번째 단계로 각 인접 시점 사이의 전역 변이(Global disparity)를 구하게 된다. 구해진 전역 변이들의 합의 중앙에 기준 시점을 정하게 된다. 기준 시점이 정해지게 되면 I-픽처(또는 P-픽처)와 P-픽처 사이에 존재하는 B-픽처의 개수를 조절하게 된다. B-픽처의 개수는 부호화하기 전에 알고 있는 카메라 파라미터를 이용하여 결정한다. 카메라 파라미터 중에서 카메라 간의 간격정보를 고려하여 B-픽처의 개수를 0~2로 조절한다.

1. 기준 시점 결정

가변형 다시점GOP 예측 구조를 위해 부호화의 기준이 되는 I-픽처의 위치를 결정하는 것이 중요하다. 기준 시점이 될 I-픽처의 위치가 중요한 이유는 두 가지가 있다. 첫째, 다시점 영상에서 변이 추정을 수행할 때 참조 영상과 부호화할 영상의 정합 영역이 많아야 변이 추정의 성능이 향상되기 때문이다. 기준 시점은 다른 시점을 예측 부호화하기 위해 참조 영상으로 사용된다. 따라서 다른 시점과 정합 영역이 많도록 기준 시점이 중앙에 위치하는 것이 부호화 효율을 높일 수 있을 것이다. 두 번째, 다시점 영상 간에는 서로 중복되는 정보를 많이 포함하고 있다. 그러므로 각 시점을 2차원 압축 표준(H.264/AVC)으로 독립적으로 부호화하는 것보다 다른 시점을 참조 영상으로 하여 변이 추정 및 보상 기법을 사용하여 부호화 하는 것이 더 효율적이다. 다시점 동영상 부호화에서는 참조 영상으로는 최대 4개

를 사용한다. 움직임 방향으로 2개, 변이 방향으로 2개의 참조 영상을 사용한다. 이때 변이 방향의 참조 영상으로 많이 쓰이게 될 기준 시점을 다른 영상과 중복되는 정보를 많이 포함하는 것으로 선택하여야 압축 성능을 높일 수 있기 때문이다.

그림 5는 "Ballroom"영상에서 기준 시점을 결정하는 과정을 간략하게 보여주고 있다. 이 과정을 자세히 설명하면 다음과 같다. 본 논문에서는 기준 시점이 될 I-픽처를 결정하기 위해 다시점 영상의 인접 영상간의 전역 변이 값을 사용하였다. N-시점의 영상이라면 구해진 인접한 영상간의 전역 변이의 개수는 N-1개가 된다. 전역 변이는 수식 (1)으로 구한다. 수식 (2)에서 SAD(Sum Absolute Difference) 항은 두개의 인접 영상간의 유사도를 표현한다. I_L 과 I_R 은 각각 인접한 두 영상의 왼쪽 영상과 오른쪽 영상으로 나타낸다. d_j 와 d_k 는 각각 수평 방향, 수직 방향의 변이 벡터의 값이고 SR_H 과 SR_V 은 수평 방향의 탐색 범위와 수직 방향의 탐색 범위를 나타낸다. 수식 (2)에서 수평, 수직 방향의 탐색 범위만큼 변이 추정을 수행하고 SAD를 최소화 시키는 d_j 를 k 번째 전역 변이 벡터 GD_k 로 정한다. 구해진 첫 번째 전역 변이 벡터 값과 가장 마지막 전역 벡터 값을 비교하여 작은 값을 가지는 전역 변이 벡터 값에 인접한 다른 전역 벡터 값을 더하고 다시 두개 값의 크기를 비교한다. 예를 들어, 첫 번째 전역 변이 벡터 값과 N-1번째 전역 변이 벡터 값을 비교하여 첫 번째 전역 변이 벡터 값이 작다면 인접한 두 번째 전역 변이 값을 첫 번째 전역 변이 벡터와 더하여 다시 N-1번째 전역 변이 값과 비교한다.

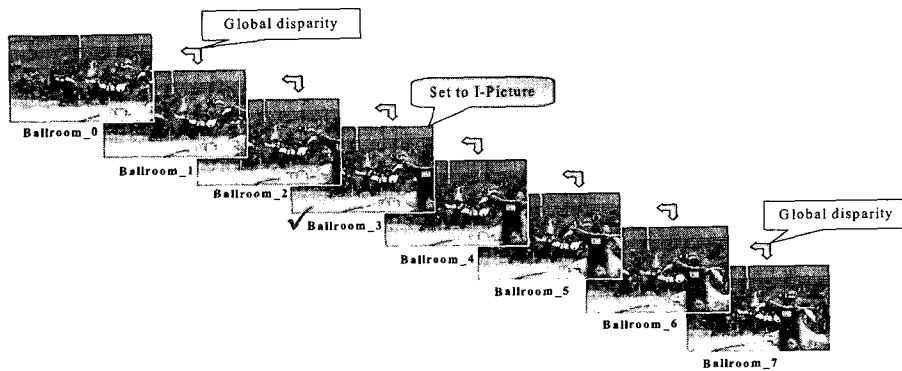


그림 5. "Ballroom"영상에서 기준 시점 결정의 예
 Fig. 5. Example of the determining process as to the base view in "Ballroom"

양쪽에서 더해지는 전역 변이 값을 비교 할 수 없을 때까지 이 과정은 계속한다. 비교 과정이 끝나면 양쪽에서 더해진 전역 변이 값의 왼쪽, 오른쪽 두개의 인덱스 값이 구해진다.

$$GD_k = \arg \min_{d_j} f_k, \quad k=1, \dots, N-1 \quad (1)$$

$$f_k = \sum_{xy \in \text{frame}} |I_L(x,y) - I_R(x+d_j, y+d_i)|, \quad d_j \in \mathcal{S}_H, d_i \in \mathcal{S}_V \quad (2)$$

두개의 인덱스 중에서 기준 시점을 정하기 위해서 두개의 영상 화소의 공분산을 이용한다. 공분산은 수식 (3)으로 구한다. 수식 (4)의 μ_L 과 μ_R 은 각각 왼쪽 영상과 오른쪽 영상의 평균 화소값을 나타내고 M 은 영상의 가로 길이, N 은 세로 길이를 나타낸다. 수식 (3)에서 위의 수식 (2)를 통해 구해진 변이 벡터 d_j 와 d_i 를 이용하여 두 영상이 정합을 이루는 위치의 화소를 곱하고 두 영상의 평균값의 곱을 빼주어 공분산을 구하게 된다. 두 영상의 상관도가 높으면 공분산 값은 큰 값을 가진다. 신뢰도 높은 참조 영상은 영상간의 상관도가 높아야 하기 때문에 위에서 설명한 두개의 인덱스 중에서 공분산 값이 큰 인덱스를 기준 시점으로 결정하게 된다.

$$Cov_{LR} = \sum_{x,y \in \text{frame}} I_L(x,y) \times I_R(x+d_j, y+d_i) - \mu_L \mu_R \quad (3)$$

$$\mu_L = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N I_L(x,y) \quad (4)$$

$$\mu_R = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N I_R(x,y)$$

위의 식에서 구한 전역 변이 합의 중앙값과 공분산 값을 이용하여 기준 시점의 영상을 I-픽처로 부호화를 수행한다^[7].

2. B-픽처의 개수 결정

기준 시점이 정해지게 되면 I-픽처(또는 P-픽처)와 P-픽처 사이에 존재하는 B-픽처의 개수를 결정해야 한다. 본 논문에서는 카메라 간의 간격(baseline distance)을 고려하여 B-픽처의 개수를 결정한다. B-픽처의 개수는 다시점 비디오 압축 효율에 큰 영향을 미친다. 카메라간의 간격을 좁게 배열하여 획득한 다시점 영상 압축할 때에 B-픽처의 개수를 적게 두면 부호화할 영상과 참조영상의 거리가 짧기 때문에 부호화 예측의 정확도는 높아지지만 영상 압축에 사용하는 비트가 많아져 효율은 떨어지게 된다. 반대로 카메라 간격이 넓은 상태로 획득한 영상에 B-픽처를 많이 두고 부호화하게 되면 예측을 위해 사용하는 참조 영상과의 거리가 멀기 때문에 부호화 예측의 정확도가 떨어지게 된다. 따라서 다시점 영상을 부호화하기 전에 카메라간의 간격이 임의의 경계 값을 넘으면 I-픽처 (또는 P-픽처)와 P-픽처 사이에 존재하는 B-픽처의 개수를 늘리고 임의의 경계 값보다 작으면 적은 수의 B-픽처를 할당한다. 그림 6은 8시점의 입력 영상이 들어왔을 때, I-픽처 (또는 P-픽처)와 P-픽처 사이에 존재하는 B-픽처의 개수가 1개 일 때의 부호화 구조를 보여준다.

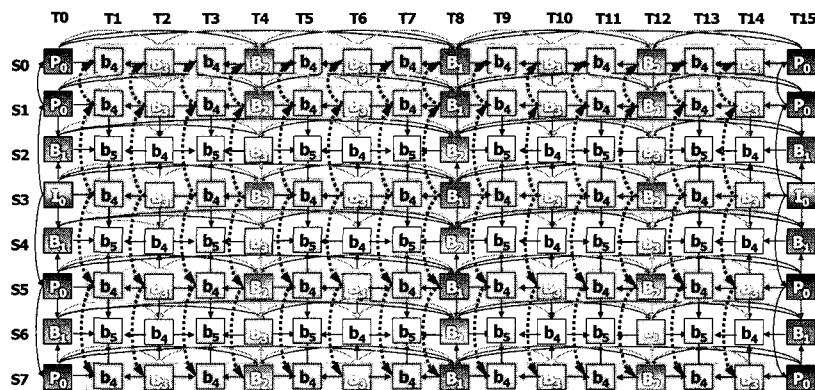


그림 6 8시점 영상에 적용된 가변형 다시점GOP 예측구조의 예
Fig. 6. An example of the proposed prediction structure

3. 가변형 다시점GOP 예측 구조

그림 6은 8시점 실험 영상에 적용된 가변형 다시점GOP 예측구조의 한 예를 보여주고 있다. 그림 6에서 기준 시점은 8시점 중에서 중앙에 위치하는 것을 보여준다. 기준 시점은 위에서 설명한 과정대로 전역 변이 합의 중간 값에 해당하는 시점으로 결정한다. 이렇게 결정된 기준 시점은 기준 시점 좌우에 존재하는 P-픽처의 참조 영상으로 사용된다. 그리고 카메라 간의 간격을 고려하여 B-픽처의 개수가 1개로 조절된 것을 알 수 있다. 기존의 다시점 동영상 부호화 참조 소프트웨어의 예측구조에서 P-픽처를 포함하는

GOP에서는 참조 시점의 변이 정보를 사용하지 않았다. 그러나 가변형 다시점GOP 예측구조에서는 P-픽처를 포함하는 GOP의 B-픽처에서 동일한 시간대에 존재하는 참조 시점의 변이 정보를 활용하는 시점간의 예측을 수행하였다. 다시점 동영상 부호화에서 사용하는 실험 영상 "Race1"은 카메라 간의 간격이 20cm이고 "Rena", "Akko&Kayo" 영상은 카메라들이 5cm 떨어져있다^[18]. "Race1"영상은 카메라 간격이 멀다고 판단하여 I-픽처 (또는 P-픽처)와 P-픽처 사이에 B-픽처의 개수를 1개로 결정한다. 그림 7은 "Race1" 영상의 참조 소프트웨어의 부호화 구조와 제안한 가변형 다시점GOP 예측 구조의 부호화 방식을 보여준다. 가변형 다시점GOP 예

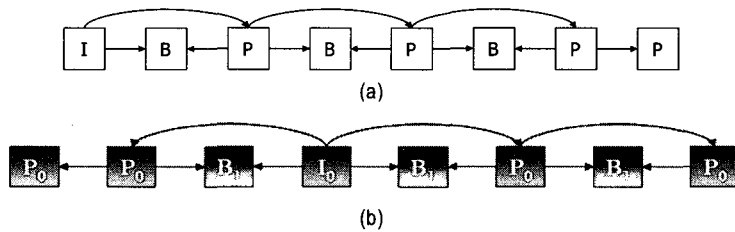


그림 7. 8시점 영상에 적용된 부호화 구조 (a) 참조 소프트웨어의 부호화 구조 (b) "Race1"영상에 적용한 가변형 다시점GOP 예측 구조
Fig. 7. Prediction structure for 8-view sequence (a) Prediction structure for reference S/W (b)Flexible GGOP prediction structure for "Race1"

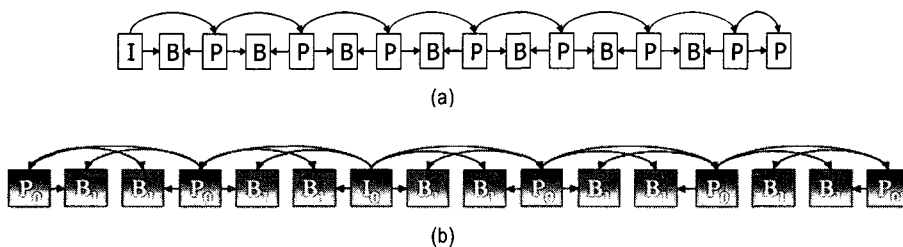


그림 8. "Rena" 영상에 적용한 부호화 구조 (a) 참조 소프트웨어의 부호화 구조 (b) 가변형 다시점GOP 예측 구조
Fig. 8. Prediction structure for "Rena" (a) Prediction structure for reference S/W (b) Flexible GGOP prediction structure for "Rena"

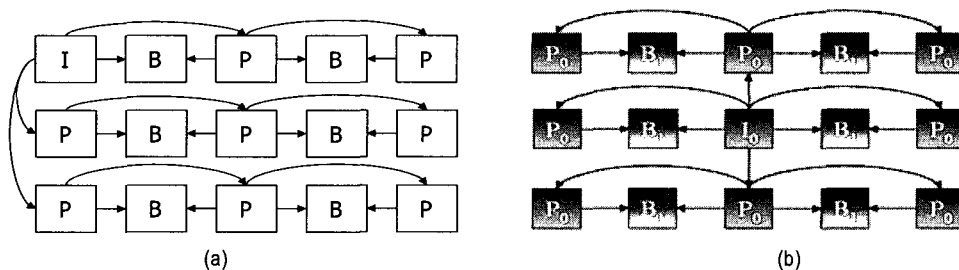


그림 9. "Akko&Kayo" 영상에 적용한 부호화 구조 (a) 참조 소프트웨어의 부호화 구조 (b) 가변형 다시점GOP 예측 구조
Fig. 9. Prediction structure for "Akko&Kayo" (a) Prediction structure for reference S/W (b) Flexible GGOP prediction structure

측 구조에서 참조 영상 사이에 존재하는 B-픽처의 개수가 1개 일 때 같은 시간의 시점 축의 각 프레임을 ...PBIBP...로 부호화하는 것을 보여 준다. 기준 시점은 전역 변이 합의 중앙에 위치하는 4번째 카메라로 획득한 영상이 된 것을 알 수 있다. 그림 8은 16 시점의 "Rena" 영상에 적용된 부호화 구조이다. 가변형 다시점GOP 예측구조의 경우에 기준 시점이 카메라 배열의 중앙 근처인 7번째 카메라 영상인 것을 알 수 있다. "Rena" 영상은 카메라 간의 거리가 좁기 때문에 그림 8과 같이 시점 축을 ...PBBIBBP...와 같이 I-픽처 (또는 P-픽처)와 P-픽처 사이에 B-픽처의 개수를 2개로 결정하여 부호화 한다^[17]. 그림 9는 카메라 구조가 2차원 배열형태인 "Akko&Kayo"에 적용된 부호화 구조를 보여준다. 참조 소프트웨어는 맨 위, 가장 왼쪽을 기준 시점으로 부호화 수행하지만 제안된 가변형 다시점GOP 예측 부호화 구조에서는 2차원 배열구조를 고려하여 정중앙에 기준 시점을 위치시켜 부호화를 수행한다.

IV. 실험결과

다시점 비디오 압축에 가변형 다시점GOP 예측 구조를 적용할 경우 부호화 효율이 개선되는지 확인하기 위하여 다시점 동영상 부호화의 참조 소프트웨어인 JSVM(Joint Scalable Video Model) 3.5상에 제안 방법을 구현하였고, 다시점 동영상 부호화의 참조 소프트웨어와 독립 시점 부호화(Anchor)방식과 성능 비교 실험을 수행하였다. 실험에 사용된 영상의 특성과 카메라 파라미터는 표 1과 같다. 참조

소프트웨어와 동등한 조건에서 결과를 비교하기 위하여 각 시점의 GOP 길이는 74차 MPEG 회의에서 정한 조건과 동일하게 실험하였다^[19]. 표 2는 각 실험 영상의 시간 축에서의 분할 구조를 나타낸다. 다시점 영상이 가지고 있는 변이 정보를 활용하여 부호화의 성능을 높이는 가변형 다시점 GOP의 성능을 평가하기 위해서 기존의 다시점 동영상 부호화를 위한 참조 소프트웨어에서 사용한 움직임 예측 구조를 동일하게 사용하였다. 그림 10은 GOP 길이가 12와 15 일 때 사용되는 움직임 축에서의 기본적인 부호화 구조를 보여준다. 실험은 하나의 영상에 서로 다른 3개의 비트율에 대해 수행하였다. 비트율은 양자화 파라미터(QP)값으로 조

표 1. 실험 영상의 영상 특성 및 카메라 파라미터

Table 1. Image properties and camera parameters of the experimental images

실험 영상	영상 특성		카메라 파라미터		
	영상 크기	프레임율	시점수	카메라 간격	카메라 배열
Race1	640 x 480	30fps	8	20cm	1D parallel
Ballroom	640 x 480	25fps	8	20cm	1D parallel
Rena	640 x 480	30fps	16	5cm	1D parallel
Akko&Kayo	640 x 480	30fps	15	5cm(수평), 20cm(수직)	2D array

표 2. 각 실험 영상의 시간축 분할 구조

Table 2. Temporal partitioning of test data sets

실험 영상	시간축 분할 구조
Race1	532 frames = 35 * GOP_15 + GOP_6
Ballroom	250 frames = 20 * GOP_12 + GOP_9
Rena	300 frames = 19 * GOP_15 + GOP_14
Akko&Kayo	300 frames = 19 * GOP_15 + GOP_14

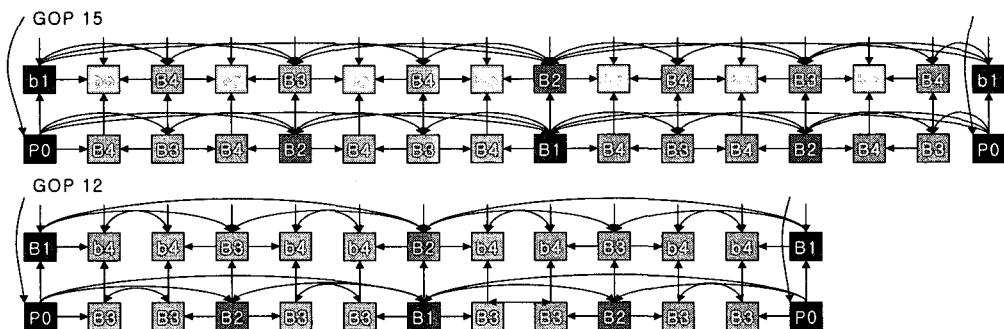


그림 10. GOP길이가 12와 15일 때 사용되는 기본적인 부호화 구조
Fig. 10. Basic structures for coding with GOP_15 and GOP_12

표3. 비트율 제어를 위한 기본 양자화 파라미터 값
Table 3. Quantization parameter settings

실험영상	BasisQP values for each bitrate		
	저비트율	중간 비트율	고비트율
Race1	29	27	25
Ballroom	34	31	29
Rena	33	28	23
Akko&Kayo	37	31	25

절하고 각 영상에 사용한 양자화 파라미터는 표 3과 같다^[20].

제안 방법의 성능을 검증하기 위해 다시점 동영상 부호화의 참조 소프트웨어와 독립 시점 부호화의 결과와 비교하였다. 독립 시점 부호화(Anchor)는 다시점 영상의 영상 간 중복성을 고려하지 않고 각 시점을 독립적으로 2차원 영상 압축 표준인 H.264/AVC로 부호화한 방법이다. 실험 결과는 비트율 대비 영상의 객관적 화질 (Bitrates vs. PSNR)로 표시하였다. 이를 위하여 각 실험 영상에 대한 저비트율과 중간 비트율 그리고 고비트율의 결과를 하나의 그래프로 도시하였다.

참조 소프트웨어는 실험 영상에 관계없이 시공간 예측 구조에서 기준 시점은 가장 왼쪽의 영상으로 고정하고 I-픽처(또는 P-픽처)와 P-픽처 사이의 B-픽처의 개수는 1개를 사용한다. 그림11~그림14는 "Race1", "Ballroom", "AkkoKayo", "Rena" 영상에 대해 제안 알고리즘으로 얻은 PSNR(Peak to Signal to Noise Ratio)값을 나타낸 것이다. 그림 11은 "Race1" 영상의 결과 그래프로 참조 소프트웨어와 제안 알고리즘이 같은 PSNR의 값일 때 가변형 다시점GOP 예측 구조를 적용한 것이 최대 7.1%의 비트를 적게 사용한 것을 알 수 있었다. 그리고 "Race1"에서 독립 시점 부호화의 결과와 비교하여 2~3.5dB의 PSNR 이득을 얻음으로써 각 시점을 독립적으로 부호화 하는 것보다 시점간의 상관도를 이용하는 것이 압축 효율을 증대시키는 것을 알 수 있었다. 그림 12는 "Ballroom" 영상의 결과 그래프로 제안한 부호화 예측 구조가 기존의 다시점 동영상의 참조 소프트웨어 보다 0.3dB 정도 우수한 성능을 보였고 독립 시점 부호화의 결과보다

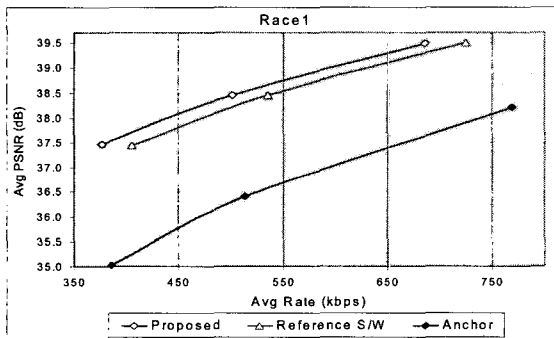


그림 11. "Race1"영상의 결과
Fig. 11. Results for "Race1"

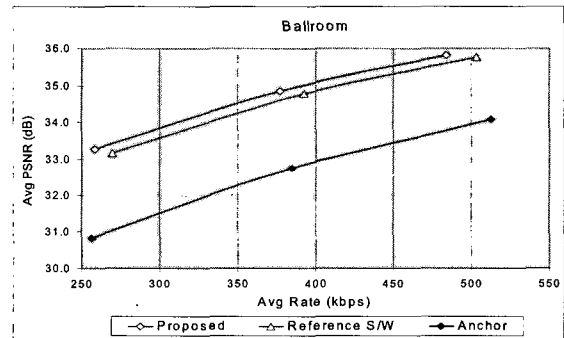


그림 12. "Ballroom" 영상의 결과
Fig. 12. Results for "Ballroom"

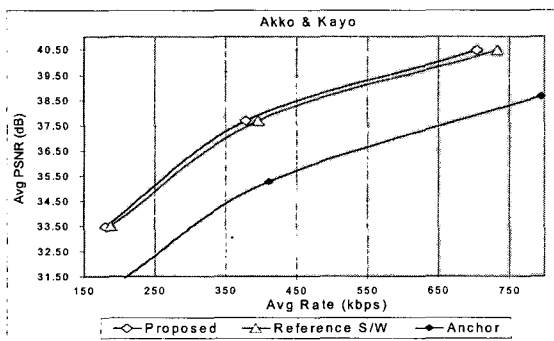


그림 13. "AkkoKayo"영상의 결과
Fig. 13. Results for "AkkoKayo"

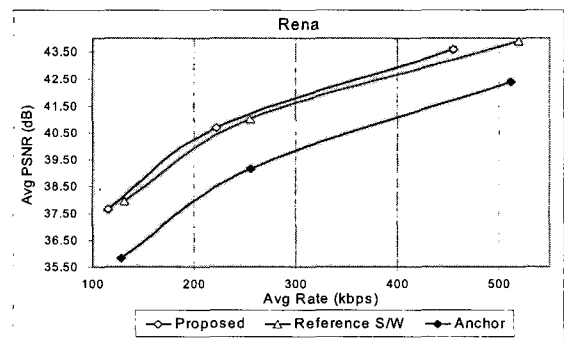


그림 14. "Rena"영상의 결과
Fig. 14. Results for "Rena"

2~2.5dB 높은 결과를 나타냈다. "Akko&Kayo" 영상의 부호화 결과도 참조 소프트웨어보다 압축 효율이 더 좋음을 그림 13을 통해 알 수 있었다. 독립 시점 부호화와의 결과 비교에서도 약 2.5~3dB의 이득을 얻어 가변형 다시점GOP 예측 구조의 부호화 방법이 다시점 영상의 특성을 활용하여 효율적인 압축을 했음을 알 수 있었다. "Race1", "Ballroom"과 "Akko&Kayo"와 같이 카메라 간의 간격이 비교적 넓고 움직임이 큰 영상에 대해서 제안 알고리즘이 참조 소프트웨어와 독립 시점 부호화 방식 보다 모두 우수한 성능을 보였다.

카메라간의 간격이 좁고 다른 영상에 비해서 상대적으로 움직임이 적은 영상의 결과를 알아보기 위해 "Rena"영상을 사용하였다. 그림 14에서 "Rena" 영상의 결과는 모든 비트율에 대해서 참조 소프트웨어와 독립 시점 부호화의 결과 보다 압축 효율이 더 좋음을 알 수 있었다. 특히 독립 시점 부호화의 결과 보다 2dB 정도의 이득을 얻음을 확인 할 수 있었다. "Race1"과 "AkkoKayo"영상의 실험 결과와 마찬가지로 "Rena"영상과 같이 카메라 간의 간격이 좁고 움직임이 상대적으로 적은 영상에서도 기준 시점을 전역 변이합의 중앙에 위치하도록 하고 카메라 간의 간격을 통해 B-픽처의 개수를 조절하는 가변형 다시점GOP 예측구조를 적용함으로써 더 우수한 성능을 보임을 확인 할 수 있었다.

V. 결 론

JVT에서 표준화가 진행중에 있는 다시점 동영상 부호화의 참조 소프트웨어는 고정된 시공간 예측 구조를 사용하고 있다. 다시점 비디오 압축에서 어떠한 영상을 기준 시점으로 정하는 것에 따라 부호화 결과는 영향을 받는다. 이러한 다시점 영상의 특성을 고려하여 본 논문에서는 여러 대의 카메라로 획득한 영상의 시점축 특성과 카메라간의 간격에 따라 가변형 다시점 GOP 예측 구조를 결정하여 부호화하는 방법을 제안하였다. 실험 결과에서 제안 방법이 기존의 참조 소프트웨어 보다 우수한 성능을 보여줌을 확인하였다. 기존 참조 소프트웨어와 비교하여 가변형 다시점GOP 예측 구조의 부호화 방법이 비슷한 PSNR일 때 최대 7.1%까지 비트량 감소를 보였고 다시점의 동영상에서의 시점간의 정보를 사용하지 않는 독립 시점 부호

화(Anchor)의 결과 보다는 2~3.5dB 정도의 높은 PSNR 이득을 얻었다. 그러나 본 논문에서 사용한 실험 영상은 MPEG 다시점 동영상 부호화(MVC)에서부터 사용한 것으로 카메라 배열이 다양하지 않고 제한된 구조로 된 것이었다. 카메라 배열과 구조가 더 다양한 영상으로 실험을 한다면 다시점 영상의 특성을 활용하여 부호화하는 가변형 다시점GOP 예측 구조방법을 통해 더 많은 이득을 얻을 수 있을 것이다.

향후 연구 과제로는 가변형 다시점GOP 예측 구조의 성능을 높이기 위한 다른 요소에 대한 연구가 필요할 것이다.

참 고 문 헌

- [1] R. Franich, R. Lagendijk and R. Horst, "Reference model for hardware demonstrator implementation," RACE DISTIMA deliverable 45/TUD/IT/DS/B/003/b1, Oct. 1992.
- [2] RACE 2045-DISTIMA, <http://www.tnt.uni-hannover.de/plain/project/eu/distima/>
- [3] A. Rauol, "State of the art of autostereoscopic displays," RACE DISTIMA deliverable 45/THO/WP4.2/DS/R/57/01. Dec. 1995.
- [4] S. Malassiotis and M. G. Strintzis, "Coding of video-conference stereo image sequences using 3D models," Signal Processing: Image Communications, vol. 9, no. 1, pp 125-135, Jan. 1997.
- [5] Berlin, PANIRAM Final Demonstrations, AC092/SIE/FinalDemo/DS/P/032/b1, Oct. 1998.
- [6] <http://www.virtue.eu.com/>
- [7] N. Hur and C. Ahn, "Experimental service of 3DTV broadcasting relay in Korea," Proc. SPIE 4864, pp 1-13, 2002.
- [8] N. Hur, G. Lee, W. You, J. Lee and C. Ahn, "An HDTV- Compatible 3DTV Broadcasting System," ETRI J. 2003.
- [9] R. Borner, "Autostereoscopic direct-view displays and rear projection for short viewing distances by lenticular method," Proc. of the First International Symposium on Three Dimensional Image Communication Technologies, Tokyo, pp 1-14, Dec. 1993.
- [10] C. V. Berkel and D. W. Parker, "Multiview 3D-LCD," Proc. SPIE 2653, pp 32-39, 1996.
- [11] ISO/IEC JTC1/SC29/WG11, "Description of Core Experiments in MVC", W7798, Bangkok, Thailand, Jan. 2006.
- [12] E. Izquierdo, Stereo Matching for Enhanced Telepresence in Three-Dimensional Videocommunications, IEEE Trans. Circuits and Systems for Video Technol., vol. 7, no. 4, pp. 629-643, Aug. 1997.
- [13] D. Tzovaras, N. Grammalidis and M. G. Strintzis, Object-Based Coding of Stereo Image Sequences Using Joint 3-D Motion/Disparity Compensation, IEEE Trans. Circuits and Systems for Video Technol., vol. 7, no. 4, pp. 312-327, Apr. 1997.
- [14] K.H. Sohn, J.R. Ryou and J. Lim, "Efficient stereoscopic video coding using joint disparity-motion estimation," Circuits, Systems and Signal Processing, vol. 23, no. 1, pp. 57-76, Jan. 2003.
- [15] G. Egnal and R.P. Wildes, "Detecting binocular half-occlusions: empirical comparisons of five approaches," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 8, pp. 1127-1133, Jun. 2000.

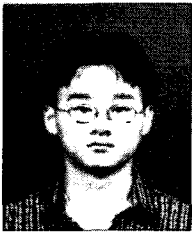
- [16] G.A. Triantafyllidis, D. Tzovaras and M.G. Strintzis, "Detection of occlusion and visible background and foreground areas in stereo image pairs," Proc. IEEE 9th International Conference on Electronics, Circuits and Systems, vol. 3, pp. 1019-1022, Sep. 2002.
- [17] ISO/IEC JTC1/SC29/WG11, "Results on CE1 for multi-view video coding", M13544, Klagenfurt, Austria, Jul. 2006.
- [18] ISO/IEC JTC1/SC29/WG11, "Preliminary Call for Proposals on Multi-View Video Coding", W7094, Busan, South Korea, Apr. 2005.
- [19] ISO/IEC JTC1/SC29/WG11, "Updated Call for Proposal on Multi-view Video Coding", N7567, Nice, France, Oct. 2005.
- [20] ISO/IEC JTC1/SC29/WG11, "Results of Core Experiment 1-D on Multiview Video Coding", M13228, Montreux, Swiss, Apr. 2006.

저 자 소 개



윤재원

- 2005년 2월 : 연세대학교 전기전자공학부 학사
- 2005년 9월 ~ 현재 : 연세대학교 전기전자공학과 석사과정
- 주관심분야 : 3차원 동영상 부호화/복호화, MPEG-2/4, H.264



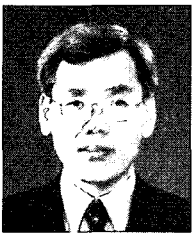
서정동

- 2004년 2월 : 연세대학교 전기전자공학부 학사
- 2006년 2월 : 연세대학교 전기전자공학과 석사
- 2006년 3월 ~ 현재 : 연세대학교 전기전자공학과 박사과정
- 주관심분야 : 3차원 동영상 부호화/복호화, MPEG-2/4, H.264



김용태

- 2000년 2월 : 연세대학교 기계전자공학부 학사
- 2002년 2월 : 연세대학교 전기전자공학과 석사
- 2002년 9월 ~ 현재 : 연세대학교 전기전자공학과 박사과정
- 주관심분야 : 3차원 동영상 부호화/복호화, 움직임 및 변이 추정, MPEG-2/4, H.26x



박창섭

- 1980년 : 한양대학교 전자공학과 학사
- 1982년 : 한양대학교 산업대학원 석사
- 2001년 : 한국과학기술원 정보 및 통신공학과 박사
- 1981년 ~ 현재 : 한국방송 방송기술연구팀
- 주관심분야 : 영상신호압축, 컴퓨터 비전, 입체TV



손광훈

- 1983년 : 연세대학교 전자공학과 공학사
- 1985년 : University of Minnesota, MSSE
- 1991년 : North Carolina State University, Ph.D
- 1993년 : 한국전자통신연구원 선임연구원
- 1994년 : Georgetown University, Post-doctor fellow
- 1995년 ~ 현재 : 연세대학교 전기전자공학부 교수
- 주관심분야 : 3차원 영상처리, 영상 통신