

일반논문-08-13-3-05

다시점 영상 부호화에서 전역 변이 벡터를 이용한 고속 모드 결정

한동훈^{a)}, 조숙희^{b)}, 허남호^{b)}, 이영렬^{a)‡}

Fast Mode Decision using Global Disparity Vector for Multi-view Video Coding

Dong-Hoon Han^{a)}, Sukhee Cho^{b)}, Namho Hur^{b)} and Yung-Lyul Lee^{a)‡}

요 약

H.264/AVC기반의 다시점 영상 부호화 기술은 시점간 상관성을 이용한 새로운 예측 방법을 이용하여 여러 대의 카메라로부터 촬영된 영상을 효율적으로 부호화하는 기술이다. 그러나 다시점 부호화 기술은 시점의 증가와 시점간 예측의 사용으로 인해 부호화 시간이 크게 증가한다. 본 논문은 부호화 시간을 줄이기 위해 다시점 영상에서 시점간에 대응하는 매크로블록 (Macroblock) 기반 영역 분할 정보와 시점간의 전역 변이 벡터를 이용한 고속 모드 결정 방법을 제안한다. 제안하는 방법은 다시점 영상 부호화 표준의 참조 소프트웨어인 joint multi-view video model (JMVM) 4.0에 비해 약 0.04dB의 화질 열화를 보이지만 전체 부호화 시간을 평균적으로 40% 단축할 수 있었다.

Abstract

Multi-view video coding (MVC) based on H.264/AVC encodes multiple views efficiently by using a prediction scheme that exploits inter-view correlation among multiple views. However, with the increase of the number of views and use of inter-view prediction among views, total encoding time will be increased in multiview video coding. In this paper, we propose a fast mode decision using both MB(Macroblock)-based region segmentation information corresponding to each view in multiple views and global disparity vector among views in order to reduce encoding time. The proposed method achieves on average 40% reduction of total encoding time with the objective video quality degradation of about 0.04 dB peak signal-to-noise ratio (PSNR) by using joint multi-view video model (JMVM) 4.0 that is the reference software of the multiview video coding standard.

Keywords : Fast Mode Decision, Multi-view Video Coding, H.264/AVC, Global Disparity Vector

a) 세종대학교 컴퓨터공학과 DMS 연구실
DMS Lab., Dept. of Computer Engineering, Sejong University

b) 한국전자통신연구원 전파방송연구단
Radio & Broadcasting Research Division, ETRI

‡ 교신저자: 이영렬(yllee@sejong.ac.kr)

※ 본 연구는 정보통신부 및 정보통신연구진흥원의 IT신성장동력핵심기술개발사업의 일환으로 수행하였음. [2007-S004-01, 무안경 개인형 3D 방송기술 개발]

1. 서론

현재 국제 표준화가 진행 중인 다시점 영상 부호화 (Multi-view video coding) 기술^[1]은 기존의 단일시점 영상 부호화 (Single-view video coding)와 달리 같은 시간을 기준으로 일렬 혹은 행렬로 배열되어있는 여러 대의 카메라로 촬영된 영상들의 효율적인 부호화를 위한 새로운 형태

의 영상 부호화 표준이다. 다시점 영상 부호화 기술은 사용자가 원하는 시점의 장면을 시청 할 수 있는 자유 시점 TV (Free viewpoint TV), 사람의 왼쪽과 오른쪽 눈에 해당하는 영상을 두 눈에 각각 보여주어 입체감을 느낄 수 있도록 하는 3차원 TV 등 차세대 방송시스템을 주도할 여러 가지 실감미디어를 위해 사용될 수 있다.

이러한 다시점 영상을 이용한 다양한 기술들을 구현하는데 있어서 가장 큰 문제점은 다수의 카메라로부터 촬영된 방대한 양의 영상 데이터이다. 일반적으로, 다시점 영상은 동일한 시간의 단일시점 영상에 비해 영상 데이터의 크기가 촬영에 사용된 카메라의 수, 영상의 해상도에 비례하여 증가한다. 이에 따라 다시점 영상에 대한 효율적인 부호화 방법이 요구되어, 다시점 영상을 처리하기 위해 가장 최근에 표준화가 완성된 단일시점 영상 부호화 방식인 H.264/AVC^[2]를 기반으로 하는 다시점 영상 부호화 기술을 국제 영상 부호화 표준 단체인 ITU-T의 VCEG (Video Coding Experts Group)과 ISO/IEC의 MPEG (Moving Picture Experts Group)에서 공동으로 국제 표준화 작업을 진행하고 있다^[3].

다시점 영상 부호화 기술의 가장 큰 특징은 서로 다른 카메라로부터 얻은 영상 간의 상관성 (Correlation)을 이용하는 것이다. 다시점 영상 부호화는 단일시점 영상 부호화 예측 구조인 화면내의 상관성을 이용한 화면내 예측 구조 (Intra prediction structure)와 화면간의 상관성을 이용한 화면간 예측 구조 (Inter prediction structure)에 추가적으로 시점간의 상관성을 이용한 시점간 예측 구조 (Inter-view prediction structure)^[4]를 사용하여 높은 압축 효율을 얻을 수 있는 기본적인 부호화/복호화 구조를 갖추었으며, H.264/AVC를 이용해 각 시점들을 독립적으로 부호화하는 동시 부호화 (Simulcast coding)방법보다 향상된 부호화 효율을 얻기 위해 추가적으로 새로운 형태의 부호화/복호화 방법들을 채택하고 있다. 하지만 이러한 다시점 영상 부호화 방식은 단일시점 영상 부호화 방식에 비해 위에서 언급된 것과 같이 다수의 카메라로 촬영된 방대한 양의 영상데이터를 부호화하는데 필요한 시간 외에 추가적으로 시점간 예측과 새로운 형태의 부호화/복호화 방법들로 인하여 부호화 시간이 더욱 증가한다.

H.264/AVC는 영상을 매크로블록 (Macroblock, MB) 단위로 나누어 부호화를 진행하는데 이때 하나의 매크로블록을 효율적으로 부호화하기 위한 윌-왜곡 최적화는 최적의 매크로블록 모드를 찾기 위해 모든 매크로블록 모드에 대하여 화면내 예측 혹은 움직임 예측/보상 (Motion Estimation/Compensation)을 이용한 화면간 예측, 정수 역현 변환 (Integer cosine transform), 양자화 (quantization), 엔트로피 부호화 (entropy coding)와 이에 대한 모든 역 과정을 진행한다. 이러한 윌-왜곡 최적화를 이용하게 되면 효율적인 부호화가 가능해지지만 부호화에 필요한 시간은 크게 증가한다. H.264/AVC를 기반으로 하는 다시점 영상 부호화도 마찬가지로 이러한 윌-왜곡 최적화를 사용할 수 있는데 시점간 예측이 사용될 경우에 다른 시점의 참조 영상 (Reference frame)에 대한 윌-왜곡 최적화가 추가적으로 진행되므로 다시점 영상 부호화에 필요한 시간은 더욱 증가하게 된다.

본 논문에서는 다시점 영상 부호화에서 시점간 예측이 사용된다는 점에 착안하여 이를 이용하여 다시점 영상을 부호화하는데 필요한 시간을 감소시키기 위한 윌-왜곡 최적화에서의 새로운 고속 모드 결정 방법을 제안한다.

II 장에서는 고속 모드 결정을 위해 사용되는 다시점 영상 부호화 기술의 특징을 소개하고, III 장에서는 다시점 영상 부호화에서 사용되는 전역 변이 벡터 (Global disparity vector)를 이용한 고속 모드 결정 방법을 제안한다. IV 장에서는 고속 모드 결정방법이 적용되지 않은 방법과의 비교 실험 결과를 제시하고 마지막으로 V 장에서는 본 논문의 결론을 끝으로 논문을 마친다.

II. 다시점 영상 부호화 기술의 특징

1. 계층적 B 화면과 예측구조

H.264/AVC는 이전의 MPEG-1/2/4, H.261/262/263 등의 단일시점 영상 부호화 표준과 비교하여 부호화 구조에 있어서 높은 유연성을 가진다. 이전의 표준에서는 부호화 순서 (coding order)와 화면 출력 순서 (display order)가 서로

관련되어 있었기 때문에 부호화 구조에 많은 제약이 있었으나 H.264/AVC에서는 이들을 완전히 분리하였으며, 이로 인해 다양한 구조로 부호화를 수행할 수 있게 되었다.

이러한 구조적 유연성을 이용하는 대표적인 예로써 계층적 B 화면 구조 (Hierarchical B picture structure)^[5]가 스케일러블 영상 부호화 기술^[6]을 통해 소개되었으며, 기존의 IBBP 구조에 비해 높은 부호화 효율을 얻을 수 있었다. 그림 1은 16장의 화면과 5개의 시간적 계층 (Temporal level)으로 구성된 GOP (Group of picture)를 가지는 계층적 B 화면 부호화 구조의 예를 나타낸다.

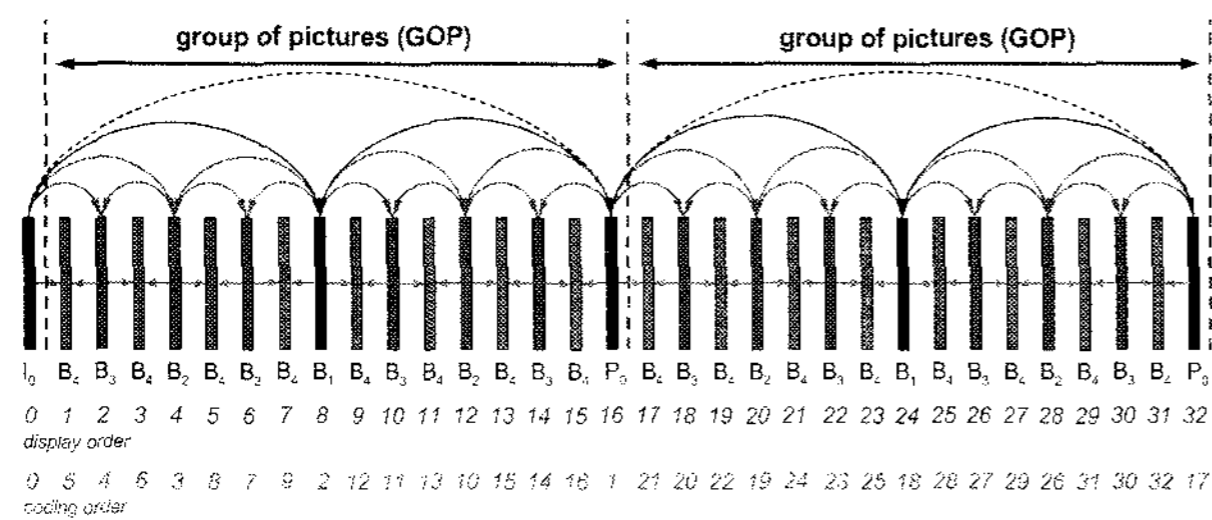


그림 1. 계층적 B 화면 부호화 구조의 예
Fig. 1 Example of hierarchical B picture structure

다시점 영상 부호화 기술은 다시점 영상을 효율적으로 부호화하기 위해서 이러한 계층적 B 화면 구조가 실험조건으로 권장되어 사용 중에 있다. 기본적인 부호화 구조^[7]는 그림 2와 같다.

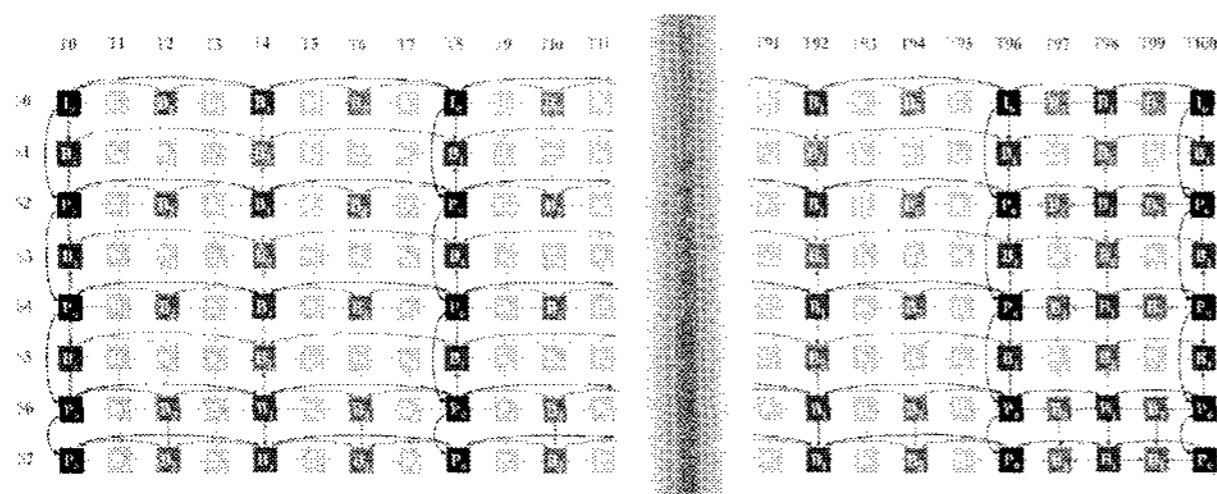


그림 2. 계층적 B 화면을 이용한 시공간적 예측구조
Fig. 2. Inter-view temporal prediction structure using hierarchical B pictures

그림 2에서 S_n은 n번째 시점을 의미하고 T_n은 n번째 시간을 의미하며 화살표는 각 화면/시점간의 참조를 의미한다.

다. 시점 방향으로는 촬영 시 사용된 구조에 맞는 다양한 구조를 이용하고 시간 방향으로는 계층적 B 화면 구조를 이용한다. 각 실험 영상들은 카메라 배치에 따라 시간 방향으로 다른 형태의 GOP 구조를 갖는다.

2. 변이 벡터

다시점 영상 부호화에서 시점간 예측이 추가됨에 따라 참조 시점에 대한 변이 예측/보상 (Disparity Estimation/Compensation)이 사용된다. 이로 인해 동시 부호화에 비해 부호화 복잡도는 크게 증가하나, 시점간 상관성을 이용하여 더욱 정확한 예측이 가능하기 때문에 기존의 화면내 예측과 화면간 예측만을 사용하는 방법보다 효율적으로 다시점 영상을 부호화 할 수 있다.

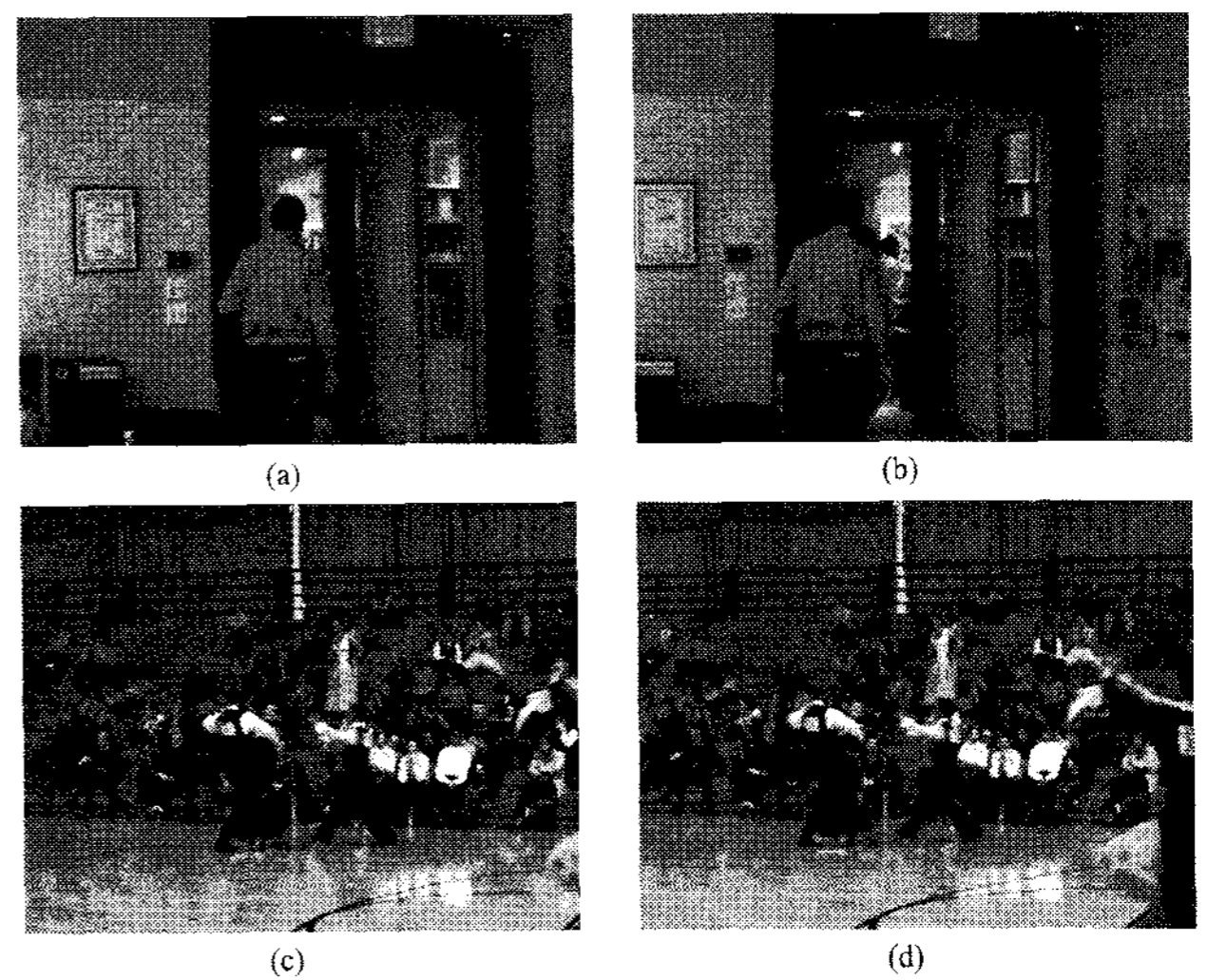


그림 3. Exit 영상과 Ballroom 영상에서 시점간 변이: (a) Exit 영상의 0번 시점에서 0번 화면, (b) Exit 영상의 2번 시점에서 0번 화면, (c) Ballroom 영상의 0번 시점에서 0번 화면, (d) Ballroom 영상의 2번 시점에서 0번 화면
Fig. 3. View disparity of Exit and Ballroom sequences, (a) 0th frame in view 0 of Exit sequence, (b) 0th frame in view 2 of Exit sequence, (c) 0th frame in view 0 of Ballroom sequence, (d) 0th frame in view 2 of Ballroom sequence

시점간 예측이 사용될 경우 각 매크로블록은 움직임 예측/보상과 유사한 방법으로 참조 시점에 대해 변이 예측/보상을 진행하여 변이 벡터를 구한다. 여기서 변이 벡터란 다

시점 영상 부호화 방식에 시점간 예측이 사용되면서 추가된 개념으로 단일시점 영상 부호화에서 사용하는 움직임 벡터와 비슷한 개념으로 사용되고 있다. 움직임 벡터는 시점내의 움직임 정보를 나타내지만 변이 벡터는 동일한 시간을 기준으로 했을 때 참조하는 다른 시점과의 변이 정보를 나타낸다. 그림 3(a)와 그림 3(b)는 다시점 영상 부호화 기술 표준화 회의에서 지정된 실험 영상인 Exit와 Ballroom의 시점간의 위치 차이를 나타낸 그림이다. 두 영상 모두 2번 시점의 카메라가 0번 시점의 카메라의 오른쪽에 위치해 있는 것을 확인 할 수 있는데 이러한 카메라 간의 위치가 변이 정보를 생성한다.

III. 전역 변이 벡터를 이용한 고속 모드 결정 방법

1. 영역 분할

H.264/AVC에서는 영상을 효율적으로 부호화하기 위한 방법으로 울-왜곡 최적화를 사용할 수 있다. 울-왜곡 최적화는 각 매크로블록을 모든 매크로블록 모드에 대해 부호화/복호화를 진행하고 그 중 최적의 매크로블록 모드를 선택하는 방법이다. H.264/AVC에서는 기존의 단일시점 영상 부호화에 비해 그림 4와 같이 움직임 예측/보상 시 가변 블록 크기(variable block size)와 각각에 해당하는 매크로블록 모드가 사용된다. H.264/AVC를 기반으로 하는 다시점 영상 부호화 기술에서도 역시 울-왜곡 최적화가 사용될 수 있으며, 움직임 예측/보상뿐만 아니라 변이 예측/보상에서도 그림 4와 같은 가변 블록 크기가 사용된다.

울-왜곡 최적화 과정에서 화면간 예측, 또는 시점간 예측이 수행될 때 일반적으로 매크로블록에 해당하는 영상의 움직임이 작거나 고정되어 있는 경우, 매크로블록의 모드는 스킵모드 (Skip mode) 혹은 다이렉트모드 (Direct mode)가 선택되거나 상대적으로 큰 크기의 블록모드가 선택되고, 반대로 영상의 움직임이 크거나 복잡할 경우에는 상대적으로 작은 크기의 블록모드가 선택된다.

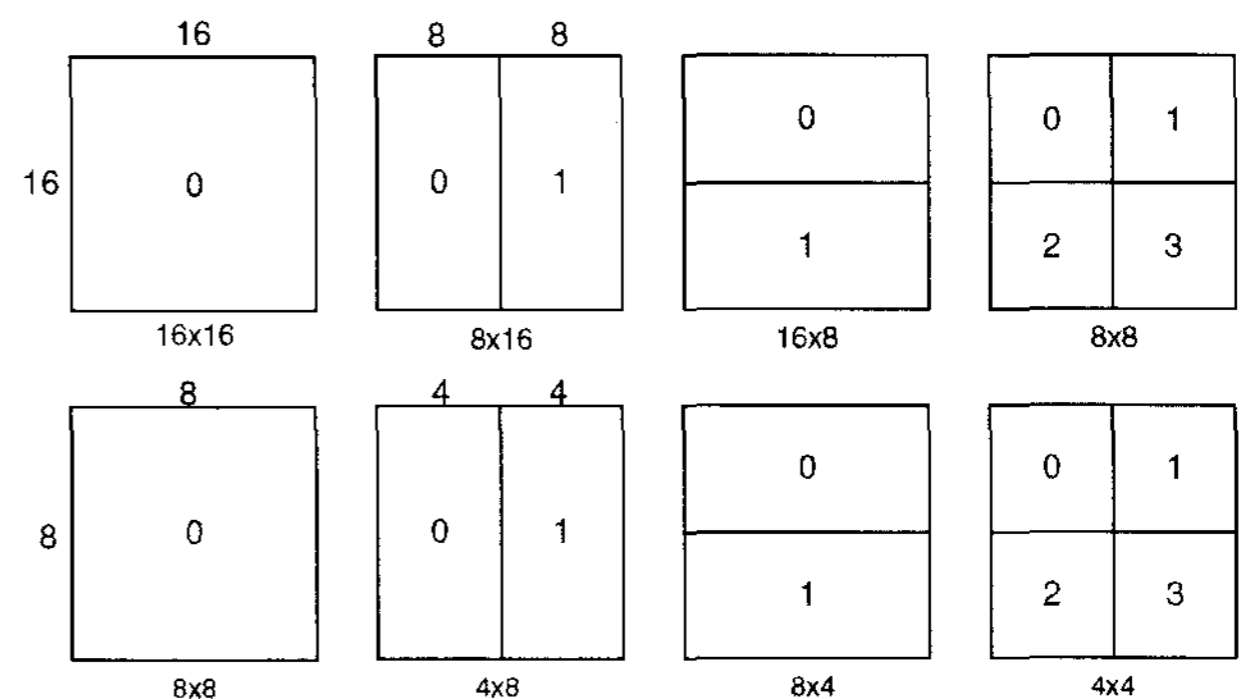


그림 4. H.264/AVC에서 가변 블록 크기
Fig. 4. Variable block size in H.264/AVC

단일시점 부호화와 마찬가지로 다시점 부호화에서도 영상의 배경영역은 움직임이 없거나 작기 때문에 일반적으로 스킵모드, 다이렉트모드, 큰 크기의 블록모드로 부호화가 되고, 배경영역을 제외한 객체영역은 움직임이 많기 때문에 상대적으로 작은 크기의 블록모드로 부호화가 될 가능성이 크다. 본 논문에서는 이러한 특징을 이용하여 매크로블록의 모드가 스킵모드, 다이렉트모드, Inter 16x16모드 중 한가지로 선택된 경우 해당 매크로블록의 영역을 배경영역이라 판단하였고, 반대로 그 외의 모드로 선택된 경우 객체영역이라고 판단하였다. 그리고 이를 이용하여 표 1과 같이 영상을 배경모드집합과 객체모드집합으로 분할하였다. 이때 표 1에서 배경모드집합에 해당하는 다이렉트모드와 Inter 16x16모드는 유도된 움직임 벡터가 1/4보다 작거나 같을 경우에 배경영역이라 판단하였다. 이 같은 방법을 통하여 참조되는 시점의 영상을 매크로블록 단위로 영역 분할 정보 (Region segmentation information)를 생성한다.

표 1. 배경모드와 객체모드의 분할

Table 1. Segmentation of background and object modes

	배경 모드	객체 모드
매크로블록 모드	P_Skip B_Skip Direct Inter 16x16	배경 모드를 제외한 모든 모드

그림 5는 다시점 영상 부호화 실험영상 중 Exit와

Ballroom영상의 영역을 배경영역과 객체영역으로 분할한 그림이다. 그림 5(b)와 그림 5(d)의 어두운 영역이 객체영역으로 판단된 영역이고 그 외의 부분이 배경영역으로 판단된 영역이다. 영역 분할 정보는 참조 시점의 영역을 판단할 수 있는 정보가 되며 2절에서 설명할 전역 변이 벡터와 함께 고속 모드 결정에 이용된다. 이들을 이용한 고속 모드 결정 방법에 대해서는 3절에서 자세하게 설명한다.

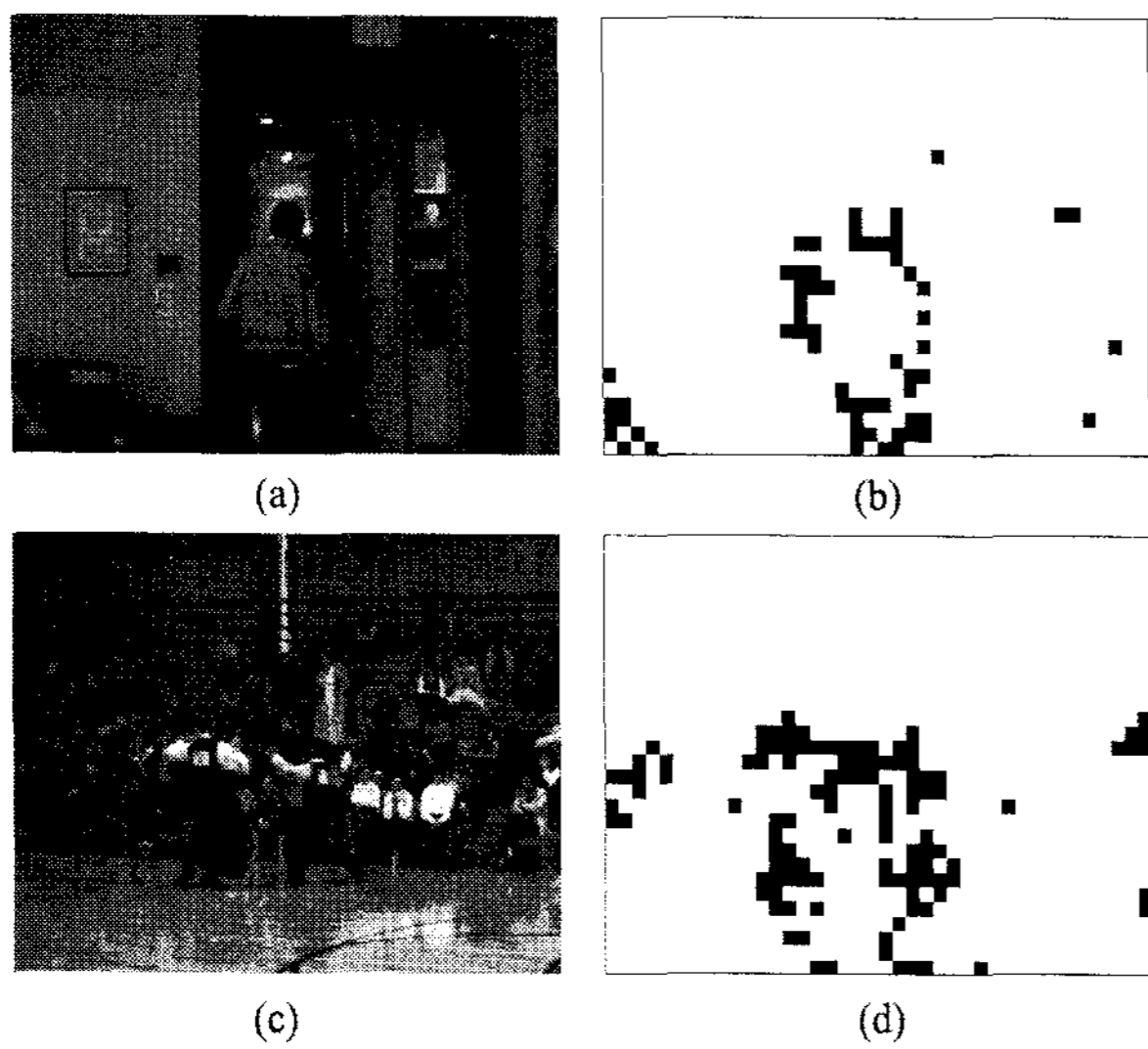


그림 5. Exit 영상과 Ballroom 영상의 영역 분할 정보, (a) Exit 영상의 0번 시점에서 4번 화면, (b) (a)의 영역 분할 정보 (c) Ballroom 영상의 0번 시점에서 4번 화면, (d) (c)의 영역 분할 정보

Fig 5. Region segmentation information of Exit and Ballroom sequences, (a) 4th frame in view 0 of Exit sequence, (b) Region segmentation information of (a), (c) 4th frame in view 0 of Ballroom sequence, (d) Region segmentation information of (c)

2. 움직임 스킵 모드와 전역 변이 벡터

전역 변이 벡터는 현재 시점 내의 한 화면과 다른 시점 내의 동일한 시간상의 한 화면간의 전체적인 변이의 차이라 할 수 있다. 전역 변이 벡터를 구하는 방법은 여러 가지가 있을 수 있으나 본 논문에서 사용하는 전역 변이 벡터는 다시점 영상 부호화에서 움직임 스킵 모드 (Motion skip mode)^[8]를 위해 계산된 전역 변이 벡터를 사용한다.

현재 다시점 영상 부호화의 표준화에 채택된 움직임 스킵

킵 모드는 현재 부호화 하는 시점 내의 한 매크로블록 위치에서 매크로블록 단위로 전역 변이 벡터에 대응되는 참조 시점의 매크로블록을 참조하여 해당 매크로블록의 모드 정보, 참조 영상 색인, 움직임 벡터를 복사하여 사용하는 방법이다. 움직임 스킵 모드가 선택될 경우 이러한 정보들을 부호화하지 않아도 되기 때문에 부호화 효율을 향상시킬 수 있다.

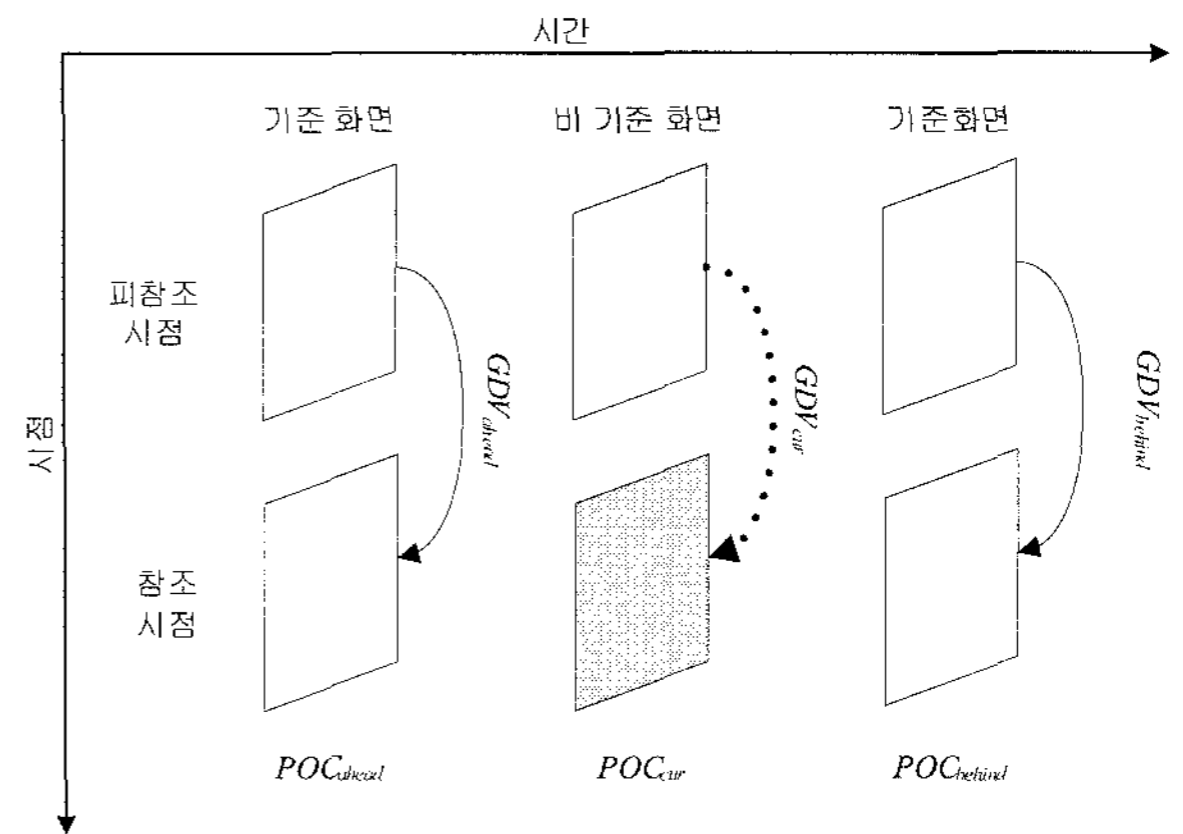


그림 6. 움직임 스킵 모드에서 전역 변이 벡터 유도 과정

Fig 6. Derivation process of global disparity vector in motion skip mode

그림 6은 움직임 스킵 모드에서의 전역 변이 벡터 유도 과정을 나타낸 그림이다. 그림 6의 피 참조 시점은 그림 2의 S0, S2와 같이 다른 시점에 의해 참조되는 시점을 의미하고 참조 시점은 그림 2의 S1, S2와 같이 피 참조 시점을 참조하여 부호화하는 시점이다. 각 시점에 대한 기준 화면은 그림 2의 T0, T8에 해당하는 화면으로 동일 시점 내에서의 화면간 예측을 사용하지 않고 비 기준 화면의 화면간 예측에서 참조화면이 되는 화면인데 이러한 기준 화면의 전역 변이 벡터는 화면 전체에 대해 움직임 벡터를 찾아내듯 SAD (sum of absolute differences)를 이용하여 매크로블록 단위로 찾아낸다.

그림 7은 Exit영상에서 기준 화면의 전역 변이 벡터를 찾는 그림이다. 이해를 쉽게 하기 위해 각 그림을 매크로블록 단위로 나누었다. 그림 7(a)는 0번 시점의 영상이고 그림 7(b)는 2번 시점의 영상, 그림 7(c)는 두 시점의 영상을 매크로블록 단위로 최적의 SAD값을 갖는 위치에 겹쳐놓은

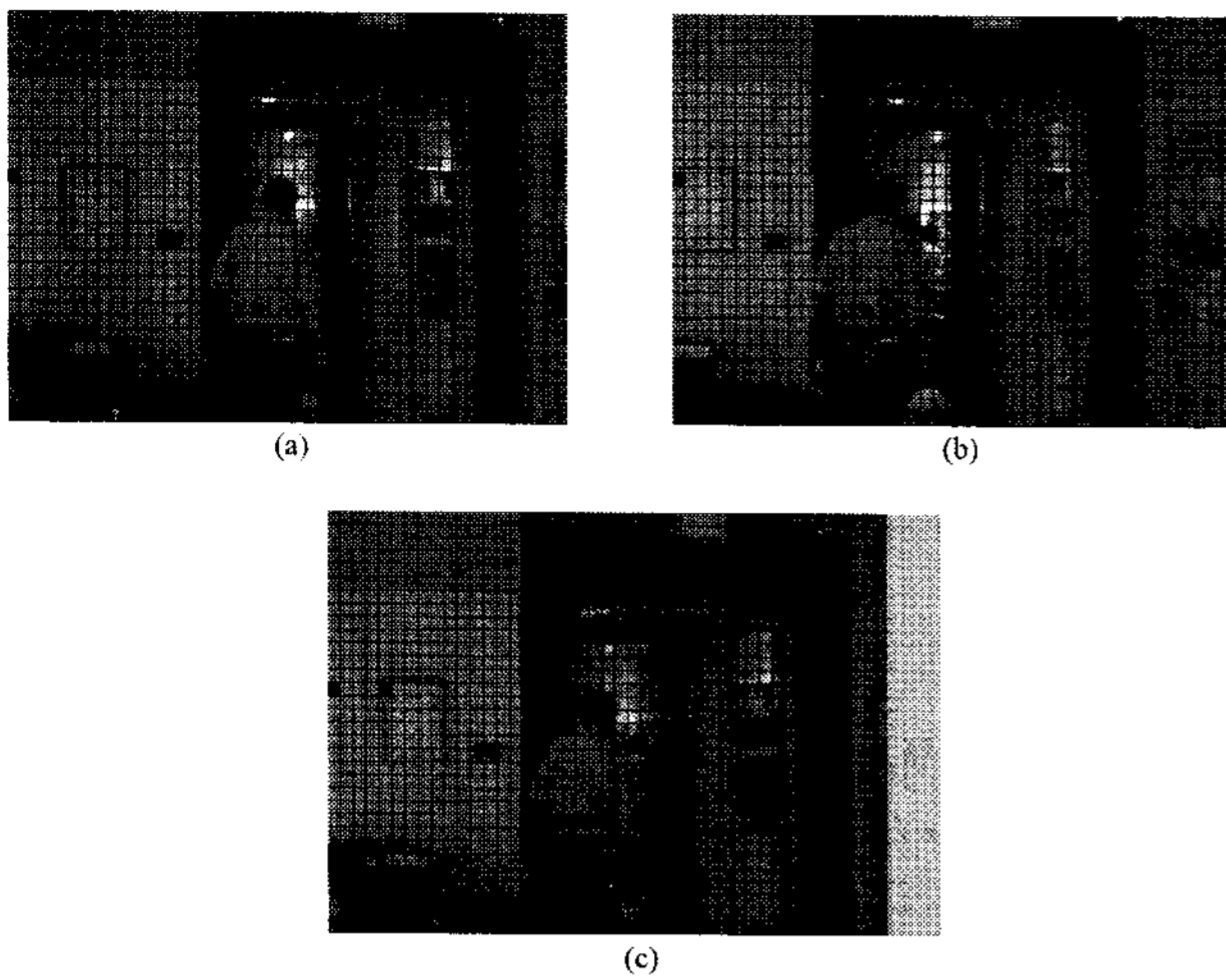


그림 7. Exit영상의 기준 화면에서 전역 변이 벡터를 찾는 예
 Fig 7. Example of finding global disparity vector at anchor picture in Exit sequence

영상이다. 그림 7(a)와 그림 7(b)의 좌측 중간에 어두운 블록의 위치는 동일한 위치인데 그림 7 (c)의 좌측을 보면 그림 7(b)의 어두운 블록이 그림 7(a)의 어두운 부분보다 오른쪽으로 네 칸 떨어진 것을 확인할 수 있고 이 정보가 곧 기준 화면에서의 전역 변이 벡터이다.

비 기준 화면의 전역 변이 벡터는 시간적으로 비 기준 화면의 앞에 위치한 기준 화면과 뒤에 위치한 기준 화면의 전역 변이 벡터를 이용하여 식 1과 같이 구해진다.

$$GDV_{cur} = GDV_{ahead} + \left[\frac{POC_{cur} - POC_{ahead}}{POC_{behind} - POC_{ahead}} \times (GDV_{behind} - GDV_{ahead}) \right] \quad (1)$$

비 기준 화면의 전역 변이 벡터는 현재 부호화 하는 화면에서 시간상으로 더 가까운 곳에 위치한 기준 시점의 전역 변이 벡터에 더 큰 영향을 받는 구조이다. 식 1에서의 GDV_{ahead} 와 GDV_{behind} 는 각각 현재 화면의 앞, 뒤에 위치한 기준 화면의 전역 변이 벡터이고 POC_{ahead} , POC_{behind} , POC_{cur} 는 시간 축을 기준으로 한 각각의 영상 출력 순서 (Picture order count) 를 나타낸다. GDV_{cur} 는 계산될 비 기준 화면의 전역 변이 벡터이다.

3. 시점간 예측을 사용하는 시점에서의 고속 모드 결정

그림 8은 제안하는 고속 모드 결정 방법의 전체적인 흐름도이다. 먼저 부호화되는 피 참조 시점의 경우 3장 1절에서 소개한 영역 분할 방법을 이용하여 화면을 배경영역과 객체영역으로 분할한 영역 분할 정보를 생성한다. 참조 시점의 경우 피 참조 시점을 참조하여 부호화를 진행하기 전에 이미 생성된 피 참조 시점의 영역 분할 정보와 움직임 스킵 모드를 위해 생성된 전역 변이 벡터를 이용하여 현재 부호화 하는 참조 시점의 영역을 예측한다.

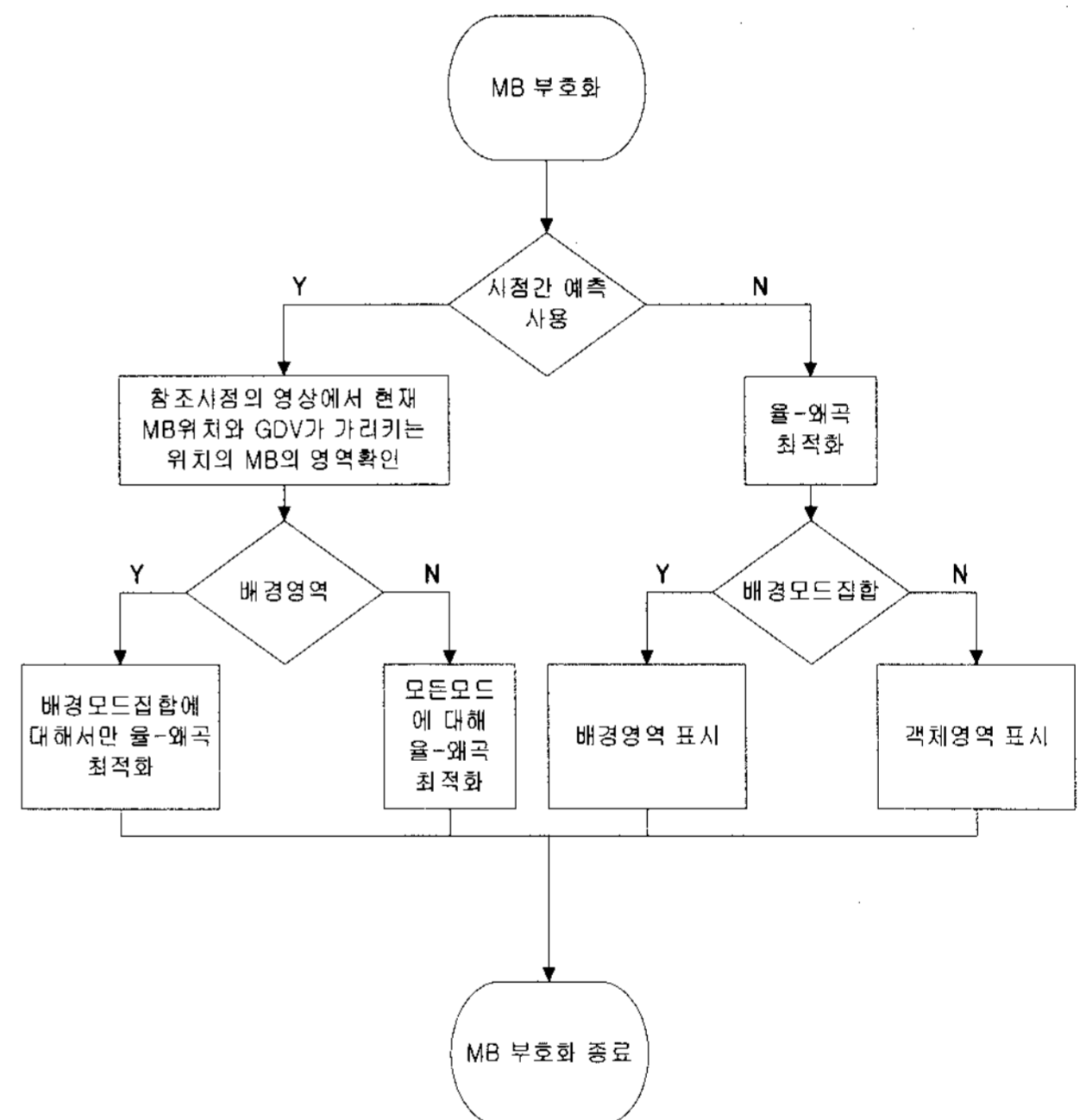


그림 8. 제안하는 방법의 전체적인 흐름도
 Fig 8. Flowchart for the proposed method

그림 9는 Exit영상에서 참조 시점의 영역을 피 참조 시점의 영역 분할 정보와 전역 변이 벡터를 이용하여 예측한 것이다. 그림 9(a)는 피 참조 시점의 영역 분할 정보로 어두운 영역이 객체영역이고 그림 9(b)는 그림 9 (a)와 3장 2절에서 찾은 전역 변이 벡터를 이용하여 예측한 참조 시점의 영역 분할 정보로 바둑판 무늬의 영역이 객체영역이다.

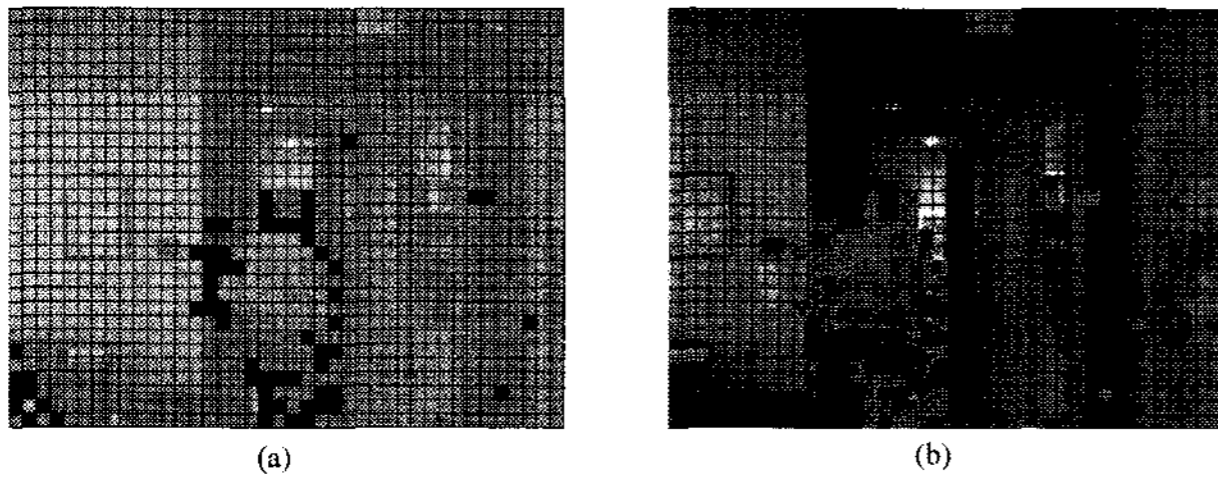


그림 9. Exit영상에서 참조 시점의 영역 예측, (a) 피 참조 시점의 영역 분할 정보, (b) 전역 변이 벡터와 (a)를 이용하여 예측한 참조 시점의 영역 분할 정보

Fig 9. Region prediction at inter-view in Exit sequence, (a) Region segmentation information of Base-view, (b) Region segmentation information of Non-base view using global disparity vector and (a)

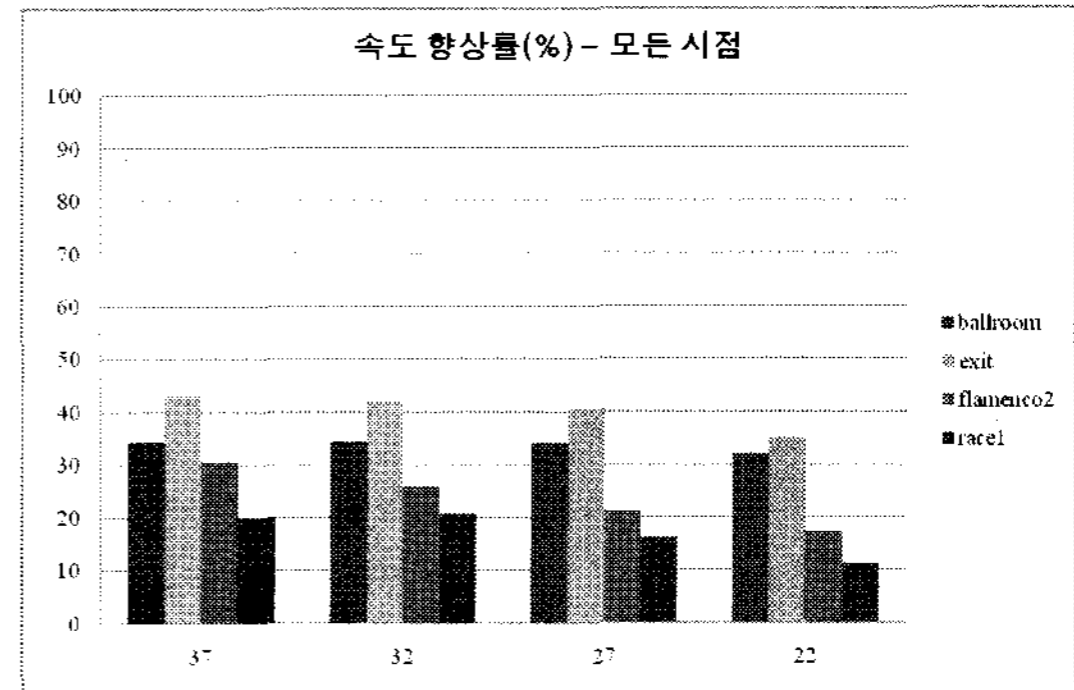
이렇게 예측한 정보를 통하여 시점간 예측과 시점내 예측을 진행하는 동안 부호화 하는 매크로블록이 배경영역인 경우는 표 1의 배경모드집합에 해당하는 네 가지 모드에 대해서만 율-왜곡 최적화를 진행하고 객체영역일 경우는 모든 모드에 대해서 율-왜곡 최적화를 진행하여 최적의 모드를 선택한다. 배경영역으로 판단된 부분에 대해서는 배경모드집합에 해당하는 네 개의 모드에 대해서만 율-왜곡 최적화를 진행하므로 율-왜곡 최적화에 필요한 시간을 크게 줄일 수 있다.

IV. 실험 결과

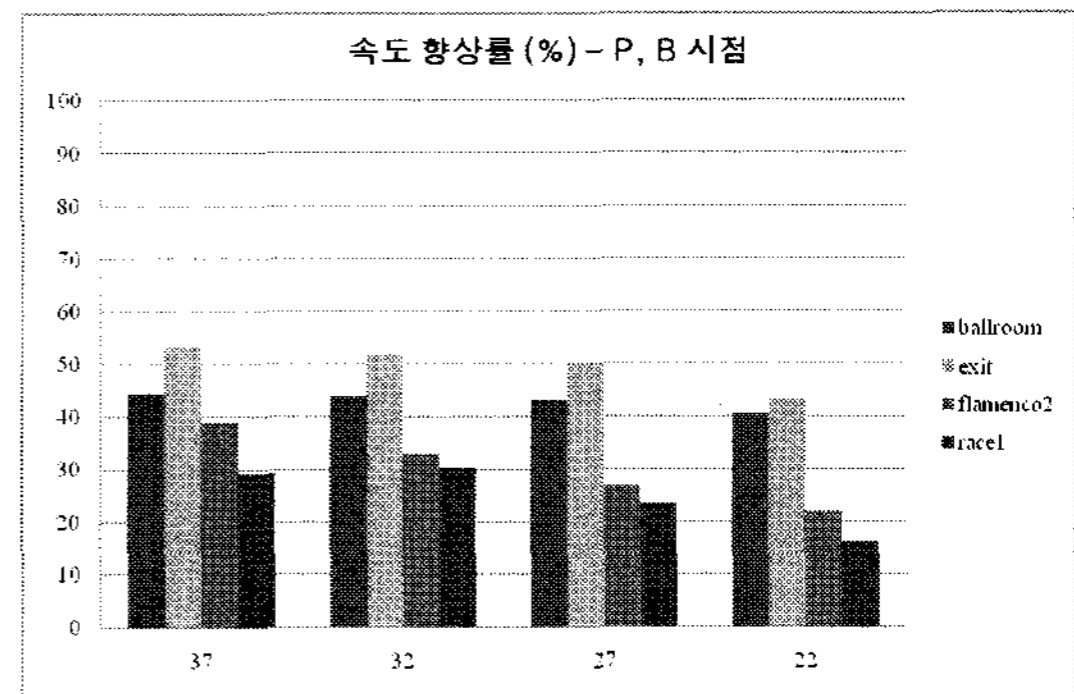
이 장에서는 다시점 영상 부호화 기술의 참조 소프트웨어인 JMVM 4.0에 본 논문에서 제안하는 고속 모드 결정 방법을 구현하여 JMVM 4.0과의 비교를 통해 제안하는 방법의 성능을 확인한다. 다시점 영상 부호화 표준화 회의에서 지정한 실험 영상 중 Ballroom, Exit, Flamenco2, Race1 이렇게 네 개의 영상으로 실험을 진행하였고 마찬가지로 다시점 영상 부호화 표준화 회의에서 지정한 실험조건^[9]으로 실험을 진행하였다. 이때 고속 모드 결정 방법의 성능 확인에 필요한 시간을 최소화하기 위해 세 개의 시점과 두 개의 GOP에 대해서만 실험을 하였다.

실험을 위해 단일시점 영상 부호화 방식에서 화면을 I, P, B로 구분한 것처럼 다시점 영상 부호화 방식에서는 시점

을 I, P, B로 나누었다. 다른 시점을 참조하지 않는 시점은 I시점, 한 시점만을 참조하는 시점은 P시점, 여러 시점을 참조하는 시점은 B시점이다. 그림 2의 S0이 I시점에 속하고 S2, S4, S6, S7이 P시점, S1, S3, S5가 B시점에 속한다.



(a)



(b)

그림 10. (a) 세 개의 시점에 대한 부호화 속도 향상률 (b) P, B시점에 대한 부호화 속도 향상률

Fig 10. (a) Speed-up ratio for three views (b) Speed-up ratio for P and B views

그림 10은 전체 부호화 속도를 비교한 것으로 세 개의 시점 모두에 대한 전체 부호화 시간이며, 여기서 가로축은 양자화 매개변수를 나타내고 세로축은 부호화 속도 향상률(%)을 나타낸다. 그림 10(a)의 그래프는 세 개의 시점 모두를 부호화 할 때의 속도를 비교한 것이고 그림 10(b)의 그래프는 시점간 예측을 사용하는 P, B시점만을 부호화 할 때의 시간을 비교한 것이다. 시점간 예측을 사용하지 않는 I시점에 대해서는 고속 모드 결정 방법이 적용이 되지 않기 때문에 고속 모드 결정 방법의 성능을 확인하기 위해서 세

개의 시점과 P, B시점에 대한 비교를 구분하여 나타내었다. 속도 비교시 속도 향상률 측정 방법은 식 2와 같다.

$$\text{Speed-up Ratio (\%)} = \frac{(TES_{JMVM} - TES_{Proposed})}{TES_{JMVM}} \times 100 \quad (2)$$

TES_{JMVM}은 JMVM 4.0을 통한 전체 부호화 시간을 나타내며, TES_{Proposed}는 제안하는 모드 결정 방법을 통한 전체 부호화 시간을 나타낸다.

제안하는 방법의 성능은 참조되는 시점의 영상에서 배경영역으로 판단되는 MB의 비율에 따라 달라지는 것을 속도 향상률 측정실험에서 확인할 수 있었다. 배경영역이 많을 경우는 제안하는 방법을 적용함으로써 얻을 수 있는 속도 향상률이 크고 적을 경우는 속도 향상률도 적다. 각 영상의 I시점에 대한 배경영역으로 판단되는 MB의 비율은 표 2와 같다.

표 2. 각 영상의 참조되는 시점에서의 배경영역으로 판단된 비율
Table 2. Ratio for background region decision of each sequence in base view

영상	양자화 매개변수	배경영역으로 판단된 MB 비율 (%)	평균
Ballroom	22	57.03	67.03
	27	66.01	
	32	70.83	
	37	74.26	
Exit	22	63.04	75.26
	27	76.27	
	32	79.71	
	37	82.00	
Flamenco2	22	53.30	62.49
	27	58.75	
	32	65.22	
	37	72.70	
Race1	22	38.81	52.22
	27	48.00	
	32	57.32	
	37	64.75	
평균			64.25

Exit 영상의 경우 모든 양자화 매개변수에 대해 배경영

역으로 판단되는 MB의 비율은 75%로 상대적으로 많은 영역이 배경영역으로 판단되고 Race1 영상의 경우는 52%로 상대적으로 적은 영역이 배경영역으로 판단된다. 이는 영상의 특성 때문인데 Race1 영상의 경우 카메라의 촬영 각도가 변함에 따라 실제 배경인 영역이 배경영역으로 판단되지 않는 영역이 많고 영상의 내용도 매우 동적이어서 배경영역으로 판단되는 부분이 적다. 반대로 Exit 영상의 경우는 카메라의 촬영 각도가 고정되어 있는데 영상의 내용도 정적이기 때문에 많은 영역이 배경영역으로 판단된다.

고속 모드 결정 방법이 적용된 P, B시점에 대한 실험에서 상대적으로 배경영역이 많은 Exit 영상의 경우, 양자화 매개변수에 따라 40%에서 50%정도의 전체 부호화 시간 속도 향상률을 보였다. 하지만 Race1 영상처럼 상대적으로 적은 배경영역을 갖는 영상은 이보다 적은 15%에서 30% 정도의 향상률을 보이며 모든 영상에 대해 평균적으로 30~40% 정도의 부호화 시간 속도 향상률을 보였다.

시점간 예측을 사용하지 않는 I시점이 하나 포함된 세 개의 시점에 대한 실험의 속도 향상률은 시점간 예측을 사용하는 시점만을 부호화 하는데 걸리는 시간에 비해 10% 정도 감소하는 것을 볼 수 있다. 이는 시점간 예측을 사용하지 않는 시점에 대해서는 고속 모드 결정방법이 적용되지 않기 때문이다. 그러나 시점간 예측을 사용하는 시점의 개수가 증가하고 이들 간의 시점간 예측구조가 복잡해질 경우에는 제안하는 방법을 통해서 보다 높은 속도 향상률을 얻을 수 있을 것으로 기대된다.

제안하는 방법과 JMVM 4.0을 비교하였을 때 약 40% 정도의 속도 향상을 얻을 수 있었지만, 부호화하는 매크로블록이 배경영역으로 선택될 시 선별된 모드에서 부호화를 수행하기 때문에 상대적으로 전체 매크로블록 모드에 대해 부호화 할 때보다 부호화 효율이 감소한다. 고속 모드 결정 방법에 따른 객관적 화질 저하는 울-왜곡 성능의 비교를 통해 나타낼 수 있는데 이를 그림 11과 표 3에서 나타내었다.

울-왜곡 성능의 비교는 객관적인 화질 측정에 사용하는 BD (Bjontegaard delta)-PSNR을^[10] 사용하여 JMVM 4.0과 제안하는 방법을 세 개의 시점과 P, B 시점에 대해 모두

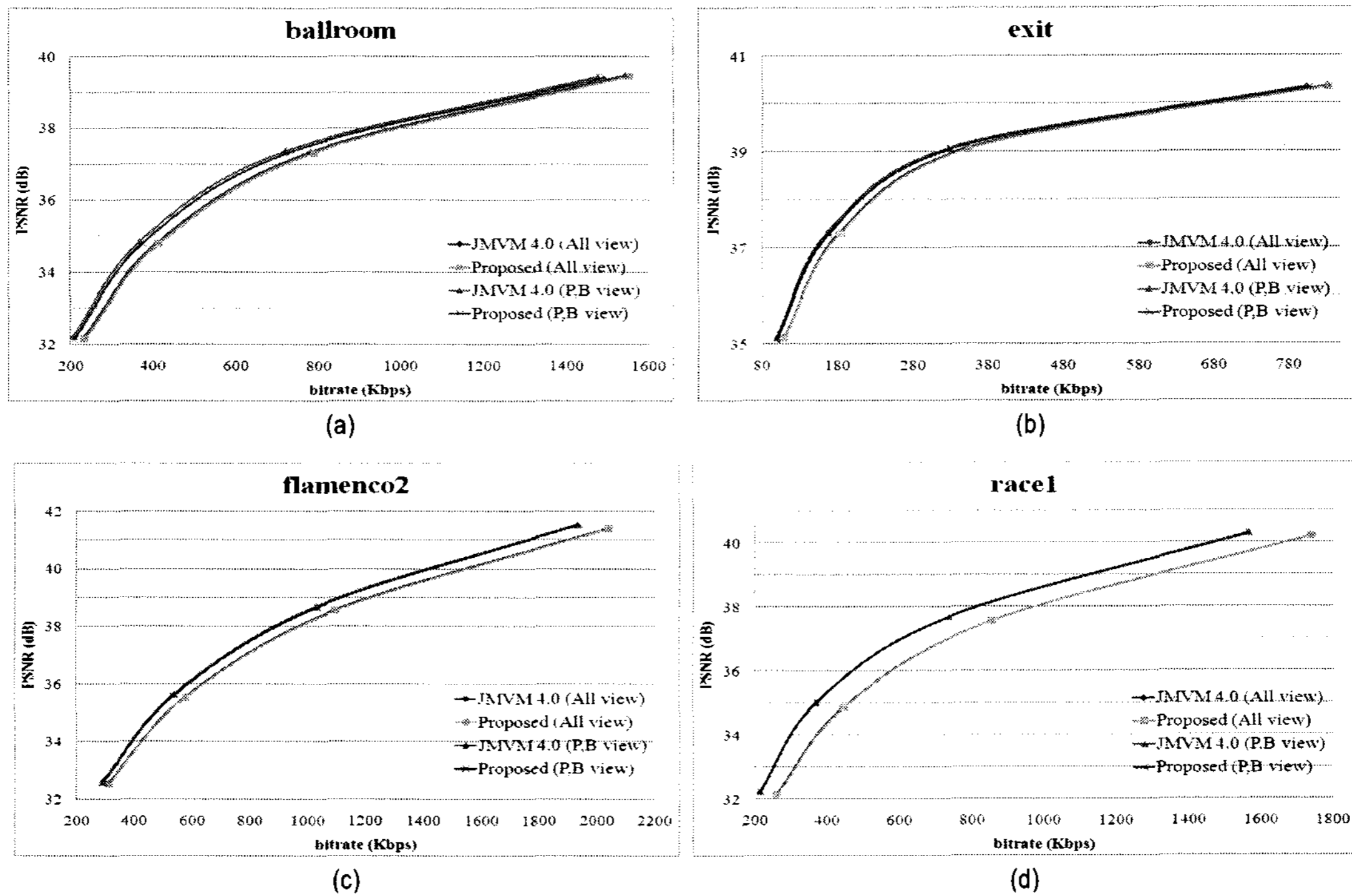


그림 11. 제안하는 방법과 JMVM 4.0과의 율-왜곡 성능 비교
 Fig 11. Rate-Distortion efficiency comparison of the proposed method vs. JMVM 4.0

표 3. 제안하는 방법과 JMVM 4.0과의 실험 결과
 Table 3. Experimental results between the proposed method and JMVM 4.0

영상	양자화 매개 변수	JMVM 4.0		Proposed (All view)			영상	양자화 매개 변수	JMVM 4.0		Proposed (All view)		
		Bitrate (kbps)	PSNR (dB)	Bitrate (kbps)	PSNR (dB)	Δ BD-PSNR (dB)			Bitrate (kbps)	PSNR (dB)	Bitrate (kbps)	PSNR (dB)	Δ BD-PSNR (dB)
Ball-room	22	1541.94	39.46	1553.60	39.44	-0.06	Ball-room	22	1476.79	39.43	1494.27	39.41	-0.09
	27	781.08	37.34	788.42	37.31			27	720.18	37.37	731.18	37.33	
	32	411.34	34.817	415.52	34.78			32	368.40	34.85	374.67	34.79	
	37	232.85	32.21	234.22	32.17			37	206.14	32.22	208.19	32.15	
Exit	22	831.60	40.35	833.46	40.34	-0.03	Exit	22	804.92	40.35	807.70	40.33	-0.05
	27	351.30	39.06	353.80	39.05			27	327.47	39.07	331.22	39.05	
	32	182.46	37.32	183.65	37.29			32	166.72	37.32	168.50	37.28	
	37	108.56	35.14	108.99	35.12			37	98.38	35.13	99.03	35.09	
Flam-enco2	22	2034.34	41.40	2037.63	41.39	-0.03	Flam-enco2	22	1930.97	41.54	1935.90	41.52	-0.04
	27	1091.81	38.56	1093.44	38.54			27	1028.76	38.69	1031.20	38.66	
	32	577.55	35.55	577.04	35.52			32	538.03	35.64	537.27	35.59	
	37	310.39	32.58	310.23	32.54			37	288.36	32.62	288.13	32.56	
Race1	22	1737.14	40.19	1745.86	40.18	-0.01	Race1	22	1564.40	40.29	1566.12	40.28	-0.02
	27	855.25	37.55	859.30	37.54			27	737.30	37.69	737.66	37.67	
	32	445.95	34.86	449.10	34.84			32	369.63	35.02	369.81	34.99	
	37	259.90	32.13	262.11	32.11			37	214.64	32.25	215.32	32.23	
평균					-0.03	평균					-0.05		

비교하였다. 4개의 실험 영상에 대해 제안하는 방법과 JMVM 4.0과 비교했을 때 평균 0.04dB 정도의 BD-PSNR 감소가 있는 것을 확인할 수 있었다.

울-왜곡 성능에서도 배경영역의 비중에 따라 울-왜곡 성능 감소율이 달라지는데 배경영역은 선별된 모드에 대해서만 부호화를 수행하게 되기 때문에 배경영역이 많은 영상의 경우 상대적으로 울-왜곡 성능이 감소한다. 각 영상의 배경영역 비율을 나타낸 표 2와 울-왜곡 성능을 비교한 그림 11과 표 3을 통해 이를 확인할 수 있다. 배경영역이 상대적으로 적게 판단된 Race1 영상의 경우 BD-PSNR의 감소는 0.01~0.02dB 이고 상대적으로 많이 판단된 Exit 영상의 경우 0.03~0.05dB 감소한다.

전체적인 실험결과를 통해 제안하는 방법은 움직임이 크고 복잡한 영상에 비해 움직임이 느리고 배경 영역의 비중이 상대적으로 높은 영상에 대해서 약간의 화질 저하만으로 더욱 빠른 부호화 속도를 보이는 것을 확인할 수 있었다.

V. 결 론

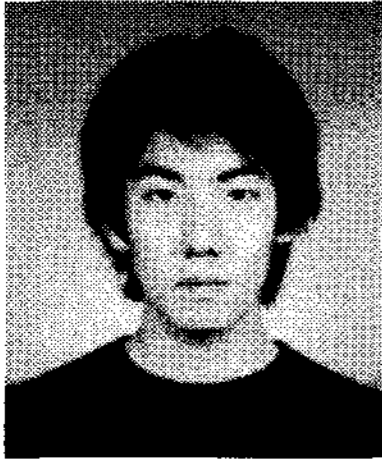
다시점 영상 부호화 기술은 시점간의 상관성을 이용하여 단일시점 영상 부호화 방식을 이용한 동시부호화 방식보다 더욱 효율적으로 다시점 영상을 부호화할 수 있지만, 추가적인 예측구조에 의해 전체 부호화 시간이 상당히 증가한다. 본 논문에서는 매크로블록 단위로 참조하는 시점을 배경영역과 객체영역으로 나누고 여기에 전역 변이 벡터를 통한 시점간의 상관성을 이용하여 현재 부호화 하는 시점의 영역을 구분하여 배경영역의 경우 선별된 모드들로만 부호화하는 방법을 제안한다. 그렇기 때문에 제안하

는 방법은 객체보다 배경이 많은 영상에서 부호화 시간을 감소시킬 수 있다. 실험 결과, 다시점 영상 부호화 표준화 회의에서 지정된 실험 영상을 사용하였을 때 제안하는 고속 모드 결정 방법을 통해 평균 BD-PSNR 0.04dB의 화질 저하만으로 전체 부호화 시간을 약 30~40% 단축할 수 있었다.

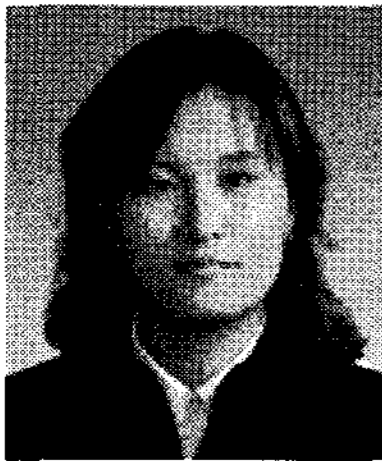
참 고 문 헌

- [1] A. Vetro, P. Pandit, H. Kimata, and A. Smolic, "Joint Draft 4.0 on Multiview Video Coding," ISO/IECJTC1/SC29/WG11andITU-TQ6/SG16,Doc.JVT-X209,Jul.2007.
- [2] ITU-T Recommendation H.264 and ISO/IEC 14496-10, Advanced Video Coding for Generic Audiovisual Services, May 2003.
- [3] "Requirements on Multi-view Video Coding v.7," ISO/IECJTC1/SC29/WG11,Doc.N8218,Jul.2006.
- [4] K. Mueller, P. Merkle, A. Smolic, and T. Wiegand, "Multiview Coding using AVC," ISO/IECJTC1/SC29/WG11,Doc.M12945, Jan.2006.
- [5] H. Schwarz, D. Marpe, T. Wiegand, "Hierarchical B pictures," ISO/IECJTC1/SC29/WG11andITU-TQ6/SG16,Doc.JVT-P014,Poznan,Poland,July2005.
- [6] J. Reichel, M. Wien, and H Schwarz, "Joint Scalable Video Model JSVM-7," ISO/IECJTC1/SC29/WG11andITU-TQ6/SG16,Doc.JVT-T202,Jul.2006.
- [7] A. Vetro, P. Pandit, H. Kimata, and A. Smolic, "Joint Multiview Video Model (JMVM) 5.0," ISO/IECJTC1/SC29/WG11andITU-TQ6/SG16,Doc.JVT-X207,Jul.2007.
- [8] H.-S. Koo, Y.-J. Jeon, and B.-M Jeon, "MVC Motion Skip Mode", ISO/IECJTC1/SC29/WG11andITU-TQ6/SG16,Doc.JVT-W081, Apr.2007.
- [9] Y. Su, A. Vetro, A. Smolic, "Common Test Conditions for Multiview Video Coding," ISO/IECJTC1/SC29/WG11andITU-TQ6/SG16,Doc. JVT-U211,Oct.2006.
- [10] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," ITU-TQ6/SG16,Doc.VCEG-M33,Apr.2001.

 저 자 소 개

**한 동 훈**

- 2007년 2월 : 세종대학교 컴퓨터공학과 (학사)
- 2007년 3월 - 현재 : 세종대학교 컴퓨터공학과 (석사)
- 주관심분야 : Video coding, Image processing, H.264/AVC, Multi-view video coding

**조 속 희**

- 1993년 2월 : 부경대학교 전자계산학과 (학사)
- 1996년 2월 : 부경대학교 대학원 전자계산학과 (석사)
- 1999년 9월 : 요코하마국립대학교 대학원 전자정보공학과 (박사)
- 1999년 11월 - 현재 : 한국전자통신연구원 전파방송연구단 방송시스템연구그룹 3DTV시스템연구팀
- 주관심분야 : Multiview video coding, MPEG-2/4 systems, 3DTV

**허 남 호**

- 1992년 2월 : 포항공과대학교 전기전자공학과 (학사)
- 1994년 2월 : 포항공과대학교 대학원 전기전자공학과 (석사)
- 2000년 2월 : 포항공과대학교 대학원 전기전자공학과 (박사)
- 2000년 4월 - 현재 : 한국전자통신연구원 전파방송연구단 방송시스템연구그룹 3DTV시스템연구팀
- 주관심분야 : 3DTV, Free-viewpoint TV

**이 영 렬**

- 1985년 2월 : 서강대학교 전자공학과 (학사)
- 1987년 2월 : 서강대학교 전자공학과 (석사)
- 1999년 2월 : 한국과학기술원 전기·전자공학과 (박사)
- 1987년 1월 - 1994년 2월 : 삼성전자 R&D 센터 Digital Media Lab.
- 1999년 3월 - 2001년 8월 : 삼성전자 R&D 센터 Digital Media Lab. 수석연구원
- 2001년 9월 - 현재 : 세종대학교 전자정보대학 컴퓨터공학부 컴퓨터공학과 부교수
- 주관심분야 : Multi-view video coding, MPEG video compression, Multimedia(MPEG) systems, H.264/AVC, Image compression, Image processing, Multimedia data(A/V/System) over IP, Embedded system for A/V transport, Transcoding, Watermarking