

논문-05-10-4-14

다시점 동영상에서의 효율적인 변이 벡터 압축 기법

김 용 태^{a)*}, 손 광 훈^{a)}

Multi-view video coding using efficient disparity vector prediction

Yongtae Kim^{a)*} and Kwanghoon Sohn^{a)}

요 약

다시점 동영상 부호화기의 성능을 향상시키기 위해서 본 논문에서는 평행식 카메라 구조에서의 효율적인 변이 벡터 예측을 이용한 부호화 방식을 제안한다. 변이 벡터는 움직임 벡터와는 달리 다시점 카메라 구조 정보로부터 예측이 가능하다. 이러한 성질을 이용하여 예측하여 구한 예측 벡터와 직접 추정된 변이 벡터와의 차이값을 부호화한다. 그러므로 변이 벡터 부호화의 성능을 향상시키기 위해서 정교한 변이 벡터의 예측이 필요하다. 기존의 벡터 예측 방식은 미리 부호화된 주위 블록의 변이 벡터를 이용하여 현재 블록의 변이 벡터를 예측하지만 제안 알고리즘은 다시점 영상간의 상관성을 이용한다. 본 논문에서는 5시점 동영상에 대해서 차벡터의 엔트로피와 절대 평균값을 구하는 실험을 수행하였다. 실험 결과를 통해서 기존의 공간적인 상관성만을 이용하여 변이 벡터를 부호화하는 방식보다 제안 알고리즘이 우수한 성능을 보임을 확인하였다. 제안 알고리즘은 기존 알고리즘과 비교하여 영상의 화질을 유지하면서 30~40%의 부호화 효율을 증가시킨다.

Abstract

To enhance the performance of multi-view sequence CODEC, an efficient disparity vector coding method for multiview sequences is proposed herein. For higher coding efficiency, we encode the differential vectors acquired by subtracting the original vectors from the predicted ones. To enhance the performance of disparity vector coding, it is essential to predict the disparity vectors accurately. The prediction by this proposed method utilizes the correlation among the multiview images, while conventional methods exploit the correlation among the causal blocks. Experiments were performed for three different 5 view sequences. We were able to confirm that the proposed method predicts disparity vectors accurately by comparing the entropy and the mean absolute values for differential vectors with conventional methods. Its performance is superior to vector coding methods used in MPEG-4 which uses only a spatial correlation. The proposed method increases the coding efficiency by a factor of 30~45% while preserving image quality.

Keyword: multiview sequence CODEC, disparity estimation, disparity vector coding

I. 서 론

이론적으로 스테레오 영상(Stereo Image)의 전송을 위해

서는 기존 영상에 비해 두 배의 대역폭이 필요하지만, 두 영상은 서로 수평으로 인간의 눈 사이 거리만큼 떨어진 곳에서 획득된 영상이므로 상호 큰 유사성(Correlation)을 가지고 있다. 그러므로 이 유사성을 이용해 데이터량을 크게 줄일 수 있는 장점이 있다. 입체감을 느끼기 위해서는 스테

a) 연세대학교 전기전자공학과*

Dept. of Electrical and Electronic Engineering, Yonsei University

레오 영상만으로 충분하지만, 다양한 시점을 제공하여 보다 현장감 있는 영상을 재생하기 위해서는 여러 위치에서 영상을 획득한 다시점 영상(Multiview Image)이 필요하다. 다시점 영상을 사용하면 시청자의 위치 이동에 따라 보는 영상이 달라지는데, 시점의 수가 많아질수록 보다 자연스러운 시점 이동이 가능해진다. 하지만 이 경우 데이터량이 시점수만큼 증가하기 때문에 효율적인 압축 방법이 필요하게 된다.

다시점 동영상의 압축시 시점수에 비례하여 영상의 정보량과 계산 복잡도가 증가하게 된다. 이 문제를 해결하기 위해 3차원 동영상 정보를 효율적으로 압축하는 알고리즘들이 제안되었다^[1-3]. 다시점 동영상 부호화기는 변이 보상 기법을 이용하여 참조 영상과 변이 벡터만을 전송하여 전체 대역폭을 줄일 수 있다.

보다 현실감 있는 영상을 전달하기 위해서 다시점 동영상 부호화기는 정해진 대역폭에서 보다 많은 시점의 영상을 전송할 수 있어야 한다. 인간의 시각 시스템은 자연스러운 입체감을 느끼기 위해서 근거리의 개인용 모니터의 경우 9시점의 다시점 영상을 필요로 한다^[4]. 기존의 압축 알고리즘들은 이렇게 많은 시점을 지원하지 못할 뿐만 아니라 매우 많은 부호화 비트를 발생시킨다. 이에 본 논문은 평행식 다시점 동영상 부호화기에서 변이 벡터를 효율적으로 압축하는 방식을 제안한다. 기존의 벡터 압축 기법은 현재 블록과 인과적인 블록들간의 공간적인 상관성을 이용한다. MPEG-1,2,4 와 H.264/AVC에서는 현재 블록의 벡터를 압축하기 위해서 이웃 블록들의 벡터들로부터 구한 예측 벡터와 현재 블록의 벡터와의 차이값을 이용한다. 그러나 제안 알고리즘은 각 시점의 같은 위치에 존재하는 변이 벡터들의 상관성을 이용하여 부호화를 수행한다. 효율적인 변이 벡터 부호화를 위해 제안 알고리즘은 매크로블록의 방향 모드 결정 과정, 변이 벡터의 예측과 보정 과정을 수행한다. 실험 결과를 통해서 제안 알고리즘은 기존 알고리즘과 비교하여 우수하게 변이 추정을 수행하며 변이 벡터 부호화에서 만족스러운 결과를 보임을 확인하였다.

본 논문의 전체적인 구성은 다음과 같다. 제 2 장에서는 다시점 카메라 구조와 전역 변이 추정 알고리즘에 대해 설명하고, 제 3 장에서는 다시점 동영상을 위한 변이 벡터 압축 알고리즘을 제안한다. 제 4 장에서는 모의실험을 통해서

제안 알고리즘의 우수성을 검증하였다. 마지막으로 제 5 장에서 결론을 맺는다.

II. 다시점 카메라 구조

본 논문에서는 그림 1과 같은 평행식 카메라 구조의 다시점 입체 영상을 대상으로 부호화를 수행한다^{[5][6]}. 평행식 카메라 구조에서는 수직 변이가 이론적으로 발생하지 않고 정합하는 변이가 가로의 성분으로만 구성된다는 장점이 있다. 또한 정합 영역이 왜곡되지 않아 그대로 정합알고리즘을 사용하여 영상을 변이 벡터로 복원할 수 있다. 또한 평행식 카메라 구조는 인간 시각 시스템 측면에서 적합한 모델이다^{[7][8]}.

변이벡터는 그림 1에서처럼 좌영상과 우영상이 매칭되는 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^[9-11]. 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. 하지만 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. 수많은 연구가 가려진 영역에 대해서 진행되어왔으며 여러 알고리즘이 개발되었지만 가려진 영역을 완벽하게 감지하고 부호화하는 것은 불가능하여 부가적인 정보를 전송해야한다^{[12][13]}.

평행식 카메라의 구조에서 다시점 영상들 중에 가장 왼쪽에 위치한 시점과 가장 오른쪽에 위치한 시점의 영상은 모든 시점의 질감 정보를 포함하고 있다. 이는 곧 두 영상

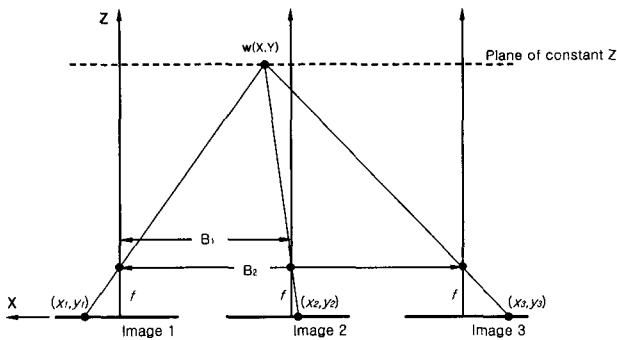


그림 1. 평행식 카메라 구조
Fig. 1. Parallel camera geometry

의 질감 정보를 미리 알고 있다면 그 두 영상의 중간 시점 영상들은 두 영상을 참조하여 복원할 수 있다는 뜻이다. 그림 2는 평행식 다시점 카메라로부터 획득된 다시점 영상의 한 예를 보여주고 있다. 그림 2 (a)에서 구는 가운데 앞부분에 위치하고 있고 왼쪽과 정육면체는 구 뒤에 위치하고 있

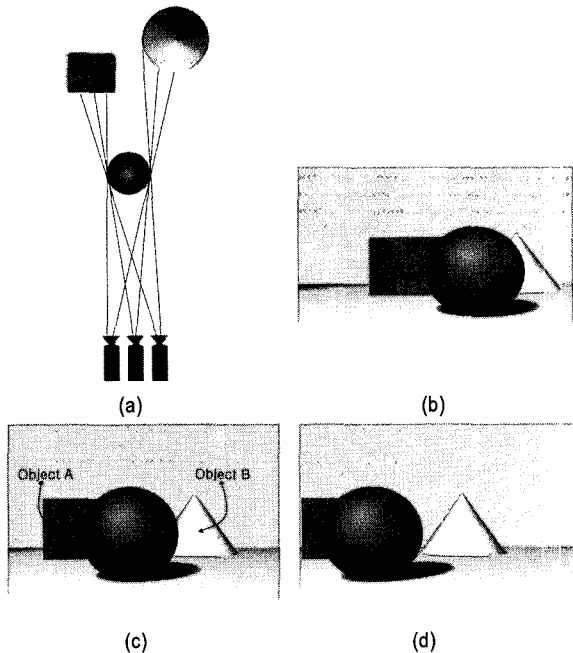


그림 2. 다시점 영상의 한 예 (a) 수직 방향의 시점 (b) 왼쪽 시점 (c) 중간 시점 (d) 오른쪽 시점
Fig. 2. A simple example of multiview images produced by MAX (a) an image from top direction (b) the leftmost view (c) the intermediate view (d) the rightmost view

다. 그림 2 (b-d)는 각각 좌영상, 중간 시점 영상, 우영상을 보여주고 있다. 중간 시점에서 객체 A는 좌영상에서 완벽하게 보인다. 그러므로 중간 시점에서 객체 A는 좌영상을 이용해서 완벽하게 복원될 수 있다. 객체 B의 경우 같은 원리로 우영상으로부터 복원될 수 있다. 그러므로 중간 시점 영상은 참조 영상을 각각 좌, 우영상으로 설정하였다면 두 영상으로부터 가려진 영역 없이 복원할 수 있다. 즉, 두 영상 중 최소 하나의 영상을 참조 영상으로 하여 가운데 중간 시점 영상을 복원하게 된다. 다시점 영상을 부호화하기 위해서 기준 시점(맨 왼쪽 시점)은 기존의 2차원 압축 표준 (MPEG-1,2,4 또는 H.264/AVC)을 이용해서 압축하고 맨 오른쪽 시점은 기준 시점을 참조 영상으로 하여 변이 추정 및 보상 기법을 사용하여 부호화를 수행한다. 좌, 우영상을 부호화한 후 이 영상들을 참조 영상을 설정하고 중간 시점들에 대해서 양방향 변이 추정을 수행하여 부호화를 수행하게 된다.

다음의 수식들은 카메라들이 완벽하게 가로방향으로 정렬되어 있어 영상 좌표계가 오직 원점의 이동만으로 구성되어 있다고 가정한다. 본 논문에서 다시점 영상들은 평행식 카메라로 획득되었기 때문에 수식 (1)을 유도한다.

$$Y - \lambda \cdot B = \lambda \cdot d \tag{1}$$

여기서 Y 는 실제 좌표계에서의 깊이 정보를 λ 는 초점 거리를 의미한다. B 와 d 는 각각 카메라 사이의 거리(Base line distance)와 다시점 간의 변이 벡터를 나타낸다. 변이 벡터는 객체의 매칭되는 지점간의 상대적인 이동 정보로 정의되는데 d_{21} 은 $x_2 - x_1$ 을 d_{31} 은 $x_3 - x_1$ 을 나타낸다. 수식 (2)는 수식 (1)로부터 얻어진다.

$$\frac{Y - \lambda}{\lambda} = \frac{B}{d} \tag{2}$$

수식 (2)에서의 왼쪽 항은 상수로만 구성되어 있는 반면, 오른쪽 항의 B 와 d 는 각 시점에 따라서 변화한다. 그러므로 B 와 d 의 비율은 같은 Y 에 대해 항상 일정하다. 여기서

B 와 d 의 비율을 α 라는 변수를 사용하여 정의한다. 이론적으로 α 의 값은 모든 시점에 대해서 동일하므로 수식 (3)을 이끌어 낼 수가 있다.

$$\alpha = \frac{B_1}{d_{21}} = \frac{B_2}{d_{31}} \quad (3)$$

α 는 깊이 정보에 대한 다른 표현이라 할 수 있다. 큰 α 값을 가지는 객체의 경우 카메라로부터 매우 멀리 떨어져 있다는 것을 의미하고 작은 α 을 가지는 객체의 경우 카메라로부터 가깝다는 것을 의미한다. 깊이 정보를 변이 벡터로 표현하지 않고 α 로 표현하는 이유는 변이 벡터의 경우 카메라의 간격에 따라서 비례하여 변동하기 때문에 변이 벡터의 크기로 깊이감을 나타내기 힘들지만 α 의 경우 같은 깊이 정보를 가질 경우 시점의 위치에 상관없이 같은 α 값을 가지기 때문에 보다 시점과 상관없이 깊이 정보를 표현할 수가 있기 때문이다.

B_1 과 B_2 의 값은 다시점 영상으로부터 직접 알아낼 수가 없다. 하지만 B_1 과 B_2 의 비율은 전역 변이 벡터로부터 계산할 수 있다. 전역 변이 벡터는 모든 매크로블록의 변이 벡터들의 평균을 통해서 구해낸다. 이 때, 가려진 영역의 매크로블록들은 제외를 시킨다. 가려진 영역을 검출은 양방향 일관성 검사를 사용하였다. 수식 (4)에서 전역 변이 벡터는 B_1 과 B_2 을 대체한다. GD_{21} 은 시점 1과 시점 2사이의 전역 변이 벡터를 의미하고 GD_{31} 은 시점 3과 시점 1사이의 전역 변이 벡터를 의미한다.

$$\begin{aligned} \alpha &= \frac{GD_{21}}{d_{21}} = \frac{GD_{31}}{d_{31}} = \dots \\ &= \frac{GD_{21}}{GD_{21} + d'_{21}} = \frac{GD_{31}}{GD_{31} + d'_{31}} = \dots \end{aligned} \quad (4)$$

B 는 실제 세계에서 mm나 cm같은 단위를 가지는 절대 거리값이다. 이 값을 pixel단위의 전역 변이 벡터로 대체함으로써 α 는 단위가 없는 깊이 정보로 표기된다. 일반적으

로 α 는 0.5와 5사이의 값을 갖는다. 전역 변이 추정은 또 다른 장점은 변이 추정의 검색 범위를 줄일 수 있다는 것이다. 기존의 방법은 변이 추정에서 영상들의 변이 추정의 상대 지점을 같은 위치로 놓고 수행한다. 그러므로 변이 추정을 수행하기 위해서 모든 가능한 변이 벡터를 포함하는 넓은 검색 범위를 필요로 하여 효율적이지 못하다. 하지만 본 논문에서는 검색 범위를 전역 변이 벡터의 주변으로 제한함으로써 작은 검색 범위에서도 정확한 변이 추정을 수행할 수 있다. 그러므로 수식 (4)에서 d_{21} 과 d_{31} 은 $GD_{21} + d'_{21}$ 과 $GD_{31} + d'_{31}$ 으로 표현될 수 있다. 이를 통해 다시점 영상의 변이 추정의 계산 복잡도를 줄일 수 있다.

III. 다시점 동영상에서의 효율적인 변이 벡터 압축 알고리즘

그림 3은 제안된 변이 벡터 부호화를 수행하는 간략한 블록도를 보여주고 있다. 첫 번째 단계에서 매크로블록의 방향 모드를 결정한다. 변이 보상 방법에 대해서 순방향, 역방향, 양방향 모드를 정의하였다. 모드를 결정한 후 각 시점에 대해서 α 값을 계산한다. 이론적으로 실제 좌표계에서 같은 깊이를 가지는 매크로블록들에 대해서는 같은 α 값을 가져야 한다. 하지만 가려진 영역과 잘못된 변이 추정에 의해서 매크로블록들은 다른 α 값을 가질 수 있다. 이

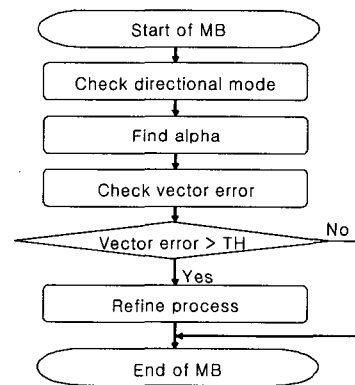


그림 3. 변이 벡터 부호화 과정 블록도
Fig. 3. Block diagram of the proposed algorithm

문제를 해결하기 위해서 각 시점에 대해서 계산된 α 값 중에서 가장 적당한 대표 α 값을 찾는 과정이 필요하다. 대표 α 값을 이용해서 예측한 벡터와 직접 추정된 벡터와의 차이 값을 구한다. 변이 벡터는 α 를 통해서 예측한 벡터와 직접 추정된 벡터와의 차벡터를 부호화하기 때문에 이 차이값이 큰 경우 많은 부호화 비트가 필요하게 된다. 큰 벡터 차이 값을 가지는 블록에 대해서는 보완과정을 통해서 새로운 변이 벡터들을 추정하여 부호화량을 감소시킨다. 벡터 차이 값이 임계값보다 작은 경우 다음 블록으로 넘어간다.

1. 방향 모드 결정

가려진 영역에 대한 추정은 심각한 열화를 일으키므로 변이 벡터의 예측은 가려진 영역을 제외한 정합이 되는 영

역에서만 수행되어야만 한다. 위에서 설명했듯이 가장 왼쪽 시점과 가장 오른쪽 시점의 영상을 참조 영상으로 사용하였다. 일반적인 경우 이 두 참조영상에서는 안 보이고 중간 시점에서만 보이는 영역은 존재하지 않기 때문에 모든 중간 시점의 매크로블록들은 이 두 참조 영상중 최소 하나의 영상으로부터 변이 보상을 수행하게 된다. 하지만 두 참조 영상을 동시에 참조하지 않고 하나씩만 하는 경우가 발생할 수도 있기 때문에 어느 쪽에서 참조했는지를 나타내는 방향 모드를 결정해야만 한다.

기존의 방법에서는 움직임 벡터의 방향 모드를 결정하기 위해서 MAD(Mean absolute difference)값을 순방향, 역방향 및 양방향에 대해서 계산한 후 가장 작은 MAD를 내는 방향을 선택하였다. 본 논문에서도 이 같은 방법을 기본적으로 사용하지만 보다 효율적인 변이 벡터의 부호화를 위

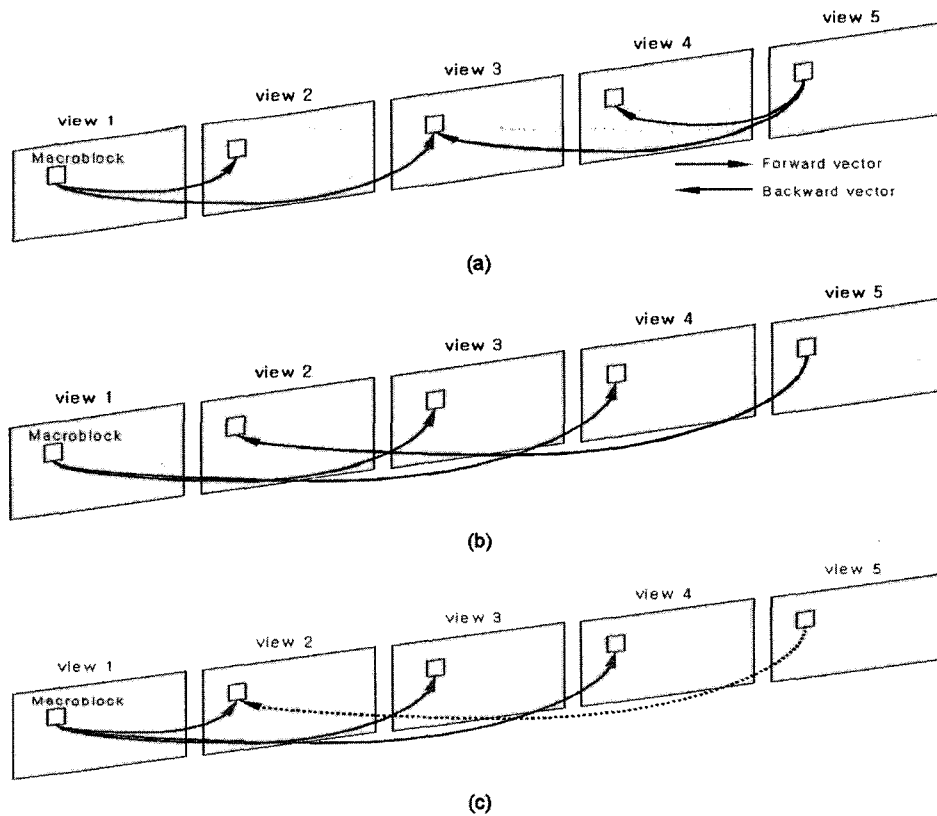


그림 4. 방향 모드 결정 (a) 정렬 모드의 예 (b) 엉킴 모드의 예 (c) 엉킴 모드의 재설정

Fig. 4. Directional mode decision (a) an example of arranged mode (b) an example of tangled mode (c) rearranging process for tangled mode

해서 매크로블록에 대해서 정렬된 방향 모드가 필요하다. 여기서 정렬된 방향 모드는 참조 영상으로부터 연속된 방향 모드로 구성되는 것으로 정의된다. 또한 정렬된 방향 모드에서 순방향 모드의 수와 역방향 모드의 수의 합은 시점의 수가 N 이라 했을 때, $N-2$ 이상이 되어야 한다. 순방향 모드의 수가 2일 경우 맨 왼쪽의 시점 1은 인접한 시점 2와 시점 3의 참조 영상이 된다. 그림 4(a)는 5시점 영상에서 정렬된 방향 모드의 한 예를 보여준다. 여기서 시점 3과 시점 4는 맨 오른쪽에 위치한 시점 5로부터 보상되어진다. 즉, 시점 2는 순방향 모드이고 시점 4는 역방향 모드, 시점 3은 순방향, 역방향 두 쪽에서 모두 보상되므로 양방향 모드이다. 이 경우 역방향 모드의 수와 역방향 모드의 수는 둘 다 2이다. 만약 모든 중간 시점들이 시점 1로부터만 참조되어진다면 순방향 모드의 수는 3이고 역방향 모드의 수는 0이다. 그러나 경우에 따라서 모든 매크로블록들의 모드들은 정렬되지 않을 수 있다. 이 경우 엉킴 모드라 정의한다. 그림 4(b)는 엉킴 모드의 한 예를 보여주고 있다. 이론적으로 평행식 카메라에서는 이런 경우가 발생할 수 없다. 그림 4(c)는 그림 4(b)의 모드들에 대한 재정렬 과정을 보여주고 있다. 재정렬하기 전에 시점 2의 매크로블록은 시점 5로부터 보상되었다. 그러나 재정렬을 통해서 시점 2의 매크로블록은 시점 1로부터 보상되게 된다. 5시점의 동영상에서 발생할 수 있는 가능한 정렬 모드는 표 1과 같이 모두 10개의 경우가 존재한다. 재정렬 과정을 수행할 때 이 10가지 경우에 대해서 가장 작은 MAD값을 가지는 경우를 선택한다.

표 1. 5시점 다시점 영상에서 가능한 방향 모드의 수
Table 1. The available combination for the rearrangement process in the case of 5-view sequences

순방향 모드의 수	역방향 모드의 수
0	3
1	2,3
2	1,2,3
3	0,1,2,3

한 시점의 방향 모드를 부호화하기 위해서는 순방향과

역방향을 정보를 보내주기 위해서 2비트의 정보가 필요하다. 즉, 중간 시점들의 방향 모드를 부호화하기 위해서는 총 $2 \times (N-2)$ 의 비트가 필요하게 된다. 그러나 각 시점의 방향 모드를 각각 부호화하는 것이 아니라 전체 다시점 영상 측면에서 순방향 모드의 수와 역방향 모드의 수만을 부호화하기 때문에 $2 \times \log_2(N-1)$ 비트만큼의 정보로 방향 모드 정보를 전송할 수 있다. 이 방식의 장점은 시점의 수에 비례하지 않는다는 것이다. 시점의 수가 선형적으로 증가하더라도 방향 정보를 전송하는데 필요한 비트는 비례하지 않는다. 즉 시점수가 증가하여도 부호화량은 상대적으로 증가하지 않는다.

2. α 값 계산

이전 단계에서는 정해진 방향 모드를 이용해서 순방향, 역방향에 대해서 각각 α 값을 계산한다. 순방향 모드나 역방향 모드가 하나만 존재할 경우 α 값은 수식 (5)과 같이 단순히 해당되는 시점의 α 값을 대표 α 값으로 설정한다.

$$\alpha = \frac{GD_{\text{corresponding view}}}{d_{\text{corresponding view}}} \quad (5)$$

순방향 모드나 역방향 모드가 일보다 큰 값을 가질 경우 수식 (6)과 같이 계산된다.

$$\alpha = \text{median}(\alpha_1, \alpha_2, \dots), \quad \alpha_1 = \frac{GD_1}{d_1}, \quad \alpha_2 = \frac{GD_2}{d_2}, \dots \quad (6)$$

일반적으로 추정 과정은 MAD를 비용 함수로 사용한다. MAD값을 최소화시키는 후보 벡터를 해당 매크로블록의 변이 벡터로 설정한다. 대부분의 경우에서 최소 MAD를 가지는 후보 벡터는 실제 좌표계에서의 깊이 정보처럼 정확한 변이 벡터가 될 수 있다. 하지만 가려진 영역과 로컬 미니멈에 빠지는 현상으로 인해 MAD의 신뢰도는 낮아질 수 있다. 잘못된 변이 벡터로 추정될 경우 해당 벡터의 α 는 잘못된 값을 가지게 된다. 만약 대표 α 값을 여러 시점의

α 값들의 평균값으로 구하게 될 경우 잘못된 α 값에 의해서 대표 α 값은 크게 달라질 수 있다. 그러므로 평균값 대신 중간값을 사용하여 여러 시점의 정보를 대표하면서 한 시점에서 크게 잘못된 α 값의 영향이 미치지 않게 하였다.

3. α 를 이용한 차벡터 계산

이전 단계에서 구한 α 값을 이용해서 구한 예측 벡터와 실제로 구한 벡터와의 차이값을 계산한다.

여기서 $v_{est}[i]$ 는 i 번째 시점에서 직접 구한 변이 벡터이고 α 값은 이전 단계에서 구한 대표 α 값이다. $\frac{GD_i}{\alpha}$ 는 수식 (3)의해서 i 번째 예측 벡터를 의미한다. V_{error} 는 직접 구한 변이 벡터와 α 를 통해서 예측한 벡터들간의 차이의 절대값이다. 변이 벡터는 예측된 벡터와 직접 구한 벡터와의 차벡터를 이용해서 부호화되므로 V_{error} 는 결국 부호화량을 의미한다. 즉, V_{error} 의 감소할수록 부호화량은 줄어들게 된다. 대표 α 값을 잘 찾는 일도 중요하지만 V_{error} 을 줄이기 위해서 후보 α 를 기반으로 하여 $v_{est}[i]$ 를 조정하는 과정도 매우 중요하다.

$$V_{error} = \sum_{i=1}^{\text{number of modes}} |v_{est}[i] - \frac{GD_i}{\alpha}| \quad (7)$$

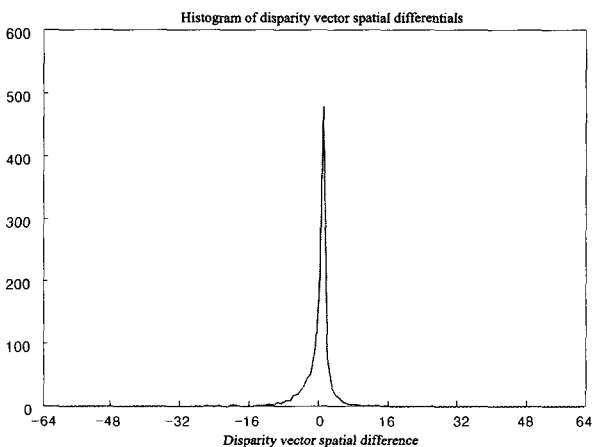


그림 5. 이웃한 블록간의 변이 벡터들의 차이값
Fig. 5. Histogram of disparity vector spatial differentials

4. 변이 벡터 보정 과정

변이 벡터 보정 과정은 V_{error} 이 정해진 임계값(TH_{vec})보다 클 경우 수행된다. V_{error} 이 작은 경우 보정 과정을 수행하지 않는다. 보정 과정은 전역 변이 추정 알고리즘과 비교하여 변이 보상에 대한 에러를 최소화시키면서 적은 벡터 차이값을 갖는 변이 벡터들을 추정한다. 대표 α 값은 보완과정을 통해서 새로 구한 변이 벡터들을 이용하여 새로 계산된다.

그림 5는 이웃한 블록의 변이 벡터간의 차이값을 보여주고 있다. 이 그림에서 이웃한 블록들의 변이 벡터들이 유사한 값을 가지는 것을 알 수 있다. 실제 세계에서 한 객체의 깊이 정보는 경계를 제외하면 대부분 일정하거나 약간의 차이를 주며 미세하게 변한다. 보정 과정에서는 이러한 이웃 블록간의 유사성을 이용한다. 이웃 블록의 대표 α 값을 현재 블록의 후보 α 값으로 설정하여 보완과정을 수행한다. 이전의 블록들의 대표 α 값은 이전 단계의 보정 과정을 거치면서 안정화되었기 때문에 신뢰도가 높다. 후보 α 값을 제공하는 이웃 블록은 그림 6과 같이 4개의 인과적인 블록으로 제한된다. 그림에서 c의 블록은 현재 블록을 의미하고 1-4번째 블록은 인과적인 이웃 블록을 의미한다. 수식 (8)을 이용해서 새로운 변이 벡터를 추정한다. 수식 (8)은 보정 과정에서의 비용 함수이다. 수식 (8-1)에서 MAD항은 현재 영상과 참조 영상간의 유사도를 표현한다.

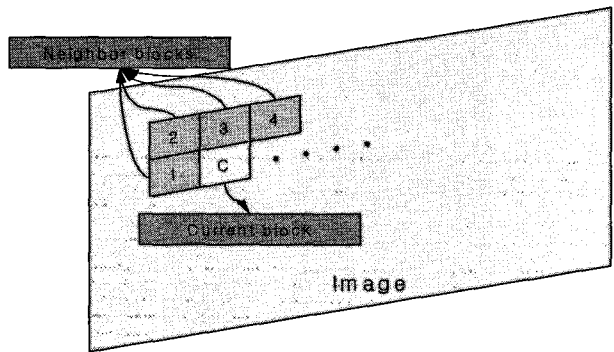


그림 6. 후보벡터를 제공하는 인과적인 이웃 블록들
Fig. 6. Blocks which offer candidate α s to refinement process

수식 (8-2)에서 후보 α 값을 이용하여 각 시점에 대해서 변이 벡터를 예측한 후 이 예측 벡터를 중심으로 보정 과정의 검색 범위를 설정한다.

$$vec_k[i] = \arg \min_{d_j} f \quad (8)$$

$$f = \sum_{x, y \in block} |I_C(x, y) - I_R(x + d_j, y)| \quad (8-1)$$

$$d_j \in \frac{GD_i}{\alpha_k} - SR_{refine}, \dots, \frac{GD_i}{\alpha_k} + SR_{refine} \quad (8-2)$$

수식 (8)에서 k 는 이웃 블록에 대한 인덱스이고 i 는 해당 시점에 대한 인덱스이다. I_C 와 I_R 은 각각 현재 영상과 참조 영상을 나타낸다. 수식 (3)에 의해서 GD_i/α_k 는 k 번째 이웃 블록의 대표 α 값으로 구한 i 번째 시점의 예측된 변이 벡터이다. 이 예측 변이 벡터를 중심으로 수식 (8-2)와 같이 SR_{refine} 만큼의 검색 범위를 설정하고 변이 추정을 수행하고 MAD를 최소화 시키는 d_j 를 $vec_k[i]$ 로 설정한다. 수식 (8)에서는 각 후보 α 값에 대해서 $vec_k[i]$ 를 추정 한 후 $vec_k[i]$ 기반으로 새로운 α 값을 계산한다. 이렇게 구한 α 를 이용해서 식 (9)를 통해 새로운 벡터 차이값을 계산한다.

$$V_{error}[k] = \sum_{i=1}^{count} |vec_k[i] - \frac{GD_k}{\alpha_k}| \quad (9)$$

수식 (10)에서는 $TotalMAD_k$ 는 해당되는 모든 시점들의 MAD의 총합이다.

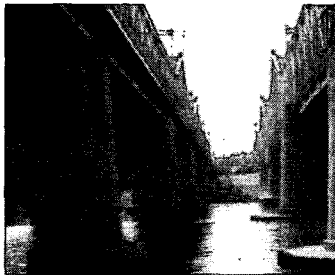
$$TotalMAD_k = \sum_{i=1}^{count} \sum_{x, y \in block} |I_C(x, y) - I_R(x + vec_k[i], y)| \quad (10)$$

$V_{error}[k]$ 가 임계값보다 작은 k 에 대해서 수식 (11)과 같이 $TotalMAD_k$ 를 구하고 가장 작은 값을 가지는 k 의 α 와 $vec_k[i]$ 를 현재 매크로블록의 대표 α 값과 변이 벡터로 설정한다.

$$(\alpha, vec[i]) = \arg \min_{(\alpha_k, vec_k[i])} TotalMAD_k \quad (11)$$

IV. 모의실험 결과 및 분석

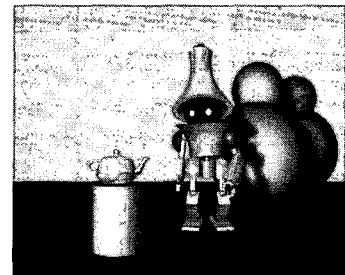
본 논문에서는 그림 7의 수평방향으로 평행한 5시점 145장의 영상에 대해서 실험을 수행하였다. 352×288 크기의 "Bridge" 영상과 320×240 크기의 "Venus" 영상은 5대의 카메라로부터 같은 시간에 획득된 동영상이고 320×240 크기의 "Robot" 영상은 3D MAX를 이용하여 생성된 가상 영상이다. 실험에 사용된 변이 추정은 8×8 크기의 블록 단위로 수행되었으며 -32~+32의 검색 범위를 할당하였다. 기존의 변이 추정 방식과 전역 변이 추정 알고리즘을 비교 알고리즘으로 하여 다양한 파라미터에서 제안 알고리즘을 실험하였다. 기존 알고리즘들의 변이 벡터는 MPEG VLC



(a) Bridge



(b) Venus



(c) Robot

그림 7. 각 실험 영상의 맨 왼쪽 시점의 첫 번째 프레임 (a) "Bridge" (b) "Venus" (c) "Robot"

Fig. 7. The experimental images of 1st frame of view #1 (a) "Bridge" (b) "Venus" (c) "Robot"

테이블을 이용하여 부호화되었다.

다시점 동영상 레벨에서는 각 시점에 대해 하나의 GD 값을 전송하고 각 매크로블록마다 방향 모드 정보를 전송한다. 이 두 정보는 실험 결과에 반영되어 있다. α 값들은 가장 오른쪽 시점 (View-5)의 변이 벡터로부터 유추할 수 있다. 식 (4)로부터 $\alpha = GD_{51}/d_{51}$ 를 얻어낼 수 있다.

GD_{51} 와 d_{51} 은 각각 시점 1과 시점 5사이의 전역 변이 벡터와 해당 매크로블록의 변이 벡터를 의미한다. 그러므로 α 값에 대한 추가적인 정보를 필요로 하지 않는다. 하지만 보정 단계에서 발생하는 차벡터에 대해서는 전송해야만 한다. 결국 전송되는 정보는 GD 값, 방향 모드 정보, 보정 단계에서 발생하는 차벡터 정보이다.

그림 8은 TH_{vec} 변화에 따른 비트율 왜곡을 보여주고 있다. TH_{vec} 가 커질수록 왜곡은 감소하고 비트율은 증가한다. 본 논문에서는 어느 정도의 왜곡을 허용하는 한도에서 발생하는 비트량을 최대한 낮추고자 노력하였다. 발생하는 비트량이 TH_{vec} 가 40이상인 경우 변화량이 거의 없으며 그 이하에서는 점진적으로 증가하는 것을 확인할 수 있다. 이 결과에서 작은 TH_{vec} 은 더 많은 블록에서 보정 단계를 수행하게 되고 그 보정 단계를 통해서 발생하는 비트량이 감소하는 것을 확인할 수 있다. TH_{vec} 가 10보다 작은 경우 PSNR은 급격히 증가하지만 10이상에서는 수렴한다. 영상의 화질만을 고려한다면 TH_{vec} 가 큰 값으로 설정하면 충분하지만 그런 경우 많은 비트량을 발생시키기 때문에 적당한 TH_{vec} 값을 결정해야만 한다. 본 논문에서는 그림 8 (a),(b)와 (c)로부터 비트율 왜곡 측면에서 10에서 20사이의 TH_{vec} 의 값을 선택한다. 그림 9는 TH_{vec} 을 10으로 설정하여 SR_{refine} 에 따른 PSNR 변화를 보여주고 있다. 발생하는 비트량은 SR_{refine} 에 비례하여 증가하지만 3이상에서는 PSNR은 수렴한다. 이 결과를 통해 좋은 화질의 영상을 얻기 위해서 큰 SR_{refine} 을 설정할 필요는 없다.

비교 알고리즘으로 전방향 탐색 알고리즘과 전역 변이추정 알고리즘을 사용하였다. 전방향 탐색 알고리즘은 정해진 검색 범위내의 모든 검색 지점에 대해서 MAD값을

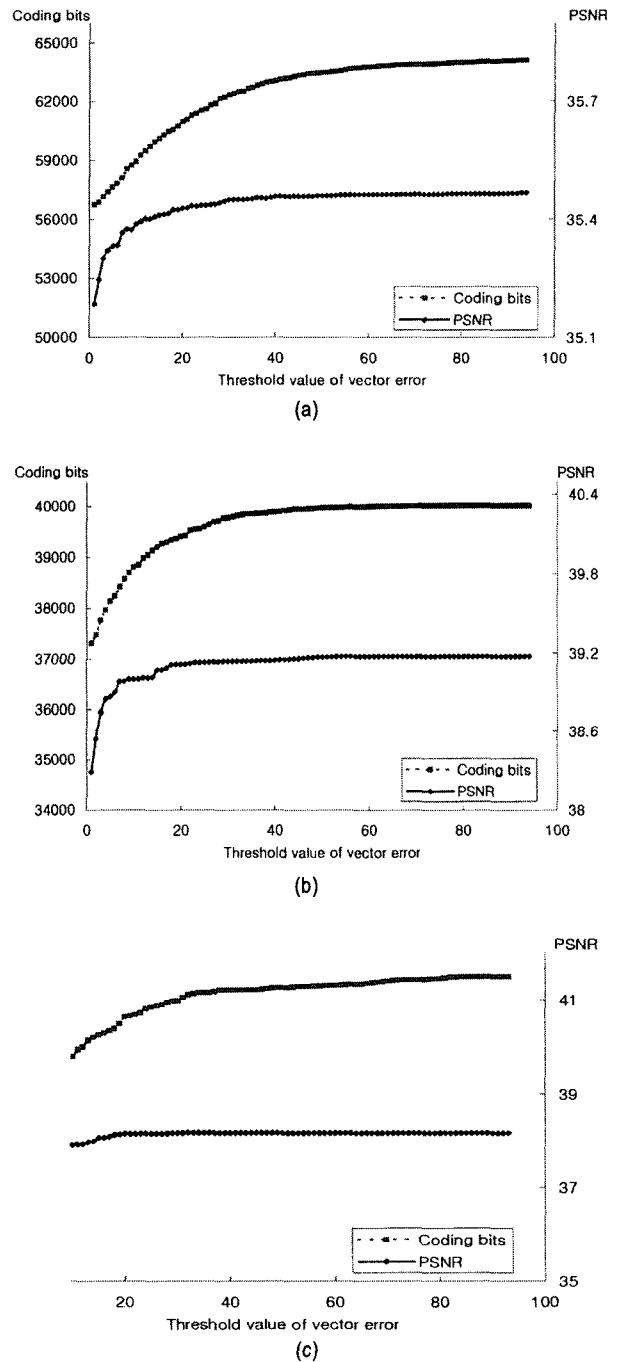


그림 8. 다양한 TH_{vec} 에서의 부호화 비트량과 PSNR (a) "Bridge" (b) "Venus" (c) "Robot"
 Fig. 8. Coding bits and PSNR for various TH_{vec} (a) "Bridge" (b) "Venus" (c) "Robot"

표 2. 차벡터의 절대값 평균과 엔트로피 비교

Table 2. Comparisons of mean absolute value and entropy for differential vectors

방법		전방향 탐색 알고리즘	전역 변이 추정 알고리즘	제안 알고리즘
절대값의 평균	Bridge	6.080	3.941	1.161
	Venus	3.267	2.190	1.340
	Robot	1.632	1.143	0.491
엔트로피	Bridge	4.460	4.298	2.268
	Venus	3.695	3.328	2.873
	Robot	2.297	2.271	1.666

표 3. 기존 알고리즘과의 성능 평가

Table 3. Comparisons of performance with the existing methods

방법		전방향 탐색 알고리즘	전역 변이 추정 알고리즘	제안 알고리즘
RMSE	Bridge	4.45	4.523	4.44
	Venus	2.808	2.821	2.819
	Robot	3.083	3.239	3.135
PSNR(dB)	Bridge	35.163	35.023	35.184
	Venus	39.162	39.124	39.129
	Robot	38.351	37.923	38.207
bpb (bits/block)	Bridge	20.58	19.82	11.94
	Venus	15.43	14.94	11.00
	Robot	11.39	11.18	6.18

계산하고 그 값이 최소인 지점을 벡터로 설정한다. 전역 변이 추정 알고리즘은 다시점 영상간의 정보를 이용하여 검색 범위의 중점을 전역 변이만큼 이동시키고 검색 범위를 제한하는 알고리즘이다. 이렇게 추정된 벡터들은 MPEG-4의 공간적인 상관성을 이용하는 방식을 이용하여 부호화된다. 표 2는 차벡터의 절대값의 평균과 엔트로피를 보여주고 있다. 절대값의 평균은 변이 벡터의 예측의 정확도를 의미하며 이 값이 작다는 것은 예측 벡터의 값이 실제 벡터와 유사하다는 것을 나타낸다. 여기서는 엔트로피는 변이 벡터를 부호화하기 위해서 필요한 비트량을 의미한다. 제안 알고리즘은 기존 알고리즘과 비교하여 절대값의 평균과 엔트로피 측면에서 가장 작은 값을 갖는다. 이 결과는 제안 알고리즘이 높은 부호화 효율을 갖는 것을 보여 준다. 표

3에서는 세 가지 변이 추정 방식들의 복원된 영상의 PSNR과 RMSE와 변이 벡터를 부호화하기 위해 소모된 비트량을 비교하였다. 비트량은 블록당 발생하는 비트량 (bpb, bits per block)을 기준으로 하였다. 전방향 탐색 알고리즘은 -32~+32의 검색범위를 전역 변이 추정 알고리즘은 전역 변이를 기준으로 -16~+16의 검색 범위를 할당하였다. 제안 알고리즘은 TH_{vec} 은 10으로 SR_{refine} =3으로 설정하였다. 표 3으로부터 제안 알고리즘이 기존 알고리즘과 비교하여 영상의 화질을 유지하면서 가장 작은 bpb를 가짐을 확인할 수 있다. 세부적으로 부호화량이 "Bridge" 영상에서 42%, "Venus" 영상에서는 28.3%, "Robot" 영상에서는 45% 감소한다. "Bridge"의 변이 맵이 "Venus"와 비교하여 매우 복잡하기 때문에 보다 큰 성능 개선이 있었다. 이는 변이 맵이 복잡한 경우 벡터간의 상관도를 저하시키기 때문에 기존 알고리즘의 부호화 성능을 감소시키기 때문이다. "Robot" 영상의 경우 인공적으로 합성된 영상이기 때문에 평활한 영역이 많고 변이 맵이 단조롭다. 하지만 기존 알고리즘들은 평활한 영역에서 잘못된 변이 벡터를 추정할 확률이 높기 때문에 올바른 벡터를 추정한 제안 알고리즘에 비해 성능이 감소한다.

V. 결론

본 논문은 다시점 동영상에서 효율적인 변이 벡터 부호화 방법을 제안하였다. 기존 방식은 각 시점내에서 공간적인 상관성만을 이용하여 부호화하기 때문에 객체의 경계나 가려진 영역에서 잘못된 예측을 통해 부호화 성능이 감소하였다. 제안 알고리즘은 시점간의 상관성을 이용하기 때문에 변이 벡터를 효율적으로 예측할 수 있어 낮은 엔트로피를 가지는 차벡터로부터 변이 벡터에 필요한 비트량을 감소시킬 수 있었다. 제안 알고리즘에 사용된 전역 변이 추정을 통해 변이 벡터의 신뢰도를 증가시켰으며 방향 모드 결정을 통해 모드를 부호화하기 위해 필요한 비트량을 감소시키고 정확한 참조 시점을 결정할 수 있었다. 또한 보정 단계를 통해 변이 보상의 왜곡을 감소시켜 화질을 유지시킬 수 있었다. 실험 결과는 제안 알고리즘이 기존 알고리즘

과 비교하여 비슷한 화질을 유지하면서 30~45%정도의 비트량을 감소시키는 것을 확인하였다. 하지만 추가적인 과정으로 인해 높은 계산 복잡도가 발생하는데 이를 낮추기 위해 향후 연구를 수행하고자 한다.

참 고 문 헌

[1] G..C Chang and W.N. Lie, Multi-View Image Compression and Intermediate View Synthesis For Stereoscopic Applications, IEEE International Symposium on Circuits and Systems, ISCAS 2000, vol. 2, pp. 277-280, May, 2000.

[2] Y. Zhang and G. Li, An efficient hierarchical disparity estimation algorithm for stereoscopic video coding, IEEE Asia-Pacific Conference on Circuits and Systems, pp. 744-747, December, 2000.

[3] D.R. Clewer, L.J. Luo, C.N. Canagarajah, D.R. Bull and M.H. Barton, Efficient multiview image compression using quadtree disparity estimation, IEEE International Symposium on Circuits and Systems, vol. 5 pp. 295-298, May, 2001.

[4] S. Pastoor and K. Schenke, Subjective Assensments of the Resolution of Vewing Directions in a Multi-Viewpoint 3D TV System, Proc. SID, vol.3, no.3, pp. 217-222, 1989.

[5] A. Tamtaou and C. Labit, Coherent disparity and motion compensation in 3DTV image sequence coding schemes, in ICASSP, pp. 2845-2848, 1991.

[6] L. Oisel, E. Memin, L. Morin and F. Galpin, One-Dimensional Dense Disparity Estimation for Three-Dimensional Reconstruction, IEEE Trans. Image Processing, vol. 12, no. 9, pp. 1107-1119, September, 2003.

[7] H. Aydinoglu and M.H. Hate, Stereo image coding : A projection approach, IEEE Trans. Image Processing, vol. 7, no. 4, pp. 506-516, April, 1998.

[8] N. Grammalidis and M.G. Strintzis, Disparity and occlusion estimation in multiocular systems and their sequences, IEEE Trans. Circuits and Systems for Video Technol. vol. 8, no. 3, pp. 329-343, June, 1998.

[9] E. Izquierdo, Stereo Matching for Enhanced Telepresence in Three-Dimensional Videocommunications, IEEE Trnas. Circuits and Systems for Video Technol., vol. 7, no. 4, pp. 629-643, August, 1997.

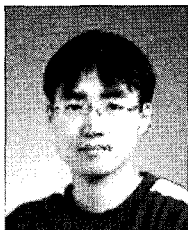
[10] D. Tzovaras, N. Grammalidis and M. G. Strintzis, Object-Based Coding of Stereo Image Sequences Using Joint 3-D Motion/Disparity Compensation, IEEE Trans. Circuits and Systems for Video Technol., vol. 7, no. 4, pp. 312-327, April, 1997.

[11] 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.

[12] 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, June, 2000.

[13] 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, September, 2002.

저 자 소 개



김 용 태

- 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



손 광 훈

- 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 전기전자공학과 부교수
- 주관심분야 : 3차원 영상처리, 영상 통신