

論文98-35S-7-8

Multi-View 영상 코딩에서 영상 개선 알고리즘

(An Algorithm for the Image Improvement in the Multi-View Images Coding)

金 導 鉉 * , 崔 東 準 * , 梁 榮 日 **

(Do-Hyun Kim, Dong-Jun Choi, and Yeong-Yil Yang)

요 약

본 논문에서는 multi-view 영상으로부터 최적의 깊이와 텍스처를 추출하는 효과적인 알고리즘을 제안하였다. 기존에 제안된 알고리즘이^[1] 깊이와 텍스처를 동시에 추출하는데 비해 제안된 알고리즘은 깊이 추출 단계와 텍스처 추출 단계의 2단계로 구성되어 있다. 정규화된 물체 공간에서 X - Y 로 형성되는 면을 삼각패치로 나누고 각 삼각패치 노드의 Z 값을 깊이 추출 단계에서 결정한 후, 두 번째 단계에서 삼각패치의 텍스처를 추출한다. 깊이 추출 단계에서는 삼각패치의 노드를 중심으로 이루어진 윈도우 영역에 대하여 블록에 기반한 변이 보상 모델을 적용함으로써 노드의 깊이를 결정한다. 텍스처 추출 단계에서는 아핀 변환에 기반한 변이 보상 모델을 이용하여 multi-view 영상으로부터 각 삼각패치의 텍스처가 추출된다. *Dragon*, *santa*, *city* 그리고 *kid*의 multi-view 영상 세트에 대해 실험한 결과 제안된 알고리즘으로 깊이와 텍스처를 추출한 후 이로부터 복원된 영상의 신호 대 잡음비(SNR)는 기존의 알고리즘으로 처리된 후 복원된 영상^[1]의 신호 대 잡음비보다 평균 4dB 정도 개선된 결과를 얻을 수 있다.

Abstract

In this paper, we propose an efficient multi-view images coding algorithm to find the optimal depth and texture from the set of multi-view images. The proposed algorithm consists of two consecutive steps, i) the *depth extraction step*, and ii) the *texture extraction step*, compared with the traditional algorithm^[1] which finds the depth and texture concurrently. The X - Y plane of the normalized object space is divided into triangular patches and the Z value of the node is determined in the first step and then the texture of the each patch is extracted in the second step. In the *depth extraction step*, the depth of the node is determined by applying the block based disparity compensation method to the windowed area centered at the node. In the second step, the texture of the triangular patches is extracted from the multi-view images by applying the affine transformation based disparity compensation method to the triangular patches with the depth extracted from the first step. Experimental results show that the SNR(Signal-to-Noise Ratio) of images encoded by our algorithm is better than that of images encoded by the traditional algorithm by the amount about 4dB for the test sets of multi-view images called *dragon*, *kid*, *city* and *santa*.

* 正會員, 慶尙大學校 航空機部品技術研究所
(GyeongSang National University, Research
Center for Aircraft Parts Technology)

** 正會員, 慶尙大學校 電子材料工學科
(GyeongSang National University, Dept. of
Electronic Materials Engineering)

接受日字:1997年12月30日, 수정완료일:1998年5月25日

I. 서론

80년대에는 음성과 영상정보 전달을 위한 미디어로 전화와 텔레비전에서의 정보 전송을 중심으로 발전하였고, 컴퓨터간에는 정보전달 수단으로 데이터전달 관련기술이 발전되어왔다. 집적회로기술, 신호처리기술 그리고 통신기술과 다중화 기술이 급속하게 발전함에 따라, 독립적으로 이루어지던 음성, 영상, 데이터가 하나의 미디어로 처리되고 전송되는 멀티미디어의 시대에 도래하게 되었다. 멀티미디어 관련 기술의 지속적인 발전은 1.2Gbps의 전송을 요구하는 고선명 텔레비전 방송을 가능하게 하였다. 이런 발전추세를 보면 고선명 텔레비전보다 더 많은 데이터량의 전송처리가 요구되는 3차원 텔레비전은 고도의 인간 감각에 좀더 다가갈 수 있는 영상을 보내어 새로운 문화를 창조해 가는 다음 세대의 미디어로 부각될 것이다.

3차원 영상코딩 기술은 스테레오 영상처리^[3]와 더불어 임의의 위치에서 보았을 때 영상을 재구성하는 것이 용이한 multi-view 영상처리에 관한 연구가 진행되었다.^[1, 2, 4] Multi-view 영상처리는 여러 장의 영상을 깊이와 텍스처로 추출하여 압축하여 저장 또는 전송한 후 복원하는 기술이다. Fujii 등은^[1] 정규화 좌표계를 사용하여 물체를 표현한 후 블록에 기반한 변이 보상 방법을 사용하여 깊이와 텍스처를 동시에 구하거나, 이항 변환에 기반한 변이 보상 모델을 사용하여 깊이와 텍스처를 구하였다. 블록에 기반한 변이 보상 모델을 사용하면, 텍스처와 깊이 정보 모두 지역적인 값에 의해 결정되나 artifact가 있는 텍스처가 얻어진다. 이항 변환에 기반한 보상 모델을 사용하면 artifact가 없는 텍스처가 구해진다. 삼각패치의 연속성에 의하여 깊이는 지역적인 성격을 잃어버린 부정확한 값이 추출된다. 본 논문에서는 깊이와 텍스처를 추출하는 단계를 분리함으로써, 깊이와 텍스처 모두 지역적인 값에 의해 결정되고 artifact가 없는 복원영상을 얻는 알고리즘을 제안한다. 정규화 공간에서 X - Y 면을 삼각패치로 나눈다. 첫 단계는 깊이 추출 단계로 각 노드를 둘러싼 사각형 영역에 대하여 블록에 기반한 변이 보상 방법을 적용함으로써 노드의 깊이를 결정한 후, 두 번째 단계에서 삼각패치의 텍스처를 이항 변환에 기반한 변이 보상 모델을 이용하여 artifact 현상이 없는 텍스처를 추출하고자 한다.

제 2 장에서는 multi-view 영상과 물체공간의 기

하학적 관계를 기술하였고, 제 3 장에서는 본 논문에서 제안한 알고리즘에 대해 기술하였다. 실험 결과를 제 4 장에 설명하였고, 제 5 장에서 결론을 내렸다.

II. Multi-view 영상과 물체공간의 기하학적 관계

여러 대의 카메라를 사용하여 물체를 촬영하거나, 한 대의 카메라를 이동시키면서 물체를 촬영하였을 때, 동일한 물체에 대하여 여러 장의 영상이 얻어진다. 이때 얻어진 여러 장의 영상을 multi-view 영상 세트라 한다. 그림 1은 *dragon* 물체를 촬영하여 얻어진 multi-view 영상 세트를 보여준다. 카메라는 $+z$ 축을 향하여 $(-2D, 0, z_0)$, $(-D, 0, z_0)$, $(0, 0, z_0)$, $(D, 0, z_0)$ 과 $(2D, 0, z_0)$ 에 놓여있다. 뷰포인트에서 $+z$ 방향으로 f 만큼 떨어진 곳에 상이 맺히게 되며, 상이 맺히는 x - y 평면을 영상 평면 x_v - y_v 라 한다.

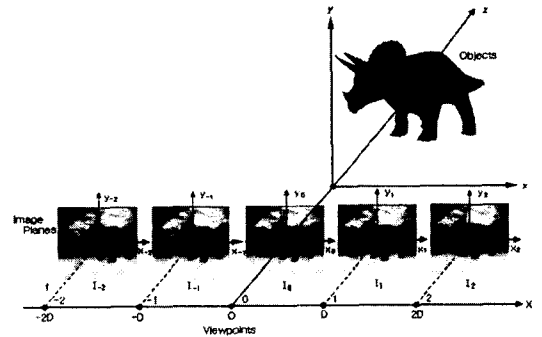


그림 1. Multi-view 영상 세트
Fig. 1. The set of the multi-view images.

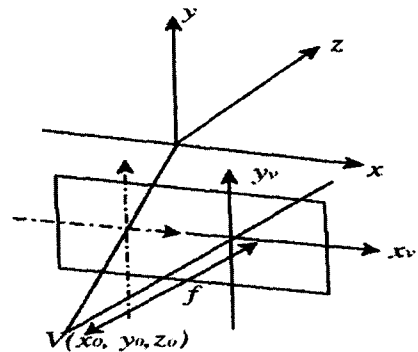


그림 2. 투영 기하학
Fig. 2. Geometry of the perspective projection.

그림 2에서 물체공간의 한 점(x, y, z)이 영상 평면 상의 점(x_v, y_v)에 투영되면, 두 점 사이는 식(1)로 표현되고, 뷰포인트와 물체공간 좌표계의 원점을 연결하였을 때 영상 평면과의 교점(x'_v, y'_v)은 식 (2)와 같이 된다. 분석을 간단히 하기 위하여 교점(x'_v, y'_v)을 영상 평면 x_v-y_v의 원점으로 변경하면 식 (1)은 식 (3)과 같이 표현된다.

$$x_v = f \frac{x - x_0}{z - z_0}, \quad y_v = f \frac{y - y_0}{z - z_0} \quad (1)$$

$$x'_v = f \frac{x_0}{z_0}, \quad y'_v = f \frac{y_0}{z_0} \quad (2)$$

$$x_v = f \frac{x - x_0}{z - z_0} - x'_v = f \frac{x - x_0}{z - z_0} - f \frac{x_0}{z_0} = f \frac{x - x_0}{z - z_0} \frac{z - z_0}{z_0} \quad (3)$$

$$y_v = f \frac{y - y_0}{z - z_0} - y'_v = f \frac{y - y_0}{z - z_0} - f \frac{y_0}{z_0} = f \frac{y - y_0}{z - z_0} \frac{z - z_0}{z_0}$$

n번째의 뷰포인트에서 얻어진 영상 I_n(m, l)로 표시할 때 영상 I_n(m, l)과 영상좌표 (x_v, y_v)는 m=k₁x_v, l=k₂y_v의 관계를 가진다. 여기서 k₁과 k₂는 3차원 TV에서 수평 방향과 수직 방향의 해상도와 연관된 상수이다. 이를 식 (3)에 대입하고, x₀=nD, y₀=0, z₀=-F 이므로 (x, y, z)와 영상 I_n(m, l)의 관계는 식 (4)와 (5)로 표현되고, X, Y, Z를 식 (6)과 같이 정의하면 식 (4)와 (5)는 식 (7)과 (8)로 표현된다.

$$m = k_1 f \frac{x - \frac{x_0}{z_0} z}{z - z_0} = k_1 f \left(\frac{x + \frac{nD}{F} z}{z + F} \right) \quad (4)$$

$$= k_1 f \frac{x}{z + F} + k_1 f \frac{nDz}{(z + F)F}$$

$$l = k_2 f \frac{y - \frac{y_0}{z_0} z}{z - z_0} = k_2 f \frac{y}{z + F} \quad (5)$$

$$X = k_1 f \frac{x}{z + F}, \quad Y = k_2 f \frac{y}{z + F},$$

$$Z = k_1 f \frac{zD}{(z + F)F} = k_1 \frac{f}{(z + F)F} z \quad (6)$$

$$m = X + nZ \quad (7)$$

$$l = Y \quad (8)$$

그림 3은 multi-view 영상의 기하학적 관계를 보여준다. 그림 3에서 왼쪽 그림은 물체공간에서 multi-view 영상을 획득하는 과정이다. 카메라는 y=0, z₀=-F 인 위치에 x축으로 D만큼 떨어져 놓여 있다. 그러므로 뷰포인트는 V(nD, 0, -F)로 표현되고, n은 정수이며 n=0 이면 z 축상의 뷰포인트를 나타낸다. n번째 영상은 뷰포인트 V(nD, 0, -F)에서 얻어진 영상을 말한다.

식 (6)의 정의에 의해 만들어진 X-Y-Z 좌표계로 형성되는 공간은 정규화 공간(normalized space)으로 정의되고, 그림 3의 오른쪽 위 부분에 나와 있다. 물체공간 좌표로 표현된 물체가 정규화 공간 좌표계로 변환되어 표현되고, 정규화 공간상의 점 (X, Y, Z)는 식 (7)과 (8)에 의하여 I_n(m, l)로 투영된다. 식 (7)은 multi-view 영상간의 좌표를 Z 픽셀(pixel) 만큼의 변이(disparity)로 정규화하는 것이다. 그림 3의 오른쪽 아래 부분에 보인 것처럼, I_n(m, l)를 n 방향에 따라 배열하여 형성되는 n-m-l 공간을 multi-view 영상 공간(multi-view image space)이라 한다. Multi-view 영상 공간의 특징을 이용해 multi-view 영상의 압축과 복원을 수행한다.

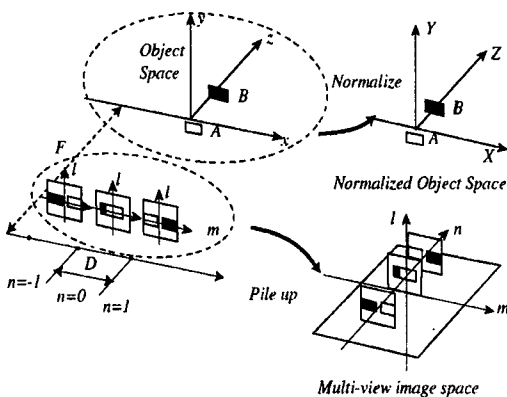


그림 3. 정규화 물체 공간과 multi-view 영상
Fig. 3. Normalized object space and multi-view images.

III. Multi-view 영상 압축 알고리즘

Multi-view 영상 처리는 여러 장의 영상에서부터 깊이와 텍스처를 추출하는 압축 단계와 깊이와 텍스처로부터 원래의 영상을 복원하는 단계로 구분된다. 1 절에서는 multi-view 영상의 압축에 사용되는 블록에 기반한 변이 보상 모델과 아핀 변환에 기반한 변이 보

상 모델에 대하여 살펴보고, 2 절에서는 제안한 알고리즘을 기술한다. 그리고 3 절에서는 블록에 기반한 변이 보상 모델을 사용하여 압축을 수행하였을 때 발생하는 artifact를 보정하는 방법에 대하여 기술한다.

1. 블록과 아핀 변환에 기반한 변이 보상 모델

이 절에서는 블록에 기반한 변이 보상 모델과 아핀 변환에 기반한 변이 보상 모델을 설명한다. 그림 4는 블록에 기반한 변이 보상 모델을 보여준다. 그림 4(a)에서 정규화 공간에서 블록 A와 B의 X와 Y좌표가 같고 Z가 다른 곳에 위치해 있다. 블록 A와 B는 식 (7)과 식(8)에 의해 multi-view 영상으로 투영되는데, 그림 4(a)의 오른쪽 그림은 왼쪽의 multi-view 영상들을 쌓아서 만들어진 multi-view 영상공간과 이를 위에서 본 단면도($m-n$ 면)을 나타낸다. 블록 A는 n 이 커질수록 상이 왼쪽으로 이동하므로 $m-n$ 면에서 음의 기울기를 갖고, 블록 B는 n 이 커질수록 상이 오른쪽으로 이동하므로 $m-n$ 면에서 양의 기울기를 갖는다. 블록 A와 B의 Z가 다른 곳에 위치하면 multi-view 공간에서 다른 기울기를 갖는 모양으로 나타난다. 그림 4(b)의 왼쪽 그림에서처럼 블록이 +Z 방향으로 움직일 때, $m-n$ 면에서 블록에 대응되는 상은 그림 4(b)의 오른쪽에 나타나 있다. 블록이 Z가

음인 곳에 있으면 음의 기울기를 갖고 Z가 양인 곳에 있으면 양의 기울기를 갖는다. 이 때 모든 Z에 대하여 각 영상에서의 블록에 대응되는 값들의 평균값과 휘도 분산값을 식 (9)와 식 (10)을 사용하여 구하고, 이 중 가장 적은 휘도 분산값을 갖는 위치에 블록을 넣는다. 그리고 평균값을 해당 블록의 텍스처로 할당한다. 이 방법으로 텍스처와 깊이를 구하는 것이 블록에 기반한 변이 보상 모델이다.

$$I_{avg}(X, Y, Z) = \frac{1}{(2N+1)} \sum_{n=-N}^N I_n(X+nZ, Y) \quad (9)$$

$$I_{var}(X, Y, Z) = \frac{1}{2N+1} \sum_{n=-N}^N (I_n(X+nZ, Y) - I_{avg}(X, Y, Z))^2 \quad (10)$$

이 방법을 사용하면 구하여진 블록들은 정규화 공간에서 그림 5(a)의 왼쪽 그림과 같이 불연속적으로 존재한다. 이로부터 multi-view 영상을 복원하면, 블록들이 연속되어 있지 않으므로 영상이 존재하지 않는 부분이 발생한다. 복원된 multi-view 영상을 쌓아서 $m-n$ 면을 살펴보면 그림 5(a)의 오른쪽과 같이 나타나는데, 영상이 존재하지 않는 부분을 artifact라 한다. 아핀 변환에 기반한 변이 보상 모델에서는 정규화 공간에서 $Z=0$ 인 X-Y 평면을 삼각패치로 나누고, 삼각 패치들은 그림 5(b)의 왼쪽 그림과 같이 연속적으로 존재한다. 그러므로 그림 5(b)의 오른쪽 그림에 보인

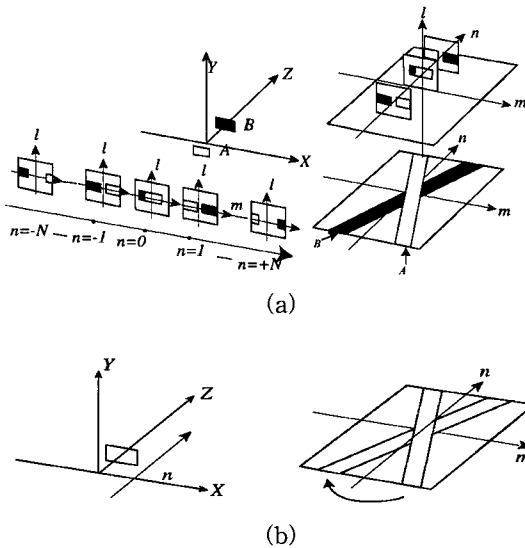


그림 4. (a) Multi-view 영상과 multi-view 공간 (b) 블록의 위치에 따른 $m-n$ 면에서 기울기 변화
Fig. 4. (a) Multi-view images and multi-view space (b) slopes according to the position of the block.

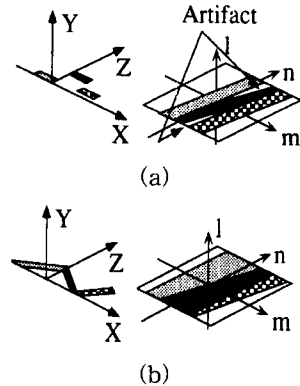


그림 5. (a) 블록에 기반한 변이 보상 방법에서 생성된 artifact와 (b) 아핀변환에 기반한 변이 보상
Fig. 5. (a) Artifacts resulting from the block-based disparity compensation, and (b) disparity compensation based on the affine transformation.

것처럼 artifact 현상이 없어지고 multi-view 영상에서 상이 맺히지 않는 부분은 발생하지 않는다. 블록에 기반한 변이 보상 모델 방법에서와 같이 각 영상에서 삼각패치에 해당되는 값들의 평균값과 휘도 분산값을 구한다. Z 값을 변화시키면서 구한 휘도 분산값들 중에서 가장 적은 휘도 분산값을 갖는 위치에 삼각패치의 노드를 고정한다. 그리고 평균값을 해당 삼각패치의 텍스처로 할당함으로써, 노드의 깊이와 텍스처를 동시에 구한다. 다음절에서는 블록에 기반한 변이 보상 모델로 깊이를 구하는 방법을 아핀 변환에 기반한 변이 보상모델의 방법과 비교 설명하였다.

2. 제안한 텍스처와 깊이 추출 방법

기존의 아핀 변환에 기반한 변이 보상 모델에서는 그림 6(a)처럼 X - Y 평면을 삼각패치로 나눈다. 여기서 소문자 a, b, c, d 는 삼각패치의 노드를 나타내고, 대문자 A, B, C, D, E, F, G 는 삼각 패치면을 나타낸다. 깊이가 정해진 두 개의 노드 a, c 로부터 깊이가 정해지지 않은 나머지 노드 d 의 깊이는 d 의 Z 를 변화시키면서 A 삼각패치에 대응되는 휘도 분산값을 구하고 최소인 값을 갖는다. d 의 깊이를 결정한 후, 정해진 a, c, d 의 정보를 이용하여 참여한 multi-view 영상의 텍스처의 평균값을 추출하여 패치 A 의 텍스처로 한다. 다음 노드 b 의 깊이를 결정은 d 의 깊이 결정과 같은 방법으로 이미 정해진 a 와 d 의 깊이와 B 삼각패치의 분산값은 이용하여 결정하고 텍스처 B 를 추출한다. 깊이와 텍스처를 동시에 구하는 아핀 변환에 기반한 변이 보상 방법은 삼각패치 노드의 깊이를 결정하는데 하나의 삼각패치 정보만을 이용하기 때문에 최적의 깊이를 얻을 수 없다. 즉, 노드 d 의 깊이는 패치 A, B, C, E, F, G 의 모든 정보를 고려해야한다. 제안한 방법에서는 깊이와 텍스처를 추출하는 단계를 분리하여, 정확한 깊이를 추출할 뿐만 아니라, 정확한 깊이를 사용해 텍스처를 추출하는 알고리즘을 제안하였다. 알고리즘 1은 제안된 알고리즘의 pseudo code 이다.

Input : Multi-view 영상 세트

Output : 깊이와 텍스처

STEP1 : 정규화 공간에서 $Z=0$ 인 X - Y 평면을 삼각패치로 나눈다. (그림 6 참조) 삼각패치 노드의 $Z=0$ 이다.

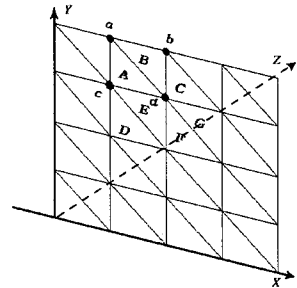
STEP2 : 각 노드에서 노드를 둘러싼은 사각형 영역

(블록)에 대하여 노드의 Z 값을 변화시키면서 블록에 기반한 변이 모델을 이용하여 텍스처 분산값이 가장 적은 노드의 깊이, Z 값을 찾는다.

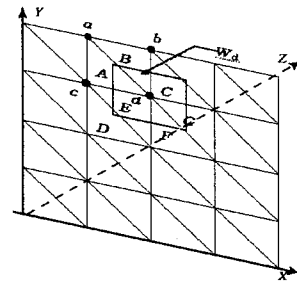
STEP3 : Multi-view 영상에서 각 삼각패치에 대응되는 텍스처들의 평균값을 구한 후, 삼각패치에 할당한다.

알고리즘 1 : 제안한 알고리즘의 슈도 코드

Algorithm 1 : The pseudo-code of the proposed algorithm.



(a)



(b)

그림 6. (a) 아핀 변환에 기반한 변이 보상을 이용한 깊이 결정 (b) 제안한 방법

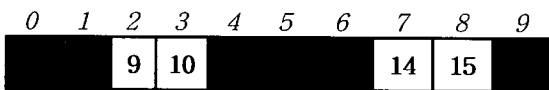
Fig. 6. Decision of depth using (a) disparity compensation based on affine transformation and (b) proposed method.

정규화 공간에서 $Z=0$ 인 X - Y 평면을 삼각패치로 나눈다. 첫 단계는 깊이 추출 단계로 그림 6(b)처럼 노드 d 깊이를 결정하기 위해서, 노드 d 를 둘러싼은 사각형 영역 W_d 에 대하여 블록에 기반한 변이 보상 방법을 적용함으로써 노드의 깊이를 결정한다. 즉 블록 W_d 에 대하여 3.1절에 기술한 방법으로 블록 W_d 가 놓일 수 있는 모든 Z 에 대하여 각 영상에서의 블록에 대응되는 값들의 평균값과 휘도 분산값을 식 (9)와 식 (10)을 사용하여 구하고, 이 중 가장 적은 휘도 분산

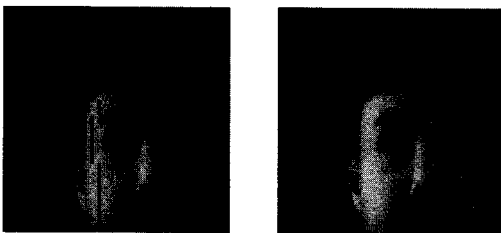
값을 갖는 위치에 블록을 넣는다. 이러한 방법은 빠른 속도로 깊이를 결정할 뿐 아니라, 노드 주위의 지역적(local)인 값만 고려하기 때문에 전역적(global)인 값을 고려하는 방법보다 더욱 정확한 깊이를 결정할 수 있는 장점이 있다. 실험에서 각 노드의 간격은 8 픽셀이며, 사용한 윈도우 W_d 의 크기는 4×4 , 8×8 과 16×16 을 사용하여 실험하였고, 16×16 에서 가장 좋은 결과를 얻었다. 첫 번째 단계에서 노드들의 위치가 결정 되었으므로, 두 번째 단계로 multi-view 영상으로부터 아핀 변환을 이용하여 각 삼각패치에 대응되는 텍스처들의 평균값을 구한 후 삼각패치에 할당한다. 삼각패치에서 노드 값이 다르기 때문에 multi-view 영상에서 대응하는 삼각형의 모양은 일정하지 않다. 임의의 위치에 바라보는 삼각패치의 모양은 아핀 변환에 의해 생성되며 아핀 변환에 기반한 변이 보상 모델을 이용하므로 artifact 현상이 없는 영상을 얻을 수 있다.

3. Artifact를 갖는 영상의 보정

블록에 기반한 변이 보상 모델^[1]을 사용하였을 때 나타나는 artifact를 다음과 같이 보정하여 보았다. 그림 7(a)은 한 줄이 10개의 픽셀로 이루어진 예를 보여준다. 위의 숫자는 해당되는 픽셀의 위치를 나타낸다. 회색은 artifact 효과로 복원 후 텍스처의 값이 할당되지 않은 픽셀이고, 흰색은 복원 후 텍스처가 구해진 픽셀을 나타내고 숫자는 픽셀의 값을 나타낸다.



(a)



(b)

(c)

그림 7. (a) 회색 픽셀은 artifact영역 (b)보정 전의 영상과 (c)보정 후의 영상

Fig. 7. (a) Shaded regions represent the pixels with artifacts. (b) Image with artifact (c) Image after interpolation.

블록에 기반한 변이 보상 모델에 의해 복원된 영상에 나타나는 artifact는 3 종류로 분리할 수 있다. 영상의 좌측에서부터 시작하는 경우(픽셀 0과 1), 영상의 우측에 값이 채워지지 않는 경우(픽셀 9)와 내부에 artifact가 존재하는(픽셀 4, 5, 6) 경우가 있다. 좌측에 존재하는 픽셀 0과 1은 처음으로 값이 나타난 픽셀 2의 값 9로 배정한다. 우측의 픽셀 9의 경우는 우측에서 처음으로 나타나는 값(픽셀 8의 값)으로 보정한다. 내부에 존재하는 연속적인 artifact에 대해서는(픽셀 4, 5, 6) 인접한 픽셀(픽셀 3과 픽셀 7)의 값을 사용하여 보정한다. 그림 7(b)은 보정하기 전의 영상이고 그림 7(c)은 보정한 후의 영상을 보여준다. 보정한 결과 평균 신호 대 잡음비는 향상되지만, 보정 흔적이 뚜렷하게 나타남을 볼 수 있다.

IV. 실험결과

본 논문에서 제안한 알고리즘을 실험하기 위하여 컴퓨터로 생성한 dragon 영상과 일본 Tsukuba 대학에서 제공한 실 영상 kid, city, santa 영상을 사용하였다. 각 영상은 9개의 256×256 그레이 영상으로 이루어져 있다. 블록에 기반한 변이 보상 모델, 아핀 변환에 기반한 변이 보상 모델 그리고 제안한 방법을 위 영상에 대하여 실험하여 신호 대 잡음비(SNR)를 비교하여 보았다.

$$SNR = 20 \log_{10} \frac{255}{\sqrt{MSE}} \tag{10}$$

여기서 MSE는 복원 영상과 원래 영상의 에러 제곱의 평균을 나타낸다.

그림 8은 dragon 영상에 대하여 각 알고리즘들을 사용하여 압축한 후 복원한 multi-view 영상 중 $n=-3, 0, 3$ 번째 영상을 보여 주고 있다. 그림 8(a)는 인코딩하기 전의 dragon 영상이고, 그림 8(b)는 블록에 기반한 변이 보상으로 복원한 dragon 영상이며, 그림 8(c)는 그림 8(b)에서 artifact를 보정한 블록에 기반한 변이 보상 모델로 복원한 dragon 영상이고, 그림 8(d)는 아핀 변환에 기반한 변이 보상 모델로 복원한 dragon 영상이며, 그림 8(e)는 제안한 방법에 의해 복원한 dragon 영상이다. Table 1은 dragon 영상에 대하여 각 알고리즘으로 압축한 후 복원하였을 때 신호 대 잡음비를 나타내며, 그림 9에서 그래프로 표



(a)



(b)



(c)



(d)



(e)

그림 8. Dragon multi-view 영상 세트 (a) 인코딩하기 전의 원 영상($n=-3, 0, 3$) (b) 블록에 기반한 변이 보상 방법으로 복원된 영상 (c) artifact를 보정한 영상 (d) 아핀 변환에 기반한 변이 보상 방법으로 복원한 영상 (e) 제안한 방법으로 복원한 영상

Fig. 8. Sets of the multi-view images called *dragon*, (a) the original images before encoding($n=-3, 0, 3$) (b), (c), (d) and (e) are the decoded images by the transformation methods based on the block matching, block matching with interpolation after decoding, affine transformation and proposed transformation, respectively.

현하였다. 제안한 알고리즘에 의해 압축된 *dragon* 복원 영상에서 평균 신호 대 잡음비는 아핀 변환에 기반한 모델보다 약 3dB 개선되었다. Table 2는 *kid* 영상에 대하여 각 알고리즘에 의한 신호 대 잡음비를 나

타낸다. Table 3은 *city* 영상에 대하여 실행하였을 때 신호 대 잡음비를 나타낸다. 실 영상인 *kid*, *city* 그리고 *santa* 영상에 대해서 제안한 알고리즘으로 영상을 압축한 후 복원한 결과 기존의 알고리즘에 비해 각각 4dB, 3dB, 그리고 4dB 개선되었다. 보정한 알고리즘이 개선된 알고리즘보다 신호 대 잡음비가 더 좋은 경우는 베이스라인이 짧고 단순한 물체일 경우이며, 복잡한 장면에서는 개선된 알고리즘의 신호 대 잡음비가 더 좋을 것을 관찰 할 수 있었다.

표 1. 변환 방법에 따른 dragon 영상의 신호 대 잡음비 비교

Table 1. Comparison of SNR (Signal-to-Noise Ratio) for the multi-view images called dragon according to /the transformation methods.

Number of image	SNR(dB)			
	블록기반 알고리즘	보정한 알고리즘	아핀 변환 알고리즘	제안된 알고리즘
-4	10	28	25	26
-3	12	30	26	28
-2	14	31	27	30
-1	17	33	29	33
0	34	34	30	34
1	17	33	30	33
2	14	31	28	31
3	13	30	26	29
4	10	28	24	27

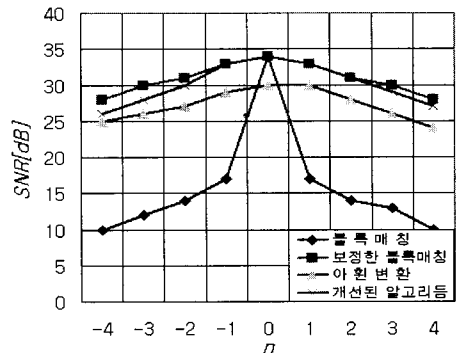


그림 9. 변환 방법에 따른 dragon 영상의 신호 대 잡음비 비교

Fig. 9. Comparison of SNR (Signal-to-Noise Ratio) for the multi-view images called dragon according to the transformation methods.

표 2. 변환 방법에 따른 kid 영상의 신호 대 잡음비 비교

Table 2. Comparison of SNR(Signal-to-Noise Ratio) for the multi-view images called kid according to the transformation methods.

Number of image	SNR(dB)			
	블록기반 알고리즘	보정한 알고리즘	아원변환 알고리즘	제안된 알고리즘
-4	15	29	24	27
-3	16	30	25	29
-2	18	32	27	30
-1	20	33	28	32
0	34	34	29	32
1	20	33	27	32
2	18	32	26	30
3	16	31	24	29
4	15	29	23	28

표 3. 변환 방법에 따른 city 영상의 신호 대 잡음비 비교

Table 3. Comparison of SNR(Signal-to-Noise Ratio) for the multi-view images called city according to the transformation methods.

Number of image	SNR(dB)			
	블록기반 알고리즘	보정한 알고리즘	아원변환 알고리즘	제안된 알고리즘
-4	15	25	24	26
-3	16	27	25	27
-2	17	28	26	29
-1	19	29	27	30
0	30	30	28	30
1	20	29	27	30
2	18	28	26	29
3	17	27	26	28
4	15	26	24	27

V. 결론

본 논문에서는 multi-view 영상 코딩에서 정규화 공간에서의 물체를 생성할 때 전체 노드의 깊이를 결정하기 위해 노드 주위의 값을 고려한 블록에 기반한 변이 보상 모델로 최적의 깊이를 결정한 후 아원 변환에 기반한 변이 보상 모델로 텍스처를 추출하는 방법을 제안하였다. 제안한 알고리즘은 Windows NT에

서 Visual C++를 사용하여 구현하였고, 9개의 영상으로 이루어진 multi-view 영상 셋에 대하여 Pentium II를 사용하여 수행하였을 때 인코딩 할 때 약 12~15초 소요되고, 디코딩시 약 2초가 소요되었다.

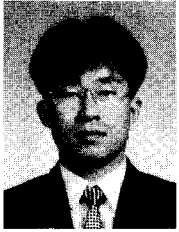
제안한 알고리즘은 기존 알고리즘^[1]과 시간과 공간에 대한 난이도(Complexity)는 같다. 평균 신호 대 잡음비는 아원 변환에 기반한 변이 보상 모델에 비해 평균 약 4dB 향상되었다. 제안한 방법은 깊이와 텍스처를 추출하는데 특징 추출이나 세그멘테이션이 필요하지 않고, 병렬성이 유지되므로 하드웨어로 구현이 용이하다.

참고 문헌

- [1] T. Fujii and H. Harashima, "Data Compression and Interpolation of Multi-View Image Set," *IEICE Trans. Inf. & Syst.*, E77-D, 9, pp. 987-995, 1994.
- [2] M. E. Lukas, "Predictive Coding of Multi-Viewpoint Image Sets," *IEEE ICASSP'86*, pp. 521-524, 1986.
- [3] M. G. Perkins, "Data Compression of Stereopairs," *IEEE Trans. Commun.* vol. 40, pp. 684-696, 1992.
- [4] H. Aydinoglu and M. H. Hayes III, "Compression of Multi-View Images," *IEEE ICIP'94*, pp. 385-389, 1994.
- [5] R. C. Bolles, H. H. Baker and D. H. Marimont, "Epipolar-Plane Image Analysis : An Approach to Determining Structure from Motion," *International Journal of Computer Vision*, 1, pp. 7-55, 1987.
- [6] S. T. Barnard and W. B. Thompson, "Disparity analysis of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 333-340, July, 1980.
- [7] U. R. Dhond and J. K. Aggarwal, "Structure from Stereo-A Review," *IEEE Trans. Systems, Man, Cybernetics*, vol. 19, no. 6, pp. 1489-1510, 1989.
- [8] T. Naemura, M. Kaneko and H. Harashima, "3-D Segmentation of Multi-view Images Based on Disparity Estimation," in *Proc. SPIE VICIP'96 : Visual Communication*, 1996.

- [10] M. D. Levine, D. A. O'Handley and G. M. Yagi, "Computer determination of depth maps," *Computer Graphics and Image Processing*, vol. 2, no. 2, pp. 131-150, 1973.
- [11] Y. Ohta and T. Kanade, "Stereo by intra-and inter-scanline search using dynamic programming," *IEEE Trans. PAMI*, vol. 7, no. 2, pp. 139-154, 1985.
- [12] T. Yasuno and T. Hamano, "Three dimensional reconstruction using homocentric spherical spatiotemporal image analysis," *Journal of Visual Communication and Image Representation*, vol. 2, no. 4, pp. 365-372, 1991.
- [13] H. Aydinoglu and M. H. Hayes III, "Compression of Multi-View Images," *IEEE ICIP '94, II*, pp. 385-389, 1994.
- [14] 김도현, 양영일, "Multi-view 영상표현을 위한 개선된 알고리즘," *제10회 영상처리 및 이해에 관한 워크숍*, pp. 3-8, 1998.

 저 자 소 개



金 導 鉉(正會員)

1970년 4월 11일생. 1996년 2월 경상대학교 전자재료공학과 졸업(학사). 1998년 2월 경상대학교 전자재료공학과 졸업(석사). 현재 경상대학교 항공기부품기술연구소 연구원.

주관심분야는 ASIC 설계와 영상처리신호임.



崔 東 準(正會員)

1969년 6월 7일생. 1992년 2월 경상대학교 전자재료공학과 졸업(학사). 1998년 2월 경상대학교 전자공학과 졸업(석사). 현재 경상대학교 항공기부품기술연구소 연구원. 주

관심분야는 영상처리신호임.

梁 榮 日(正會員)

1983년 2월 경북대학교 전자공학과 졸업(학사). 1985년 한국과학기술원 전기 및 전자공학과 졸업(석사), 1989년 한국과학기술원 전기 및 전자공학과 졸업(박사). 1990년~현재 경상대학교 전기전자공학부 부교수. 1994년 1월~1995년 1월 UC, Irvine 교환교수. 경상대학교 항공기부품기술연구소 연구원. 주관심분야 VLSI&CAD, 영상신호처리 등.