

일반논문-08-13-3-01

특이값 분해에 기반한 3차원 메쉬 동영상의 SNR 계층 부호화

허준희^{a)}, 김창수^{b)†}, 이상욱^{a)}

SNR Scalable Coding of 3-D Mesh Sequences Based on Singular Value Decomposition

Jun-Hee Heu^{a)}, Chang-Su Kim^{b)†}, and Sang-Uk Lee^{a)}

요약

본 논문은 특이값 분해에 기반하여 다양한 화질을 지원하는 3차원 메쉬 동영상의 SNR 계층 부호화 기법을 제안한다. SVD는 메쉬 동영상을 적은 수의 기저 벡터들과 특이값들로 표현하여 부호화 성능을 높일 수 있다. 본 논문에서는 비트 평면 부호화를 적용한 후 각 이진화 단계와 화질 사이의 관계를 정량적으로 유도한다. 유도된 관계식을 이용하여 비트량-왜곡 측면에서 최적화된 부호화 순서를 정의한다. 또한 시공간 영역의 잉여 정보를 효율적으로 제거하는 예측 기법을 제시한다. 모의 실험을 통하여 제안하는 알고리즘이 다양한 SNR을 지원하며 기존의 기법에 비해 향상된 비트량-왜곡 성능을 발휘함을 보인다.

Abstract

We propose an SNR-scalable coding algorithm for three-dimensional mesh sequences based on singular value decomposition (SVD). SVD achieves a coding gain by representing a mesh sequence with a small number of basis vectors and singular values. First, we introduce a bit plane coding scheme and derive a quantitative relationship between each bit plane and the reconstructed image quality. Using the relationship, we develop a rate-distortion (RD) optimized coding algorithm. Moreover, we propose prediction techniques to exploit the spatio-temporal correlations in real mesh sequences. Simulation results demonstrate that the proposed algorithm provides significantly better RD performance than conventional SVD coders.

Keyword : SVD, Mesh, Sequence, Scalable Coding

1. 서론

멀티미디어 기술의 발전에 따라 3차원 데이터는 게임, 영화, 애니메이션 등 다양한 분야에서 널리 활용되고 있다. 3차원 데이터를 표현하는 다양한 방법들이 개발되었으며, 그

중 가장 많이 쓰이는 방법은 꼭지점(vertex), 선분(edge), 그리고 면(face)으로 3차원 형상을 표현하는 메쉬(mesh) 표현 방법이다. 그러나 메쉬를 사용하여 복잡하고 세밀한 3차원 물체를 표현하기 위해서는 방대한 양의 데이터가 필요하다. 더욱이 3차원 정지 영상의 연속된 형태로 이루어진 3차원 메쉬 동영상의 경우 너무 많은 정보량을 요구하므로 저장 및 전송에 어려움이 있다. 따라서 3차원 메쉬 동영상의 효율적 압축 기법의 개발은 매우 중요한 연구 분야이다.

3차원 메쉬 동영상은 일반적으로 맥스(MAX)나 마야(MAYA)와 같은 애니메이션 소프트웨어를 사용하여 만들

a) 서울대학교 전기컴퓨터공학부, 뉴미디어통신공동연구소
Signal Processing Lab., School of Electrical Engineering and INMC,
Seoul National University, Seoul, Korea
b) 고려대학교 전기전자전파공학부*
School of Electrical Engineering, Korea University, Seoul, Korea*
† 교신저자 : 김창수(changsukim@korea.ac.kr)

어진다^[1,2]. 이와 같은 동영상에서는 모든 프레임이 동일한 개수의 꼭지점과 동일한 연결 정보를 갖는다. Lengyel은 이와 같은 3차원 메쉬 동영상의 압축 기법을 처음으로 제안하였다^[3]. MPEG-4에서는 애니메이션을 만드는 방법 중 하나인 키 프레임 애니메이션(key frame animation) 기법에 기반한 압축 알고리즘을 사용한다^[6,7]. 또한 2차원 동영상 압축과 유사하게 현재 프레임을 이전 프레임으로부터 움직임 보상을 한 후 그 오차만 압축하는 알고리즘들도 제안되었다^[4,5].

다른 접근 기법으로써 특이값 분해(SVD: singular value decomposition)를 이용한 메쉬 동영상 압축 알고리즘도 제안되었다. Alexa와 Müller가 처음 SVD에 기반한 주성분 분석(PCA: principal components analysis)을 통해 3차원 동영상을 압축하였고^[8], 이후 Karni와 Gotsman은 PCA와 선형 예측 기법을 결합하여 시간적인 잉여 정보를 제거함으로써 압축 성능을 향상시켰다^[9]. 마지막으로 Heu 등은 SVD 기저 벡터들의 에너지에 따른 차등 비트 할당을 통해 사용할 기저 벡터의 개수를 자동으로 결정하며, 비트량-왜곡(RD: rate-distortion) 성능을 향상시키는 기법을 제안하였다^[10]. 그러나 이와 같은 SVD 기반의 알고리즘들은 모두 단일 전송률만을 지원한다. 반면에 최근에는 휴대폰, PDA, PC, 디지털 TV 등 다양한 해상도와 계산 성능을 갖는 복호기들이 사용되고 있다. 이와 같은 다양한 복호기들을 지원하기 위해서는 하나의 비트열(bitstream)로부터 서로 다른 전송률과 화질을 갖는 동영상의 복원을 가능하게 하는 계층 부호화(scalable coding) 기법의 개발이 필수적이다.

본 논문은 SVD에 기반한 3차원 동영상 데이터의 SNR (signal-to-noise ratio) 계층 부호화 기법을 제안한다. 대부분의 논문에서와 같이, 우리는 3차원 데이터의 공간 좌표 정보에 대한 부분을 고려한다. 우선 SNR 계층 부호화를 지원하기 위한 비트 평면 부호화(bit-plane coding) 기법을 제안하며, 시공간 상의 잉여정보를 효과적으로 제거하기 위한 예측 기법을 제안한다. 모의 실험을 통하여 제안하는 알고리즘이 기존의 알고리즘에 비하여 효과적인 압축 성능을 발휘함을 확인한다.

본 논문은 다음과 같이 구성된다. 2장에서 SVD와 이를 사용한 기존의 메쉬 부호화 알고리즘을 살펴보고, 3장에서 제안하는 알고리즘을 설명한다. 4장에서는 모의 실험을 통

하여 제안 알고리즘의 성능을 비교 평가한다. 마지막으로 5장에서 결론을 맺는다.

II. 기존의 SVD에 기반한 메쉬 동영상 부호화 알고리즘

1. 특이값 분해 (SVD)

SVD는 행렬 A 를 아래와 같이 분해한다.

$$A = U \Sigma V^T \quad (1)$$

여기서 U 와 V 는 정규직교(orthonormal) 행렬이고, Σ 는 대각(diagonal) 행렬로써 특이값이라 불리는 양수 σ_i 들로 구성된다. 본 논문에서는 σ_i 들이 크기의 내림 차순으로 정렬되었다고 가정한다. 행렬 A 를 다음 식과 같이 구성 요소의 합으로 표현할 수 있다.

$$A = \sum_i A_i = \sum_i \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad (2)$$

여기서 \mathbf{u}_i 와 \mathbf{v}_i 는 각각 U 와 V 의 i 번째 열 벡터를 의미하며 특이벡터라고 불린다. i 번째 구성 요소 $A_i = \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ 의 에너지는 σ_i^2 로 주어진다. 그림 1은 "Face" 동영상에 해당하는

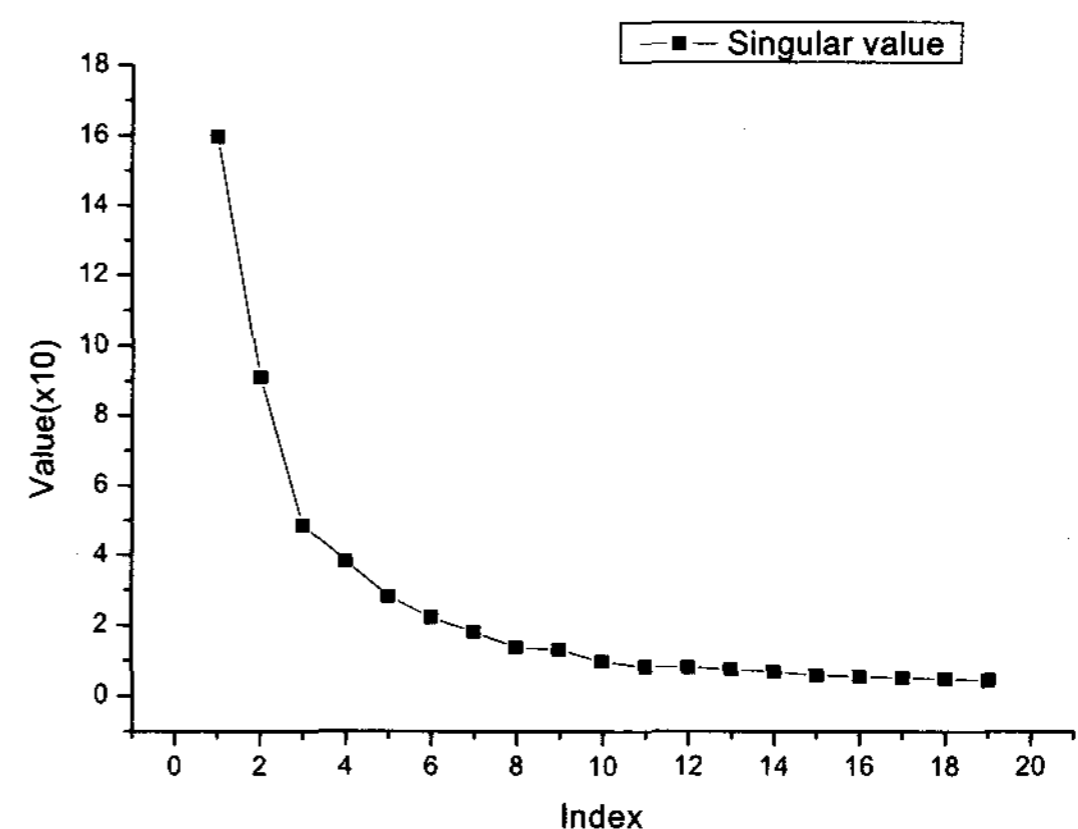


그림 1. "Face" 동영상에 해당하는 행렬의 특이값 크기.
Fig. 1. The singular value of "Face" sequence.

행렬의 특이값 크기를 도시하는데, 특이값들이 빠르게 감소하여 처음 몇 개를 제외하고는 매우 적은 값을 가짐을 확인할 수 있다. 이와 같은 경향은 다른 메쉬 동영상에서도 확인할 수 있는 일반적인 특징이다. 이는 행렬 A의 에너지가 대부분 처음 몇 개의 구성 요소에 집중되는 것을 의미한다. 그러므로 아래와 같이 에너지가 작은 특이 벡터들을 버리고 처음 k개의 특이벡터만을 사용해도 원본 A와 유사한 복원 결과를 얻을 수 있다.

$$A \approx \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_k \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & L & 0 \\ 0 & \sigma_2 & L & 0 \\ M & M & O & M \\ 0 & 0 & L & \sigma_k \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_k^T \end{bmatrix} \quad (3)$$

이와 같이 행렬 A를 k개의 특이값과 2k개의 특이벡터로 표현함으로써 부호화 이득을 얻을 수 있다.

2. SVD에 기반한 메쉬 동영상 부호화 기법

Alexa와 Müller는 SVD를 사용한 메쉬 동영상 부호화 기

법을 처음 소개했다^[8]. 그들은 메쉬 동영상을 다음과 같은 행렬 A로 구성한다.

$$A = \begin{bmatrix} p_{11} & L & p_{1f} \\ M & O & M \\ p_{n1} & L & p_{nf} \end{bmatrix} \quad (4)$$

A의 각 열은 프레임을 표현하며 f는 프레임의 개수를, n은 한 프레임 내의 꼭지점의 개수를 나타낸다. 즉 p_{ij} 는 j번째 프레임의 i번째 꼭지점을 표현한다. 그림 2는 "Bounce Ball" 동영상에 대한 SVD 결과를 보여준다. 왼쪽 특이 벡터는 공간적인 기저 벡터를 표현하는 반면에 오른쪽 특이 벡터는 시간적인 기저 벡터를 나타낸다. 따라서 우리는 왼쪽 특이 벡터를 공간 기저 벡터라고 정의하고, 오른쪽 특이 벡터를 시간 기저 벡터라 정의한다. 메쉬 동영상의 압축은 식 (3)과 같이 작은 특이값에 해당하는 기저 벡터들을 생략함으로써 수행한다.

Karni와 Gotsman은 Alexa와 Müller 기법의 성능 향상을 위해서 특이값과 시간 기저 벡터로 구성된 행렬 $C = \sum V^T$

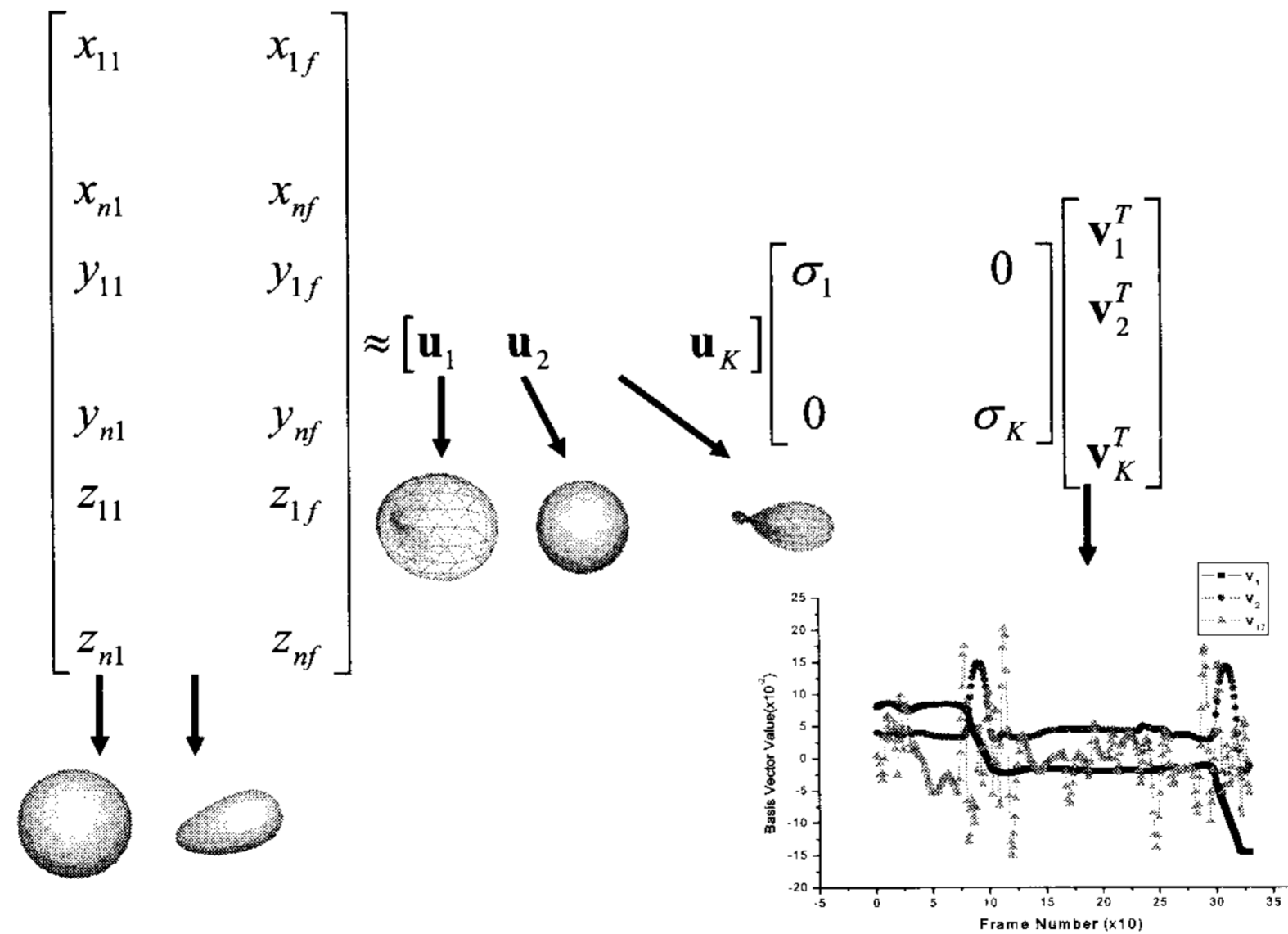


그림 2. "Bounce Ball" 동영상의 특이값 분해
Fig. 2. SVD of "Bounce Ball" sequence

의 원소 c_i 에 다음과 같은 선형 예측 부호화를 적용하였다^[9].

$$\hat{c}_i = a_0 + \sum_{k=1}^m a_k c_{i-k} \quad (5)$$

여기서 예측 계수 a_0, a_1, \dots, a_m 은 평균 제곱 예측 오차 (mean square prediction error)를 최소화하도록 결정되며 결정된 계수값은 부가 정보로써 복호기에 전송된다. 부호기는 예측된 오차를 전송하며, 복호기에서는 동일한 예측을 수행한 후 전송 받은 예측 오차를 더하여 원 신호를 복원한다.

Heu 등은 SVD에 차등 양자화 스텝 크기(quantization step size)를 적용하여 비트량-왜곡 성능을 향상시키는 알고리즘을 제안하였다^[10]. 그들은 우선 SVD에 기반한 메쉬 동영상의 부호화 기법의 비트량-왜곡 모델을 개발한 후, 라그랑지 등식 (Lagrangian equation) $J = D + \lambda R$ 을 적용하여 다음과 같이 기저 벡터들의 양자화 스텝 크기 사이의 최적화된 관계를 도출하였다.

$$q_i = \frac{\sigma_1}{\sigma_i} q_1 \quad (6)$$

여기서 q_i 는 i 번째 기저 벡터의 양자화 스텝 크기이다. 이를 통하여 다른 SVD에 기반한 기법에 비해서 동일한 비트량을 사용할 때 더 좋은 화질을 제공할 수 있었다. 그러나 이상의 기존 기법들은 계층 부호화를 지원하지 못하고 단일 전송을 압축(single-rate coding)만을 수행한다는 단점을 갖는다.

III. 제안 알고리즘

앞 장에서 언급한 바와 같이 SVD를 통하여 메쉬 동영상을 특이값, 공간 기저 벡터 및 시간 기저 벡터로 분해할 수 있다. 본 장에서는 이상의 분해된 정보를 효과적으로 압축하는 3차원 동영상의 SNR 계층 부호화와 함께 시간, 공간적 잉여 정보를 제거하기 위한 시공간 기저 벡터의 예측

기법을 제안한다.

1. 비트 평면 부호화

우선 SVD에 기반한 메쉬 동영상의 비트 평면 부호화 기법을 제안한다. 복호기에서 복원되는 메쉬 동영상의 화질 또는 SNR은 특이값에 따라 u_i 와 v_i 의 전송 순위를 정함으로써 효과적으로 조절할 수 있다. 우리는 SNR 계층 부호화를 위해 비트 평면 부호화 기법을 제안한다. 비트 평면 부호화에서는 데이터 x 를 다음과 같이 이진화 한다.

$$x = s + tb^{(1)} + \frac{t}{2}b^{(2)} + \dots + \frac{t}{2^{r-1}}b^{(r)}, \quad b^{(r)} \in \{0,1\}. \quad (7)$$

여기서 s 와 t 는 x 의 범위에 따라 변하는 상수이며, 구체적으로 x 의 범위는 $[s, s+2t)$ 로 주어진다. 또한 $b^{(p)}$ 은 p 번째 비트로써 0 혹은 1의 값을 가진다.

본 연구에서는 기저 벡터 u_i, v_i 를 (7)의 이진화 공식에 의해 r 개의 비트 평면 $u_i^{(p)}$ 와 $v_i^{(p)}$ 로 분해한다($1 \leq p \leq r$).

[10]에서와 같이, 기저 벡터들의 원소들이 범위 $[s, s+2t)$ 안에서 균일한 분포를 갖는다고 가정하면, 첫 번째 비트 평면 $u_i^{(1)}$ 와 $v_i^{(1)}$ 을 보냈을 때의 복원된 동영상의 평균 왜곡은 특이값 σ_i 에 의해 가중화된 $\sigma_i^2 \frac{t^2}{12}$ 가 됨을 유도할 수 있다. 즉

큰 특이값에 해당하는 기저 벡터들의 오차는 복원된 동영상의 화질 왜곡에 더욱 심각한 영향을 미친다. 이때 그림 3은 [10]에서 제안한 왜곡 모델을 사용하여 예측한 결과와 실제 양자화 왜곡을 비교한 그래프이다. 균일한 분포를 가진다고 가정한 이후 정의한 왜곡 모델이 실제와 유사함을

알 수 있다. 다음 비트 평면이 전송되면 $\frac{3}{4} \left(\sigma_i^2 \frac{t^2}{12} \right)$ 의 왜곡이 감소하여 평균 왜곡은 $\frac{1}{4} \left(\sigma_i^2 \frac{t^2}{12} \right)$ 이다. 일반적으로 p 번째 비트 평면들까지 보냈을 때의 복원된 동영상의 평균 왜곡은 다음 식으로 주어진다. 즉 한 비트 평면이 전송됨에 따라 왜곡은 $\frac{1}{4}$ 의 비율로 감소한다. 이와 같이 i 번째 기저

벡터의 원소들이 범위 $[s, s+2t)$ 안에서 균일한 분포를 갖는다고 가정하면, 첫 번째 비트 평면 $u_i^{(1)}$ 와 $v_i^{(1)}$ 을 보냈을 때의 복원된 동영상의 평균 왜곡은 특이값 σ_i 에 의해 가중화된 $\sigma_i^2 \frac{t^2}{12}$ 가 됨을 유도할 수 있다. 즉

큰 특이값에 해당하는 기저 벡터들의 오차는 복원된 동영상의 화질 왜곡에 더욱 심각한 영향을 미친다. 이때 그림 3은 [10]에서 제안한 왜곡 모델을 사용하여 예측한 결과와 실제 양자화 왜곡을 비교한 그래프이다. 균일한 분포를 가진다고 가정한 이후 정의한 왜곡 모델이 실제와 유사함을

알 수 있다. 다음 비트 평면이 전송되면 $\frac{3}{4} \left(\sigma_i^2 \frac{t^2}{12} \right)$ 의 왜곡이 감소하여 평균 왜곡은 $\frac{1}{4} \left(\sigma_i^2 \frac{t^2}{12} \right)$ 이다. 일반적으로 p 번째 비트 평면들까지 보냈을 때의 복원된 동영상의 평균 왜곡은 다음 식으로 주어진다. 즉 한 비트 평면이 전송됨에 따라 왜곡은 $\frac{1}{4}$ 의 비율로 감소한다. 이와 같이 i 번째 기저

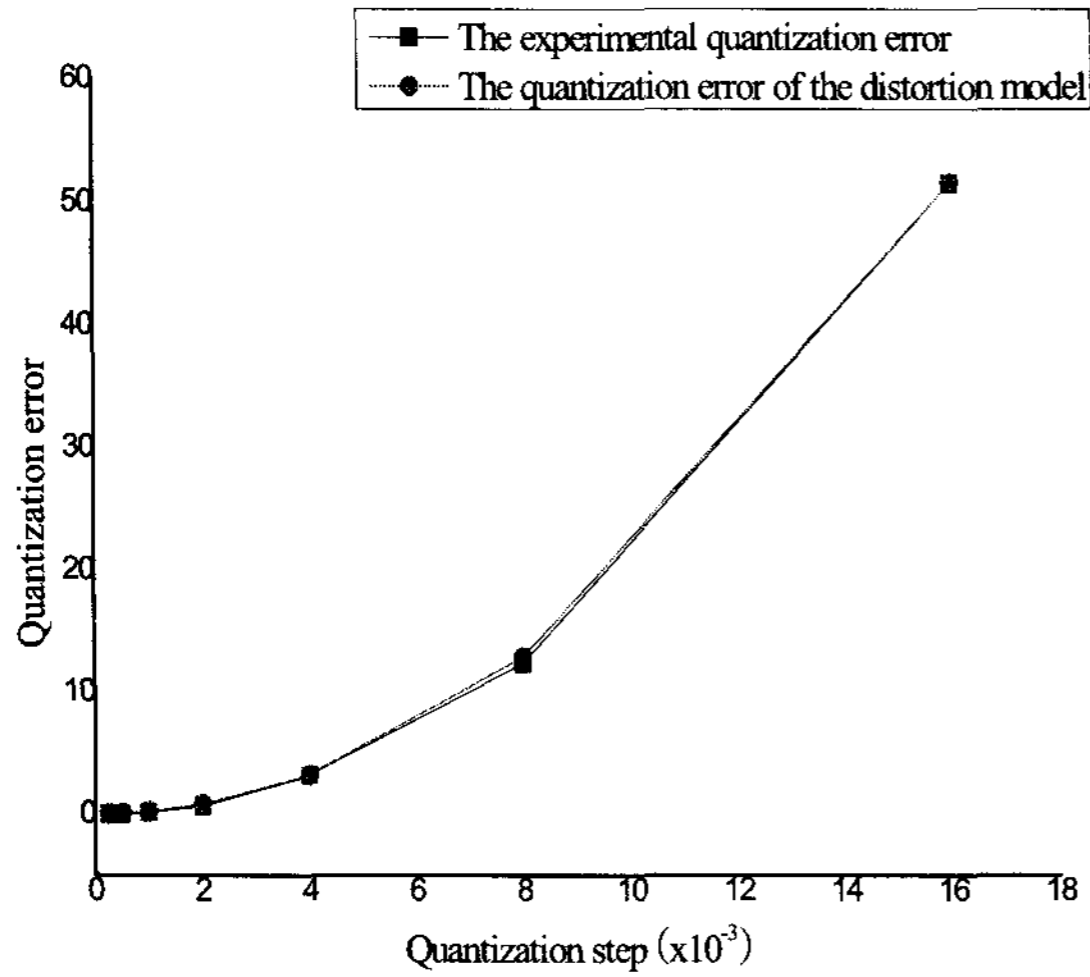


그림 3. Heu 등^[10]의 기법에서 제안된 왜곡 모델과 실제 양자화 왜곡 비교
 Fig. 3. The comparison predicted distortion by Heu *et al.* to experimental distortion

$$D_i^{(p)} = \frac{1}{4^{p-1}} \left(\sigma_i^2 \frac{t^2}{12} \right). \quad (8)$$

벡터의 p 번째 비트 평면 $\mathbf{u}_i^{(p)}$ 와 $\mathbf{v}_i^{(p)}$ 을 전송할 때 복원된 동영상의 평균 왜곡이 줄어드는 정도를 정량적으로 계산할 수 있다. 따라서 본 연구에서는 모든 비트 평면들 중에서 복호된 동영상의 평균 왜곡을 최대한 감소시키는 비트 평면들을 찾아 전송하는 과정을 반복함으로써 SNR 계층 부호화를 수행한다.

그림 4은 "Bounce Ball" 동영상에 제안하는 비트 평면 부호화를 적용하였을 때, 복호기가 전송된 비트들을 받음

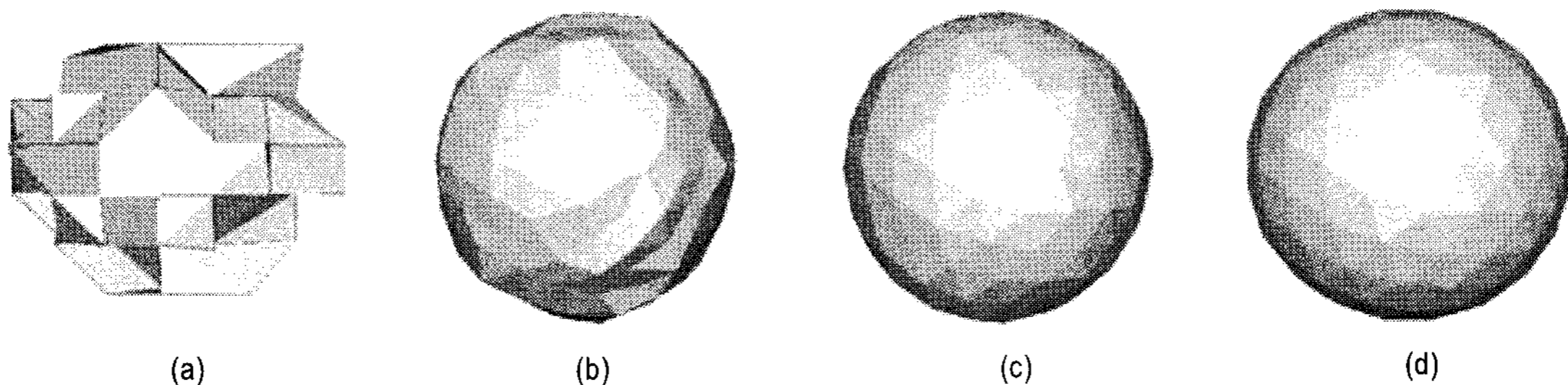


그림 4. "Bounce Ball" 동영상의 1번째 프레임의 SNR 계층 부호화.
 Fig. 4. The SNR scalability of the first frame of the "Bounce Ball" sequence.

에 따라 점진적으로 좋은 화질을 갖는 프레임이 복원됨을 도시한다. 그림 5는 "Bounce Ball" 동영상에 대해서 Heu 등의 기법^[10]과 제안하는 SNR 계층 부호화 기법의 비트량-왜곡 성능을 비교한다. 제안하는 알고리즘이 ^[10]과 거의 유사한 성능을 발휘함을 확인할 수 있다. 그러나 ^[10]은 단일 전송을만을 지원하는 반면에, 제안하는 알고리즘은 하나의 비트열로부터 다양한 화질의 동영상을 복원할 수 있다는 장점이 있다.

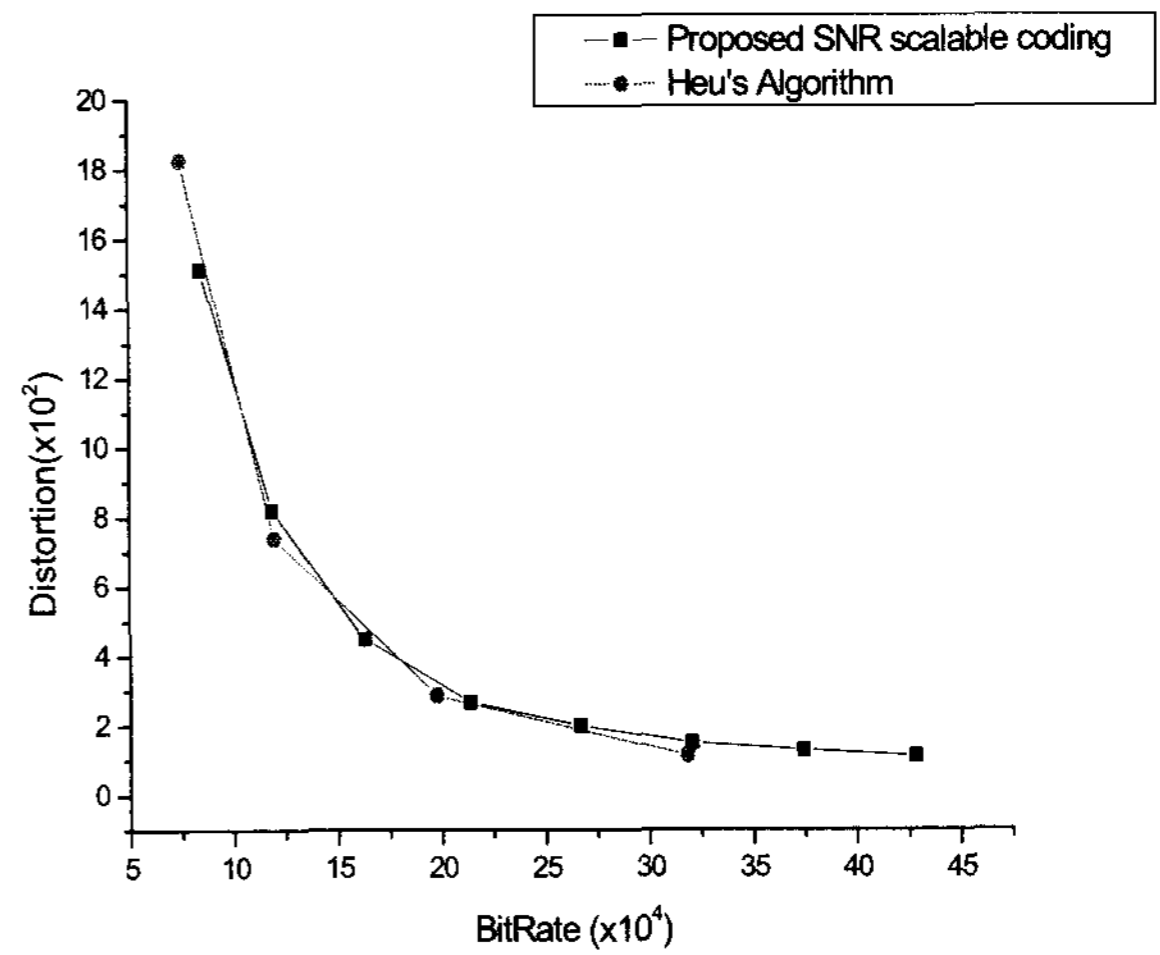


그림 5. Heu 등의 기법^[10]과 제안하는 비트 평면 부호화의 비트량-왜곡 성능 비교

Fig. 5. The RD curves of Heu *et al.* and proposed algorithm

2. 시간 기저 벡터의 예측 기법

일반적으로 동영상에서 인접한 프레임 사이에는 높은 상

관성이 존재한다. II 절에서 살펴봤듯이 3차원 메쉬 동영상은 SVD에 의해 식 (3)과 같이 근사(approximation)될 수 있다. 이때 근사된 j 번째 프레임 A_j 는 다음과 같이 주어진다.

$$A_j = \sum_{i=1}^k \sigma_i \mathbf{u}_i v_{i,j} \quad (9)$$

여기서 $v_{i,j}$ 는 시간 기저 벡터 v_i 의 j 번째 원소를 의미한다. 따라서 동영상에 존재하는 프레임 사이의 상관성은 시간 기저 벡터에서도 존재한다. 이를 이용하여 그림 6과 같이 기저 프레임 A_j 와 A_{j+2} 를 정의하고, 사이의 프레임 A_{j+1} 을 다음과 같이 선형 보간법(linear interpolation)에 의해 예측함으로써 시간 잉여 정보를 효율적으로 제거할 수

있다.

$$A_{j+1}^o = \frac{1}{2}(A_j + A_{j+2}) = \frac{1}{2} \sum_{i=1}^k \sigma_i \mathbf{u}_i (v_{i,j} + v_{i,j+2}) \quad (10)$$

기저 프레임에 해당하는 시간 기저 벡터 $v_{i,j}$ 와 $v_{i,j+2}$ 에는 III.1 절에서 서술한 비트 평면 부호화를 직접 적용하고, 세밀화 프레임에 해당하는 $v_{i,j+1}$ 는 (10)의 예측을 수행한 후 예측 오차에 비트 평면 부호화를 적용하여 이진 부호화한다.

그림 7의 예에서 기저 프레임 A_1 에 해당하는 시간 기저 벡터 $v_{i,1}$ 에 임계치 t 로 시작하는 비트 평면 부호화를 적용하면, 첫 번째 비트 $v_{i,1}^{(1)}$ 을 전송할 경우 $\sigma_i^2 \frac{t^2}{12}$ 의 평균 왜곡

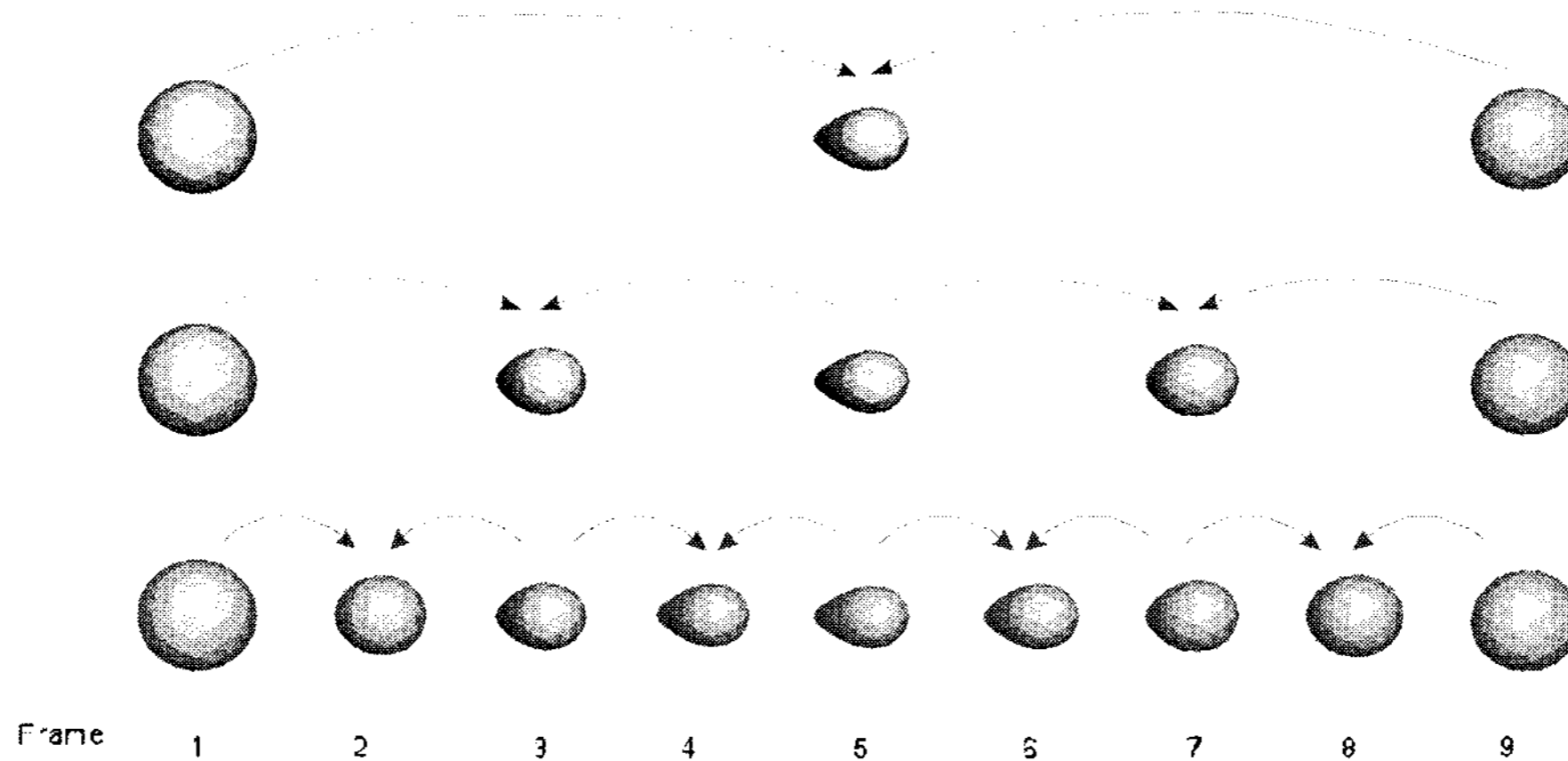


그림 6. 시간 기저 벡터의 예측
Fig. 6. The prediction method of temporal basis vectors

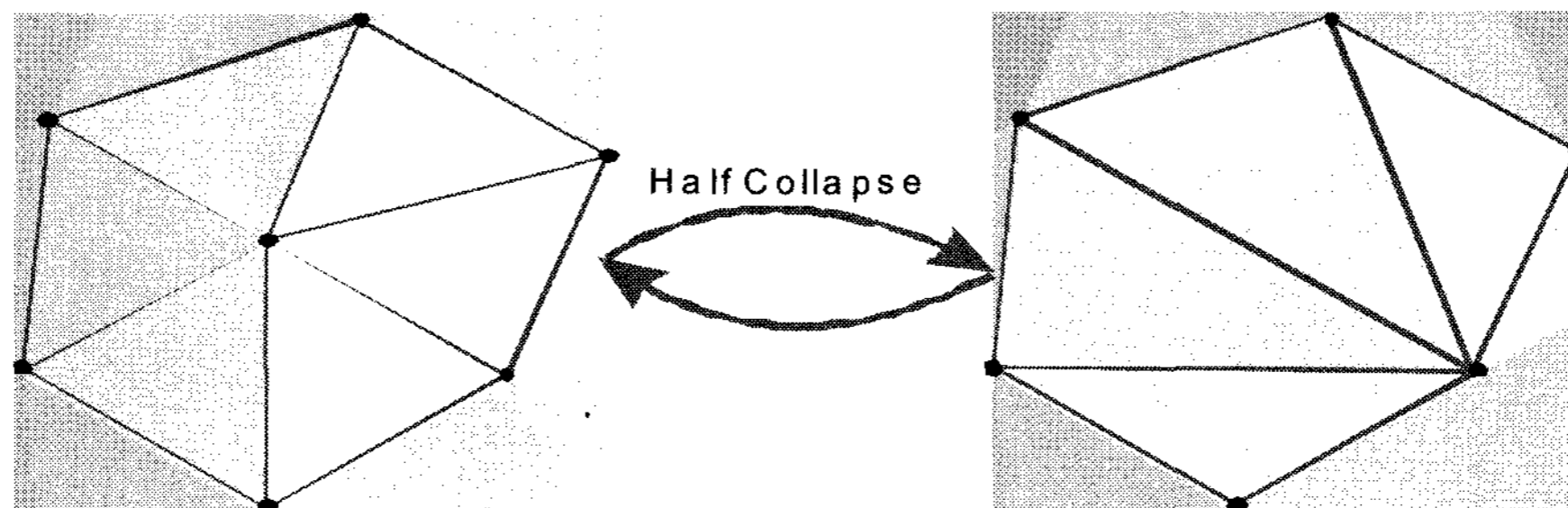


그림 7. 하프 컬랩스를 이용한 간략화.
Fig. 7. Example of the half collapse.

이 복원 데이터에 발생한다. 그러나 기저 프레임 A_1 은 그림 6에서 볼 수 있듯이 다른 중간 프레임들의 예측에 사용되므로 A_1 의 왜곡은 중간 프레임들로 전파된다. 그러므로 그림 6의 예에서 첫 번째 비트 $V_{i,1}^{(1)}$ 를 전송할 경우 전체 동영상에 반영되는 왜곡은 다음과 같이 주어진다. 우선 A_1 은 $\sigma_i^2 \frac{t^2}{12}$ 의 왜곡을 갖는다. 또한 A_5 는 A_1 과 A_9 를 사용하여 선형 보간법에 의해 예측 되므로 A_1 에 발생한 왜곡이 $\sigma_i^2 \frac{t^2}{12} \cdot \frac{1}{2}$ 로 파생된다. 이와 같이 순차적으로 예측 되는 프레임들에 왜곡이 전파되어 전체 동영상에 나타나는 왜곡은 다음 식과 주어진다.

$$D_i^{(1)} = \sigma_i^2 \frac{t^2}{12} + \sigma_i^2 \frac{t^2}{12} \cdot \frac{1}{2} + (\sigma_i^2 \frac{t^2}{12} + \sigma_i^2 \frac{t^2}{12} \cdot \frac{1}{2}) \frac{1}{2} + \sigma_i^2 \frac{t^2}{12} \cdot \frac{1}{2} \cdot \frac{1}{2} \quad (11)$$

이에 반해 마지막 세밀화 단계에 해당하는 프레임 A_2 의 시간 기저 벡터의 경우에는 $V_{i,2}^{(1)}$ 에 의해 $\sigma_i^2 \frac{t^2}{12}$ 의 왜곡만이 발생하며 다른 프레임으로 왜곡이 전파되지 않는다.

이와 같이 시간 기저 벡터의 예측 기법이 사용될 경우에 각 비트 평면이 전체 동영상의 왜곡에 미치는 영향을 계산할 수 있다. 제안하는 알고리즘은 이런 정량적인 데이터에 기반하여 비트량-왜곡 성능을 최적화 하는 전송 순서를 결정하여 전송한다.

3. 공간 기저 벡터의 예측

본 연구에서는 공간상의 잉여 정보를 제거하기 위하여 메쉬 웨이블릿 변환^[11]을 이용한다. 제안하는 공간 기저 벡터의 예측을 위해서 우선 하프 컬랩스(half collapse)를 사용하여 간략화(simplification) 한다. 하프 컬랩스는 그림 7에 보인 바와 같이 두 개의 꼭지점을 하나로 결합할 때, 새로운 꼭지점을 생성하지 않고 기존의 꼭지점을 사용한다. 또한 [12]에 제안된 배치(batch) 알고리즘을 적용하여 각 단계에서 여러 개의 꼭지점들을 동시에 간략화하는 과정을 수행한다.

이상의 간략화 과정 또한 SVD의 기저 벡터들과 특이값들로 표현할 수 있다. 우선 j 번째 꼭지점의 생략에 의한 간략화는 해당 공간 기저 벡터 $u_{1,j}, u_{2,j}, \dots, u_{k,j}$ 를 생략함으로써 실행할 수 있다. 생략된 꼭지점을 p_0 이라고 하면 이를 그림 8에 도시한 바와 같이 주변의 꼭지점들을 이용하여 우산 예측(umbrella prediction)할 수 있다.

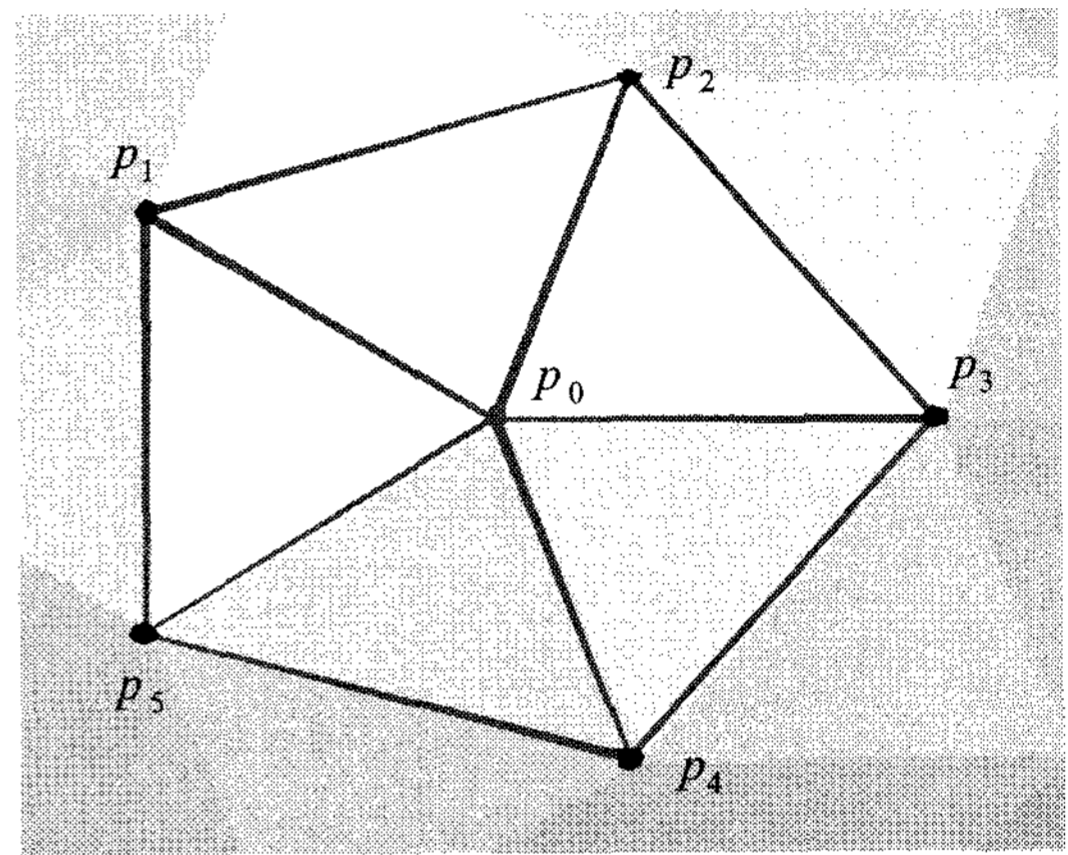


그림 8. 공간 계층 부호화에 사용되는 우산 예측(umbrella prediction).
Fig. 8. The umbrella prediction method.

$$\hat{p}_0 = \frac{1}{5} \sum_{i=1}^5 p_i \quad (12)$$

배치 간략화 과정에서 발생하는 기저 메쉬 데이터는 비트 평면 부호화를 직접 적용하고, 세밀화 메쉬는 예측 오차에 비트 평면 부호화를 적용한다. 시간 기저 벡터의 예측과 유사하게 기저 메쉬와 세밀화 메쉬에 해당하는 데이터는 서로 다른 중요성을 가진다. 즉 각 꼭지점이 예측에 많이 사용될수록 그 왜곡은 전체 동영상에 더 심각한 영향을 미치게 된다. 이러한 영향도를 식 (11)과 유사한 방법으로 수치를 계산할 수 있다. 그러나 간략화 과정에서 사용되는 예측 구조가 이웃 꼭지점의 개수에 영향을 받고, 꼭지점들의 연결 구조가 정형화 되어 있지 않으므로 왜곡이 전파되는 정확한 수치를 구하기 위해서는 매우 많은 계산량이 필요하다. 따라서 정확한 수치 대신에 실험을 통하여 평균적인 대표값을 사용한다.

4. 부가 정보

부호기는 복호기에 기저 벡터들의 비트 평면 이외에 다음의 정보를 추가로 전송한다. 우선 특이값을 전송한다. 또한 메쉬의 연결 정보 및 간략화 정보를 전송한다. 추가로 전송되는 정보들은 전체 데이터 크기에 비하여 매우 미미한 데이터를 차지한다.

이와 같은 부가 정보를 수신 받으면 복호기는 부호기와 동일한 방법으로 각 비트 평면의 중요도를 계산하여 현재 수신 받는 비트 평면이 어떤 기저 벡터의 몇 번째 비트 평면에 해당하는지를 결정할 수 있다.

IV. 모의 실험 결과

본 절에서는 제안하는 SNR 계층 부호화 기법의 성능을 기존의 SVD에 기반한 기법^[10]과 비교한다. 실험에 사용된 동영상은 329 프레임과 162 꼭지점으로 이루어진 "Bounce Ball" 과 1,000 프레임, 539 꼭지점으로 구성된 "Face" 동영상이다.

SVD에 의해 분해된 기저 벡터의 경우 정규직교 하기 때문에 식 (7)의 상수 t 는 0.5로 설정한다. 객관적인 비교를 위해 왜곡은 아래와 같이 평균 자승 오차 (mean squared error)를 사용한다.

$$MSE = \frac{|A - \hat{A}|^2}{3n \cdot f} \quad (13)$$

이때, \hat{A} 는 복원 데이터이고, n 은 꼭지점의 개수를 의미하며 f 는 프레임의 개수를 의미한다.

그림 9와 그림 10은 각각 "Bounce Ball"과 "Face"에 대한 비트량-왜곡 성능을 제시한다. 제안하는 기법이 기존의 기법^[10]에 비하여 우수한 성능을 발휘함을 확인할 수 있다. 특히 적은 비트량을 전송할 경우 제안하는 알고리즘이 기존의 기법에 비하여 훨씬 좋은 성능을 발휘하는 데 이는 제안하는 기법이 기저 벡터를 간략화한 후 예측함으로써 잉여

정보를 효과적으로 제거하기 때문이다.

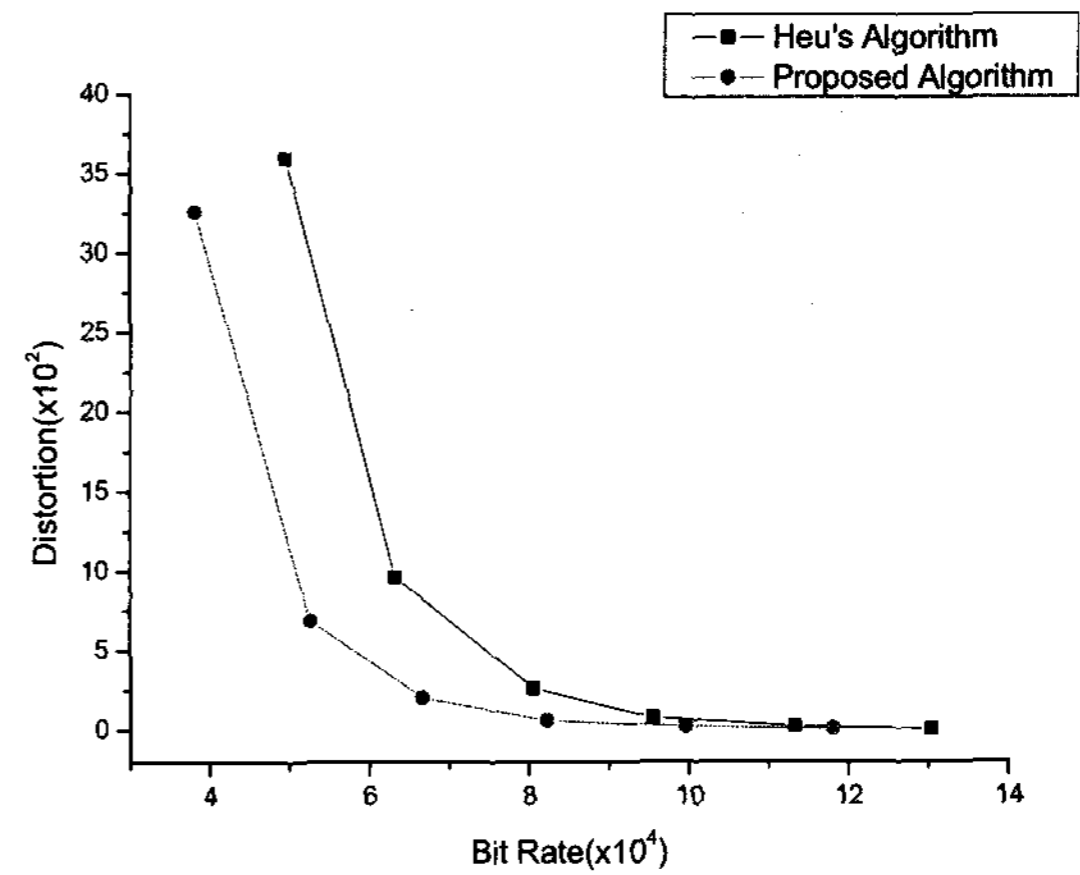


그림 9. "Bounce Ball" 동영상에 대한 비트량-왜곡 성능 비교.
Fig. 9. The RD performance of "Bounce Ball" sequence.

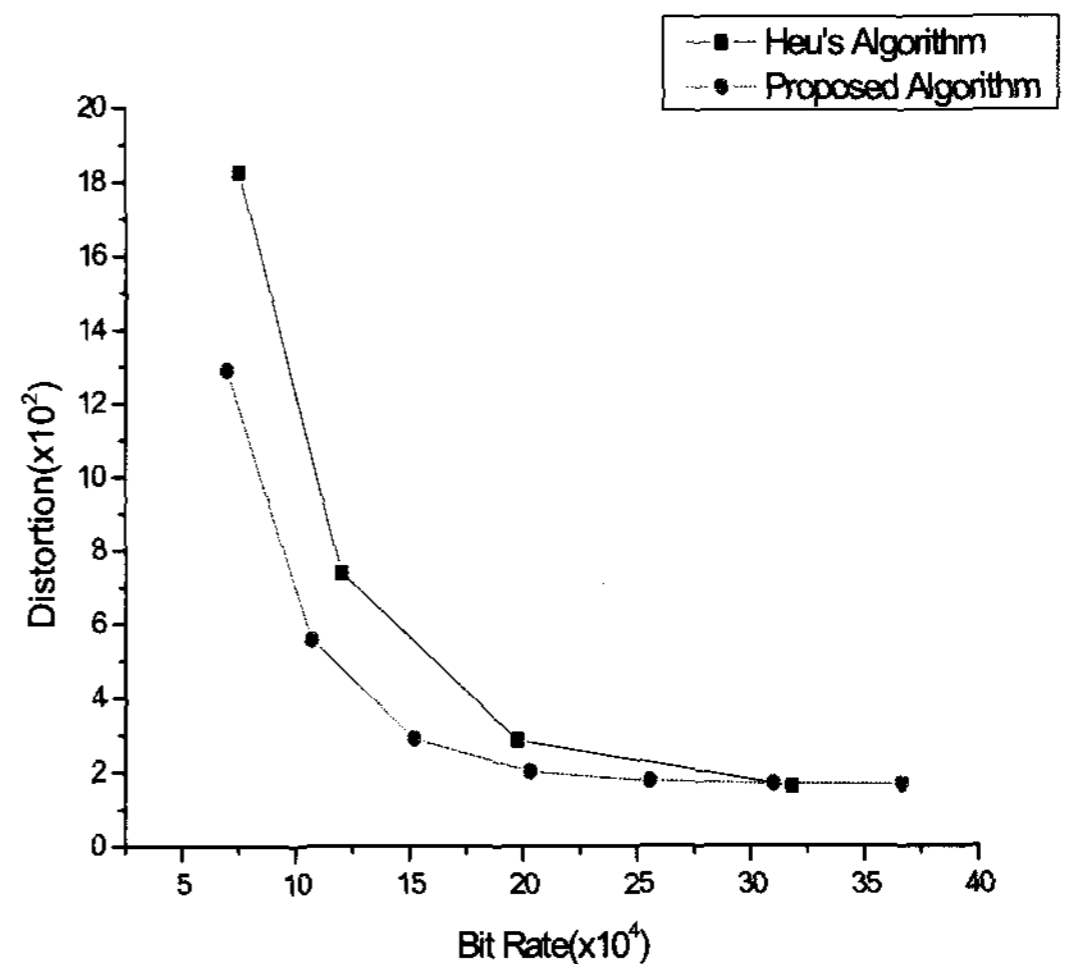


그림 10. "Face" 동영상에 대한 비트량-왜곡 성능 비교.
Fig. 10. The RD performance of "Face" sequence.

또한 기존의 기법이 단순히 SVD만을 적용하여 단일 전송을 지원하는 것에 비하여, 제안하는 다양한 화질을 하나의 비트열로부터 점진적으로 제공할 수 있는 장점이 있다.

V. 결론

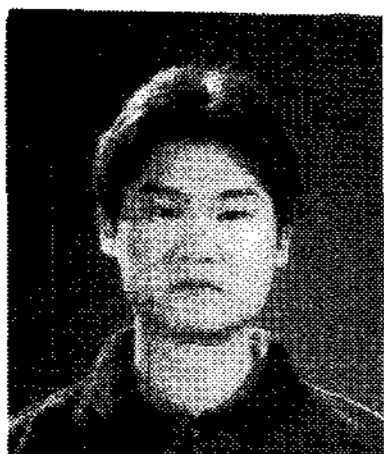
본 연구에서는 SVD에 기반한 3차원 메쉬 동영상의 효율

적인 계층 부호화 기법을 제안하였다. 우선 각 기저 벡터들과 화질 사이의 관계를 살펴보고 각 기저 벡터를 시공간 예측을 통하여 기저 단계와 세밀화 단계들로 분해하였다. RD 측면에서 우수한 부호화 성능을 얻기 위해 특이값에 따른 기저 단계 비트 평면과 세밀화 평면의 중요도를 구하였다. 제안하는 알고리즘은 이러한 중요도에 따라 부호화 순서를 정의함으로써, 기존의 기법에 비해서 우수한 비트량-왜곡 성능을 발휘하였다. 실험 결과에서 볼 수 있듯이, 약 70,000 비트량에서 2.5배의 우수한 MSE를 나타내고 있다. 또한 제안하는 알고리즘이 계층적 복호화를 지원함으로써 게임, 애니메이션 등에 많은 활용이 가능할 것으로 기대한다.

참 고 문 헌

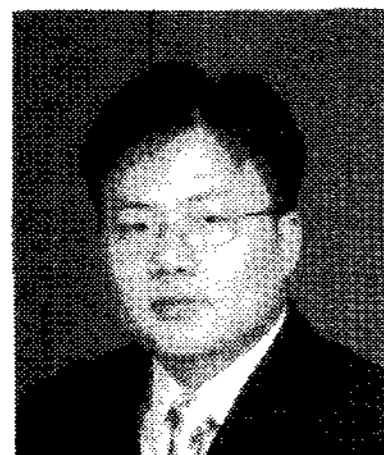
[1] Autodesk, [http:// www.autodesk.com/3dsmax](http://www.autodesk.com/3dsmax).
 [2] Autodesk, [http:// www.autodesk.com/maya](http://www.autodesk.com/maya).
 [3] J. E. Lengyel, "Compression of time-dependent geometry," in Proc. ACM Symposium on Interactive 3D Graphics, pp. 89-96, July 1999.
 [4] J. H. Yang, C. S. Kim, and S. U. Lee, "Compression of 3D triangle mesh sequences based on vertex-wise motion vector prediction," IEEE Trans. Circuits and System for Video Technology, vol. 12, no. 12, pp. 1179-1184, Dec. 2002.
 [5] L. Ibarria, and J. Rossignac, "Dynapack: Space-time compression of the 3D animations of triangle meshes with fixed connectivity," in Proc. Eurographics/SIGGRAPH Symp. Computer Animation 2003, pp. 126-135, July 2003.
 [6] E. S. Jang, D. K. Kim, S. Y. Jung, M. J. Han, S. O. Woo, and S. J. Lee, "Interpolator Data Compression for MPEG4 Animation," IEEE Trans. Circuits and System for Video Technology, vol. 14, no. 7, pp. 989-1008, July 2004.
 [7] ISO/IEC 14 772-1, The Virtual Reality Modeling Language, 1997.
 [8] M. Alexa, and W. Müller, "Representing animations by principle components," Eurographics 2000, pp. 291-301, June 2000.
 [9] Z. Karni, and C. Gotsman, "Compression of soft-body animation sequences," Computers & Graphics, vol. 2004, no. 28, pp. 25-34, June 2004.
 [10] J. -H. Heu, C. - S. Kim, J. - H. Yang, and S. - U. Lee, "Compression of 3D mesh sequences based on R-D optimized principal component analysis," IEE Electronics letters, vol. 42, no. 14, pp. 799-800, July 2006.
 [11] I. Guskov and A. Khodakovsky, "Wavelet compression of parametrically coherent mesh sequence," in Proc. SIGGRAPH, pp. 183-192, Aug. 1994.
 [12] R. Pajarola and J. Rossignac, "Compressed progressive mesh," IEEE Trans. Visualization and Computer Graphics, vol. 6, no. 1, pp. 79-93, Jan. 2000.

저 자 소 개



허 준 회

- 2003년 2월 : 홍익대학교 전기공학부 졸업(학사)
- 2005년 2월 : 서울대학교 전기공학부 졸업(석사)
- 2005년 3월 ~ 현재 : 서울대학교 전기공학부 박사과정
- 주관심분야 : 영상 신호 처리, 3D 그래픽스



김 창 수

- 1994년 2월 : 서울대학교 전기공학부 졸업(학사)
- 1996년 2월 : 서울대학교 전기공학부 졸업(석사)
- 2000년 8월 : 서울대학교 전기공학부 졸업(박사)
- 2000년 7월 ~ 2001년 11월 : University of Southern California, Visiting Scholar
- 2001년12월 ~ 2003년 7월 : 서울대학교 전기공학부 박사후연구원
- 2003년 8월 ~ 2003년 7월 : The Chinese University of Hong Kong, Assistant Professor
- 2005년 8월 ~ 현재 : 고려대학교 전기전자전파공학부, 부교수
- 주관심분야 : 영상처리, 비디오 부호화, 3차원 모델 부호화

저 자 소 개



이 상 옥

- 1973년 2월 : 서울대학교 졸업(학사)
- 1976년 5월 : Iowa State University 석사
- 1980년 2월 : University of Southern California 박사
- 1983년 3월 ~ 현재 : 서울대학교 전기공학부 정교수
- 2007년 1월 ~ 현재 : IEEE Fellow
- 주관심분야 : 신호 처리, 비디오 부호화, 컴퓨터 비전