2000年度 第13回 信號處理合同學術大會 論文集 第13卷 1號

스켈레톤을 이용한 삼차원 체적소 데이터의 부호화

송인욱, 김창수, 이상욱

서울대학교 전기공학부

A Voxel Data Compression Using Skeleton

In-Wook Song, Chang-Su Kim, Sang-Uk Lee

School of Electrical Engineering
Seoul National University

e-mail : iwook@ipl.snu.ac.kr, cskim@ieee.org, sanguk@sting.snu.ac.kr

요 약

3차원 영상은 2차원 영상에 비해 대이터량이 매우 방대하다. 따라서 3차원 데이터를 효율적으로 압축하는 것은 매우 중요하다. 현재까지 대부분의 연구는 데이터량이 체적소(voxel)에 비해 월등히 적은 메쉬(mesh)를 기반으로 하여 이루어져 왔다. 하지만, 메쉬를 이용한 데이터 압축의 경우 체적소에 비해 데이터 자체의 규칙성이 떨어져 체적소를 이용한 압축에 비해 압축 효율이 낮다. 그러고, 체적소 데이터를 이용할 경우, 이를 스켈레톤화 하여 데이터량을 더욱 줄일 수 있다. 따라서 본 논문에서는 3차원 체적소 데이터의 규칙성과 스켈레톤을 이용한 압축 기법을 제안할 것이다.

그림 1: 체적소 영상

제 1 절 서론

3차원 영상의 경우 메쉬(mesh) 기반의 부호화 기법과 체적소(voxel) 기반의 부호화 기법이 제안되고 있으나, 대부분의 경우 메쉬 기반으로 활발한 연구가 진행되어왔다[1, 2, 3]. 또한, 데이터 취득과 저장시 야기되는 하드웨어적 문제로 대부분의 3차원 영상이 메쉬로 이루어져 있다. 또한 일반적으로 메쉬 데이터량이 체적소에 비해 월등히 적어서 빠른 처리와 부호화가 가능하다. 그러나, 그림 1과 그림 2를 비교하면 체적소 기반의 3차원 영상이 메쉬 기반의 영상에 비해 우수한 화질을 보여주고, 체적소의 많은 데이터량에도 불구하고 효율적이 부호화 기법으로 메쉬에 비해 높은 압축률을 얻을 수 있으며, 특히 하드웨어의 발달로 체적소 데이터의 빠른 처리가 가능해짐에 따라 체적소 기반의 부호화 기법이 점차 연구되는 추세이다[4, 5].

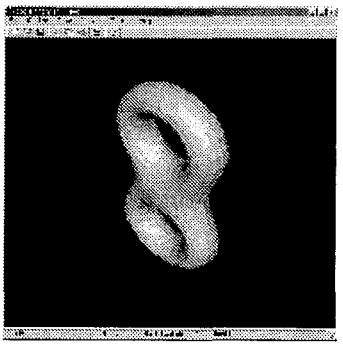


그림 2: 메쉬 영상

체적소를 이용한 3차원 영상의 표현에서 가장 문제가 되는 것은 방대한 데이터 양이다. 비록 하드웨어 상 많은 발전으로 이를 처리하는 것이 가능해졌다고 하지만, 이를 압축하는 것은 매우 중요한 문

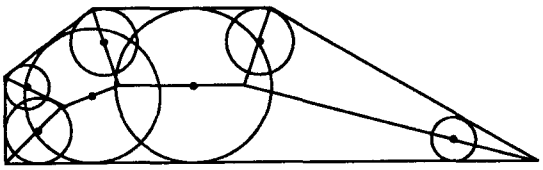


그림 3: Definition 1에 의한 스켈레톤

제이다. 메쉬 데이터와는 달리 체적소 데이터의 경우 삼차원 격자의 규칙성을 이용할 수 있어 압축 효율을 매우 높일 수 있다. 하지만 현재까지 체적소를 이용한 3차원 데이터 표현 및 부호화에 관한 연구는 매우 미비한 상태이다.

본 논문에서는 스켈레톤을 이용한 체적소 데이터 부호화 기법을 제안한다. 3차원 영상을 스켈레톤으로 표현했을 경우 영상을 표현하는데 필요한 체적소의 수가 매우 줄어든다. 실제로 3차원 물체를 스켈레톤화 할 경우 필요한 체적소의 개수를 1/2에서 1/10정도로 줄일 수 있다. 하지만, 스켈레톤의 경우 각 체적소에 스켈레톤 반지름이라는 부가 정보가 필요로 한다. 각 스켈레톤 체적소당 이 정보가 필요해서 부가 정보로 인한 비트율 증가가 매우 커지게 된다. 하지만 디스크 C(n)을 이용하여 스켈레톤을 생성했을 경우 자손이 되는 스켈레톤 점들은 반지름이 같다는 점[6]을 이용하면 대부분의 체적소에 스켈레톤 반지름 정보를 필요로 하지 않게 된다. 즉, 부가 정보로 인한 비트율의 증가는 감소시킬 수 있다.

이 논문은 다음과 같이 구성되어있다. 제2절에서는 스켈레톤의 개요에 대해 기술하고 제3절에 제안한 알고리듬에 대해서 기술하고 제4절에 모의실험을 통한 성능을 나타내고, 제5절에 결론을 제시한다.

제 2 절 스켈레톤의 개요

스켈레톤은 Blum[7] 에 의해 처음 제안 되었다. Blum에 의해 제한된 스켈레톤은 매우 직관적이여서 수많은 수학적 모델이 나오게 되었다. 그 중 가장 보편적으로 사용되는 정의는 다음과 같다.[그림 3,4]

Definition 1 *주어진 형상 $X \subset \mathbf{R}^2$에 대해서 X에 내접하는 디스크를 최대 디스크(maximal disc)라 하자. 이 때, 형상 X의 스켈레톤은 모든 최대 디스크의 중심점 들의 집합이다.*

[8]에서 Lantuéjoul은 기하학적으로 열린 형상 $X \subset \mathbf{R}^2$의 스켈레톤 $S(X)$가 이진 형태 연산(morphological operation)으로 계산될 수 있음을 증명하였다. 반지름이 $r > 0$인 최대 디스크(maximal discs)의 중심점들을 포함하는 집합 $S_r(X)$는,

Formula 1 (Lantuéjoul's formula)

$$S_r(X) = X \ominus rB - [(X \ominus rB) \circ drB] \qquad (1)$$

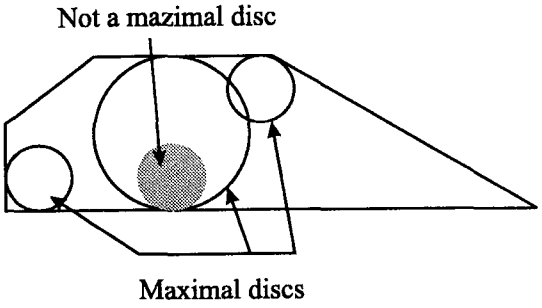


그림 4: 최대 디스크(Maximal Disc)

이다. 여기서 rB 와 drB는 각각 반지름이 r인 위상학적으로 열린 디스크이고, 무한소 반지름이 dr인 위상학적으로 닫힌 디스크이다. 형상 X의 스켈레톤 $S(X)$는 $r > 0$인 모든 $S_r(X)$의 합집합이다.

[9]에서 Serra는 Lantuéjoul's formula의 이산 형식을 나타내었다. [10]에서는 이를 generalized-step skeleton이라 명명하였다.
유클리디안 공간 E에서 $\{B(n)\}_{n\in\mathbf{N}}$ 기하학적으로 열린 형상자라 하자. 이 형상자를 이용하여 $\{A(n)\}_{n\in\mathbf{N}}$을 아래와 같은 방법으로 생성해 낸다.

$$\begin{cases} A(0) = (0,0) \\ A(n+1) = A(n) \oplus B(n), \quad n = 0,1,2,\cdots. \end{cases} \qquad (2)$$

$\{A(n)\}$은 스켈레톤의 정의(Definition 1)에서 말한 "디스크"를 의미한다.

Generalized-step skeleton $S_n(X)$은,

$$S_n(X) = X \ominus A(n) - [X \ominus A(n)] \circ B(n), \quad n = 0,1,\cdots. \qquad (3)$$

이다. 집합 $S_n(X), n \in \mathbf{N}$, 과 (3)은 스켈레톤 하위 집합과 Lantuéjoul 공식을 일반화 한 것이다. $S_n(X)$는 연속 영상과 이산 영상에 모두 적용될 수 있다. 식(3)으로부터 구해진 스켈레톤의 복원 공식은,

$$X \circ A(k) = \bigcup_{n \geq k} S_n(X) \oplus A(n) \qquad (4)$$

이다.
간략화를 위해 X_n, Y_n을 각각,

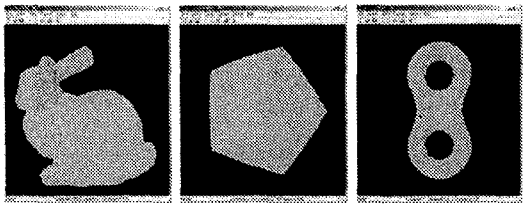
$$X_n \equiv X \ominus A(n) \qquad (5)$$

$$Y_{n+1} \equiv X_{n+1} \oplus B(n) \qquad (6)$$

이라하면,
generalized-step skeleton S_n은,

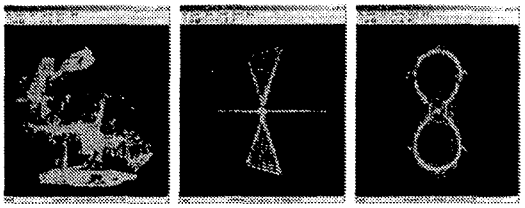
$$S_n = X_n - Y_{n+1} \qquad (7)$$

이다. S_n은 식(3)의 이산 형태이다.



(a)bunny (b)icosahedron (c)eight

그림 5: 실험에 사용된 영상



(a)bunny (b)icosahedron (c)eight

그림 6: 실험에 사용된 영상의 스켈레톤

제 3 절 제안한 방법

스켈레톤을 이용해 형상을 복원하는데는 각 스켈레톤 체적소의 반지름 정보가 필요하다. 이들 반지름 정보 중 이전 반지름 정보와 스켈레톤 체적소 사이의 관계를 이용해 찾아낼 수 있는 것들이 있다. 식 (5)에서 $X_0 = X$이다. 식(7)에서 $X_n = S_n + Y_{n+1}$이다. $n = 0$인 경우, $X_0 = S_0 + Y_1$이다. 따라서, 원 형상 X를 복원하기 위해서는 반지름이 0인 스켈레톤 체적소 정보 S_0와 Y_1정보가 필요하다. Y_1은 X_1으로 부터 알아낼 수 있다[식(6)]. 식(7)에서 $S_N = X_N - Y_{N+1}$이다. 여기서 $Y_{N+1} \equiv 0$이다[식(6)]. 따라서 반지름이 N인 스켈레톤 정보로 부터 형상 X를 순차적으로 복원할 수 있다. 각 반지름이 같은 스켈레톤 정보를 순차적, 즉 반지름이 큰 것 부터 차례로 부호화 한다면 최대 반지름 스켈레톤 체적소 정보만 부호화 하면 나머지 스켈레톤 체적소의 반지름 정보는 부호화 하지 않아도 형상 X의 복원이 가능하다. 일반적인 내부가 채워진 3차원 형상 X에 대해 반지름이 $3 \times 3 \times 3$인 디스크로 스켈레톤을 생성하는 것에 대해 생각해 보자. 내부가 채워져 있으므로 반지름이 n인 스켈레톤 체적소는 Y_{n+1}과 8-connected 되어 있다. 그리고, Y_{n+1}은 X_{n+1}을 dilation한 것이다. 따라서 스켈레톤 체적소 좌표 정보만 부호화를 한 후 최대 반지름 좌표만을 부호화 해 준다. 즉, 복호기 에서 알 수 있는 정보는 스켈레톤 체적소 좌표 정보와 이들 중 최대 반지름을 갖는 체적소의 좌표이다. 이를 통해 복호화 하는 과정은 다음과 같다. 식(7)에서 $X_N = S_N$이므로 식(6)에서 Y_N을 알 수 있다. S_{N-1}은 Y_N과 8-connected 되어 있으므로 스켈레톤 좌표 정보만으로 S_{N-1}을 알아낼 수 있다. S_{N-1}과 Y_N을 알고 있으므로 X_{N-1}을 식(7)으로부터 알 수 있다. 이 과정을 재귀적으로 반복하면 X_0, 즉 형상 X를 복원 할 수 있다. 스켈레톤 체적소를 "1"그렇지 않은 체적소를 "0"으로 표현하면 영상 X를 스켈레톤화 했을 경우 "0"이 매우 많은 영상을 얻게 된다. 따라서 이를 래스터 스캔 방식으로 스캔을 했을 경우 연속적인 "0"이 매우 많은 비트열이 나타나게 된다. 따라서 이를 산술부호화[11] 할 경우 매우 좋은 성능을 나타내게 된다.

제 4 절 실험 결과

실험은 $128 \times 128 \times 128$ 해상도의 영상에 대해서 수행하였다. 이는 메쉬 데이터의 경우 7bit resolution을 가진 영상과 같은 해상도이다. 실험에서 사용된 영상은 그림 5과 같다. 그리고, 이들 각각의 스켈레톤은 그림 6와 같다. 이들 두 그림을 비교해 보면 알 수 있듯이 원영상에 비해 스켈레톤 데이터의 양이 매우 적음을 알 수 있다. 체적소 개수만을 단순히 비교하면 스켈레톤 체적소의 개수가 원영상의 체적소 개수의 $1/2 \sim 1/10$정도이다. 이들 각각의 영상에 대한 압축율을 보면 표1 와 같다. 표1에서 알 수 있듯이 압축율이 매우 높다. 이는 7bit resolution을 가지는 메쉬 데이터를 부호화[3] 하는 경우에 대략 0.13bpv(bits per vertex)인 것을 가만하면 매우 높은 압축율 임을 알 수 있다.

표 1: 실험 결과

3D Image	bpv(bits per voxel)
bunny	0.04
eight	0.02
icosahedron	0.03

제 5 절 결론

이 논문에서 스켈레톤을 이용한 3차원 체적소 데이터 부호화 기법을 제안하였다. 스켈레톤 간의 연관관계를 이용한 반지름 정보 추출 방법과 원영상에 비해 스켈레톤 영상의 체적소 수가 적음을 이용하였다. 체적소 데이터의 가장 큰 단점은 해상도가 높아 지게 되면 메쉬 데이터에 데이터량이 매우 크게 늘어나는 점이다. 따라서 높은 해상도에서는 메쉬 데이터를 이용하고 낮은 해상도에서는 체적소 데이터를 이용하는 것이 바람직하다.

참고 문헌

[1] M Deering , " Geometry Compression, Computer Graphics," Proc. SIGGRAPH '95, pp. 13-20, Aug 1995.

[2] S. Gumhold and W. Strasser, "Real Time Compression of Triangle Mesh Connectivity," Proc. ACM SIGGRAPH '98, pp. 133-140, July 1998

[3] G. Taubin and J. Rossignac, "Geometric Compression through Topological Surgery," ACM Trans. Graphics, vol. 17, no. 2, pp. 84-115, Apr. 1998.

[4] Girod, B.; Eisert, P.; Magnor, M.; Steinbach, E.; Wiegand, T. "3-D image models and compression: synthetic hybrid or natural fit?" Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on , Volume: 2 , 1999 Page(s): 525 -529 vol.2

[5] Eisert, P.; Steinbach, E.; Girod, B. "Multi-hypothesis, volumetric reconstruction of 3-D objects from multiple calibrated camera views" Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on , Volume: 6 , 1999

[6] R. Kresch, "Morphological image representation for coding appli-cations," D.Sc. dissertation, Dept. Electr. Eng., Technion-Israel Ins. Technol., Haifa, Israel, June 1995. Available WWW: http://www-sipl.technion.ac.il /publications /thesis /renato /renato.html

[7] H. Blum, "A transformation for extracting new descriptors of shape," in Models for the Perception of Speech and Visual Forms, W. Wathen-Dunn, Ed. Cambridge, MA: MIT Press, 1967.

[8] C. Lantuéjoul, "La squelettization et son applicationaux mesures topologuiques des mosaiques polycristallines," Ph. D dissertation, School of Mines, Paris, France, 1978

[9] J. Serra, Ed., Image Analysis and Mathematical Morphology. New York: Academic, 1982

[10] D. Schonfeld and J. Goutsias, "Morphological representation of discrete and binary images," Trrans. Signal Processing, vol. 39, pp. 1369-1379 June 1991

[11] P.G. Howard and J. S. Vitter, "Arithmetic coding for data compression," in Proc. IEEE, vol 82, pp. 857-865, June 1994.