

객체 적응적인 정점 기반 윤곽선 부호화 기법

조성중, 홍민철, 한헌수
숭실대학교 정보통신전자공학부
서울시 동작구 상도5동 1-1

A Vertex Based Coding Technique Adaptive to Object's Shape

Sung-Joong Cho, Min-Cheol Hong, Hernsoo Hahn
School of Electronics Engineering, Soongsil University
E-mail : chosj@visionlab.ssu.ac.kr

Abstract

This paper presents a new approach to the vertex based shape coding technique. The conventional approaches encode objects using a spline method with the same distortion coefficients. The proposed approach, however, classifies the objects based on the object's features, and then applies different distortion values depending on the classified object types. Using this pre-classifying technique, this paper reduces the bit rate and the computational complexity necessary for the encoding process. The performance of the proposed method has been proved by experiments on the various sample images.

I. 서론

객체기반의 영상 부호화의 가장 큰 특징이라면 블록화 현상(blocking effect)[1]을 줄여 좋은 화질의 동영상 부호화 할 수 있다는 것과 객체의 추출 및 삽입이 용이하여, 영상의 재사용성을 높일 수 있다는 것이다.[2] MPEG-4 규격은 객체를 기반으로 하는 대표적인 동영상 부호화의 예이다. 영상객체를 부호화하기 위해서는 객체의 움직임, 질감, 형상을 가장 적절한 방법을 이용하여 부호화해야 한다.[3]

본 논문은 정점을 기반으로 하는 객체 형상 부호화에 중점을 두고 있다. 기존의 정점기반의 형상부호화 방법은 일률적인 계수값과 spline 방법을 적용하여 부호화하고 있다. 본 논문에서는 객체의 성질을 구분하

여 부호화 하므로써 비트량과 부호화 계산량을 줄일 수 있고 결과적으로 전체적인 부호화 효율을 높일 수 있는 방법을 소개하고 있다. 객체는 직선 성분이 많아서 직선에 민감하지 않은 인공물과 불규칙한 곡선으로 이루어져 있어 직선이 주관적인 시각에 매우 잘 띄는 자연물로 나뉘게 되고, 각각에 서로 다른 왜곡치를 적용하여 정점을 찾음으로써 비트량을 줄일 수 있으며, 계산량이 다른 과정에 비해 매우 큰 스플라인 근사화를 적용적으로 사용함으로써 전체적인 계산 시간량을 줄일 수 있다. 실험을 통하여 여러 객체에 적용하여 그 효율성을 확인하였다.

II. 윤곽선 부호화(Shape Coding)

객체의 형태는 화질과 직결되며 객체의 부호화에 있어서 기본이 되는 가장 중요한 요소이다. 형태정보는 정보량이 가장 많을 뿐만 아니라 인간의 시각적 환경에 가장 민감하게 적용되는 부분이다. 특히 동영상에서 영상 객체를 인식하는 데는 형태정보에 대부분을 의지하게 된다. 그래서 형태정보에 대한 부호화 및 처리 과정에 대한 다양한 방법들이 연구되어져 오고 있다.[3][4]

2.1 내용기반 영상 부호화의 개념

내용 기반 영상 부호화는 다음과 같이 적용된다. 생성된 객체의 형상, 질감, 움직임 정보를 추출하여 각각에 적합한 최적 부호화 방식을 이용하여 부호화 한 뒤 모든

정보를 일괄하여 보내는 것이다. 이러한 과정은 그림 1에서 소개하고 있다.

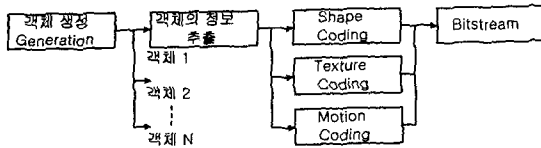


그림 1. 내용기반의 영상부호화를 위한 과정.

2.2 정점 기반의 윤곽선 부호화(Vertex Based Shape Coding)

정점 기반의 형태 부호화의 방법은 다음과 같다. 객체의 윤곽선을 이루는 점들 사이에서 가장 먼 거리를 가진 두 정점(V_1, V_2)을 찾아내고 그 두 점 사이의 윤곽선 중에서 수직 최장거리(d_{max})를 갖는 정점을 찾게 된다. 기존의 정점과 새로운 정점은 새로운 두 개의 선분을 만들게 되고, 새로운 선분에 대하여 앞의 과정을 되풀이하게 된다. Index k 에서의 d_{max} 값은 두 점(V_k, V_{k+1})로 이루어진 선분과 이 사이의 윤곽선에 해당하는 구간의 점 $c_{k,i}$ 의 수직 직선거리의 최대값을 나타낸다.

$$d_{max}(k) = \max_{0 \leq i < I} d(V_k V_{k+1}, c_{k,i})$$

이렇게 하여 이미 정해진 왜곡치(d_{max^*})이내에 들어올 때까지 ($d_{max} < d_{max^*}$) 이 과정을 되풀이하여 정점을 찾게 된다.

자연적인 영상에서는 그림 2에서 보는 것과 같이 다각형 왜곡이 확연히 드러남을 알 수 있다. 다각형 근사화를 통해 자연스러운 영상을 만들기 위해서는 미세한 d_{max^*} 값을 이용해야 한다. 자연스러운 영상을 위해서는 B-Spline 근사화 방식을 사용하게 되는데, B-Spline 방식은 3개의 제어점(control points)을 이용하여 2차식을 만든 후 객체의 외곽선을 근사화 시킨다. 이 방법의 경우에 쓰이는 왜곡치는 원 영상과 만들어진 영상의 최장 수직 거리 값인 d_{max} 뿐만이 아니라 mean square error 등의 통계적 왜곡치를 쓰기도 한다.

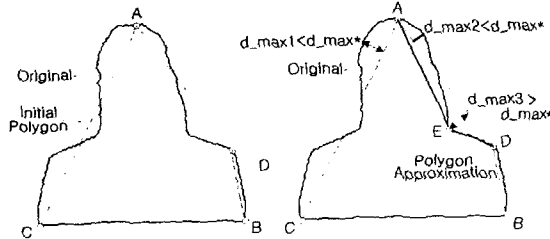


그림 2. 다각형 근사화.

원영상과의 에러값들을 최소로 하면서 곡선으로 근사화 시키는 방법에는 많은 방법들이 있지만, 계산값들을 고려할 때 많은 높은 차수의 함수를 사용할 수는 없다. 현재 제어점 3개로 2차식을 통하여 각 정점들을 근사화 하는 B-Spline 방식을 사용하고 있다. 그림 3은 Spline을 이용하여 근사화된 윤곽선을 보여준다.

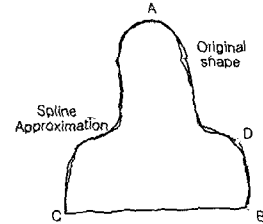


그림 3. Spline 근사화.

III. 제안하는 정점 부호화 방법

본 논문에서는 shape의 정보량을 객체의 성질에 따라 적용적으로 부호화하므로써 기존의 방법보다 주관적 화질의 열화를 최소로 하면서 적은 정보량으로 효과적인 부호화를 할 수 있는 방법을 제안한다. 다음 그림 4는 기존의 정점 기반의 부호화 방식을 나타내고 있다.[6][7]



그림 4. 기존의 정점 기반 윤곽선 부호화.

객체는 인위적으로 만들어져서, 그 외곽선의 형태가 직선과 정형적인 곡선의 형태를 취하는 인공물과, 불규칙하게 이루어지는 자연물로 구분할 수 있다. 예를 들자면 전자에 빌딩, 자동차 등을 후자에 사람, 나무, 새 등을 들 수가 있다.

이러한 객체를 분류의 단계는 형태 부호화하기 전 단계에서 선행되어야 한다. 분류된 객체는 주관적인 화질의 범위를 넘지 않는 범위에서 객체에 적합한 왜곡치 범위를 적용함으로써 효율적인 부호화가 가능하다.

다시 말해서, 왜곡이 커져도 주관적인 화질의 저하가 없는 객체(인공물)를 분류하므로 더 작은 비트량을 이용하여 전과 유사한 화질을 유지할 수 있다. 그림 5는 제안된 정점 기반의 부호화 방식을 나타내고 있다.

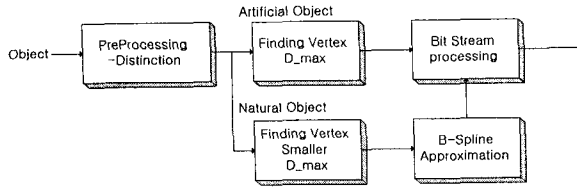


그림 5. 제안된 정점기반의 윤곽선부호화.

3.1 전처리 과정

윤곽선의 요소 중에서 윤곽선의 전체 길이의 1% 이상이 되는 직선의 요소를 구하고 그 직선 요소들의 전체 평균을 구하여 객체의 직선화도 값으로 삼는다. 윤곽선을 트래킹하여 직선에 해당되는 점들의 합을 구할 수 있다. 이미 윤곽선 추출 단계에서 윤곽선의 인덱스와 좌표값을 알고 있기 때문에 복잡한 절차 없이 직선화도를 구할 수 있다.

3.2 Adaptive d_max

인공물 이미지의 경우는 왜곡치를 크게 하여도 주관적인 화질의 열화가 심하지 않다. 이유는 직선요소가 주관적인 시각에 많이 띄지 않기 때문이며, 반대로 자연적인 이미지의 경우 기본적으로 d_max 값을 미세하게 조절해야 해야 주관적인 화질을 유지할 수 있다. 본 실험에서는 인공물은 d_max 3 pixel 자연물인 경우 1 pixel로 정했다. d_max 값을 정한 배경에는 기존의 논문에서 적용해온 기준을 따랐다. CIF 비디오 포맷에서 자연스러운 영상을 위해서는 보통 1.2-1.4 pixel 값을 사용하였다. 이러한 선 실험 기준치와 실험자의 주관적인 판단을 통하여 위와 같은 기준치를 구했으며, 이 값은 더 나은 판단 과정을 거치면서 바뀔 수 있다.

3.2 Spline 근사화와 계산량 (computational complexity)

각 과정의 계산량은 알고리즘에 비추어 쉽게 산정된다. 전처리 과정에서는 $O(N)$ 이 된다. N은 윤곽선의 포인트 개수이다. 다각형 근사화(Hotter's Algorithm)[3]의 단계에서 요구되는 complexity는 $O(N^2)$ 이며, Spline Approximation (schuster & Katsagellos)[1]는 왜곡치인 d_max를 구할 경우 평균 $N/2$ 개의 픽셀에 대해 계산을 하게 된다.(최대값을 찾는 것은 평균적으로 중간 지점에서 찾게 되기 때문이다) 시작점을 찾고, 왜곡치를 구하는 일련의 과정이 포함된 이 과정의

complexity는 $O(N^4)$ 이 된다. Spline 방식의 계산량은 다른 과정들을 무시 할 만큼 매우 크다.

IV. 영상 객체를 이용한 실험

먼저 Akyo 와 Plane의 두 영상을 통하여 본 논문의 방법을 적용해 보았다.

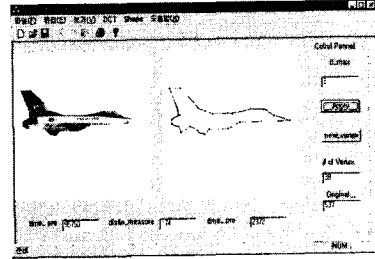


그림 6. Plane : d_max* = 3.
직선화도 : 3.558(%)

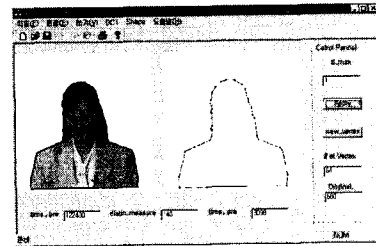


그림 7. Akyo : d_max* = 1.
직선화도 : 2.794(%)

	d_max*	Number of Vertices	Original Boundary	PSNR (dB)
Akyo (Natural)	3	30	680	35.03
	1.4	64	680	40.63
Plane (Artificial)	3	38	537	28.41
	1.4	50	537	32.33

표 1. 왜곡치를 달리 했을 때 두 객체의 개수 비교.

위 표에서 d_max*값을 달리 했을 경우 찾아지는 정점의 개수와 PSNR 값의 실험치를 나타내고 있다. 정점의 개수는 부호화 비트량의 정도를 나타내고 최소의 정점의 수로 윤곽선을 효율적으로 표현하는 주관적인 방법이 필요하다.

Akyo 와 Plane 영상에서 객체의 등장율이 50%라 가정했을 때 10.5 % bit 가량의 비트량의 이득을 가져올 수 있었다. 계산량에 있어서는 Spline 방식이 다른 과정에 비해 비교할 수 없을 만큼의 많은 계산량을 보이기 때문에 객체의 등장율이 50%라 가정했을 경우

계산량은 반으로 줄어드는 것을 알 수가 있다. Worst case인 경우에도 기존의 방법 이상의 비트량이나 계산량을 요구하지는 않는다. 객체에 따라 주관적인 화질의 열화가 있을 수 있으나, 얻어지는 효율에 비한다면 감수 할 수 있는 손실이며, trade-off를 통해 최선의 기준 왜곡치와 판별 기준치를 찾는 것도 앞으로 남은 과제일 것이다.

다음은 동영상 객체에서 추출된 다양한 객체들을 이용하여 직선화도를 구하고 객체를 판별하는 실험을 하였다. 실험 결과 평균 직선화도가 3%가 넘는 객체일 경우 인공물로, 3% 이내의 값을 갖는 객체를 자연물 객체로 분류할 수 있었다. 그림 8은 실험에 사용된 다양한 객체와 그 윤곽선을 나타냈으며, 각각의 직선화도를 표시하고 있다.

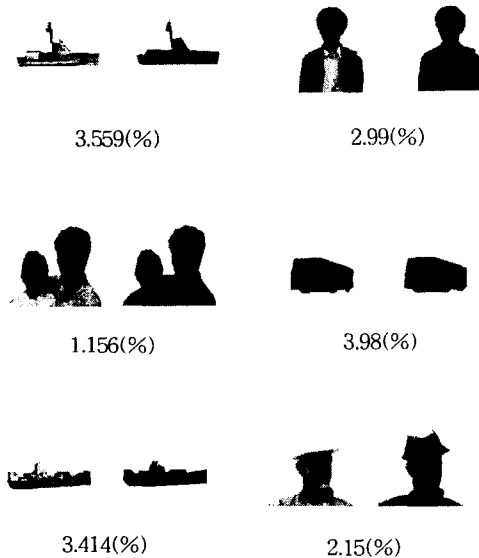


그림 8. 다양한 객체와 직선화도.

V. 결론

본 논문에서는 기존의 내용기반의 동영상 부호화에 사용되어지는 윤곽선 부호화에 적용될 수 있는 적용적인 윤곽선 부호화에 대하여 논하였다. 이 방법은 기존의 방법에 비해 계산량 및 비트량에 있어 부호화 효율을 증가시킬 수 있었다. 이것은 MPEG4의 실제적인 부호화 처리과정에 삽입이 가능할 것으로 사려되며, 객체에 의미와 정보를 부여할 수 있었던 중요한 시도라고 할 수 있다. 이러한 접근은 영상을 검색하고 데이터베이스화 할 수 있는 MPEG7과의 중간적인 형

태이고, 영상을 포함한 객체의 부호화 기법에 있어서 새로운 시도가 될 것이다.

Reference

- [1] H. Musmann, M. Hotter, and J. Ostermann, "Object-oriented analysis-synthesis coding of moving images", *Signal Process*, vol. 1, pp. 117-138, Oct. 1989.
- [2] M. Kunt, A. Ikononopoulos, and M. Kocher, "Second generation image coding technique", *Proc. IEEE*, vol. 73, pp. 549-575, Apr. 1985.
- [3] Aggelos K. Katsaggelos, Lisimachos P. Kondi, Fabian W. Meier, Jörn Ostermann, Guido M. Schuster, "Mpeg-4 and Rate-Distortion-Based Shape-Coding Technique", *Proc. IEEE*, vol. 86, no. 6, June 1998.
- [4] S. Lee, D. Cho, S. Son, E. Jang, and J. Shin, "Binary shape coding using 1-D distance values from baseline", in *Proc. Special Session on Shape Coding, ICIP97*, Santa Barbara, CA, pp. 508-511, 1997.
- [5] N. Brady, F. Bossen, and N. Murphy, "Context-based arithmetic encoding of 2D shape sequences", in *Proc. Special Session on Shape Coding, ICIP97*, Santa Barbara, CA, pp. 504-507, 1997.
- [6] K. J. O'Connell, "Object-adaptive vertex-based shape coding method", *IEEE Trans. Circuit Syst. Video Technol.*, vol. 7, pp. 251-255, Feb. 1997.
- [7] Michael J. DeHaemer, Jr. and Michael J. Zyda, "Simplification of Objects Rendered by Polygonal Approximations", *Computer & Graphics*, vol. 15, no. 2, Great Britain: Pergamon Press, pp. 175-184, 1991.