

# QP(Quadratic Programming) 방법을 이용한 객체단위의 영상압축 알고리즘

최 유 태, 이 상 업, 광 대 호, 김 시 내, 송 문 호  
고려대학교 전자공학과  
전화 : 02-927-0560 / Fax : 02-927-0544

## Object Based Image Compression Using QP (Quadratic Programming) Method

Yu Tae Choi, Sang Youp Lee, Dae Ho Kwak, Kim Si Nae, and S. Moon Ho Song  
Dept. of Radio Sciences and Engineering, Korea University  
E-mail : ytchoi@bessel.korea.ac.kr

### Abstract

The object level image compression is a useful technology for reducing the necessary data and manipulating individual objects. In this paper, we propose a new image object compression algorithm that uses the quadratic programming (QP) method to reduce the compressed data. The results indicate the superiority of the proposed QP based algorithm over the low pass extrapolation (LPE) method of MPEG-4.

### I. 서론

동영상 압축 기술은 멀티미디어 통신이나 차세대 개인용 휴대통신 시스템과 HDTV 등을

구현하는 데 반드시 필요로 하는 핵심적인 기술이다. 음성이나 텍스트의 경우와는 달리 동영상을 전송하는 경우에는 상대적으로 넓은 대역을 필요로 한다. 이런 이유로 인해서 동영상 통신 시스템의 보편화 과정에서는 동영상 압축 기술의 발전을 반드시 필요로 하고 있다. 그 간의 영상압축 기술은 MPEG-1, MPEG-2 H.261, H.263 등의 표준을 거치면서 현재 MPEG-4에 까지 이르게 되었다. 이 MPEG-4가 기존의 영상압축 표준과 구별되는 가장 큰 특징은 영상압축 과정이 객체단위로 이루어진다는 점이다. 다시 말하면 영상을 전송할 때 이미지에서 배경, 인물, 물건 등이 각각 독립적인 객체로서 따로따로 전송을 하게 된다.

본 논문의 핵심내용은 QP(Quadratic Programming) 방법을 토대로한 객체 단위의 영상 압축 알고리즘이다. 더 나아가 제안하는 QP 방식을 기존의 객체기반 압축방법 중 MPEG-4에서 시험모델로 채택된 LPE(Low Pass Extrapolation)방식과 비교 분석하여 QP

방법의 우수성을 보일 것이다.

## II. Quadratic Programming Method

QP 방법은 비 선형적인 방정식의 해를 찾는 한 방법으로 일반적인 QP 방정식의 형태는 다음과 같다.

$$\text{minimize } \frac{1}{2} X^T Q X \quad (1)$$

$$\text{subject to } a_i^T X = b_i, \quad i \in E$$

$$c_i^T X \leq d_i, \quad i \in I$$

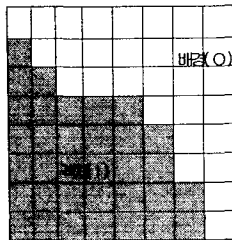
$Q$  : symmetric and positive semidefinite matrix

$E, I$  : Index Sets for equality and inequality constraints

여기에서는  $a$ 와  $b$ 를 통하여 적절한 constraint를 줌으로써  $X^T Q X$ 를 최소화하는  $X$ 를 찾게 된다.

## III. QP의 영상에의 적용

영상의 압축방법에는 오랜동안 사용되어온 DCT 기반의 영상압축과 최근에 활발히 연구되어지고 있는 wavelet에 기반을 둔 영상압축방법 등이 있다. QP 방법은 DCT에 기반을 둔 압축방법으로 8 x 8 블록을 기준으로 압축을 수행한다. 다시 말하면 기존의 영상 전체에 걸쳐서 행하는 압축방법과는 달리 객체기반 압축 방법은 객체내부와 객체 외부는 원래의 이미지 값이 그대로 전송되므로 압축을 수행하지 않으며 객체의 경계면에서만 압축을 수행한다.



<그림 1> 영상에서의 객체와 배경

영상의 경계면에 있는 한 8 x 8 블록을  $f$ 라 하고, 경계면을 중심으로 객체의 내부와 외부로

<그림 1>과 같이 나눈다. 위의 배경(O)에 적절한 값을 채워넣은 후 DCT basis matrix와의 내적을 통하여 주파수 도메인의 고주파의 값을 배경 pixel의 개수만큼 "0"으로 만들면 Run Length Coding 후 출력될 데이터의 양을 훨씬 줄일 수 있을 것이다. 즉 같은 영상정보를 전송하더라도 적은 양의 데이터만으로 전송이 가능할 것이다.

주파수 영역의 성분을 zig-zag scan 하기 위한 lookup matrix(L)를 아래와 같이 정의한다.

$$L = \begin{bmatrix} 0 & 1 & 5 & 6 & 14 & 15 & 27 & 28 \\ 2 & 4 & 7 & 13 & 16 & 26 & 29 & 42 \\ 3 & 8 & 12 & 17 & 25 & 30 & 41 & 43 \\ 9 & 11 & 18 & 24 & 31 & 40 & 44 & 53 \\ 10 & 19 & 23 & 32 & 39 & 45 & 52 & 54 \\ 20 & 22 & 33 & 38 & 46 & 51 & 55 & 60 \\ 21 & 34 & 37 & 47 & 50 & 56 & 59 & 61 \\ 35 & 36 & 48 & 49 & 57 & 58 & 62 & 63 \end{bmatrix}$$

DCT basis matrix( $\Phi$ )는 위의 lookup matrix를 이용하여 64개의 matrix로 정의 할 수 있다.

$$\Phi_{L(u,v)}(i,j) = a_u a_v \cos \frac{\pi(2i+1)u}{16} \cdot \cos \frac{\pi(2j+1)v}{16} \quad (u,v,i,j=0,1,2, \dots, 7) \quad (2)$$

$$a_u = \begin{cases} 1/\sqrt{8}, & u=0 \\ \sqrt{2}/8, & 1 \leq u \leq 7 \end{cases}, \quad a_v = \begin{cases} 1/\sqrt{8}, & v=0 \\ \sqrt{2}/8, & 1 \leq v \leq 7 \end{cases}$$

위의 basis matrix 는 주파수 영역에서의 각 영상의 DCT 계수를 의미한다.

DCT basis matrix와 영상의 내적은 다음과 같다.

$$\langle f, \Phi_t \rangle = F_t, \quad 0 \leq t \leq 63 \quad (3)$$

$f$  : original image

$F_t$  :  $f$ 의 DCT의  $t$ 번째 DCT 계수

위 식 (3)을 영상의 내부(I)와 외부(O)로 나누어 표현을 하면,

$$F_t = \langle f_i, \Phi_{iI} \rangle + \langle f_o, \Phi_{iO} \rangle$$

$f_i$  : 객체내부의 pixel만 선택한  $(64-N) \times 1$  vector

$f_o$  :  $N \times 1$  배경 pixel

$\Phi_{iI}$  :  $t$  번째 DCT matrix  $\Phi_t$ 에서 객체 내부의 pixel값을 lexicographically order 한  $(64-N) \times 1$  vector

$\Phi_{iO}$  :  $t$  번째 DCT matrix  $\Phi_t$ 에서 객체 외부의 pixel값을 lexicographically order 한  $N \times 1$  vector

$$\langle f_o, \phi_{i_0} \rangle + \langle f_i, \phi_{i_1} \rangle = \vec{0}, \quad 63 - N < i \leq 63 \quad (4)$$

위 식을 vector 화 하면

$$\phi \vec{x} + \vec{s} = \vec{0} \quad (5)$$

$$\phi = \begin{bmatrix} \phi_{63|0}^T \\ \phi_{62|0}^T \\ \vdots \\ \phi_{63-N+1|0}^T \end{bmatrix}$$

$\vec{x}$ : 구하고자 하는 객체외부 픽셀, 즉  $f_o$

$\vec{s}$ : 객체내부에서의 내적 값, 즉  $\langle f_o, \phi_{i_0} \rangle$

위 식 (5)를 만족하는  $\vec{x}$ 는 정확히 한 개 또는 경우에 따라서 많이 있을 수 있다. 그러므로, 여기에서는 다음과 같이 QP 방법을 사용하여 "smooth" 한 객체외부의 pixel 값 즉,  $\vec{x}$ 를 구하고자 한다.

$$\text{minimize } \vec{x}^T Q \vec{x} \quad (6)$$

$$\text{subject to } |\phi \vec{x} + \vec{s}| \leq 0.5 * \vec{m} \quad (7)$$

$\vec{m}$ : Quantization step size vector

$Q$ : Symmetric Positive Semidefinite matrix

위에서 정의한  $Q$ 는 영상에서 객체를 중심으로 픽셀 값이 밖으로 퍼져나가는 효과를 좋게 하기 위해 다음과 같은 matrix를 사용하였다.

$$Q = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

위 식 (7)을 만족하는 X를 찾기 위해서는 일정한 조건의 constraint를 지정해 주어야 한다. 부등식의 constraint를 주게 되면 constraint 안의 여러 개의 해를 찾게 된다. DCT 기반의 영상압축에서는 주파수 영역에서 양자화를 하게 되는 데 식 (6)의 조건을 이용하여 식 (7)와 같이 constraint를 지정하였다.

조건이 성립한다. 식 (7)에서 constraint를 Quantization step size 의 반으로 설정하였는데 그 이유는 DCT 한 후의 값에 quantization matrix로 나눌 때 그 값이 0.5보다 작으면 정수 값 만을 취하는 방법에 의해 "0" 으로 mapping 되므로 최소한 배경 픽셀 수만큼의 고주파 성분을 "0"으로 취하고, 파일의 크기도 줄일 수 있다. 즉, 위의 식 (6)과 식 (7)을 만족하는 X의

해를 을 찾아 배경에 대입하면 주파수 영역에서는 "0" 의 값을 갖는 주파수 성분이 적어도 배경의 개수만큼은 존재하게 된다. 여기서 식 (1)에서도 설명하였지만 부등식이기 때문에 여러 개의 해가 존재하게 된다. 하지만 "0" 에 가까운 해는 또 하나의 객체기반 압축방법의 하나인 zero padding 방법과 같게되기 때문에 선택한 해는 "0"에서 되도록 양의 방향으로 멀리 떨어져 있는 해를 이 논문에서는 선택하였다.

#### IV. Simulation 및 LPE 방법과의 결과 비교

이 논문에 사용된 QP 방법을 사용한 이미지 압축을 테스트하기 위해 <그림 2> (가)와 같은 정지영상을 가지고 먼저 객체를 추출하기 위해 마스크 영상<그림 2> (나)를 먼저 추출하였다.

<그림 2> (나)의 마스크 이미지는 압축시 영상에서의 객체와 배경의 구분을 위해 사용되는 이미지이다. 마스크 영상으로 객체를 구분한 후 배경과 객체는 그대로 두고 배경과 객체의 경계의 8 x 8 블록에서만 QP 알고리즘을 적용하여 배경의 픽셀값을 구한 후의 영상은 <그림 3>과 같다. <그림 3>를 보면 경계 블럭을 중심으로 처리된 배경이미지를 발견 할 수



(가)

(나)

<그림 2> Original 이미지(가), 마스크된 이미지(나)

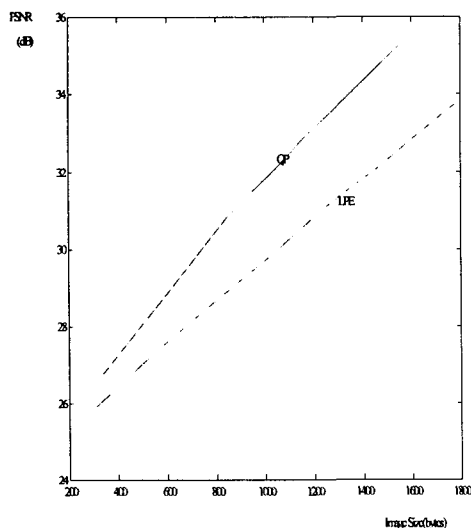
## V. Conclusion

이상에서 살펴 본 바와 같이 본 논문에서 사용된 QP 알고리즘은 기존의 객체기반의 영상압축방법에 기반을 두고 MPEG-4에서 test model로 사용되어지고 있는 LPE 방법에 비해 좋은 성능을 보여주고 있다. 비록 본 논문에서는 정지 영상에 대해서만 적용된 결과이지만 좀 더 심도 있는 연구를 통해 Sequence 영상에 대해서도 좋은 결과를 얻을 수 있으리라고 생각한다.



<그림 3> QP 방법으로 압축된 영상

있을 것이다. 이와 같이 QP를 이용하여 JPEG 영상에 적용하여 원 영상과 PSNR을 구하여 Rate-Distortion Curve를 그려보았다. 비교는 기존의 MPEG-4에서 테스트 모델로 사용되어지고 있는 방식인 LPE와의 비교한 결과를 그림에 표시했으며 사용한 tool은 matlab을 사용했으며 <그림 4>는 시뮬레이션 결과를 보여준다. <그림 4>에서 보여 주고 있듯이 QP를 이용한 압축 방법이 같은 데이터 크기일 경우 크게 2dB 까지 PSNR이 좋은 것으로 나타나고 있다.



<그림 4> QP 와 LPE의 Rate-Distortion Curve

## 감사의 글

본 논문은 학술진흥재단의 연구비 지원으로 작성되었습니다.

## 참고문헌

- [1] P.E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "QP-Based Methods for Large Scale Nonlinearly Constrained Optimization," *Nonlinear Programming*, Academic Press, New York, 1981.
- [2] D. LeGall, MPEG: "A video compression standard for multi-media", *Commun ACM*, vol. 34, pp47-58, April 1991.
- [3] ITU-T Recommendation H.261 Video coding for low bitrate communication, Dec. 1995.
- [4] T. Sikora and B. Makai, Shape-Adaptive DCT for Generic Coding for Video, *IEEE Trans. on Circuit and Systems for Video Tech.*, Vol. 5, No 1., Feb. 1995