이산코사인변환 기반 이미지 압축 핵심 알고리즘 시각적 재구성

진찬용 · 남수태
원광대학교(융복합창의연구소)

A Visual Reconstruction of Core Algorithm for Image Compression Based on the DCT (discrete cosine transform)

Chan-yong Jin · Soo-tai Nam
Wonkwang University (Institute of Convergence and Creativity)
E-mail : cy85366@wku.ac.kr

요 약

JPEG은 가장 널리 사용되는 표준 이미지 압축기술이다. 본 논문에서는 이미지 압축 알고리즘을 소개하고 압축 및 압축 해제의 각 단계를 서술하고자 한다. 이미지 압축은 디지털 이미지를 데이터 압축을 적용하는 과정이다. 이산여현변환은 시간 도메인에서 주파수 도메인으로 변환하는 기술이다. 먼저, 이미지는 8 by 8 픽셀 블록으로 분할하게 된다. 둘째, 위에서 아래로 왼쪽에서 오른쪽으로 진행하면서 DCT가 각각의 블록에 적용하게 된다. 셋째, 각 블록은 양자화를 통해 압축을 진행한다. 넷째, 이미지를 구성하는 압축된 블록의 배열은 크게 줄어든 공간에 저장된다. 끝으로, 원하는 경우 이미지는 역 이산여현변환 (IDCT)을 사용하는 프로세스인 압축 해제를 통해 재구성하게 된다.

ABSTRACT

JPEG is a most widely used standard image compression technology. This research introduces the JPEG image compression algorithm and describes each step in the compression and decompression. Image compression is the application of data compression on digital images. The DCT (discrete cosine transform) is a technique for converting a time domain to a frequency domain. First, the image is divided into 8 by 8 pixel blocks. Second, working from top to bottom left to right, the DCT is applied to each block. Third, each block is compressed through quantization. Fourth, the array of compressed blocks that make up the image is stored in a greatly reduced amount of space. Finally if desired, the image is reconstructed through decompression, a process using IDCT (inverse discrete cosine transform).

키워드

압축 알고리즘, 데이터 압축, 이산코사인변환, 이미지, 양자화

Ⅰ. INTRODUCTION

Recently, compression technology has been actively studied in the field of multimedia. Since the amount of data of the multimedia is large, a lot of storage space and bandwidth are required for storing and transmitting the data. Therefore, it can be said that the use of the compression technique is essential for storing and transmitting at a minimum cost without distorting the image data. Compression of a video signal means compressing data using properties or characteristics of the video signal. There are two types of image signal correlation: spatial correlation and time correlation. Spatial correlation is the similarity between adjacent pixels in the screen.

The time correlation refers to the similarity between pixels in the same position of the current screen and the previous screen. Only the spatial correlation is used for the encoding of the still image, and both the spatial correlation and the temporal correlation are used for coding the motion picture.

Ⅱ. DISCRETE COSINE TRANSFORM

The following is a general overview of the JPEG process. We can get a more comprehensive understanding of the process by looking closely at the JPEG method. First, the image is divided into 8 by 8 pixel blocks. Second. working from top to bottom left to right, the DCT is applied to each block. Third, each block is compressed through quantization. Fourth, the array of compressed blocks that make up the image is stored in a greatly reduced amount of space. Finally if desired, the image is reconstructed through decompression, a process using IDCT (inverse discrete cosine transform). The DCT equation (1, 2) computes the i,jth entry of the DCT of an image. Compression algorithm schemes include predictive coding, transformation coding and vector coding.

$$D(i,j) = \frac{1}{\sqrt{2N}} C(i) C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x,y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right] \quad 1$$

$$C(u) = \left\{ \begin{matrix} \frac{1}{\sqrt{2}} & \text{if } u = 0 \\ 1 & \text{if } u > 0 \end{matrix} \right\} \quad 2$$

p(x,y) is the x,yth element of the image represented by the matrix p. N is the size of the block that the DCT is done on. The equation calculates one entry (i,jth) of the transformed image from the pixel values of the original image matrix. In order to the standard 8 by 8 block which JPEG compression uses, N equals 8 and x and y range from 0 to 7. Thus D(i,j) would be such as in Equation (3).

$$D(i,j) = \frac{1}{4} C(i) C(j) \sum_{x=0}^{7} \sum_{y=0}^{7} p(x,y) \cos\left[\frac{(2x+1)i\pi}{16}\right] \cos\left[\frac{(2y+1)j\pi}{16}\right] \quad 3$$

Because the DCT uses cosine functions, the result matrix depends on horizontal, diagonal, and vertical frequencies. Thus an image black with a lot of change in frequency has a very random looking resulting matrix, while an image matrix of just one color, has a resulting matrix of a large value for the first element and zeroes for the other elements. To obtain the matrix through equation (1), use the following equation. For an 8 by 8 block it results in this matrix T. The first row (i = 1) of the matrix has all the entries equal to $1/\sqrt{8}$ as expected from Equation (4).

$$T_{ij} = \left\{ \begin{matrix} \frac{1}{\sqrt{N}} & \text{if } i > 0 \\ \sqrt{\frac{2}{N}}\, cos\left[\frac{(2j+1)i\Pi}{2N}\right] & \text{if } i > 0 \end{matrix} \right\} \quad 4$$

0.3536	0.3536	0.3536	0.3536	0.3536	0.3536	0.3536	0.3536
0.4904	0.4157	0.2778	0.0975	-0.0975	-0.2778	-0.4157	-0.4904
0.4619	0.1913	-0.1913	-0.4619	-0.4619	-0.1913	0.1913	0.4619
0.4157	-0.0975	-0.4904	-0.2778	0.2778	0.4904	0.0975	-0.4157
0.3536	-0.3536	-0.3536	0.3536	0.3536	-0.3536	-0.3536	0.3536
0.2778	-0.4904	0.0975	0.4157	-0.4157	-0.0975	0.4904	-0.2778
0.1913	-0.4619	0.4619	-0.1913	-0.1913	0.4619	-0.4619	0.1913
0.0975	-0.2778	0.4157	-0.4904	0.4904	-0.4157	0.2778	-0.0975

Table 1. Orthogonal matrix [T]

III. CONCLUSIONS

The discrete Fourier transform is a technique of converting from the time domain to the frequency domain. The discrete cosine transform is based on the discrete Fourier transform. Figure 1 shows the result of visual reconstruction of the discrete Fourier transform using Excel.

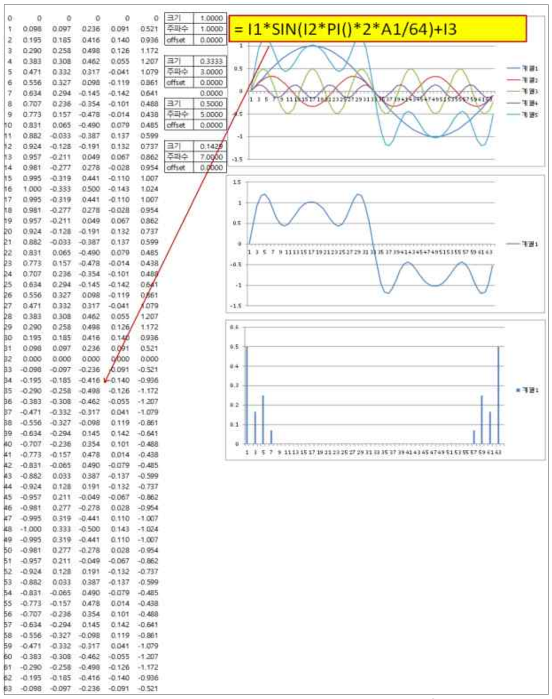

Figure 1. The discrete Fourier transform using Excel of Microsoft

*Please note that references are omitted.