Journal of Information Engineering and Applications
ISSN 2224-5758 (print) ISSN 2224-896X (online)
Vol 1, No.3, 2011
www.iiste.org

3D- Discrete Cosine Transform For Image Compression

Anitha S[1*] Dr. B. S.Nagabhushana[2]

1. Research Scholar, Dr MGR Educational and Research Institute , University, Chennai
2. Lead consultant,KPIT Cummins Infosystems Limited Bangalore, India

* E-mail of the corresponding author: anithasd@yahoo.com

Abstract

Image compression addresses the problem of reducing the amount of data required to represent a digital image called the redundant data. The underlying basis of the reduction process is the removal of redundant data. The redundancy used here is pschychovisual redundancy. 3D-DCT video compression algorithm takes a full-motion digital video stream and divides it into groups of 8 frames. Each group of 8 frames is considered as a three-dimensional image, which includes 2 spatial components and one temporal component. Each frame in the image is divided into 8x8 blocks (like JPEG), forming 8x8x8 cubes. Each 8x8x8 cube is then independently encoded using the 3D-DCT algorithm: 3D-DCT, Quantizer, and Entropy encoder. A 3D DCT is made up of a set of 8 frames at a time. Image compression is one of the processes in image processing which minimizes the size in bytes of a graphics file without degrading the quality of the image to an unacceptable level. The reduction in file size allows more images to be stored in a given amount of disk or memory space.

Keywords: 2D DCT, 3D DCT, JPEG, GIF, CR

1. Introduction

Interest in image processing methods stems from two principal application areas, improvement of pictorial information for human interpretation for autonomous machine perception. An image may be defined as a two-dimensional function,f(x,y) where x and y are spatial coordinates and (x,y) is called the intensity of f are all finite, discrete quantities , the image is called a digital image. Transform theory plays a fundamental role in image processing, as working with the transform of an image instead of the image itself may give us more insight into the properties of the image. Two dimensional transforms are applied to image enhancement, restoration, encoding and description and also vitally contributes for image compression. There are several different ways in which image files can be compressed. For Internet use, the two most common compressed graphic image formats are the JPEG format and the GIF format. The JPEG method is more often used for photographs, while the GIF method is commonly used for line art and other images in which geometric shapes are relatively simple.

2. 3D Discrete Cosine Transform (*DCT*)

Interest in image processing methods stems from two principal application areas, improvement of pictorial information for human interpretation for autonomous machine perception. An image may be defined as a two-dimensional function,f(x,y) where x and y are spatial coordinates and (x,y) is called the intensity of f are all finite, discrete quantities , the image is called a digital image. Transform theory plays a fundamental role in image processing, as working with the transform of an image instead of the image itself may give us more insight into the properties of the image. Two dimensional transforms are applied to image enhancement, restoration, encoding and description and also vitally contributes for image compression. There are several different ways in which image files can be compressed. For Internet use, the two most common compressed graphic image formats are the JPEG format and the GIF format. The JPEG method is more often used for photographs, while the GIF method is commonly used for line art and other images in which geometric shapes are relatively simple.

Fig 1 shows an overview block diagram of the compression and decompression process

First a sequence of 4 or 8 subsequent frames is collected and partitioned into three-dimensional video-cubes. The size of the cubes in our implementation is chosen to be 8*8 pixels in the spatial domain. On the time axis a depth of either 4 or 8 frames is being used. While deeper cubes lead to higher compression ratios, they also increase the latency of the compression process and the memory requirements.

Secondly, the cubes are transformed into the frequency domain, using a three-dimensional cosine transform. The energy compaction property of the cosine transform concentrates most of the information in a few low-frequency components. An example for the probability distribution after the transformation is shown in Fig 2. Subsequent quantization selectively removes information that is less relevant to the observer. This mostly affects the higher-frequency components. The quantization process determines the tradeoff between quality of the result and compression ratio. As quantization has an impact on spatial as well as on temporal components of the sequence, an additional degree of freedom which allows trading quality for compression ratio without affecting the computational complexity is attained. A detailed analysis of optimized quantization matrices based on rate-distortion theory and experiments can be found in for the 2D case, while describes a statistical approach for the 3D case. Other proposals have been made for 3D-DCT quantization schemes. However they mostly aim at the compression of 3D images as found in medical imaging applications and apply different criteria for the quality measurement of the result.

Finally, compression is performed in our approach with a zero run-length encoder and subsequent variable-length encoding. For the reconstruction process the above steps are just inverted and carried out in reverse order.

The similarity of the encoding and decoding process is one of the interesting properties of the 3D-DCT compression since it greatly facilitates the implementation of a joint en-decoder.

3 System Description

This section gives a more detailed description of the parts of the system outlined above. It also describes some extensions of the basic algorithm that lead to improved compression results. Tradeoffs are presented that greatly simplify an efficient implementation while only slightly degrading the overall performance.

3.1 Color-space Conversion

As the human visual system is less sensitive to color (chrominance) than to contrast (luminance), the original RGB image is initially converted into the YCbCr color space, separating the luminance component (Y) of the signal from the two chrominance components (Cb and Cr). This separation allows to apply a more aggressive quantization to the less sensitive color components of the image sequence.

3.2 Diferential Prediction

Maintaining Differential encoding is the simplest possible prediction scheme. It proves especially efficient for scenes with highly detailed static backgrounds. Since the 3D-DCT intrinsically decorrelates successive frames within individual video cubes, differential encoding within cubes is inappropriate. However, since the transform does not operate across cube boundaries, differential prediction may be used to remove some of the redundant information in subsequent cubes. Hereto we employ the last frame of the known cube as a predictor for each frame in the following cube. As the same predictor is used for every frame, this process only affects the temporal DC-component in the frequency domain. Especially in relatively static scenes these components carry the highest amount of information. The prediction process removes redundant information from the DC-plane and therefore greatly improves the achievable compression ratio. A major problem of combining differential encoding with a lossy compression algorithm is the fact that the original predictor is not perfectly known in the receiver. A general solution to this problem is to reconstruct the predictor in the transmitter after quantization, just as it is seen by the receiver, and to use this reconstruction instead of the original frame. This is depicted in Figure 3.The drawback of this approach is the need to concurrently run encoding and decoding in the transmitter, which nearly doubles complexity.

3.3. DC-Prediction:

After the transformation into the frequency domain, the information is concentrated in the low frequency coefficients. Looking across cube boundaries in the spatial domain, it becomes obvious that the DC-components of adjacent cubes can be highly correlated, especially in scenes with large monotonous areas. This correlation is again not exploited by the 3D-DCT due to its limited scope. A good prediction of the DC-component of some cube could be obtained from a combination of the DC-components of all its neighboring cubes. However, with respect to the implementation, only the previous cube to the left is used as a predictor, with the DC-component of the first cube in every row being encoded with its absolute value. This corresponds to the feed-back prediction scheme in Figure 3. Thus, error propagation is avoided for the DC-components. In this case the introduced overhead is marginal.

3.4. Quantization:

The quantization step allows the tradeoff between quality and compression ratio. While in the 2D-case quantization only offers two degrees of freedom to trade spatial blur for compression ratio, it offers three degrees of freedom in the 3D case. Very little work has been done so far in exploring the sensitivity of the human visual system to quantization induced distortions of the temporal frequency components. In our approach we simply use an exponential degradation along the spatial as well as along the temporal axis as quantization pattern. Another approach to this problem is described in [3] where statistical properties of the original signal are used to derive the quantization pattern.

An additional degree of freedom is provided by independently quantizing the luminance and chrominance channels, which can be treated independently. In general the color information in the Cb and Cr channels can be compressed much more than the contrast information, without having a noticeable effect on the observer.

3.5. Streaming:

The streaming process converts the 3D data structure of the video cube into a linear data stream. The stream order is essential for the subsequent zero run-length encoder. To optimize its efficiency the stream order has to maximize the average number of subsequent zeros. As the statistical properties of each component within a cube are known, the optimum sequence can be determined by sorting the components with their zero probability in ascending or descending order. If latency is not an issue, the streaming process can be delayed until all cubes of the sequence are compressed. The optimized stream order can then be applied to the whole sequence. This, again, increases the compression ratio and facilitates the use of component-specific error protection mechanisms for the transmission over lossy channels and the use of variable data rate channels. However, since the latency penalty and the hardware effort to store the whole sequence are very high, in this implementation each cube is streamed out individually. suggests diagonal planes as good approximation of an optimized stream order.

An appropriate way to reorder the components of the three color channels has to be found. Due to more aggressive quantization of the two chrominance channels they usually have a higher probability of becoming zero than the corresponding component in the luminance channel.

3.6. Run-Length (RL) Encoding:

The first step in the actual compression process is zero RL encoding. It combines runs of successive zeros into 2-tuples of a leading zero and the number of zeros in the run. With the knowledge of the statistics of the non-zero components, the encoding of the RL descriptors can be optimized for the subsequent variable-length encoding. The symmetric statistical distribution of the DCT components in Figure 2 encourages the use of a symmetrical encoding of the RL-descriptor: encoding even runs with a positive number and odd runs with a negative number. Shorter runs will be encoded with shorter symbols, while longer

RLdescriptors are used for longer runs.

3.7. Entropy-Encoding:

A standard Huffman encoder is being used for the variable-length encoding. As the probability distribution of the signal is nearly symmetrical, a fully sign-symmetrical code-book is chosen. Its size is limited to a range of [-255,+255] to maintain a reasonable size. All values that exceed the range of the code-book are encoded with a unique prefix code, followed by their actual value. Our experiments show that the optimum code-book is relatively independent of the type of scene that is being compressed, while it is clearly subject to the quantization level used. The en/decoder therefore contains three built-in code-book ROMs for low, medium and high compression ratios. An additional sequential decoder with a programmable code-book can be used when the system does not have to run at its full speed. It is mainly intended for evaluation purposes. The limited range of the code-book has no noticeable influence on the compression efficiency.

4.RESULTS

Consider a matrix of an original image before applying the 3D DCT, which is given by

$$A(,:,:,) = \begin{bmatrix} 149 & 56 & 9 & 16 & 8 & 1 & 10 & 6 \\ 107 & 34 & 2 & 10 & 4 & 0 & 6 & 0 \\ 55 & 11 & 0 & 9 & 2 & 2 & 5 & 0 \\ 19 & 5 & 5 & 10 & 5 & 6 & 8 & 0 \\ 6 & 8 & 12 & 10 & 5 & 8 & 8 & 2 \\ 5 & 11 & 13 & 8 & 5 & 7 & 7 & 6 \\ 5 & 11 & 9 & 4 & 7 & 7 & 6 & 8 \\ 3 & 9 & 4 & 1 & 8 & 8 & 6 & 11 \end{bmatrix}$$

The resultant matrix after applying 3D DCT is given as below,

$$B(,:,:1) = \begin{bmatrix} 17 & 18 & 15 & 7 & 3 & 1 & 0 & 0 \\ 13 & 20 & 13 & 7 & 4 & 1 & 0 & 0 \\ 7 & 9 & 7 & 3 & 1 & 0 & 0 & 0 \\ 5 & 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The compression ratio for transmitting one set of images , that is, 8 frames of data after applying 3D DCT is found to have a compression ratio(CR) of 27.

5.Conclusion

Every day, an enormous amount of information is stored, processed and transmitted digitally. Methods of compressing data prior to storage and/or transmission are of significant practical and commercial interest. A 3D DCT is made up of a set of 8 frames at a time. Image compression using 3D DCT minimizes the size in bytes of a graphics file without degrading the quality of the image to an unacceptable level, psychovisual redundancy is acheived. The reduction in file size allows more images to be stored in a given amount of disk or memory space. It also reduces the time required for images to be sent over the Internet or downloaded from Web pages. For Internet use, the two most common compressed graphic image formats are the JPEG format and the GIF format. The JPEG method is more often used for photographs. At some later time, the image is retrieved to reconstruct the original image or an approximation of it. The CR for a 3D DCT is found to be approximately 27 , which implies that this technique has a better CR as compared to a JPEG using 2D DCT.

References

R.C.Gonzalez and R.E.Woods,“Digital Image Processing,”*Reading, MA*: Addison-Wesley,1992.

A. Ramaswamy and W. B. Mikhael, “A mixed transform approach for efficient compressionof medical images,” *IEEE Trans. Med. Imag.*, vol.12, no. 4, pp. 803-811, 1996.

A.K.Jain,“A sinusoidal family of unitary transforms,” *IEEE Trans. Pattern Anal. Machine* Intell., vol. PAMI-1, pp.356-365,Oct. 1979.

G.Strang and T.Nguyen,“Wavelets and Filter Banks,” *Cambridge*, MA: Wellesley-Cambridge,1996.

Shen-Chuan Tai, Yung-Gi Wu, and Chang-Wie Lin, “An Adaptive 3D Discrete Cosine TransformCoder for Medical Image Compression,”*IEEE Trans. On Information Technology in Biomedicine*, Vol.4, No.3, pp. 259-263,Sep.2000.

Y. G.Wu and S. C. Tai, “Medical image compression using 2*2 discrete cosine transform,” *Opt. Eng.*, vol.37,no.5,pp.1539-546, May 1998.

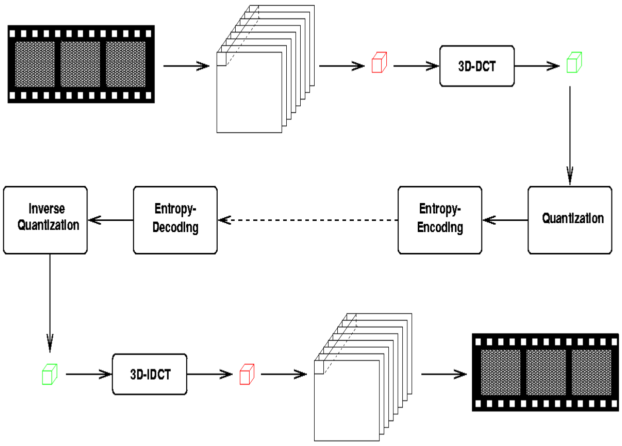
Y. G. Wu and S. C. Tai, “Medical ImageCompression by Discrete Cosine Transform Spectral Similarity Strategy”, *IEEE Trans. Information Technology in Biomedicine*, vol.5, no.3, pp. 236-243, Sep.2001.

J.Wang and H.K.Huang,“Medical Image Compression by Using Three Dimensional Wavelet Transformation,” *IEEE Trans. Medical Imaging*, vol.15, no.4, pp. 547-554, Aug.1996.

Yung-Gi Wu, “Medical Image Compression by Sampling DCT Coefficients”, IEEE Trans. Information Technology in medicine, vol.6, No.1, pp. 86-94, Mar.2002.M.Nelson and J.L.Gaily, *The Data Compression Book*, 2nd ed. New York: M & T Books, 1996.

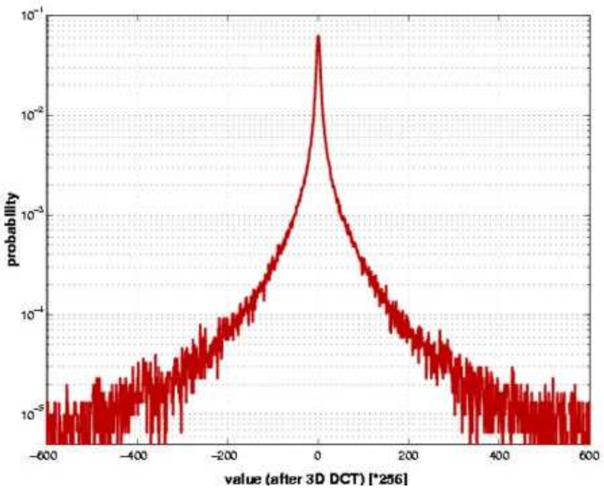
Y.-L. Chan and W.-C. Siu, “Variable Temporal-Length 3-D Discrete Cosine Transform Coding,” *IEEE Transactions on Image Processing*, Vol. 6, No. 5, pp. 758-763, May 1997.

Figure 1



Block diagram of the compression and decompression process

Figure 2



Probability distribution after 3D-DCT

This academic article was published by The International Institute for Science, Technology and Education (IISTE). The IISTE is a pioneer in the Open Access Publishing service based in the U.S. and Europe. The aim of the institute is Accelerating Global Knowledge Sharing.

More information about the publisher can be found in the IISTE's homepage: http://www.iiste.org

The IISTE is currently hosting more than 30 peer-reviewed academic journals and collaborating with academic institutions around the world. **Prospective authors of IISTE journals can find the submission instruction on the following page:** http://www.iiste.org/Journals/

The IISTE editorial team promises to the review and publish all the qualified submissions in a fast manner. All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Printed version of the journals is also available upon request of readers and authors.

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar