



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

Review on Conversion of Various Block Sizes in Compressed Domain using DCT

Vaishali Choudhary¹, Ketki Bhakare², Swati Dhopte³

Lecturer, Department of Computer Technology, RGCER (NYSS), Nagpur, Maharashtra, India¹

Lecturer, Department of Computer Science and Engineering, DBACER (SVSS), Nagpur, Maharashtra, India²

Assistant Professor, Department of Computer Science and Engineering, MITCOE, Pune, Maharashtra, India³

ABSTRACT: Image transforms are extensively used in image processing and image analysis. Transform is basically a mathematical tool, which allows us to move from one domain to another domain. Transforms play a significant role in various image processing applications such as image analysis, image enhancement, image filtering and image compression.

Nowadays, almost all digital images are stored in compressed format in order to save the computational cost and memory. To save the memory cost, all the image processing techniques like feature extraction, image indexing and watermarking techniques are applied in the compressed domain itself rather than in spatial domain.

In this paper, composition of a block from all of its sub-blocks and vice versa, directly in DCT domain is obtained and showed that the result of both operations are same and the computational complexity of the proposed algorithm is lower than that of the existing ones.

Keywords: discrete cosine transform, compression, JPEG, region of interest, frequency domain.

I. INTRODUCTION

A. IMAGE TRANSFORMS

Transform is basically a mathematical tool, which allows us to move from one domain to another domain. Transforms play a significant role in various image processing applications such as image analysis, image enhancement, image filtering and image compression. Nowadays, almost all digital images are stored in compressed format in order to save the computational cost and memory. The transforms do not change the information content present in the signal. Transform give information about the frequency contents in an image.

B. NEED OF COMPRESSED DOMAIN OPERATIONS

In order to save the computation and memory cost, it is desirable to have image processing operations such as feature extraction, image indexing, and pattern classifications implemented directly in the DCT domain. The results reveal that DCT coefficients of any block can be directly obtained from the DCT coefficients of its sub-blocks and that the inter-block relationship remains linear.

It is useful in extracting global features in compressed domain for general image processing tasks. For compression purpose we have used DCT (Discrete Cosine Transform). The discrete cosine transform (DCT) is one of the most widely used tools in digital signal processing. All existing digital media compression standards—JPEG, MPEG, H.26x—are mainly based on DCT. Recently, direct image processing in the DCT domain, rather than in the pixel domain, has been developed in such areas as feature extraction, watermarking and error concealment. It is often the case that different applications require different block sizes, such as 4×4 blocks in image indexing, 8×8 blocks in JPEG, and 16×16 macro-blocks in MPEG.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

II. RELATED WORK

Qionghai Dai [1] have proposed the inter-transfer between different DCT block sizes for multidimensional images which is useful in applications such as transcoding between compression standards. Besides two dimensional (2-D) image and three-dimensional (3-D) image (video), there are four or more dimensional images, e.g., light field and lumigraph. Therefore, multidimensional DCT, given its associated sub-blocks of DCT coefficients, is also useful in some applications for these high dimensional images. The traditional way of computing DCT coefficients for a signal block from its sub-blocks is to employ inverse discrete cosine transform (IDCT) for each sub-block to recover data to the time (or spatial) domain first and then to apply DCT again to obtain DCT coefficients with required sizes. It is obvious that this method is inefficient. In 1991, Kou *et al.* proposed a direct computation method that slightly reduced the number of multiplications and additions compared with the traditional method. Recently, Skodras also proposed a direct method to compute one-dimensional (1-D) N -point DCT from two adjacent $\frac{N}{2}$ -point sub-blocks, and the computation cost mainly lies in an $\frac{N}{2}$ -point DCT, $\frac{N}{2}$ -point IDCT, and $\frac{N}{2}$ multiplications between DCT and IDCT. Although it was suitable for the case of blocks and sub-blocks with various sizes, it is unable to use any existing fast DCT algorithms, and thus, its computational cost still remains high.

In this paper, an efficient algorithm for computing the N -point DCT for a signal block, given its two adjacent blocks in the DCT domain, is proposed. The result is most efficient compared with other existing algorithms for 1-D DCT-to-DCT computation. In this paper, an efficient algorithm for computing the N -point DCT for a signal block, given its two adjacent blocks in the DCT domain, is proposed. The result is most efficient compared with other existing algorithms for 1-D DCT-to-DCT computation.

Athanassios [2] have proposed the direct calculation of the N DCT coefficients, when the two adjacent $\frac{N}{2}$ sets of DCT coefficients. According to author, signals, images and video are compressed from their early stages, i.e., just after their acquisition. One such example is a system for processing documents. The images are scanned at high resolution and are immediately compressed by hardware or software, in order to save memory space. Another example comes from the digital video area, where video signals are compressed when transmitted over networks or stored in databases.

A similar situation holds for the speech signals. However, in all these cases, it is likely that the signals will have to be processed before being displayed, transmitted, printed, etc. Some of the frequently used processing functions are scaling, filtering, rotation, and translation. Implementing these functions in the compressed domain is advantageous from the computational complexity point of view, as well as the image quality and the memory usage (space and number of accesses). This is because the transition to the time or spatial domain and the recompression of the data are avoided. All existing compression standards—JPEG, MPEG1, MPEG2, H.26x—are based on the discrete cosine transform (DCT), which is applied on blocks of data of certain length. Specifically, the problem that is studied is the direct calculation of the N DCT coefficients, when the two adjacent $\frac{N}{2}$ sets of DCT coefficients are given. According to this, one has first to recover the data to the time (or spatial) domain by calculating the inverse DCT of the two $\frac{N}{2}$ coefficient sequences and then to retransform them by means of an $-N$ point DCT. The proposed approach finds direct application to image browsing, where it may be sufficient to initially deliver an image of lower resolution to the user. Then, based on user's response, the higher resolution image could be provided by the server.

According to Chan Yul Park and Nam Ik Cho [3], with the advent of H.264 video coding standard and its potential to be a major coding scheme in many applications, the need for the conversion of existing MPEG-2 video into this new standard is ever growing. However, unlike the conversion between the former video coding standards, the conversion between the H.264 and former standards is not easy because the new standard use different transform, i.e., integer transform. More precisely, the source video stream need not be fully decoded to pixel level in the case of transcoding between MPEGs, because the DCT coefficients can be reused. Hence the transform domain transcoding is generally faster than the pixel domain transcoding. However, for the conversion of DCT coefficients into IT coefficients, additional transform matrix multiplications are needed. As a result, the transform domain transcoding is not efficient at all. Hence, a fast conversion algorithm between two transform coefficients is required for the efficient transform domain transcoder. But we could find only two algorithms related with the conversion between DCT and IT. One is the algorithm proposed by Xin that multiplies a conversion matrix to the DCT coefficient matrix. The cascade of this conversion matrix and DCT matrix is similar to the integer DCT, and the number of computations is not reduced. The



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

other has proposed by Shen which uses the factorized form of 8×8 DCT matrix. Multiplications in the process of matrix multiplications are replaced by additions and shifts by using Merhav's method. But this conversion method has disadvantage that there are much computational errors because of approximation.

So, in this paper, authors have proposed a more efficient algorithm for the conversion from the DCT coefficients into IT coefficients. The algorithm is based on the decomposition of the conversion matrix into cascades of sparse matrices that need fewer computations. They also have proposed a modified 4×4 DCT and use it instead of IT in order to reduce the computational complexity. Two cases of transcoding are considered based on these basic ideas, i.e., transcoding into the same and half resolutions. For the former, the 8×8 DCT coefficients are converted into four 4×4 IT coefficients, and the lower band coefficients among 8×8 DCT coefficients are converted into 4×4 IT coefficients for the latter case. The number of computations and resulting video quality of each algorithm for each case are compared and it is shown that the proposed algorithm has very little approximation errors while reducing much computational complexity.

In paper, by Shih-Fu Chang [4], as effective techniques for *image indexing/searching* are required for large visual information systems (such as image databases and video servers). In addition to traditional methods that allow users to search images based on keywords, image query by example and feature-based image search provide powerful tools to complement existing keyword-based search techniques. Usually, prominent image features (such as texture, shape, color, and object motion) are extracted and stored as side information. Then, similarity retrieval is performed, based on the comparison of the features associated with each image in the database. Another important image technology for general multimedia systems and applications is *image manipulation*.

On a desktop video editing system, users would like to have general tools for image geometrical transformation, image filtering, multi-image composition, and video segment cut-and-paste. In a networked video application, users may want to subscribe multiple image/video sources from different locations and combine them to a single displayable format. In a multi-point video conferencing application, a network device such as a video bridge may receive multiple video sources and generate multiple video streams of various forms to different end users. As mentioned above, there is a general need for efficient image searching and manipulation techniques for multimedia applications. However, these techniques are usually pursued independently of the design of image compression algorithms. Most of today's image compression methods are concerned mainly with optimization of signal distortion, bit rate, and coding complexity.

This paper has given an overview of research on compressed domain techniques for image video indexing and manipulation. Authors have given examples of visual feature extraction, image matching, image manipulation, and video editing in the compressed domain.

In paper, by Juan R. Hernández [5], since the development achieved during the last decades by both image processing techniques and telecommunication networks have facilitated the current explosion of applications that considerably simplify the acquisition, representation, storage, and distribution of images in digital format. Since protocol and network design is oriented to digital data delivery, most contents providers are rapidly transforming their archives to a digital format. However, all these advances have also made it possible to produce digital copies identical to the original with the greatest ease. In addition, unauthorized manipulation or reuse of information has become so common that this emerging creative process has been put in danger. Although cryptography is an effective tool against the illegal digital distribution problem, it has to be coupled with specialized hardware in order to avoid direct access to data in digital format, something that it is not only costly but would also reduce the marketing possibilities for the service provider. It is in this scenario where watermarking techniques can prove to be useful. A digital watermark is a distinguishing piece of information that is adhered to the data that it is intended to protect.

In this paper, authors have considered invisible watermarks that protect rights by embedding ownership information into the digital image in an unnoticeable way. This imperceptibility constraint is attained by taking into account the properties of the human visual system (HVS), which in turn helps to make the watermark more robust to most types of attacks. Watermarking, like cryptography, needs secret keys to identify legal owners; furthermore, most applications demand extra information to be hidden in the original image (steganography).

This information may consist in ownership identifiers, transaction dates, serial numbers, etc., that play a key role when illegal providers are being tracked. Closely related to the embedment of this information is the extraction (*watermark decoding*) process whenever in possession of the secret key. In most cases of interest, there will be a certain probability of error for the hidden information which can be used as a measure of the performance of the system. Clearly, this probability will increase with the number of information bits in the message.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 1, January 2014

In this paper, authors have dealt with the analysis of watermarking methods in the discrete cosine transform (DCT) domain, which is widely used in compression applications and consequently in digital distribution networks. They have assumed that the original image is not present during the watermark detection and information decoding processes, since this would be hard to manage when huge quantities of images need be compared by intelligent agents searching the net for unauthorized copies.

The main goal of this paper is to present a novel analytical framework that allow to assess the performance of a given watermarking method in the DCT-domain. New results have been achieved by resorting to a theoretical model of the problem that permits to derive optimal detector and decoding structures. With those results, it is possible to determine how much information one can hide in a given image for a certain probability of error or what is the probability of correctly deciding that a given image has been watermarked.

III. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a concept of general spatial relationship between the DCT coefficients of a block and that of its sub-blocks. The results of survey reveal that there exists a concise and linear relationship between the DCT of a block and that of its sub-blocks. We hope that a substantial savings in computing cost can be achieved in comparison with those in the pixel domain, which is especially useful when image processing is carried out directly in the DCT domain.

We can extend this theory by applying proposed technique to videos since most digital images and video sequences today are both stored and transmitted in compressed form.

We can extend this theory of inter-transfer of DCT coefficients, for insertion of captions and logos in compressed domain itself.

REFERENCES

1. Qionghai Dai(2005), 'Fast Algorithms for Multidimensional DCT-to-DCT Computation between a Block and its Associated Sub-blocks', IEEE transactions on signal processing, VOL. 53, NO. 8.
2. Athanassios N. Skodras(1999), 'SAWSDL: Direct Transform to Transform Computation', IEEE Signal Processing Letters, VOL. 6, NO. 8.
3. Chan Yul Park and Nam Ik Cho, 'A Fast Algorithm for the Conversion of DCT Coefficients to H.264 Transform Coefficients', Seoul National University San 56-1, Shilim-dong, Kwanak-gu, Seoul, 151-742, Korea.
4. Shih-Fu Chang, 'Compressed-Domain Techniques for Image Video Indexing and Manipulation', Department of Electrical Engineering and Center for Telecommunications Research Columbia University, New York, NY 10027.
5. Juan R. Hernández, *Associate Member, IEEE*, Martín Amado, and Fernando Pérez-González(2000), 'DCT Domain Watermarking Techniques for Still Images: Detector Performance Analysis and a New Structure', IEEE Transactions on Image Processing, VOL. 9, NO. 1.
6. Guocan Feng and Jianmin Jiang(2001), 'Image Spatial Transformation in DCT Domain', IEEE.