Graph-based Non-Rectangular Transform Coding with Application to Video Compression

K.A.S.T. Senarath
(M.Sc Candidate)
Department of Electrical and Computer Engineering,
University of Manitoba

March 21, 2023

1 Introduction

In order to increase video quality for very low bit-rate applications, the separation of the video content into segments or objects has emerged as an important technique which is known as Video Compression. The state of art methods for video compression involves partitioning images into rectangular blocks and using transform coding on each block. However, it is widely recognized that dividing images based on texture into irregularly shaped regions and utilizing shape-adaptive transforms that are optimized for local textures can produce more concise representations of images than rectangular ones [1]. Although this method provides superior compression for textures, it poses a challenge in efficiently communicating intricate and irregular geometric shapes that arise from the texture-based segmentation of images to the decoder.

2 Background

Contemporary video codecs divide video frames into rectangular regions of different sizes through structured partitions and then use transform coding on these rectangular blocks to minimize the amount of additional data required in signaling region shapes [2]. An alternative approach would be to combine adjacent blocks within the partition to create regions that are a combination of rectangular shapes, resulting in more efficient transform coding. Additionally, these combined regions can be signaled to the decoder with high efficiency [3]. Although researchers have previously explored creating non-rectangular transforms, these transforms have mostly been developed through orthogonalization of transforms intended for square regions. Consequently, they may not be the most effective option for compressing non-rectangular shapes. One potential method for creating the most efficient transforms for non-rectangular image regions involves using a graph representation of image signals [4]. An earlier study explored creating a transform using basic 4-connected graphs with weights of one [5]. However, a crucial criterion for image transforms used in coding is energy compaction. But the optimality of the transforms developed in [5] has not been analyzed in terms of energy compaction. It is worth noting that when the signal support is square, this results in the 2D-DCT transform, which gradually approaches the KLT transform - the most efficient transform possible in theory.

3 Proposed Research

The proposed research for video compression will investigate several aspects related to the effectiveness of transforms used in video compression. Firstly, the study aims to determine the similarity between the transforms obtained from a 4-connected unit-weight graph and the KLT of the signal, with reference to the eigenvectors of the correlation matrix. This will provide insight into the optimality of the 4-connected unit-weight graph as a transform and how closely it approximates the KLT. And also, the research aims to evaluate the energy compaction effectiveness of the transform derived from the 4-connected unit-weight graph when compared to the conventional shape-adaptive DCT [1]. This will help to determine whether the proposed transform is better suited for video compression than the traditional method. Furthermore, the study seeks to explore whether the transforms can be improved by using graph weights optimized to the characteristics of the image. This can be achieved by fitting a Gauss-Markov random field with spatially invariant edge weights to the image segment to be coded, and encoding only a few edge weights as side-information for each image segment. This approach can potentially reduce the overhead due to transform signaling. Ultimately, the proposed research will provide insights into the effectiveness of different transforms for video compression, and suggest potential improvements to enhance compression efficiency.

References

[1] T. Sikora, "Low complexity shape-adaptive dct for coding of arbitrarily shaped image segments," *Signal Processing: Image Communication*, pp. 381 – 395, Nov. 2015.

[2] B.Bross, J.Chen, J.R.Ohm, G.J.Sullivan, and Y.K.Wang, "Developments in international video coding standardization after avc, with an overview of versatile video coding (vvc)," *Proceedings of the IEEE*, vol. 109, no. 9, Sept. 2021.

[3] D. Marpe, H. Schwarz, S. Bosse, B. Bross, P. Helle, T. Hinz, H. Kirchhoffer, H. Lakshman, T. Nguyen, S. Oudin, M. Siekmann, K. Suhring, M. Winken, and T. Wiegand, "Video compression using nested quadtree structures, leaf merging, and improved techniques for motion representation and entropy coding," *IEEE Trans. Circuits Syst. Video Technol*, vol. 20, no. 12, pp. 1676–1687, 2010.

[4] E. M. G. Cheung and Ed, "Graph spectral image processing," *ISTE and John Wiley*, 2021.

[5] G. Fracastoro, F. Verdoja, M. Grangetto, and E. Magli, "Superpixel-driven graph transform for image compression," *IEEE ICIP*, pp. 2631–2635, 2015.