DOI: 10.21917/ijivp.2022.0390

DEEP GENERATIVE DISCRETE COSINE TRANSFORM FOR SPECTRAL IMAGE PROCESSING

L. Mary Florida[1], S. Vadhana Kumar[2], A.J. Deepa[3], M.L. Sworna Kokila[4] and S. Brilly Sangeetha[5]

[1]Department of Mathematics, St. Xavier's Catholic College of Engineering, India
[2]Department of Computer Science and Engineering, ILAHIA College of Engineering and Technology, India
[3]Department of Computer Science and Engineering, Karunya Institute of Technology and Science, India
[4]Department of Computer Science and Engineering, Ponjsely College of Engineering, India
[5]Department of Computer Science and Engineering, IES College of Engineering, India

Abstract

The ever-increasing number of publications and applications in the field of cross-spectral image processing has led to the area receiving greater focus than it previously had. In cross-spectral frameworks, the data from hyperspectral bands is blended with the data from other spectral bands in order to provide responses that are more robust to particular obstacles. Cross-spectral processing could be useful for a variety of applications, including dehazing, segmentation, calculating the vegetation index, and face identification, to name just a few of them. The availability of cross-and multi-spectral camera sets on the market, such as smartphones, multispectral cameras for remote sensing, or multi-spectral cameras for automotive systems or drones, has spawned an increased number of applications for these cameras. In this paper, we develop a deep generative discrete cosine transform for possible image processing for the enhancing the quality of images. This is conducted to improve the prediction or classification ability by the classifiers on hyperspectral images. The models are validated with various machine learning classifiers. The results of simulation shows that the proposed method higher degree of accuracy than the existing methods.

Keywords:

Deep Generative Model, Discrete Cosine Transform, Spectral Image Processing

1. INTRODUCTION

Everyone now has the ability to shoot images whenever and wherever they like thanks to the proliferation of smartphones that come equipped with high-quality digital cameras. In addition, users of social media platforms such as Facebook, Instagram, and Snapchat are presented with a plethora of compelling reasons to share the private images they have stored on their devices. Sharing images is a significant component of the revenue model of social media platforms, and numerous businesses make use of this feature as a marketing tactic. It should not come as a surprise that the number of images taken each year continues to rise. During the time period of 2000–2005, when film image was still widely used, an estimated 80 billion images were taken all over the world. As a direct result of digital image and the widespread use of mobile phones, there will be more than 1.4 trillion images taken in the year 2020. As a consequence of this, it is estimated that 7.4 trillion images will be retained in the year 2020, and it is anticipated that this number will climb to 9.3 trillion images in the year 2022.

The transition from analogue to digital image led to the research of image compression techniques for use with digital files. A digital image is made up of these individual intensity values that are collected together. It hard to believe that even with today advanced storage technologies, only three 1280 x 720 grayscale photos could fit on an old 2.88MB floppy disc. Floppy discs were the most common type of storage medium used in the 1980s and 1990s. It hard to believe that even with today advanced storage technologies, Due to the fact that there are three colour channels and, as a result, 24 bits per pixel, a floppy disc can only hold a single colour image that is 1280 by 720 pixels.

Even though data storage technology has advanced greatly over the past few decades and storage mediums have risen by many magnitudes, image sizes have continued to increase despite these developments. Compact cellphones include cameras that are capable of storing 324 megabytes of data and have a resolution of 108 megapixels. In order to transmit and receive those images, you will need to have a mobile internet connection. As a consequence of this, research into the process of photo compression is still important and ongoing.

Images are compressed in order to cut down on the number of bits that are required to encode them, which allows for a more space-effective method of storing them. Because the majority of images contain a lot of redundant data in their initial representation, this is something that can be done with most of them. The majority of photos have some kind of structure to them, as opposed to simply being a jumbled assortment of intensity levels. As a consequence of this, pixels that are geographically close to one another frequently share connections. A method of data compression that is considered to be good eliminates information that is superfluous or redundant and encodes the data that is left over in a manner that is both effective and efficient. In order to achieve the desired compression ratio in practise, it is often necessary to delete material that is nonredundant as well as data that is relevant.

Image transform coding has become one of the most widely used techniques for the compression of images (ITC). By reorganising and decoding the raw data, it is possible to extract this information in a time-efficient manner. The Discrete Cosine Transform (DCT) developed by Ahmed, Natarajan, and Rao is the image compression transform that is utilised the majority of the time [2]. The transformation is related to the Fourier series, but instead of using discrete values, only real numbers are employed in the transformation. JPEG, which stands for Joint photographic Experts Group, is a digital image format that was developed in 1992 by the JPEG. The DCT served as the foundation for the development of JPEG.

Because the majority of frequencies in the frequency domain are unimportant to the image at hand, the transform in question is related to Fourier analysis. Another consequence of this is a representation that is devoid of any kind of spatial localization.

The JPEG committee was responsible for the design of JPEG 2000, and it was first used in the year 2000 [9]. It uses the discrete wavelet transform (DWT), as stated in [2], in order to represent signals utilising a wavelet-generated orthonormal basis. A multiresolution analysis is the result of the dilation and translation of a mother wavelet in order to establish an orthonormal foundation. This analysis was invented by Mallat. In the years that followed, other multiscale techniques such as wedgelets, curvelets, contourlets, bandelets, and shearlets [7] as well as the easy path wavelet transform (EPWT) [1], [2] were all suggested as potential solutions [3]. The DWT coefficients are used by the EPWT to generate a graph, which ultimately results in a high degree of correlation between the various data points displayed on the graph. As a consequence of this, it is an easy fit with the graphing approaches that are discussed further on.

The adaptive thinning approach was presented in [4]. This algorithm offers a means to deconvolve an image without altering it in any way. A procedure known as thinning is applied to the data in order to provide a sparse selection of representative pixels. Using these meticulously chosen pixels, an estimate of each image is calculated using the Delaunay triangulation method. Before being used for image compression in [5], the technology had its beginnings in [6] and was initially created for the purpose of landscape modelling. Its purpose is to act as a stand-in for a vast quantity of bivariate scattered data. [6] and [10] both present examples of how pixel selection, post-processing methods, and effective contextual coding have all been made to be more effective [8]. In [9], the best N-term approximation rates for the relevant classes of piecewise linear horizon functions and regular functions are presented. The utilised adaptive thinning on film in order to gain a deeper understanding of the technique. In, it was demonstrated that trivariate horizon functions possess optimal N-term approximation rates.

2. DISCRETE COSINE TRANSFORM

When applied to random processes with certain correlational properties, the DCT is a useful approximation for the KLT.The distance between the DCT and the KLT is what we are attempting to ascertain at this time. An approximation of the Q matrix can be obtained by deriving the DCT-II matrix Φ through the process of eigenvalue decomposition.

In the process of disassembling M into its component pieces, E ought to presumably assume a nearly diagonal position, given that:

$$\Phi^{-1}*E*\Phi \text{ (or } E = \Phi*M*\Phi^{-1}\text{).}$$

If the element E was diagonal, then the previously random process S, which had a covariance matrix M, would be converted into an uncorrelated random process Φ, as specified by the KLT. To compress and decompress signals similarly to the KLT, we need a signal s, also known as a realisation of S.

Therefore, to get started, we need to partition the S covariance matrix M into the components $\Phi^{-1}*E*\Phi$. The columns of are cleaned up by removing the least significant diagonal entries in E using a $N\times(N\text{-}r)$ diagonal matrix called V.

It is expected that one percent of E diagonal entries will originate from the very smallest diagonal entries. Now, a signal that has a high degree of correlation, Φ, needs to be converted into a signal that has no correlation at all by utilising $\Phi^{-1}*s$. Because of this, $(V^T*\Phi^{-1})$ will throw out the entries in $\Phi^{-1}*s$ that are the least significant. The decompression matrix for these data is represented by the V, and our compression matrix will be written as $(V^T*\Phi^{-1})$. Based on the assumption that our future signal possesses a known covariance matrix, the compression matrix $T=V^T*\Phi^{-1}$ and the decompression matrix $TT=\Phi*V$ would be predetermined in a real system, which is not modelled; however, these matrices would be determined in advance in a hypothetical system. After multiplying the input signal s by T, sending the product (which now has $N\text{-}r$ samples instead of N samples), and receiving the product, the reconstructed signal is obtained (which is then multiplied by $TT*T*$s). All of this may be compressed and decompressed using the formula

$$\Phi*V*V^T*\Phi^{-1} = \Phi*V`*\Phi^{-1}, \text{ where } V' = V*V^T.V'$$

This formula is capable of describing the entire process. In the diagonal entries of V', the columns that were thrown away are represented by the value 0, while the columns that were kept are represented by the value 1 in those same diagonal entries.

3. DEEP GENERATIVE DCT

The discrete cosine transform, also known as the DCT, is one of the image processing methods that sees the most action and is employed in both the denoising and compression of digital images.

Because the DCT has the potential to compress energy, only a small number of its coefficients are required to encode the information contained in an image. Because of this enhanced fit for DCT, it is possible to say that DCT is better suited for image compression applications than other methods. Because DCT is linear and invertible, separating the transformation coefficients is a simple and straightforward process. Creating a meaningful data structure that can be utilised to extract information with more precision can be accomplished with the help of transformation coefficients that are independent of one another. As a consequence of this change, the interpixel redundancy as well as the interband redundancy have been removed. In this research, the hyperspectral cube is encoded using the DCT coefficients obtained from a three-dimensional DCT. The pixel vector can also be subjected to a 2-D DCT transformation in order to generate a 3-D DCT. This is the two-dimensional discrete cosine transform with a size of $M\times N$.

A deep generative DCT, and its coefficients are utilised in order to do an analysis on an image of a low-dimensional hyperspectral spectrum (coefficients of DCT). One of the most essential characteristics of DCT is its ability to condense the energy of a image into a small number of coefficients. As a direct consequence of this, DCT coefficients are widely used in the pattern recognition process as features.

In the hyperspectral image with a low dimension, the DCT coefficients of each pixel can be directly concatenated to form the feature vector for the pixel (m, n).

Convert the deep generative feature vector that is based on DCT to the final concatenated cube $f\in R^{H\times W\times j}$, which can be written as:

$$f=(f_1,f_2,\ldots,f_j)$$

4. RESULTS AND DISCUSSION

In this part, experiments were conducted on three standard datasets, including Indian Pines, Pavia University, and the Salinas dataset, in order to evaluate the efficacy of the suggested technique. These datasets were chosen because they are considered to be representative of real-world situations.

36. In every test, a personal computer outfitted with 16 gigabytes of random-access memory (RAM), a processor running at 2.70 gigahertz (GHz), and the most recent version of MATLAB 2018a was employed.

In order to determine whether or not the DGDCT method is effective, it was evaluated against more conventional methods of feature extraction. The SVM-PCA, the ICDA, and the LDA were pitted against one another, despite the fact that all three have received a considerable amount of attention. In this inquiry, the transform-based method of feature extraction known as the DGDCT is also taken into consideration. Because the initial hyperspectral image is used immediately for the classification phase in this method, the feature extraction step is skipped. This allows the method to be more efficient.

Table.1. Accuracy

Features	SVM-PCA	ICDA	LDA	Proposed
10	89.58	90.10	90.93	91.77
20	89.45	89.97	90.80	91.64
30	87.18	87.68	88.49	89.31
40	78.91	79.37	80.10	80.84
50	89.91	90.44	91.27	92.12
60	88.15	88.66	89.48	90.31
70	89.60	90.12	90.95	91.80
80	88.13	88.64	89.45	90.29
90	89.57	90.09	90.92	91.77
100	89.90	90.42	91.25	92.10

Table.2. Precision

Features	SVM-PCA	ICDA	LDA	Proposed
10	90.48	91.01	91.85	92.70
20	90.15	90.67	91.51	92.36
30	89.47	89.99	90.82	91.67
40	80.83	81.29	82.04	82.81
50	89.91	90.44	91.27	92.12
60	86.09	86.59	87.38	88.20
70	90.34	90.87	91.70	92.56
80	88.13	88.64	89.45	90.29
90	89.94	90.47	91.30	92.15
100	90.67	91.20	92.04	92.90

Table.3. Recall

Features	SVM-PCA	ICDA	LDA	Proposed
10	89.58	90.10	90.93	91.77
20	89.20	89.71	90.54	91.38
30	87.94	88.45	89.27	90.10
40	86.95	87.46	88.26	89.09
50	88.67	89.19	90.01	90.84
60	88.53	89.04	89.86	90.70
70	89.75	90.27	91.10	91.95
80	87.37	87.87	88.68	89.51
90	89.48	90.00	90.83	91.67
100	90.38	90.90	91.74	92.59

Table.4. F-measure

Features	SVM-PCA	ICDA	LDA	Proposed
10	89.88	90.40	91.23	92.08
20	90.34	90.86	91.70	92.55
30	90.13	90.65	91.49	92.34
40	81.59	82.06	82.82	83.59
50	87.05	87.56	88.37	89.19
60	87.03	87.53	88.34	89.16
70	89.90	90.42	91.25	92.10
80	89.27	89.78	90.61	91.45
90	89.10	89.62	90.45	91.29
100	90.12	90.64	91.48	92.33

The KKA, KKB, and KFR datasets, each of which contains 112 images, are utilised in this investigation. Each of these datasets is referred to as a set. Pigment patches and hyperspectral paintings are used to create the images that are included in the datasets. These paintings serve as both the foreground and the background of the images. The level of complexity of the foreground objects utilised in each different dataset is what differentiates them from one another. The majority of the stuff that makes up KKA and KKB backgrounds comes from pigment patches. After that, the foregrounds of each patch show a single homogenous zone or two homogenous zones, depending on the number of patches. The background and foreground of KFR are composed of pigment patches and a painting, respectively.

DGDCT optimises design by its own nature, making it more general, as opposed to reducing the number of operations without first optimising the algorithms involved. In spite of the fact that DGDCT is a deep learning algorithm, there are still many unanswered questions concerning the risks that are associated with the involvement of humans in applications of this kind. As far as we are aware, there is currently no treatment that can be considered effective for them. While the medically measured skin parameter ranges are used to verify that assessment results are accurate, a random subset of the individuals is used to prevent evolution from reaching a local maximum. According to the results of extensive testing, the fitness values of the DGDCT skin parameters are lower than those of the whole, and the maps that are generated from them have a better aesthetic effect than those that were evolved using the pure GA (328 270 multi-spectral image pixel samples). When all of these factors are taken into account, the newly suggested solution is an excellent option for dealing with the issue at hand.

The expanding number of applications in fields such as environmental, industrial, and cultural heritage research is driving an ever-increasing demand for the metrological processing of hyperspectral images. In this particular work, the performance evaluations of nonlinear image processing methods that are based on ordering have been the primary focus of our attention and efforts. Two distinct classes have been established for the criteria. Examples of quantitative examinations of measurement errors and bias include median filtering and gradient detection, for instance. The processing of data to guarantee that it corresponds to projected theoretical qualities and physical sense is the basis for the second set of standards. It is necessary for the hyperspectral sensor to comply with physical requirements for data in order for it to be able to link the processing results with the characteristics of the material or surface.

The results of several of the assessments carried out in the course of this work have shown that additional spectral datasets are required for performance evaluations of this kind. ' Due to the lack of spatio-chromatic complexity in the dataset, it was not possible to illustrate the theoretical conditions that define the limitations of an ordering relation. These strategies were successful in achieving high scores on the aforementioned criteria because they made use of fundamental datasets. In spite of this, the CRA ordering method is the first one to meet all of the required limits, as the results of this study make clear. In addition to that, it does very well in quantitative evaluations. As a result, it is suitable for application in the processing and interpretation of hyperspectral images used in metrology.

5. CONCLUSION

In comparison, applications of graph spectral techniques to image processing have garnered considerably less attention in recent years. This is in contrast to DGDCT signal processing for vast data networks, which has gained a great deal of interest. The primary focus of this work is the application of DGDCT spectral algorithms for the purposes of image reduction, restoration, filtering, and segmentation. Because digital images naturally exist on a discrete 2D grid, one of the fundamental challenges of graph-based image processing is choosing the appropriate graph to represent the image structure for the graph-based tools that operate on top of it. This is a challenge that must be overcome before graph-based image processing can be implemented. Because of the description of the graph, there is an additional amount of side information that must be coded for compression. In edge restoration, filtering, and segmentation processes, edge weights communicate either the local signal similarity or a higher-level context (such as saliency). The subsequent steps consist of building task-specific graph topologies, such as those for image enhancement, while also attempting to strike a balance between concerns regarding performance and computational complexity.

REFERENCES

[1] Naoto Yokoya, Claas Grohnfeldt and Jocelyn Chanussot. "Hyperspectral and Multispectral Data Fusion: A Comparative Review of the Recent Literature", *IEEE Geoscience and Remote Sensing Magazine*, Vol. 5, No. 2, pp. 29-56, 2017.

[2] Renwei Dian, Shutao Li, Leyuan Fang, Ting Lu and Jose M. Bioucas-Dias, "Nonlocal Sparse Tensor Factorization for Semiblind Hyperspectral and Multispectral Image Fusion", *IEEE Transactions on Cybernetics*, Vol. 50, No. 10, pp. 4469-4480, 2019.

[3] Xuelong Li, Yue Yuan and Qi Wang, "Hyperspectral and Multispectral Image Fusion via Nonlocal Low-Rank Tensor Approximation and Sparse Representation", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 59, No. 1, pp. 550-562, 2020.

[4] Mishra, P., & Herrmann, I. (2021). GAN meets chemometrics: Segmenting spectral images with pixel2pixel image translation with conditional generative adversarial networks. *Chemometrics and Intelligent Laboratory Systems*, *215*, 104362.

[5] Fei Ma, Feixia Yang, Ziliang Ping and Wenqin Wang, "Joint Spatial-Spectral Smoothing in a Minimum-Volume Simplex for Hyperspectral Image Super-Resolution", *Applied Sciences*, Vol. 10, No. 1, pp. 237-256, 2019.

[6] D. Hong, N. Yokoya, J. Chanussot and X. Zhu, "An Augmented Linear Mixing Model to Address Spectral Varialbilty for Hyperspectral Unmixing, Geography", *IEEE Transactions on Image Processing*, Vol. 54, No. 3, pp. 1-17, 2018.

[7] Naoto Yokoya, Takehisa Yairi and Akira Iwasaki, "Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 50, No. 2, pp. 528-537, 2012.

[8] Li Sun, Kang Zhao and Ziwen Liu, "Enhancing Hyperspectral Unmixing With Two-Stage Multiplicative Update Nonnegative Matrix Factorization", *IEEE Access*, Vol. 7, pp. 171023-171031, 2019.

[9] R. Dhaya, "Hybrid Machine Learning Approach to Detect the Changes in SAR Images for Salvation of Spectral Constriction Problem", *Journal of Innovative Image Processing*, Vol. 3, No. 2, pp. 118-130, 2021.

[10] C. Yu, Y. Liu and Z. Hu, Z. (2022). Multi-branch Feature Difference Learning Network for Cross-Spectral Image Patch Matching. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 87, pp. 1-14, 2022.