ALL-OPTICAL KARHUNEN-LOEVE TRANSFORM USING MMI COUPLERS FOR IMAGE PROCESSING APPLICATIONS

Bui Thi Thuy[1], Dang The Ngoc[2], Le Trung Thanh[3*]
[1]*Hanoi University of Natural Resources and Environment*
[2]*Posts & Telecommunications Institute of Technology (PTIT)*
[3]*International School (VNU-IS), Vietnam National University, Hanoi*

ARTICLE INFO

Received: 13/8/2022
Revised: 27/8/2022
Published: 29/8/2022

KEYWORDS

All-optical signal processing
Data compression
Image processing
Karhunen-Loeve transform
Multimode interference

ABSTRACT

In this study, we present a method for all-optical image compression using all-optical Karhunen-Loeve transform (KLT). The KLT was designed in all-optical domain using only one multimode interference (MMI) coupler. The Karhunen-Loeve transform (KLT) is very attractive for image processing applicaons due to its computing efficiency, residual correlation, and rate distortion criteria benefits. The restricted multimode interference coupler with suitable locations of input and output ports was used to realize the KLT. The numerical simulations were applied to design and analyze the new hardware architecture of the KLT. The image compression has succesfully realized on the all-optical KLT within the RGB visible wavelength with high fabrication tolerance ± 2µm in the MMI length. The signal procesisng is performed in all-optical domain and can be integrated into a single chip within an AI camera. The proposed method can process image signals directly in optical domain, so it can improve signal processing speed and reduce the energy consumption.

BỘ BIẾN ĐỔI KLT TOÀN QUANG ỨNG DỤNG CHO XỬ LÝ ẢNH

Bùi Thị Thùy[1], Đặng Thế Ngọc[2], Lê Trung Thành[3*]
[1]*Trường Đại học Tài nguyên và Môi trường Hà Nội*
[2]*Học viện Công nghệ Bưu chính Viễn thông*
[3]*Trường Quốc tế, Đại học Quốc gia Hà Nội*

THÔNG TIN BÀI BÁO

Ngày nhận bài: 13/8/2022
Ngày hoàn thiện: 27/8/2022
Ngày đăng: 29/8/2022

TỪ KHÓA

Xử lý tín hiệu toàn quang
Nén dữ liệu
Xử lý dữ liệu
Biến đổi KLT
Giao thoa đa mode

TÓM TẮT

Trong nghiên cứu này, chúng tôi đề xuất một phương pháp mới để nén ảnh trong miền toàn quang sử dụng kỹ thuật biến đổi KLT. Biến đổi KLT được thực hiện chỉ dùng một cấu trúc giao thoa đa mode MMI trong miền toàn quang. Biến đổi Karhunen-Loeve (KLT) rất hấp dẫn đối với các ứng dụng xử lý hình ảnh do hiệu quả tính toán, tương quan dư và lợi ích của tiêu chí biến dạng tỷ lệ. Cấu trúc giao thoa sử dụng hiệu ứng giao thoa giới hạn với việc thiết kế vị trí cổng vào và ra phù hợp để tạo được biến đổi KLT. Kỹ thuật nén ảnh toàn quang sử dụng KLT đã được thiết kế thành công trong dải bước sóng RGB nhìn thấy với độ chính xác cao, dải sai số chế tạo ±2 µm trong chiều dài MMI. Việc xử lý tín hiệu trong miền toàn quang như vậy có thể tích hợp được trên một chip đơn trong các camera trí tuệ nhân tạo trong tương lai. Phương pháp được đưa ra trong nghiên cứu này xử lý tín hiệu ảnh trực tiếp trong miền quang, do đó nâng cao tốc độ xử lý và giảm công suất tiêu thụ.

DOI: https://doi.org/10.34238/tnu-jst.6360

* Corresponding author. *Email: thanh.le@vnu.edu.vn*

1. Introduction

Signal transforms have attracted significant interest for use in data compression, image processing, and other signal processing applications. Among several signal transforms, the Karhunen-Loeve transform (KLT) is regarded as the best due to its computing efficiency, residual correlation, and rate distortion criteria benefits.

Orthogonal signal transforms such as Fourier transform, discrete cosine transform and discrete sine transform are very useful in signal processing and communication systems [1]. Among many transforms, the Karhunen-Loeve transform is well-known for data compression and filtering and known to be optimal in sense that they yield uncorrelated data, simplifying succeeding operations.

While discrete wavelet transform (DWT) is used for image compression, KLT is used for image decorrelation; that is, KLT is employed within compression methods of many images with a high degree of mutual correlation, such as frames of medical imaging and video hyperspectral imagery [2]. In recent years, several attempcaats have been undertaken to compress such data sets as effectively as possible. The goal is to create a data representation that simultaneously takes into consideration both the benefits and drawbacks of KLT for the most effective compression based on optimum decorrelation. In all cases, the KLT is used to decorrelate in the spectral domain. All images are first decomposed into blocks, and each block uses its own KLT instead of one single matrix for the whole image.

The objective of image compression is to save a picture in a form that uses fewer bits to encode than the original image. This is conceivable because photos in their "raw" form include a significant amount of duplicate data. The majority of images are not comprised of random intensity changes. Every visual picture has some type of structure. Consequently, there is some association between adjacent pixels. If it is possible to discover a reversible transformation that eliminates duplication by decorrelating the data, a picture may be stored more effectively. The linear transformation that does this is the Karhunen-Loève Transform.

In the event of low resolution and low bit rate compression, this approach is inferior than the standard JPEG. In the case of high quality photos at high bit rates, however, the quantity of side information becomes relatively tiny in comparison to the amount of primary information. In addition, the capture of multiview photos is becoming more significant in the present day, particularly for enhancing the value of virtual reality (VR) and augmented reality (AR) systems.

The KLT, also known as Principal Components Analysis (PCA), is a method often used to reduce multidimensional data sets to lower dimensions for the purposes of analysis, compression, or classification. PCA (Principal component analysis) requires the calculation of eigenvalue decomposition or singular value decomposition of a data collection, often after mean centering. However, it should be emphasized that utilizing KLT for tasks like as pattern recognition or image processing might be difficult since it handles data as one-dimensional while they are in reality two-dimensional. Because of this, almost all established methods utilize some kind of dimensionality pre-reduction, in many instances ignoring the spatial relationships between pixels. Two-dimensional transformation based on PCA is one of the available options (and KLT).

For high speed signal procesisng, it is expect to process signals in all-optical domain [3]. In the literature, there are some method for all-optical signal processing. Most of them are based on optical fibers or lens [4]. Another important approach for realizing all-optical orthogonal transforms is to use multimode interference structures due to their advantages of compactness, good fabrication tolerances and ease of integration. The realizations for such Haar transform, Hadamard transform, discrete Hartley transform, DFT and the discrete cosine transform using MMI structures have been reported in the literature [5] - [8].

In this paper, it is shown further that the realization of the KLT for image processing applications in all-optical domain is possible. We use a NxN restricted multimode interference structure to realize the KLT. Image data is directly processed in the all-optical domain and it does not need to be converted to digital electronic signals. We found that an analytical expression for the transfer matrix of the MMI coupler can be derived, which is the characteristic matrix of the KLT if phase shifters are placed at the input and output ports of the MMI structure. The proposed devices are then verified and designed optimally using the numerical simulation tools.

2. Theory of KLT transform for image compression using optical MMI coupler

The conventional MMI coupler has a structure consisting of a homogeneous planar multimode waveguide region connected to a number of single mode access waveguides [9]. Figure 1 shows a structure of a rectangular NxN MMI coupler, where W_{MMI} and L_{MMI} are the width and length of the MMI coupler, respectively. The MMI region is sufficiently wide to support a large number of lateral modes (in the y direction).

Figure 1. *The structure of a 4x4 RI-MMI coupler for the KLT transform*

The access waveguides are identical single mode waveguides with width W_a and they run parallel to the z axis on a constant pitch $p = W_{MMI}/(N+1)$. The center line of the ith waveguide is at x=ip (i=1,2,…,N). The electrical field inside the MMI coupler can be expressed by [10]

$$E(x,z) = \exp(-jkz)\sum_{m=1}^{M} E_m \exp(j\frac{m^2\pi}{4\Lambda}z)\sin(\frac{m\pi}{W_{MMI}}x) \tag{1}$$

where $k = 2\pi n/\lambda$, λ is the operating wavelength, n is the waveguide refractive index and M is the total number of guided modes in the MMI coupler, E_m is the summation coefficients. We have the orthogonal set relating the internal modes field to the outer input-output field

$$V_{ir} = \frac{2}{\sqrt{2N+2}}\sin(\frac{ir\pi}{N+1}) \tag{2}$$

where V_{ir} is the element on row i and column r of a matrix $\mathbf{V_N}$, which relates the propagation modes inside the waveguide to the output field.

It is assumed that the length of the MMI coupler is set to $L_{MMI} = 2\Lambda/(N+1)$, where $\Lambda = nW_{MMI}^2/\lambda$. If the common phase term in equation (1) is not considered, the i^{th} propagating modes will experience different phase shift of $r^2\pi/(2N+2)$ and the matrix $\mathbf{V_N^T}$ is then multiplied by a diagonal matrix with the diagonal elements

$$b_{rr} = \exp(j\frac{r^2}{2N+2}) \tag{3}$$

The total transfer matrix of the waveguide from input to the output ports now can be calculated by

$$\mathbf{M} = \mathbf{VBV^T} \tag{4}$$

This equation can be rewritten by

$$M_{uv} = 2j\frac{\exp(j\frac{\pi}{4})}{\sqrt{2N+2}}\sin(\frac{\pi uv}{N+1})\exp(-j\frac{\pi(u^2+v^2)}{2N+2}) \quad (5)$$

If the phase shifters are added to the input ports and output ports of the MMI structure as shown in Fig. 2, the total transfer matrix can be calculated by

$$\mathbf{T} = \mathbf{D_{out}MD_{in}} \quad (6)$$

Where $\mathbf{D_{in}}$ and $\mathbf{D_{out}}$ are the matrices indicating the contribution of the input and output phase shifters arrays. If the phase shifter are set to be $\pi i^2/(2N+2)$, i=1, 2,…,N, at the input and output waveguides, the total transfer matrix **T** can be computed by

$$T_{uv} = 2j\frac{\exp(j\frac{\pi}{4})}{\sqrt{2N+2}}\sin(\frac{\pi uv}{N+1}) \quad (7)$$

This is the KLT if the phase constant coefficient $j\exp(j\frac{\pi}{4})$ is neglected. The matrix of the KLT transform formed from the 4x4 MMI structure above can be calculated by

$$M_{KLT} = \begin{bmatrix} 0.3717 & 0.6015 & 0.6015 & 0.3717 \\ 0.6015 & 0.3717 & -0.3717 & -0.6015 \\ 0.6015 & -0.3717 & -0.3717 & 0.6015 \\ 0.3717 & -0.6015 & 0.6015 & -0.3717 \end{bmatrix} \quad (8)$$

It can be expressed in the form of the normalized power as follows

$$M_{KLT} = \begin{bmatrix} 0.1382 & 0.3618 & 0.3618 & 0.1382 \\ 0.3618 & 0.1382 & 0.1382 & 0.3618 \\ 0.3618 & 0.1382 & 0.1382 & 0.3618 \\ 0.1382 & 0.3618 & 0.3618 & 0.1382 \end{bmatrix} \quad (9)$$

The KLT transform refers to multi-resolution approximation expressions [2]. In practice, multi-resolution analysis is performed using 4 channel filter banks (for each level of decomposition) consisting of a low-pass and a high-pass filter, and each filter bank is sampled at half the rate (1/2 down sampling) of the previous frequency. By repeating this method, any order of wavelet transform is achievable. The down sampling method maintains the scaling parameter (equal to 1/2) during consecutive wavelet transforms, which improves computer execution. In the case of an image, filtering is carried out independently by filtering the lines and columns. The part at each scale is decomposed recursively and illustrated in Figure 2.

The KLT begin with the covariance matrix of the vectors x generated between values of pixel with similar allocation in all arranged sub-blocks of the matrix as shown in Figure 3. The covariance matrix is $C_x = E\{(x-m_x)(x-m_x)^T\}$; where $x = (x_1, x_2, x_3, x_4)^T$ is the correlated original set, T is the transpose, m_x is the mean vector and E is the expected value of the argument. As the result, the KLT is expressed by:

$$X = V^T(x - m_x) \quad (10)$$

where $X = (X_1, X_2, X_3, X_4)^T$ is one of the decorrelated transformed vector set and V is the matrix columns and the eigenvectors of C_x. As an example, we work on the image of "camera man" of 512-by-512 pixels with sub-blocks of 4x4 pixels formed from the 4x4 MMI coupler in all-optical domain.

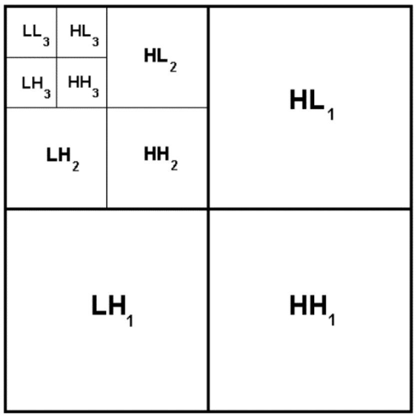


Figure 2. *Data preparation for the image (L-low pass and H-high pass)*

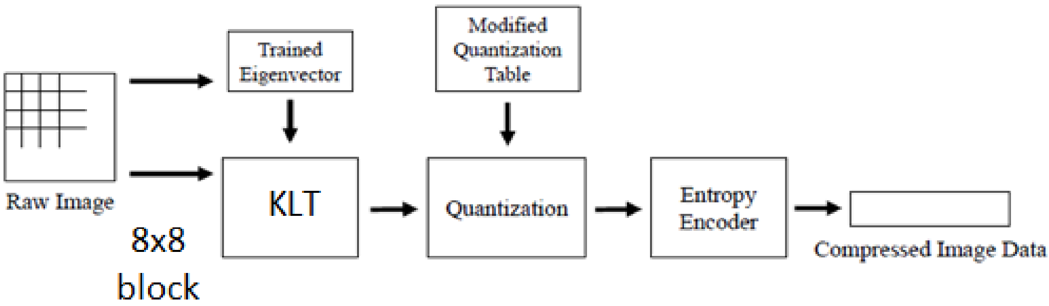


Figure 3. *Principle of the image compression using the KLT transform*

3. Simulation results and discussion

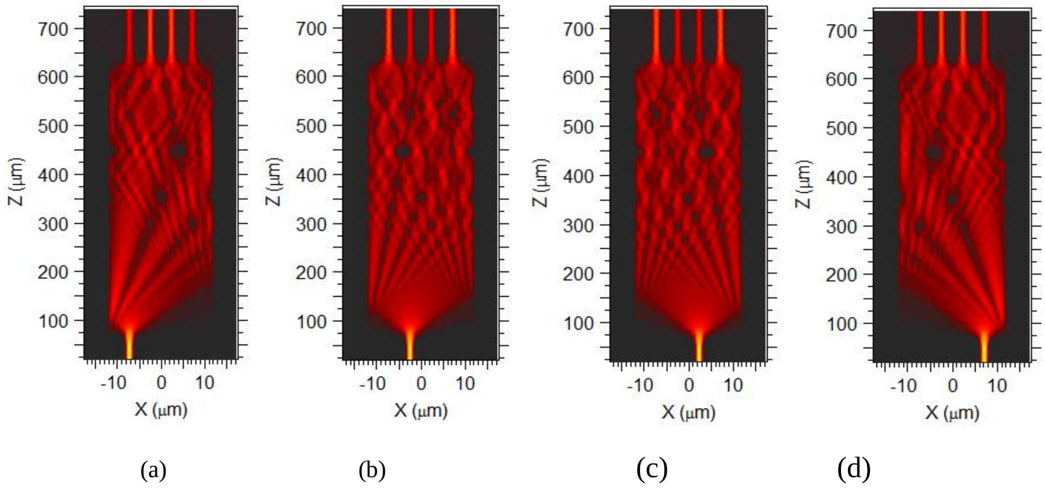
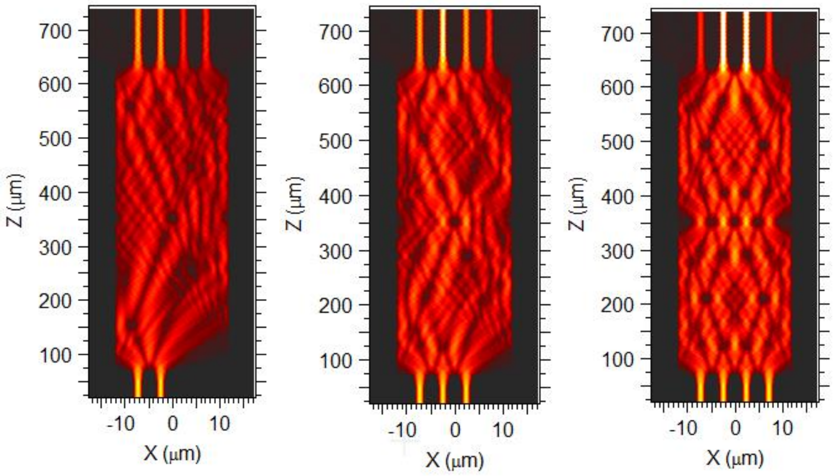


Figure 4. *Simulation result of the KLT for different image data*

The numerical simulation results for the 4x4 KLT transform are shown in Figure 5. In this study, we use the Si3N4 platform using CMOS technology fabrication process for optical device design. The optimal length and width of the 4x4 MMI coupler calculated to be 567 µm and 24 µm. In this design we use the Si3N4 material, which can work with the color images with RGB wavelength range. Figure 4(a), (b), (c) and (d) shows for input data $(x_0x_1x_2x_3)^T$ = (1000), (0100),

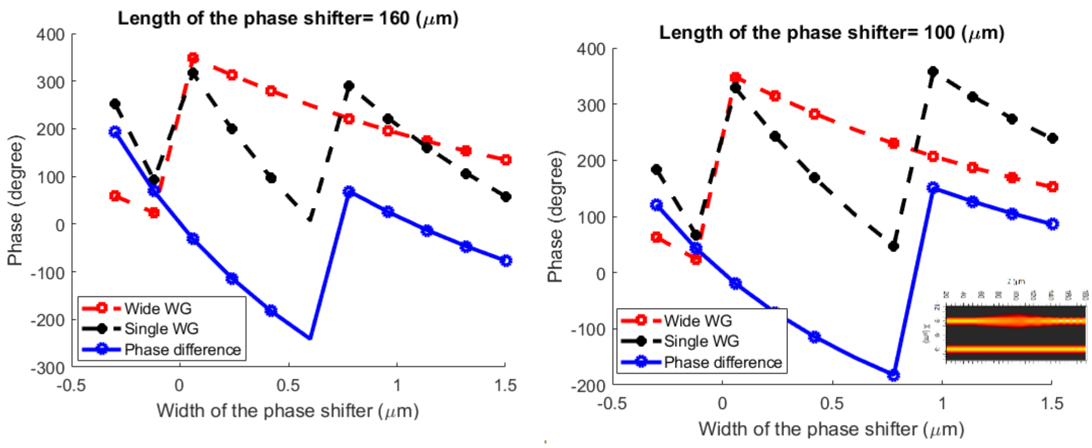
(0010) and (0001), respectively. Figure 5 presents the signal processing over the KLT transform with input data (1100), (1110) and (1111), respectively. Here, the power is represented for gray level of pixel in the image. The output amplitudes and phases at the output ports are calculated by using the numerical methods.



(a) (b) (c)

Figure 5. *Simulation result of the KLT for different image data with 2, 3 and 4 inputs*

Here we also design the phase shifters using the wide waveguides used at the input and output ports of the MMI structure to create the KLT transform as shown in Figure 1. The phase shifter can be achieved by selecting the suitable width and length of the waveguide used for the phase shifting region. Figure 6(a) and (b) show the phase shift obtained from two waveguides.



(a) (b)

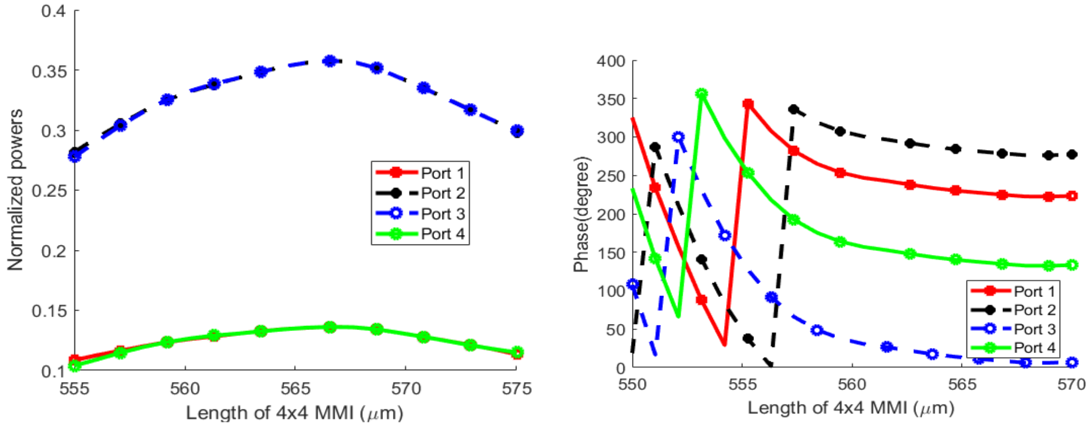
Figure 6. *Phase shifts achieved from the wide waveguide (a) with a length of 160μm and (b) length of 100μm*

Based on the presented BPM simulations, it obviously showed that (8) accurately predict the wave propagation characteristics of the MMI structure with phase shifters and the KLT can be accurately realized.

For all-optical signal processing, the KLT formed from the 4x4 MMI structure of Figure 1 can be expressed by the following matrix:

$$H_{KLT} = \begin{bmatrix} \alpha_{11}e^{j\pm\delta} & \alpha_{21}e^{j\pm\delta} & \alpha_{31}e^{j\pm\delta} & \alpha_{41}e^{j\pm\delta} \\ \alpha_{12}e^{j\pm\delta} & \alpha_{22}e^{j\pm\delta} & \alpha_{32}e^{j\pm\delta} & \alpha_{42}e^{j\pm\delta} \\ \alpha_{13}e^{j\pm\delta} & \alpha_{23}e^{j\pm\delta} & \alpha_{33}e^{j\pm\delta} & \alpha_{43}e^{j\pm\delta} \\ \alpha_{14}e^{j\pm\delta} & \alpha_{24}e^{j\pm\delta} & \alpha_{34}e^{j\pm\delta} & \alpha_{44}e^{j\pm\delta} \end{bmatrix} \quad (13)$$

Where α_{ij} and δ are variations in amplitudes and phases at the output ports. The fabrication tolerance has effects on these values. In the next section, we show that our proposed architecture can provide a very low fabrication tolerance. As a result, the KLT transform can be implemented accurately. The bock of 4 bits in image signals represented at input x1, x2, x3, x4. The processed signals after the KLT are presented at output ports X1, X2, X3, X4.



(a) (b)

Figure 7. (a) *Normalized power and (b) Phase at output ports of the KLT transform*

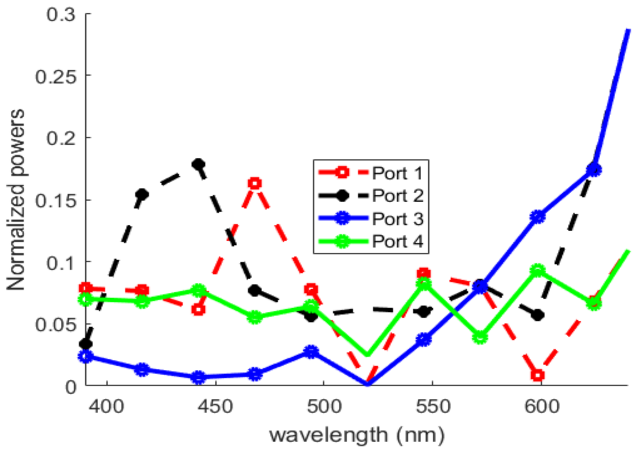
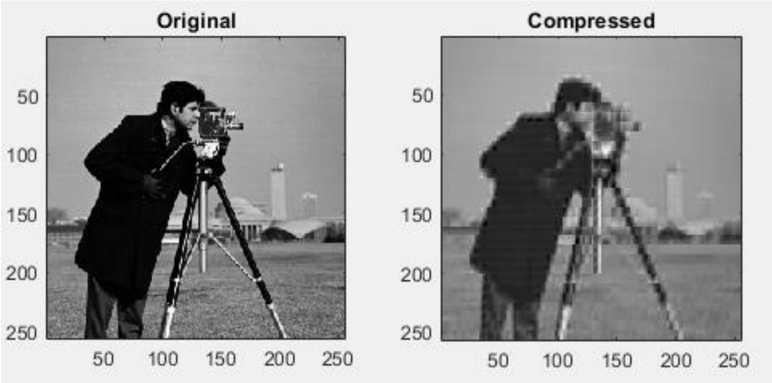


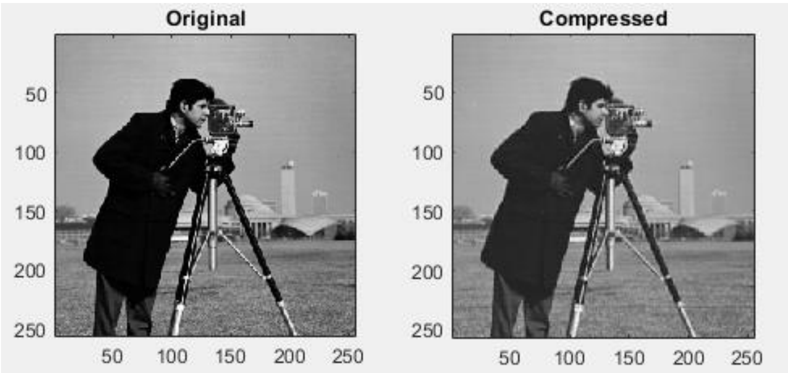
Figure 8. *Normalized power at ports 1-4 over the wavelength RGB*

The normalized powers at output ports 1-4 when input signal is at 1, 2, 3 and 4 are shown in Figure 7(a). Figure 7(b) show the phases at the output ports of the KLT device based on 4 x 4 MMI coupler at different MMI lengths. The simulations show that the length variation of ±2 µm is still keep the output powers unchanged. Figure 8 shows the variation of the power within the visible wavelength range of R = 532 nm, G = 635 nm and B = 405 nm. The fluctuation of the normalized powers is in the range of 0-0.2. This means that the fabrication tolerance of the proposed structure is high. The current CMOS fabrication technology for VLSI industry is feasible.

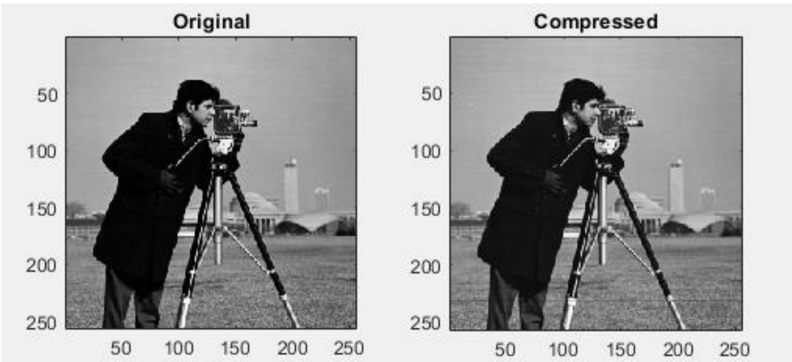
Next we undertake the simulations for input image of the "camera man", 256x256 in size at different compressed ratios of 10%, 20%, 70% with the optical KLT Transform architecture designed above. The simulation results are shown in Figure 9. The simulation results show that the compressed images are obtained from the all-optical KLT transform.



(a) "Camera man" image, CR=70%



(b) "Camera man" image, CR=20%



(c) "Camera man" image, CR=10%

Figure 9. *Original and compressed images*

4. Conclusion

We have proposed a method for realizing all-optical Karhunen-Loeve transform using multimode interference structures on an SOI platform. The designs of the proposed devices have been performed using the transfer matrix method and the beam propagation method. This all-optical approach for the KLT realization can be useful for all-optical signal processing applications such as data compression, filtering and coding. The all-optical KLT based on only one 4x4 MMI coupler for image compression. The KLT has been successfully designed in the Si3N4 material platform which is suitable for VLSI and FPGA circuits. This all-optical approach for the KLT realization can be useful for all-optical high speed and real-time image processing applications such as data compression, filtering and coding. The method can also be useful for the integration of the fast image processing into the AI camera in the future.

REFERENCES

[1] K. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Academic Press, 2007.

[2] A. P. Radünz, F. M. Bayer, and R. J. Cintra, "Low-complexity rounded KLT approximation for image compression," *Journal of Real-Time Image Processing,* vol. 19, no. 1, pp. 173-183, 2022, doi: 10.1007/s11554-021-01173-0.

[3] F. P. Sunny, E. Taheri, M. Nikdast, and S. Pasricha, "A Survey on Silicon Photonics for Deep Learning," *J. Emerg. Technol. Comput. Syst.*, vol. 17, no. 4, 2021, Art. no. 61, doi: 10.1145/3459009.

[4] A. A. Fashi, M. H. V. Samiei, C. Pinho, and A. L. Teixeira, "Photonic Integrated Chip on TriPleX Platform for Realizing Optical Haar Transform and Compression in the Visible Spectrum," *IEEE Journal of Quantum Electronics*, vol. 57, no. 5, pp. 1-10, 2021, doi: 10.1109/JQE.2021.3102845.

[5] L. W. Cahill and T. T. Le, "Photonic Signal Processing using MMI Elements," in *10th International Conference on Transparent Optical Networks (ICTON 2008)*, Athens, Greece, 22-26 June, 2008.

[6] L. W. Cahill and T. T. Le, "Optical Signal Processing Based on Planar Waveguides," in *VII International Workshop on Optical Waveguide Theory and Numerical Modelling*, Einhoven, The Netherlands, 2008.

[7] L. Cahill and T. Le, "The design of signal processing devices employing SOI MMI couplers," in *Paper 7220-2, Integrated optoelectronic devices (OPTO 2009), Photonics West, Proceedings of the SPIE*, San Jose Convention Center, San Jose, California, USA, 24 - 29 January 2009.

[8] A. R. Gupta, K. Tsutsumi, and J. Nakayama, "Synthesis of Hadamard Transformers by Use of Multimode Interference Optical Waveguides," *Appl. Opt.*, vol. 42, pp. 2730-2738, 2003.

[9] T.-T. Le, *Multimode Interference Structures for Photonic Signal Processing: Modeling and Design*. Germany: Lambert Academic Publishing, May 2010.

[10] J. M. Heaton and R. M. Jenkins, "General matrix theory of self-imaging in multimode interference(MMI) couplers," *IEEE Photonics Technology Letters*, vol. 11, no. 2, pp. 212-214, Feb. 1999.