Dictionary based Image Compression via Sparse Representation

Arabinda Sahoo¹, Pranati Das²

¹Department of ECE, ITER, Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, India ²Department of Electrical Engineering, IGIT, Sarang, Odisha, India

Article Info ABSTRACT

Article history:

Received Feb 27, 2017 Revised May 25, 2017 Accepted Jun 15, 2017

Keyword:

Dictionary learning Image compression Intra prediction K-SVD Sparse representation Nowadays image compression has become a necessity due to a large volume of images. For efficient use of storage space and data transmission, it becomes essential to compress the image. In this paper, we propose a dictionary based image compression framework via sparse representation, with the construction of a trained over-complete dictionary. The overcomplete dictionary is trained using the intra-prediction residuals obtained from different images and is applied for sparse representation. In this method, the current image block is first predicted from its spatially neighboring blocks, and then the prediction residuals are encoded via sparse representation. Sparse approximation algorithm and the trained overcomplete dictionary are applied for sparse representation of prediction residuals. The detail coefficients obtained from sparse representation are used for encoding. Experimental result shows that the proposed method yields both improved coding efficiency and image quality as compared to some state-of-the-art image compression methods.

> Copyright © 2017 Institute of Advanced Engineering and Science. All rights reserved.

1964

Corresponding Author:

Arabinda Sahoo, Departement of Electronics and Communication Engineering, ITER, Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, India. Email: arubabu123@yahoo.co.in

1. INTRODUCTION

Images compression [1] has always been important, and nowadays it becomes essential due to the huge requirement of image storage and transfer. Over the last two decades, numerous and diverse image compression methods [2-6] have been proposed. The most widely used methods are based on transform-based coding. Based on transform-based approaches by far many image compression standards like JPEG [2] and JPEG2000 [3] have been developed. The widely used JPEG and JPEG2000 standards utilize discrete cosine transform (DCT) and wavelet transform to compressively represent images. However, JPEG and JPEG2000 do not take spatial correlation of neighboring blocks into consideration for more compact representation of images [7]. In recent years, intra-prediction [8], [9] schemes serve as a promising direction, which use neighboring blocks in the compression process for more compression. The idea of intra-prediction is to predict the unknown image block based on the knowledge of decoded neighboring blocks. A good prediction reduces the overall coding rate.

However, these compression methods suffer some limitations, because they are not able to efficiently compress some specific classes of images. They are not able to sparsely represent complex characteristics of an image. To overcome these limitations, sparse representations [10] has been evolving in recent years. Sparse representation is a very effective tool for compressing a large variety of images. This is possible due to the fact that the images can be sparse or compressible with respect to some basis or dictionary [11]. Thus, sparse representation provides a potential for effective image compression. An image is compressible or not, it depends on the dictionary, so the design of dictionary is vital in sparse representation.

In contrast to fixed DCT and wavelet dictionary, the latest trend of image compression techniques is extended to design trained dictionaries [12]. Numerous dictionary based image compression methods have been proposed for sparse representation. In several recently published works, the use of learned over-complete dictionaries in image compression has shown promising results at low bit rate as compared to fixed dictionaries. The initial proposal towards dictionary based image compression was proposed by Bryt and Elad [13]. They proposed an algorithm for the compression of facial image based on the learned K-SVD [14] over-complete dictionary. Though this method outperforms JPEG and JPEG2000 but this method is limited to compressing facial images. In [15], the author proposed an image compression method based on the iteration-tuned dictionary (ITD). In this scheme, the dictionary consists of a layer structure with each layer trained for a specific class of images and carries a separate dictionary matrix. This method is shown to outperform K-SVD over-complete dictionary method but it is employed to compress a specific class of images. In [16], the author proposed an adaptive dictionary design method for the fingerprint image compression. They used a set of fingerprint image to learn a dictionary. In [17], the author proposed a dictionary based method for compressing surveillance image.

Experimental result shows these above proposed algorithms outperform JPEG and JPEG2000. However, most of the above mentioned compression scheme is either related to facial images or to some specific class of images, while there is a lack of research on general arbitrary images. So, the main challenge of above schemes is the compression of general arbitrary images. To address this, in this paper we proposed a novel image compression scheme for arbitrary images. In this scheme, an efficient trained over-complete dictionary is integrated into the intra-prediction framework. The conventional transform-domain representation of intra-prediction scheme is replaced by a trained over-complete dictionary. We trained a dictionary offline using the residuals obtained from intra-prediction. This dictionary can sparsely represent the complex characteristics of the residual block. The coefficients and indices of appropriate dictionary is shared at both encoder and decoder, only coefficients and indices of dictionary elements need to be encoded, which compress the image significantly. Experimental results on the arbitrary images shows that the proposed method yields both improved coding efficiency and image quality as compared to JPEG and JPEG 2000.

The rest of the paper is organized as follows. In Section 2, we present some preliminaries on sparse representation and dictionary design. The proposed image compression method is introduced in section 3. Section 4 illustrates experimental results and discussion. Finally, Section 5 concludes the paper.

2. PRELIMINARIES

In sparse representation [18], a signal $y \in \mathbb{R}^n$ can be represented by a linear combination of a small number of signals known as atoms taken from an over-complete dictionary $\mathbf{D} \in \mathbb{R}^{n \times k}$. It is called sparse representation as it employs only a few number dictionary atoms or elements to represent the signal. A signal is said to be compressible if it can be represented by few dictionary atoms. Mathematically sparse representation can be expressed as:

$$y = \mathbf{D}x \tag{1}$$

The solution vector $\mathbf{x} \in \mathbf{R}^k$ contains the coefficients of the signal \mathbf{y} . Where \mathbf{D} is one $\mathbf{n} \times \mathbf{k}$ matrix with $\mathbf{n} < \mathbf{k}$ called over-complete dictionary and each column of \mathbf{D} is called an atom. If \mathbf{y} and each atom of \mathbf{D} are treated as a signal then \mathbf{y} can be represented as a linear combination of atoms of \mathbf{D} . This linear combination can be expressed as solution vector \mathbf{x} . Due to over complete nature of \mathbf{D} , an infinite number of solution exist for \mathbf{x} . In last decade, various sparse approximation algorithms have been proposed to find out the sparse solution for \mathbf{x} . The sparse approximation algorithm always aims to represent \mathbf{y} in terms of minimum number atoms. Mathematically, this can be expressed as solving Equation (1) such that the solution \mathbf{x} contains minimum number of non-zero elements. A signal is said to be compressible if the number of non-zero elements in \mathbf{x} is very less as compared to number of elements of \mathbf{y} . The sparse representation of \mathbf{y} may be either exact $\mathbf{y} = \mathbf{D}\mathbf{x}$ or approximate, $\mathbf{y} \approx \mathbf{D}\mathbf{x}$ satisfying $|| \mathbf{y} - \mathbf{D}\mathbf{x} ||_p \le \delta$. Where \mathbf{p} is the norm. Typical norms used in the approximation are 1, 2 or ∞ . Normally $\mathbf{p} = 2$ is taken in image compression. Mathematically, the sprarsest representation of either Equations (2) or (3)

$$\min_{\boldsymbol{x}} \| \boldsymbol{x} \|_{\boldsymbol{0}} \quad \text{subject to } \boldsymbol{y} = \boldsymbol{D}\boldsymbol{x} \tag{2}$$

$$\min_{x} \| x \|_{0}$$
 subject to $\| y - \mathbf{D}x \|_{2} \le \delta$ for some $\delta \ge 0$

(3)

or

where $\|.\|_{0}$ is the l_{0} norm, represents the number of non-zero elements in solution vector x. Some of the well known algorithms used to find sparse representation are: Matching pursuit (MP) [19], Orthogonal Matching Pursuit (OMP) [20] and Complimentary matching pursuit (CMP) [21]. In this paper, we focus on OMP algorithm due to its efficiency.

The dictionary based image compression can also be effectively modeled by Equations (2) and (3). In image compression, we consider a set of **k** image blocks of size **n** pixels with n < k, ordered lexicographically as column vectors of dictionary **D**. A column vector **y** is obtained from an image block of size **n** pixels. In sparse representation, the problem is to find out the solution vector **x** which will represent **y** with least number of dictionary elements. Indeed, compression of image patch **y** can be achieved by transmission of nonzero elements of vector **x**, by specifying their coefficients and indices.

The dictionary plays an important role in a successful image compression modeling via sparse representation. Am image is compressible if it can be represented by few number of dictionary elements. The dictionary can either be chosen as a prespecified set of images or designed by adapting its contents to fit a given set of images. The objective of dictionary design is to train the dictionary which able to represent a signal set sparsely [12]. Given an image set $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^N$, dictionary design aims to find the best dictionary \mathbf{D} that gives rise to sparse solution for each \mathbf{y}_i . In other words, there exists D, such that solving Equation (2) for each \mathbf{y}_i gives a sparse representation \mathbf{x}_i . The minimization problem to find the best dictionary for sparse representation of \mathbf{Y} in the given sparsity constraint \mathbf{T}_0 can be represented by:

$$\min_{\mathbf{D},\mathbf{X}} \| \mathbf{Y} - \mathbf{D}\mathbf{X} \|_{2} \quad \text{subject to } \forall i, \| X_{i} \|_{0} \leq \mathbf{T}_{0}$$

$$\tag{4}$$

The dictionary is trained to provide a better representation of the actual signal when the number of dictionary elements used to represent it is less than or equal to T_0 .Various algorithms have been developed to train over complete dictionaries for sparse signal representation. The K-SVD algorithm [14] is very efficient and it works well with different sparse approximation algorithm. K-SVD algorithm iteratively updates the dictionary atoms to better fit the data. In this paper, we focus on K-SVD algorithm to train the dictionary and OMP algorithm for sparse representation.



Figure 1. Detailed block diagram of the proposed method

3. PROPOSED METHOD

The proposed image compression method consists of four main processes: intra prediction, dictionary training, sparse representation, and coding. The block diagram of the proposed compression method is shown in Figure 1.

3.1. Intra Prediction

Intra prediction [8] exploits spatial correlation within one image. In this process, the current block is predicted by using the boundary pixel values of previously reconstructed neighboring blocks. As shown in the Figure 2, the boundary pixel values of neighboring blocks as shown in shaded boxes are copied into the current block pixels along a specified direction indicated by the mode. Eight directional modes and a DC mode, which is almost same as the nine intra-prediction modes employed in H.264 standard [9].



Figure 2. (a) Intra Prediction, (b) 8-Directional mode + DC Mode

In this proposed method, the image is divided into blocks of size $8 \ge 8$. The nine intra-prediction modes (0-8) are applied over each $8 \ge 8$ block and residual error is calculated for each prediction mode. The residual error is the difference between the pixel value of the current block and pixel value of the predicted block. The best prediction mode for a current block is selected based on the minimum residual error. The mode number **M** and the residual error block are transmitted for encoding.

3.2. Dictionary Training

The dictionary is trained off-line using prediction residual samples resulting from a wide variety of images. We first divide different images into blocks of size 8 x 8 and then 9 intra- prediction modes (0-8) are applied over each 8 x 8 block. The predicted block is subtracted from the current block to generate residual blocks. A set of 8×8 prediction residual blocks for different modes are selected to train the dictionary. To train the dictionary we employed K-SVD algorithm [14]. During dictionary training, in each iteration, K-SVD algorithm updates the dictionary elements by optimizing minimization problem given in Equation(4). K-SVD algorithm iteratively updates the dictionary is used for sparse representation of the residual block during image coding.

3.3. Sparse Representation

OMP algorithm is employed for the sparse representation of residual block. OMP algorithm selects the appropriate dictionary element to represent each residual image block. In each iteration, OMP selects the best linear combination of dictionary elements by minimizing Equation (3). The same process is continued and the algorithm terminates when the residual error of the reconstructed signal is equal to or less than a specified value. However, the number of OMP iterations may not exceed T_0 . Once the algorithm terminates the coefficients C and indices P of appropriate dictionary element is transmitted for encoding.

3.4. Coding

After sparse representation, the coefficients C and indices P of dictionary element, and the prediction mode M are encoded [22]. The coefficients are uniformly quantized followed by entropy coding.

The indices are encoded with fixed length codes whose sizes are $\log_2 k$, where k is the number of dictionary elements. The prediction mode numbers are encoded with fixed length codes whose sizes are 4 bits. The encoder of the proposed compression method is shown in Figure 1(a).

Since the dictionary is shared at both encoder and decoder, only intra- prediction mode number, indices, and coefficients of the dictionary elements are transmitted to the decoder. The decoder generates the residual block from the knowledge of dictionary, coefficients, and indices. The decoder then predicts the block based on mode number and combines with the residual block to reconstruct the block. The decoder of the proposed compression method is shown in Figure 1(b).

4. EXPERIMENTAL RESULTS

In this section, we conducted several experiments in order to evaluate the performance of the proposed compression method. The proposed algorithm is applied over several images. The compression efficiency and quality of the reconstructed image are compared with several other competitive image compression techniques.

4.1. Intra Prediction

In the experiment, we used 100 images from the Berkeley segmentation database as our training set. Nine modes of 8 x 8 intra prediction are applied and intra-prediction residual for each image is generated. A set of 45000 blocks of size 8 x 8 are randomly selected from residual images to train a dictionary. We selected 5000 residual blocks from each nine modes to form our learning set. Examples of intra-predicted image and residual image are demonstrated in Figure 3. A random collection of such training blocks for different mode is shown in Figure 4.



(a) Original Image



(b) Predicted Image

Figure 3. Intra prediction of second image



(c) Residual Image



(a) Mode 3



(b) Mode 6





(c) Mode 8

4.2. K-SVD Dictionary

We used 45000 residual blocks obtained from intra-prediction as our learning set. The K-SVD algorithm is applied to train an over-complete dictionary of size 64 x 512, where the number of rows 64 represents the number of pixels in a block, and the number of columns 512 represents the number of dictionary elements. In the training process, 100 number of K-SVD iterations was set as the primary stopping criterion. Sparsity constraint (T_0 =4) was set as another stopping criterion. Example of a trained dictionary is demonstrated in Figure 5.



Figure 5. Example of dictionary trained on 8x8 residual blocks

4.3. Image Compression

The K-SVD Dictionary resulting from prediction residuals is used for sparse representation. In the experiment, the residual error (δ =0.2) and the number of OMP iterations equal to 4 was set as stopping criterion of OMP algorithm. The OMP algorithm stops when the residual error is less than or equal to 0.2, otherwise it stops after 4 iterations. In maximum, four dictionary elements are required to encode each residual block.



Figure 6. Rate-distortion curves of different methods

Dictionary based Image Compression via Sparse Representation (Arabinda Sahoo)

In order to evaluate our proposed compression method, we compare our method with JPEG and JPEG2000. All experiments are performed using MATLAB. The performance of proposed method is evaluated by taking different standard test images. The quantitative evaluation of our proposed method is accomplished using two image quality metrics: PSNR and SSIM (Structural Similarity Metric) [23]. Figure 6 shows comparison of rate-distortion curve for 4 standard test images. Table 1 shows PSNR comparison for different images at two different bit-rates. The proposed method yields around PSNR gain of 3 dB compared to JPEG, and PSNR gain of 0.3 dB compared to JPEG 2000. The performance in terms of average PSNR and SSIM are shown in Table 2. The results are averaged over 9 standard test images and best results are bolded. Subjective assessment of one image is shown in Figure 7 at bit-rate 0.2. The results show that our proposed method outperforms JPEG and JPEG2000 in terms of all quality metrics, including PSNR and SSIM.

two different off futes. Dest festilis are boliced									
]	PSNR(dB) at bit-ra	te 0.2 bpp	PSNR(dB) at bit-rate 1 bpp					
Images	JPEG	JPEG2000	Proposed method	JPEG	JPEG2000	Proposed method			
Barbara	23.81	27.29	27.60	34.46	37.15	37.52			
Sailboat	26.56	29.12	29.53	34.44	36.81	37.01			
Baboon	22.48	25.24	25.45	27.98	30.65	30.81			
couple	25.72	28.48	28.80	34.25	36.79	37.10			
Hill	26.78	29.88	30.21	33.62	36.42	36.84			
Jet plane	29.56	31.92	32.28	39.65	41.88	42.20			
Lena	29.43	33.02	33.46	37.88	40.40	40.82			
Lighthouse	25.82	28.39	28.82	35.50	38.42	38.83			
Peppers	29.98	32.49	32.88	37.55	38.38	38.72			
Average	26.68	29.54	29.89	35.03	37.43	37.76			

Table 1. PSNR (dB) comparison of the proposed method with JPG and JPEG2000 for a set of 9 test images at two different bit-rates. Best results are bolded



(a) Original image

(b) JPEG

(c) JPEG2000

(d) Proposed method

Figure 7. Subjective assessment

Table 2. Performance comparison in terms of two image quality metrics, PSNR (dB) and SSIM at six
different bit-rates. The results are averaged over 9 test images. Best results are bolded

Quality Metric	JPEG	JPEG2000	Proposed method	JPEG	JPEG2000	Proposed	
						method	
Bit-rate 0.2 bpp				Bit-rate 0.4 bpp			
PSNR	26.68	29.54	29.89	28.78	31.82	32.14	
SSIM	0.748	0.771	0.782	0.762	0.844	0.861	
Bit-rate 0.6 bpp				Bit-rate 0.8 bpp			
PSNR	31.12	33.82	34.22	33.40	36.42	36.74	
SSIM	0.824	0.906	0.915	0.874	0.921	0.931	
Bit-rate 1 bpp			Bit-rate 1.2 bpp				
PSNR	35.03	37.43	37.76	36.43	38.55	38.87	
SSIM	0.909	0.941	0.949	0.918	0.951	0.954	

5. CONCLUSION

In this paper, we presented a dictionary based intra-prediction framework for image compression. K-SVD algorithm is used in order to train a dictionary. We trained the dictionary with a variety of residual blocks obtained from intra-prediction and then used this dictionary for sparse representation of an image. OMP algorithm, fixed length coding, and entropy coding have employed for encoding. Different coding results based on a set of test images are presented to compare the performance of the proposed method with the existing methods. Experimental result shows that the proposed method outperforms JPEG and JPEG2000.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Gagan Rath for his support and guidance throughout this work.

REFERENCES

- [1] Salomon, "Data Compression: The Complete Reference", third ed., Springer, New York, 2004.
- W. Pennebaker, J. Mitchell, "JPEG Still Image Data Compression Standard", Kluwer Academic Publishers, 1993.
 C. Christopoulos, A. Skodras, T. Ebrahimi, "The JPEG2000 Still Image Coding System: An Overview", IEEE
- JPEG2000 Standart", TELKOMNIKA Indonesian Journal of Electrical Engineering, vol. 16, no. 1, pp. 83-90, Oct. 2015.
- [5] R. Chourasiya, A. Shrivastava, "A Study of Image Compression Based Transmission Algorithm using SPIHT for Low Bit Rate Application", *Bulletin of Electrical Engineering and Informatics*, vol. 2, no. 2, pp. 117-122, Jun. 2013.
- [6] D. Narmadha, K. Gayathri, K. Thilagavathi, N. Sardar Basha, "An Optimal HSI Image Compression using DWT and CP", *International Journal of Electrical and Computer Engineering*, vol. 4, no. 3, pp. 411-421, Jun. 2014.
- [7] M. Stirner, G. Seelmann, "Improved Redundancy Reduction for JPEG Files", Proc. of Picture Coding Symposium (PCS 2007), Lisbon, Portugal, Nov. 7-9, 2007.
- [8] J. Yang, B. Yin, Y. Sun, N. Zhang, "A block- Matching based Intra Frame Prediction H.264/AVC", ICME, 2006.
- [9] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, A. Luthra, "Overview of the H.264/AVC Video Coding Standard", *IEEE Trans. Circuits Syst. Video Technology.*, vol. 13, no. 7, pp. 560-576, Jul. 2003.
- [10] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, S. Yan, "Sparse Representation for Computer Vision and Pattern Recognition", Proc. IEEE, vol. 98, no. 6, pp. 1031–1044, Jun. 2010.
- [11] R. Rubinstein, A. M. Bruckstein, M. Elad, "Dictionaries for Sparse Representation Modeling", Proc. IEEE, vol. 98, no. 6, pp. 1045–1057, June 2010.
- [12] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in Proc. 26th Annu. Int. Conf. Mach. Learning., 2009, pp. 689-696.
- [13] O. Bryt, M. Elad, "Compression of Facial Images using the K-SVD Algorithm", J. Vis. Commun. Image Representation. vol. 19, no. 4, pp. 270-282, 2008.
- [14] M. Aharon, M. Elad, A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *IEEE Trans. Signal Processing.*, vol. 54, no. 11, pp. 4311-4322, Nov. 2006.
- [15] J. Zepeda, C. Guillemot, E. Kijak, "Image Compression using Sparse Representations and the Iteration-tuned and Aligned Dictionary", *IEEE J. Sel. Topics Signal Processing*, vol. 5, no. 5, pp. 1061-1073, Sep. 2011.
- [16] G. Shao, Y. Wu, A. Yong, X. Liu, T. Guo, "Fingerprint Compression based on Sparse Representation", *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 489-501, Feb. 2014.
- [17] Jing-Ya Zhu, Zhong-Yuan Wang, Rui Zhong, Shen-Ming Qu, "Dictionary based Surveillance Image Compression", J. Vis. Commun. Image R. vol. 31, no. 7, pp. 225-230. Jul. 2015.
- [18] G. Rath, A. Sahoo, "A Comparative Study of some Greedy Pursuit Algorithms for Sparse Approximation", *EUSIPCO2009*, pp.398-402, Aug. 2011.
- [19] S. Mallat, Z. Zhang, "Matching Pursuits with Time Frequency Dictionaries", *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397-3415, Dec. 1993.
- [20] Y.C.Pati, R.Rezaiifar M, P.S.Krishnaprasad, "Orthogonal Matching pursuit: Recursive Function Approximation with Application to Wavelet Decomposition", In Proc. 27th Conf. on Sig. Sys. and Comp., vol.1, Nov. 1993.
- [21] G. Rath, C. Guillemot, "On a Simple Derivation of the Complimentary Matching Pursuit", *Signal processing*, vol. 90, no. 2, pp.7 02-706, Feb. 2010.
- [22] A. Bovik, Handbook of Image and Video Processing, 2nd ed.San Francisco, CA, USA: Academic, 2005.
- [23] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Processing.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

BIOGRAPHIES OF AUTHORS



Arabinda Sahoo. Currently, he is working as an Assistant Professor in the department of Electronics & Communication Engineering at ITER, Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, India. He has completed his B.E from Utkal University, Bhubaneswar, Odisha, India and M.Tech from National Institute of Technology, Rourkela, India. His research interests include signal and image processing, image compression and sparse representations.



Pranati Das. Currently, she is working as an Associate Professor in the department of Electrical Engineering at Indira Gandhi Institute of Technology, Sarang an Autonomous Institute of Government of Odisha, India. She has completed her M.Tech & PhD from Indian Institute of Technology, Kharagpur, India and Engineering graduation from Sambalpur University, Odisha, India. She has published number of research and conference papers in both national and international forum. Her area of research interests are Image Processing and Signal Processing.