A Lossless Image Compression Based on Support Vector Machine and Adaptive Predictor Combination

Chandrabhan Ghuge
M.Tech Scholar
computer science & Engg.
PCST, Bhopal, India
chandra.ghuge@gmail.com

Saurabh Mandloi
Asst.Prof.
computer science & Engg.
PCST, Bhopal, India
saurabhmamdloi@gmail.com

Parmalik Kumar
Asst.Prof.
computer science & Engg.
PCST, Bhopal, India
parmalik.kumar@patelcollege.com

ABSTRACT

Lossless image compression is need of efficient transmission of multimedia data and also improved the efficiency of memory storage. In image compression technique basically image coding technique plays an important role for speed up the image compression process. For the process of image coding various methods are used by some authors such as adaptive prediction and combination of prediction. in this paper proposed a new adaptive predication technique based on classification mechanism based on support vector machine. The support vector machine is regressive classification technique. The regression classification technique approach finds the similar and dissimilar class of consecutive pixel value of image. The difference class of image performs encoding technique for image compression. For the validation of proposed algorithm used MATLAB computational software and used some standard image such as cameraman, Barbara, and evaluate some standard parameter such as compression ratio and PSNR value of image .our empirical result shows better performance in compression of BAPC and other technique.

Keywords: *Lossless Image Compression, Adaptive Prediction, SVM*

I. INTRODUCTION

Lossless image compression provides the great facility of scientist researcher as well as many organizations plays a role in image processing. The process of lossless image compression reduces the image data size and increases the efficiency of bandwidth and storage memory area. The image compression task done by two components one is predictive coding and another one is entropy based coding technique. the predictive coding technique is very efficient coding technique for image compression. Journey of researcher various predictive coding technique are used. All these technique increases the speed of coding technique and increase the compression ratio value. The Image compression coding is to store the image into bit-stream as compact as possible and to display the decoded image in the monitor as exact as possible. Now consider an encoder and a decoder, when the encoder receives the original image file, the image file will be converted into a series of binary data, which is called the bit-stream. The decoder then receives the encoded bit-stream and decodes it to form the decoded image. If the total data quantity of the bit-stream is less than the total data quantity of the original image, then this is called image compression. The goal of lossless image compression is to represent an image signal with the smallest possible number of bits without loss of any information, thereby speeding up transmission and minimizing storage requirements. The number of bits representing the signal is typically expressed as an average bit rate (average number of bits per sample for still images, and average number of bits per second for video). The goal of lossy compression is to achieve the best possible fidelity given an available communication or storage bit rate capacity or to minimize the number of bits representing the image signal subject to some allowable loss of information. In this way, a much greater reduction in bit rate can be attained as compared to lossless compression, which is necessary for enabling many real time applications involving the handling and transmission of audiovisual information. Predictor combination is one of the key techniques in lossless image compression. By taking a combination of candidate predictors, predictor combination schemes can produces better predictive performance than any one of the candidate predictors alone. The well known Median Edge Detection predictor and the Gradient Adaptive Predictor are low complexity examples of predictor selection and can be considered as an extreme case of predictor combination. History-Based Blending (HBB) was one of the early attempts to formulate prediction as a problem of predictor combination. In theory, it can be used to combine any number of sub-predictors. The HBB algorithm penalizes candidate predictors that perform poorly in prior predictions by optimizing the combination coefficients using a least squares approach. The average compression performance of HBB was demonstrated to be better than CALIC. The use of support vector machines (SVMs) in an image compression algorithm was first presented in [3]. This method used SVM to directly model the color surface. The parameters of a neural network (weights and Gaussian centers) were transmitted so that the color surface could be reconstructed from a neural network using these parameters. SVM learning was used to directly model the color surface. In the algorithm presented in this paper, we apply SVM learning to an image after mapping the image into the frequency domain. The above section discuss introduction of lossless image compression using predictors and support vector machine. In section II we describe adaptive predictors, support vector machine and Structure reference. In section III proposed algorithm. In section IV discuss experimental result analysis. In section and finally conclude in section V.

II. ADAPTIVE PREDICTORS

The adaptive prediction combination address the following problem: given a set of L predictors p ={pi }iL=1, what is the optimal prediction for the unknown pixel x[1]. A linear additive model for each predictor is defined as

x = pi + ei

Where ei is called the prediction error for the i^{th} predictor. It is assumed that ei follows a Gaussian distribution with zero mean and variance σi2 which is denoted by the general notation P(ei) = N (ei |0,σi2). It is further assumed that errors from different predictors are independent. The optimum prediction xˆ can then be found by maximizing the likelihood function

$$\hat{x} = \arg\max_{x} \prod_{i=1}^{L} \mathcal{N}(p_i|x, \sigma_i^2) = \sum_{i=1}^{L} \alpha_i p_i$$

……..(1)

Where

$$\alpha_i = \frac{1}{S}\frac{1}{\sigma_i^2} \quad \text{.....(2)}$$

$$\hat{\sigma}_i^2 = \frac{1}{K}\sum_{k=1}^{K} e_i^2(k) \quad \text{......(3)}$$

Where the notation ei (k) represents the prediction error for a pixel at location index k, made by the ith predictor. The index k is defined with respect to the current pixel x. The K pixels form a causal neighborhood of the pixel x.

III. SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a novel machine learning method based on statistical learning theory developed by V.N.Vapnik, and it has been successfully applied to numerous classification and pattern recognition problems such as text categorization, image recognition and bioinformatics[19]. SVM can be used for pattern recognition, regression analysis and principal component analysis. The achievements of SVM in training have Platt's the sequential minimal optimization method. These methods are directed at the training process, and not related to classification process. In the process of SVM training, all the samples are used. So it has no effect on the speed of the classification. Lee and others propose a method of reduction SVM training time and adding the speed of training, reduced support vector machines. The method in the training process is not used in all the samples but by randomly selecting one of the subsets to train, which is through reducing the scale of training to achieve the objective of speeding up the training pace. At the same time, because of the reduction of the support vector quantity, the speed of classification is improved to some degree. However, due to the loss of some support vector classification, precision has declined, especially when the number of support vector is so many that the accuracy of its classification will decline. Burges put forward a way of increasing the speed of Classification ,which does not use the support vector in the category function but use a reduction of vector set, which is different from the standard vector set .That is neither training samples nor support vector but it is the transformation of the special vector. The method achieved certain results, but in the process of looking for the reduction of the vector collection, the cost of calculation paid is too large to widely use in practice. The concept of SVM is to transform the input vectors to a higher dimensional space Z by a nonlinear transform, and then an optical hyper plane which separates the data can be found. This hyper plane should have the best generalization capability. As shown in Figure 1, the black dots and the white dots are the training dataset which belong to two classes. The Plane H series are the hyperplanes to separate the two classes. The optical plane H is found by maximizing the margin value $2/\|w\|$. Hyperplanes H_1 and H_2 are the planes on the border of each class and also parallel to the optical hyperplane H. The data located on H_1 and H_2 are called support vectors.

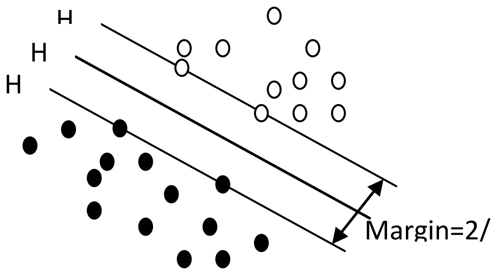

Figure 1 The SVM binary classifications

For training data set $(x_1, y_1),...,(x_l, y_l), y_i \in \{-1,1\}$, to find the optical hyper plane H, a nonlinear transform, $Z = \Phi(x)$, is applied to x, to make x become linearly dividable. A weight w and offset b satisfying the following criteria will be found:

$$\begin{cases} w^T z_i + b \geq 1, & y_i = 1 \\ w^T z_i + b \leq -1, & y_i = -1 \end{cases} \quad (1)$$

i.e.

$$y_i(w^T z_i + b) \geq 1, \qquad i = 1,2,...,l \quad (2)$$

Assume that the equation of the optical hyperplane H (Fig.1) is $w_0{}^T z + b_0 = 0$, then the distance of the data point in any of the two classes to the hyperplane is:

$$\rho(w,b) = \min_{x|y=1} \frac{z^T w}{\| w \|} - \max_{x|y=-1} \frac{z^T w}{\| w \|} \quad (3)$$

A w_0 is to be found to maximize

$$\rho(w_0, b_0) = 2 / \| w_0 \| = 2 / \sqrt{w_0{}^T w_0} \quad (4)$$

Then the search of the optimal plane H turns to a problem of a second order planning problem.

$$\min_{w,b} \Phi(w) = \frac{1}{2}(w^T w) \quad (5)$$

Subject to

$$y_i(w^T z_i + b) \geq 1, \qquad i = 1,2,...,l \quad (6)$$

If the sample data is not linearly dividable, find the minimum value of

$$\Phi(w) = \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i \quad (7)$$

Whereas ξ can be understood as the error of the classification and C is the penalty parameter for this term. By using Lagrange method, the decision function of

$$w_0 = \sum_{i=1}^{l} \lambda_i y_i z_i \quad (8)$$

will be

$$f = \text{sgn}[\sum_{i=0}^{l} \lambda_i y_i (z^T z_i) + b] \quad (9)$$

From the functional theory, a non-negative symmetrical function $K(u,v)$ uniquely define a Hilbert space H, K is the rebuild kernel in the space H:

$$K(u,v) = \sum_i \alpha \varphi_i(u) \varphi_i(v) \quad (10)$$

This stands for an internal product of a characteristic space:

$$z_i{}^T z = \Phi(x_i)^T \Phi(x) = K(x_i, x) \quad (11)$$

Then the decision function can be written as:

$$f = \mathrm{sgn}[\sum_{i=1}^{l} \lambda_i y_i K(x_i, x) + b] \quad (12)$$

The development of a SVM image classification model depends on the selection of kernel function K. There are several kernels that can be used in Support Vector Machines models. These include linear, polynomial, radial basis function (RBF) and sigmoid function:

$$K(x_i, x_j) = \begin{cases} x_i^T x_j & Linear \\ (\gamma x_i^T x_j + coefficien\ t)^{\deg ree} & Polynomial \\ \exp(-\gamma \mid x_i - x_j \mid^2 & RBF \\ \tanh(\gamma x_i^T x_j + coefficien\ t) & Sigmoid \end{cases}$$

(13)

The RBF is by far the most popular choice of kernel types used in Support Vector Machines. This is mainly because of their localized and finite responses across the entire range of the real x-axis.

Improper kernel function might generate poor performance. Currently there is no effective "learning" method to choose a proper kernel function for a specific problem. The selection is decided by the experiment result at this time. In our proposed system, two kernel functions are tested: Radial Basis Function-RBF and Polynomial Function.

$$K_{poly}(x_1, x_2) = (x_1 * x_2 + 1)^p \quad (14)$$

$$K_{RBF}(x_1, x_2) = \exp(-p \parallel x_1 - x_2 \parallel^2)$$
(15)

Due to its better performance, RBF was chosen as the kernel function in the model.

IV. PROPOSED METHODOLOGY

In this section describe the proposed algorithm for image compression using adaptive predictor and support vector machine. Adaptive predictor technique generates the set of pixel matrix for the process of image matrix. The generated matrix estimate the value of mean and variance of image pixel the variance of image pixel added the process of classification technique. in this process of classification support vector machine play an important role for classification technique. The support vector machine reduces the redundancy pixel in fashion of adaptive predicator. The algorithm divided into three sections given below.

Section first

1. Input the digital image
2. Apply linear adaptive predictor and find the resolution of matrix.
3. Estimate the mean and variance of matrix.
4. First phase of processing the variance value assigned as 0
5. Second phase the mean value assigned the 0
6. The mean and variance value pass through support vector machine.

Section two

1. Assigned the mean and variance value of image

- ➢ Transform data to the format of an SVM that is X is original data R is transform data such that Xi∈ R^d here d is dimension of data.
- ➢ Conduct scaling on the data α = $\sum_{i=1}^{m} \sum_{j=1}^{n} sim(Xi, xj) \therefore m * k$ here α is scaling factor and m is total

data point and k is total number of instant and sim find close point of data.

- Consider the RBF kernel K(x; y) H(x)= exp($(-(\delta - c)2$ / $(r2)$ this is kernel equation of plane.
- Use cross-validation to 2nd the best parameter C and
- Employ the best factor C and to train the complete training set Ro=$\alpha\frac{1}{p}\sum_{i=1}^{p} min\ (xi - yi)$ where Ro is learning parameter of kernel function.
- Generate formatted data.

2. Step of predictors.

1. Process the vector generated by matrix.
2. For all the classes are represented
Let us consider class of features c1, c2, c3…………………….cn
BEGIN
Find class with no features
C=¢
Find class at Max cross product rate
C=R*X^d
Find the class at half cross product
REPEAT
Pointer= False
Find the intervals of hyper plane

If the end condition is met
Pointer = True
If the first period has improved results we should
Use this, otherwise the other
Find the class evaluation after cross product class
Instances middle times
UNTIL pointer= False
END
3. Multiply all the classes with the best factor obtained;
4. Predictors are classified.
Section three

1. The classified predictors are assigned to code matrix
2. Image compressed
3. Find C.R value
4. Find PSNR value
5. Exit

V. EXPERIMENTAL RESULT ANALYSIS

The proposed compression algorithms are implemented using MATLAB software. For the performance of the proposed algorithm used medical CT image as well as on a sequence of images. The sequence used is grey-scale MRI images taken from local hospitals. The images in these sequences are of dimension 512*512 with 8-bit grey-scale image. for the empirical evaluation used following parameter. For the validation of proposed algorithm compared with SOM algorithm JPEG algorithm and composite algorithm. Used two different set of image in terms of CT image and MRI image [17].

$$PSNR=10\log_{10}\left(\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(F(i,j))^2}{\sum_{i=1}^{N}\sum_{j=1}^{N}(f(i,j)-F(i,j))^2}\right)$$
…………… (1)

Where R is maximum fluctuation in the cover image=512

$$C.R = \frac{ORIGINAL\ IMAGE\ SIZE}{COMPRESSED\ IMAGE\ SIZE} \dots\dots(2)$$

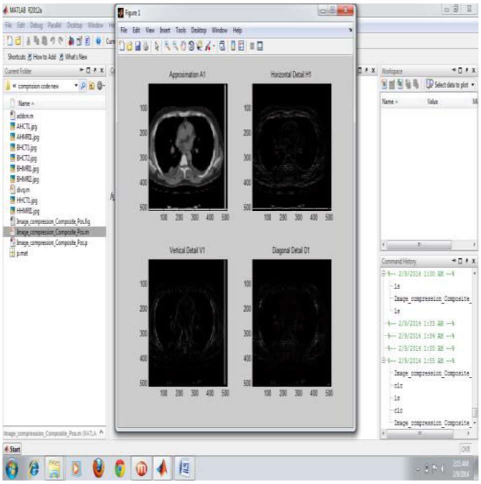

Figure1shows that processing of adaptive predictor process for spatial pixel extraction for the process of mean and variance

Compression method	PSNR	Compressed ratio	Image Type
APC-ML	49.485	8.884	HEAD MRI
BAPC	52.457	14.373	HEAD MRI
PROPOSED METHOD	52.981	15.876	HEAD MRI

Table 1 show the PSNR and Compressed ration value of all method applied on Head MRI image.

For Head Front CTA image resolution 512* 512.

APC-ML	37.43	4.746	Head front CTA
BAPC	43.59	7.826	Head front CTA
PROPOSED METHOD	46.21	12.664	Head front CTA
APC	46.67	13.983	Head front CTA

Table 2 shows the PSNR and compressed ration value of all method applied on Head front CTA image

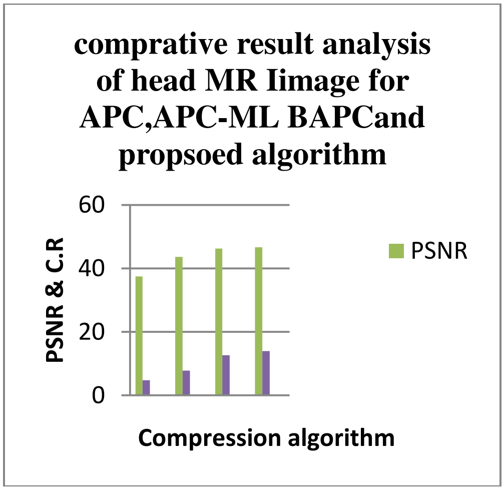

figure 2 the comprative analysis of compression ratio and PSNR value of MRI image. in case of proposed algorithm the value of PSNR and C.R are inceases in compression of other algorithm.

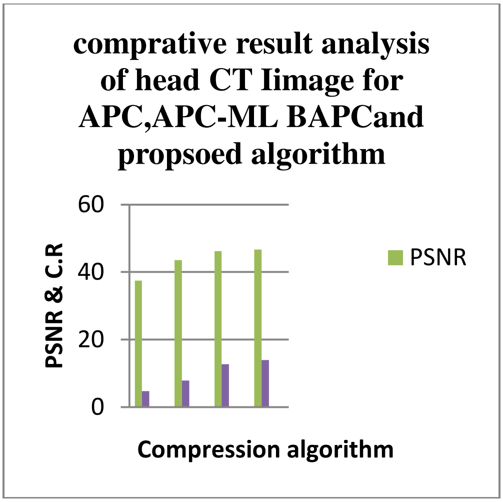

figure 3 the comprative analysis of compression ratio and PSNR value of CT image. in case of proposed algorithm the value of PSNR and C.R are inceases in compression of other algorithm.

VI. CONCLUSION AND FUTURE WORK

In this paper proosed a reduction of non structrued pixel value for image compression. the all adaptive pridictors divided into two group redudant structure and non redudant structure pixel with the help of structure refrence method. for the collection of similar and disimilar pixel used support vector machine. support vector machine process define the fitness constraints according to the diffrence value of structure refrence pixel. if diffrence of pixel is zero value assigned 1 and the diffrence value get value assined 0. both similar and disimilry pixel collect in two different unit and paases throgh code matrix. after that image is compressed. our emperical evalatuon of PSNR and comprssion ratio shows that better perfromance instead of other method used in experimental process. the seaching of redudant pixel sturcture conume more time and increase the computional time , in future reduces the computional time for compression.

REFERENCES

[1] Andrew Martchenko, Guang Deng "Bayesian Predictor Combination For Lossless Image Compression" Ieee Transactions On Image Processing, Vol. 22, 2013. Pp 5263-5270.

[2] Wang Xing-Yuan, Wang Yuan-Xing, Yun Jiao-Jiao "An Improved Fast Fractal Image Compression Using Spatial Texture Correlation" Chin. Phys. B Vol. 20, 2011. Pp 1-11.

[3] Taha Mohammed Hasan, Xingqian Wu " An Adaptive Fractal Image Compression" Ijcsi International Journal Of Computer Science Issues, Vol. 10, Issue 2, No 1, 2013. Pp 98-110.

[4] Jie Yang "Multiple Description Wavelet-Based Image Coding Using Iterated Function System" Hindawi Publishing Corporation Mathematical Problems In Engineering Volume 2013, Article Id 924274. Pp 1-13.

[5] R. E. Chaudhari, S. B. Dhok "Review Of Fractal Transform Based Image And Video Compression" International Journal Of Computer Applications, Volume 57– No.19, November 2012. Pp 23-32.

[6] Guo Hui And Wang Yongxue "Wavelet Packet And Neural Network Basis Medical Image Compression" Ieee 2011. Pp 234-240.

[7] Jie Yang "Multiple Description Wavelet-Based Image Coding Using Iterated Function System" Hindawi Publishing Corporation Mathematical Problems In Engineering Volume 2013, Article Id 924274. Pp 1-13.

[8] A.Benjamin Joseph , Dr. R. Baskaran "Edge Preserved Image Compression Technique Using Wavelet And Edge Based Segmentation" International Journal Of Computer Science & Informatics, 2011. Pp 80-85.

[9] N. Senthilkumaran, J. Suguna "Neural Network Technique For Lossless Image Compression Using X-Ray Images" International Journal Of Computer And Electrical Engineering, Vol.3, 2011. Pp 17-23.

[10] S .S. Panda, M.S.R.S Prasad, Mnm Prasad, Ch Skvr Naidu " Image Compression Using Back Propagation Neural Network" International Journal Of Engineering Science & Advanced Technology, Vol-2, 2012. Pp 74-78.

[11] Ilya Sutskever, James Martens, Geoffrey Hinton "Generating Text With Recurrent Neural Networks" International Conference On Machine Learning, 2011. Pp 1-8.

[12] A.Benjamin Joseph , Dr. R. Baskaran "Edge Preserved Image Compression Technique Using Wavelet And Edge Based Segmentation" International Journal Of Computer Science & Informatics, 2011. Pp 80-85.

[13] N. Senthilkumaran, J. Suguna "Neural Network Technique For Lossless Image Compression Using X-Ray Images" International Journal Of Computer And Electrical Engineering, Vol.3, 2011. Pp 17-23.

[14] S .S. Panda, M.S.R.S Prasad, Mnm Prasad, Ch Skvr Naidu " Image Compression Using Back Propagation Neural Network" International Journal Of Engineering Science & Advanced Technology, Vol-2, 2012. Pp 74-78.

[15] Ilya Sutskever, James Martens, Geoffrey Hinton "Generating Text With Recurrent Neural Networks" International Conference On Machine Learning, 2011. Pp 1-8.

[16] Hsiu-Niang Chen , Kuo-Liang Chung , Jian-Er Hung "Novel Fractal Image Encoding Algorithm Using Normalized One-Norm And Kick-Out Condition" Image And Vision Computing, Elsevier Ltd. 2010. Pp 518-525.

[17] Yue Wu, Yicong Zhou, Joseph P. Noonan, Sos Agaian, C. L. Philip Chen "A Novel Latin Square Image Cipher" Ieee Transactions On Information Forensics And Security, 2013. Pp 1-26.

[18] Chandan Singh Rawat, Sukadev Meher "A Hybrid Image Compression Scheme Using Dct And Fractal Image Compression" The International Arab Journal Of Information Technology, Vol. 10, No. 6, 2013. Pp 553-562.

[19] Jonathan Robinson And Vojislav Kecman" Combining Support Vector Machine Learning With The Discrete Cosine Transform In Image Compression" In Ieee Transactions On Neural Networks, Vol. 14, No. 4, July 2003