

Watermarking of Grayscale Images in DCT Domain Using Least-Squares Support Vector Regression

Vikash Chaudhary, Anurag Mishra, Rajesh Mehta, Monika Verma, Ram Pal Singh, and Navin Rajpal

Abstract—In this paper, we have used Least-Squares Support Vector Regression (LS-SVR) method, which is a reliable and robust method for regression analysis, for grayscale image watermarking in DCT domain. This method offers several advantages unlike conventional SVRs which is perceived as a minimization problem with linear inequality constraints and has solution to quadratic programming (QP) problem. Due to this reason, the problem solution becomes computationally costly. On the other hand, the solution to the LS-SVR algorithm may be obtained by solving a system of linear equations instead of solving a QP problem and therefore it consumes less time. In this case, the LS-SVR algorithm embeds a given binary watermark in three different grayscale images in a short time span. PSNR values indicate good quality of the signed images. The watermarks are also extracted and the computed values of $SIM(X, X^*)$ correlation parameter indicate that the extraction process is quite successful.

Index Terms—Grayscale image watermarking; least-squares support vector regression; PSNR.

I. INTRODUCTION

Several methods of image watermarking based on sampling datasets such as Artificial Neural Networks (ANNs) are prominently implemented using linear and nonlinear function estimation and approximation [5,13]. However, the ANN based approaches has many inherent problems such as presence of multiple local minima's, non-convex quadratic minimization and has the drawback of over fitting. Recently, support vector machine (SVM) based on statistical leaning theory has been developed by Vapnik et. al [13,14] at AT&T lab. This machine learning method has been widely and successfully applied for pattern classification and later extended for regression and function approximation, financial time series forecasting [3, 4], text classification [9] etc. Several papers have been reported wherein conventional Support Vector Regression (SVRs) are used for image watermarking [6, 7]. However, the SVM is perceived as a

minimization problem with linear inequality constraints and has a solution to quadratic programming (QP) problem. Due to this reason, the problem solution becomes computationally costly. On the other hand, the solution to the Least-Squares Support Vector Regression (LS-SVR) algorithm [12] may be obtained by solving a system of linear equations instead of solving a QP problem and therefore it consumes less time. In addition to this, the LS-SVR has been developed and successfully applied to regression of various non linear datasets [12]. In this paper, we have, for the first time, demonstrated the use of LS-SVR based regression for embedding and extraction of a binary image as watermark in three different grayscale images using Cox's formula [10]. The watermarked images are evaluated by computing PSNR which shows high quality of watermarked images. The watermarks are also extracted from the signed images and a comparison between the embedded and recovered watermarks is done using $SIM(X, X^*)$ correlation parameter in all three cases. High numerical values of the $SIM(X, X^*)$ coefficient indicate that the extraction process is quite successful and overall the watermarking algorithm finds good practical applications.

This paper is organized as follows: Section II gives a detailed mathematical description of the Least Squares Support Vector Regression algorithm. Section III describes experimental details of watermark embedding and extraction algorithms. Section IV presents the results and its analysis. Finally, the results are concluded in Section V and Section VI gives references used in this work.

II. LEAST-SQUARES SUPPORT VECTOR REGRESSION (LS-SVR)

The formulation of the LS-SVR consists of equality constraints instead of inequality constraints which is an integral part of standard SVR formulation [11, 12].

Let us assume a training image I of l sample such that $I \in \{x_i, y_i\}_{i=1, \dots, l}$ with the i^{th} input datum $x_i \in R^n$ and the i^{th} output datum $y_i \in R$. The purpose of the support vector regression model is to construct the regression function $f: R^n \rightarrow R$, $f(x)$, is defined as:

$$y = f(\tilde{x}) = \langle w', \phi(x) \rangle + b \quad (1)$$

where the nonlinear mapping function $\phi(\cdot)$ maps the input data space into higher dimensional feature space. In LS-SVR, the following optimization problem is formulated [13]

Manuscript received August 27, 2012; revised October 16, 2012. This work was supported in part by University Grant Commission, Ministry of Human Resource Development, Government of India, New Delhi, India under Grant No. F.6-620/2011(TG).

Vikash Chaudhary is with Department of Computer Science, BN College, University of Delhi, Delhi, India (vikasch09@yahoo.co.in).

Anurag Mishra is with Department of Physics and Electronics, DDUC, University of Delhi, Delhi, India.

Rajesh Mehta is with Department of Computer Science, Amity Engineering College, GGSIP University, Delhi, India.

Monika Verma is with MJKP School, Delhi, India.

Ram Pal Singh is with the Department of Computer Science, DDUC, University of Delhi, Delhi, India (rprana@gmail.com)

Navin Rajpal is with University School of Information Technology, GGSIP University, Delhi, India.

$$\min_{(w,e)} J(w,e) = \frac{1}{2} w^t w + \frac{c}{2} \sum_{i=1}^l e_i^2 \quad (2)$$

subject to

$$y_i = w^t \phi(x_i) + b + e_i, i = 1, \dots, l \quad (3)$$

Using dual theory, the Lagrangian function can be constructed with Equations (2) and (3) as

$$L(w,b,e,\alpha) = J(w,e) - \sum_{i=1}^l \alpha_i \{w^t \phi(x_i) + b + e_i - y_i\}$$

where $y = (y_1, \dots, y_l)$, $e = (1, \dots, 1)^t$ and Lagrange multipliers $\alpha = (\alpha_1, \dots, \alpha_l)^t$. The condition for optimality may be obtain from Lagrangian function as

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial w} = 0 \rightarrow \sum_{i=1}^l \alpha_i \phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^l \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \frac{c}{2} e_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \rightarrow w^t \phi(x_i) + b + e_i = 0 \end{array} \right. \quad (4)$$

for $i = 1, \dots, l$. After eliminating w and e_i , we can get the solution by the system of linear equations

$$\begin{bmatrix} 0 & e^t \\ e & Q + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (5)$$

and kernel function satisfying Mercer's condition

$$Q_{i,j} = \phi^t(x_i)\phi(x_j) = k(x_i, x_j), i, j = 1, \dots, l \quad (6)$$

s applied here and set $A = Q + C^{-1}I$. For A is symmetric and positive definite, all eigen values of A must be non-zero positive value, i.e., A^{-1} must exists. By solving the system of linear equations (5) we get

$$\alpha = A^{-1}(y - be) \text{ and } b = \frac{e^t A^{-1} y}{e^t A^{-1} I} \quad (7)$$

Substituting the value of w in Equation (1) after obtaining it from first of Equations (4) and using Eqn. (6), the nonlinear regression function (the output of LS-SVR) can be obtain as:

$$y = f(\tilde{x}) = \sum_{i=1}^l \alpha_i k(x, x_i) + b \quad (8)$$

where α_i and b are the solution of Equations (5). In LS-SVR approach, the proper kernel function selection is utmost important for its success. The kernel function can be a polynomial kernel of order p , an RBF kernel or a hyperbolic kernel. In our study, we have selected the widely used radial basis function (RBF) $k(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$ as a

kernel function because it offers less numerical difficulty due to the simple reason that it has less number of parameters to compute [5]. The kernel parameter γ is known as kernel bandwidth and is a constant. The LIBSVM Matlab code is used to implement the LS-SVR model [8].

III. EXPERIMENTAL DETAILS

Three standard 8-bits / pixel grayscale images of size 256 x 256 pixels each are taken as cover images - Lena, Cameraman and Goldhill. The watermark is a 32 x 32 pixel size binary image (1 bit / pixel). The cover object is first divided into 8x8 pixel blocks which are transformed into frequency domain using DCT algorithm. The block coefficients are first preprocessed by applying a zig - zag scan and the first 21 low frequency coefficients are selected barring the DC coefficient from each block. A dataset is made using the selected block coefficients for all 1024 blocks of the cover image. For each row vector of the data set, after having selected block coefficients of a particular block, we calculate their mean and it is used as a label for the corresponding row vector of the dataset. This dataset so formed is used to train the LS-SVR algorithm for following optimization parameters: $(C, \gamma, \epsilon) = (10^6, 10^{-4}, 10^{-2})$.

A. Embedding Watermark in Cover Image

Once the LS-SVR algorithm is trained, it gives as its output the predicted values for the mean calculated. The predicted output values $f(\tilde{x})$ corresponding to each block is kept in a separate data file which is further used to embed the binary watermark according to the formula given in Equation (9).

$$V' = V(1.0 + \alpha \times f(\tilde{x})) \quad (9)$$

where V is the maximum of the 21 coefficients for a specific block within the dataset, $f(\tilde{x})$ is the output of the LS-SVR algorithm and V' is the corresponding modulated DCT coefficient of the dataset. Fig. 1 depicts the watermark embedding procedure. The watermark embedding algorithm is given as Listing 1.

Listing 1: Algorithm Watermark Embedding

- Divide the cover image into 8x8 size blocks and transform them in frequency domain using DCT.
- Apply zig - zag scan to all 63 coefficients of each block barring DC coefficient.
- Select first 21 coefficients from each block and keep them in the row vector of the dataset prepared to train LS-SVR algorithm.
- Calculate mean of all coefficients of a row vector and make it a label, thereby making it a complete dataset of 1024 x 22 size.
- Train the LS-SVR algorithm using this dataset and obtain the output values for all blocks of the cover image.
- Embed the watermark in the cover image using the computed output values.
- Compute Inverse DCT (IDCT) to obtain signed image.

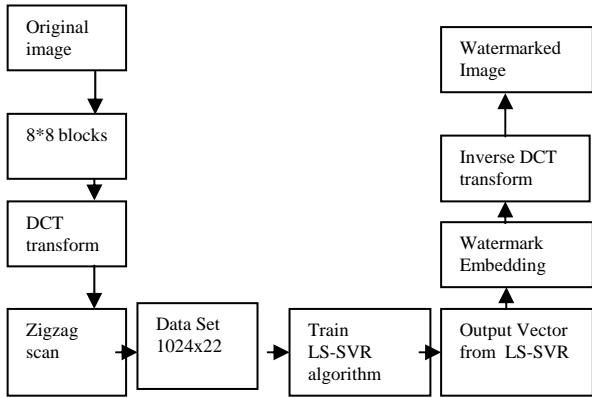


Fig. 1. Watermark embedding procedure

B. Extraction of Watermarks from Signed Images

Fig. 2 depicts the watermark extraction procedure according to Cox’s algorithm [10]. In this procedure, first of all, the DCT of all blocks from both the original and the signed images is computed. Secondly, the computed coefficients which are used in the embedding process are subtracted and the watermark is recovered. Let the original and recovered watermarks be denoted by

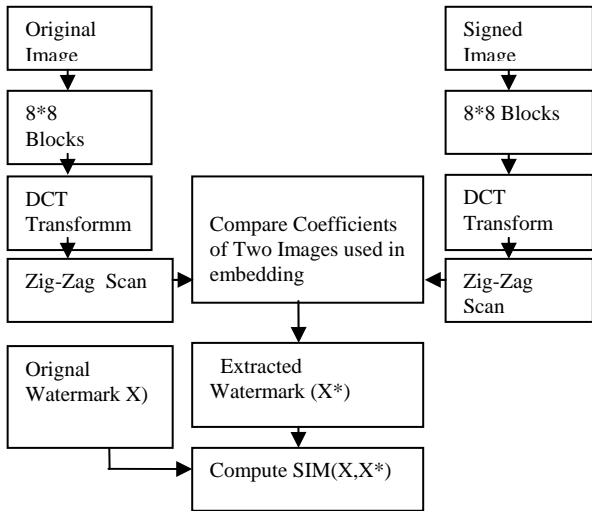


Fig. 2. Watermark extraction procedure

X and X^* respectively. Further, a comparison check is performed between X and X^* using the statistical similarity correlation parameter given by equation (13).

$$SIM(X, X^*) = \frac{\sum_{i=1}^n (X - X^*)}{\sum_{i=1}^n \sqrt{(X^* \cdot X^*)}} \quad (10)$$

The watermark extraction algorithm is given as Listing 2.

Listing 2: Algorithm Watermark Extraction

- Compute DCT of all blocks of both original and signed images.
- Subtract those computed coefficients of the original image from the signed image which are used in embedding process.
- Recover the watermark.
- Compare recovered watermark with the original mark using $SIM(X, X^*)$ parameter.

IV. RESULTS AND DISCUSSION

Figs. 3(a-c) depict the original images – Lena, Cameraman and Goldhill. Fig. 3(d) depicts the binary image used as the watermark to be embedded in images of Figs. 3(a-c). Figs. 4(a-c) depict the signed images obtained after watermark embedding. The respective PSNR values obtained from these images are written over these images. Figs. 5(a-c) depict the extracted watermarks obtained from images of Figs. 4(a-c). Their respective $SIM(X, X^*)$ correlation parameter is computed and its values are written beside these images. The perceptible quality of the signed images is very good and this is also indicated by high computed values of the PSNR. The extraction process is also quite successful. This is indicated from the high computed values of the $SIM(X, X^*)$ correlation parameter.

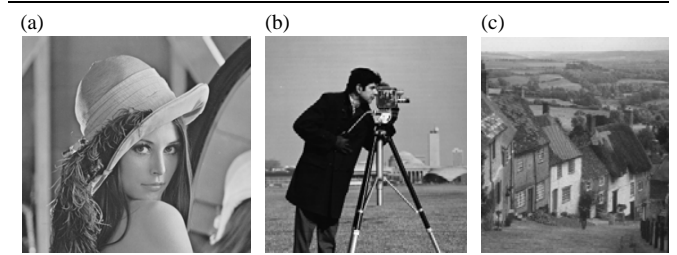


Fig. 3(a-c). Original Images – Lena, Cameraman and Goldhill



Fig. 3(d). Binary Image used as a watermark.



Fig. 4(a-c). Signed Images – (a) Lena, (b) Cameraman and (c) Goldhill

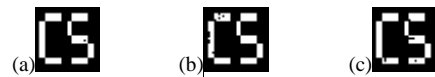


Fig. 5(a-c). Extracted watermarks obtained from Figs. 4(a-c).

V. CONCLUSIONS

A new method of digital image watermarking using Least-Squares Support Vector Regression (LS-SVR) is proposed in this paper. LS-SVR has been successfully applied to imperceptibly embed a binary image as a watermark in three 8-bits / pixel grayscale images in low frequency coefficients in the DCT domain. The perceptible quality of these images is very good and is quantified by PSNR. The watermarks embedded in these images are extracted using Cox’s algorithm. The embedded and the extracted watermarks are compared and similarity correlation coefficients are computed for all three extracted watermarks. The high numerical values of $SIM(X, X^*)$ is indicative of

good extraction process and overall the algorithm finds good practical applications.

ACKNOWLEDGMENT

The research presented in this paper was completely supported by University Grant Commission, Ministry of Human Resource Development, Government of India, New Delhi, India with Grant No. F.6-620/2011(TG).

REFERENCES

[1] R. Mehta, A. Mishra, R. Singh, and N. Rajpal, "HVS Based Digital Image Watermarking Algorithm Implemented using Back Propagation Neural Networks," in *Proceedings of International Conference on Data Management*, Ghaziabad, India, pp.206 – 213, March 11-12, 2010.

[1] Cheng-Ri Piao, Wei-Zhong Fan, Dong-Min Woo and Seung-Soo Han, "Robust Digital Image Watermarking Algorithm Using BPN Neural Networks," pp. 285 – 292.

[2] S. Balasundaram and Rampal Singh, "On Finite Newton Method for Support Vector Regression," *Journal of Neural Computing and Applications*, 2010, vol. 19, pp. 967 - 977.

[3] L. J. Cao, "Support Vector Machines Experts for Time Series Forecasting," *Neurocomputing*, 2003, vol. 51, pp.321-339.

[4] C. Campbell, "Kernel methods: a survey of current techniques," *Neuralcomputing*, 2002, 48, pp. 63-84.

[5] Xiuli Lv, Hongyu Bias, Yufei Yang, "Color Image Watermarking Using Support Vector Regression," in *Proceedings of 2nd International Congress on Image and Signal processing (CISP '09) 2009*.

[6] Rui-min Shen, Yong-gang Fu, Hong-tao Lu, "A Novel Image Watermarking Scheme Based on Support Vector Regression," *Journal of Systems and Software*, 2005, vol. 78, no. 1, pp. 1 - 8.

[7] C. C. Chang and C. J. Lin, "LIBSVM: A Library for Support Vector Machines," 2001: [Online]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

[8] H. Lodhi, J. Shawe-Taylor, N. Cristianini and C. Watkins, "Text Classification using String Kernels," *Neural Information Processing Systems*, 2000,13.

[9] Ingemar J. Cox, Joe Kilian, F. Thomson Leighton, and Talal Shamoan, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Transactions on Image Processing*, 1997, vol. 6, no. 12, pp.1673 - 1687.

[10] J. A. K. Suykens and J. Vandewalle, "Chaos control using least squares support vector machines," *International Journal of Circuit Theory and Applications*, 1999, vol. 27, no. 6, pp.605 - 615.

[11] J. A. K. Suykens and J. Vandewalle, "Least Squares Support Vector Machine Classifiers," *Neural Process Letter*, 1999, vol. 9, no. 3, pp. 293 – 300.

[12] V. N. Vapnik, "Statistical Learning Theory," John Wiley & Sons, New York, 1998.

[13] V. N. Vapnik, "The Nature of Statistical Learning Theory," 2nd edition, Springer, New York, 2000.

[14] O. L. Mangasarian and D. R. Musicant, "Large Scale Kernel Regression via Linear Programming," *Machine Learning*, 2002, vol. 46, pp. 255-269.



Vikash Chaudhary received M.Tech degree in computer science and engineering in 1998 and currently Assistant Professor in department of computer science, BN College, University of Delhi, Delhi. He is currently pursuing PhD in computer science and his research interests include multimedia digital watermarking/information hiding, image processing.



Anurag Mishra received M.Sc and Ph.D in electronic science in 1989 and 1995 respectively from University of Delhi, Delhi, India. He is currently Associate Professor in electronic science with department of physics and electronic science, DDUC, University of Delhi, Delhi, India and his current research interests include, watermarking/information hiding, image processing and soft-computing, computational intelligence, embedding system and VLSI.



Ram Pal Singh received M.Sc (Electronic Science) and M.Tech in computer science and engineering in 1992 and 1998 respectively and the Ph.D degree with School of Computer and System Sciences, Jawarlal Nehru University, New Delhi, India, in 2008. He joined as Post Doc fellow at Institute for Informatics, Free University, Berlin, Germany under Erasmus Mundus program belong to European Commission in 2009. He is currently Associate Professor in computer science with Department of Computer Science, DDUC, University of Delhi, Delhi, India and his current research interests include, machine learning, intelligent watermarking/information hiding, image processing and soft-computing.