Appropriate Variance Mean for Sorting Based on a Reversible Watermarking Algorithm

Chaiyaporn Panyindee

***Abstract*—Data Sorting is one of the key goals of the Reversible Watermarking (RW) Algorithm which has excellent performance in terms of reducing distortion, and was developed over a long period of time. The paper presents an improvement of the sorting technique using adaptive two-level variance mean (ATLVM). The combination of the average value with the new design of the appropriate expansion variance is applied to achieve increased capacity and to reduce the number of operators. The appropriate expansion variance allows the PEs to sort closer to the ideal. The performance of the proposed reversible watermarking scheme is evaluated using different images and is compared with conventional methods.**

Index Terms—**Data Sorting, reversible watermarking (RW), adaptive two-level variance average (ATLVM).**

I. Introduction

IN digital watermarking schemes, the host image will be distorted after the data embedding and the original image cannot be recovered. But in some applications such as medical image sharing, image trans-coding and multimedia archive management, any distortion due to data embedding is unacceptable since it is necessary to preserve the original image. Therefore, a solution called reversible watermarking (RW) is presented, in which the host image can be recovered completely after data embedding. RW is one scheme which combines several techniques to ensure that reversibility. It was developed over a long period of time. Most techniques are based on difference expansion (DE) [1], histogram shifting (HS) [2], prediction-error expansion (PEE) [3], and data sorting [4], etc. These approaches have the same aim are increasing the embedding capacity while keeping the distortion low.

The DE scheme is an important work related to RW that was presented by Tian [1]. In Tian's method, the host image is divided into pixel pairs, and the difference value of two pixels in a pair is expanded to hide one data bit. This method can provide an embedding rate of up to 0.5 bits per pixel (BPP) and it significantly outperforms previous compression works. In particular, Tian employed a location map to record all expandable locations, and non-expandable locations. The location map was applied in many algorithms of RW. It is usually huge in size and should be compressed. Even if location maps are compressed, they occupy a part of the payload. Thus, the size of the compressed location map determines the efficiency of a method. Tian's method has been improved in various ways. An important improvement was presented by Thodi and Rodriguez [2]. They achieved a significant improvement by incorporating DE with HS. In addition, instead of the difference value, Thodi and Rodriguez [3] introduced the utilization of the prediction-error for expansion embedding since this can better exploit local correlations within neighboring pixels. Sachnev *et al.* [5] improved the performance of the PEE by using sorting. Sachnev's algorithm was worth mentioning in that the location map is usually not necessary and even when it is necessary the size of this map is negligible. Therefore, sorting is a good choice among existing approaches of RW. Kang *et al.* [6] enhanced a sorting technique that exploits a broader image area to achieve a better result. Kotvicha *et al.* [7] have improved this parameter of the sorting process by using expanded variance means (EVM) which can increase the efficiency the embedding capacity and raise PSNR values. Panyindee *et al.* [8] introduced a genetic algorithm to optimize EVM parameters.

In this paper, we focus on a sorting scheme based on an EVM algorithm. There are two advantages of data sorting which can be explained as follows: The first is to decrease image distortion as PEs are used for embedding data that are rearranged based on the magnitude of their EVM values. Thus, when data is embedded using a histogram shifting scheme, a low PE value would be more advantageous than a high PE value in terms of distortion. The second advantage is to reduce the size of the location map. High PE values cause more maps than low PE values. However, the processing time is very important and is proportional to the number of operators. Note that a large number of operators are used to calculate EVM according to an extension variance. Therefore, we proposed an improvement of the sorting technique using ATLVM instead of the EVM value. The new design of the appropriate expansion variance is used to increase the embedding capacity and reduce the number of operators for the simulations. Results were compared with conventional sorting.

The organization of this paper is as follows: Section II discusses the related works. In Section III, the proposed method is presented in detail. Section IV presents the experimental results. Section V concludes the paper.

II. Related Works

Data sorting techniques are widely used in RW schemes and are added as a part of the scheme to enhance embedding

Manuscript received September 20, 2015.

C. Panyindee is with the Department of Computer Engineering, Faculty of Engineering, Rajamangala University of Technology Rattanakosin (RMUTR), Nakhon Pathom, 73120, Thailand (phone: +662-889-4585-7; fax: +662-889-4585-7; e-mail: chaiyaporn.pan@gmail.com, Chaiyaporn.pan@rmutr.ac.th)

capacity and reduce the size of the location map. Kamstra and Heijmans [4] greatly improved the DE technique. They proposed a method to reduce location map size by sorting pairs according to correlation measures to facilitate compression. However, sorting is possible only when pixels are independent. In other words, embedding data into one pixel should not affect the other pixels. Some methods cannot use sorting, such as Thodi and Rodriguez's method [3]. Their work produces dependent pixels, where embedding data to one pixel changes PEs of other pixels. Note that Sachnev *et al.*'s [5] approach is divides an image into two layers, (i.e., Dot layer and Cross layer respectively), while the rhombus predictor is used which requires independent pixels (see Figure 1). Thus, the pixels can be rearranged by sorting according to the correlation of the neighbouring pixels. Local variance $\mu_{i,j}$ for each pixel can be calculated from the four neighbouring pixels $v_{i,j-1}$, $v_{i+1,j}$, $v_{i,j+1}$, and $v_{i-1,j}$ as follows:

$$\mu_{i,j} = \frac{1}{4}\sum_{k=1}^{4}(\Delta v_k - \Delta\hat{v}_k)^2 \tag{1}$$

where $\Delta v_1 = |v_{i,j-1}-v_{i-1,j}|$, $\Delta v_2 = |v_{i-1,j}-v_{i,j+1}|$, $\Delta v_3 = |v_{i,j+1}-v_{i+1,j}|$, $\Delta v_4 = |v_{i+1,j}-v_{i,j-1}|$, $\Delta v_k = (\Delta v_1+\Delta v_2+\Delta v_3+\Delta v_4)/4$. $\mu_{i,j}$ calculated by using (1) to achieve suitable sorting for improving the performance of data embedding. Local variance $\mu_{i,j}$ has two features as follows: First, this value remains unchanged after data hiding. Second, this value is directly proportional to the magnitude of PE. For example, a small variance indicates a small magnitude of PE. Kotvicha *et al.* [7] have developed a sorting process by using averaging of the local variance of the neighbourhood pixels; this value is Expanded Variance Mean (EVM) which is used as a parameter in sorting. EVM can be calculated by Equations (2) and (3) as follows:

$$\tilde{\mu}_{i,j}^{ex} = \frac{\left(\sum_{l=-\frac{ex}{2}}^{\frac{ex}{2}}\mu_{i+l,j+l} + \sum_{k=1}^{\frac{ex}{2}}\sum_{l=-\frac{ex}{2}}^{\frac{ex}{2}-2k}(\mu_{i+l,j+l+2k}+\mu_{i+l+2k,j+l})\right)}{\frac{ex^2}{2}+ex+1} \tag{2}$$

and

$$\tilde{\mu}_{i,j}^{ex} = \frac{\left(\sum_{l=-\frac{(ex-1)}{2}}^{\frac{(ex-1)}{2}}\mu_{i+l,j+l} + \sum_{k=1}^{\frac{(ex-1)}{2}}\sum_{l=-\frac{(ex-1)}{2}}^{\frac{(ex+1)}{2}-2k}(\mu_{i+l,j+l+2k}+\mu_{i+l+2k,j+l})\right)}{\frac{ex^2}{2}+2ex-\frac{3}{2}} \tag{3}$$

Their method achieved the best performance among the aforementioned algorithms. However, the improved results have an effect on the high complexity algorithm that has an impact on processing time.

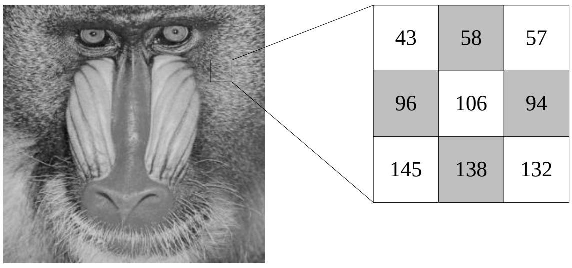


Figure. 1: A window (3x3) extracted from the Baboon image.

III. The Prosed Scheme

Consider a problem for sorting of the previous work [5]. Let the four-neighbouring pixels be used to calculate the local variance (see Figure 1 showing rough texture for Baboon image). Note that the magnitude of PE of the four-neighbouring pixels does not indicate there is a small variance. Therefore, this problem is solved by using ATLVM. The new area of adjustable averaged variance is computed as follows:

$$\mu_{i,j}^{k_1} = \frac{\sum_{n=1}^{9}\mu_n}{9} \tag{4},$$

$$\mu_{i,j}^{k_2} = \frac{\sum_{n=1}^{25}\mu_n}{25} \tag{5}$$

and

$$\mu_{i,j}^{k} = \frac{\sum_{k=1}^{2}\mu_n}{n} \tag{6}$$

where ATLVM $\mu_{i,j}^{k}$ is the appropriate averaged variance value of the pixels for an image. k is instead a level of the new expansion variance. n is the number of variances, which is used to calculate ATLVM. There are two available k value (i.e., k_1, k_2 for our work) in order to achieve the best possible PSNR and also the minimum time cost, as it is necessary to find and use an appropriate k value based on the required payload. Such an appropriate k value can be found by fine-tuning. In Figure 3 shows two different sequences of PE values of the Baboon and Barbara testing images calculated by the rhombus predictor. In Figure 3 (a) and (c), both images are sorted by using their local variances. In Figure 3 (b) and (d), two sets of pixels are sorted by using our method. Note that after comparing the sequence of PE sorted by local variance and by ATLVM, the results show that the sequence of PE, sorted by ATLVM value, is better organized; i.e., close to ideal (from the low PE value to the high PE value). A problem of the EVM technique is that it is considered to reduce the complexity of the scheme. Note

ISBN: 978-988-19253-8-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

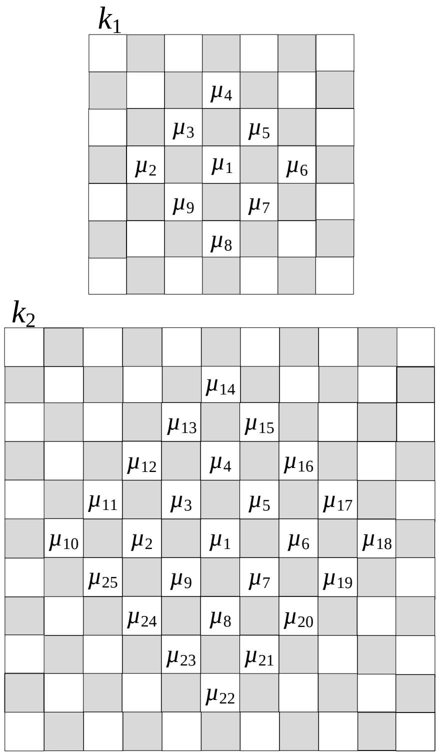


Figure 2: The number of pixels that are used for each k value.

that their method involves repeating until the maximum level of the expansion variance which is not necessary. We use two levels which is sufficient for predicting PE.

Embedding Algorithm:

1) Separate pixels in an image into Cross and Dot according to the Cross embedding scheme [5].
2) Preserve the first 36 bits of Cross and sort the remaining pixels according to ATLVM, starting from the 37th pixel.
3) Collect the original LSB of the first 36 bits and include them as a part of the payload.
4) Find appropriate threshold values (T_n, T_p) and classify pixels into the expandable set E and shiftable set S, and create the location map L according to [5].
5) Embed payload using prediction-error histogram shifting (PEHS) and calculate PSNR, according to [3].
6) Repeat steps 1-6 by modifying the k value to calculate ATLVM, according to [7].
7) Select the k value that optimized PSNR to use in data embedding.

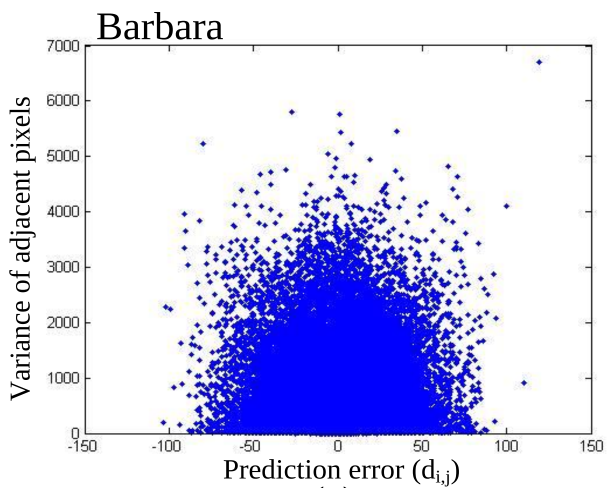
Recovering data, threshold values, and payload size for the Dot embedding scheme or Cross embedding scheme should be accomplished through the decoder. Computations for the Dot and Cross decoding schemes are similar. Thus, to simplify the explanation of the double decoding scheme, we only describe the Cross decoding scheme.

Decoding Algorithm:

1) See Step 1 of the embedding algorithm.
2) Extract header from the first 36 bits of Cross; thus we obtain threshold values (T_n, T_p), and the k value to calculate ATLVM.
3) Sort Cross data by starting at the 37th pixel using ATLVM.
4) Apply the extracting algorithm sequentially from step 3, according to [7].

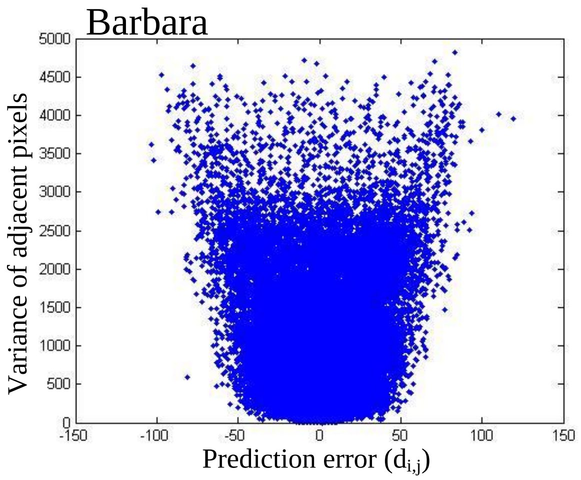
IV. Experimental Results

The proposed scheme is compared with the three methods of Kamstra and Heijmans [4], Sachnev *et al.* [5] and Kotvicha *et al.* [7] using typical 512×512 grayscale images (i.e., Baboon, and Barbara shown in Figure 4). The baboon test image is downloaded from http: //sipi.usc.edu/database, while Barbara is obtained from http://decsai.ugr.es/cvg/dbimagenes/g512.php. All the simulations in this paper were carried out on a typical PC with an Intel(R) Core(TM) i7-3720QM CPU @ 2.60 GHz that has 8 GB memory, running Windows 7 (64-bit). Results are shown in Table I. For both images, the proposed algorithm is superior to other methods, although Kamstra and Heijmans' method shows good performance when the payload is small. However, the comparison for small payloads (see Table I) shows that the proposed scheme is better. Sachnevs' algorithm exploits sorting with several efficient well-known existing techniques in combination, such as histogram shift, double embedding scheme, etc. Their work provides superior results when compared to Kamstra and Heijmans method. Kotvicha *et al.* take advantage of the average value of increasing the number of pixels for calculating EVM which varies depending on the image's variation and payload size. This tool allows more accurate sorting. Their method achieved the best performance among the aforementioned algorithms. The fact is that increasing the number of pixels for each expansion variance results in a higher time cost of processing, especially when larger payloads are embedded (see Table II). Our method provides better improvements in terms of reducing the time cost because exploiting a suitable number of pixels for calculating the expansion variance ensure the number of operators is small.

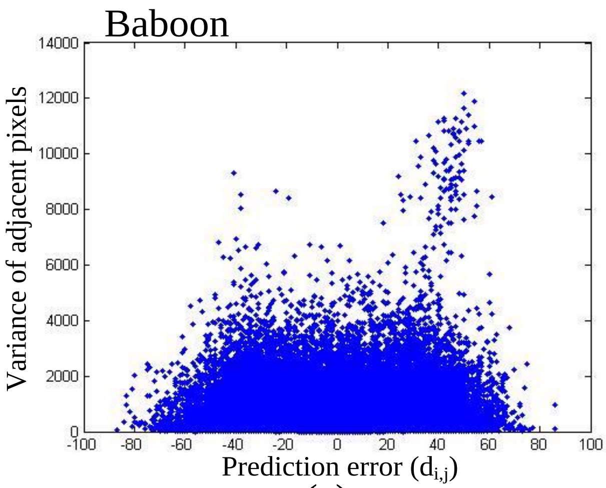


(a)

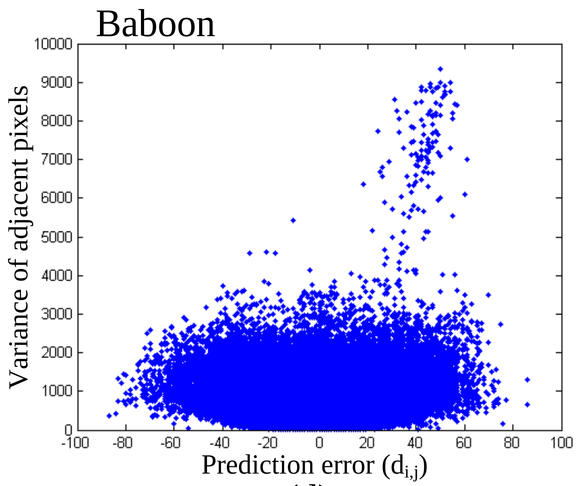
ISBN: 978-988-19253-8-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)



(b)



(c)



(d)

Figure 3: Sequences of pixels sorted. (a) and (c) The sequence of sorted PE by using local variance (b) and (d) The sequence of sorted PE by using ATLVM.

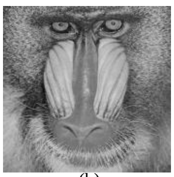

(a)

(b)

Figure 4: Tested images (a) Barbara, (b) Baboon.

TABLE I

Comparison of PSNR between the proposed algorithm and the two-methods of Kamstra and Heijmans [4] and Sachnev *et al.* [5] for small capacities.

Payload (bits/pixel)	PSNR(dB)					
	Barbara			Baboon		
	[4]	[5]	Proposed	[4]	[5]	Proposed
1311/0.005	62.90	67.03	68.05	58.10	63.14	64.67
3932/0.015	58.76	62.50	63.25	53.44	58.34	59.70
6554/0.025	56.25	60.22	60.86	50.31	55.67	57.11
9175/0.035	54.69	58.62	59.31	49.16	53.32	55.37

TABLE II

Comparison of PSNR and running time between the EVM scheme [7] and the ATLVM scheme based on the algorithm [5] for 0.8 BPP.

Image	0.8BPP			
	PSNR (dB)		Time (s)	
	[7]	Proposed	[7]	Proposed
Barbara	40.56	40.96	445.95	12.64
Baboon	31.74	31.98	982.63	21.90

V. CONCLUSION

Our In this paper, an improved data sorting approach based on [7] is presented to reduce the work complexity. The number of operators used for processing is an important part of the RW process for reducing the time cost. We enhance performance of the expansion variance for each embedding and each image by using ATLVM. Two levels of the expansion variance are used which is sufficient for determining the appropriate parameters to provide the optimal results, especially for small capacities and high variation images. Experimental results show that the proposed method has better results compared to conventional sorting.

ACKNOWLEDGMENT

This work was supported by the Faculty of Engineering, Rajamangala University of Technology Rattanakosin (RMUTR).

REFERENCES

[1] J. Tian, “Reversible data embedding using a difference expansion,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 8, pp. 890–896, August 2003.

[2] D. M. Thodi and J. J. Rodriguez, “Reversible watermarking by prediction-error expansion,” in Proceedings of the 6th IEEE Southwest Symposium on Image Analysis and Interpretation, Lake Tahoe, vol. 6, pp. 21–25, March 2004.

[3] D. M. Thodi and J. J. Rodriguez, “Expansion embedding techniques for reversible watermarking,” IEEE Transactions on Image Processing, vol. 16, no. 3, pp. 721–730, March 2007.

[4] L. H. J. Kamstra and A. M. Heijmans, “Reversible data embedding into images using wavelet techniques and sorting,” IEEE Transactions on Image Processing, vol. 14, no.12, pp. 2082–2090, December 2005.

[5] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y. Q. Shi, “Reversible watermarking algorithm using sorting and prediction,” IEEE Transaction on Circuit System and Video Technology, vol. 19, no. 7, pp. 989-999, July 2009.

[6] S. Kang, H. J. Hwang, H. J. Kim, “Reversible Watermark Using an Accurate Predictor and Sorter Based on Payload Balancing,” ETRI Journal, vol. 34, pp. 410-420, June 2012.

[7] C. Kotvicha, P. Sanguansat and M. L. K. Kasemsa, “Expand Variance Mean Sorting for Reversible Watermarking,” International Journal of Computer and Communication Engineering, vol. 1, no. 3, September 2012

[8] C. Panyindee, and C. Pintavirooj, “Reversible Watermarking Using GaussianWeight Prediction and Genetic Algorithm,” in Proc. IMECS, vol.1, pp. 457-461, Mar. 2013.