Paper: RTIP-D-18-00402, Improved reversible data hiding based on PVO and adaptive pairwise embedding

Dear Editors and Reviewers,

Thank you very much for giving us an opportunity to revise our manuscript and thanks for reviewers' comments. These comments are all valuable and very helpful for revising and improving our paper. We have studied these comments carefully and have made corrections. Our responses to reviewers' comments are listed below. Here, reviewers' comments are marked in red.

**************** Responses to Reviewer 1 ****************

Q1-1: For adaptively searching of the optimal mapping, the parameter K is preset as 2 here, maybe you can give more specific data to illustrate for choosing region $[0,2] \times [0,2]$ for optimization.

In the proposed method, there are totally 1996 2D mappings that can be selected with the local region $[0,2] \times [0,2]$. This is a trade-off between the embedding performance and the running time cost. However, if we set $K = 3$, i.e., taking local region $[0,3] \times [0,3]$ for generating 2D mappings for embedding performance optimization, there are millions of mappings and it costs too much time to embed data which is not suitable for real-time applications.

****

Q1-2: Some typo errors should be corrected with carefully proofreading. Such as: "with the same block sizes" → " with the same block size", "this improeved" → " this improved ", "can get respectively an increase of PSNR " → "can respectively get an increase of PSNR ", "more than one pixels" → "more than one pixel".

Thanks a lot for your comments. These typos have been corrected in the revised paper.

**************** Responses to Reviewer 2 ****************

Q2-1: This method has to use the methods described in [10] and [29] in the references, but the methods were not explained in the related works.

The method pairwise PEE [10] has been briefly reviewed in the revised paper, please see Section 2.3 (marked in red in the revised paper). The method PVO-K [29] is just mentioned in the experimental part, and it is compared with the proposed one since it is a recently proposed PVO-based embedding method. In fact, our method is not designed based on PVO-K, our idea is mainly based on the original PVO and the pairwise PEE embedding.

****

Q2-2: In this manuscript, to calculate the block complexity the difference between the vertical and horizontal differences of two adjacent pixels has been used. Please explain why this method was used. Is it the best choice? Why and why not?

Actually, there are several methods for calculating the block complexity. Once we use the block selection strategy, the embedding performance can be then improved. The method we choose here is not necessarily optimal. But, in our experience, this block selection manner generally works well, and we then simply adopt this way. We think it is a good question to investigate the optimal way for block selection, this is maybe for the future work.

****

Q2-3: It might happen that overflow and underflow issue occurs during the process of data embedding. But no discussion in the problem is presented. Please explain how to solve this problem if it occurs, and mark the additional information.

Thanks a lot for your comments. This is done. A new paragraph for overcoming the overflow/underflow issue is added in the revised paper, please see the last paragraph of Section 3 (marked in red in the revised paper).

****

Q2-4: How much memory is needed in this method to record the optimized table and different parameters?

This is done. The proposed method requires 8.65 MB of memory to store all 2D mappings. Please see the red part of the first paragraph of Section 4 in the revised paper.

****

Q2-5: How to design an optimal 2D mapping for each test image? Please elaborate on the rules and procedures and list all combinations of experiments.

Actually, in the proposed method, the optimal 2D mapping is obtained by exhaustive search. In the left column of page 6, second paragraph, we have explained that "We will exhaustively search all the 2D mappings and find the optimal one such that it can provide the required embedding capacity while the embedding distortion is minimized ...".

****

Q2-6: How long does it take to create an optimal 2D mapping? For a new image, how much time will it take to get the optimal 2D mapping of the image? Please compare and analyze the operation time of other methods as well.

The proposed method is implemented by using C++ on Visual Studio 2017 with a personal laptop (ThinkPad X280). It costs less than 0.01s to generate all 1996 2D mappings. Moreover, the embedding process can be completed within five seconds for a given embedding capacity. Please see the red part of the first paragraph of Section 4 and Table 1 in the revised paper.

****

Q2-7: In Table 3 and Table 4, whether it compares with the best result of the comparison items, please describe the parameters of the comparison.

Yes, in Table 3 and Table 4 (it is numbered Table 2 and Table 3 in the revised paper), it is the best embedding result for each comparison method. In the revised paper, the embedding parameters of these comparison methods are also listed in Table 2 and Table 3.

****

Q2-8: The (a) and (b) diagrams in Fig. 5 is not clearly indicated. Please indicate the direction of b=0 and b=1. There are errors in the position (1, -2). Please correct them.

Thanks a lot for your comments. the directions of b=0 and b=1 have been indicated in Fig. 5(a) and Fig. 5(b). Due to the symmetry, only the first quadrant is marked. And the marked error of position (1, -2) is corrected.

**************** Responses to Reviewer 3 ****************

Q3-1: When compared with other PVO-like schemes (see Tables 3 and 4). it is observed that the improvement is limited (note: the improvement is only 0.5 dB as compare with Ou et al.'s scheme).

Thanks a lot for your comments. Actually, the existing PVO solutions such as Ou et al.'s method have achieved good experimental performance. Compared with the other PVO-based methods shown in Table 3 and 4 (it is numbered Table 2 and Table 3 in the revised paper) and Figure 7, over 0.5 dB boost in average is already a good improvement since the previous RDH schemes perform already rather well. We think that RDH is already a mature technique and its significant performance increase is hard to achieve. On the other hand, we think that PVO is somewhat a new methodology for RDH and it is essentially different with the previous predictors.

****

Q3-2: To enhance the limited contribution, a theoretical analysis and proof that the proposed scheme is better than previous schemes should be given.

Thanks a lot for your comments. Generally, modifying higher dimensional reversible mapping and using adaptive optimization on local region can achieve better performance due to the larger solution space for optimization. The theoretical discussion on the reversible embedding is already discussed in our previous work [42].

**************** Responses to Reviewer 4 ****************

Q4-1: This paper proposed to utilize the correlation between the Max and Min prediction-errors in each block for reversible data hiding. However, the experimental results demonstrate that the proposed method promote the data embedding performance less, where, the PSNR gains few improvement compared with the traditional schemes. So, is the computation cost affordable to obtain so small data embedding capacity improvement?

Thanks a lot for your comments. In fact, the current PVO-based schemes achieved good performance compared with the traditional PEE-based RDH schemes especially for low embedding capacities. On the other hand, we think that PVO is somewhat a new methodology for RDH and it is essentially different with the previous predictors. It can be seen in Figure 7 that 0.5 dB boost in average is already a good improvement since the previous RDH schemes perform already rather well. Yes, we agree that RDH is already a mature technique and its significant performance increase is hard to achieve.

****

Q4-2: The paper proposed to modify the generated 2D histogram for data embedding, the method employed in this scheme should be discussed more deeply, it would be better to give an example.

In the right column of page 4, we show an example of embedding 10,000 bits data into image Baboon as shown in Figure 5(a) and Figure 5(c). In this example, two different mappings of the improved PVO method ([34] in the revised paper) and Pairwise-PEE 2D mapping ([18] in the revised paper) are respectively utilized for embedding. And in the left column of page 6, the second paragraph explains that adaptive optimization on local region strategy further leverages correlation of pixels to further improve the final performance.

****

Q4-3: In the proposed reversible data embedding, how to achieve an optimal data embedding coefficients, that is, how to determine the coefficients such as n1,n2 and T in the process data embedding?

In the proposed method, the complexity threshold T is determined as the smallest one such that the embedding capacity can be satisfied with the generated 2D histogram. Please see the the fifth sentence in the right column of page 4. And for n_1 and n_2, the embedding procedure is implemented several times for different block size $n_1, n_2 \in \{2, 3, 4, 5\}$, and the best embedding result is taken as our final embedding result. Please see the the third paragraph in the left column of page 6.

****

Q4-4: In Fig.5(b), 8 pixels pairs are employed for reversible data embedding, are the data embedded in 8-ary form? Does method decrease the data embedding capacity? Or, as the object pixel pairs would be changed much, does this method bring more image distortion?

Actually, during the embedding procedure, all the pixels are modified for data embedding according to the 2D mapping in order. Among the various available 2D mappings, traditional 1D mapping is one special case belong to them. So, when we exhaust all the 2D mappings to find the optimal one, the lower limit of the maximum embedded capacity is at least the same as before. And with a given embedding capacity, by the proposed method, we focus on finding the optimal 2D mapping that minimizes the embedding distortion while satisfying the capacity. Thus, the proposed method reduce the embedding distortion than previous methods.

****

Q4-5: the run-time efficiency of the proposed scheme is preferred to be discussed in this paper, as the Journal is about the real-time image processing.

Thanks a lot for your comments. For the proposed method, less than 0.01s is cost to generate all 2D mappings. And the embedding process is completed within five seconds. The discuss about run-time efficiency is added in first paragraph of section 4 (marked in red in the revised paper), and shown in the form of table (as seen in table 1 in revised paper) as well.

Best regards,
Haorui Wu, Xiaolong Li, Yao Zhao, and Rongrong Ni