Robust JPEG steganography based on the robustness classifier.

Tashkent State Technical University.
Department: radio vices and systems.
Senior Lecturer: ***Korotkova Larisa Aleksandrovna.***
2nd year student: ***Yuldasheva Diyora Ravshanovna.***

Introduction.

Steganography is the technology of hiding secret messages in multimedia files, such as images, audio, or videos, without introducing the trace of modification. Its counterpart, steganalysis, is the technique of detecting whether the testing multimedia files have been modified to carry secret messages. Currently, the most advanced steganography algorithm is adaptive steganography. The latest steganalysis algorithms are based on high-dimensional feature sets and the ensemble classifier or deep learning methods, which are able to detect adaptive steganography. Adaptive steganography usually contains two parts: the cost function and the information embedding algorithm. The design of the cost function is related to the impact of embedding to its counterpart, steganalysis, so that the pixels or coefficients that are hard to detect by steganalysis after modification have low costs. The design of the cost function could be divided into two disciplines. One is defined empirically by assigning lowcost values to pixels or coefficients in the complex content areas without considering the specific steganalysis features. Another method tries to attack a specific

steganalysis by setting the pixels or coefficients that weaken detection power as small cost. After the cost function is calculated, the information embedding algorithm, the syndrome trellis codes (STC) is used to realize information embedding with modifications introducing minimal costs.

Preliminaries.

In this section, we give the preliminaries of this paper, which include the JPEG compression process, the STC embedding, and the error-correcting codes. The bold letters refer to matrices and vectors, and the non-bold letter with subscripts refers to an element of matrices or vectors. The symbols utilized in this paper are listed in Table 1. The side information representing the position of message embedding blocks is called the protocol message.

Symbol	Meaning
$\boldsymbol{X}$	Quantized DCT coefficients
$\boldsymbol{D}$	Unquantized DCT coefficients
$\boldsymbol{Q}$	Quantization table
$\boldsymbol{S}$	Rounded spatial pixels restored from DCT coefficients

Table 1. The list of symbols used and their meanings.

Proposed method.

In this paper, with the output of the robustness classifier, a robust steganography based on the robust block selection

and the protocol message embedding is proposed. The framework of the proposed method is shown in Fig. 2.

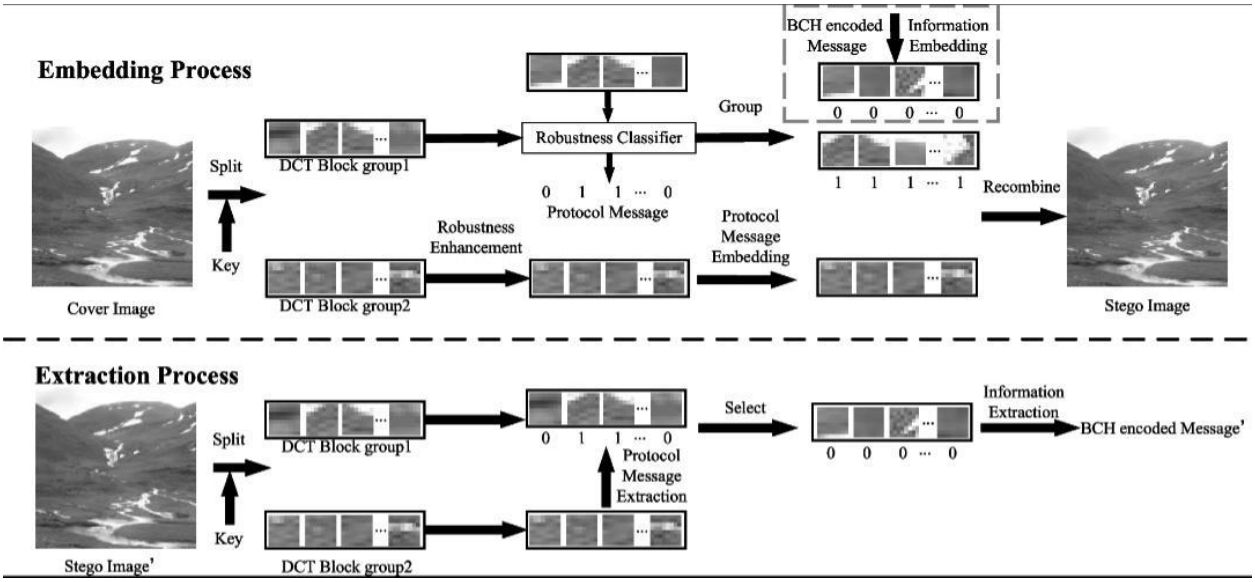

Fig. 2. The framework of the proposed method.

The embedding process can be described as follows. First, using a secret key shared by the steganographer and the receiver, all the DCT blocks in cover images are randomly divided into two disjoint groups, which are named group1 and group2 separately. The length of the two groups is equal, which means each group contains half the number of all DCT blocks in cover images. Then, the DCT blocks in group1 are inputted into the trained classifier that can classify whether the input block is robust. The bit representation is 1 if the block is predicted as non-robust and 0 if it is predicted as robust. The bits result predicted by the classifier are concatenated into byte streams as protocol messages, which will be embedded into the DCT blocks in group2. To increase the success possibility of protocol message restoration, the DCT blocks in group2 experience the robustness

enhancement operation that improves the robustness. After error-correcting code encoding, the protocol messages are embedded into robustness-enhanced DCT blocks in group2 with the general robust steganography. The DCT blocks in group1 will be further grouped by the prediction results of the classifier. The secret message is encoded by the error-correcting code encoder and randomly permutated. Then in group1, they are embedded into the DCT blocks predicted robust by the classifier with adaptive steganography. Finally, after embedding secret messages and protocol messages, the DCT blocks in group1 and group2 are combined to generate stego images. The extraction process is described as follows. Firstly, the receiver receives the stego image recompressed by channels. Then, the DCT blocks of recompressed stego images are divided into two groups using the same key in the embedding process. The protocol message is extracted from DCT blocks in group2, and the prediction results of DCT blocks in group1 can be restored by the extracted protocol message. The DCT blocks in group1 with the bit representation as 0 are selected to form the extraction domain. After performing the STC extraction and inverse permutation, the error-correcting code encoded message is extracted. The message can be restored by performing error-correcting code decoding. Aiming to improve the robustness, general robust steganography usually introduces more distortion compared with non-robust steganography. The reason that we divide DCT

blocks into two groups for embedding secret messages and protocol messages separately is as follows. The length of the protocol message is fixed and is related to the size of an image. In our methods, when the payload is large, the adaptive steganography which is more secure is used to embed the message. The robustness is ensured by selecting the robust DCT blocks as the embedding domain by a trained classifier. The length of the protocol message remains the same, and the protocol message is embedded by general robust steganography. Compared with only using general robust steganography for embedding, the combination of two embedding processes can improve security performance when embedding large-sized secret messages. At the same time, through a trained classifier, we select robust DCT blocks as the embedding domain and perform a robust enhancement method. The robust performance is improved, and a higher success extraction rate can be achieved.

List of used literature.

1. C.-H. Cheng, Y.-F. Huang, H.-C. Chen, T.-Y. Yao, Neural network-based estimation for OFDM channels.
2. E. Balevi, J.G. Andrews, Deep learning-based channel estimation for high-dimensional signals (2019).