

# Applying LR Cube Analysis to JSteg Detection

Kwangsoo Lee, Changho Jung, Sangjin Lee, HyungJun Kim, and Jongin Lim

Center for Information Security Technologies, Korea University, KOREA  
kslee@cist.korea.ac.kr

**Abstract.** JSteg is a steganographic method for JPEG images that can be viewed as LSB steganography. Recently, we proposed a new principle to detect the use of LSB steganography in digitized vector signals and showed how to apply it to gray-scale images. In this paper, we discuss how to apply it to JPEG images and show some experimental results for the JSteg-like algorithm.

**Introduction:** JSteg-like algorithm [1] is a steganographic method for JPEG images. It works by replacing the LSBs (least significant bits) of quantized DCT (discrete cosine transform) coefficients excepting the values of '0' and '1' with message bits. Thus, the JSteg-like algorithm can be viewed as the LSB steganography with an exception for usable DCT coefficients. Recently, we proposed the LR cube analysis to detect the use of LSB steganography in RGB color images [2]. It used RGB vectors as basic units and took advantage of the local correlation of these vectors. In [3], we described its general framework to use  $n$ -dimensional vectors as basic units for the applications to various types of digital signals and showed how to apply it to gray-scale images. So far, we have been working on the task whether the LR cube analysis can be applied to detect the JSteg-like algorithm for JPEG images; the answer is yes. This paper shows the method of how to apply the LR cube analysis for detecting the usage of JSteg-like algorithm.

**Method Description and Experimental Results:** Fig. 1 and Fig. 2 show the two basic ideas of the LR cube analysis. Fig. 1 represents the two types of comparable sets in  $Z^3$  under the LSB embedding, where the  $\delta$  is a positive odd integer and the two-sided arrows represent the possible changes of the vectors by LSB embedding. Given a vector signal, the sets can be classified into their complexities as shown in Fig. 2, where the complexity of a set means the number of different vectors contained in the set. We observed that the left and right  $\delta$ -cubes are similarly distributed on their complexities before the LSB embedding, but the left (or right)  $\delta$ -cubes' complexities tend to decrease (or increase) after the LSB embedding, respectively. The LR cube analysis measures a dissimilarity between their complexities by a well-known hypothesis test: the  $\chi^2$ -test.

For the application of the LR cube analysis to JPEG format, we have investigated various kinds of vector samplings from the DCT coefficients. As a result, we developed the  $(x, y; i, j)$  sampling;  $x$  and  $y$  indicate the numbers of joint blocks in  $x$  and  $y$  directions, and also,  $i$  and  $j$  indicate the numbers of joint

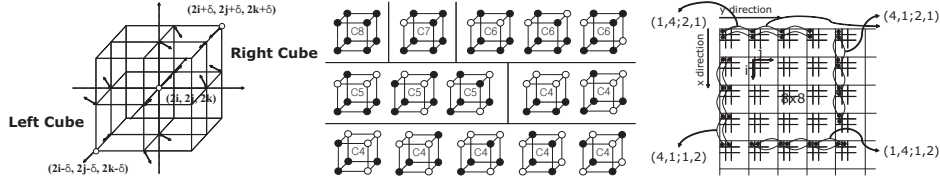


Fig. 1. LR cube model

Fig. 2. Cube complexities

Fig. 3. Vector sampling

coefficients in  $i$  and  $j$  directions of each jointed JPEG block, respectively. The vectors drawn by the  $(x, y; i, j)$  sampling have the dimension  $n = xyij$ . Fig. 3 illustrates some of the efficient vector sampling.

The proposed method has been tested on the 518 uncompressed JPEG images of CBIR image database of Washington University [4] with their compressed images of 75%, and 50% qualities in order to take into consideration the quantization effect on the DCT coefficients. We embedded random messages into Y component data of the images by JSteg-like algorithm, where the message lengths are 5% and 10% of the capacity of each image.

We drew the 8-dimensional vectors from the Y component data by the  $(1,4; 1,2)$  and  $(1,4;2,1)$  samplings, and set the  $\delta$  equal to 3. The below table shows the detecting performances of our method for each image set, when the decision thresholds of p-values were set to have the false positives about 0% and 5%. The most interesting result is that our method highly detected the low-rate embedding by the JSteg-like algorithm, even if it did not show any false detections.

Image quality	Decision threshold	Embedding rate			Decision threshold	Embedding rate		
		0%	5%	10%		0%	5%	10%
100%	0.99	0.000	0.602	0.940	0.58	0.464	0.830	0.995
75%	0.91	0.000	0.450	0.987	0.40	0.046	0.857	0.998
50%	0.89	0.000	0.419	0.960	0.47	0.043	0.757	0.990

**Acknowledgements:** This research was supported by the Ministry of Information and Communication (MIC), Korea, under the Information Technology Research Center (ITRC) support program supervised by the Institute of Information Technology Assessment (IITA).

## References

1. D. Upham, <http://www.funet.fi/pub/crypt/steganography/jpeg-jsteg-v4.diff.gz>.
2. K. Lee, C. Jung, S. Lee, and J. Lim, "Color Cube Analysis for Detection of LSB Steganography in RGB Color Images", ICCSA 2005, LNCS vol. 3481, Springer-Verlag, Berlin Heidelberg, 2005, pp. 537–546.
3. K. Lee, C. Jung, S. Lee, and J. Lim: "New Steganalysis Methodology: LR Cube Analysis for the Detection of LSB Steganography", accepted to the *7th Information Hiding Workshop*, Barcelona, Spain, June 6-8, 2005.
4. <http://www.cs.washington.edu/research/imagedatabase/groundtruth>.