EURASIP Journal on Image and Video Processing
a SpringerOpen Journal

RESEARCH Open Access

Performing scalable lossy compression on pixel encrypted images

Xiangui Kang[1,3*], Anjie Peng[1], Xianyu Xu[2] and Xiaochun Cao[3]

Abstract

Compression of encrypted data draws much attention in recent years due to the security concerns in a service-oriented environment such as cloud computing. We propose a scalable lossy compression scheme for images having their pixel value encrypted with a standard stream cipher. The encrypted data are simply compressed by transmitting a uniformly subsampled portion of the encrypted data and some bitplanes of another uniformly subsampled portion of the encrypted data. At the receiver side, a decoder performs content-adaptive interpolation based on the decrypted partial information, where the received bit plane information serves as the side information that reflects the image edge information, making the image reconstruction more precise. When more bit planes are transmitted, higher quality of the decompressed image can be achieved. The experimental results show that our proposed scheme achieves much better performance than the existing lossy compression scheme for pixel-value encrypted images and also similar performance as the state-of-the-art lossy compression for pixel permutation-based encrypted images. In addition, our proposed scheme has the following advantages: at the decoder side, no computationally intensive iteration and no additional public orthogonal matrix are needed. It works well for both smooth and texture-rich images.

Keywords: Image compression; Image encryption; Lossy compression; Image reconstruction; Scalable coding

1. Introduction

Compression of encrypted data draws much attention in recent years due to the security concerns in a service-oriented environment such as cloud computing [1,2]. The traditional way of securely and efficiently transmitting redundant data is to first compress the data to reduce the redundancy then encrypt the compressed data. At the receiver side, decryption is performed prior to decompression. However, in some application scenarios (e.g., sensor networking), a sender may first perform encryption with a simple cipher and then send it to a network provider. The network provider always has the interest to reduce the rate. It is desirable to be able to compress the encrypted data without the key to reduce the security concerns. At the receiver side, joint decryption and decompression will be used to reconstruct the original data.

It has been proved in [1] that the overall system performance of such approach can be as good as the conventional approach, that is, neither the security nor the compression efficiency will be sacrificed by performing compression in the encrypted domain. Two practical approaches to lossless compression of encrypted binary images and to lossy compression of encrypted Gaussian sequence are also presented in [1]. In the first approach, the original binary image is encrypted by adding a pseudorandom string; the encrypted data are compressed by finding the syndromes with respect to a low-density parity-check (LDPC) code [3]. In the second approach, the original data are encrypted by adding an iid Gaussian sequence, and the encrypted data are quantized and compressed as the syndromes of a trellis code. In [4], compression of encrypted data for both memoryless sources and sources with hidden Markov correlation using LDPC codes is also studied. A study [5] introduces a few methods for lossless compression of

* Correspondence: isskxg@mail.sysu.edu.cn
[1]School of Information Science and Technology, Sun Yat-Sen University, Guangzhou, Guangdong 510006, China
[3]State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China
Full list of author information is available at the end of the article

encrypted grayscale and color images by employing LDPC codes to various bit planes and exploiting the spatial and cross-plane correlation among pixels. In [6], Liu et al. proposed to decompose the encrypted image in a progressive manner, and the most significant bits in the higher levels are compressed using rate-compatible punctured turbo codes. The decoder can observe a low-resolution version of the image, study the local statistics based on it, and use the statistics to estimate the content in the higher levels. Another study [7] presents some algorithms for compressing encrypted data and demonstrates blind compression of encrypted video by developing statistical models for source data and extending these models to video. All of the works mentioned above use the distributed source coding (DSC) technique. However, a frequent backward channel communication is needed for the joint decryption and decoding at the receiver, and thus, large delay maybe of concern. So, DSC-based methods may not be a desirable choice in some practical network transmission scenarios.

There are a few works on lossy compression for encrypted data. In [8], the authors introduce a compressive sensing technique to achieve lossy compression of encrypted image data, and a basis pursuit algorithm is appropriately modified to enable joint decompression and decryption. In the state-of-the-art work [2], a lossy compression and iterative reconstruction for permutation-based encrypted image is proposed. However, when using the permutation-based encryption, only the pixel positions are permuted, but the pixel values are not masked in the encryption phase, which means that the histogram of the encrypted image will remain the same as the original image, revealing some significant information. Iterative reconstruction may have a hard time to converge for a texture-rich image. An additional public orthogonal matrix of huge size is needed for the decompression at the receiver side. If the size of the to-be-compressed image is 512×512, then the size of the public orthogonal matrix is about 512×512×512×512. Each target rate requires a distinct public orthogonal matrix. The huge public orthogonal matrix results in huge storage space requirement and computational load. Note that such a public orthogonal matrix cannot be used in the compression for pixel-value encrypted image. There is another lossy compression and iterative reconstruction for encrypted image proposed in the state-of-the-art work [9]. The encryption method of the image is by making a modulo-256 addition on the original pixel values with pseudorandom numbers. The scheme is scalable, and it performs very well with iteration.

In this paper, we propose a scalable lossy compression scheme for images having their pixel value encrypted with a standard stream cipher. At the receiver side, a decoder performs a content-adaptive interpolation prediction based on the decrypted partial information, and the received bit plane information serves as the side information to facilitate accurate image reconstruction. The experimental results show that our proposed scheme achieves much better performance than the existing lossy compression scheme for pixel-value encrypted image and achieves similar performance as the state-of-the-art lossy compression for permutation-based encrypted images.

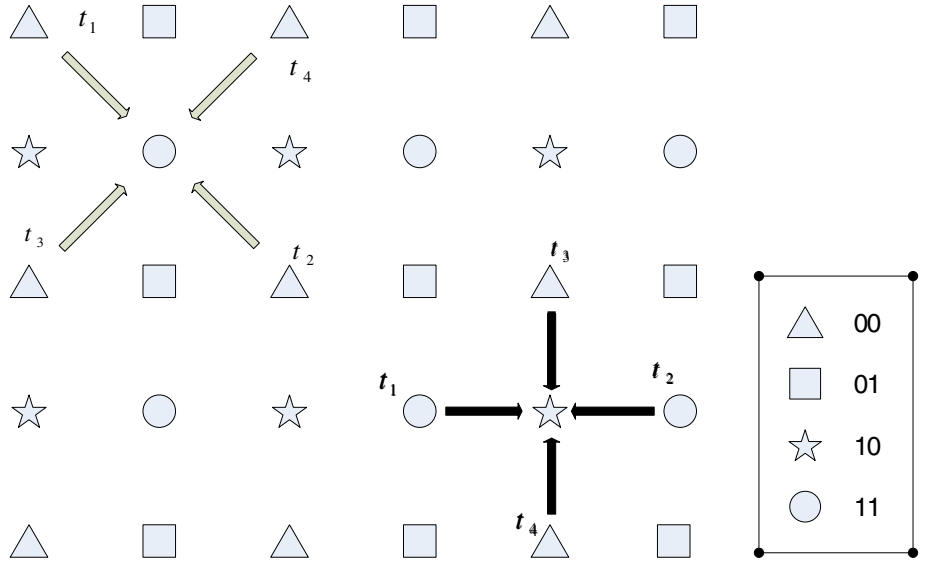

Figure 1 Illustration of the CAI prediction.

The rest of this paper is organized as follows: in 'The proposed scalable compression scheme' section, we describe the proposed compression scheme for encrypted images in detail. The 'Experimental results' section shows the experimental results with comparison to the state-of-the-art works. The conclusion is made in the 'Conclusions' section.

2. The proposed scalable compression scheme

We assume the images have been encrypted by applying a standard stream cipher to the pixel values in the spatial domain. Even though the pixel value has been encrypted, the resulting encrypted data preserve some of the inherent property of the original image, e.g., the spatial relationship of pixels and the bit plane structure and their relative importance. This leads us to adopt a multi-resolution and bit plane-based scalable approach for the compression. The basic idea is to package and transmit a downsampled version of the encrypted image as the base layer, then selectively transmit additional bit plane information from another downsampled version (with a different spatial offset) of the encrypted image to facilitate the interpolation/reconstruction of the higher resolution image at the receiver. This process can be recursively applied in a multi-layer structure. This results in an embedded, compressed, and encrypted bitstream, where the bitstream can be cut off flexibly to meet a target bit rate constraint without requiring complex communication/negotiation between the encoder and decoder as was the case in some prior work that used DSC, e.g., [1,3-7]. In the following, we describe our proposed scheme in a two-layer scenario.

Suppose the size of an original 8-bit grayscale image is $N_1 \times N_2$. It is encrypted with a standard stream cipher, resulting in an encrypted image E.

To compress, we downsample the encrypted image by a factor of two in both dimensions and generate four sub-images, denoted as E_{00}, E_{01}, E_{10}, and E_{11}. Here, the first digit '1'(or '0') denotes that the horizontal offset for downsampling is 1 (or 0), the second digit '1' (or '0') denotes that the vertical offset is 1 (or 0). As shown in Figure 1, each icon is a pixel. We use the four icons to distinguish the four sub-images after downsampling. When they are decrypted and decompressed, they are denoted as 00, 01, 10, and 11 sub-images, respectively.

The uncompressed E_{00} sub-image will be transmitted to the decoder. Some of the E_{11} sub-image's bit planes will be transmitted, too, according to the target bit rate. The target bit rate (R) per information source bit can be calculated by:

$$R = 0.25 + 0.25 \times N/8, \tag{1}$$

where N is the number of bit planes of sub-image E_{11} to be transmitted. For example, if N = 2, it means two bit planes of sub-image E_{11} are transmitted. Let $b_8 b_7 b_6 \ldots b_2 b_1$ denote the eight bit planes, and $b_7 b_6$ are transmitted. The compression rate is 0.25 + 0.25 × 2/8= 0.3125. The decoder reconstructs the 00 sub-image by decrypting the E_{00} sub-image and also obtains b_7,b_6 of the 11 sub-image by decryption. Here, according to our observation, b_8 can be recovered with little error by decompression, so the b_8 bit plane of sub-image 11 is always transmitted only when all of b_7,b_6,...,$b_1$bit planes of sub-image 11 are transmitted, that is, only when N = 8. N = 8 means that sub-image 11 is transmitted.

For every pixel in the 11 sub-image, there are four neighboring pixels $\boldsymbol{t} = [t_1, t_2, t_3, t_4]$ in the 00 sub-image as shown in the top left of Figure 1. We predict the 11 sub-image using the 00 sub-image with the context-adaptive interpolation (CAI) scheme proposed in [6]. In this work, we propose to use the received bit plane values of sub-image 11 as the side information to facilitate the estimation of the image edge information in the context adaptive interpolation thus improving the prediction. For the 10 sub-image, there are also four neighboring pixels in the 00 and 11 sub-images as shown in the bottom right of Figure 1. So, when the receiver obtains sub-image 00 and sub-image 11, the 10 sub-image (and 01 sub-image) can be predicted by the conventional CAI (please refer to [6] for a detailed description of the conventional CAI). In the following, we only present the improved CAI prediction of the 11 sub-image with the received bit plane information as the side information.

Let 0 be a pixel in the 11 sub-image which is to be predicted and $\boldsymbol{t} = [t_1, t_2, t_3, t_4]$ be the vector of its four neighboring pixels (please refer to the top left of Figure 1). The preliminary prediction of pixel 0 with CAI [6] is:

$$\mathrm{pred}_0 = \begin{cases} \mathrm{mean}(\mathbf{t}) & (\max(\mathbf{t}) - \min(\mathbf{t}) \le 20) \\ (t_1 + t_2)/2 & (|t_3 - t_4| - |t_1 - t_2| > 20) \\ (t_3 + t_4)/2 & (|t_1 - t_2| - |t_3 - t_4| > 20) \\ \mathrm{median}(\mathbf{t}) & \mathrm{otherwise} \end{cases} \tag{2}$$

Table 1 The PSNR (dB) of the reconstructed images using our scheme

N	0	1	2	3	4	5	6	7	8
Rate per info bit	0.25	0.28	0.31	0.34	0.38	0.41	0.44	0.47	0.5
Lena	34.7	35.0	36.0	37.8	38.9	39.4	39.7	39.8	40.2
Baboon	30.2	30.9	31.9	34.4	36.0	36.6	36.7	36.8	36.8
Man	30.4	30.9	31.2	32.6	33.5	33.9	34.0	34.0	34.9
Hill	30.3	30.5	30.8	32.8	33.8	34.2	34.4	34.4	35.1

In Equation 2, the local region is classified into four types: smooth, horizontally edged, vertically edged, and other median-related edge. With the received bit plane values of sub-image 11, we can match the bit plane values of pred_0 with the received bit plane value. If they match with each other, we accept the preliminary prediction value; otherwise, we find a better-matching prediction using the image edge directions other than the four local regions considered in Equation 2.

The decoder will receive E_{00} sub-image and some bit planes of the E_{11} sub-image. After decryption, the decoder will get 00 sub-image and some bit planes of the 11 sub-image. We denote the decimal value of the bit planes which are transmitted and decrypted as w. Take $N = 2$ for example, b_7b_6 of the 11 sub-image was considered. If $b_7b_6 = (10)_2$, $w = 2$. $w \in [0, 2^N - 1] = [0, M\text{-}1]$. Let Δ be the stepsize corresponding to the most significant bit plane of the side information. In this paper, we adopt $\Delta = 2^7$ when $N < 8$ in our scheme. Define the matching distance d as follows:

$$d = \text{floor}\left(\text{mod}\left(\frac{\text{pred}_0}{\frac{\Delta}{M}}, M\right)\right) - w, \tag{3}$$

where mod( ) is the modulation operation. As b_7b_6 was known, we calculate the distance d between w and the decimal value of the same bit planes of pred_0. The distance can be used to judge whether the pred_0 matches well. If the distance is large, such that:

$$M/4 < |d| < 3 \times M/4, \tag{4}$$

we consider that pred_0 does not match well. Then, other two prediction values pred_1 and pred_2, which correspond to other image edge directions, will compete with the preliminary prediction value pred_0 for the best match with the side information:

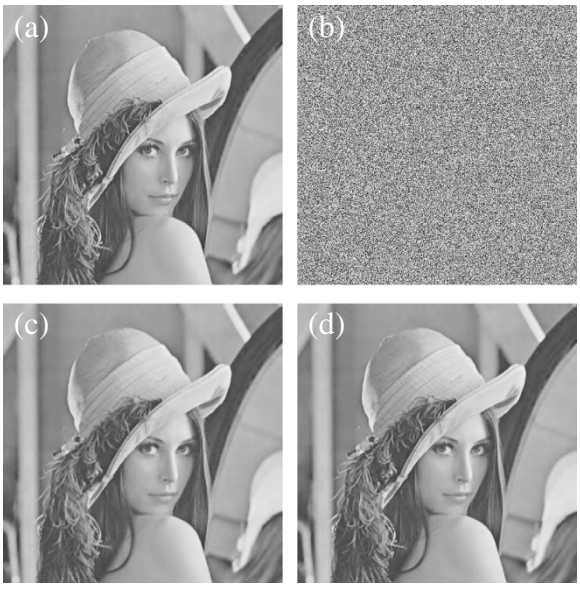

Figure 2 The proposed compression scheme is applied on Lena image. (a) Original Lena image; **(b)** Encrypted Lena image; **(c)** Decompressed Lena image with $R = 0.31$(PNSR = 36.0 dB); **(d)** Decompressed Lena image with $R = 0.44$ (PNSR = 39.7 dB).

$$\text{pred}_1 = \frac{\text{sum}(\mathbf{t}) - \max(\mathbf{t})}{3}, \tag{5}$$

$$\text{pred}_2 = \frac{\text{sum}(\mathbf{t}) - \min(\mathbf{t})}{3}, \tag{6}$$

where sum( ) denotes the summation operation, and max( ) and min( ) denote taking maximum and minimum operation, respectively. We find the best match by seeking the minimum value of $\min(|d|, M - |d|)$ among the three prediction values pred_0, pred_1 and pred_2 and obtain the final best matching prediction pred. Finally, with the side information, the corresponding prediction value r can be further refined to be:

$$r = \begin{cases} floor(\text{pred}/\Delta) \times \Delta + w \times \Delta/M - \Delta + \Delta/M - 1, if d < -M/2 \\ floor(\text{pred}/\Delta) \times \Delta + w \times \Delta/M + \Delta, if d > M/2 \\ floor(\text{pred}/\Delta) \times \Delta + w \times \Delta/M + mod(pred, \Delta/M), otherwise \end{cases} \tag{7}$$

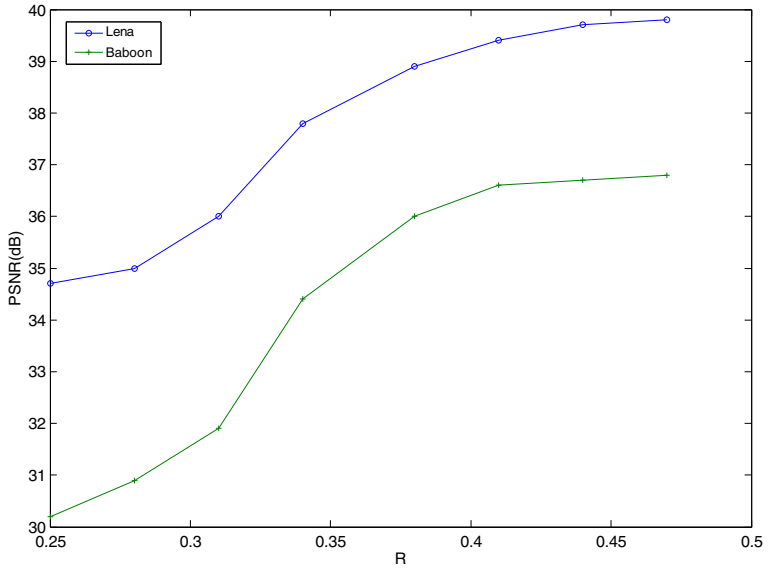

Figure 3 PSNR of reconstructed images with respect to bit rates.

where floor(pred/Δ) × Δ is the value of the b_8 of the prediction in the pixel; $w \times \Delta/M$ is the value of the bit planes transmitted in the pixel.

3. Experimental results

In this section, we will examine the performance of our proposed method and also compare it with the existing state-of-the-art works. The proposed compression scheme is applied on a variety of images with different sizes. We will show the test results for four selected standard images which have varying texture contents. The test images used here are Lena, Baboon, Man, and Hill. All the test images are 8-bit grayscale images of 512×512. Results for a two-layer decomposition structure are presented.

Table 1 shows the peak signal-to-noise ratio (PSNR) of the decompressed image with varying bit rates (bit rate per information source bit). The bit rate is determined by N, the number of transmitted bit planes of the E_{11} sub-image. Higher rate leads to higher PSNR.

Figure 2a is the original image of Lena with a size of 512×512. The encrypted image of Lena is shown in Figure 2b. Let $N = 2$, the corresponding bit rate $R = 0.31$, the PSNR of the reconstructed Lena is 36.0 dB (Figure 2c); let $N = 6$, the bit rate $R = 0.44$, the PNSR of the

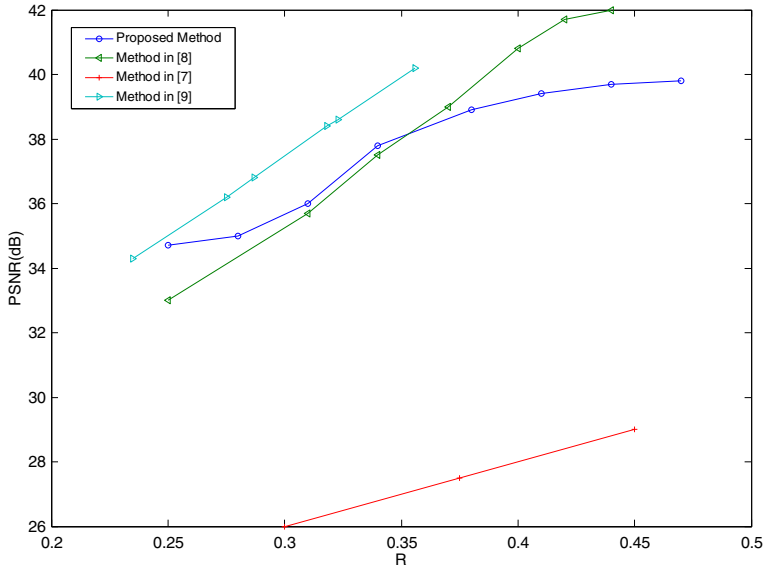

Figure 4 Comparison results on Lena image.

reconstructed Lena is 39.7 dB (Figure 2d). It is observed that the decompressed images (Figure 2c,d) are very similar to the original image, and there is no visible compression artifact. With N increasing from 2 to 6, the PNSR increases significantly, but with N increasing from 6 to 7, the PNSR increases slowly because the least significant bit plane b_0 has little effect on the pixel value. Higher rate leads to better quality of the reconstructed image. Figure 3 illustrates the PSNR of the reconstructed images with the varying rates when Lena and Baboon are used. It shows that the proposed scheme works for both smooth and texture-rich images.

There are few existing works on the lossy compression of the pixel-value encrypted image. We compare our proposed method with the method in [8], which applies compressive sensing technique to compress the encrypted image. Figure 4 shows that our method achieves much better performance than the method in [8] on the same Lena image. With the bit rate changing from 0.25 to 0.44, the PSNRs of our method are all higher than 34 dB, while the PSNRs of the method in [8] are lower than 30 dB.

We also compare our method to the methods in [2] and [9]. Our proposed method achieves similar performance as the method in [2] for the pixel-value unencrypted image (Figure 4). A public orthogonal matrix is used in [2] to disperse the estimation error in the permutation-based encrypted domain. Note that such a public orthogonal matrix cannot be used in the pixel value-encrypted domain. Compared to the most recent method [9], our method is a little worse than the method in [9], but our method is not an iterative method and both methods in [2] and [9] are iterative ones and thus may have the issue of convergence for a texture-rich image and possible intensive computation.

4. Conclusions

In this paper, we propose a lossy compression scheme for pixel-value encrypted images. The main contributions are as follows:

1. At the receiver side, the received bit plane information serves as the side information to facilitate the estimation of image edge information thus making the image reconstruction more precise. The more bit planes are transmitted, the higher quality of the reconstructed image.
2. The experimental results show that our proposed scheme achieves much better performance than the existing lossy compression scheme for pixel-value encrypted images, and also achieves similar performance as the state-of-the-art lossy compression on the pixel permutation-based encrypted images.
3. Compared to the state-of-the-art work, our proposed scheme also has the following advantages: at the decoder side, no computationally intensive iteration and no additional public orthogonal matrix are needed. The scheme can be applied to both smooth and texture-rich images.

In the future, we will also extend our work to compression of encrypted video.

Competing interests
The authors declare that they have no competing interests.

Acknowledgment
This work was supported by NSFC (Grant nos. 61070167, U1135001), 973 Program (Grant no. 2011CB302204), the Research Fund for the Doctoral Program of Higher Education of China (Grant no. 20110171110042), and NSF of Guangdong province (Grant no. s2013020012788).

Author details
[1]School of Information Science and Technology, Sun Yat-Sen University, Guangzhou, Guangdong 510006, China. [2]China Telecom Corporation Limited Shantou Branch, Shantou, Guangdong 515000, China. [3]State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China.

Received: 9 October 2012 Accepted: 7 May 2013
Published: 21 May 2013

References

1. M Johnson, P Ishwar, VM Prabhakaran, D Schonberg, K Ramchandran, On compressing encrypted data IEEE Trans. Signal Process **52**(10), 2992–3006 (2004)
2. X Zhang, Lossy Compression and Iterative reconstruction for Encrypted Image. IEEE Trans. on Inf. Forensic Secur **6**(1), 53–58 (2011)
3. RG Gallager, *Low Density Parity Check Codes* (Ph.D. dissertation, MIT, 1963)
4. D Schonberg, SC Draper, K Ramchandran, *On blind compression of encrypted correlated data approaching the source entropy rate, in Proceedingsof the 43rd Annual Allerton Conference* (Allerton, IL, 2005)
5. R Lazzeretti, M Barni, *Lossless compression of encrypted grey-level and color images, in Proceedings of the 16th European Signal Processing Conference (EUSIPCO 2008)* (Lausanne, Switzerland, 2008)
6. W Liu, W Zeng, L Dong, Q Yao, Efficient compression of encrypted grayscale images. IEEE Trans. Image Process. **19**(4), 1097–1102 (2010)
7. D Schonberg, SC Draper, C Yeo, K Ramchandran, Toward compression of encrypted images and video sequences. IEEE Trans. Inf. Forensic Secur **3**(4), 749–762 (2008)
8. A Kumar, A Makur, Lossy compression of encrypted image by compressing sensing technique, in *Proceedings of the IEEE Region 10 Conference. (TENCON 2009)*, 2009, pp. 1–6
9. X Zhang, G Feng, Y Ren, Z Qian, Scalable coding of encrypted images. IEEE Trans. Image Process. **21**(6), 3108–3114 (2012)

doi:10.1186/1687-5281-2013-32
Cite this article as: Kang *et al.*: **Performing scalable lossy compression on pixel encrypted images.** *EURASIP Journal on Image and Video Processing* 2013 **2013**:32.