Accurate estimation of primary quantisation table with applications to tampering detection

Yuanyuan Peng and Benyong Liu

There are many ways to accurately estimate the low-frequency region of the primary quantisation table in a double compressed JPEG image with different quality factors. However, they cannot accurately estimate the high-frequency region. An image compression model to describe the relationship between the primary compression and the second compression is used to present a simple approach to accurately estimate the whole primary quantisation table, and based on this, an efficient method for detection of image tampering involving JPEG recompression is proposed. The effectiveness of the proposed algorithms is tested by experiments and the experimental results are presented.

Introduction: Digital images are often tampered with for different purposes, with different degrees of adverse effects on society. Therefore, blind detection of image tampering becomes more and more important in modern society. In tampering involving JPEG recompression, it is generally assumed in the related work that the quality factor of the primary compression is only a small ratio of the second one [1, 2], and many methods are proposed for estimation of the primary quantisation table, such as the JPEG error analysis method [3], the method proposed in [4] using the histogram of the DCT coefficients and its improved versions [5, 6] and a similar but more effective method recently proposed in [7] using the likelihood map of the DCT coefficients.

In this Letter, we set the relationship between the quantisation steps of the primary compression and the second compression, and based on this relationship, present a novel approach to accurately estimate the whole primary quantisation table. In addition, using the accurate estimation, we propose an efficient method for the detection of image tampering involving JPEG recompression. The experimental results show good performance of the proposed algorithms.

Image compression model: Denoted by u_1 and u_2, the DCT coefficients of an image before the primary quantisation and after the second quantisation, respectively. The 'image compression model' may be described as follows:

$$Q_{q1q2}(u_1) = \left[\left[\frac{u_1}{q_1}\right]\frac{q_1}{q_2}\right] = u_2 \tag{1}$$

where q_1 and q_2 are the primary quantisation step and the second quantisation step, Q_{q1q2} is the relationship between u_1 and u_2. Hence

$$u_2 - 0.5 \le \left[\frac{u_1}{q_1}\right]\frac{q_1}{q_2} < u_2 + 0.5 \tag{2}$$

Therefore

$$\frac{q_2}{q_1}(u_2 - 0.5) \le \left[\frac{u_1}{q_1}\right] < \frac{q_2}{q_1}(u_2 + 0.5) \tag{3}$$

or equivalently

$$eA \le \left[\frac{u_1}{q_1}\right] < eB \tag{4}$$

where e, A and B denote q_2/q_1, $u_2 - 0.5$ and $u_2 + 0.5$, respectively. It is observed that B minus A is equal to 1. Since $[u_1/q_1]$ is an integer, when $e \le 1$, we have

$$\lceil eA \rceil = \left[\frac{u_1}{q_1}\right] = \lfloor eB \rfloor \tag{5}$$

where $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ denote the ceiling and floor functions, respectively.

Quantisation table estimation: To better estimate the quantisation table of the forgery image, since $\lceil eA \rceil = \lfloor eB \rfloor$, we define a function as follows:

$$u = \lceil eA \rceil - \lfloor eB \rfloor \tag{6}$$

From the JPEG image header, we can obtain the second compression quantisation table. In fact, assume that q_1 changes from 1 to 100 and let q_1' be an estimated quantisation step, since q_2 is already known, we know that when $q_1'=q_1$, the values of u are all 0, so we can define a measure as follows:

$$p\{u = 0|q_1\} = \frac{\text{num}(u = 0)}{\text{num}(u)} \tag{7}$$

where num is a function for counting the element number in its variable when the primary quantisation step is q_1. Since p in (7) is a function of q_1, and it reaches maximum when $q_1'=q_1$. However, a lot of DCT coefficients in the high-frequency region are quantised to 0, and hence many points may lead p to maximum and they all may be taken as the candidate points, but only the points with values bigger than q_2 are rational and their medium value is taken as q_1'.

Forgery region location: After the primary quantisation table is obtained, the following blocking artefact measure may be adopted to locate the tampered region of a forged image

$$B(i) = \sum_{k=1}^{64}\left|D(k) - Q(k)\text{round}\left(\frac{D(k)}{Q(k)}\right)\right| \tag{8}$$

where $B(i)$ is the estimated blocking artefact for the testing block i, $D(k)$ is the DCT coefficient at position k of the testing block and $Q(k)$ is the estimated DCT quantisation step.

Experimental results: In our experiments, we set the primary and second compression quality factors Q_1 and Q_2 as 80 and 90, respectively. The results by the proposed method are compared with the method in [7].

When the image is doubly compressed with quality factors $Q_1 = 80$ and $Q_2 = 90$, or by using $q_1 = 5$ and $q_2 = 3$ as the first and second quantisation steps, respectively, where $q_1 = 5$ is at the position (3, 2) of the low-frequency region of the primary quantisation table, then the $p(q_1)$ function obtained from (7) is as depicted in Fig. 1*a*, and the estimated q_1' is 5. Similarly, Fig. 1*b* shows the $p(q_1)$ function when the image is doubly compressed with quality factors $Q_1 = 80$ and $Q_2 = 90$, but using $q_1 = 27$ and $q_2 = 14$ as the first and second quantisation steps, respectively, where $q_1 = 27$ is at the position (5, 5) of the high-frequency region of the primary quantisation table, and the estimated q_1' is 28. In comparison, the $p(q_1)$ functions of single compression with $Q = 80$, or $q_1 = 5$ and $q_1 = 27$, are depicted in Figs. 1*c* and *d*, respectively. In these conditions, q_1' does not exist since its value is required to be bigger than the present known quantisation step where Table 1 (*a*) is the real quantisation table with quality factor $Q = 80$, and Table 1 (*b*) and Table 1 (*c*) are the estimated versions of the method proposed in [7] and our method, for the recompressed Lena image.

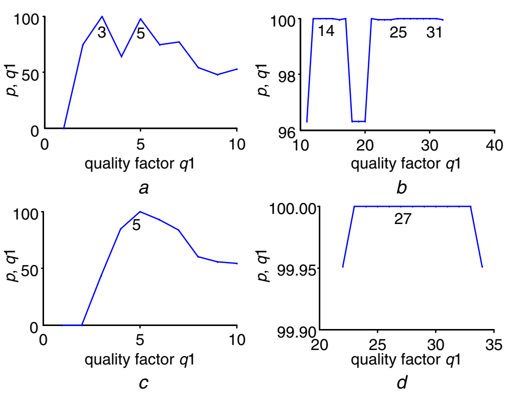

Fig. 1 *$p(q_1)$ functions with q_1 = 5 and q_2 = 3 at position (3,2) of quantisation table (Fig. 1a); q_1 = 27 and q_2 = 14 at position (5,5) of quantisation table (Fig. 1b); $p(q_1)$ functions of single compression with Q = 80, or q_1 = 5 and q_1 = 27 (Figs. 1c and d)*

Fig. 2 shows the forgery region locating results using the block artefact measure defined in (8), where Figs. 2*a* and *b* denote the normal image and forged image. Furthermore, when the block artefact measure defined in (8) is applied to the estimated quantisation tables of Fig. 2*b* using the method proposed in [7], the detected tampering regions are as shown in Fig. 2*c*; the corresponding results of our method are shown in Fig. 2*d*.

Table 1: Real quantisation table with quality factor $Q = 80$ (a), and estimated versions using method in [7] (b) and proposed method (c)

6	4	4	6	10	16	20	24	6	4	4	6	10	16	21	22	6	4	4	6	10	16	20	24
5	5	6	8	10	23	24	22	5	5	6	8	10	23	22	17	5	5	6	8	10	23	23	22
6	5	6	10	16	23	24	22	6	5	6	10	16	22	22	17	6	5	6	10	16	23	28	22
6	7	9	12	20	35	32	25	6	7	9	12	19	26	25	19	6	7	9	12	20	34	32	24
7	9	15	22	27	44	41	31	7	9	15	20	22	34	32	23	7	9	15	22	28	44	42	30
10	14	22	26	32	42	45	37	10	14	20	21	25	32	79	28	10	14	22	26	32	42	46	36
20	26	31	35	41	48	48	40	18	20	25	26	47	35	41	49	20	26	32	34	42	48	48	40
29	37	38	39	45	40	41	40	57	49	39	43	34	55	1	1	28	38	38	40	44	40	40	40
a								*b*								*c*							

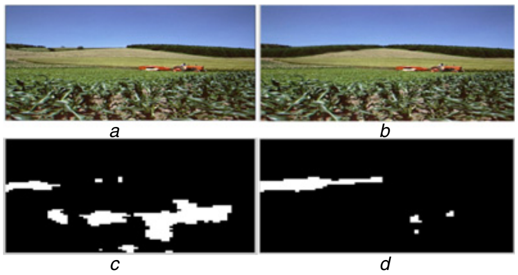

Fig. 2 *Normal image (Fig. 2a); forgery image (Fig. 2b); tampered region detection results using method in [7] (Fig. 2c); our method estimate (Fig. 2d)*

To conclude, the proposed method outperforms the method in [7], especially in accurate estimation of the high-frequency region of the quantisation table. From the experimental results, we can see that the accurate estimation is very useful for locating the tampered region.

Conclusion: Blind detection of a doctored JPEG image operation based on accurate estimation of the whole quantisation table is proposed in this Letter, where we set an image compression model to obtain the relationship between the primary compression and the second compression, define a $p(q_1)$ function to estimate the quantisation step, and provide a block artefact measure to detect the tampered region of a forgery image.

Acknowledgments: This research was supported by grants of the National Natural Science Foundation of China (60862003), the Project of the International S&T Cooperation Programme of China (2009DFR10530), the Doctoral Fund of the Ministry of Education of China (20095201110002), the Industrial-Key-Problem-Tackling Programme of Guizhou S&T Bureau (Qian Ke He GY Word (2010) 3054) and the Graduate Innovation Fund of Guizhou University (2013044).

11 July 2013

doi: 10.1049/el.2013.2218

One or more of the Figures in this Letter are available in colour online.

Yuanyuan Peng and Benyong Liu (*College of Computer Science and Information/Institute of Intelligent Information Processing, Guizhou University, Guiyang 550025, People's Republic of China*)

E-mail: csc.byliu@gzu.edu.cn

References

1 Popescu, A.C., and Farid, H.: 'Statistical tools for digital forensics'. Proc. 6th Int. Workshop on Information Hiding, Toronto, Canada, 2004

2 Farid, H.: 'A survey of image forgery detection', *IEEE Signal Process.*, 2009, **2**, (26), pp. 622–645

3 Luo, W., Huang, J., and Qiu, G.: 'JPEG error analysis and its applications to digital image forensics', *IEEE Trans. Inf. Forensics Sec.*, 2010, **5**, (3), pp. 480–491

4 Lukas, J., and Fridrich, J.: 'Estimation of primary quantization matrix in double compressed JPEG images'. Digital Forensic Research Workshop, Cleveland, OH, USA, August 2003

5 Bianchi, T., Rosa, A.D., and Piva, A.: 'Improved DCT coefficient analysis for forgery localization JPEG images'. Proc. ICSSSP 2011, Prague, Czech Republic, May 2011, pp. 2444–2447

6 Lin, S., Chang, M., and Chen, Y.: 'A passive-blind forgery detection scheme based on content-adaptive quantization table estimation', *IEEE Trans. Circuits Syst. Video Technol.*, 2011, **21**, (4), pp. 421–434

7 Bianchi, T., and Piva, A.: 'Image forgery localization via block-grained analysis of JPEG artifacts', *IEEE Trans. Inf. Forensics Sec.*, 2012, **7**, (3), pp. 1003–1017