TWO-STEP PROCEDURE TO DETECT MODIFIED JPEG IMAGES

A.M. Kapusta
Research Institute for Applied Problems of Mathematics and Informatics
Minsk, BELARUS
e-mail: `kapusta@bsu.by`

Abstract

In this paper a two-step procedure to detect modified JPEG images is introduced. It reduces the number of false positives with slight decreasing of the detection rate.

1 Introduction

Consider a problem of processing a large amount of digital images and detecting modified ones. JPEG image modification is carried out changing least significant bits (LSB) of quantized discrete cosine transform (DCT) coefficients. The modifications are uniformly scattered over the whole image.

Classical classifiers with significance level α have a drawback. The number of false positives (type I errors) will be extremely large. Reducing of the significance level is unacceptable due to the sharp decreasing of the modified images detectability.

As the dimensionality of JPEG image space $\mathcal{X}$ is extremely large, in practice, individual images from $\mathcal{X}$ are represented using a simplified model. One possibility is to project each image $x \in \mathcal{X}$ onto a space of a much smaller dimension formed by "features", that, in some sense, captures everything important about the image x.

In sections 2-3 two approaches to construct informative features are presented. The two-step procedure is presented in section 4. Numerical results illustrating the proposed two-step procedure are presented in section 5.

2 DCT features

Introduce a feature set to describe macroscopic characteristics of JPEG images [1]. Suppose that a JPEG image is represented with a DCT coefficient array d_{ij}^k, $i, j = 1, \ldots, 8$, $k = 1, \ldots, n_B$, where n_B is the total number of DCT coefficient 8×8-blocks in the image.

Denote the global histogram of all $8 \times 8 \times n_B$ DCT coefficients as $H = (H_L, \ldots, H_R)$, where $L = \min\limits_{i,j,k} d_{ij}^k$, $R = \max\limits_{i,j,k} d_{ij}^k$; and let h_r^{ij}, $r = L, \ldots, R$, denotes the individual histogram of values d_{ij}^k.

Define the dual histogram for a fixed DCT coefficient value d as follows

$$g_{ij}^d = \sum_{i,j=1}^{8} \sum_{k=1}^{n_B} \delta(d, d_{ij}^k), \tag{1}$$

where $\delta(u,v) = 1$, if $u = v$, and 0 otherwise.

Let I_r and I_c denote the vectors of block indices while scanning the image by rows and by columns respectively. As a measure capturing inter-block dependencies the variation V is used.

$$V = \frac{\sum\limits_{i,j=1}^{8} \sum\limits_{k=1}^{|I_r|-1} |d_{ij}^{I_r^{(k)}} - d_{ij}^{I_r^{(k+1)}}| + \sum\limits_{i,j=1}^{8} \sum\limits_{k=1}^{|I_c|-1} |d_{ij}^{I_c^{(k)}} - d_{ij}^{I_c^{(k+1)}}|}{|I_r| + |I_c|}. \tag{2}$$

As a measure of discontinuities along DCT coefficient 8×8-block boundaries blockiness measures B_α, $\alpha = 1, 2$, are used. The blockiness measures are calculated from the decompressed JPEG image as follows

$$B_\alpha = \frac{\sum\limits_{i=1}^{\lfloor (M-1)/8 \rfloor} \sum\limits_{j=1}^{N} |x_{8i,j} - x_{8i+1,j}|^\alpha + \sum\limits_{j=1}^{\lfloor (N-1)/8 \rfloor} \sum\limits_{i=1}^{M} |x_{i,8j} - x_{i,8j+1}|^\alpha}{N\lfloor (M-1)/8 \rfloor + M\lfloor (N-1)/8 \rfloor}, \tag{3}$$

where $x_{i,j}$ are grayscale values of decompressed JPEG image, M and N are image dimensions.

The probability distribution of pairs of neighboring DCT coefficients is described by a co-occurance matrix C, that defined as

$$C_{st} = \frac{\sum\limits_{k=1}^{|I_r|-1} \sum\limits_{i,j=1}^{8} \delta(s, d_{ij}^{I_r^{(k)}})\delta(t, d_{ij}^{I_r^{(k+1)}}) + \sum\limits_{k=1}^{|I_c|-1} \sum\limits_{i,j=1}^{8} \delta(s, d_{ij}^{I_c^{(k)}})\delta(t, d_{ij}^{I_c^{(k+1)}})}{|I_r| + |I_c|}. \tag{4}$$

3 Markov features

Introduce a feature set to describe microscopic characteristics of JPEG images [3]. The feature calculation starts by forming the matrix $F(u,v)$ of absolute values of DCT coefficients in the image. The DCT coefficients in $F(u,v)$ are arranged in the same way as pixels in the image by replacing each 8×8-block of pixels with the corresponding block of DCT coefficients. Next, 4 difference arrays are calculated along 4 directions: horizontal, vertical, diagonal, and minor diagonal (further denoted as $F_h(u,v)$, $F_v(u,v)$, $F_d(u,v)$, $F_m(u,v)$ respectively)

$$\begin{aligned} F_h(u,v) &= F(u,v) - F(u+1,v), \\ F_v(u,v) &= F(u,v) - F(u,v+1), \\ F_d(u,v) &= F(u,v) - F(u+1,v+1), \\ F_m(u,v) &= F(u+1,v) - F(u,v+1). \end{aligned} \tag{5}$$

From these difference arrays 4 transition probability matrices M_h, M_v, M_d, M_m are constructed

$$
\begin{aligned}
M_h(i,j) &= \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v}\delta(F_h(u,v)=i,F_h(u+1,v)=j)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v}\delta(F_h(u,v)=i)},\\
M_v(i,j) &= \frac{\sum_{u=1}^{S_u}\sum_{v=1}^{S_v-2}\delta(F_v(u,v)=i,F_v(u,v+1)=j)}{\sum_{u=1}^{S_u}\sum_{v=1}^{S_v-1}\delta(F_v(u,v)=i)},\\
M_d(i,j) &= \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v-2}\delta(F_d(u,v)=i,F_d(u+1,v+1)=j)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta(F_d(u,v)=i)},\\
M_m(i,j) &= \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v-2}\delta(F_m(u+1,v)=i,F_m(u,v+1)=j)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta(F_m(u,v)=i)},
\end{aligned} \qquad (6)
$$

where S_u and S_v denote the dimensions of the image.

4 Two-step procedure

Adopt multiple hypothesis testing procedure in [2] to the problem of processing a large amount of digital images. We aim to provide a tiny level of false positives (type I errors) without significant loss of detection accuracy.

Define two criteria C_1 and C_2 to test null hypothesis $\mathcal{H}_0$, that image isn't modified. Alternative hypothesis $\mathcal{H}_1$ states that image is modified. Suppose, that criteria decisions for the most of the processed images take the same value, but take different decisions for some images, thus they are not equivalent.

Propose a two-step procedure to test $\mathcal{H}_0$.

Step 1. Testing $\mathcal{H}_0$ using criterion C_1 with significance level α_1.

Step 2. Clarifying the decision using witness criterion C_2 with significance level α_2 to sift false positives of the Step 1.

Suggest one-tailed criteria with normally distributed statistics under the null hypothesis. In [2] an upper bound α_+ for type I error α^* of two-step procedure was estimated

$$\alpha^* \le \alpha_+ = \alpha_1\alpha_2 + \sum_{k=1}^{+\infty} \frac{4p^{(2k-1)}(\Delta_1(\alpha_1))p^{2k-1}(\Delta_2(\alpha_2))}{(2k)!\rho_i^{2k}}, \qquad (7)$$

where $p^{(k)}(\Delta)$ is the k-th derivative of standard normal probability density function at the point Δ, ρ is the correlation of criteria statistics.

It can be seen, that suggested type I error of the proposed two-step procedure has order $\alpha_1\alpha_2$.

5 Results

For the experiments training and examination sets of JPEG images was prepared. Training set consists of 3000 empty images and 3000 modified images. Examination set consists of 5000 empty images and 5000 modified images. Initial feature sets were reduced to exclude most correlated features. The remaining sets of 100 DCT features and 53 Markov features was used for the experiments.

Several classifiers were trained. The first classifier utilizes macroscopic DCT feature set. The second one utilizes microscopic Markov feature set. The third one utilized the united set of 153 features. The last classifier is based on the two-step procedure utilizing DCT features on the first step and Markov features on the second step.

Classification accuracy was estimated using both training sample re-classification and examination sample classification. False positives rate and modified images detection rate are presented in the table 1.

Table 1: Classification accuracy

Classification procedure	Training set		Examination set	
	false postitives	detection rate	false positives	detection rate
Based on DCT features	1.26%	99.48%	3.16%	97.44%
Based on Markov features	1.72%	99.48%	2.82%	97.88%
Based on both features	1.13%	99.66%	2.46%	98.68%
Two-step procedure	0.26%	98.92%	0.36%	96.32%

As can be seen from the results presented in the table 1, two-step procedure can be effectively used for the problem of classifying a large amount of digital images. It provides considerably smaller number of false positives almost preserving detection accuracy.

This result can be reached only using little correlated detection criteria statistics. In the paper two image feature sets of different nature with relatively small correlation were constructed for the criteria. Numerical results confirm the suggestions and illustrate the proposed two-step procedure.

References

[1] Fridrich J., Pevny T. (2007). Merging Markov and DCT features for multi-class JPEG analysis. *Proceedings SPIE Electronic Imaging*. Vol. **6505**, pp. 3 1-3 13.

[2] Nikitina I.S. (2011). Analysis of the properties of multiple hypothesis testing procedures. *International congress on Computer Science: Information Systems and Technologies*. Vol. **1**, pp. 123-128.

[3] Shi Y.Q., Chen C., Chen W. (2006). A Markov process based approach to effective atacking JPEG. *Lecture Notes in Computer Science*. Vol. **4437**, pp. 249-264.