Revealing the Traces of Histogram Equalisation in Digital Images

Ravi Shankar(22EE65R01)
Vikas Goyal(22MM61R07)
Pankaj Kumar Bajiya(22EE65R12)

Keywords: Histogram Equalisation(HE), Image Resolution(IR), Discrete Fourier Transform(DFT)

Abstract

Generally HE technique is used to improve image contrast, but now a days it is also used to temper the images to propagate false information and existing image forensics techniques may fails to detect HE operation and highly compressed and low resolution images. therefore, the propose algorithm can detect HE operation even if the image is saved in JPEG format after equalisation where most of the existing algorithms fail.

1 Introduction

1.1 Motivation

In the automated era of information war, it is very easy to use image editing tools for even an inexperienced person to temper the images to propagate false information to mislead the target.Now a days a bigger section of population shares images and videos in very high volume to various social media platforms, websites etc. which may be used to propagate the wrong information through this type of content. The main aim of the digital image forensics to detect the nature of manipulation done in image. Since it is impossible to detect all type of forgeries with a single method, a number of methods to detect image alteration under different scenarios.

There are two types of image authentication approaches classified as active and passive. Most popular example of active approach is digital watermarking, the watermark is embedded into the image at the time of image generation or before any processing, however the passive approach does not need any prior decision. The passive forensic analysis has played an important role in popular research field because of its applicability and practicality.

Recently due to wide acceptability, the forensic issues related JPEG compressed images have received a great attention. many algorithms have been developed to determine whether a given JPEG image is doubly compressed or not. This method consider double compression as a possible indication of image forgery, however sometimes image may be simply decompress and compress again without any manipulation just to satisfy the image size for uploading purpose. Many forensic developed a dodge existing forensic method but at the same time there has been continuous effort to develop algorithm to cope up with anti-forensic method.

It is observed that generally image forgery operations are used some filtering and image enhancement techniques to hide the traces of image manipulation to give a realistic view. So this report proposed methods to recognise, whether the given image is manipulated or not before the compression.

1.2 Histogram equalisation

HE is a one of the most widely used contrast enhancement technique, which transforms the input histogram (normalised histogram is the probability distribution function) into uniformly distributed output histogram. The histogram of processed image is uniformly spread in the entire dynamic range of pixel intensities [39]. The transformation function for HE is defined mathematically as

$$y_k = T\left(x_k\right) = \sum_{i=0}^{k} p_x x\left(i\right) = \sum_{i=0}^{k} \frac{n_i}{n}; k = 0, 1, ..L - 1$$

(1)

Here, variable x represents the pixel intensities of input image that is to be enhanced and y is the output value for each input x to the transformation T, ni is the number of pixels with ith intensity, n is the total number of pixels in the image and L is the number of possible intensity levels (L=256 for grey-scale images).

1.3 Effect of compression on histogram contour

In order to reduce bandwidth requirement for transmission over the Internet/mobile networks or to satisfy the file-size limitations for uploading, many times images need to be highly compressed. This is generally achieved through the use of large step-size quantiser, which causes unsmoothness in the contour of image histogram. Since JPEG is widely used for compression of images shared over social media, we consider JPEG compressed images to study the effect of compression (or quantisation) on image histogram.In JPEG compression, DCT coefficients are quantised by dividing them with corresponding elements of a predefined quantisation matrix Q. The elements of Q matrix are scaled according to the QF. For a specified QF, elements of matrix Q are generally large for high-frequency coefficients compared to the DC and lowfrequency coefficients.

To achieve a high degree of compression, more and more high-frequency DCT coefficients are quantised to zero value according to

$$Y\left(u,v\right)=round\left(\frac{F\left(u,v\right)}{q\left(u,v\right)}\right)=0;if\ q\left(u,v\right)>2F\left(u,v\right). \tag{2}$$

where Y(u, v) is the quantised version of DCT coefficient F(u, v) and q(u, v) is the corresponding element of the quantisation matrix. As indicated in (2), Y(u, v) will be rounded to zero if q(u, v) is greater than twice the magnitude of F(u, v). Since at lower value of QF (high degree of compression), most of the elements in quantisation matrix are of large magnitude, many DCT (mainly high frequency) coefficients are likely to be quantised to zero value. Due to loss of these high frequency coefficients, the reconstructed image appears to be relatively smoother (containing only low-frequency components) compared to images compressed at higher QF. That is, in a highly compressed image, the majority of pixels within the neighbourhood of a given pixel are likely to be of similar intensity.

2 Proposed method for identifying histogram equalised images

In this section, we propose a very generic method to detect HE process in image, which is based on the frequency domain features of the images histogram and is able to reveal the traces of HE even if the equalised image is JPEG compressed.Refer Figure 1.

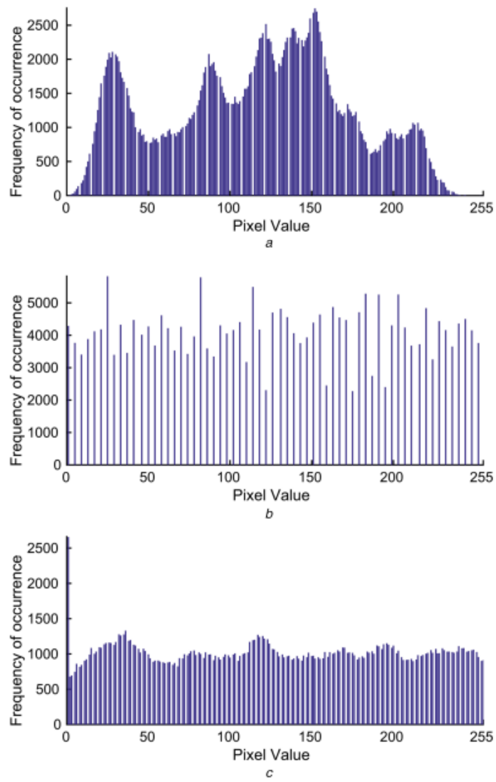
2.1 Discrete Fourier transform (DFT) of image histogram

An image histogram is the representation of frequency of occurrences of grey-scale intensities in the image, and can be considered as a 1-D discrete signal (with intensity as an independent variable and count of intensity as a dependent variable). The information about the degree of changes (or frequency of abrupt changes) in the histogram can be obtained by transforming histogram in frequency domain. For an 8-bit grey-scale image I, with pixel intensities varying in the range n = 0, 1, 2, . . . , 255, and if H(n) represents count of nth intensity level, then N point (here N= 256) DFT of the histogram H can be defined as

$$X\left(k\right)=\sum_{n=0}^{N-1}H\left(n\right)exp\left(-\frac{j2\pi nk}{N}\right);k=0,1..N-1 \tag{3}$$

where N represents the number of equally spaced points in the interval [0, 2] on a unit circle in the Z-plane. Since H(n) is a real valued discrete sequence, its DFT X(k) will be a complex quantity and will exhibit complex conjugate symmetry. Thus, for a real Npoint discrete time 1D signal H(n), from (3), it can be easily verified that

$$X\left(N-k\right)=X^{*}\left(k\right)=X\left(-k\right) \tag{4}$$



t

Figure 1. (a) Original image, (b) Its equalised version, (c) Its equalised plus JPEG compressed version

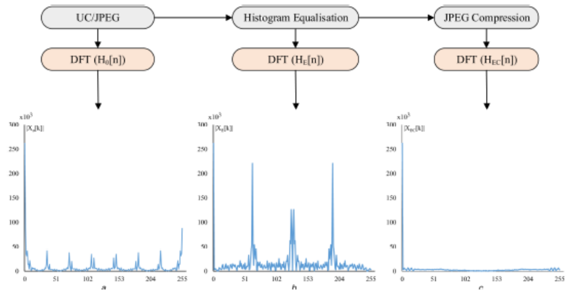


Figure 2. (a) Magnitude plot of the DFT of greyscale histogram of an original unequalised image. (b) Magnitude plot of the DFT of greyscale histogram of the equalised version of the image. (c) Magnitude plot of the DFT of greyscale histogram of an equalised then compressed image

$$|X\left(N-k\right)|=|X\left(k\right)| \tag{5}$$

Therefore, from (4) and (5), it can be inferred that for an even N (which is the case here), apart from a DC term X(0), the magnitude of X(k), i.e. $|X(k)|$ will have only N/2 unique values which are symmetrically located on both sides of k=N/2. The proposed algorithm uses frequency domain characteristics of histogram of given image to detect HE operation on it. Refer Figure 2.

2.2 Proposed algorithm

In order to describe the proposed algorithm, consider that an unequalised original image I_o (which may be either uncompressed or JPEG compressed) has undergone the HE process followed by JPEG compression, resulting in images IE and IEC, respectively. Let H_o, H_E and H_{EC} are the histograms of I_o, I$_E$ and I_{EC} images, respectively. We believe that exploiting the frequency domain information inherent in the processed (HE process in our case) images, the nature of the process applied to generate the corresponding image can be identified. Let $X_o(k)$, $X_E(k)$ and $X_{EC}(k)$ are DFT coefficients of H_o, H_E andH_{EC} obtained using (3), respectively.

The magnitude spectrum of histograms of a typical unequalised (uncompressed or JPEG compressed) $(|X_o(k)|)$, HE image $(|X_E(k))$ and JPEG compressed equalised image $(|X_{EC}(k)|)$ are shown in Figure 2a–c, respectively. In each $|X(k)|, k=0$ represents the corresponding DC term and can be written as $|X\left(0\right)|=\sum_{n=0}^{N-1}H\left(n\right)$, where H is the histogram whose DFT is to be evaluated. It may be noted that $|X(0)|$ is equal to the total number of pixels in the image, which will remain the same before and after the HE and/or JPEG compression.

We propose to identify the HE process in an image by observing the difference between DC coefficient $(|X(0)|)$ and weighted sum of unique AC coefficients $\left(|X\left(k\right)|; k\,\epsilon\left\{1,128\right\}\right)$ of DFT of its histogram. As evident from Fig. 1a, the histogram of an unequalised image likely to have random envelope, with many abrupt changes in its shape, and therefore corresponding $\left(|X\left(k\right)|; k\,\epsilon\left\{1,128\right\}\right)$ are likely to have larger values, which can be verified in Figure 1(a). Since it is expected that the HE process generates an image having almost uniformly distributed histogram, and therefore the values of $\left(|X\left(k\right)|; k\,\epsilon\left\{1,128\right\}\right)$ are likely to be zero (or very small value). The similar observations can be made for an image which is JPEG compressed after HE as shown in Figure 1(c). This is because, when an equalised image is JPEG compressed, resulting in redistribution of pixel intensities in its histogram; however, the overall contour of the histogram still remains almost uniform.Therefore, in histogram equalised images, the AC coefficients are expected to be much smaller as compared to DC coefficient. However, in practice the histogram of equalised image is not perfectly uniform and often contains sudden jumps or breaks, which may sometime results in large value of mid and high frequency components in its DFT. The AC coefficients should be weighted in such a way that low frequency components are emphasised more than the mid and high frequency components. Based on above discussions and using the DFT coefficients of the histograms, we propose that the DC normalised difference between DC and weighted sum of AC coefficients of DFT of the image histogram can be used to identify whether the given image is histogram equalised or not. For this purpose, we define a new parameter ξ as

$$\xi=\frac{|X(0)|-\delta}{X(0)}$$

where δ is the weighted sum of magnitude of AC coefficients as defined in above equation and weights are decaying exponential with as a decay factor

$$\delta=\sum_{k=1}^{128}|X\left(k\right)|exp\left(-\beta\left(k-1\right)\right)$$

The purpose of is to deemphasise mid- and high-frequency components. Experiments show that = 1 is appropriate.

2.3 Algorithm Used

Initialization:
$I \leftarrow Input\ Image$
$H \leftarrow Histogram$
$X \leftarrow DFT\ of\ H$
$for\ each\ input\ image\ I\ do$
$for\ each\ grayscale\ value\ P\epsilon\ 0,1,2.......,255\ do$
$|\ H_I\left(P\right)=\ count\ occurance\ of\ P$
end
$X_I \leftarrow DFT\left(H_I\right)$
$compute\ \xi\ using\ equation\ given\ below$
$if\ \xi \geq \xi_{Th}\ then$
$|\ I\ :\ Histogram\ Equalized$
end
end

3 Results and discussion

For the proposed method, the threshold values obtained are 0.8473, 0.8437, 0.8139, 0.8325 0.8235, 0.8327, 0.8576 for image size 512 × 384, 256 × 256, 128 × 128, 64 × 64, 32 × 32, 16 × 16 and 8 × 8, respectively. It must be noted that we have performed evaluation on small-size images generated by both central cropping and down sampling of the original images. Centrally cropped (CC) images, in contrast to down sampled images (DS), does not preserve the overall structure of the actual image. However, since in copy and move forgery, a cropped portion of the image is generally used to hide or add some data.

4 References

Zeeshan Akhtar, Ekram Khan: "Revealing the traces of histogram equalisation in digital images ",(IET Image processing journal 2018)