# Statistics- and Spatiality-based Feature Distance Measure for Error Resilient Image Authentication

Shuiming Ye[1,2], Qibin Sun[1], and Ee-Chien Chang[2]

[1] Institute for Infocomm Research, A*STAR, Singapore, 119613
[2] School of Computing, National University of Singapore, Singapore, 117543
{Shuiming, Qibin}@i2r.a-star.edu.sg, Changec@comp.nus.edu.sg

**Abstract.** Content-based image authentication typically assesses authenticity based on a distance measure between the image to be tested and its original. Commonly employed distance measures such as the Minkowski measures (including Hamming and Euclidean distances) may not be adequate for content-based image authentication since they do not exploit statistical and spatial properties in features. This paper proposes a feature distance measure for content-based image authentication based on statistical and spatial properties of the feature differences. The proposed statistics- and spatiality-based measure (*SSM*) is motivated by an observation that most malicious manipulations are localized whereas acceptable manipulations result in global distortions. A statistical measure, kurtosis, is used to assess the shape of the feature difference distribution; a spatial measure, the maximum connected component size, is used to assess the degree of object concentration of the feature differences. The experimental results have confirmed that our proposed measure is better than previous measures in distinguishing malicious manipulations from acceptable ones.

**Key words:** Feature Distance Measure, Image Authentication, Image Transmission, Error Concealment, Digital Watermarking, Digital Signature.

## 1 Introduction

With the wide availability of digital cameras and image processing software, the generation and manipulation of digital images are easy now. To protect the trustworthiness of digital images, image authentication techniques are required in many scenarios, for example, applications in health care.

Image authentication, in general, differs from data authentication in cryptography. Data authentication is designed to detect a single bit change whereas image authentication aims to authenticate the content but not the specific data representation of an image [1], [2]. Therefore, image manipulations which do not change semantic meaning are often acceptable, such as contrast adjustment, histogram equalization, and compression [3], [4]. Lossy transmission is also considered as acceptable since errors under certain level in images would be tolerable and acceptable [5]. Other manipulations that modify image content are classified as malicious manipulations, such as object removal or insertion. Image authentication is desired to be robust to acceptable manipulations, and necessary to be sensitive to malicious ones.

In order to be robust to acceptable manipulations, several content-based image authentication schemes have been proposed [6], [7], [8]. These schemes may be robust to one or several specific manipulations, however, they would classify the image damaged by transmission errors as unauthentic [9]. Furthermore, content-based image authentication typically measures authenticity in terms of the distance between a feature vector from the received image and its corresponding vector from the original image, and compares the distance with a preset threshold to make a decision [10], [11]. Commonly employed distance measures, such as the Minkowski metrics [12] (including Hamming and Euclidean distances), may not be suitable for robust image authentication. The reason is that even if these measures are the same (e.g., we cannot tell whether the question image is authentic or not), the feature difference patterns under typical acceptable modifications or malicious ones may be still distinguishable (feature differences are differences between the feature extracted from the original image and the feature extracted from the testing image). That is to say, these measures do not properly exploit statistical or spatial properties of image features. For example, the Hamming distance measures of Fig. 1(b) and Fig. 1(d) are almost the same, but yet, one could argue that Fig. 1(b) is probably distorted by malicious tampering since the feature differences concentrate on the eyes.

The objective of this paper is to propose a distance measure based on statistical and spatial properties of the feature differences for content-based image authentication. The proposed measure is derived by exploiting the discernable patterns of feature differences between the original image and the distorted image to distinguish acceptable manipulations from malicious ones. Two properties, the kurtosis of the feature difference distribution and the maximum connected component size in the feature differences, are combined to evaluate the discernable patterns. We call the proposed measure statistics- and spatiality-based measure (*SSM*) since it considers both global statistical properties and spatial properties. Many acceptable manipulations, which were detected as malicious modifications by previous schemes based on Minkowski metrics, were correctly verified by the proposed scheme based on *SSM*. To illustrate how the proposed *SSM* can improve the performance of image authentication scheme, we applied it in a semi-fragile image authentication scheme [13] to authenticate images damaged by transmission errors. The proposed error resilient scheme obtained better robustness against transmission errors in JPEG or JPEG2000 images and other acceptable manipulations than the scheme proposed in [13].

# 2 Proposed Statistics- and Spatiality-based Measure (*SSM*) for Image Authentication

Content-based or feature-based image authentication generally verifies authenticity by comparing the distance between the feature vector extracted from the testing image and the original with some preset thresholds [14]. The distance metric commonly used is the Minkowski metric $d(X, Y)$ [12]:

$$d(X,Y) = (\sum_{i=1}^{N} |x_i - y_i|^r)^{1/r} \qquad (1)$$

where $X$, $Y$ are two $N$ dimensional feature vectors, and $r$ is a Minkowski factor. Note that when $r$ is set as 2, it is actually Euclidean distance; when $r$ is 1, Manhattan distance (or Hamming distance for binary vectors).

However, the Minkowski metric does not exploit statistical or spatial properties of image features. Therefore, the image authentication scheme based on Minkowski metric may not be suitable to distinguish the tampered images (e.g., small local objects removed or modified) from the images by acceptable manipulations such as lossy compression. On the other hand, we found that even if the Minkowski metric distances are the same, the feature difference under typical acceptable manipulations and malicious ones are still distinguishable especially in the case that the feature contains spatial information such as edges or block DCT coefficients. Therefore, the Minkowski metric is not a proper measure for content-based image authentication.

## 2.1 Main Observations of Feature Differences

Many features used in content-based image authentication are composed of localized information about the image such as edges [3], [6], block DCT coefficients [1], [10], [13], highly compressed version of the original image [7], or block intensity histogram [11]. To facilitate discussions, we let $x_i$ be the feature value at spatial location $i$, and $X$ be an $N$-dimension feature vector, for example, $N = W \cdot H$ when using edge feature ($W$ and $H$ are the width and height of the image). We define the feature difference vector $\delta$ as the difference between feature vector $X$ of the testing image and feature vector $Y$ of the original image:

$$\delta_i = |x_i - y_i| \qquad (2)$$

where $\delta_i$ is the difference of features at spatial location $i$.

After examining many discernable feature difference patterns from various image manipulations, we could draw three observations on feature differences:

1. The feature differences by most acceptable operations are evenly distributed spatially, whereas the differences by malicious operations are locally concentrated.

2. The maximum connected component size of the feature differences caused by acceptable manipulations is usually small, whereas the one by malicious operation is large.
3. Even if the maximum connected component size is fairly small, the image could have also been tampered with if those small components are spatially concentrated.

These observations are supported by our intensive experiments and other literatures mentioned previously [6], [9]. Image contents are typically represented by objects and each object is usually represented by spatially clustered image pixels. Therefore, the feature to represent the content of the image would inherit some spatial relations.

A malicious manipulation of an image is usually concentrated on modifying objects in image, changing the image to a new one which carries different visual meaning to the observers. If the contents of an image are modified, the features around the objects may also be changed, and the affected feature points tend to be connected with each other. Therefore, the feature differences introduced by a meaningful tampering typically would be spatially concentrated.

On the contrary, acceptable image manipulations such as image compression, contrast adjustment, and histogram equalization introduce distortions globally into the image. The feature differences may likely to cluster around all objects in the image, therefore they are not as concentrated locally as those by malicious manipulations. In addition, many objects may spread out spatially in the image, thus the feature differences are likely to be evenly distributed with little connectedness. The distortion introduced by transmission errors would also be evenly distributed since the transmission errors are randomly introduced into the image [18].

The above observations not only prove the unsuitability of Minkowski metric to be used in image authentication, but also provide some hints on how a good distance function would work: it should exploit the statistical and spatial properties of feature differences. These observations further lead us to design a new feature distance measure for content-based image authentication.

### 2.2 Proposed Feature Distance Measure for Image Authentication

Based on the observations discussed so far, a feature distance measure is proposed in this section for image authentication. The distance measure is based on the differences of the two feature vectors from the testing image and from the original image. Two measures are used to exploit statistical and spatial properties of feature differences, including the kurtosis (*kurt*) of feature difference distribution and the maximum connected component size (*mccs*) in the feature difference map. Observation (1) motivates the uses of the kurtosis measure, and observation (2) motivates the uses of the *mccs* measure. They are combined together since any one of the above alone is still insufficient, as stated in observation (3).

The proposed Statistics- and Spatiality-based Measure (*SSM*) is calculated by sigmoid membership function based on both *mccs* and *kurt*. Given two feature vectors $X$ and $Y$, the proposed feature distance measure $SSM(X,\, Y)$ is defined as follows:

$$SSM(X,Y) = \frac{1}{1 + e^{\,\alpha(mccs \cdot kurt \cdot \theta^{-2} -\, \beta)}} \tag{3}$$

The measure $SSM(X,\, Y)$ is derived from the feature difference vector $\delta$ defined in Eq. (2). The *mccs* and *kurt* are obtained from $\delta$, and their details are given in the next few paragraphs. $\theta$ is a normalizing factor.

The parameter $\alpha$ controls the changing speed especially at the point $mccs \cdot kurt \cdot \theta^{-2} = \beta$. $\beta$ is the average $mccs \cdot kurt \cdot \theta^{-2}$ value obtained by calculating from a set of malicious attacked images and acceptable manipulated images. In this paper, the acceptable manipulations are defined as contrast adjustment, noise addition, blurring, sharpening, compression and lossy transmission (with error concealment); the malicious tampering operations are object replacement, addition or removal. During authentication, if the measure $SSM(X,\, Y)$ of an image is smaller than 0.5 (that is, $mccs \cdot kurt \cdot \theta^{-2} < \beta$, the image is identified as authentic, otherwise it is unauthentic.

**Kurtosis.** Kurtosis describes the shape of a random variable's probability distribution based on the size of the distribution's tails. It is a statistical measure used to describe the concentration of data around the mean. A high kurtosis portrays a distribution with fat tails and a low even distribution, whereas a low kurtosis portrays a distribution with skinny tails and a distribution concentrated towards the mean.

Therefore, it could be used to distinguish feature difference distribution of the malicious manipulations from that of the acceptable manipulations.

Let us partition the spatial locations of the image into neighborhoods, and let $N_i$ be the $i$-th neighborhood. That is, $N_i$ is a set of locations that are in a same neighborhood. For example, by dividing the image into blocks of 8×8, we have a total of $W \cdot H/64$ neighborhoods, and each neighborhood contains 64 locations. Let $D_i$ be the total feature distortion in the $i$-th neighborhood $N_i$:

$$D_i = \sum_{j \in N_i} \delta_j \tag{4}$$

We can view $D_i$ as a sample of a distribution $D$. The *kurt* in the Eq. (3) is the kurtosis of the distribution $D$. It can be estimated by:

$$kurt(D) = \frac{\sum\limits_{i=1}^{N} (D_i - \mu)^4}{Num\ \sigma^4} - 3 \tag{5}$$

where $Num$ is the total number of all samples used for estimation. $\mu$ and $\sigma$ is the estimated mean and standard deviation of $D$, respectively.

**Maximum Connected Component Size.** Connected component is a set of points in which every point is connected to all others. Its size is defined as the total number of points in this set. The maximum connected component size (*mccs*) is usually calculated by morphological operators. The isolated points in the feature difference map are first removed and then broken segments are joined by morphological dilation. The maximum connected component size (*mccs*) is then calculated by using connected components labeling on the feature map based on 8-connected neighborhood. Details can be found in [15].

**Normalizing Factor.** Since images may have different number of objects, details as well as dimensions, normalization is therefore needed. Instead of using traditional normalization (i.e., the ratios of the number of extracted feature points to image dimension), we employ a new normalizing factor $\theta$ as:

$$\theta = \frac{\mu}{W \cdot H} \tag{6}$$

where $W$ and $H$ are the width and height of the image respectively. $\mu$ is the estimated mean of $D$, same as that in Eq.(5). The normalized factor $\theta$ makes the proposed measure more suitable for natural scene images.

It is worth noting that the two measures *mccs* and *kurt* should be combined together to handle different malicious tampering. Usually tampering results in three cases in terms of the values of *mccs* and *kurt*: (1) the most general case is that tampered areas are with large maximum connected size and distributed locally (Fig. 1(b)). In this case, both *kurt* and *mccs* are large; (2) small local object is modified such as a small spot added in face (Fig. 2(a)). In this case, the *mccs* is usually very small, but *kurt* is large; (3) tampered areas are with large maximum connected size but these areas are evenly distributed in the whole image (Fig. 2(c)). In this case, the *mccs* is usually large, but *kurt* is small. Therefore, it is necessary for *SSM* to combine these two measures so that *SSM* could detect all these cases of malicious modifications.

## 3 Application of *SSM* to Error Resilient Image Authentication

Image transmission is always affected by the errors due to channel noises, fading, multi-path transmission and Doppler frequency shift [16] in wireless channel, or packet loss due to congestion in Internet [17]. Therefore, error resilient image authentication which is robust to acceptable manipulations and transmission errors is desirable. Based on the proposed feature distance measure, an error resilient image authentication scheme is proposed in this section.

The proposed error resilient scheme exploits the proposed measure in a generic semi-fragile image authentication framework [8] to distinguish images distorted by transmission errors from maliciously modified ones. The experimental results support that the proposed feature distance measure can improve the performance of the previous scheme in terms of robustness and sensitivity.

### 3.1 Feature Extraction for Error Resilient Image Authentication

One basic requirement for selecting feature for content-based image authentication is that the feature should be sensitive to malicious attacks on the image content. Edge-based features would be a good choice because usually malicious tampering will incur the changes on edges. And edge may also be robust to some distortions. For instances, the results in [18] show that high edge preserving ratios can be achieved even if there are uncorrectable transmission errors. Therefore, the remaining issue is to make the edge more robust to the defined acceptable manipulations. Note that this is main reason why we employ the normalization by Eq. (6) to suppress those "acceptable" distortions around edges.

In [19], a method based on fuzzy reasoning is proposed to classify each pixel of a gray-value image into a shaped, textured, or smooth feature point. In this paper we adopt their fuzzy reasoning based detector because of its good robustness.

### 3.2 Image Signing

The image signing procedure is outlined in Fig. 3. Binary edge of the original image is extracted using the fuzzy reasoning based edge detection method [19]. Then, the edge feature is divided into 8×8 blocks, and edge point number in each block is encoded by error correcting code (ECC) [8]. BCH(7,4,1) is used to generate one parity check bit (PCB) for ECC codeword (edge point number) of every 8×8 block. The signature is generated by hashing and encrypting the concatenated ECC codewords using a private key. Finally, the PCB bits embedded into the DCT coefficients of the image. In our implementation, the PCB bits are embedded into the middle-low frequency DCT coefficients using the same quantization based watermarking as in [13].

Let the total selected DCT coefficients form a set $\mathbf{P}$. For each coefficient $c$ in $\mathbf{P}$, it is replaced with $c_w$ which is calculated by:

$$c_w = \begin{cases} Qround(c/Q), \text{ if LSB}(round(c/Q)) = w \\ Q\left(round(c/Q) + sgn\left(c - Qround(c/Q)\right)\right), \text{ else} \end{cases} \tag{7}$$

where $w$ (0 or 1) is the bit to be embedded. Function *round(x)* returns the nearest integrate of $x$, *sgn(x)* returns the sign of $x$, and *LSB(x)* returns the least significant bit of $x$. Eq. (7) makes sure that the LSB of the coefficient is the same as the watermark bit.

Note that embedding procedure should not affect the feature extracted, since the watermarking procedure would introduce some distortions. In order to exclude the effect of watermarking from feature extraction, a compensation operator $C_w$ is adopted before feature extraction and watermarking:

$$\begin{cases} I_c = C_w(I) \\ I_w = f_e(I_c) \end{cases} \tag{8}$$

$$C_w(I) = IDCT\left\{IntQuan\left(d_i, 2Q, \mathbf{P}\right)\right\} \tag{9}$$

where $d_i$ is the $i$-th DCT coefficient of $I$, and IDCT is inverse DCT transform. $f_e$(I) is the watermarking function, and $I_w$ is the final watermarked image. The IntQuan($c$, $\mathbf{P}$, $Q$) function is defined as:

$$IntQuan\,(c, Q, \mathbf{P}) = \begin{cases} c, \text{ if } c \notin \mathbf{P} \\ Q \text{ round}(c/Q), \text{ else} \end{cases} \tag{10}$$

$C_w$ is designed according to the watermarking algorithm, which uses $2Q$ to prequantize the DCT coefficients before feature extraction and watermarking. That is, from Eq. (7), (9) and (10), we can get $C_w(I_w) = C_w(I)$, thus $f_e(I_w) = f_e(I)$, i.e., the feature extracted from the original image $I$ is the same as the one from the watermarked image $I_w$. This compensation operator ensures that watermarking does not affect the extracted feature.

### 3.3 Image Authenticity Verification

The image verification procedure can be viewed as an inverse process of the image signing procedure, as shown in Fig. 4. Firstly, error concealment is carried out if transmission errors are detected. The feature of image is extracted using the same method as used in image signing procedure. Watermarks are then extracted. If there are no uncorrectable errors in ECC codewords, the authentication is based on bit-wise comparison between the decrypted hashed feature and the hashed feature extracted from the image [8]. Otherwise, image authenticity is calculated by the *SSM* based on differences between the PCB bits of the re-extracted feature and the extracted watermark. Finally, if the image is identified as unauthentic, the attacked areas are then detected.

**Error Concealment.** Given an image to be verified, the first step is to conceal the errors if some transmission errors are detected. For wavelet-based images, edge directed filter-based error concealment algorithm proposed in [18] is adopted. For DCT-based JPEG images, a content-based error concealment proposed in [20] is used.

It is efficient and advisable to apply error concealment before image authentication since the edge feature of the error-concealed image is much closer to the original one than that of the damaged image [18], [20]. As a result, the content authenticity of the error concealed image is higher than that of the damaged image, which is validated in our experiments of the error resilient image authentication.

**Image Content Authenticity.** Given an image to be verified, we repeat feature extraction described in image signing procedure. The corresponding PCB bits ($PCB_W$) of all 8×8 blocks (one bit/block) of the image are extracted from the embedded watermarks. Then the feature set extracted from the image is combined with the corresponding PCBs to form ECC codewords. If all codewords are correctable, we concatenate all codewords and cryptographically hash the result sequence. The final authentication result is then concluded by bit-by-bit comparison between these two hashed sets. If there are uncorrectable errors in ECC codewords, image authenticity is calculated based on the proposed distance measure. The two feature vectors in the proposed measure are $PCB_W$ from watermarks and the recalculated PCB bits ($PCB_F$) from ECC coding of the re-extracted image feature set. If the distance measure between $PCB_W$ and $PCB_F$ is smaller than 0.5 ($SSM(PCB_W,\ PCB_F)$ <0.5), the image is authentic. Otherwise, the image is unauthentic.

**Feature Aided Attack Location.** If the image is verified as unauthentic, the tampered areas will be detected. Attack location is an important part of the authentication result since the detected attacked areas give the users a clear figure where the image has been possibly tampered with. The diagram of our feature aided attack location algorithm is shown in Fig. 5. The attack areas are detected using information from watermarks and image feature. The difference map between $PCB_W$ and $PCB_F$ is calculated, and then morphological operations are used to compute connected areas, with isolated pixels and small connected areas removed. After these operations, the difference map is masked with the union of the watermark and feature. The masking operation can refine the detected areas by concentrating them on the objects in the tampered image or in the original image. The areas in the difference map which do not belong to any object (defined by edge feature) are removed, which may be a false alarm of some noises.

It is worth noting that the authentication result of our scheme is friendly to users. Since human perceptivity treats image as a combination of objects, some objects may be region of interest (ROI) to users. If the image fails to pass the authentication, our scheme provides possible attacked areas which concentrate on objects. If these detected areas are not the user's ROI, further decision can be made by the user on a case by case basis. Finally, this scheme can also provide a degree of authenticity (by *SSM* measure) to the user which gives the user a confidence on the trustiness of the image.

## 4 Experimental Results and Discussions

In this Section, the proposed *SSM* is evaluated by experiments, compared with Minkowski metrics and our previous results [13]. In our experiments, JPEG and JPEG2000 image formats were used. Testing images include *Actor, Barbara, Bike, Airplane, Fruits, Girl, Goldhill, Lena, Mandrill, Monarch, Pepper, Woman*, and so on. The dimensions of these images differ among 512×512, 640×512, 640×800, and 720×576. *Daubechies* 9/7 wavelet filter is used for the wavelet transformation which is used in JPEG2000 standard [21]. The parameters $\alpha$ and $\beta$ in Eq.(3) were set to 0.5 and 48.0, respectively.

### 4.1 Feature Distance Measure Evaluation

The observations present in Section 2, which are the basis of the proposed *SSM*, were investigated first in our experiments. Edge detected by [19] was selected as feature in our evaluations. Fig. 6 shows the histogram of edge difference and their respective probability density estimates of noisy, error concealed, damaged and maliciously tampered images. We can find that the distribution of feature differences between malicious tampered image and the original image have a much longer tail than that of the error-concealed image. The damaged, error-concealed and noisy images all have smaller right tails. These results support our observations that the maliciously tampered image has a different pattern of feature differences from that of the acceptable manipulations.

Some acceptable distortions and malicious attacks were introduced into the original images for robustness evaluation. The proposed *SSM* was compared with *Hamming* (Minkowski Metric with $r$=1 for binary feature) as shown in Fig. 7. Pratt's *Figure of Merit* (*FoM*) [22] was also used for comparison, since it is commonly used at measuring image similarity based on edges, which is defined as:

$$FoM = \frac{1}{\max\left(N_O,\, N_C\right)} \sum_{i=1}^{N_C} \frac{1}{1 + \lambda \times di^2} \tag{11}$$

where $N_C$ and $N_O$ are the number of detected and original edge pixels, respectively. $d_i$ is the *Euclidean* distance between the detected edge pixel and the nearest original edge pixel, and $\lambda$ is a constant typically set to 0.1. Fig. 7(a) shows the experimental results of the proposed *SSM* for image *Lena* after JPEG compression, and Fig. 7(b) shows the experimental results for Gaussian noisy images. These figures show that the *Hamming* and *FoM* distances are almost linear to the compression level or Gaussian noise strength. On the contrary, there were some sharper changes (such as the circled points in Fig. 7) in *SSM* curves which may be good choices for authenticity threshold. As an image can be considered as points in a continuous space, it is typically difficult to set up a sharp boundary between authentic and unauthentic images [10]. This intrinsic fuzziness makes the content-based authentication design challenging and, likely, ad hoc in most cases [10]. Therefore, the sharper change of authenticity based on the proposed measure around threshold may lead to a sharper boundary between the surely authentic and unauthentic images, which is desirable for image authentication.

Fig. 8 shows the comparison results of different distance measures in terms of their discernable abilities. In Fig. 8(a), the last three columns are images maliciously tampered from the original portrait image *Lena*, by enlarging the eyes, modifying multiple objects in the image, and adding a small spot on the face. The others are images from acceptable manipulations including Gaussian noise introduction, auto contrast adjustment, sharpening, and lossy transmission (with error concealment). Fig. 8(b) shows results of image *Bike* with much stronger edges than image *Lena*. The last three columns of Fig. 8(b) are images tampered by deleting the saddle, modifying multiple objects (changing logo at the left top, modifying the display of the clock at right top, and deleting the saddle), and adding a small spot in the center of the right circle. Note that the *SSMs* were all below 0.5 for acceptable manipulations and all above 0.5 for maliciously attacked images. On the contrary, the Hamming and Figure of Merit (*FoM*) measures of maliciously attacked images were among the range of acceptable manipulations especially the measures of the attacked image in which there was a small local object changed (last column). The results show that the proposed *SSM* was able to distinguish the malicious manipulations from acceptable ones, i.e., identify lossy transmission as acceptable, and was sensitive to malicious manipulations. On the contrary, the Hamming and *FoM* measures were not sensitive to small localized object modification. The results indicate that the proposed *SSM* is more suitable for content-based image authentication than Hamming and *FoM* measures.

### 4.2 *SSM*-based Error Resilient Image Authentication Scheme Evaluation

**Robustness to Transmission Errors and other Acceptable Manipulations.** The transmission errors in wireless networks were simulated based on the *Rayleigh* model [20] which is commonly used for wireless networks. Fig. 9(b) is an example of wavelet-based images damaged by transmission errors, and Fig. 9(c) is its error-concealed result. Fig. 9(d) is a DCT-based image damaged by transmission errors, and Fig. 9(e) is its error concealed result. The *SSM* values of image Fig. 9(c) and Fig. 9(e) are 0.134 and 0.250, i.e., the error-concealed images are both authentic.

With the set of images produced, an average peak signal-to-noise ratio (defined by *PSNR*) of 44.46 dB (Table 1) was obtained which is above the usually tolerated degradation level of 40 dB [23] and much better than the average 33.45dB in [13]. It is also better than the 42.47 dB obtained by the paper [23]. The quantization table used in these experiments is JPEG recommended quantization table of Q50. These results indicate the embedding procedure did not introduce visual artifacts in the images.

Table 2 shows the evaluation results of the system robustness of the proposed error resilient image authentication scheme based on the proposed *SSM*. *PSNR* and *SSM* measures of the images damaged by transmission errors with different bit error rate (BER) $10^{-4}$ and $2\times10^{-4}$. The corresponding *PSNR* and *SSM* of the error-concealed images are also listed in this table. 60% of the damaged images at BER $10^{-4}$ and 100% at BER $2\times10^{-4}$ in our experiments were verified as unauthentic. On the contrary, all error-concealed images were verified as authentic. These results indicate that our proposed scheme could obtain a good robustness to transmission errors. Note that on the contrary, the authentication scheme [23] was not robust to transmission errors. These results further confirm that it is effective and advisable for error concealment to be applied before image authentication. The reason that the authenticities of the recovered images were better than those of the damaged images may be the image quality improvement by using error concealment on the damaged images [18], [20]. For example, the recovered image had much better objective qualities than the damaged images (evaluated by *PSNR*). This quality improvement made features of the error-concealed images closer to those of the original images than damaged images, so that the image authenticities (evaluated by *SSM*) of the error-concealed images were much larger than the damaged images.

Our scheme was also tested with other acceptable manipulations such as image contrast adjustment, histogram equalization, compression and noises addition. The results are shown in Table 3, with the parameter for each manipulation. The *SSM* values of these images were all less than 0.5, i.e., all these images can pass the authentication. These results validate that the proposed scheme is not only designed to be robust to transmission errors, but also robust to general acceptable manipulations.

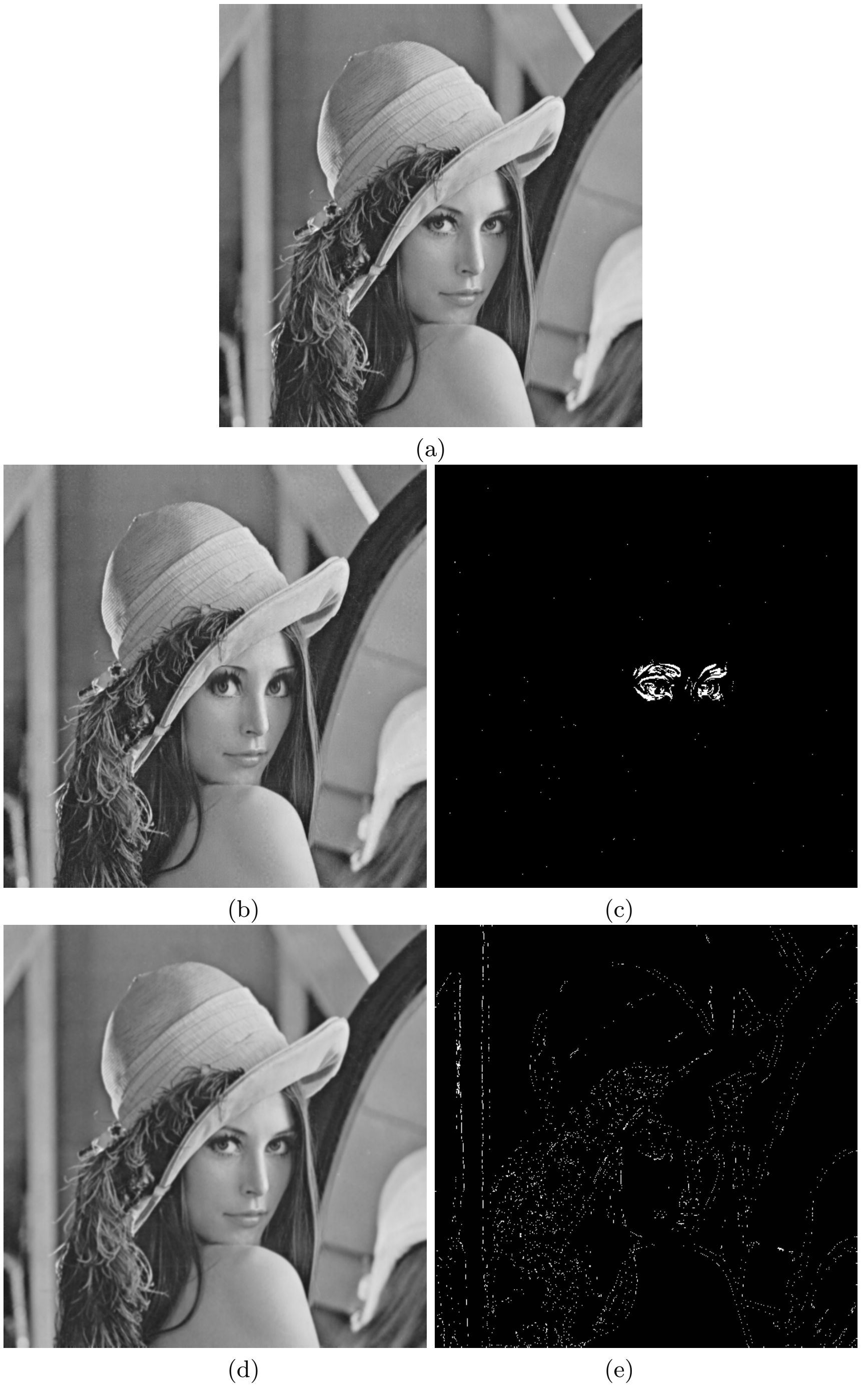
**Sensitivity to Malicious Content Tampering.** An important aspect of our *SSM*-based authentication scheme is that it is sensitive to the malicious content tampering. For that reason, we tampered the previous watermarked *Bike* and *Lena* images and tested the ability of our system to detect and highlight the attacked areas. All the attacked images were detected and possible attacked areas were located. The attack location results are shown in Fig. 10.

These results indicate that the ability of our system to detect tampering is good even in the presence of multiple tampered areas (Fig. 10(e)), or noises (Fig. 10(a)), or very small area modified (Fig. 10(c)). Furthermore, the attack detection result of our scheme is friendly to users. If the image fails to pass the authentication, our scheme provides detected attacked areas which concentrate on objects. Further authentication decision can be made by the user with the aid of attack detection results.

## 5 Conclusions

A new feature distance measure based on statistical and spatial properties of the feature differences for content-based image authentication is proposed in this paper. The use of the typical patterns of feature differences by acceptable image manipulations and malicious content modifications did help improve system performance. Many acceptable manipulations which were detected as malicious modifications in the previous schemes were correctly classified into authentic images in the scheme based on *SSM*. The results also indicate that the statistical and spatial properties of the image feature are helpful and useful in distinguishing acceptable image manipulations from malicious content modifications. Moreover, the results would lead to a better understanding of the role of statistics and spatial properties of feature differences for detecting digital forgeries. The scheme was further evaluated under transmission errors.

The proposed feature distance measure is quite general and can be used in many other content-based authentication schemes provided that the features contain spatial information.



(a)

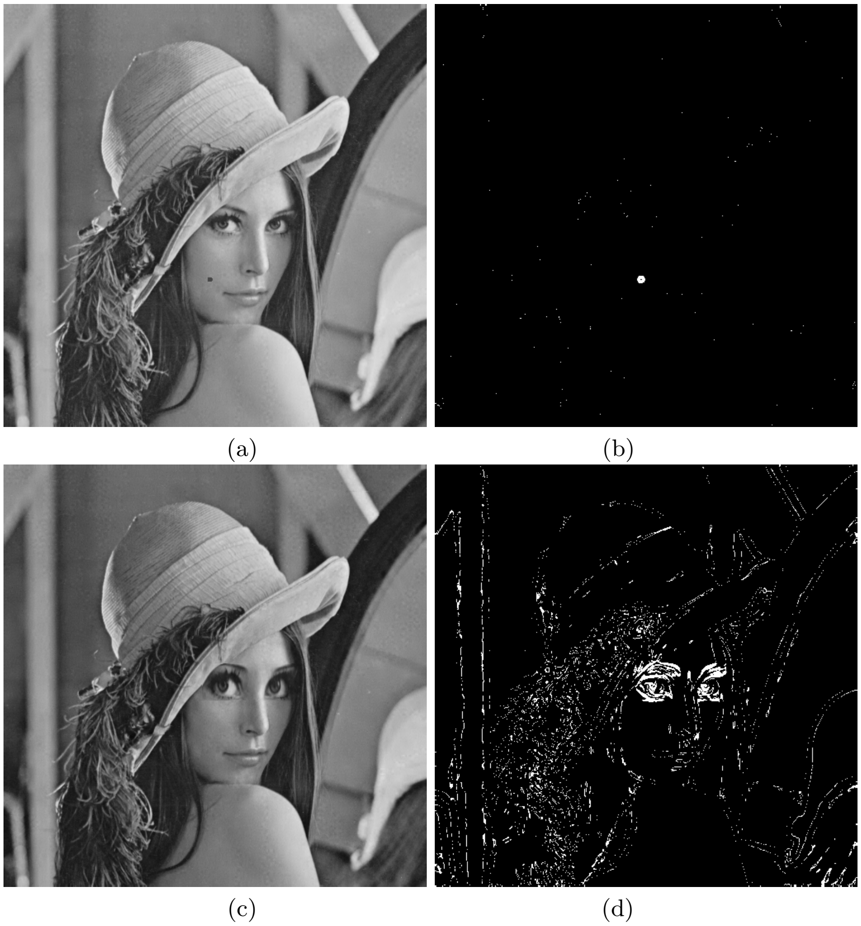
(b)

(c)

(d)

(e)

**Fig. 1**. Discernable patterns of edge feature differences caused by acceptable image manipulation and malicious modification: (a) original image; (b) tampered image; (c) feature difference of (b); (d) blurred image (by Gaussian 3×3 filter); (e) feature difference of (d)



(a) (b)

(c) (d)

**Fig. 2**. Cases that required both *mccs* and *kurt* to work together to successfully detect malicious modifications: (a) small object tampered (*kurt*: large; *mccs*: small); (b) feature differences of (a); (c) large object tampered with global distortions (*kurt*: small; *mccs*: large); (d) feature differences of (c)

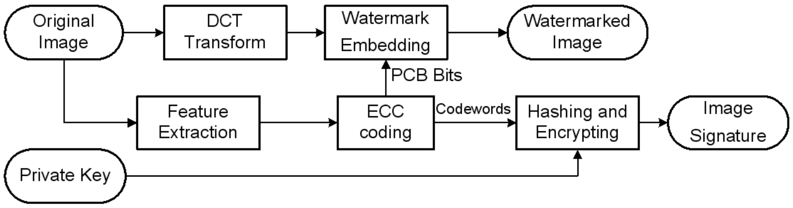


**Fig. 3**. Signing process of the proposed error resilient image authentication scheme

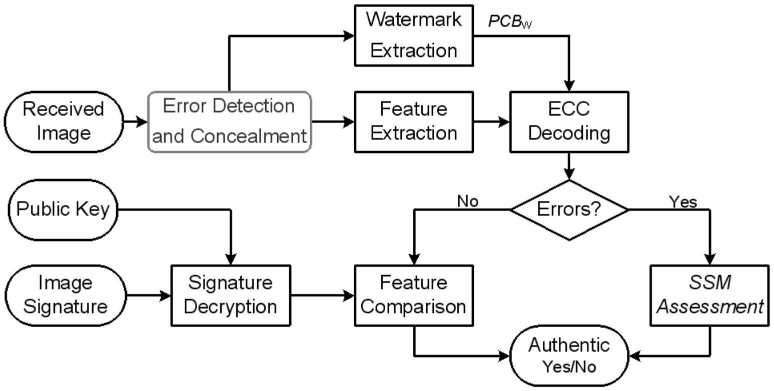


**Fig. 4**. Image authentication process of the proposed error resilient image authentication scheme

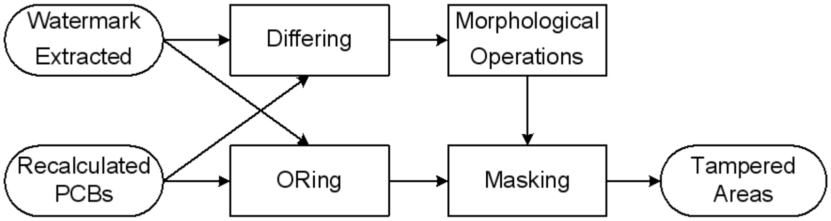
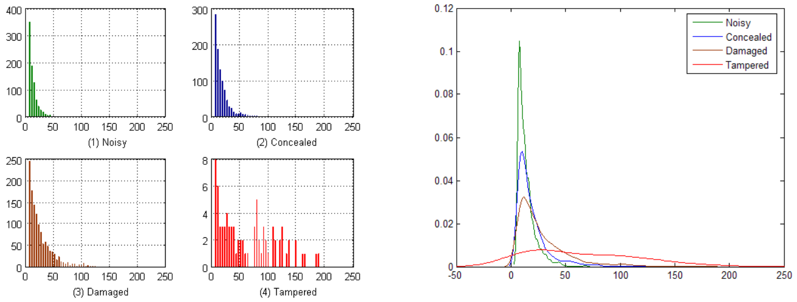
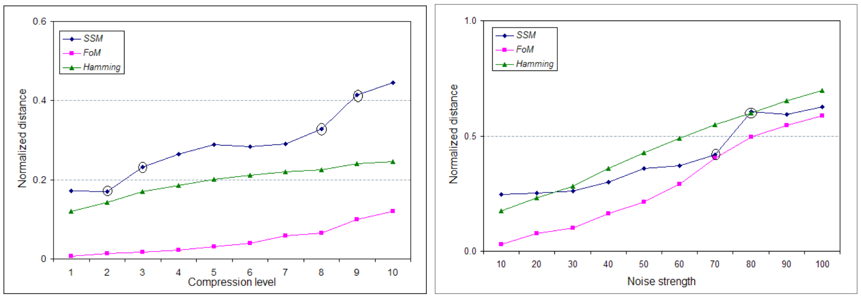


**Fig. 5**. Feature aided attack location process



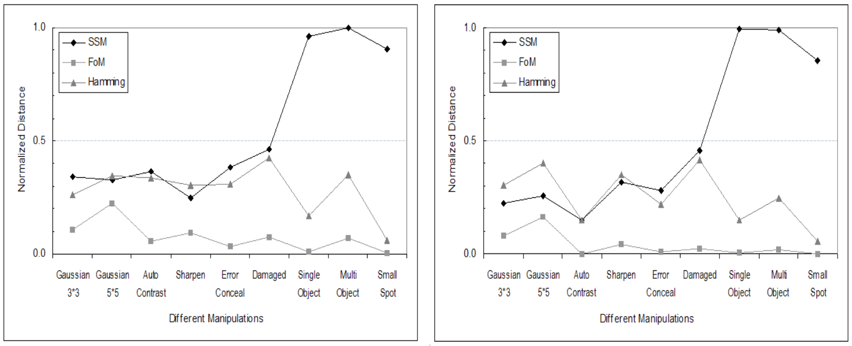
(a) (b)

**Fig. 6**. Different patterns of edge difference distribution: (a) histograms of edge differences; (b) probability density estimation



(a) (b)

**Fig. 7**. Distance measures comparison: (a) for JPEG compressions (b) for *Gaussian* noises

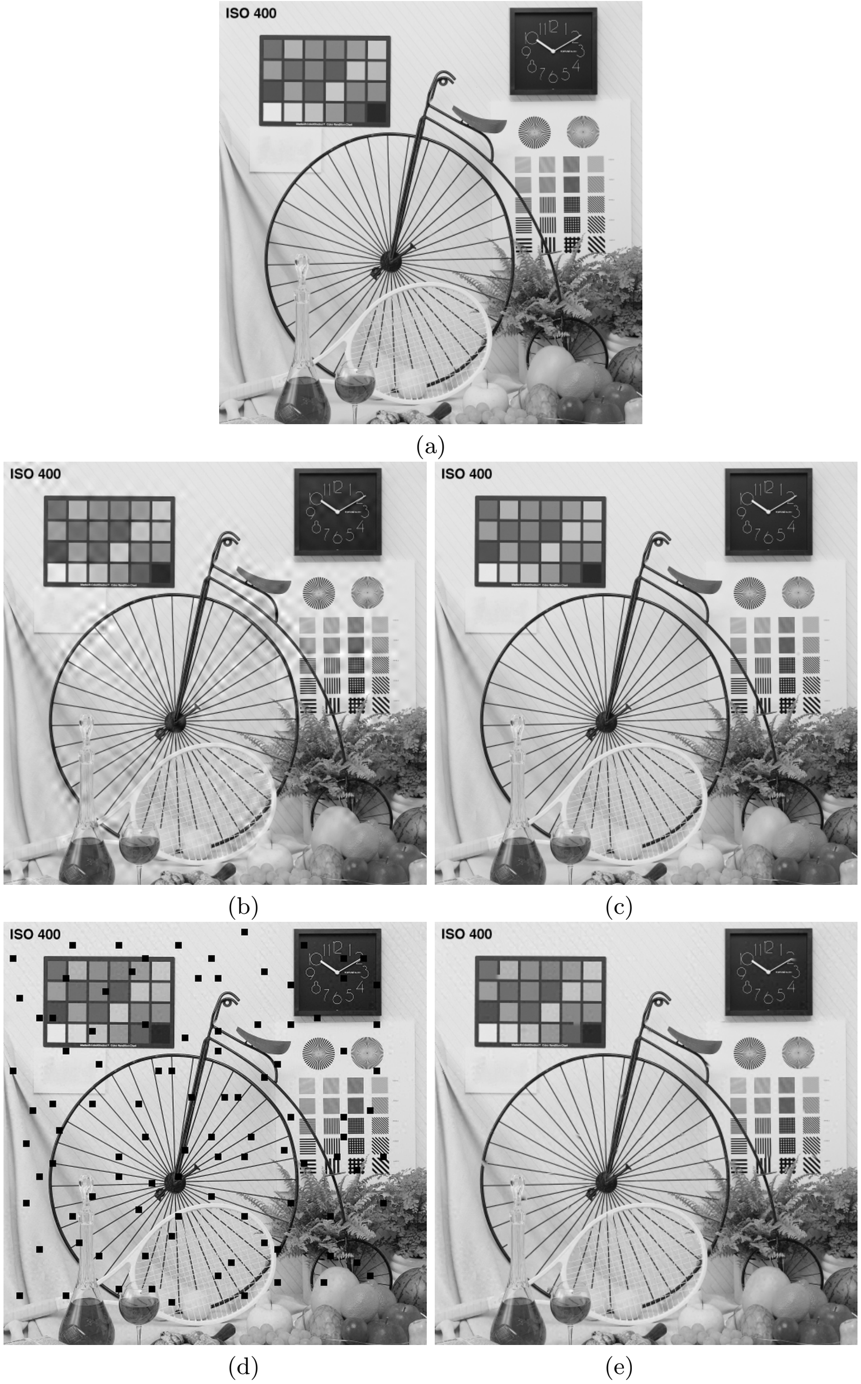


(a) (b)

**Fig. 8**. Comparison of distinguish ability of different distance measures: only the proposed measure can successfully distinguish malicious manipulations from acceptable ones: (a) Results of image *Lena*; (b) Results of image *Bike*

**Table 1**. Comparison of objective quality decrease introduced by watermarking: *PSNR*(dB) of watermarked images

| *PSNR* | *Barbara* | *Bike* | *Airplane* | *Girl* | *Goldhill* | *Lena* | *Mandrill* | *Monarch* | *Pepper* | *Woman* |
|---|---|---|---|---|---|---|---|---|---|---|
| Proposed | 44.17 | 44.40 | 44.56 | 44.39 | 44.32 | 44.60 | 44.14 | 44.75 | 44.46 | 44.79 |
| Ref. [13] | 32.90 | 29.91 | 32.01 | 34.20 | 34.07 | 36.11 | 32.38 | 30.43 | 35.53 | 36.98 |
| Ref. [23] | 42.72 | / | 43.15 | / | / | / | / | / | / | / |



(a)

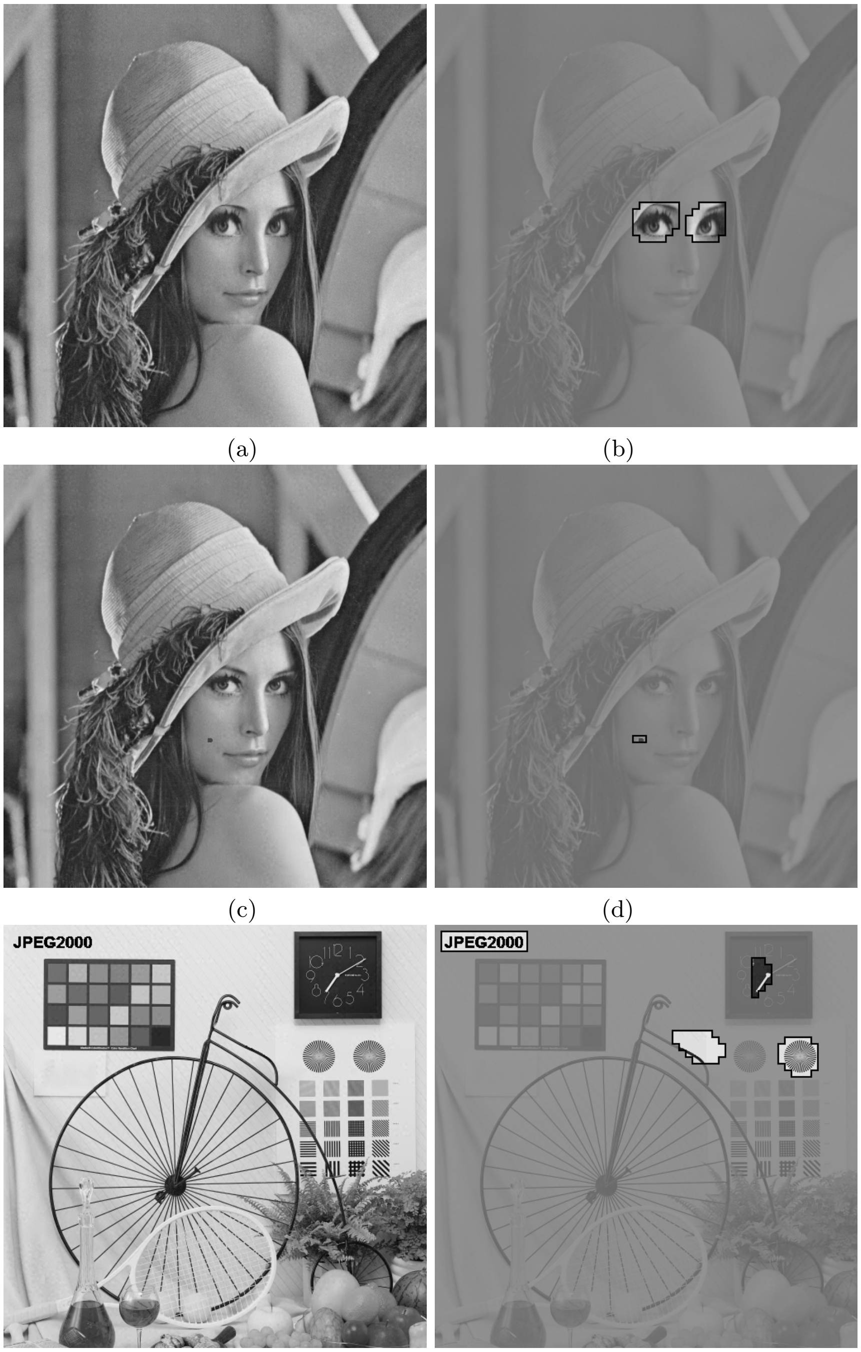
(b)

(c)

(d)

(e)

**Fig. 9**. Robustness against transmission errors: (a) original image (b) damaged image (wavelet based); (c) error concealed result of (b); (d) damaged image (DCT based); (e) error concealed result of (c)



(a) (b) (c) (d) (e) (f)

**Fig. 10**. Detected possible attack locations which are concentrated on objects in images: (a) noisy tampered image *Lena* (0.995); (b) attacked areas detected of (a); (c) *Lena* with small spot added (0.569); (d) attacked areas detected of (c); (e) attacked image *Bike* (logo modified, time modified, saddle deleted, and circle copied/pasted) (0.995); (f) attacked areas detected of (e)

**Table 2**. Authentication performance improvement by error concealment: *PSNR* (dB) and *SSM* of damaged images and error-concealed images (BER1:$10^{-4}$; BER2:$2\times10^{-4}$)

| Images | | *Actor* | *Bike* | *Chart* | *Flight* | *Fruits* | *Hotel* | *Lake* | *Lena* | *Pepper* | *Woman* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Damaged *PSNR* | BER1 | 30.78 | 31.26 | 33.95 | 32.41 | 33.68 | 33.87 | 31.39 | 33.31 | 33.07 | 35.50 |
| | BER2 | 25.87 | 25.76 | 28.51 | 26.05 | 27.81 | 26.71 | 25.68 | 30.34 | 27.74 | 30.72 |
| Damaged *SSM* | BER1 | 0.948 | 0.939 | 0.707 | 0.297 | 0.794 | 0.365 | 0.143 | 0.391 | 0.729 | 0.989 |
| | BER2 | 0.812 | 0.999 | 0.987 | 0.951 | 0.942 | 0.568 | 0.883 | 0.638 | 0.865 | 0.955 |
| Recovered *PSNR* | BER1 | 38.03 | 41.76 | 41.11 | 41.03 | 39.90 | 42.40 | 38.54 | 40.21 | 41.25 | 42.96 |
| | BER2 | 32.06 | 34.99 | 34.74 | 34.06 | 31.68 | 33.26 | 31.64 | 36.03 | 33.85 | 36.84 |
| Recovered *SSM* | BER1 | 0.158 | 0.134 | 0.141 | 0.035 | 0.204 | 0.067 | 0.057 | 0.345 | 0.089 | 0.329 |
| | BER2 | 0.220 | 0.099 | 0.446 | 0.072 | 0.406 | 0.045 | 0.280 | 0.059 | 0.182 | 0.015 |

**Table 3**. Robustness against acceptable image manipulations

| Manipulations | Histogram Normalizing | Brightness Adjustment | Contrast Adjustment | JPEG Compression | JPEG2000 Compression |
|---|---|---|---|---|---|
| Parameter | Auto | -40 | Auto | 10:1 | 1bpp |
| *SSM* | 0.159 | 0.159 | 0.262 | 0.017 | 0.057 |

## References

1. C. Y. Lin and S.F. Chang, "A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation", *IEEE Transaction on Circuits and Systems of Video Technology*, Vol.11, pp. 153-168, 2001.
2. E. Martinian, G. W. Wornell, and B Chen, "Authentication With Distortion Criteria", *IEEE Transaction on Information Theory*,Vol.51, No. 7, pp. 2523-2542, July 2005.
3. J. Dittmann, A. Steinmetz, and R. Steinmetz, "Content-based Digital Signature for Motion Pictures Authentication and Content-fragile Watermarking", *IEEE International Conference on Multimedia Computing and Systems*, Vol.2, pp.209-213, 1999.
4. B.B. Zhu, M.D. Swanson, and A.H. Tewfik, "When Seeing Isn't Believing: Multimedia Authentication Technologies", *IEEE Signal Processing Magazine*, Vol. 21, No. 2, pp.40-49, Mar 2004.
5. Y. Wang, J. Ostermann, and Y.Q. Zhang, "Video Processing and Communications", *Prentice Hall*, 2002.
6. M.P. Queluz, "Authentication of Digital Images and Video: Generic Models and a New Contribution", *Signal Processing: Image Communication*, Vol.16, pp. 461-475, January 2001.
7. E.C. Chang, M.S. Kankanhalli, X. Guan, Z.Y. Huang, and Y.H. Wu, "Robust Image Authentication Using Content-based Compression", *ACM Multimedia Systems Journal*, Vol. 9, No. 2, pp. 121-130, 2003.
8. Q. Sun and S.F. Chang, "Semi-fragile Image Authentication using Generic Wavelet Domain Features and ECC", *IEEE International Conference on Image Processing (ICIP)*, Rochester, USA, Sep. 2002.
9. S. Ye, Q. Sun, and E.C. Chang, "Error Resilient Content-based Image Authentication Over Wireless Channel", IEEE International Symposium on Circuits and Systems, Japan, 2005.

10. C. W. Wu, "On the Design of Content-Based Multimedia Authentication Systems", *IEEE Transactions on Multimedia*, Vol. 4, No. 3, pp.385-393, September 2002.
11. M. Schneider and S.F. Chang, "A Robust Content-based Digital Signature for Image Authentication", in *Proceedings of International Conference on Image Processing (ICIP)*, Vol.3, pp.227 - 230, 1996.
12. B. Li, E. Chang, and Y. Wu, "Discovery of a Perceptual Distance Function for Measuring Image Similarity", *ACM Multimedia Journal Special Issue on Content-based Image Retrieval*, Vol. 8, No. 6, pp.512-522, 2003.
13. Q. Sun, S. Ye, L.Q. Lin, and S.F. Chang, "A Crypto Signature Scheme for Image Authentication over Wireless Channel", *International Journal of Image and Graphics*, Vol. 5, No. 1, pp.1-14, 2005.
14. J. L. Cannons and P. Moulin, "Design and Statistical Analysis of a Hash-Aided Image Watermarking System", *IEEE Transaction on Image Processing*, Vol. 13, No. 10, October 2004, pp. 1393-1408.
15. R. Jain, R. Kasturi and B. G. Schunck, "Machine Vision", *McGraw Hill*, New York, 1995.
16. V. Erceg and K. Hari, "Channel Models for Fixed Wireless Applications", *IEEE 802.16 Broadband Wireless Access Working Group*, 2001.
17. P. Golle and N. Modadugu, "Authenticating Streamed Data in the Presence of Random Packet Loss", In *Proceedings of the Symposium on Network and Distributed Systems Security*, pp.13-22, 2001.
18. S. Ye, Q. Sun, and E.C. Chang, "Edge Directed Filter based Error Concealment for Wavelet-based Images", *IEEE International Conference on Image Processing*, Singapore, 2004.
19. W. Chou, "Classifying Image Pixels into Shaped, Smooth and Textured Points", *Pattern Recognition*, Vol. 32, No. 10, pp.1697-1706, 1999.
20. S. Ye, X. Lin and Q. Sun, "Content Based Error Detection and Concealment for Image Transmission over Wireless Channel", *IEEE International Symposium on Circuits and Systems* (*ISCAS*), Thailand, May 2003.
21. M. Boliek (ed.), "JPEG 2000 Final Committee Draft", *ISO/IEC FCD*1.5444-1, Mar. 2000.
22. Y. Yu and S. T. Acton, "Speckle Reducing Anisotropic Diffusion", *IEEE Transaction on Image Processing*, Vol. 11, No. 11, Nov. 2002, pp.1260-1270.
23. A.H. Paquet, R.K. Ward, and I. Pitas, "Wavelet Packets-based Digital Watermarking for Image Verification and Authentication", *Signal Processing*, Special Issue on Security of Data Hiding Technologies, Vol. 83, No. 10, pp.2117-2132, October 2003.