Blind forensics in medical imaging based on Tchebichef image moments
Hui Huang, Gouenou Coatrieux, Huazhong Shu, Limin Luo, Christian Roux

To cite this version:

HAL Id: hal-00737203
https://hal.science/hal-00737203
Submitted on 19 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Blind Forensics in Medical Imaging Based on Tchebichef Image Moments

H. Huang, G. Coatrieux, H.Z. Shu, L.M. Luo, Ch. Roux

Abstract—In this paper, we present a blind forensic approach for the detection of global image modifications like filtering, lossy compression, scaling and so on. It is based on a new set of image features we proposed, called Histogram statistics of Reorganized Block-based Tchebichef moments (HRBT) features, and which are used as input of a set of classifiers we learned to discriminate tampered images from original ones. In this article, we compare the performances of our features with others proposed schemes from the literature in application to different medical image modalities (MRI, X-Ray ...). Experimental results show that our HRBT features perform well and in some cases better than other features.

I. INTRODUCTION

The sophisticated and low cost tools of the digital age facilitates creation and handling of digital images. In healthcare, image manipulations encompass global image processing like contrast enhancement, brightness adjustment, lossy compression (e.g. JPEG)… or any combinations of them. They are used by the physician during his or her interpretation of the image and also in order to facilitate image sharing like in telemedicine applications. Because these processes may induce loss of important information that supports the diagnosis, one can no longer take the authenticity of images for granted. It is thus important being capable to detect such image modifications especially to inform the practitioner the image he or she observes has been processed, and also when the image comes to medico-legal evidence.

Image forensics techniques and methodologies for validating the authenticity of an image have recently attracted attention. They can, in principle, reconstitute the set of processing operations the image undergone. They allow us to make statements about the image veracity and give clues about the nature of the performed manipulations. Forensic methods can be distinguished in two classes according to the access to a priori knowledge on the original image. For example, some solutions store or watermark within the image some digests or signatures which by comparison with the recomputed ones states about the image integrity [1]. Even though watermarking is an effective tool for verifying image integrity and origins [2], its application requires that a watermark is embedded during image creation, limiting its use to applications where the digital object generation mechanisms have built-in watermarking capabilities. Majority of images captured today are not watermarked. Therefore, there is a strong need to resort to blind image forensic techniques. These techniques work without a priori information and aim at identifying characteristics or evidences left by most image modifications.

A large group of “blind forensic” methods are based on classifier mechanisms. Classifiers are built on some image features to recognize or identify image modification footprints. Avcibas et al. [3] developed a detection scheme that discriminates “doctored” images from original ones (i.e. not tampered) based on the training of a classifier with Image Quality Metrics (IQM). Farid et al. [4] developed a detection scheme trained to recognize statistical footprints of image modifications within the High Order Wavelet statistics (HOWs). Similarly, Bayram et al. [5] proposed to make use of Binary Similarity Measures (BSM), derived from the correlation between the image bit planes as well as binary texture characteristics within one bit plane. It is important to notice that these methods use image features originally proposed for steganalysis whose purpose is to detect stego-images (i.e. image with secret message embedded). In fact, steganography acts similarly to image modification. It affects more or less the image content while not carrying out perceptible distortions into the image [5]. Recently, for the steganalysis purpose, Liu et al. [6] proposed a set of image features designed with Histogram statistical properties of the Reorganized Block-based Discrete cosine transform coefficients (HRBD), that outperforms HOWs features. Notice that HRBD has not been yet experimented for image authenticity purpose. This participates also to the originality of this work. It is expected that the features which perform well for steganalysis will also work better than other image features for detecting image modification.

The objective we pursue in this work is to detect if an image undergone or not some “global” image modification like lossy JPEG compression, filtering and so on. For that purpose, we introduce a new set of image features based on Tchebichef moments, features extracted from the image following the same strategy as Liu et al. [6]. DCT coefficients and Tchebichef moments have similar properties. However, Tchebichef moments have shown better performances than DCT coefficients in some image processing applications [7]. Thus similarly to HRBD, we will talk about HRBT. We use both set of features to learn classifiers capable to discriminate original images from modified ones. In this works, we...

The rest of this paper is organized as follows. Section 2 remembers the basic principles of blind forensic mechanisms for digital image. Section 3 details the proposed method while section 4 provides some experimental results. Conclusions are given in section 5.

II. BLIND FORENSICS FOR IMAGE

A. Basic Principles

Whatever its extent, any image processing will leave evidences of its manipulation like some image features statistical irregularities. One approach proposed to catch up these evidences is based on the learning of classifiers whose input are image features impacted by image modification. Once the classifier learned, one just has to extract these features from the image under investigation, and to provide them to the classifier for analysis and decision. Efficiency of such an approach largely depends on: i) the design of proper image features and, ii) the way the classifier is built.

In the following, in order to only evaluate features’ efficiency, we used Support Vector Machines [8] for the design of classifiers. SVM is one of the most popular supervised classification methods due to its superior classification performances in many applications.

B. Image Features for Blind Forensics

Features we are looking at should reflect changes in an image while being content independent. Image modifications we wish to detect don’t impact image main structures but rather image details. In this section, we present some common features proposed for modification detection.

1) Image Quality Metrics (IQM)

In [3], Avcıbas et al. proposed to compute IQMs on the signal of difference between the image under investigation and a reference image. In fact, they show that it is possible to find distance measurements between images that can isolate signal changes regardless of image content. In particular, they demonstrate that the first two statistical moments of the angular correlation and of Czekanowski similarity measures satisfy such independency constraint. The reader may refer to [3] for more details. Nevertheless, because in the following experiments, we only work with gray scale images of N × N pixels, the angular correlation Avcıbas et al. measure between color components is constant and cannot be exploited. Thus IQMs we consider in our experiments are the two first order statistics of Czekanowski similarity measure. Considering an image under investigation $f(i,j)$ and a reference image $g(i,j)$, these features can be defined as:

$$d_1 = \frac{1}{N^2} \sum_{i,j=0}^{N-1} (\chi_{ij} - \mu_{\chi})^2$$

$$d_2 = \frac{1}{N^2} \sum_{i,j=0}^{N-1} (v_{ij} - \mu_{v})^2$$

where $\chi_{ij} = \frac{f(i,j)g(i,j)}{f(i,j)g(i,j) + g(i,j)}$, $\mu_{\chi} = \frac{1}{N^2} \sum_{i,j=0}^{N-1} |\chi_{ij}|$, and $v_{ij} = \frac{f(i,j)g(i,j)}{f(i,j)g(i,j) + g(i,j)}$, $\mu_{v} = \frac{1}{N^2} \sum_{i,j=0}^{N-1} |v_{ij}|$.

2) High Order of Wavelet Statistic (HOWs)

Farid et al. benefit from the wavelet decomposition to isolate image details [4]. They propose to discriminate modifications through their impact on the statistical distribution of wavelet coefficients and of their prediction error within wavelet sub-bands. If $V_p(i,j)$, $H_p(i,j)$ and $D_p(i,j)$ denote the coefficients in the vertical, horizontal and diagonal sub-bands at position $(i,j)$ of resolution $p$ respectively, then the estimator of the coefficient $V_p(i,j)$ is:

$$[V_p(i,j) = \omega_1 |V_p(i-1,j)| + \omega_2 |V_p(i+1,j)| + \omega_3 |V_p(i,j-1)| + \omega_4 |V_p(i,j+1)| + \omega_5 |V_p(i-1,j-1)| + \omega_6 |V_p(i-1,j+1)| + \omega_7 |V_p(i+1,j-1)| + \omega_8 |V_p(i+1,j+1)|]$$

where weight $\omega_i \in \mathbb{R}$. Statistical measurements – the mean, the variance, the skewness and the kurtosis – on wavelet coefficients and on their error of prediction in each sub-band constitute the HOWs features. Readers can refer to [4] for more details on the computation of weight ($\omega_i$) and features.

3) Binary Similarity Measures (BSM)

These measures have been proposed considering the fact that image modifications may induce correlation variations between and within bit planes. These changes can be pointed out by statistical features extracted from image bit planes. The measures are based on the comparison of binary texture statistics between bit plane pairs of the images (e.g. 5-6, 6-7, 7-8 bit planes). Due to the limited size of the paper, we cannot present details of BSM features. The reader may refer to [5] for their description.

4) Histogram Statistics of Reorganized Block-based DCT Coefficients (HRBD)

These features have been proposed for the steganalysis purpose but not yet experimented for image integrity verification. Harmsen et al. [9] and Shi et al. [10] show up that the center of mass and the higher order statistic moments of the Discrete Fourier Transform of the DCT coefficients Histogram (DFTH) vary after information embedding. Different set of such features have been proposed, Shi et al. [10] make use of the HDFT defined as:

$$M_n = \frac{\sum_{k=0}^{N/2} k^n |H(k)|}{\sum_{k=0}^{N/2} |H(k)|}$$

where $H(k)$, $|H(k)|$ denote the DFT at the frequency $k$ and its magnitude respectively. Wang et al. [11] improves steganalysis performances working on filtered version of the DFT extracting three distinct features $f_i$, $i = 1 \ldots 3$

$$f_1 = \frac{\sum_{k=0}^{N/2} |H(k)| \cdot \sin (\pi k / K)}{\sum_{k=0}^{N/2} |H(k)|}$$

$$f_2 = \frac{\sum_{k=0}^{N/2} |H(k)| \cdot \sin^2 (\pi k / K)}{\sum_{k=0}^{N/2} |H(k)|}$$

$$f_3 = \frac{\sum_{k=0}^{N/2} |H(k)| \cdot \sin (\pi k / K)}{\sum_{k=0}^{N/2} |H(k)|}$$

Liu et al. suggest extracting these both set of features from the image and from its difference with its predicted version. They compute them on the DFTH of DCT coefficients regrouped or reorganized in independent “sub-bands” of an L-scale wavelet-like tree [6]. Because we follow the same
reorganization procedure with Tchebichef moments, we
detail it in section III. Anyway, the prediction algorithm used
in the sequel is expressed as [12]:

\[
e(x, y) = \begin{cases} 
\min(a, b) & c \leq \min(a, b) \\
\max(a, b) & c \geq \max(a, b) \\
\end{cases} 
\]

(8)

where \( e \) is the prediction error of the pixel \( x \); \( \hat{x} \) is the predicted
value of \( x \); \( a, b, \) and \( c \) are the pixels surrounding \( x \) (see Fig. 1).

\[
\begin{array}{ccc}
  & x & \\
  a & & c \\
\end{array}
\]

Fig. 1. The neighborhood of the prediction

III. PROPOSED IMAGE MOMENT BASED FEATURE

A. Image moment

Introduced by Hu [13], one general image moment \( \Psi_{nm} \)
is defined with a basis function \( \psi_{nm}(x, y) \), and an image
intensity function \( f(x, y) \) such as:

\[
\Psi_{nm} = \iint \psi_{nm}(x, y)f(x, y)dxdy, n, m = 0, 1, 2, \ldots
\]

(9)

Different types of moments can be defined depending on \( \psi_{nm} \). Among these moments, Tchebichef moments are the
simplest discrete orthogonal moments, widely used in digital
image processing such as image reconstruction and pattern
recognition.

The \((n + m)^{th}\) order Discrete Tchebichef moment \( T_{nm} \) is
defined as [7]:

\[
T_{nm} = \frac{1}{\rho(n, N)\rho(m, N)}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \psi_{nm}(x, y)f(x, y)
\]

(10)

\[
t_n(x) = n! \sum_{k=0}^{n} (-1)^{n+k} \binom{n}{k} X^k
\]

(11)

\[
\rho(n, N) = (2n)! \binom{N+n}{2n+1}
\]

(12)

It is possible to derive the DCT basis from the discrete
Tchebichef moments [14]. Some experimental results given in [7]
illustrate a close similarity between performances of DCT and Tchebichef moments for lossy image compression and
image reconstruction. Sometimes Tchebichef moments provide
better reconstruction performances. This motivates
us to consider Tchebichef moments as image features.

B. Proposed features

Considering the intrinsic properties of Tchebichef
moments, we adapted the strategy of Liu et al. [6] to extract
our set of features.

In a first time, the image of size \( N \times N \) is divided into small
blocks of \( n \times n \) pixels - in the sequel \( n = 8 \) - and for each,
Tchebichef moments up to \( ((n-1) + (n-1))^{th}\) order
\( T_{pq}, p = 0, \ldots, n-1; q = 0, \ldots, n-1 \) are computed leading
to \( n \times n \) moments values. In Fig. 2, \( T_{pq} \) is at the coordinate
position \( (p+1, q+1) \).

In a second time, each block of \( n \times n \) moments is partitioned
into an \( L \)-scale wavelet-like tree (see Fig. 2a), within a
\( 3L + 1 \) subbands decomposition, where \( n = 2^L \).

At last, moments belonging to the same sub-bands in each
block are clustered forming an \( L \)-scale coefficients tree for the
whole image. As shown in Fig. 2b, the “sub-band” \( G_0 \)
regroups \( T_{00} \) moments of all the blocks in the image,
moments that are placed at the same relative positions in the
image.

\[
\begin{array}{c}
G_0 \quad G_1 \quad G_2 \quad \ldots \quad G_{3L-1} \\
\end{array}
\]

Fig. 2. Reorganization of block-based Tchebichef moment for an image of
\( N \times N \) pixels. (a) 8x8 pixel block partitioned as the 3-scale wavelet-like tree.
(b) sub-band construction based on 8x8 block to form a L-scale wavelet like
tree of global dimension \( N \times N \).

Our HRBT features are then computed within
\( \{G_0, G_1, \ldots, G_{3L}\} \) sub-bands of the test image and its
discriminant image with equations (4) – (7).

IV. EXPERIMENTAL RESULTS

A. Image test sets

We have used four test sets of images issued from different
medical image modalities (see Fig. 3):
- Magnetic resonance images (MRI) of the head: 145
  images of 256 \times 256 pixels and 12-bit depth.
- Two kinds of X-Ray images: 162 mammograms of
  4740x3540 pixels encoded on 12 bits and 135 CT images
  of 512x512 pixels of 12-bit depth.
- Ultrasound images (Echo): 52 images of 576 \times 690 pixels
  and of 8-bit depth.

\[
\begin{array}{cccc}
(a) & (b) & (c) & (d) \\
\end{array}
\]

Fig. 3. Image samples of our test sets: (a) MRI (b) CT image (c) X-Ray image
(d) Echography

The modifications we have considered in the experiments
are: contrast and brightness adjustment, Gaussian filtering,
rotation, compression and histogram equalization. Table I
gives the parameters used for each of these modifications.

<table>
<thead>
<tr>
<th>Modification</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imaging</td>
<td></td>
</tr>
<tr>
<td>MRI</td>
<td>(a) 8x8, (b) 4x4, (c) 16x16, (d) 32x32</td>
</tr>
<tr>
<td>CT</td>
<td>(a) 8x8, (b) 4x4, (c) 16x16, (d) 32x32</td>
</tr>
<tr>
<td>X-Ray</td>
<td>(a) 8x8, (b) 4x4, (c) 16x16, (d) 32x32</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>(a) 8x8, (b) 4x4, (c) 16x16, (d) 32x32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Filtering</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI</td>
</tr>
<tr>
<td>CT</td>
</tr>
<tr>
<td>X-Ray</td>
</tr>
<tr>
<td>Ultrasound</td>
</tr>
</tbody>
</table>

Our HRBT features are then used to classify these images.

B. Performances Evaluation

In this work our goal is only to distinguish original images
from modified ones. To achieve this goal, we have learned
different SVM based detectors (e.g. original (org.) vs. JPEG,
org. vs. filtering …) using our image features. We also built
SVM classifiers with the features proposed in [3]-[6]. With
this strategy, an image is declared unauthentic, if at least one
of our classifiers notifies it.
TABLE II DETECTION RATE ACHIEVED BY DIFFERENT SVM CLASSIFIERS WITH DIFFERENT SETS OF IMAGE FEATURES (IQM, HOWS, BSM, HRBD, HRBT)

<table>
<thead>
<tr>
<th></th>
<th>IQM</th>
<th>HOWs</th>
<th>BSM</th>
<th>HRBD</th>
<th>HRBT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI</td>
<td>79.46</td>
<td>81.02</td>
<td>81.26</td>
<td>82.53</td>
<td>83.14</td>
</tr>
<tr>
<td>Brighten adj. (IQM)</td>
<td>99.70</td>
<td>97.07</td>
<td>86.09</td>
<td>96.09</td>
<td>97.93</td>
</tr>
<tr>
<td>hist. equ.</td>
<td>54.31</td>
<td>44.79</td>
<td>73.33</td>
<td>63.12</td>
<td>61.11</td>
</tr>
<tr>
<td>Deviation of Gauss.</td>
<td>99.22</td>
<td>99.91</td>
<td>93.81</td>
<td>99.54</td>
<td>99.48</td>
</tr>
<tr>
<td>Scaling</td>
<td>52.22</td>
<td>42.53</td>
<td>52.72</td>
<td>63.78</td>
<td>62.78</td>
</tr>
<tr>
<td>rotation (IQM)</td>
<td>47.96</td>
<td>33.58</td>
<td>53.77</td>
<td>73.90</td>
<td>79.90</td>
</tr>
<tr>
<td>Brighten rate (IQM)</td>
<td>89.97</td>
<td>74.88</td>
<td>74.26</td>
<td>97.90</td>
<td>97.90</td>
</tr>
<tr>
<td>Brighten adj. (HRBD)</td>
<td>100</td>
<td>74.94</td>
<td>99.68</td>
<td>97.90</td>
<td>97.90</td>
</tr>
<tr>
<td>hist. equ.</td>
<td>50.12</td>
<td>40.52</td>
<td>49.30</td>
<td>98.15</td>
<td>98.15</td>
</tr>
<tr>
<td>Deviation of Gauss.</td>
<td>52.04</td>
<td>39.14</td>
<td>53.89</td>
<td>98.15</td>
<td>98.15</td>
</tr>
<tr>
<td>Scaling</td>
<td>52.04</td>
<td>39.14</td>
<td>53.89</td>
<td>98.15</td>
<td>98.15</td>
</tr>
<tr>
<td>rotation (HRBD)</td>
<td>47.96</td>
<td>33.58</td>
<td>53.77</td>
<td>73.90</td>
<td>79.90</td>
</tr>
<tr>
<td>Brighten rate (HRBD)</td>
<td>89.97</td>
<td>74.88</td>
<td>74.26</td>
<td>97.90</td>
<td>97.90</td>
</tr>
<tr>
<td>Brighten adj. (HRBT)</td>
<td>100</td>
<td>74.94</td>
<td>99.68</td>
<td>97.90</td>
<td>97.90</td>
</tr>
<tr>
<td>hist. equ.</td>
<td>50.12</td>
<td>40.52</td>
<td>49.30</td>
<td>98.15</td>
<td>98.15</td>
</tr>
<tr>
<td>Deviation of Gauss.</td>
<td>52.04</td>
<td>39.14</td>
<td>53.89</td>
<td>98.15</td>
<td>98.15</td>
</tr>
<tr>
<td>Scaling</td>
<td>52.04</td>
<td>39.14</td>
<td>53.89</td>
<td>98.15</td>
<td>98.15</td>
</tr>
<tr>
<td>rotation (HRBT)</td>
<td>47.96</td>
<td>33.58</td>
<td>53.77</td>
<td>73.90</td>
<td>79.90</td>
</tr>
<tr>
<td>Brighten rate (HRBT)</td>
<td>89.97</td>
<td>74.88</td>
<td>74.26</td>
<td>97.90</td>
<td>97.90</td>
</tr>
<tr>
<td>Brighten adj. (HRBT)</td>
<td>100</td>
<td>74.94</td>
<td>99.68</td>
<td>97.90</td>
<td>97.90</td>
</tr>
<tr>
<td>hist. equ.</td>
<td>50.12</td>
<td>40.52</td>
<td>49.30</td>
<td>98.15</td>
<td>98.15</td>
</tr>
</tbody>
</table>

We consider the detection rate as performance indicator, which is the number of modified and original images correctly detected versus the number of tested images. The results obtained for different detectors are given in Table II.

From the results in Table II, we find out that HRBD and HRBT features can be used for image modification detection and that they have similar performances. HRBT perform slightly better than HRBD except for JPEG, contrast and brightness adjustment. This can be explained by the fact that JPEG quantizes DCT coefficients and that HRBT evolves slightly or strongly the image content. However, they are only slightly better than HRBD features in a few cases, not in all cases as expected. Even so, our proposed method widens the application of Tchebichef moments.

V. CONCLUSION

In this paper, we have proposed to use Tchebichef moments to built image features (HRBT) following the strategy of Liu et al. (HRBD) for the purpose of verifying medical image integrity. Originally proposed for the steganalysis of natural images, we show that this approach can also be used for blind detection of global image modification. It outperforms other solutions based on HOWs, IQM and BSM features. Obtained results on different medical image modalities show that our features based on Tchebichef moments perform well on modifications that either affect slightly or strongly the image content. However, they are only slightly better than HRBD features in a few cases, not in all cases as expected. Even so, our proposed method widens the application of Tchebichef moments.

REFERENCES