A Rough Geometric-Distortion Resilient Watermarking Algorithm Using Fuzzy C-Means Clustering of SURF Points

Saeid Fazli*, Masoumeh Moeini

* Research Institute of Modern Biological Techniques, University of Zanjan, Daneshgah Blvd, Zanjan-38791, Zanjan, Iran.
* Corresponding author. Tel.: +982433051; E-mail address: fazli@znu.ac.ir, m.moieni@znu.ac.ir

Abstract

Keywords: SURF Points, Fuzzy C-Means Clustering, Delaunay Triangulation, Rough Geometric Distortion, Watermarking.

In recent times, due to great developments in computer and internet technology, multimedia data i.e. images, audio and video have found wide applications. Digital watermarking is one of the best solutions to prevent illegal copying, modifying and redistributing multimedia data. Robustness is one of the most important qualities of watermarking. Basically, robustness is the watermarking technique’s tolerance to common geometric distortions (such as cropping, scaling, and spatial transformation). A robust watermark should be able to survive all those distortions. Watermark can be done in visible and invisible manner. Visible watermark places a watermark over an image in such way that it is visible to all the users about the owner of the image and invisible watermarking is done in such a way that watermark content is hidden from the user. In this paper, we present invisible watermarking methods applied on digital images for copyright protection. This paper presents a novel robust watermarking algorithm for rough geometric distortions. In this algorithm, speed up robust features points extracted from standard images. Then these points were clustered with fuzzy C-means clustering. The membership function matrices contain the grade of membership of each data point in each cluster. The coordinates for each cluster center are squeezed for employment as a feature points. At last; the set of extracted feature points is used to divide the image into triangular patches, to prepare resistant region for watermark embedding. Delaunay triangulation was used for this shaping. Experimental results support the potential of the proposed algorithm for resilient in rough geometric distortion.

Accepted: 27 April 2015 © Academic Research Online Publisher. All rights reserved.

1. Introduction

In recent years, all business applications are moving towards the digital era because of great development in latest technologies such as in the area of communication, networked multimedia system, digital data storage etc. Also from the last two decades use of internet is rapidly increased in business environment towards achievement of effectiveness, convenience and Security by introducing the digitization in their work. It was estimated that in 1993 the Internet will carry only 1% of the information however by 2000 this figure had grown to 51%, and by 2013 more than 97 % information was carried away across the globe [1]. This digital data includes text, images, audio, video and software which are transferred over open public network, hence there is need to protect this data. There are many techniques that are available...
for protection of this digital data, such as encryption (cryptography), authentication and time stamping. Also there is another method that improved the protection of digital data by merging a low level signal directly into the digital data. This low level signal is known as watermark, that uniquely identities the ownership and provides the security to the digital data and can be easily extracted [2-3].

The process of embedding the watermark into a digital data is known as Digital Watermarking. It is a process of embedding unremarkable logos or labels or information data or pattern into the digital data [4]. Figure 1 depicts the simple example of encoding of a mark into the host image. The mark is a binary image that has been uniquely generated by the watermarking system. The next step is to insert the watermark by watermark embedder. There are a variety of watermark embedding methods recommended in the literature [5].

A complete watermarking system is exhibited in Figure 2, which contains watermark embedder and watermark extractor. The inputs to the embedder are images and mark, which is to be embedded into the original image. The output of watermark embedder is watermarked image.

In general, watermarking techniques require several properties such as robustness [6]. It is one of the most important qualities of watermarking. Basically, robustness is the watermarking technique’s tolerance to common geometric distortions (such as cropping, scaling, and translation). A robust watermark should be able to survive all those distortions, because watermarked image may go through certain intentional or unintentional distortions in the real world. As a result, a watermark detection scheme should be robust in finding and verifying the embedded watermark under possible distortions. The robustness of the watermark to common image processing and geometric attacks is important to the copyright marking system [6]. Geometric distortions are very difficult to tackle because they can make the verification task unreliable by inducing synchronization errors between the extracted and original watermarking positions during the detection process [7].

Suppose that X is the original image and M is the mark to be embedding. In digital watermarking system a embedding function F() takes X and M as a input values and gives X i.e. watermarked data as a output. X is obtained as equation 1.

\[ X' = F(X,M) \] (1)
The embedding algorithm is considered as robust if watermark is embedded in away such that it can survive even if the watermarked data $X'$ goes through several attacks. Attacked image was shown with $X'$. During extraction process, the extraction function $E()$ is defined as:

$$\tilde{M} = E(X'', M) \tag{2}$$

Where $\tilde{M}$ is retrieved watermark, $\tilde{M}$ should be as close to $M$ as possible. If $X'$ is not manipulated/modified with attacks, the extracted watermark should be same as $M$, but if $X'$ is modified, $\tilde{M}$ should still match $M$ to give clear judgment of the existence of watermark. For fragile watermarking, after even the slight manipulation to $X$, extracted $\tilde{M}$ would be totally different from $M$. But in robust watermarking, after even the rough manipulation to $X$, extracted mark would be as close as $M$.

The goal of this paper is to present a novel rough geometry attacks resilient watermarking algorithm. The proposed algorithm is based on fuzzy C-means clustering of SURF feature points. So feature point detection is a one of critical step in this paper.

The remainder of this paper is ordered as follows: Literature review of feature point detection is presented in section (2). Brief introduction of Speed Up Robust Feature (SURF) points, fuzzy C-means clustering and Delaunay triangulation were introduced in section (3), (4) and (5) respectively. Section (6) shows the experimental result of the proposed algorithm. Finally concluding remarks are included in Section (7).

2. Literature Review

Feature point is one of the best candidates for determining watermark inserting region. Feature Point Detectors, the most widely used detector probably is the Harris corner detector second-moment matrix. However, Harris corners are not resister for scale-invariant attacks [10].

Lindeberg proposed the idea of automatic scale selection [11]. This permits determining feature points in an image, each with their own scale. He experimented with both the determinant of the Hessian matrix as well as the Laplacian (which corresponds to the trace of the Hessian matrix) to detect blob-like structures in pixels pattern.
For increasing speed of algorithm, Lowe [12] approximated the Laplacian of Gaussian (LoG) by a Difference of Gaussians (DoG) filter. By studying the existing detectors results and from published comparisons [13, 14], we could conclude that (1) Hessian-based detectors are more stable and repeatable than other counterparts. This detector is stronger than Harris-based detectors. (2) Approximations like the DoG can increase speed at a low cost in terms of lost accuracy.

Our approach in this paper, use a novel detector-descriptor scheme, named SURF, which introduced in [15]. SURF is a detector and a high-performance descriptor points of feature point in an image where the image is transformed. Up to now, it be used in computer vision tasks like object recognition or 3D reconstruction, and rarely used for other task. We propose utilize SURF points for image watermarking.

3. SURF Points

SURF is a new scale and rotation-invariant interest point detector and descriptor introduced in literature [15]. It approximates or even outperforms previously proposed schemes with respect to repeatability, robustness and distinctiveness yet could be computed and compared much fast.

In this research, the stable SURF key points of watermarked image are hired to estimate rough geometric distortion factors and correct the watermarked image, including JPEG compression, image cropping distortion, translation distortion and scale distortion. After watermark embedding, the feature points and their corresponding descriptors of the watermarked image are extracted and saved for make decision of algorithm resilient.

In the process of watermark extraction, the attacked watermarked image’s key points are extracted, and then the key points matching are started. Finally, the matching points will be listed and the distortion factor (MSE) will be calculated.

3.1. Image integral

An Integral of image I(q), is an image where for each point grey level q(x,y), stores the sum of all pixels in a rectangular area between a and b. Equation (3) shows this calculation.

\[
p(x,y) = I(q(x,y)) = \sum_{i=a}^{b} \sum_{j=b}^{j=b} q(i,j) \quad (3)
\]

Image integral allow for the fast implementation of convolution filters.

3.2. Constructing the Hessian Matrix

For given a integral of image I(q(x,y)) = p(x, y); the Hessian matrix H(p, σ) in location of x at scale σ, is defined as equation (4).

\[
H(p, \sigma) = \begin{bmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{bmatrix} \quad (4)
\]

where L (p, σ) is the convolution of the Gaussian second order derivative \( \frac{\partial^2}{\partial x^2} g(\sigma) \) with the image integral (p). in point x. \( L_{xx}(p, \sigma), L_{xy}(p, \sigma) \) and Gaussian filter equation were shown in equation (5), (6) and (7) respectively.

\[
L_{xx}(p, \sigma) = \frac{\partial^2}{\partial x^2} g(\sigma) \ast p(x, y) \quad (5)
\]

\[
L_{xy}(p, \sigma) = \frac{\partial^2}{\partial x \partial y} g(\sigma) \ast p(x, y) \quad (6)
\]

\[
g(\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (7)
\]

Gaussian filter can blur corners and edges of image in the bad manner and it can be cause to lost important features of image. So selection of σ is
The 9 × 9 box filters in figure 3 are approximations for Gaussian second order derivatives with σ = 0.5. The color of each region depicts filter mask’s weights (-2<weight<1). Notice that each weight can be decimal. So calculation process will be heavy and take long time for convergence.

\[ \frac{\partial^2}{\partial x^2} g(0.5) \]

\[ \frac{\partial^2}{\partial x \partial y} g(0.5) \]

Fig. 3: Gaussian second order derivatives in x-direction and xy-direction

For decrease calculation load and time, approximate second order Gaussian derivatives, can be evaluated, and speed up features detection will be provided. The approximation Gaussian second order derivatives in x-direction and xy-direction were shown in figure 4. These weights will be integer (-2,-1,0,1).

\[ \cong \frac{\partial^2}{\partial x^2} g(0.5) \]

\[ \cong \frac{\partial^2}{\partial x \partial y} g(0.5) \]

Fig. 4: approximation Gaussian second order derivatives in x-direction and xy-direction

The Hessian determinant is calculated according to the equation (4). This equation was shown in equation (8).

\[ \text{DOH}(p, \sigma) = \frac{D_{xx}D_{yy} - (D_{xy})^2}{\sigma^2} \]  

(8)

where DOH is Determinant OF Hessian matrix, Dxx, Dyy and Dxy denote the weighted box filter approximations in the x, y and xy-directions in figure (4).

The final approximation for speed up algorithm was utilized in equation (8). It is kept simple for computational efficiency. Equation (9) was revealed this approach.

\[ \text{DOH}_{\text{Approx}}(p, \sigma) = D_{xx}D_{yy} - ((0.9)D_{xy})^2 \]  

(9)

3.3. SURF Points

The DOH of each pixel was counted as a SURF point of pixel. But Value greater than or equal to 1.6. Specifies SURF point at the location of where a feature point was detected. Circle was used for depicting SURF points. Radius of circle was characterized of SURF point strength of detected feature, and center of circle was described the location of this point. Equation (10) exhibits SURF point’s derivation form location of (t1, t2) and strength of (r).

\[ (x - t1)^2 + (y - t2)^2 = (r)^2 \]  

(10)

Figure (5) displays the SURF point’s location and figure (6) shows the strength of standard Lena, living room and Pirate image. Notice that large circles in the figure (6) exhibit durability of points and small circles reveal less durability of detected feature points.
4. Fuzzy C-means Clustering

Clustering of numerical data forms the essential steps of many classification and system modeling algorithms, like watermarking. The purpose of clustering is to identify natural groupings of data from a large data set to manufacture a concise representation of a point’s behavior and distribution in whole image.

![Fig. 5: SURF point's location](image1)

![Fig. 6: SURF point's strength](image2)

Fuzzy c-means clustering (FCM) is a data clustering technique in which a dataset is grouped into \( n \) clusters with each data point in the dataset belonging to every cluster to a certain degree. For example, a certain SURF point that lies close to the center of a cluster will have a high degree of belonging or membership to that cluster and another SURF point that lies far away from the center of a cluster will have a low degree of belonging or membership to that cluster.

It starts with an initial guess for the cluster centers, which are intended to mark the mean location of each cluster. The initial guess for these cluster centers is most likely incorrect. Next, algorithm assigns every SURF point a membership grade for each cluster. By iteratively updating the cluster centers and the membership grades for each data point. Fuzzy clustering iteratively moves the cluster centers to the right location within a SURF set. This iteration is based on minimizing an objective function that represents the distance from any given SURF point to a cluster center weighted by that SURF point’s membership grade.

Fuzzy C-means clustering is based on minimization of the equation (11) objective function [16].

\[
J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} U_{ij}^m \|X_i - c_j\|^2 \quad 1 \leq m < \infty \quad (11)
\]
where \( m \) is any real number greater than 1, \( U_{ij}^m \) is the degree of membership of \( X_i \) in the cluster \( J \), \( X_i \) is the \( i \)-th of \( d \)-dimensional measured image matrix, \( c_j \) is the \( d \)-dimension center of the cluster, and \( \| \ast \| \) is any norm expressing the similarity between any measured data and the center.

Fuzzy partitioning is carried out through an iterative optimization of the objective function shown equation (11), with the update of membership \( U_{ij}^m \) and the cluster centers \( c_j \) by equation (12) and (13) respectively.

\[
U_{ij} = \frac{1}{\sum_{k=1}^{C} \left( \frac{\| X_i - c_j \|}{\| X_i - c_k \|} \right)^{\frac{2}{m-1}}} 
\]

\[
c_j = \frac{\sum_{i=1}^{N} U_{ij}^m X_i}{\sum_{i=1}^{N} U_{ij}^m} 
\]

This iteration will stop when \( \{ |U_{ij}^{k+1} - U_{ij}^k| \} < \epsilon \), where \( \epsilon \) is a termination criterion between 0 and 1, and \( k \) is the iteration steps. This procedure converges to a local minimum or a saddle point of \( J_m \).

SURF points are bound to each cluster by means of a Membership Function, which symbolizes the fuzzy behavior of this algorithm. To do that, we simply have to build an appropriate matrix named “U” whose factors are numbers between 0 and 1, and show the degree of membership between data and centers of clusters. At this paper we identify 20 clusters in proximity of the 90 SURF points. Figure 7 shows Membership Function of this fuzzy clustering and figure 8 displays these cluster’s center location of the Lena image.

5. Delaunay Triangulation

Extracted cluster’s centers are used to divide the image into elementary patches, which are then warped into a standard geometry. The watermark is inserted into the image with the standard geometry and then the watermarked image is unwrapped to obtain the watermarked version of the elementary patch. While the standard geometry may be of any shape and size, it is advantageous to use triangles perspective warping transformations.

Obviously, elementary patches formed by the tessellation of the image may have the same number of peaks as the standard shapes have. Our goal in the analysis of the requirements for this module of the watermarking process is primarily on the repeatability of the image-segmentation into elementary blocks, and the distortions in the
embedded watermark information established by the process of unwrapping elementary blocks from the standard shapes and sizes.

In particular, we evaluate the case where a given image is segmented using Delaunay triangulation and warped using affine-transformations.

Figure 9 explains Delaunay triangulation of cluster's centers of Lena image and figure 10 displays Delaunay triangulation of cluster's centers of living room image. The proposed algorithm can be summarized as figure 11.

6. Experimental Result

In this section, we evaluate the performance of the proposed watermarking scheme. The experimental results are analyzed with different scenarios while computing the proposed technique in this paper. So we consider geometric distortion of images and its effect on cluster centers' points and Delaunay triangular.

JPEG compression is one of the essential reasons for the success of the internet and must be taken into account when designing an image watermarking system. JPEG compression will attempt to remove the perceptual unimportant elements from an image and may transform the imperceptible watermark undetectable. Image compressions are considered one of the strongest enemies of digital watermarking techniques today. So finding the effect of JPEG compression in the embedded mark will be good step in evaluation of proposed algorithm.
Mean Square Error (MSE) of each image in various attacks was simulated. It is defined average squared difference between to reference image clusters’ centers and distorted image clusters centers. It is calculated by the equation (14).

\[
MSE = \frac{1}{N} \left[ \sum_{i=0}^{x} \sum_{j=0}^{y} (c(i,j) - c'(i,j))^2 \right]
\] (14)

Where \(N\) is number of cluster, \(x\) and \(y\) are coordinate of cluster, \(c(i,j)\) is the cluster center of the original image and \(c'(i,j)\) is the cluster center of the attacked image.

MSE is also calculated for SURF points’ locations and strength. Equation (14) is improved for equation (15) for SURF location and curved for equation (16) for SURF strength.

\[
MSE = \frac{1}{M} \left[ \sum_{i=0}^{x} \sum_{j=0}^{y} (L(i,j) - L'(i,j))^2 \right]
\] (15)

\[
MSE = \frac{1}{M} \left[ \sum_{i=0}^{x} \sum_{j=0}^{y} (S(i,j) - S'(i,j))^2 \right]
\] (16)

Where \(M\) is number of SURF points, \(x\) and \(y\) are coordinate of point, \(L(i,j)\) is the point center of the original image and \(L'(i,j)\) is the point center of the attacked image. \(S(i,j)\) is the point strength of the original image and \(S'(i,j)\) is the point strength of the attacked image.

Figure 12 shows some rough geometric attack of Living room standard image, their Cluster centers and Delaunay Triangulation of these images. At this figure; it is clear that rough geometric distortion cannot alter cluster centers point and Delaunay triangular in serious matter.

![Fig. 12: Cluster centers and Delaunay Triangulation of rough attacked living room image](image-url)
The JPEG compression resistance test presented in this paper is performed on 10 standard images over 10 different JPEG quality factors ranging from 100% to 10%. The cluster centers MSE value is calculated for each individual image as illustrated in table 1.

Table 1: MSE value for JPEG compression

<table>
<thead>
<tr>
<th>JPEG Quality</th>
<th>Cluster centers MSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>0.05</td>
</tr>
<tr>
<td>90%</td>
<td>0.12</td>
</tr>
<tr>
<td>80%</td>
<td>0.32</td>
</tr>
<tr>
<td>70%</td>
<td>0.49</td>
</tr>
<tr>
<td>60%</td>
<td>0.64</td>
</tr>
<tr>
<td>50%</td>
<td>0.89</td>
</tr>
<tr>
<td>40%</td>
<td>1.04</td>
</tr>
<tr>
<td>30%</td>
<td>2.07</td>
</tr>
<tr>
<td>20%</td>
<td>2.54</td>
</tr>
<tr>
<td>10%</td>
<td>5.98</td>
</tr>
</tbody>
</table>

Resistance of proposed algorithm for JPEG compression is plotted in figure 13. The illustration in figure 13 is the direct result of individual JPEG operations for each image on watermarking points, and shows the advantage of the proposed algorithm.

It clearly shows that our proposed watermarking scheme is resistant to JPEG compressions since robust feature points of the SURF points and fuzzy clustering. Fuzzy clustering SURF points watermarking scheme is resistant to JPEG compression quality factor of 94% or above.

Table 2 summarizes our watermarking points MSE results under various rough geometric attacks, which depict in figure 12.

As a result in table 1 and table 2, we claim that our proposed watermarking algorithm is resistant to all geometric attacks and JPEG compression in the range of MSE factor 0.00 to 5.98. In other words it works extremely well on the standard images with low MSE factor.

Table 2: MSE value for geometric attacks

<table>
<thead>
<tr>
<th>Attack</th>
<th>Cluster centers MSE</th>
<th>SURF Center MSE</th>
<th>SURF Strength MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial transformation</td>
<td>2.34</td>
<td>5.4</td>
<td>3.77</td>
</tr>
<tr>
<td>Symmetry mapping</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>20% cropping scaling</td>
<td>1.02</td>
<td>1.66</td>
<td>2.45</td>
</tr>
<tr>
<td></td>
<td>3.34</td>
<td>4.56</td>
<td>5.79</td>
</tr>
</tbody>
</table>

7. Conclusions

A new digital watermarking system; called fuzzy clustering SURF points, is discussed as a method for protecting copyright for digital images. In the proposed algorithm, SURF points obtained from standard images. Then these points were clustered with fuzzy C-means clustering. Membership function matrices were squeezed for this clustering. The coordinates for each cluster center are used for...
employment as a feature points. At final step; the set of extracted feature points is utilized to divide the image into triangular patches, to prepare resistant region for watermark embedding. Delaunay triangulation was used for this shaping.

Different tests demonstrate the repeatability and accuracy of this algorithm. Based on MSE calculation, we assert that our proposed watermarking algorithm is resistant to all geometric attacks and JPEG compression in the range of MSE factor 0.00 to 5.98. Experimental results reveal that the proposed algorithm is powerful for watermark usage.

References