Analysis of feature-based geometry invariant watermarking

Mehmet U. Celik[*a], Eli Saber[b], Gaurav Sharma[b], A.Murat Tekalp[a]

[a] University of Rochester, Rochester, NY, 14627-0126
[b] Xerox Corporation, 800 Phillips Road, Webster, NY, 14580

ABSTRACT

Ensuring robustness against geometrical manipulations is a challenging task for watermarking algorithms. Watermarking algorithms that depend on the inherent geometry of the image for synchronization fail ungracefully due to de-synchronization through geometric manipulations. A class of algorithms tries to overcome this problem using feature-based synchronization markers. These are essentially composed of three building blocks: feature point extraction, elementary patch formation and registration to a standard geometry, and watermarking of elementary patches. Extracted feature points are bound to the image content, thus may be used as robust synchronization markers. This set of points is used to divide the image into elementary patches, which are warped into a standard geometry ensuring the exact synchronization during insertion and extraction. Once the synchronization is provided, any watermarking technique may reliably operate on elementary patches.

In this paper, we present an analysis of feature-based geometry invariant watermarking algorithms. A discussion of the requirements on each building block is followed by potential solutions to meet these requirements. Furthermore, we present theoretical and practical limitations of these solutions via examples. In particular, segmentation based feature point extractors and triangulation based elementary patch formations are evaluated.

Keywords: Watermarking, geometry invariance, feature extraction, registration, warping.

1. INTRODUCTION

Digital imaging has seen tremendous growth over the last decade, wherein now it is possible to find and download a large number of images from the world wide web in a matter of seconds. Entire digital libraries can be located on the web, where the user has the capability to assess, copy, and re-distribute many images easily and upon demand. In order to protect and preserve the owner's rights, a number of watermarking algorithms have been proposed in the literature [1,2]. Robust watermarking algorithms, which are able to withstand a large degree of image manipulations, may be utilized in copyright enforcement and access-control, where semi-fragile ones with a certain degree of robustness are valuable for authentication and content verification.

Powerful publicly available software packages, such as Adobe PhotoShop or PaintShop Pro, and specialized software tools that particularly aim to disable/remove watermarks, such as Stirmark [3], extend the range of possible image manipulations a watermarking algorithm should withstand. The set of manipulations ranges from compression and filtering to geometrical manipulations.

1.1. Effects of Geometrical Manipulations on Watermarking Schemes

A majority of the watermarking algorithms proposed in the literature operate on a principle analogous to spread-spectrum communications. A pseudo-random sequence, which is generated using a secret key, is inserted into the image. During extraction, the same pseudo-random sequence is correlated with the estimated pattern extracted from the image. The watermark is said to be present if the computed correlation exceeds a chosen threshold value. Among this general class of watermarking schemes, there are several variations that include choice of specific domain for watermark insertion, e.g. spatial, DCT, wavelet, etc; and enhancements of the basic scheme to improve robustness and reduce visible artifacts. The computed correlation depends on the alignment of the pattern regenerated and the one extracted from the image. Thus proper synchronization of the two patterns is critical for the watermark detection process. Typically, this synchronization is provided by the inherent geometry of the image, where pseudo-random sequences are assumed to be placed on the same image geometry. When a geometric manipulation is applied to the watermarked image, the underlying geometry is distorted, which

[*] Correspondence: Email: celik@ece.rochester.edu; WWW: http://www.ece.rochester.edu/~celik; Telephone: 716 275 2628

often results in the de-synchronization and failure of the watermark detection process. The geometric manipulations can range from simple scaling and rotation or aspect ratio changes to more complicated random geometric distortions as applied by Stirmark.

Different methods have been proposed in literature to reduce/prevent algorithm failure modes in case of geometric manipulations. For non-oblivious watermarking schemes, where the original image is available at the detector, the watermarked image may be registered against the original image to provide proper synchronization [4]. For oblivious watermarking schemes, where the original image is not available at the detector, proposed methods include use of the Fourier-Melin transform space that provides rotation, translation, scale invariance [5], and watermarking using geometric invariants of the image such as moments [6] or cross-ratios [7]. Hartung et al [8] have also proposed a scheme that divides the image into small blocks and performs correlation for rotations and translations using small increments, in an attempt to detect the proper synchronization

In this paper, a class of oblivious image watermarking algorithms, which attempt to obtain geometry-invariant oblivious watermarks using image-dependent feature-based synchronization markers, is analyzed. In Section 2, we introduce the framework for such algorithms. In Section 3, we present a detailed analysis of the framework. We end in Section 4 with concluding remarks on this class of algorithms.

2. FEATURE-BASED GEOMETRY INVARIANT WATERMARKS

The algorithms in this class can be represented in the framework shown in the block diagram of Fig. 1. The three building blocks of the watermarking scheme are:

- feature point extraction,
- elementary patch formation and registration to a standard geometry,
- watermarking of elementary patches

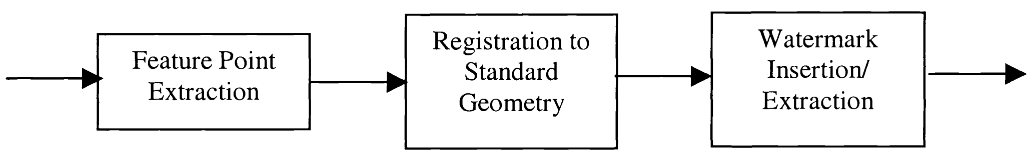

Figure 1. Building blocks for the feature-based geometry invariant watermarking.

While a discussion of the individual blocks is deferred until the next section, the complete watermark insertion and extraction procedure may be summarized as follows:

In the *feature point extraction* step, synchronization markers which are bound to the image content, and thus are robust to geometrical manipulations as well as content preserving signal processing operations, are obtained. This set of feature points is utilized to divide the image into elementary patches in the *elementary patch formation* step. It is crucial to obtain the same partitioning of the image every time, since these elementary patches are used to insert and extract watermarks. The *registration to a standard geometry* step warps these elementary patches to a standard shape. Typically, the warping transformation for registration will have some degree(s) of freedom corresponding to the different topology preserving point correspondences between the elementary patches and the standard geometry. While the orientation causing least degree of deformation may be chosen during watermark insertion, it may be necessary to check all possible transformations during extraction.

Once the elementary patches are registered to a standard shape, hence a uniform synchronization is provided, any watermarking algorithm may reliably operate on elementary patches. After watermark insertion, warping operation may be reversed, so that watermarked image patch is registered to its original location in the image. An excellent example of this class of algorithms is presented by Bas et al in [9], where they utilize warping of pre-defined triangular patterns.

3. ANALYSIS OF FEATURE-BASED GEOMETRY INVARIANT WATERMARKS

The functional partitioning of geometry-invariant watermarking into the modular blocks shown in Figure 1 allows independent evaluation of different candidates for each of the blocks. In addition, the partitioning is also advantageous in that it allows an analysis of the requirements for each component individually. Such an analysis is presented in the next two sections.

3.1. Feature Point Extraction

As stated earlier, the *feature point extraction* module extracts synchronization markers which are robust to geometrical manipulations and to other content-preserving signal processing operations. It is a challenging task to find feature point extractors which produce repeatable results under the broad range of image processing operations which can be typically employed to attack watermarking schemes. Among known feature detectors, potential candidates for the feature point extraction are those based on image segmentation and those based on direct localization of image feature points such as edge and corners.

3.1.1. Segmentation-Based Feature Point Extraction

Segmentation based feature point extraction algorithms utilize a segmentation of the image for the determination of the feature points. For instance, the centroid of each region identified through segmentation may be selected as a feature point. Regions in the segmentation map are expected to be invariant to common image processing operations such as lossy compression and contrast enhancements. Moreover, each region will be effected by the geometric manipulation as the whole image, resulting in feature points that migrate along with the embedded watermark in case of a manipulation. In this section, we consider the characteristics of two candidates for segmentation based feature point extraction.

3.1.1.1. Segmentation based on Gibbs Random Fields & Effects of local operators

As an example, we used a Gibbs Random Field (GRF) based color image segmentation algorithm [10]. GRF based algorithms are known to provide a segmentation of the image into spatially contiguous regions. The centroids of the regions are selected as feature points. For example, the segmentation of two differently scaled versions of the "Lena" image is shown in Fig. 2. While the segmented regions in either case are spatially contiguous and in reasonable agreement with a hand segmentation, the number of segmented regions obtained for the two cases is different (compare for e.g., the black regions), moreover some of the regions determined in the segmentation are significantly modified. The algorithm is therefore not robust to scale changes.

Figure 2. Segmentation of Lena256 image using a GRF based algorithm. (on right 75% scaling)

GRF based segmentation provides a good example illustrating the limitations of the class of feature detectors that depend on local operators for the determination of feature points. When the image is subjected to a geometric operation that alters the pixels under the "operator mask", the results of the feature point detection may also be altered resulting in a non-repeatable localization of the feature points and consequent failure of watermark detection. Scale and aspect ratio changes specially tend to effect local operators significantly, as the effective area under the fixed size operator deviates and the assumed statistics for pixel neighborhoods are also altered. Nonetheless, the effect of this phenomenon to the final result of the feature point detector may not be significant, as for the majority of the regions in the GRF example. Scaling to a fixed image size prior to segmentation may also be used as an alternative to combat this problem, however, in that case, cropping of the image can result in similar problems with repeatability.

3.1.1.2. Segmentation based on Color-Clustering using the *K-means* algorithm

Instead of using local operators, one may instead employ a global approach in the determination of the feature points. A global scheme is less susceptible to repeatability problems caused by change in the pixels under the operator mask, since a significant modification of the image may be required to significantly influence the global operations. As an example of a global approach, we consider a feature point extraction based on clustering of colors in the image. The image colors may be assigned to a small number of clusters determined from the image using for instance the "K-means" (or generalized LBG) algorithm. The centroids of the pixel locations corresponding to each color cluster may then be used as the feature points. Note that the algorithm does not guarantee spatially contiguous clusters. In our tests, an 8-cluster K-means algorithm is used. The algorithm was initialized using the 8 highest peaks of a coarse color histogram of the image as cluster centroids. Figure 3, illustrates the feature points locations of extracted using k-means clustering scheme for the two differently scaled versions of the "Lena" image. The deviation of feature point locations is at most one pixel for the majority of the clusters, while in two cases the deviation is two pixels. However, considering possible truncation errors, the deviation of the feature points are negligible for scaling of the "Lena" image.

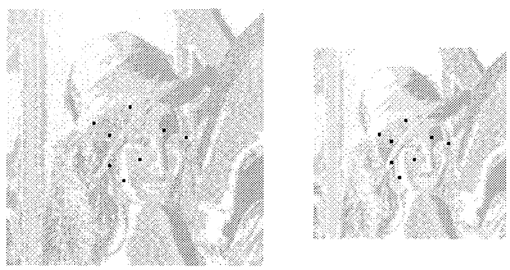

Figure 3. Spatial centroids of each cluster for Lena256 image. (on right 75% scaling)

Stability of the feature point locations for various manipulations of the "Lena" image is indicated in Table 1. The table lists the root-mean-squared (RMS) deviation of the feature point locations for the manipulated image with respect to the expected location based on the feature point locations for the original image. The results indicate that feature point extraction using k-means algorithm has satisfactory performance under geometrical manipulations and high quality JPEG compression. However, manipulations that have significant effect on the image histogram, e.g. histogram equalization, adversely effect the stability of the feature points. Note that for this particular case, the histogram equalization also dramatically changes the image from the original.

Manipulation	RMS Deviation (pixels)
Scaling (75%)	1.2
Aspect (4:3)	0.8
Rotation (30°)	1.7
Histogram Equalization	> 15
JPEG (Q75)	1.6
JPEG (Q30)	2.2

Table 1. Stability of feature points by "K-means" algorithm. (Lena image)

Another criterion that may be used for evaluation of the feature point extraction module is the distribution of feature points over the image or the region of interest. An irregular distribution may effect the performance of the overall algorithm since it may result in shapes that have to be significantly altered in order to register them to a standard geometry. It is self evident that the distribution of feature points depends more on the image under test rather than the particular algorithm employed for the feature point extraction. In order to demonstrate this effect, the feature points for the "Peppers" image, extracted using the K-means clustering approach and shown in Figure 4, may be contrasted with the feature points for the "Lena" image shown earlier in Fig. 3. While the feature points for the "Lena" image are well distributed over the center of the image which constitutes the natural region of interest, the distribution of colors in the Peppers image results in feature points that are clustered in a very close proximity at the center.

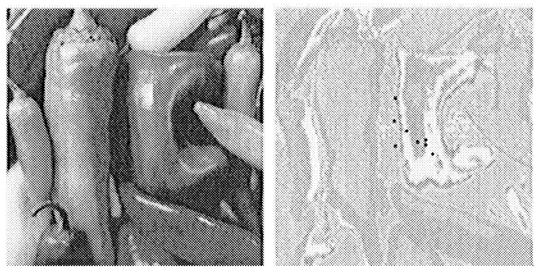

Figure 4. Peppers image and spatial centroids of each cluster.

3.1.2. Edge Based Feature Point Extraction

Edge and corner detectors are popular in image processing [11,12] and may be employed for feature point extraction. While they often provide good agreement with a human operator demarcation of edges and corners, these detectors employ local operators in a manner similar to the GRF segmentation algorithm described earlier in this paper. They are therefore subject to similar limitations of repeatability caused by changes of the pixel values under the operator mask. In addition, these operators typically provide a large set of candidate points, from which a smaller subset is selected as feature points based on the strength of the detector response and other criteria. A common technique is to select local maxima of the edge/corner-detector responses in a defined local neighborhood for each feature point. The change of the local neighborhood for the maxima search may also effect the repeatability of these algorithms.

The analysis presented above placed significant emphasis on obtaining repeatability from the feature extraction algorithms. However, it may be noted that the overall performance of the watermarking algorithm may only be modestly impacted by the inaccuracies in repeatability of the feature point extraction procedure. A consistent development of the other modules may compensate for these inaccuracies, or survival of a group of points may be sufficient for reliable watermark extraction thereby trading-off capacity in favor of reliable watermark detection.

3.2. Formation and Registration of Elementary Patches

The set of extracted feature points is used to divide the image into elementary patches, which are then warped into a standard geometry. As indicated earlier, the watermark is embedded into the image with the standard geometry and then the image is unwarped to obtain the watermarked version of the elementary patch. While the standard geometry may be of any shape, it is advantageous to use triangles or quadrilaterals in accordance with 6-point affine or 8-point perspective warping transformations, respectively. Clearly, elementary patches formed by the tessellation of the image should have the same number of vertices as the standard shapes. Our focus in the analysis of the requirements for this module of the watermarking process is primarily on the repeatability of the image-tessellation into elementary blocks, and the distortions in the embedded watermark information introduced by the process of unwarping elementary blocks from the standard shapes. In particular, we evaluate the case where a given image is tessellated using Delaunay triangulation and warped using affine-transformations.

Given a set of points in a plane, their Delaunay triangulation is the straight-line dual of the corresponding Voronoi diagram. For a given set of points, the results of Delaunay triangulation are independent of scale changes and rotation. However, geometric manipulations that modify the relative positions of the points in the set may result in a different tessellation. An example of such a manipulation is the aspect ratio change. Figure 5 shows the effect of anamorphic scaling along the vertical direction on Delaunay tessellation for a hypothetical example, where the triangles in the two cases represent the regions obtained from the Delaunay triangulation. Note that these regions do not correspond because the Delaunay triangulation procedure is not independent of anamorphic scaling.

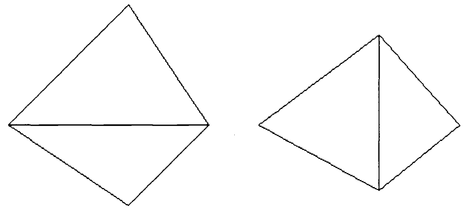

Figure 5. Effect of aspect ratio change on Delaunay triangulation.

Warping of triangles into a standard triangle (for instance an isosceles right-triangle) may be done via affine-transformations. An affine transformation may be represented by a vector equation as:

$$\mathbf{x'} = \mathbf{A}\mathbf{x} + \mathbf{b} \qquad (1)$$

where **x** and **x'** are the coordinates of original and warped points respectively, **A** is a 2x2 matrix that governs the rotation and scaling aspects, and **b** is a 2x1 translation vector. The image value at each new pixel location **x'** may then be calculated by interpolation. The necessary affine transformation parameters can be calculated using the corner correspondences of the triangles. The corresponding unwarping transformation that is applied to the standard shaped watermark block is the inverse of the warping transformation:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{x'} - \mathbf{A}^{-1}\mathbf{b} \equiv \mathbf{A}^{-1}\mathbf{x'} - \mathbf{b'}$$

The inserted watermark information is also subjected to this unwarping transformation. A frequency domain analysis of this unwarping transformation is useful for understanding the effect of this unwarping transformation on the inserted watermark information. While the term **b'** obviously contributes a mere change in phase, as a result of the unwarping transformation, the frequency coordinates are warped by a transformation which corresponds to $(\mathbf{A})^T$. While the exact impact of this coordinate transformation is dependent on value of the matrix **A**, one can see by means of an example that this transformation is likely to create aliasing effects in the inserted watermark. Figure 6 shows impact of an example unwarping transformation on the frequency spectrum of the watermarked image in the standard triangle. The two triangles on the left represent the standard triangle and the unwarped elementary triangle obtained from the standard triangle by a scaling of 0.85 and rotation of 45 degrees. On the right, corresponding changes in frequency domain are shown. Shaded regions are points of aliasing under the assumption that the original patch has non-zero spectrum all over the frequency plane and that the same sampling resolution is used for the two triangles. Note that while the elementary triangle is again warped to the standard triangle by the transformation **A** in the watermark extraction process (assuming a repeatable tessellation of the watermarked image), the aliased information cannot be recovered.

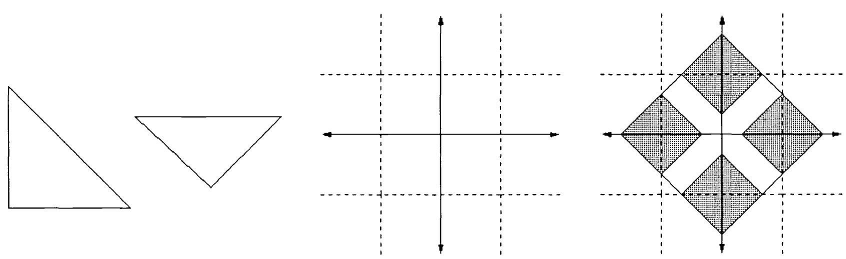

Figure 6. An affine transformation (on the left) and its effect on frequency domain (on the right). Potential aliasing is shown by shaded regions.

In Figure 7, we demonstrate this effect on a binary pseudo-noise pattern shown in 7(a). This pattern is warped into an arbitrary triangle as shown in 7(b). Warped pattern is then warped back to the original shape. Aliasing effects can be seen in 7(c), while 7(d) is the resulting difference from the original noise pattern. It is clearly seen that for the high frequency noise pattern given, performance of a common correlation detector will be much lower than expected. Designing low frequency noise patterns will definitely help to improve the performance.

Figure 7. Warping of a noise pattern. (a) Original noise pattern (b) Warped pattern (c) Warped back pattern (d) Difference between original and warped back patterns.

3.3. Watermark Insertion and Extraction

Watermark insertion and extraction module is responsible of embedding a pseudo-random pattern into the image in an appropriate domain, while improving robustness and imperceptibility via models of human visual system. An explicit analysis of such algorithms is beyond the scope of this paper, however we may elaborate on the design criteria to improve overall performance.

In Section 3.1. we have shown that locations of the feature points may deviate in a certain range. Therefore, a criterion of the watermark insertion algorithm is graceful degradation of performance with respect to small synchronization errors. Practically, low frequency watermarks may provide gradual degradation of performance. Low frequency watermarks should also be preferred due to the aliasing effects related to the warping of the patterns as explained in Section 3.2.

4. CONCLUSION

In this paper, we have analyzed a framework for feature-based geometry invariant watermarking algorithms. While the framework depends on the repeatability and accuracy of the feature point detectors, most of the algorithms fail to ensure repeatability under a broad range of image processing operations. On the other hand, tessellation of an image is in general repeatable, unless the relative geometry of the image is changed in an unusual fashion. Finally, warping of the image or watermark patch may introduce aliasing, if either the watermark insertion module is not properly designed or deformation is too strong. However, despite the outlined deficiencies, some of which may be avoided by careful design, this feature-based framework is a promising method for obtaining geometry invariant watermarks.

REFERENCES

1. M.D. Swanson, M. Kobayashi, and A.H. Tewfik, “Multimedia data-embedding and watermarking technologies”, Proceedings of the IEEE, vol. 86, no 6, pp. 1064-1087, June 1998.
2. F. Hartung, and M. Kutter, “Multimedia watermarking techniques”, Proceedings of the IEEE, vol. 87, no 7, pp. 1079-1107, July 1999.
3. Stirmark Package, “http://www.cl.com.uk/~fapp2/watermarking/stirmark”.
4. Q. Sun, J. Wu, R. Deng, “Recovering modified watermarked image with reference to original image”, in Proceedings of SPIE: Security and Watermarking of Multimedia Contents, Vol. 3657, Jan 1999, San Jose.
5. J.J.K.O Ruanaidh and T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking", Signal Processing 66, pp. 303-317, May 1998.
6. M. Alghoniemy, A. Tewfik, “Image watermarking by moment invariants”, in Proceedings of ICIP, 2000.
7. R. Caldelli, M. Barni, F. Bartolini, A. Piva, "Geometric-invariant robust watermarking through constellation matching in the frequency domain", in Proceedings of ICIP, 2000.
8. J.S.F. Hartung, B. Girod, “Spread spectrum watermarking: Malicious attacks and counter-attacks”, in Proceedings of SPIE: Security and Watermarking of Multimedia Contents, Vol. 3657, Jan 1999, San Jose.

9. P. Bas, J.M. Chassery, and B. Macq, "Robust watermarking based on the warping of pre-defined triangular patterns", in Proceedings of SPIE: Security and Watermarking of Multimedia Contents II, vol. 3971, Jan. 2000.
10. E. Saber, A. M. Tekalp, G. Bozdagi, "Fusion of color and edge information for improved segmentation and edge linking", Image and Vision Computing 15, pp. 769-780, 1997.
11. C. Harris, M. Stephen, "A combined corner and edge detector", in 4th Alvey Vision Conference, pp 147-151, 1988.
12. S.M. Smith, J.M. Brady, "Susan- a new approach to low level image processing", International Journal of Computer Vision, vol. 23, pp. 45078, May 1997.