IMAGE AUTHENTICATION USING AUTOMATIC DETECTED ROIs

Huajian Liu, Hichem Sahbi, Lucilla Croce Ferri, Martin Steinebach

Fraunhofer IPSI - Integrated Publication and Information Systems, Darmstadt, 64293, Germany
Email: {liu, sahbi, ferri, steinebach}@ipsi.fraunhofer.de

ABSTRACT

We propose in this paper a new approach to verify the integrity of digital images using automatic detection of "regions of interest (ROIs)". The goal is to achieve a strong protection of important parts of an image and to ensure a positive integrity verification of its slightly modified content. The modifications could also be a content change, but this change must be outside the regions of interest. The approach combines digital watermarking techniques with a method for automatic detection of regions of interest.

1. INTRODUCTION

Recognition of manipulations in images is addressed by digital watermarks for authentication. Existing methods can fall into three classes: fragile, content-fragile and semi-fragile watermarking algorithms [1]. One bit change in a pixel value already destroys a fragile watermark, while content-fragile watermarking is sensitive to any slight change of the data content. Semi-fragile watermarking is moderately robust to common signal processing; the difficulty of this approach is related to its limited possibility to differentiate between malicious data manipulations and common postproduction editing processes. Most of the proposed watermarking schemes are not robust to geometric transformations and cropping, resulting from common image processing. Even though slight geometric transformations and cropping preserve the image content, it is not certain that the watermark will be detectable, so the verification of integrity may fail.

Most of the existing watermarking methods for image authentication embed watermarks into the whole image without taking into account the underlying semantic content [2-4]. However, for many applications, some portions contain more important information than the rest of the image. These portions are usually referred to as the "regions of interests" (ROIs).

The ROI concept is well studied in the JPEG2000 standard for still image compression [5]. The standard supports different progressive decoding modes including the ROI functionality. In this context, a ROI is a part of the image with an arbitrary shape, which is compressed with a high quality at any decoding bit-rate than the rest of the image. In our work the ROI concept is not limited to the JPEG2000 images. Indeed, a ROI is defined as an object of interest, which can be automatically detected by the object retrieval algorithm [6] (cf. section 4.1) and extensively protected, so the integrity of the whole image only depends on the integrity of such objects.

Geometric transformations, adding or removal of information outside the ROIs (e.g. time stamp) should not make the content unauthentic, even though such manipulations change the image content. Additional protections for ROIs are required in order to distinguish between ROIs and background alterations.

In this paper we propose an approach, which ensures integrity verification of slightly modified digital images. The modifications could be content changes, but these changes are accepted only outside the ROIs. As we go through different sections of this paper, a ROI refers to an object of interest in which the user needs more protection, for instance logos, faces, cars, buildings, etc. Faces are particular objects, which need a higher level of protection as the multimedia content, involving more and more well-known person faces, is always subject to falsification and illegal copy redistributions.

The proposed authentication scheme combines digital watermarking techniques with automatic detection of faces as ROIs. Again, faces plays an important role in many security and integrity protection applications but our concept can be extended to other objects.

The paper is organized as following: section 2 briefly discusses related work in ROI based watermarking and a survey in object retrieval. We illustrate in section 3 our approach based on the embedding and the verification scheme. A discussion on advantages and limitations of the proposed method is followed in section 4 and we conclude in section 5.

2. RELATED WORK

In this section, we review the most representative studies in digital watermarking based on ROIs. We give also a brief survey in object retrieval since the proposed watermarking approach is motivated by automatic detection of the ROIs.

2.1. Image watermarking with ROI

Different approaches for ROI-based watermarking exist in the literature. For copyright protection of ROI, Park and Han [7] proposed a method based on image segmentation, which embeds encoded watermark into a face region as a semantic important part. Su et al. [8] proposed two ROI-watermarking schemes based on wavelets. In this approach the watermark information is embedded mainly in the ROI. The two schemes require either the knowledge of the original image and ROI or user's selection of ROI during the watermark retrieval.

For content integrity protection, Lie et al. [9] proposed a fragile watermarking approach for JPEG2000 images. A fragile watermark is embedded into the ROI and the relations of ROI wavelet coefficients are extracted as features and embedded into the ROB (region of background) in a robust way. However, if the ROIs are visually similar in an image database, the relation of wavelet coefficients could not be able to identify different ROIs. Furthermore, as the ROI position is assumed predefined and transmitted by side information, the scheme is limited to JPEG2000 images and vulnerable to format conversion. In addition, only one ROI is defined in all above mentioned watermarking schemes and the robustness to geometric attacks, e.g. rotation, is also questionable.

2.2. Object retrieval

Object retrieval is an important issue for several applications including multimedia search engines, video surveillance and driver assistance systems. One of the main distinguishing factors in object retrieval methods is the degree of semantic used for training and classification. Methods based on low level semantic consider a scene as a collection of regions or blobs, which are found using segmentation algorithms [10] based on low level features (color, shape, texture, etc.). In this mode the user is interested in a particular region or object in a scene, clicks on that region and the system will find automatically the corresponding objects in the database using low level features and the nearest neighbor classifier [11].

The second mode is based on high level semantic and attempts to capture the variation of a particular class of object using either geometrical or statistical models. Statistical learning models try to capture the statistical variation in the object class using a representative training set of images belonging to the object of interest [12,13]. Geometrical approaches use some a priori knowledge about the structure of the targeted object. Training consists in extracting some invariant characteristics of the object such as differential invariants and exploiting the geometrical relationship between these invariants using graphs [14], hash-tables [15], etc., in order to discriminate the object of interest from the others.

The main limitation in object retrieval is the difficulty to extend these methods to some classes of objects with particular behaviors including deformations such as cats, or some fractals such as trees. In the context of image watermarking, one important issue is object normalization, which helps the watermark to be robust to some transformations. When objects are fractal or show non-linear deformations, 3D variations of the pose and occlusions, the normalization process becomes intractable.

3. PROPOSED AUTHENTICATION SCHEME

The proposed scheme is based on three steps: face detection, embedding watermarks into each detected face and into the background and their authentication. These steps are discussed in the following subsections.

3.1. Face detection

Faces are particular semi-rigid objects subject to 2D-3D pose and photometric variations and to some non-linear deformations related to expression. Face detection is one of the most challenging problems in face analysis and there is as yet no solution with performances comparable to humans' both in precision and speed. Many methods for face detection are discussed in the literature including neural networks [16], support vector machines [17], graph matching [14], skin color learning [18] and coarse-to-fine processing [19].

The used method for face detection [6] is based on a tree-structured network of support vector machines designed for efficient computation. This hierarchy serves as a platform for pose modeling and efficient search of faces where most of a scene area is rejected efficiently using simple classifiers, while regions containing faces and face-like structures are rejected using more complex and dedicated classifiers.

A scene is processed at four different scales and at some locations. For each of those locations, we process the surrounding data in order to extract the 16 x 16 low frequency wavelet coefficients. The global hierarchical detector declares these coefficients as a face, if and only if, there is at least one complete chain of positive responses from the root classifier to a leaf classifier. The parameters of the face, i.e., the location, the orientation and the scale are given from the leaf cells.

3.2. ROIs Authentication

An authentication loop is introduced in order to give extra protection to the detected ROIs, which is similar to the

Fig 1. Face detection: left: original, right: after rotation

authentication loop we used for MPEG-4 [20]. The loop contains watermarks embedded into each ROI. The watermark information in each ROI contains the underlying identifier (ID), the total number of ROIs in the image (denoted as MaxID) and the current ROI position relative to the next ROI, e.g. a vector pointing to the next ROI. Any manipulation, like adding, moving or deleting a ROI, will break the authentication loop or make a leak and can be detected by verifying the loop's integrity.

Since the detected ROIs have quite similar visual contents, e.g. human faces, it is difficult to identify different ROIs by extracting unique features from ROIs. Therefore, in our scheme, we use different pseudo-random sequences to identify different faces, which are generated by secret keys. The watermark information is modulated first by the random sequence before embedding into faces. To ensure the face integrity, existing semi-fragile watermarking methods [4] can be used to embed the modulated sequence into the detected faces. Malicious manipulations of a face will destroy its watermark information and make it unauthentic.

A human face ROI is defined as the rectangle defined by the detected two eyes and mouth positions, as shown in Fig.1. Furthermore, in order to guarantee the robustness to rotation attacks, the watermark should be embedded along the pre-defined ROI orientation, e.g. the tilt of a face. Hence, before embedding watermarks into ROIs, the face rectangles should be normalized first.

3.3. Verification of ROIs and background

In order to verify the relationship between a specific ROI and the background, spread spectrum watermarks [21] are embedded into the background in a robust way, which are highly correlated to the corresponding sequences embedded into each ROI but independent from each other. This makes it impossible to copy a marked and authentic face into another image, but moving a face in the same image will not break its authenticity. Hence, every robust watermark should be embedded into the background by the order of starting from the areas near to the corresponding face and also along the pre-defined face orientation. This ensures the verification of the face location in case of cut & paste attacks in the same marked image and can also detect the false alarms of face detection as shown in Fig.1. The spiral embedding approach proposed in [22] can be applied.

In conclusion, we protect the membership of a face to its specific image by the relationship existing in the watermark between the specific face and the background, while we protect the integrity of faces in an image by the authentication loop. The combination of both protection mechanisms can detect the most common malicious manipulations.

4. DISCUSSION

The proposed approach addresses the issue of ensuring the integrity of selected image content, also after some cropping and/or geometric transformations. Different application fields are faced with this security challenge, e.g. medical databases, video surveillance, passport and identity control applications, visual communication, such as e-learning and video conferencing.

Using an automatic detection method for the identification of the faces as ROIs, we are not forced in storing ROI related information into the image. At the same time we can gain the synchronization points needed for the retrieval of the watermark. As shown in Fig.1, after slight rotation, the faces can still be correctly located. Thus, the watermark retrieval can be easily synchronized by normalizing the face rectangles again.

We briefly discuss the advantages of our approach, summarized as following:

- Multiple ROIs in an image: even though each detected face is considered a separate ROI, it is possible to identify more ROIs belonging to the same image up to the limit due to the capacity of the embedding algorithm. Their integrity can be checked separately and the spatial order of the ROIs can be verified.
- Automatic detection of ROI: from the practical point-of-view, this prevents the user from a tedious manual selection of the ROIs and makes the approach very attractive for large database applications. Furthermore, no-extra storage of ROI location and size are expected from the watermark, saving capacity and solving also the synchronization problem after possible geometrical transformations. Of course, the user can also manually select or mask faces, which are detected automatically.
- Secure for cut & paste attacks in the same image and from other images: since the embedded information provides a relationship between the ROI and its background, it is not possible to cut an entire face from an image and to paste it into the same or another one.
- Secure for adding and removal attacks: it is not possible to add or remove any face from an image, without breaking its authentication loop, due to the information embedded into the faces.

- Robust to rotation and cropping attacks: after these manipulations, it is still possible to verify the integrity of an image positively, if the face regions have not been manipulated.

The main limitations of the approach are:

- It is difficult to embed the watermarks in the background when the ratio between the size of the ROIs and the whole image is high. This can result for instance from close up faces.
- Robustness to image scaling depends on the applied watermarking algorithms, since the proposed authentication scheme is not intrinsically robust against scaling.
- Limited types of ROIs automatically detectable: It is known that deformable, fractal and complex ROIs containing more than one object are difficult to detect automatically.

5. CONCLUSION

We introduced in this paper an image authentication scheme tolerant to distortions caused by common image processing and based on automatic detection of regions of interest. The proposed scheme provides extra protection to the important ROIs of an image, for instance faces. The concept of authentication loop is introduced to verify the integrity of faces, so our approach detects changes, moving, adding and removal of faces from the same and/or another watermarked image combined with a robust embedding of watermarks in the background.

Finally, the proposed ROI watermarking can find lots of applications in medical databases, video surveillance, passport and identity control applications, visual communication, such as e-learning and video conferencing.

6. REFERENCES

[1] J. Dittmann, “Content-fragile Watermarking for Image Authentication”, Security and Watermarking of Multimedia Contents III, Proceedings of SPIE, Vol. 4314, pp. 175 - 184, 2001

[2] D.A. Winne, et al. “Digital Watermarking in Wavelet Domain with Predistortion for Authenticity Verification and Localization”, Proceedings of SPIE Security and Watermarking of Multimedia Contents IV, San Jose, California, vol. 4675, January, 2002.

[3] E. T. Lin, C. I. Podilchuk and E. J. Delp, “Detection of Image Alterations Using Semi-Fragile Watermarks”, Proc. of SPIE Inter. Conf. on Security and Watermarking of Multimedia Contents II, vol. 3971, Jan. 2000.

[4] H. Kang, J. Park, “A Semi-fragile Watermarking Using JND”, Proc. of STEG2003, pp.127-131, Japan, July, 2003.

[5] JPEG 2000 part 2 Final Committee Draft, ISO/IEC JTC1/SC20 WG1 N2000, December 2000.

[6] H. Sahbi, D. Geman, N. Boujemaa, “Face Detection Using Coarse-to-fine Support Vector Classifiers”, In Proc. of the IEEE Inter. Conf. on Image Processing, pp. 925–928, 2002

[7] M.C. Park, S.K. Han “A Digital Image Watermarking Using Region Segmentation”, International Conference on Circuits/Systems Computers and Communications (ITC-CSCC 2002), Phuket, Thailand, 2002

[8] P. Su, H. Wang, C.J. Kuo, “Digital Watermarking in Regions of Interest”, IS&T Image Processing/Image Quality/Image Capture Systems (PICS), Savannah, Georgia, April 25-28, 1999

[9] W. Lie, T. Hsu, G. Lin, W. Ho, “Fragile Watermarking for JPEG-2000 Images”, 16th IPPR Conf. on Computer Vision, Graphics and Image Processing (CVGIP 2003), pp. 823 – 826, August. 2003.

[10] H. Frigui and R. Krishnapuram, “Clustering by Competitive Agglomeration”, In Pattern Recognition, 30(7), 1997.

[11] C. Carson, M. Thomas, S. Belongie, J. Hellerstein, J. Malik, “Blobworld: A System for Region-based Image Indexing and Retrieval”, Third Inter. Conf. on Visual Information Systems (Visual’99), 1999.

[12] B. Moghaddam, A. Pentland, “Probabilistic visual learning for object recognition”, In IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):696–710, 1997.

[13] H. Schneiderman, T. Kanade, “Object Detection Using the Statistics of Parts”, International Journal of Computer Vision, 2002.

[14] T. Leung, M.C. Burl, P. Perona, “Finding Faces in Cluttered Scenes Using Random Labelled Graph Matching”, In Proceedings of the International Conference on Computer Vision, pages 637–644, 1995.

[15] Y. Lamdan, J.T Shwartz, H. J. Wolfson, “Object Recognition by Affine Invariant Matching”, In Proc. of Computer Vision and Pattern Recognition, pp. 335–344, 1998.

[16] H. Rowley, S. Baluja, T. Kanade, “Neural Network-based Face Detection”, In IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(1):23–38, 1998.

[17] E. Osuna, R. Freund, F. Girosi, “Support Vector Machines: Training and Applications”, CBCL Paper N 144/AI Memo N 1602,Massachusetts Institute of Technology, Cambridge, MA., 1997.

[18] R.L. Hsu, M. Abdel-Mottaleb, A.K. Jain, “Face Detection in Color Images”, In Proceedings of the IEEE Inter. Conference on Image Processing, pages 1046–1049, 2001.

[19] F. Fleuret, D. Geman, “Coarse-to-fine Visual Selection”, In Inter. Journal of Computer Vision, 41(2):85–107, 2001.

[20] A. Lang, S. Thiemert, E. Hauser, H. Liu, F.A.P. Petitcolas, “Authentication of MEPG-4 Data: Risks and Solutions”, Proc. of SPIE Security and Watermarking of Multimedia Contents V, pp. 453 - 461, 2003.

[21] I.J. Cox, et al. “Secure Spread Spectrum Watermarking for Multemedia”, IEEE Trans. Image Processing, Vol. 6, No. 12, Dec. 1997.

[22] A. Wakatani: “Digital Watermarking for ROI Medical Images by Using Compressed Signature Image”; 35th Hawaii International Conference on System Sciences (HICSS-35), Island of Hawaii, January 7-10, 2002.