COMBATING GEOMETRICAL ATTACKS IN A DWT BASED BLIND VIDEO WATERMARKING SYSTEM

C.V. Serdean[1], M.A. Ambroze[1], M. Tomlinson[1] and G. Wade[2]

[1] Department of Communication & Electronic Engineering, University of Plymouth, Plymouth, PL4 8AA, UK
{C.Serdean | M.Ambroze | M.Tomlinson}@plymouth.ac.uk
[2] Department of Electrical & Computer Engineering, University of Newcastle, Callaghan, NSW 2308, Australia
gwade@hartley.newcastle.edu.au

Abstract: *This paper describes a high capacity blind video watermarking system invariant to geometrical attacks such as shift, rotation, scaling and cropping. A spatial domain reference watermark is used to obtain invariance to geometric attacks by employing image registration techniques to determine and invert the attacks. A second, high capacity watermark, which carries the data payload, is embedded in the wavelet domain according to a human visual system (HVS) model. This is protected by a state-of-the-art error correction code (Turbo code). The proposed system is invariant to scaling up to 180%, rotation up to 70^0, and arbitrary aspect ratio changes up to 200% on both axes. Furthermore, the system is virtually invariant to any shifting, cropping, or combined shifting and cropping.*

Key words: *Wavelet, Video watermarking, Geometrical invariance, Fourier-Mellin transform*

1. INTRODUCTION

One of the most difficult problems in digital video watermarking is watermark recovery in the presence of geometric attacks like frame shift, cropping, scaling, rotation, and change of aspect ratio, especially when some of these are combined together. Although for uncompressed video the geometric attacks tend to be less severe compared to those for image watermarking, the recovery problem is compounded for video since it must be carried out blind due to the difficulty of storing the original. In this case, for the typical spread spectrum (SS) watermarking system, blind retrieval is performed via cross-correlation between the marked video and the secret pseudo-noise (PN) sequence used to spread the watermark at the embedding stage. Because geometric attacks destroy the

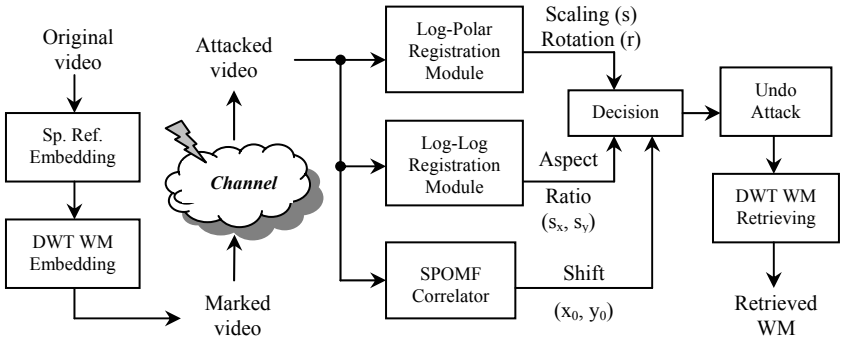

Fig. 1. Block schematic of the geometric invariant video watermarking system

synchronisation, some form of sliding correlation must be used in order to ensure resynchronisation i.e. multiple cross-correlations over a specified search space. Unfortunately the search space grows very quickly, making it difficult to recover the watermark in a reasonable time. Given that retrieval in a video context must be done in near real time, the computational problem is very significant. One way to overcome this is to use image registration techniques for resynchronisation.

This paper combines the advantages of an algorithm based on the Fourier-Mellin transform (FMT) image registration techniques [2], with watermarking in the Discrete Wavelet Transform (DWT) domain. The idea is to first undo geometric attacks using the FMT approach and an additional spatial reference watermark used only for registration purposes. Once the attack parameters are determined, the geometric attacks are undone and the resulting frame is passed to the main watermark decoder. The main watermark, which carries multi-bit data, is inserted in the DWT domain and capacity is maximised by embedding based on a HVS model. The complete system can be regarded as a noisy communications channel and so is protected by turbo coding. The net result is a system that can withstand severe geometric attack, the limiting attack being defined by a threshold yielding a false detection probability of 10^{-8}, and capacity being defined by a BER of 10^{-8}.

2. COMBATING GEOMETRIC ATTACK USING LOG-POLAR AND LOG-LOG MAPPING

When registering two images, the noise is relatively small, and so the correlator usually performs very well. The problem is more difficult for video watermarking since the original video frame is not available. To overcome this, we use a SS reference watermark. The PN sequence used to embed this watermark corresponds to the "original image" and the watermark embedded in the unsynchronised marked video corresponds to a noisy "attacked image" (the video itself represents the noise). In the proposed system we embed two different watermarks. The first is a 1-bit watermark used only for geometric reference, and is embedded in the spatial domain. The second, multi-bit watermark is used for the data payload, and is embedded in the DWT domain.

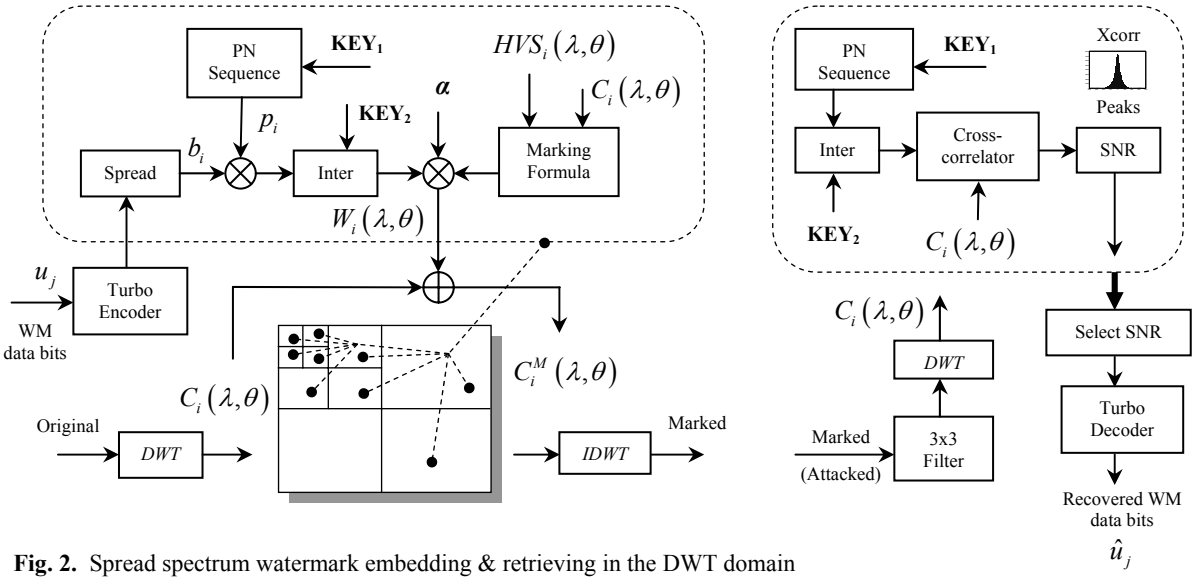

Fig. 2. Spread spectrum watermark embedding & retrieving in the DWT domain

The geometric invariance can be achieved by using the FMT to convert rotation and scale to spatial shifts [2], which are then easily recovered by a Symmetrical Phase Only Matched Filter (SPOMF) [3]. A log-polar map permits recovery over a wide range of scale changes, rotation, or even combined scale-rotation attack. If a log-log map is used, then it is possible to recover arbitrary aspect ratio changes.

Fig. 1 shows a schematic of the proposed system. The decision block determines if the reference watermark is present (to within a desired false detection probability), and if present it automatically determines the attack parameters. The two watermarks are embedded in different domains in order to minimise crosstalk, and each watermark is embedded at the full strength dictated by its own HVS model. The reference watermark is embedded in the spatial domain using SS, together with a simple visual model that inserts a stronger watermark in those regions where it is less easily observed. The same reference watermark is embedded in all the frames in order to increase the SNR at the correlator input via frame averaging. As a result, the registration takes place only once, and not for each separate frame. This is possible because attacks must be identical for each frame in order to avoid temporal artefacts.

3. THE DWT VIDEO WATERMARKING SCHEME

The hierarchical property of the DWT offers the possibility of analysing a signal at different resolutions (levels) and orientations. This multiresolution analysis gives both space and frequency localisation, and different orientations extract different features of the frame, such as vertical, horizontal, and diagonal information. Generally speaking, edges and textures will be represented by large coefficients in the high frequency sub-bands, and they are well localised within the sub-band. For watermarking, we selected the Antonini 7.9 wavelet, as being one of the best wavelets available for image compression [4]. Few advantages of DWT are: the multiresolution property provides both local and global spatial support, it is compatible with the HVS, there are no blocking artefacts, and it has lower computational cost and better energy compaction properties than the FFT and DCT. The information capacity of the channel is maximised by embedding the main payload according to an HVS model, and through the use of turbo coding. The embedding and retrieving of the watermark are shown in Fig. 2. The security of such a system relies in the secret watermarking keys, K_1 and K_2.

The hierarchical nature of the DWT is exploited by inserting a self-contained watermark in each sub-band, i.e. all payload bits are inserted into each sub-band. The watermark is embedded using amplitude modulation [3, 4]. This marks more heavily the high frequency

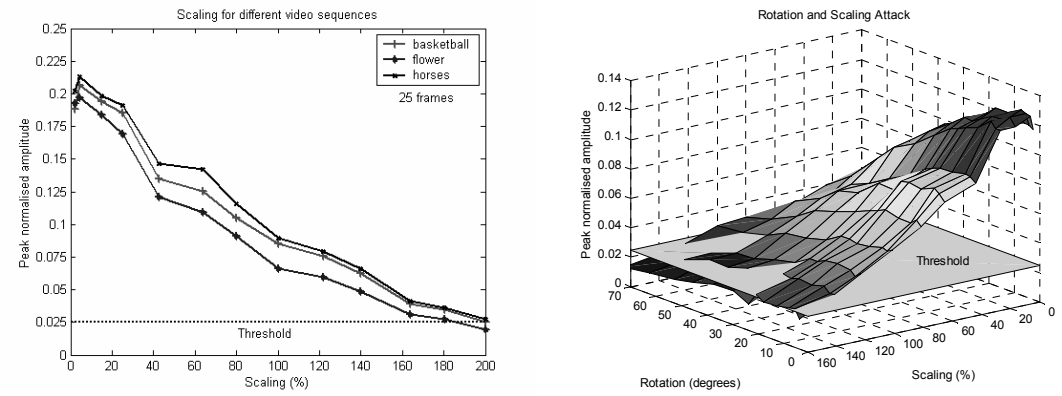

Fig. 3. Performance of the system for scaling (a) and for rotation combined with scaling attack (b)

sub-bands and the largest coefficients, since modification of these coefficients is less likely to incur visible artefacts and incorporates HVS model and media dependence [3, 4].

For retrieval, the video sequence is filtered prior to correlation in order to improve the performance of the correlator. It is advantageous to have a self-contained watermark in each sub-band, since a SNR can be determined for each sub-band as an indicator of sub-channel quality. Different types of attack affect different levels and orientations in different ways, and so it is always possible to select an optimal sub-band via SNR. Correlation is therefore performed separately for each sub-band, obtaining a set of cross-correlation peaks (one peak for each embedded data bit) for each sub-band. A SNR is then computed for each set of cross-correlation peaks, and retrieval is carried out for the sub-band with the highest SNR.

4. RESULTS

Fig.3a shows how the system performs for different percentages of scaling, when 25 frames are averaged in order to improve the robustness of the system. Slightly better results are obtained in case of rotation. The threshold value of 0.025 which can be observed in both figures, guarantees a false detection probability better than 10^{-8} when the correlation peak exceeds it [3]. For a false detection probability of 10^{-8}, the proposed system is invariant to scaling in the range –50% to 180%, to rotation up to 70^0, and to any arbitrary aspect ratio changes in the range –100% to 200% on both axes. Furthermore, the system is virtually invariant to any shifting, cropping, or combined shifting and cropping. When rotation is combined with scaling (Fig.3b), up to 120% scaling and up to 20^0 rotation can be tolerated. The system therefore exceeds the EBU recommendation [1] for all these attacks.

5. CONCLUSIONS

Robustness to geometric attack is one of the most important requirements for a watermarking system, and an approach based on the FMT and log-polar/log-log representations of the video frames has been developed. This is combined with the advantages of the DWT, HVS-based marking, and turbo coding to produce a very robust, high capacity video watermarking system. It outperforms many current schemes in terms of geometric invariance and channel capacity.

REFERENCES

[1] Cheveau, L., Goray, E. and Salmon R. 'EBU Technical Review – March 2001'.

[2] Reddy, B.S. & Chatterji, B.N., 'An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration', *IEEE Trans. Image Proc.,* Vol.5, No.8, August 1996.

[3] Serdean, C.V., Ambroze, M.A., Tomlinson, M., & Wade, G., 'DWT Based High Capacity Blind Video Watermarking, Invariant to Geometrical Attacks', submitted to *IEE Proc. Vision, Image & Signal Processing*, in December 2001.

[4] Serdean, C.V., Ambroze, M.A., Tomlinson, M., & Wade, G., 'Protecting Intellectual Rights: Digital Watermarking in the Wavelet Domain', *IEEE Int. Workshop 'Trends & Recent Achievements in IT'*, 16-18 May 2002, Cluj-Napoca, Romania, Invited Paper.