



Robust Video Watermarking of Hybrid Based Techniques

Salam Abdunabi Thajeel

Computer Science Department, College of Education, University of Al-Mustansiriyah, Baghdad, Iraq.

Abstract

This paper adopted a new video watermarking scheme based on hybrid transforms to achieve the watermarking requirements, that is, robustness, imperceptibility, and security. Where In our system first of all, the original video will be divided into frames to transform it using the Slantlet Transform into four sub-bands (i.e., LL, LH, HL, and HH). Secondly, The HL sub-band will be chosen for further decomposition using Contourlet Transform (CT), and then the low sub-band of CT will be selected to decompose it to fixed size non overlapping blocks to employ the DCT (Discrete Cosine Transform) on each block. lastly, to improve the security and robustness, the watermark logo is scrambled using AT(Arnold transformation) as a prelude to embed it's on the each transformed non-overlapping blocks. Experimental results show that the proposed system achieves good imperceptibility and high resistance against various attacks.

Keywords: Digital Image Watermarking, Contourlet transform, Arnold transform, Copyright protection.

علامة مائية موثوقة للفيديو اعتمادا على تقنيات هجينه

سلام عبد النبي ثجيل

قسم علوم الحاسبات ، كلية التربية، الجامعة المستنصرية، بغداد، العراق.

الخلاصه

في بحثنا هذا اعتمدنا طريقة جديدة لغرض تضمين العلامة المائية داخل الفيديو بالاعتماد على عدة تحويلات لغرض تحقيق متطلبات العلامة المائية والتي هي المتانة، وعدم التحيز، والأمن. حيث في نظامنا أولاً، سيتم تقسيم الفيديو الأصلي إلى مجموعه من الاطارات Frame ومن ثم باستخدام تحويل سلانليت(SLT) يتم تحويلها إلى أربعة نطاقات فرعية(LL, LH, HL, HH). ثانياً، سيتم اختيار النطاق الفرعي HL لتنفيذ المزيد من التحلل باستخدام تحويلات الكونتورليت (CT). ثم سيتم اختيار النطاق الفرعي المنخفض من CT لتحليله إلى كتل ثابتة غير متداخلة الحجم لاستخدام تحويل جيب التمام المنفصل (DCT) على كل كتلة. وأخيراً ولغرض تحقيق امان ومتانة افضل يتم تشفير العلامة المائية باستخدام تحويل Arnold قبل ان يتم تضمينها على الكتل غير متداخلة. وتظهر النتائج التجريبية أن النظام المقترح يحقق قابلية حساسة جيدة ومقاومة عالية ضد مختلف الهجمات. اظهرت النتائج العملي أن النظام المقترح يحقق شفافية افضل ومقاومة عالية ضد مختلف انواع الهجمات

1. Introduction

In today's digital era, images and videos serve as the primary vehicle for information transfer. Digital multimedia are highly capable of easy articulation, dissemination, and storage, thereby making

them avenues for information transfer, especially with the advent of information technology and communication devices[1].

At present, the availability of graphical editing software eases the duplication or tampering of any digital media by illegitimate users. Therefore, copyright and content protection have become a critical issue in the digital world[2]. Digital watermarking is a widely used protection technique where confidential information which are called a watermark is insert into any types of authentic multimedia such as: images, video, text, and audio. The techniques of Watermarking can ordinarily categorized into four different types according on the data types to be watermarked: (i) text, (ii) image, (iii) video, and (iv) audio[3].

With the technological advancements, relevant studies have dealt with image and video watermarking[4]. In addition, the proposed techniques for digital watermarking can be categorized to two non-identical types based on the domain type used to hide the watermark in a host image: spatial domain and transform (frequency) domain. Watermarking is more expeditious in the spatial domain than in the transform domain. However, the spatial domain is not robust and is prone to any attacks occur on image. By contrast, the transform domain is most durability against image processing attacks and is more capable of embedding watermarks [5]. Thus, frequency domain watermarking, which is more capable of withstanding image and video processing and geometrical attacks, was adopted to overcome the flaws of spatial domain techniques[3]. The most frequently utilized transforms are Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), and Discrete Wavelet Transform (DWT) [5].

Lee et al. [6] introduced a new video watermarking technique depend on 3D DFT in which an image's perceptual model in the DFT domain is designed and used to insert video watermarking. This method achieves good imperceptibility.

Sun et al.[7] Proposition a DCT-based on video watermarking mechanism that uses low-frequency AC coefficients in each DCT block to insert watermarking. Such method is prone to noise-based attacks. PredaandVizireanu [4] proposed a DWT-based watermarking technique. After the decomposition of DWT, the binary watermark is inserted in wavelet coefficients of the midst wavelet sub-bands. This method demonstrates high perceptual quality and high resistance to video processing attacks. Many algorithms have been developed recently; several hybrid techniques combined more than one transform to improve the performance of watermarking methods [5, 8].

An imperceptible and robust video watermarking scheme that utilizes DWT and Principal Component Analysis (PCA) was also developed. Where after the application of the wavelet DWT, watermarking was embedded in the principal components of low-frequency wavelet coefficients. Although this technique can withstand spatial attacks, it is prone to temporal attacks, such as frame swapping, dropping, and removal[9].

Agarwal et al. [10] introduced a novel video watermarking scheme that combines DWT and DCT. The DCT will be applied on the midst-frequency bands (LH and HL) of the wavelet transform. The watermarking will then be embedded in middle-frequency components of the DCT. However, this method requires the original watermark during extraction.

Therefore, this study proposes a robust video watermarking that adopts the features of Slantlet Transform (SLT), Contourlet Transform (CT), Arnold transform (AT), and DCT to achieve high robustness, imperceptibility, and security.

The remainder of this paper is organized like follows. Section 2 presented the theoretical background of AT, SLT, CT, and DCT. Section 3 include discusses the steps of the embedding and extraction processes of the proposed scheme. Section 4&5 describes the experimental results with comparisons versus the schemes in previous works and the conclusions .

2. Preliminaries

This section presents different terminologies used in the proposed method.

2.1 Arnold Transform

AT, also called Arnold's cat map, is a scrambling technique that is widely used to alter an image into an entirely different and senseless image. The pixel locations are changed; if the pixel locations are changed several times, the image would appear messy. Thus, it is widely used to hide information [11]. Similar to [8,10,12], in our system will be also used AT. For an image with size $N \times N$, the AT can be obtained using the following equation[11]:

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \text{mod} M \quad (1)$$

Where M indicates the image size (watermark image), (x, y) denotes the pixel coordinate in the original image, and (\hat{x}, \hat{y}) represents the coordinates after scrambling the image. The original image from scrambling image can then be obtained by applying the inverse AT as follows[11]:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} \text{mod} M \quad (2)$$

Where (x, y) represents the coordinates of the descrambled image, (\hat{x}, \hat{y}) refers to the pixel coordinate of the scrambled image, and M represents the image size.

2.2 Contourlet Transform

CT is a geometric transform of digital image [13]. Unlike DWT, CT can effectively represent smooth contours in different directions of a frame or image. This process is done by using a double-filter bank, namely, Laplacian pyramid (LP), and a directional filter bank (DFB). LP selects the point discontinuities, while DFB is used to connect the point discontinuities to linear structures. Figure- 1 shows that LP decomposes an input image into low-pass and band-pass images, whereas DFB decomposes each band pass image. Moreover, CT obtains 2^N directional sub-bands after each level. For example, eight sub bands (2^3) are obtained when an input image is decomposed into three levels. CT was developed as an amelioration over wavelet[14]. Previous studies adopted this transform[15-17], because it is more capable of picking up smooth contours and directional edges compared with other transforms. In the current study, the low-pass image of the Contourlet decomposition was used to embed the watermarking and thereby achieve robustness[17].

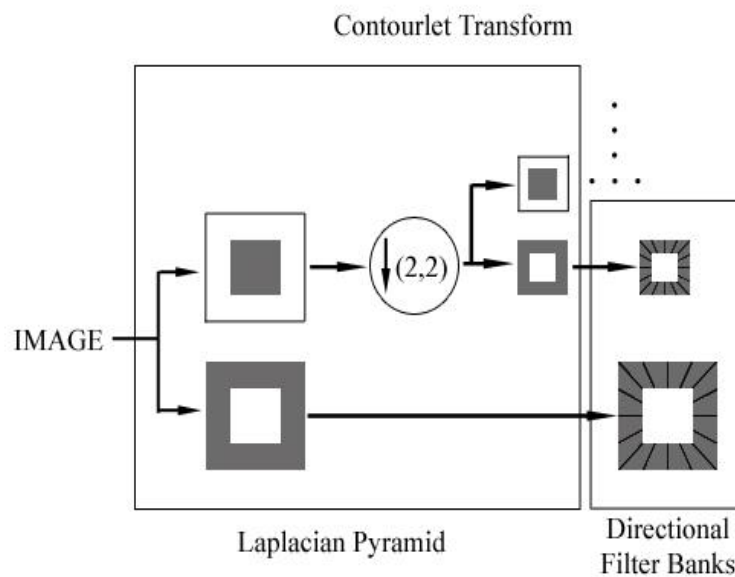


Figure 1-Block diagram of the CT double-filter bank. First the input image is split into a low pass & a band pass image by Laplacian pyramid, and then each band pass is further decomposed by DFB[17].

2.3 Slantlet Transform

SLT is an amended version of DWT with two-zero moments and time localization improvement[18]. It is used as a filter bank that has a parallel structure. It uses parallel processing in which different filters have been configured for each scale alternatively of filter repetition for each level. The length of SLT filters is less than those of an equivalent DWT. Thus, SLT has been used in several applications and has achieved the best results compared with the other schemes. Like DWT, SLT decomposition of an image in a 2D environment is divided into four parts: Low-Low (LL), Low-High (LH), High-Low (HL), and High-High (HH), as shown in Figure-2. L and H signify the low- and high-frequency bands, respectively. Each band carries a different image information. The image's low-frequency band component marked as LL keeps the original image information. By contrast, medium- and high-frequency bands, namely, LH, HL, and HH, carry information related to the edge,

contour, and other details of the image. Important information in the image is represented by high coefficients. Thus, the HL decomposition is used to embed the watermarking to achieve high robustness[19].

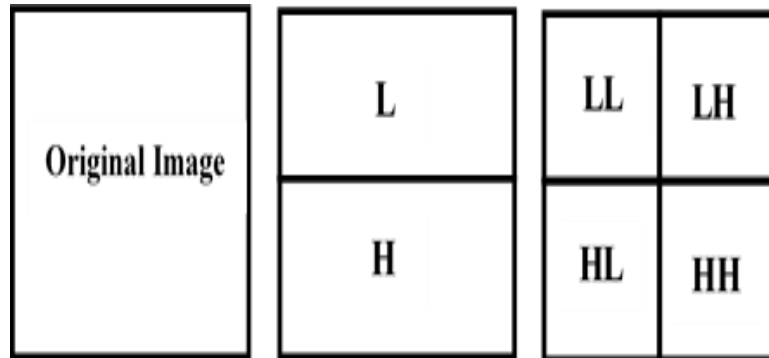


Figure 2- 2D STL decomposition schemes. The input image is first broken into low and high frequency bands, subsequently the output further separated into four parts LL, LH, HL and HH respectively[19].

2.4 Discrete Cosine Transform

DCT translates an image into its frequency components. In image processing, DCT de-correlates image data. DCT can set the image data into the maximum possible small number of DCT coefficients without any distortion. It can also maintain separability and symmetry. A 2D DCT of an input image may be defined via the next equation:

$$C(u,v) = a(u)a(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos \left[\frac{\pi(2x+1)u}{2N} \right] \cos \left[\frac{\pi(2y+1)v}{2n} \right]. \quad (3)$$

Where U & v denote to discrete frequency variables of (x, y) pixel index, while n refer to size of input image (8 in this study). DCT divides the image into three frequency bands, that is, high, low, and middle frequency. Low- and high-frequency components are ticklish for human visual systems and can withstand various attacks, respectively to [20]. ADCT's middle-band frequency will prevent watermarking from being influenced by compression and noise attacks. Therefore, using middle-frequency bands implies a compromise between robustness and imperceptibility[20], [21].

3. Proposed System

The proposed scheme for the video watermarking compose of two main steps which are: watermark embedding and extraction. In the subsequent subsections we will illustrated these two steps.

3.1 Embedded watermarking based on SLT, CT, and DCT

The next steps represent the proposed watermark embedding algorithm:

- The original video will be partition to frames. Subsequently, the SLT matrix will be applied on every frame to transformed it's to four sub-bands (i.e., LL, LH, HL, and HH). after that the sub-band (HL) will be then selected for purpose of implementing additional decomposition utilize CT also
- A first-level CT will be applied on the sub-bands selected through the prior stride. Where the low sub-band (directional sub-bands) of the CT will be used to embed the watermark on it.
- The selected low sub-band of CT will be splitted to 8×8 fixed partition scheme of the non-overlapping blocks. DCT will then be performed for every block.
- The grayscale watermarking will be scrambled using AT for N times (where N represents the secret key).The watermarking will then be reconstructed to the vector consisting of zeros and ones.
- Two unconnected pseudorandom series (i.e., PN_S_0 and PN_S_1) will be created. The first pseudorandom series helps in embedding the bits "0" of the scrambled watermarking image, whereas the second pseudorandom series will be used to ingrain the watermark bits"1". The pseudorandom series total size is equivalent to the elements quantity in the midst-frequency band of the 8×8 DCT blocks.

- In the middle-band frequency of the 8×8 DCT blocks, the coefficients of the selected sub-band will be modified by inserting two random sequences (i.e., PN_S_0 and PN_S_1) with the gain factor k depending on the following equations:

$$f(X') = \begin{cases} 0, & X' = X + \alpha \times PN - 0 \\ 1, & X' = X + \alpha \times PN - 1 \end{cases} \quad (4)$$

Where X refer to the coefficient of the midst -frequency band of the 8×8 DCT blocks, and X' indicated to the matrix coefficient of midst-frequency band of the DCT when watermark was inserting. To achieve the preferably tradeoff between both of visibility and robustness, an appropriate gain factor must be selected.

- Upon accomplishment of the embedding operation, reverse DCT will be performed on all 8×8 blocks.
- ICT is performed utilizing the elected sub-bands from each level, followed by the executed of inverse SLT. All the subsequent watermarked frames obtained will then be combined to achieve a secure watermarked video.

3.2 Watermark extraction based on SLT, CT, and DCT

The algorithm used to extract watermarking Composed from the following strides:

- SLT will be applied to decompose the watermarked video obtained from the embedding algorithm to obtain four non-overlapping multi-resolution sub-bands: LL, HL, LH, and HH. The HL sub-band will then be chosen for further decomposition using CT.
- CT will be applied to the chosen sub-bands in the previous stride. Directional sub-bands will be chosen to extract the watermark from it.
- The directional sub-bands will be divided into 8×8 blocks. DCT will be applied on each block.
- The same technique used during watermark embedding will be used to create two types of the pseudorandom series, namely, PN-0 and PN-1.
- The correlation amount between the midst, -frequency band of the DCT block will be calculated and compared with the random sequences (i.e., PN0 and PN1). If a correlation amount with PN-0 is more than PN-1, this mean watermark bits will be "0". Else, the watermark bits will be "1".
- To reconstruct the original watermarking image using the extracted watermark bits this is done by applying the inverse AT for N times (where N is a period of AT).

4. Experimental Results and Comparative Analysis

For performance evaluation, the simulation of the proposed method of video watermarking system was carry out through MATLAB 2016a. Thus, several experiments were conducted using five samples with the same dimensions (i.e., akiyo, foreman, salesman, container, Miss_am) as an original test cover video. The first frame of these samples is presented in Figure-3. The performance of the proposed watermarking algorithm is evaluated using three standard measurements :Peak Signal to Noise Ratio (PSNR), Mean Squared Error (MSE) for measuring imperceptibility, and the Normalized Cross Correlation (NCC) for determining robustness[6-8].

NCC is a metric commonly used to evaluate the degree of similarity (or dissimilarity) between two images or frames [22]. It measures the correlation between the original and the extracted watermark after an attacks. Ideally, the NCC value should be equal to 1[7], which indicates similarity between the original and extracted watermarks. The NCC value is usually agreeable even when it is equal or more than 0.75. NCC can be obtained using the following equation[7] :

$$NCC = \frac{\sum_{p=0}^{N-1} \sum_{q=0}^{M-1} w(m, n) \times \tilde{w}(m, n)}{\sqrt{\sum_{p=1}^{N-1} w(m, n) \times w(m, n) \quad \sum_{q=0}^{M-1} \tilde{w}(m, n) \times \tilde{w}(m, n)}} \quad (5)$$

Where $w(m, n)$ and $\tilde{w}(m, n)$ represent the original and watermarked frames of videos, respectively. MSE is quantified parameter that measures the error value after averaging the squares of two images. MSE can be obtained using the following

$$MSE = \frac{1}{m * n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (6)$$

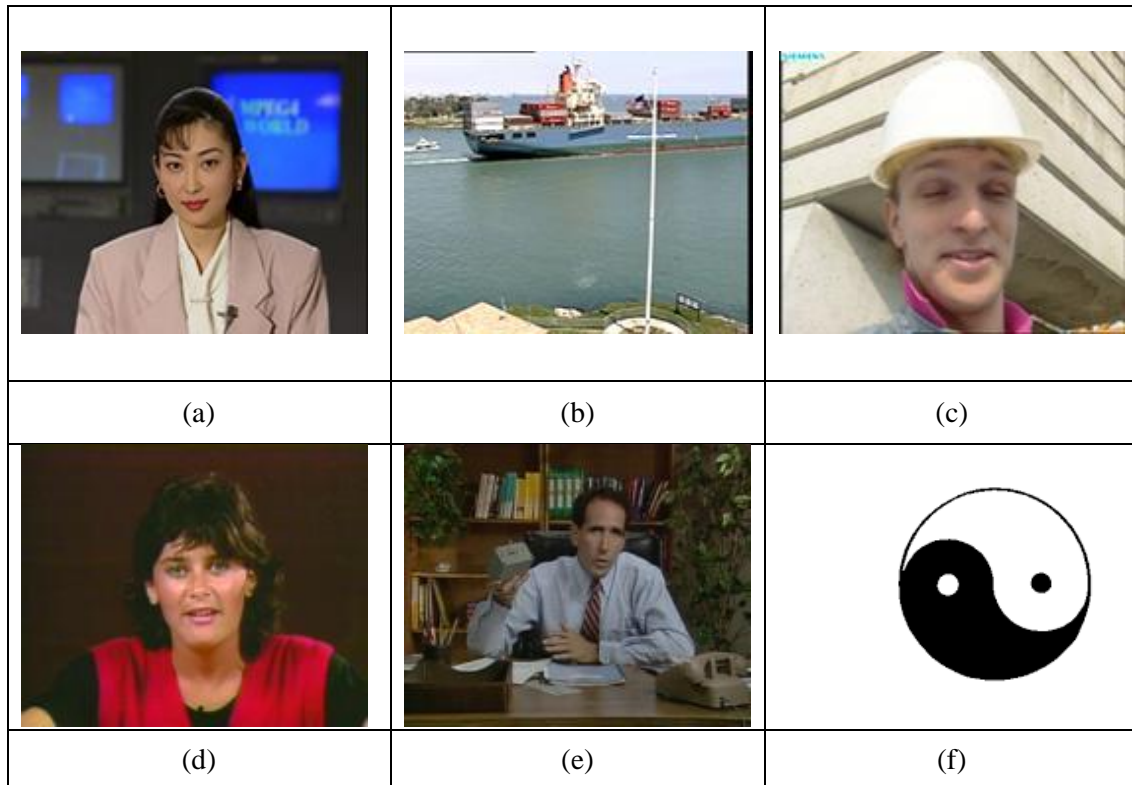


Figure 3- (a-e)First frames of the original videos, and (f) original watermark.

Where I and K represent the original and watermarked frames of videos, respectively. MSE is inversely related to PSNR[23]. Both parameters are determined between original and watermarked frames to check for the degradation caused by watermarking. PSNR refers to the ratio of the maximum value of a frame or image's pixels and the MSE. It is commonly used to measure the quality of a reconstructed watermarked frame or image.

The proposed algorithm was validated against various image processing attacks that ordinarily happen in a frame of a video. where Gaussian noise ,salt and pepper noise, median filtering, sharpening as well as histogram equalization is used for the analysis of robustness performance of proposed system by exposing all video watermarking frames with the same attack .The performances are summarized in Table-1.

Moreover, the performance of the algorithm versus different temporal attacks was measured. Where the temporal operations, such as frame averaging, frame swapping and fram removal have ability to undermine efficiency the video watermarking through destruction of video synchronization. Therefore, although the visual quality of the video is not affected by this type of attack, this type of attack causes the destruction of the watermark. So it is advisable to be evaluate any proposed system of video watermarking versus temporal attacks[24].Table-2 represent the experimental results in case of temporal attacks

Table 1- Proposed video watermarking against frame based attacks

Type of attack	NC	MSE	PSNR
Without attack	0.9952	0.8472	48.456
Gaussian noise (0 mean and 0.01 variance)	0.9765	72.974	29.3217
Salt and pepper noise (density 0.05)	0.9834	48.7630	30.6753
Gamma correction	0.9233	203.76	24.856
Median filtering (3×3)	0.9785	1.712	46.1327
Sharpening	0.9342	10.023	37.5745
Histogram equalization	0.9712	46.45	21.876

Table 2- Proposed video watermarking against temporal attacks

Type of attack (10 %)	NC	MSE	PSNR
Frame averaging	0.9323	22.645	31.5732
Frame swapping	0.9545	1.532	46.231
Frame removal	0.9478	0.8215	47.9603

Tables-(1, 2) shows that the proposed watermarking algorithm's indices are highly robust against almost all types of attacks. Its performance will decline in the presence of histogram equalization and gamma correction. Nonetheless, the overall performance of the algorithm is excellent. The results of the proposed scheme are compared with those of relevant works on spatial and temporal domain attacks, as shown in Tables-(3,4), respectively. Where compared our proposed method with three distinct watermarking system found in literature.

From Table-3 It easy to notice clearly that the our system has the highest NC value comparing with NC values of the all methods exists when the video watermarking exposed to Gaussian noise ,salt and pepper noise ,median filtering and histogram equalization attacks . From Table-4, yet again, the proposed system performed superbly even in case the video watermarking have been exposed to temporal domain attacks. The conducted experimental results refer the proposed algorithm outperforms the other schemes and has better NC against all types of attacks

Table 3- Comparison of NC for proposed video watermarking with previous work for spatial domain attacks

Spatial attacks	Proposed work	Deshpande et al. [24]	Agarwal et al. [10]	Sinha et al.[9]
Without attack	0.9952	0.9658	0.9785	0.9680
Gaussian noise (0 mean and 0.01 variance)	0.9765	0.8993	0.9154	0.9250
Salt and pepper noise (density 0.05)	0.9834	0.9156	0.9357	0.9250
Median filtering (3×3)	0.9785	0.9345	0.9526	0.9490
Histogram equalization	0.9712	0.9456	0.9563	0.9497

Table 4- Comparison of NC for proposed video watermarking with previous work for temporal domain attacks

Temporal attacks	Proposed work	Deshpande et al. [24]	Agarwal et al. [10]	Sinha et al. [9]
Frame swapping	0.9856	0.9634	0.9796	0.9556
Frame removal	0.9674	0.8756	0.8863	0.7958

5. Conclusion

This study proposed a new video watermarking scheme depend on SLT, CT, and DCT. Where the main contribution to the proposed method is the linkage between security, robustness and imperceptibility, this has been done by the combined attributes of all transformation techniques which have been used. Where the use of the higher decomposition level of the SLT with directional sub-bands of the CT led to increases the imperceptibility of the suggested system, resulting in better performance in terms of robustness.as well as use of DCT is useful in terms of compaction and energy, which helped to decrease the invisibility feature during watermark embedding. In order to improve the security and robustness, the watermark was first scrambled using AT and then inserted it as a spread spectrum pattern. According to the results of the experiments obtained, which were evaluated using of NCC, MSE and PSNR, as well as compared with previous works, it easy to conclude that Our approach gives encouraging results versus various types of attacks (spatial and temporal).Consequently, the proposed system demonstrated it is useful and appropriate for both of copyright protection well as to content authentication applications

References

1. Thajeel, S.A.-N. and Sulong, G. **2015**. A novel approach for detection of copy move forgery using completed robust local binary pattern. *J Inf Hiding Multimed Signal Process*, 2015. **6**(2): 351-364.
2. Shao, Z., et al., **2015**. Combining double random phase encoding for color image watermarking in quaternion gyrator domain. *Optics Communications*. **343**: 56-65.
3. Langelaar, G.C., I. Setyawan, and R.L. Lagendijk, R.L. **2000**. Watermarking digital image and video data. A state-of-the-art overview. *IEEE Signal Processing Magazine*, **17** (5).
4. Preda, R.O. and D.N. Vizireanu, D.N. **2010**. A robust digital watermarking scheme for video copyright protection in the wavelet domain. *Measurement*, **43**(10): 1720-1726.
5. Niu, X. and Sun. S. **2000**. A new wavelet-based digital watermarking for video. in 9th IEEE digital signal processing workshop. Citeseer.
6. Lee, Y.-Y., H.-S. Jung, and S.-U. Lee. **2003**. 3D DFT-based video watermarking using perceptual models. in Circuits and Systems, *2003 IEEE 46th Midwest Symposium on*. IEEE.
7. Sun, T., et al. **2009**. A novel differential energy video watermarking based on Watson visual model. in Electronic Commerce and Security, ISECS'09. Second International Symposium on. 2009. IEEE.
8. Makbol, N.M. and Khoo, B.E. **2013**. A Hybrid Robust Image Watermarking Scheme Using Integer Wavelet Transform, Singular Value Decomposition and Arnold Transform. in International Visual Informatics Conference. 2013. Springer.
9. Sinha, S., et al. **2011**. Digital video watermarking using discrete wavelet transform and principal component analysis. *International Journal of Wisdom Based Computing*, **1**(2): 7-12.
10. Agarwal, A., R. Bhadana, and S. Chavan, S. **2011**. A robust video watermarking scheme using DWT and DCT. *IJCSIT) International Journal of Computer Science and Information Technologies*, **2**(4): 1711-1716.
11. Arnold, M., et al. **2014**. A phase-based audio watermarking system robust to acoustic path propagation. *IEEE Transactions on Information Forensics and Security*, **9**(3): 411-425.
12. Saikrishna, N. and Resmipriya, M. **2016**. An Invisible Logo Watermarking Using Arnold Transform. *Procedia Computer Science*, 2016. **93**: 808-815.

13. Do, M.N. and Vetterli, M. **2005**. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on image processing*, 2005. **14**(12): 2091-2106.
14. Po, D.-Y. and M.N. Do. **2003**. Directional multiscale modeling of images using the contourlet transform. in *Statistical Signal Processing, 2003 IEEE Workshop on*. IEEE.
15. Su, Q., et al. **2016**. A novel blind color image watermarking based on Contourlet transform and Hessenberg decomposition. *Multimedia Tools and Applications*, 2016: 1-21.
16. Khalighi, S., P. Tirdad and Rabiee, H. **2010**. A contourlet-based image watermarking scheme with high resistance to removal and geometrical attacks. *EURASIP Journal on Advances in Signal Processing*, **2010**(1): 1.
17. Akhaee, M.A., S.M.E. Sahraeian, and F. Marvasti, *Contourlet-based image watermarking using optimum detector in a noisy environment*. *IEEE Transactions on Image Processing*, 2010. **19**(4): 967-980.
18. Selesnick, I.W. **1999**. The slantlet transform. *IEEE transactions on signal processing*, 1999. **47**(5): 1304-1313.
19. Mohammed, R.T. and Khoo, B.E. **2012**. Image watermarking using slantlet transform. in *Industrial Electronics and Applications (ISIEA), 2012 IEEE Symposium on*. 2012. IEEE.
20. Deb, K., et al. **2012**. Combined DWT-DCT based digital image watermarking technique for copyright protection. in *Electrical & Computer Engineering (ICECE), 2012 7th International Conference on*. IEEE.
21. El Bireki, M.F.M., M. Abdullah, and Ali Abdrhman, M. **2016**. Digital Image Watermarking Based On Joint (DCT-DWT) and Arnold Transform. *International Journal of Security and Its Applications*, 2016. **10**(5): p. 107-118.
22. Li, L., et al. **2015**. AN H. 264/AVC HDTV watermarking algorithm robust to amcorder recording. *Journal of Visual Communication and Image Representation*, 2015. **26**: 1-8.
23. Joshi, A., R. Patrikar, and V. Mishra, **2012**. *Real time implementation of digital watermarking algorithm for image and video application*: INTECH Open Access Publisher.
24. Deshpande, N. **2010**. Robust DCT based video watermarking algorithms for assorted watermarks. in *Signal Processing Systems (ICSPS), 2nd International Conference on*. 2010. IEEE.