Journal of King Saud University – Computer and Information Sciences (2016) **28**, 199–210

King Saud University

Journal of King Saud University – Computer and Information Sciences

www.ksu.edu.sa
www.sciencedirect.com

A blind video watermarking scheme resistant to rotation and collusion attacks

Amlan Karmakar [a], **Amit Phadikar** [a,*], **Baisakhi Sur Phadikar** [b], **Goutam Kr. Maity** [c]

[a] *Department of Information Technology, MCKV Institute of Engineering, Liluah, Howrah 711204, India*
[b] *Department of Computer Science and Engineering, MCKV Institute of Engineering, Liluah, Howrah 711204, India*
[c] *Dept. of Electronics and Communication Engineering, MCKV Institute of Engineering, Liluah, Howrah 711204, India*

Received 6 November 2013; revised 7 May 2014; accepted 4 June 2014
Available online 2 November 2015

KEYWORDS

Video watermarking;
DCT;
Complex Zernike moments;
Rotation attack;
Collusion attacks;
Rayleigh fading

Abstract In this paper, Discrete Cosine Transform (DCT) based blind video watermarking algorithm is proposed, which is perceptually invisible and robust against rotation and collusion attacks. To make the scheme resistant against rotation, watermark is embedded within the square blocks, placed on the middle position of every luminance channel. Then Zernike moments of those square blocks are calculated. The rotation invariance property of the Complex Zernike moments is exploited to predict the rotation angle of the video at the time of extraction of watermark bits. To make the scheme robust against collusion, design of the scheme is done in such a way that the embedding blocks will vary for the successive frames of the video. A Pseudo Random Number (PRN) generator and a permutation vector are used to achieve the goal. The experimental results show that the scheme is robust against conventional video attacks, rotation attack and collusion attacks.

1. Introduction

Digital watermarking is a well-established technique to protect the Intellectual property and digital copyright of the multimedia information such as image, audio or video. It also secures the integrity of digital contents at the time of transferring those through the internet.

In recent times, Moving Picture Experts Group (MPEG) video standard has widespread applications in internet streaming, digital High-definition (HD) handy-cams as well as in mobile phones. So many researchers are working in the field of digital video watermarking (Al-Taweel et al., 2010; Guojuan and Rang-ding, 2009; Chao et al., 2008). Many video-watermarking algorithms were proposed by many researchers both in the spatial (Al-Taweel et al., 2010) or in the temporal domain (Chao et al., 2008) as well as in the transform or frequency domain (Chao et al., 2008; Al-Taweel et al., 2009; Jing, 2009; Hartung and Girod, 1998) over the last few years.

At the time of developing a video-watermarking algorithm, the researchers should concentrate on the two most important things, i.e., imperceptibility and robustness (Phadikar, 2013). Imperceptibility refers to perceptual transparency, i.e., the

* Corresponding author. Tel.: +91 (33) 2654 9315; fax: +91 (33) 2654 9318.
E-mail addresses: amlan.karma@gmail.com (A. Karmakar), amitphadikar@rediffmail.com (A. Phadikar), baisakhi.sur@gmail.com (B.S. Phadikar), goutam123_2005@yahoo.co.in (G.Kr. Maity).
Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksuci.2014.06.019

algorithm embeds the watermark in video frames in a fashion that the quality of the video frames is not perceptually affected. Robustness refers to the ability of the watermark extraction process to extract the watermark successfully from the video frames, even if the quality of the frames is degraded by various types of intentional or unintentional attacks.

In most of the video-watermarking schemes it is seen that the discussion on robustness is mainly focused on temporal attacks like frame dropping, frame inserting, frame rate changes, etc. (Chao et al., 2008; Al-Taweel et al., 2009; Hartung and Girod, 1998; Liang, 2009). However, very few number of video watermarking schemes are robust against geometrical attacks (Al-Taweel et al., 2010; Guo-juan and Rang-ding, 2009; Jing, 2009; Wei et al., 2010; Bahrushin et al., 2009; Wang et al., 2008; Xu et al., 2008) especially rotation of video frames in any random angle. The rotation attack destroys the synchronization of watermark extraction as it changes the pixel positions of the watermarked video frame. So there is a great challenge for the researches to develop an algorithm which is robust against rotation attack. Zernike moments have a rotation invariance property. One can find out the angle of rotation from the phase information of the Zernike moments. So it is widely used in the field of image as well as video watermarking.

Besides temporal and geometrical attacks, another type of attack for watermarked video is collusion attack (Su et al., 2002). In some situations, it is possible for an attacker to obtain multiple watermarked data. The attacker can often exploit this situation to remove watermarks without knowing the watermarking algorithm. This kind of attack is known as collusion attack. The collusion attack is a different kind of attack and very few numbers of video watermarking schemes can resist collusion attack (Kanócz et al., 2009; Saxena and Gupta, 2007). Moreover, in the literature, very few works are found so far which address both rotation and collusion attacks in a single platform.

In this paper, a blind video watermarking algorithm is proposed in the DCT domain. The scheme embeds watermark before MPEG-4 encoding and extracts the watermark after decoding. So, the scheme can be applied to both uncompressed video or to any compression method (before and after compression). In the present scheme, the watermark information is embedded to every frame. It offers two-fold advantages, namely (1): it increases robustness against frame dropping and frame insertion, (2): one can identify individual frames so that integrity of the video is checked. The Complex Zernike moment is used to make the scheme robust against rotation attack. At the same time, the design of the watermark embedding algorithm is made in such a way that robustness is achieved against collusion attacks, including a Rayleigh fading wireless channel.

The rest of the paper is organized as follows: Section 2 describes the details of Discrete Cosine Transformation (DCT). Section 3 focuses on Complex Zernike moments. Section 4 discusses Type-1 and Type-2 collusion attacks. Section 5 highlights the related works already done in this area. Section 6 explores the algorithm for embedding and extracting watermark. Section 7 presents performance evaluation of the proposed scheme and Section 8 describes conclusions and the scope of future work.

2. Discrete Cosine Transformation (DCT)

Transform coding is a valuable component of contemporary image and video processing applications. Like other transforms, Discrete Cosine Transform (DCT) attempts to decorrelate the image data. It has an excellent energy compaction property for highly correlated images. DCT can be expressed in separable format and exhibits decrease in the entropy of an image. After de-correlation, each transform coefficient can be encoded independently without losing compression efficiency (Khayam, 2003).

The one dimension DCT is defined as:

$$C(u) = \alpha(u)\sum_{x=0}^{N-1} f(x)\cos\left[\frac{\pi(2x+1)u}{2N}\right] \tag{1}$$

For $u = 0, 1, 2 \ldots N-1$.

The inverse transformation is defined as:

$$f(x) = \sum_{u=0}^{N-1} \alpha(u)C(u)\cos\left[\frac{\pi(2x+1)u}{2N}\right] \tag{2}$$

For $x = 0, 1, 2 \ldots N-1$.

where,

$$\alpha(u) = \sqrt{\frac{1}{N}},\ \text{For } u = 0 \text{ and } \alpha(u) = \sqrt{\frac{2}{N}}\ \text{For } u \neq 0 \tag{3}$$

From Eq. (1) it is understood that for $u = 0$,

$$C(u=0) = \sqrt{1/N}\sum_{x=0}^{N-1} f(x)$$

Thus, the first transform coefficient is the average value of the sample sequence. In literature, this value is referred to as DC coefficient. All other transform coefficients are called AC coefficients. However, in image or video processing applications 2-D DCT is used instead of 1-D DCT and it is the extension of Eq. (1).

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) \times \cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \tag{4}$$

For $u, v = 0, 1, 2 \ldots N-1$ and $\alpha(u)$ and $\alpha(v)$ are defined in Eq. (3). The inverse transformation is defined as:

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u,v)\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \tag{5}$$

For $x, y = 0, 1, 2 \ldots N-1$.

In case of video, the temporal redundancy is to be exploited to provide a better compression. Hence, adjacent pixels in consecutive frames show very high correlation. Consequently, this correlations can be exploited to predict the value of a pixel from its respective neighbors. Moreover, most of the image and video data are still available in DCT compressed form. So the scheme is more suitable for real time implementation.

3. Complex Zernike moments

Complex Zernike moments (Guo-juan and Rang-ding, 2009) are constructed using a set of complex polynomials which forms a complete orthogonal set over unit disk of $(x^2 + y^2) \leqslant 1$. The set of such polynomials can be defined as:

$$A_{mn}(x,y) = A_{mn}(r,\theta) = R_{mn}(r)e^{\mathrm{in}\theta} \tag{6}$$

where, $r = \sqrt{x^2 + y^2}$, $\theta = \tan^{-1}(\frac{y}{x})$ and $j = \sqrt{-1}$. The symbols r and θ are defined over the unit disk and $R_{mn}(r)$ is the orthogonal radial polynomial. The symbol n is an integer (positive or negative) depicting the angular dependence, or rotation subject to the conditions $m - |n| =$ even and $|n| \leqslant m$. The orthogonal radial polynomial $R_{mn}(r)$ can be expressed as:

$$R_{mn}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s F(m, n, s, r) \tag{7}$$

where,

$$F(m, n, s, r) = \frac{(m-s)!}{s!\left(\frac{m+|n|}{2} - s\right)!\left(\frac{m-|n|}{2} - s\right)!} r^{m-2s} \tag{8}$$

Based on the Eqs. (6)–(8), the Zernike moment of order n with repetition 'm' of a digital image $f(x, y)$ is defined as:

$$Z_{mn} = (m+1)/\pi \sum_x \sum_y f(x, y)[A_{mn}(x, y)]^* \tag{9}$$

where, $(x^2 + y^2) \leqslant 1$.

In Eq. (9) $A_{mn}(x, y)^*$ represents the complex conjugate of $A_{mn}(x, y)$. To calculate the Zernike moments, the image is first mapped to the unit disk using polar coordinates, where the center of the image is the origin of the unit disk. The pixels that fall outside the unit disk are not used in calculation.

The absolute value of Zernike moments is rotation invariant as reflected in mapping of the image to the unit disk. The rotation of shape around the unit disk is expressed as a phase change. If φ is the angle of rotation of the original image $f(x, y)$, Z^R_{mn} is the Zernike moments of rotated image and Z^R_{mn} is the Zernike moments of the original image then the Zernike moments of the rotated image can be expressed as:

$$Z^R_{mn} = Z^R_{mn} e^{-\mathrm{jn}\varphi} \tag{10}$$

Hence, the rotation of an image only affects its phase angle. One can easily estimate the rotation angle by the phase information of the Zernike moments. Suppose the phase angle of the original image is phase(Z_{mn}) and the rotated image is phase(Z^R_{mn}) then the angle of rotation φ can be computed as:

$$\varphi = \frac{\text{phase}(Z^R_{mn}) - \text{phase}(Z_{mn})}{n} \tag{11}$$

where, $m \neq 0$.

4. Collusion attacks

In some cases, it is possible for an attacker to obtain multiple watermarked copies. The attacker can often exploit this situation to remove watermarks, without knowing the watermarking algorithm. An attack that needs several watermarked copies is known as collusion attacks (Su et al., 2002; Kanócz et al., 2009; Saxena and Gupta, 2007). There are two basic types of collusion attacks:

Type-1: In this type of collusion attack, attacker obtains several copies of the same work, with different watermarks. Here, the attacker tries to find out the video frames which are similar in nature. Hence, frames belonging to the same scene have a high degree of correlation. The attacker then separates various scenes of the video. Then statistical average of the neighboring frames is done to mix the different marks together and computes a new unmarked frame. Type-1 collusion attack can only be successful if successive frames are different enough. Type-1 is depicted in Fig. 1.

Type-2: In this type of attack, the attacker obtains several different copies that contain the same watermark and studies them to learn about the algorithm. Then several copies are averaged by the attacker. If all copies have the same reference pattern added to them, then this averaging operation would return something that is closed to the pattern. Then, the average pattern can be subtracted from the copies to generate an unmarked video. Fig. 2 depicts Type-2.

5. Related works, research gap and motivation of the present work

In the past few years, many researchers proposed several video (Asghar and Ghanbari, 2012) watermarking techniques both in the spatial (Al-Taweel et al., 2010) and frequency domain (Hartung and Girod, 1998). However, the main goal of the researchers is to develop an imperceptible algorithm that is robust against different types of video attacks. Hartung and Girod (1998) proposed a DCT based blind video watermarking algorithm for uncompressed and compressed video using spread spectrum modulation. The watermark energy is spread over all of the pixels of each of the I-frames. To achieve good visual quality of watermarked video a drift-compensation signal is added into the P-frame in addition to watermark signal. Chao et al. (2008) proposed a DCT based temporal synchronous video watermarking technique for MPEG-4 video. A video can be considered as a series of scenes. A scene is a series of temporal continual and perceptually similar frames. In case of temporal synchronization watermarking scheme, watermark is embedded according to features of the scenes. Here, the same watermark is embedded into frames of a single scene, while different watermarks are embedded into perceptually different scenes. Moreover, bit-rate control is introduced to prevent increment of size of the watermarked video and drift-compensation is used to maintain the visual quality of the watermarked video. In the above video watermarking techniques (Hartung and Girod, 1998; Chao et al., 2008), the robustness of the watermarked video is tested against different types of temporal attacks as well as noise attacks. However, those schemes are unable to show the robustness against geometrical attacks, such as rotation of the video in any random angle and cropping. There are two major ways to handle rotation attack, one is geometrical correction and the other is to embed the watermark in geometrical invariant points. Al-Taweel et al. (2009) proposed a DCT based video watermarking technique based on (Hartung and Girod, 1998). The scheme shows robustness against Joint Photographic Experts Group (JPEG) compression, noise attacks and geometrical attacks such as rotation, scaling and cropping. However, the correlation between original watermark and the extracted watermark is relatively small.

L. Jing proposed a novel video watermarking algorithm (Jing, 2009). In this scheme, the watermark is embedded into DCT coefficients of luminance component of the video. Initially, geometric features of one frame are extracted by Harris corner detection. Then the watermark composed of information part and symbol part, is embedded into the video frames. To extract the watermark, geometric attacks are

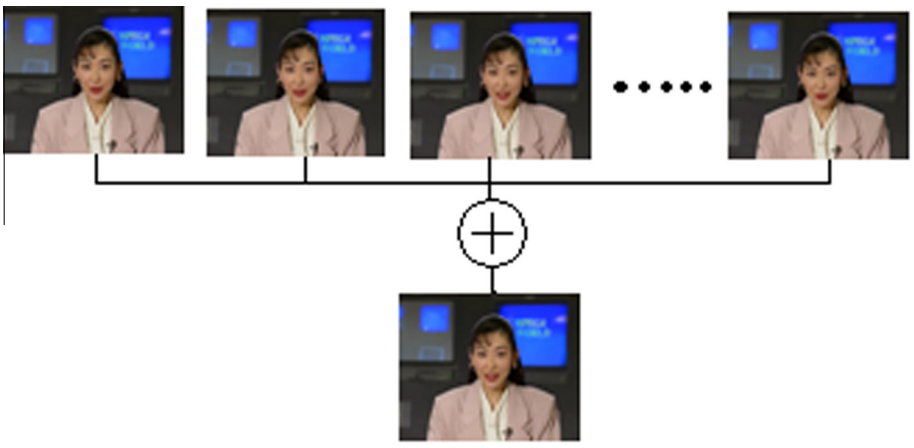

Figure 1 Collusion attack of Type-1 (averaging neighboring frames to obtain a new unmarked frame).

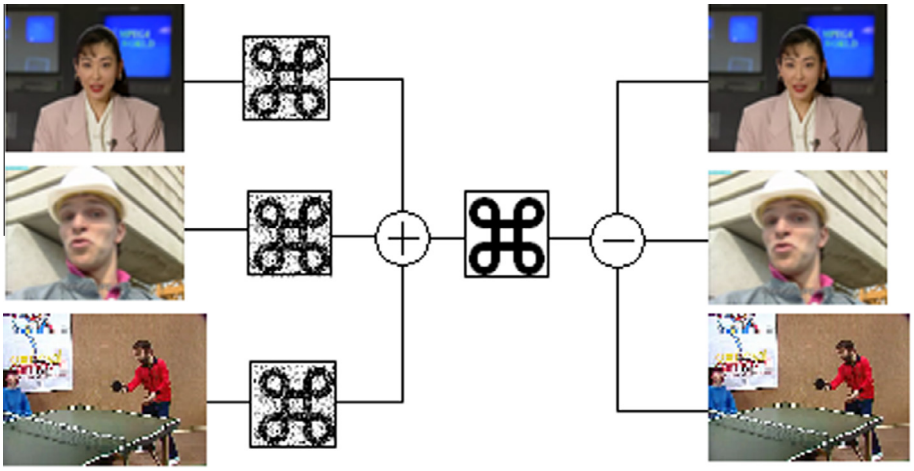

Figure 2 Collusion attack of Type-2 (estimating the watermark and subtracting it from the marked videos to obtain the unmarked videos).

corrected with the help of geometric features. Guo-juan and Rang-ding (2009) proposed a blind video watermarking scheme that is highly robust against rotation attack. Complex Zernike moments and Singular Value Decomposition (SVD) are combined to develop such robust algorithm. In detection, the possible rotation attack is corrected through Zernike moments. Kanócz et al. (2009) proposed a frame by frame video watermarking scheme that is robust against collusion attacks. Robustness of the scheme is shown against both Type-1 and Type-2 collusion attacks as well as other types of temporal attacks like frame dropping, inserting etc. However, the algorithm is non-blind and not robust against geometrical attacks.

Most of the works discussed so far focus on individual properties while designing a video watermarking algorithm. This is pretty reasonable as different applications demand fulfillment of different properties. However, it is not unusual to expect that one single video watermarking scheme may be effective for this dual purpose. The present study is such an attempt to address both rotation and collusion attacks in a single platform. The significant contributions are performance study in Rayleigh fading wireless channel, collusion and rotation attacks. The scheme can also unambiguously distinguish parties involved in collusion operation and innocent users.

6. Proposed scheme

The scheme is divided into two major parts, i.e., watermark embedding and watermark extraction. The block diagrams of the proposed watermark embedding and extraction processes are described in Figs. 3 and 4, respectively.

6.1. Watermark embedding

- In the proposed scheme, the watermark information is embedded in every frame of the video to make the algorithm robust against frame dropping and frame insertion.
- YC_bC_r (luminance, chrominance-red, chrominance-blue) color model is chosen instead of RGB (red, green, blue) color model and the luminance parts of the video are considered as the target embedding area. This is due to the fact that the RGB color space representation has the

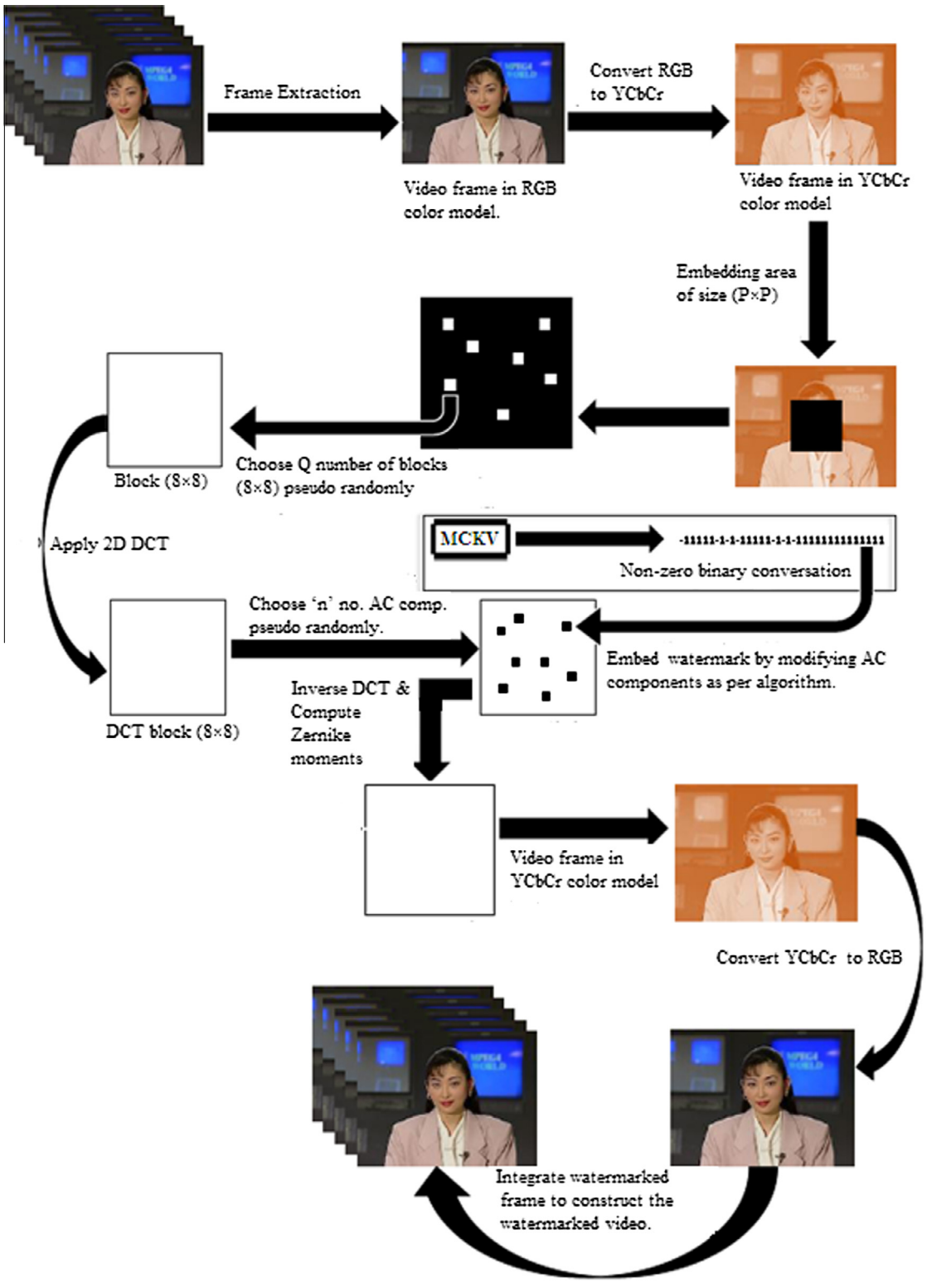

Figure 3 Watermark embedding.

most correlated components, while the YC_bC_r color components are the least correlated components (Dharwadkar et al., 2012). The correlated RGB components are not suitable to embed the watermark. In RGB color space the perceived color quality of a video frame is dependent on all components. Thus, embedding watermark bits into one component independently of the other RGB components is not the best choice. On the other hand the YC_bC_r permits to extract uncorrelated components and it favors the separation of the achromatic part from the chromatic parts of the color image. To achieve high robustness and large embedding capacity, the proposed scheme uses the least

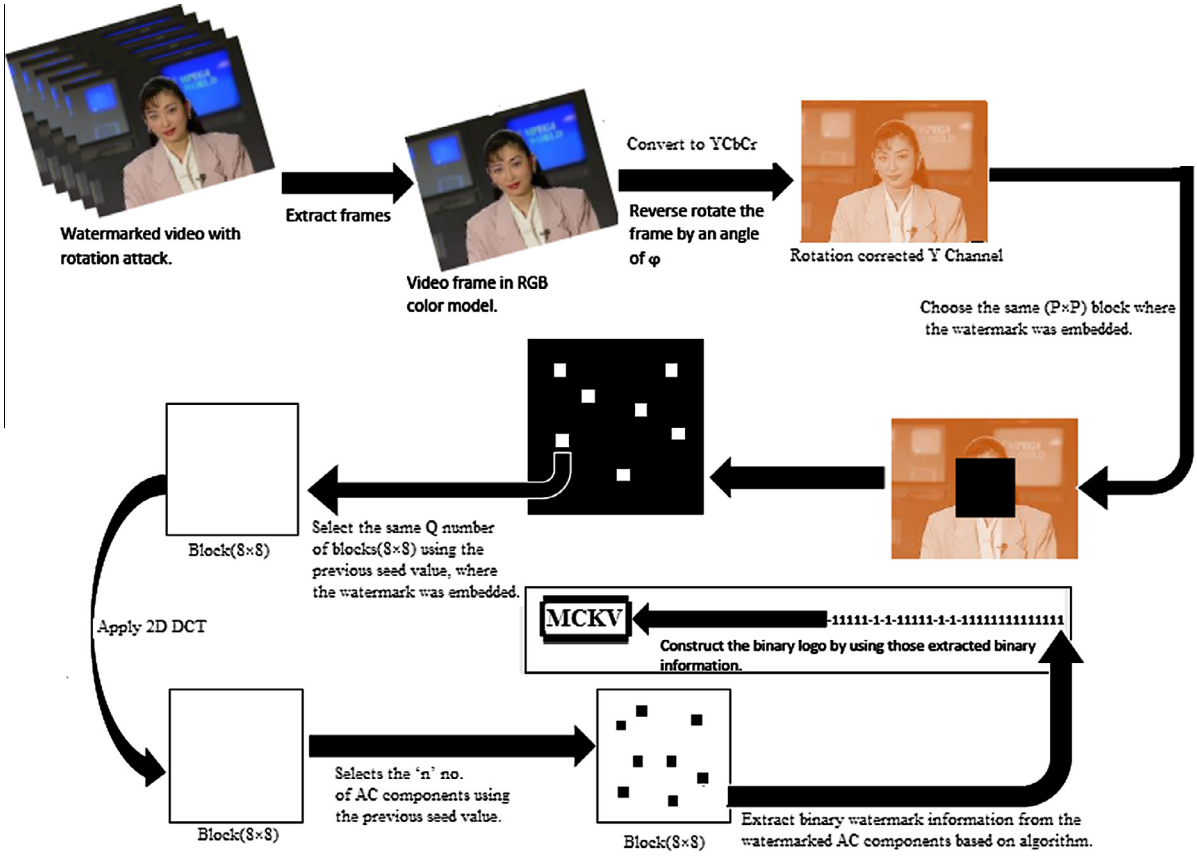

Figure 4 Watermark extraction.

correlated YC_bC_r components of the color image. The color image is represented by Y, C_b and C_r components. The present scheme uses luminance component for data embedding to make the scheme robust against compression. This is due to the fact that during compression no down sampling is done in the luminance component (Phadikar, 2010). That means the loss in luminance component during compression is less than the chrominance component. So the watermark bits can be detected effectively if the bits are embedded in the luminance component.

- A binary logo of size ($M \times N$) is chosen as the watermark logo. Hence, every bit of the binary logo is either "0" or "1". To make a non-zero sequence, all the "0"s are converted to "−1"s.
- To make the scheme robust against rotation attack, luminance block's center is chosen as the center. A square area of size ($P \times$ P) is chosen as the embedding area where the unit circle has to be placed at the time of computation of Complex Zernike moments. The size of ($P \times P$) is multiple of (8×8). This is due to the fact that the present scheme uses a block of size (8×8) to embed a watermark bit.
- To make the scheme robust against collusion attacks, a Pseudo Random Number (PRN) is used to select the embedding blocks (8×8). It also ensures that the embedding blocks will vary for the successive frames. To generate the seed value of the PRN generator a permutation vector is used and is kept as a secret key (K). Here, permutation vector is a user defined secret key (K) that is used as an input (seed value) to the PRN generator. At the time of extraction of the watermark bit, the same permutation vector is used to regenerate the seed as well as the selection of the embedding blocks.
- To compromise between fidelity and robustness, a single watermark bit is embedded into a block of size (8×8) and the whole watermark is embedded within a group of frames. The number of frames within a group depends on the size of the watermark logo.

The embedding process is described as follows:

Step 1: Extract the frames from the original video one by one.

Step 2: Change the color model of the extract frame from RGB to YC_bC_r according to Eq. (12).

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.16875 & -0.33126 & 0.5 \\ 0.5 & -0.41869 & -0.08131 \end{pmatrix} \times \begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{12}$$

Step 3: A square block of size ($P \times P$) is chosen in each luminance component's center which is considered as the target embedding area. Data embedding into the central area of the frame makes the scheme resistant against cropping. The block ($P \times P$) is divided into non-overlapping sub-blocks of size (8×8). It is seen that the watermark is embedded in the central area of a frame to make the scheme robust against cropping. Moreover, the scheme is also secured due to the following reasons, (1) the blocks (8×8) which are used for data embedding

are selected pseudo randomly (key dependent), (2) the 'n' number of AC coefficients which are modulated during data embedding are also selected pseudo randomly (key dependent). Without the above information, it is hard to extract the watermark by unauthorized user into its exact form which in turn makes the scheme secured.

Step 4: Then Q number of distinct (8×8) blocks are selected pseudo-randomly, where the watermark information is to be embedded.

Step 5: Apply 2D DCT on each selected blocks (8×8).

Step 6: Select 'n' number of AC components pseudo randomly for each DCT block (8×8). Those 'n' number of AC coefficients of a block (8×8) are used for embedding one bit of watermark. The same mechanism of Step-4 is used for secret key management of PRN generator. The modification of AC components is done according to the following rule:

```
if (W(k) = = 1)
  do {
  if (mod(C(i, j), δ) ⩽ α)
    Cw(i, j) = C(i, j) − mod(C(i, i), δ) − α
  else
  Cw(i, j) = C(i, j) − mod(C(i, j), δ) + γ
  endif
  } until 'n' number of AC coefficients are considered.
elseif (W(k) = = −1)
  do{
  if (mod(C(i, j), δ) ⩾ γ)
    Cw(i, j) = C(i, j) − mod(C(i, j), δ) + ε
  else
    Cw(i, j) = C(i, j) − mod(C(i, j), δ) + α
  endif
  }until 'n' numbers of AC coefficients are considered.
endif
```

where, $W(k)$ is the watermark bit to be embedded, $C(i, j)$ is the original AC component, $C_w(i, j)$ is the watermarked AC components. Here, α, β, γ, ε and δ are the embedding strength and considered as global constants. The relation between these global constants is $\beta = 2\alpha$, $\gamma = 3\alpha$, $\delta = 4\alpha$ and $\varepsilon = 5\alpha$. The values are taken based on the large number of experimentation, so that the fidelity and robustness of the watermarked video is with an acceptable range. One may use optimization tools like Genetic Algorithm (GA) for the selection of better value of the above parameters for more optimum results. The function mod$(C(i, j), \delta)$ is the remainder of $C(i, j)$ and δ.

Step 7: Compute the Zernike moments $Z_{m,n}$ for some specific values of 'm', 'n' and save those values as keys.

Step 8: Repeat Steps 1–6 until all the video frames are considered.

Step 9: Convert YC_bC_r to RGB according to Eq. (13) and merge all frames to construct the watermarked video.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1.0 & 0 & 1.402 \\ 1.0 & -0.34413 & -0.71414 \\ 1.0 & 1.772 & 0 \end{pmatrix} \times \begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} \tag{13}$$

6.2. Watermark extraction

The watermark extraction is the reverse process of embedding. In this connection, it is to be pointed out that the scheme is "blind" which means that the scheme does not require the cover (original data), and the embedded watermark for extraction of the watermark signal. However, the extraction algorithm has to know the seed of the random number generator, to determine the position of the embedding block. It is to be noted that the watermark bits are embedded into blocks (8×8) which are selected pseudo randomly to increase security of the proposed scheme. The watermark extraction is described as follows:

Step 1: Frames are extracted from the watermarked video one by one.

Step 2: Change the color model of the extract frame from RGB to YC_bC_r according to Eq. (12).

Step 3: Compute the Zernike moments $(Z^R_{m,n})$ of the rotated watermarked video frame (Y channel) for the same values of 'm' and 'n' that was used during embedding. Then rotation is corrected using Eq. (11).

Step 4: A square block of size $(P \times P)$ is selected from the luminance component's center which was used as the embedding area. Then the block is divided into non-overlapping sub-blocks of size (8×8).

Step 5: Same Pseudo Random Number (PRN) generator is used to select the same Q number of (8×8) blocks, where the watermark information was embedded.

Step 6: Apply 2D DCT on each selected blocks (8×8).

Step 7: Select the same 'n' number of AC components pseudo randomly which was modified at the time of embedding and the watermark bit is detected according to the following rule:

```
do {
  if(mod(Cw(i, j), δ) > β)
    Wblock(b) = 1
  else
    Wblock(b) = −1
  endif
  } until 'n' numbers of AC coefficients are considered.
```

where $1 \leqslant b \leqslant n$ and $W_{block}(b)$ are the extracted watermark bit and $C_w(i, j)$ is the watermarked AC components. Then final decision of extracted watermark bit $W(k)$ from a block(8×8) is taken depending on the maximum occurrence of 1 or -1 in $W_{block}(b)$.

Step 8: Repeat Steps 2–7 until all watermark bits are extracted.

7. Performance evaluation

The proposed algorithm is implemented in MATLAB 7.0. The experiment is performed on Intel Core 2 Duo CPU under Windows XP. Three standard videos viz. "akiyo.mp4", "foreman.mp4" and "tennis.mp4" are used for the experiment. The size of binary logo is (50×25). During simulation, we have used 150 frames of each video. The size of the square luminance block, where the watermark is embedded, is taken as (176×176). The value of Q, mentioned in Step 4 of the

embedding process is calculated according to the size of the watermark logo, and in this experiment the value of Q is 50. The value of the global constant (α) mentioned in Step 6 of the embedding process is 13.

This study uses the Peak–Signal-to-Noise-Ratio (PSNR) and the Mean-Structure-Similarity-Index-Measure (MSSIM) Wang et al. (2004) as distortion measures for the watermarked video frame, whereas the relative entropy distance (Kullback Leibler distance) (Maity et al., 2004) is used as measure of security (ε). The high PSNR and MSSIM values of the watermarked video frame and low security values indicate better imperceptibility and security of the hidden data, respectively. PSNR is defined as:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right) \tag{14}$$

Here, MAX is the maximum possible pixel value of the video frame. Mean Square Error (MSE) is defined as:

$$\mathrm{MSE} = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[X(i,j) - X'(i,j)] \tag{15}$$

where X is the coefficients of the original video frame and X' is the coefficients of the watermarked video frame. M and N are the height and width of the video frame respectively.

MSSIM is defined as follows (Wang et al., 2004):

$$\mathrm{MSSIM}(P,\bar{P}) = \frac{1}{M'}\sum_{j'=1}^{M'}\mathrm{SSIM}(P_{j'},\overline{P_{j'}}) \tag{16}$$

where

$$\mathrm{SSIM}(P,\overline{P}) = \left[l(P,\overline{P})\right]^{\delta}.[c(P,\overline{P})]^{\beta}\cdot[s(P,\overline{P})]^{\gamma} \tag{17}$$

where P and $\overline{P}$ are the reference and the distorted image/frame signals, respectively; $P_{j'}$ and $\overline{P}_{j'}$ are the image/frame contents at the j'-th local window; and M' is the number of local windows in the image/video frame. The functions $l(P,\overline{P})$, $c(P,\overline{P})$ and $s(P,\overline{P})$ are the luminance comparison, the contrast comparison and the structure comparison functions, respectively. The symbols δ, β, and γ $\{\delta, \beta, \gamma > 0\}$ are the parameters used to adjust the relative importance of the components.

Let the random variables R and S represent the original and the watermarked video frame, respectively. The Kullback Leibler distance $D(p||q)$ is defined as follows (Maity et al., 2004):

$$D(p||q) = \sum_{x\in X}p(x)\log\frac{p(x)}{q(x)} = E_p\log\frac{p(X)}{q(X)} \tag{18}$$

$$\text{with } 0\log\frac{0}{q} = 0,\ p\log\frac{p}{0} = \infty$$

where $p(X)$ and $q(X)$ denote the probability distribution functions of the random variables R and S, respectively. The symbol E_p represents the expectation with respect to the joint distribution p. The value is always non-negative or zero (if $p(X) = q(X)$). If $D(p||q) \leqslant \varepsilon$, the security value may be assumed to be ε.

We calculate the Normalized Cross Correlation (NCC) between the original watermark image (W) and the decoded watermark image ($\hat{W}$) to quantify the visual quality of the extracted watermark. The NCC is defined by:

$$\mathrm{NCC} = \frac{\sum_i\sum_j W_{ij}\hat{W}_{ij}}{\sum_i\sum_j (W_{ij})^2} \tag{19}$$

Fig. 5 shows the original and watermarked video frame along with the PSNR, MSSIM and ε values. Fig. 6(a) shows the original watermark image, while Fig. 6(b) shows the extracted watermark from all the watermarked video. Without the true key, the extracted signature looks like noise (Fig. 6(c)), which demonstrates that our scheme is sensitive to key (K) and hence is secured. The numerical values show that the quality of watermarked video frame is high and also the watermarked video frame is secured. Table 1 compares the data imperceptibility performance of the proposed technique with the related works (Al-Taweel et al., 2010; Guo-juan and Rang-ding, 2009; Saxena and Gupta, 2007) for the same watermark payload. The numerical values in Table 1 are obtained as the average value of three independent experimentations conducted over three videos with varied image/frame characteristics. The relatively high values of PSNR (dB) and MSSIM indicate that better invisibility of the hidden data is achieved by the proposed technique.

To test the robustness of the proposed watermarking scheme, some typical signal processing operations are performed. Robustness is the measure of immunity of the watermark. Robustness of the proposed scheme is shown in Tables 2–5. The high NCC values in Table 2, Table 3, Table 4 and Table 5 depict that the scheme is not only robust to frame dropping, frame insertion and frame cropping but also rotation of frame in any angle. Table 6 shows the NCC against geometric attacks such as scaling and translation. For image/frame scaling operation, before watermark extraction, the attacked images/frames are rescaled to the original size. It is seen that proposed algorithm can successfully resist attacks like scaling and translation. Moreover, it is observed that the proposed technique provides a superior performance compared to the techniques proposed in Al-Taweel et al. (2010), Guo-juan and Rang-ding (2009).

To study the performance for Type1-collusion operation, we simulate fading like operation on watermarked video. We call this operation as intelligent collusion operation. For collusion attack on continuous media such as audio and video, the estimation of time varying weights become important which is analogous to different gains in fading channels. Fading in wireless mobile channel means unpredictable variation in received signal strength due to vector sum of multiple copies of the same message signal received over variable path lengths (Maity and Mukherjee, 2009). In (Cha and Kuo, 2009), a novel robust Multi Carrier Code Division Multiple Access (MC-CDMA) based fingerprinting scheme against time-varying collusion attack, which is similar to fading operation, is proposed. The algorithm uses novel communication tool sets, namely, multicarrier approach for codeword generation (Hadamard-Walsh codes are used), time varying channel response for colluder weight estimation and Maximal Ratio Combining (MRC) detector. It is quite reasonable to accept fading operation as an intelligent collusion-like as colluders would develop an average watermarked video through variable weights instead of equal weight to remove their identities.

To simulate collusion operations, six different watermarks (see Fig. 7) are embedded in the host video and six different watermarked videos are obtained. We test anti-collusion performance of the proposed algorithm by transmitting first (in fact it is random selection from whole set) five (e.g.) watermarked data using MC-CDMA (Maity and Mukherjee, 2009) through Rayleigh fading channel. Transmission is

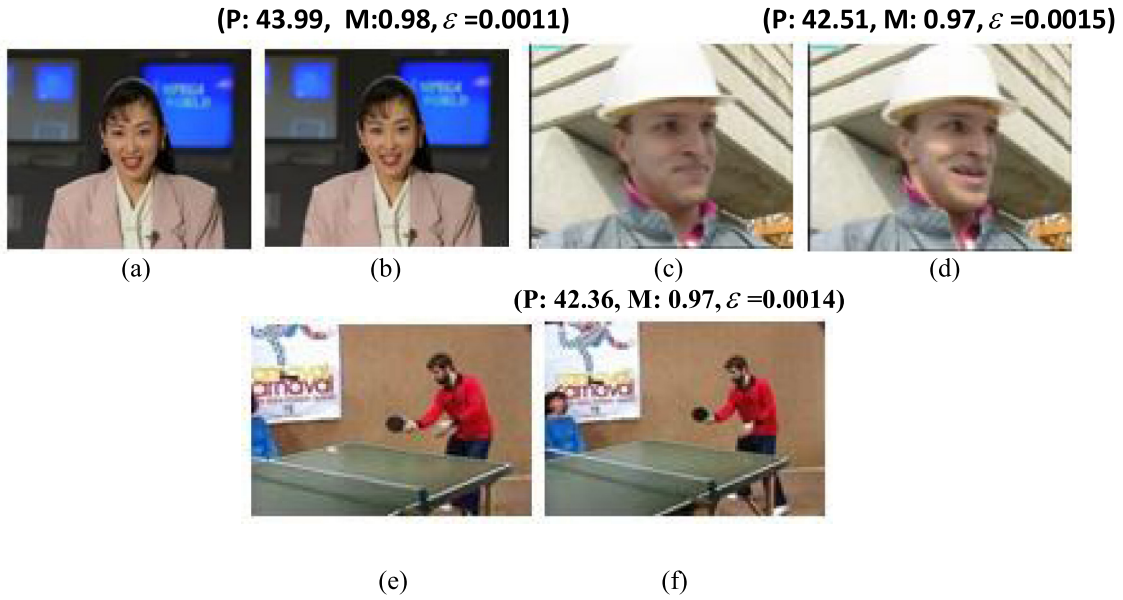

Figure 5 (a) Original frame of Akiyo, (b) Watermarked frame of Akiyo, (c) Original frame of Foreman, (d) Watermarked frame of Foreman, (e) Original frame of Tennis, (f) Watermarked frame of Tennis. (P, M, ε) above each image represents the PSNR (in dB), MSSIM and security values of the watermarked video frame.

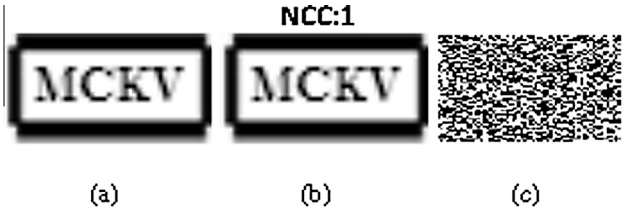

Figure 6 (a) Original watermark image (50 × 25), (b) Extracted watermark (50 × 25) from all watermarked video with NCC = 1, (c) Extracted watermark using fake key.

studied at different Signal to Noise Ratio (SNR) values varying from 10 dB to 18 dB. The resultant received watermarked video frames are then averaged. Transmission of watermarked video through Rayleigh fading channel followed by averaging operation is one way of implementing intelligent collusion operation. In radio mobile communication, the small value of SNR represents that the channel is under deep fade, while high value of SNR represents the reverse. In the present scenario, high and low SNR values represent light and heavy collusion operations, respectively.

Table 1 The PSNR (dB) and MSSIM values of the watermarked images for the proposed technique and related works (Al-Taweel et al., 2010; Guo-juan and Rang-ding, 2009; Saxena and Gupta, 2007). Watermark payload is 1250 bits.

	PSNR	MSSIM
Al-Taweel et al. (2010)	41.20	0.97
Saxena and Gupta (2007)	39.99	0.95
Guo-juan and Rang-ding (2009)	42.12	0.97
Proposed technique	43.99	0.98

Fig. 8(a)–(e) show watermarked video frames after time varying collusion attacks with different weight factors (transmitting each watermarked video frames through fading channel at different SNR values) and then averaged. Table 7 shows the BER (bit error rate) values for different watermarks extracted from the colluded video. Low BER values indicate that the scheme is robust to fading like collusion operation. It is also seen that the BER values for columns 2–6 are quite low compared to the BER values in column 7. The low values of BER clearly indicate that the parties having watermarks in Fig. 7(a)–(e) are identified as colluders. It is also quite clear from the numerical values of BER that parties involved in collusion operation would unambiguously be identified from the innocent users (second set) even at a low SNR value (10 dB). That is, at heavy collusion operation, although the separating zone, that is, the difference between two sets of BER values, are reduced with decrease in SNR values. Similar results are also obtained if different combinations of watermarked images in the set are used in collusion operations.

We also test anti-collusion performance of the proposed algorithm by directly averaging the watermarked video frame, but without channel fading. We call this operation as non-intelligent collusion operation. Table 8 shows the results for non-intelligent collusion operation. High NCC values in Table 8 indicate that the scheme is also robust to non-intelligent collusion operation. Fig. 9 shows the colluder identification performance of the proposed scheme.

Table 2 Robustness against frame drop.

Dropped frames (%)		0%	5%	10%	20%
Akiyo	NCC values	1.000	0.997	0.976	0.962
Foreman		1.000	0.965	0.942	0.920
Tennis		1.000	0.989	0.962	0.931

Table 3 Robustness against rotation attack.

Rotation angle	Akiyo	Foreman	Tennis
10°	0.992	0.978	0.985
20°	0.976	0.989	0.976
30°	0.973	0.987	0.987
45°	0.996	0.985	0.974
50°	0.997	0.997	0.981
60°	0.987	0.974	0.988

Table 4 Robustness against frame cropping.

Cropping percentage	10	20	30	40	50
NCC (Akiyo, Foreman, Tennis)	1.00	1.00	1.00	1.00	1.00

Table 5 Robustness against frame insertion.

Inserted frames (%)		0%	5%	10%	20%
Akiyo	NCC values	1.000	0.986	0.967	0.956
Foreman		1.000	0.974	0.967	0.965
Tennis		1.000	0.937	0.936	0.925

Table 6 The NCC values of the extracted watermark for scaling and translation operations.

	Scaling (down sampling by 50%)	Translation
Al-Taweel et al. (2010)	0.73	0.68
Guo-juan and Rang-ding (2009)	0.64	0.59
Proposed	0.91	0.76

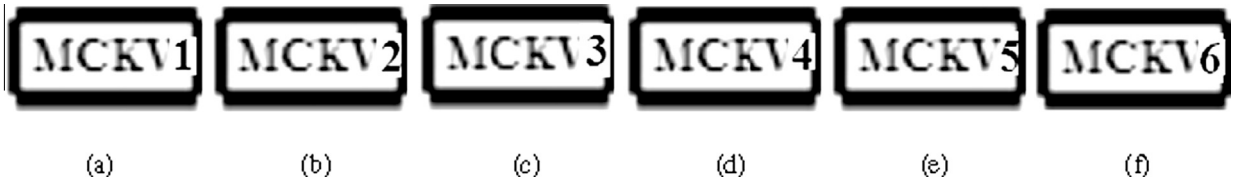

Figure 7 Embedded watermark (50 × 25) for collusion-based attack.

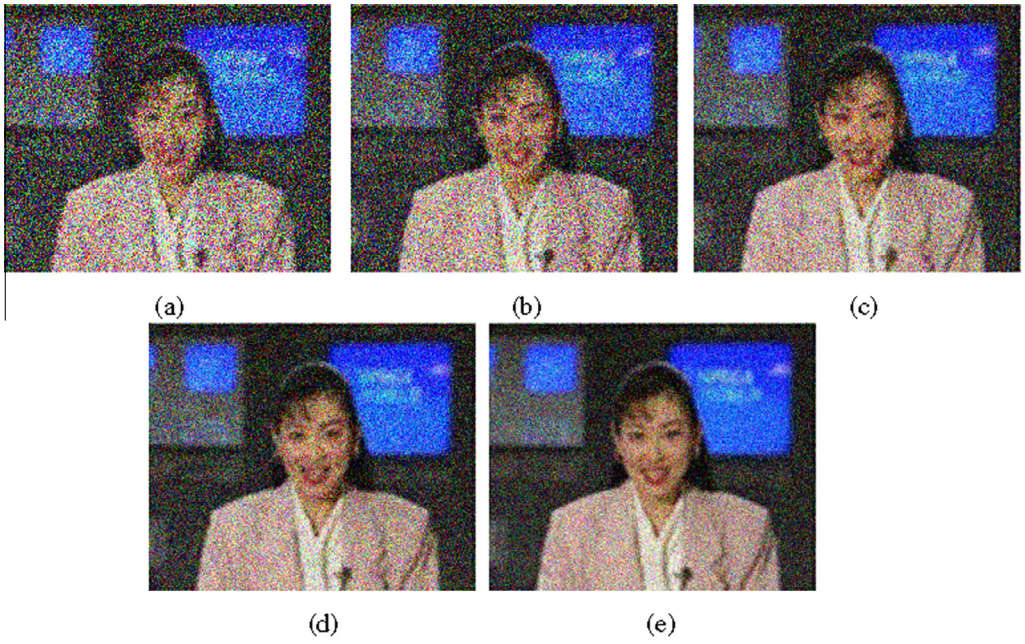

Figure 8 Results for different channel SNR, (a): SNR = 10 dB, (b): SNR = 12 dB, (c): SNR = 14 dB, (d): SNR = 16 dB, (e): SNR = 18 dB.

Table 7 BER value of the extracted watermark.

Watermarks in (Fig. 7)	*a*	*b*	*c*	*d*	*e*	*f*
	Parties involved in collusion operation					Innocent users
SNR 18	0.107	0.109	0.076	0.092	0.069	0.578
SNR 16	0.115	0.112	0.088	0.101	0.076	0.652
SNR 14	0.125	0.125	0.097	0.113	0.089	0.556
SNR 12	0.156	0.214	0.171	0.203	0.164	0.498
SNR 10	0.242	0.257	0.269	0.245	0.183	0.711

Table 8 Robustness against collusion attack of Type-1.

No. of averaged frames		2	5	7	10
Akiyo	NCC values	0.996	0.956	0.885	0.810
Foreman		0.994	0.954	0.884	0.806
Tennis		0.994	0.954	0.883	0.802

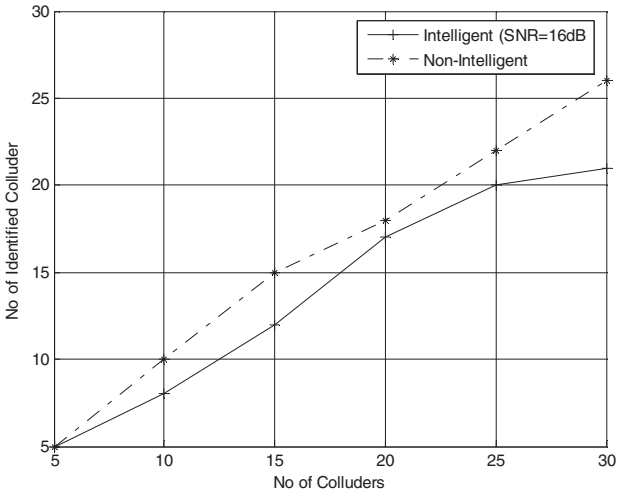

Figure 9 Colluder identification performance.

Table 9 Robustness against collusion attack of Type-2. PV1: Permutation Vector 1 (used in the video-Akiyo). PV2: Permutation Vector 2 (used in the video-Foreman). PV3: Permutation Vector 3 (used in the video-Tennis).

Videos		Extracted using PV1	Extracted using PV2	Extracted using PV3
Akiyo	NCC	1.000	0.5536	0.5534
Foreman	values	0.5464	1.000	0.5426
Tennis		0.5544	0.5256	1.000

We also test anti-collusion performance of the proposed algorithm for Type-2 collusion. Table 9 shows the results for Type-2 collusion operation. High NCC values in Table 9 indicate that the scheme is also robust to Type-2 collusion operation.

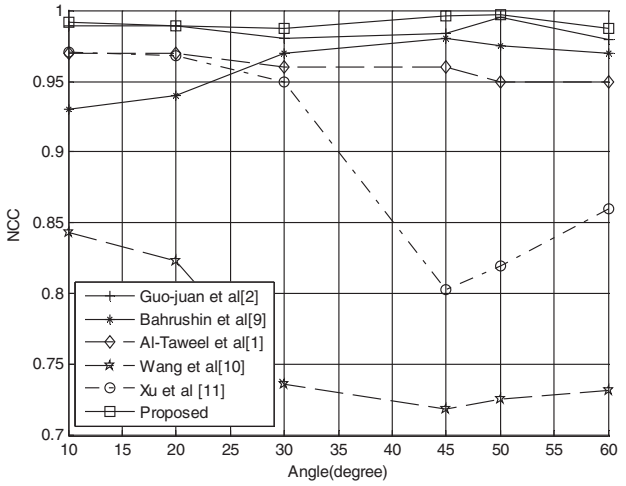

Figure 10 Comparative performance against rotation attack.

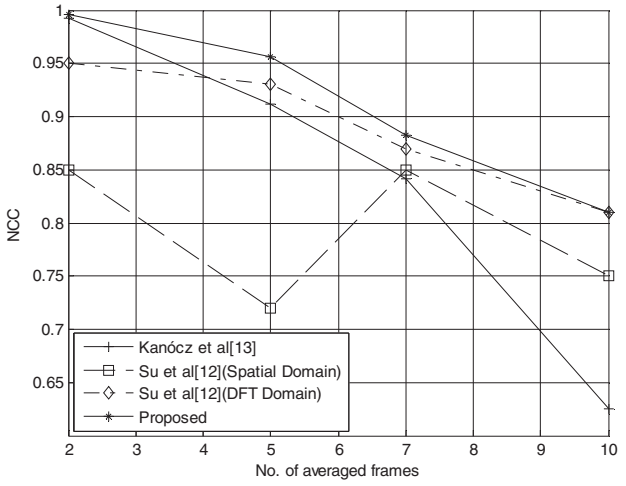

Figure 11 Comparative performance against collusion attack.

The robustness performance of the proposed method is also compared with previously reported works (Al-Taweel et al., 2010; Guo-juan and Rang-ding, 2009; Bahrushin et al., 2009; Wang et al., 2008; Xu et al., 2008; Su et al., 2002; Kanócz et al., 2009) to demonstrate the performance comparison. In the time of comparison of our system with the previously proposed ones, we have used reference implementations provided by the authors. It is observed from the results of Figs. 10 and 11 that the proposed method offers better gain in terms of NCC, which is due to the use of Complex Zernike moments, Pseudo Random Number (PRN) generator and permutation vector which are used to select varieties of embedding square luminance blocks (P × P) for the successive frames of the video. In Fig. 10 it is also observed that Wang et al. (2008) returns poor performance than others related work. This is due to the fact that the scheme proposed by Wang et al. (2008) used even/odd embedding. So the scheme is less robust than others. As expected, in Fig. 11, we have seen that as the number of frames being combined increases, the NCC value decreases. It is also seen that there is a dip in the curve for Su et al. in Fig. 11. This is due to the fact that in their experimentation, every 5th frame was extracted to form the actual set of test frames. So for every 5th frame there is a valley in the graph.

We also examine the time taken in one whole procedure of the proposed video watermarking scheme to depict the computational complexity. Our scheme is in the DCT domain which is fast and demands computation load in O (n log n) operations where n indicates the signal length. Feig (1990) also pointed out; it only takes 54 multiplications to compute DCT for a block of size (8 × 8). So the scheme requires only 1250 × 54 multiplications to embed the watermark logo (50 × 25). Moreover, most of the image and video data are still available in DCT compressed form. So the scheme is more suitable for real time implementation. The execution time of the proposed scheme is studied and it is seen that the scheme takes on an average 126.1 s (watermark embedding: 83.42 s; watermark extraction: 42.68 s) for both watermark embedding and extraction processes. The simulation is conducted on Pentium IV, 2.80 GHz processor, with 512 MB RAM using MATLAB 7 version.

8. Conclusion

In this paper, a DCT based rotation attack resistant blind video watermarking technique is proposed. Rotation invariance property of the Complex Zernike moments is used to achieve the goal, while to make the scheme robust against collusion, design of the scheme is done in such a way that the embedding blocks will vary for the successive frames of the video. The experimental results show that the scheme provides robustness against rotation of video in any angle, collusion attack of Type-1 and Type-2 and conventional video attacks, including a Rayleigh fading wireless channel. Future work can be concentrated on further performance improvement of the proposed scheme, as well as the development of hardware implementation of the proposed scheme through field programmable gate array (FPGA).

Acknowledgments

The authors gratefully acknowledge the valuable and critical comments and suggestions made by the anonymous reviewers for technical quality improvement of the article.

References

Al-Taweel, S.A.M., Sumari, P., Alomari, S.A.K., 2010. Robust video watermarking algorithm using spatial domain against geometric attacks. Int. J. Comput. Sci. Inf. Secur. 8 (2), 51–58.

Guo-juan, X., Rang-ding, W., 2009. A blind video watermarking algorithm resisting to rotation attack. In: Proc. of. IEEE International Conference on Computer and Communications Security, Hong Kong, 111–114.

Chao, C., Tie-gang, G., Li-zong, L., 2008. A compressed video watermarking scheme with temporal synchronization. In: Proc. of IEEE Congress on Image and Signal Processing, Sanya, China, 605–612.

Al-Taweel, S.A.M., Sumari, P., Alomari, S.A., Hussain, A.J.A., 2009. Digital video watermarking in discrete cosine transform domain. J. Comput. Sci. 5 (8), 536–543.

Jing, L., 2009. A novel scheme of robust and blind video watermarking. In: Proc. of IEEE International Forum on Information Technology and Applications, Chengdu, 430–434.

Hartung, F., Girod, B., 1998. Watermarking of uncompressed and compressed video. Signal Processing, Elsevier 66 (3), 283–301.

Liang, H., 2009. Research on the MPEG-2 video watermarking scheme based on spread spectrum technology. In: Proc. of IEEE International Conference on Computer Engineering and Technology, Singapore, 408–411.

Wei, H., Jin-guang, S., Zhong-xu, Y., Di, Y., 2010. Video watermarking scheme based on normalization of pseudo-zernike moment. In: Proc. of the IEEE International Conference on Measuring Technology and Mechatronics Automation, Washington DC, USA, 1080–10820.

Bahrushin, A., Kim, H.J., Tsoy, R., Lopatin, K., 2009. A video watermarking scheme resistant to synchronization attacks. J. Ubiquitous Convergence Technol. 3 (1), 35–40.

Wang, Z., Ye, X., Xiao, N., 2008. Robust watermarking based on norm quantization singular value decomposition and Zernike moments. In Proc: of the IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, 1005–1008.

Xu, D., Wang, R., Wang, J., 2008. Object-based watermarking scheme against geometrical attacks. In: Proc. IEEE Int. Conference Neural Networks & Signal Processing, China, 255–258.

Su, K., Kundur, D., Hatzinakos, D., 2002. A novel approach to collusion-resistant video watermarking. In: Proceedings of SPIE Security and Watermarking of Multimedia Contents, vol. 4675, 491–502.

Kanócz, T., Tokár, T., Levický, D., 2009. Robust frame by frame video watermarking resistant against collusion attacks. In: Proc. of the IEEE International Conference on Radioelektronika, Bratislava, 99–102.

Saxena, V., Gupta, J.P., 2007. Collusion attack resistant watermarking scheme for colored images using DCT. IAENG Int. J. Comput. Sci. 34 (2), 1–7.

Khayam, S.A., 2003. The discrete cosine transform (DCT) – theory and application: information theory and coding, <http://www.lokminglui.com/DCT_TR802.pdf>.

Dharwadkar, N.V., Kulkarni, G.K., Melligeri, T.Y., Amberker, B.B., 2012. The image watermarking scheme using edge information in YCbCr Color Space. In: Proc of 3rd International Conference on Information Security and Artificial Intelligence, Singapore, 127–133.

Phadikar, A., Maity, Santi P., 2010. Quality access control of compressed color images. Int. J. Electron. Commun. 64, 833–843.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error measurement to structural similarity. IEEE Trans. Image Process. 13, 1–14.

Maity, S.P., Nandy, P., Das, T.S., Kundu, M.K., 2004. Robust Image Watermarking using Multiresolution Analysis. In: Proc. of the IEEE INDICON, India, pp. 174–179.

Maity, S.P., Mukherjee, M., 2009. Subcarrier PIC scheme for high capacity CI/MC-CDMA system with variable data rates. In: Proc. of IEEE Mobile WiMAX'09, Canada, pp. 135–140.

Cha, B.H., Kuo, C.J., 2009. Robust MC-CDMA-based fingerprinting against time-varying collusion attacks. IEEE Trans. Inf. Forensics Secur. 4 (3), 302–317.

Phadikar, A., 2013. Multibit quantization index modulation: a high-rate robust data- hiding method. J. King Saud Univ. Comput. Inf. Sci. 25 (2), 163–171.

Asghar, M.N., Ghanbari, M., 2012. MIKEY for keys management of H.264 scalable video coded layers. J. King Saud Univ. Comput. Inf. Sci. 24 (2), 107–116.

Feig, E., 1990. A fast scaled DCT algorithm. In: Proc. of SPIE Image Processing Algorithms and Techniques, vol. 1224, pp. 2–13.