ADAPTIVE VIDEO WATERMARKING WITH PERCEPTUAL GUARANTEE AND EFFICIENCY OPTIMIZATION

Fei Zhang, Hongxia Wang[†]*, Mingze He, Ling Yang, and Jinhe Li*

School of Cyber Science and Engineering, Sichuan University, Chengdu, China;
Key Laboratory of Data Protection and Intelligent Management (Sichuan University),
Ministry of Education, China.

ABSTRACT

Existing video watermarking embeds robust watermarks in each frame of the video for copyright protection and tracking. However, just as any content written on a blank paper is easily perceived, embedding watermarks in the texture-poor frames impairs imperceptibility. Common geometric attacks such as scaling and rotation pose a significant challenge to the existing video watermarking. Image watermarking based on moments is robust against geometric attacks. However, moment-based watermarking is difficult to migrate to the video due to its lack of perceptual guarantee and high computational cost. In this paper, we propose an adaptive video watermarking scheme by exploring the relationship between moments and video textures, which can adaptively select texture-rich frames to embed watermarks for perceptual guarantee. Furthermore, we utilize the properties of moment calculation in videos to optimize efficiency. Extensive experiments show that the proposed method can achieve better imperceptibility than existing methods while maintaining strong robustness.

Index Terms— Adaptive video watermarking, moment, robust, imperceptibility

1. INTRODUCTION

In the past decades, watermarking schemes have been extensively studied [1–3]. Video watermarking is essential for protecting video copyright and tracking [4]. Recently, Liu *et al*. [5] used the distortion function [6] to achieve texture adaption. Yang *et al*. [7] embedded watermark in the high-frequency component of the discrete cosine transform (DCT) can resist intense compression. This work mainly focuses on compression attacks. Some researchers have reported that deep neural networks (DNN) can be applied in video watermarking [8, 9]. Due to the limitation of computing capability, the current research is limited to small-resolution videos only. Huan *et al*. [10] proposed a method for adaptively selecting the decomposition level of dual-tree complex wavelet transform (DT CWT) based on the video resolution to obtain the best performance. Compared to non-adaptive video watermarking [7, 11, 12], adaptive video watermarking offers superior performance. However, existing watermarks lack an adaptive mechanism and have limited robustness against geometric attacks. [10].

The moment-based watermarking approach has strong robustness against geometric attacks. Polar harmonic Fourier moments (PHFMs) [13] and Bessel-Fourier moments (BFMs) [14, 15] are commonly used moments. However, compared with image watermarking, moment-based video watermarking is less studied. The main challenges of applying moments to video watermarking are as follows. First, the video is composed of multiple frames and the picture content of each frame will be different with the lapse of video playing time [16, 17]. Therefore, the visual quality and robustness of watermarks with the same embedding mechanism on different video frames will be different. The direct application of existing moment-based image watermarking techniques to video watermarking cannot guarantee the imperceptibility of each frame. Second, the computational cost of moments is large [18], and video watermarking requires high efficiency for embedding (especially for live scenarios).

Inspired by this, we propose an adaptive video watermarking scheme, which can adaptively filter texture-rich frames to embed watermarks for perceptual guarantee. Moreover, unlike images, video has a large amount of data. In order to meet the requirements of embedding watermarks when the video is played in real-time, we have optimized the method of calculating moments for fast calculation. Thanks to the adaptive mechanism and the geometric invariance of moments, we further extend the robustness and imperceptibility scope of existing schemes. The main contributions of this paper are two-fold:

- We design an adaptive watermarking scheme by exploring the relationship between moments and video textures, which can adaptively select texture-rich video

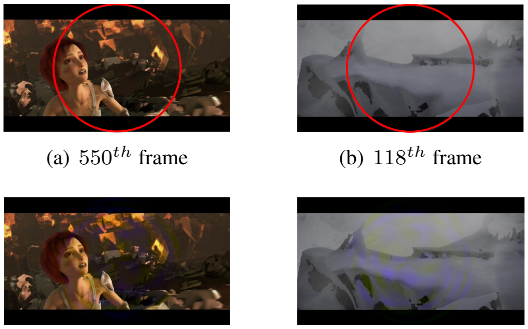
This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grants 62272331 and 61972269 and Sichuan Science and Technology Program under Grant 2022YFG0320, the Key Laboratory of Data Protection and Intelligent Management, Ministry of Education, Sichuan University, and the Fundamental Research Funds for the Central Universities under Grant SCU2023D008. [†]Corresponding author: hxwang@scu.edu.cn. (Hongxia Wang)

frames to embed watermarks for perceptual guarantee.

- We optimize the efficiency based on the independence of the moment basis function and video content, which makes the moment technique practically applicable to video watermarking.

2. MOTIVATION

Applying the existing moment-based image watermarking algorithm to video suffers from the following challenges. First, if each frame of the video is treated as a static image and embedded using a fixed embedding strategy, which has limited imperceptibility when used for frames with different textures. Second, the time complexity of the moments computation is large, which can make the algorithm less efficient if the moments are computed frame by frame. [13] obtained the best performance among current moment-based image watermarking schemes. What will happen if we directly migrate the scheme in [13] to video? Now, we do a test and observe the experimental results. Without losing generality, we take the standard video 'Sintel trailer' as an example, which has a resolution of 1280×720 and a frame rate of 25 fps, containing 1253 frames. The human visual system (HVS) is not sensitive to the chromaticity U-channel, which contributes to the imperceptibility of the watermark. Therefore, we embed 63 bits of the watermark in the center circular area of the U-channel of each frame (YUV 4:2:0) using the scheme in [13]. Fig. 1 shows a comparison of the imperceptibility of the watermark embedded by [13] in frames with different texture richness (relatively rich texture in 550^{th} frame and relatively poor texture in 118^{th} frame). Note that here we embed watermarks of the same capacity and the same intensity in the U-channel of each frame. It is clear that the watermark signal will be more easily perceived by the human eye in the more texture-poor 118^{th} frame. Besides that, the embedding time of [13] is up to 1.52 hours. This is unacceptable for practical applications. Thus, an adaptive mechanism should be explored so that the watermark is embedded in the texture-rich frames for perceptual guarantee and optimizes the efficiency of the scheme as much as possible.



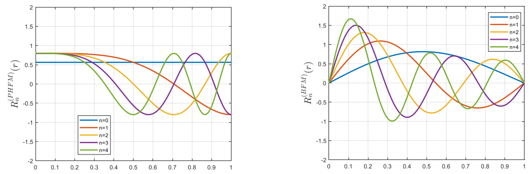
(a) 550^{th} frame

(b) 118^{th} frame

(c) Watermarked (a)

(d) Watermarked (b)

Fig. 1. Comparison of the imperceptibility of embedding the watermark [13] on frames with different texture richness.



(a) PHFMs

(b) BFMs

Fig. 2. Illustrations of radial basis functions of unit disk-based orthogonal moments (PHFMs and BFMs).

3. ADAPTIVE VIDEO WATERMARKING

We propose an adaptive video watermarking scheme, which obtains the texture features of video frames based on moments and adaptively selects texture-rich frames for watermark embedding and extraction.

3.1. Calculation of Moments

A YUV 4:2:0 video $\boldsymbol{M}$ with frame number F and a resolution of $p \times q$ can be expressed as $\boldsymbol{M} = \{\boldsymbol{m}_k, k \in \mathbb{N}_F\}$. $\boldsymbol{m}_k$ denotes the k^{th} frame of the video, which can be represented as $\boldsymbol{m}_k = \{\boldsymbol{y}_k, \boldsymbol{u}_k, \boldsymbol{v}_k\}$. $\boldsymbol{y}_k$, $\boldsymbol{u}_k$, and $\boldsymbol{v}_k$ are the Y, U, and V components of the k^{th} frame. The resolution of U channel is $1/4$ of Y channel, i.e. $\boldsymbol{u}_k = \{u_k(i,j), i \in \mathbb{N}_{p/2}, j \in \mathbb{N}_{q/2}\}$. We embed the watermark in the circular region $\boldsymbol{u}_k^c$ in the center of the U component. Unit disk-based orthogonal moments have better geometric invariance. Therefore, we need to represent $\boldsymbol{u}_k^c$ in unit polar coordinates, i.e. $\boldsymbol{u}_k^c = \{u_k^c(r,\theta), r \in [0,1), \theta \in [0, 2\pi)\}$. The order n and repetition m video frame moment $\boldsymbol{C} = \{C_{n,m}, (n,m) \in \mathbb{Z}^2\}$ is defined as the inner product $\langle \boldsymbol{u}_k^c, \boldsymbol{R}_n, \boldsymbol{A}_m \rangle$ of the $\boldsymbol{u}_k^c$ on the domain Ω:

$$C_{n,m} = \langle \boldsymbol{u}_k^c, \boldsymbol{R}_n, \boldsymbol{A}_m \rangle = \iint_{\Omega} u_k^c(r,\theta) R_n^*(r) A_m^*(\theta) r dr d\theta, \quad (1)$$

where the asterisk $*$ denotes the complex conjuate. Angular basis function $A_m(\theta) = exp(\boldsymbol{j}m\theta)(\boldsymbol{j} = \sqrt{-1})$ and radial basis function $R_n(r)$ could be of any form.

3.2. Adaptive Watermark Embedding

In the field of pattern recognition, moment features are often used to provide a semantic description of image content, so they can also be used to measure the texture richness of a video. Fig. 2 gives the radial basis functions of PHFMs and BFMs as an example. Similar to the frequency coefficients in traditional orthogonal transforms such as DCT, the obtained amplitudes of each order and each repetition moment represents the distribution of the current frame content at different frequencies. When the order n is fixed, the moments whose repetitions are the opposite of each other ($\pm m$) have the same amplitude and are conjugate to each other. The moments of zero repetition are similar to the direct current (DC) coefficient in the DCT transform, which is related to the average intensity in the video frame. Therefore, we can express the texture feature TF of a video frame in terms of the amplitude of the positive repetition moments, i.e.

$$TF=\sum_{n}\sum_{m\in\mathbb{N}+}\langle \boldsymbol{u}_k^c,\boldsymbol{R}_n,\boldsymbol{A}_m\rangle=\sum_{n}\sum_{m\in\mathbb{N}+}C_{n,m}. \qquad (2)$$

We compare texture feature TF of the video frame with a threshold $T_{emb}\in\mathbb{R}^+$. When $TF>T_{emb}$, the texture-rich frames are adaptively selected for watermark embedding, while for the texture-poor frames, we keep them unchanged to ensure imperceptibility. So that we have a set of the texture-rich frame. We assume that $\boldsymbol{w}=\{w_{n,m},n,m\in\mathbb{N}^+\}$ is the binary watermark. The $\boldsymbol{C}=\{C_{n,m},n,m\in\mathbb{N}^+\}$ in $\boldsymbol{u}_k^c$ are calculated by Eq. (1). The quantization equation is as follows

$$|C_{n,m}^w|=\begin{cases} Q(|C_{n,m}|)+3/4\Delta, & w_{n,m}=1\\ Q(|C_{n,m}|+1/4\Delta)+1/4\Delta, & w_{n,m}=0,\end{cases} \qquad (3)$$

where $Q(x)=floor(x/\Delta)\cdot\Delta$ and $|C_i^w|$ denotes the watermarked amplitude of moment. And the corresponding watermarked moment denotes as C_i^w. The watermarked U-channel center region $(\boldsymbol{u}_k^c)^w$ is obtained by

$$(\boldsymbol{u}_k^c)^w=\boldsymbol{u}_k^c+(\sum_{n}\sum_{m}C_{n,m}^w\boldsymbol{R}_n\boldsymbol{A}_m-\sum_{n}\sum_{m}C_{n,m}\boldsymbol{R}_n\boldsymbol{A}_m), \qquad (4)$$

We adaptively replace the original central region $\boldsymbol{u}_k^c$ with the watermarked central region $(\boldsymbol{u}_k^c)^w$ to obtain a watermarked U-channel $\boldsymbol{u}_k^w$. Finally, combine the $\boldsymbol{y}_k$, $\boldsymbol{u}_k^w$, and $\boldsymbol{v}_k$ to get the watermarked frame $\boldsymbol{m}_k^w$. Traversing all frames to get the final watermarked video $\boldsymbol{M}^w$.

3.3. Adaptive Watermark Extraction

Given a watermarked video $\boldsymbol{M}^w$, we obtain its U-channel frames and extract the texture feature (TF) described in Section 3.2. Then, we compare TF with a threshold $T_{ext}\in\mathbb{R}^+$. When $TF>T_{ext}$, the watermarked frames can be adaptively obtained, and then we extract the watermark from these watermarked frames. The extraction equation is as follows

$$\hat{w}_{n,m}=\begin{cases} 1, & mod(|C_{n,m}^w|,\Delta)\geq 1/2\Delta\\ 0, & otherwise,\end{cases} \qquad (5)$$

where $\hat{\boldsymbol{w}}=\{\hat{w}_{n,m},n,m\in\mathbb{N}^+\}$ is the extracted watermark.

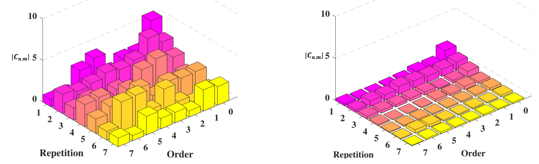
4. EFFICIENCY OPTIMIZATION

Among the unit disk-based orthogonal moments, the computational time complexity of BFMs is the highest [18]. Without losing generality, we use BFMs for the implementation of the video watermarking scheme. The time complexity of calculating moments is related to the definition of the radial basis functions $\boldsymbol{R}_n$. For BFMs, $\boldsymbol{R}_n^{BFM}$ is defined as follows

$$\boldsymbol{R}_n^{BFM}(r)=\frac{1}{2\pi a_n}J_v(\lambda_n r), \qquad (6)$$

where v is a real constant and $J_v(\cdot)$ is the Bessel function of the first kind [19]. $a_n=[J_{v+1}(\lambda_n)]^2/2$ is the normalization constant. λ_n is the n-th zero of $J_v(x)$.

Assume that the highest order $n_{max}=K$, maximum repetition $m_{max}=K$, and the number of a video frame is F. According to Eq. (1), $\mathcal{O}(F\times K^2\times(q/2)^2)$ Bessel function $J_v(\cdot)$, multiplication, and addition are required to calculate all BFMs. In fact, we analyze Eq. (1) to find that $\boldsymbol{R}_n$, $\boldsymbol{A}_m$, $\boldsymbol{u}_k^c$ are independent of each other. Therefore, we calculate $\boldsymbol{R}_n$ and $\boldsymbol{A}_m$ only once and reuse it when calculating the BFMs for different frames. After optimization, the time complexity is $\mathcal{O}(K^2\times(q/2)^2)$ $J_v(\cdot)$ and $\mathcal{O}(F\times K^2\times(q/2)^2)$ multiplication and addition. Thus, the number of calculations of $J_v(\cdot)$ is effectively reduced.



(a) $|C_{n,m}|$ of Fig. 1 (a) (b) $|C_{n,m}|$ of Fig. 1 (b)

Fig. 3. Comparison of the amplitudes of BFMs with positive repetition for 550^{th} frame and 118^{th} frame.

5. EXPERIMENTAL EVALUATION

5.1. Experimental Setup

Data. We evaluate our method on nine videos with a resolution of 1280×720 and a frame rate of 25 fps, ranging from smooth to complex from the test set stated in [10,20]: *Vidyo1*, *Vidyo4*, *Kristen and Sara*, *In to Tree*, *Stockholm*, *Mobcal*, *Shields*, *Park Joy* and *Sintel trailer*.

Parameter. The maximum order and repetition of BFMs $n_{max}=m_{max}=7$, and the quantization step length $\Delta=3.6$. The embedding threshold $T_{emb}=50$ and the extraction threshold $T_{ext}=75.2$. All experiments were performed on a PC with a 3.6 GHz Intel Core i7 CPU and 32GB RAM.

Comparison. We compare our scheme with four existing state-of-the-art schemes including two recent non-moment video watermarking [7, 10], and two moment-based image watermarking [13,21] that treats video sequences as a series of static images and applies them. For a fair comparison, we embed the same 63 bits watermark message for every frame in [7,13,21] and our scheme. For [10], only a 1-bit watermark is embedded in each frame due to the limitation of its scheme.

Table 1. Comparison of the average PSNR, SSIM, capacity, and embedding time. The best and second-best results are highlighted in bold and underlined, respectively.

Schemes	non-moment		moment-based		Ours
	[10]	[7]	[13]	[21]	
PSNR (dB)	43.59	40.28	47.77	47.63	**48.74**
SSIM	0.90	0.85	0.98	0.98	**0.99**
Capacity (bit)	1	**63**	**63**	**63**	**63**
Time (s)	0.42	**0.03**	4.38	2.98	0.19

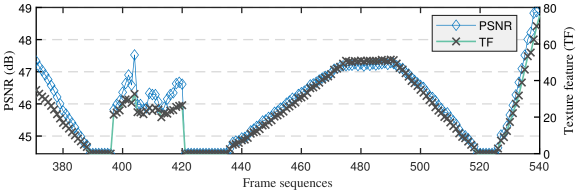


Fig. 4. Relationship between imperceptibility and texture.

5.2. Effectiveness of Adaptive Mechanism

We take the 'Sintel trailer' as a video example. Fig. 3 shows the amplitudes of BFMs with positive repetition for texture-rich 550^{th} frame and texture-poor 118^{th} frame. It can be seen that the amplitudes are larger for 550^{th} frame, while they tend to be close to zero in 118^{th} frame. Watermark is directly embedded in every U component without an adaptive mechanism. TF and PSNR curves of different frames are partially shown in Fig. 4. It is obvious that there is a strong correlation between imperceptibility and texture. It is consistent with the results of the analysis in Section 3.2. Then, we use the adaptive mechanism to obtain a watermarked video. We then perform H.264/AVC encoder compression (quantification parameter, i.e. QP, is equal to 40) and a scaling 0.5 attack on the watermarked video to obtain the attacked video. Fig. 5 shows the TF curves from 100^{th} frame to 600^{th} frame of the original video, watermarked video and attacked video. It can be seen that the adaptive mechanism eliminates texture-poor frames (e.g., 245^{th} frame is subtitled with an all-black background, and 430^{th} frame is black screen transitions). Besides, the extractor can accurately determine which frames contain watermarks and the extraction process is blind.

5.3. Imperceptibility and Computational Cost

The average (peak signal-to-noise ratio) PSNR, (structural similarity index) SSIM, the embedding capacity per frame, and the average embedding time per frame for the five algorithms are given in Table 1. For a fair comparison, in our scheme, we only calculate the PSNR, SSIM, and embedding time of the frames with embedded watermarks. It can be seen that our scheme obtains the highest PSNR and SSIM. The adaptive mechanism can effectively improve imperceptibility. Besides, compared to existing moment-based methods, our scheme takes less average embedding time, which is due to the optimization algorithm in Section 4.

Table 2. Comparison of the average BER(%). The best and second-best results are highlighted in bold and underlined, respectively.

Attacks	non-moment		moment-based		Ours
	[10]	[7]	[13]	[21]	
H.264 (QP=28)	1.15	**0.10**	0.46	0.37	0.42
H.264 (QP=40)	8.07	**0.53**	4.34	5.30	4.03
Scaling 1.1	5.70	N/A	1.28	1.22	**1.11**
Scaling 1.3	6.81	N/A	1.20	1.10	**1.08**
Scaling 0.5	32.63	N/A	3.20	2.41	**1.72**
Scaling 0.8	8.00	N/A	1.52	1.53	**1.34**
Rotation 10°	13.26	N/A	1.41	1.43	**1.26**
Rotation 30°	42.26	N/A	1.36	1.42	**1.25**
Flipping (Hor.)	1.52	N/A	0.96	0.87	**0.86**
Flipping (Ver.)	5.89	N/A	0.97	**0.88**	0.90
FRC (fps=22)	N/A	**0.09**	0.93	0.88	0.82
FRC (fps=17)	N/A	**0.13**	0.88	0.85	0.80
Combined attack	N/A	N/A	13.82	14.51	**10.34**

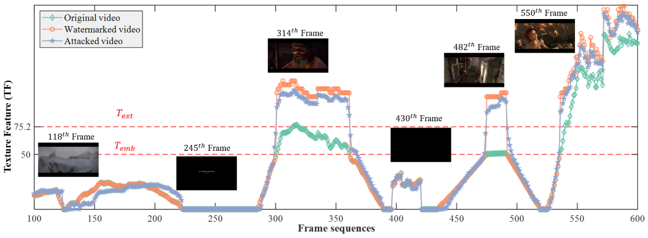


Fig. 5. Adaptive embedding and extraction process.

5.4. Robustness

This subsection compares the robustness of existing methods with attack types including 1) H.264/AVC compression, 2) Geometric attacks including scaling, rotation, and flipping, 3) Frame rate conversion (FRC), and 4) Combined attack: Here, the video is compressed with H.264/AVC (QP = 40), then the resolution is scaled 0.5. It is also horizontally flipped and rotated by 30 degrees, and eventually, the frame rate is altered to 17 fps. The comparison results in terms of average bit error ratio (BER) are given in Table 2. It can be clearly seen that the robustness of the proposed scheme is stronger than the existing moment and non-moment-based schemes due to the introduction of the adaptive embedding mechanism.

6. CONCLUSION

This paper presents an adaptive video watermarking scheme by exploring the relationship between moments and video textures. Our scheme can adaptively select texture-rich frames for embedding and has excellent efficiency. Experimental results show that our scheme has better imperceptibility and stronger robustness. Also, the adaptive mechanism mentioned in this paper can be extended to other moments. In the future, we will introduce motion properties of videos into adaptive video watermarking.

7. REFERENCES

[1] M. Asikuzzaman, M. J. Alam, A. J. Lambert, and M. R. Pickering, "A blind watermarking scheme for depth-image-based rendered 3D video using the dual-tree complex wavelet transform," in *2014 IEEE International Conference on Image Processing (ICIP)*, pp. 5497–5501, IEEE, 2014.

[2] Y. Huang, B. Niu, H. Guan, and S. Zhang, "Enhancing image watermarking with adaptive embedding parameter and PSNR guarantee," *IEEE Trans. Multimedia*, vol. 21, no. 10, pp. 2447–2460, Oct. 2019.

[3] Z. Ma, W. Zhang, H. Fang, X. Dong, L. Geng, and N. Yu, "Local geometric distortions resilient watermarking scheme based on symmetry," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 12, pp. 4826–4839, Dec. 2021.

[4] F. Zhang, H. Wang, L. Yang, and M. He, "Robust blind video watermarking by constructing spread-spectrum matrix," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, IEEE, 2022.

[5] Q. Liu, S. Yang, J. Liu, P. Xiong, and M. Zhou, "A discrete wavelet transform and singular value decomposition-based digital video watermark method," *Appl. Math. Model.*, vol. 85, pp. 273–293, May. 2020.

[6] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP J. Inf. Secur.*, vol. 2014, no. 1, pp. 1–13, Jan. 2014.

[7] L. Yang, H. Wang, Y. Zhang, J. Li, P. He, and S. Meng, "A robust DCT-based video watermarking scheme against recompression and synchronization attacks," in *International Workshop on Digital Watermarking (IWDW),* LNCS 13180, pp. 149–162, Springer, 2022.

[8] X. Luo, Y. Li, H. Chang, C. Liu, P. Milanfar, and F. Yang, "Dvmark: A deep multiscale framework for video watermarking," *arXiv preprint arXiv:2104.12734*, 2021.

[9] Y. Gao, X. Kang, and Y. Chen, "A robust video zero-watermarking based on deep convolutional neural network and self-organizing map in polar complex exponential transform domain," *Multimed. Tools and Appl.*, vol. 80, no. 4, pp. 6019–6039, Oct. 2021.

[10] W. Huan, S. Li, Z. Qian, and X. Zhang, "Exploring stable coefficients on joint sub-bands for robust video watermarking in DT CWT domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 4, pp. 1955–1965, Apr. 2022.

[11] F. Madine, M. A. Akhaee, and N. Zarmehi, "A multiplicative video watermarking robust to H. 264/AVC compression standard," *Signal Process. Image Commun.*, vol. 68, pp. 229–240, Spet. 2018.

[12] H. Huang, C. Yang, and W. Hsu, "A video watermarking technique based on pseudo-3-D DCT and quantization index modulation," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 4, pp. 625–637, Dec. 2010.

[13] B. Ma, L. Chang, C. Wang, J. Li, X. Wang, and Y. Shi, "Robust image watermarking using invariant accurate polar harmonic Fourier moments and chaotic mapping," *Signal Process.*, vol. 172, p. 107544, Jul. 2020.

[14] Z. Liu and H. Wang, "A novel speech content authentication algorithm based on Bessel–Fourier moments," *Digit. Signal Process.*, vol. 24, pp. 197–208, Jan. 2014.

[15] G. Gao and G. Jiang, "Bessel-Fourier moment-based robust image zero-watermarking," *Multimedia Tools Appl.*, vol. 74, no. 3, pp. 841–858, Oct. 2015.

[16] Y. Chen, H. Wang, H. Wu, Z. Wu, T. Li, and A. Malik, "Adaptive video data hiding through cost assignment and STCs," *IEEE Trans. Depend. and Secur.*, vol. 18, no. 3, pp. 1320–1335, Jun. 2021.

[17] Y. Chen, H. Wang, K. Choo, P. He, Z. Salcic, D. Kaafar, and X. Zhang, "A distortion drift-based cost assignment method for adaptive video steganography in the transform domain," *IEEE Trans. Depend. and Secur.*, vol. 19, no. 4, pp. 2405–2420, Aug. 2022.

[18] S. Qi, Y. Zhang, C. Wang, J. Zhou, and X. Cao, "A survey of orthogonal moments for image representation: Theory, implementation, and evaluation," *ACM Comput. Surv.*, vol. 55, no. 1, pp. 1–35, Nov. 2021.

[19] B. Xiao, J. Ma, and X. Wang, "Image analysis by Bessel–Fourier moments," *Pattern Recognit.*, vol. 43, no. 8, pp. 2620–2629, Aug. 2010.

[20] M. Asikuzzaman, M. J. Alam, A. J. Lambert, and M. R. Pickering, "Imperceptible and robust blind video watermarking using chrominance embedding: a set of approaches in the DT CWT domain," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 9, pp. 1502–1517, Jul. 2014.

[21] C. Wang, X. Wang, and Z. Xia, "Geometrically invariant image watermarking based on fast radial harmonic Fourier moments," *Signal Process. Image Commun.*, vol. 45, pp. 10–23, Jul. 2016.