Received September 27, 2020, accepted October 20, 2020, date of publication November 4, 2020,
date of current version November 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3035912

Robust Light Field Watermarking by 4D Wavelet Transform

AMIR ANSARI, GENARO SAAVEDRA, AND MANUEL MARTINEZ-CORRAL
Department of Optics, University of Valencia, E-46100 Burjassot, Spain

Corresponding author: Amir Ansari (amir.ansari@uv.es)

This work was supported in part by the Ministerio de Ciencia, Innovacion y Universidades, Spain, under Grant RTI2018-099041-B-I00, and in part by the Generalitat Valenciana, Spain, under Grant PROMETEO/2019/048. The work of Amir Ansari was supported by the Predoctoral contract from EU H2020 Program under Grant 676401 (MSCA).

ABSTRACT Unlike common 2D images, the light field representation of a scene delivers spatial and angular description which is of paramount importance for 3D reconstruction. Despite the numerous methods proposed for 2D image watermarking, such methods do not address the angular information of the light field. Hence the exploitation of such methods may cause severe destruction of the angular information. In this paper, we propose a novel method for light field watermarking with extensive consideration of the spatial and angular information. Considering the 4D innate of the light field, the proposed method incorporates 4D wavelet for the purpose of watermarking and converts the heavily-correlated channels from RGB domain to YUV. The robustness of the proposed method has been evaluated against common image processing attacks.

INDEX TERMS Digital watermarking, light field, plenoptic image, 4D wavelet, DCT, Gaussian noise, JPEG compression, median filtering, JPEG 2000, integral imaging screen.

I. INTRODUCTION

The capabilities of communication infrastructures have continuously increased over the past few decades. Although the development of the internet architecture facilitates data exchange, it may also lead to a higher risk of copyright infringement. This threat has attracted a lot of attention to copyright protection [1], [2]. In [2] Poort *et al.* reported that more than 60% of audio-visual books, music and other digital content have been acquired illegally. Their survey covers a very diverse range of population in six European, three Asian and three south American nations. This example shows how rampant the copyright breach is around the world. This is only one of the plentiful applications of digital watermarking. In literature, the generally accepted use of digital watermarking refers to embedding hidden information into the host signal which may be either audio, image or video [3]. Image watermarking techniques may be used for copyright protection [4], authentication [5], error detection/correction [6] and other applications [7]. Regardless of the purpose of the watermarking, a compromise is usually held among three major requirements namely transparency, robustness and capacity [8]. Transparency signifies that the similarity between the watermarked and host image is such that one should find no noticeable visual difference. Robustness refers to the probability of an accurate extraction of the embedded watermark if the watermarked image is exposed to some attack. In this context, an attack refers to an operation/event causing the alteration of the image pixels. Finally, it is highly desired that the watermark conveys as much information as possible while retaining the watermark capacity. Fig. 1 illustrates the compromise among the three requirements. The utopian watermarking in Fig. 1 is an ideal situation where all the three requirements are met perfectly, even though reaching such point will be extremely complicated for any given watermarking method.

Depending on the resistance of the watermarking techniques against image processing attacks, they are divided into robust, fragile and semi-fragile. The robust watermarking refers to a method in which the embedded watermark survives from a wide range of attacks. In contrast, the fragile watermark will react to even the smallest modification of the watermarked image. The semi-fragile watermarking methods provide good robustness against some attacks and

The associate editor coordinating the review of this manuscript and approving it for publication was Dezhong Peng.

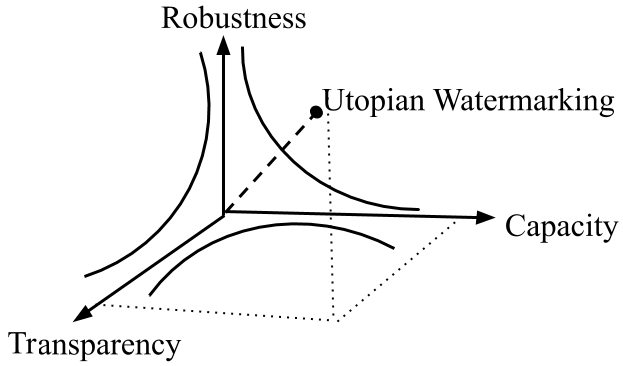

FIGURE 1. The compromise of desired watermarking characteristics (vastly motivated by [9]).

are vulnerable against others [10], [11]. A typical application of the robust watermarking is ownership protection [4], [11] and fragile watermarking is widely used to verify the image authenticity and tamper detection [12], [13].

The watermarking can also be fulfilled in either signal domain or transform domain. While the former alters the values of the signal samples, the latter embeds the watermark into transform coefficients. The signal-domain methods commonly imply less computational complexity [14], [15]. Despite the computational burden of the methods carried out in the transform domain, such methods usually deliver significantly higher robustness [16]. Among the plentiful transforms that can be utilized for image watermarking, some popular ones are the discrete cosine transform (DCT) [17]–[19], the discrete wavelet transform (DWT) [20]–[23], contourlet [24], [25], curvelet [26], [27], ridgelet [28], principal components analysis (PCA) [29] and singular value decomposition (SVD) [30]. There are some other methods employing both signal and transform domains which are used to be known as hybrid methods in the literature [31], [32].

If the host image is not required for the watermark extraction, then the watermarking method is known as blind [33]. Otherwise, the watermarking method is known as non-blind. Reversible watermarking refers to a scheme in which the host image can be recovered after the watermark extraction [34]. Watermarking can also be applied into the chrominance or luminance components but watermark insertion into the luminance component provides more robustness. However, the human visual system (HVS) is extremely sensitive to the luminance information [35].

On the other hand, even though it originates from Leonardo da Vinci's works, the mathematical formulation of the plenoptic function was debuted in 1991 [36]. The plenoptic function is a perfect description of the scene which can hardly be attained; because it is practically very difficult (if ever possible) to observe a scene from every viewpoint, every wavelength and for infinite time span. Regardless, provided that the plenoptic function of a scene is available, one can perfectly and loselessly reconstruct the 3D scene. The terms *plenoptic function* and *light field* (LF) represent the very

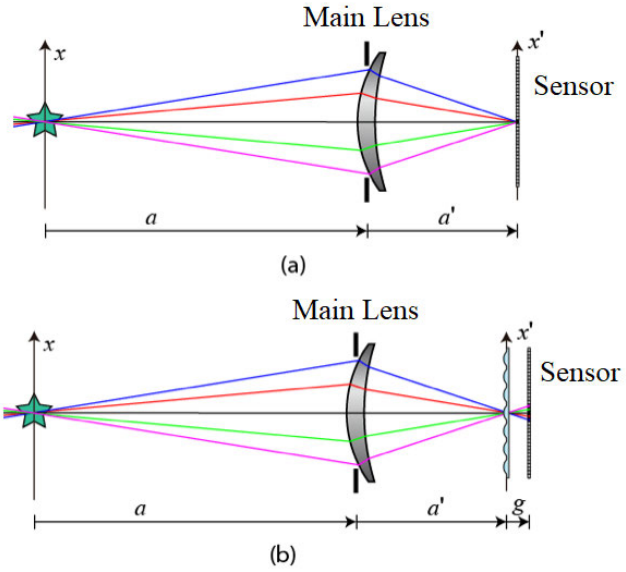

FIGURE 2. a) summation of the rays falling onto a pixel of a conventional 2D camera b) Capturing of individual rays by LF camera.

same concept and are used interchangeably in the literature. As is apparent from Fig. 2, the common 2D camera captures the summation of all rays passing through the same point. Conversely, the LF camera captures the individual rays traveling through the same point. Obviously, the former causes a substantial loss of angular information whereas the latter preserves angular information which is very helpful for the 3D reconstruction at the expense of spatial resolution reduction [37].

The 2D watermarking methods have in principle no specific mechanism to preserve angular information to deal with LF images. Even if the common 2D watermarking method is adoptable to LF data which has more dimensions, there is still a huge risk of ruining the LF angular information. As far as we know, despite the plentiful number of research papers addressing 2D image watermarking, LF watermarking is seldom addressed [38]–[41]. There are also some few works e.g. in 3D object watermarking [42], free-view video watermarking [43], [44] which still do not cover LF watermarking.

Regarding the practical implementation, DCT is usually applied into small blocks and is computationally much more affordable than the transforms as wavelet. Nevertheless, some previous works have demonstrated that the benefits of the wavelet transform may be worth the higher computational complexity [22], [45], [46]. Taking the higher dimensions of the LF into account, its prodigious redundancy cannot be decreased by a 2D wavelet along all the dimensions. Inspired by the promising advantages of the 2D wavelet transform in 2D image watermarking, in this paper we propose a new approach utilizing 4D wavelet for LF watermarking. As will be detailed later on, LF has a great redundancy along all four dimensions in its representation. This indicates the greatest importance of applying a 4D wavelet transform

ensuring the maximum decorrelation along all dimensions. Mangor *et al.* [47] have demonstrated the advantages of 4D wavelet transform in LF compression. To the best of our knowledge, the 4D wavelet transform has never been used for LF watermarking so far. Therefore, we managed to go beyond our previous method [41] which is based on DCT and boost the performance of our LF watermarking scheme.

Thus, in this paper we propose a novel method for LF watermarking. The previous works have shown the advantage of watermark insertion into the luminance component [48], [49]. The aim of this paper is to propose a robust watermarking method taking advantage of the abundant spatial, angular and inter-channel correlation of the LF. In our earlier work [41] the watermark was inserted into the blue channel to minimize the visual distortion. In contrast, in this paper we use a color space conversion for this purpose. Our other contribution is to employ 4D wavelet for spatial and angular decorrelation. Another contribution is to use DCT and 4D wavelet jointly to increase the robustness of the proposed method. Additionally, the impact of the watermarking platform on the 3D perception of the image produced by a LF monitor has been investigated. To the best of our knowledge, this is the first work in LF watermarking which employs 4D wavelet and experimentally verifies the 3D perception of the watermarked LF. In order to evaluate the utility of our method, we check its robustness against common image processing attacks like median filtering, Gaussian noise, JPEG and JPEG 2000 compression. The remainder of this paper is organized as follows. The proposed method and the new contributions are elaborated in section II. The experimental setup and the analysis of the results are discussed in section III. Finally, the conclusions and the future work are drawn in section IV.

II. THE PROPOSED METHOD

A. DCT VS WAVELET

The DCT is widely used in watermarking and compression applications [17]–[19], [50]–[52]. Since this paper is about LF watermarking, the meticulous details of compression standards are beyond the scope of our discussion. More details on compression can be found on [50], [53]. The typical approach of the DCT-based methods is to divide the entire image into non-overlapping blocks and compress every single block independently [50]. The low computational complexity of encoding the non-overlapping blocks, makes it very attractive for numerous methods such as lossy JPEG [50], MPEG-4 [53] and H.264 [54]. The underlying presumption of all these standards is the independence of adjacent blocks. However it is well known that the neighboring pixels of the natural scenes are *significantly correlated* in the spatial domain [50]. Such enormous correlation among the proximate pixels located in the adjacent blocks calls substantial doubt into the hypothesis of independent blocks [55]. Conversely, the wavelet-based coding applies the wavelet transform to the whole image (or to the sub-tiles of big ones). In our previous work, we addressed the DCT-based watermarking of LF [41]. The wavelet transform is computationally more expensive than DCT but the major advantage is elimination of blocking artifacts [56]. However, it is noticeable that the main source of artifacts caused by DCT is the inaccuracy of the hypothesis of independent blocks in spatial domain. This hypothesis is fairly true in transform domain, as the transform coefficients are much more decorrelated than the pixels of spatial domain. That is why we investigated the joint usage of wavelet and DCT to gain the advantages of both transforms. Regarding the typical application of channel decorrelation in most coding standards, it makes a lot of sense to take advantage of substantial channel correlation of RGB images. The widespread usage of chrominance and luminance components in image watermarking and image/video compression standards intrigued us to investigate the impact of using the luminance component of the LF for the purpose of watermarking [50], [53], [54], [57], [58]. Watermarking the luminance channel is considered to reach utmost robustness in comparison to the chrominance channel. The experimental results fully agree with the superiority of unified DCT-wavelet platform. As mentioned in section I, the transparency is one of the fundamental milestones of every watermarking scheme and great care should be taken before any modification of the luminance component. Taking that into account, we also employed the conversion of the color space to obtain a more robust watermarking feature.

B. LIGHT FIELD REPRESENTATION AND PROBLEM STATEMENT

The pixels of common 2D RGB images are usually pinpointed by three variables namely row, column and channel number. As shown in Fig. 3, more dimensions are required to locate LF pixels. Throughout this paper, we would use the term "elemental image" (EI) when referring to the individual images (triple rectangles) shown in Fig. 3. The first two components of the position vector of the illustrated pixel in Fig. 3, indicate the $EI(s_0, t_0)$ where the pixel lies. One can truly infer that each EI is indeed a common 2D image obtained by subsampling the LF. In this paper we use the following notation to address the LF pixels

$$LF(s, t, u, v, ch) \tag{1}$$

Table 1 enumerates the components of this vector position. Consequently, the elemental image s_0, t_0 is represented as

$$EI(s_0, t_0) = LF(s_0, t_0, :, :, :) \tag{2}$$

To achieve complete consistency with notation of [41], we define microimages μI of u_0th row and v_0th column as

$$\mu I(u_0, v_0) = LF(:, :, u_0, v_0, :) \tag{3}$$

where $\mu I(u_0, v_0)$ refers to all the pixels in the row u_0 and column v_0 of all the EIs.

Unsurprisingly, there is a great amount of visual similarity between the neighboring EIs horizontally, vertically and

TABLE 1. Components of the position vector in the LF.

s	The row where the EI lies among other views.
t	The column where the EI lies among other views.
u	The row of the EI in which the pixel is located.
v	The column of the EI in which the pixel is located.
ch	The channel number (from zero to two for RGB images).

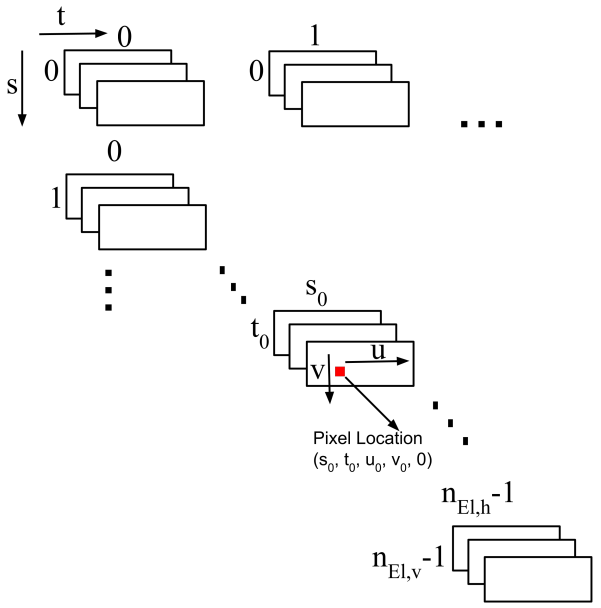

FIGURE 3. A light field image of a 3D scene is composed from a matrix of 2D elemental images taken from equidistant positions and, storing a different perspective of the scene. In the figure we illustrate the pixel location on a light field image.

diagonally. For example the $EI(1, 1)$ will be much more correlated with $EI(m, n)$ $0 \leq m, n \leq 2$ than $EI(10, 10)$. Through this paper, the terms *intercorrelation* and *intracorrelation* will be used to refer to the correlation among the neighboring EIs and that within the neighboring pixels of the same EI. Evidently, there is a major intercorrelation and intracorrelation among the LF pixels. As we embed the watermark into a single channel, without loss of generality we will not use the fifth component of (1) from now on. The continuous wavelet transform (CWT) is mathematically defined as [59]

$$W_f^{\psi}(a;\tau) = \int_{-\infty}^{+\infty} \psi_{a;\tau}(t) f(t) dt \tag{4}$$

where $\psi_{a;\tau}(t)$ is

$$\psi_{a;\tau}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-\tau}{a}\right) \tag{5}$$

a and τ are often known as *scale* and *shift*. Turning now to the LF, as mentioned in section II-A, the LF is inherently 4D data and it stands to reason to use 4D wavelet for decorrelation purpose. The generalization of (4) to 4D space is as

$$W_f^{\psi}(a,\tau) = \int_{s,t,}\int_{u,v}\int_{=}\int_{-\infty}^{s,t,\,u,v\,=\,+\infty} \psi_{a;s_0,t_0,u_0,v_0}(s,t,u,v) LF(s,t,u,v)\, ds\, dt\ \ du\ \ dv \tag{6}$$

where

$$\psi_{a;s_0,t,u_0,v_0}(s,t,u,v) = \frac{1}{\sqrt{|a|^4}} \psi\left(\frac{s-s_0}{a}, \frac{t-t_0}{a}, \frac{u-u_0}{a}, \frac{v-v_0}{a}\right) \tag{7}$$

s_0, t_0, u_0 and v_0 are the shift along the quadruple dimensions of the space used to parameterize LF. Due to the separable property of the wavelet transform, we have

$$\begin{aligned}\psi_{a;s_0,t,u_0,v_0}(s,t,u,v) &= \frac{1}{|a|^2} \psi\left(\frac{s-s_0}{a}\right)\psi\left(\frac{t-t_0}{a}\right)\psi\left(\frac{u-u_0}{a}\right)\psi\left(\frac{v-v_0}{a}\right) \\ &= \frac{1}{|a|^2} \psi_{a;s_0}(s)\psi_{a;t_0}(t)\psi_{a;u_0}(u)\psi_{a;v_0}(v)\end{aligned} \tag{8}$$

As we are interested in the scales of the form $1/2^n$ for $n \in \mathbb{N}$, the above equation simplifies to

$$\psi_{j;m,n,p,q}(s,t,u,v) = 2^{j/2}\psi(2^j s - m, 2^j t - n, 2^j u - p, 2^j v - q) \tag{9}$$

Similar to a for 1D wavelet (5), m, n, p and q are the shifts along the quadruple dimensions used for LF representation (1). Mallat introduced multiresolution decomposition of the signals using the wavelet transform which is a very effective method to calculate (4) [60]. Eq. 4 reminds us of mathematical definition of convolution. According to [60], [61], the wavelet coefficients are obtained by a filtering operation. Hence for a 1D signal with M samples (from 0 to $M-1$)

$$f(t) = \frac{1}{\sqrt{M}} \sum_k W_f^{\phi}(j_0;k)\phi_{j_0;k}(t) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty}\sum_k W_f^{\psi}(j;k)\psi_{j;k}(t) \tag{10}$$

where $\phi_{j_0;k}(t)$ and $\psi_{j;k}(t)$ are orthogonal sets and are known as scaling and wavelet functions respectively [61]. For the 4D LF, (10) generalizes to

$$\begin{aligned}LF(s,t,u,v) = {} & \frac{1}{\sqrt{n_{El,h}\, n_{El,v}\, n_{\mu I,h}\, n_{\mu I,v}}} \\ & \Bigg(\sum_{s_0=0}^{n_{El,h}-1}\sum_{t_0=0}^{n_{El,v}-1}\sum_{u_0=0}^{n_{\mu I,h}-1}\sum_{v_0=0}^{n_{\mu I,v}-1} \\ & LLLL(j_0;s_0,t_0,u_0,v_0)\phi_{j_0;s_0,t,u_0,v_0}(s,t,u,v) \\ & + \sum_{ABCD}\sum_{j=j_0}^{\infty}\sum_{s_0=0}^{n_{El,h}-1}\sum_{t_0=0}^{n_{El,v}-1}\sum_{u_0=0}^{n_{\mu I,h}-1}\sum_{v_0=0}^{n_{\mu I,v}-1} \\ & ABCD(j;s_0,t_0,u_0,v_0)\psi^{ABCD}_{j;s_0,t_0,u_0,v_0}(s,t,u,v)\Bigg)\end{aligned} \tag{11}$$

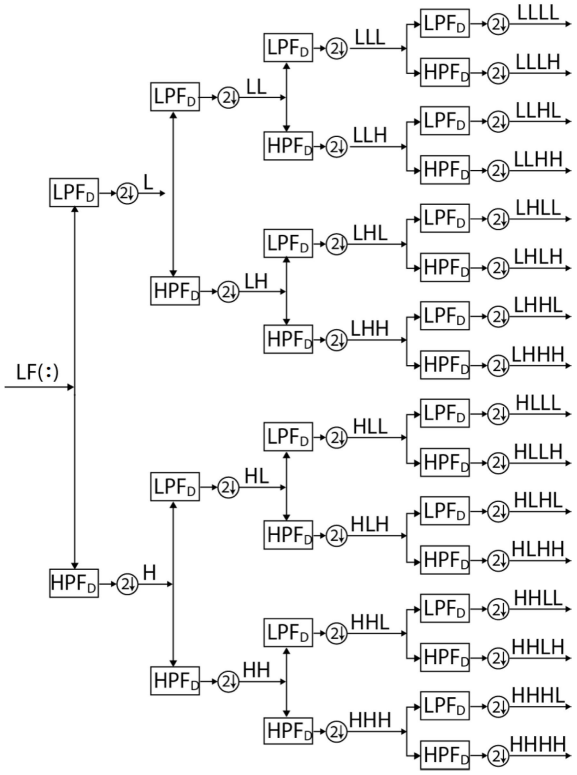

FIGURE 4. 4D wavelet decomposition by filterbanks.

in which

$$A, B, C, D \in \{L, H\}, ABCD \neq LLLL \tag{12}$$

Accordingly, 4D wavelet decomposition can be fulfilled using filterbanks. The low-frequency component *LLLL* represents low frequency information along all elemental and angular directions (s, t, u, v)

$$LLLL(s_0, t_0, u_0, v_0) = \frac{1}{\sqrt{n_{El,h}\, n_{El,v}\, n_{\mu I,h}\, n_{\mu I,v}}} \sum_{s=0}^{n_{El,h}-1} \sum_{t=0}^{n_{El,v}-1} \sum_{u=0}^{n_{\mu I,h}-1} \sum_{v=0}^{n_{\mu I,v}-1} LF(s, t, u, v)\phi_{j_0;s_0,t_0,u_0,v_0}(s, t, u, v) \tag{13}$$

The details are obtained by

$$ABCD(j; s_0, t_0, u_0, v_0) = \frac{1}{\sqrt{n_{El,h}\, n_{El,v}\, n_{\mu I,h}\, n_{\mu I,v}}} \sum_{s=0}^{n_{El,h}-1} \sum_{t=0}^{n_{El,v}-1} \sum_{u=0}^{n_{\mu I,h}-1} \sum_{v=0}^{n_{\mu I,v}-1} LF(s, t, u, v)\psi^{ABCD}_{j;s_0,t_0,u_0,v_0}(s, t, u, v) \tag{14}$$

in which A, B, C and D are defined according to (12).

Equations 11, 13 and 14 can also be realized by filterbanks [60], [61]. Fig. 4 shows the implementation of the 4D wavelet decomposition we employed for LF watermarking. LPF_D and HPF_D in Fig. 4 represent lowpass and highpass filters used for decomposition. The outputs of every filter is downsampled by a factor of two. That is why the decomposed LF has different dimensions with the input LF in spatial domain. As will be mentioned later on, we have used a LF with $16 \times 16 \times 300 \times 300$ pixels, so the dimensions of the decomposed LF will be $8 \times 8 \times 150 \times 150$. Fig. 5 illustrates such decomposition. It is noticeable that we have used color LF but as the 4D wavelet is only applied into Y component, we have converted all the elemental views of Fig. 5 to grayscale images. The LF of Fig. 5 shows the central views of the utilized LF. The LF originally has $16 \times 16 \times 300 \times 300$ pixels but we have selectively shown only 8×8 of the *EI*s. Fig. 5(a) shows the host LF while Fig. 5(b) shows the 4D wavelet decomposition. Even though the filtering order does not matter, we have chosen to apply the wavelet transform to rows and columns of the *EI*s and then to rows and columns of the μIs respectively. Obviously, there is no obligation to use the same LPF_D and HPF_D along all dimensions, but the orthogonality should always be held to ensure a perfect reconstruction. Accordingly, the subbands are upsampled by a factor of two before convolving with reconstruction filters. The reconstruction procedure is illustrated in Fig. 6. Similarly, LPF_R and HPF_R in Fig. 6 represent lowpass and highpass filters used for reconstruction. Fig. 4 and Fig. 6 give an intuitive insight of the decomposition and reconstruction scheme we employed to fulfill (11), (13) and (14). However, regardless of whether these equations are calculated directly or implemented by filterbanks, the result will be the same.

C. THE EMBEDDING PROCEDURE

Fig. 7 shows the embedding procedure. In our previous work [41], we used DCT and SVD to watermark the LF. As mentioned earlier in section II-A, it sounds very plausible to use DCT and DWT jointly. The plausibility of using both transforms is based on the fact that the assumption of independent blocks essentially holds in the wavelet domain. Using DCT and the wavelet transform together provides the benefits of both transforms. To improve the robustness of the watermarking scheme even further, the redundancy of the RGB components of the LF has been substantially decreased by converting the input LF from RGB to YUV. Although it is theoretically possible to decorrelate the pixel vectors of RGB domain by other transforms (e.g. wavelet, DCT, etc), filtering such a short sequence with three samples does not seem logical. The authors must emphasize that every single block in Fig. 7 contributes to the robustness of the proposed method and our experiments thoroughly approve that removing any of them can deteriorate the performance. Another important point about color space convertion is the downsampling of chrominance. In most compression schemes the chrominance is often downsampled e.g. as 4:2:2 [50], [53]. However, as the main purpose of this paper is not compression, no downsampling has been performed on chrominance components. For the reasons of watermark robustness, the luminance component is chosen for watermark insertion.

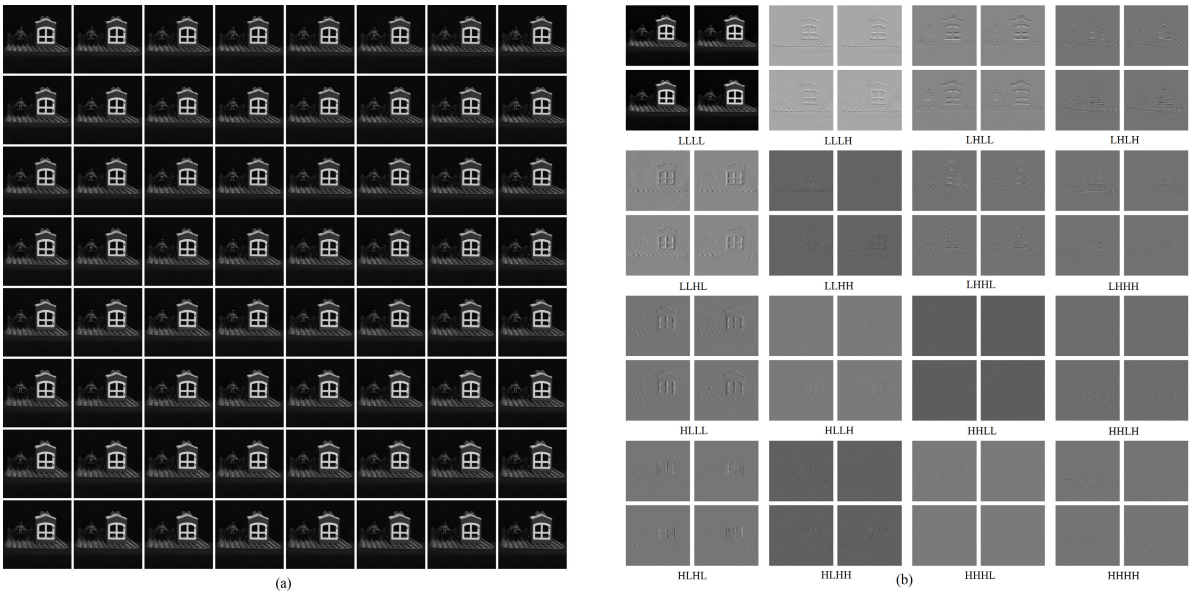

FIGURE 5. 4D wavelet decomposition by filterbanks. The input LF has 8 × 8 × 300 × 300 pixels. The subbands of the decomposed LF has dimensions of 4 × 4 × 150 × 150. All the elemental views are converted to grayscale images.

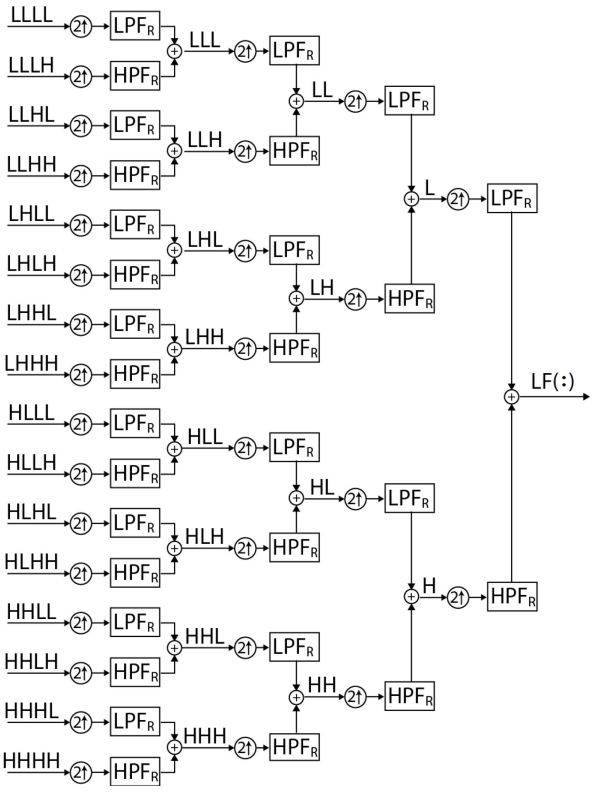

FIGURE 6. 4D wavelet Reconstruction by filterbanks.

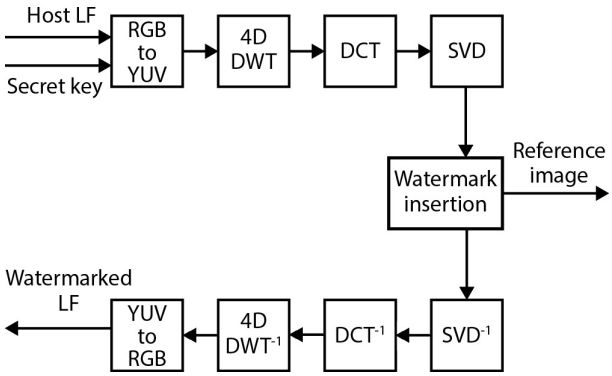

FIGURE 7. The embedding procedure.

As can be seen from Fig. 7, the 4D wavelet decomposition is applied into the luminance component. Depending on the secret key, DCT is applied to the *LLLL* coefficients that are chosen to form a block. Every block will carry one bit of a binary embedded watermark. Following the same method of [41], the generated block to carry watermark bit (i, j) is noted as $LLLL\_blk_{ij}$. Without loss of generality, suppose each block has *BlockSize* rows and *BlockSize* columns. Obviously, there is no compulsion to make such an assumption. The main reason of this assumption is solely for notational convenience.

The DCT of Matrix $A_{M\times N}$ is defined as

$$B_{ij} = \alpha_i \alpha_j \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos\left(\frac{\pi(2m+1)i}{2M}\right) \times \cos\left(\frac{\pi(2n+1)j}{2N}\right) \quad 0 \le m \le M-1,\ 0 \le n \le N-1 \tag{15}$$

in which

$$\alpha_i = \begin{cases} \frac{1}{\sqrt{2}}, & i = 0 \\ 1, & 1 \le i \le M-1 \end{cases}$$

$$\alpha_j = \begin{cases} \frac{1}{\sqrt{2}}, & j = 0 \\ 1, & 1 \le j \le N-1 \end{cases}$$

Afterward the DCT coefficients of each block are arranged in zigzag order

$$blk\_LLLL\_zz_{ij} = zigzag(DCT(LLLL\_blk_{ij})) \tag{16}$$

where

$$zigzag\left(\left[A_{M\times N}\right]\right) = [a_{00}, a_{01}, a_{10}, a_{20}, a_{11}, a_{12}, \ldots, a_{M-1,N-1}]^T \tag{17}$$

The zigzag arranged coefficients are then factorized by SVD. Each matrix $A_{M\times N}$ is rewritten as

$$A_{M\times N} = U_{M\times M}\Sigma_{M\times N}V^T_{N\times M} \tag{18}$$

The columns of U and the rows of V^T are known as the left and right singular vectors respectively. Σ is referred to as a matrix of singular values. Regarding (16) the first n_dct coefficients of $blk\_LLLL\_zz_{ij}$ are chosen

$$blk\_LLLL\_zz\_wm_{ij} = \begin{bmatrix} blk\_LLLL\_zz_0 \\ blk\_LLLL\_zz_1 \\ \vdots \\ blk\_LLLL\_zz_{n\_dct-1} \end{bmatrix} \tag{19}$$

Then the obtained vector is factorized using SVD

$$blk\_LLLL\_zz\_wm_{ij} = U\Sigma V^T \tag{20}$$

The entries of Σ are ordered in descending order. Provided that Σ is diagonal, the first top left entry, σ_{ij} is used for the insertion of the watermark bit (i, j) (wm_bit_{ij})

$$\sigma_{wm\_ij} = \begin{cases} \sigma_{ij} + GF, & wm\_bit_{ij} = 1 \\ \sigma_{ij} - GF, & wm\_bit_{ij} = 0 \end{cases} \tag{21}$$

where σ_{wm_ij} stands for the largest singular value of the watermarked block and GF is the watermark strength. The value of σ_{ij} is stored as the ith row and jth column of the reference image. Hence, the reference image has the same number of rows and columns as the embedded watermark. These values will be used in the extraction procedure. Then the first n_dct components of the DCT are yielded as

$$blk\_LLLL\_zz\_dct\_rec_{ij}\Big|_{n\_dct-1}^{0} = U\Sigma_{wm}V^T \tag{22}$$

where $x\Big|_{N_2}^{N_1}$ refers to the sequence

$$x_{N_1}, x_{N_1+1}, x_{N_1+2}, \ldots, x_{N_2}\ (N_1 < N_2)$$

As the other DCT coefficients have remained unchanged, these coefficients will not be modified in the watermarked LF

$$blk\_LLLL\_zz\_dct\_rec_{ij}\Big|_{(BlockSize\times BlockSize)-1}^{n\_dct} = blk\_LLLL\_zz_{ij}\Big|_{(BlockSize\times BlockSize)-1}^{n\_dct} \tag{23}$$

As the watermark has basically been inserted in $LLLL$, all the other subbands are identical to those obtained through (14). The corresponding coefficients of the $LLLL$ subband can be obtained by

$$blk\_LLLL\_wm_{ij} = DCT^{-1}(zigzag^{-1}(blk\_LLLL\_zz\_dct\_rec_{ij})) \tag{24}$$

in which $blk\_LLLL\_wm_{ij}$ is the block carrying watermarked $LLLL$ coefficients and $zigzag^{-1}$ is the inverse of the $zigzag$ function described in (17). The inverse DCT of the Matrix $B_{M\times N}$ is defined as

$$A_{ij} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\alpha_m\alpha_n B_{ij}\cos\left(\frac{\pi(2i+1)m}{2M}\right) \times \cos\left(\frac{\pi(2j+1)n}{2N}\right) \quad 0 \le m \le M-1,\ 0 \le n \le N-1 \tag{25}$$

in which α_m, α_n are defined identical to α_i, α_j of (15). Finally, after embedding the watermark in all the specified locations, the Y component of the watermarked LF can be obtained by the inverse 4D wavelet

$$Y_{wm} = IDWT4D(LLLL_{wm}, ABCD) \quad A, B, C, D \in \{L, H\}, ABCD \neq LLLL \tag{26}$$

where $IDWT4D(:)$ is the inverse 4D wavelet which can be calculated using filterbanks (Fig. 6) or (11). $LLLL_{wm}$ is the approximation subband obtained after watermark insertion. The last step to accquire the watermarked LF is to convert the coefficients of luminance-chrominance to color domain. As the U, V components of the host LF have not been modified, the same components will be used to get the watermarked LF

$$\begin{bmatrix} R_{wm} \\ G_{wm} \\ B_{wm} \end{bmatrix} = \begin{bmatrix} T_{00} & T_{01} & T_{02} \\ T_{10} & T_{11} & T_{12} \\ T_{20} & T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} Y_{wm} \\ U \\ V \end{bmatrix} \tag{27}$$

D. THE EXTRACTION PROCEDURE

The extraction procedure is quite similar to the embedding procedure. Fig. 8 shows the extraction procedure. The conversion of the color space, 4D wavelet, DCT and SVD are performed in the same fashion as described in section II-C. The watermark is extracted using the comparison of the reference image and the singular values of the watermarked image

$$wm\_ext_{ij} = \begin{cases} 1, & \sigma_{ij} > ref\_img_{ij} \\ 0, & \sigma_{ij} < ref\_img_{ij} \end{cases} \tag{28}$$

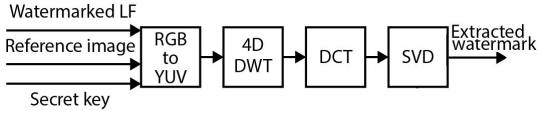

FIGURE 8. The extraction procedure.

III. THE EXPERIMENTAL RESULTS

A. THE EVALUATION METRICS

In order to evaluate the performance of the proposed method objectively, some numeric metrics have been employed. The first metric, is the Peak Signal to Noise Ratio (PSNR) which is very well known in the literature. For the LF, it is defined as [41]

$$PSNR = 10\log_{10}\left(\frac{n\_ch\, n_{el,h}\, n_{el,v}\, n_{\mu I,h}\, n_{\mu I,v}}{MSE}\right) \quad (29)$$

where

$$MSE = \sum_{n=0}^{n\_ch-1}\sum_{s=0}^{n_{el,h}-1}\sum_{t=0}^{n_{el,v}-1}\sum_{u=0}^{n_{\mu I,h}-1}\sum_{v=0}^{n_{\mu I,v}-1}\Big(I(n,s,t,u,v) - I_W(n,s,t,u,v)\Big)^2$$

where $I(:)$ and $I_W(:)$ are the host and the watermarked LF respectively. Throughout this section, the subscripts I and W will be used to refer to the host and watermarked image respectively e.g. $PSNR_W$ stands for the $PSNR$ of the watermarked image.

The similarity between the embedded and extracted watermark is measured by bit error rate (BER)

$$BER = \frac{\sum_{i=0}^{N_b-1}\sum_{j=0}^{N_b-1}(wm\_bit_{ij} \oplus wm\_ext_{ij})}{N_b^2} \quad (30)$$

where wm_bit_{ij} and wm_ext_{ij} are the embedded and extracted watermark bits in the ith row and jth column respectively. $\oplus$ is the bitwise exclusive or. N_b is the number of the rows and columns of the embedded watermark. Apparently there is no necessity to set the same number of the rows and columns for the watermark. However, such an assumption is only for more convenience and the embedded watermark may have an arbitrary number of rows and columns. To make more sense of percentage-wise BER, through this paper we normalize the BER from 0% ($BER = 0$) to 100% ($BER = 1$). A BER of 100% signifies that *all* the watermark bits have been extracted falsely.

Both aforementioned metrics are popular in the literature to measure the performance of the watermarking scheme. Nevertheless, some previous works have shown that the $PSNR$ may be misleading, i.e. delivering inflated results for a severely degraded image or a low value for an image with minor degradation [62]. Alternatively, it is commonly asserted that the mean structural similarity $MSSIM$ is a solid quality metric as it is mainly based on HVS characteristics. $MSSIM$ is defined as

$$MSSIM(I, I_W) = \frac{(2\mu_I\mu_W + C_1)(2\sigma_{I,I_W} + C_2)}{(\mu_I^2 + \mu_W^2 + C_1)(\sigma_I^2 + \sigma_W^2 + C_2)} \quad (31)$$

in which μ_I and μ_W are the average intensity and σ_I^2, σ_W^2 are the variance of the host and watermarked LF. Accordingly, σ_{I,I_W} is the covariance of the host and watermarked LF. C_1 and C_2 are some constant figures chosen depending on the LF content. Letting $C_1 = 0, C_2 = 0$ will reduce (31) to universal quality index (UQI). $MSSIM$ incorporates the luminance, contrast and obviously structural similarity. This fully justifies the common belief of consistency of $MSSIM$ and HVS.

B. OBJECTIVE AND SUBJECTIVE PERFORMANCE

As stated in section I, LF watermarking is a very recent topic and is seldom addressed in the literature. In order to make a fair comparison to previous work in LF watermarking, we used the same experimental setup and the same LF of [41]. The interested reader may refer to [41] for more details on capturing and preprocessing the utilized LF.

As mentioned earlier in section I, the capacity is one of the triple requirements that every single watermarking scheme should meet. The proposed method can deliver good robustness and transparency for watermark dimensions of 32×32 which is 16 times higher than the 8×8 watermark of our previous work [41]. All the tests that will be addressed later on, have also been performed for a watermark with dimensions of 8×8 and the graphs exhibited a very similar trend.

Since SVD provides a perfect decorrelation, one may propose that SVD will suffice to achieve the watermarking task. We previously have shown the invalidity of such a hypothesis [41]. As this assumption may seem plausible to our readers, we raise this point and highlight the importance of all the building blocks of Fig. 7 as a whole. Fig. 9 shows the central EI view of both the host and watermarked LF. As is evident from Fig. 9, the watermarked image is visually indiscernible from the host one and no visual difference can be detected. For a better comparison, the same region of both images is magnified demonstrating any possible degradation to the low (the smooth black part of the window) and high (edges, e.g. where the color changes from black to white on the window) frequency information of the LF. Letting $GF = 80$ and $n\_dct = 6$ results in $PSNR$ of 54.77 dB. Even though the achieved $PSNR$ figure is lower than [41], it is still far out of the HVS perceptible range to yield a measurable difference with the host LF. To assert this claim, one can simply flip all the least significant bits(LSB) of pixels of the popular Lena image. Even though such alteration is hardly (if ever) noticeable, the $PSNR$ of the altered image will fall to 48.13 dB. So, achieving a $PSNR$ higher than 54 dB can visually be considered to yield a watermarked image identical to the host image. Eq. 29 shows the logarithmic inherent of $PSNR$. Therefore if an image with $PSNR$ of 48.13 dB seems identical to the reference image, there is certainly no concern in regards to the imperceptibility of a watermarked image with a $PSNR$ of 54.77. This statement is approved by a $MSSIM$ figure higher than 0.99. Achieving a high value of $PSNR$ and $MSSIM$ signifies the visual and numerical fidelity of the watermarked LF.

It is well-known in the literature that even a severely degraded image may have a high $PSNR$ figure. Nevertheless, based on the shown watermarked image and magnified regions, it seems safe to assume that the proposed method

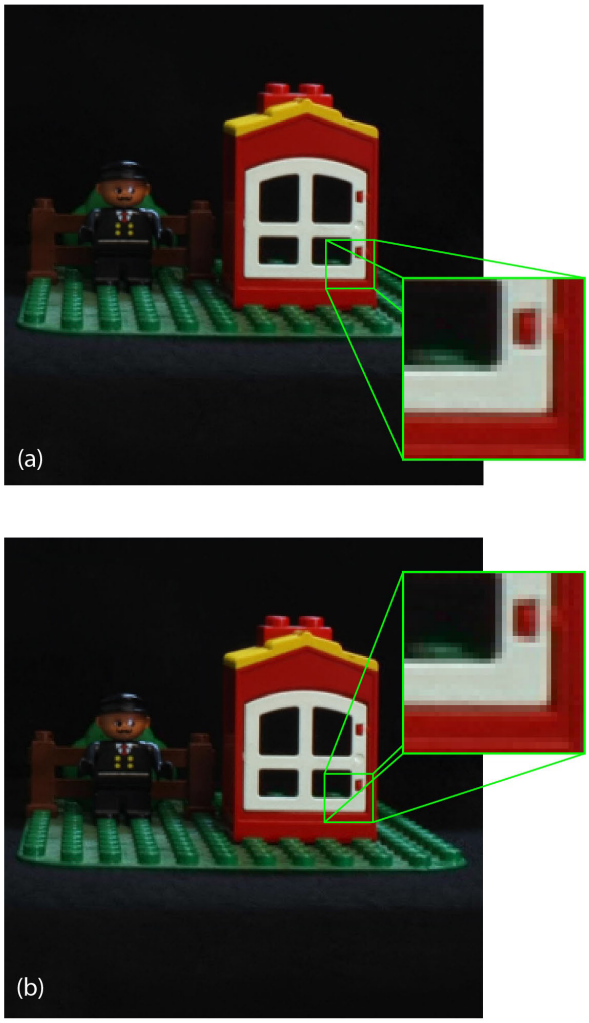

FIGURE 9. The central *EI* view of a) The host LF b)The watermarked LF.

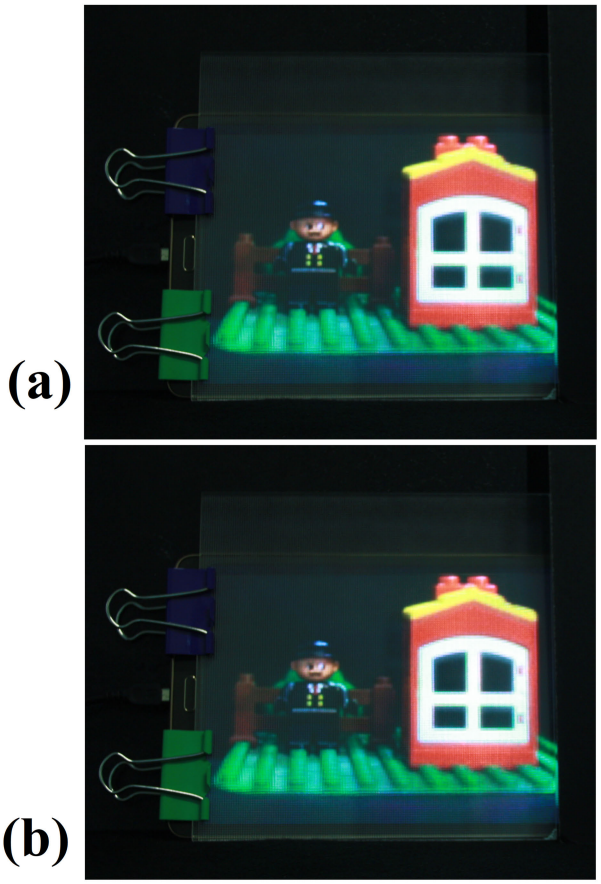

FIGURE 10. LF projection to the integral imaging monitor. a) The host LF b) The watermarked LF.

does not introduce any noticeable difference in the content of the watermarked image. As outlined previously in section II-B, a prominent feature of LF is the great parallax it provides for 3D representation. Incorporating a vast amount of *angular* information along with dense spatial information facilitates a better and more realistic 3D perception for the spectators observing the reconstructed scene from different views. This is also true when the same spectator changes his position while observing the scene. Therefore, considerable care must be taken to ensure that the LF's angular information does not deteriorate. To verify the influence of the proposed watermarking method on 3D reconstruction of the LF, both the host and watermarked LF are projected to an integral imaging monitor. Apparently, both reconstructed scenes look identical which is a solid confirmation of the transparency of the proposed method earlier approved by high figures of *PSNR* and *MSSIM*. Fig. 10 shows the projection of the host and watermarked LF to an integral imaging monitor. Not even a minor negative effect on parallax of the watermarked LF was found. Both reconstructed scenes, either with host or watermarked LF, are indistinguishable. The objective and subjective transparency of the proposed method is further corroborated by this observation, thus meaning that the proposed method fully preserves the angular information of the watermarked LF. To show the parallax of the host and watermarked LF, separate videos are prepared. The video for the host and watermarked LF are respectively available online.

Unlike most of the watermarking schemes in the literature, an encouraging characteristic of the proposed method and [41] is that both promise an error-free watermark extraction. Provided that GF$<$236, despite the watermark strength, the embedded watermark will always be extracted perfectly from the watermarked LF. However, $GF = 236$ is a highly extreme configuration and practically, such a high value of GF will never be used. Even for $GF = 275$, the *BER* equals 1.46% which is essentially negligible and can safely be tolerated. Fig. 11 shows the embedded and extracted watermark. As is apparent from Fig. 11, the embedded and extracted watermark are identical. Throughout this paper we will use the term *SVD method* for referring to the watermark insertion using SVD while excluding all other transforms. Fig. 12 illustrates the *PSNR*, *BER* and *MSSIM* of the proposed method, SVD and [41]. The error-free extraction of the embedded watermark is fully confirmed by Fig. 12(b) and remains zero

FIGURE 11. a) The embedded watermark b)The extracted watermark.

despite the value of GF. According to Fig. 12(c), for the GF values less than 176, the $MSSIM$ will pertain higher than 0.99. The ordinate axis of Fig. 12(c) has been greatly magnified to highlight the difference, but the differences are definitely negligible from a practical point of view. Even for an extreme case of $GF = 235$, the $MSSIM$ will not fall below 0.98 which is only 0.02 from the ideal similarity ($MSSIM = 1$).

C. ROBUSTNESS ANALYSIS

1) GAUSSIAN NOISE

The robustness of the proposed method is verified against Gaussian noise. Fig. 13 shows the central EI view of noisy watermarked LF. The noise power is chosen among $\sigma_n^2 \in \{100, 225, 625, 1225\}$. Fig. 13 shows how dominantly the noise affects the watermarked image. As is obvious from Fig. 14 (a) and (b), the SVD method may work for more moderate noise attacks but once noise power increases, the SVD method collapses. We previously showed the impressive robustness of [41] against Gaussian noise if a watermark of dimensions 8×8 is embedded. However, when it comes to a larger watermark of 32×32, the proposed method considerably outperforms [41] in more intense attacks of Gaussian noise. When the noise power increases from 100 to 225, there will be a noticeable gap between the performance of SVD and the proposed method, where SVD never attains $BER = 0$. For a very intense attack with a noise power of 1225, even [41] does not attain $BER = 0$ while the proposed method still can deliver BER of 0.

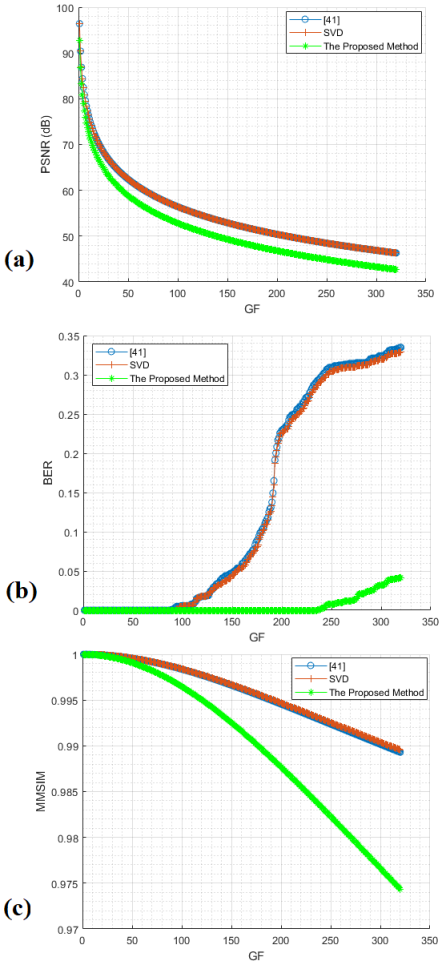

FIGURE 12. a)*PSNR* b) *BER* c) *MSSIM* of the watermarked LF.

2) JPEG COMPRESSION

Another common attack is JPEG compression which is frequently used in different platforms. Fig. 15 shows the central EI view of watermarked images with $qf = 25\%, 50\%, 75\%$ and 100%. As is evident from Fig. 16, the proposed method has a distinct advantage over [41] in terms of robustness against JPEG compression. Letting $qf = 5\%$, the compression distorts the watermarked image so severely that neither [41] nor the proposed method can extract the embedded watermark successfully (However, our simulations show that even for an aggressive qf of 5%, the proposed method is still marginally better than the two other methods). Once qf rises to 25%, the superiority of the proposed method emerges. When $qf = 25\%$, [41] and SVD reach a BER no lower than 31.15% while the proposed method achieves $BER = 8.69\%$. Lletting $qf = 50\%$, the proposed method gains a competitive advantage over [41]. If $GF = 80$, the proposed method delivers $BER = 11.71\%$ whereas [41] provides $BER = 27.05\%$. However, with $qf = 50\%$, the BER of the proposed method will finally converge to zero. Letting $qf = 75\%$, the BER of the proposed method converges to zero even faster and for

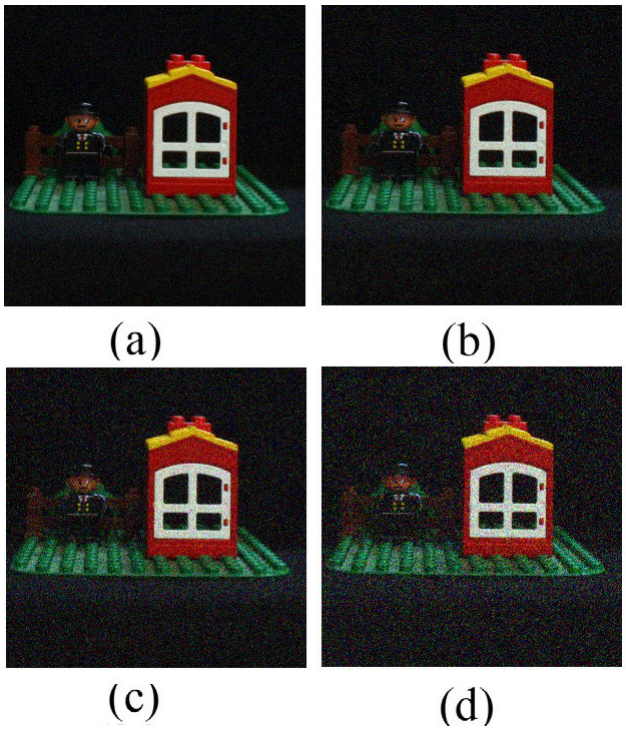

FIGURE 13. The central *EI* view of the noisy watermarked LF with noise power of a) 100 b) 225 c)625 d)1225.

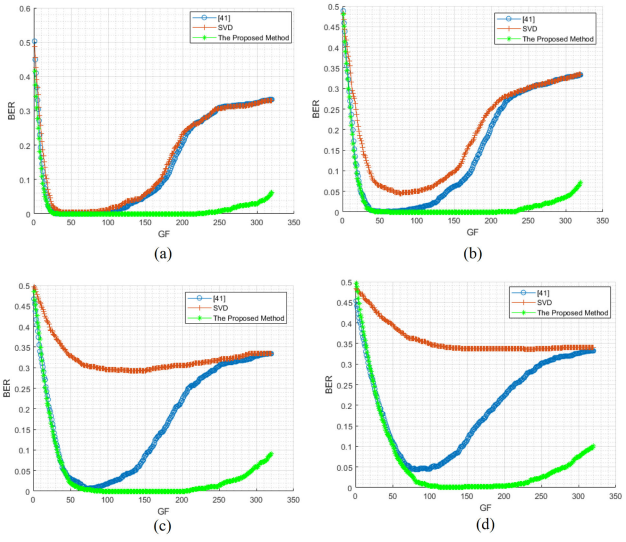

FIGURE 14. The *BER* of the noisy watermarked images attacked by Gaussian noise with noise power of a) 100 b) 225 c) 625 d) 1225.

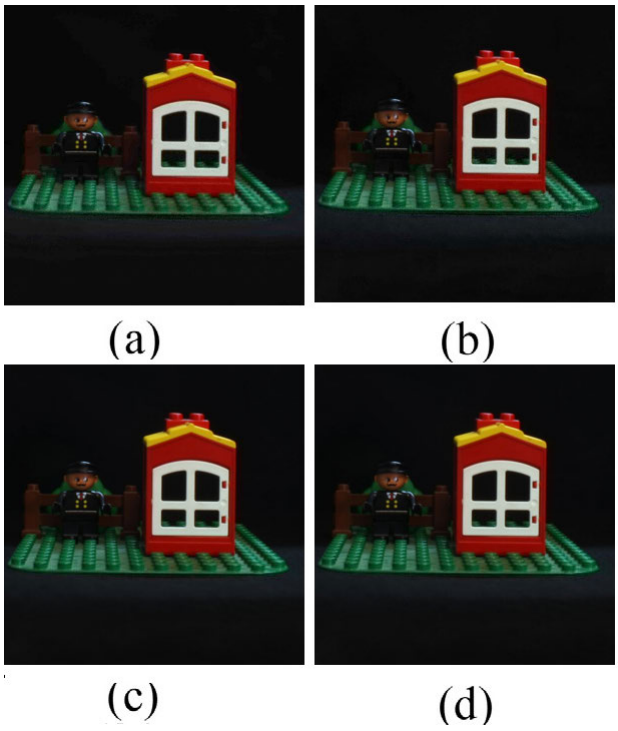

FIGURE 15. JPEG compression of the watermarked LF (central *EI* view) a) $qf = 25\%$ b) $qf = 50\%$ c) $qf = 75\%$ d) $qf = 100\%$.

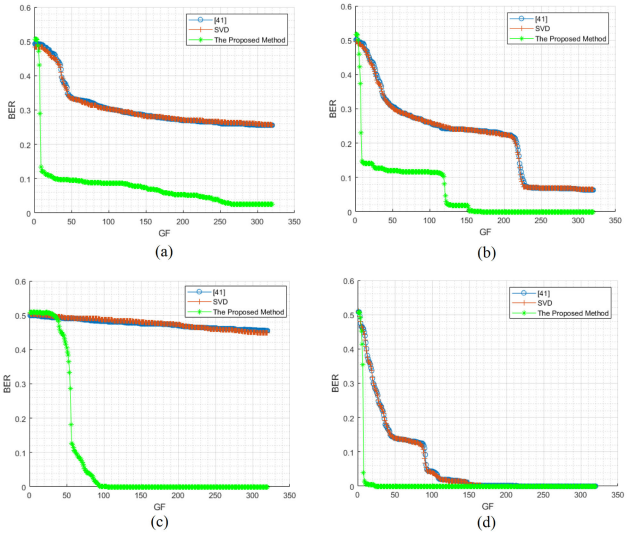

FIGURE 16. The *BER* of the watermarked image compressed with JPEG a) $qf = 25\%$ b) $qf = 50\%$ c) $qf = 75\%$ d) $qf = 100\%$.

$GF = 80$ we have $BER = 4.00\%$. To our surprise, [41] offers almost a constant *BER* regardless of GF value. The best *BER* figure obtained by [41] is 48.34%. As the distortion of JPEG compression with $qf = 100\%$ is significantly less noticeable (if ever), it is used far more frequently than other quality factors. For $qf = 100\%$, the *BER* of the proposed method rapidly converges to zero such that for $10 < GF$ we have $BER < 1\%$ and with an increment of GF to 24 the *BER* converges to zero. As is obvious from Fig. 16, the proposed method outperforms the SVD method and [41] in terms of robustness against JPEG compression. The absolute superiority of the robustness of the proposed method against JPEG compression can reasonably be contributed to the 4D wavelet transform and conversion of color space.

3) MEDIAN FILTERING

Another likely attack to occur, is median filtering. In comparison to [41], the proposed method also delivers more robustness against median filtering. As can be seen from Fig. 17,

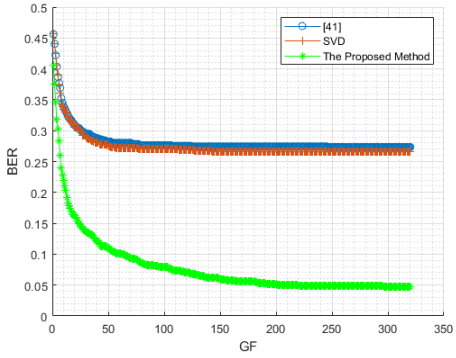

FIGURE 17. The *BER* of the watermarked image exposed to the median filter.

the *BER* of the proposed method falls to 8.59% while that of [41] reaches 27.64%

4) JPEG 2000

When it comes to compression, JPEG 2000 is unanimously recognized to provide much higher quality images for the same compression ratio [56]. As JPEG 2000 employs a wavelet transform for compression, we investigated the effect of this compression scheme on our wavelet-based watermarking platform. The compression ratio *CR* is set to 10, 20, 30 and 40. It is worth mentioning that in JPEG compression a higher *qf* produces a higher quality compression output. Contrary to JPEG, in JPEG 2000 a higher *CR* is associated with higher compression and higher amount of loss. As can easily be seen from Fig. 18, the distortion caused by JPEG 2000 is significantly less than JPEG. This assertion holds also when *CR* keeps rising up to 40. Fig. 19 indicates that for low or moderate *CR* values ($CR = 10$ and even 20), all three methods provide good robustness and *BER* converges to zero quite fast. Turning to severer *CR*, e.g. 30, the supremacy of the proposed method becomes highlighted. According to Fig. 19(c), for $CR = 30$ the *BER* of the proposed method converges to zero faster than [41] and for $GF = 24$, *BER* falls to zero while the *GF* of [41] should be no less than 44 to attain $BER = 0$. The superiority of the proposed method becomes even more observable when *CR* increases to 40. As is clear from Fig. 19(d), the *BER* of the proposed method falls to zero enormously faster than SVD and [41]. The proposed method achieves a BER $= 0$ for $GF = 24$ while SVD and [41] require a *GF* of at least 61 to converge to $BER = 0$.

D. INTEGRITY ANALYSIS

As stated in section II-C, the proposed method incorporates multiple transforms, namely the 4D wavelet, DCT, color conversion along with SVD factorization. In section III-C we showed the inferior results of the SVD method. Hence, we do not discuss the SVD method in this section. Specifically, the poor performance of such a method may cause the other curves to seem following the same trend. Not addressing the SVD method helps the ordinate axis to have a finer scale to compare other methods.

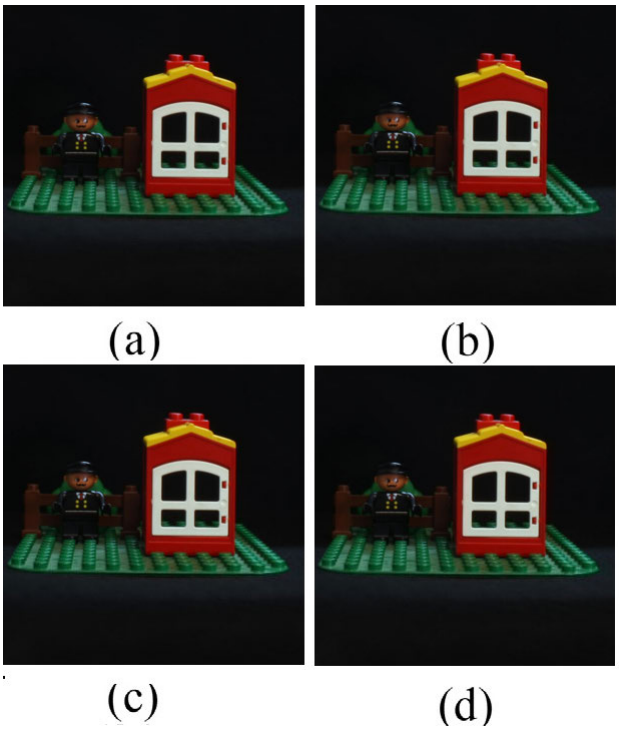

FIGURE 18. Compression of the watermarked images with JPEG 2000 (central *EI* view) a) $CR = 10$ b) $CR = 20$ c) $CR = 30$ d) $CR = 40$.

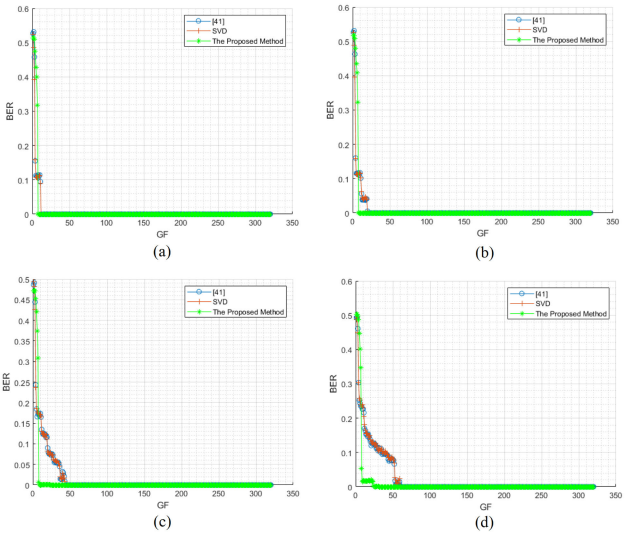

FIGURE 19. The *BER* of the watermarked image compressed with JPEG 2000 a) $CR = 10$ b) $CR = 20$ c) $CR = 30$ d) $CR = 40$.

A question of interest is whether it is necessary to utilize all the aforementioned transforms jointly. In other words, is it possible to achieve the same performance without one or more of these transforms? We would like to emphasize that each of these transforms reduces the redundancy in a dimension which is not fulfilled by other transformations.

- The color conversion from RGB to YUV serves to avert the substantial intercorrelation of color channels. Such decorrelation cannot be achieved by applying a

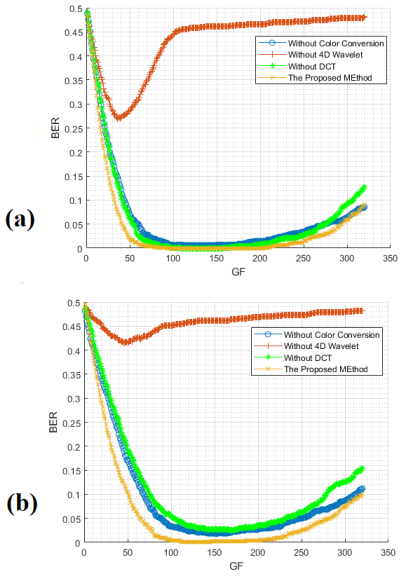

FIGURE 20. The *BER* of the watermarked LF exposed to Gaussian noise with noise power of a) 625 b) 1025.

transform either to the individual EIs (4D wavelet, or DCT) or to the regions of the (transformed) EIs.

- The 4D wavelet is used to decrease the tremendous correlation among EIs. Such correlation occurs horizontally, vertically and diagonally. For example $EI(i, j)$ is highly correlated with $EI(i \pm k, j)$, $EI(i, j \pm k)$ and $EI(i \pm k, j \pm k)$ where $k \in \mathbb{N}$. No matter if the transform is applied to an entire EI or some regions, the intercorrelation of individual EIs cannot be reduced by applying a transform to the triple channels of the EIs. Hence, the 4D wavelet significantly decreases the spatial and angular correlation of the LF which is definitely not possible to achieve by color conversion.
- SVD is used to extract a robust watermarking feature. The usage of SVD is inevitable since it has to be used in both the embedding and extraction procedure.

Clearly, each transformation allows reduction in the correlation along a specific dimension. Excluding any of these transformations (specifically the 4D wavelet) will downgrade the performance of the proposed method. In other words, the lose of performance through the removal of any transform cannot be compensated by the other remaining transforms. In order to verify the contribution of each transform, an experiment was conducted. The insertion and extraction of the watermark was performed excluding one transform while retaining all others. Thus, the simulation was carried out first without the color conversion. Another simulation was performed without the 4D wavelet. Finally another simulation was performed without using DCT. Fig. 20 shows the results of these simulations against Gaussian noise. For minor or even moderate attacks, the performance of the aforementioned methods may be quite similar. However, once the attack becomes intenser, the importance of the integrity is more pronounced. Letting the noise power of Gaussian noise

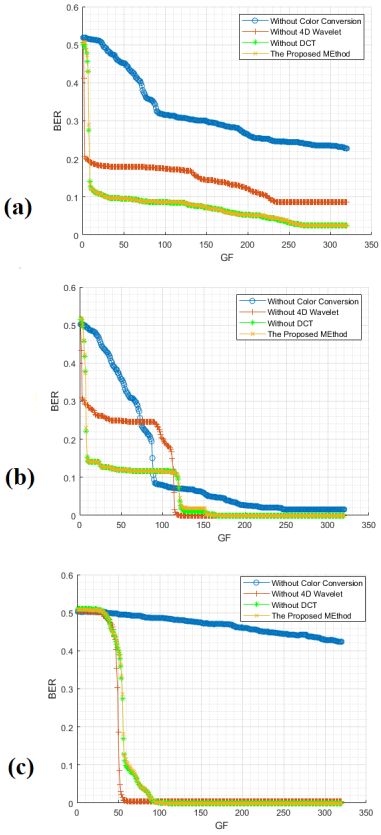

FIGURE 21. The *BER* of the watermarked LF compressed by JPEG of a) $qf = 25\%$ b) $qf = 50\%$c) $qf = 75\%$.

equal 100, 225, the performance of different methods were quite similar. Letting noise power equal 625, the importance of using all transformations becomes more important. When the watermarked LF is exposed to an aggressive attack of Gaussian noise with a noise power of 1225, the superiority of the proposed method is more emphasized. Letting $GF = 80$, the removal of color conversion and DCT will increase the *BER* by 4.88% and 6.48% respectively, whereas removal of the 4D wavelet will cause the *BER* to rise 42.28%.

Fig. 21 shows the robustness of the investigated methods against JPEG compression. As anticipated, in the case of JPEG compression with $qf = 100\%$, the distortion is minimized and no noticeable difference was observed among the mentioned methods. More drastically, once qf decreases to 75%, then the advantage of the proposed method becomes clear. Likewise, for JPEG compression with $qf = 50\%, 25\%$, the difference between the proposed method and the reduced implementation without color conversion or the 4D wavelet considerably increases. Regarding the median filtering, it is clear from Fig. 22 that the removal of DCT only downgrades the performance as compared with the proposed method. However, the removal of color conversion will cause the *BER* of the median filtered watermarked LF to rise up to 10% in case of $GF = 80$. The removal of the 4D wavelet has an even more detrimental effect on the robustness and introduces

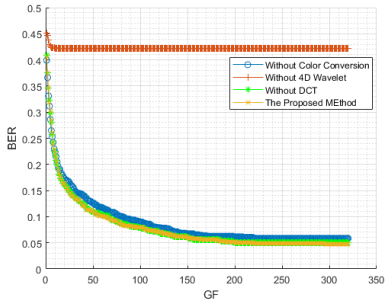

FIGURE 22. The *BER* of the watermarked LF passed through a median filter.

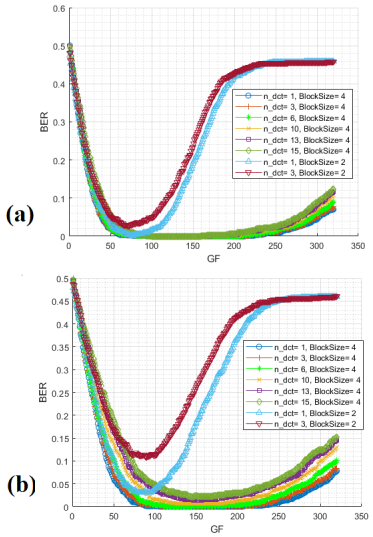

FIGURE 23. The effect of block size and number of DCT coefficients on robustness of the proposed method against Gaussian noise a) $\sigma_n^2 = 625$ b) $\sigma_n^2 = 1225$.

a *BER* of 42.19%. As for compression with JPEG 2000, no discernible difference was found except that the removal of color conversion decreases the rate of convergence of *BER* to zero.

To summarize this section, excluding any of the transformations would deteriorate the performance of the proposed method against Gaussian noise. The usage of DCT minimizes the *BER* of watermark extraction against Gaussian noise and the lowest *BER* is attained if all the transformations are used jointly. Excluding the 4D wavelet will be fatal to the robustness of the proposed method against all the attacks. In the case of median filtering, the impact of removing the 4D wavelet will be very extreme. Even though the removal of DCT does not cause major error against JPEG and JPEG 2000, it is vital to make the proposed method robust against Gaussian noise. Hence, the best performance is achieved when color conversion, DCT and 4D wavelet are all used.

E. THE INFLUENCE OF BLOCK SIZE AND NUMBER OF DCT COEFFICIENTS

For a bigger chosen block size, a higher percentage of the host LF will carries the watermark information. Hence, the alteration of *BlockSize* will affect the performance of the proposed

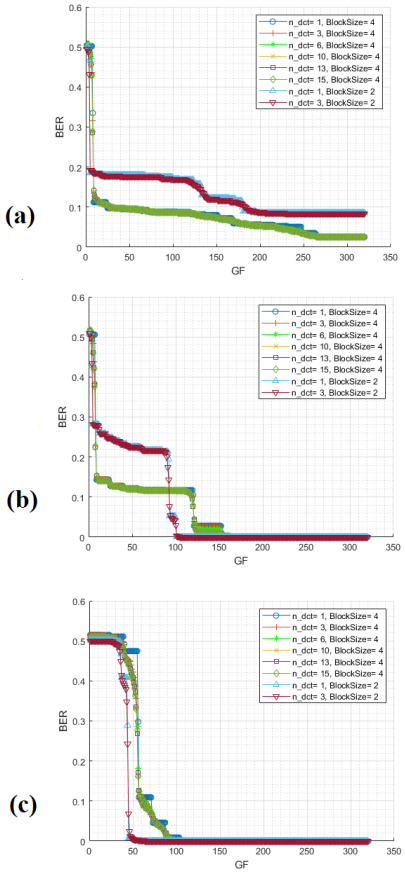

FIGURE 24. The effect of block size and number of DCT coefficients on robustness of the proposed method against JPEG compression a) *qf* = 25% b) *qf* = 50% c) JPEG compression *qf* = 75%.

method. On the other hand, *n_dct* in (19) lies in the range of 1 to *BlockSize* × *BlockSize*. The number of chosen DCT coefficients will adjust the number of high frequency components accommodating the watermark. Modifying a large portion of high frequency content of the LF compromises the transparency of the proposed method as well as creating undesirable visual distortion (e.g. considerable degradation of the edges). In order to address the effect of block size and number of DCT coefficients, all the experiments of sections III-B and III-C have been done with $BlockSize = 2, 4$ and $n\_dct = 1, 3, 6, 10, 15$. No significant difference was found for minor to moderate attacks. However, if the watermarked LF is exposed to more aggressive attacks, *BlockSize* and *n_dct* contributes considerably to the performance of the proposed method. For the sake of brevity, only the graphs which are dissimilar to those of section III-C are illustrated.

According to Fig. 23(a), embedding the watermark in 2×2 blocks will heavily deteriorate the robustness of the proposed method against intensive Gaussian noise. For Gaussian noise of power $\sigma_n^2 = 100, 225$, the *BER* graphs had a similar trend while for $\sigma_n^2 = 625$ some differences were observed. In the case of $\sigma_n^2 = 1225$ the differences would be noticeable, highlighting the effects of *BlockSize* and *n_dct*. Fig. 23(a) and Fig. 23(b) shows that embedding the watermark in 2×2

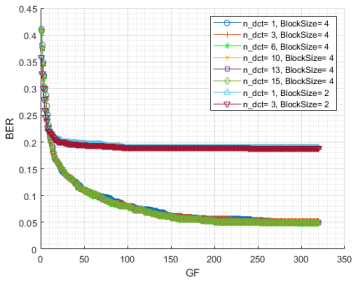

FIGURE 25. The effect of block size and number of DCT coefficients on robustness of the proposed method against median filtering.

blocks will severely reduce the robustness of the proposed method against Gaussian noise. $BlockSize = 2$ generally caused a higher *BER* specifically when n_dct is set to 3. That is to say, if the blocks are configured to have 2×2 coefficients and three of them are chosen to carry the watermark, then the Gaussian noise will have a very adverse effect on watermark extraction. If only one DCT coefficient is chosen, the gap with $BS = 4, n\_dct = 3$ becomes smaller. Regardless, letting $n\_dct = 1$ will heavily degrade the robustness of the propsed method against other attacks. This trend is fully in line with our expectation of the vulnerability of high frequency components against Gaussian noise, i.e. embedding the watermark in higher frequency components will result in lower robustness against Gaussian noise [41]. As expected, for the larger blocks of 4×4, the excessive value of $n\_dct = 10, 15$ will have a huge negative effect on robustness of the proposed method against Gaussian noise such that the graph concerning $ndct = 15$ will never converge to the value of $BER = 0$. In the case of $n\_dct = 6$, there is still a good degree of robustness against Gaussian noise, but it converges to a minimum *BER* of zero more slowly than $n\_dct = 3$. Based on our observation for $BlockSize = 2, 4$, even though the DCT is applied to wavelet coefficients, embedding the watermark into high frequency components will lower the robustness against Gaussian noise.

As per JPEG compression of $qf = 25\%$ and $GF = 80$, the *BER* introduced by $Blocksize = 2$ is 8.89% bigger than $Blocksize = 4$. As can be seen from Fig. 24(a), this gap pertains to other values of *GF* and starts decreasing when *GF* approaches excessive values. As shown in Fig.24(b), letting $qf = 50\%$ and $GF = 80$, regardless of the number of DCT coefficients $Blocksize = 2$ would yield a *BER* 10.06% higher than that of $Blocksize = 4$. In the case of $Blocksize = 2$, for *GF* values bigger than our ideal one ($GF = 80$) *BER* converges to zero more rapidly. Nevertheless, the poor performance of $Blocksize = 2$ against Gaussian noise, JPEG compression with $qf = 25\%$ and median filtering leaves it far behind $Blocksize = 4$. Fig. 24(c) shows that since JPEG compression with $qf = 75\%$ causes minor compression loss, the block dimension may become important and the *BER* introduced by $BlockSize = 4$ is 4.00% more than $BlockSize = 2$. As expected, JPEG compression with $qf = 100\%$ delivers such a good quality result, that regardless of $BlockSize$ and n_dct, the *BER* rapidly converges to zero. Generally, no significant difference was observed with the variation of n_dct for this given attack. In regards to JPEG 2000, no noticeable difference was found between $BlockSize = 2, 4$.

It is obvious from Fig. 25 that letting $BlockSize = 2$ will heavily degrade the performance of the proposed method in terms of robustness against median filtering. Provided that $BlockSize = 2$, *BER* will remain higher than 18.84%. In contrast, letting $BlockSize = 4$ will greatly improve the robustness such that *BER* will even converge to zero. Letting GF<90 and $n\_dct = 6$, the *BER* falls below 8.20%. Taking the aforementioned statements into account, it seems $BlockSize = 4, n\_dct = 6$ makes a good compromise between robustness, transparency and capacity.

IV. CONCLUSION AND FUTURE WORKS

This paper proposes a new method for LF watermarking. Using the 4D wavelet transform, the extensive intercorrelation and intracorrelation of the LF has been employed to improve the watermarking performance. The mathematical representation of both embedding and extraction procedures have been detailed and the performance of the proposed method is assessed objectively and subjectively. For $GF = 80$, the *PSNR* of the watermarked LF will be 54.77 dB which is far too high to cause any visual difference for HVS. Even for the lowest values of *GF* the watermark can be extracted free of errors. The proposed method highly preserves the 3D perception of the watermarked LF and exhibits promising robustness against Gaussian noise, JPEG compression, median filtering and JPEG 2000. The contribution of individual transforms used for watermarking has been investigated. The effect of the block size and number of DCT coefficients on watermarking performance has also been examined. It was shown that using block dimensions that are too small, will degrade the performance as well as using too few or too many DCT coefficients. Even though the proposed method employs the 4D wavelet transform for LF watermarking, it may also be used in other applications of LF realms, e.g. visualization, compression, denoising, etc. Another interesting approach to LF watermarking is the usage of the shearlet transform as it has already exhibited promising results in LF reconstruction [63]. Multi-level decomposition is another interesting experiment to conduct LF watermarking. The interested researchers can also consider adapting countourlet, curvelet, ridgelet or any other transform for development of the proposed method.

ACKNOWLEDGMENT

The authors wish to thank Przemyslaw Kopyzki and Nicolo Incardona for their kind support in the capturing and projection of the LF to the integral imaging monitor.

REFERENCES

[1] J. P. Quintais and J. Poort, "The decline of online piracy: How markets-not enforcement-drive down copyright infringement," *Amer. U. Int. L. Rev.*, vol. 34, no. 4, p. 807, 2018, Art. no. 5.

[2] J. Poort, J. Quintais, M. A. van der Ende, A. Yagafarova, and M. Hageraats, "Global online piracy study," Inst. Inf. Law, Univ. Amsterdam, Amsterdam, The Netherlands, 2018.
[3] A. Soltani Panah, R. Van Schyndel, T. Sellis, and E. Bertino, "On the properties of non-media digital watermarking: A review of state of the art techniques," *IEEE Access*, vol. 4, pp. 2670–2704, 2016.
[4] X. W. Li and I. K. Lee, "Robust copyright protection using multiple ownership watermarks," *Opt. Express*, vol. 23, no. 3, pp. 3035–3046, Feb. 2015.
[5] P. Wah Wong and N. Memon, "Secret and public key image watermarking schemes for image authentication and ownership verification," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1593–1601, Oct. 2001.
[6] J. Lee and C. Sun Won, "A watermarking sequence using parities of error control coding for image authentication and correction," *IEEE Trans. Consum. Electron.*, vol. 46, no. 2, pp. 313–317, May 2000.
[7] M. Barni, "What is the future for watermarking? (Part II)," *IEEE Signal Process. Mag.*, vol. 20, no. 6, pp. 53–59, Nov. 2003.
[8] M. Khalil and A. Adib, "Audio watermarking with high embedding capacity based on multiple access techniques," *Digit. Signal Process.*, vol. 34, pp. 116–125, Nov. 2014.
[9] J. Zain and M. Clarke, "Security in telemedicine: Issues in watermarking medical images," in *Proc. Sci. Electron., Technol. Inf. Telecommun.*, Tunisia, North Africa, 2005, p. 4.
[10] J.-P. Boyer, P. Duhamel, and J. Blanc-Talon, "Scalar DC–QIM for semifragile authentication," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 4, pp. 776–782, Dec. 2008.
[11] R. Liu and T. Tan, "An SVD-based watermarking scheme for protecting rightful ownership," *IEEE Trans. Multimedia*, vol. 4, no. 1, pp. 121–128, Mar. 2002.
[12] C.-J. Cheng, W.-J. Hwang, H.-Y. Zeng, and Y.-C. Lin, "A fragile watermarking algorithm for hologram authentication," *J. Display Technol.*, vol. 10, no. 4, pp. 263–271, Apr. 2014.
[13] A. Ansari, A. Dorado, G. Saavedra, and M. Martinez-Corral, "Plenoptic image watermarking to preserve copyright," *Proc. SPIE*, vol. 10219, May 2017, Art. no. 102190A.
[14] M. K. Arnold, M. Schmucker, and S. D. Wolthusen, *Techniques and Applications of Digital Watermarking and Content Protection*. Norwood, MA, USA: Artech House, 2003.
[15] N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal Process.*, vol. 66, no. 3, pp. 385–403, May 1998.
[16] Q. Cheng and T. S. Huang, "Robust optimum detection of transform domain multiplicative watermarks," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 906–924, Apr. 2003.
[17] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Process.*, vol. 66, no. 3, pp. 357–372, May 1998.
[18] W. C. Chu, "DCT-based image watermarking using subsampling," *IEEE Trans. Multimedia*, vol. 5, no. 1, pp. 34–38, Mar. 2003.
[19] S. Liu, Z. Pan, and H. Song, "Digital image watermarking method based on DCT and fractal encoding," *IET Image Process.*, vol. 11, no. 10, pp. 815–821, Oct. 2017.
[20] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 783–791, May 2001.
[21] Z. Dawei, C. Guanrong, and L. Wenbo, "A chaos-based robust wavelet-domain watermarking algorithm," *Chaos, Solitons Fractals*, vol. 22, no. 1, pp. 47–54, Oct. 2004.
[22] T.-S. Nguyen, C.-C. Chang, and X.-Q. Yang, "A reversible image authentication scheme based on fragile watermarking in discrete wavelet transform domain," *AEU-Int. J. Electron. Commun.*, vol. 70, no. 8, pp. 1055–1061, Aug. 2016.
[23] R. Mehta, N. Rajpal, and V. P. Vishwakarma, "Robust image watermarking scheme in lifting wavelet domain using GA-LSVR hybridization," *Int. J. Mach. Learn. Cybern.*, vol. 9, no. 1, pp. 145–161, Jan. 2018.
[24] M. Rabizadeh, M. Amirmazlaghani, and M. Ahmadian-Attari, "A new detector for contourlet domain multiplicative image watermarking using bessel k form distribution," *J. Vis. Commun. Image Represent.*, vol. 40, pp. 324–334, Oct. 2016.
[25] Q. Su, G. Wang, G. Lv, X. Zhang, G. Deng, and B. Chen, "A novel blind color image watermarking based on contourlet transform and hessenberg decomposition," *Multimedia Tools Appl.*, vol. 76, no. 6, pp. 8781–8801, Mar. 2017.
[26] R. S. R. Channapragada and M. V. Prasad, "Watermarking techniques in curvelet domain," in *Computational Intelligence in Data Mining*, vol. 1. New Delhi, India: Springer, 2015, pp. 199–211.
[27] W.-H. Kim, S.-H. Nam, J.-H. Kang, and H.-K. Lee, "Robust watermarking in curvelet domain for preserving cleanness of high-quality images," *Multimedia Tools Appl.*, vol. 78, no. 12, pp. 16887–16906, Jun. 2019.
[28] C. R. S. Rao and M. V. Prasad, *Digital Watermarking Techniques in Curvelet and Ridgelet Domain*. Berlin, Germany: Springer, 2016.
[29] V. S. Verma, A. Bhardwaj, and R. K. Jha, "A new scheme for watermark extraction using combined noise-induced resonance and support vector machine with PCA based feature reduction," *Multimedia Tools Appl.*, vol. 78, pp. 1–22, May 2019.
[30] I. A. Ansari, M. Pant, and C. W. Ahn, "SVD based fragile watermarking scheme for tamper localization and self-recovery," *Int. J. Mach. Learn. Cybern.*, vol. 7, no. 6, pp. 1225–1239, Dec. 2016.
[31] J. Wang and S. Lian, "On the hybrid multi-watermarking," *Signal Process.*, vol. 92, no. 4, pp. 893–904, Apr. 2012.
[32] C. Song, S. Sudirman, and M. Merabti, "A robust region-adaptive dual image watermarking technique," *J. Vis. Commun. Image Represent.*, vol. 23, no. 3, pp. 549–568, Apr. 2012.
[33] W.-H. Lin, Y.-R. Wang, S.-J. Horng, T.-W. Kao, and Y. Pan, "A blind watermarking method using maximum wavelet coefficient quantization," *Expert Syst. Appl.*, vol. 36, no. 9, pp. 11509–11516, Nov. 2009.
[34] V. Sachnev, H. Joong Kim, J. Nam, S. Suresh, and Y. Qing Shi, "Reversible watermarking algorithm using sorting and prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 7, pp. 989–999, Jul. 2009.
[35] Y. Wang, J. Ostermann, and Y.-Q. Zhang, *Video Processing and Communications*, vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.
[36] E. H. Adelson and J. R. Bergen, *The Plenoptic Function and the Elements of Early Vision*, vol. 2. Cambridge, MA, USA: Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology, 1991.
[37] C. Heinze, S. Spyropoulos, S. Hussmann, and C. Perwass, "Automated robust metric calibration algorithm for multifocus plenoptic cameras," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 5, pp. 1197–1205, May 2016.
[38] A. Koz, C. Çığla, and A. A. Alatan, "Watermarking for light field rendering," in *Proc. IEEE 15th Eur. Signal Process. Conf.*, Sep. 2007, pp. 2296–2300.
[39] P. Paudyal, F. Battisti, A. Neri, and M. Carli, "A study of the impact of light fields watermarking on the perceived quality of the refocused data," in *Proc. 3DTV-Conf., True Vis. Capture, Transmiss. Display 3D Video (3DTV-CON)*, Jul. 2015, pp. 1–4.
[40] Y. Lu, S. You, W. Zhang, B. Yang, R. Peng, and S. Zhuang, "Watermarking scheme for microlens-array-based four-dimensional light field imaging," *Appl. Opt.*, vol. 55, no. 13, pp. 3397–3404, 2016.
[41] A. Ansari, S. Hong, G. Saavedra, B. Javidi, and M. Martinez-Corral, "Ownership protection of plenoptic images by robust and reversible watermarking," *Opt. Lasers Eng.*, vol. 107, pp. 325–334, Aug. 2018.
[42] S. Kishk and B. Javidi, "3D object watermarking by a 3D hidden object," *Opt. Express*, vol. 11, no. 8, pp. 874–888, 2003.
[43] A. Koz, C. Cigla, and A. A. Alatan, "Watermarking of free-view video," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1785–1797, Jul. 2010.
[44] A. Koz, C. Cigla, and A. A. Alatan, "Free-view watermarking for Free-view television," in *Proc. Int. Conf. Image Process.*, Oct. 2006, pp. 1405–1408.
[45] A. A. Reddy and B. N. Chatterji, "A new wavelet based logo-watermarking scheme," *Pattern Recognit. Lett.*, vol. 26, no. 7, pp. 1019–1027, May 2005.
[46] A. Shaamala, S. M. Abdullah, and A. A. Manaf, "Study of the effect of DCT and DWT domains on the imperceptibility and robustness of genetic watermarking," *Int. J. Comput. Sci. Issues*, vol. 8, no. 5, p. 220, 2011.
[47] M. A. Magnor, A. Endmann, and B. Girod, "Progressive compression and rendering of light fields," in *Proc. Vis., Modeling, Vis.*, 2000, pp. 199–204.
[48] A. K. Parthasarathy and S. Kak, "An improved method of content based image watermarking," *IEEE Trans. Broadcast.*, vol. 53, no. 2, pp. 468–479, Jun. 2007.
[49] P. S. Huang, C.-S. Chiang, C.-P. Chang, and T.-M. Tu, "Robust spatial watermarking technique for colour images via direct saturation adjustment," *IEE Proc.-Vision, Image Signal Process.*, vol. 152, no. 5, pp. 561–574, 2005.
[50] G. K. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 30–44, Apr. 1991.
[51] A. B. Watson, "Image compression using the discrete cosine transform," *Math. J.*, vol. 4, no. 1, p. 81, 1994.

[52] B. E. Wohlberg and G. de Jager, "Fast image domain fractal compression by DCT domain block matching," *Electron. Lett.*, vol. 31, no. 11, pp. 869–870, May 1995.
[53] Z. Xiong, K. Ramchandran, M. T. Orchard, and Y.-Q. Zhang, "A comparative study of DCT- and wavelet-based image coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 5, pp. 692–695, Aug. 1999.
[54] G. J. Sullivan and T. Wiegand, "Video compression-from concepts to the H.264/AVC standard," *Proc. IEEE*, vol. 93, no. 1, pp. 18–31, Jan. 2005.
[55] S. Liu and A. C. Bovik, "Efficient DCT-domain blind measurement and reduction of blocking artifacts," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 12, pp. 1139–1149, Dec. 2002.
[56] D. Santa-Cruz, R. Grosbois, and T. Ebrahimi, "JPEG 2000 performance evaluation and assessment," *Signal Process., Image Commun.*, vol. 17, no. 1, pp. 113–130, Jan. 2002.
[57] 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, Sep. 2014.
[58] M. Kutter and S. Winkler, "A vision-based masking model for spread-spectrum image watermarking," *IEEE Trans. Image Process.*, vol. 11, pp. 16–25, 2002.
[59] J. Franco, G. Bernabé, J. Fernández, and M. E. Acacio, "A parallel implementation of the 2D wavelet transform using CUDA," in *Proc. 17th Euromicro Int. Conf. Parallel, Distrib. Netw. Process.*, 2009, pp. 111–118.
[60] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
[61] S. Mallat, *A Wavelet Tour of Signal Processing*. Amsterdam, The Netherlands: Elsevier, 1999.
[62] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[63] S. Vagharshakyan, R. Bregovic, and A. Gotchev, "Light field reconstruction using shearlet transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 1, pp. 133–147, Jan. 2018.

AMIR ANSARI received the master's degree in electrical engineering-communication systems from the Shiraz University of Technology, Iran, in 2014. He is currently pursuing the Ph.D. degree with the University of Valencia. During his master's, he was mainly focused on remote sensing, image processing, and video coding. He also has experience in hardware design for GPS-based remote positioning and telemetry systems. He joined the 3D Imaging and Display Group (3DID), University of Valencia, in September 2016. His research interests include (but are not limited to) light field/hyperspectral image processing/watermarking/compression/fusion, hardware design, and communication systems.

GENARO SAAVEDRA received the M.Sc. and Ph.D. *(cum laude)* degrees in physics from the University of Valencia, Spain, in 1990 and 1996, respectively. He is currently a Full Professor and has been the Co-Leader of the 3D Imaging and Display Laboratory, University of Valencia, since 1999. His research interests are optical diffraction, plenoptic/integral imaging techniques, and 3D high-resolution optical microscopy. He has supervised nine Ph.D. theses (two of them honored with the Best Thesis Award) on these topics. He has published about 120 technical articles on these topics in major journals. He has contributed in more than 80 conference proceedings, including more than 50 invited/keynote presentations. He has filed 15 patents on 3D display and imaging techniques.

MANUEL MARTINEZ-CORRAL was born in Spain, in 1962. He received the Ph.D. degree in physics from the University of Valencia, in 1993, which honored him with the Ph.D. Extraordinary Award. He is currently a Full Professor of Optics with the University of Valencia, where he co-leads the "3D Imaging and Display Laboratory." His teaching experience includes lectures and supervision of laboratory experiments for undergraduate students on geometrical optics, optical instrumentation, diffractive optics, and image formation. He gave lectures on diffractive optics for Ph.D. students. He is a Fellow of the SPIE since 2010 and has been a Fellow of the OSA since 2017. His research interests include microscopic and macroscopic 3D imaging and display technologies. He has supervised 17 Ph.D. students (three honored with the Ph.D. Extraordinary Award) on these topics, published over hundred and twenty technical articles in major journals (which have been cited more than 2.900 times, H-index=28), and pronounced over seventy invited and keynote presentations in international meetings. He is the Co-Chair of the Three-Dimensional Imaging, Visualization, and Display Conference within the SPIE meeting in Defense, Security, and Sensing. He has been the Topical Editor of the IEEE/OSA JOURNAL OF DISPLAY TECHNOLOGY, and is the Topical Editor of the *Applied Optics* (OSA).