A Random Matrix Approach to the Forensic Analysis of Upscaled Images

David Vázquez-Padín, *Member, IEEE,* Fernando Pérez-González, *Fellow, IEEE,*
and Pedro Comesaña-Alfaro, *Senior Member, IEEE*

Abstract—The forensic analysis of resampling traces in upscaled images is addressed via subspace decomposition and random matrix theory principles. In this context, we derive the asymptotic eigenvalue distribution of sample autocorrelation matrices corresponding to genuine and upscaled images. To achieve this, we model genuine images as an autoregressive random field and we characterize upscaled images as a noisy version of a lower dimensional signal. Following the intuition behind Marčenko-Pastur law, we show that for upscaled images, the gap between the eigenvalues corresponding to the low-dimensional signal and the ones from the background noise can be enhanced by extracting a small number of consecutive columns/rows from the matrix of observations. In addition, using bounds provided by the same law for the eigenvalues of the noise space, we propose a detector for exposing traces of resampling. Finally, since an interval of plausible resampling factors can be inferred from the position of the gap, we empirically demonstrate that by using the resulting range as the search space of existing estimators (based on different principles), a better estimation accuracy can be attained with respect to the standalone versions of the latter.

Index Terms—Image forensics, Marčenko-Pastur law, random matrix theory, resampling detection, resampling estimation.

I. Introduction

THE ubiquity of the Internet together with the prevalence of multimedia technology in our society have fostered the usage of digital images in our daily life. Nowadays, digital images serve as a means of communication and can be used to disseminate knowledge in e-learning platforms, to inform people across the media, or as an instrument to share experiences in social networks, among other applications. Their growing importance has unavoidably been tied to the development of editing software tools that were originally devised to enhance image quality, but that now enable an unskilled person to manipulate any captured content. This state of affairs has boosted a major concern about the reliability of digital images and, as a response, a number of techniques have emerged in the last years to automatically analyze the authenticity or integrity of a given image in a blind way [1].

D. Vázquez-Padín, F. Pérez-González, and P. Comesaña-Alfaro are with the Department of Signal Theory and Communications, University of Vigo, 36310 Vigo, Spain (e-mail: dvazquez@gts.uvigo.es; fperez@gts.uvigo.es; pcomesan@gts.uvigo.es)

The main idea behind the study of resampling traces to unveil tampering arises from the fact that credible forgeries are often performed by geometrically adapting the new added contents to the original scene. Such adaptation may require the utilization of spatial transformations that involve the use of resampling operations which inherently leave characteristic traces that are not typically present in genuine contents.

The forensic analysis of resampled signals has been largely investigated in recent years and many different directions have been explored to detect resampling traces from an observed image block [1, Sect. 5.1.2]. Despite this variety of approaches, most of the available detectors share a common processing structure. Generally, as a first step, a residue signal from the observed image is extracted, where resampling artifacts are observable. Depending on the foundations of each method, this signal can be obtained either by a global predictor [2], by prefiltering the image with a fixed linear filter [3] or a derivative one (e.g., [4], [5], [6], [7]), or by computing the difference of predictor coefficients stemming from adjacent rows/columns of the image [8]. Then, given that the resampling artifacts vary periodically, a postprocessing step is applied in the frequency domain to detect the presence of spectral peaks (related to these periodicities) and, finally, a decision on whether the observed block has been resampled is rendered. Some approaches (e.g., [9], [10], [11]), avoid the postprocessing part in the frequency domain and check if a set of candidate resampling factors satisfies the underlying linear structure induced by the application of a resampling process. Other detectors, instead, bypass the frequency-based analysis by making use of a Support Vector Machine (SVM) to take the final decision. For example, in [12] a set of features is gathered from the singular value decomposition of the observed image, while in [13] the normalized energy density for different window sizes of the image is used.

Although initial works studying resampling traces were mainly focused on resampling detection, several approaches have also been oriented towards the identification of the particular geometric transformation undergone by an image. For instance, the estimation of the scaling factor applied to an image has been analytically investigated in [14] together with the estimation of the applied rotation angle, which has been further assessed in [15] and [16]. Later, a more general solution has been derived in [17], where a frequency-based analysis is used to estimate an arbitrary affine transformation. On the other hand, the authors of [18] have figured out how to solve some of the ambiguities that show up when facing the resampling factor estimation in the frequency domain and,

more recently, other research works such as those in [19] and [20], have been able to identify the particular interpolation kernel used during the resampling process.

It is clear from the above description that at some point, most of the mentioned techniques perform an analysis in the frequency domain for resampling detection or resampling factor estimation. However, as indicated in [11], frequency-based analyses present some drawbacks impairing the performance of the proposed approaches when a reduced number of samples is available. On the other hand, exhaustive-search detectors (e.g., [9]), may suffer from high positive rates due to the need of a large set of candidates to cover all plausible resampling factors. Similarly, SVM-based approaches (e.g., [12], [13]), need to build up a large and representative number of training samples which might be difficult to obtain.

To overcome these deficiencies, in this paper we follow up the new direction hinted in our previous work [21], where the particular signal-plus-noise structure of upscaled images has been highlighted. In that case, the distinct evolution of the singular values of genuine images with respect to their upscaled counterpart, enabled the definition of a statistic to distinguish them. To deepen the understanding of such distinct evolution, we resort to Random Matrix Theory (RMT), which provides useful tools for modeling the behavior of the eigenvalues and singular values of random matrices [22]. This theory finds applications in many different fields, such as signal processing and communications, where a considerable volume of works has been recently published [23], which attracted our interest.

In our case, RMT offers a proper way to numerically derive the asymptotic eigenvalue distribution of the $N \times N$ sample autocorrelation matrix corresponding to a genuine image or its upscaled version, as $N \to \infty$. Moreover, landmark contributions from this theory, such as the Marčenko-Pastur law [24], suggest us to process submatrices containing $K < N$ consecutive columns/rows from the $N \times N$ matrix of observations as a means to sharpen the underlying signal-plus-noise structure in upscaled images. Coupling this property with the bounds provided by the same law to the eigenvalues of the noise space, a new resampling detector is proposed. On the other hand, focusing on the gap that marks the transition between the signal subspace and the noise, we propose an alternative form to accomplish resampling factor estimation. By simply combining our strategy with existing approaches (e.g., [6], [14]), we achieve a substantially better estimation accuracy, because their advantages complement each other.

In this paper, we do not theoretically examine the downscaling case (i.e., a reduction in image size), because its analysis is formally quite different. Nevertheless, for the sake of completeness, in Sect. VI-A we do test the proposed approach in a downscaling scenario and compare its performance against the state-of-the-art detector in [8]. Hence, this work should be considered as a first step towards the complete characterization of the resampling problem in terms of RMT, leaving the analysis of downscaling for a future work.

The remainder of the paper is organized as follows. In Sect. II, the resampling process is formally described, introducing the signal-plus-noise decomposition of upscaled images. Drawing on the technical report that complements this

TABLE I
SUMMARY OF NOTATION

| Symbol	Meaning
$\mathbf{X}$	Matrix representing a genuine image
$\mathbf{H}$	Matrix containing coefficients of an interpolation kernel
$\mathbf{Y}$	Matrix representing an upscaled image
$\mathbf{W}$	Random matrix containing i.i.d. white noise entries
$\mathbf{Z}$	Matrix representing an upscaled and quantized image
ξ	Resampling factor ($\xi > 1$)
k_w	Width of the interpolation kernel ($k_w \in 2\mathbb{Z}^+$)
Δ	Step size of the uniform scalar quantizer
ρ	Correlation coefficient of the autoregressive model
β	Aspect ratio of an $N \times K$ matrix, i.e., $\beta \triangleq \frac{K}{N}$
P	Position of the gap in the evolution of the eigenvalues
For an arbitrary $N \times N$ matrix $\mathbf{A}$:	
$A_{i,j}$	(i,j)-th element of matrix $\mathbf{A}$, with $i,j \in \{0, \dots, N-1\}$
σ_A^2	Variance of the entries of $\mathbf{A}$
$\mathbf{A}_K$	$N \times K$ submatrix with K consecutive columns from $\mathbf{A}$
$\mathbf{\Sigma}_{A_K}$	Renormalized sample autocorrelation matrix of $\mathbf{A}_K$
$\lambda_i(\mathbf{\Sigma}_{A_K})$	i-th eigenvalue of $\mathbf{\Sigma}_{A_K}$, with $i = 1, \dots, N$
$\lambda_-(\mathbf{\Sigma}_{A_K})$	Smallest eigenvalue of $\mathbf{\Sigma}_{A_K}$
$\lambda_+(\mathbf{\Sigma}_{A_K})$	Largest eigenvalue of $\mathbf{\Sigma}_{A_K}$
$f_{\mathbf{\Sigma}_{A_K}}(\lambda)$	Probability density function of the eigenvalues of $\mathbf{\Sigma}_{A_K}$
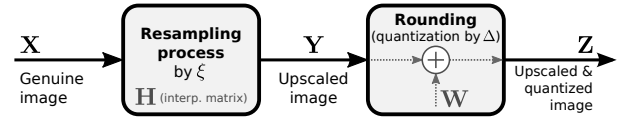
Fig. 1. Block diagram of the model we use for the resampling process.

work [25], Sect. III characterizes the eigenvalue distribution for genuine and upscaled images. In Sect. IV, the proposed detector and the strategy developed for estimating the resampling factor are specifically detailed, while their experimental validation is carried out in Sect. V. A robustness analysis of the proposed detector in more realistic scenarios is tackled in Sect. VI. Finally, the conclusions are summarized in Sect. VII.

Notation: Given a matrix $\mathbf{X}$ of size $N \times K$, $K \leq N$, its sample autocorrelation matrix is $K^{-1}\mathbf{X}\mathbf{X}^T$. Instead, we will often work with the renormalized sample autocorrelation $\mathbf{\Sigma}_X \triangleq N^{-1}\mathbf{X}\mathbf{X}^T$. The i-th eigenvalue of an arbitrary matrix $\mathbf{A}$ is denoted by $\lambda_i(\mathbf{A})$, whereas its i-th singular value is $\sigma_i(\mathbf{A})$. Both $\lambda_i(\mathbf{A})$ and $\sigma_i(\mathbf{A})$ are sorted in descending order. For the sake of clarity, the notation used throughout the paper is summarized in Table I together with an illustration of the considered resampling model in Fig. 1.

II. RESAMPLING PROCESS AND PROBLEM FORMULATION

Let $\mathbf{X}$ be the matrix representation of an image with a single color channel. Unless otherwise stated, we will assume that $\mathbf{X}$ has not been subject to any prior resampling operation, so we will refer to it as a genuine or natural image. The resampling operation is assumed to be linear, so each pixel value in the resampled image is computed by linearly combining a finite set of neighboring samples coming from the genuine image $\mathbf{X}$. The process of resampling involves two main steps: the definition of the resampling grid with the new pixel locations and the computation of the intensity values in those new locations. For the first operation, we assume an affine

transformation of the original axes $(u,v) \in \mathbb{Z}^2$ to $\mathbf{A}(u,v)^T + \mathbf{b}$, where $\mathbf{A}$ is a matrix that embodies the linear transformation (e.g., scaling, rotation, etc.) and $\mathbf{b}$ represents the translation vector. In this paper, we will consider that the resampling operation uniformly upscales each dimension of the genuine image $\mathbf{X}$ by a resampling factor $\xi > 1$, so $\mathbf{A} = \xi \mathbf{I}_2$, where $\mathbf{I}_2$ is the 2×2 identity and ξ is rational, i.e., $\xi \triangleq \frac{L}{M}$ with L and M relatively prime natural numbers. In addition, a homogeneous translation through $\mathbf{b} \triangleq (\phi, \phi)^T$ is generally applied with $\phi = \frac{1}{2}\left(\frac{M}{L} + 1\right)$, such that the sampling points of the resampled image are centered with respect to the grid of the original image. A more detailed analysis on the role of ϕ in the pixel grid alignment is given in [26].

For the second step, the calculation of the intensity values in the new resampled grid can be implemented by using different two-dimensional (2D) interpolation kernels. Here, we study symmetric linear kernels that are *separable* (i.e., they can be interpreted as the successive application of the same one-dimensional (1D) function across each dimension of the grid), given that they are the most common in practice. We leave out more complex interpolation algorithms, such as adaptive or non-linear, since their use is typically constrained to perform demosaicing and they are rarely employed to resize images.

Summarizing the steps described above, each (i,j)-th pixel value $Y_{i,j}$ from an upscaled image can be obtained as

$$Y_{i,j} = \sum_{u,v \in \mathbb{Z}} h\left(i\frac{M}{L} + \phi - u\right) h\left(j\frac{M}{L} + \phi - v\right) X_{u,v}, \quad (1)$$

where $X_{u,v}$ denotes the (u,v)-th element of the genuine image $\mathbf{X}$ and $h(\cdot)$ is the 1D impulse response of the underlying continuous interpolation kernel, whose width is denoted by $k_w \in 2\mathbb{Z}^+$. Given a genuine image $\mathbf{X}$ of size $U \times V$, our analysis will be performed on a block $\mathbf{Y}$ of size $N \times N$ from the upscaled image by ξ, i.e., with $N \leq \xi \cdot \min(U,V)$. For the sake of readability, we will assume that N is a multiple of L and also that both blocks $\mathbf{X}$ and $\mathbf{Y}$ are aligned so that their upper-left corner corresponds to the $(0,0)$ element. Hence, using (1), the $N \times N$ block $\mathbf{Y}$ can be written in matrix form:

$$\mathbf{Y} = \mathbf{H}\hat{\mathbf{X}}\mathbf{H}^T, \quad (2)$$

where the entries of $\mathbf{H}$ are given by

$$H_{i,j} = h\left(i\frac{M}{L} + \varphi - j\right), \quad (3)$$

with $\varphi \triangleq \phi + k_w/2 - 1$, such that $\hat{\mathbf{X}}$ only contains the $R \times R$ samples of $\mathbf{X}$ that are involved in the computation of $\mathbf{Y}$, and $R \triangleq N\frac{M}{L} + k_w$, i.e., $\hat{X}_{u,v} = X_{u,v}$, for $u, v = 0, \ldots, R-1$.

It is important to note that, by construction, the size of $\mathbf{H}$ is $N \times R$, so we will assume that for sufficiently large N, $\mathbf{H}$ is full column rank (i.e., $\text{rank}(\mathbf{H}) = R$), which is true for all kernels we tested; for a more detailed discussion about the conditions that $\mathbf{H}$ must satisfy to have $\text{rank}(\mathbf{H}) = R$, the reader is referred to [27].

As a last step, after all the pixels of the resampled image have been obtained, we assume that the interpolated values are rounded or, in general, quantized by a uniform scalar quantizer with step size Δ to fit a particular bit depth. If we denote by $\mathbf{Z}$ the resulting quantized image, we can model it as

$$\mathbf{Z} = \mathbf{Y} + \mathbf{W}, \quad (4)$$

where matrix $\mathbf{W}$ stands for the quantization noise. We assume an additive white noise model in (4), such that the entries of $\mathbf{W}$ are i.i.d. $\mathcal{U}\left[-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$ with zero mean and variance $\sigma_W^2 = \frac{\Delta^2}{12}$. It is commonly accepted that this assumption holds whenever the probability density function (pdf) of the input signal is smooth and its variance is much larger than the square quantization step, which is generally the case for real images. Note, however, that this model will not be valid for image blocks containing flat regions (e.g., they may appear when capturing uniform areas or when undesired artifacts, such as saturation, occur). Hence, these special cases will be dealt with specifically when deriving the proposed approaches.

A. Subspace Decomposition to Expose Resampling Traces

In this section, we present somewhat informally the main ideas in the paper. Recall that $\text{rank}(\mathbf{H}) = R \approx N/\xi$. Thus, the rank of $\mathbf{Y}$ in (2) will be at most R, so at least $N - R$ columns (or rows) will be linearly dependent. In other words, the columns/rows from an $N \times N$ unquantized upscaled image (considered as N-dimensional vectors) will lie in an R-dimensional subspace. On the other hand, the noise $\mathbf{W}$ in (4) is expected to span the full space. Thus, through a Singular Value Decomposition (SVD) it is possible to extract the signal subspace and, upon determining its dimensionality, detect the presence of resampling traces. This is in fact the idea that we followed in [21], where the SVD was used to solve this forensic problem. Prior to that, Kao *et al.* proposed in [9] a different way to detect the signal subspace: by first computing a series of zeroing masks from a set of candidate resampling factors (each mask contains a basis vector of the orthogonal complement of a candidate signal subspace), an exhaustive search over this set is performed, and the presence of the signal subspace is finally exposed whenever a candidate matches the applied resampling factor. This idea works reasonably well, but the detector performance strongly depends on the selected set of candidates, which is unknown in a realistic forensic scenario.

On the other hand, the heuristic approach derived in [21] performs remarkably well detecting traces of resampling, but the same idea cannot be successfully exploited to estimate the applied resampling factor, because the ordered singular values from a square image block $\mathbf{Z}$ as in (4) do not show a clear break between the significant components and the noisy ones. Moreover, a measure of saturated pixels is needed in [21] to deal with linearly dependent regions, but it is not sufficiently robust to recognize non-saturated uniform regions, which also hinders the detectability of resampling traces using the SVD.

To circumvent these problems, we focus on the analytical characterization of the singular value distribution of an image block, by modeling $\mathbf{Z}$ as a random field. This allows us to draw connections with the fascinating theory of random matrices [28], [29]. As it can be checked in [22] and [23], there exist a significant number of powerful theorems that ensure the convergence of empirical eigenvalue distributions to deterministic functions. In fact, our key idea comes from the surprising fact that taking from $\mathbf{Z}$ a submatrix $\mathbf{Z}_K$ of size $N \times K$, $K \leq N$, helps to discriminate the signal and noise

subspaces. This property is illustrated in the most-celebrated result from RMT, Marčenko-Pastur Law (MPL) [24], which is stated in the following theorem.

Theorem (Marčenko-Pastur Law). *Let $\mathbf{A}$ be an $N \times N$ random matrix whose entries are i.i.d. zero-mean random variables with variance σ_A^2 and consider a submatrix $\mathbf{A}_K$ of size $N \times K$, $K \leq N$, with K consecutives columns from $\mathbf{A}$. As $K, N \to \infty$ with $\frac{K}{N} \to \beta \in (0, 1]$, the eigenvalues of the renormalized sample autocorrelation matrix $\mathbf{\Sigma}_{A_K} \triangleq N^{-1} \mathbf{A}_K \mathbf{A}_K^T$ are distributed according to the following probability density function*

$$f_{\mathbf{\Sigma}_{A_K}}(\lambda) =$$

$$\begin{cases} \frac{\sqrt{\left(\lambda_+ (\mathbf{\Sigma}_{A_K}) - \lambda\right)\left(\lambda - \lambda_- (\mathbf{\Sigma}_{A_K})\right)}}{2 \pi \sigma_A^2 \lambda \beta}, & \text{if } \lambda_- (\mathbf{\Sigma}_{A_K}) \leq \lambda \leq \lambda_+ (\mathbf{\Sigma}_{A_K}) \\ 0, & \text{otherwise} \end{cases},$$

with a point mass $(1 - \beta)$ at the origin if $\frac{K}{N} \to \beta < 1$. The limits of the domain of $f_{\mathbf{\Sigma}_{A_K}}(\lambda)$ are given by

$$\lambda_\pm (\mathbf{\Sigma}_{A_K}) = \sigma_A^2 \left(1 \pm \sqrt{\beta}\right)^2.$$

For our purposes, the most important consequence of this law is that the eigenvalues of $\mathbf{\Sigma}_{Z_K} \triangleq N^{-1} \mathbf{Z}_K \mathbf{Z}_K^T$ cluster around the variance σ_Z^2 as the matrix aspect ratio β converges to zero, i.e., as $K/N \to 0$.

If, for large N, the rank of $\mathbf{Y}$ were approximately N/ξ, for submatrix $\mathbf{Y}_K$ the rank would be approximately K/ξ. Therefore, from the eigendecomposition of $\mathbf{\Sigma}_{Z_K}$, we count on finding approximately K/ξ leading eigenvalues corresponding to the underlying interpolated signal and the remaining portion of nonzero eigenvalues will correspond to the noise space. Moreover, by taking a relatively small number of columns K in $\mathbf{Z}_K$, we can expect the signal eigenvalues to cluster around σ_Z^2, while the noise eigenvalues will do so around σ_W^2. Hence, such clustering will make more evident the existence of a low-dimensional signal subspace when resampling is in place. Note that the above discussion easily extends to the case where K rows (instead of columns) are extracted from $\mathbf{Z}$, i.e., we would only need to redefine $\mathbf{Z}$ by transposing the original.

To illustrate whether these findings may also apply in a practical setting with real images, we plot the eigenvalues against their indices (i.e., the so-called "scree plot" [30]) of $\mathbf{\Sigma}_{Z_K}$ for two different aspect ratios K/N. In order to highlight the main differences with respect to the non-interpolated case, we compare the scree plots obtained when $\mathbf{Z}_K$ is constructed from an upscaled image by $\xi = 2$ with a linear kernel and when $\mathbf{Z}_K$ comes from a genuine image. For this analysis, we use the image block shown in Fig. 2(a).

In the first case, we work with the full $N \times N$ block $\mathbf{Z}$ with $N = 512$, which leads us to the same scenario studied in [21]. As can be seen in Fig. 2(b), when $K = N$, only an "S-shape" can be discerned in the evolution of the eigenvalues for the upscaled image with a very smooth drop around $i = 230$ (i.e., nearby K/ξ) which barely hints the transition between the low-dimensional signal and the noise space.

TABLE II
IMPULSE RESPONSE AND WIDTH OF SEVERAL INTERPOLATION KERNELS.

Kernel Type	Impulse Response
Linear ($k_w = 2$)	$h(t) = \begin{cases} 1 - \lvert t \rvert, & \text{if } \lvert t \rvert \leq \frac{k_w}{2} \\ 0, & \text{otherwise} \end{cases}$
Catmull-Rom ($k_w = 4$)	$h(t) = \begin{cases} \frac{3}{2} \lvert t \rvert^3 - \frac{5}{2} \lvert t \rvert^2 + 1, & \text{if } \lvert t \rvert \leq \frac{k_w}{4} \\ -\frac{1}{2} \lvert t \rvert^3 + \frac{5}{2} \lvert t \rvert^2 - 4 \lvert t \rvert + 2, & \text{if } \frac{k_w}{4} < \lvert t \rvert \leq \frac{k_w}{2} \\ 0, & \text{otherwise} \end{cases}$
B-spline ($k_w = 4$)	$h(t) = \begin{cases} \frac{1}{2} \lvert t \rvert^3 - \lvert t \rvert^2 + \frac{2}{3}, & \text{if } \lvert t \rvert \leq \frac{k_w}{4} \\ -\frac{1}{6} \lvert t \rvert^3 + \lvert t \rvert^2 - 2 \lvert t \rvert + \frac{4}{3}, & \text{if } \frac{k_w}{4} < \lvert t \rvert \leq \frac{k_w}{2} \\ 0, & \text{otherwise} \end{cases}$
Lanczos ($k_w = 6$)	$h(t) = \begin{cases} \text{sinc}(t) \text{sinc}\left(\frac{t}{3}\right), & \text{if } \lvert t \rvert < \frac{k_w}{2} \\ 0, & \text{otherwise} \end{cases}$

In the second case, the submatrix $\mathbf{Z}_K$ is obtained by extracting the first $K = 64$ columns from the observed block $\mathbf{Z}$. Contrary to the former case, we can observe in Fig. 2(c) that now a clear gap in the evolution of the eigenvalues for the upscaled image becomes apparent. The gap marks the transition at $i = 33$ (i.e., very close to K/ξ) between the signal subspace and the background noise. This is a consequence of the compaction of the eigenvalues of $\mathbf{\Sigma}_{Z_K}$ predicted by MPL. On the other hand, since the underlying signal-plus-noise structure is generally not present in the non-interpolated case, there is no visible gap for genuine images.

The above discussion is formalized in Sect. III with a stochastic model for $\mathbf{X}$ and a characterization of the eigenvalue distribution for genuine and upscaled images. Further, in Sect. IV, we explain how the described eigenvalue compaction property (fostered by the elimination of columns/rows from the observed block $\mathbf{Z}$), can be exploited to expose resampling traces. Although the proposed solution is not optimal due to the loss of information after discarding samples from $\mathbf{Z}$, the design of an optimal approach still remains an open problem and we conjecture that it will probably require an exhaustive search strategy, which is eventually less convenient in practice.

B. Interpolation Kernels

The forthcoming analysis is applicable to any 2D linear interpolation kernel of separable nature; however, in order to illustrate some interesting properties and the different behavior among individual kernels, we take as examples the most common types of kernels available in any editing software tool, such as linear, cubic, and Lanczos (i.e., truncated sinc). From the family of cubic filters described in [31], we select two well-known cardinal splines: the Catmull-Rom spline and the cubic B-spline, both having uniformly spaced knots. As Lanczos kernel, we take a three-lobed Lanczos-windowed filter defined as in [32]. Table II gathers the 1D impulse response $h(t)$, $t \in \mathbb{R}$ of each separable kernel together with its width. These kernels will also be employed in the experimental validation of the proposed approaches for resampling detection and estimation.

III. EIGENVALUE CHARACTERIZATION FOR GENUINE AND UPSCALED IMAGES

In Sect. II-A, we have briefly outlined how to take advantage of the distinct evolution of the eigenvalues stemming from
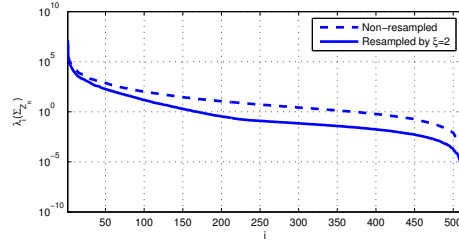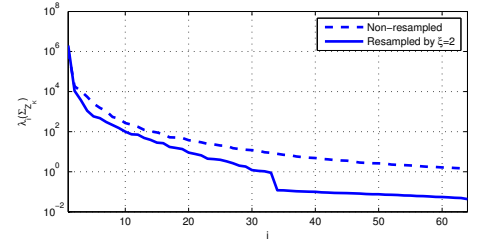
(a) Natural image (1000×1000)   (b) $K = N = 512$ (c) $K = 64$, $N = 512$

Fig. 2. Scree plot comparison of $\boldsymbol{\Sigma}_{Z_K}$ when $\mathbf{Z}_K$ is obtained by extracting the first K columns from the $N \times N$ centered block $\mathbf{Z}$ of the image in (a) after resampling it by $\xi = 2$ with a linear kernel (solid lines), and when no resampling operation is applied (dashed lines). In (b), $K = N$, and in (c) $K < N$. Note that in (c) we only depict the nonzero eigenvalues of $\boldsymbol{\Sigma}_{Z_K}$, i.e., the first K eigenvalues $\lambda_i\left(\boldsymbol{\Sigma}_{Z_K}\right)$ with $i = 1, \ldots, K$.

upscaled images to unveil the presence of resampling traces. Now, to devise a method capable of detecting these traces from an observed image (which might be upscaled or not) and further estimate the applied resampling factor (where applicable), here we formally characterize the statistical distribution of the eigenvalues coming from genuine and upscaled images. As noted above, RMT provides tools to derive deterministic functions that describe the asymptotic pdf of the eigenvalues corresponding to random matrices, however, prior to their application, a stochastic representation of genuine images is required to model the dependencies that arise in real images.

The empirical distribution of the eigenvalues corresponding to a genuine image can be approximated by using a causal two-dimensional autoregressive (2D-AR) model with a single correlation coefficient ρ, as explained in [25]. The generated 2D-AR random field (denoted by $\mathbf{X}$) as a model of an $N \times N$ genuine image can be expressed in matrix form as

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{U}^T, \qquad (5)$$

where $\mathbf{S}$ is an $(N + Q - 1) \times (N + Q - 1)$ random matrix with i.i.d. elements following $\mathcal{N}(0, \sigma_S^2)$, and $\mathbf{U}$ is a Toeplitz matrix of size $N \times (N + Q - 1)$, with Q denoting the length of the truncated infinite impulse response of the AR model. Hence, matrix $\mathbf{U}$ is fully described as

$$U_{i,j} = \begin{cases} \rho^{Q-1-(j-i)}, & \text{if } (j-i) = 0, \ldots, Q-1 \\ 0, & \text{otherwise} \end{cases},$$

where the value of Q is generally taken as $Q \geq N$ to minimize modeling errors due to truncation.

Fig. 3 reports the practical suitability of this model by illustrating how well the evolution of the eigenvalues coming from a real image can be fit using the above random field $\mathbf{X}$.

Adhering to the model in (5) for genuine images, we can apply a procedure proposed by Tulino and Verdú in [23, Theorem 2.43] to analytically compute the asymptotic eigenvalue distribution of $\boldsymbol{\Sigma}_X$ as $N \to \infty$, which affords a better understanding of the empirical eigenvalue distribution arising in real images. Likewise, by adding the resampling process described in Sect. II, the behavior of the eigenvalues coming from upscaled images can also be determined. Interested readers are referred to our technical report in [25], where the process for calculating these asymptotic eigenvalue distributions is comprehensively detailed.
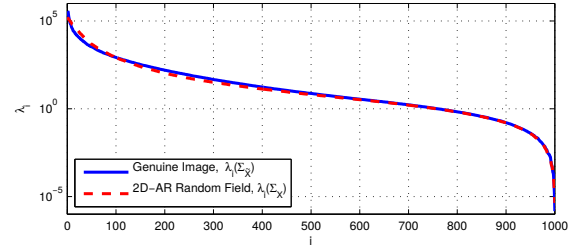

Fig. 3. Comparison of the scree plot of $\boldsymbol{\Sigma}_{\tilde{X}}$ with $\tilde{\mathbf{X}}$ a zero-mean version of the depicted image in Fig. 2(a), against the scree plot of $\boldsymbol{\Sigma}_X$ with $\mathbf{X}$ the 2D-AR random field in (5), using $\rho = 0.945$, $\sigma_S^2 = 22.34$, and $Q = N = 1000$.

In the following, we describe the effect on the evolution of the eigenvalues when K out of N columns/rows from the $N \times N$ matrix under study (either genuine or upscaled) are used to compute the sample autocorrelation matrix. The conclusions drawn from this analysis will lay the foundations of our strategies for resampling detection and estimation.

A. Eigenvalue Distribution for Genuine Images

Adopting the $N \times N$ random field $\mathbf{X}$ in (5) as a model for genuine images, we analyze the eigenvalues of $\boldsymbol{\Sigma}_{X_K} \triangleq \frac{1}{N}\mathbf{X}_K\mathbf{X}_K^T$, when $\mathbf{X}_K$ is an $N \times K$ matrix made up of $K \leq N$ consecutive columns from $\mathbf{X}$. The particular effect on the evolution of the eigenvalues of $\boldsymbol{\Sigma}_{X_K}$ for different matrix aspect ratios $\beta = \frac{K}{N}$ is explained next by graphically examining their asymptotic pdf $f_{\boldsymbol{\Sigma}_{X_K}}(\lambda)$, which can be analytically obtained as detailed in [25, Sect. 1.1].

Fig. 4 shows the eigenvalues of $\boldsymbol{\Sigma}_{X_K}$, i.e., $\lambda_i\left(\boldsymbol{\Sigma}_{X_K}\right)$ with $i = 1, \ldots, N$, together with their pdf $f_{\boldsymbol{\Sigma}_{X_K}}(\lambda)$ for three different values of β, which yield different aspect ratios for the submatrices $\mathbf{X}_K$ as depicted in the leftmost column pictures. The middle column compares the distinct evolution of the eigenvalues, where it can be observed that the number of zero eigenvalues (i.e., $N - K$), increases as $\beta = \frac{K}{N}$ decreases, which is reflected in $f_{\boldsymbol{\Sigma}_{X_K}}(\lambda)$ with a point mass at $\lambda = 0$, whose magnitude is accordingly $(1 - \beta)$ for $\beta < 1$. In addition, the pdfs displayed in the rightmost column indicate through their smallest and largest eigenvalues, i.e., $\lambda_-\left(\boldsymbol{\Sigma}_{X_K}\right)$ and $\lambda_+\left(\boldsymbol{\Sigma}_{X_K}\right)$, that the nonzero eigenvalues of $\boldsymbol{\Sigma}_{X_K}$ tend to be compacted around the variance of $\mathbf{X}$ when β decreases, as predicted by MPL for random matrices with i.i.d. entries. This finding is very important because it generalizes
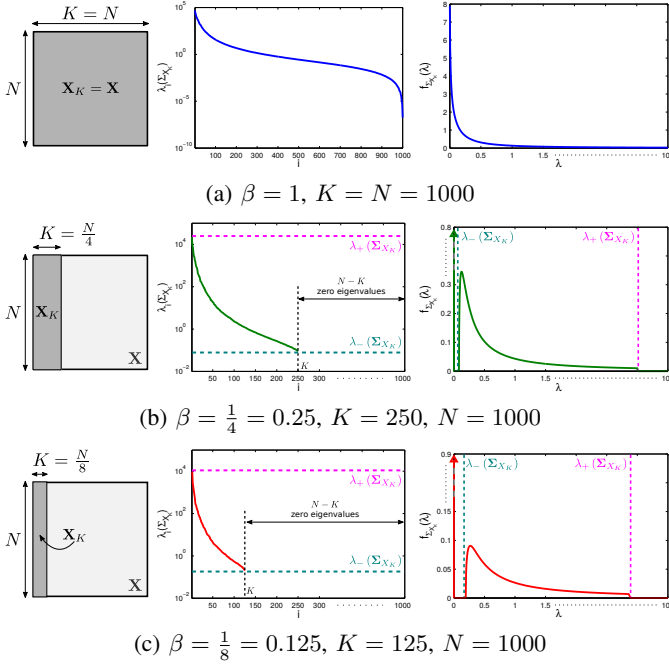
Fig. 4. Graphical representation of the distinct behavior of the N eigenvalues $\lambda_i\left(\boldsymbol{\Sigma}_{X_K}\right)$, when $\mathbf{X}_K$ has different aspect ratios β (leftmost column). Scree plots of $\boldsymbol{\Sigma}_{X_K}$ are displayed in the middle column. The corresponding pdfs are depicted in the rightmost column. AR model: $\rho = 0.97$, $\sigma_S^2 = 1$, $Q = N$.

Fig. 5. Graphical representation of the distinct behavior of the N eigenvalues $\lambda_i\left(\boldsymbol{\Sigma}_{Y_K}\right)$ when two different interpolation kernels are employed. The effect of pixel quantization in $\lambda_i\left(\boldsymbol{\Sigma}_{Z_K}\right)$ is shown in (c). Scree plots are displayed in the middle column and the corresponding pdfs are depicted in the rightmost column. AR model: $\rho = 0.97$, $\sigma_S^2 = 1$, $Q = N$. Quantization step: $\Delta = 1$.

the intuition behind this law to stochastic representations of genuine images. In the following subsection, we will show that this property still holds after an interpolation.

B. Eigenvalue Distribution for Unquantized Upscaled Images

In line with the resampling process described in Sect. II, an upscaled image can be represented by an $N \times N$ matrix $\mathbf{Y}$ generated as in (2) and using the model introduced in (5) for $\hat{\mathbf{X}}$. Accordingly, the same procedure by Tulino and Verdú in [23, Theorem 2.43] can be applied to analytically compute the eigenvalue distribution of $\boldsymbol{\Sigma}_{Y_K} \triangleq \frac{1}{N} \mathbf{Y}_K \mathbf{Y}_K^T$ when $\mathbf{Y}_K$ is constructed by extracting $K \leq N$ consecutive columns from matrix $\mathbf{Y}$, as we explain in [25, Sect. 1.2]. Besides helping us understand the effect of different aspect ratios $\beta = \frac{K}{N}$ on the evolution of the eigenvalues of $\boldsymbol{\Sigma}_{Y_K}$, the knowledge of the asymptotic pdf $f_{\boldsymbol{\Sigma}_{Y_K}}(\lambda)$ in this case allows us to describe the behavior of the eigenvalues as a function of the applied resampling factor and the employed interpolation kernel.

Fig. 5(a)-(b) represents the eigenvalues of $\boldsymbol{\Sigma}_{Y_K}$, i.e., $\lambda_i\left(\boldsymbol{\Sigma}_{Y_K}\right)$ with $i = 1, \ldots, N$, together with the corresponding pdf $f_{\boldsymbol{\Sigma}_{Y_K}}(\lambda)$ for a fixed value of β, but assuming an upscaling operation by $\xi = 2$ and two different interpolation kernels: B-spline and Lanczos.[1] A crucial difference with respect to the case of genuine images (see Fig. 4(c)), is that now $\mathbf{Y}_K$ has rank P strictly smaller than K, i.e., $P \approx K/\xi$, due to the linear dependencies induced by the resampling process. Therefore, only the first P leading eigenvalues are different from zero and the remaining $N - P$ are zero, which yields in $f_{\boldsymbol{\Sigma}_{Y_K}}(\lambda)$ a point mass at $\lambda = 0$ of magnitude $(1 - \beta\xi^{-1})$.

[1]We refrain from showing graphical results for Catmull-Rom and Linear kernels since their behavior is similar to the Lanczos one (cf. [25, Sect. 1.2]).
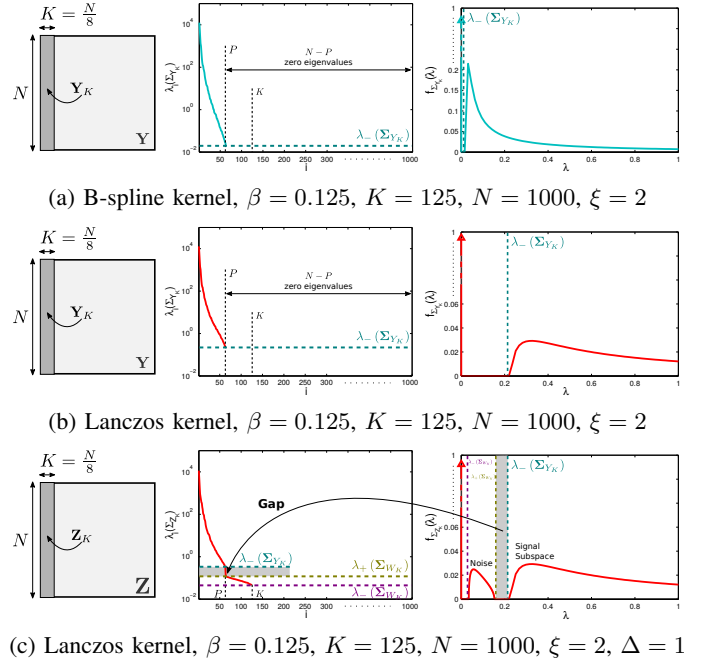
Focusing now on the shape of the pdfs, it is evident that the eigenvalues are squeezed in the same way as with genuine images, which confirms that the compaction of eigenvalues as β decreases still holds for upscaled images. Similarly, in [25, Sect. 1.2], we show that the larger the resampling factor ξ is, the more compacted the eigenvalues become. Both properties will prove to be important in the following section for distinguishing the signal subspace from the background noise. Finally, the particular effect of each interpolation kernel also becomes apparent since, for instance, the B-spline kernel has its eigenvalues more concentrated towards zero than the Lanczos one. This difference on the magnitude of the smallest eigenvalue $\lambda_-\left(\boldsymbol{\Sigma}_{Y_K}\right)$ among the different kernels will unavoidably result in different performance when tackling resampling detection and estimation, as we will see next.

C. Eigenvalue Distribution for Upscaled & Quantized Images

The effect of further quantizing the pixels of an upscaled image is examined here through the model in (4) and by analyzing the eigenvalues of $\boldsymbol{\Sigma}_{Z_K}$, whose definition is given in Sect. II-A. Drawing on MPL and Tulino and Verdú's theorem, we will analytically support the empirical findings discussed in Sect. II-A from the example in Fig. 2(c).

Fig. 5(c) depicts the eigenvalues $\lambda_i\left(\boldsymbol{\Sigma}_{Z_K}\right)$ and their pdf, so as to visually compare the main differences with respect to the unquantized case shown in Fig. 5(b). Given that $\mathbf{Y}_K$ has rank P, in $\boldsymbol{\Sigma}_{Z_K}$ there are P leading eigenvalues corresponding to the signal subspace, but the following $K - P$ are no longer zero, because they now correspond to the background noise. The remaining $N - K$ eigenvalues are zero, as reflected in

$f_{\boldsymbol{\Sigma}_{Z_K}}(\lambda)$ with a point mass at $\lambda = 0$ of magnitude $(1 - \beta)$. From Weyl's inequality applied to the singular values [29, Exercise 1.3.22], we know that if $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ matrices, $m \leq n$; then $\sigma_{i+j-1}(\mathbf{A}+\mathbf{B}) \leq \sigma_i(\mathbf{A}) + \sigma_j(\mathbf{B})$ for all $i, j \geq 1$ such that $i + j - 1 \leq m$. Hence $\lambda_{P+j}(\boldsymbol{\Sigma}_{Z_K}) \leq \lambda_j(\boldsymbol{\Sigma}_{W_K})$, for all $j = 1, \ldots, N - P$. In particular, we can write an upper bound for the first eigenvalue belonging to the noise space:

$$\lambda_{P+1}(\boldsymbol{\Sigma}_{Z_K}) \leq \lambda_1(\boldsymbol{\Sigma}_{W_K}) \to \sigma_W^2 (1 + \sqrt{\beta})^2, \quad (6)$$

where $\lambda_+(\boldsymbol{\Sigma}_{W_K})$ denotes the limiting upper bound on the eigenvalues of $\boldsymbol{\Sigma}_{W_K}$ given by MPL. Now we turn our attention to the last eigenvalue from the signal subspace $\lambda_P(\boldsymbol{\Sigma}_{Z_K})$. Resorting again to Weyl's inequality, we can derive a lower bound for $\lambda_P(\boldsymbol{\Sigma}_{Z_K})$, which is of interest for us to model the transition between the signal subspace and the noise space. To this end, we write $\sigma_{i+j-1}(\mathbf{A}) - \sigma_j(-\mathbf{B}) \leq \sigma_i(\mathbf{A} + \mathbf{B})$, leading us to the following lower bound for $\lambda_P(\boldsymbol{\Sigma}_{Z_K})$:

$$\lambda_P(\boldsymbol{\Sigma}_{Z_K}) \geq \lambda_P(\boldsymbol{\Sigma}_{Y_K}) - \lambda_1(\boldsymbol{\Sigma}_{W_K}) \geq \sigma_S^2 \lambda_-(\boldsymbol{\Sigma}_{Y_K}) - \lambda_+(\boldsymbol{\Sigma}_{W_K}). \quad (7)$$

Combining (6) and (7) we confirm the existence of an asymptotic gap between $\lambda_P(\boldsymbol{\Sigma}_{Z_K})$ and $\lambda_{P+1}(\boldsymbol{\Sigma}_{Z_K})$, i.e.,

$$\frac{\lambda_P(\boldsymbol{\Sigma}_{Z_K})}{\lambda_{P+1}(\boldsymbol{\Sigma}_{Z_K})} \geq \frac{\sigma_S^2 \lambda_-(\boldsymbol{\Sigma}_{Y_K}) - \lambda_+(\boldsymbol{\Sigma}_{W_K})}{\lambda_+(\boldsymbol{\Sigma}_{W_K})} \to \left(\frac{\sigma_S^2}{\sigma_W^2}\right) \frac{\lambda_-(\boldsymbol{\Sigma}_{Y_K})}{(1+\sqrt{\beta})^2} - 1, \quad (8)$$

which marks the transition between the signal subspace and the noise, as shown in Fig. 5(c). The limiting bound in (8) monotonically increases with the signal-to-noise ratio $\left(\frac{\sigma_S^2}{\sigma_W^2}\right)$, which is intuitively appealing. Moreover, it also increases as $\beta \to 0$, since $\lambda_-(\boldsymbol{\Sigma}_{Y_K})$ and $(1+\sqrt{\beta})^2$ respectively increases and decreases as $\beta \to 0$. Finally, for a fixed value of β, the gap becomes smaller as both the resampling factor ξ and the correlation coefficient ρ get closer to 1, as reported in [25, Sect. 1.2]. Analyzing the gap with respect to the interpolation kernel we can conclude that the smallest gap will be obtained through the B-spline kernel, while the largest bound will be achieved by the Lanczos kernel. A significant gap is also expected for the Catmull-Rom kernel since it keeps a similar behavior to the Lanczos filter. Finally, the Linear kernel is halfway between Catmull-Rom and B-spline.

IV. RESAMPLING DETECTION AND ESTIMATION

One obvious concern regarding the asymptotic eigenvalue analysis in Sect. III would be its applicability in practical cases where the observed $N \times N$ image block has a finite size. However, it is a well-known fact from RMT that, even for small N, the asymptotic limit turns out to be an excellent approximation (cf. [23], [29]). Therefore, all the results from the preceding section can be applied to the problems of resampling detection and resampling factor estimation.

In the definition of our hypothesis test, we make the following assumptions: under the null hypothesis, i.e., $\mathcal{H}_0$, the observed matrix $\mathbf{Z}$ comes from a genuine image that has not been resampled and has negligible quantization noise, i.e., it follows the model introduced in Sect. III-A; while, under the alternative hypothesis, i.e., $\mathcal{H}_1$, the observed matrix $\mathbf{Z}$ stems from a genuine image that has been upscaled and later quantized, thus containing quantization noise as in Sect. III-C.

In order to easily understand the behavior of the eigenvalues of $\boldsymbol{\Sigma}_{Z_K}$ with $K < N$ in a finite realistic setting, we outline an example covering the two hypotheses that uses the image in Fig. 2(a). Under $\mathcal{H}_0$, the observed matrix $\mathbf{Z}$ is taken from the 512×512 centered block of the image, while under $\mathcal{H}_1$, $\mathbf{Z}$ is built from the 512×512 centered block of its upscaled version by $\xi = 2$ (using the kernels in Table II and $\Delta = 1$). In both cases, $\mathbf{Z}_K$ contains the first $K = 64$ columns from $\mathbf{Z}$.

Fig. 6 collects the scree plots of $\boldsymbol{\Sigma}_{Z_K}$ for each of these cases. As expected, under $\mathcal{H}_1$ we observe a clear gap in the evolution of the eigenvalues that marks the transition between the signal subspace and the noise, while under $\mathcal{H}_0$ there is no apparent break. Examining the magnitude of the gap for each kernel, we can see that it follows the theoretical findings exposed at the end of Sect. III-C. On the other hand, the gap position found in Sect. II must be broadened, because its exact location depends on the alignment of $\mathbf{Z}_K$ with the resampled grid, the values of ϕ and k_w, and the relation between K and ξ, which are unknown in a realistic forensic setting. According to (1), we can determine a range where the position of the gap P will lie for any $\mathbf{Z}_K$ (regardless of the alignment and ϕ), i.e.,

$$k_w + \left\lfloor (K-1)\frac{M}{L} \right\rfloor - 1 \leq P \leq k_w + \left\lceil (K-1)\frac{M}{L} \right\rceil + 1, \quad (9)$$

where, in these conditions, the position of the gap only depends on k_w and the ratio between K and the applied resampling factor $\frac{L}{M}$, which defines the resulting bounds. Note also that narrower bounds can be achieved for $\phi = 0$. As can be checked in Fig. 6, the position of the gap for all the interpolation kernels is consistent with the theoretical range in (9) when $k_w = 2$.[2] Therefore, by estimating the position of the gap, we can provide a feasible interval where the applied resampling factor $\xi = \frac{L}{M}$ should lie in.

Focusing now on the bounds given by MPL for the quantization noise, i.e., $\sigma_W^2 (1 \pm \sqrt{\beta})^2$, depicted as well in Fig. 6, we can also confirm that even working with matrices with a relatively small value of K and N, the trailing eigenvalues of $\boldsymbol{\Sigma}_{Z_K}$ under $\mathcal{H}_1$ fall within the predicted range. In addition, the magnitude of the smallest nonzero eigenvalue in the resampled case is significantly smaller than its genuine counterpart, thus suggesting that the smallest nonzero eigenvalue of $\boldsymbol{\Sigma}_{Z_K}$ can be used to distinguish between both hypotheses.

A. Description of the Resampling Detector

Based on the previous discussion, the definition of the test statistic for resampling detection will rely on the limiting bounds for the smallest nonzero eigenvalue of $\boldsymbol{\Sigma}_{Z_K}$ under each hypothesis, which corresponds to the K-th eigenvalue $\lambda_K(\boldsymbol{\Sigma}_{Z_K})$. Hence, for a fixed value of K (or equivalently β), we build a vector $\boldsymbol{\Lambda}$ collecting these eigenvalues from a set of V different submatrices $\mathbf{Z}_K$ belonging to the observed matrix $\mathbf{Z}$. Notice that we take different submatrices to sort out regions that may contain linear dependencies leading to rank-deficient

---

[2]We fix $k_w = 2$ even for the kernels with a larger width (e.g., Catmull-Rom, B-spline, or Lanczos), because their smallest eigenvalues (which correspond to the contributions of the filter outside the main lobe, i.e., for $|t| > 1$) fall below the quantization noise level for practical values of σ_S^2/σ_W^2.
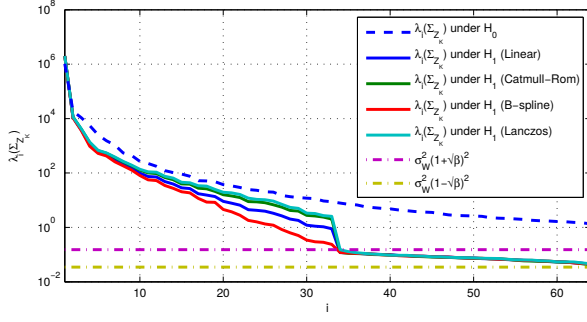
Fig. 6. Scree plot comparing the eigenvalues of $\boldsymbol{\Sigma}_{Z_K}$ under $\mathcal{H}_0$ and $\mathcal{H}_1$, for different interpolation kernels and using the genuine image shown in Fig. 2(a). Notice that only the nonzero eigenvalues are depicted for $K = 64$ and $N = 512$ (i.e., $\beta = \frac{1}{8}$). Under $\mathcal{H}_1$, we use $\xi = 2$ and $\sigma_W^2 = \Delta^2/12$ ($\Delta = 1$).

matrices. In particular, we first take each sliding $N \times K$ submatrix $\mathbf{Z}_K$ of $\mathbf{Z}$, and then we repeat the same process after transposing $\mathbf{Z}$, finally obtaining a total of $V \triangleq 2(N - K + 1)$ submatrices. Denoting by $\lambda_K\left(\boldsymbol{\Sigma}_{Z_K}^{(v)}\right)$ the K-th eigenvalue of $\boldsymbol{\Sigma}_{Z_K}^{(v)}$ corresponding to the v-th submatrix, the vector of eigenvalues is defined as

$$\boldsymbol{\Lambda} \triangleq \left(\lambda_K\left(\boldsymbol{\Sigma}_{Z_K}^{(0)}\right), \quad \lambda_K\left(\boldsymbol{\Sigma}_{Z_K}^{(1)}\right), \quad \cdots, \quad \lambda_K\left(\boldsymbol{\Sigma}_{Z_K}^{(V-1)}\right)\right)^T.$$

To avoid possible rank-deficient matrices $\boldsymbol{\Sigma}_{Z_K}^{(v)}$, we define a set $\mathcal{S}$ containing the indices of those vector components from $\boldsymbol{\Lambda}$ whose magnitude is smaller than the limiting lower bound under $\mathcal{H}_1$, i.e., $\lambda_-(\boldsymbol{\Sigma}_{W_K}) = \sigma_W^2(1 - \sqrt{\beta})^2$. The set is defined as follows

$$\mathcal{S} \triangleq \left\{v \in \{0, \ldots, V - 1\} \; : \; \lambda_K\left(\boldsymbol{\Sigma}_{Z_K}^{(v)}\right) < \sigma_W^2(1 - \sqrt{\beta})^2\right\}. \quad (10)$$

Accordingly, the test statistic is defined as:

$$\kappa \triangleq \begin{cases} \min\limits_{v \in \{0, \ldots, V-1\}} \Lambda_v, & \text{if } |\mathcal{S}| = 0 \\ \operatorname*{median}\limits_{v \in \{0, \ldots, V-1\} \setminus \mathcal{S}} \Lambda_v, & \text{if } 1 \leq |\mathcal{S}| < V, \\ \lambda_0, & \text{if } |\mathcal{S}| = V \end{cases} \quad (11)$$

where Λ_v denotes the v-th component of $\boldsymbol{\Lambda}$, $|\mathcal{S}|$ is the cardinality of $\mathcal{S}$, and λ_0 represents the magnitude of the smallest eigenvalue among all $\lambda_i\left(\boldsymbol{\Sigma}_{Z_K}^{(v)}\right)$ that are larger than the limiting lower bound $\lambda_-(\boldsymbol{\Sigma}_{W_K}) = \sigma_W^2(1 - \sqrt{\beta})^2$. Notice that the last two cases in (11) are set to discard the linear dependencies that may naturally arise in an image block as mentioned in Sect. II-A. When there are some rank-deficient submatrices, as in the second case in (11), we take the median instead of the minimum eigenvalue to avoid a possible bias in the decision towards upscaling, since, in this case, the block under analysis will probably contain several "nearly" rank-deficient matrices even without being resampled. The pseudocode of our resampling detection approach is shown in [25, Alg. 1].

In case one would like to set the detector to operate at a fixed false alarm rate, an empirical threshold T_κ must be first calculated, and then images with $\kappa < T_\kappa$ are declared as upscaled, and else genuine.

B. Description of the Resampling Estimator

As pointed out above, the estimation of the applied resampling factor ξ can be performed by determining the position of the gap $\lambda_P(\boldsymbol{\Sigma}_{Z_K})/\lambda_{P+1}(\boldsymbol{\Sigma}_{Z_K})$ between the signal subspace and the background noise. From a theoretical point of view, the interval containing the actual position of the gap is given by (9) for a fixed value of K, and so the range for ξ becomes

$$\frac{K - 1}{(P - k_w) + 1} \leq \xi \leq \frac{K - 1}{(P - k_w) - 1}. \quad (12)$$

Although this interval provides a rough approximation to ξ, the lack of resolution can be mitigated to some extent by intersecting the obtained intervals using different values of K.

Following a similar procedure to the one exposed in Sect. IV-A, for a fixed value of K, we process a set of V submatrices $\mathbf{Z}_K$ from $\mathbf{Z}$ to obtain an estimator of P. For each v-th submatrix we compute the eigenvalues of $\boldsymbol{\Sigma}_{Z_K}^{(v)}$ and analyze the evolution of the sequence

$$\Psi_v[i] \triangleq \frac{\lambda_i\left(\boldsymbol{\Sigma}_{Z_K}^{(v)}\right)}{\lambda_{i+1}\left(\boldsymbol{\Sigma}_{Z_K}^{(v)}\right)}, \quad \forall i \in \{1, \ldots, K - 1\}.$$

Given that the values of $\Psi_v[i]$ for the first leading eigenvalues are typically high, the set of indices for i is shortened to $\mathcal{I} = \{\lfloor K/\xi_{\max} \rfloor, \ldots, K - 1\}$, so as to easily capture the peak that signals the transition between subspaces. Notice that by doing so, we limit the maximum resampling factor that can be estimated to $\xi_{\max}$. As before, to avoid the contributions of possible rank-deficient matrices, we restrict the analysis to those matrices $\boldsymbol{\Sigma}_{Z_K}^{(v)}$ with $v \notin \mathcal{S}$, where $\mathcal{S}$ is the set of indices in (10). On the other hand, since the gap can occasionally be misdetected as, for instance, with the B-spline kernel which yields gaps of small magnitude (cf. Sect. III-C and Fig. 6), we quantify the strength of the gap, by averaging the sharpness of the peaks in $\Psi_v[i]$ for the available submatrices, i.e.,

$$\mu \triangleq \frac{1}{V - |\mathcal{S}|} \sum_{v \in \{0, \ldots, V-1\} \setminus \mathcal{S}} \max_{i \in \mathcal{I}} \Psi_v[i] / \operatorname*{median}_{i \in \mathcal{I}} \Psi_v[i].$$

When μ is smaller than an empirical threshold T_μ, we assume that the transition between subspaces is not so evident and we rely on the index where the eigenvalues cross the upper bound in (6), i.e., $\sigma_W^2(1 + \sqrt{\beta})^2$, to infer the applied resampling factor. Having this in mind, we build up a vector $\mathbf{p}$ with components $p_v, \; v = 0, \ldots, V - 1$, defined as

$$p_v \triangleq \begin{cases} \arg\max\limits_{i \in \mathcal{I}} \Psi_v[i], & \text{if } v \notin \mathcal{S} \text{ and } \mu \geq T_\mu \\ i_v, & \text{if } v \notin \mathcal{S} \text{ and } \mu < T_\mu, \\ 0, & \text{if } v \in \mathcal{S} \end{cases}$$

where $i_v \triangleq \arg\min\limits_{i \in \{1, \ldots, K\}} \left|\lambda_i\left(\boldsymbol{\Sigma}_{Z_K}^{(v)}\right) - \sigma_W^2(1 + \sqrt{\beta})^2\right|$ is the index of the closest eigenvalue to the upper bound in (6). Note that the nonzero components of $\mathbf{p}$ contain candidate values of P, so its estimation can be drawn from

$$\hat{P} \triangleq \arg\max_{i \in \{0, \ldots, K\}} h(i, \mathbf{p}), \quad (13)$$

where the function $h(i, \mathbf{p})$ counts the number of times that a candidate value i appears in vector $\mathbf{p}$. In case $\hat{P} = 0$, this

would indicate that the observed block is mostly flat (e.g., due to saturation) and no reliable estimation can be obtained.

Given that the index i_v is less precisely related to the applied resampling factor than the actual position of the gap, the derived interval from the estimate in (13) also depends on μ. When a clear gap is obtained, i.e., $\mu \geq T_\mu$, the feasible interval in (12) is directly taken. Conversely, when $\mu < T_\mu$, we modify the lower bound from (12) to a more conservative value, because in this case the trailing eigenvalues converge faster to the variance of the quantization noise, thus having

$$1 < \xi \leq \frac{K-1}{(P-k_w)-1}. \tag{14}$$

According to this, we denote by $\hat{\xi}_K^{(l)}$ the lower bound of the possible interval for ξ obtained either from (12) or (14) for a particular value of K. In a similar way, the upper bound is denoted by $\hat{\xi}_K^{(u)}$. When $\hat{P} = 0$, we set $\hat{\xi}_K^{(l)} = 1$ and $\hat{\xi}_K^{(u)} = \xi_{\max}$. The pseudocode to obtain this interval for a particular K is summarized in [25, Alg. 2]. As previously stated, by taking the intersection of the obtained intervals for different values of K, we can narrow the bounds of the resulting range. So, assuming a set $\mathcal{K}$ with distinct values of K, the final interval is

$$\max_{K \in \mathcal{K}} \left(\hat{\xi}_K^{(l)} \right) \leq \hat{\xi} < \min_{K \in \mathcal{K}} \left(\hat{\xi}_K^{(u)} \right). \tag{15}$$

Note that in practice, some of the intervals may not intersect, so the finally selected interval will be the one resulting from the largest number of intersections.

In a practical setting, the interval in (15) will be generally wide since small values of K are needed to easily localize the gap (recall that the gap increases as K becomes smaller). This would be a problem if we could only rely on the interval in (15) to estimate ξ. However, the interval for ξ can be translated into a range of frequencies where the spectral peak due to resampling should be located. Therefore, for any frequency-based resampling estimator (e.g., [6] or [14]), the combination of both approaches will consequently limit the search space to find the spectral peak. Moreover, since the principles used in each approach are fundamentally different, a gain is to be expected from their combination. Indeed, a first desirable byproduct of this combination is that the provided interval resolves the well-known ambiguity that appears when estimating resampling factors in the ranges $(1, 2)$ and $[2, +\infty)$ with frequency-based methods. In Sect. V-B, the combined approach (with [6] and [14]) will be experimentally validated.

V. EXPERIMENTAL RESULTS

The experimental validation of the proposed methods for resampling detection and estimation is first addressed following the resampling process described in Sect. II. To further evaluate the reliability of our detector when facing real forensic scenarios, we examine in Sect. VI its robustness in case of downscaling and after some common postprocessing operations or targeted attacks (as the ones proposed in [33]).

The set of experiments is conducted over a total of 1317 raw images captured by different Nikon cameras belonging to the Dresden Image Database [34]. Prior to applying the resampling operation, all raw images are demosaiced using the so-called

"variable number of gradients" method [35] (which is available in the `dcraw` software tool with option `-q 1`). Given its extensive use, we employ the image processing tool `convert` from ImageMagick's software to perform each full-frame resampling operation. As interpolation kernels, we select the ones most commonly available (cf. Table II). To assess the most suitable upscaling factors avoiding the introduction of visible distortions, we constrain the set of resampling factors to the interval $[1.05, 2]$ uniformly sampled with step 0.05.

The analysis of resampling traces is performed over the green channel of each image under study by processing its central square block $\mathbf{Z}$ of size $N \times N$. To test our detector in a practical scenario where the tampered regions might be small, we process image blocks with $N = 32$. On the other hand, to give an accurate estimation of the applied upscaling factor, a larger block is processed by taking $N = 128$.

A. Performance Analysis of the Detector

The performance of the proposed detector is measured in terms of the Area Under the Curve (AUC) corresponding to the Receiver Operating Characteristic (ROC), and the True Positive Rate (TPR), or detection rate, at a fixed False Acceptance Rate (FAR). For comparison, the same tests are applied to three state-of-the-art detectors: the "SVD-based" detector derived in [21], the "LP-based" detector proposed in [8] (where LP stands for Linear Predictor), and the "ZM-based" detector from [9] (where ZM stands for Zeroing Mask). We configure our detector to work with submatrices of small aspect ratio $\beta = 0.2812$ (i.e., $K = 9$, yielding $V = 48$), because it makes the separation between the smallest eigenvalue under each hypothesis more evident. For [21] we set $\xi_{\min} = 1.05$, for [8] we fix a neighborhood of 3 rows/columns, and finally, for [9] we take the tested set of upscaling factors as the set of candidates[3] and we adapt the computation of the zeroing masks to handle more kernels than the Linear one with $\phi = 0$.

The values of AUC for each tested upscaling factor ξ are computed by applying the four detectors on both the genuine and the upscaled images. However, when deriving the detection rates for each ξ, the database is randomly split in two disjoint sets, where $1/3$ of the images (i.e., a total of 439) are exclusively used to empirically determine the thresholds of each detector, and the remaining 878 images are used for testing. Upon deriving the thresholds at a fixed FAR $\leq 1\%$, all detectors are applied on the upscaled versions of the images from the testing set to compute the detection rates for each ξ.

Fig. 7(a) shows the obtained AUC values for each detector. Since the SVD-based detector relies on the same principle of subspace decomposition than our detector, the performance of the two approaches in terms of AUC is similar, achieving better results for B-spline and Linear interpolation kernels than for Catmull-Rom and Lanczos. Both methods outperform the LP-based detector, which shows difficulties when dealing with small block sizes, and also the ZM-based detector, which suffers from high FARs as N gets smaller. This is the reason

---

[3] This selection of the set of candidates clearly benefits the approach in [9], since a larger set would notably increase the false acceptance rate.
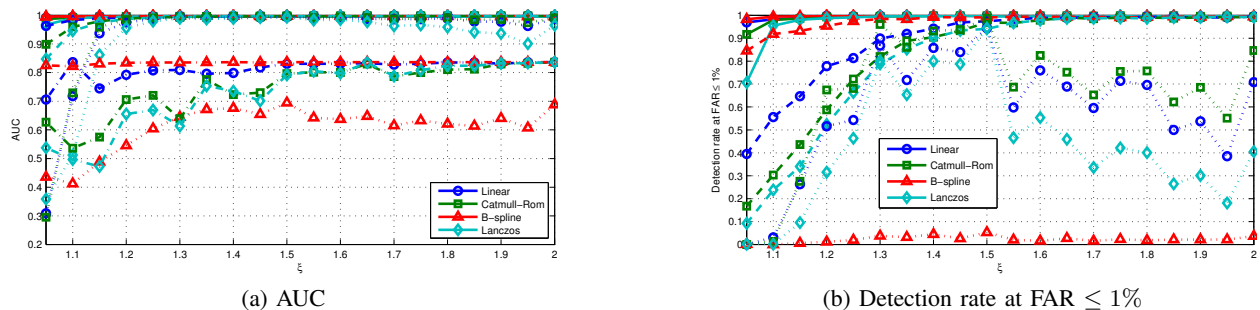
(a) AUC

(b) Detection rate at FAR $\leq 1\%$

Fig. 7. Evaluation of our detector (solid lines) against the SVD-based [21] (dashed lines), the LP-based [8] (dotted lines), and the ZM-based [9] (dash-dotted lines) detectors in terms of AUC and detection rate for 32×32 blocks. No detection rates are shown for the ZM detector, since the obtained FAR was about 34%.

TABLE III
$\overline{\text{TPR}}$ VALUES AT FAR $\leq 1\%$ FOR DIFFERENT $N \times N$ BLOCK SIZES. BOLD NUMBERS INDICATE THE BEST RESULTS.

Kernel Types:	Linear				Catmull-Rom				B-spline				Lanczos			
Detectors:	Ours	SVD	LP	ZM	Ours	SVD	LP	ZM	Ours	SVD	LP	ZM	Ours	SVD	LP	ZM
$N=32$	**0.9946**	0.8897	0.5882	0.9383*	**0.9911**	0.8296	0.6727	0.8032*	**0.9966**	0.9749	0.0226	0.9942*	**0.9781**	0.8067	0.4325	0.7873*
$N=64$	**0.9983**	0.9541	0.9402	0.9157†	**0.9979**	0.9108	0.9390	0.8341†	**0.9985**	0.9923	0.3812	0.9954†	**0.9960**	0.8924	0.9138	0.9240†
$N=128$	**1.0000**	0.9842	0.9793	0.7086	**1.0000**	0.9544	0.9708	0.8176	**1.0000**	0.9985	0.8194	0.9964	**0.9999**	0.9387	0.9524	0.8569
$N=256$	**1.0000**	0.9924	0.9988	0.7038	**1.0000**	0.9764	0.9988	0.7936	**1.0000**	0.9990	0.9293	0.9986	**1.0000**	0.9640	0.9924	0.9110
$N=512$	**1.0000**	0.9838	0.9996	0.7013	**1.0000**	0.9518	0.9997	0.7343	**1.0000**	0.9993	0.9455	0.9969	**1.0000**	0.9337	0.9986	0.8722

*FAR$\approx 34\%$, †FAR$\approx 18\%$

(a) Forged image
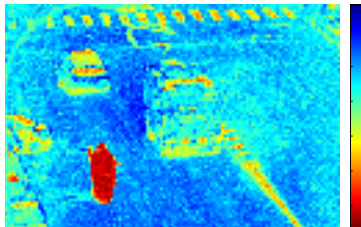
(b) log-κ map

(c) Binary mask

Fig. 8. Example of forgery detection using the proposed resampling detector. The forged image in (a) has been created by adding the policeman on the bottom-left corner. The log-magnitude of the statistic κ given by (11) is painted in (b) for each 32×32 block from the green channel of the forged image. Warmer colors indicate smaller values. The intersected binary mask (after applying threshold T_κ to each color channel), is shown in (c).

why the detection rates for the latter method were not included in Fig. 7(b), as the respective FAR was about 34%.

In Fig. 7(b) it can be seen that the proposed technique clearly improves the detection performance of the other two methods when $1.05 \leq \xi \leq 1.5$, which is particularly interesting in practice for exposing slightly transformed regions. In fact, this range was known to be especially challenging for existing resampling detectors when dealing with small image blocks. For our detector, the compaction of eigenvalues for $K < N$ elicits this notable improvement, making it possible to achieve a FAR $\leq 1\%$ with a threshold $T_\kappa = 0.0827$ (i.e., approximately one-half of the upper bound in (6): $\sigma_W^2(1+\sqrt{\beta})^2 = 0.1952$). Related to this, the fast convergence towards zero of the eigenvalues when using the B-spline kernel (see Fig. 5(a)), improves the separability between upscaled and genuine images, thus providing always the best performance.

To save space representing the obtained results in different experiments and to facilitate comparisons with state-of-the-art detectors, we use a meaningful measure that summarizes the performance of each detector under each setting. This overall digest, denoted by $\overline{\text{TPR}}$, reflects the average rate of detection

across all the tested upscaling factors for a given detector:

$$\overline{\text{TPR}} \triangleq \frac{1}{|\Xi|} \sum_{\xi \in \Xi} \text{TPR}(\xi), \qquad (16)$$

where $\text{TPR}(\xi)$ represents the obtained true positive rate (at a fixed FAR) for the tested upscaling factor ξ, which in our case belongs to the set $\Xi \triangleq \{1.05, 1.1, \ldots, 2\}$.

As an example, the first row of Table III contains the $\overline{\text{TPR}}$ values corresponding to the curves represented in Fig. 7(b) for $N = 32$. Additionally, we also include in Table III the $\overline{\text{TPR}}$ values when each detector works with blocks of a larger size, i.e., for $N \in \{64, 128, 256, 512\}$. From these results, we can appreciate that larger block sizes generally increase detection rates, and also that our approach keeps yielding superior detection results with respect to the rest of detectors. It is important to note that the values of FAR achieved by the ZM-based detector strongly depend on the size of the block under analysis (a FAR close to 1% is not reachable for $N = 32$ or $N = 64$, while this is not an issue for a larger N). These FAR values may even get worse in a real forensic scenario since no prior knowledge on the range of plausible upscaling
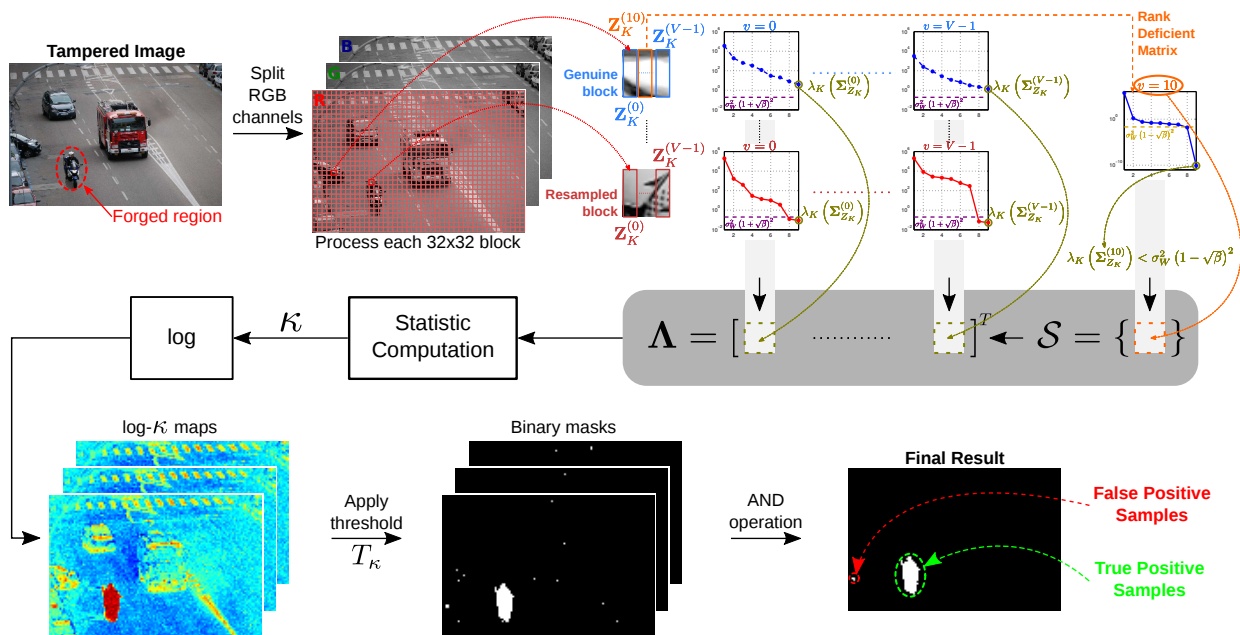
Fig. 9. Description of the intermediate steps involved in the whole analysis of a forged image to obtain the result shown in Fig. 8.

factors is available, and a larger number of candidates implies an increase on the detector FAR.

To further illustrate the good separability between genuine and upscaled regions afforded by our detector, we test it with a hand-made forgery in which only a portion of a demosaiced image has been resampled. To this end, we plot a log-magnitude map of the proposed test statistic (11) for each 32×32 block of the image, which compellingly highlights potential geometrically transformed regions. Fig. 8(b) depicts the resulting map after processing the green channel of the forged image in Fig. 8(a). Applying the obtained empirical threshold T_κ for demosaiced images, we then construct a binary map for each color channel. Finally, by computing the logical AND operation between each of the three resulting binary masks, the suspected region is exposed in Fig. 8(c). The accuracy of the result further supports the robustness of the threshold T_κ, since the dataset from which it was determined is totally independent from the forged image. Fig. 9 illustrates the whole process, showing each intermediate step.

Finally, for the sake of computation time comparison, we compute the average time that each method takes to process the central 32×32 block of the 1317 images from the database. All simulations are performed using MATLAB R2013a on a desktop computer,[4] obtaining (in seconds): 0.0066 for our method, 0.0003 for the SVD-based detector [21], 0.0108 for the LP-based one [8], and 0.2074 for the ZM-based method [9]. From these results, we can see that the two approaches based on subspace decomposition (i.e., the SVD-based and ours) are more computationally efficient than the others. The SVD-based approach achieves the best performance because, unlike ours, only the calculation of one SVD per block is needed. Nevertheless, the performance of our detector could be further improved by resorting to iterative eigenvalue algorithms (cf.

[36, Section 4.4]), so as to reduce the whole eigenvalue decomposition to the computation of the smallest eigenvalue of a given matrix, which is the one actually employed in our statistic (11).

B. Performance Analysis of the Estimation Strategy

In this section, we study the viability of using the described estimation strategy in Sect. IV-B as a complementary technique for resampling factor estimation through its combination with existing approaches to the same problem. In particular, we analyze the performance of the combination of the proposed strategy and two state-of-the-art frequency-based estimators: the "2D frequency-based method" described in [14], and the "1D frequency-based method" in [6]. The evaluation is conducted in terms of Mean Squared Error (MSE) and percentage of correct resampling factor estimates, which is determined according to the criterion $|\hat{\xi} - \xi| < 0.05$, where $\hat{\xi}$ is the estimated factor and ξ the true one. For each combination, the obtained results are compared against its standalone version.

The tests are performed over blocks of size $N \times N$ ($N = 128$) coming from the full-frame upscaled versions of the 1317 demosaiced images in the database. For our strategy, we restrict ourselves to $\xi_{\max} = 2.1$, employ the set $\mathcal{K} = \{29, \ldots, 33\}$, fix $k_w = 2$ and the threshold $T_\mu = 2.5$, and shorten the number of processed submatrices to $V = 20$ (i.e., we only process the first 10 shifted submatrices in each direction). For the 2D frequency-based method, a set of 9 lags and a spectral window of size 11×11 are chosen.

The procedure to combine our strategy with the referred methods is simple and can be synthesized as follows: first, the interval for ξ is determined by following the process described in Sect. IV-B. In the meantime, each frequency-based method separately computes its statistic in the Fourier domain. Then, the obtained interval in (15) is translated to the corresponding

[4]Intel Core i5-2500K @ 3.30GHz, 8GB RAM, running Ubuntu 14.04.

(a) MSE (2D frequency-based method)

(b) Percentage of Correct Resampling Factor Estimation

(c) MSE (1D frequency-based method)

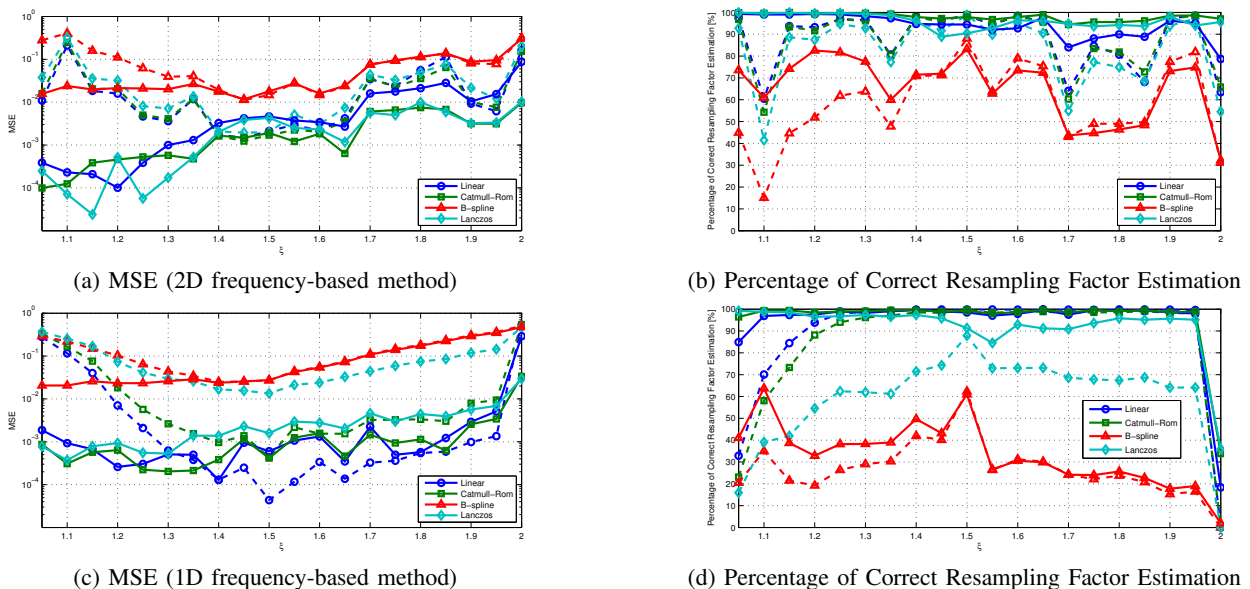(d) Percentage of Correct Resampling Factor Estimation

Fig. 10. Evaluation in terms of MSE and percentage of correct resampling factor estimation of the combined methods (solid lines) against their standalone versions (dashed lines). The upper panels contain results from the 2D frequency-based method [14] and the lower ones gather the outcomes from the 1D approach [6].

frequencies in the 2D and 1D domains, as respectively required by each method. Finally, the estimate $\hat{\xi}$ is determined from the position of that spectral peak with largest magnitude in the calculated range of frequencies.

The upper panels of Fig. 10 depict the obtained results when combining our estimator with the 2D frequency-based method. From Fig. 10(a), we can observe how the MSE is significantly reduced for all the interpolation kernels, except for the B-spline, which achieves the smallest gain as expected. With this particular kernel, the upscaling factor is only well approximated when the interpolation is performed over images with high frequency content. The significant MSE reduction translates into an important gain in the percentage of correct resampling factor estimates, as shown in Fig. 10(b). It is interesting to notice how the provided interval helps improve the estimation accuracy for resampling factors close to 1.

The lower panels of Fig. 10 contain the corresponding results for the combination of our estimator with the method in [6]. In general, the combination of both methods yields better results, significantly so for the Lanczos kernel. However, in comparison with the previous combined approach, considerably large values of MSE are now obtained for resampling factors close to 1. This means that the statistic from the 1D frequency-based method does not provide a clear peak in the expected range for resampling factors approaching 1. Similarly, the gain from the combination of estimators is negligible for $\xi = 2$, since no spectral peak is available at that upsampling rate for the 1D frequency-based method (a fact that was pointed out in [14]).

Note that we have constrained this viability study to the methods in [14] and [6], but the proposed strategy can be easily combined with others. For instance, following the same idea, we can use our estimator to relax the exhaustive search required by Popescu and Farid in [2] and we can even go beyond frequency-based techniques, such as the methods in

[9], [10], and [11], where the interval in (15) can be used to narrow the set of candidate resampling factors to a few values, heavily reducing the computation time.

VI. Robustness Analysis of Resampling Detection

For evaluating the robustness of our resampling detector in more realistic scenarios, we test our detector sensitivity to different pre/postprocessing operations using the same dataset described in Sect. V. The comparison with state-of-the-art approaches is carried out in terms of $\overline{\text{TPR}}$ (cf. (16)), but constraining the analysis to the LP-based and ZM-based detectors. We leave out the SVD-based detector because it shows a similar behavior to our approach, but with worse performance. Due to space constraints, we only report results for blocks of size $N \times N$ with $N = 512$, which can be seen as upperbounds on each detector's performance, since smaller blocks generally lead to worse detection results. In this case, we configure our detector to work with $K = 51$ and $V = 20$.

A. Multiple Resampling and Downscaling

We have conducted several experiments in the presence of multistage resamplings, such as different upscalings followed by multiple downscalings and vice versa. As long as the global resampling factor (which results from the product of each intermediate rate) is larger than one, the performance of our detector is similar to that obtained with an analogous upscaling operation by the global factor. Therefore, we do not report these results here. Interestingly, rounding errors originated in each resampling stage do not strongly affect the performance.

As noted in Sect. I, our approach based on RMT and subspace decomposition has not been originally conceived to detect image downsizing. However, given the potential use of downscaling operations as a way to hide traces of tampering, here we investigate the performance of our detector in case
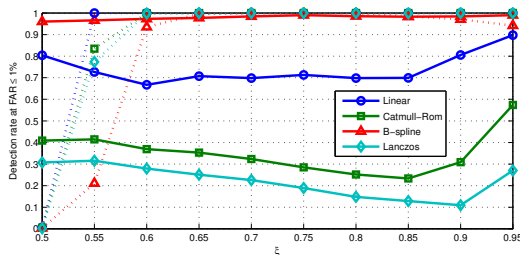
Fig. 11. Detection rates at FAR $\leq 1\%$ for downscaling. Comparison among our detector (solid lines) and the LP-based detector [8] (dotted lines). $N = 512$.

of downsizing with a set of factors in the range $[0.5, 0.95]$ (uniformly sampled with step 0.05). To perform full-frame downscalings we use the tool `convert`, which automatically applies an antialiasing filter.

Fig. 11 shows the detection rates obtained with our detector and the LP-based approach after downscaling using the kernels in Table II. The ZM-based detector is not tested because, under downscaling, the signal subspace spans the whole N-dimensional vector space and it is not possible to obtain zeroing masks. While in most cases the LP-based approach yields the best performance, we remark that our detector shows superior results for $\xi = 0.5$ and mainly for the B-spline kernel, whose influence on the fast convergence of the eigenvalues towards zero boosts the distinction between downscaled and genuine images.

As a conclusion, even without a specific model for the downscaling process, the large potential of RMT translates into acceptable detection rates for some kernels, so RMT properties should be further investigated to detect image downsizing.

B. Common Postprocessing Operations

The behavior of our detector is evaluated here when common postprocessing operations like JPEG compression, rotation, or linear filtering are applied after upscaling.

1) JPEG re-compression: we analyze the case where demosaiced images are first JPEG compressed with a quality factor QF_1, then upscaled by $\xi \in \Xi$, and finally re-compressed with the same factor, i.e., $QF_2 = QF_1 = QF$.[5] It is known that most of the resampling detectors fail after moderate JPEG compression [33], so we only consider $QF \in \{90, 95, 98, 100\}$.

The obtained $\overline{TPR}$ values for each tested detector are collected in Table IV(a) and are classified as a function of the distinct interpolation kernels. Although the JPEG re-compression scenario might be the most common in practice, all detectors start failing at $QF \leq 95$. Our detector generally yields higher detection results for quality factors above 95, but when the applied lossy compression is strong enough to render the subspace decomposition unreliable, it suffers from severe performance loss.

We have also tested the case where genuine images are not JPEG compressed and the compression is only applied after upscaling. Under this setting, all detectors ended up discerning uncompressed from compressed images instead of exposing

resampling traces (in other words, their performance was no longer dependent on the applied ξ, but on the selected JPEG quality factor), so we do not report the results.

2) Rotation: we examine the resampling detectability when demosaiced images are upscaled and later rotated by an angle Θ in the range $[5, 45]$ sampled at equidistant steps of $10°$.[6]

Table IV(b) collects the obtained $\overline{TPR}$ values which show that the performance of the tested detectors regularly decreases as the rotation angle increases. An exception to this behavior occurs with the LP-based detector that is only able to detect resampling traces for $\Theta = 45°$, where subtle periodic correlations among adjacent rows and columns are still preserved. On the other hand, our detector achieves the best detection results by means of the implicit search towards all directions performed by the eigenvalue decomposition, which always seeks the most suitable underlying subspace.

3) Linear Filtering: we evaluate to what extent linear filtering can blur resampling traces when it is applied after upscaling by $\xi \in \Xi$ a demosaiced image. The analysis is focused on a 3×3 Gaussian kernel with $\sigma_g \in \{0.5, 1, 1.5\}$.

The obtained results are gathered in Table IV(c) and they reveal the high degree of robustness of our detector for this type of filtering. Unlike the LP-based method which suffers from moderate blurriness (i.e., for $\sigma_g \leq 1$) and the ZM-based approach which only shows difficulties for $\sigma_g = 0.5$, our detector always achieves the best performance for all combinations of interpolation kernels and values of σ_g.

C. Targeted Attacks Against Resampling Detection

As can be presumed from the example in Fig. 6, a possible attack to deceive our detector would be to add random noise to the samples of the upscaled image. By doing so, the resulting magnitude of the trailing eigenvalues arising from the background noise of the attacked image will increase and may finally reach the expected magnitude of those arising from genuine images. This might impede a reliable resampling detection; however, the addition of noise will unavoidably degrade the quality of the upscaled images and consequently raise suspicions about their authenticity, thus making this attack mostly ineffective. A more subtle way to mislead our detector could be conceived by using non-linear or content-adaptive interpolation techniques (such as those commonly employed in demosaicing) to perform any resampling operation. This would yield a potential anti-forensic technique against our detector.

Despite the existence of this type of ad hoc attacks, we find particularly interesting the analysis of our detector's robustness against existing countermeasures, such as the ones presented in [33], which were specifically designed to hide resampling traces. In the following, we describe the parameters we use in the three types of attacks that were proposed in [33, Sect. V].

1) Nonlinear filtering: a median filter of size 3×3 is applied after the upscaling process.
2) Geometric Distortion (GD): the resampling process is manipulated by introducing a geometric distortion [33,

---

[5]All compressions are performed in MATLAB with the `imwrite` function.

[6]MATLAB `imrotate` function with Linear kernel is used for all rotations.

TABLE IV
$\overline{\text{TPR}}$ VALUES AT FAR $\leq 1\%$ FOR DIFFERENT PRE/POSTPROCESSING OPERATIONS. BOLD NUMBERS INDICATE THE BEST RESULTS.

Kernel Types:	Linear			Catmull-Rom			B-spline			Lanczos		
Detectors:	Ours	LP	ZM	Ours	LP	ZM	Ours	LP	ZM	Ours	LP	ZM
(a) JPEG re-compression												
QF = 90	0.1183	0.0001	**0.1663**	0.0908	0.0003	**0.3212**	0.1824	0.0009	**0.6896**	0.0829	0.0009	**0.4876**
QF = 95	**0.4473**	0.0189	0.2503	**0.3966**	0.0314	0.3250	0.7399	0.0110	**0.8263**	0.4151	0.0223	**0.5027**
QF = 98	0.9551	**0.9672**	0.7113	0.9298	**0.9618**	0.7023	**0.9990**	0.8972	0.9948	**0.9568**	0.9532	0.8196
QF = 100	**1.0000**	0.9948	0.7928	**1.0000**	0.9962	0.7890	**1.0000**	0.8039	0.9987	**1.0000**	0.9637	0.9008
(b) Rotation												
$\Theta = 5°$	**0.9897**	0.0000	0.8650	**0.9771**	0.0000	0.8789	**0.9998**	0.0002	0.9992	**0.9839**	0.0002	0.9435
$\Theta = 15°$	**0.9708**	0.0001	0.8629	**0.9494**	0.0001	0.8647	**0.9991**	0.0002	**0.9991**	**0.9612**	0.0002	0.9282
$\Theta = 25°$	**0.9445**	0.0001	0.8517	**0.9141**	0.0001	0.8325	**0.9970**	0.0002	0.9955	**0.9206**	0.0002	0.8822
$\Theta = 35°$	**0.9120**	0.0003	0.8119	**0.8640**	0.0001	0.7705	**0.9887**	0.0002	0.9744	**0.8595**	0.0002	0.8080
$\Theta = 45°$	**0.9200**	0.4619	0.8072	**0.8666**	0.4653	0.7540	**0.9870**	0.2044	0.9688	**0.8528**	0.3223	0.7937
(c) Linear filtering												
$\sigma_g = 0.5$	**1.0000**	0.9808	0.9387	**1.0000**	0.9701	0.9324	**1.0000**	0.8700	0.9999	**1.0000**	0.9554	0.9744
$\sigma_g = 1$	**1.0000**	0.3395	**1.0000**	**1.0000**	0.4080	**1.0000**	**1.0000**	0.0253	**1.0000**	**1.0000**	0.3528	**1.0000**
$\sigma_g = 1.5$	**1.0000**	0.1093	0.9998	**1.0000**	0.1916	0.9999	**1.0000**	0.0023	**1.0000**	**1.0000**	0.1004	**1.0000**
(d) Targeted attacks												
Median filtering	**0.8785**	0.7285	0.4700	**0.8257**	0.5895	0.4802	**0.9259**	0.1038	0.7197	**0.7949**	0.3605	0.5345
GD without edge modulation	**1.0000**	0.0010	0.0637	**1.0000**	0.0009	0.0916	**1.0000**	0.0031	0.1485	**1.0000**	0.0010	0.1305
GD with edge modulation	**1.0000**	0.8771	0.1340	**1.0000**	0.8858	0.2261	**1.0000**	0.1241	0.3829	**1.0000**	0.6286	0.3169
Dual-path approach	0.7718	**0.8313**	0.2314	0.7240	**0.8315**	0.2396	0.8423	0.0264	0.4744	**0.7010**	0.4331	0.2921

Eq. 13], whose strength is controlled by σ. In addition, an extension of the same idea with edge modulation is also considered [33, Eq. 14]. In both cases, $\sigma = 0.4$.

3) Dual-path approach: this attack combines median filtering with the above geometric distortion [33, Eq. 16]. In this case, we use a 5×5 median filter and $\sigma = 0.4$.

Table IV(d) reports $\overline{\text{TPR}}$ values obtained by the three detectors in each scenario. On the one hand, these results reflect that the application of nonlinear operations (such as the median filter), either before or after the upscaling process slightly affects the performance of our detector. On the other hand, geometric distortions with the selected strength $\sigma = 0.4$ do not impair its reliability. The reason why our detector is resilient to geometric distortions is that our test statistic does not only rely on the linear dependencies induced by the resampling process, but also on the level of noise that remains after the subspace decomposition. Under these targeted attacks, the transition between subspaces becomes less evident (preventing a correct resampling factor estimation), but the level of noise is still consistent with the one from undistorted upscaled images.

As a conclusion, since our method faces the resampling detection problem from a different perspective, a fusion of the outputs of the discussed detectors or other existing ones that already prevent resampling anti-forensics techniques, e.g., [37], might be considered to improve the overall detectability in these realistic scenarios.

VII. CONCLUSIONS

The forensic analysis of resampling traces has been investigated in this paper using new tools and concepts from RMT. We have proven that there is a large potential in the use of RMT coupled with the subspace decomposition approach for resampling detection. As our approach is radically different from existing ones, it also contributes new information. In fact, we have shown that the combination of our scheme with other techniques yields substantial improvements in the accuracy of resampling factor estimates. Remarkably, the computational complexity of the proposed approaches is low and could be further reduced by using iterative eigenvalue computation algorithms, thus bringing the possibility to design efficient implementations.

Although this work can be deemed as a first step towards the characterization of resampled signals in terms of their eigenvalue distribution, greater advances need to be made in the analysis of their behavior in more complex scenarios. In this direction, our incipient robustness analysis in realistic scenarios shows promising results, but it also reveals that nonlinear operators must be further investigated. Moreover, the analysis of image downscaling within this random matrix framework needs to be addressed separately to check whether processing the three color channels might be useful. Further research is also required to better model the asymptotic eigenvalue distribution from genuine images that contain demosaicing traces.

REFERENCES

[1] A. Piva, "An overview on image forensics," *ISRN Signal Processing*, 2013, Article ID 496701, 22 pages, 2013.

[2] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 758–767, Feb. 2005.

[3] M. Kirchner, "Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue," in *ACM Workshop on Multimedia and Security (MM&Sec)*, Sep. 2008, pp. 11–20.

[4] A. C. Gallagher, "Detection of linear and cubic interpolation in JPEG compressed images," in *Canadian Conference on Computer and Robot Vision (CRV)*, May 2005, pp. 65–72.

[5] S. Prasad and K. Ramakrishnan, "On resampling detection and its application to detect image tampering," in *IEEE International Conference on Multimedia and Expo (ICME)*, Aug. 2006, pp. 1325–1328.

[6] B. Mahdian and S. Saic, "Blind authentication using periodic properties of interpolation," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 529–538, Sep. 2008.

[7] G. K. Birajdar and V. H. Mankar, "Blind method for rescaling detection and rescale factor estimation in digital images using periodic properties of interpolation," *AEU - International Journal of Electronics and Communications*, vol. 68, no. 7, pp. 644 – 652, Jul. 2014.

[8] M. Kirchner, "Linear row and column predictors for the analysis of resized images," in *ACM Workshop on Multimedia and Security (MM&Sec)*, Sep. 2010, pp. 13–18.

[9] Y. T. Kao, H. J. Lin, C. W. Wang, and Y. C. Pai, "Effective detection for linear up-sampling by a factor of fraction," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3443–3453, Aug. 2012.

[10] D. Vázquez-Padín and P. Comesaña, "ML estimation of the resampling factor," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, Dec. 2012, pp. 205–210.

[11] D. Vázquez-Padín, P. Comesaña, and F. Pérez-González, "Set-membership identification of resampled signals," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, Nov. 2013, pp. 150–155.

[12] R. Wang and X. Ping, "Detection of resampling based on singular value decomposition," in *International Conference on Image and Graphics (ICIG)*, Sep. 2009, pp. 879–884.

[13] X. Feng, I. J. Cox, and G. Doërr, "Normalized energy density-based forensic detection of resampled images," *IEEE Transactions on Multimedia*, vol. 14, no. 3, pp. 536–545, Jun. 2012.

[14] D. Vázquez-Padín, C. Mosquera, and F. Pérez-González, "Two-dimensional statistical test for the presence of almost cyclostationarity on images," in *IEEE International Conference on Image Processing (ICIP)*, Sep. 2010, pp. 1745–1748.

[15] W. Wei, S. Wang, X. Zhang, and Z. Tang, "Estimation of image rotation angle using interpolation-related spectral signatures with application to blind detection of image forgery," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 507–517, Sep. 2010.

[16] C. Chen, J. Ni, and Z. Shen, "Effective estimation of image rotation angle using spectral method," *IEEE Signal Processing Letters*, vol. 21, no. 7, pp. 890–894, Jul. 2014.

[17] S.-J. Ryu and H.-K. Lee, "Estimation of linear transformation by analyzing the periodicity of interpolation," *Pattern Recognition Letters*, vol. 36, pp. 89–99, Jan. 2014.

[18] S. Pfennig and M. Kirchner, "Spectral methods to determine the exact scaling factor of resampled digital images," in *International Symposium on Communications Control and Signal Processing (ISCCSP)*, May 2012, pp. 1–6.

[19] T. R. Goodall, I. Katsavounidis, Z. Li, A. Aaron, and A. C. Bovik, "Blind picture upscaling ratio prediction," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1801–1805, Dec. 2016.

[20] Y. Su, X. Jin, C. Zhang, and Y. Chen, "Hierarchical image resampling detection based on blind deconvolution," *Journal of Visual Communication and Image Representation*, Jan. 2017.

[21] D. Vázquez-Padín, P. Comesaña, and F. Pérez-González, "An SVD approach to forensic image resampling detection," in *European Signal Processing Conference (EUSIPCO)*, Sep. 2015, pp. 2067–2071.

[22] D. Paul and A. Aue, "Random matrix theory in statistics: A review," *Journal of Statistical Planning and Inference*, vol. 150, pp. 1–29, Jul. 2014.

[23] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.

[24] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 457–483, Apr. 1967.

[25] D. Vázquez-Padín, F. Pérez-González, and P. Comesaña Alfaro, "Derivation of the asymptotic eigenvalue distribution for causal 2D-AR models under upscaling," arXiv:1704.05773 [cs.CR], Apr. 2017.

[26] J. Kodovský and J. Fridrich, "Effect of image downsampling on steganographic security," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 5, pp. 752–762, May 2014.

[27] Y. Li and Z. Ding, "Blind channel identification based on second order cyclostationary statistics," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Apr. 1993, pp. 81–84.

[28] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*. Springer, 2009.

[29] T. Tao, *Topics in Random Matrix Theory*. American Math. Soc., 2012.

[30] J. T. Mardia, K. V. Kent and J. M. Bibby, *Multivariate Analysis*. Academic Press, 1979.

[31] D. P. Mitchell and A. N. Netravali, "Reconstruction filters in computer-graphics," in *Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, Aug. 1988, pp. 221–228.

[32] A. S. Glassner, *Graphics Gems*. Academic Press, 1990.

[33] M. Kirchner and R. Böhme, "Hiding traces of resampling in digital images," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 4, pp. 582–592, Dec. 2008.

[34] T. Gloe and R. Böhme, "The Dresden Image Database for benchmarking digital image forensics," in *ACM Symposium on Applied Computing (SAC)*, Mar. 2010, pp. 1584–1590.

[35] E. Chang, S. Cheung, and D. Y. Pan, "Color filter array recovery using a threshold-based variable number of gradients," *Proc. SPIE*, vol. 3650, pp. 36–43, Mar. 1999.

[36] J. W. Demmel, *Applied Numerical Linear Algebra*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1997.

[37] A. Peng, Y. Wu, and X. Kang, "Revealing traces of image resampling and resampling antiforensics," *Advances in Multimedia*, vol. 2017, Article ID 7130491, 13 pages, 2017.

David Vázquez-Padín (S'10-M'16) received the M.S. degree in telecommunications engineering and the Ph.D. degree from the University of Vigo, Vigo, Spain, in 2008 and 2016, respectively.

He has been working at the Signal Theory and Communications Department in the University of Vigo since 2008 as an Associate Researcher. For a period of six months in 2012 he visited the Department of Information Engineering at University of Florence, Italy.

His research interests include multimedia forensics, digital watermarking and video coding.

Fernando Pérez-González (M'90-SM'09-F'16) received the Telecommunication Engineer degree from the University of Santiago, Santiago, Spain in 1990, and the Ph.D. degree in telecommunications engineering from the University of Vigo, Vigo, Spain, in 1993.

In 1990, he became an Assistant Professor with the School of Telecommunication Engineering, University of Vigo. From 2007 to 2010, he was Program Manager of the Spanish National R&D Plan on Electronic and Communication Technologies, Ministry of Science and Innovation. From 2009 to 2011, he was the Prince of Asturias Endowed Chair of Information Science and Technology, University of New Mexico, Albuquerque, NM, USA. From 2007 to 2014, he was the Executive Director of the Galician Research and Development Center in Advanced Telecommunications. He has been the Principal Investigator of the University of Vigo Group, which participated in several European projects, including CERTIMARK, ECRYPT, REWIND, NIFTY, and WITDOM.

He is currently a Professor in the School of Telecommunication Engineering, University of Vigo, Vigo, Spain, and a Research Professor in Information Science and Technology, University of New Mexico, Albuquerque, NM, USA. He has coauthored more than 60 papers in leading international journals and 160 peer-reviewed conference papers. He has coauthored several international patents related to watermarking for video surveillance, integrity protection of printed documents, fingerprinting of audio signals, and digital terrestrial broadcasting systems. His research interests include the areas of digital communications, adaptive algorithms, privacy enhancing technologies, and information forensics and security. Prof. Pérez-González was an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS (20052009) and the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY (20062010). He is currently is an Associate Editor of the LNCS Transactions on Data Hiding and Multimedia Security, and the EURASIP International Journal on Information Security.

Pedro Comesaña-Alfaro (M'08-SM'15) received the Telecommunication Engineer degree from the University of Vigo, Vigo, Spain in 2002 and the Ph. D. degree in telecommunications engineering from the same institution in 2006. Dr. Comesaña made research stays at the Technische Universiteit Eindhoven (The Netherlands, 2004), University College Dublin (Ireland, 2006), University of Siena (Italy, 2007-2008), University of New Mexico (Albuquerque, NM, USA, 2010-2011), and State University of New York at Binghamton (NY, USA, 2015). Since 2012 Dr. Comesaña is an Associate Professor at the University of Vigo. His research interests lie in the areas of multimedia security and digital communications. He has coauthored several international patents related to watermarking for video surveillance, and fingerprinting of audio signals. Dr. Comesaña has co-authored over 50 papers in international journals and peer-reviewed conferences; he was recipient of IEEE-WIFS 2014 Best Paper Award. Currently, Dr. Comesaña serves as Associate Editor of IEEE Transactions on Circuits and Systems for Video Technology, IEEE Signal Processing Letters, IET Information Security, and Journal of Visual Communication and Image Representation; he was a member of the IEEE SPS Information Forensics and Security-Technical Committee, Technical co-Chair of ACM IH&MMSec 2015, Area Chair of IEEE ICIP 2015, Tutorials Chair of IEEE WIFS 2012, and Publications Chair of IH&MMSec 2016.