

AN SVD APPROACH TO FORENSIC IMAGE RESAMPLING DETECTION

David Vázquez-Padín, Pedro Comesaña, Fernando Pérez-González

Signal Theory and Communications Department, University of Vigo, Vigo 36310, Spain

ABSTRACT

This paper describes a new strategy for image resampling detection whenever the applied resampling factor is larger than one. Delving into the linear dependencies induced in an image after the application of an upsampling operation, we show that interpolated images belong to a subspace defined by the interpolation kernel. Within this framework, by computing the SVD of a given image block and a measure of its degree of saturated pixels per row/column, we derive a simple detector capable of discriminating between upsampled images and genuine images. Furthermore, the proposed detector shows remarkable results with blocks of small size and outperforms state-of-the-art methods.

Index Terms— digital image forensics, resampling detection, singular value decomposition, tampering detection

1. INTRODUCTION

Digital images are nowadays a substantial medium for communicating with people and also one of the most effective ways to easily transmit information. With the current growth of Internet usage from mobile devices, a captured image may get the power of instantly distribute breaking information by simply sharing it in any social network. Consequently, as important as the instantaneous communication is, the authenticity of the sent information should also be quickly validated.

In this regard, during last years, a lot of techniques have been designed to automatically analyze the authenticity or integrity of a given image in a blind way [1]. Although most of the shared images are JPEG compressed, there is an increasing tendency to give support to raw image formats. For example, several image hosting websites support the upload of raw images and also some of the high-end smartphones bring the opportunity of capturing raw images. Therefore, tampering detection in uncompressed images will probably be a valuable tool in the near future.

Resampling detection as a means of unveiling forgeries is very appropriate for uncompressed images. Usually, when

image splicing is performed, one of the pasted regions has to be geometrically adapted to the background scene. The applied spatial transformation (e.g., resizing, rotation, etc.) will make use of a resampling operation which will unavoidably leave characteristic dependencies among neighboring samples that are not typically present in a genuine image. In the literature, we can find several techniques that are able to expose forgeries by detecting inconsistencies in such characteristic resampling traces [2–7].

Some of the proposed techniques work with a residue signal obtained by a global predictor [2], a fixed linear kernel [4], or a derivative filter [3]. The statistics of this signal reflect the underlying periodic correlations due to image resampling, so those schemes exploit such properties in the frequency domain. Alternatively, the study of the quantization applied after the resampling operation (i.e., due to rounding) has also been exploited to estimate the resampling factor in [7].

In this case, we are more interested in the approaches that study local linear dependencies induced by linearly interpolated spatial transformations [2–6]. A first attempt to characterize these linear dependencies through the Singular Value Decomposition (SVD) of a resampled image has been carried out in [6] by Wang and Ping, resorting to an SVM classifier to perform the detection of resampling. A deeper understanding of the linear correlations originated locally has been described by Kirchner in [5], where a local predictor per each row/column of the image is computed. By analyzing in the frequency domain the differences between the obtained predictor coefficients, Kirchner provided an effective resampling detector, especially for downsampling. However, as indicated in [7], the frequency analysis presents some drawbacks impairing the performance of the detector when a reduced number of samples is available and also when a regular structure or a periodic pattern is present in the image under analysis.

In relation to the last two works, and motivated by their shortcomings, here we investigate the local linear dependencies introduced once an upsampling process is applied to an image. As a result, we propose a very simple method that relies on the calculation of the SVD of a given block from an image, without resorting to an SVM classifier and being able to produce suitable results by processing blocks of small size. Note that we do not address the downsampling process in this paper, but we provide some insights about how the proposed detector could be adapted to deal with downscaled images.

This work was partially funded by the Spanish Government and the European Regional Development Fund (ERDF) under projects TACTICA, by the Galician Regional Government and ERDF under projects “Consolidation of Research Units” (GRC2013/009), REDTEIC (R2014/037) and ATLANTIC, and by the EU 7th Framework Programme under project NIFTY (HOME/2012/ISEC/AG/INT/4000003892).

The remaining of the paper is organized as follows: the theoretical analysis of the linear dependencies in upsampled images and the basic idea behind the proposed detector are explained in Section 2. The formal definition of the developed detector is tackled in Section 3, while the experimental results are treated in Section 4. Finally, conclusions and future work are reported in Section 5.

2. PROBLEM MODELING

Let us define a digital image with a single color channel as a $P \times Q$ matrix \mathbf{F} with elements $F_{p,q}$ and indices $p \in \{0, \dots, P-1\}$ and $q \in \{0, \dots, Q-1\}$. The values of each element $F_{p,q}$ are discrete quantities whose range is determined according to the image bit depth.

The resampling operation is assumed to be linear, so each pixel value in the resampled image \mathbf{G} is computed by linearly combining a finite set of neighboring samples coming from the original image. We consider that the applied resampling factor ξ uniformly scales each dimension of the original image and we define it as $\xi \triangleq \frac{L}{M}$, i.e., the ratio between the upsampling factor $L \in \mathbb{N}^+$ and the downsampling factor $M \in \mathbb{N}^+$. The application of this resampling operation involves two main steps: the definition of the resampling grid with the new pixel locations, and the computation of the values in those locations. In a single expression, each pixel value $G_{i,j}$ of the resampled image can be obtained as follows:

$$G_{i,j} = \sum_{k=0}^{P-1} \sum_{l=0}^{Q-1} h\left(i\frac{M}{L} + \delta - k\right) h\left(j\frac{M}{L} + \delta - l\right) F_{k,l}, \quad (1)$$

where δ denotes a shift between the two sampling grids¹ and $h(\cdot)$ represents the impulse response of an interpolation kernel whose length is denoted by N_h .

The analysis in the remainder of this section is applicable to any linear kernel regardless of its length or impulse response; however, for non-linear kernels, the forthcoming modeling could only be understood as an approximation.

Notice that after computing all the pixels of the resampled image, its values must fit the original resolution or bit depth of the input image. Therefore, as a last step, the resampled values must be quantized to the original precision, having

$$R_{i,j} = Q_{\Delta}(G_{i,j}), \quad (2)$$

where $R_{i,j}$ denotes each element of the quantized resampled image \mathbf{R} and $Q_{\Delta}(\cdot)$ represents a uniform scalar quantizer with step size Δ .

From the resampling operation shown in (1), it can be easily checked that any resampled value $G_{i,j}$ is calculated using exactly the same interpolation weights as for $G_{i+k_1L, j+k_2L}$ with $k_1, k_2 \in \mathbb{N}$. More precisely, for any $k_1, k_2 \in \mathbb{N}$, the

¹In MATLAB's function `imresize` and also in the tool `convert` from ImageMagick's software, the shift corresponds to $\delta \triangleq \frac{1}{2} \left(1 + \frac{M}{L}\right)$.

column-ordered vector $\mathbf{y} \in \mathbb{R}^{(L+2)^2}$ that is built up from an $L \times L$ block of the resampled image starting at sample G_{k_1L, k_2L} and adding a surrounding border of one pixel, is computed through the linear combination of samples from the column-ordered vector $\mathbf{x} \in \mathbb{R}^{(M+N_h)^2}$, which, in turn, is set up from an $M \times M$ block of the original image starting at sample F_{k_1M, k_2M} and adding a border of $\lfloor N_h/2 \rfloor$ pixels. Notice that the amount of border pixels to add to each block depends on the shift δ introduced between the original and the resampled grid.

According to this observation, each column-ordered vector from the resampled signal \mathbf{y} is obtained through the following linear transformation

$$\mathbf{y} = \mathbf{H}\mathbf{x}, \quad (3)$$

where \mathbf{H} represents the interpolation matrix containing the weights of the interpolation kernel. From the linear transformation in (3), it is clear that each column-ordered vector \mathbf{y} belongs to an $(M+N_h)^2$ -dimensional subspace of $\mathbb{R}^{(L+2)^2}$ generated by matrix \mathbf{H} . We will use \mathcal{Y} for denoting this subspace, which is defined as

$$\mathcal{Y} \triangleq \left\{ \mathbf{w} \in \mathbb{R}^{(L+2)^2} : \mathbf{w} = \mathbf{H}\mathbf{s}, \mathbf{s} \in \mathbb{R}^{(M+N_h)^2} \right\}.$$

However, note that due to the quantization applied after performing the resampling operation in (2), the observed vectors \mathbf{z} of length $(L+2)^2$ (i.e., starting at sample R_{k_1L, k_2L} with $k_1, k_2 \in \mathbb{N}$ and adding a surrounding border of one pixel), are a perturbed version of the interpolated ones. As long as the statistical distribution of the input signal is smooth and its variance is much larger than the square quantization step Δ^2 , each vector \mathbf{z} can be modeled as

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{n},$$

where the new variable \mathbf{n} is a random vector, whose components are i.i.d. with zero mean and variance $\sigma_n^2 = \frac{\Delta^2}{12}$ (i.e., the mean and variance of the quantization noise).

Based on this model, by stacking K vectors \mathbf{z} into a $K \times (L+2)^2$ matrix, we obtain an observation matrix \mathbf{Z}_K that can be represented in terms of its singular value decomposition, having

$$\mathbf{Z}_K = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

where $\mathbf{U} \in \mathbb{R}^{K \times K}$ is a unitary matrix whose columns are the left-singular vectors of \mathbf{Z}_K ; $\mathbf{\Sigma} \in \mathbb{R}^{K \times (L+2)^2}$ is a rectangular diagonal matrix whose diagonal elements σ_i (with $i \in \{0, \dots, (L+2)^2 - 1\}$), are known as the singular values of \mathbf{Z}_K which are sorted in descending order; and, finally, $\mathbf{V} \in \mathbb{R}^{(L+2)^2 \times (L+2)^2}$ is a unitary matrix with the right-singular vectors of \mathbf{Z}_K .

From this decomposition, it is expected that the $(M+N_h)^2$ dominant right-singular vectors of \mathbf{Z}_K (i.e., those corresponding to the largest singular values) span the signal subspace \mathcal{Y} (induced by the resampling operation), while the remaining ones span the noise subspace (caused by the rounding operation), i.e., for $i \geq (M+N_h)^2$ we have $\sigma_i \approx \sqrt{K\sigma_n^2}$, as $K \rightarrow \infty$.

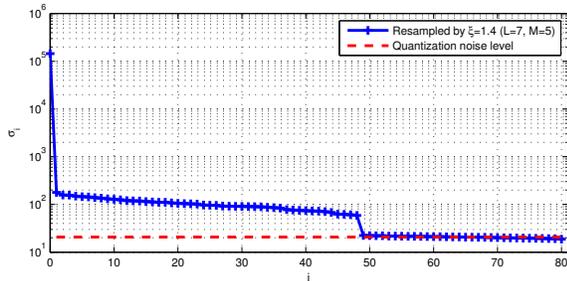


Fig. 1. Evolution in logarithmic scale of the singular values of matrix \mathbf{Z}_K built from a block of size 512×512 from the green channel of a quantized resampled image without demosaicing traces. $\xi = \frac{7}{5} = 1.4$ and linear kernel.

In Fig. 1, we show an example of the evolution of the singular values when matrix \mathbf{Z}_K is built from a block of size 512×512 of an image resampled by $\xi = \frac{7}{5} = 1.4$ with a linear kernel ($N_h = 2$). It is easy to see how the first $(M + N_h)^2 = 49$ singular values (out of $(L + 2)^2 = 81$) have a magnitude above the quantization noise level $\sqrt{K\sigma_n^2}$ (with $K = 5184$ and $\Delta = 1$), as it was anticipated.

Note that the proposed scheme assumes L to be known at the detector; of course, this does not hold in real forensic scenarios. As a plausible solution, an iterative procedure could be considered covering all the possible values of L , but this would significantly increase the computational burden. Therefore, we simplify the process of resampling detection by directly resorting to the singular value decomposition of an image block. Then, we explore when the singular values vanish in such a manner that the signal subspace is discernible from the noise subspace.

2.1. Practical solution

Let us define \mathbf{Z} as a matrix gathering pixel intensity samples from a block of size $N \times N$ of a quantized resampled image \mathbf{R} under test. Due to the presence of noise (e.g., rounding errors after resampling), we will initially assume that matrix \mathbf{Z} has N non-zero singular values, i.e., \mathbf{Z} has full rank, but at the end of this section, a discussion on how to manage rank-deficient matrices will also be introduced.

The rows (or columns) of \mathbf{Z} can be treated as N points belonging to an N -dimensional space. In the previous analysis, we have seen throughout the application of the SVD to \mathbf{Z}_K that in a resampled image, the vectors of length $(L + 2)^2$ can be represented by $(M + N_h)^2$ dimensions, yielding a dimensionality reduction of a factor around ξ^2 (i.e., the applied up-sampling factor in each direction). By considering now each row/column of \mathbf{Z} as a vector of length N , it will be possible to represent the set of N points with a smaller number of dimensions $k \approx \frac{N}{\xi}$, since each single row/column has been resampled solely by ξ .

This will be reflected in the calculation of the SVD of \mathbf{Z} , where only the first $k \approx \frac{N}{\xi}$ singular values will have a con-

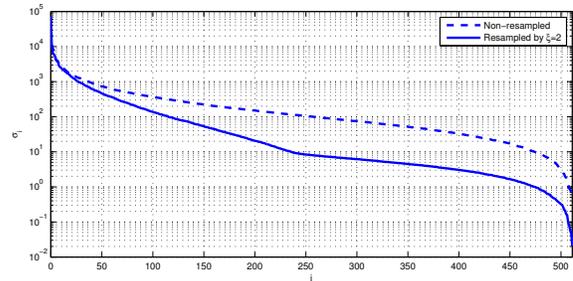


Fig. 2. Evolution in logarithmic scale of the singular values of an image block \mathbf{Z} of size 512×512 . Dashed lines represent the results from a non-resampled image and solid lines correspond to an image resampled with $\xi = 2$ and linear kernel.

siderably larger magnitude than the rest. Conversely, if we take a block \mathbf{Z} from a never-resampled image, there will not exist such characteristic linear dependency between neighboring samples and all singular values will have a magnitude significantly larger than the noise level in a resampled image.

Fig. 2 draws a comparison in terms of singular values in two different cases, i.e., when matrix \mathbf{Z} is built from a block of size 512×512 from the green channel of a non-resampled image without demosaicing traces (dashed line) and from its resampled version, obtained by using $\xi = 2$ and a linear interpolation kernel (solid line). In both cases, matrix \mathbf{Z} has $N = 512$ non-zero singular values, but as it can be checked in Fig. 2, the magnitude of the singular values from the resampled image drops more sharply than the corresponding one coming from the non-resampled image. For the resampled image, the number k of singular values significantly larger than the noise level approaches $\frac{N}{\xi}$, so in this particular case $k \approx \frac{512}{2} = 256$. This is due to the fact that approximately half of the samples of each row/column of \mathbf{Z} can be computed in this resampled case as a linear combination of the remaining ones. For non-resampled images, there should only be a significant drop-off in indices close to the rank of the matrix under analysis.

Since we are assuming the applied resampling factor to be larger than one, we can state that typically the total amount of variance of the input signal explained by those singular values with indices smaller than $i \approx \frac{N}{\xi} - 1$ for any image resampled by ξ , will be larger than for a non-resampled image. As a consequence, the magnitude of the singular values at such index is also expected to be smaller. This fact indicates that a statistic accounting for the magnitude of a singular value at the correct position can be discriminative for detecting the application of a resampling operation in a given image block.

As indicated above, any matrix \mathbf{Z} built from an image block is assumed to have N non-zero singular values; however, in practice, undesirable artifacts as pixel saturation may arise, thereby removing part of the noise due to rounding and yielding singular values with negligible magnitude. The presence of saturation and the possibility of having some linear

dependency between samples will affect the expected evolution of the singular values of \mathbf{Z} , producing two possible outcomes:

1. The number of non-zero singular values is substantially smaller than N . This rarely happens unless several rows/columns of matrix \mathbf{Z} are completely saturated.
2. The number of non-zero singular values is close to N , but their magnitude vanishes more sharply than usual. This takes place when linear dependencies are present among the rows/columns of \mathbf{Z} , and can be boosted by saturations.

Therefore, the detector will have to deal with the degree of saturation borne by any image block under analysis and it will also probably consider a means for deciding when a computed singular value is negligible. In the next section, the adopted detection strategy is described.

3. PROPOSED DETECTOR FOR $\xi > 1$

From the analysis carried out on the previous section, it is clear that by exploiting the magnitude of the singular values at a certain index we can derive a hypothesis test for image resampling detection. In the definition of our hypothesis test, the observed data consists of an $N \times N$ matrix \mathbf{Z} containing a block of samples coming from one of the color channels of a digital image which may have been resampled or not. Under the null hypothesis, i.e., \mathcal{H}_0 , we assume that the observed data \mathbf{Z} has never been resampled; while, under the alternative hypothesis, i.e., \mathcal{H}_1 , we assume that the observed data has been resampled by any factor greater than one.

The definition of the test statistic ρ will depend on the degree of saturation that the image block under analysis may have experienced, as pointed out in the last part of Section 2.1. By denoting γ_{row} (respectively, γ_{column}) as the quotient between the total number of saturated pixels in a block (i.e., pixels equal to 255 for an 8-bit depth image) and the number of rows (respectively, columns) that contain at least one saturated pixel, the degree of saturation s , is defined as

$$s \triangleq \frac{1}{N} \max \{ \gamma_{\text{row}}, \gamma_{\text{column}} \}.$$

On the other hand, to determine which of the computed singular values are negligible, we use a tolerance level $\epsilon = 2^{-52}$ (i.e., the spacing of double precision numbers in MATLAB). Moreover, we define a variable $r \leq N$ that represents the total number of singular values above this tolerance level. Accordingly, the proposed test statistic is:

$$\rho \triangleq \begin{cases} 0, & \text{if } r < 0.1N, \\ \log(\sigma_{\nu - \lfloor 0.05N \rfloor}), & \text{if } s \geq 0.45 \text{ and } r > 0.95N, \\ \log(\sigma_{\nu-1}), & \text{otherwise,} \end{cases} \quad (4)$$

where $\nu = \lfloor \frac{r}{\xi_{\min}} + 0.5 \rfloor$ represents the rounded version of the maximum number of significant dimensions that could be achieved by a resampled image with any $\xi \geq \xi_{\min}$. Hence,

ξ_{\min} is the minimum resampling factor that can be detected by our detector. Notice that the first two cases contemplated in (4) have been heuristically derived and are settled to avoid the two effects caused by pixel saturation discussed at the end of Section 2.1.

Assuming all the particularities for obtaining the test statistic, we expect to find larger values of ρ for non-resampled images, thus accepting the hypotheses according to the following conditions:

$$\mathcal{H}_0 : \rho > T,$$

$$\mathcal{H}_1 : \rho \leq T,$$

where T is a predefined threshold. Several experiments are performed next to study the validity of the proposed approach.

4. EXPERIMENTAL RESULTS

The designed detector is tested over all the uncompressed images belonging to the Dresden Image Database [8] and stemming from Nikon cameras (a total of 1317 images). To perform each full-frame resampling operation, the image processing tool `convert` from ImageMagick's software is used. As interpolation kernels, we select those that are commonly available in any image processing tool, namely: *Linear*; from the family of cubic filters we choose *Catmull-Rom* and *B-spline*; and, finally, a three-lobed *Lanczos*-windowed kernel. The employed discrete set of resampling factors is defined in the interval $[1.05, 2]$ (sampled with a step size of 0.05), given that these are the most appropriate upsampling factors to avoid the introduction of visible distortions.

Since our main objective is to unveil tampered regions (which might be small) through the detection of resampling inconsistencies, we are interested in studying the achieved performance of our detector with blocks of small size, thus leading us to process $N \times N$ image blocks \mathbf{Z} with $N = 32$. The analysis of resampling traces is then carried out by taking the center 32×32 block of the green channel from each image under study. The evaluation of the performance of the proposed detector is conducted in terms of AUC (Area Under the Receiver Operating Characteristic (ROC) curve) and detection rate at a fixed False Alarm Rate (FAR), i.e., concretely at $\text{FAR} \leq 1\%$. All results are compared with the detector proposed by Kirchner in [5], because this method outperforms Popescu and Farid's detector [2], which is usually considered to be the most reliable detector.

The performance analysis is twofold: firstly, images without demosaicing traces are processed,² and, secondly, demosaiced images are tested. In each of these cases, both detectors must also be applied on all the non-resampled images in order to fix the detection thresholds (i.e., T). For our test statistic we take $\xi_{\min} = 1.05$, and a neighborhood size $K = 3$ for Kirchner's detector, as specified in [5].

²Specifically, non-resampled images without demosaicing traces are constructed by getting access to the output of the camera sensor (through the image processing tool `dcraw`) and then picking always the same-positioned green pixel from the two available samples in each 2×2 Bayer pattern.

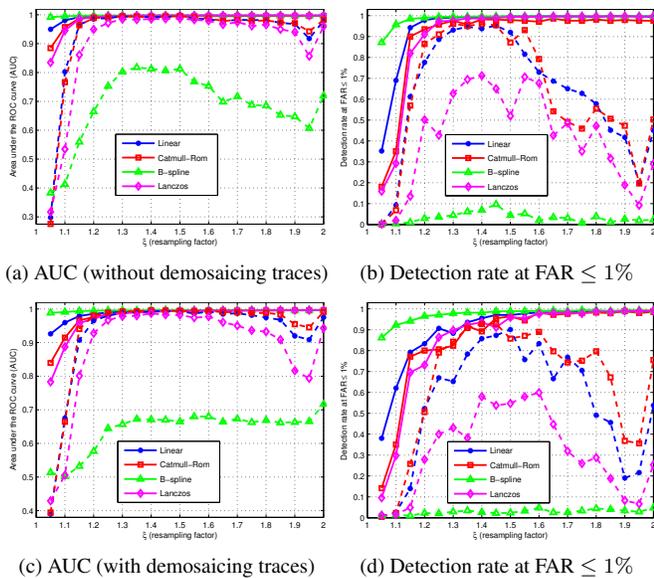


Fig. 3. Evaluation of the proposed detector (solid lines) against Kirchner’s detector (dashed lines) in terms of AUC and detection rate for blocks of size 32×32 . The first row contains the results from images without traces of demosaicing, while the second row is for images with demosaicing traces.

The first row of Fig. 3 shows the performance of the proposed approach when testing images without demosaicing traces. From these results, we can state that our method shows better performance with B-spline and Linear interpolation kernels than with Catmull-Rom and Lanczos, which commonly get the worst results. Our detector presents some difficulties with resampling factors close to one, i.e., for $1.05 \leq \xi \leq 1.2$, whereas less issues arise when the resampling factor approaches 2. Although not being reported, additional experiments have been performed increasing the size of the block (e.g., with $N = 128$), obtaining values of $\text{AUC} \geq 0.998$ for all tested filters and $\xi \geq 1.1$.

An interesting aspect is that our detector shows a strong gain with respect to Kirchner’s when images are resampled with the B-spline kernel, regardless of the value of ξ . For instance, in Fig. 3(b), Kirchner’s detection rate is below 0.1 for all tested ξ , while our detector shows a detection rate almost always above 0.9 (excepting $\xi = 1.05$). Therefore, the proposed SVD-based analysis has proven to be very convenient for B-spline resampling detection. A second stimulating feature is that our detector needs a very small set of samples (i.e., 32×32 pixels) to work remarkably well, while this particular size starts to be a problem for Kirchner’s detector.

The second row of Fig. 3 collects the results arising from images with demosaicing traces. By comparing the achieved outcomes in this case with respect to the previous ones (i.e., without traces of demosaicing), it becomes apparent that our detector works better when it has to distinguish purely non-resampled (i.e., non-demosaiced) images against their upsam-

pled version. The reason is that when non-resampled images exhibit demosaicing traces, there exist unavoidable linear dependencies which affect the expected value of the statistic ρ for genuine images. Usually, these linear correlations caused by the demosaicing process are not so strong as the ones introduced by the resampling operation (mainly because current demosaicing algorithms are adaptive and, commonly, non-linear), but this will harm to some extent the idea behind the use of the SVD as a means to distinguish linearly correlated data against uncorrelated data.

Apart from this global lost in performance, the behavior of our detector is almost identical to the one discussed for images without demosaicing traces. In general, all the experimental results show that our detector is a reliable solution for image resampling detection.

5. CONCLUSIONS

In this paper, a simple strategy for resampling detection has been derived. The proposed detector only needs to compute the SVD of a given image block and a measure of its degree of saturated pixels per row/column, for discerning upsampled images from genuine ones. The achieved performance is promising and when compared with Kirchner’s state-of-the-art method, our detector outperforms it.

As future work, we are planning to apply the same idea for detecting resampling operations by factors smaller than one, but jointly exploiting the traces left by the demosaicing process and the downsampling in the three color components of a digital image.

REFERENCES

- [1] A. Piva, “An overview on image forensics,” *ISRN Signal Processing*, vol. 2013, 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, 2005.
- [3] 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, 2008.
- [4] M. Kirchner, “Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue,” in *ACM Multimedia and Security Workshop*, 2008, pp. 11–20.
- [5] M. Kirchner, “Linear row and column predictors for the analysis of resized images,” in *ACM Multimedia and Security Workshop*, 2010, pp. 13–18.
- [6] R. Wang and X. Ping, “Detection of resampling based on singular value decomposition,” in *5th International Conference on Image and Graphics*, 2009, pp. 879–884.
- [7] D. Vázquez-Padín, P. Comesaña, and F. Pérez-González, “Set-membership identification of resampled signals,” in *5th IEEE International Workshop on Information Forensics and Security*, 2013, pp. 150–155.
- [8] T. Gloe and R. Böhme, “The Dresden Image Database for benchmarking digital image forensics,” in *ACM Symposium on Applied Computing*, 2010, vol. 2, pp. 1584–1590.