Reliable Demosaicing Detection for Image Forensics

Quentin Bammey*, Rafael Grompone von Gioi[†], Jean-Michel Morel[‡] CMLA, cole Normale Suprieure Paris-Saclay, Cachan, France

Email: *quentin.bammey@cmla.ens-cachan.fr, [†]grompone@cmla.ens-cachan.fr, [‡]morel@cmla.ens-cachan.fr

Abstract—Visually plausible image forgeries are easy to create even without particular knowledge or skills. However, most forgeries unknowingly alter the underlying statistics of an image. In particular, demosaicing artefacts created by the camera are usually destroyed or modified when the image is tampered. Most of the literature focus on detecting where these traces are destroyed, and generally do it in a way that still requires a visual interpretation. We introduce an *a contrario* method which detects global demosaicing parameters, and then checks for regions of the image which are inconsistent with these parameters. Detections are guaranteed in the form of a number of false alarms (NFA), which enables the user to control the false positive rate. Such a guarantee is a useful complement to existing methods, and enables inclusion into fully automatic image authentication processes. The source code and an online demo are provided with the article.

Index Terms—image forgery, forgery detection, forgery, CFA interpolation, CFA, color filter array, demosaicing, demosaicking, filter estimation, linear estimation, a contrario, tampering, artefact detection, Bayer matrix

I. INTRODUCTION

Images are frequently presented as proofs of events or facts. However, it is nowadays easy, even for novices, to alter an image and its meaning while keeping it visually plausible. This has given rise to a plethora of forged images, used in social media for fake news and misinformation, or in academic papers where images are sometimes used as proofs of results in experiments. It is therefore of utmost importance to design reliable methods detecting forgeries in images [1].

Although image forgeries are varied in nature, different kinds of clues can be used to detect them. In this paper, we focus on detecting demosaicing artefacts left by the camera, and their use in finding and localising forgeries. Most cameras can only sample one colour at each pixel, and have to interpolate the missing information with neighbouring pixels. The most common sampling pattern is the Bayer matrix (see Fig. 1), which samples two green pixels for one red and one blue pixel.

Missing colour values must be interpolated on each pixel from its neighbourhood. This process is known as demosaicing, or Colour Filter Array (CFA) interpolation [2]. It leaves artefacts that can be detected to get information about the

This research is funded by the French Ministère des armées Direction



$R_{0,0}$	$G_{0,1}$	$R_{0,2}$	$G_{0,3}$	
$G_{1,0}$	$B_{1,1}$	$G_{1,2}$	$B_{1,3}$	
$R_{2,0}$	$G_{2,1}$	$R_{2,2}$	$G_{2,3}$	
$G_{3,0}$	$B_{3,1}$	$G_{3,2}$	$B_{3,3}$	
:	:	:	:	۰.
•				•

Fig. 1: Top left portion of a CFA image obtained with a Bayer matrix ${}^{R}_{G}{}^{G}_{B}$. At each pixel, a single color (either red, green, or blue) is sampled.

demosaicing process in the image, which can in turn reveal forgeries by finding regions inconsistent with the global image.

The simplest interpolation method is known as bilinear demosaicing, where each missing colour is interpolated as the mean of adjacent pixels sampled at the same colour. Although the traces left by this method are simple to detect, most commonly used demosaicing algorithms are more sophisticated and thus harder to detect [2].

Popescu and Farid [3] proposed to estimate the demosaicing method through a linear filter with an expectationmaximisation algorithm, which alternately computes a probability map of each pixel being interpolated and finds the linear filter best explaining the interpolated pixels. They use the Fourier transform of the probability map to check for periodicities of period 2, which reveal the existence of CFA interpolation. Bayram et al. expand on this method in [4] to identify the camera model. By separating the smooth and nonsmooth regions of the image in the EM algorithm, they enable a better separation of the demosaicing algorithms.

Chen and Stamm [5] tackle the problem of camera model identification in another way. They note that building parametric models for demosaicing can be hard as the algorithms used are often complex and proprietary. They decide instead to use an ensemble of submodels, each carrying parts of the demosaicing artefacts, and merge them with standard classifying algorithms to identify the camera.

Most methods of forgery detection that make use of CFA artefacts to find forgeries do it by looking for places where CFA interpolation artefacts have been destroyed through tampering. However, getting reliable detections through this is made difficult by the fact that CFA interpolation traces can be naturally absent in parts of an image for structural reasons. Instead, we focus on a similar, but different problem. When a copy-move forgery is made, if the copied part has undergone a demosaicing too, the copied object should present CFA

artefacts. When pasted onto the image, there is a $\frac{3}{4}$ probability that the position of the Bayer Matrix will be different in the forged region and in the global image. With this in mind, we are interested first in detecting the global CFA pattern of an image, and subsequently the local CFA patterns.

To do that, Kirchner [6] uses the fact that bilinear interpolation is a good estimation of demosaicing algorithms in a large part of the image. For each possible CFA pattern of the image, they apply a pseudo-inverse of this demosaicing algorithm to the corresponding subgrid and generate an error map. The most likely pattern is then identified as the one with the lowest error. Choi et al. [7] noticed that interpolated pixels are usually an average of part of their neighbours; as a consequence they tend to have more often intermediate values than originally sampled pixels. They use this strategy to identify the correct Bayer matrix position, and to detect colour modifications in images by detecting when the green channel is transposed to the red and blue channels [8].

State-of-the-art methods already provide decent results for CFA pattern classification, even for small windows, provided the JPEG compression is not too drastic. However, the results of such a classification are not sufficient when it comes to detecting forged images. Indeed, an incorrect CFA diagnose on a single image block is enough to consider an image as falsified. Thus, a method with a 95% accuracy on 128×128 blocks would risk detecting a falsification nine times out of ten in a standard, authentic image of size 1024×768 . In other words, controlling the false alarms rate over the many blocks of an image is necessary to detect forgeries. In a preliminary work [9], we expanded the method of [6] to use information about the demosaicing algorithm. By estimating linear filters for demosaicing in all four possible grid positions, we could generate error maps that were closer to the used method. This provided a simple way of controlling the number of false positives as we could then be certain that the found pattern was correct and that the image had effectively gone through demosaicing. On top of that, we added a statistical verification of our results by making blocks in the image vote on the most likely position. However, natural variations in the estimation of filters could induce cases where one of the Bayer patterns would seem better than others even in the absence of any demosaicing. Furthermore, block voting implied a sensitivity loss making it harder to find forgeries with small area.

In this article, we propose a new method to reliably detect the correct position of the Bayer matrix. We start by applying a high-pass filter tailored to highlight the difference between original and interpolated pixels. We then use a statistical test to compare the samples of the possible grid positions, in order to know which positions are significantly impossible in the image. Comparing the results in the global image and across smaller windows enables us to detect and localise forged regions in images.

An online demo for this method is available at https://bit.ly/2UklsAR, and the source code can be downloaded at https://bit.ly/2GYSLGv.

II. METHOD

A. The a contrario paradigm for reliable detections

One of the key goals of our method is not only to find which of the four possible patterns ${}^{R}_{G}{}^{G}_{B}$, ${}^{G}_{B}{}^{R}_{G}$, ${}^{G}_{R}{}^{B}_{G}$ and ${}^{B}_{G}{}^{G}_{R}$ have been used, but to know how confident we are in our detection so as to control the number of false alarms.

Let $x_p, p \in \{ {}^{R}_{G} {}^{G}_{B}, {}^{G}_{B} {}^{R}_{G}, {}^{G}_{R} {}^{B}_{G}, {}^{G}_{R} {}^{R}_{G} {}^{P}_{G} {}^{R}_{R} {}^{G}_{R} {}^{P}_{R} {}^{G}_{R} {}^{P}_{R} {}$

Assuming that an image has been demosaiced, the position of the Bayer matrix can be obtained by taking the mean of all x_p and selecting the highest one. However, it may be that the image has not been demosaiced, or that the traces of demosaicing can no longer be found, for example because of post-processing effects or because the image is too small. How then can we be certain that the values our algorithm returns are correct?

In order to get a reliable detection, we make use of the *a* contrario paradigm [10], [11]. The approach is based on the non-accidentalness principle, according to which an observed geometric structure is perceptually meaningful only when its expectation is low under random background model. Detection thresholds can then control the expected number of false detections in this background, or *a contrario* model H_0 . An observed structure is validated only when a test rejects the H_0 hypothesis. The detection threshold must take into account our multiple testing, as in Gordon et al. [12]. Assume that N_T tests are performed and that a variable u is observed at each one. We desire to set a threshold τ such that the sought structure is validated when $u \geq \tau$. Following the *a contrario* methodology, we define the Number of False Alarms (NFA) of a candidate by

$$NFA = N_T \mathbb{P}_{H_0}(U \ge u), \tag{1}$$

where $\mathbb{P}_{H_0}(U \ge u)$ is the *p*-value of observing under random hypothesis H_0 a value *U* as large or equal to *u*. The candidate is validated whenever NFA $< \varepsilon$, where ε is a predefined accepted mean number of false detections. This yields an implicit value for the threshold τ . Accepting all detections with a NFA score of $\varepsilon = 10^{-3}$, we should expect an average of one false positive for 1,000 images.

For CFA detection, we compare two samples x_1 and x_2 , each having n non-negative values. We determine whether the first sample has significantly larger values than the second. For this, we compute the Mann-Whitney U statistic [13],

$$u = \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbb{1}_{\{x_1^i > x_2^j\}}$$
(2)

where $\mathbb{1}$ is the indicator function and x_p^k is the k-th value of sample p. The value of u belongs to $[0, n^2]$. This value is zero when all samples of x_1 are smaller than the samples of x_2 . Conversely, $u = n^2$ when all samples of x_1 are larger than the ones of x_2 . We can now define a natural background model for our statistical test. Its null hypothesis H_0 is that all samples in both X_1 and X_2 were independently drawn from the same distribution, which must be correct if no demosaicing has been performed. Thus, $\mathbb{P}_{H_0}(X_1^i > X_2^j) = \mathbb{P}_{H_0}(X_1^i < X_2^j) = \frac{1}{2}$ for any i, j. This defines the Mann-Whitney U test [13] with the corresponding random variable U. For large samples, U is approximately normally distributed, which allows a simple computation of the p-value $\mathbb{P}_{H_0}(U \ge u)$.

To sum up, given two samples x_1 and x_2 , the associated U statistic is computed with eq. 2. Then, the NFA value is given by eq. 1 and by the *p*-value of the Mann-Whitney test. Finally, the sample x_1 is declared significantly larger than x_2 if NFA $< \varepsilon$.

B. Detecting the possible CFA patterns

The first thing to do is to find which pixels are sampled and which have been interpolated. An easy way to do this is to apply a differential operator like the discrete Laplacian, which highlights extremal values and thus pixels likely to have been interpolated. Following the example of [6] or [3] or our preliminary work in [9], it is also possible to apply in each of the four possible grid positions a demosaicing algorithm – either fixed as in [6] or estimated as in [3] – and to compare the residuals.

One could simply compute a heat map with a differential operator – which can be implemented as a linear convolution followed by a point-wise absolute value operator. Yet this would not take into account that, depending on their positions on the Bayer matrix, the pixel subset from which interpolation is performed may vary. On the other hand, using a demosaicing algorithm in four positions must be done with caution. Preliminary experiments suggest that using a fixed demosaicing algorithm, such as the bilinear algorithm, does not yield good results when it is too different from the algorithm that was used to process the image, and an estimation of the demosaicing algorithm is prone to bias towards one of the four possible patterns.

As the possible patterns will be compared in pairs, we follow another approach that gives us both the simplicity and impartiality of differential operators and the ability to take different interpolation cases into accounts. Each test is a comparison between two patterns, pixels that are not sampled are thus interpolated using values from a known subset of pixels.

For example, if we compare the CFA patterns ${}^{R}_{G}{}^{G}_{B}$ and ${}^{G}_{B}{}^{R}_{G}$, we know the considered red values can only be interpolated from horizontal neighbours, blue values from vertical neighbours, and green values can be interpolated from all four adjacent neighbours, because none of the other pixels would be original in any of the two considered patterns. Furthermore, we note that when an interpolation can be done in two directions

simultaneously, many algorithms can decide to interpolate in only one direction, mainly to prevent interpolating across a strong edge. In order to mimic this behaviour, we do the interpolation simultaneously and then separately in the two directions, then for each pixel we take the result with the lowest absolute residual.

This leads us to define not one, but four different heat maps that will be used in the different comparisons. Let I be an image of shape $(X, Y, 3)^1$, we define the horizontal and (resp. vertical) heat map $H_h[x, y, c]$ (resp. $H_v[x, y, c]$) as the absolute difference between I[x, y, c] and its two horizontal (resp. vertical) neighbours. They correspond to the interpolation of red or blue pixels on green-sampled locations. The straight cross heat map $H_c[x, y, c]$ is equal to either the absolute difference between I[x, y, c] and the mean of its two horizontal neighbours or its two vertical neighbours or both, whichever of the three possibilities yields the lowest result. It corresponds to the interpolation of green pixels. Finally, the diagonal cross heat map $H_d[x, y, c]$ is equal to the absolute difference between I[x, y, c] and the mean of its two or four diagonal neighbours following either or both of the diagonals, whichever of the three possibilities yields the lowest result. It corresponds to the interpolation of red pixels on blue-sampled locations, and vice versa.

With these four heatmaps, we now describe how to reliably compare different CFA positions. We start by assuming that each of the four CFA patterns ${}_{G}{}_{B}{}_{G}{}_{B}{}_{G}{}_{G}{}_{R}{}_{G}{}_{G}{}_{B}{}_{G}{}_{a}$ and ${}_{G}{}_{R}{}_{G}{}_{a}$ are possible, and we look for CFA patterns that are significantly impossible. A grid position is considered significantly impossible if it is inferior to another grid position and the NFA score of the comparison between those too positions is below the set threshold. For each pair of CFA patterns $i, j \in \{{}_{G}{}_{B}{}_{G}{}_{G}{}_{B}{}_{G}{}_{G}{}_{R}{}_{G}{}_{G}{}_{G}{}_{R}{}_{G}{}_{R}{}_{G}{}_{R}{}_{G}{}_{R}{}_{G}{}_{R}{}_{R}{}_{G}{}_{R}{}_{R}{}_{G}{}_{R}{}_{R}{}_{G}{}_{R}{}_{R}{}_{G}{}_{R}{}_{R}{}_{R}{}_{G}{}_{R}{}_{R}{}_{R}{}_{G}{}_{R}{}_{R}{}_{R}{}_{G}{}_{R}{}_{R}{}_{R}{}_{G}{}_{R}{$

The first case is when both positions share the same green pattern, ie. if we are either comparing ${}_{G}{}_{B}{}_{B}^{G}$ with ${}_{G}{}_{R}{}_{R}^{G}$ or ${}_{B}{}_{G}{}_{R}^{R}$ with ${}_{R}{}_{G}{}_{B}^{G}$. Then we can only perform pattern comparisons using the red and blue channels. Since the interpolation between these two grids is done in diagonal, We look at the red and blue pixels of H_{d} . We average over each 2×2 block the red and blue pixels corresponding to CFA pattern *i*, and we do the same for *j*. We then compare the two samples and multiply its score by 6 since all pairs are compared. If the comparison is coherent, in other words if its score is below the set NFA threshold, then the identified weaker of the two grids is marked as impossible.

The second case is when both positions do not share the same green pattern, ie. if ${}^{R}_{G}{}^{G}_{B}$ or ${}^{B}_{G}{}^{G}_{R}$ is compared with ${}^{G}_{B}{}^{R}_{G}$ or ${}^{G}_{R}{}^{B}_{G}$. We then can perform pairwise comparisons in all three channels. As most demosaicing algorithms start by interpolating the green channel, and use its values to demosaic the other two channels, we consider the green channel separately: For the green channel, we look at the pixels

¹where the last dimension represents the colour channels.

of H_c and compare those that correspond to an original pixel in *i* to those that correspond to an original pixel in *j*. For the red and blue channel, we also compare those that correspond to each position, but we look in different heatmaps depending on the compared patterns: If we are comparing ${}^{R}_{G}{}^{G}_{B}$ with ${}^{G}_{B}{}^{R}_{G}$ or ${}^{B}_{G}{}^{R}_{R}$ with ${}^{G}_{R}{}^{B}_{G}$ we look at H_h for the red channel and H_v for the blue channel. In the other two cases, we look at H_v for the red channel and H_h for the blue channel. Each comparison score is multiplied by 6 as above and by two since green values are treated differently. Then, if the comparison of the green channel is significant, and if at least one of the two red and blue channels is significant and coherent with the green channel, the weaker of the two patterns is marked as impossible.

C. Finding forgeries

The main use of demosaicing artefacts is to find forgeries in an image. This is done by partitioning the heat map H in small windows. As with the global image, we decide in each window which grid positions are possible and which are significantly not, with the difference that we multiply the comparison scores by the number of windows before thresholding on the NFA score, as we want to control the number of false alarms in each image, and not just in each window.

If a window is inconsistent with the global image, ie. if there is no grid position which is possible both in the global image and the window, then a forgery has been identified and localised. An example of this result can be found in Fig. 2a. In a genuine image, there should also be at least one grid position which remains possible throughout all windows. If this is not the case, then a forgery has been identified. The windows having an impossible grid configuration may contain a forgery. An example of this result can be found in Fig. 2c.

Since grid detection is usually easier and thus more precise in the global image than in smaller windows, inconsistent grids is the primary way of detecting forgeries. However, if an image is fabricated from several images of similar sizes, or if an object recovering an important part of the image itself has been copy-moved, the CFA grid position of the forged part of the image may have been accepted as possible in the global image. In such cases, impossible grid configurations could detect forgeries not found by the former.

III. EXPERIMENTS

As a sanity check of our method, we confirm the absence of significant detections in images which have not been demosaiced. We used a set of 19 high-quality images – available in the linked online demo – that have been downsampled by a factor of 8 to remove all traces of demosaicing or JPEG encoding, as well as images of uniform and normal noise of different sizes: for each kind of noise, we made 10 images of size 128×128 , 10 of size 256×256 , 10 of size 512×512 and 10 of size 1024×1024 , for a total of 80 noise images. The dataset was tested uncompressed as well as with JPEG compression of qualities 100, 99, 98, 95, 90, 80, 70, 60, 50, 30 and 10. As expected, no traces of demosaicing were found







(b) Original image



(c) Regions where the respectively ${}^{R}_{G}{}^{G}_{B}$, ${}^{G}_{B}{}^{R}_{G}$, ${}^{G}_{R}{}^{B}_{G}$ and ${}^{B}_{G}{}^{G}_{R}$ Bayer matrix positions are deemed possible (in green) or significantly impossible (in red).

Fig. 2: Example of forgery detection. Images from [14].

by our algorithm. Choi et al.'s algorithm [7] was designed to always select one CFA grid; naturally, when applied to images which have not been demosaiced, it gives false detections in all cases. The NFA formulation allows the proposed method to make this decision automatically.

For the next experiment we demosaiced our 19 images with bilinear (BI) demosaicing, adaptive inter-channel correlation (AICC) demosaicing [15], and countour stencils (CS) demosaicing [16]. We tested the images uncompressed and compressed with the qualities above in Fig. 3a, as well as uncompressed but cropped from the centre into images of size 256×256 , 128×128 , 64×64 , 32×32 and 16×16 in Fig. 3b.

Before computing the accuracy of the detection, we rejected all results of our algorithm whose NFA was above $1 \cdot 10^{-3}$, as we would usually do. This means that even if the correct grid was found by our algorithm, we only accept the result if it was significant. We compared the resulting accuracy with Choi et al.'s algorithm [7]. In addition, we also use our method without requiring channel coherence: when two compared positions share their green sampled pixels, we compare them through the mean of their red and blue channels, and when they do not share their green channel, we base our decision on those pixels only. This variation is labelled as "proposed, classify".

We note that every time an incorrect pattern was found with our algorithm, the NFA associated with the detection was above 1, which proves the reliability of our algorithm.

The proposed method correctly detects most of the CFA patterns even at JPEG qualities as low as 70 or for images as small as 32×32 pixels.

Although the accuracy of [7] is usually higher, our results are actually more useable: though less patterns are detected, we can be confident on those that are identified. On the contrary, since [7] always output a grid, we can have no certainty in its output. When no NFA threshold is required, our method becomes stronger than Choi. This compensate the



(a) Accuracy of detections for varying JPEG qualities



(b) Accuracy of our detections for varying crop sizes

Fig. 3: Comparison of our method with Choi [7]

inevitable loss of accuracy we have when we require reliable detections, and enables us to still have many detections even when we need to be confident on our results.

The ability of the proposed algorithm to find the CFA pattern even in small grids and compressed images, and most importantly its reliable control of the false positives rate, make it particularly suited for forgery detection.

IV. CONCLUSION

In this article, we constructed a method highlighting the difference between original and interpolated pixels in demosaiced images. This method does not require an estimation of the CFA interpolation algorithm, but still uses knowledge on the specifics of demosaicing. We then explained how this method could be used to find not only traces of demosaicing, but also to get information on the position of the Bayer matrix. Inconsistencies about this information is a very strong clue of tampering in images, especially since the proposed approach yields a strong control of the false positives rate.

However, neither the proposed method nor the state of the art take into account the extensive inter-channel correlation that takes place during demosaicing. Without taking this entanglement into account, it is either necessary to accept the results of the strongest channels – which leads to strong classification scores but prevents all control of the false alarms rate – or to only accept results that are coherent across channels – effectively enabling a control of the false alarms rate, but rendering the method unable to decide on many windows.

Conjointly using the three colour channels to detect the CFA pattern is difficult, mostly because the existing demosaicking algorithms behave differently. But this approach will be necessary to improve the detection of the CFA pattern.

REFERENCES

- [1] H. Farid, Photo Forensics. The MIT Press, 2016.
- [2] X. Li, B. Gunturk, and L. Zhang, "Image demosaicing: A systematic survey," in VCIP, 2008.
- [3] A. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *Trans. Sig. Proc.*, vol. 53, no. 10, pp. 3948–3959, Oct. 2005.
- [4] S. Bayram, H. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on CFA interpolation," *ICIP*, 2005.
- [5] C. Chen and M. C. Stamm, "Camera model identification framework using an ensemble of demosaicing features," in WIFS, 2015, pp. 1–6.
- [6] M. Kirchner, "Efficient estimation of CFA pattern configuration in digital camera images," in *Media Forensics and Security*, 2010.
- [7] C.-H. Choi, J.-H. Choi, and H.-K. Lee, "CFA pattern identification of digital cameras using intermediate value counting," in *MM&Sec*, 2011, pp. 21–26.
- [8] C.-H. Choi, H.-Y. Lee, and H.-K. Lee, "Estimation of color modification in digital images by cfa pattern change," *Forensic Science International*, vol. 226, no. 1, pp. 94 – 105, 2013.
- [9] Q. Bammey, R. Grompone von Gioi, and J. Morel, "Automatic detection of demosaicing image artifacts and its use in tampering detection," in *MIPR*, 2018, pp. 424–429.
- [10] A. Desolneux, L. Moisan, and J.-M. Morel, "Meaningful alignments," *IJCV*, vol. 40, no. 1, pp. 7–23, 2000.
- [11] —, From Gestalt Theory to Image Analysis: A Probabilistic Approach. Springer, 2008.
- [12] A. Gordon, G. Glazko, X. Qiu, and A. Yakovlev, "Control of the mean number of false discoveries, Bonferroni and stability of multiple testing," *Ann. Appl. Stat.*, vol. 1, no. 1, pp. 179–190, 2007.
- [13] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The annals of mathematical statistics*, pp. 50–60, 1947.
- [14] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, "An evaluation of popular copy-move forgery detection approaches," *Trans. Inf. Forensic Secur.*, vol. 7, no. 6, pp. 1841–1854, Dec 2012.
- [15] J. Duran and A. Buades, "A Demosaicking Algorithm with Adaptive Inter-Channel Correlation," *IPOL*, vol. 5, pp. 311–327, 2015.
- [16] P. Getreuer, "Image Demosaicking with Contour Stencils," *IPOL*, vol. 2, pp. 22–34, 2012.