Int J Comput Vis (2011) 92: 71–91
DOI 10.1007/s11263-010-0403-1

Digital Image Forgery Detection Based on Lens and Sensor Aberration

Ido Yerushalmy · Hagit Hel-Or

Received: 31 August 2009 / Accepted: 21 October 2010 / Published online: 6 November 2010

Abstract A new approach to detecting forgery in digital photographs is suggested. The method does not necessitate adding data to the image (such as a *Digital Watermark*) nor require other images for comparison or training. The fundamental assumption in the presented approach is the notion that image features arising from the image acquisition process itself or due to the physical structure and characteristics of digital cameras, are inherent proof of authenticity and they are sensitive to image manipulation as well as being difficult to forge synthetically. Typically, such features do not affect image content nor quality and are often invisible to the inexperienced eye. The approach presented in this work is based on the effects introduced in the acquired image by the optical and sensing systems of the camera. Specifically, it exploits image artifacts that are due to *chromatic aberrations* as indicators for evaluating image authenticity.

Keywords Image forgery · Camera based forgery detection · Chromatic aberration · Lens artifacts · Purple blooming · Lateral chromatic aberration

1 Introduction

Digital photography produces images that can be easily edited using simple and widely accessible software. Along with images acquired by digital cameras, the field of Computer Graphics enables the generation of highly realistic images. This progress allows not only the enhancement of photographs or generation of realistic animations, but unfortunately also the creation of forged images. This has reduced the reliability of digital images and escalated copyrights issues—leading to the necessity of image authentication.

Digital Watermarks (Wolfgang and Delp 1996) enable verification of authenticity of an image. However, it requires special processing at the time of acquisition as well as when image authenticity is tested. This work suggests a new method of testing whether an image has been forged or tampered. The method does not necessitate adding data to the image (such as a *Digital Watermark*) nor require other images for comparison. The fundamental assumption in the presented approach is the notion that image features arising from the image acquisition process itself or due to the physical structure and characteristics of digital cameras, are inherent indicators of authenticity and they are sensitive to image manipulation as well as being difficult to forge synthetically. Typically, such features do not affect image content nor quality and are often invisible to the inexperienced eye. The approach presented in this work is based on the effects introduced in the acquired image by the optical and sensing systems of the camera. Specifically, it exploits image artifacts that are due to *chromatic aberrations*, as indicators for evaluating image authenticity. Chromatic aberrations in digital images arise from physical and optical sources during the image acquisition process. A plethora of aberrations intermix to produce chromatic (and spatial) artifacts that are barely perceived and often disregarded by the naive observer (see Sect. 3). Yet these effects may serve as authenticity indicators as they are subtle, difficult to reproduce artificially and abundantly common in digital images ranging from simple every-day cameras to high-end cameras and professional

I. Yerushalmy · H. Hel-Or (✉)
Department of Computer Science, University of Haifa, Mount Carmel, Haifa, Israel
e-mail: hagit@cs.haifa.ac.il

I. Yerushalmy
e-mail: ido.yerushalmy@gmail.com

lenses. The interactions and intermixing of the aberration effects within the digital image undermine the use of any one model to describe the aberration effects within the image. On the other hand the aberration effects are characterized by a widely dispersed distribution of local effects throughout the image and are supported by easily detected and measurable chromatic features that can be evaluated locally in the image. The approach suggested in this paper, attempts to robustly combine the numerous cues of chromatic aberration detected throughout an image in order to evaluate authenticity of an image.

2 Previous Work

The field of image authenticity has evolved in four different directions, each relying on different assumptions and techniques.

2.1 Embedding Additional Data

One of the most common techniques to ensure the authenticity of images is *Watermarking*. This method assimilates an invisible structure (the watermark) into a digital image. The structure and its contents assure, with high probability, that image content has not been affected since the watermark was added, allowing to detect forged images and to assure copyrights are not violated. Digital watermarking involves explicitly adding the invisible structure when the image is to be sealed, and applying a decryption algorithm when the authenticity of the image is tested. Such a method is clearly distinct from the algorithm suggested here, and so will not be elaborated on beyond this point. Further details on work following this course of progress can be found in Wolfgang and Delp (1996) and many others.

2.2 Physical Constraints Based Methods

Simple forgery detection, based on naïve physical constraints on the scene, may involve detection of inconsistencies of lighting direction (shadows) (Johnson and Farid 2005), color balancing, intelligent reasoning etc. (e.g. Fig. 1). These inconsistencies, however, are easily avoided even by a novice forger while usually hard to detect automatically using software. Computer Vision based algorithms that detect lighting inconsistencies, for example, must introduce strong assumptions on the scene (e.g. that all surfaces of interest are Lambertian). Intelligent reasoning typically requires segmenting most of the image into meaningful objects and classifying them correctly as a prior to any further analysis. Such a task is highly complex and demanding and is currently feasible only for a limited set of objects (e.g. faces, cars, airplanes etc...) (Schroff et al. 2008).

Fig. 1 Intelligent reasoning can discover the forgery (image generated by funadium, at: http://www.flickr.com/photos/funadium/1331284420 and distributed under the Creative Commons license: http://creativecommons.org/licenses/by-nc-sa/2.0/deed.en)

2.3 Statistics Based Methods

Statistics based methods are a common way to detect forged images (Wolfgang and Delp 1996; Keren 2002; Cutzu et al. 2003; Szummer and Picard 1998; Lyu and Farid 2005; Lyu et al. 2004). Several algorithms have been specifically developed to address the authenticity problem (Wolfgang and Delp 1996; Lyu and Farid 2005; Lyu et al. 2004), others are classifiers that aid in pointing out images that are not what they claim to be (e.g. an indoor image with outdoor characteristics Szummer and Picard 1998). The statistics based methods rely on special characteristics that appear in most natural images acquired by a camera. The basic approach in these studies is to extract feature vectors from the image, then, use either simple measures (such as the Euclidean distance) or more advanced machine learning techniques such as *Support Vector Machines* (SVM) (Vapnik 1995), to distinguish between genuine and forged images. In the case of machine learning, a training database of authentic and forged images is assumed. These methods enable, for example, the discrimination between authentic images acquired by a camera and computer graphics generated images. This is based on the fact that the latter is usually smoother in large patches of the image and on the other hand sharper at the edges, introducing a stronger magnitude of high spatial frequency in those regions (Lyu and Farid 2005). The distinguishing feature vectors use first and second order statistics such as mean, variance, skew and kurtosis which are calculated for every sub-band of the wavelet decomposition of the image. The value and characteristics of these features are expected to capture the uniqueness of each class of images. Other statistics based studies attempt to distinguish between images taken indoors and outdoors (Szummer and Picard 1998), paintings and photographs (Cutzu et al. 2003) and even classify paintings according to the artist (Keren 2002).

Obviously, the main drawback of any statistics based method is that it is heavily dependant upon the training set used, the features extracted and the classifying algorithm. These techniques are often limited in that they perform well only on test images that are very similar to the training set or, if tuned differently, have many false positives. Furthermore, a well trained and knowledgeable forger may directly incorporate these statistically expected characteristics of images into the forgery process (e.g. smoothing edges of forged regions).

2.4 Detection Without Additional Data

The most challenging of the forgery detection approaches is that which relies on a single image where the source is unknown. Although this would be the most common case in which forgery detection is required, it is the most difficult, as it can not rely on watermarks, nor on statistical priors from a training set. The study presented in this work falls in this category of forgery detection. A frequent form of forgery involves replacing parts of an image with a copy of another part of the same image. This is performed in an attempt to occlude un-wanted regions or to intensify a phenomenon. Figure 2 presents an example.

The forgery detection method presented by Fridrich et al. (2003) deals specifically with this kind of forgery. In their method, the entire image is segmented into small, non-overlapping regions, each reshaped into a row vector that is then treated as a number. The regions are ordered according to their numbers using any known sorting algorithm. The method then searches for consecutive regions that have the same (or very similar) values. If such exist it indicates, with high probability, that one region was copied from the other. The method, however, generates false alarms for large uniform areas such as sky regions. Moreover, it does not distinguish between the original and copied region. The method presented in this work addresses, among others, this type of forgery and is capable of distinguishing the original from the copied regions as will be further detailed in Sect. 4.

In the study presented in this paper, it is asserted that in order to truly detect forgery, including those created by experienced forgers, one must rely on inherent characteristics of an image that originate in the image acquisition process itself. In the case of this study the effects of the digital acquisition process, along with analog optical effects are taken into account (as will be elaborated in Sect. 4). Very few studies have taken this approach.

One such method is based on the fact that almost all digital sensors used in contemporary cameras use a *Color Filter Array* (CFA) (Wolfgang and Delp 1996). Thus, during acquisition, every pixel receives only a single color-channel value (red, green or blue). To produce the final image, the raw data undergoes an interpolation process. There are several methods to perform the interpolation, and the forgery detection algorithm suggested by Popescu and Farid (2005) attempts to detect which method was used, based on data from the entire image. Assuming that only a relatively small part of the image was distorted, this algorithm can find the values in each of the color channels that do not agree with the general interpolation scheme and mark those regions as suspected forgery.

A different method is based on the most common image compression technique—*JPEG* (Wang and Farid 2006). The JPEG image format involves loss of information as a result of, among others, quantization of the frequency domain (DCT) coefficients. When the image is decompressed, the quantified frequency coefficients are restored. If the decompressed image is now edited and re-compressed in JPEG format, using different quantization steps (e.g. a different JPEG compressor was used, or the image was edited or cropped), a specific repetitive spatial pattern may emerge. This in turn can indicate a forged image.

In another study, more directly related to the approach suggested in this paper, Johnson and Farid (2006) consider chromatic aberration effects due to the optical system of the

Fig. 2 Forged image published by Reuters. Smoke from the original (*left*) image was duplicated to create a more dramatic view (*right*). Reuters had to publicly apologize for the forgery (image was taken from the Jerusalem Post newspaper published on August 7th, 2006. Additional review can be found at the National Press Photographers Association (NPPA) site http://www.nppa.org/news_and_events/news/2006/08/reuters.html)

camera. However their approach is restricted to a specific form of aberration—*Lateral Chromatic Aberration* (LCA). This allows the authors to model the chromatic aberration effects in the image as a spatial scaling of the blue (short wavelength) channel of the image with respect to the green (middle wavelength) channel (see explanation in Sect. 3). A similar model is assumed for the red (long wavelength) channel. A brute force algorithm is thus used to find the center and magnitude of scaling between the 2 image channels by maximizing the correlation between the green channel and scaled blue channel as a function of the center location and magnitude of the scaling evaluated over the entire image. In the second stage of the process the image is segmented into non-overlapping regions, and the channel scaling parameters (locus and magnitude) are determined independently in each region. Regions whose parameters differ significantly from those found in the global search are marked as suspected forgery. Thus, this method allows detecting forgeries, based on significant spatial displacements between color channels within image regions. However, in practice, the Lateral Chromatic Aberration is typically confounded with a plethora of additional aberrations such as *Axial Chromatic Aberration*, *Purple Blooming Aberration* and others (see Sect. 3) that affect the image chromatically. Thus the assumptions on which the authors rely on often do not hold in images resulting in incorrect behavior (see Fig. 6).

3 Lens Chromatic Aberration

Optical systems, even in contemporary cameras, suffer from imperfections, causing a variety of aberrations in the final image including chromatic aberrations, spatial blurring and geometric distortions. Most artifacts are un-noticeable to the human eye but nevertheless can still be detected by computers. In this study, we focus on chromatic aberrations, though it has been observed (Smith 2007) that spatial and geometric aberrations such as Spherical aberration, Coma aberration and Astigmatism aberration (Ray 2002; Jenkins and White 1976; Pedrotti et al. 2006) have a secondary affect on chromatic aberrations.

A plethora of chromatic aberrations intermix to produce the chromatic artifacts found in digital images (an excellent review can be found in van Walree 2009, see also Smith 2007; Ray 2002; Pedrotti et al. 2006). Both lens and camera sensors are considered as the source of these aberrations. Aberrations due to the camera optics (lens) are traced back to Snell's law that explains the refraction of light on the boundary between two media: $n * \sin(\theta) = n_f * \sin(\theta_f)$ where θ is the angle of incidence, θ_f the angle of refraction and n and n_f are the refractive indices of the two media. Glass has a different refractive index for every spectral wavelength, causing a single polychromatic ray of light

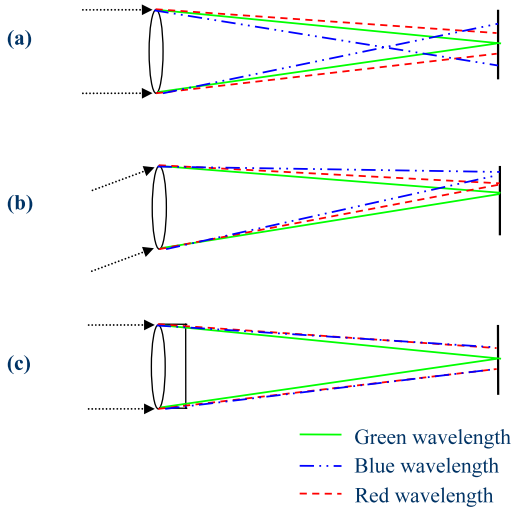

Fig. 3 Due to differences in refraction indices, light of different wavelengths passing through a lens, do not converge on the image plane. (**a**) Axial Aberration occurs when different wavelengths converge at different depths from the lens. (**b**) Lateral Aberration occurs when different wavelengths converge at different points on the image plane. (**c**) Achromatic Doublet—a negative focal length lens can be added to focus the *blue* and *red* wavelengths, however residual aberration remains

that enters a camera lens to focus at different points on the acquisition plane (depending on wavelength) as shown in Fig. 3. *Axial Chromatic Aberration* (also known as *Longitudinal Aberration*) occurs when light impinging on the lens parallel to the optic axis, refracts so that different wavelengths focus at different focal planes as shown in Fig. 3a. This has an effect of differentiating blur between the wavelength channels often referred to as a difference in the size of the *circle of confusion* between the different wavelength images. *Lateral Chromatic Aberration* (LCA) (also known as *Transverse Aberration*) is caused by the fact that long (red), middle (green) and short (blue) wavelengths are not focused by the lens at the same point in the image plane when the source light is off the optical axis as shown in Fig. 3b. The visual effect of this aberration in an image is shown in Fig. 4. The LCA can be viewed as having an enlarging effect on the blue (short) and red (long) wavelength images compared to the green (middle) wavelength image.

Technically, LCA can be viewed as an expansion-contraction effect of the 3 image channels (R, G, B) about the image center. High end lenses attempt to reduce the magnitude of the chromatic aberration effect by combining positive and negative focal length lenses to produce an Achromatic Doublet (see Fig. 3c) which unites two wavelengths at a common focus point. However other wavelengths are still out of focus producing a residual aberration called Secondary Spectrum. Thus, even expensive lenses still suffer from chromatic aberrations. There is a discussion as to which of the chromatic aberrations is the more dominant (see van Walree 2009 and

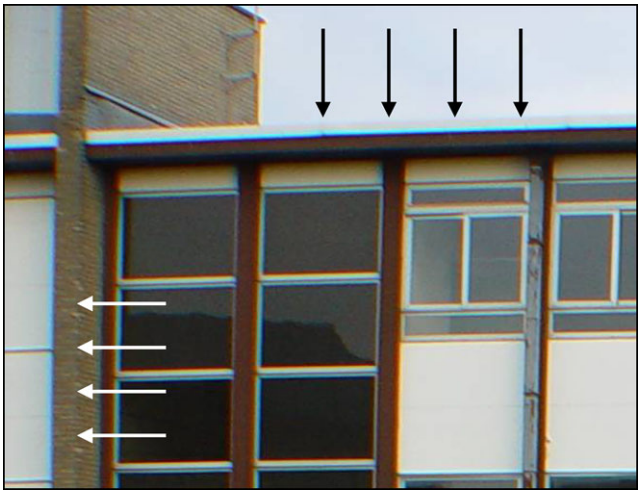

Fig. 4 (Color online) Enlargement of the *top left* section of an image acquired with a Sony 707 using a Sakar wide-angle converter. The LCA effect can be seen as the *bright blue* stripes along the edges of the building, marked by *arrows* (contrast was added to amplify the aberration) (image is courtesy of Tom Niemann from ePaperPress.com: http://www.epaperpress.com/ptlens)

citations within) however it is accepted that they coexist and interact in their effects in the image.

In addition to the chromatic aberrations due to imperfections of the lens, the camera's *Charge Coupled Device* (CCD) sensors interact with the lens and produce additional aberrations including the *Purple Fringing Aberration* (PFA), (Fig. 5) which appears in the form of a blue-purple halo near the edges of objects in the image. The source of PFA is controversial (see Parr 2006) though it is typically attributed to the following:

1. PFA is characterized by blurriness at high contrast edges. This may be due to CCD sensors, which tend to suffer from electron overflow to adjacent photodiode cells when an extremely high rate of photons per second hits the sensor (Ochi et al. 1997). Thus, bright areas in the scene may result in a blur near the edges.
2. In addition to visible light, CCD sensors are often sensitive to infrared wavelength as well. To overcome this problem, the sensor is coated with an appropriate filter which, however, does not completely block these wavelengths. Thus some energy may reach the sensor and cause edge-blur similar to that produced due to LCA. In fact, the Secondary Spectrum aberration mentioned above is most significant in the ultra-violet and infrared wavelengths, which (similar to LCA) focus slightly beyond the focal plane (Rudolf 1992).
3. Each CCD sensor cell consists of a photon sensitive surface and an electronic circuit. To increase light absorption, each cell is covered with a micro lens (Daly 2001). Light impinging on the edges of the sensor cell may refract on the micro lens and affect neighboring cells.

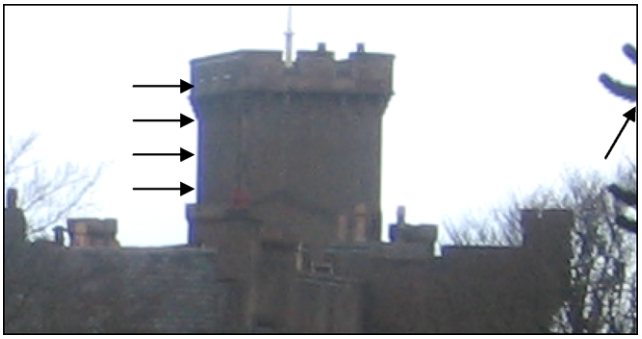

Fig. 5 (Color online) An enlarged part of an image acquired by a Canon Power-Shot A520 camera. Purple Blooming is noticeable at the edges marked by the *arrows*

PFA expresses several unique characteristics that can not be explained by any one aberration, specifically the PFA encompasses LCA yet differs from it. The most significant difference is that while both types of aberrations frequently appear near edges in the image, PFA does not exhibit an expansion-contraction transformation as does the LCA. In the latter, the blue color channel of the image is magnified about the center, with respect to the green channel. In PFA, the blue-purple halo appears on the distal side of bright objects (or on the proximal side of dark objects) relative to the image center, while the opposite side remains almost unchanged (see Fig. 6). It should be noted that the purple-blue halo is more prominent than the blue halo of the LCA due to the additional factors described above, which intensify the aberration. Figure 7 shows a schematic example of LCA versus PFA.

Referring to the possible sources for PFA (mentioned above) one can observe that as the contrast between objects increases, the PFA increases as well. This can be explained by the increase in overflow to neighboring cells (bullets 1, 3 above) due to the large amount of light that reaches the sensor. Since this aberration causes the values of neighboring pixels to change, it is more visible across edges where pixels have a large difference in values. The LCA (which affects the PFA as well) is also more noticeable at high contrast edges, since the expanded long/short wavelengths images (blue and red) are displaced and visually affect neighboring pixels, across the edge. When edge contrast is low this effect is less visible.

Furthermore, the PFA effect is more prominent at image periphery, specifically in wide angle shots (Born and Wolf 1999). This is due to increased diffraction at the lens edge and a shallower angle of contact of the light rays with the lens or micro lens as well as increased imperfection at the lens with deviation from the center of focus (see also Fig. 19).

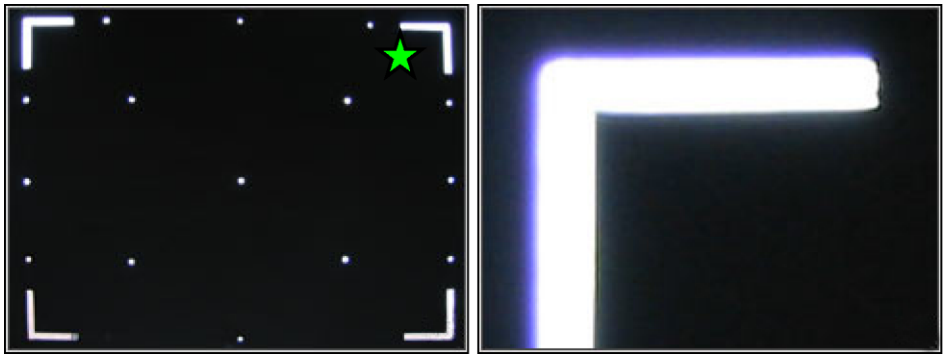

Fig. 6 (Color online) An image of bright regions on a dark background acquired using a Canon S330 digital CCD camera (*left*) (DpReview website, http://www.dpreview.com/reviews/canondigitalixus330/page11.asp). Enlargement (*right*) of *the top left* corner clearly shows the *blue-purple* halo on the outer side of the bright region. *The green star* depicts the incorrectly calculated image center according to the FD algorithm presented in Johnson and Farid (2006)

4 Forgery Detection Based on Local Indicators of Chromatic Aberration (PFA)

As described above, one of the most important characteristics of the PFA is that the purple-blue halo is directional, namely it appears on the distal side of bright objects (or on the proximal side of dark objects) relative to the image center (see Fig. 6). In the suggested approach this characteristic is exploited to locate the image center. Furthermore, regions of the image in which PFA does not point to the common image center can be marked as suspected forged areas. In Fig. 8 the arrows indicate the "direction" of the purple halo. The approach is detailed in the next two sections.

Two additional characteristics of the PFA aberration are exploited in the proposed approach:

1. PFA increases in strength with distance from the image center. Figure 9 (top) shows an image with edges having the same intensity contrast. As the distance from the image center (located to the right of the region in the figure) increases, so does the intensity of the aberration (corresponding to size of the diamond head).
2. Due to the factors affecting PFA, it becomes more acute in regions of higher intensity change over a short distance. This implies a stronger effect of purple blooming along edges with increased difference in intensity (contrast), for example white object on a black background. Figure 9 (bottom) shows three patches, equally distant from the image center, with varying degrees of contrast relative to the background. As the contrast increases, the aberration becomes stronger.

Taking both characteristics into account, a reliability measure for every PFA event can be introduced. The presented approach attempts to detect all PFA events in the image and determine their directions together with such a measure of reliability. The combined information from all detected PFA directions is used to determine the image center. If a region contains inconsistent PFA directions, it can be concluded that either local image noise has affected the results, or that the region is forged. The measure of reliability assists in overcoming such noise, allowing the algo-

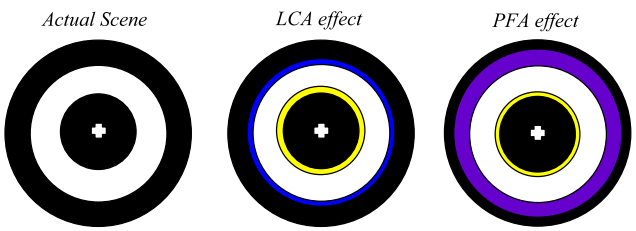

Fig. 7 (Color online) Schematic diagram that depicts the difference between LCA and PFA (image center is depicted by '+'). While LCA is a pure expansion-contraction effect of the Blue vs Green channels, PFA is characterized by a *blue-purple* halo on the distal side of bright objects (sometimes also accompanied by a minor *yellow* tint on the opposite side, due to the LCA effect)

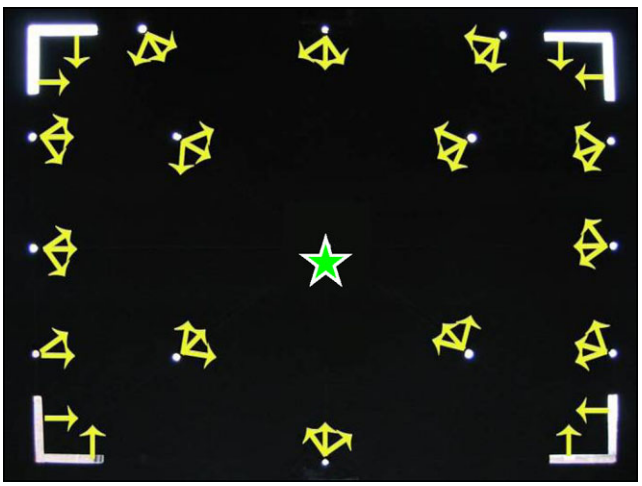

Fig. 8 Image from Fig. 6, with *arrows* indicating the "direction" of the PFA. The image center can then be found (depicted by the *star*)

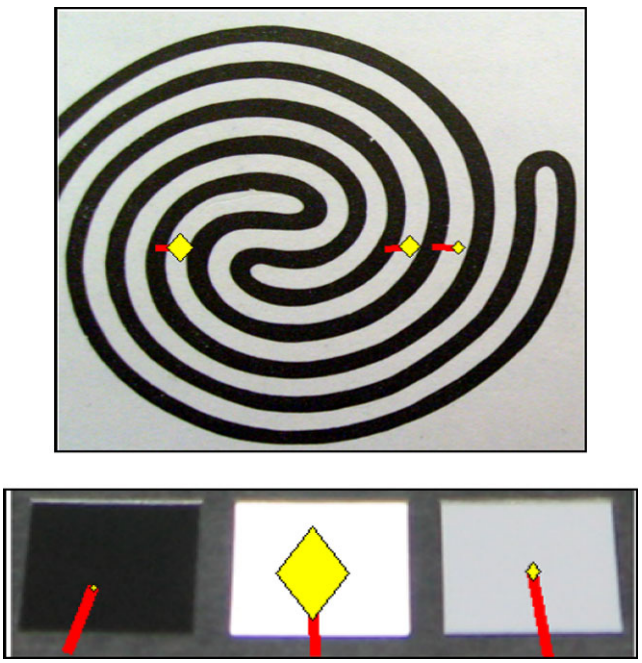

Fig. 9 *Top*: Aberration strength increases with distance from image center. An enlarged region of an image is shown (image center is located to the right of the region). Three edges with equal color contrast were tested for PFA using the suggested algorithm and marked with a *diamond* headed arrow whose size corresponds to the strength of aberration. The *arrows* point in the direction of the image center. The edge most distant from the center was calculated to have more than 4 times the aberration of the most central edge. *Bottom*: Aberration strength increases with contrast. An enlarged region of an image is shown (image center is above the region) with three patches equally distant from the image center. The *diamond-headed arrows* point in the direction of the image center. The size of the diamond corresponds to the strength of aberration

rithm to perform well even on JPEG compressed images (see Sect. 6).

5 Algorithm

Using the PFA properties mentioned above, one can locate the presumed image center. Regions that do not point to the geometric center of the image suggest that they are not original and have been tampered or completely implanted there after the actual scene was acquired by the camera. Figure 10 shows an example of an authentic digitally acquired image, with image center correctly detected according to the PFA.

To conclude whether a given image is authentic or forged, the following algorithm is implemented:

1. Identify edges with PFA.
2. Determine PFA direction for each detected PFA event.
3. Assign a reliability measure according to aberration strength, edge contrast and distance from the geometric center of the image.

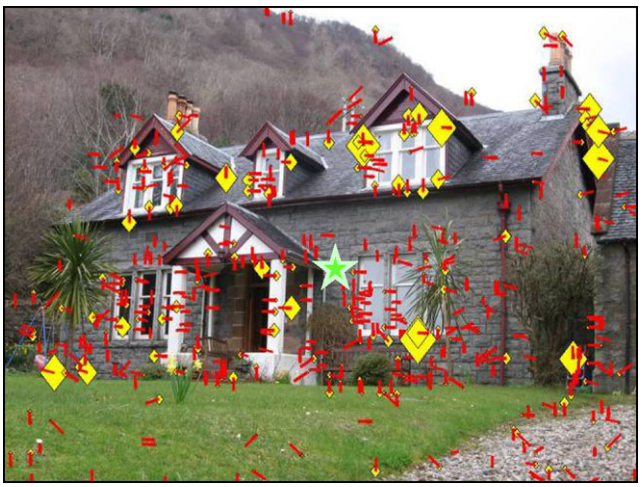

Fig. 10 An authentic image acquired by a Canon A520 digital camera. *Yellow diamond* headed *arrows* indicate a selection of PFA events detected by the proposed method. Diamond size corresponds to strength of aberration. *The star* depicts the detected image center

4. From the collection of PFA directions robustly determine the center of the image (x_0, y_0).
5. Re-evaluate the PFA directions to determine regions in which these directions are inconsistent with the evaluated image center. These regions are marked as suspected tampered regions in the image.

Details of these steps follow.

5.1 Identifying PFA Events

PFA can be found along edges, especially where high contrast exists. To identify the aberration indicators, all significant edges in the image are first detected. This stage can be performed using any known algorithm (e.g. Canny edge detector, Gonzalez and Woods 2002). Every edge pixel is evaluated to determine if it is a viable PFA event. For each edge pixel, the transition of color across the edge is analyzed in the xyY color space (Wyszecki and Styles 1982) (Fig. 11).

It is assumed that in natural images the change in color across the boundary between different color regions is linear in chromaticity. Thus, a pixel sequence across a boundary should contain a linearly varying mixture of the two bordering colors, (perhaps with a change of luminosity). In the xyY color space, this assumption translates to a linear transition between the two color points in the chromaticity $(x–y)$ plane (see Fig. 12), and perhaps a change in the luminance (Y) parameter. In the case of PFA this assumption is violated, since a blue-purple hue affects the edge. This will cause the transition in the chromaticity plane to behave non-linearly with a bias towards the blue-purple region of the plane.[1] Figure 12 shows an example of the transition in chro-

[1] Note that the non-linearity in color values introduced by the Gamma-factor of the camera is insignificant compared to the PFA effect.

maticity plane across a non-aberrated edge (Fig. 12 left) and across an edge with PFA (Fig. 12 right). Pixel values along a segment perpendicular to the edge were uniformly sampled (their chromaticity marked as * in the plots). The chromaticity of the two bordering color regions is marked with a red square and a black dot in the chromaticity plane, corresponding to the start and end points of the sampled pixel sequence, respectively. The non aberrated edge (left) displays a linear chromaticity transition, while the PFA edge (right) displays a deviation towards the blue-purple region.

A PFA event is determined at an edge pixel if the chromaticity values across the edge pixel deviate significantly towards the blue-purple chroma. This is evaluated by estimating the PFA strength as described below. It should be noted that due to the presence of LCA aberration together with PFA (see Sect. 4), some images also display the complement of a blue-purple halo, namely yellow tinted edges. These effects will appear at objects' edges on the side opposing the purple halo (i.e. on the proximal side of bright objects). For these edges a similar non-linear transition will be observed in the chromaticity plane with deviation towards the yellow region. However, this deviation is usually much weaker than the blue-purple deviation due to PFA, which is intensified by CCD associated aberrations (see Sect. 3).

5.2 Determining PFA Direction

As explained in Sect. 4, the PFA effect is displayed as a blue-purple halo on the distal side of bright objects and on the proximal side of dark objects. Thus, the intensity gradient across the edge implies the direction towards the image center. We define the PFA direction as a unit vector $\bar{N}$ perpendicular to the edge in the direction corresponding to increasing intensity across the edge. In case of aberrations that produce yellow artifacts at the edge (corresponding to deviation towards yellow tones in the chromaticity plane) the PFA direction will be reversed.

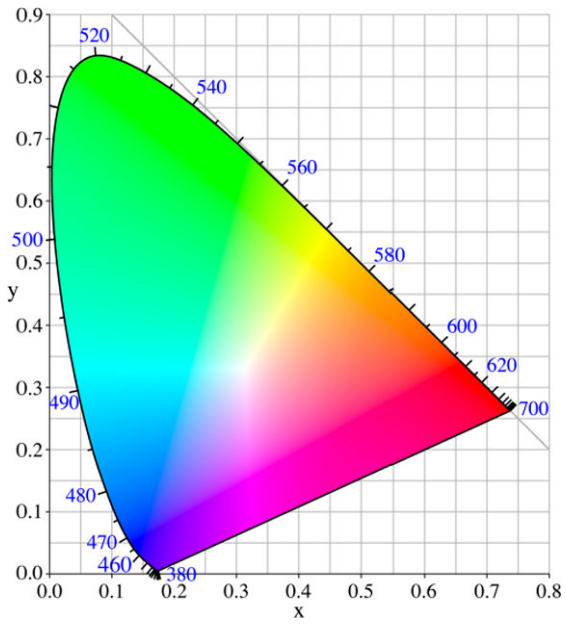

Fig. 11 (Color online) Chromaticity plane of the xyY color space. The *purple-blue region* is located at 450 nm. Image from Wikipedia, the free encyclopedia, http://en.wikipedia.org/wiki/Xyz_color_space

5.3 Determining PFA Strength

The strength of a PFA event is dependant on the magnitude and direction of the chromatic deviation in the chromaticity plane (Fig. 12). To determine PFA strength at an edge, the following is specified: given the sequence of pixel chromaticities across an edge $\{c_a = c_1, c_2, \ldots, c_n = c_b\}$ in xy coordinates, the expected linear transition between endpoint colors c_a, c_b is represented as a segment in the chromaticity plane (gray segments in Fig. 12). In an ideal world, where no blur is introduced in the optical system, we may expect the chromaticity of the sequence to be clustered tightly around c_a and c_b, however, in reality this does not occur in acquired images. Furthermore, the transition rate between chromaticity values of the two endpoints can not be assumed (and is often not uniform). Since no prior information about the expected location of the analyzed points can be assumed, except for the fact that they should appear along the expected linear transition segment, a point-to-segment distance is calculated for each pixel chromaticity value and the point with maximum distance is determined (yellow diamond in Fig. 12).

The vector in the chromaticity plane between the expected transition segment (gray segment) and the most distant point of the sequence is projected onto the blue-purple direction, determined as the unit vector from mid segment to the blue-purple wavelength. This wavelength is set to be 450 nm ($x = 0.2$, $y = 0.1$ in the chromaticity plane, see Fig. 11). The magnitudes of the projections along the blue-purple direction are normalized to the [0..1] range:

$$\delta = \left(\frac{s - s_{\min}}{s_{\max} - s_{\min}} \right) \tag{1}$$

where s is the projected value, $s_{\max}$ ($s_{\min}$) is the maximum (minimum) value found in the image. The normalized projection value, δ, is defined as the PFA strength. If aberration is present at the edge then the PFA strength is expected to be large.

5.4 Assigning a Reliability Measure

The reliability measure quantifies the consistency of the calculated PFA strength with the expected values. As described above, the PFA is expected to increase in strength with distance from image center as well as with increase in edge contrast. These characteristics are used to define a measure

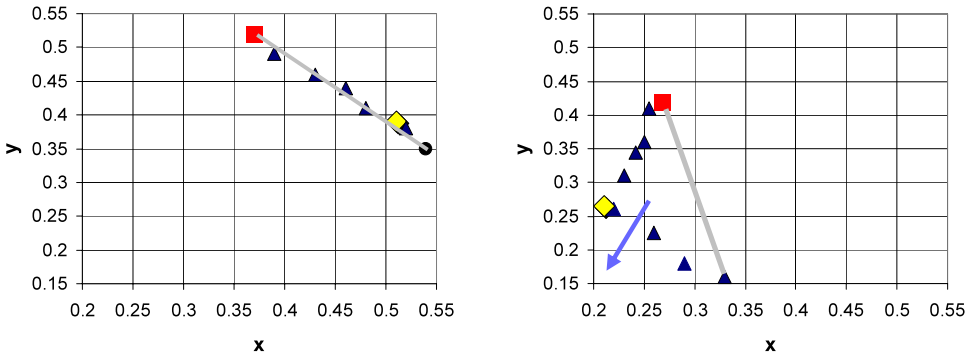

Fig. 12 (Color online) PFA analysis in the xy-chromaticity plane. The chromaticities of image pixel colors across an edge boundary are shown for a non-aberrated edge (*left*) and for an edge with Purple Blooming (*right*). The start and end points of the pixel sequence are marked with a *red square* and a *black dot*. The *gray line* is the expected linear transition between the two end points. *The left plot* shows an almost linear transition between colors on both sides of the edge. *The right plot* shows a deviation towards the blue-purple values (*blue arrow* direction) of the chromaticity plane. The *yellow diamond* depicts the furthest point from the expected segment

of reliability for each detected PFA region. If, for example, a strong aberration was detected along an edge with a small contrast difference, it may be suspected that image noise has affected the calculations; accordingly the reliability should be low and it should significantly reduce the effect of the aberration on the center calculation process (see (3)). Edge contrast t associated with a PFA event is defined as the absolute difference in intensity across the edge and the distance d of the PFA event is the linear distance to the image center. The reliability of a PFA region is then defined as:

$$\rho = \left(\frac{t - t_{\min}}{t_{\max} - t_{\min}}\right) * \left(\frac{d - d_{\min}}{d_{\max} - d_{\min}}\right) \tag{2}$$

where $t_{\max}$ ($t_{\min}$) is the maximum (minimum) edge contrast found in the image and $d_{\max}$ ($d_{\min}$) is the maximal (minimal) distance to the image center.

5.5 Calculating the Location of the Image Center

Using the collection of PFA events, the location of the image center can now be estimated. The values for each PFA event, required for this task are: the *direction* ($\vec{N}$), the *strength* (δ) and the *reliability* (ρ) of the PFA event as described above.

The aberration, and accordingly PFA direction, is always perpendicular to the analyzed edge. Thus it does not necessarily indicating the direction towards the image center. This problem, called the "Aperture Problem", is well known in Computer Vision, where local movement of objects can be obtained only in the direction perpendicular to their boundary edge, creating a "normal flow" map (Jahne 2005). In Computer Vision the "scene flow" (the actual movement of the objects), rather than "normal flow" is often required. To obtain a good estimate, an "optical flow" map is generated, which assigns a movement direction to every pixel in the image (Horn and Schunck 1981), resulting in a smooth field consistent with the normal flow data. This process is usually computationally demanding and requires dense data in the "normal flow" map to produce an accurate result. In the context of this work, generating an "optical flow" map can be useful in the process of locating the image center, however it is computationally difficult and overkill since in practice only the location of the image center is sought, and interpolated flow information calculated per pixel is not required. Thus, a different method was chosen, which is based on Focus Of Expansion (FOE) detection (Negahdaripour and Horn 1989). This method assumes that the underlying flow field displays a centralized flow (reflecting a backwards motion of the camera within the scene). The method attempts to locate the FOE point based on the PFA directions associated with the PFA events, thus, determining a central image point which is consistent with the centralized flow. Similar to the "optical flow" approach, the chosen technique relies on the fact that the "normal flow" map is relatively uniform across the image. The main differences between "optical flow" detection and FOE calculation is that the former assigns an evaluated pointing direction to every pixel in the image (usually based on a smoothness assumption), while the latter algorithm does not. In addition, the FOE method uses a strong assumption regarding the type of movement in the image (expansion relative to a single center), while an "optical flow" process has no prior assumptions, causing it to perform more calculations and perhaps even generate a false movement model. Thus, we can benefit from a faster algorithm while obtaining accurate results. The method suggested in Negahdaripour and Horn (1989) for FOE detection is adopted here.

Determining the FOE based on the PFA "flow field" takes into account both the strength and the reliability of the detected PFA events. Let $X_i = (x_i, y_i)$ be the edge pixel coordinates of the i-th detected PFA event and let $X_0 = (x_0, y_0)$

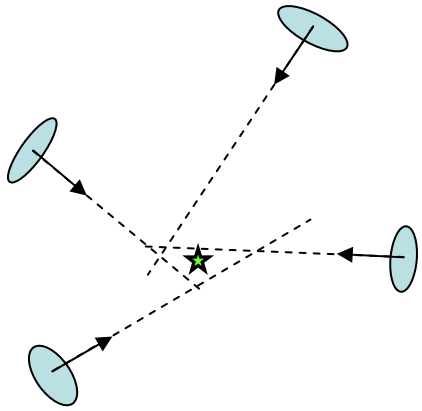

Fig. 13 Finding the point X_0 (marked by *star*) which minimizes the squared perpendicular distance to the lines associated with PFA events (marked by *ellipsoids* with associated *arrows*). See (3)

be the image center to be determined. The point X_i is associated with a line passing through the point in the direction $\vec{N}$. The line is of the form: $a_i x + b_i y + c_i = 0$ with (a, b) normalized to unit vector. The process finds the point X_0 which minimizes the squared perpendicular distance to all the lines, normalized by the PFA strength and the reliability factor (see Fig. 13):

$$X_0 = (x_0, y_0) = \arg\min_{x,y} \sum_i \frac{(a_i x + b_i y + c_i)^2}{(1 - \delta_i \rho_i + \varepsilon)^2} + R_i(x, y) \tag{3}$$

where δ_i is the PFA strength, ρ_i its reliability as described above and $(\varepsilon \to 0)$ is used to avoid singularity. $R_i(x, y)$ is a penalty factor associated with the i-th PFA event that is assigned to location (x, y) if it is inconsistent with the i-th PFA direction (i.e. (x, y) is "behind" the i-th edge relative to the PFA direction):

$$R_i(x, y) = \begin{cases} \alpha K_i, & \vec{N}_i \bullet (X - X_0) < 0 \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where $\bullet$ denotes inner product, K_i is the normalized point-to-segment distance between the i-th edge and the tested center and α is a scale factor. The penalty is added, and not multiplied by the remaining parameters, to avoid a scenario in which the algorithm converges to an image center located "behind" any PFA edge and thus is inconsistent with the PFA effect. In all experiments α was set to 5×10^5. A multi resolution descent search is used to minimize the function and determine the coordinates of the image center.

The above method produces good results, based on the normal flow map. However, there are two concerns that must be taken into account. First, the map may contain outliers due to noise in the image (e.g. due to JPEG compression). To overcome this, a robust-regression method, the Median Outlier Filter (Rousseeuw and Leroy 2003), is used. The data

Fig. 14 (Color online) Calculating image center. The *yellow dots* show convergence of centers calculated before outlier removal. The *green dots* depict centers after outliers have been removed. *Magenta colored dot* (at top right cornel of the image) depicts the final calculated center. Original image from www.worth1000.com

points proximal to the median (70% of the points) are considered valid, while the rest are marked as outliers. In the process of center detection, the weighted distances of the normal-flow vectors relative to the calculated center (based on aberration strength, reliability and penalty), are the input to the filter. The process is run iteratively; outliers are removed at each iteration, until the calculated center location stabilizes.

The second concern is that the reliability measure assigned to every PFA event (2) depends on the distance to the image center. However the image center is unknown and is computed iteratively using (3) and using the outlier removal process discussed above. This issue is specifically significant in cases of cropped images when the actual image center may differ significantly from the geometric center of the image (see below). To overcome this concern, the reliability measure is initially computed based on the geometric center of the image. At each iteration, the reliability measure is updated based on the image center found during the previous iteration. The process iterates until the calculated center location stabilizes. An example is shown in Fig. 14, where the calculated centers are shown initiating from the geometric center and converging to the found image center (magenta). Yellow dots depict the location of the center before outliers have been removed. At the next stage, outliers are

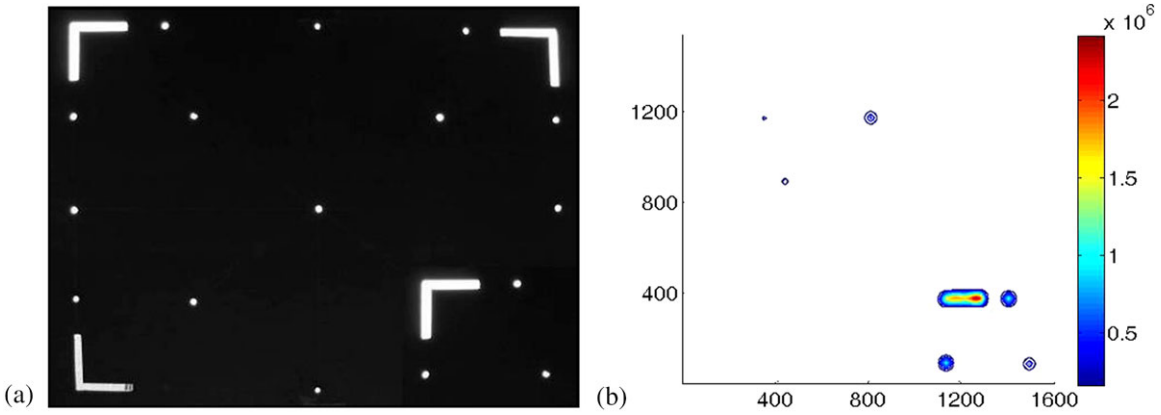

Fig. 15 (**a**) Is a forged version of Fig. 6. (**b**) Is a map of the weighted-distances of PFA events. Note that the map clearly indicates the forged region

excluded (as described above) and the process is performed once again (depicted in green dots). The final center location was found at the right side of the image due to strong PFA pointing rightwards.

5.6 Detecting Traces of Forgery

Since the PFA appears in many edges, the PFA direction map forms a "normal flow" map which is usually sufficiently detailed to include data in both original and forged regions, if such exist. Analysis of the PFA direction map as well as the calculated center, allows the detection of image regions that have been tampered since their acquisition by the camera.

PFA regions are analyzed to determine inconsistencies between their direction and the calculated center. These regions, when removed, will maintain a consistent flow. To identify these regions, the weighted distance from the calculated center of the image (as described in (3)) is calculated independently for each PFA event. The penalty factor (as described in (4)) assures that PFA events pointing to an illegal direction have a much higher weighted distance than others. This implies that such PFA events can be distinguished from others (using a threshold), even if the latter do not point directly to the calculated center due to the aperture problem. A map of these values is formed, from which regions highly suspicious of forgery are detected. An example of such a map is shown in Fig. 15. Note that in the case of forgery involving duplicated image regions, (as described in Sect. 2), only the duplicated region will be marked as suspicious, since it alone will be inconsistent with the geometrical center, allowing the system to distinguish the source region from its copies (see Fig. 17). This type of forgery was described and handled by Fridrich et al. (2003). However, the proposed algorithm is advantageous in that it allows distinguishing the source of the forgery from its copies. Additionally, it is able to detect copy-move forgeries even when the copy is magnified or reduced in size, as can be seen in Fig. 17.

Fig. 16 A genuine image acquired by a Canon 400d. *Star* depicts the calculated center

Another scenario is where an image center is detected at a distance from the geometrical center and yet is accompanied by a consistent PFA direction map. This is a unique constellation that may indicate that the image has been cropped from its original size. According to the evaluated image center, an estimate of the original image size can be deduced by the algorithm (see Fig. 20). This capability, although unable to restore the missing data, allows the user to obtain a sense of the original size of the image and what portion was cropped.

Finally, if the calculated image center was found to be located near the geometric center with strong supporting

Fig. 17 Example of copy-paste forgery detection using the proposed algorithm. The man in the original image (*left*) was duplicated to create a forged image (*right*). *Shaded box* marks the suspicious region detected by the proposed algorithm

PFA events, and there is no significant region in which PFA events are inconsistent with the globally evaluated image center, then the algorithm assumes that there is no forgery.

Additional discussion on the abilities and drawbacks of the algorithm can be found in Sect. 7.

6 Test Results

In this section the performance of the presented algorithm is tested on various types of images, all compressed using standard JPEG format. Images were collected from copyright free online sources and from personal collection. No camera calibration is assumed nor is any knowledge on the camera parameters used to acquire the images. Images included indoor and outdoor images, portraits, scenic views, urban scenery and more. Images were of size ranging between 1536×2048 and 640×480. Primary camera models were Samsung S630, Canon SD200 and Canon Power-Shot A520.

6.1 Examples

Specific examples are shown to demonstrate the capabilities of the proposed algorithm. Figure 16 shows the results of applying the proposed algorithm on a genuine image. The image center was calculated up to 4% deviation of the true center. Another such example is shown in Fig. 10, where the deviation from the true image center was 5%. Figure 17 (left) shows an example of a genuine image, acquired by a Nikon E4600 camera. The image was edited to include a copy of the original person (Fig. 17 right). The algorithm was able to detect the suspicious section, while correctly discriminating between the original and its copy. Figure 18 (as well as Fig. 31) show additional forged images taken from www.worth1000.com with the forged region detected. In Fig. 18, the typewriter and the screen were pasted from 2 different images. The center of the original typewriter image was located at the top-center of the typewriter. Since the screen has very strong PFA events pointing rightwards, the image center was shifted to that direction, causing the typewriter's PFA events to display inconsistencies. Note that no suspicious regions are found around the outer edges of the typewriter, since it was very carefully cropped from its original image, thus removing the traces of PFA. However, inside the typewriter itself, detectable PFA traces are found.

Figure 19 shows the effect of distance from center of the forged region. The figure displays a sequence of 9 images. The images on the diagonal contain one object each, these are the originals. The off diagonal images have been edited so that the objects are placed systematically at different eccentricities in the image. The images are overlaid with markings representing the forged regions and colored according to the strength of deviation from the predicted. As expected, the eccentric regions of the image have the strongest affect.

Figure 20 displays the capabilities of the approach on cropped images. Figure 20 (left) shows an image where 33% of the right part was removed. The location of the calculated center (green star) was accurately determined. The fact that most PFA vectors are consistent with the center, which is not aligned with the geometric center, suggests that the image was cropped. Dashed lines mark the estimated full image frame according to the calculated center. The proposed algorithm was able to restore both size and shape of the uncropped image with a success rate of 93% (see (6)). Figure 20 (right) shows a similar example with image cropped by 53%. The proposed algorithm was able to determine the center of the un-cropped image (located beyond the bound

Fig. 18 *Left*: A forged image published in www.worth1000.com. Portions of the typewriter were detected as forged (marked in a *red-shaded* box), as they did not agree with the calculated center, affected by the pasted screen. *Right*: Center detection for each part of the image separately (marked by a *green star*). The detected center indicates the location of each part in its original image

Fig. 19 (Color online) Effect of eccentricity on forgery detection. Images on the diagonal are authentic, while others are synthetically forged. Forgery traces found are marked by colored boxes—*yellow* for weak forgery traces, darker *red* for strong traces. As objects are further displaced from original location, PFA inconsistency with the calculated center increases, thus intensifying forgery traces

of the image) and was able to restore both size and shape of the un-cropped image with a success rate of 85%. Additional results can be found in Sect. 6.2. In the following subsections more rigorous experiments are described, that systematically evaluate the algorithm.

6.2 Comparison Testing

The study of Johnson and Farid (2006) is closely related to the method proposed in this paper. In their work, forgery is detected based on LCA and relies on the expansion-contraction characteristics associated with this type of chromatic aberration (see Sect. 3). In this section we display comparison results that follow the experiments described above. Each test was run, with the same images as input, using both the suggested algorithm as well as the FD algorithm presented in Johnson and Farid (2006). The same outlier removal method used in the proposed algorithm was applied to the FD algorithm to assure a fair comparison.

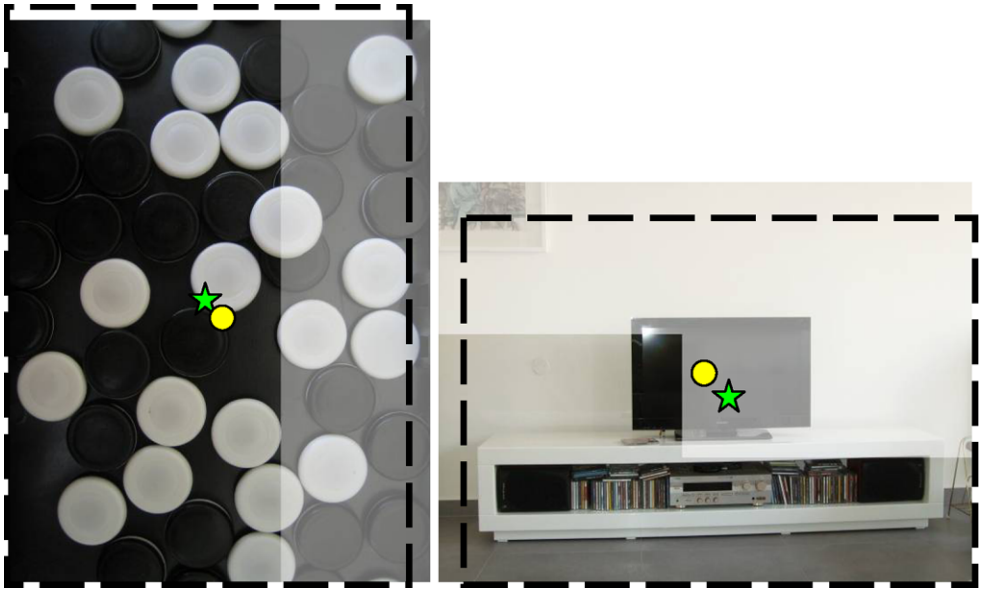

Fig. 20 Detecting cropped images. *Left*: Test image with 33% removed from the right side (shown as de-saturated). The proposed algorithm was able to restore the original size and shape with a 93% success rate, marked with a *dashed line* (see (6)). The *yellow dot* marks the genuine image center and the *green star* the calculated center using the proposed algorithm. *Right*: 53% of the image was cropped, including the center. The algorithm was able to restore original size and shape with accuracy of 85%

Fig. 21 Three samples of authentic images from the test set

Authenticity Testing A set of 45 JPEG compressed images were used to test the basic capability of the algorithms, namely, to correctly identify the geometric center in images for which no editing was performed following acquisition. Examples from the set are shown in Fig. 21. To quantify the level of accuracy of the center detection, an *Angular-Error* measure was used, as defined in Johnson and Farid (2006) so as to be consistent with Johnson and Farid (2006) and allow more effective comparison. The angular error for a given pixel is defined as the angle between the directional vector from the pixel to the geometric center of the image and the direction from the pixel to the center calculated by the algorithm as follows:

Let $X_0 = (x_0, y_0)$ define the coordinates of the center calculated by the algorithm, and let $X_g = (x_g, y_g)$ be coordinates of the geometric center. The angular error, $\theta(x, y)$, for pixel $X = (x, y)$ is defined as:

$$\theta(x, y) = \cos^{-1}\left(\frac{(X_g - X) \bullet (X_0 - X)}{\|(X_g - X)\| \cdot \|(X_0 - X)\|}\right) \tag{5}$$

The *Average Angular Error*, $\bar{\theta}$, for an image is defined as the average over all pixel angular errors.

Figure 22 presents comparison results between FD and the proposed algorithm for the authenticity experiment. The results show a significant improvement when using the latter: 51% of the images were analyzed with an average angular-error of up to 15 degrees using the proposed algorithm, while only 9% of the images reached this accuracy using the FD algorithm. When considering a larger, but still acceptable, value of 25 degrees one can notice that 87% of the images were classified within this limit using the proposed algorithm, while only 33% using the FD algorithm. Possible explanations for this difference are discussed below.

Cropped Images In the following experiment the ability of the algorithm to detect cropped images was tested. This type of forgery arises when a region of the original image contained unwanted data that was removed following acquisition. In this test 120 cropped images were analyzed by the proposed and FD algorithms, in order to test their ability to detect cropping and to restore the original image size. The images were produced automatically based on a set of 30 authentic JPEG compressed images from various cameras. Each authentic image had 33% eliminated (from the top, bottom, right or left portions of the image), producing 4 different cropped images. For each image the algorithms calculated the presumed image center and determined the size and shape of the un-cropped image (see Fig. 20). The size-

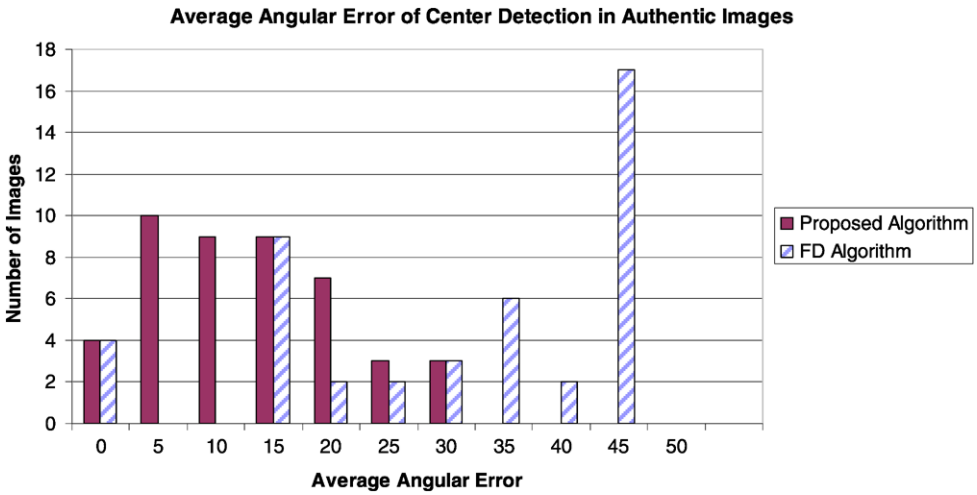

Fig. 22 Performance comparison between proposed and FD algorithms over a set of authentic images. Histogram of the average angular error between detected and correct image center is shown

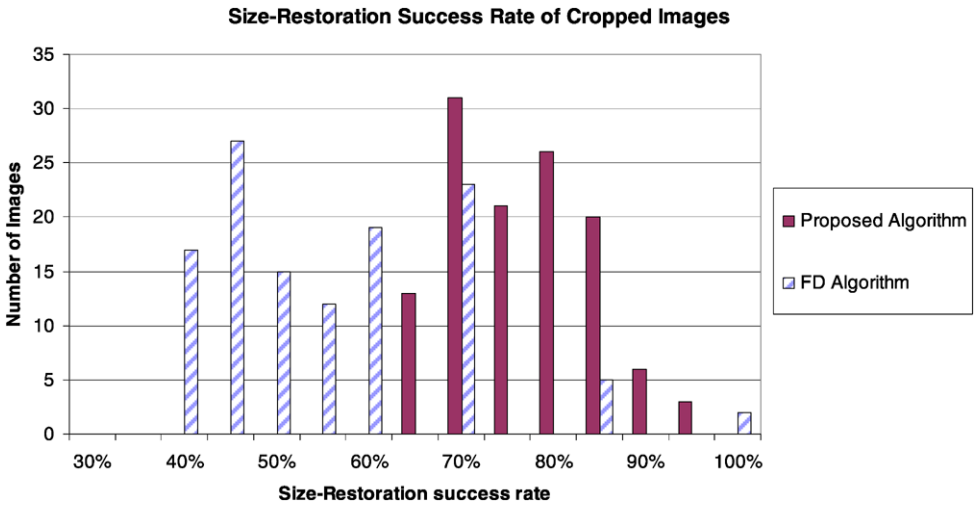

Fig. 23 Performance comparison between the proposed and FD algorithms over a set of cropped images. Histogram of success rate in size-restoration of the cropped images is shown

restoration success rate was evaluated using the average-precision measure (Muller et al. 2001):

$$Coverage = \frac{(area(\mathrm{Im}_{org}) \cap area(\mathrm{Im}_{calc}))}{(area(\mathrm{Im}_{org}) \cup area(\mathrm{Im}_{calc}))} \tag{6}$$

Figure 23 summarizes the results of both algorithms on cropped images. The proposed algorithm shows size-restoration success rate between 70% and 95% for over 89% of the images, with a total average of 80% success rate. The FD algorithm shows poorer results with average success rate of 53%.

Forged Images In another experiment, copy-paste forgery was tested (see Fig. 17). This forgery involves replacing a part of the original image by another part taken from the same or different image. A set of 45 forged images was created by replacing 4 patches in each image with patches extracted from a different image. Each patch size was 4% of the total image size. This test sets a major challenge to the algorithm due to the small size of the forged region.

Forgery detection was applied on the images and regions suspected as forged were marked. For the presented algorithm forged regions were marked using the following settings:

1. PFA events for which the PFA direction formed more than 90° angle with the calculated center were marked as suspected forgery.
2. Smoothing was applied to the map of weighted distances from the PFA events to the calculated center of the image to remove remaining noise.
3. Filtered PFA events with a weighted distance greater than a predefined threshold were marked as suspected forged regions. This threshold was set to be 0.0014% of the maximum penalty factor.

The image was segmented into regions the same size as the forged patches. An authentic region marked as forged is considered a false positive while a forged patch marked as such was considered a true positive. An example is shown in Fig. 24. An original image with 4 pasted regions is shown,

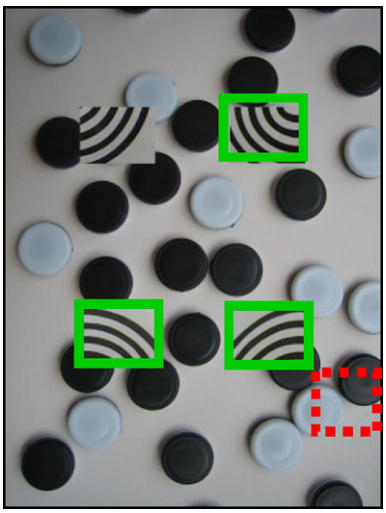

Fig. 24 An example of a synthetically forged image, with 4 patches inserted from a different image. The proposed algorithm result is shown: 3 true-positives are marked in *green*; the false-positive is marked in *red* (*dashed*)

with the regions detected as suspected forgery marked. Note that a false positive window was detected in the bottom-right corner due to image noise.

Testing for copy-paste forgery was performed for the FD and the proposed algorithm. Similar to the proposed algorithm, the FD algorithm evaluated every image region for forgery. A region was considered forged if the calculated center was not within 4% of the image center. Results for the proposed method showed a very low false-positive rate (6%), accompanied by close to 70% true-positive rate, FD suffers from over-detecting regions in the image as suspected forgery. Consequently it has a false-positive rate of 89%, with true-positive rate of 96%.

The advantage shown by the proposed algorithm over FD algorithm may be attributed to two facts: First, the FD algorithm assumes a specific chromatic aberration (LCA) and attempts to fit a model to the data. However, as discussed above, most images show aberration effects that arise from a mixture of sources and a variety of aberrations. The proposed algorithm has the ability to cope with these effects based on numerous clues collected from the entire image. Second, the proposed algorithm analyzes the image prior to the global estimation of the image center, allowing it to base its calculations only on valid and relevant data (i.e. edges with significant PFA events). The FD algorithm, on the other hand, uses the entire image data to calculate the center while unable to discard the irrelevant sections, leading to reduced accuracy.

6.3 Additional Examples

This section presents additional images, which demonstrate the different capabilities of the proposed algorithm.

Authentic Images Figure 25 displays examples of center detection in authentic images. The calculated center lies near the geometric image center. The accuracy is evaluated as average angular error as described in (5). The yellow dot depicts the geometric image center and the green star depicts the calculated image center.

Cropped Images Figure 26 displays examples of crop detection. In each image, the de-saturated region was cropped from the original image. The yellow dot depicts the original image center and the green star depicts the calculated image center based solely on the cropped image. The dashed lines depict the presumed image size and shape, based on the calculated center. As described in the paper, the accuracy measure is depicted by the size restoration rate (see (6)).

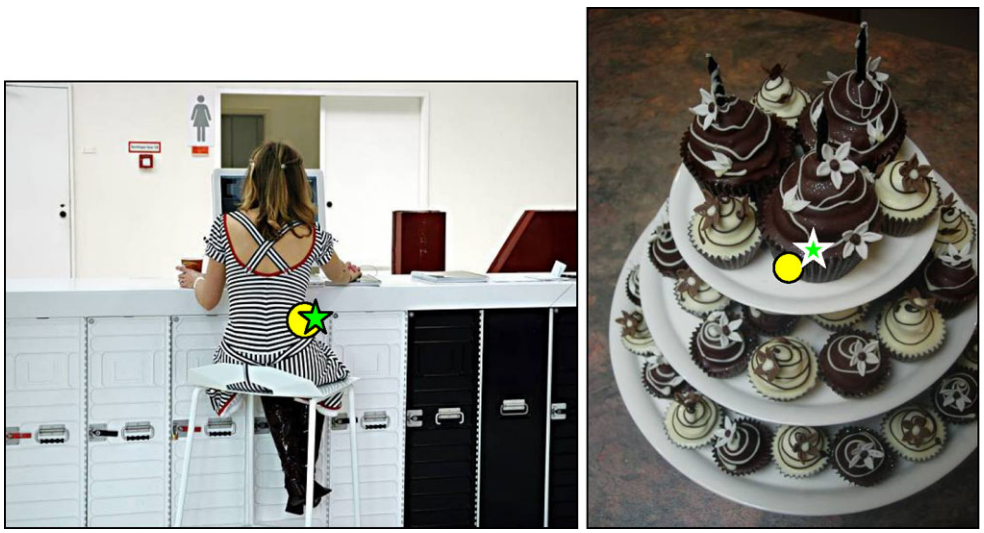

Fig. 25 (Color online) Center Detection of authentic images. The *yellow dot* depicts the geometric image center and the *green star* depicts the calculated image center. Average angular errors are: 5.71 degrees (*top*) and 8.9 degrees (*bottom*)

Fig. 26 (Color online) Crop detection. The *yellow dot* depicts the original image center and the *green star* depicts the calculated image center. Size restoration rates are: 95% (*top*) and 91% (*bottom*)

Fig. 27 (Color online) Forgery Detection. *Left*: Original image. *Right*: Image was forged by changing the *top right corner*. The suspected region correctly detected by the algorithm is marked in *green*

Forgery Detection Figure 27 shows an additional example of forgery detection. To identify forged regions, the weighted distance from the calculated center of the image (as described in (3)) is calculated independently for each PFA vector. Regions with a large weighted distance are indicated as suspected to be forged and are marked in green.

7 Discussion

Some of the results of the above tests share a common characteristic that indicates the general strengths and weaknesses of the algorithm. Across all tests, it can be seen that images with stronger contrast generate better results (see Fig. 30). This can be seen in Fig. 28 (top) which shows the distribution of angular error as a function of image contrast measured as sum of squared gradients in the image. This is not surprising, as PFA appears more strongly over edges with greater contrast as described in Sect. 4 (see Fig. 28 bottom). In addition, as the sharpness of the image decreases, there are fewer distinct boundaries between objects in the image. Since this is the grounds on which the PFA is sought (see Sect. 5), the possible number of PFA events decreases as well as their reliability, and the analysis may become degraded.

Furthermore, the number of PFA events detected in an image and their distribution within the image affect the results of the algorithm. The location of PFA events is highly dependent on image content. In our tests, PFA events ranged in number between 500 and 12,000. Distribution of PFA events was found to be uneven. Figure 29 shows 2 examples of histograms representing the percent of PFA events found in the given image per angle about the image center and per eccentricity from the center. Note that PFA distribution per eccentricity is confounded with PFA strength which increases with eccentricity and is thus more easily detected.

The weaknesses of the algorithm beyond the statistics of PFA events in the image, involves the specifics of the forgery. First, the method is not sensitive to symmetric cropping about the image center as shown in Fig. 31 (left). Additionally, insertion of a region extracted from a source image into the same relative location in another image, results in a forged image that is difficult to detect by the proposed method. An example is shown in Fig. 31 (right) where sev-

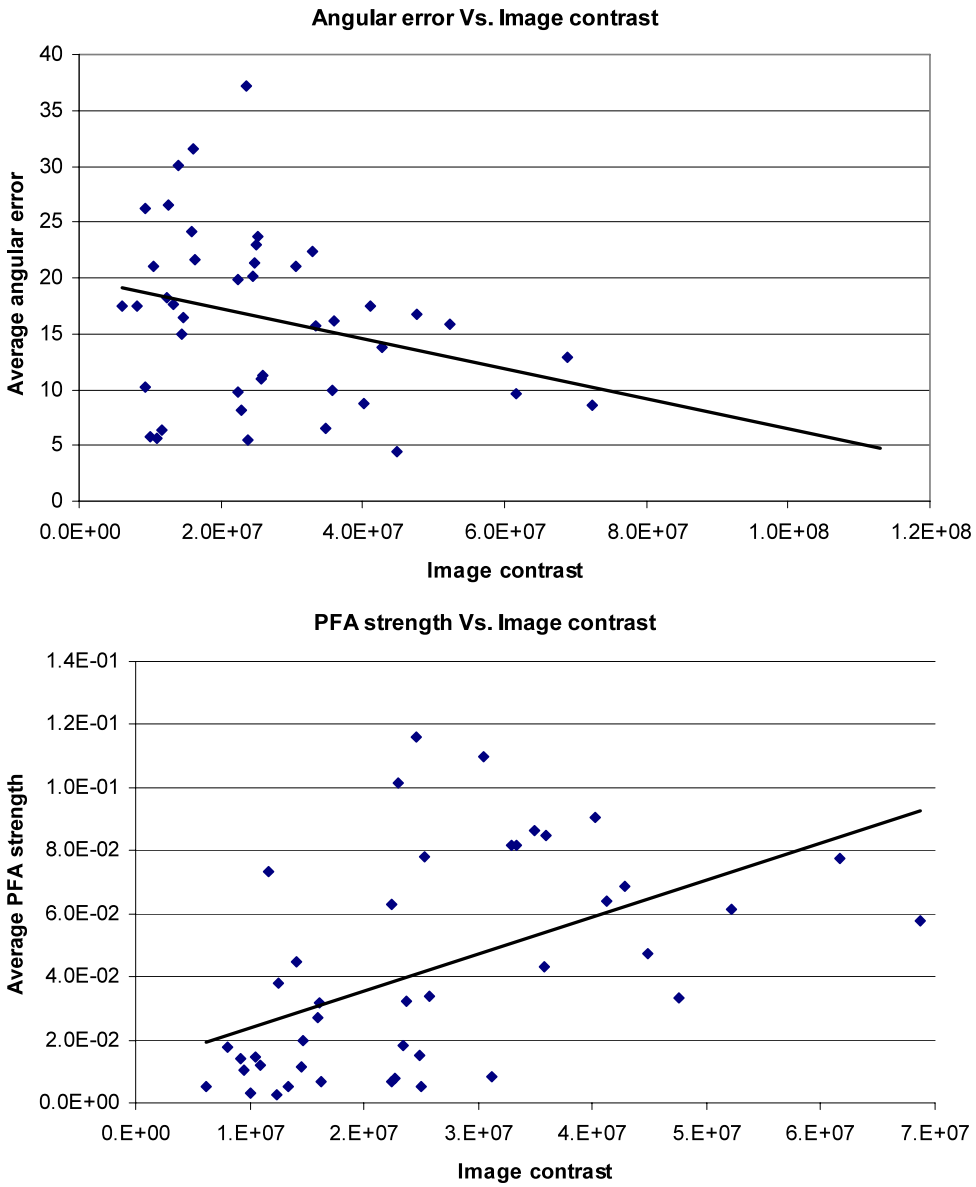

Fig. 28 *Top*: Distribution of angular error as a function of image contrast. As contrast increases, the angular error decreases since additional PFA events are found. *Bottom*: Average PFA strength of all events in an image, as a function of image contrast. As described in Sect. 4, PFA is stronger across edges with greater contrast. In both graphs, a linear trend-line was added for emphasis purposes. Data based on a set of 43 authentic images

eral ducks have been pasted into the image. Most of them, however were pasted in the same relative locations as in the source image, and thus were not detected by the algorithm. However one of the ducks was pasted in a different relative location in the image and thus produced inconsistencies in the PFA direction which were detected by the algorithm. These drawbacks arise from the fact that PFA is based on local indicators each of which are evaluated relative to the image center.

8 Conclusions and Further Research

The ease in which digital images are forged today has significantly undermined their reliability. This affects not only the level of trust people are willing to give to images they come across, but also undermines the usage of photos as legal evidence in the court of law. Two main approaches exist when attempting to authenticate an image, differentiated mainly by the requirement to add data to the image. While methods which add data (e.g. watermarking) give relatively good results, their main drawback is the need to process the image upon acquisition, requiring special hardware or software in the camera. In this work a new approach was presented which is able to verify authenticity of images and detect forged regions without additional data. Furthermore, it uses no additional information, including statistical assumptions or camera specifications. The algorithm is based on features, inherent to the camera acquisition process, namely, chromatic aberrations due to camera lens and CCD sensors, which allow detection of forgery given no prior data. The approach does not assume a specific aberration (such as LCA

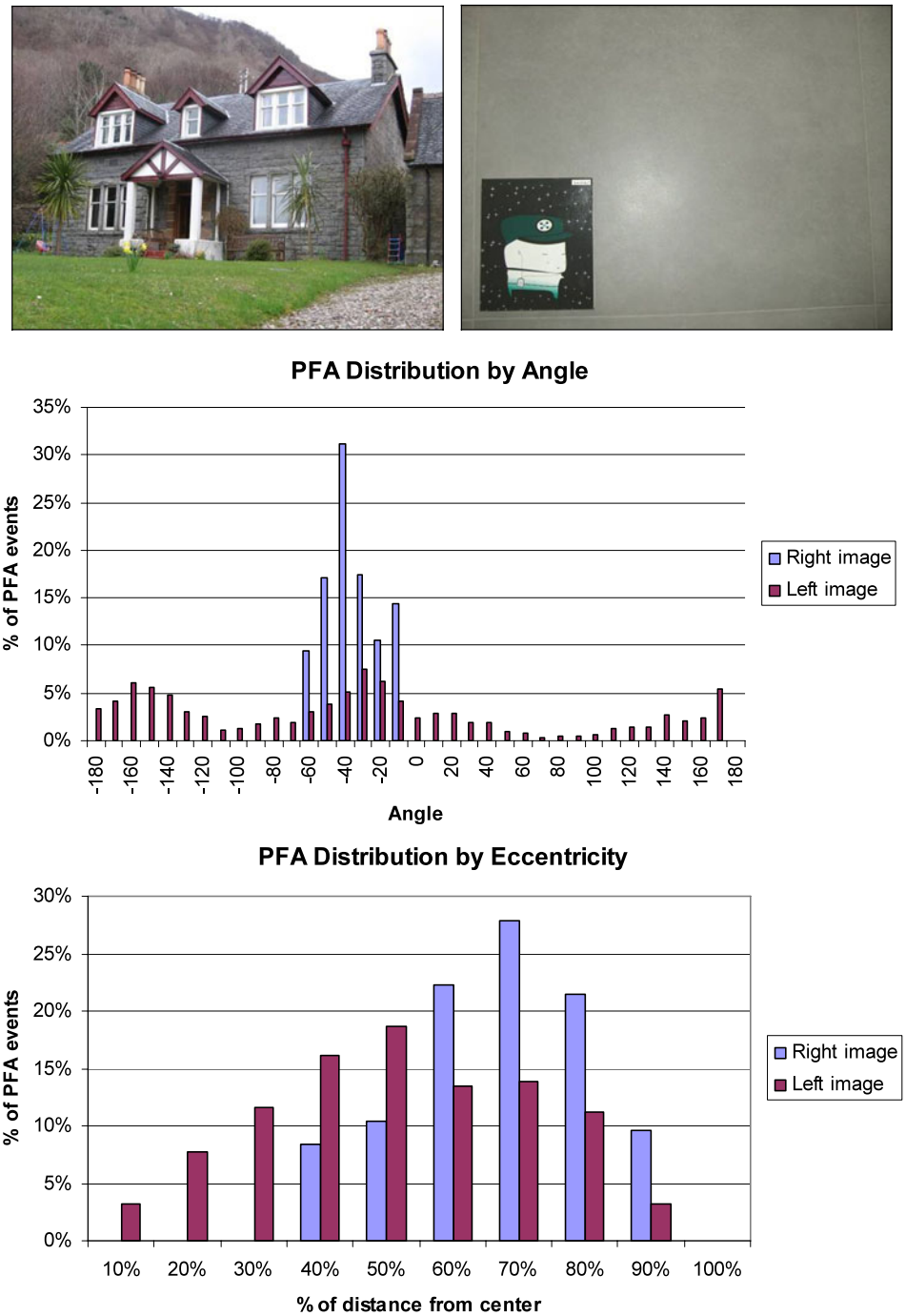

Fig. 29 Distribution of PFA events. *Middle row*: Percent of PFA events found in the *images above*, per angle about the image center. *Bottom Row*: Percent of PFA events found in the *images above*, as a function of eccentricity from image center (percentage relative to maximum possible distance from image center)

in Johnson and Farid 2006) nor does it attempt to model the complex intermixing of various aberrations in the image. Rather it relies on the fact that aberration cues have specific local characteristics that can be easily detected and that these cues are abundant in the image. Thus, the proposed method enables coping with a wider variety of aberrations and image noise.

The algorithm was tested on three groups of images and was shown to be capable of: (a) testing an image for authenticity, (b) identifying cropped images and restoring the size and shape of the original and (c) handling images that have been forged by means of copy-paste.

The copy-paste forgery tests (see Sect. 6), have shown that in addition to detecting and marking suspicious regions in the image, the process also correctly distinguishes between the source of the forgery and its copy, if it resides in the same image.

Further advances in this course of research may include integration of the algorithm with other methods of its class. For example, use the proposed algorithm in conjunction

Fig. 30 Contrast effects. *Left*: Image with high contrast and clear boundaries generated a good result (average angular error of 8.47 degrees). *Right*: Blurred image with a relatively small number of objects. The algorithm performed moderately well with almost 20 degrees of average angular error

Fig. 31 (Color online) *Left*: A cropped image about its center. Desaturated section was removed, but the algorithm failed to detect the forgery, since PFA events agreed on the detected center, depicted by *green star*. *Right*: A forged image from www.worth1000.com. Several ducks were pasted onto the original image in the same relative locations as in the source image and thus not detected by the algorithm. However, suspected regions (marked in *red*) were detected on one of the ducks since it was reflected prior to pasting into the image and thus produced inconsistencies in the PFA directions

with the JPEG related algorithm described in Wang and Farid (2006). While the proposed method is not dependant upon specific models, it does suffer from sensitivity to very low contrast. The second algorithm, while having the disadvantage of relying upon models which may not be generic enough, can better handle images with increased blur.

References

Born, M., & Wolf, E. (1999). *Principles of optics, electromagnetic theory of propagation, interference and diffraction of light*. Cambridge: Cambridge University Press.

Cutzu, F., Hammoud, R., & Leykin, A. (2003). Estimating the photorealism of images: distinguishing paintings from photographs. In *Proc. IEEE conference on computer vision and pattern recognition*.

Daly, D. (2001). *Microlens arrays*. Boca Raton: CRC Press.

Fridrich, J., Soukal, D., & Lukáš, J. (2003). Detection of copy-move forgery in digital images. In *Proc. digital forensic research workshop*, Cleveland, OH.

Gonzalez, R. C., & Woods, R. E. (2002). *Digital image processing*. New York: Prentice Hall.

Horn, B. K. P., & Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, *17*, 185–203.

Jahne, B. (2005). *Digital image processing*. Berlin: Springer.

Jenkins, F. A., & White, H. E. (1976). *Fundamentals of optics* (4th ed.). New York: Mcgraw-Hill.

Johnson, M. K., & Farid, H. (2005). Exposing digital forgeries by detecting inconsistencies in lighting. In *Proc. ACM multimedia and security workshop*, New York (pp. 1–9).

Johnson, M. K., & Farid, H. (2006). Exposing digital forgeries through chromatic aberration. In *Proc. ACM multimedia and security workshop*, Geneva, Switzerland.

Keren, D. (2002). Painter identification using local features and naive Bayes. In *Proc. international conference on pattern recognition (ICPR)* (Vol. II, pp. 474–477). Berlin: Springer.

Lyu, S., & Farid, H. (2005). How realistic is photorealistic? *IEEE Transactions on Signal Processing*, *53*(2), 845–850.

Lyu, S., Rockmore, D., & Farid, H. (2004). A digital technique for art authentication. In *Proc. National Academy of Sciences*.

Muller, H., Muller, W., Squire, D. M., et al. (2001). Performance evaluation in content-based image retrieval: overview and proposals. *Pattern Recognition Letters*, *22*, 593–601.

Negahdaripour, S., & Horn, B. K. P. (1989). A direct method for locating the focus of expansion. *Computer Vision Graphics and Image Processing*, *46*, 303–326.

Ochi, S., Lizuka, T., Sato, Y., Hamasaki, M., Abe, H., Narabu, T., Kato, K., & Kagawa, Y. (1997). *Charge-coupled device technology*. Boca Raton: CRC Press.

Parr, R. (2006). Digital photography FAQ. http://www.cs.duke.edu/~parr/photography/faq.html.

Pedrotti, F. L., Pedrotti, L. M., & Pedrotti, L. S. (2006). *Introduction to optics*. Reading: Addison-Wesley.

Popescu, A. C., & Farid, H. (2005). Exposing digital forgeries in color filter array interpolated images. *IEEE Transactions on Signal Processing*, *53*(10), 3948–3959.

Ray, S. F. (2002). *Applied photographic optics* (3rd ed.). Boston: Focal Press.

Rousseeuw, P. J., & Leroy, A. M. (2003). *Robust regression and outlier detection*. New York: Wiley.

Rudolf, K. (1992). *Optics in photography*. Bellingham: SPIE.

Schroff, F., Criminisi, A., & Zisserman, A. (2008). Object class segmentation using random forests. In *Proc. British machine vision conference*.

Smith, W. J. (2007). *Modern optical engineering* (4th ed.). New York: McGraw-Hill.

Szummer, M., & Picard, W. (1998). Indoor-outdoor image classification. In *Proc. IEEE intl workshop on content based access of image and video databases*.

van Walree, P. (2009). Photographic optics collection. http://toothwalker.org/optics.html.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. Berlin: Springer.

Wang, W., & Farid, H. (2006). Exposing digital forgeries in video by detecting double MPEG compression. In *Proc. ACM multimedia and security workshop*, Geneva, Switzerland.

Wolfgang, R. B., & Delp, E. J. (1996). A watermark for digital images. In *Proc. IEEE intl conference on image processing*.

Wyszecki, G., & Styles, W. S. (1982). *Color science: concepts and methods, quantitative data and formulae* (2nd ed.). New York: Wiley.