An Image Recapture Detection Algorithm Based on Learning Dictionaries of Edge Profiles

Thirapiroon Thongkamwitoon, *Student Member, IEEE*, Hani Muammar, *Member, IEEE*,
and Pier-Luigi Dragotti, *Senior Member, IEEE*

Abstract—With today's digital camera technology, high-quality images can be recaptured from an liquid crystal display (LCD) monitor screen with relative ease. An attacker may choose to recapture a forged image in order to conceal imperfections and to increase its authenticity. In this paper, we address the problem of detecting images recaptured from LCD monitors. We provide a comprehensive overview of the traces found in recaptured images, and we argue that aliasing and blurriness are the least scene dependent features. We then show how aliasing can be eliminated by setting the capture parameters to predetermined values. Driven by this finding, we propose a recapture detection algorithm based on learned edge blurriness. Two sets of dictionaries are trained using the K-singular value decomposition approach from the line spread profiles of selected edges from single captured and recaptured images. An support vector machine classifier is then built using dictionary approximation errors and the mean edge spread width from the training images. The algorithm, which requires no user intervention, was tested on a database that included more than 2500 high-quality recaptured images. Our results show that our method achieves a performance rate that exceeds 99% for recaptured images and 94% for single captured images.

Index Terms—Image forensics, recapture detection, image acquisition, aliasing, blurriness, dictionary learning, K-SVD.

I. INTRODUCTION

DIGITAL cameras today are capable of delivering high resolution images with pleasing colour and tone reproduction at relatively low cost to the consumer. Moreover, with the widespread availability of high quality colour ink-jet printers and liquid crystal display (LCD) devices, images can be easily reproduced by recapturing the printed or displayed image with a digital camera. If a high quality digital camera is used, such as a DSLR, and the image is recaptured from a good quality print or a high resolution LCD monitor, then a recapture with high fidelity can be obtained.

Traditionally, photographs have been associated with a high degree of authenticity and were considered difficult to forge. With the advent of digital photography image tampering is now commonplace and can easily be performed using commercial, widely available, image editing software [3]. In practice, unless an attacker is highly skilled, imperfections in the forged image may be present and the attacker may attempt to conceal them by recapturing the forged image from an LCD monitor. By recapturing the image, an additional level of authenticity, typically associated with a single captured image, is introduced into the forgery making it more difficult to detect. For this reason this paper focuses on the problem of detecting whether a given image was recaptured with a digital still camera from an image displayed on an LCD monitor or whether it was a single capture of a natural scene. The methods presented in this work aim to be robust and reliable and can be applied to unverified sources, such as the internet.

A. Related Work

The problem of detecting recaptured images from printed material, such as photographic paper or magazines, has been addressed in the literature [4], [5]. The methods identify a recaptured print from its specularity or from the dithering patterns applied by the printer. Several researchers have considered the recapture detection of images both from prints and from LCD monitors [6], [7]. They develop detectors based on several features associated with recaptured images including the non-linearity of the tone response curve, the spatial distribution of the specularity in the image, image contrast, colour, chromaticity and sharpness. The detection of recaptured images from LCD monitors has been equally addressed in the literature, for example by Cao and Kot [8], where a detector for recaptured detection based on some features of recaptured images is proposed. The fine texture pattern sometimes present in recaptured images is detected by computing Local Binary Pattern (LBP) features at multiple scales. The loss of detail in the recaptured image, due to the relatively low display resolution of the monitor compared to the camera's image sensor, is detected by computing a multi-scale wavelet decomposition where the mean and standard deviation of the absolute wavelet coefficients are used as features. The apparent increase in saturation in colours of the recaptured image is detected using 21 different colour features. Finally, the output of the individual detectors is fed

into a trained probabilistic SVM classifier. Yin and Fang [9] detect images recaptured from an LCD monitor by analysing noise features and by detecting traces of double Jpeg compression using the MBFDF algorithm [10]. To estimate the noise features the image was denoised using three different discrete wavelet transforms and statistical features such as mean, variance, skewness and kurtosis were computed from the histogram of the extracted noise residual. Their experimental results suggest that the proposed features perform well for detecting photographic copying from LCD monitors. Ke et al. [11] train a support vector machine to classify recaptured images from LCD monitors with a 136 dimension feature set. Their descriptors are based on blurriness, texture, noise and colour features. They apply their method to a dataset of recaptured images taken with smart-phone cameras and claim a detection rate of 97.2% when the features are combined. The images used in their dataset [12] are low in resolution and quality, however, due to the smart-phone cameras used to perform the recapture.

There are several other situations where recapture from LCD monitors has been analysed. Ng et al. [13] have addressed the problem of classifying photographic images and photo-realistic computer graphics (PRCG) images. They demonstrated that PRCG images that were recaptured from LCD monitors were more difficult to distinguish from originally captured photographic images. In 'face spoofing' attacks, a person attempts to bypass a face authentication system by impersonating (or masquerading) as another person who is authorized by the system. An image or video of the valid person may be displayed on a portable tablet computer and presented to the authentication system where it is recaptured by the system's digital camera. This type of face spoofing has been addressed in [14] and [15]. Recently there have been efforts to detect recapture of videos [16]–[18]. These methods use features which are unique to video systems.

B. Contribution and Paper Outline

In this paper we provide a comprehensive overview of some of the most common traces, or features, that are introduced in an image when it is recaptured from an LCD monitor. We then concentrate on aliasing and blurriness as these are probably the two features that are most commonly encountered in recaptured images. Moreover, the remaining features are highly scene dependent and can be difficult to extract. Our first contribution is then to show how aliasing can be eliminated by properly configuring the recapture settings. Aliasing, also commonly referred to as image moiré, is frequently introduced in a recaptured image due to the sampling of the LCD monitor pixel structure. We model the LCD monitor pixel grid as a 2D square wave pattern and we use the model to determine the capture distance and lens aperture setting as a function of the monitor and image sensor pixel pitch values. This results in a high quality recaptured image that is free from visible aliasing artefacts. We then use this method to create a database of high quality single capture and alias-free recaptured images using a wide range of consumer digital cameras which we later

use for testing and benchmarking of the proposed algorithm. Our second contribution is a method that uses the edge blurriness and distortion introduced by the recapture process as a feature to detect if a given image has been recaptured from an LCD monitor. We show that the edges found in single and recaptured images can be fully characterised by their line spread function (LSF). We then describe how sets of elementary atoms that provide a sparse representation of LSFs can be learned using the K-SVD dictionary learning method [19]. Specifically, a single-capture dictionary is created from a training set of single captured images and a second one from recaptured images. We also compute an edge spread width from the line spread function of the image and combine this feature with the dictionary approximation errors to train an SVM classifier. We classify a query image as single or recaptured depending on its location relative to the SVM hyperplane. Our approach has led to an automatic algorithm which is robust in that it is applicable to a wide range of naturally occurring images taken under different conditions and has an average success rate greater than 96%.

The paper is organized as follows. An overview of common features that occur in recaptured images is provided in Section II. In Section III a method for recapturing images that are free from aliasing is presented. The proposed algorithm for recapture detection is described in Section IV. Details of the image recapture database are provided in V and in Section VI we describe the experimental procedure for training and testing the proposed algorithm and provide the classification results. Finally, we conclude in Section VII.

II. Features of Recaptured Images

In this section we provide an overview of some of the more common features found in images that have been recaptured from LCD monitors. We assume that lens geometric distortion, such as barrel or pincushion distortion, has been minimised, that distortion due to the capture geometry has been eliminated and that the individual monitor pixels are not resolved by the recapture camera. We also assume that there are no specular reflections from the monitor front panel due to ambient light sources.

A. Aliasing

Aliasing is sometimes introduced in digital camera images when the scene is insufficiently band-limited or contains detail with very high spatial frequencies [20]. In cameras that are equipped with a Colour Filter Array (CFA) [21] the colour channels are normally sampled at frequencies that are lower than the native frequency of the image sensor. The green channel of a Bayer CFA can be described by a quincunx lattice arrangement and has a frequency response equivalent to the native 'unfiltered' sensor only in the horizontal and vertical directions. The red and blue channels are sampled on a rectangular lattice at one half the frequency of the native sensor. The diagram in Fig. 1 shows the Nyquist boundaries and replication points for the red (R), green (G) and blue (B) colour channels of the Bayer CFA.
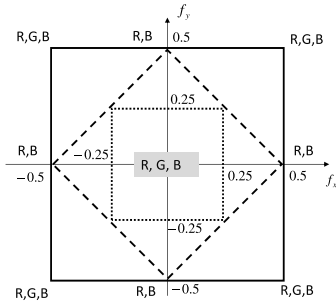
Fig. 1. The Nyquist boundary for the unfiltered sensor (solid), green (dashed) and red/blue channels (dotted) of the Bayer CFA. Replication points for the red (R), green (G) and blue (B) channels are shown. All frequencies are in cycles/pixel.

Most camera manufacturers fit optical anti-aliasing filters [22] to band-limit the high frequency components in the scene and prevent aliasing. However, the cut-off frequency of the filter is normally set above the Nyquist frequency to preserve the camera response at frequencies in the range 30-80% of the Nyquist frequency. The recapture of an image displayed on the screen of an LCD monitor is, therefore, highly likely to introduce aliasing due to the high frequency periodic pattern of the monitor pixel grid structure. Indeed, casually recaptured still images or videos of LCDs are often characterised by the presence of aliasing artefacts, also referred to as colour moiré, over the visible region of the display. These artefacts are very difficult to eliminate through post-processing. Therefore, aliasing can be used as a feature for detecting recaptures. When aliasing artefacts are present in the recaptured image, the 2D DFT of the noise residual is likely to exhibit peaks in the 2D spectrum. Detection of these peaks allows the identification of recaptured images [2], [15]. To avoid that, in Section III, we develop a recapture method that avoids aliasing.

B. Blurriness

Naturally occurring scenes contain a wide range of edges that vary in contrast and sharpness. When a scene is acquired with a digital camera, a certain level of blur, or distortion, is introduced into the image by the acquisition device. This occurs despite the fact that the image was correctly focussed by the camera at the time of capture. Imperfections in the lens, such as spherical aberration can introduce blur, as can diffraction. The latter is introduced when the diameter of the lens aperture is very small (due to a large aperture setting). Additional distortion may be introduced by processing carried out internally in the camera, such as sharpening, contrast enhancement or CFA demosaicing. The blur characteristics may, to a large extent, be considered unique to the camera at the time of acquisition. One way of characterising the blur is with the point spread function (PSF) of the capture device. In practice measuring the PSF of a device is not easily achieved and the line spread function (LSF) is used instead. By definition, a line spread function is a 1-D function corresponding to the first derivative of the edge spread function (ESF) [23]. A model of the blurring pattern introduced by a camera can be estimated by measuring, and statistically combining, the first derivative of edge profiles from a slanted edge test target captured with the camera [24].

The process of displaying an image on a monitor and recapturing it with a second digital camera increases the level of blurring relative to the originally captured image. The largest contributor to the increase in blur seen in the recaptured image is the drop in spatial resolution of the image due to the LCD monitor. Each stage of the image acquisition chain introduces a unique pattern of distortion into the image. The loss in sharpness and increase in distortion, such as ringing, that is introduced in the edge when it is captured, displayed and recaptured propagates through the chain and is present in the final image. The edges in the image, therefore, contain useful information, that can provide vital clues which enable us to reliably detect whether an image has been originally captured or whether it has been recaptured from a monitor display. For this reason, the algorithm described in this paper makes extensive use of this feature.

C. Noise

The two main sources of noise associated with images captured with a digital camera at normal and high levels of scene illumination are temporal noise, comprising mainly of shot noise, and fixed pattern noise which is dominated by Photo Response Non-Uniformity (PRNU) noise. The distribution of image noise in the recaptured image will be predominantly influenced by the noise characteristics of the recapture camera, the brightness setting of the LCD monitor, the capture distance and the scene content. The noise characteristics of the camera used to capture the original scene are likely to be also present in the recaptured image, but they will be band-limited due to the blurring effect introduced by the recapture process discussed in Section II-B. The unique PRNU fingerprint of a camera's image sensor has been shown to be a highly successful tool for identifying the source camera from an image or a set of images [25]. The method has been applied successfully to detect the presence of the PRNU pattern in a scan of a printed image [26]. However, very small levels of rotation of the print are enough to significantly reduce the detection performance, since misalignment is introduced between the PRNU pattern and the recapture device (the scanner). This limits the applicability of their approach to our application since successful identification of the original capture device would require very low levels of misalignment between the LCD monitor pixel grid and the camera's image sensor, which in practice, would be very difficult to achieve. Thus, image noise is not considered as a reliable feature for recapture detection.

D. Contrast, Colour and Illumination Non-Uniformity

Almost all digital cameras and LCD monitor devices today, support the sRGB colour encoding specification [27]. In addition to specifying the gamut of colours that can be represented, the sRGB specification also describes forward and reverse non-linear tone transformation curves. In an ideal image capture display environment, the overall system tone

response between input scene intensities and output display intensities at the monitor is linear. In practice, digital cameras apply a tone response function that deviates slightly from the ideal sRGB response, but is intended to provide a 'more pleasing' image that is slightly higher in contrast. In a recapture image chain where the response function of both the original and recapture cameras deviates from the sRGB specification as described above, the overall recaptured image is likely to appear higher in contrast relative to the single captured image. There may be some noticeable loss of detail and clipping of pixel values in the light and dark regions in the recaptured scene when compared with the original capture. For image contrast to be used as a feature for recapture detection, a reliable, scene content independent method for the recovery of the global scene contrast or tone response function is required. There exist methods in the literature [28], [29] however, they are dependent on scene content and may not, therefore, provide a reliable tool for recapture detection.

Colour related artefacts that may be present in recaptured images include colour balance errors, such as unwanted tints affecting the whole image, and increased colour saturation. Colour balance errors in a recaptured image can be minimised by calibrating the display monitor and by presetting the white point of the recapture camera to the LCD monitor white point before recapture. Thus, colour balance errors present in the recaptured image will have likely been introduced by the original camera that was used to capture the scene and not during the recapture process. The increase in colour saturation present in the recaptured image is likely to be due to the increase in overall image contrast as described above. Colour differences between original and recaptured images from LCD monitors are likely to be highly dependent on device characteristics and settings. Furthermore, reliable extraction of colour features is highly dependent on scene content.

The transition to LED backlighting from cold cathode fluorescent (CCFL) backlight in LCD monitors has resulted in improved colour gamut, dynamic range and display non-uniformity. However, as monitor display sizes have increased, obtaining even backlight illumination remains a challenge for some low cost display devices. A luminance gradient may be noticeable in recaptured images containing large regions that are low in texture or detail. Identification of the luminance gradient would enable recaptured images to be detected. However, the accuracy of detection is likely to be highly dependent on scene content, and, therefore, luminance gradient is also not considered as a reliable feature for recapture detection.

III. MINIMIZING ALIASING IN RECAPTURED IMAGES

In this section, we introduce a method to recapture images without introducing visible aliasing. We do this by modelling the luminance signal generated by the projection of the monitor pixel structure on the image sensor of the recapture camera as an infinite 2D square wave pattern [2]. The signal is then sampled on a rectangular lattice. We recover the frequency of the dominant alias pattern in the recaptured image and
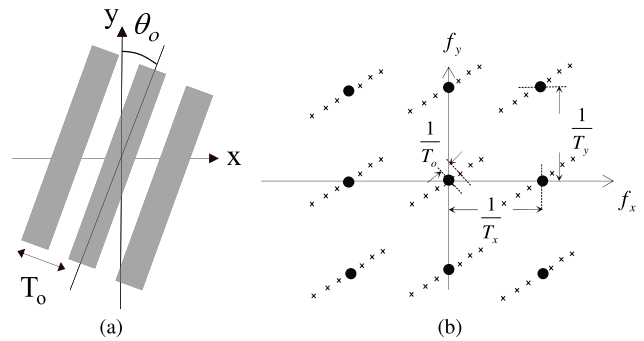
Fig. 2. The 2D square wave model of the monitor pixel grid in the vertical direction and its Fourier transform. The monitor grid is rotated by θ_o relative to the image sensor. (a) 2D square wave. (b) Fourier transform.

use this to determine the lens focal length, capture distance and lens aperture needed to minimise the perceived level of aliasing. We note that existing solutions for recapturing images from display monitors without aliasing are primarily geared towards display quality testing and often require special optical hardware [30], [31].

A. Modelling the Recapture of an LCD Screen

Each pixel, in a modern LCD monitor, comprises three long and narrow vertically oriented sub-pixel elements that are covered with a red, green and blue filter, respectively. The monitor pixel pitch, T_m is the distance between two successive sub-pixel elements of the same colour. The pixels in most LCD monitors are square and we can, therefore, assume that the vertical and horizontal pixel pitch dimensions are the same. The pixel pitch of the monitor projected on the camera's image sensor during recapture is given by $T_o = mT_m$, where m is the optical magnification of the recapture camera lens. In the following analysis we make the assumption that the geometric distortion due to the camera lens is negligible and that the monitor is fronto-parallel to the image sensor plane. We also assume that the image sensor pixels are square, ensuring that the sensor horizontal and vertical pixel pitch dimensions are the same. During recapture, the luminance component of the projected monitor pixel grid is periodic in the horizontal and vertical dimensions with period, T_o. The monitor pixel grid can, therefore, be modelled in either dimension by an infinite 2D square wave pattern as shown in Fig. 2a.

To determine the frequency spectrum, $G(f_x, f_y)$, of the square wave we take the Fourier transform of its Fourier series [32]:

$$G(f_x, f_y) = \sum_{n_o=-\infty}^{\infty} C_{n_o} \delta \left(f_x - \frac{n_o}{T_o} \cos \theta_o, f_y - \frac{n_o}{T_o} \sin \theta_o \right),$$

(1)

where, $C_{n_o} = \sin(\pi n_o/2)/(\pi n_o)$ for a square wave and $n_o = -\infty \ldots -1, 0, 1 \ldots \infty$. By convolving the spectrum of the continuous square wave with the camera's sampling system, we obtain the spectrum of the sampled square wave. For an image sensor with vertical and horizontal pixel pitch dimensions represented by T_x and T_y respectively,

the locations of the delta functions in the frequency spectrum are given by the vectors:

$$(f_x, f_y) = \left(\frac{n_o}{T_o} \cos\theta_o + \frac{n_x}{T_x}, \frac{n_o}{T_o} \sin\theta_o + \frac{n_y}{T_y} \right), \qquad (2)$$

where n_x and $n_y = -\infty \ldots -1, 0, 1 \ldots \infty$ and θ_o is the rotation of the monitor pixel grid relative to the image sensor axis. A schematic diagram showing the locations of the delta functions described by Equation (2) is shown in Fig. 2b. The square wave spectrum is replicated at n_x/T_x and n_y/T_y. For a monochrome image sensor (without a colour filter array), the Nyquist boundary limit is given by $[1/2T_x, 1/2T_y]$.

B. Eliminating Visible Aliasing From the Recaptured Image

Aliasing will be visible in the recaptured image when the delta functions associated with the replicated spectra (spectral orders) that are high in energy lie within the Nyquist boundary of the sensor. Let us assume that the monitor pixel grid is perfectly aligned to the image sensor axis (i.e. $\theta_o = 0$), and that each projected monitor pixel pitch spans k image sensor pixels, where k is some positive integer or fractional value (i.e. $T_o = kT_x$). Then for an image sensor with a Bayer CFA pattern, the positions of the delta functions in the green channel are given by:

$$(f_x, f_y)_g = \left(\frac{n_o}{kT_x} + \frac{n_x}{T_x}, \frac{n_y}{T_y} \right). \qquad (3)$$

For the red and blue channels the positions of the delta functions are given by:

$$(f_x, f_y)_{r,b} = \left(\frac{n_o}{kT_x} + \frac{n_x}{2T_x}, \frac{n_y}{2T_y} \right). \qquad (4)$$

Thus the spectrum of the square wave is present at DC and is replicated at n_x/T_x and n_y/T_y in the green channel and at $n_x/2T_x$ and $n_y/2T_y$ in the red and blue colour channels.

One possibility is to cancel the aliasing in the recaptured image. This has been explored in the past [2] and was validated on a range of cameras and monitors. Although it was shown to produce good quality recaptured images free from visible aliasing artefacts, with some cameras the method leads to low frequency aliasing, or chroma moiré, due to the small variations in image magnification across the image sensor introduced by the lens geometric distortion.

A more robust solution to minimising the level of perceived aliasing in the recaptured image is to maximise the frequency of the aliasing pattern in the green channel, while at the same time minimising the amplitude of the alias pattern in the red and blue channels. The amplitude of the aliasing in the green channel is then attenuated by introducing an appropriate level of blur into the recapture resulting from the effect of the diffraction due to the lens aperture. The advantage of this approach is that the alias pattern is easier to eliminate by controlled blurring since the colour fringes and textures introduced by the interaction of the red and blue channels with the green channel due to the CFA interpolation process, described in section II-A, are eliminated.

The amplitude of the aliasing in the red and blue channels is minimised by recapturing the LCD such that the first
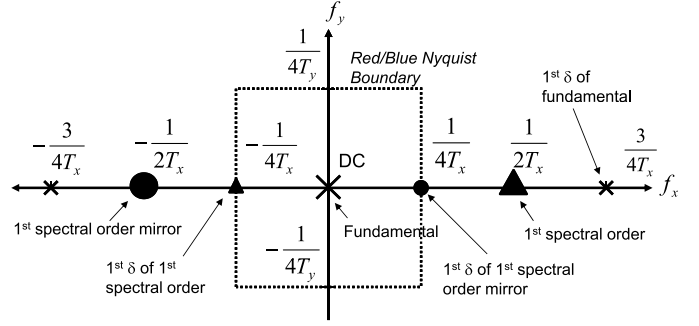

Fig. 3. A schematic diagram illustrating the positions of delta functions in the frequency domain for the recaptured LCD monitor. The first δ-function of the first spectral order (and its mirror) lies on the red/blue channel Nyquist boundary ($\pm 1/(4T_x)$), while the first δ-function associated with the fundamental lies at $\pm 3/(4T_x)$.

delta function of the first spectral order (and its mirror) are situated on the Nyquist boundary of the red and blue channels ($\pm 1/(4T_x)$) as shown in the schematic diagram in Figure 3. Note that in the diagram the centres of the fundamental and first spectral orders are shown with larger symbols than the corresponding first harmonic. The alias pattern in the green channel will have a frequency of $1/2T_x$.

To determine the capture parameters, we first need to determine the ratio of the projected monitor pixel pitch to image sensor pixel pitch, k. Note that this ratio is equivalent to the number of image sensor pixels spanned by a single projected monitor pixel. We consider the positions of the delta functions in the x-dimension only of the red and blue channels as given in Equation (4). The first replica occurs at $n_x = 1$ and, for $n_y = 0$, is described by the set of frequencies:

$$(f_x, f_y)_{r,b} \Big|_{n_y=0}^{n_x=1} = \left(\frac{n_o + k}{kT_x}, 0 \right). \qquad (5)$$

To determine the ratio, k, where the first harmonic of the first replica in the red/blue channel lies at $-1/(4T_x)$, we solve for k in Equation (5) with $n_o = -1$. This gives us a solution of $k = 4/3$. Thus each monitor pixel should project to 4/3 sensor pixels.

The distance between the camera sensor and the LCD monitor, d, can be expressed using Gaussian optics [33], as a function of lens focal length, f, and magnification, m, as[1]:

$$d = \frac{f(1-m)^2}{m}. \qquad (6)$$

From the relationship between monitor and sensor pixel pitch dimensions, $T_o = mT_m$, and using the fact that we are imposing $T_o = kT_x$, the projected monitor pixel pitch can be described by the equality $mT_m = kT_x$. Rearranging in terms of m we obtain $m = kT_x/T_m$. The subject distance, obtained by substituting m into Equation (6) and simplifying, is given as:

$$d = f \left(\frac{kT_x}{T_m} + \frac{T_m}{kT_x} + 2 \right). \qquad (7)$$

[1]In Gaussian optics image magnification is, by convention negative if the projected image size is smaller than the object size. We have considered image magnification as a positive value in this paper despite the fact that the projected size of the image on the sensor is smaller than the object. Therefore, Equation (6) has been written to take this into account.
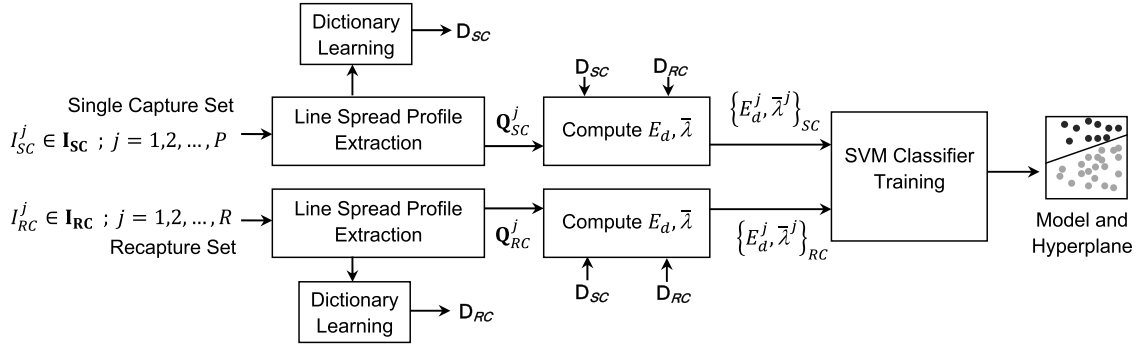
Fig. 4. Diagram showing an overview of the training process for our proposed algorithm. Following the dictionary learning process, the learned dictionaries, $\mathbf{D_{SC}}$ and $\mathbf{D_{RC}}$, are used to compute a pair of parameters $\{E_d, \bar{\lambda}\}$ for each training image. The classifier is then trained using all pairs of parameters $\{E_d, \bar{\lambda}\}$ which are labelled according to the class of training images.

To compute the required capture parameters, we substitute $k = 4/3$ in Equation (7) to give:

$$d = f\left(\frac{4T_x}{3T_m} + \frac{3T_m}{4kT_x} + 2\right). \tag{8}$$

One setting that can affect the perceived visibility of aliasing in a recaptured image is the lens aperture setting of the recapture camera. Using a high lens f-number (or small lens aperture) during image acquisition introduces blurring into the image, due to diffraction, that cannot be corrected by improving the lens design or reducing aberrations. In practice, increasing the lens f-number attenuates the amplitude of the aliasing pattern in the recaptured image, whereas using a small lens f-number (or a large aperture diameter) results in more vivid aliasing.

To eliminate the aliasing pattern from the recaptured image we increase the level of blurring due to diffraction in the recaptured image by reducing the diameter of the aperture of the recapture camera lens. We determine the aperture needed by considering the projected monitor pixels as point sources of light and we apply the Rayleigh criterion [33] to determine the required aperture value using the distance between the projected monitor pixels on the image sensor. The Rayleigh criterion is achieved when the minimum of the point spread function corresponding to the first point source of light falls on the maximum of the point spread function of the second point light source. For a camera system where the distance from the lens to the sensor is approximately equal to the focal length of the lens, the separation, r, between the projected monitor pixels on the sensor plane such that the Rayleigh criterion is satisfied is given by:

$$r = 1.22\lambda F \tag{9}$$

where λ is the wavelength of the captured light and F is the lens f-number. Since the relationship between the monitor pixel pitch as projected on the sensor, T_o, and the sensor pixel pitch, T_x, is given by $T_o = 4T_x/3$, substituting T_o for r in Equation (9) and rearranging in terms of aperture, F, yields:

$$F = \frac{4T_x}{3.66\lambda}. \tag{10}$$

When an image is recaptured at the aperture value given by F, the level of visible aliasing is reduced due to the increased overlap of adjacent monitor pixel PSFs. This is achieved without introducing an objectionable level of blur into the recaptured image. In practice, we found that the aperture value computed using the Rayleigh criterion, assuming $\lambda = 540$ nm, was insufficient to fully eliminate perceived levels of aliasing in the recaptured images. We conducted an experiment in which images were recaptured from an LCD monitor with a range of manually selected aperture settings over a wide range of cameras and we found that the mean aperture value that visually eliminated aliasing in the recaptured image was higher than the aperture value determined using the Rayleigh criterion by a factor of 1.257. The relationship between the selected aperture value and camera sensor pixel pitch used during recapture can, therefore, be summarised as:

$$F = 2544.02 \times T_x \tag{11}$$

Images recaptured using the alias frequency maximization method described in this section with the aperture settings determined using the above relationship contained no visible aliasing patterns. Some examples of single and recaptured images taken from the recapture database are shown in Fig. 12.

IV. RECAPTURE DETECTION USING EDGE PROFILES

A. Overview of Our Proposed Method

In Sections I and II we argued that the blurring distortion in an image can provide us with vital clues that can be used to determine whether an image was originally captured or whether it was recaptured from an LCD monitor. In this section we propose a method for image recapture detection based on the blurriness of edges. We stated, in Section II-B, that, for practical reasons, the line spread function of an edge was easier to determine than the edge spread function. For this reason, we base our algorithm on the *line spread profile* of an edge and not the edge spread profile.

The proposed algorithm consists of a training stage, in which a support vector machine (SVM) classifier is trained with known images, and a detection stage where the trained classifier is used to classify a given image. A diagram of the classifier training process is shown in Fig. 4. Two sets of known images are used: a set of single capture images, $\mathbf{I}_{SC}$, and a set of recaptured images $\mathbf{I}_{RC}$. The images in each
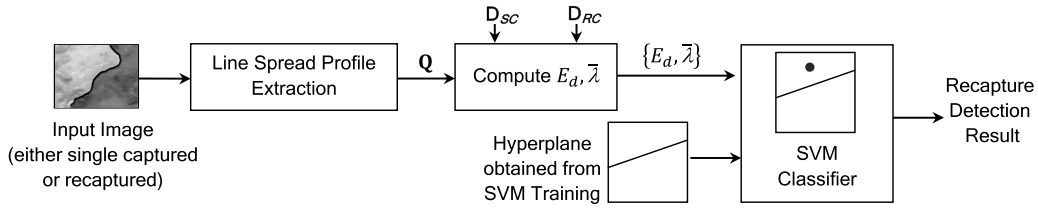
Fig. 5.   Overview working diagram of the classification scheme of our proposed recapture detection algorithm.
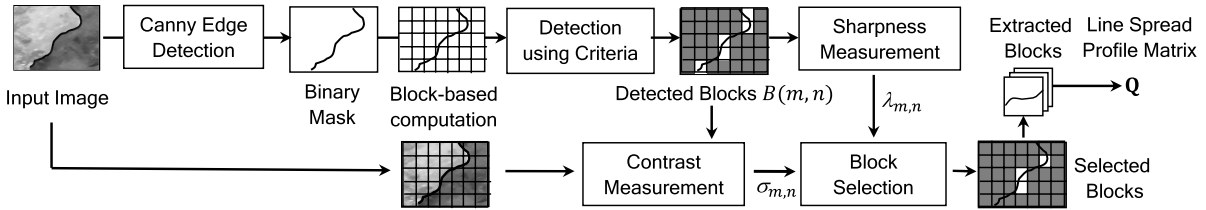

Fig. 6.   Working diagram of the proposed automatic block-based edge detection algorithm.

set are indexed with the superscript j and originate from a wide range of known cameras. The number of the images in each set, P and R, may differ. The first step of the classifier stage is to determine a set of edge profiles from each image in each set that represent the sharpest edges found in the image. The first derivative of the edge profiles is then taken to determine a corresponding set of line spread profiles for the image. Thus, for a given image from the set of single capture training images, a matrix $\mathbf{Q}_{SC}^{j}$, is generated in which each column of the matrix corresponds to an extracted line spread profile. The equivalent matrix for an image from the recaptured set is $\mathbf{Q}_{RC}^{j}$.

Two over-complete dictionaries are constructed by training using the K-SVD approach [19]. The first over-complete dictionary, $\mathbf{D}_{SC}$ is trained using the set of single captured images and the second, $\mathbf{D}_{RC}$, using the set of recaptured images. Each dictionary is trained to provide an optimal sparse representation of the line spread profiles extracted from the training set of images.

To characterize the differences between the line spread profiles of originally captured and recaptured images, we introduce two parameters related to edges: a sparse representation error E_d and an average line spread width $\bar{\lambda}$. These parameters were chosen because they provide a concise but informative description of the differences between the line spread profiles of original and recaptured images. The first metric, E_d, represents the difference in the errors, E_{SC} and E_{RC}, between the extracted line spread profiles and their sparse representations determined using the dictionaries, $\mathbf{D}_{SC}$ and $\mathbf{D}_{RC}$, respectively. The rationale being that $E_{SC} < E_{RC}$ if the image considered was original and $E_{SC} \geq E_{RC}$ if the image was a recaptured image. The value of E_d is determined by taking the differences between E_{SC} and E_{RC}. The second metric, $\bar{\lambda}$, provides a description of the width of an extracted line spread profile. Large values of $\bar{\lambda}$ correspond to blurry edges, while small values to sharp edges.

For each image, j, in the training set of single and recaptured images, I_{SC}^{j} and I_{RC}^{j}, a pair of parameters, $\{E_d^j,\ \bar{\lambda}^j\}_{SC}$ and $\{E_d^j,\ \bar{\lambda}^j\}_{RC}$ respectively, are obtained.

The parameter pairs are collected on an image by image basis and the set of parameter pairs is then used to train a 2-dimensional SVM classifier. When the training procedure is complete a hyperplane that optimally separates the two sets of images based on their values E_d and $\bar{\lambda}$ is determined.

A diagram of the detection stage is shown in Fig. 5. For any given single or recaptured image, a line spread profile matrix, $\mathbf{Q}$, is obtained using the same method that was applied to the training images during classification. The parameters, E_d and $\bar{\lambda}$ are calculated using the trained dictionaries, $\mathbf{D}_{SC}$ and $\mathbf{D}_{RC}$. The parameters are fed to the trained classifier and are classified as single or recaptured based on their location in the E_d, $\bar{\lambda}$ feature coordinate space relative to the SVM hyperplane.

The method for extracting the line spread profile is described in Section IV-B. In Section IV-C we describe the dictionary learning procedure. A detailed description of the line spread width parameter, λ, and of the classifier training and recapture detection procedure are provided in Section IV-D.

B. Automatic Edge Detection and Feature Extraction

The diagram in Fig. 6 illustrates how our proposed algorithm extracts line spread profiles from edges found in the image. Firstly, the query image is converted to greyscale and all edges contained in the image are detected using a Canny Edge Detector [34]. Edge profiles are extracted locally. Therefore, the query image is divided into a number of non-overlapping square blocks $B(m, n)$ of size $W \times W$ with $W = 16$ pixels. Here m and n are the vertical and horizontal indices of the block respectively.

For each block we first check whether it contains a horizontal or near horizontal sharp single edge. We then rotate the block by 90° to see whether it contains vertical or near vertical edges. The block selection procedure is implemented by examining the binary mask of the block and counting the number of columns, η, containing only one non-zero value. The block will be detected only when the condition $\eta \geq \beta W$ is satisfied where β has been set experimentally to $\beta = 0.6$.
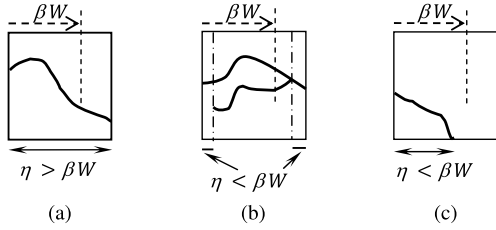
Fig. 7. Examples of blocks with a binary mask of edges that are detected (a) and discarded (b and c). The block in Figure (a) satisfies all our selection criteria. Blocks in Figure (b) and c) do not qualify because the majority of columns in the block shown in Figure (b) contain double edges and in Figure (c) the number of columns containing an edge is less than βW.

An example of a block that meets the selection criteria is shown in Fig. 7a and two examples of blocks that fail to meet the selection criteria are shown in Fig. 7b and Fig. 7c.

The detected blocks, $B(m, n)$, shown in Fig. 6, are then ranked according to their sharpness and edge contrast. This enables us to select regions that are in focus and that contain the most prominent edge features. Block sharpness is determined using the technique described in Section IV-D in which the average width $\bar{\lambda}_{m,n}$ of line spread profiles of the blocks is estimated. The contrast of a block is measured by computing the block-based variance, $\sigma_{m,n}$, of the input image at the detected block. Next, suitable blocks are chosen based on the distributions of $\bar{\lambda}_{m,n}$ and $\sigma_{m,n}$ built over all the detected blocks. Only blocks whose average width, $\bar{\lambda}_{m,n}$, falls within the narrowest 10% of the detected block widths and whose value of $\sigma_{m,n}$ falls within the largest 20% of computed values are selected.

For selected blocks, let $\mathbf{Y} \in \mathbb{R}^{W \times W}$ be a matrix which represents the grey scale values of a block. Each column, $\mathbf{y}_i; i = 1, 2, \cdots, W$, of the matrix $\mathbf{Y}$ may, therefore, be considered to represent an edge profile of the image. We determine a normalized line spread profile, $\mathbf{q}_i$, by evaluating $\mathbf{q}_i = \mathbf{y}_i^{(1)}/\|\mathbf{y}_i^{(1)}\|_2$ where $\mathbf{y}_i^{(1)}$ is the first derivative of $\mathbf{y}_i$. The differentiated edge profile is normalized in order to standardize the feature.

Our feature vector, $\mathbf{q}_i$, now contains the line spread profile at column i of the input block. The spread profile, $\mathbf{q}_i$, is then cropped and centred before zero-padding is applied in order to maintain a length of W elements. Once the line spread profiles for all the selected blocks in the image have been determined, a line spread profile matrix, $\mathbf{Q} \in \mathbb{R}^{W \times M}$ is formed by concatenating the total M line spread profiles, $\mathbf{q}_i$, from all the selected blocks. This feature matrix is used for training and testing purposes.

C. Dictionary Learning Algorithm

The objective of dictionary learning is to obtain two overcomplete dictionaries, $\mathbf{D}_{SC}$ and $\mathbf{D}_{RC}$, that provide an optimal sparse representation of line spread profiles from single captured and recaptured images, respectively. Dictionary training can be used as a tool to learn the characteristics of the distortion patterns present in edges found in most naturally occurring images. The key insight being that the descriptions in single capture and recaptured

images are fundamentally different due to the sharpness degradation introduced by the recapture process.

The first step in dictionary training is to determine the training feature matrices, $\mathbf{S}_{SC}$ and $\mathbf{S}_{RC}$, for single captured and recaptured images, respectively. For each set of training images, $\mathbf{I}_{SC}$ and $\mathbf{I}_{RC}$, the set of line spread profiles, $\mathbf{Q}_{SC}^j$ and $\mathbf{Q}_{RC}^j$, is constructed using the method described in Section IV-B. The superscript, j, denotes the individual images contained in each training set. The training feature matrices, $\mathbf{S}_{SC}$ and $\mathbf{S}_{RC}$, are determined by concatenating horizontally the extracted line spread profiles matrices, $\mathbf{Q}_{SC}^j$ and $\mathbf{Q}_{RC}^j$, over all the training images in each respective set. Thus, the resulting training feature matrix, $\mathbf{S} \in \mathbb{R}^{W \times N}$, contains N training line spread profiles $\mathbf{q}_i \in \mathbb{R}^{W}$, where $i = 1, 2, \cdots, N$, and $N \gg W$.

Given the training feature matrix $\mathbf{S}$, the goal of dictionary training is to obtain the best dictionary, $\mathbf{D} \in \mathbb{R}^{W \times K}$, that provides an optimal sparse representation for all the line spread profiles in the training matrix $\mathbf{S}$, that is

$$\min_{\mathbf{D},\mathbf{X}} \|\mathbf{S} - \mathbf{D}\mathbf{X}\|_F^2 \quad \text{subject to } \forall i, \|\mathbf{x}_i\|_0 \leq L, \quad (12)$$

where $\mathbf{X} \in \mathbb{R}^{K \times N}$ is built from the column vectors $\mathbf{x}_i$ used to represent the feature $\mathbf{q}_i$ and $i = 1, 2, \ldots, N$. The notation $\|\mathbf{A}\|_F^2$ refers to the Frobenius norm, which is defined as $\|\mathbf{A}\|_F^2 = \sum_{ij} |A_{ij}|^2$. The constant L is the maximum number of atoms permitted. The choice of L is generally a trade off between approximation precision and sparsity and we discuss its selection later in this section.

Our dictionary is designed using the K-SVD learning approach [19]. The K-SVD method is an iterative learning scheme based on two important steps for each round of computation: sparse coding and dictionary update.

In sparse coding, given an initial dictionary $\mathbf{D}$, $\mathbf{X}$ is chosen such that each of its columns $\mathbf{x}_i$ provides the best L-sparse representation of $\mathbf{q}_i$. Specifically:

$$\min_{x_i} \|\mathbf{q}_i - \mathbf{D}\mathbf{x}_i\|_2^2 \quad \text{subject to } \|\mathbf{x}_i\|_0 \leq L \quad (13)$$

In practice, this is achieved using the orthogonal matching pursuit (OMP) algorithm [35] which is known to provide near-optimal sparse coding. Next, given $\mathbf{X}$, $\mathbf{D}$ is updated so as to achieve

$$\min_{\mathbf{D}} \|\mathbf{S} - \mathbf{D}\mathbf{X}\|_F^2. \quad (14)$$

In K-SVD, the dictionary atoms are updated, one column at a time, at the kth column index, where $k = 1, 2, \ldots, K$. The residual error in (14) is computed using only the training profiles that use the kth atom for approximation. Next, the atom which minimizes the residual error can be obtained using a singular value decomposition (SVD) approach. We replace the kth column with this new atom. The process is then repeated for all K columns. Given the new $\mathbf{D}$, a new $\mathbf{X}$ is found by sparse coding and the process is repeated. As a result, the training error is reduced over several iterations and the dictionary $\mathbf{D}$ has been trained to fit all training profiles in $\mathbf{S}$.
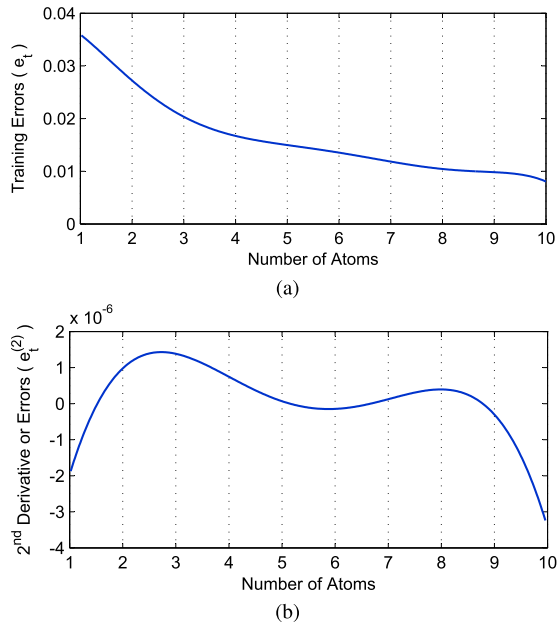
(a)

(b)

Fig. 8. (a) The root-mean-squared error from K-SVD training using single captured images over 160 iterations when the number of atoms used is varied from $L = 1, 2, 3, \ldots, 10$. (b) The optimal number of atoms used is obtained by observing the number of atoms at which the errors begin to converge. This can be estimated from the number of atoms that correspond to the peak of the second derivative of the training error. From our experiment, the optimal value $L = 3$.

By training two dictionaries, $\mathbf{D}_{SC}$ and $\mathbf{D}_{RC}$, using the training feature matrices $\mathbf{S}_{SC}$ and $\mathbf{S}_{RC}$, we ensure that the patterns from line spread profiles extracted from single captured and recaptured images will have been learned. Each dictionary provides an optimal sparse representation of the line spread features from each class of image.

We now discuss the selection of the optimal number of atoms, L, in the dictionary. Since each dictionary was trained using the specific blurring patterns from a given class of images, only one dictionary will provide a good sparse approximation of line spread profiles from the query image. We, therefore, require a value for L that is large enough to provide a good approximation. However, if too many atoms are used, the algorithm is unable to discriminate between the two image classes since both dictionaries are now able to provide good approximations.

To determine the optimal value for L the idea is that the approximation error, $e_t(L)$, decreases with L. However, once the essential information of the signal has been captured, $e_t(L)$ will stop decaying rapidly since the algorithm is now capturing noise and non-discriminative information. This transition point can be detected by finding the peak of the second derivative of e_t. The effect on the training error, e_t, when the number of atoms used for representation is varied is shown in Fig. 8a. The optimal number of atoms is then calculated from the peak of the second derivative of the error function. From Fig. 8b we can determine that the peak of the second derivative for our training sets occurs at approximately $L = 3$.

D. Classification for Recapture Detection

We extract two important feature parameters, $\bar{\lambda}$ and E_d, from the line spread profile matrix $\mathbf{Q}$ which we use
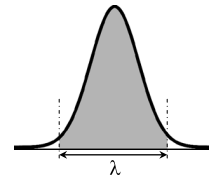

Fig. 9. The criteria for the calculation of the width λ of the spread function. The width is the minimum distance that allows the shape of the edge spread function to be approximated using an estimate of the energy spectral density.

for classification. The first parameter, $\bar{\lambda}$, which models the average line spread profile width, is computed as follows: the value of λ is defined as the distance that allows 95% of the spectral energy of the spread function to be captured and is represented by the grey area shown in the Fig. 9.

The parameter λ_i for a given line spread profile, $\mathbf{q}_i$, is, in practice, computed using the iterative algorithm shown in Fig. 10a. In our experiments, we used a block size $W = 16$ and all the line spread profiles $\mathbf{q}_i$ were interpolated by 4x to increase the number of data points to 64. The spectral energy $E_{\mathbf{q}_i}$ of the given line spread profile $\mathbf{q}_i$ is first computed. Starting from the middle of the spread function, the algorithm then computes the spectral energy E_ω of the part of the spread function with a span $\omega = 1$. The span, ω, is increased by one data point each time that E_ω is computed until the condition $E_\omega/E_{\mathbf{q}_i} > 0.95$ is satisfied. The width, λ_i, of the spread function, $\mathbf{q}_i$, is given by the value of the smallest span, ω, that fulfils the condition.

To determine $\bar{\lambda}$ we compute the spread widths, λ_i, for all the line spread profiles $\mathbf{q}_i \in \mathbb{R}^W$ taken from $\mathbf{Q} \in \mathbb{R}^{W \times M}$ and by then taking the average. Note that a blurred edge generally has a wider spread function compared to a sharp edge. Thus, the value of $\bar{\lambda}$ computed from a recaptured image is expected to be greater than the value obtained from the equivalent single capture image.

The second parameter used is the difference of approximation errors E_d. The value of E_d is used to compare the abilities of the two dictionaries, $\mathbf{D}_{SC}$ and $\mathbf{D}_{RC}$, to provide a sparse representation of line spread profiles from a query image. Given a line spread profile matrix $\mathbf{Q}$ from an unknown image, we define an approximation error using a dictionary trained from single captured images $\mathbf{D}_{SC}$ as $E_{SC} = \|\mathbf{Q} - \mathbf{D}_{SC}\mathbf{X}_1\|_F^2$, where $\mathbf{X}_1$ is the corresponding coefficients matrix computed using the orthogonal matching pursuit algorithm with the dictionary $\mathbf{D}_{SC}$. In the same way, a representation error using a dictionary trained from recaptured images is given by $E_{RC} = \|\mathbf{Q} - \mathbf{D}_{RC}\mathbf{X}_2\|_F^2$.

The approximation errors, E_{SC} and E_{RC}, describe how well each dictionary fits the line spread profile matrix $\mathbf{Q}$. To perform recapture classification we compare E_{SC} with E_{RC}. A query image is classified as recaptured if $E_{SC} \geq E_{RC}$. Otherwise it is considered as single captured. We define the difference of approximation errors (E_d) as follows:

$$E_d = E_{SC} - E_{RC}$$
$$= \|\mathbf{Q} - \mathbf{D}_{SC}\mathbf{X}_1\|_F^2 - \|\mathbf{Q} - \mathbf{D}_{RC}\mathbf{X}_2\|_F^2, \quad (15)$$

Fig. 10b summarizes how we compute the feature E_d given the trained dictionaries $\mathbf{D}_{SC}$, $\mathbf{D}_{RC}$, and a query image $\mathbf{I}_Q$.
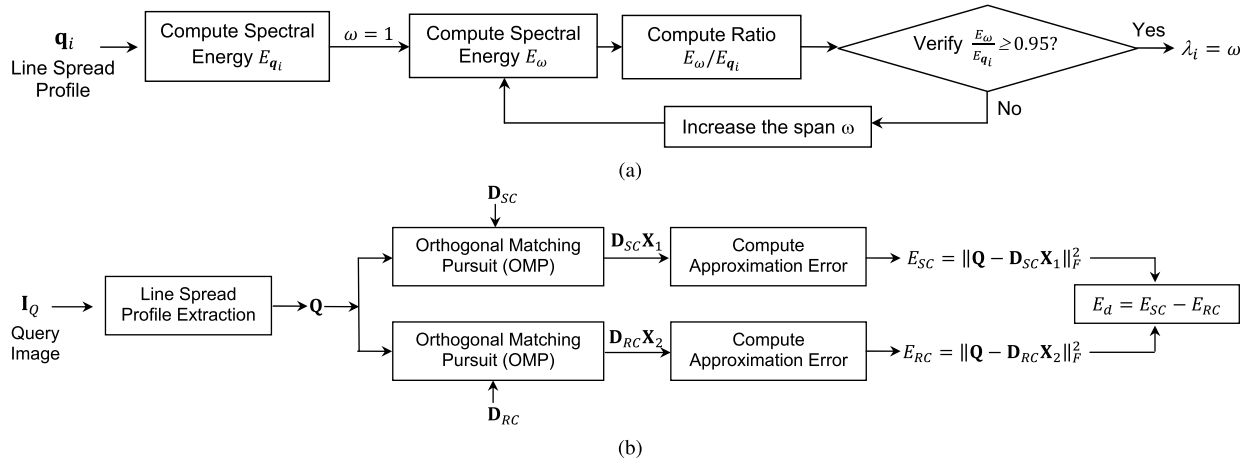
(a)

(b)

Fig. 10. Diagrams showing how to compute the features (a) λ_i given a line spread profile $\mathbf{q}_i$ and (b) E_d given the trained dictionaries $\mathbf{D}_{SC}$, $\mathbf{D}_{RC}$, and a query image $\mathbf{I}_Q$.
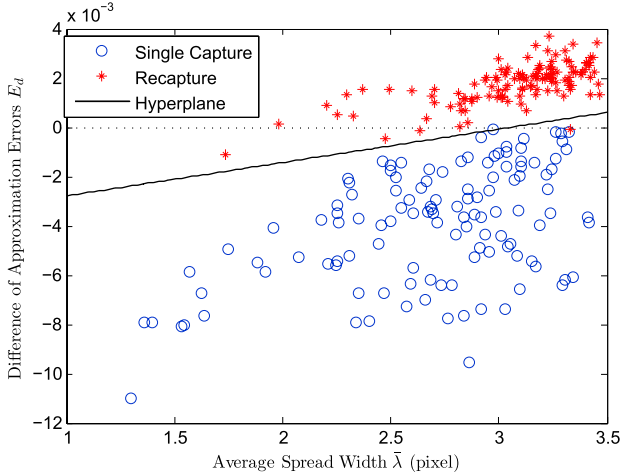

Fig. 11. A plot of the distributions of features extracted from training images with the average width of spread function ($\bar{\lambda}$) shown on the horizontal axis and the difference of representation error (E_d) on the vertical axis. The hyperplane for recapture classification was obtained using SVM training and is defined as the line that separate the features from the recaptured (star) and single captured (circle) images with minimum classification error. A query image is classified based on the coordinate location of the feature pair, $\bar{\lambda}, E_d$, determined from the image.

One way that Equation (15) may be interpreted is that the image is more likely to be single captured if E_d is negative. If E_d is positive the image is more likely to be recaptured.

A set of pairs of parameters, $\bar{\lambda}$ and E_d, extracted from each image in the training set is formed. The parameter pairs are then labelled as either single or recaptured, depending on the class of training images used. They are then used for classifier training. Fig. 11 is a plot in the feature coordinate space, $(\bar{\lambda}, E_d)$, of the parameter pairs from all the images used for training purposes. A support vector machine (SVM) classifier was trained and the classification hyperplane (solid line in Fig. 11) was generated.

From Fig. 11, we can observe that the difference in representation error, (E_d), can be used as a feature to effectively distinguish between single and recaptured images. The majority of images in the single and recaptured groups were

separated correctly by the criterion $E_d = 0$ (dotted line). However, the hyperplane obtained from the SVM training process (solid line) resulted in better classifier performance since the mean width of the edge spread function, $\bar{\lambda}$, was taken into account. We have assumed that the distribution of features from the training images is typical of the type of images we are likely to encounter on a daily basis. The trained classifier is, therefore, used for our recapture detection algorithm.

We note that other no-reference blur metrics that operate on the whole image [36], [37] could be used instead of $\bar{\lambda}$. However, we have observed through numerical simulations that these metrics are less effective at classifying single and recaptured images than $\bar{\lambda}$, especially when combined with other discriminative features such as E_d. For this reason we have decided to use $\bar{\lambda}$ in our work.

V. DATABASE OF RECAPTURED IMAGES

A database [38] of images recaptured from an LCD monitor was developed for the purposes of testing and evaluating the performance of the recapture detection algorithm described in Section IV. The recapture database comprised 1035 single capture images taken using nine different cameras. Each camera was used to capture 115 images. Out of each set of 115 images, 35 images contained scenes that were common over all nine cameras. Thus, the total number of images containing common scenes was 315. Each image in the set of common single captured images was then recaptured using eight different cameras. This resulted in a total of 2520 recaptured images. The database has been made publicly available in order that it can be used as a common database for researchers in the field of image forensics who wish to benchmark their algorithms. Currently available image databases include the 'Dresden Image Database' [39]. This is probably the most well known still image database for forensic applications, but it does not include any recaptured images.

A. Image Capture and Display Equipment

Nine cameras were used to photograph the original scenes and eight to recapture the images from the LCD monitor.

TABLE I
A TABLE OF THE DIGITAL CAMERAS USED
IN THE RECAPTURE DATABASE

Camera make and model	Year	MPixels	Original capture	Recapture
Kodak V550 (silver)	2005	5	Yes	No
Kodak V550 (black)	2005	5	Yes	No
Kodak V610	2006	6	Yes	No
Nikon D40	2006	6	Yes	No
Panasonic TZ10	2010	12.1	No	Yes
Nikon D3200	2012	24.2	No	Yes
Canon 60D	2012	18	No	Yes
Nikon D70s	2004	6	Yes	Yes
Panasonic TZ7	2009	10	Yes	Yes
Canon 600D	2012	18	Yes	Yes
Olympus E-PM2	2012	16.1	Yes	Yes
Sony RX100	2012	20.2	Yes	Yes

Five of the cameras used to carry out the original captures were used for recapture thereby resulting in a total of twelve cameras in the database. A list of cameras used, their specifications, and usage is shown in Table I. They include six compact digital cameras with fixed zoom lenses, five digital single lens reflex (DSLR) cameras with interchangeable lenses and one compact camera with interchangeable lenses. With the exception of the three Kodak cameras and the Panasonic TZ7, all cameras provided both automatic and manual exposure settings. The two Kodak V550 models are equivalent in specification and differ only in their finish. They are indicated as silver and black in Table I. All the images were recaptured from an NEC MultiSync EA232WMi 23″ IPS LCD monitor with LED backlighting and a resolution of 1920×1080 pixels.

B. Original Scene Capture

The database comprises mainly natural scenes photographed indoors and outdoors under different types and levels of illumination. Some examples of originally captured scenes (top row) and the recaptured images (bottom row) are shown in Fig. 12. A significant proportion of the images were taken outdoors under sunny or overcast conditions. Those taken indoors were acquired mostly under natural illumination, but also included a scene with a MacBeth Colorchecker test chart captured under natural illumination and using the camera's internal flash, where available. Each scene in the database was photographed once by each of the test cameras under equivalent, or nearly equivalent, illumination conditions. This allowed for a one to one correspondence between a scene and each test camera. All the cameras were set to automatically select the exposure setting, ISO and white balance setting, with the exception of the Macbeth test chart scene where different ISO settings were selected. The database contains 115 images per camera giving a total of 1035 single captured images over all 9 cameras.

C. Recapture

A high priority when developing the recapture database was that the recaptured images would be high in perceived quality and finely recaptured. All image recaptures were conducted in a darkened room, to eliminate unwanted reflections from the monitor and the surrounding environment. The single captured images were prepared for display by resizing them using a bicubic interpolation kernel to the pixel dimensions of the NEC monitor. The images were then displayed at the native resolution of the monitor. The camera used to recapture the images from the LCD monitor was placed on a sturdy tripod. Before recapturing the images, the LCD monitor was calibrated to the sRGB standard with $\gamma = 2.2$ and a monitor white point luminance of 240 cd/m^2. The lens focal length of each camera was set to a value that minimised the level of geometric distortion introduced in the recaptured image as much as was practically possible.

The monitor to camera distance was determined by applying the alias frequency maximisation method described in III-B. The procedure for determining the capture distance and lens aperture setting is described with the aid of the following example in which a Canon 600D camera was used to recapture images from the NEC monitor. The image sensor in the Canon 600D camera has a pixel pitch of 4.306 52 μm. The pixel pitch of the NEC monitor is 0.2650 mm. For a lens focal length of 30 mm, Equation (8) was used to obtain a capture distance of 1445.2 mm. To eliminate visible aliasing from the recaptured image due to aliasing in the green channel, Equation (11) was used to obtain an aperture setting of $f/11$.

After setting the camera to monitor distance, the camera's image sensor was aligned with the plane of the monitor faceplate. The camera's ISO setting was manually set to a value that did not introduce excessive levels of image noise in the recaptured images, and the camera was allowed to select the exposure automatically. To eliminate colour balance errors by the recapture camera, the camera's white point was preset by estimating it from a white patch displayed on the monitor. The recaptured images were cropped to remove the LCD monitor surround.

We conducted a psychophysical experiment to study how well human observers were able to discriminate between original captured images and images recaptured from an LCD monitor using our method. A total of 100 images, comprising 50 originally captured and 50 recaptured images, were selected at random from our dataset for use in the subjective test.

A website was created that enabled observers to login and conduct the test online. All the images used in the subjective test were resized to the same dimensions and their Exif metadata removed. A total of 40 people differing in age and career background participated in the test. During the test, observers were shown the set of 100 images in random order and were asked to classify each image as originally captured or recaptured. The tests were unsupervised and no time limit was set on completing the test. However, observers were instructed to maintain a fixed viewing distance throughout the test.

A confusion matrix detailing the statistics obtained from the subjective test is shown in Table II. The test results show that the average success rate for observers in classifying the images was only 61.31%. This level of precision is fairly close to the accuracy that would be obtained in a random choice, binary classification experiment. In addition, the results
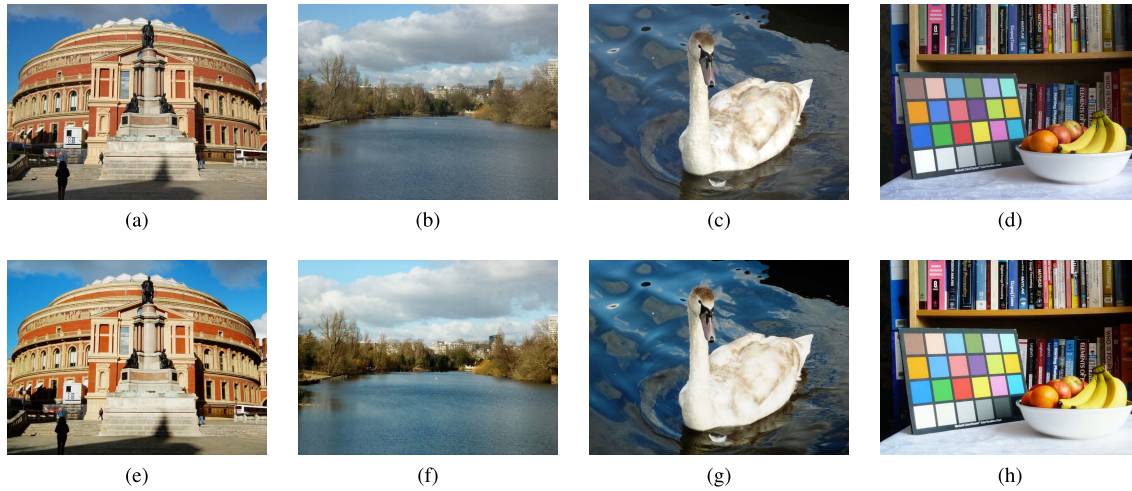
Fig. 12.   Images from the recapture database showing examples of originally captured and recaptured scenes. (a) Architectural scene. (b) Natural scene. (c) Wildlife scene. (d) Indoor scene. (e) Recaptured architectural scene. (f) Recaptured natural scene. (g) Recaptured wildlife scene. (h) Recaptured indoor scene.

TABLE II

THE SUCCESS RATES WHEN OBSERVERS WERE ASKED
TO DETECT THE RECAPTURED IMAGES BASED ON
THE PERCEIVED QUALITY OF THE IMAGE

Image Type	No. of images	Subjective Classification Statistics		
		Original Captured	Recaptured	Success Rate (%)
Original Captured	50	33.87	16.13	67.75
Recaptured	50	22.57	27.43	54.86
Overall	100			61.31

suggest that observers have a greater tendency to classify recaptured images as originally captured. This is highlighted by the differences in success rates between originally captured images (67.75%) and recaptured images (54.86%). Most importantly, it supports our claim that to detect recaptured images in our dataset by visual inspection is a difficult task.

VI. EXPERIMENTAL RESULTS

In this section the performance evaluation of the proposed recapture detection algorithm is presented. The experiments have been divided into three sections: A) experimental design and data training, B) recapture detection performance and C) universality of the algorithm.

A. Experimental Design and Parameter Setting

Our algorithm was tested by applying it to the set of single and recaptured images from the database described in Section V. The single and recaptured images were initially divided into two groups; one for training and the other for testing purposes. Each set of 115 single capture images from the nine cameras used to capture the original scenes was partitioned into 15 images for training and 100 images for performance evaluation. This resulted in a total of 135 images for training and 900 images for testing. As mentioned in Section V, 35 original captured images, per camera,

were selected for image recapture. The recaptured images in the database were taken over 72 different single/recapture camera combinations resulting in a total of 2520 images. These were divided into 1080 images for training and 1440 images for testing. The set of training images was further reduced to 216 images by randomly selecting three images from each recapture camera combination. This was done to eliminate any bias introduced during training due to large differences in the number of single and recaptured images. All the images were resized to a width of 2048 pixels because of differences in the size of the images resulting from variations in the camera image sensor pixel counts. The image height during resizing was set to a value that preserved the original aspect ratio of the image.

The K-SVD software library that is publicly provided by the authors [19] was used to train our dictionaries. The training feature matrices $\mathbf{S}_{SC}$ and $\mathbf{S}_{RC}$ were built from the edge spread features $\mathbf{q_i}$ with length $W = 16$ collected using the blocks selected from single captured and recaptured images respectively. One in every four spread features were used in order to reduce the size of the training data. Before being used for K-SVD dictionary training each spread feature was interpolated by $4\times$ to increase the number of data points to 64. The initial set of atoms was constructed from the line spread functions of the nine single capture cameras and 63 different line spread functions determined from randomly selected image recapture camera combinations. The dictionaries were trained over 160 iterations of the K-SVD algorithm at which point the errors converged to a predetermined minimum value.

B. Performance Evaluation for Recapture Detection

The algorithm was tested with the test image set described in Section VI-A and the results are shown in Table III. In the top half of the table the results are presented according to the cameras used for single capture. A total of 854 out of the 900 single captured images were correctly classified corresponding to a 94.89% true negative rate (TN). In the

TABLE III

RECAPTURE DETECTION PERFORMANCE OF THE ALGORITHM
ACCORDING TO THE CAMERA USED FOR ORIGINAL
CAPTURE OR RECAPTURE

Camera Model	No. of images	Classification Results		Performance (%)
		Single	Recapture	
Single Capture Database				
V550 (Black)	100	86	14	86
V550 (Silver)	100	83	17	83
V610	100	94	6	94
D40	100	97	3	97
D70s	100	97	3	97
EOS 600D	100	99	1	99
RX100	100	100	0	100
E-PM2	100	99	1	99
TZ7	100	99	1	99
Average				**94.89**
Recapture Database				
D3200	180	0	180	100
D70s	180	1	179	99.44
EOS 600D	180	2	178	98.89
EOS 60D	180	1	179	99.44
RX100	180	0	180	100
E-PM2	180	7	173	96.11
TZ7	180	0	180	100
TZ10	180	1	179	99.44
Average				**99.03**
Overall Accuracy	2340			**97.44**

TABLE IV

THE COMPARISON OF PERFORMANCE OF THE ALGORITHMS
IN CLASSIFYING ORIGINAL AND RECAPTURED IMAGES

Method	Number of Features	Success Rate (%)	
		Original Captured	Recaptured
MSWS+LBP+Colour Features [8]	129	83.67	92.02
Higher-order Wavelet Statistics [41]	216	87.56	90.04
Proposed Method	2	94.89	99.03

support vector machine (PSVM) trained using a combination of 3 types of features including Local Binary Feature (LBP), Multi-Scale Wavelet Statistics (MSWS), and Color Features. We implemented the algorithm based on the paper in [8] and used this to extract a total of 129 features from a given image including 54 LBP features, 54 MSWS features and 21 colour features. The features extracted from the training images were used to train the PSVM classifier using the LIBSVM tools following the technical guideline in [40].

In addition, we tested a recapture detection algorithm based on features extracted from Higher-order Wavelet Statistics [41], [42] as this algorithm is often used as a benchmark when comparing the performance of recapture image detectors. We extracted 216 features using the software provided by the authors and used the features to train a PSVM classifier. The results for performance comparison based on the same set of test images used in Section VI-B are presented in Table IV.

We then compared the classification success rate for original captured and recaptured images using the different methods. Our results show that the proposed method outperforms the existing techniques in classifying original and recaptured images when tested with our dataset. Our technique has a higher true positive (TP) rate of 99.03% compared to a TP rate of 92.02% for the method of Cao and Cot [8] and 90.04% for Higher-order Wavelet Statistics [41].

The detection performance on originally captured images of the proposed method is 94.89%. This is higher than the detection performance of methods [8] and [41] which are 83.67% and 87.56% respectively. In addition our method uses only 2 features extracted from each given image compared to 129 and 216 features used in the methods [8] and [41], respectively. Our method, however, requires block-based edge extraction and dictionary learning to obtain a smaller number of features.

As mentioned in Section V-C, care was taken to ensure that the recaptured images in the database did not contain any visible aliasing and were of high perceived quality. Furthermore, the tiny patterns formed by the surface texture of the LCD screen were, in general, not easily detected by visual inspection. The methods in [8] and [41] both rely on the presence of these patterns for recapture detection and it is likely that the performance of these methods was reduced as a result. In particular, the methods that use LBP features were likely to be affected the most. The absence of

bottom half of the table the results are presented in groups according to the camera used to recapture the images. Performance is consistently good, spanning from a 96.11% true positive (TP) rate for the E-PM2 to a 100% success rate for the Nikon D3200, Sony RX100 and Panasonic TZ7 models. This leads to an overall accuracy of 97.44%.

It is interesting to note the drop in detection rate for single captured images that were originally captured using the Kodak cameras. One possible reason is that the images captured using these cameras are, on average, less sharp than the images obtained by the other 'single capture' cameras. This may be due to the fact that they are all budget compact cameras that, with the exception of the Nikon D70s DSLR camera, are production models released three or more years earlier than the other cameras used in this test (see Table I). The newer cameras in the set are likely to have benefited from advances in image sensor technologies and lens design enabling them to generate significantly sharper images than earlier camera models.

C. Performance Comparison With Existing Methods

Next we compare the performance of the proposed algorithm with state-of-the-art techniques for detecting recaptured images using our alias-free recapture dataset. Our comparison was conducted primarily against the method of Cao and Kot [8] because it is one of the earliest and most referenced techniques for detecting recaptured images from LCD monitors. It has also been tested on a set of finely recaptured images and is, therefore, most comparable to our method. Classification relies on a probabilistic

TABLE V

GROUP CONFIGURATION FOR IMAGES USED TO TEST
THE UNIVERSALITY OF THE ALGORITHM

Image Set	Single Capture Camera	Recapture Camera
Training	Kodak V550 Black Case Kodak V610 Nikon D40 Canon EOS 600D	Nikon D3200 Canon EOS 600D Olympus - E-PM2 Panasonic TZ10
Testing	Kodak V550 Silver Case Nikon D70s Olympus - E-PM2 SONY RX100 Panasonic TZ7	NikonD70s Canon EOS 60D SONY RX100 Panasonic TZ7

these features did not affect the performance of the proposed algorithm.

Many images in the database contain a wide range of textures and edges that vary in sharpness and contrast. For example, an image may contain regions that are low in contrast or out of focus. The method in [8] computes texture features while in [41] the statistics of wavelet coefficients are calculated globally. These methods are likely to perform well when the image contains a high level of detail, however, their detection performance will be reduced if the scene is dominated by blurred or low contrast regions. The proposed algorithm, on the other hand, discards regions containing out of focus or low contrast edges and selects only the regions containing the sharpest edges. Images containing large variations in texture and edge sharpness are therefore less likely to impact the proposed algorithm.

D. Universality of the Algorithm

In the previous section we assumed that the edge profiles of a wide range of image capture devices and recapture chains are known. In practice, the range of devices and image chains that are characterised may be limited. In this section, the performance of the algorithm was tested when applied to images acquired from unknown capture devices. It is anticipated that, in practice, the edge profiles of single and recaptured images are fundamentally different, and that recapture detection is possible even when a small range of devices and chains representative of the population has been characterised. An experiment was devised in which the algorithm was trained with single and recaptured images taken with a different set of cameras to that used to capture the images used for testing. Table V illustrates the configuration of the training and testing image sets. Note that the cameras used for single capture and recapture differ for the training and testing sets. A total of 60 single captured images and 108 recaptured images were used for training. For testing purposes, 500 single captured images and 720 recaptured images were used.

The new hyperplane resulting from the training of the K-SVD dictionary and SVM, was used for performance evaluation. The recapture detection results are shown in Table VI. A recapture detection rate, or true positive (TP) rate, of 99.31% was obtained over all cameras. For the detection of single captured images, a true negative (TN) rate of 93% was obtained. This results in an overall accuracy of 96.72%.

TABLE VI

RECAPTURE DETECTION PERFORMANCE OF THE
ALGORITHM FOR UNIVERSALITY TESTING

Camera Model	No. of images	Classification Results		Performance (%)
		Single	Recapture	
Single Capture Database				
V550 Silver	100	76	24	76
D70s	100	93	7	93
E-PM2	100	99	1	99
RX100	100	100	0	100
TZ7	100	97	3	97
Average				**93.00**
Recapture Database				
D70s	180	2	178	98.89
EOS 60D	180	1	179	99.44
RX100	180	0	180	100
TZ7	180	2	178	98.89
Average				**99.31**
Overall Accuracy	1220			**96.72**

We can conclude, therefore, that the overall performance of the algorithm has been maintained despite using images from different cameras for training and testing.

The experimental work conducted in this section has shown that the proposed algorithm is robust to a wide range of images and can be applied successfully to a wide range of capture devices. Furthermore, it is able to cope with query images from unknown sources and image chains.

VII. CONCLUSION

In this paper the problem of detecting images recaptured from LCD monitors has been addressed. A comprehensive overview of the most commonly encountered features in recaptured images concluded that aliasing and blurriness can be extracted most reliably from the image. We then showed how aliasing can be eliminated by properly setting the capture distance, recapture camera lens aperture and focal length. Our approach then focused on a solution for recapture detection based on edge blurriness and distortion. The line spread functions of selected edges were used to train single capture and recapture dictionaries following the K-SVD approach. An SVM classifier was then built to detect recaptured images using dictionary approximation errors and mean edge spread width. Our algorithm, which does not require user intervention, achieved an average recapture detection performance that exceeded 99% for recaptured images and 94% for single capture images. Finally, we have developed a database comprising more than 2500 high quality recaptured images using the method described in this paper. The database will be made publicly available for research purposes and we anticipate that it will form a valuable resource for the development and benchmarking of forensic methods for recapture image detection.

REFERENCES

[1] T. Thongkamwitoon, H. Muammar, and P. L. Dragotti, "Identification of image acquisition chains using a dictionary of edge profiles," in *Proc. 20th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2012, pp. 1757–1761.

[2] H. Muammar and P. L. Dragotti, "An investigation into aliasing in images recaptured from an LCD monitor using a digital camera," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Vancouver, BC, Canada, May 2013, pp. 2242–2246.

[3] H. Farid, "Seeing is not believing," *IEEE Spectr.*, vol. 46, no. 8, pp. 44–51, Aug. 2009.

[4] H. Yu, T.-T. Ng, and Q. Sun, "Recaptured photo detection using specularity distribution," in *Proc. 15th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2008, pp. 3140–3143.

[5] J. Yin and Y. Fang, "Markov-based image forensics for photographic copying from printed picture," in *Proc. 20th ACM Int. Conf. Multimedia (MM)*, Nara, Japan, 2012, pp. 1113–1116.

[6] J. Bai, T.-T. Ng, X. Gao, and Y.-Q. Shi, "Is physics-based liveness detection truly possible with a single image?" in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Jun. 2010, pp. 3425–3428.

[7] X. Gao, T.-T. Ng, B. Qiu, and S.-F. Chang, "Single-view recaptured image detection based on physics-based features," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2010, pp. 1469–1474.

[8] H. Cao and A. C. Kot, "Identification of recaptured photographs on LCD screens," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2010, pp. 1790–1793.

[9] J. Yin and Y. Fang, "Digital image forensics for photographic copying," *Proc. SPIE*, vol. 8303, pp. 1–7, Feb. 2012, Art. ID 83030F.

[10] B. Li, Y. Q. Shi, and J. Huang, "Detecting doubly compressed JPEG images by using mode based first digit features," in *Proc. IEEE 10th Workshop Multimedia Signal Process.*, Oct. 2008, pp. 730–735.

[11] Y. Ke, Q. Shan, F. Qin, and W. Min, "Image recapture detection using multiple features," *Int. J. Multimedia Ubiquitous Eng.*, vol. 8, no. 5, pp. 71–82, 2013.

[12] X. Gao, B. Qiu, J. Shen, T.-T. Ng, and Y. Q. Shi, "A smart phone image database for single image recapture detection," in *Proc. 9th Int. Workshop Digit. Watermarking (IWDW) Revised Select. Papers*, 2011, pp. 90–104.

[13] T.-T. Ng, S.-F. Chang, J. Hsu, L. Xie, and M.-P. Tsui, "Physics-motivated features for distinguishing photographic images and computer graphics," in *Proc. 13th Annu. ACM Int. Conf. Multimedia*, 2005, pp. 239–248.

[14] N. Kose and J. Dugelay, "Classification of captured and recaptured images to detect photograph spoofing," in *Proc. Int. Conf. Informat., Electron. Vis. (ICIEV)*, May 2012, pp. 1027–1032.

[15] A. da Silva Pinto, H. Pedrini, W. Schwartz, and A. Rocha, "Video-based face spoofing detection through visual rhythm analysis," in *Proc. 25th SIBGRAPI Conf. Graph., Patterns Images (SIBGRAPI)*, Aug. 2012, pp. 221–228.

[16] J.-W. Lee, M.-J. Lee, T.-W. Oh, S.-J. Ryu, and H.-K. Lee, "Screenshot identification using combing artifact from interlaced video," in *Proc. 12th ACM Workshop Multimedia Secur. (MM&Sec)*, Rome, Italy, 2010, pp. 49–54.

[17] P. Bestagini, M. Visentini-Scarzanella, M. Tagliasacchi, P. L. Dragotti, and S. Tubaro, "Video recapture detection based on ghosting artifact analysis," in *Proc. 20th IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2013, pp. 4457–4461.

[18] M. Visentini-Scarzanella and P. L. Dragotti, "Video jitter analysis for automatic bootleg detection," in *Proc. IEEE 14th Int. Workshop Multimedia Signal Process. (MMSP)*, Sep. 2012, pp. 101–106.

[19] M. Aharon, M. Elad, and A. Bruckstein, "*K*-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[20] R. Palum, "Optical antialiasing filters," in *Single-Sensor Imaging: Methods and Applications for Digital Cameras* (Image Processing Series), R. Lukac, Ed., 1st ed. Boca Raton, FL, USA: CRC Press, Sep. 2008, ch. 4, pp. 105–135.

[21] J. Adams, K. Parulski, and K. Spaulding, "Color processing in digital cameras," *IEEE Micro*, vol. 18, no. 6, pp. 20–30, Nov./Dec. 1998.

[22] T. Koyama, "Optics in digital still cameras," in *Image Sensors and Signal Processing for Digital Still Cameras*, J. Nakamura, Ed. Boca Raton, FL, USA: CRC Press, 2006, pp. 21–51.

[23] I. A. Cunningham and A. Fenster, "A method for modulation transfer function determination from edge profiles with correction for finite-element differentiation," *Med. Phys.*, vol. 14, no. 4, pp. 533–537, 1987.

[24] *Photography—Electronic Still Picture Imaging—Resolution and Spatial Frequency Responses*, ISO Standard 12233:2014, 2014.

[25] M. Goljan, J. Fridrich, and T. Filler, "Large scale test of sensor fingerprint camera identification," *Proc. SPIE*, vol. 7254, pp. 1–12, Feb. 2009, Art. ID 72540I.

[26] M. Goljan, and J. Fridrich, "Camera identification from cropped and scaled images," *Proc. SPIE*, vol. 6819, p. 68190E, Feb. 2008.

[27] *Multimedia Systems and Equipment—Colour Measurement and Management—Part 2-1: Colour Management—Default RGB Colour Space—sRGB*, document IEC 61966-2-1:1999, 1999.

[28] T.-T. Ng and M.-P. Tsui, "Camera response function signature for digital forensics—Part I: Theory and data selection," in *Proc. 1st IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2009, pp. 156–160.

[29] T.-T. Ng, "Camera response function signature for digital forensics—Part II: Signature extraction," in *Proc. 1st IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2009, pp. 161–165.

[30] T. Hatada and F. Saitoh, "Moiré reduction method for LCD captured image by using two different focused images," *IEEJ Trans. Electron., Inf. Syst.*, vol. 128, no. 2, pp. 326–327, 2008.

[31] J. Mochizuki and T. Asano, "Study on reduction methods of moiré noise in image acquisition of periodic gray-scale pattern," *Syst. Comput. Jpn.*, vol. 33, no. 1, pp. 27–37, Jan. 2002.

[32] J. C. Krumm and S. A. Shafer, "Sampled-grating and crossed-grating models of moire patterns from digital imaging," *Opt. Eng.*, vol. 30, no. 2, pp. 195–206, Feb. 1991.

[33] W. J. Smith, *Modern Optical Engineering: The Design of Optical Systems*, 4th ed. New York, NY, USA: McGraw-Hill, 2008.

[34] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[35] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Annu. Asilomar Conf. Signals, Syst., Comput.*, Nov. 1993, pp. 40–44.

[36] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, Apr. 2009.

[37] N. D. Narvekar and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2678–2683, Sep. 2011.

[38] *Recapture Image Database*. [Online]. Available: http://www.commsp.ee.ic.ac.uk/~pld/research/Rewind/Recapture/, accessed Oct. 24, 2014.

[39] T. Gloe and R. Böhme, "The 'Dresden image database' for benchmarking digital image forensics," in *Proc. ACM Symp. Appl. Comput.*, 2010, pp. 1584–1590.

[40] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," Dept. Comput. Sci., National Taiwan Univ., Taipei, Taiwan, Tech. Rep., 2003. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvmtools/

[41] H. Farid and S. Lyu, "Higher-order wavelet statistics and their application to digital forensics," in *Proc. IEEE Workshop Statist. Anal. Comput. Vis.*, Jun. 2003, p. 94.

[42] S. Lyu and H. Farid, "How realistic is photorealistic?" *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 845–850, Feb. 2005.

Thirapiroon Thongkamwitoon received the Ph.D. degree in electrical and electronic engineering from Imperial College London, London, U.K., in 2014. While pursuing the Ph.D. degree, he was a Student Researcher on the Reverse Engineering of Audio-Visual Content Data project granted by the Future and Emerging Technologies Program of the European Commission from 2011 to 2013.

He received the bachelor's and master's degrees in electrical engineering from Chulalongkorn University, Bangkok, Thailand, in 2002 and 2005, respectively. He was with Aeronautical Radio of Thailand Ltd., Bangkok, from 2005 to 2009, before he was awarded the Ph.D. Scholarship from the Royal Thai Government. He is currently a Senior Engineer with the Office of the National Broadcasting and Telecommunications Commission, Bangkok, Thailand.

Dr. Thongkamwitoon was a recipient of the Educational Award for Excellence from the Anglo-Thai Society in 2013, and was nominated for the best presentation award from the IEEE Italy Chapter Summer School on Signal Processing. His research interests include image forensics, image processing, and sampling theory.

Hani Muammar (M'07) received the B.Sc. (Hons.) degree in electrical and electronic engineering from University College, Cardiff, U.K., in 1984, the M.Sc. degree in digital electronics from King's College, London, U.K., in 1986, and the Ph.D. degree in image processing from the University of Southampton, Southampton, U.K., in 1990.

He was a Senior Imaging Scientist with Kodak European Research and Development Laboratories, Kodak Ltd., Harrow, U.K., from 1990 to 2005, and a Senior (Lead) Imaging Scientist with Kodak European Research, Kodak Ltd., Cambridge, U.K., from 2006 to 2009. He has been a Consultant for the consumer photographic imaging industry, and was a Senior Lecturer in Imaging Science with the University of Westminster, London, from 2010 to 2011. From 2011 to 2014, he was a Research Associate with Imperial College London, London. His research interests include digital image acquisition path modeling and simulation, image enhancement algorithm development, and image quality assessment and metrics.

Dr. Muammar has been the U.K. Expert to ISO/TC42 WG18, Electronic Still Picture Imaging, since 2003, and is a Project Leader for ISO 15739, Photography-Electronic Still-Picture Imaging-Noise Measurements. He is a member of the Imaging Science Group Committee of the Royal Photographic Society and the Institution of Engineering and Technology.

Pier-Luigi Dragotti is currently a Professor of Signal Processing with the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K. He received the Laurea degree (*summa cum laude*) in electrical and electronic engineering from the University of Naples Federico II, Naples, Italy, in 1997, the master's degree in communications systems from the Swiss Federal Institute of Technology of Lausanne (EPFL), Lausanne, Switzerland, in 1998, and the Ph.D. degree from EPFL, in 2002. He has held several visiting positions at different universities and research centers. He was a Visiting Student with Stanford University, Stanford, CA, USA, in 1996, a Researcher with the Mathematics of Communications Department, Bell Labs, Lucent Technologies, Murray Hill, NJ, USA, in 2000, and a Visiting Scientist with the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2011. He was the Technical Cochair of the European Signal Processing Conference in 2012, an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2006 to 2009, and an elected member of the IEEE Image, Video and Multidimensional Signal Processing Technical Committee from 2008 to 2013. He is currently an elected member of the IEEE Signal Processing Theory and Method Technical Committee. He was a recipient of the ERC Starting Investigator Award for the project RecoSamp. His work includes sampling theory, wavelet theory and its applications, image and video compression, image-based rendering, and image super-resolution.