

Forgery Authentication in Extreme Wide-Angle Lens Using Distortion Cue and Fake Saliency Map

Huazhu Fu and Xiaochun Cao

Abstract—Distortion is often considered as an unfavorable factor in most image analysis. However, it is undeniable that the distortion reflects the intrinsic property of the lens, especially, the extreme wide-angle lens, which has a significant distortion. In this paper, we discuss how explicitly employing the distortion cues can detect the forgery object in distortion image and make the following contributions: 1) a radial distortion projection model is adopted to simplify the traditional captured ray-based models, where the straight world line is projected into a great circle on the viewing sphere; 2) two bottom-up cues based on distortion constraint are provided to discriminate the authentication of the line in the image; 3) a fake saliency map is used to maximum fake detection density, and based on the fake saliency map, an energy function is provided to achieve the pixel-level forgery object via graph cut. Experimental results on simulated data and real images demonstrate the performances of our method.

Index Terms—Distortion constraint, fake saliency map, forgery detection, lens distortion, wide-angle lens.

I. INTRODUCTION

GENERALLY, distortion is considered as a significant problem in many computer vision applications [1]–[5]. While a number of methods have been developed to correct the distortion for striving extreme accuracy [6]–[9], little work takes into account that the distortion is an important internal property of the lens. Especially for the extreme wide-angle lens (or omnidirectional lens), which involves a significant distortion, the distortion could provide an useful geometrical constraint overall the image. In this paper, we show that the distortion cues could be used to detect the photographic composites in a single image of a wide-angle lens.

Thanks to the extreme field of view (FOV), the wide-angle lens is widely used to monitor the indoor/outdoor environment. Sometimes the surveillance video is able to play an important part for confirmation of responsibility of some unusual event. Hence, the determination the authentication of monitoring image (the frame of video) is essential for avoiding perjury. Furthermore, with the advancement of technology, a number of high quality, but cheap and widely available, wide-angle

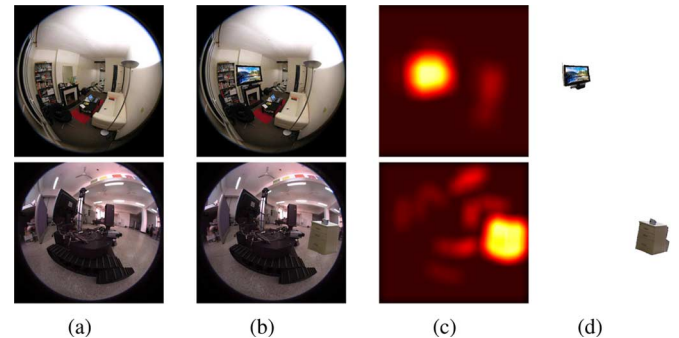


Fig. 1. Forgery detection in fisheye image using our method. (a) Original fisheye images. (b) Forged images. (c) Fake saliency map of our method. (d) Pixel-level forgery segmentation results.

lenses have appeared. More and more people are fond of using the wide-angle lens to capture photos/video [10], [11]. Same as the general image, the wide-angle image is also easily manipulated and altered by amateurs using the powerful digital image processing tools, that indicates serious vulnerabilities and reduced credibility of the digital images [12]. Fig. 1(b) exemplifies two forged images captured by fisheye lens, which is a typical wide-angle lens. The forged objects are manipulated by copying and pasting into the original images of Fig. 1(a).

In the past few years, the field of forgery detection has emerged to authenticate digital images using various constraints [13]: 1) pixel-based techniques [14] that have been proposed to detect cloned, resampled, spliced images; 2) format-based techniques [15], [16] that detect tampering in lossy image compression; 3) camera-based techniques [17]–[21] that exploit artifacts introduced by the camera lens, sensor or on-chip postprocessing; 4) physically based techniques [22], [23] that model and detect anomalies using physical rules; and 5) geometric-based techniques [24], [25] that make use of geometric constraints which are recovered from perspective views. The proposed approach of this paper falls into the fifth category, i.e., geometric techniques. As mentioned previously, the strong radial distortions of a wide-angle lens is as a double-edged sword. On one hand, the regular geometric-based techniques for a general lens are not suitable for the wide-angle lens any more. On the other hand, the radial distortion also provides some new geometric constraints, which could be used as the bottom-up cues for the image authentication. In this paper, the forgery object is assumed to contain the straight lines in the space, that presents commonly in the real world. We discuss how to detect the forgery lines on the object using the distortion constraint. Furthermore, we also propose a fake saliency map to segment the detected object. Fig. 1(c) and (d) shows the fake saliency map and segmentation results using our method.

Manuscript received January 20, 2012; revised March 29, 2012; accepted April 10, 2012. Date of publication April 19, 2012; date of current version July 09, 2012. This work was supported in part by the National Natural Science Foundation of China under Grant 60905019, in part by Tianjin Key Technologies R&D program under Grant 11ZCKFGX00800, and in part by Tsinghua-Tencent Joint Laboratory for Internet Innovation Technology, and by SKL of CG&CAD. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alex ChiChung Kot.

The authors are with the School of Computer Science and Technology, Tianjin University, Tianjin 300072, China (e-mail: hzfu@tju.edu.cn; xcao@tju.edu.cn).

Color versions of some of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIFS.2012.2195492

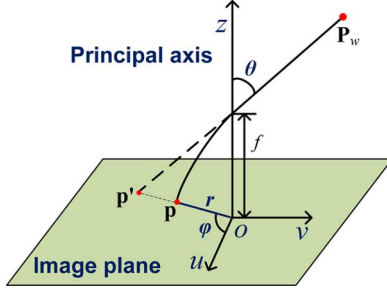


Fig. 2. Traditional capture ray-based model of wide-angle lens, where \mathbf{P}_w is a 3-D point in space, \mathbf{p}' is projected point on image plane based on pinhole model, \mathbf{p} is the projected point under radial distortions. f is focal length, z is principal axis. θ and ϕ denote the incidence angle and the directional angle, respectively.

A. Related Works

The difference between the ideal pinhole lens and the wide-angle lens is shown in Fig. 2, where \mathbf{P}_w is a 3-D point in space, \mathbf{p}' is the projected point based on the pinhole model, and \mathbf{p} is the projected point under the radial distortions. f is the focal length and z is the principal axis. θ and ϕ denote the incidence angle and the directional angle, respectively. The projected point \mathbf{p} has a radial distortion comparing to the projected point \mathbf{p}' . The perspective projection of an ideal pinhole camera can be described by the following formula:

$$r(\theta) = f \tan(\theta) \quad (1)$$

where θ is the incidence angle between the principal axis and the incoming ray $\mathbf{P}_w\mathbf{p}'$ and r is the distance between the projected point and the principal point. The wide-angle lens instead is usually designed to obey one of the following projections [26]–[28]:

$$\begin{aligned} r_1(\theta) &= f_1 \theta \quad (\text{Equidistance}) \\ r_2(\theta) &= 2f_2 \sin(\theta/2) \quad (\text{Equisolid angle}) \\ r_3(\theta) &= f_3 \sin(\theta) \quad (\text{Orthogonal}) \\ r_4(\theta) &= 2f_4 \tan(\theta/2) \quad (\text{Stereographic}) \end{aligned} \quad (2)$$

where this radial distortion model $r(\theta) = fg(\theta)$ is often called captured ray-based model. A typical wide-angle lens is the fisheye lens, which captures a semispherical view at one time. With the extreme FOV achieving 360 degrees, a fisheye lens is usually used to study the radial distortion [7]–[9]. In this paper, we focus on the fisheye lens without loss of generality, it easily derives more general results than other wide-angle lenses.

Unified model, integrating all radial distortion models in (2) into one frame, is more suitable for image processing, because the distortion model is not available in an uncalibrated image. In [29], the unified theory for all central catadioptric systems is provided. With this model, all of the catadioptric cameras are isomorphic to projective mappings from the sphere to a plane with a projection center on the perpendicular to the plane. In [30], the authors extend the unified model to include a fisheye lens by substituting the projection on a sphere by a projection on a quadric surface. In [31], the unified model for catadioptric cameras via a spherical projection is equivalent to the captured

ray-based models with radial distortions and it can be directly employed to model fisheye lens. However, these unified models contain two disadvantages: first, these models need more than two parameters, which are unlikely to be accurately calibrated from one forgery image; second, these models are invalid with the incidence angle $\theta = \pi/2$, which is a significant property of the fisheye lens. Instead, we simplify the unified model and introduce a radial distortion projection (RDP) model, which is equivalent to other captured ray-based models in (2) with only one parameter and a simpler formula.

The distortion of a wide-angle lens provides a very important geometric constraint, that the straight line in space is projected into a great circle on the projection sphere [32], [33]. We introduce two bottom-up cues based on this geometric constraint and employ them to judge the untrustworthy likelihoods of the candidate lines. In our method, the candidate lines are extracted with manual assistance: firstly the curves are detected automatically using edge detection and then the candidate lines are selected supervised, representing by three points: two endpoints and one midpoint. The reason for selecting the candidate lines manually is that it is still highly unlikely, if not impossible, for the state-of-the-art algorithms to recognize the straight world lines well under the distortion.

B. Overview and Contributions

In this paper, a forgery detection method based on lens distortion is proposed. Firstly, the untrustworthy likelihood of each candidate line is computed, by measuring the distortion cues. Then a fake saliency map is generated to express the position of the forged object. Finally, an energy function is employed to segment the forgery object via graph cut. To the best of our knowledge, this is the first work to detect forgery object via distortion constraint. The main contributions of this paper are as follows.

- 1) Radial distortion projection (RDP) model is adopted to simplify the traditional captured ray-based models, with only one parameter and a simpler formula. In the RDP model, the straight world line is projected into a great circle on the viewing sphere.
- 2) Geometric constraint of line in the RDP model is presented. This constraint is carried out on the viewing sphere of the RDP model and validated without rectifying the image. Two bottom-up cues based on distortion constraint, the volume cue and the distance cue, are defined to measure untrustworthy likelihood of the candidate line.
- 3) A fake saliency map using the untrustworthy likelihood is employed to maximum fake saliency density and locate the forgery object. The energy minimization-based segmentation is used to extract the forgery object. The fake saliency map could be widely used in other forensic methods.

This paper is organized as follows: Section II introduces our RDP model, and in Section III, the geometric constraint of line and the distortion cues are derived. Section IV describes our fake saliency map and gives the energy function for segmenting the forgery object. Experimental results and discussion are given in Section V. Finally, Section VI presents some concluding remarks.

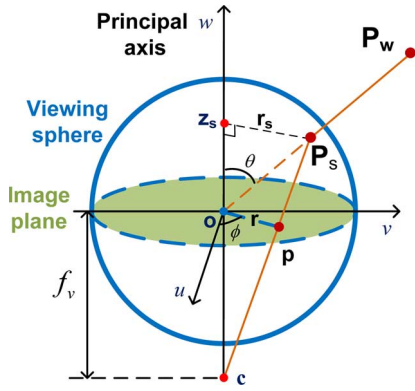


Fig. 3. RDP model for wide-angle lens, where \mathbf{P}_w is the point in space, \mathbf{P}_s is spherical projection point on viewing sphere, \mathbf{p} is projected point on the image plane, and f_v is virtual focal length.

II. RADIAL DISTORTION PROJECTION MODEL

For the forgery detecting, an RDP model is modified based on unified model [29], [31]. The RDP model describes a projection mapping from 3-D space points to 2-D distorted image points through a unit sphere, which is called a viewing sphere, as shown in Fig. 3. The virtual optical center \mathbf{c} and origin \mathbf{o} are termed optical center and principal projection center (image center), respectively. The virtual optical center \mathbf{c} has coordinates $[0, 0, -f_v]^T$ with the virtual focal length $f_v > 0$. The image formation is modeled as follows.

Step 1) Considering a 3-D point \mathbf{P}_w with Cartesian coordinate system $\mathbf{P}_w = [X_w, Y_w, Z_w]^T$ with respect to viewing sphere. The spherical projection point $\mathbf{P}_s = [x_s, y_s, z_s]^T$ of \mathbf{P}_w on the viewing sphere surface has the form

$$\mathbf{P}_s = \frac{\begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}}{\|\mathbf{P}_w\|} = \begin{bmatrix} \sin(\theta) \cos(\phi) \\ \sin(\theta) \sin(\phi) \\ \cos(\theta) \end{bmatrix} \quad (3)$$

where $\|\mathbf{P}_w\| = \sqrt{X_w^2 + Y_w^2 + Z_w^2}$, and (ϕ, θ) denote the 2-D spherical coordinates of the spherical point \mathbf{P}_s , ϕ is the directional angle and θ is the incidence angle.

Step 2) The spherical point \mathbf{P}_s is projected on the normalized image plane as the homogeneous coordinates $\mathbf{p} = [u_p, v_p, 1]^T$. This projection is a projective transformation respected to the virtual optical center \mathbf{c} , as in the pinhole model. The homogeneous coordinates of \mathbf{p} is obtained

$$\mathbf{p} = \mathbf{K}_s \mathbf{P}_s \quad (4)$$

where \mathbf{K}_s is the camera intrinsic matrix given by

$$\mathbf{K}_s = \begin{bmatrix} f_v & & \\ & f_v & \\ & & 1 \end{bmatrix}$$

with the virtual focal length f_v .

A. Back-Project Formulation

The back-project formulation of the lens model is more useful, since the image coordinate is the only input in practice.

$\mathbf{p} = [u, v, 1]^T$ denotes the point on the image plane, where the origin is set as the image center \mathbf{o} . The radius r of the point \mathbf{p} is calculated with $r = \sqrt{u^2 + v^2}$. For simplified calculating, we combine the image point \mathbf{p} and radius r to an intermediate variable $\tilde{\mathbf{p}} = [u, v, r]^T$. The back-project formulation of the RDP model is given by

$$\begin{aligned} \mathbf{P}_s &= \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} = \begin{bmatrix} \sin(\theta) \cos(\phi) \\ \sin(\theta) \sin(\phi) \\ \cos(\theta) \end{bmatrix} \\ &= \frac{1}{r} \begin{bmatrix} \sin(\theta) & & \\ & \sin(\theta) & \\ & & \cos(\theta) \end{bmatrix} \begin{bmatrix} u \\ v \\ r \end{bmatrix} = \frac{1}{r} \mathbf{K}_b \tilde{\mathbf{p}} \quad (5) \end{aligned}$$

where $\mathbf{P}_s = [x_s, y_s, z_s]^T$ is the unit directional vector of point on the viewing sphere surface, and \mathbf{K}_b is defined as the back-project matrix, $\phi = \arctan(v/u)$ denotes the directional angle and θ is the incidence angle obtained by

$$\theta = \begin{cases} 2 \arctan(r), & \text{if } f_v = 1 \\ 2 \arctan\left(\frac{f_v - \sqrt{f_v^2 + r^2 - f_v^2 r^2}}{r(f_v - 1)}\right), & \text{if } f_v \neq 1 \end{cases} \quad (6)$$

With (5), the image coordinate is transformed to the unit spherical coordinate on the viewing sphere.

B. Some Highlights

There are some highlights of the RDP model, as follows.

1) The RDP model is equivalent to the other captured ray-based model in (2) with only one parameter f_v . The RDP model could be formulated as $r = f_v g(\theta)$, according to the similar triangles $\triangle \text{ocp} \cong \triangle \text{cz}_s \mathbf{P}_s$, as follows:

$$\frac{r}{\sin \theta} = \frac{f_v}{f_v + \cos \theta} \Rightarrow r = \frac{f_v \sin \theta}{f_v + \cos \theta}. \quad (7)$$

Hence, the relation between the RDP model and the captured ray-based model is obtained by

$$\hat{f}_v = \arg \min_{f_v} \sum_{\theta \in [0, \pi/2]} \left\| f_g g(\theta) - \frac{f_v \sin \theta}{f_v + \cos \theta} \right\|^2 \quad (8)$$

where $f_q g(\theta)$ is the captured ray-based model in (2).

2) Our RDP model releases the parameter of projection plane distance in unified model [29], [31], which makes the RDP model more focused on the distortion constraint. Simultaneously, the simplified back-project formulation only back-projects the point to the viewing sphere without correction. Moreover, our RDP model is also valid with the incidence angle $\theta = \pi/2$, which is invalid in the unified models [30], [31]. In the RDP model, the point of $\theta = \pi/2$ is projected on the edge of image as the horizon circle with $r = 1$ in (7).

3) Thanks to the spherical projection step in the RDP model, the straight world line is projected into a great circle on the viewing sphere and into a conic on the image plane, which provides a geometric constraint on the viewing sphere. Detection of the forgery on the viewing sphere without being corrected could reduce the proportional error caused by scale changing. For illustrating this scale changing, we assume the distorted image is rectified into a corrected plane,

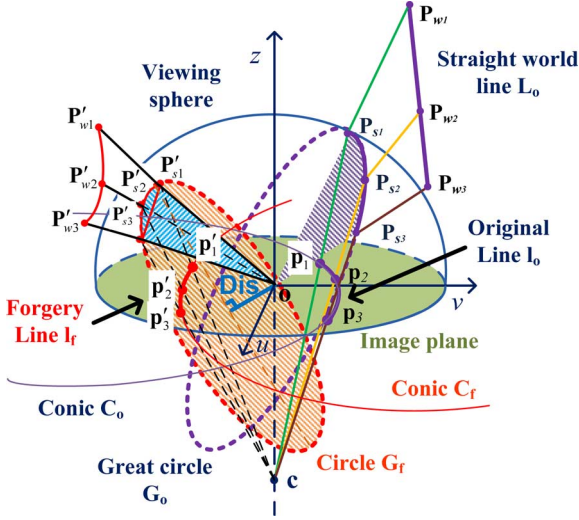


Fig. 4. Distortion constraint of line in RDP model. A world line is projected to a circle on viewing sphere and then projected to a conic on image plane. Purple lines denote projection of a straight world line (original), and red lines denote the nonstraight line (forgery). Blue shadow denotes volume cue, and variable Dis denotes distance cue.

which is parallel to the image plane. Some corrected samples are shown in Fig. 8(b). It is clear that the point with larger incidence angle θ is placed into the infinite position in corrected image. Hence, a short line on the image plane with a large incidence angle is projected to a near incidence line on the corrected plane. This scale changing can easily lead to the measure error, which is more serious with the incidence angle θ tending to $\pi/2$. In contrast, the scaling is normalized on the viewing sphere, avoiding the proportional error caused by scale changing.

- 4) As the unified model, the RDP model is suited for the catadioptric lens, with the virtual focal length $f < 0$. Moreover, the RDP model could be also used for other wide-angle lens. A simple scheme is that the f_v is obtained by

$$f_v = r_w \frac{\cos \theta_w}{\sin \theta_w - r_w} \quad (9)$$

where r_w denotes the image dimension and θ_w is the FOV of the wide-angle lens.

III. BOTTOM-UP CUES BASED ON DISTORTION CONSTRAINT

In this section, we describe the distortion constraint of the line in the RDP model and then present two bottom-up cues, the volume cue and the distance cue, to determine the straight line.

A. Distortion Constraint of the Line in RDP Model

In the RDP model, the straight world line is first projected to a great circle on the viewing sphere and then projected to a conic on the image plane. Fig. 4 shows this projection, where the purple lines L_o denote the projection of a straight world lines $P_{w1}P_{w2}P_{w3}$. The conic C_o is fitted from the projected points (p_1, p_2, p_3) of purple world line. The points (p_1, p_2, p_3) represent two endpoints and one midpoint of the candidate

line l_o and are back-projected into points (P_{s1}, P_{s2}, P_{s3}) , respectively. The points (P_{s1}, P_{s2}, P_{s3}) define a great circle G_o , which is projected to the straight line L_o in the space with the center o of the viewing sphere. In contrast, the points (p'_1, p'_2, p'_3) lie on the forgery line, which could be presented as a conic section C_f on the forgery object. The three points are back-projected to points $(P'_{s1}, P'_{s2}, P'_{s3})$ on the viewing sphere, which defines another circle G_f . Note this circle G_f is not a great circle on the viewing sphere, because the forgery line is not satisfying this geometric constraint any more. In other words, the circle generated by the forgery line determines a plane section, which cuts the viewing sphere without passing the center o of the view sphere, as shown as the red dotted circle in Fig. 4. Thanks to this constraint of line, there are two effective bottom-up cues, the volume cue and the distance cue, which could be used to distinguished the forgery line.

B. Volume Cue

With the geometrical constraint, if three points (p_1, p_2, p_3) on the image plane are projected from the straight world line, their back-projected points (P_{s1}, P_{s2}, P_{s3}) on the viewing sphere follow the geometrical constraint as

$$\text{Volume}_\Delta = \det(P_{s1} P_{s2} P_{s3}) = 0 \quad (10)$$

where Volume_Δ is the volume of tetrahedron made up by the center o of viewing sphere and points (P_{s1}, P_{s2}, P_{s3}) of the great circle G_o , as shown as the purple shadow in Fig. 4. Statistically, the forgery line is highly unlikely to satisfy this geometrical constraint, as the blue shadow in Fig. 4. So the volume cue is defined as a bottom-up cue w^V

$$w^V = \frac{1}{\|l_f\|^3} \|\det(P'_{s1} P'_{s2} P'_{s3})\| \quad (11)$$

where $\|l_f\|$ is the length of the forgery line l_f in the image. Fig. 5(b) shows that the curves of the volume cue of each candidate line with different virtual focal length of the RDP model, where the volume cues of original lines are close to zero and the cues of forgery lines are more likely to be away from the zero. But in the experiments, the volume cue of the forgery line, which locates away from the image center, is not distinguished clearly with the cues of original lines, such as the right-most candidate line in the second sample of Fig. 5(b). Note that with the virtual focal length $f = 0$, the RDP model degenerates and all the distortion values equal to zero, because the projective transformation of the Step 2 in Section II-A is invalidated and the incidence angle of every pixel vanishes identically in (6).

C. Distance Cue

Same as the volume cue, the distance between the center o and the plane, which is defined by three projected points P_{s_i} on the viewing sphere, is also considered as a cue. Suppose the matrix $M = [P, P_1, P_2, P_3]$ is composed of a general point P and three points P_i on the view sphere which define the plane Π . Since $\det(M) = 0$ for points on Π we can read off the plane coefficients as [34]

$$\Pi = (D_{234}, -D_{134}, D_{124}, -D_{123})^T \quad (12)$$

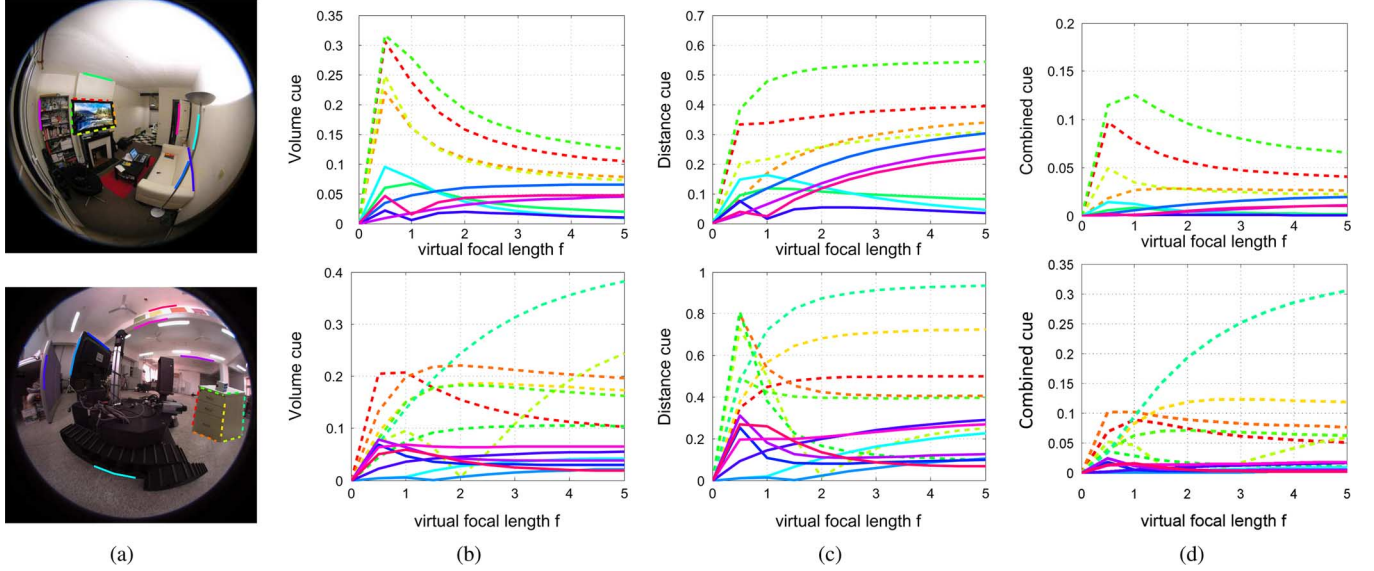


Fig. 5. Distortion cues for each candidate line with different virtual focal length. (a) Input images, where candidate lines are labeled by different colors. Forgery lines are labeled by dashed lines, and original lines are labeled by solid lines. (b)–(d) Volume cue w^V , distance cue w^D , and combined cue w^C with different virtual focal length, respectively.

where D_{jkl} is the determinant formed from the jkl rows of matrix $[\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3]$. Note in this matrix, the \mathbf{P}_i is represented by homogeneous coordinates as $\mathbf{P}_i = [x_i, y_i, z_i, 1]^T$, so they define a plane and the dimension of matrix $[\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3]$ is 4×3 . The distance between the plane Π and any 3-D point $\mathbf{P}_0 = [x_0, y_0, z_0]^T$ is computed by

$$\text{Dis} = \frac{|x_0 D_{234} - y_0 D_{134} + z_0 D_{124} - D_{123}|}{\sqrt{D_{234}^2 - D_{134}^2 + D_{124}^2}}. \quad (13)$$

As in the previous description, the original line in the distorted image, which is projected from the straight world line, is formed by the great circle \mathbf{G}_L on the viewing sphere. So the center \mathbf{o} of the viewing sphere should be on the plane defined by the great circle \mathbf{G}_L and has $\text{Dis} = 0$, shown as the purple dotted circle in Fig. 4. Oppositely, the forgery line cannot be back-projected into the great circle and has $\text{Dis} > 0$, shown as the red shadow in Fig. 4. So we define the distance between the plane generated by line and center \mathbf{o} of viewing sphere as a bottom-up cue w^D

$$w^D = \frac{|D_{123}|}{\sqrt{D_{234}^2 - D_{134}^2 + D_{124}^2}}. \quad (14)$$

Note that the center \mathbf{o} of the viewing sphere locates on the origin $[0, 0, 0]^T$ of frame, so the first three terms on the nominator in (13) is removed in (14). Fig. 5(c) shows that the curves of the distance cue of each candidate line with different virtual focal length of the RDP model. Compared with the volume cue, the distance cue could be useful to classify the candidate lines, which locates away from the image center. But in the experiments, we find the distance cue is short of robustness against the noise and has a lower precision than the volume cue. The analysis of robustness against the noise is described in detail in Section V-B3.

D. Combining Volume and Distance Cues

To achieve a better forensic detection, it is helpful to combine two distortion cues. They can be simply combined for a combined cue w^C as follows:

$$w^C = \log(w^V w^D + 1) \quad (15)$$

where w^V is computed in (11) and w^D is computed in (14). Fig. 5(d) shows the curves of the combined cue with different virtual focal length of the RDP model, where the forgery lines are separated with the original lines.

IV. FAKE SALIENCY MAP

Segmentation of forgery objects in an image remains a challenging problem, as the forged object (or detected region) is usually noisy and incomplete. In this section, we introduce a fake saliency map to maximum fake detection density. Firstly, the untrustworthy likelihood is employed to initialize seed. And then the fake saliency map is generated by the untrustworthy likelihood to express the location of forgery object. Finally, the forgery object is segmented by minimizing the energy function via graph cut, which achieves a pixel-level segmentation. Note that our fake saliency map is not constrained to a specific choice of the forensic methods, and here we still continue the distortion cues to describe this method.

A. Untrustworthy Likelihood

Thresholding is a common way to make a decision for authentication, however, which potentially leads to wrong labeling for the candidate near the thresholding edge. To avoid this “hard” decision, we prefer a “soft” fake saliency value, called untrustworthy likelihood, for each candidate. First, a training dataset is used to generate the distribution, which contains 4000 synthetic original lines and 4000 synthetic forgery lines. The synthetic data is described in detail in Section V-B. The distortion cues of these lines are divided into disjoint bins. Fig. 6 shows

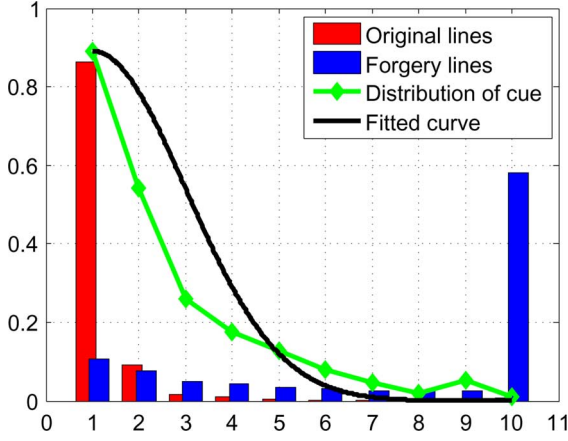


Fig. 6. Statistical histogram of combined cues on synthetic data set, where green curve denotes distribution of untrustworthy likelihood and black curve denotes fitted curve using (18). The x-axis denotes bin's number, and y-axis denotes frequencies. Red/blue bars denote proportion of total original/forgery lines that fall into each bin.

the statistical histogram of combined cue on the synthetic data set, where x-axis denotes the bin's number, y-axis denotes the frequencies, and the red/blue bars denote the proportion of total original/forgery lines that fall into each bin. $w_i, i = 1 \dots N$ denotes the bin of distortion cue, where N is the total number of the bins ($N = 10$, for example in Fig. 6). Notice in practice, the largest combined cues of the original lines are smaller than 0.01, so the statistical histogram is limited in the range $[0, 0.01]$ of the combined cue, and the combined cue $w^C > 0.01$ falls into the last bin w_N . With this histogram, the distribution curve of distortion cue is obtained by

$$p(x|w_i) = \frac{N_O(i)}{N_O(i) + N_F(i)} \quad (16)$$

where $p(x|w_i)$ is a prior probability of x given curve value w_i , $N_O(i)$ ($N_F(i)$) is the number of original (forgery) lines in bin i . The green curve in Fig. 6 shows the distribution curve of a combined cue.

For brevity's sake, we use a Gaussian kernel to fit this distribution curve

$$G(x|\alpha, \beta) = \alpha \exp\left(-\frac{x^2}{2\beta^2}\right) \quad (17)$$

where α is called hyperparameter, which controls the distribution of model parameters, and β is the precision parameter corresponding to the inverse variance of the distribution. The α can be inferred from the curve of data set with $\hat{\alpha} = p(x|w_i = 0)$. The precision parameter β is determined by the least square error as

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^N (p(x|w_i) - G(x|\hat{\alpha}, \beta))^2. \quad (18)$$

Having determined the parameter $\hat{\alpha}$ and $\hat{\beta}$, we can now make predictions for distortion cue of any candidate line. The fitted curve using $\hat{\alpha} = 0.89$ and $\hat{\beta} = 0.2$ is shown as the black curve in Fig. 6.

At last, we define the untrustworthy likelihood U with this distribution as follows:

$$U = 1 - \hat{\alpha} \exp\left(-\frac{w^2}{2\hat{\beta}^2}\right) \quad (19)$$

where w is the distortion cue. Obviously, the untrustworthy likelihood of the forgery line should be higher than that of the original line. The untrustworthy likelihood is employed to compute the initial weight in fake saliency map. Fig. 7(b) shows the examples of untrustworthy likelihood.

B. Fake Saliency Map

Our goal of forgery detection is to achieve a pixel-level segmentation rather than line detection. There are several works [35], [36] to obtain regions or objects from contours. However, these methods are not suitable for our forensic detection, for two reasons: First, the potential fake lines are not necessarily object boundaries; second, the untrustworthy likelihood could not present sufficient information to determine which side of the line is forgery. In a statistical point of view, however, the region between the fake lines are more likely to be forgery. Based on this observation, we borrow the concept of saliency, which originally denotes the fixation points or regions that a human viewer would focus on at the first glance [37]–[39]. The main difference is that our goal is to detect a forgery object rather than the visual attention. In addition, we use the untrustworthy likelihood instead of the appearance cues in visual attention.

Given an image \mathbf{I} , $\mathbf{L}^* = [l_i]$ denotes the set of candidate lines and U_i denotes untrustworthy likelihood of line l_i . To cover the forgery object better, a center of gravity with the high untrustworthy likelihood is introduced. Let $\mathbf{G}(c_g, r_g)$ denote the center of gravity, where the c_g denotes the location and r_g denotes the radius of the center. The center of gravity is obtained by

$$c_g = \frac{1}{\|\mathbf{L}^*\|} \sum_{l_i \in \mathbf{L}^*} l_i U_i. \quad (20)$$

Fig. 7(b) shows the examples of an initial saliency map with the candidate lines and square center of gravity, where the heat map presents the untrustworthy likelihood. Then, we dilate each line to extend the untrustworthy likelihood to neighbors of the line, as shown in Fig. 7(c). Next, for each pixel $x_k \in \mathbf{I}$, a fake saliency weight S_k is defined as follows:

$$S_k = \frac{1}{\|\mathbf{L}^* \cap \mathbf{G}\|} \sum_{l_i \in \mathbf{L}^* \cap \mathbf{G}} U_i \exp\left(-\frac{dis(x_k, l_i)}{\gamma^2}\right) \quad (21)$$

where \mathbf{G} is the center of gravity computed in (20), $dis(x_k, l_i)$ denote the Euclidean distance between pixel x_k and l_i , and γ is a scale parameter, which is set as $Radius/40$ pixels in experiments, and $Radius$ denotes the radius of the image. Fig. 7(d) shows the final fake saliency map $\mathbf{S} = [S_k]$.

C. Energy Function of Segmentation

In this section, we utilize a bilayer segmentation method that estimates the forgery and original pixels of the input image by minimizing an energy function. A naive solution is thresholding the fake saliency map, Fig. 7(e) shows the thresholding result with $\mathbf{S} > 0.5$ for the example. However, the extracted area

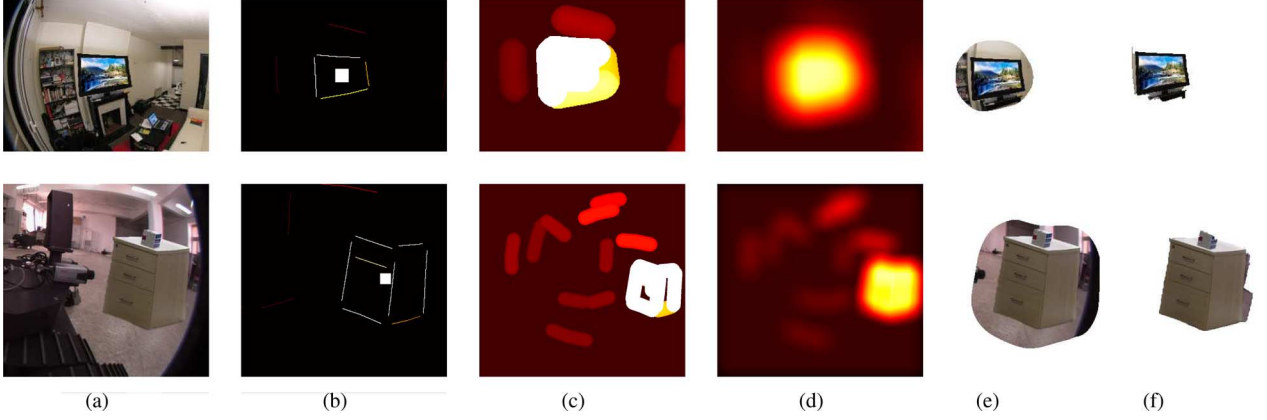


Fig. 7. Procedure of generating fake saliency map. (a) Zoom-in version of forgery regions of Fig. 1(b). (b) Initial saliency map with candidate lines and square center of gravity, where heat map presents untrustworthy likelihood. (c) Saliency map with dilation procedure. (d) Final fake saliency map. (e) Forgery segmentation results using threshold, e.g., $S > 0.5$. (f) Forgery segmentation via our energy function.

involves some background in the neighborhood of the object. Therefore, instead of considering pixels independently and segmenting the fake saliency maps by simple thresholding, we formulate the problem as a min-cut/max-flow problem on a graph and find the assignment of values to every pixel, as either forgery or original label. The energy function E is defined as

$$E(x_i) = \sum_{i \in I} S_i(x_i) + \sum_{\{p,q\} \in \mathcal{N}} V_{p,q}(x_p, x_q) \quad (22)$$

where $x = \{x_i | i \in I\}$ denotes the binary-valued label of image I , for hard segmentation $x_i \in \{0, 1\}$, with zero for original and one for forgery, S_i is the fake saliency weight of pixel i in (21), which is charged for a pixel that is labeled to forgery, $V_{p,q}$ is the smoothness penalty, which measures the cost of assigning different labels to two neighboring pixels, \mathcal{N} is a set of pairs of adjacent pixels. The weight $V_{p,q}$ of smoothness term is given by [40]

$$V_{p,q} = \frac{\lambda \exp(-\beta \|z_p - z_q\|^2)}{\text{dis}(p, q)} \quad (23)$$

with $\text{dis}(\cdot, \cdot)$ denoting the Euclidean distance between two pixels in the image, $\beta = (2\langle (z_p - z_q)^2 \rangle)^{-1}$, where z_x is the RGB color appearance of pixel x , and $\langle \cdot \rangle$ denotes expectation over the image and λ is, respectively, the weight for the contrast sensitive term. The segmentation can be estimated as a global minimum

$$\mathbf{x}^* = \arg \min E(x_i). \quad (24)$$

This function can be justified by the well known Markov Random Fields (MRF) formulation, which is submodular and can be solved via finding a cut by the min-cut/max-flow algorithm [41], [42]. Fig. 7(f) shows the segmentation results with minimizing the energy function (22).

V. EXPERIMENTS

We test our method in three parts. First, the evolution of our RDP model is given. Then the simulation and real images experiments of our forensic method are presented. In synthetic simulations, we evaluate the influences of different cues in the RDP model and the robustness with respect to the noise. In real image tests, the performances are evaluated on synthesized forgery

image dataset and visually plausible images. Third, the applications of fake saliency map in the other methods is shown. Furthermore, some degenerate cases of our method are discussed. Unless otherwise specified, the image center is set as the geometric center of the image and the virtual focal length $f = 1$ in the RDP model. The resolution of the image is 800×800 .

A. RDP Model

As discussed in Section II-B, our RDP model is equivalent to other captured ray-based models. In this experiment, we assume the FOV of ideal fisheye lens is $\pi/2$ and the projection point is projected on the normalized image plane. Thus the f_g could be calculated by $f_g = 1/g(\pi/2)$. And the behavior of the different projections and the RDP model is illustrated in Fig. 8(a). The star points denote the curves of the captured ray-based model in (2), and the curves show the best fits of the RDP model with the corresponding virtual focal length \hat{f}_v . Interestingly, almost all the captured ray-based models can be approximated well by the RDP model by a fixed virtual focal length f_v . Moreover, Fig. 8(b) shows the image correction using our RDP model, where the distortion is rectified well with our RDP model. Noting that the last two samples are wide-angle images, our RDP model focuses on image distortion and is not limited by the FOV of the lens.

B. Simulation

In the simulation, we evaluate the performances of our method versus different factors in geometric method. These factors include: the bias of the image center, the virtual focal length of the RDP model, and the noise. The synthetic data set is configured as: firstly, a set of straight world lines is generated randomly. The length of world line is limited in $[0.2, 0.8]$ for a unit radius of distorted image, i.e., $R = 1$, and the orientation is in the range of $[0, \pi]$. Then these lines are projected into the image plane using traditional unified models in (2). These lines are considered as the original lines. The set of forgery lines is generated by moving these original lines randomly into the new positions inside the image. Because the synthetic line is generated randomly, we employ increasing the total number of synthetic lines to replace repeating independent trials in statistics. And there are four types of traditional unified model

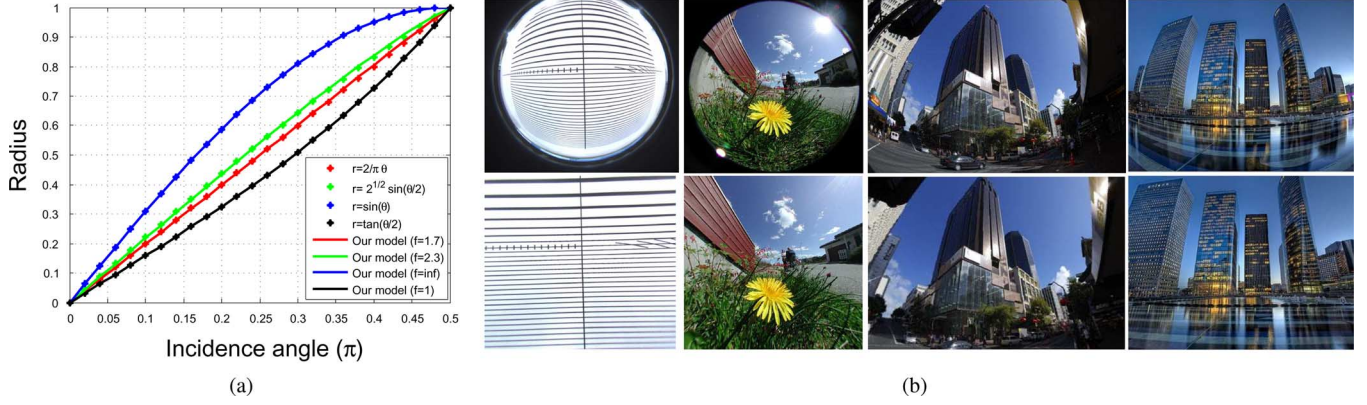


Fig. 8. (a) Relations of captured ray-based models and RDP model. Star points denote curves of captured ray-based models in (2). Curves show the best fits of RDP model with corresponding virtual focal length f_v . (b) Distortion corrected image using our RDP model, where first two samples are fisheye images, and last two samples are wide-angle images.

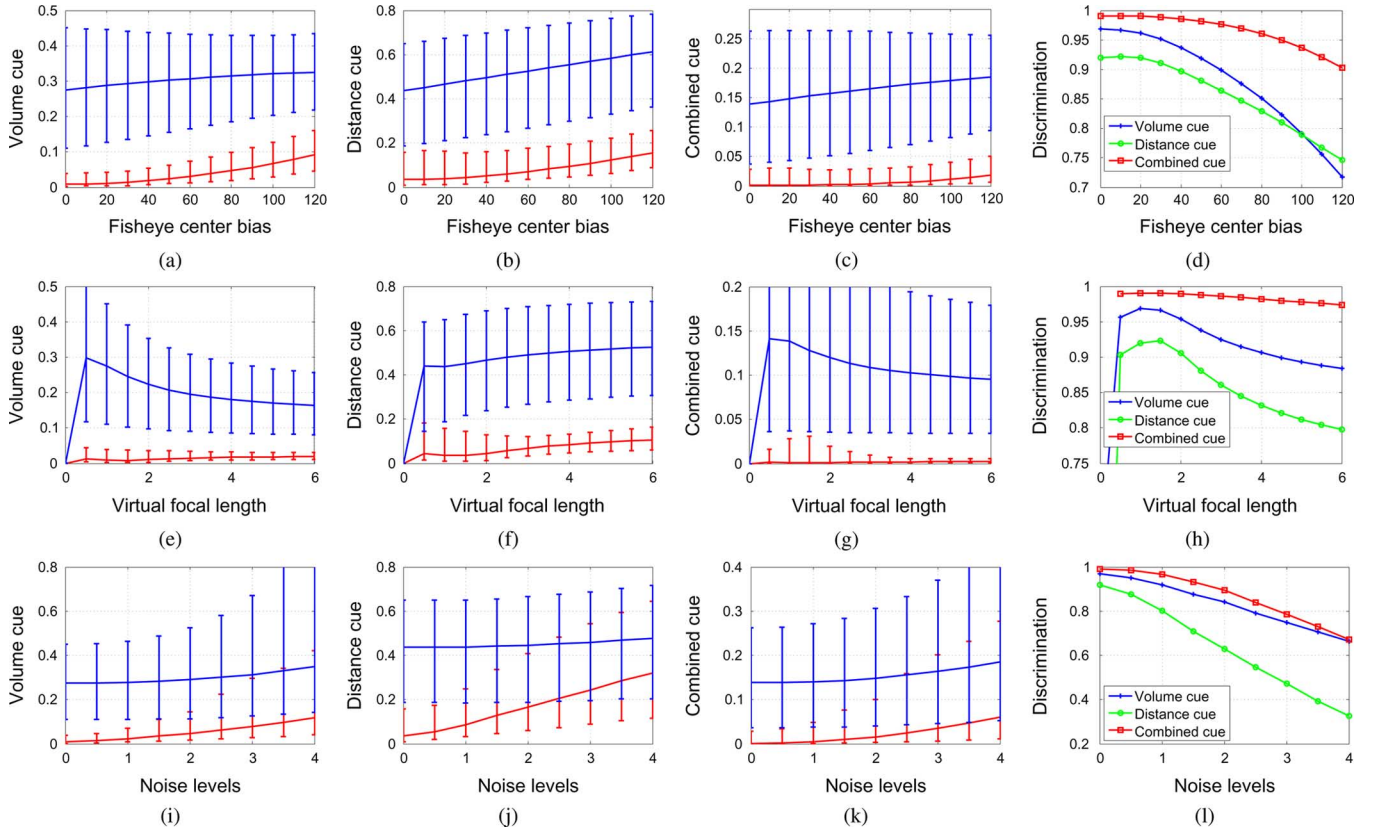


Fig. 9. Performances of computation of distortion cues with respect to various factors. (a)–(c) Performance with respect to varied image center's bias using volume cue, distance cue and combined cue, respectively. Blue bar denotes simulation on forgery data set, and red bar denotes simulation on original data set. (d) Discrimination Score of each forensic cue on the image center's bias. (e)–(g) Performance with respect to varied virtual focal length. (h) Discrimination Score of each bottom-up cue on virtual focal length. (i)–(k) Performance with respect to varied noise's levels. (l) Discrimination Score of each bottom-up cue on noise levels.

in (2), we generate 10 000 original lines and 10 000 forgery lines for each traditional model, and get a synthetic data set contained 40 000 original lines and 40 000 forgery lines.

1) *Performance With Respect to Bias of Image Center:* For forgery detecting in an uncalibrated image, the distortion center is often assumed to be the geometric center of the radial distorted image, which is not accurate in the practice [7]. In the first simulation, we evaluate the influence of the image center's bias on the bottom-up cues: the volume cue, the distance cue, and the combined cue. The Gaussian noise with zero-mean and

standard deviation δ from 0 to 120 pixels is added to the image center, which is used to present the bias of the image center. Fig. 9(a)–(c) shows the performed and the averaged values of the volume cue, the distance cue and the combined cue, respectively. In this paper, the standard deviations of error bar are divided into two parts to present, which are computed with greater than means and less than means. The red lines are results on original lines, and the blue lines denote the forgery lines. Notice that the errors do not go to zero, because the original lines are generated by four kinds of the traditional wide-angle models,

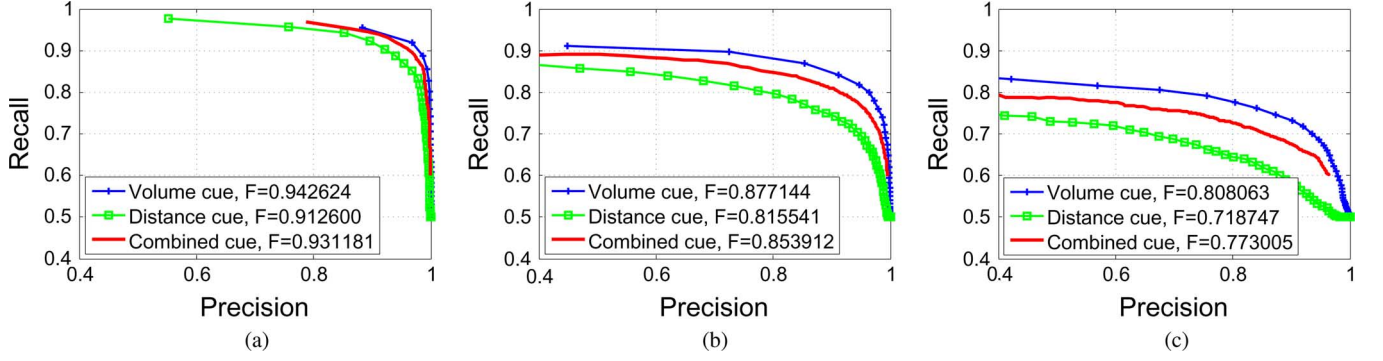


Fig. 10. Precision/Recall curves with respect to varied noise's levels, where blue, green, and red curves denote volume cue, distance cue and combined cue, respectively. (a) Noise $\delta = 0$. (b) Noise $\delta = 2$. (c) Noise $\delta = 5$.

and the virtual focal length f in this experiment is fixed to 1. The trends of the cues increase with the bias level of image center, but they still keep the capability to classify the candidate lines.

For better illustrating the capability of forensic classifying power, we introduce the discrimination Score for the bottom-up cue as:

$$\text{Score} = \frac{|\text{mean}(w_f) - \text{mean}(w_o)|}{\text{mean}(w_f)} \quad (25)$$

where the $\text{mean}(w)$ denotes the mean of bottom-up cue w , w_f and w_o are the distortion cues on the forgery and original lines. A larger discrimination Score is preferred for distinguishing the forgery from the candidate lines. Fig. 9(d) shows the discrimination scores of each bottom-up cue on the image center's bias, where the blue, green, and red lines denote the volume cue, the distance cue, and the combined cue, respectively. We observe that the discrimination curve does not decrease linearly as the increasing of the image center's bias, because the discrimination describes the value distance of the cue between original and forgery lines, which are both influenced by the image center's bias. Fig. 9(d) shows the combined cue has the best discrimination in all cues as desired.

2) *Performance With Respect to Virtual Focal Length:* Another parameter in the RDP model is the virtual focal length f . In Section II-B, we demonstrate that the RDP model is equivalent to the other captured ray-based models in (2) with only one parameter f . But for blind detection, the information of the lens is not available, and the virtual focal length f is assumed to a fixed value. In the second simulation, therefore, we evaluate the influence of the virtual focal length f to the bottom-up cues. Fig. 9(e)–(g) shows the error bars of the volume cue, the distance cue and the combined cue, respectively. The analysis is focused to the range $f \in [0, 6]$, because the curves of computed values with $f > 6$ tend towards stability. As described in Section III-B, all the cue values equal to zero with the virtual focal length $f = 0$. The experiment shows that the bottom-up cues classify the candidate lines clearly with respect to different virtual focal lengths. For the original data set, the range of fluctuation of the distance cue is larger than the other two cues, but it still under 0.2. Fig. 9(h) shows the discrimination Score, where the combined cue has the best performance with Score maintained above 0.97. The discrimination of combined cue has the

highest value in the cues and achieves the highest Score = 0.99 at $f = 1$, then descends with the f increasing. So, we usually fix the virtual focal length $f = 1$ for the RDP model in the experiments, which performs best and has a simpler back-project θ computed equation in (6) than other values.

3) *Performance With Respect to the Noise:* In the last simulation, we study the robustness of our forensic cues with respect to the noise. In our method, the candidate line is represented by three points, two endpoints and one midpoint, so a Gaussian noise varied in the range from 0 to 4 pixels is added on these points. Fig. 9(i) shows the result of the volume cues, where the volume cue on original lines increases from 0.0085 to 0.1204 with the noise level from 0 to 4 pixels. The original and forgery lines are separated clearly with the noise levels less 1.5 pixels. Fig. 9(j) shows the result of the distance cue, where the error of original lines increases almost linearly as the increase of noise level. The original and forgery lines could not be distinguished accurately when the noise level is larger than one pixel, and the distance cue is short of robustness against the noise. Fig. 9(k) shows the simulation result of the combined cue, where the bar of forgery lines dose not reach the original's until the noise level $\delta = 1$ pixels. Fig. 9(l) shows the discrimination Score, where the combined cue still holds on the largest discrimination.

The Precision/Recall curves are also computed to show the performance of the bottom-up cues more clearly. Fig. 10 shows the results of our methods with respect to the noise level $\delta = 0, 2, 5$ pixels, respectively. The blue, green and red curves denote the volume cue, the distance cue, and the combined cue, respectively. The adjusting thresholds used for generating the curve range from 0 to 1 for volume cue and distance cue, from 0 to 0.5 for combined cue. The F-measure is also presented using

$$F = \frac{2\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (26)$$

In Fig. 10(a), the volume cue has the best F-measure 0.9426 compared to the other cues without noise. The F-measure of the combined cue is close to the volume cue's and reaches 0.93, as expected. When the noise increases, as shown in Fig. 10(b) and (c), it seems that the performance of the volume cue is better than the other cues. But in fact, the performance of volume cue is weak for some forgery lines, which locate close to the border of the distorted image, as described in Section III-B.

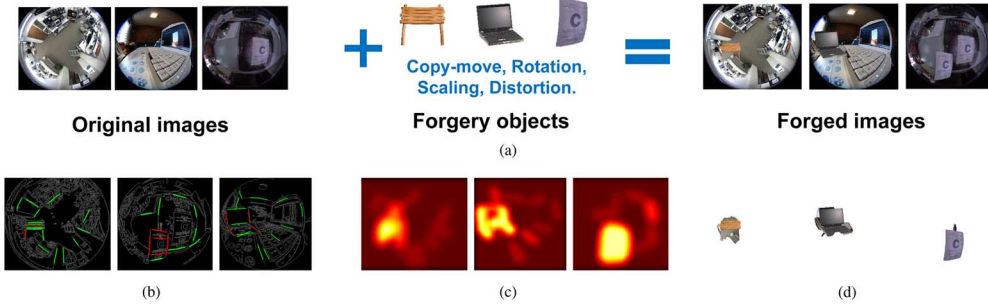


Fig. 11. Three examples of synthesized forgery images used in our experiments. (a) Procedure of generating synthesized forgery images. (b) Line detection results of these samples, where candidate lines with combined cue $w^C > 0.02$ are labeled by red and other lines are labeled by green. (c) Fake saliency map of these samples. (d) Segmentation results using our method.

C. Real Images

1) *Performance With Respect to Synthesized Forgery Images:* In this experiment, we test our method on a synthesized forgery image dataset, which is generated using 11 source images and 13 target images. Fig. 11(a) shows the procedure of generating the synthesized forgery images. Firstly, the original images are collected either by the authors or from the image hosting websites, such as Flickr or Google. The forgery objects are next extracted manually from the source images, which include pinhole and wide-angle images. Then the forgery object is pasted onto the target image. During the tampering, some operations provided in most photo editing tools, such as distorting, resizing, stretching and rotating, are used randomly on the forgery objects. At last, the synthesized forgery image dataset is generated. In the experiment, there are total 1582 lines extracted from 100 images of dataset, which includes 648 forgery lines.

Fig. 11(b) shows the line detection results of the samples in Fig. 11(a), where the candidate lines with combined cue $w^C > 0.02$ are labeled [the untrustworthy likelihood $U \approx 0.45$ with $w^C = 0.02$ in (19)]. The forgery line detection has the misdetection, such as the green line on the forgery object in Fig. 11(b). Fig. 12 presents the Precision/Recall curves of our method on synthesized forgery image dataset, where the blue, green, and red curves denote the volume cue, the distance cue and the combined cue, respectively. The adjusting thresholds used for generating the curve range from 0 to 1 for volume cue and distance cue, from 0 to 0.5 for combined cue. For the real images, the performances of three distortion cues are similar, and the combined cue performs best of all with the F-measure to 0.77.

Compared to the line detection, the fake saliency map achieves a better performance, which locates a forgery object and neglects the misdetection of the line. Fig. 11(c) shows the fake saliency map could locate the forgery object. The forgery object is segmented on a pixel-level by our minimizing the energy function (22), as shown in Fig. 11(d).

2) *Performance With Respect To Visually Plausible Images:* In this experiment, our method is tested on the images with visually plausible composites. The first row of Fig. 13 shows the original images, and the second row shows the forgery images, which are generated by a series of operations to make the generated images more visually pleased. The third row shows the forgery line detection using threshold, where the candidate lines with combined cue $w^C > 0.02$ are labeled by red and the others

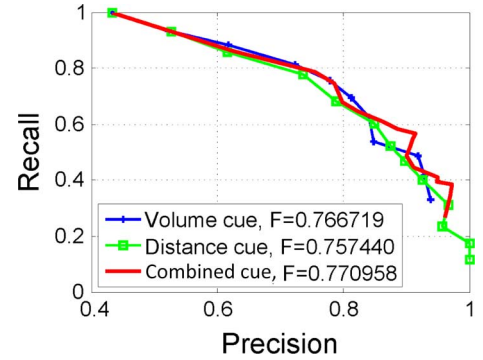


Fig. 12. Precision/Recall curves of our method on synthesized forgery image dataset, where blue, green, and red curves denote volume cue, distance cue and combined cue, respectively.

lines are labeled by green. The last two rows of Fig. 13 show the fake saliency maps and segmentation of forgery objects using our method, which employs the max-flow/min-cut algorithm in [43] to minimize the energy function.

The first three samples of Fig. 13 are the typical rectangle forgery objects, where these straight lines in space are considered as the candidate lines for our method. Moreover, these straight lines are the boundaries enclosing the forgery object, and the forgery object could be segmented based on these detected lines. The fourth sample contains two forgery objects, which are segmented accurately by our method. The last two samples in Fig. 13 are the nonrectangle objects, which contain straight lines on themselves. The same as the rectangle object, the forgery lines of nonrectangle object are detected well and the forgery objects are segmented using the fake saliency map by our method.

Table I shows the qualitative results in the experiment, where the first column is the forgeries in the test images of Fig. 13, the second column shows the numbers of candidate (ground truth forgery) lines, the third column shows the number of detected lines using combined cue $w^C > 0.02$, corresponding the third row of Fig. 13, and the fourth column is the recall of the forgery line detection. The recall reaches average 80% using thresholding with the combined cue. In the third row of Fig. 13, the green line on the forgery object are the misdetections, such as the left edge of the carpet in the third sample and the top edge of window in the fourth sample. As a geometric-based method, our method is not suitable for judging the forgery line inserted

TABLE I

QUALITATIVE RESULTS OF FIG. 13, WHERE FIRST COLUMN IS FORGERIES IN FIG. 13, SECOND COLUMN SHOWS NUMBERS OF CANDIDATE (GROUND TRUTH FORGERY) LINES, THIRD COLUMN SHOWS NUMBER OF DETECTED LINES, AND FORTH COLUMN IS RECALL RATE. FIFTH AND SIXTH COLUMNS SHOW FAKE OBJECT SEGMENTATION USING THRESHOLDING AND OUR GRAPH-BASED METHOD. LAST COLUMN SHOWS PERCENTAGE OF IMPROVEMENT

Forgery object	The number of lines			Segmentation misclassification error		
	Candidates (forgery)	With ($w^C > 0.02$)	Recall	Threshold with ($S > 0.5$)	Our method	Improvement
Poster	11 (4)	3	75%	0.1068	0.0044	95.9%
Sign	11 (5)	4	80%	0.1858	0.1021	45.0%
Carpet	12 (4)	3	75%	0.1802	0.0350	80.6%
Window	14 (9)	7	77.8%	0.0860	0.0206	76.5%
Building	11 (3)	3	100%	0.0425	0.0118	72.24%
Robot	10 (6)	5	83.3%	0.1279	0.0428	66.5%
Avg.	12 (5)	4	80%	0.1215	0.0361	70.29%

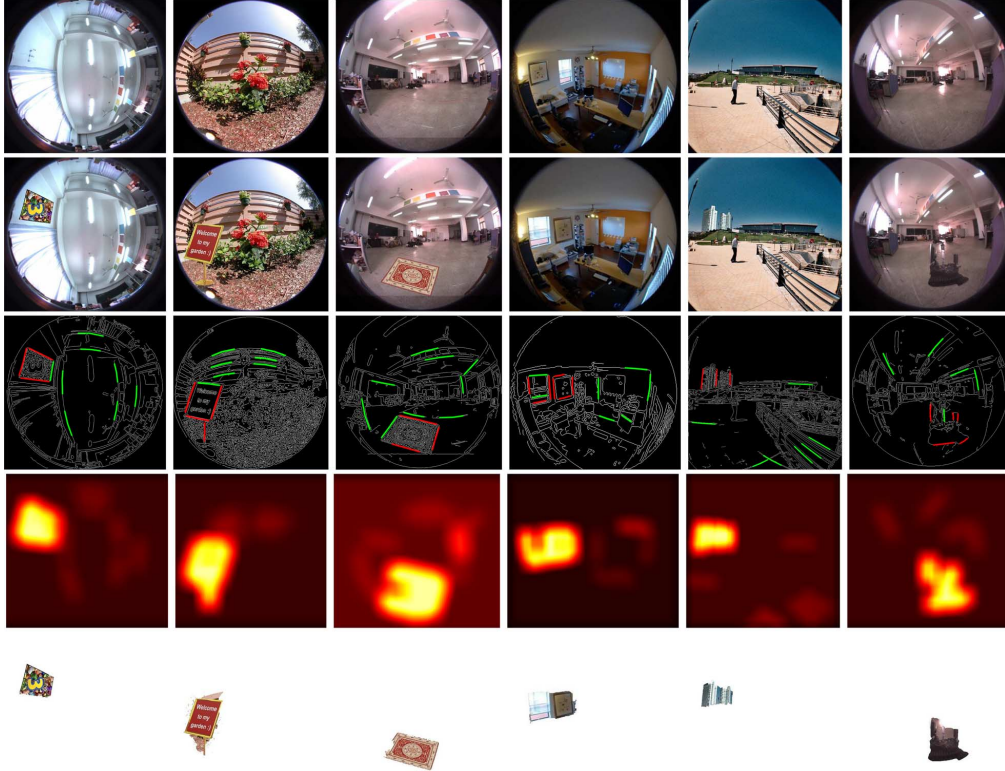


Fig. 13. Forensic detection results of our method. First row shows original images and second row is forgery images used to test. Third row shows forgery line detection using threshold, where candidate lines with combined cue $w^C > 0.02$ are labeled by red and others lines are labeled by green. Fake saliency maps present in the fourth row and last row show segmentation results.

at almost perfect positions with correct distortion. Fortunately, we find it is difficult to forge all the lines on the forgery objects to coincide the distortion discipline, specifically by copying and pasting manipulation. Therefore, the other forgery lines, except degenerated lines, are detected correctly by our method. The forgery object could be segmented by the fake saliency map, which is generated according to the detected forgery lines. The fourth row of Fig. 13 shows that the fake saliency map locates the forgery object accurately. With the experiment, the advantage of the fake saliency map is not only used to achieve an pixel-level detection, but also able to refine the misdetection case.

The fifth column of Table I shows the fake object segmentation using the thresholding with ($S > 0.5$), and the sixth column shows the segmentation using our graph-based method

with the energy function (22). The error term is the misclassification error, which is defined as the percentage of misclassified pixels with respect to ground truth. The improvement of our method is obvious, and the last column of Table I shows the percentage of improvement between thresholding and our method, where the average improvement reaches 70.29%.

D. Application of Fake Saliency Map

As described in Section IV-B, our fake saliency map is not limited in distortion cue, it could be used in other common digital images. In this experiment, we extend our fake saliency map to the method [25]. The method [25] is based on the planar homography constraint, and segments of the forgery object using graph cut with online feature/parameter selection. Here, we retain the homography constraint to provide the initial seed as the

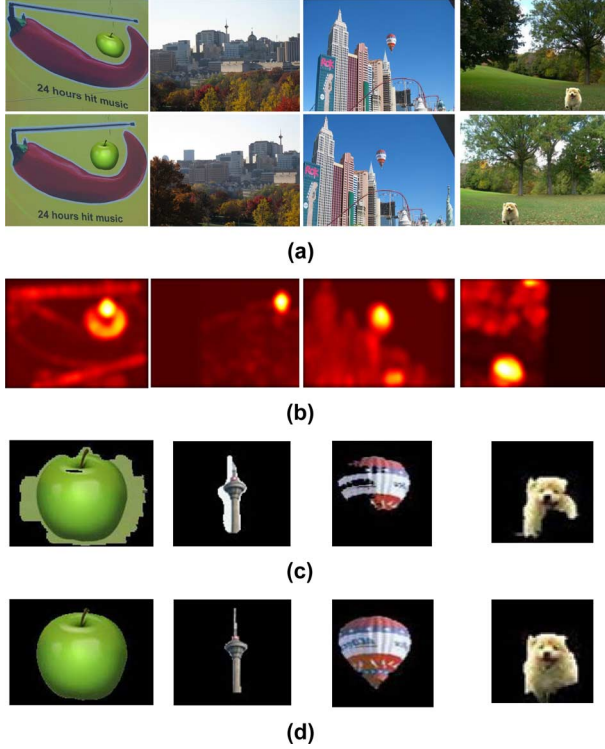


Fig. 14. Application of our saliency map. (a) Input image pairs. (b) Our fake saliency map based on planar homography constraint. (c) Segmentation using [25]. (d) Segmentation using our method.

U in (21), and use energy function (22) based on fake saliency map to replace the graph cut with online feature/parameter selection in [25]. Fig. 14 shows the comparison results, and from top to bottom they are: input image pairs, our fake saliency map, segmentation using [25], and segmentation using our method. It can be seen that the method [25] leads undersegmentation (the first two examples) and oversegmentation (the last two examples), since method [25] targets a “hard” mask for segmentation. In contrast, our fake saliency map offers a “soft” map to generate energy function (22), which extracts the better accurate object, as shown in Fig. 14(d).

The other common forged way is region duplication [44]–[46], which is a simple and effective operation to copy and paste to a different location in the same image, with possible geometrical and illumination adjustments. To detect such modifications, we employ the scale invariant features transform (SIFT) with RANSAC to provide the initial seed as the U in (21), and use a fake saliency map to generate a fake mask and segment the fake object. Fig. 15 shows the duplication detection and segmentation, where from top to bottom they are: original images, fake images, SIFT matching, fake saliency map, and segmentation using our method. Our method detects and segments the duplicated regions in the test image.

E. Discussion

The first case is the misdetection of the forgery line, when we rely only on distortion cues. When the forgery line fully satisfies the projection model of the lens, our distortion cue would fail. A typical misdetection scenario is that the forgery straight line is crossing the image center, which is discussed later. Fig. 16(a)

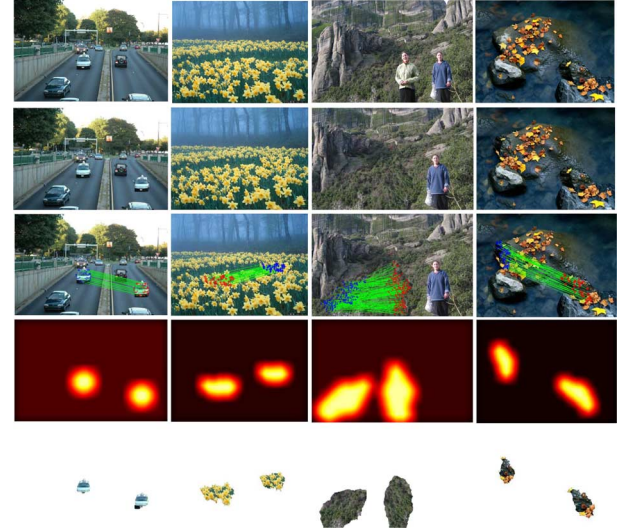


Fig. 15. Application of our saliency map. From top to bottom: original images, fake images, SIFT matching, fake saliency map, and segmentation using our method.

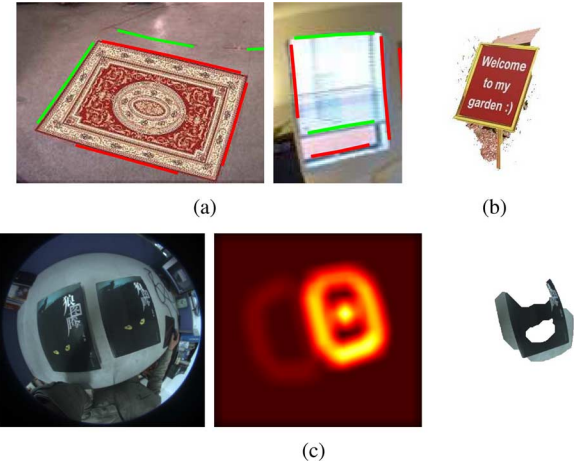


Fig. 16. Failure cases of our method. (a) Misdetection of forgery line, where missed forgery lines are labeled by green. (b) Failure performance of segmentation, where original background is not segmented well. (c) Failure case caused by oversized object.

shows the samples of this degenerated case, where the misdetected lines satisfy the geometry constraint and are labeled by the green. As described in Section V-C2, however, it is difficult to forge all the lines on the forgery object to coincide the non-linear distortion. Therefore, the forgery object could be located and segmented using our fake saliency map.

The other case of degeneration exists on the procedure of segmentation. Since our fake saliency map is generated based on forged seeds (such as lines in this paper), and for covering the region, a series of operations is used as described in Section IV-B. Some background near the forgery seed also has a high saliency weight in (21). When the appearance of them is similar (e.g., the complex background and weak boundary), some details of the forgery object could not be segmented accurately. Fig. 16(b) shows the samples of this case, where the segmented object involves the original parts. However, the main part of the forgery object is extracted by our method.

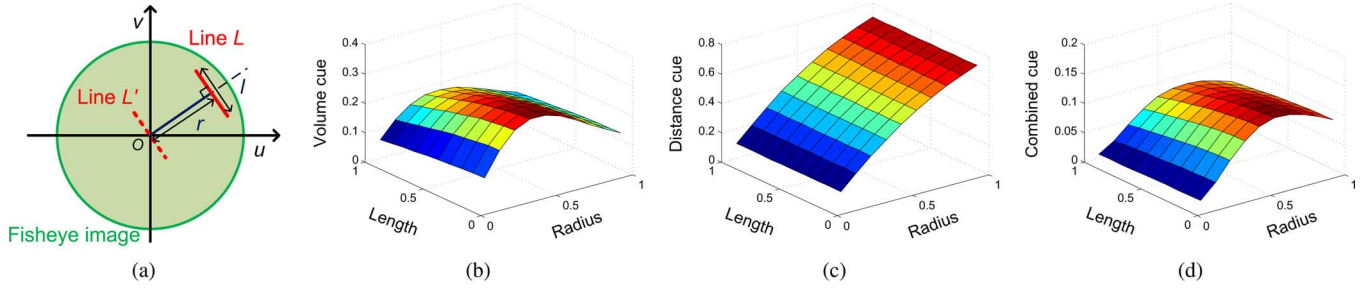


Fig. 17. (a) Configuration of synthetic line simulation. (b)–(d) Performances of distortion cues with respect to varied length l and radius r of the line.

Fig. 16(c) shows other segmentation failures, where the forgery object is oversized to be covered by the fake saliency map. The region between two forgery lines could not be filled by our method, and only the boundary of the forgery object could be extracted by our segmentation.

The last case is weak distortion of the lens. Since our detection method is based on distortion constraint, the region near the image center only provides mild distortion, which weakens the distortion constraint and leads our method to be degenerated. Simultaneously, this region near the image center can be treated as the case for general lens with a weak radial distortion. For better illustrating this weakly distorted case, we design a new simulation, where a synthetic line L is generated in a distorted image with image radius $R = 1$, as shown in Fig. 17(a). In this experiment, two parameters of the line are employed, the length l of the line and the radius r of the line with the distorted center o . We ignore the radial angle of the line and consider the line is straight in the image and perpendicular to the radial direction. Obviously, without any distortion, this synthetic line is a forgery line unless it crosses the image center as the line L' . Moreover, with the radius r increasing, the distortion is more stronger and the cues of the synthetic line are more likely to deviate from the original's cue. The ranges of length l and radius r are limited in $[0.1, 1]$ and $[0, 1]$, respectively. Fig. 17(b) shows the result of the volume cue. With the $\|1_f\|^3$ in (11), the volume cue is not influenced by the length l . But the volume cue does not increase monotonously as the function of the radius r , where the volume cue reaches the peak at $r = 0.5$, and then descends with the radius r increasing. With respect to the volume cue, the distance cue increases almost linearly as the radius r increases, as shown in Fig. 17(c). Fig. 17(d) shows the result of a combined cue, where the combined cue avoids the degeneration caused by the large radius r and inherits the robustness to noise and good performance from the volume cue. Unfortunately, the combined cue is still degenerated to zero with the radius $r = 0$, and it observes that performance of our method is degenerated for the weak distortion case.

VI. CONCLUSION

In this paper, we have proposed a geometric-based method for detecting a forgery object via the lens distortion. The geometry constraint on the viewing sphere of radial distortion projection model is employed to support two bottom-up cues. Furthermore, a fake saliency map is generated by untrustworthy likelihood, to segment the forgery object. The experiments demonstrated the performance of our method in both simulated data and real

images. The applications showed that potential usage of our fake saliency map in other common digital images.

REFERENCES

- [1] F. Devernay and O. Faugeras, "Straight lines have to be straight," *Mach. Vision Appl.*, vol. 13, no. 1, pp. 14–24, 2001.
- [2] D. Claus and A. Fitzgibbon, "A rational function lens distortion model for general cameras," in *CVPR*, 2005, vol. 1, pp. 213–219.
- [3] J. Kannala and S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 8, pp. 1335–1340, Aug. 2006.
- [4] H. Liu, O. Javed, G. Taylor, X. Cao, and N. Haering, "Omni-directional surveillance for unmanned water vehicles," in *Proc. Eighth Int. Workshop Visual Surveillance*, Marseilles, France, 2008.
- [5] H. Fu, Z. Cao, and X. Cao, "Embedded omni-vision navigator based on multi-object tracking," *Mach. Vision Appl.*, vol. 22, no. 2, pp. 349–358, 2011.
- [6] A. W. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in *CVPR*, 2001, vol. 1, pp. 125–132.
- [7] R. Hartley and S. B. Kang, "Parameter-free radial distortion correction with center of distortion estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1309–1321, Aug. 2007.
- [8] J. Tardif, P. Sturm, M. Trudeau, and S. Roy, "Calibration of cameras with radially symmetric distortion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 9, pp. 1552–1566, Sep. 2009.
- [9] Z. Kukelova and T. Pajdla, "A minimal solution to radial distortion autocalibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2410–2422, Dec. 2011.
- [10] B. Mičušik and T. Pajdla, "Estimation of omnidirectional camera model from epipolar geometry," in *CVPR*, 2003, vol. 1, pp. 485–490.
- [11] J. Wei, C. Li, S. Hu, R. Martin, and C. Tai, "Fisheye video correction," *IEEE Trans. Vis. Comput. Graphics*, to be published.
- [12] S. Lyu and H. Farid, "How realistic is photorealistic?," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 845–850, Feb. 2005.
- [13] H. Farid, "Image forgery detection," *IEEE Signal Process. Mag.*, vol. 26, no. 2, pp. 16–25, Feb. 2009.
- [14] A. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of re-sampling," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 758–767, Feb. 2005.
- [15] Z. Fan and R. L. de Queiroz, "Identification of bitmap compression history: JPEG detection and quantizer estimation," *IEEE Trans. Image Process.*, vol. 12, no. 2, pp. 230–235, Feb. 2003.
- [16] W. Luo, J. Huang, and G. Qiu, "JPEG error analysis and its applications to digital image forensics," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 3, pp. 480–491, Jun. 2010.
- [17] M. Johnson and H. Farid, "Exposing digital forgeries through chromatic aberration," in *Proc. 8th ACM Workshop Multimedia and Security*, 2006, pp. 48–55.
- [18] A. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3948–3959, Oct. 2005.
- [19] A. Swaminathan, M. Wu, and K. Liu, "Nonintrusive component forensics of visual sensors using output images," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 1, pp. 91–106, Feb. 2007.
- [20] H. Cao and A. Kot, "Accurate detection of demosaicing regularity for digital image forensics," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 4, pp. 899–910, Aug. 2009.
- [21] A. Swaminathan, M. Wu, and K. Liu, "Digital image forensics via intrinsic fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 101–117, Feb. 2008.

- [22] M. Johnson and H. Farid, "Exposing digital forgeries in complex lighting environments," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 450–461, Jun. 2007.
- [23] Q. Liu, X. Cao, C. Deng, and X. Guo, "Identifying image composites through shadow matte consistency," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 1111–1122, Jun. 2011.
- [24] M. Johnson and H. Farid, Metric measurements on a plane from a single image Department of Computer Science, Dartmouth College, Hanover, NH, Tech. Rep., 2006.
- [25] W. Zhang, X. Cao, Y. Qu, Y. Hou, H. Zhao, and C. Zhang, "Detecting and extracting the photo composites using planar homography and graph cut," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 3, pp. 544–555, Jun. 2010.
- [26] R. Benosman and S. B. Kang, *Panoramic Vision: Sensors, Theory, and Applications*. New York: Springer Verlag, 2001.
- [27] C. Geyer and K. Daniilidis, "Catadioptric projective geometry," *Int. J. Comput. Vision*, vol. 45, no. 3, pp. 223–243, 2001.
- [28] C. Hughes, P. Denny, M. Glavin, and E. Jones, "Equidistant fish-eye calibration and rectification by vanishing point extraction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2289–2296, Dec. 2010.
- [29] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," in *ECCV*, 2000, pp. 445–461.
- [30] X. Ying and Z. Hu, "Can we consider central catadioptric cameras and fisheye cameras within a unified imaging model?," in *ECCV*, 2004, vol. 3021, pp. 442–455.
- [31] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet, "A generic fisheye camera model for robotic applications," in *Proc. Int. Conf. Intelligent Robots Systems*, 2007, pp. 1683–1688.
- [32] X. Ying and Z. Hu, "Catadioptric camera calibration using geometric invariants," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 10, pp. 1260–1271, Oct. 2004.
- [33] X. Ying, Z. Hu, and H. Zha, "Fisheye lenses calibration using straight-line spherical perspective projection constraint," in *ACCV*, 2006, pp. 61–70.
- [34] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Oxford, U.K.: Cambridge University Press, 2004.
- [35] V. Ferrari, L. Fevrier, F. Jurie, and C. Schmid, "Groups of adjacent contour segments for object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 1, pp. 36–51, Jan. 2008.
- [36] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "From contours to regions: An empirical evaluation," in *CVPR*, 2009, pp. 2294–2301.
- [37] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Dec. 1998.
- [38] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 353–367, Feb. 2011.
- [39] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in *CVPR*, 2011, pp. 409–416.
- [40] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut": Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004.
- [41] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.
- [42] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, Feb. 2004.
- [43] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, Sep. 2004.
- [44] X. Pan and S. Lyu, "Region duplication detection using image feature matching," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 4, pp. 857–867, Aug. 2010.
- [45] C. Zhang, X. Guo, and X. Cao, "Duplication localization and segmentation," *Advances in Multimedia Information Processing*, vol. 6297, pp. 578–589, 2010.
- [46] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, and G. Serra, "A SIFT-based forensic method for copy/move attack detection and transformation recovery," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 1099–1110, Jun. 2011.



Huazhu Fu received the B.S. degree in mathematical science from Nankai University, in 2006, and the M.E. degree in electromechanical engineering from Tianjin University of Technology, in 2010. He is currently working toward the Ph.D. degree at the School of Computer Science and Technology, Tianjin University.

He has been working as a Research Assistant in the Computer Vision Lab since April 2010. His research interests include: image forensic, segmentation, and multiple image correspondence.

He received the Best Thesis Award from Tianjin University of Technology, in 2010.



Xiaochun Cao received the B.E. and M.E. degrees, both in computer science, from Beihang University (BUAA), Beijing, China. He received the Ph.D. degree in computer science from the University of Central Florida, Orlando.

After graduation, he spent about three years at ObjectVideo Inc. as a Research Scientist. Since August 2008, he has been with Tianjin University, China, where he is currently a Professor of Computer Science. He has authored and coauthored over 50 peer-reviewed journal and conference papers.

In 2004 and 2010, Dr. Cao was the recipient of the Piero Zamperoni best student paper award at the International Conference on Pattern Recognition. He has been in the organizing and the technical committees of several international colloquia. His dissertation was nominated for the university-level award for Outstanding Dissertation.