FORGERY DETECTION FROM SHADOW INCONSISTENCIES IN 3D-SENSOR IMAGES

Azmi Haider, Hagit Hel-Or

University of Haifa, Department of Computer Science, Haifa, 31905 Israel

ABSTRACT

Image Forgery detection, aims at finding evidence of image manipulation. It is widely studied in 2D images and videos. However, with the increase in availability and use of 3D cameras (cameras with depth sensors), it has become necessary to deal with forgery in depth images as well. In this paper, we show that inherent characteristics of the 3D camera, together with physical rules and models of the scene can be exploited for detecting forgery and manipulations in 3D images as well as determining the source camera.

Index Terms— Image Forensic, Image Forgery, 3D Cameras, depth sensors, depth images

1. INTRODUCTION

With the wide availability of image processing and manipulation software, tampering and abuse of images have become ubiquitous. This has consequential effects as images are often used as legal evidence, in surveillance systems, in medical records, as news items and on social media where their influence is at times alarming. It is thus unsurprising that the field of Image Forensic, and with it the notion of Image Forgery and its detection, has become of significant importance.

The field of digital image forensics must keep up with the ever changing technologies. Recent years has seen an increase in the use and availability of 3D cameras (cameras with depth sensors). They are already in use in medical applications, security systems, cinematography, art production and many more. These cameras output a stream of depth images in which pixel values represent the distance from camera. The notion of forgery detection should be expanded to deal with this new emerging media as its necessity in the near future for judicial issues, copyright and ownership is unquestionable.

In this paper, we show that inconsistencies of the scene together with characteristics of 3D cameras can be exploited for detecting forgery and manipulations in depth images.

2. BACKGROUND

2.1. 2D image Forgery and its detection

Image forgery detection and image authentication aim at finding evidence of image manipulation following its acquisition by the imaging sensor. Often, these algorithms are able to point out the type of manipulation that was performed and its location within the image. Image forgery detection, typically aims at one of the following:

- **Image authentication** - in which evaluation is performed to verify that no modification has been introduced in the image. Output is a measure of authenticity, often a binary output - authentic or not.
- **Image forgery detection** - in which the goal is to determine whether the original image has been manipulated (copy-paste, cropping, tone manipulation and more). Outcome typically includes the type of forgery detected as well as the suspected image regions.
- **Image signature and camera source identification** - in which the source of the image, namely, the specific camera used to acquire the image is determined, or distinguished from other cameras.

We focus on passive image forgery detection that does not rely on watermarks [1] or inherent markings within the image. The passive approaches can be categorized based on the assumptions they rely on:

1. **Physical laws expressed in the scene** - Rules of physics should be preserved under their projection into the image. Inconsistencies in these rules form a basis for forgery detection such as: object size, [2], lighting direction [3] and shadows [4].
2. **Statistics of the source images** - Statistics of local features in the image can be used to detect various manipulations including Copy-Paste forgery as well as detecting source camera. Approaches are based on color features, DCT and Wavelet coefficients, moments , local feature descriptors and more. See review of these approaches in [5, 6].
3. **Inherent characteristics of the camera** - Camera components, and imaging pipeline leave identifying signatures in the image which can be used to detect forgery as well as determine the source camera. Examples include radial distortion [7], vignetting [8], chromatic aberration [9], sensor noise patterns[10], Camera Response Function [11], demosaicing [12], White Balancing [13] and JPEG compression [14].

Whereas image forgery detection is well studied in 2D (see [15, 16, 17] for surveys), its concern with depth images is novel. An initial study in this field was presented in [18] where 3D Camera sensor noise was exploited to determine copy-paste forgery as well as determine source camera. The studies thus based on statistics of the source images (with underlying cause due to camera characteristics).

In this paper we propose an image forgery detection approach that relies on physical rules and models of the scene combined with the inherent characteristics of the 3D camera.

2.2. 3D Imaging

In this research we discuss image forgery detection methods that are specifically targeted towards images acquired by 3D cameras. This is a novel field of research and will most likely become an important field as 3D consumer cameras become popular and depth images become ubiquitous.

Depth sensing by 3D cameras

The output of a 3D camera is a sequence of depth images (often referred to as 2.5D images) where pixel values represent distance from the camera. 3D cameras differ in the method by which the image is acquired (Figure 1):

- Stereo imaging [19, 20] - consists of two 2D cameras positioned along a baseline that simultaneously capture two views of the scene. Using point correspondences between the views, depth can be computed.
- Structured light [21, 22] - an IR pattern is projected onto the scene and is captured by a calibrated IR imaging sensor. Point correspondences between IR projector and IR sensor are computed and triangulation is used to compute the depth.
- Time of flight (ToF) [23, 24] - An IR beam is projected onto the scene and an IR sensor captures the reflected light. Distance to points in the scene are computed by measuring the time difference between projected and reflected IR beams.

3. SHADOWS IN 3D IMAGES

3D sensing systems are based on two components: two RGB sensors in stereo cameras,an IR projector and IR sensor in structured light and ToF cameras. A shadow forms in a depth image when one component sees a portion of the scene while it is blocked from view by the second component (Figure 2). This typically happens at edges of objects in the scene. With ToF cameras, shadows form in the resulting depth image when the IR sensor sees a portion of the scene that is blocked for the IR projector. 3D cameras typically report shadows in the output depth images as pixels with 0 depth (thus showing up as black pixels in depth images). See examples in Figure 4. We show that shadow size is a function of object distance from sensor, background distance from camera and camera configuration.

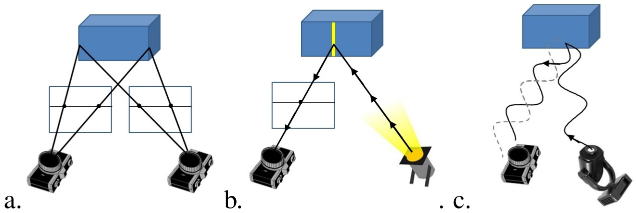

Fig. 1. Depth Sensing by 3D cameras. a) Passive stereo b) Structured light c) Time of Flight.

3.1. Shadow size

Consider the scene and camera configuration shown in Figure 2 viewed as a projection onto the x-z plane of the camera coordinate system. The scene contains an object with a background object behind it. A 3D camera records the scene. Denote by b the camera baseline, i.e. the distance between the 2 camera components (IR projector and IR sensor). Let d_o be the distance from camera sensor to the object and d_b be the distance from camera sensor to the background. Let s denote the shadow width. We consider only the horizontal geometry along the x-axis). Using simple geometry and triangle similarity, we have:

$$\frac{s}{d_b - d_o} = \frac{b}{d_o} \qquad (1)$$

It can be seen that the shadow width is independent of the horizontal positioning of the object. Thus shadow width remains the same when the object is moved horizontally in the scene. However, when the object changes position in depth from d_{o1} to d_{o2} (Figure 3) shadow width $s1$ changes to $s2$ given by :

$$S_2 = \frac{d_{o1}}{d_{o2}} \frac{(d_b - d_{o2})}{(d_b - d_{o1})} * S_1 \qquad (2)$$

Since the camera output is a 2D projection of the scene onto an image where each pixel value denotes depth, we consider the size of the object's projection and that of its shadow. Using a standard camera model where object size is inversely related to its distance d_o from the camera sensor we have that the size $w1$ and $w2$ of two objects at distance d_{o1} and d_{o2} from the camera are related by:

$$w_2 = \frac{d_{o1}}{d_{o2}} * w_1 \qquad (3)$$

Finally, we map world units to image pixel units to allow measurements in the camera output image. Let α be the horizontal field of view angle of the camera, N_x be the horizontal resolution of the image (i.e. the number of pixels per row of the camera sensor), and d be a distance from the camera sensor in millimeters (Figure 3). the conversion between millimeters and pixels at depth d is given by:

$$1\ [pixel] = \frac{2d}{N_x} * tan\frac{\alpha}{2}\ [millimeter] \qquad (4)$$

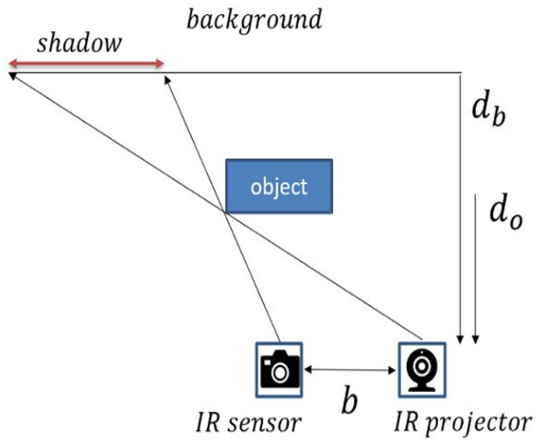

Fig. 2. Shadow is the area seen by the IR sensor with no patterns projected by the IR projector.

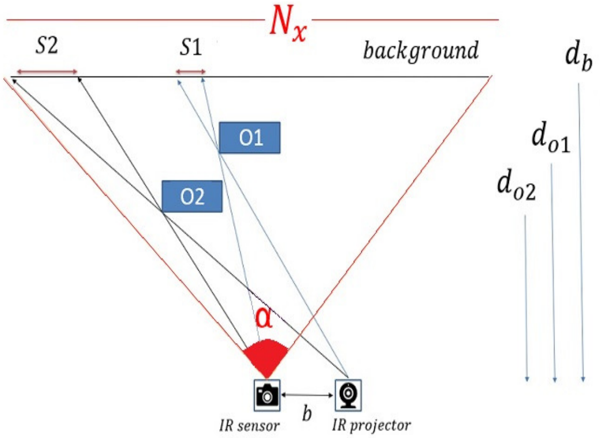

Fig. 3. Shadow size depends on object distance from camera.

To verify these calculations, we used an Intel D435 camera, and captured scenes with different object and background distances. The shadow width was measured and compared with the expected value calculated from Equations 1 and 4 using the camera parameters ($b = 50mm$, $N_x = 1280$ and $\alpha = 90°$). Table 1 shows the comparison results. It was found that the mean error between measured and estimated is 2.167 pixels.

4. FORGERY DETECTION FROM SHADOWS

We show that forgery can be detected in depth images by measuring shadow width. Specifically, forgery is suspected when measured shadow width in the image does not match the theoretically expected value. In the following we describe several principles of forgery detection.

4.1. Shadow direction based forgery detection

For most depth sensing cameras, the IR projector is positioned to the right of the IR sensor, thus object shadows are formed primarily on the left edges of objects. Object shadows appearing on other sides of the object indicate the image has been tampered with.

d_b	d_o	estimated	measured
730	450	27	32
880	600	18	22
1030	750	12	15
830	330	59	62
950	450	38	42
1100	600	24	24
1250	750	17	17
930	330	63	63
1050	450	41	42
1200	600	26	25
1350	750	19	20

Table 1. Shadow size measurement vs calculated.

4.2. Shadow size based forgery

4.2.1. Scaling depth values

The simplest form of depth image forgery is to scale the depth values of the object pixels. In this type of forgery the object and the shadow size do not change. However, the scene parameters (d_o, d_b) do, thus affecting the calculated shadow size. Figure 4 shows an example. The original image (Figure 4a) shows an object at distance 900mm with background at distance 1340mm. Shadow width is measured at 12pixels where the calculated width, given the camera parameters is 11.7 pixels. Figure 4b shows the forged image where depth values of the object were scaled to 500mm. In this case the shadow width remains 12pixels whereas the calculated width is 40.11 pixels. Clearly, implying forgery.

4.2.2. Scaling depth values with scale in size

The simple forgery of depth value scaling does not take object size into account and so produces object size inconsistency which in itself can hint at forgery [2]. However, even when corrected for object size, shadow width is still inconsistent. Figure 4c shows the object with depth values scaled to 500mm and object together with shadow increased in size consistently with depth using Equation 3. Shadow width is thus scaled to 25 pixels which is still inconsistent with the calculated with of 40.11 pixels and forgery is detected here as well. This inconsistency in shadow width even with object re-sizing can be derived directly from Equations 1-3.

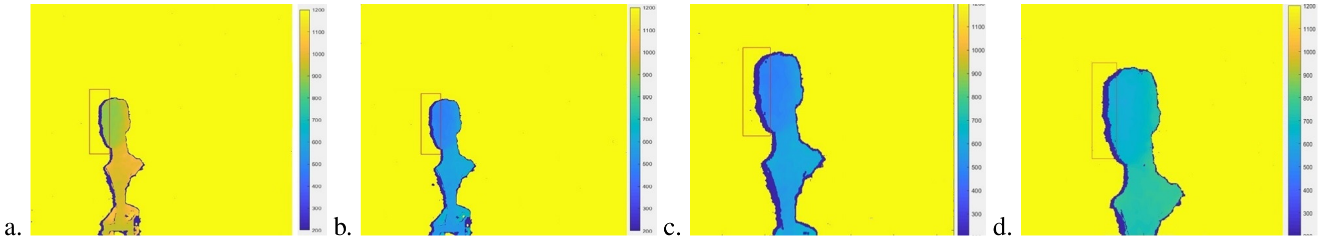

Fig. 4. Forged depth images. a) Original. Object distance is 900mm, background distance is 1340mm (shadow calculated:11.675, measured:12) b) Depth values scaled to 500mm (shadow calculated:40.11, measured:12). c) Depth values scaled to 500mm + object re-sized (shadow calculated:40.11, measured:25). d) Depth values scaled to 500mm + object re-sized to correct for shadow width (shadow calculated:40.11, measured:40).

4.2.3. Scaling to correct Shadow width

Finally, if the object and shadow are re-sized to obtain the shadow width as computed for the scaled depth value (40.11mm), as shown in Figure 4d, we find inconsistency in object size with real world size. In the example, reverse-projecting the head in the image (from pixels to millimeters) shows a head size of 230mm (while real human head is on average ∼140mm).

4.3. Shadow size as a source camera identifier

Shadow size can also be used as a source camera identifier since cameras differ in build, specifically in their baselines and fields of view which in turn, affect shadow size. Furthermore, same camera with multiple resolution modes will generate different shadow sizes for the same scene, allowing the detection of the specific resolution. Figure 5 shows a scene captured by 3 structured light cameras: Intel D415, Intel D435 and Kinect V1. Camera parameters are : $\{b = 55mm, \alpha = 70°, N_x = 1280\}$, $\{b = 50mm, \alpha = 90°, N_x = 1280\}$ and $\{b = 75mm, \alpha = 57.8°, N_x = 320\}$ respectively.

To test for camera source identification, we collected a set of depth images using the 3 cameras with the object at different depths. For each image, the measured shadow width was compared with the 3 possible shadow widths calculated using each of the cameras' parameters. The closest to the measured shadow width indicated the source camera. Results were 100m% correct as shown in Table 2.

d_b	d_o	measured	D435	D415	KinectV1
930	330	63	63	98	85
1050	450	42	41	63	55
1350	750	20	19	30	26
1410	940	10	19	29	12
1810	1125	5	11	17	7
1100	600	36	24	38	16
1030	750	18	11	18	7
1350	750	29	18	30	12

Table 2. Camera Source detection.The measured shadow size is compared with the 3 shadow sizes calculated using each camera's parameters. True source camera is marked in gray.

5. CONCLUSION

In this paper, we exploited characteristics of the 3D camera, together with scene parameters to determine shadow width within depth images. We show that width depends on scene parameters such as distance of object and background, and on camera parameters such as baseline, angle of view and resolution of the camera sensor. Thus, manipulation of object depth, or manipulation of object size can be detected. Additionally, due to differences in camera parameters, shadow size can be used to determine source camera. This study is one of the first to deal with forgery detection in depth images. The importance in dealing with forgery detection in this new emerging media is significant due to its use in judicial issues, security, medical imaging, art and more.

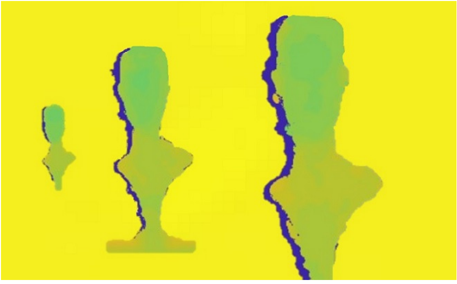

Fig. 5. Same object captured from three cameras at the same position. KinectV1, Intel D435, Intel D415 respectively.

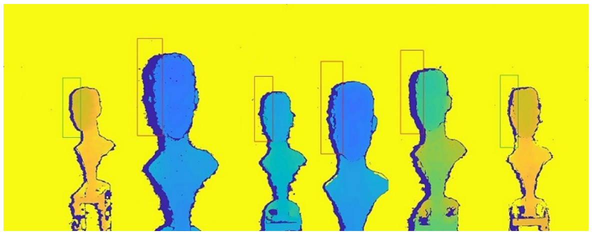

Fig. 6. Which is real and which is forged?

6. REFERENCES

[1] Ingemar Cox, Matthew Miller, Jeffrey Bloom, Jessica Fridrich, and Ton Kalker, *Digital Watermarking and Steganography*, Morgan Kaufmann Publishers Inc., San Francisco, 2 edition, 2008.

[2] Lin Wu, Xiaochun Cao, Wei Zhang, and Yang Wang, "Detecting image forgeries using metrology," *Machine Vision and Applications*, vol. 23, no. 2, pp. 363–373, 2012.

[3] Eric Kee and Hany Farid, "Exposing digital forgeries from 3-d lighting environments," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, Dec 2010, pp. 1–6.

[4] Eric Kee, James F. O'Brien, and Hany Farid, "Exposing photo manipulation with inconsistent shadows," *ACM Transactions on Graphics*, vol. 32, no. 4, pp. 28:1–12, 2013.

[5] Xunyu Pan, "Digital image forensics with statistical analysis," in *Handbook of Digital Forensics of Multimedia Data and Devices*, Anthony T. S. Ho and Shujun Li, Eds. Springer-Verlag, 2015.

[6] Vincent Christlein, Christian Riess, Johannes Jordan, Corinna Riess, and Elli Angelopoulou, "An evaluation of popular copy-move forgery detection approaches," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1841–1854, 2012.

[7] H. R. Chennamma and Lalitha Rangarajan, "Image splicing detection using inherent lens radial distortion," *International Journal of Computer Science Issues*, vol. 7, no. 6, pp. 149–158, 2010.

[8] Siwei Lyu, "Estimating vignetting function from a single image for image authentication," in *ACM Workshop on Multimedia and Security*, 2010, pp. 3–1.

[9] Ido Yerushalmy and Hagit Hel-Or, "Digital image forgery detection based on lens and sensor aberration," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 71–91, 2011.

[10] Lit-Hung Chan, Ngai-Fong Law, and Wan-Chi Siu, "A two dimensional camera identification method based on image sensor noise," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2012, pp. 1741–1744.

[11] Yu-Feng Hsu and Shih-Fu Chang, "Image splicing detection using camera response function consistency and automatic segmentation," in *IEEE International Conference on Multimedia and Expo*, 2007.

[12] Alin C. Popescu and Hany Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3948–3959, 2005.

[13] Zhonghai Deng, Arjan Gijsenij, and Jingyuan Zhang, "Source camera identification using auto-white balance approximation," in *IEEE International Conference on Computer Vision (ICCV)*, Nov 2011, pp. 57–64.

[14] Giuseppe Valenzise, Marco Tagliasacchi, and Stefano Tubaro, "Revealing the traces of jpeg compression anti-forensics," *Information Forensics and Security, IEEE Transactions on*, vol. 8, no. 2, pp. 335–349, 2013.

[15] Anthony T. S. Ho and Shujun Li, *Handbook of Digital Forensics of Multimedia Data and Devices*, Wiley-IEEE Press, 2015.

[16] Alessandro Piva, "An overview on image forensics," *ISRN Signal Processing*, vol. 2013, no. 496701, 2013.

[17] Hany Farid, "A survey of image forgery detection," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 16–25, 2009.

[18] Noa Privman-Horesh, Azmi Haider, and Hagit Hel-Or, "Forgery detection in 3d-sensor images," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.

[19] Daniel Scharstein and Richard Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, pp. 7–42, 2002.

[20] Lynne L. Grewe and Avinash C. Kak, "Stereo vision," in *Handbook of Pattern recognition and Image Processing: Computer Vision*, Tzay Y. Young, Ed., pp. 239–317. Academic Press, 1994.

[21] Jason Geng, "Structured-light 3d surface imaging: a tutorial," *Advances in Optics and Photonics*, vol. 3, no. 2, pp. 128–160, 2011.

[22] Barak Freedman, Alexander Shpunt, Meir Machline, and Yoel Arieli, "Depth mapping using projected patterns, us patent number 20080240502," 2012.

[23] Sergi Foix, Guillem Aleny, and Carme Torras, "Lock-in time-of-flight (ToF') cameras: A survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917–1926, 2011.

[24] Miles Hansard, Seungkyu Lee, Ouk Choi, and Radu Horaud, *Time of Flight Cameras: Principles, Methods, and Applications*, Springer-Verlag London, 2012.