Forgery Detection in Depth Images

Azmi A. Haider

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE MASTER'S DEGREE

University of Haifa
Faculty of Social Sciences
Department of Computer Science

January, 2020

Forgery Detection in Depth Images

By: Azmi A. Haider
Supervised by: Dr Hagit Hel-Or

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE MASTER'S DEGREE

University of Haifa
Faculty of Social Sciences
Department of Computer Science

January, 2020

Approved by: _____ Date: _____
(Supervisor)
Approved by: _____ Date:_____
(Chairperson of M.A Committee)

Acknowledgement

I would like to express my sincere appreciation to my supervisor Prof. Hagit Hel-Or for her constant guidance and support during this research. I am truly grateful.

I would also like to express my heartfelt gratitude to my colleagues at Intel Haifa, Mr. Yair Findling, Mr. Shlomo Marciano, Mr. Or Machlav, Dr. Ido Nissenboim, for their moral support and understanding during this time. Without which I would not have been able to complete my research and intern at the same time.

Last but not least, I would like to thank my family and friends for putting up with me for this long. Especially my mother.

Contents

Forgery Detection in Depth Images

Azmi Haider

Abstract

The field of Image Forensic, and with it the notion of image forgery and its detection, is widely studied in 2D images and videos. With the increase in availability and use of cameras with depth sensors, it has become necessary to consider forgery detection in depth-images as well. In this research, we present an introductory study of forgery detection in depth-images. Specifically, we show that noise statistics in depth-images can be exploited for camera source identification, image forgery detection. We further show that scene illumination can be used to detect forgery. Finally, we show that inherent characteristics of the camera mechanics can be exploited to determine image forgery from sensor-based shadows.

List of Tables

List of Figures

1 Introduction

The art of image forgery is over a century old and 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 criminal investigation, surveillance systems, medical records, and 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 depth cameras (cameras with depth sensors which outputs a stream of depth images in which pixel values represent the distance from camera). This technology is already in use in medical applications, security systems, cinematography, autonomous navigation systems and many more. Thus, the notion of forgery detection should be expanded to deal with this new emerging media, as its upcoming necessity for judicial issues, copyright and ownership is unquestionable. In this research we present several image forgery detection methods that are specifically targeted towards depth images without the use of RGB images. The proposed methods rely on camera sensor characteristics, scene characteristics, and camera build.

We show how the depth camera's sensor noise can be exploited to determine copy-paste forgery as well as determine source camera and even depth reconstruction from noise. We further expanded this approach to determine forgery using noise statistics to detect inconsistencies in scene illumination. In addition, this method was also applied to a real life application of 3D face anti-spoofing. Another approach we propose relies on physical rules and models of the scene combined with the inherent characteristics of the depth camera. We considered sensor shadows within the scene as a basis for forgery detection.

To the best of our knowledge this is the first rigorous study in the field of Forgery detection applied to depth images This work was published and presented in CVPR2018 workshop: The Bright and Dark Sides of Computer Vision and is currently under evaluation in ICCP2020.

2 Background

2.1 Depth imaging

The outputs of 3D cameras are typically videos or image sequences where each frame is represented as a depth image which we term 3D image (although often referred to as 2.5D images) where pixel values indicate distance from the camera (see Figure 1).

Similar to 2D cameras, 3D camera components include optics, sensors and imaging pipeline [2], however, these are tuned to obtain 3D data. Also included are additional components, unique to depth sensing (such as IR projectors and phase detectors). 3D cameras differ in the method by which the 3D image is acquired (Figure 2):

- Stereo imaging [3, 4] - a passive imaging system comprised of two or more 2D cameras positioned along a common baseline that simultaneously capture two views of the scene. Following correspondence of points between the two views, depth (distance from camera baseline) can be computed. Cameras used in this research and based on this depth acquisition technique are Intel D435 camera and ZED camera (Figure 3).

- Structured light (Projected-light sensors) [5, 6, 7, 8] - an IR pattern is projected onto the scene and forms a unique visual code for each surface point. The observed pattern points are captured by a calibrated IR imaging sensor. Correspondence between IR projector and IR sensor is computed using stereo matching methods and triangulation is used to compute the 3D position of each surface point [4]. Cameras used in this research and based on this depth acquisition technique are Intel D415 camera and Kinect V1 camera (Figure 3).

- Time of flight (ToF) [9, 7] - An IR wave is projected onto the scene and an IR sensor captures the reflected light wave. By measuring the difference between the projected and reflected IR waves, the distance to points in the scene can be computed. Two types of ToF cameras are used: Continuous wave modulation - in which the frequency of the projected IR wave is varied and the phase delay is measured

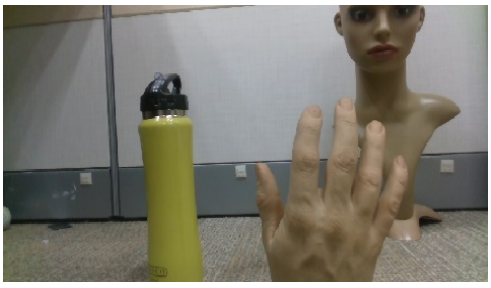
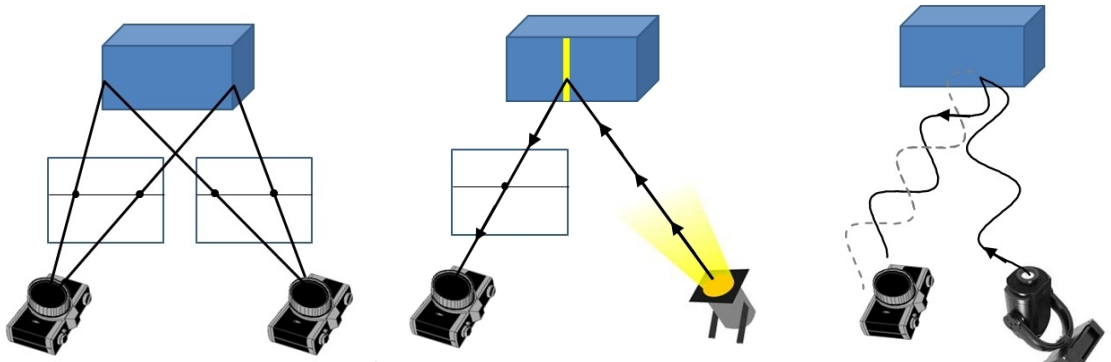
a.  b. 

Figure 1: a) RGB image b) Depth image - darker pixels indicate distances closer to the camera (captured with Intel D435)

.



a. b. . c.

Figure 2: Depth Sensing by 3D cameras. a) Passive stereo b) Structured light c) Time of Flight.

to evaluate depth. Pulse light modulation - in which a very short pulse of light is projected and the time till its return is measured. This approach avoids dealing with phase and thus overcomes issues of phase ambiguity (see below). A camera used in this research and based on this depth acquisition technique is Kinect V2 (Figure 3).

Figure 3: Cameras used. First row: Intel D415, D435. Second row: Kinect V2, V1. Third row: ZED.

2.2 2D image forensics

Image forensic is the field of 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 forensic in 2D images has been widely studied (see [10, 11, 12] for surveys).

Image forensic detection, is often aimed 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 that have been manipulated.
- **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 [13] or inherent markings within the image. The passive approaches can be categorized based on the assumptions they rely on:

1. **Physical rules and models of the scene** - Laws of physics and nature should be preserved under their projection into the image. Inconsistencies and breaking of these rules form a basis for forgery detection. Examples include inconsistencies in size of objects in the image [14, 15, 16], inconsistency of lighting directions within the image [17, 18, 19], shadow inconsistencies [20, 21, 22] and reflection inconsistencies [23, 24].
2. **Statistics of the source images** - Certain image statistics should be observed in acquired images. Forgery Detection is based on these statistics. These methods typically aim to detect Copy-Move forgery in which a portion of the original image has been copied from one region of the image and pasted into another region. Statistics of local features within the image are extracted and then analyzed, projected or sorted to efficiently extract regions suspected of being repetitive. Approaches are based on projecting raw pixel data [25], extracting color features from the raw data [26], analyzing DCT and Wavelet transform coefficients [27, 28, 29], using moments [30, 31], extracting local features such as SIFT and SURF [32, 33] and many more. Using image statistics and multiple image features, source camera of an image can also be detected [34, 35]. See review of these approaches in [36, 37].

3. **Inherent characteristics of the camera** - The acquisition system itself, namely the camera, its components and the imaging pipeline within, have been shown to leave identifying signatures within the image which can be used to detect forgery as well as determine the source camera. Each component of the camera provides unique indicators within the image. The optic system of a camera introduces geometric distortions, aberrations and lighting flow artifacts in an image. These are exploited for forgery detection based on radial distortion [38, 39], vignetting [40] and chromatic aberration [41, 42, 43]. Image sensors introduce artifacts in the image due to patterns of sensor noise, specifically fixed noise pattern which appear consistently in all images of a given camera yet differ among cameras. The noise pattern can be considered a camera signature and is exploited to determine source camera as well as to detect forgeries such as copy-move [44, 45, 46, 47].

   Dust within the sensor system can also be exploited to detect forgery [48]. Finally, the image processing pipeline that transforms sensor output to the final digital image, is a highly complex module, containing many sub-modules and algorithms and varies greatly between cameras. Numerous studies in forgery detection have been based on effects of this module on the end image. These studies mostly aim at determining the source camera of a given image, though some extend to detecting copy-move or spliced images. Examples include methods based on the Camera Response Function [49, 50, 51, 52], the Demosaicing technique [53, 54, 55, 56], the White Balancing method [57] and on JPEG compression artifacts [58, 59, 60, 61].

For a general review of 2D image forgery detection techniques see [10, 11, 62, 12].

2.3 3D image forensics

Forgery detection and copyright verification will soon become an important issue in depth images as they are now in 2D images. Since there is a variety of depth cameras, and new and improved technologies are continuously introduced, we set our objective to supply a set of tools capable of testing forgery in depth images acquired under different cameras and under various camera settings and scene conditions. We restrict our analysis to consumer, low-cost and readily available cameras, as these are more susceptible to forgery attacks.

We consider the 3 classes of forgery detection: image authentication, image forgery detection and source camera identification (as described in Section 2.2) in the context of depth images.

We note, in this research, forgery detection is done using only depth images. RGB information can only improve our results. however, our main idea was to the capabilities of depth information.

2.4 Noise in depth images

A central aspect of our study of depth image forgery detection relies on the noise inherent in the depth camera output. Noise and errors in depth images are dependent on numerous parameters including the acquisition method used by the camera, the physical parameters of the camera (e.g. baseline for stereo and structured imaging, the space and time resolution of the phase modulating ray in the ToF cameras), the analysis algorithms used in the pipeline (correspondence methods, error correction, phase analysis), as well as scene characteristics such as position and depth of objects in the scene and scene lighting. Noise may be considered as arising from 4 sources:

Noise Inherent to the camera parameters - Camera build and technical specs such as Focal length, field of view, quality of lenses, affect image quality and consequently the noise in the depth image. In active acquisition systems, the quality of the projected IR light, including its intensity and collimation (Figure 4) affects image noise and in TOF cameras the quality of the IR signal modulation is a significant factor for noise level. Camera build parameters such as the baseline between and alignment of cameras for stereo based systems and the camera to projector distance in structured light based systems also strongly affect image depth quality [8, 5, 63, 6, 64].

Noise Inherent to depth measuring methods - Stereo and structured light rely on correspondence points at which relatively accurate depth measurements are obtained. Between these points, interpolation is used which inherently introduces depth errors [8, 5, 6]. ToF approaches can obtain depth measurements at every pixel location thus avoiding interpolation errors. ToF methods based on phase are inherently prone to phase ambiguity and demodulation errors [7] which result in erroneous depth estimates.

Noise due to scene characteristics - Both scene illumination and object positioning within the scene strongly affect accuracy of depth estimation. It has been shown that depth cameras do not perform well under strong ambient illumination, specifically outdoor lighting (see Figure 5) compared to indoor lighting. This is mainly due to the fact that natural light contains IR components that interfere with the IR lighting used by the depth cameras. Furthermore the IR light of the camera is typically of very low intensity and is over powered by the high intensity outdoor lighting [65, 66].

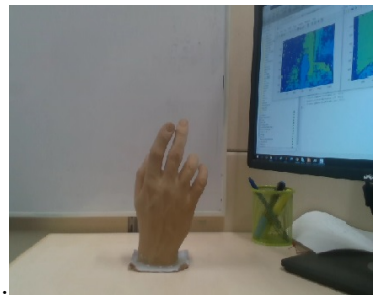
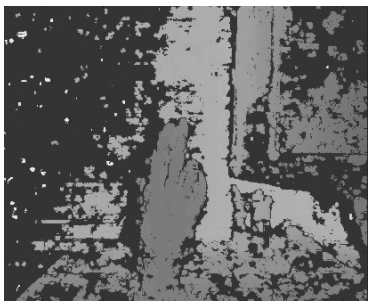
a.  b.  c. 

Figure 4: Noise is dependent on camera laser power. a) RGB image b) Low laser power c) High laser power.

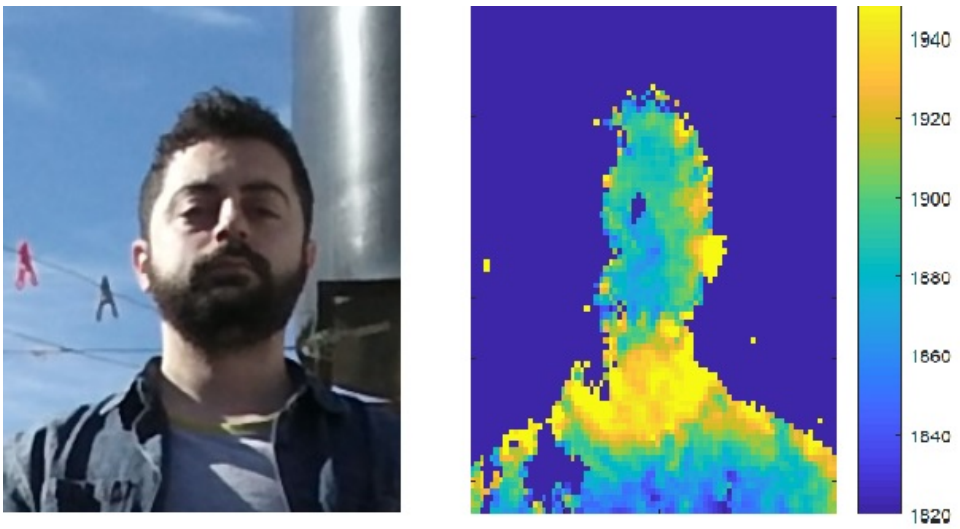


Figure 5: Strong sun light results in loss of pixels on the left side of the face.

Inaccuracy of estimated depth, often termed Axial Noise, has been shown to increase quadratically with distance of objects in the scene from the camera [67, 68, 69, 70, 71, 72, 73, 74]. It has been suggested that this is due to the relation between disparity and depth in the stereo and structured light cameras and to IR amplitude attenuation with distance in the ToF based methods. Lateral Noise increases linearly (in the x and y directions) and very extreme at the edge of the camera's field of view possibly due to lens distortion [67, 68, 71, 72, 73, 74]. Furthermore several studies have shown a radial ripple like noise that extends laterally [67, 68, 73, 74].

Temporal or Vibrating Noise measured as variance in depth values across frames at a specific pixel on a given surface, grows quadratically with depth [67, 68]. It is stronger at depth edges, shadows, specular surfaces and when motion exists in the scene. It has been shown to appear in vertical stripe patterns[67, 68, 74]. When object (or camera) motion is involved, motion blur in depth cameras results in depth over or underestimation near depth edges [75, 76, 77].

Shadow Noise arises in occluded areas of the scene into which the IR projections of the camera can not reach, thus erroneous 3D measures are obtain, if at all [78, 67]. Shadow and lateral noise increases at strong depth edges [78, 67, 79, 80] possibly due to difficulty in triangulation at these locations or due to erroneous reflected light.

Noise due to object characteristics - Studies show that color and brightness of objects affect depth estimation [69, 7] (Figure 6), however, others maintain that it does not [67]. Since materials differ in the rate of IR absorption and thus affect depth estimation [67], it is possible that color of objects is confounded with its material giving rise to the confusion.

Reflective properties of specular surfaces cause errors in depth estimation [78, 67] and depth estimation is erroneous or unavailable for transparent objects [67]. It has also been shown that temperature of the scene and its objects within strongly affect the accuracy

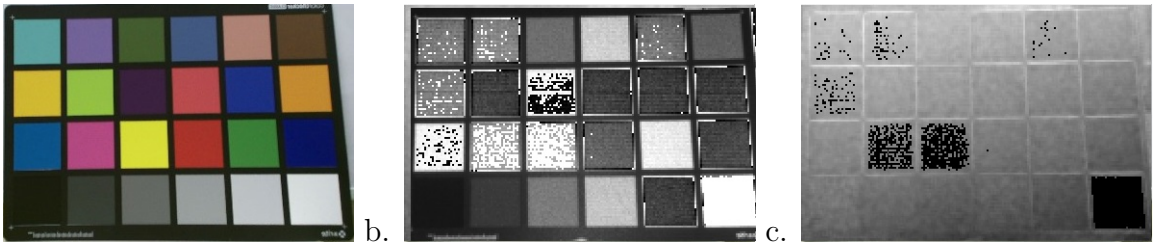


Figure 6: The effect of color in Kinect V2 cameras. A flat surfaced color checkerboard is viewed. a) RGB image b) IR image. c) Depth image.

of depth estimation [9]. It has also been shown that depth errors depend not only on the characteristics of an object but on those of the surrounding objects thus inter-reflection and light scatter from neighboring objects may give rise to erroneous depth estimates [69, 81].

As mentioned above, noise patterns in 2D images, can be considered a camera signature and is exploited to determine source camera as well as to detect forgeries such as copy-move [44, 45, 46, 47]. In this study, image noise is considered as well, albeit in depth-images. In contrast with 2D image forgery detection, where spatial noise is exploited, we use temporal noise and consider both lateral and axial noise in our noise modeling.

Figure 7a shows the depth response at a target pixel acquired by a KinectV2 camera [82] over 50 frames. We define noise as the deviation from the mean depth response at a pixel across a period of time. Figure 7b shows the histogram of deviations from the mean depth of the pixel for the given time period. *Noise magnitude* is taken as the variance of the depth values[1] and *noise variance* is defined as the variance of the absolute of deviation values. For the example in Figure 7, noise magnitude is $\mu = 0.15$, and noise variance is $\sigma^2 = 0.035$.

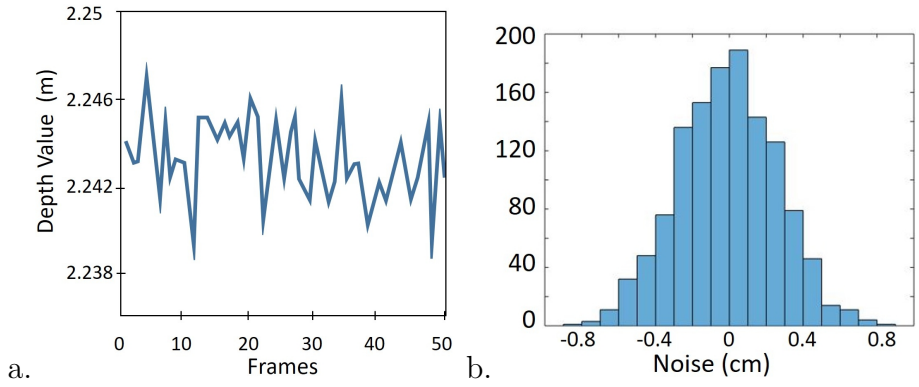


Figure 7: a) Depth response at a pixel acquired by KinectV2. b) Histogram of noise at the pixel.

[1]Taking the norm 1 variance, (Mean of Absolute), values of the deviations showed similar results in our forgery detection analysis.

3 Noise Data collection

For our noise-based forgery detection analysis, we collected a set of noise measurements by placing a target board at a lattice of positions varying in depth (z-value) between 120cm and 400cm at 40cm intervals and in horizontal positioning (x-positions) at intervals of 40cm extending horizontally from the center of the camera's view field up to 280cm on either side (at this stage we disregard vertical positioning). The number of target locations was dependent on the field of view of each camera. A schematic diagram of target positions is shown in Figure 8. At each target position, a 300 frame recording of the cardboard target was performed. The target formed a region of at least 20x20 pixels in each of the acquired images at a constant vertical position in all images. Acquisition was performed under this setup using cameras of 3 types: KinectV1 [83] (structured light), KinectV2 [82] (time of flight), ZED [84] (stereo). To exploit noise for forgery detection, we collected noise statistics at each target position, including: noise distribution (histogram), noise mean and variance. These measures were normalized and concatenated to form the sample's feature vector.

Similar to [67, 74], we found that noise magnitude increased with depth and with increasing deviation from the center along the horizontal direction. Figure 9 shows the noise magnitude as a function of depth (z pos) and as a function of horizontal position (x-pos) relative to the center. We exploit these characteristics and demonstrate various applications in depth image forensic, including source camera identification, forgery detection, motion path recovery and real vs fake face discrimination.

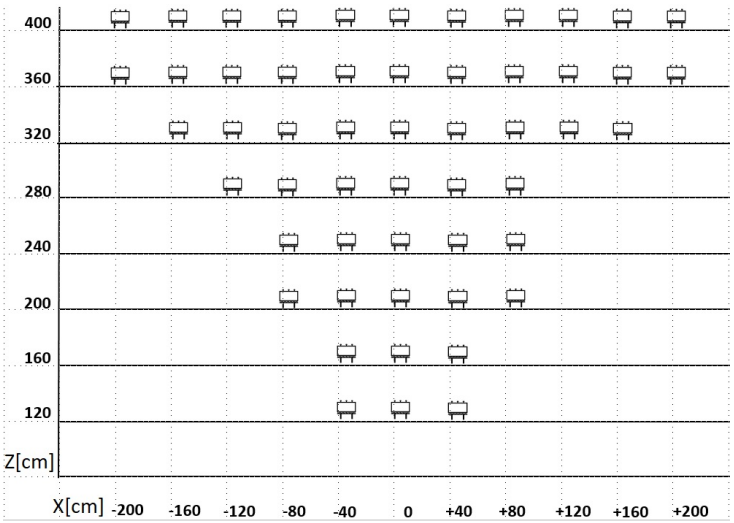


Figure 8: Map of target positions for data collection by KinectV2 (53 target positions). X is the distance from the center of camera's field of view. Z is the distance from the camera.

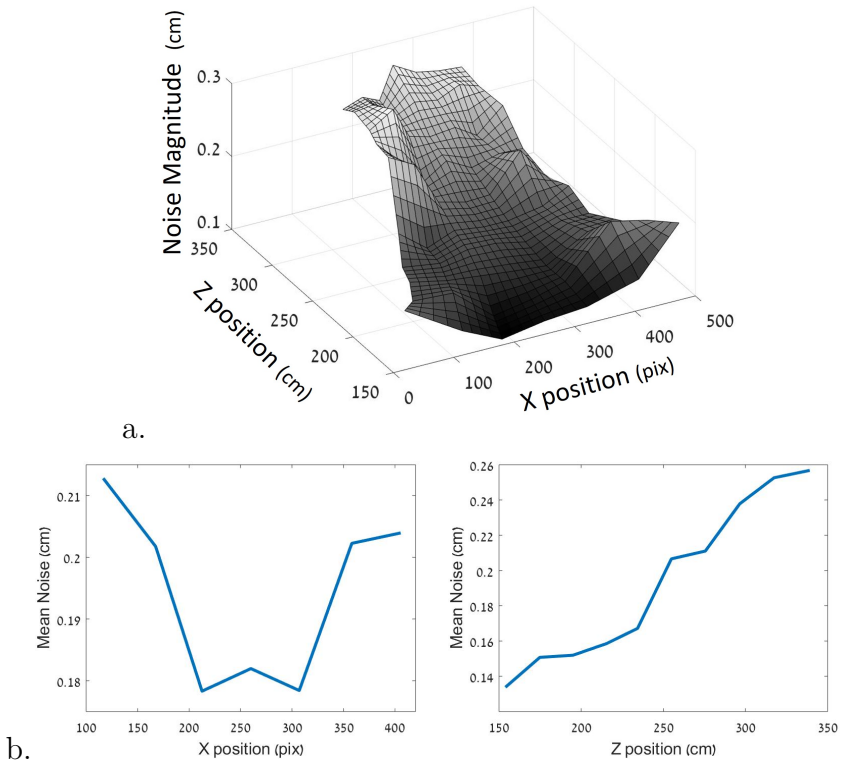


Figure 9: a) Noise magnitude as a function of x position and depth (z-pos). Noise increases with depth and with horizontal deviation from center. b) Mean Noise as a function of x position (left) and of depth (z-pos)(right).

4 Noise-based forgery detection

4.1 Source camera identification

We show that source camera identification can be performed reasonably well from noise statistics. Noise measurement data was collected, as described above, for three different types of cameras: KinectV1 (structured light), KinectV2 (time of flight) and ZED (stereo). Per each type of camera, several units were used to collect the data.

To exploit noise for source camera type detection, three data clusters were defined, one for each camera type (KinectV1,KinectV2, ZED) based on data collected from one camera unit of each type. Testing was performed on input collected from cameras not used in the training data. KNN was used to determine the source camera. Success rates are shown in Table 1. It can be seen that Kinect-V2 are very reliable in identifying source camera whereas it was found that KinectV1 data is often confused as arising from the ZED camera.

We also tested for source camera identification in the wild. A collection of 6 depth image sequences were collected from public databases and from home units [85]. Sequences were cropped to 300 frames. The sequences were analysed by randomly selecting 300 patches of size 20x20x300 and testing each for camera source. The resulting camera source was determined as the majority voting of the image patches. Table 2 shows the

Camera Unit	**% Correct Camera Type Identification**
KinectV1 (unit #1) (Training)	90
KinectV2 (unit #1) (Training)	98
ZED (training)	96
KinectV1 (unit #2)	74
KinectV1 (unit #3)	75
KinectV2 (unit #2)	92
KinectV2 (unit #3)	95

Table 1: Camera Source Identification Results

results. All examples showed over %50 of the patches correctly identified the camera, implying correct camera source identification for all sequences.

In the next test, we created 300 forged image sequences by copying random patches from a source image sequence to a random position in the target image of a different camera, within the target region (see example in Figure 10). The target was then tested by scanning over all image patches and determining source camera. If the source was found to differ from target camera it was marked as forged, and an attempt was made to detect correct camera source. Results showed that 100% of the forged images were able to detect the forged region, but only a 60% were able to determine the correct source camera of the forged region.

In order to test for detection of the specific source camera unit, we tested the ability of distinguishing between specific camera units in a set of 3 KinectV2 cameras. Noise data was collected from 3 different KinectV2 cameras and unit specific data clusters were learned. A collection of 30 additional patches were collected using these 3 cameras and their noise statistics were extracted and used to classify to one of the camera units. Classification results were compared with the true source camera. We find that only 40% of the patches were able to correctly detect the specific camera unit. This poor result is not surprising as the above results (Table 1) show that the noise statistics of a single camera well defines the noise distribution of other cameras of the same type.

Camera	**Source**	**Correctly Classified**
KinectV1	[85]	92%
KinectV1	[85]	65 %
KinectV1	[85]	68 %
KinectV2	Apt - private Cam13	95 %
KinectV2	Studio - Private Cam2	87%
KinectV2	Studio - Private Cam11	86 %
KinectV2	Office - Private Cam15	97 %

Table 2: Camera Source Identification results in the wild

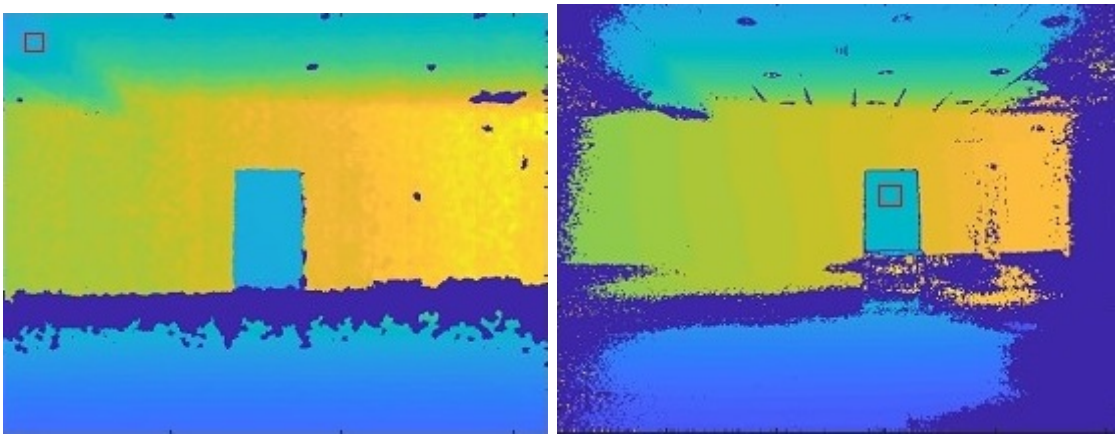


Figure 10: Forged image used for source target detection. A patch was copied from a KinectV1 sequence into a KinnectV2 sequence, (marked as square). Based on noise statistics, the forged area was detected and correct source (KinectV1) was successfully deduced.

4.2 Copy-paste and depth change forgery detection

Copy-paste forgery involves pasting an image region from a source image into a given image. In depth-images this type of forgery inherently copies the noise content of the region as well. When forging a depth-image sequence, both spatial and axial noise are copied. Another type of forgery is the *depth-change* forgery where the depth map is altered to create a forged image with the object or region at the correct xy position but at an altered depth (e.g. by constant shift of z-values). In both types of forgeries we advocate that forgery can be detected by determining that the noise associated with the given depth is incorrect. To show this, we take the test to the extreme in the sense that we completely disregard the given depth values and consider only the noise. We show that depth and x-position can be estimated from the noise alone up to a certain success rate.

A multi-class SVM classifier [86] was built based on the noise data collected from a KinectV2 camera (as described above). with depth (z-values) sampled at 30cm intervals between 1.40cm and 3.50cm distance from camera and x-positions sampled at 30cm intervals symmetrically about the the scene center position. The classes represented all possible xz-positions (81 classes). Verification was performed using N-fold analysis. Forgery test-

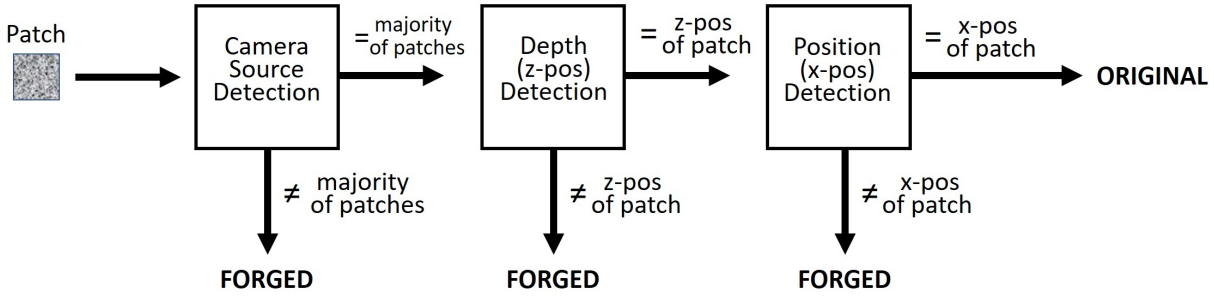


Figure 11: Forgery Detection flow diagram.

	Correct x-z prediction	Avg z-value error	Avg x-pos error
Closest match	73%	9.6cm	38.5cm
2-closest	92%	1.54cm	6.6cm
3-closest	97%	0.31cm	1.84cm
4-closest	99%	0.09cm	0.95cm

Table 3: Depth (z-value) and X-position prediction from noise.

ing was performed on 2025 test patches at all z-positions and all x-positions. Per test patch, classification to the closest xz-value position was calculated based on the patch noise statistics alone. Table 3 (first row) shows the resulting success rate of detecting the correct xz-position. The average distance error (in cm) between the classified and correct xz-position are given as well. It can be seen that success rate is not very high, implying that noise statistics are confounded between xz-positions. However the average distance error is not unreasonable. We extend our test to include the second-best classification, in which case the xz-position of the two best classes, which is closest to the true position is taken as the classification result. Table 3 lists the success rate and error distance for the second, third and fourth best classifications. It can be seen that success rate rises quickly to high values with very low average distance errors.

We tested for combined camera source identification (Section 4.1) and xz-position detection. As in Section 4.1, noise measurement data was collected in our lab, from 3 camera types (KinectV1,KinectV2, ZED). Data from a single camera unit of each type was used to learn a cluster hierarchy as shown in Figure 11, containing three data clusters, one for each camera type, subdivided into 8 sub-clusters associated with the possible z-values and further subdivided into 11 clusters associated with x-positions. Testing was performed in a cascade: first, source camera was identified, then for the successful detection, z-position was determined and for the successful cases x-position was determined. Table 4 shows the results. As can be seen success rates are high for both source identification as well as positioning.

	KinectV2	KinectV1	ZED
Source Camera Identification	98%	100%	90%
Depth Prediction (z-value)	91%	90%	99%
X-position Prediction	82%	99%	98%

Table 4: Camera Source Identification and x-z value prediction. Classification is performed in a cascading manner.

4.3 3D masks anti-spoofing

We exploit the noise in depth images to distinguish between real faces and masked faces worn by attackers in a face recognition system. For this purpose we used the 3D Mask Attack Data-set [1] which contains 17 real face sessions and 17 mask sessions (Figure 12). Each session consists of 5 videos consisting of 300 frames each (both RGB and depth, of which we used only the depth frames). As a pre-processing step, we cropped a patch of size 98x98 around the face in all videos. For each video, we generated 30 noise images as feature vectors. Each pixel in the noise image is the variance of the pixels at that location across 10 consecutive depth images. Thus, a video of 300 depth images is converted to 30 noise images and a session containing 5 videos has in total 150 noise images.

Our training set consisted of 14 real sessions, and 14 mask sessions resulting in a total of 4200 noise images. The test set consists of the remaining subjects: 3 real and 3 fake sessions totalling 900 noise images. We note that the test set is of subjects not participating in the training at all. Using a coarse Gaussian SVM (kernel scale 390, 5-fold cross validation), we were able to achieve an accuracy of 95.11% on the test set. We were able to increases the accuracy to 96.67% by voting on noise images in the same video. As mentioned earlier, each video consists of 30 noise images. For each of the 30 noise images in a video, we perform a prediction. Then based on the majority of 30 predictions we decide if the video as a whole is of a real face or a mask.

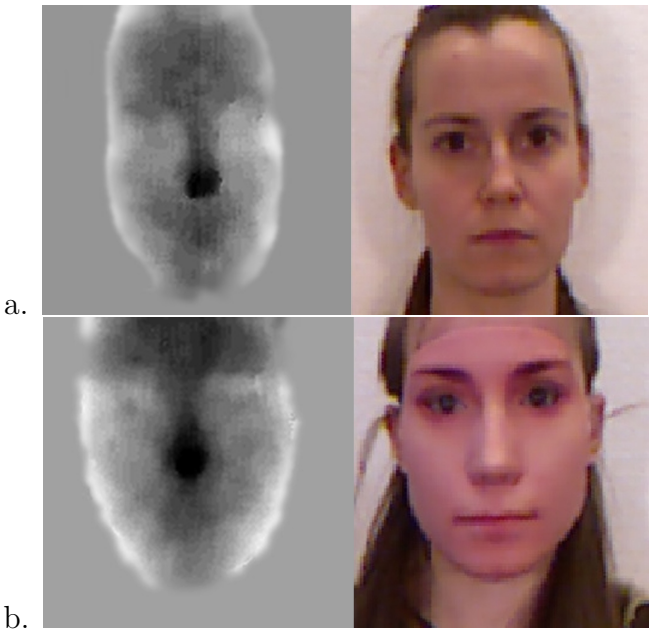


Figure 12: Examples from the 3D Mask Attack Data-set [1]. Depth image (left) and RGB image (right) of a) Real face. b) A person wearing a mask.

5 Scene illumination-based forgery detection

Another approach to detecting forgery in depth images is by examining physical rules of nature in the scene, e.g. consistency in scene illumination. Similar to 2D RGB images, detecting inconsistencies in light direction within a single image, is considered a sign of forgery [17, 18, 19].

As mentioned in Section 4, depth cameras do not perform well under outdoor lighting mainly due to the fact that natural light contains IR components that interfere with the IR lighting used by depth cameras.[65, 66]. Figure 13 shows this effect. A subjects face was captured with the sunlight source on left, using a Kinect V2 camera. The RGB image is shown on the left and the noise of the captured depth sequence (variance per pixel across all frames, as defined in Section 4) is shown on the right. It can be seen that noise values are larger on the right side of the face compared to the left side corresponding to higher concentration of light rays on the right side of the face.

The effect of illumination on noise in depth images can be further exploited to determine the source direction of the sunlight (relative to the object). We captured a depth sequences using a Kinect V2 camera, of a doll under sunlight every hour from 6am to 6pm. 100 depth frames were captured at each recording and the noise image was created. Figure 14 shows noise images captured every hour from 6am (top-left) to 6pm (bottom-right). Higher noise values can be seen shifting from the right side of the doll to the left side of the doll with change in sun position (specifically at the side, neck and nose of the doll). Figure 15 displays the noise ratio between the right side and the left side of the doll's face. For example, at 7am the noise ratio was 2.1, i.e. the noise on the right was much greater than on the left, indicating that sun light direction was from the right. At 1pm, the ratio was close to 1, i.e. equally distributed on both sides of face, and indeed sun light direction was from above. Additionally, the strong sun light combined with a reflective surface (the plastic material of the doll), may produce very strong reflections, and a large increase in noise, which causes saturation of the camera's IR sensors and ultimately results in burned out pixels. This can be seen in Figure 14 where missing pixels at the top of the doll's head can be seen as moving with the sun position. These pixels correspond to the top of the doll's head where sun rays were most intense.

This characteristic noise distribution can be exploited for forgery detection in depth images. Consider the example shown in Figure 16. Three objects are positioned on a table in the scene. The noise image of the depth sequence of this scene is shown in Figure 16b. For each object, we examined the ratio between its right and left sides. Ratios are 0.6,1.4,0.5 for the three objects respectively from left to right. Ratios indicating that two objects share the same lighting direction while the middle object has a different lighting direction. This inconsistency of lighting direction in the scene indicates some form of image manipulation. The inconsistency in lighting direction can actually be seen in the RGB image. However as we emphasized above, this paper aims at displaying the ability of detecting forgery from the depth signal alone. See Discussion.

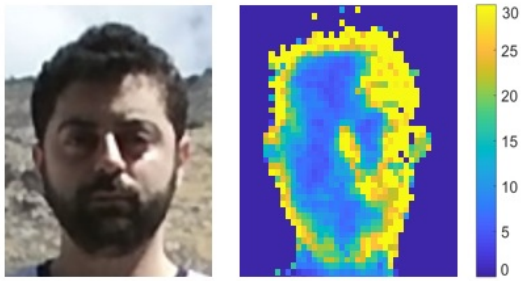


Figure 13: RGB image of a subject in sunlight (left), and corresponding depth noise image (right). Greater noise can be seen on the right side indicating the sun direction.

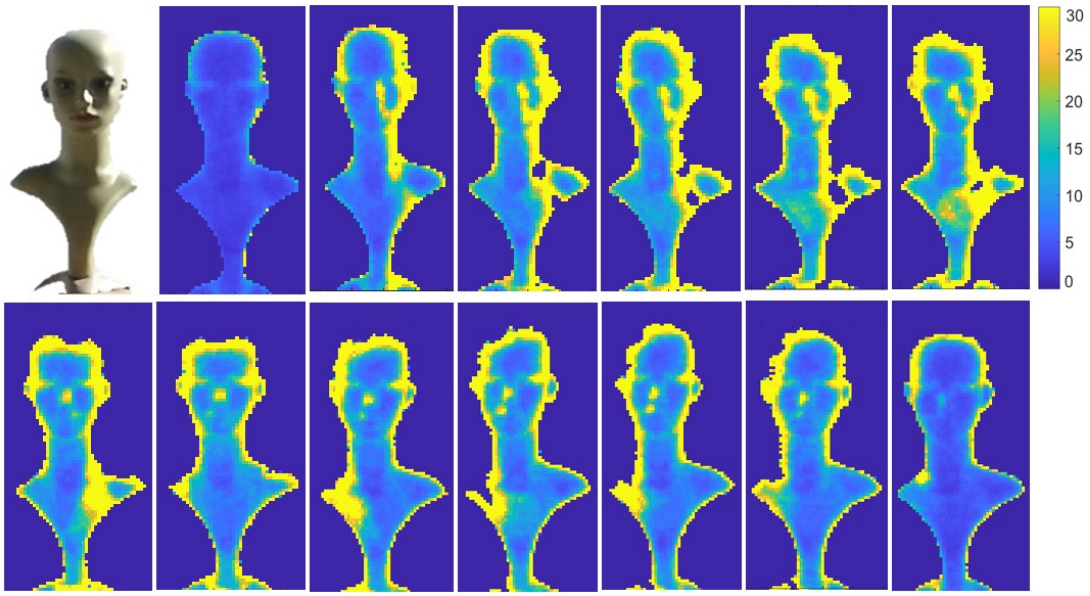


Figure 14: RGB image of doll (top left)and Noise images of the doll captured by a KinectV2 camera, under sunlight every hour from 6am (top left) to 6pm (bottom right).

6 Shadow inconsistencies-based forgery detection

In this section we show that inherent camera structure together with scene characteristics can be exploited to detect forgery in depth images and determine source camera. Depth 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 18a). This typically happens at edges of objects in the scene. With active 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. Depth cameras typically report shadows in the output depth images as pixels with 0 depth. We show that shadow size is a function of object distance from the camera (Figure 17a-b), background distance from the camera (Figure 17c-d) and camera configuration. We then show how inconsistencies in shadow size can be exploited for image forensic.

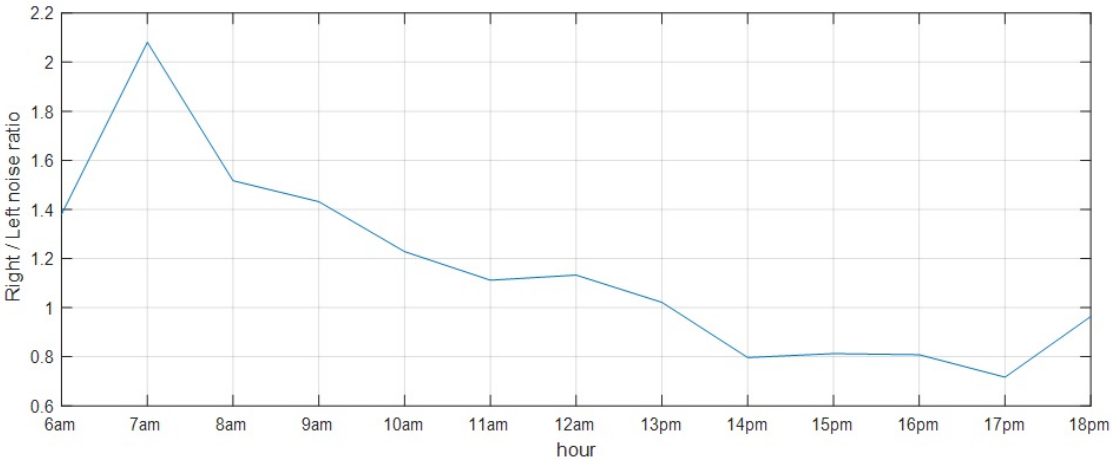


Figure 15: Ratio between right and left sides of a face according to hour of day.

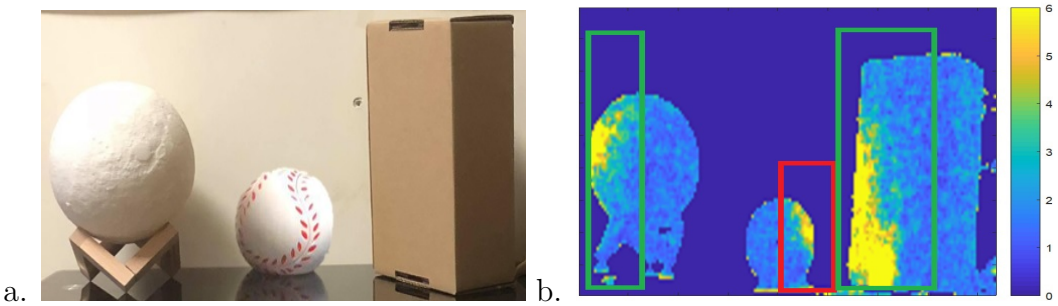


Figure 16: Example of illumination based forgery detection. a. RGB image of a scene with 3 objects. b) Noise image of the scene captured using a depth camera. Larger noise values can be seen on the right of the middle object (red rectangle) and on the left of the other 2 objects (green rectangles).

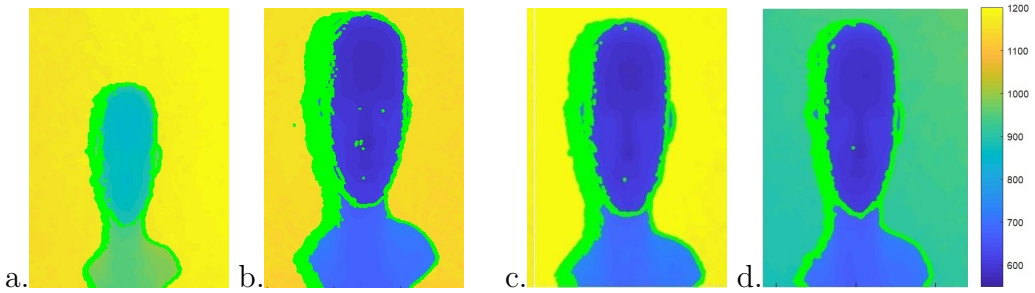


Figure 17: a-b) Shadow size depends on object distance from camera. Distance of background from camera is 1150mm in both images. Object distance from camera is a. 770mm. b. 450mm. As can be seen on the left side of the object, the smaller the distance from the camera the wider the shadow. c-d) Shadow size depends on distance of object from the background. Object's distance from camera is 600mm in both images. The distance of background from the camera is a. 1200mm. b. 880mm. As can be seen on the left side of the object, the greater the distance to the background, the wider the shadow. Shadow pixels (0 valued pixels) are marked in green. Images were captured using an Intel D415 depth camera.

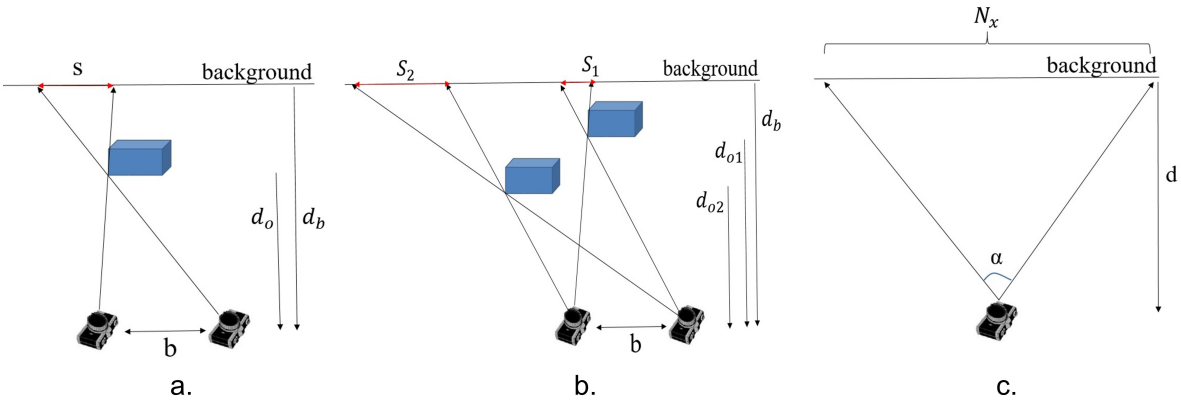


Figure 18: a) Shadow in depth image is the area seen by the left (e.g. IR sensor) while blocked from view by the right component (e.g. IR projector). b) Shadow size depends on object distance from camera. c) conversion from world units to image units.

6.1 Shadow size

Consider the scene and camera configuration shown in Figure 18a viewed as a projection onto the x-z plane of the camera coordinate system. The scene contains an object with a background object or plane behind it. A depth camera records the scene. Denote by b the camera baseline, i.e. the distance between the 2 camera components. 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:

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

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, it changes when the object changes position in depth. Figure 18b shows two depth positions d_{o1}, d_{o2}. Using the relation in equation (1) and equating for

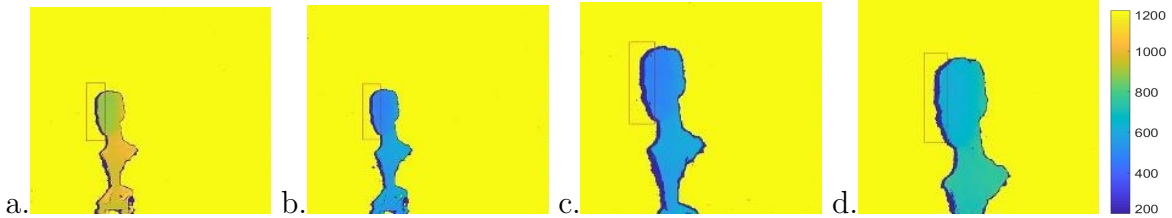


Figure 19: Forged depth images. a) Original. Object distance is 900mm, background distance is 1340mm (shadow calculated:11.675, measured:12) b) Depth values scaled to 500mm (shadow calculated:40.11, measured:12). c) Depth values scaled to 500mm + object re-sized (shadow calculated:40.11, measured:25). d) Depth values scaled to 500mm + object re-sized to correct for shadow width (shadow calculated:40.11, measured:40).

the baseline b, we obtain the relation between shadow widths $s1$ and $s2$:

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

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 projection sizes $w1$ and $w2$ of two objects of same physical size at distances d_{o1} and d_{o2} from the camera are related by:

$$w_2 = \frac{d_{o1}}{d_{o2}} \cdot w_1 \tag{3}$$

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 18c). the conversion between millimeters and pixels at depth d is given by:

$$N_x\,[pixel] = 2d \cdot tan(\frac{\alpha}{2})\,[millimeter] \tag{4}$$

To verify these calculations, we used an Intel D415 depth camera, and captured 30 natural office scenes with different objects and backgrounds distances (300 depth images). Shadow widths were measured in each image and compared with the expected value calculated from Equations (1) and (4) using the camera parameters ($b = 55mm$ and $\alpha = 70°$ under multiple resolutions). The average absolute difference between real and estimated width over all 300 examples was $\mu = 1.23$ pixels with variance of $\sigma^2 = 1.21$. Implying consistency of the computed shadow width values.

6.2 Shadow Inconsistency Based Forgery Detection

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.

6.2.1 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 opposite side of the object indicates that the image has been tampered with.

6.2.2 Shadow size based forgery

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 change, thus affecting the calculated shadow size. Figure 19 shows an example. The original image (Figure 19a) shows an object at distance 900mm with background at distance 1340mm. Shadow width is measured at 12 pixels where the calculated width, given the camera parameters is 11.7 pixels. Figure 19b shows a 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 suspected forgery.

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 [15]. However, even when corrected for object size, shadow width is still inconsistent. Figure 19c shows the object with depth values scaled to 500mm and the object itself together with it's shadow scaled to have the new depth, using Equation 3. Shadow size in this example 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.

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 19d, we still find inconsistency in object size with respect to the real world size. In the example, reverse-projecting the head in the image (from pixels to millimeters using equation 4) shows a head size of 230mm (while real human heads average approximately ∼140mm).

6.3 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.

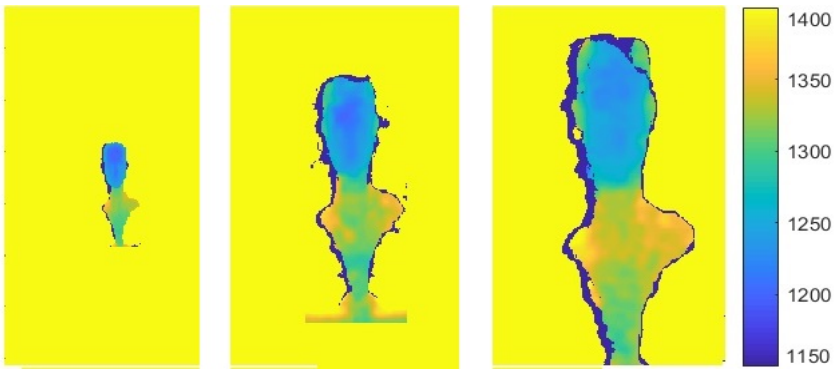


Figure 20: Same object captured by three different depth cameras at the same position. Left to right: KinectV1, Intel D435 and Intel D415.

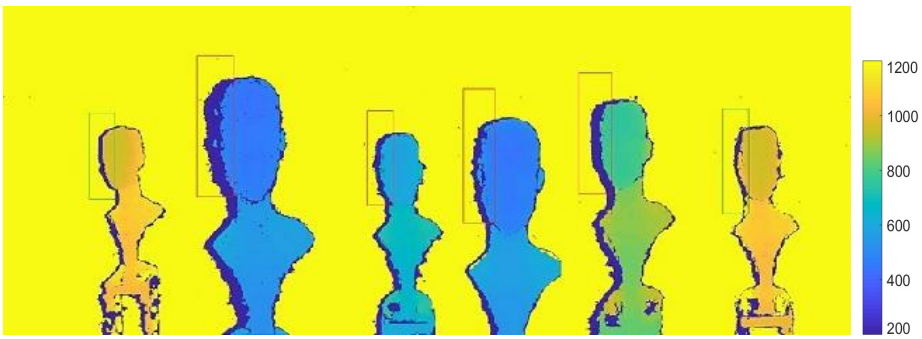


Figure 21: Which is real and which is forged?

Furthermore, the same camera with multiple resolution modes will generate different shadow sizes for the same scene, allowing the detection of the specific resolution. Figure 20 shows the same 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 shadow width as an indicator for source camera identification, we collected a set of 300 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 value closest to the measured shadow width indicated the source camera. Table 5 shows the results. Each row depicts the object's distance to camera d_o, the background distance to camera d_b, the measured shadow width (in pixels) and the calculated shadow width (in pixels). The shaded entries depicts the true source camera, as can be seen it is the closest to the measured width.

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

Table 5: 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.

7 Discussion and future work

In this research, we exploited characteristics of the depth camera, together with scene parameters and image statistics to detect forgery within depth images.
First, we showed that noise statistics in depth images can be exploited for camera source identification. Both in the lab and from data collected "in the wild" we showed that it is possible to determine the source camera very reliably. When an image was forged by copy-pasting patches from different image sources, we were able to refute image authenticity and determine that the image contained an invalid patch. However determining the correct camera source of the forged region was successful only in part. We further showed that beyond authenticity and camera source identification, noise statistics allowed us to determine whether the patch is forged and originated from a different location. Specifically, we were able to determine the correct 3D positioning (depth and x-position) of a patch from its noise statistics alone.
In addition, We showed that scene characteristics such as s in the depth image can be used to detect forgery since width depends on scene parameters such as the 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 through shadow size measurements. 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. We present and analyze a set of tools that can be used against forgery. Further studies can use machine learning tools to analyze noise statistics, while taking into account additional factors such as scene illumination, object color and material. However, the methods presented are not fool proof. Attackers who have knowledge about noise distribution and shadow formation can generate fake noise with the expected distribution or zero out the shadow areas so that shadow widths match that which is expected at the new forged depth.

References

[1] Nesli Erdogmus and Sébastien Marcel, "Spoofing in 2d face recognition with 3d masks and anti-spoofing with kinect," september 2013.

[2] Efthimia Bilissi and Michael Langford, *Langford's Advanced Photography*, Focal Press, Oxford, 2011.

[3] 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.

[4] 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.

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

[6] Asla M. Sa, Esdras Filho, Paul Cezar Carvalho, and Luiz Velho, "Coded structured light for 3d-photography: An overview," *RITA - Revista de Inform?tica Te?rica e Aplicada (Journal of Theoretical and Applied Informatics)*, 2002.

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

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

[9] 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.

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

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

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

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

[14] Wei Zhang, Xiaochun Cao, Yanling Qu, Yuexian Hou, Handong Zhao, and Chenyang Zhang, "Detecting and extracting the photo composites using planar homography and graph cut," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 544–555, 2010.

[15] 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.

[16] Micah K. Johnson and Hany Farid, "Detecting photographic composites of people," in *Digital Watermarking*, Y.Q. Shi, H.-J. Kim, and S. Katzenbeisser, Eds. 2008, Lecture Notes in Computer Science, pp. 19–33, Springer.

[17] Micah K. Johnson and Hany Farid, "Exposing digital forgeries in complex lighting environments," *IEEE Transactions on Information Forensics and Security*, vol. 2, pp. 450–461, 2007.

[18] 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.

[19] Wei Fan, Kai Wang, Francois Cayre, and Zhang Xiong, "3d lighting-based image forgery detection using shape-from-shading," in *Signal Processing Conference (EUSIPCO)*, Aug 2012, pp. 1777–1781.

[20] Wei Zhang, Xiaochun Cao, Jiawan Zhang, Jigui Zhu, and Ping Wang, "Detecting photographic composites using shadows," in *IEEE International Conference on Multimedia and Expo ICME*, June 2009, pp. 1042–1045.

[21] Qiguang Liu, Xiaochun Cao, Chao Deng, and Xiaojie Guo, "Identifying image composites through shadow matte consistency," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1111–1122, 2011.

[22] 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.

[23] Hany Farid and Mary J. Bravo, "Image forensic analyses that elude the human visual system," 2010.

[24] James F. O'Brien and Hany Farid, "Exposing photo manipulation with inconsistent reflections," *ACM Transactions on Graphics*, vol. 31, no. 1, pp. 4:1–11, 2012.

[25] Alin C. Popescu and Hany Farid, "Exposing digital forgeries by detecting duplicated image regions," Tech. Rep. TR2004-515, Department of Computer Science, Dartmouth College, 2004.

[26] Weiqi Luo, Jiwu Huang, and Guoping Qiu, "Robust detection of region-duplication forgery in digital image," in *International Conference on Pattern Recognition ICPR*, 2006, vol. 4, pp. 746–749.

[27] Jessica Fridrich, David Soukal, and Jan Lukas, "Detection of copy-move forgery in digital images," in *Digital Forensic Research Workshop*, Aug 2003.

[28] Vivek Kumar Singh and R.C. Tripathi, "Fast and efficient region duplication detection in digital images using sub-blocking method," *International Journal of Advanced Science and Technology*, vol. 35, pp. 1355–1370, 2011.

[29] Khayrul Bashar, Keiji Noda, Noboru Ohnishi, Hiroaki Kudo, Tetsuya Matsumoto, and Yoshinori Takeuchi, "Wavelet-based multiresolution features for detecting duplications in images," in *Conference on Machine Vision Application*, 2007, pp. 264–267.

[30] Vivek Kumar Singh and R.C. Tripathi, "Fast rotation invariant detection of region duplication attacks even on uniform background containing digital images," in *International Multi-Conference on Information Processing*, 2015.

[31] Seung-Jin Ryu, M. Kirchner, Min-Jeong Lee, and Heung-Kyu Lee, "Rotation invariant localization of duplicated image regions based on zernike moments," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 8, pp. 1355–1370, Aug 2013.

[32] Hailing Huang, Weiqiang Guo, and Yu Zhang, "Detection of copy-move forgery in digital images using SIFT algorithm," in *Workshop on Computational Intelligence and Industrial Application*, Dec 2008, vol. 2, pp. 272–276.

[33] Xu Bo, Wang Junwen, Liu Guangjie, and Dai Yuewei, "Image copy-move forgery detection based on surf," in *International Conference on Multimedia Information Networking and Security (MINES)*, Nov 2010, pp. 889–892.

[34] Sevinc Bayram, Ismail Avcibas, Bulent Sankur, and Nasir Memon, "Image manipulation detection," *Journal of Electronic Imaging*, vol. 15, no. 4, pp. 041102, 2006.

[35] Min-Jen Tsai, Cheng-Liang Lai, and Jung Liu, "Camera/mobile phone source identification for digital forensics," in *IEEE International Confernece on Acoustics, Speech, and Signal Processing*, 2007, vol. 2, pp. II–221–II–224.

[36] 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.

[37] 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.

[38] Kai San Choi, Edmund Y. Lam, and Kenneth K. Y. Wong, "Automatic source camera identification using the intrinsic lens radial distortion," *Optics Express*, vol. 14, no. 24, pp. 11551–11565, 2006.

[39] 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.

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

[41] Micah K. Johnson and Hany Farid, "Exposing digital forgeries through chromatic aberration," in *ACM Multimedia and Security Workshop*, 2006, pp. 48–55.

[42] Tran Van Lanh, Sabu Emmanuel, and Mohan S.Kankanhalli, "Identifying source cell phone using chromatic aberration," in *IEEE International Conference on Multimedia and Expo*, 2007.

[43] 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.

[44] Jan Lukas, Jessica Fridrich, and Miroslav Goljan, "Digital camera identification from sensor noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, 2006.

[45] Chang-Tsun Li, "Source camera identification using enhanced sensor pattern noise," in *IEEE International Conference on Image Processing*, 2009.

[46] Mo Chen, Jessica Fridrich, Miroslav Goljan, and Jan Lukas, "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 74–90, 2008.

[47] 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.

[48] Ahmet Emir Dirik, Husrev Taha Sencar, and Nasir Memon, "Digital single lens reflex camera identification from traces of sensor dust," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 539–552, 2008.

[49] 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.

[50] Zhouchen Lin, Rongrong Wang, Xiaoou Tang, and Heung-Yeung Shum, "Detecting doctored images using camera response normality and consistency," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.

[51] Hany Farid, "Blind inverse gamma correction," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1428–1433, 2001.

[52] Tian-Tsong Ng, Shih-Fu Chang, and Qibin Sun, "Blind detection of photomontage using higher order statistics," in *International Symposium on Circuits and Systems*, 2004, pp. 688–691.

[53] Chang-Hee Choia, Hae-Yeoun Leeb, and Heung-Kyu Lee, "Estimation of color modification in digital images by CFA pattern change," *Forensic Science International*, vol. 226, no. 1-3, pp. 94–105, 2013.

[54] 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.

[55] Andrew C. Gallagherc and Tsuhan Chen, "Image authentication by detecting traces of demosaicing," in *IEEE Computer Vision and Pattern Recognition Workshop*, June 2008, pp. 1–8.

[56] Hong Cao and Alex C. Kot, "Accurate detection of demosaicing regularity for digital image forensics," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 4, pp. 899–910, 2009.

[57] 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.

[58] Shuiming Ye, Qibin Sun, and Ee-Chien Chang, "Detecting digital image forgeries by measuring inconsistencies of blocking artifact," in *IEEE International Conference on Multimedia and Expo*, 2007, pp. 12–15.

[59] Archana V. Mire, S. B. Dhok, Neha J. Mistry, and P. D. Porey, "Resampling detection in digital images: A survey," *International Journal of Computer Applications*, vol. 84, no. 8, pp. 24–29, 2013.

[60] Jan Lukas and Jessica Fridrich, "Estimation of primary quantization matrix in double compressed jpeg images," in *Digital Forensic Research Workshop*, Aug. 2003.

[61] 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.

[62] Husrev Taha Sencar and Nasir Memon, *Digital Image Forensics - There is More to a Picture than Meets the Eye*, Springer-Verlag New York, 2012.

[63] Fabio Remondino and David Stoppa, *TOF Range-Imaging Cameras*, Springer-Verlag Berlin, 2013.

[64] Reinhard Koch Andreas Kolb Marcin Grzegorzek, Christian Theobalt, *Time-of-Flight and Depth Imaging. Sensors, Algorithms and Applications*, Springer-Verlag Berlin, 2013.

[65] Mohit Gupta, Qi Yin, and Shree K. Nayar, "Structured light in sunlight," in *IEEE International Conference on Computer Vision (ICCV)*, Dec 2013.

[66] Wajahat Kazmi, Sergi Foix, Guillem Alenya, and Hans J?rgen Andersen, "Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: Analysis and comparison," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 88, pp. 128 – 146, 2014.

[67] Tanwi Mallick, Partha Pratim Das, and Arun Kumar Majumdar, "Characterizations of noise in kinect depth images: A review," *IEEE Sensors Journal*, vol. 14, no. 6, pp. 1731–1740, 2014.

[68] Michael Riis Andersen, Thomas Jensen, Pavel Lisouski, Anders Krogh Mortensen, and Micheal Hansen, "Kinect depth sensor evaluation for computer vision applications," Tech. Rep. ECE-TR-6, Dept. Eng. Electr. Comput. Eng., Aarhus Univ. Aarhus, Denmark, Tech., 2012.

[69] Dragos Falie and Vasile Buzuloiu ., "Noise characteristics of 3d time-of-flight cameras," in *International Symposium on Signals, Circuits and Systems, ISSCS*, 2007, vol. 1, pp. 1–4.

[70] Kourosh Khoshelham and Er Oude Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *IEEE Sensors Journal*, vol. 12, no. 2, pp. 1437–1454, 2012.

[71] Chuong Nguyen, Shahram Izadi, and David Lovell, "Modeling kinect sensor noise for improved 3d reconstruction and tracking," in *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT)*, 2012, pp. 524–530.

[72] Jae-Han Park, Yong-Deuk Shin, Ji-Hun Bae, and Moon-Hong Baeg, "Spatial uncertainty model for visual features using a kinect sensor," *Sensors*, vol. 12, no. 7, pp. 8640–8662, 2012.

[73] Jan Smisek, Michal Jancosek, and Tomas Pajdla, "3d with kinect," in *Consumer Depth Cameras for Computer Vision - Research Topics and Applications*, Anthony T. S. Ho and Shujun Li, Eds. Springer-Verlag, New York, 2013.

[74] Benjamin Choo, Michael Landau, Michael DeVore, and Peter A. Beling, "Statistical analysis-based error models for the microsoft kinect depth sensor," *Sensors*, vol. 14, no. 9, pp. 17430–17450, 2014.

[75] Stephan Hussmann, Alexander Hermansk, and Torsten Edeler, "Real-time motion artifact suppression in TOF camera systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 5, pp. 1682–1690, 2011.

[76] Marvin Lindner and Andreas Kolb, "Compensation of motion artifacts for time-of-flight cameras," in *Dynamic 3D Imaging*, Andreas Kolb and Reinhard Koch, Eds., vol. 5742, pp. 16–27. Springer Berlin Heidelberg, 2009.

[77] Seungkyu Lee, "Time-of-flight depth camera motion blur detection and deblurring," *IEEE Signal Processing Letters*, vol. 21, no. 6, pp. 663–666, 2014.

[78] Elise Lachat, Helene Macher, Marie-Anne Mittet, Tania Landes, and Pierre Grussenmeyer, "First experiences with kinect v2 sensor for close range 3d modelling," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 93–100, 2015.

[79] Yu Yu, Yonghong Song, Yuanlin Zhang, and Shu Wen, "A shadow repair approach for kinect depth maps," in *Asian Conference on Computer Vision*, 2013, ACCV'12, pp. 615–626.

[80] Malcolm Reynolds, Jozef Dobos, Leto Peel, Tim Weyrich, and Gabriel J Brostow, "Capturing time-of-flight data with confidence," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011, pp. 945–952.

[81] James Mure-Dubois and Heinz Hugli, "Real-time scattering compensation for time-of-flight camera," in *Proceedings of the ICVS Workshop on Camera Calibration Methods for Computer*, 2007.

[82] "Microsoft, kinect for xbox one (kinect version 2)," https://www.xbox.com/en-US/xbox-one/accessories/kinect.

[83] "Microsoft, kinect (kinect version 1)," https://blogs.msdn.microsoft.com/kinectforwindows /2012/01/09/starting-february-1-2012-use-the-power-of-kinect-for-windows-to-change-the-world/.

[84] "ZED stereo camera by Stereolabs ," https://www.stereolabs.com/.

[85] Kevin Lai, Liefeng Bo, Xiaofeng Ren, and Dieter Fox, "A large-scale hierarchical multi-view RGB-D object dataset," 2011, pp. 1817–1824.

[86] Chih-Wei Hsu and Chih-Jen Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.

זיופים בתמונות עומק

עזמי חידר

תקציר

הרעיון של זיוף תמונה וגילויו, נחקר נרחב בתמונות ובווידיאו דו-ממדיים. עם העלייה בזמינות ושימוש במצלמות עם חיישני עומק (מצלמטת עומק), הפך להיות נחשב לשקול זיהוי זיוף גם בתמונות עומק.

במחקר זה אנו מציגים מחקר מבוא לגילוי זיוף בתמונות עומק. באופן ספציפי, אנו מראים כי ניתן לנצל סטטיסטיקות רעש בתמונות עומק לצורך זיהוי מקור מצלמה, וזיהוי זיוף בתמונות. אנו מראים עוד שאפשר להשתמש בתאורת סצנות לגילוי זיוף. לבסוף, אנו מראים כי ניתן לנצל את המאפיינים הגלומים במכניקת המצלמה כדי לקבוע זיוף תמונה מהצללים מבוססי חיישן.

זיופים בתמונות עומק

מאת: עזמי חידר
בהנחיית: ד"ר. חגית הל-אור

עבודת גמר מחקרית (תיזה) המוגשת כמילוי חלק מהדרישות לקבלת תואר "מוסמך האוניברסיטה"

אוניברסיטת חיפה
הפקולטה למדעי החברה
החוג למדעי המחשב

ינואר , 2020

זיופים בתמונות עומק

עזמי חידר

עבודת גמר מחקרית (תיזה) המוגשת כמילוי חלק מהדרישות לקבלת תואר ״מוסמך האוניברסיטה״

אוניברסיטת חיפה
הפקולטה למדעי החברה
החוג למדעי המחשב

ינואר , 2020