Forgery detection in 3D-Sensor Images

Noa Privman-Horesh, Azmi Haider, Hagit Hel-Or

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

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. Since 3D cameras (cameras with depth sensors) are becoming increasingly commonplace, it is of importance to introduce the notion of forgery detection in depth-images. In this paper, 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 and even depth reconstruction from noise.

1. Introduction

With the wide availability of image processing and manipulation software, tampering and abuse of images have become abundant. Unfortunately, this gives rise to serious consequences as images are often used as legal evidence, in forensics investigations, in medical records, as news items that reach millions of people, 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.

In recent years the use of low cost 3D cameras (cameras with depth sensors) has increased tremendously. They are already in use in medical applications, cinematography, art production and most importantly security systems and biometric based recognition systems. The output of these cameras, which we term *depth-images*[1] have become commonplace and are expected to be even more so now that 3D cameras are integrated into cellular phones. The increased usage of depth images, raises serious concerns of authenticity, and reliability of the data, especially in the context of biometric screening and identification and reliability of medical data and diagnosis. The notion of forgery detection in depth-images is novel and has yet to be seriously researched and developed, however, its necessity in the near future for security as well as for judicial issues, is clear.

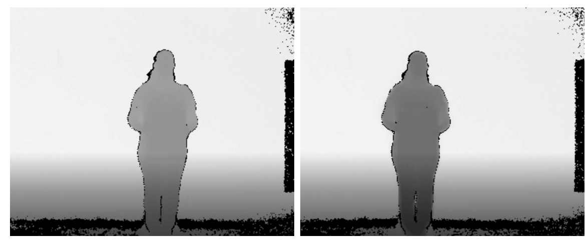


Figure 1. Forgery detection in depth-images - Which one is fake?

The main goal of the paper is to introduce the idea of forgery detection in 3D sensors (Figure 1) and present an initial set of results and insights in this area. To the best of our knowledge this is the first paper to deal with this issue.

This research was supported by grant no 1455/16 from the Israeli Science Foundation and by a grant from the Center for Cyber Law and Policy (CCLP), at the University of Haifa.

[1] We distinguish between depth-images, often referred to as 2.5D images, and 3D data often referred to as 3D cloud of points or 3D model of a scene. We do not deal with the latter in this paper.

2. Background

2.1. 2D image forgery and its detection

Image forgery detection and image authentication typically aim to find irregular characteristics of an image or an unexpected footprint of the camera or acquisition device in order to provide a reliable measure of suspected forgery. In many cases the forgery detection algorithms are even able to point out the suspected forged region within the image and the forgery type. Image forgery in 2D images has been widely studied. The various methods of 2D image forgery detection, are 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.

- **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.

A distinction is made between passive forgery detection and detection based on actively embedding digital signatures in the image in the form of Digital Watermarks [6]. In this study we focus on passive forgery detection that rely only on the depth-image content iteself.

The passive approaches to image forgery detection fall into 3 categories based on the assumptions they rely on:

1. **Physical rules and models of the scene** - Examples include inconsistencies of object size, lighting directions, shadow inconsistencies, and reflection inconsistencies.

2. **Statistics of the source images** - Either from raw pixel data, DCT and Wavelet transform coefficients, moments, or local features such as SIFT and SURF.

3. **Inherent characteristics of the camera** - Such as lens based detections, detection based on image sensors and sensor noise and detection based on the camera's image processing pipeline.

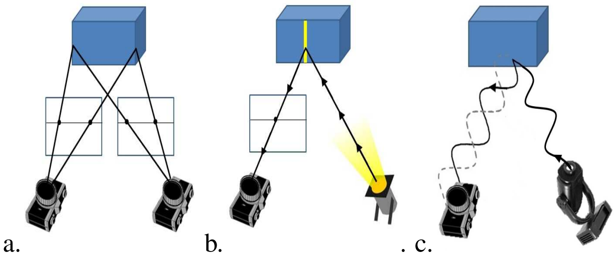
For a general review of 2D image forgery detection techniques see [16, 26, 9].

2.2. Depth imaging

In this research we discuss image forgery detection methods that are specifically targeted towards depth-images. 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.

2.3. Depth sensing by 3D cameras

The outputs of 3D cameras are typically videos or image sequences where each frame is represented as a depth image (often referred to as 2.5D images) with pixel values indicating distance from the camera (see Figure 1). Similar to 2D cameras, 3D camera components include optics, sensors and an imaging pipeline, however, these are tuned to obtain depth data. Also included are additional components, unique to depth sensing (such as IR projectors and phase detectors). 3D cameras differ in the technology used to acquire the depth-images (Figure 2):



a. b. . c.

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

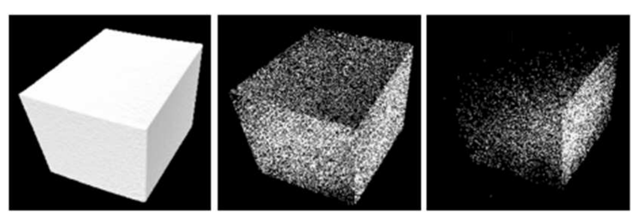
- Stereo imaging [29, 13] - a passive imaging system comprised of two or more 2D cameras positioned along a common baseline that simultaneously capture two views of the scene. Following correspondence of points between the two views, depth (distance from camera baseline) can be computed.

- Structured light (Projected-light sensors) [11, 28] - an IR pattern is projected onto the scene forming 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.

- Time of flight (ToF) [10, 15] - 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.

2.4. Noise in depth-images

A central aspect of our study of depth-image forgery detection relies on the noise inherent in the 3D 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 and more), the analysis algorithms used in the pipeline (correspondence methods, error correction, phase analysis and more), as well as scene characteristics such as position and depth of objects in the scene and scene lighting. We



a. b. c.

Figure 3. Illumination strength increases noise. a) Low illumination b) Medium illumination c) High illumination. (From [14]).

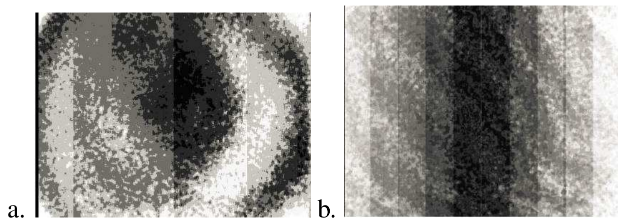


Figure 4. a) Axial noise shows radial pattern across flat surface, increases quadratically with depth. b) Temporal noise in flat surface, noise increases laterally and forms stripes. (From [24])

consider noise as arising from 4 sources:

Noise Inherent to the camera parameters -
Accuracy of the 3D data is dependent on physical parameters of the cameras and thus affect the signal to noise of the system's output [11, 27]. Focal length, field of view, quality of lenses, all contribute to these affects. Active depth measurements are dependent on the quality of the IRED and its projected IR light, including its intensity and collimation. Camera specific parameters include baseline between cameras for stereo based systems and camera to projector distance in structured light based systems. ToF based systems are dependent on the modulation quality of the IR signal.

Noise Inherent to depth measuring methods -
Stereo and structured light rely on correspondences at which points relatively accurate depth measurement is obtained. Between these points, interpolation is used which inherently introduces depth errors [11, 28]. Tof approaches obtain depth measurements at every pixel location thus avoiding interpolation errors, however they are inherently prone to phase ambiguity and demodulation errors [15] which result in erroneous depth estimates.

Noise due to scene characteristics -
Both scene illumination and object positioning within the scene affect accuracy of depth estimation. 3D cameras do not perform well under strong ambient illumination (Figure 3), specifically outdoor lighting. This is mainly due to the fact that natural light contains IR components that interfere with the camera's IR source. Furthermore the camera's IR is typically of very low intensity and is overpowered by the high intensity outdoor lighting [14, 18]. Inaccuracy of estimated depth, often termed Axial Noise, has been shown to increase quadratically with distance of objects in the scene from the camera [24, 8, 19, 25, 5]. 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 Temporal Noise increases linearly (in the x and y directions) and is very extreme at the edge of the camera's field of view possibly due to lens distortion [24, 25, 5]. Furthermore several studies have shown a radial ripple like noise that extends laterally (Figure 4a) [24, 5]. Shadow and lateral noise increase at strong depth edges (Figure 5) [20, 24, 30] possibly due to difficulty in triangulation at these locations or due to erroneous reflected light. Lateral noise appears in vertical stripe patterns (Figure 4b) [24, 5]. When object (or camera) motion is involved, motion blur in 3D cameras results in depth over or underestimation near depth edges [23, 22].

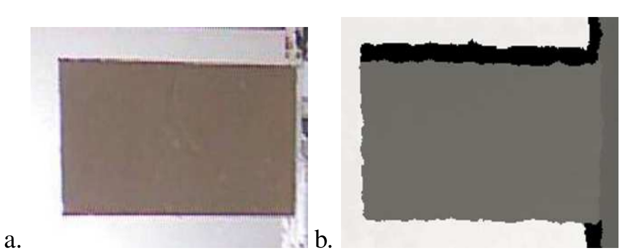


Figure 5. Shadow Noise. a) RGB image b) depth-image, shadow can be seen on the top. (From [24])

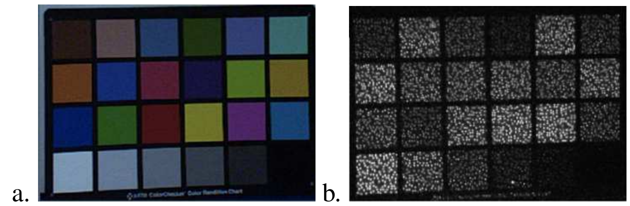


Figure 6. Noise is dependent on object color. a) RGB image b) depth-image. (From [15]).

Noise due to object characteristics - Studies show that color and brightness of objects affect depth estimation (Figure 6) [8, 15] however others maintain that it does not [24]. Since materials differ in their IR absorption and thus affect depth estimation [24], it is possible that the color of objects is confounded with its material giving rise to the confusion.

3. Image forgery detection in depth-images

In this paper we aim to show that detection of forgery is possible and viable in depth-images.

We consider the 3 classes of forgery detection: image authentication, image forgery detection and source camera identification (as described in Section 2.1) in the context of depth-images. There are various approaches that can be used for 3D forgery detection and of course one may fall back to forgery detection in the 2D camera typically incorporated in the 3D camera system. In this study we chose to focus on the statistics of depth-images as acquired by 3D cameras. We restrict our analysis to consumer, low-cost and readily available cameras, as these are more susceptible to forgery attacks.

3.1. Noise based forgery detection

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. In this study, we too consider image noise, albeit in depth-images. In contrast with 2D image forgery detection, where spatial noise is exploited, we use temporal noise and

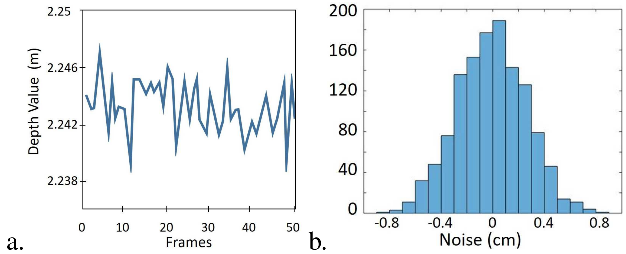


Figure 7. a) Depth response at a pixel acquired by KinectV2. b) Histogram of noise at the pixel.

consider both lateral and axial noise (Section 2.4) in our noise modeling.

Figure 7a shows the depth response at a target pixel acquired by a KinectV2 [1] camera over 50 frames. We define noise as the deviation from the mean depth response within 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 raw depth values and *noise variance* is taken as the variance of the absolute of deviation values. For the example in Figure 7, noise magnitude is $\underline{\mu} = 0.15$, and noise variance is $\sigma^2 = 0.035$.

3.1.1 Data Collection

For our 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 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 [2] (structured light), KinectV2 [1] (time of flight), ZED [3] (stereo). To exploit noise for forgery detection, we collected noise statistics at each target position, including: noise distribution (histogram), noise mean and variance and higher order statistics of skew, and kurtosis. These measures were normalized and concatenated to form the sample's feature vector.

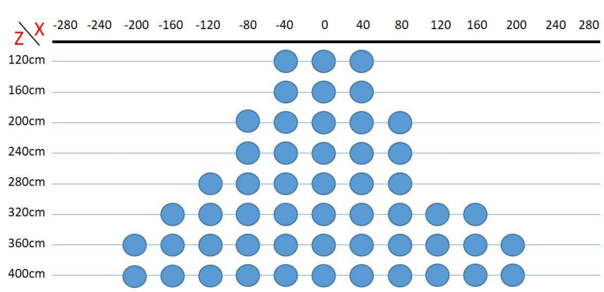


Figure 8. Map of target positions for data collection by KinectV2 (52 target positions).

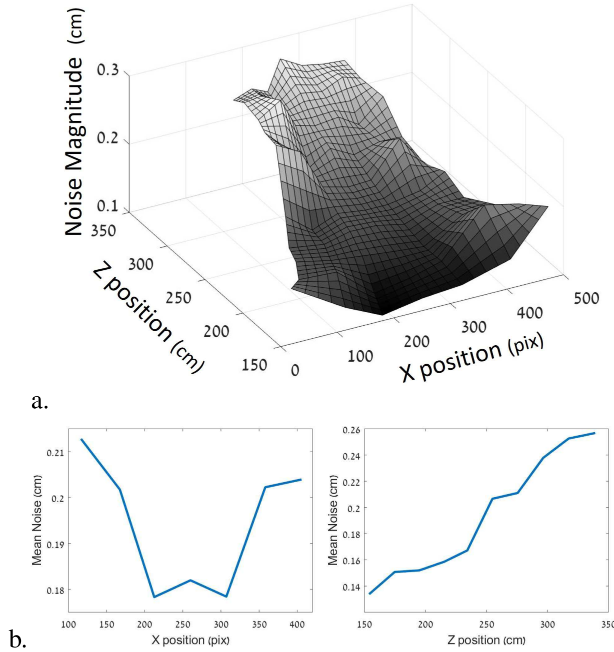


Figure 9. a) Noise magnitude as a function of x position and depth (z-pos). Noise increases with depth and with horizontal deviation from center. b) Mean Noise as a function of x position (left) and of depth (z-pos)(right).

Similar to [24, 5], we found that noise magnitude increased with depth and with increasing deviation from the center along the horizontal position. Figure 9 shows the noise magnitude as a function of depth (z pos) and as a function of horizontal position (x-pos) relative to center. We exploit these characteristics for forgery detection.

3.1.2 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 new inputs not used in the training data (new patches from the same cameras and data from different cameras). Nearest Neighbor was used to determine the source camera. Success rates are shown in Table 1. It can be seen that Kinect-V2 are

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

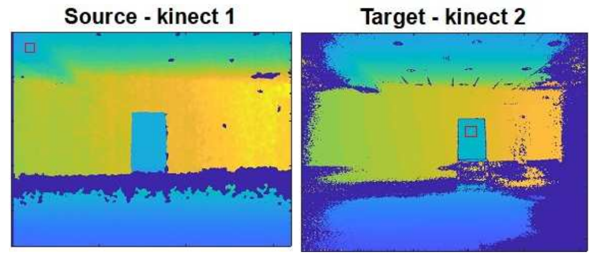


Figure 10. Forged image used for source target detection. A patch was copied from a KinectV1 sequence into a KinnectV2 sequence, (marked as square). Based on noise statistics, the forged area was detected and correct source (KinectV1) was successfully deduced.

very reliable in identifying source camera whereas it was found that KinectV1 data is often confused as arising from the ZED camera.

The high success rate is not surprising, given that the 3 types of cameras use different technologies each producing unique noise signatures.

We also tested for source camera identification in the wild. A collection of 6 depth-image sequences were collected; 3 from home units and 3 randomly selected from a public database [21]. Sequences were cropped to 300 frames. The sequences were analysed by randomly selecting 300 flat patches of size 20x20 and testing each for camera source. The resulting camera source was determined as the majority voting of the image patches. Table 2 shows the 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 of size 20x20x300 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 overlapping image patches and determining source camera. If the source was found to differ from the target camera it was marked as forged, and an attempt was made to detect the correct camera source. Results showed that for all forged images, all patches that overlapped some portion of the forged region were detected. However, in only 60% of these cases, the correct source camera of the forged patch was determined. Most misses occurred when the overlap of the patch with the forged region was less than 50%.

Camera	Source	Correctly Classified
KinectV1	[21]	92%
KinectV1	[21]	65 %
KinectV1	[21]	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

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. Thus, to be able to detect specific cameras, we consider pixel points having ”extreme statistics” namely, defective pixels, as described in the following section.

3.1.3 Defective Pixels

The sensor manufacturing process is such that pixels of a CCD or CMOS sensor may be defected (termed hot pixels, dead pixels or burnt pixels). In an early study [12], defects of CCD pixels in 2D cameras were examined and used for source camera identification. Similarly in [7], pixel defects due to dust were exploited for camera source identification.

3D cameras show similar pixel defects which, here too can be used to determine source camera identification as well as copy-paste forgery. We considered 6 KinectV2 cameras and tested for their defective pixels which were expressed as undefined valued pixels. Each camera showed a unique and distinct pattern (Figure 11). To validate, a collection of 1000 frames of depth-images were randomly chosen from sequences acquired by the 6 cameras. Source camera was correctly detected in 100% of the trials. However, since correcting for these defective pixels is easily performed, this does not form a very reliable forgery detection method.

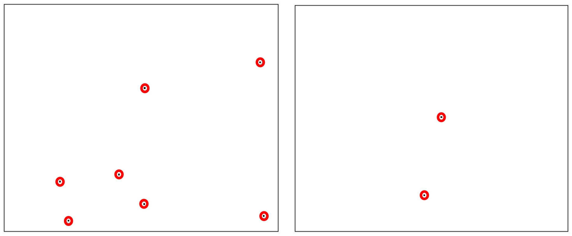


Figure 11. Every Kinect cameras unit shows a unique pattern of defective pixels that produce constant zero-valued pixels.

3.1.4 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 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 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 [17] was built based on the noise data (feature vector per sample) collected from a KinectV2 camera (as described in Section 3.1.1) with depth (z-values) sampled at 30cm intervals between 140cm and 350cm distance from camera and x-positions sampled at 30cm intervals symmetrically about the the scene center position. The classes represented all the sampled xz-positions (81 classes). Forgery testing was performed on 2025 new 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 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. Rows 2 and beyond reflect success rates when the N-closest matches are considered and the closest of the set is taken as the resulting x-z positions.

	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.

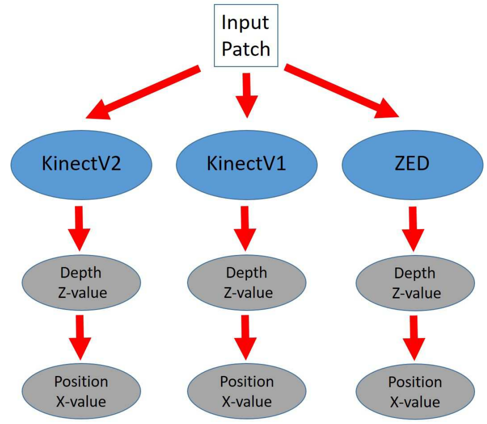


Figure 12. A cascade of clusters used for camera source identification, depth value prediction and x-position prediction.

Results show that when considering the 2 closest matches, the correct x-z of a patch is detected at over 92% success rate and a low average error of 1.5cm and 6.6 cm for x and z distances respectively. The lower success rate of 73% for the closest match indicates that there is some confusion between neighboring x-z positions (for the sampling positions used in our experiment).

We further tested for combined camera source identification (Section 3.1.2) and xz-position detection. As in Section 3.1.2, 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 12, containing three data clusters, one for each camera type, subdivided into 6 subclusters associated with the possible z-values and further subdivided into 9 clusters associated with x-positions. Testing was performed on inputs collected from cameras not used in the training data. Testing was performed in a cascade: first, source camera was identified, then for the successful detections, z-position was determined and for the successful cases x-position was determined. Table 4 shows the results.

	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.

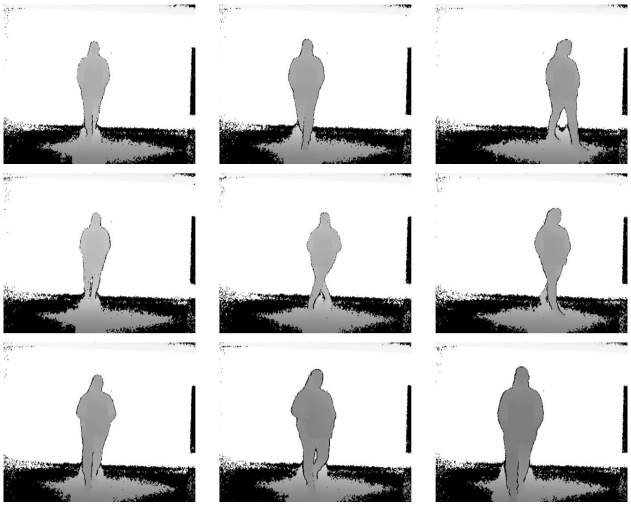


Figure 13. Frames from depth-image sequences of human motion along a path.

3.1.5 Movement Path from Noise

As a further demonstration of the strength of noise analysis for image authentication, we show that within the above capabilities of detection of x-z position from noise we are able to reconstruct the motion path of a moving object (human) using only the noise of the sequence.

Let $P_i = [x_i, z_i], \quad i = 1 \ldots k$ denote the sequence of positions along the path P. For each position i, we can determine $\hat{P}_i = (\hat{x}_i, \hat{z}_i)$, the most likely x,z position, from its noise statistics by classification to one of the trained xz clusters. However, classification is error prone (see Section 3.1.4). Thus to determine the path we incorporate the reliability of the classification given as the negative loss function [4] as well as a smoothness of path term. The data term is defined as:

$$D = \sum_{i}^{k} dist(P_i, \hat{P}_i) * NegLoss_i$$

Where $dist$ calculates the distance between 2 positions xz (correcting for difference in units) and $Negloss_i$ is the negative loss in classifying P_i to $\hat{P}_i$. The smoothness term is given by:

$$S = \sum_{i}^{k} [dist(P_i, P_{i+1}) - dist(P_i, P_{i-1})]^2$$

The reconstructed path is then given by:

$$P_{final} = argmin_P\{D + \lambda * S\}$$

where λ is a weighting scalar (good results were achieved with $\lambda = 10^3$) .

We collected several sequences of a human moving in the xz space as captured by a KinectV2 camera (Figure 13).

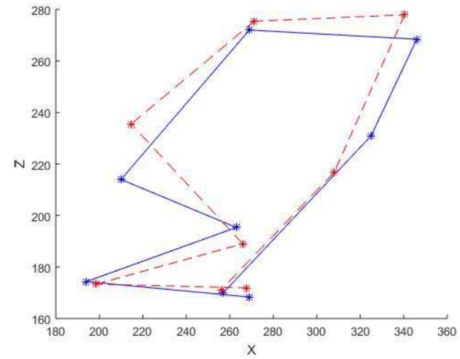


Figure 14. Reconstruction of human motion path from noise statistics. The original path is in solid blue, the reconstructed path dashed in red.

For several key positions along the path, 25-50 frames were extracted and patches within the figure were manually cropped from which the noise statistics was collected. The ground truth was determined as the average depth of the collected data. The above procedure was then applied to reconstruct the path. Although the subject was in motion, the small number of frames reduced motion blur effects enough so that together with the smoothness constraint allowed the use of the original noise model collected in the lab to be used with good path reconstruction results. Figure 14 shows an example of a motion path (solid blue) and the reconstructed path (dashed red). The reconstruction error, measured as the average distance between positions of the path P and positions of the reconstructed path P_{final}:

$$\sum_{i}^{k} dist(P_{final}, P)$$

was found to be 10.66 cm. Original path length is 497.66 cm and reconstructed path length 484.62 cm. Figure 15 shows results for a collection of shorter paths performed by

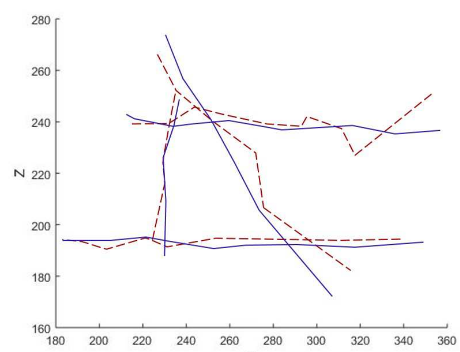


Figure 15. Reconstruction of a collection of human motion paths from noise statistics. The original path is in solid blue, the reconstructed path dashed in red.

the moving human. The average reconstruction error in this case was found to be 14 cm.

4. Discussion

In this paper, we presented an introductory study of forgery detection in depth-images acquired by 3D cameras. In this study we focused on noise statistics in these images and exploited their characteristics for forgery detection. Specifically, 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. Success rates were high when considering the second best position estimation. For single best position success rate was lower implying some confusion between neighboring x-z positions (for the sampling positions of 30 cm used in our experiment). However, this still enables the reconstruction of the motion path of a subject in the x-z space from the noise statistics alone. This implies that copy-paste forgery of people in motion can be detected in depth-images. We note that as in many 2D forgery detection cases, here too, copy-pasting a patch to an x-z location equivalent to the patch source location, will not be detected using the approaches suggested here.[2]

Although, introductory, this paper aims to show that detecting forgery in depth-images is indeed feasible. Further studies will use learning tools to analyze noise statistics, while taking into account additional factors such as scene illumination, object color and material. In addition, we plan to expand the number and types of cameras we are testing as well considering additional bases for forgery detection.

Further studies of image forensic in this media is quickly becoming a necessity, considering 3D sensors are becoming increasingly commonplace, and issues of authenticity, copyright, and forgery prevention will become inevitable.

[2]Referring back to Figure 1: the fake image is on the right - the depth values for the human figure were manually decreased.

References

[1] Microsoft, Kinect for Xbox One (Kinect Version 2). https://www.xbox.com/en-US/xbox-one/accessories/kinect.

[2] 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/.

[3] ZED stereo camera by Stereolabs . https://www.stereolabs.com/.

[4] E. L. Allwein, R. E. Schapire, and Y. Singer. Reducing multiclass to binary: A unifying approach for margin classifiers. *Journal of Machine Learning Research*, 1(Dec):113–141, 2000.

[5] B. Choo, M. Landau, M. DeVore, and P. A. Beling. Statistical analysis-based error models for the microsoft kinect depth sensor. *Sensors*, 14(9):17430–17450, 2014.

[6] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker. *Digital Watermarking and Steganography*. Morgan Kaufmann Publishers Inc., San Francisco, 2 edition, 2008.

[7] A. E. Dirik, H. T. Sencar, and N. Memon. Digital single lens reflex camera identification from traces of sensor dust. *IEEE Transactions on Information Forensics and Security*, 3(3):539–552, 2008.

[8] D. Falie and V. B. . Noise characteristics of 3d time-of-flight cameras. In *International Symposium on Signals, Circuits and Systems, ISSCS*, volume 1, pages 1–4, 2007.

[9] H. Farid. A survey of image forgery detection. *IEEE Signal Processing Magazine*, 26(2):16–25, 2009.

[10] S. Foix, G. Aleny, and C. Torras. Lock-in time-of-flight (ToF') cameras: A survey. *IEEE Sensors Journal*, 11(9):1917–1926, 2011.

[11] J. Geng. Structured-light 3d surface imaging: a tutorial. *Advances in Optics and Photonics*, 3(2):128–160, 2011.

[12] Z. J. Geradts, J. Bijhold, M. Kieft, K. Kuroki, and N. Saitoh. Methods for identification of images acquired with digital cameras. In *Proc. of SPIE: Enabling Technologies for Law Enforcement and Security*, volume 4232, pages 505–512, 2001.

[13] L. L. Grewe and A. C. Kak. Stereo vision. In T. Y. Young, editor, *Handbook of Pattern recognition and Image Processing: Computer Vision*, pages 239–317. Academic Press, 1994.

[14] M. Gupta, Q. Yin, and S. K. Nayar. Structured light in sunlight. In *IEEE International Conference on Computer Vision (ICCV)*, Dec 2013.

[15] M. Hansard, S. Lee, O. Choi, and R. Horaud. *Time of Flight Cameras: Principles, Methods, and Applications*. Springer-Verlag London, 2012.

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

[17] C.-W. Hsu and C.-J. Lin. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Networks*, 13(2):415–425, 2002.

[18] W. Kazmi, S. Foix, G. Alenya, and H. J. 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*, 88:128 – 146, 2014.

[19] K. Khoshelham and E. O. Elberink. Accuracy and resolution of kinect depth data for indoor mapping applications. *IEEE Sensors Journal*, 12(2):14371454, 2012.

[20] E. Lachat, H. Macher, M.-A. Mittet, T. Landes, and P. Grussenmeyer. First experiences with kinect v2 sensor for close range 3d modelling. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 93–100, 2015.

[21] K. Lai, L. Bo, X. Ren, and D. Fox. A large-scale hierarchical multi-view RGB-D object dataset. pages 1817–1824, 2011.

[22] S. Lee. Time-of-flight depth camera motion blur detection and deblurring. *IEEE Signal Processing Letters*, 21(6):663–666, 2014.

[23] M. Lindner and A. Kolb. Compensation of motion artifacts for time-of-flight cameras. In A. Kolb and R. Koch, editors, *Dynamic 3D Imaging*, volume 5742, pages 16–27. Springer Berlin Heidelberg, 2009.

[24] T. Mallick, P. P. Das, and A. K. Majumdar. Characterizations of noise in kinect depth images: A review. *IEEE Sensors Journal*, 14(6):1731–1740, 2014.

[25] J.-H. Park, Y.-D. Shin, J.-H. Bae, and M.-H. Baeg. Spatial uncertainty model for visual features using a kinect sensor. *Sensors*, 12(7):8640–8662, 2012.

[26] A. Piva. An overview on image forensics. *ISRN Signal Processing*, 2013(496701), 2013.

[27] F. Remondino and D. Stoppa. *TOF Range-Imaging Cameras*. Springer-Verlag Berlin, 2013.

[28] A. M. Sa, E. Filho, P. C. Carvalho, and L. 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.

[29] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47:7–42, 2002.

[30] Y. Yu, Y. Song, Y. Zhang, and S. Wen. A shadow repair approach for kinect depth maps. In *Asian Conference on Computer Vision*, ACCV'12, pages 615–626, 2013.