



HAL
open science

Using Internal Depth to Aid Stereoscopic Image Splicing Detection

Mark-Anthony Fouche, Martin Olivier

► **To cite this version:**

Mark-Anthony Fouche, Martin Olivier. Using Internal Depth to Aid Stereoscopic Image Splicing Detection. 8th International Conference on Digital Forensics (DF), Jan 2012, Pretoria, South Africa. pp.319-333, 10.1007/978-3-642-33962-2_22 . hal-01523699

HAL Id: hal-01523699

<https://inria.hal.science/hal-01523699v1>

Submitted on 16 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Chapter 22

USING INTERNAL DEPTH TO AID STEREOSCOPIC IMAGE SPLICING DETECTION

Mark-Anthony Fouche and Martin Olivier

Abstract Splicing is a common image manipulation technique, where parts of multiple images are combined to create a new composite image. Commercial image editing software enables almost anyone to splice images and create fake photographs. This paper investigates how the relationship between object distance and internal depth can aid in detecting spliced stereoscopic images. An equation is derived for predicting the distance at which an object loses internal depth. Experiments with stereoscopic images indicate that the analysis of this depth information can assist in detecting image splicing.

Keywords: Image forensics, stereoscopic images, splicing detection

1. Introduction

Stereoscopy is a technique that is used to create an illusion of depth on a flat surface. This is achieved by displaying two perspectives of the same scene separately to the left and right eyes [22]. The perspectives can be taken by a single or multiple image capturing devices. Together, the two perspectives form a stereoscopic (3D) image. Stereoscopic images have applications in a variety of fields, including astronomy [14], education [23], environmental forensics [8], medical sciences [21], robotics [12] and entertainment [18].

Capturing digital stereoscopic images has become very popular due to recent advances in digital image capturing devices such as digital cameras [9], digital video cameras [17], handheld game consoles [1] and smartphones [13]. The widespread use of stereoscopic images makes it important to be able to detect image tampering.

Considerable research has been conducted in the area of digital image forensics [5, 7, 16]. However, little research has focused on how depth information from stereoscopic images can be used to detect splicing. In a previous paper [6], we showed that disparity maps can be used to aid stereoscopic splicing detection, but the results were limited to non-stereoscopic to stereoscopic splicing. To our knowledge, no published research exists on using depth information for stereoscopic image splicing detection.

This paper investigates how the relationship between an object's distance and internal depth can aid in detecting spliced stereoscopic images. The internal depth of an object is computed as the distance between the nearest and farthest visible points on the object with regard to the camera. In digital stereoscopic images, a point can be represented by a single pixel. The distance of a point is computed as the difference in the horizontal pixel positions of the point in the left and right images of a stereoscopic image pair. At a certain distance from the camera any two points of an object would appear to have the same distance because the difference in pixel positions is less than one pixel. In this case, the object is considered to have no internal depth. A stereoscopic image containing an object with internal depth at a position where it should not have internal depth is an anomaly that may indicate splicing. Likewise, an image containing an object with no internal depth at a position where it should have internal depth is also anomalous and may indicate splicing.

This paper derives an equation for computing the distance at which an object loses its internal depth. The equation is tested on a set of stereoscopic images containing chairs with target signs in the front and back to express internal depth. The chairs were placed at different distances and photographed. The internal depths of the chairs were measured and the results compared with the expected values using the derived equation. A chair was then spliced from a distance where it had internal depth to an area where it should not have internal depth and vice versa. The internal depth of each spliced chair was compared with the expected internal depth. The experimental results indicate that the distance at which an object loses internal depth can aid in the detection of stereoscopic image splicing.

2. Background

This section discusses stereoscopic images and image splicing techniques.



Figure 1. Stereoscopic image pair.

2.1 Stereoscopic Images

Stereoscopic (3D) images are image pairs of the same scene taken from slightly different positions. Objects at different distances from a camera have different horizontal placements in the left and right images of a stereoscopic pair. The closer the object is to the camera, the greater the difference in the horizontal placements. This enables depth to be perceived when one image is projected on one eye and the other image on the other eye. Figure 1 shows an image pair taken with a digital stereoscopic camera.

Digital stereoscopic images can be split into separate image files or stored in a single image file. An example of the latter is multi-picture object (MPO) files [2].

Stereoscopic images can be viewed in various ways, some of them require a user to wear special glasses. One technique converts an image to an anaglyph, which requires glasses with different colored lenses [4]. Another technique uses a special screen with a double refresh rate. The left and right images are alternated on the screen and viewed with shutter glasses. The shutter glasses block the vision of the eye for which the current image should not be displayed [19]. A third technique superimposes the two images, each of which is projected with a different polarity. Polarized glasses are then used to filter each image for the correct eye [10].

Stereoscopic images can also be viewed without glasses. One technique involves autostereoscopic displays [3, 11] that are present in many modern stereoscopic image capturing devices [1, 9, 13, 17].

2.2 Image Splicing Techniques

Splicing is a common image manipulation technique where multiple images are combined to create a new composite image. Stereoscopic im-

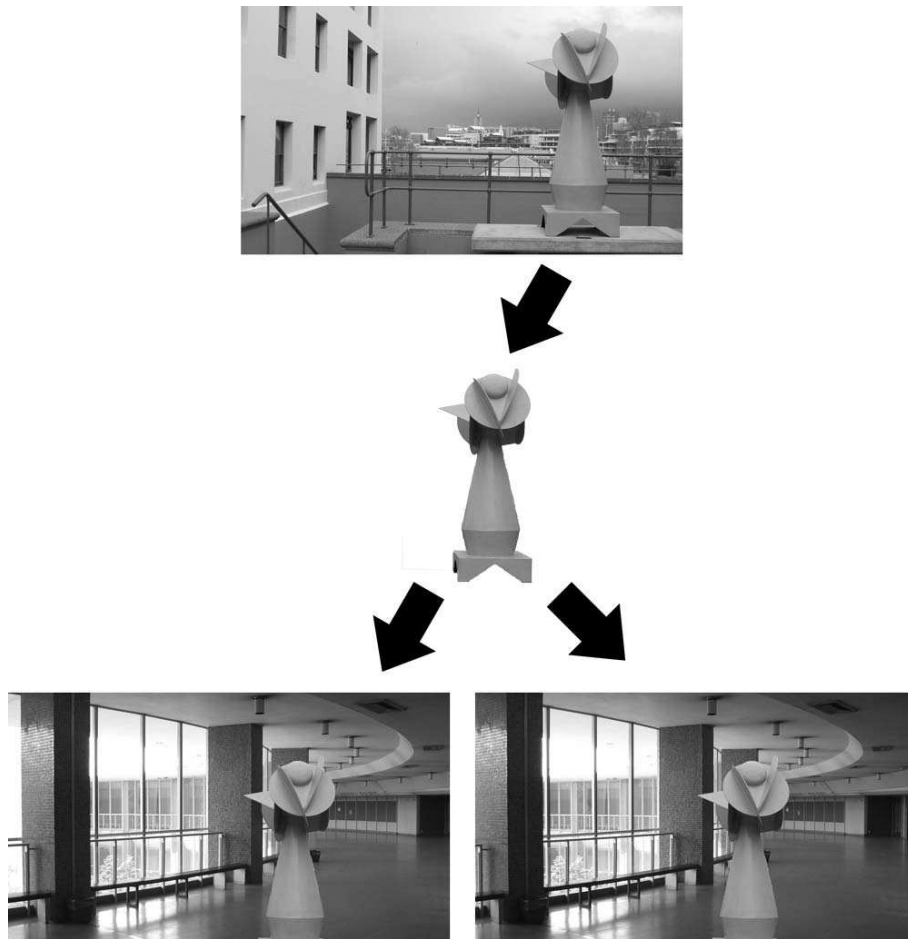


Figure 2. Non-stereoscopic to stereoscopic splicing.

age splicing has two forms: (i) non-stereoscopic to stereoscopic splicing; and (ii) stereoscopic to stereoscopic splicing.

Non-Stereoscopic to Stereoscopic Splicing. In non-stereoscopic to stereoscopic splicing, a part of a non-stereoscopic image is combined with a stereoscopic image. The horizontal placement of the non-stereoscopic image part should be adjusted in both images of the stereoscopic image pair, so that the non-stereoscopic part appears to be at the desired distance from the camera. Figure 2 illustrates this approach.

When the non-stereoscopic to stereoscopic splicing approach is used, the non-stereoscopic part of the image tends to appear “flat” because it

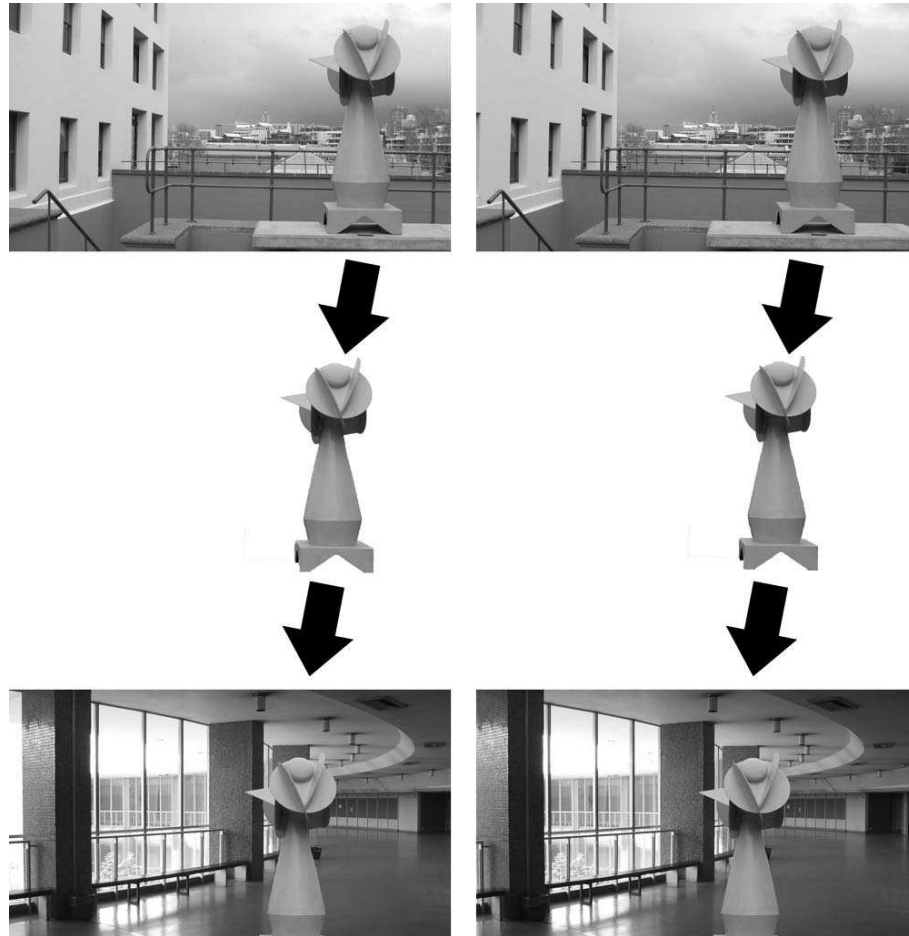


Figure 3. Stereoscopic to stereoscopic splicing.

has no internal depth. One way to address this problem is to adjust the horizontal positions of various sections of the non-stereoscopic part of the image. This could lead to gaps in the image that must be filled via image cloning or another technique. Time and patience are required, but if done correctly, the procedure can yield similar results as stereoscopic to stereoscopic splicing.

Stereoscopic to Stereoscopic Splicing In stereoscopic to stereoscopic splicing, parts of different stereoscopic images are combined. To maintain internal depth, the left and right images are combined separately. Every modification in the left image is duplicated in the right image. Figure 3 illustrates this approach.

Stereoscopic to stereoscopic splicing maintains the internal depth of the spliced images, but the splicing can still lead to irregularities. At a certain distance objects lose internal depth. When an object with internal depth is spliced and placed beyond this distance, it might still appear to have internal depth when internal depth is not expected. It could also happen that a spliced object has no internal depth in a stereoscopic image, but is placed at a point where internal depth is expected. These irregularities can indicate that the image has been spliced. In order to detect these irregularities, it is necessary to determine the distance at which an object starts to lose internal depth. The next section derives an equation for determining this distance.

3. Stereoscopic Distance Measurements

An object has no internal depth if the calculated distance of any two points on the object is the same. The following equation from Mrovlje and Vrancic [15] can be used to calculate the distance (D) from the camera of a single point in a digital stereoscopic image:

$$D = \frac{Bx_0}{2 \tan\left(\frac{\varphi_0}{2}\right)(x_R - x_L)} \quad (1)$$

In the equation, B is the horizontal distance between the lenses of the camera or cameras used to capture the stereoscopic image; x_0 is the horizontal image resolution in pixels; φ_0 is the viewing angle of the camera; and $(x_R - x_L)$ is the number of horizontal pixels that the point differs in the left and right images of a stereoscopic image pair. Note that Equation (1) assumes that the viewing directions of both lenses are parallel.

According to Equation (1), the distance of a point is dependent on the number of horizontal pixels by which the point differs in the two images of the stereoscopic image pair. When two points have the same difference in pixels, the points are the same distance from the camera. If the two points are the nearest and farthest points on an object with regard to the camera position, then the object has no internal depth. Deriving an equation for the distance at which an object loses internal depth requires the horizontal pixel difference of two points to be less than a pixel.

Let D be the distance of the closest point on an object with a depth of ΔD . Then, the distance of the farthest point on the object is $(D + \Delta D)$. Let $(x_R - x_L)$ be the number of pixels that the point at D differs in the left and right images of a stereoscopic image pair. Let $(x_{R2} - x_{L2})$ be the number of pixels that the point at $(D + \Delta D)$ differs in the left and

right images of a stereoscopic image pair. Then, the smallest number of pixels (P) needed to represent the internal depth for the object is given by:

$$P = (x_R - x_L) - (x_{R2} - x_{L2}) \quad (2)$$

Using Equation (1), we obtain:

$$D + (\Delta D) = \frac{Bx_0}{2 \tan(\frac{\varphi_0}{2})(x_{R2} - x_{L2})} \quad (3)$$

Substituting Equation (2) in Equation (3) yields:

$$D + (\Delta D) = \frac{Bx_0}{2 \tan(\frac{\varphi_0}{2})((x_R - x_L) - P)} \quad (4)$$

Upon simplifying Equation (4) using Equation (1), we obtain:

$$0 = [2P \tan(\frac{\varphi_0}{2})]D^2 + [2P(\Delta D)\tan(\frac{\varphi_0}{2})]D + [-B(\Delta D)x_0] \quad (5)$$

D can be solved using the quadratic equation:

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (6)$$

Since we are only interested in the positive distance in front of the camera, solving and simplifying D yields:

$$D = \frac{(\Delta D)}{2} \left[\left(\sqrt{1 + \frac{2Bx_0}{P(\Delta D) \tan(\frac{\varphi_0}{2})}} \right) - 1 \right] \quad (7)$$

where ΔD is the object depth; B is the difference in the horizontal camera lens positions; x_0 is the horizontal pixel resolution; φ_0 is the camera viewing angle; and P is the maximum internal depth. The distance D where $P = 1$ is the approximate distance at which an object starts losing internal depth.

4. Experimental Methodology

An experiment was conducted to test the accuracy of Equation (7), which specifies the distance at which an object loses internal depth. The experiment also tested the results obtained upon applying the equation to stereoscopic image splicing detection.

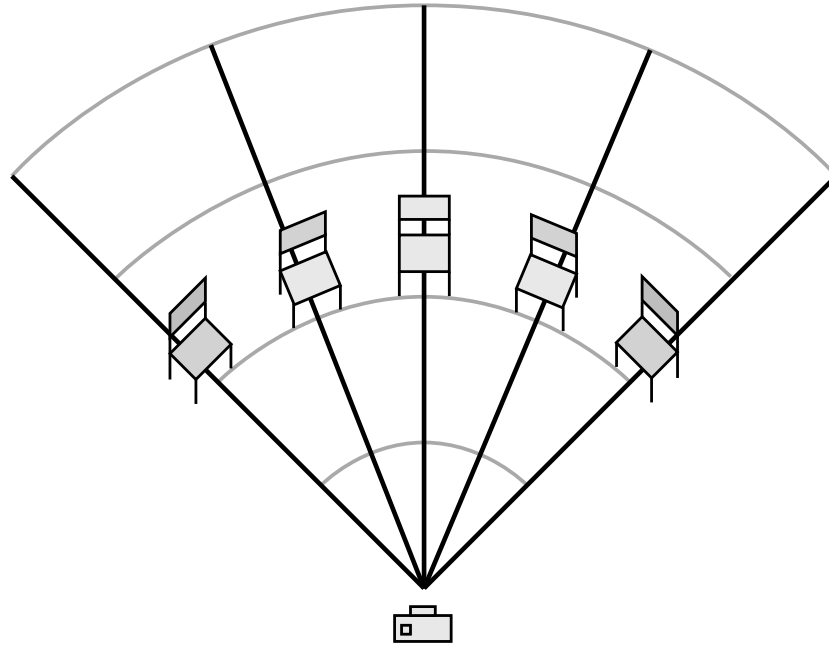


Figure 4. Chair placement illustration.

4.1 Distance Prediction

Stereoscopic images were created of a scene comprising five chairs. The five chairs were set up with two targets each. One target was placed in front of each chair and one against the back of each chair. The distance between the targets was 37cm, representing the internal depth of each chair.

The five chairs were placed in an arc in front of the camera position (Figure 4). The angle between each adjacent chair was 10° , so that the chairs would fit in the camera frames. The chairs were moved and photographed at 5m, 10m, 15m and 20m from the camera position. These distances were chosen because, according to Equation (7), the chairs lose internal depth somewhere in this range. Figure 5 shows a photograph of the chairs placed at 5m.

Two cameras were used to photograph the chairs. The first was a Fujifilm FinePix REAL 3D W3; this camera was used to take 2D and 3D photographs. The second camera used was a Sony Cybershot DSC-P43. The cameras were chosen because they provided different variables for testing Equation (7).



Figure 5. Chair placement photograph.

When the chairs were photographed with a single lens, the camera was moved on a horizontal line at 0mm, 25mm, 50mm, 75mm and 100mm. Stereoscopic images were created using the 0mm image as the left image and the 25mm, 50mm, 75mm and 100mm images separately as the right image in each stereoscopic image pair. Stereoscopic images were created at different lens distances to accommodate the fact that the distance of the lenses differs in different stereoscopic image capturing devices.

Table 1 provides information about the cameras used in the experiment. Table 2 specifies the distance at which the chairs are expected to lose internal depth.

After the chairs were photographed at the four distances, the images were downloaded. For each image, the horizontal pixel distance between the targets on each chair was measured separately and recorded. The measurements were taken in pixel units.

Table 1. Camera information.

Make Model	Sony		Fujifilm	
	DSC-P43		FinePix REAL 3D W3	
Mode	2D (Single Lens)		2D (Single Lens)	3D (Double Lenses)
Resolution	2304 × 1728		3584 × 2016	3584 × 2016
Focal Length	5mm		6.3mm	6.3mm
Sensor	1/2.7" CCD		1/2.3" CCD	1/2.3" CCD × 2
Sensor Size	5.37 × 4.04mm ²		6.16 × 4.62mm ²	6.16 × 4.62mm ²
Angle of View	56.47°		52.11°	45°
Inter-Lens Dist.	-		-	75mm

Table 2. Distance at which chairs ($\Delta D = 0.37m$) lose internal depth.

Make Model	Sony				Fujifilm				
	DSC-P43				FinePix REAL 3D W3				
Mode	2D				2D				3D
Image Width	2304				3584				3584
Angle of View	56.47°				52.11°				45°
Distance of Lenses (mm)	25	50	75	100	25	50	75	100	75
Predicted Distance (m)	4.27	6.12	7.53	8.73	5.64	8.05	9.9	11.5	10.8

4.2 Splicing Detection

Two stereoscopic images of the chairs were used to test splicing detection. Both the images were taken with the Fujifilm FinePix REAL 3D W3 camera in the 3D mode.

Images of the chairs at 5m and 15m were used in the experiment. This is because, according to Equation (7) and the information in Table 1, the distance at which a chair with a depth of 37cm starts to lose internal depth is approximately:

$$D = \frac{(0.37)}{2} \left[\left(\sqrt{1 + \frac{2(0.075)(3584)}{(0.5)(0.37) \tan(\frac{45}{2})}} \right) - 1 \right] = 10.8m \quad (8)$$

Two tests of stereoscopic to stereoscopic splicing were conducted. First, a chair at 5m was spliced at a distance of 15m. Second, a chair at 15m was spliced at a distance of 5m. The spliced images were scaled to be the same size as the non-spliced chairs. The internal depths of the spliced chairs were measured and compared. All the measurements were taken in pixel units.

Table 3. Average measured internal depth of chairs at different distances.

Make Model	Sony DSC-P43				Fujifilm FinePix REAL 3D W3					
Mode	2D				2D				3D	
Distance of Lenses (mm)	25	50	75	100	25	50	75	100	75	
Average Measured Internal Depth of the Five Chairs in Pixels										
Chair Distance from Camera	5m	*** 0.6	1.0 ***	1.8 ***	2.2 ***	1.0 ***	1.4 ***	2.6 ***	3.4	4.8
	10m	0.2	0.6	0.6	0.6	0	0	0.6	0.8 ***	1.6 ***
	15m	0	0	0	0	0	0	0.4	0.6	0.6
	20m	0	0	0	0	0	0	0	0	0
*** Distance chairs should loose internal depth according to Table 2										

5. Experimental Results

This section discusses the experimental results.

5.1 Distance Prediction

Table 3 provides the average measured internal depths of the five chairs for the various cameras, camera settings and chair distances. For example, the average measured distance of 4.8 pixels was obtained for the chairs at 5m in the image taken with the Fujifilm camera in the 3D mode. The asterisks (***) indicate the point where the chairs should lose internal depth according to Equation (7) and Table 2. All the pixel distances above the asterisks are expected to be more than 1 pixel and those below are expected to be less than 1 pixel.

Table 3 indicates that Equation (7) accurately predicts the point at which an object loses its internal depth in most cases. The only exception is the image taken by the Fujifilm camera in the 2D mode. When the distance of the lenses is 100mm and the chairs are at 10m, the average internal depth is expected to be more than 1 pixel. However, the measured value of 0.8 pixels in Table 3 is less than expected. A possible explanation is camera lens distortion [20] and the position of the chairs relative to this distortion.

It is also necessary to note the internal depth values at 20m. All these values are zero, which means that all the chairs had no internal depth

at this point. A value less than one can be chosen for P in Equation (7) to obtain the point at which an object always has no internal depth. Calculating the appropriate value for P is beyond the scope of this paper.

5.2 Splicing Detection

This section discusses the results of splicing stereoscopic images before and after the point at which an object loses internal depth.

Splicing an Object with Internal Depth to No Internal Depth.

In this experiment, the center chair was spliced from a distance of 5m to a distance of 15m. The spliced chair was scaled to be the same size as the non-spliced chair in the same image. Before the spliced image was scaled, the internal depth was 5 pixels. The non-spliced chair had no internal depth.

After splicing and scaling, the non-spliced chair still had no internal depth, but the spliced chair had an internal depth of 2 pixels. Thus, the spliced chair had internal depth at a point where internal depth should not be present. This indicates that a spliced object can be detected if the object has internal depth beyond the distance predicted by Equation (7).

Splicing an Object with No Internal Depth to Internal Depth.

In this experiment, the center chair was spliced from a distance of 15m to a distance of 5m. The spliced chair was scaled to be the same size as the non-spliced chair in the same image. Before the spliced image was scaled, there was no internal depth. The non-spliced chair at 5m had an internal depth of five pixels.

After splicing and scaling, the internal depth of the non-spliced chair was still five pixels. However, the internal depth of the spliced chair was measured as two pixels compared with its original value of zero pixels.

The graphs in Figure 6 explain the reason for the increased internal depth. The internal depth was measured in pixels using markers such as the darkest point on an edge. When the darkest point on an edge is estimated using the graphs, the distance of the point shows an increase in internal depth. The first graph shows that the difference in pixel distance between the left (L) and right (R) edge is less than one pixel, and, thus, there is no internal depth. When the edges are scaled, the difference in internal depth is scaled to become more than one pixel. Thus, when an image has an internal depth of less than one pixel, scaling can cause it to have an internal depth of more than one pixel.

The experimental results thus indicate that an image that is spliced from a distance where it was measured to have no internal depth may

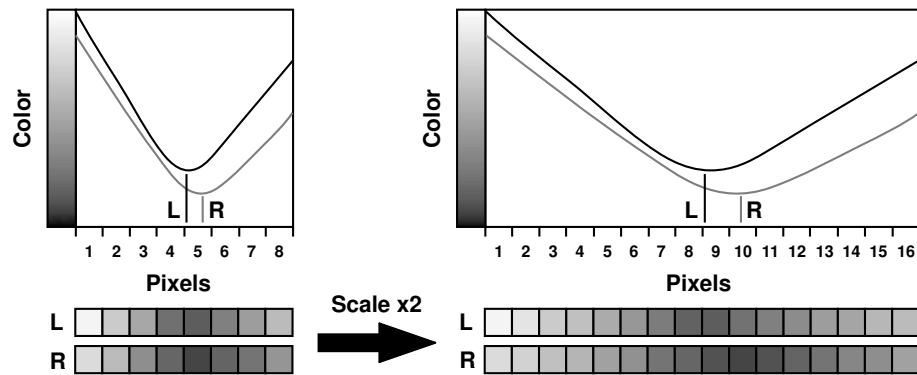


Figure 6. Increasing the scale increases the measured internal depth.

have internal depth after scaling. However, the internal depth of a spliced object can be significantly less than its internal depth when it is not spliced.

6. Conclusions

Understanding the relationship between the distance of an object and its internal depth can aid in the detection of spliced stereoscopic images. Experimental results demonstrate that the derived equation provides good estimates of the distance at which an object loses internal depth. The experimental results also show that it is possible to detect an object that is scaled and spliced from an area with internal depth to an area without internal depth. Moreover, an object that is spliced from an area without internal depth to an area with internal depth can potentially gain internal depth if scaled, but it may have a significantly smaller internal depth than expected. These results demonstrate that stereoscopic image splicing can be detected by comparing the physical depth of an object with its computed internal depth.

Our future research will continue to analyze the derived equation and its variables. Also, our research will investigate the effects of camera distortion and the position of a spliced object on internal depth.

References

- [1] M. Anderson, Can Mario make the leap? [Tools and Toys], *IEEE Spectrum*, vol. 48(6), pp. 24–26, 2011.
- [2] Camera and Imaging Products Association, CIPA DC-007 Standard: Multi-Picture Format, Tokyo, Japan, 2009.

- [3] N. Dodgson, Autostereoscopic 3D displays, *IEEE Computer*, vol. 38(8), pp. 31–36, 2005.
- [4] E. Dubois, A projection method to generate anaglyph stereo images, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. 1661–1664, 2001.
- [5] H. Farid, Image forgery detection, *IEEE Signal Processing*, vol. 26(2), pp. 16–25, 2009.
- [6] M. Fouche and M. Olivier, Detecting non-stereoscopic to stereoscopic image splicing with the use of disparity maps, *Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists*, pp. 271–274, 2011.
- [7] R. Granty, T. Aditya and S. Madhu, Survey of passive methods for image tampering detection, *Proceedings of the International Conference on Communication and Computational Intelligence*, pp. 431–436, 2010.
- [8] W. Grip, R. Grip and R. Morrison, Application of aerial photography and photogrammetry in environmental forensic investigations, *Environmental Forensics*, vol. 1(3), pp. 121–129, 2000.
- [9] M. Harris, Your next camera will shoot 3-D [Tools and Toys], *IEEE Spectrum*, vol. 48(1), p. 22, 2011.
- [10] S. Kim and E. Kim, A new liquid crystal display based polarized stereoscopic projection method with improved light efficiency, *Optics Communications*, vol. 249(1-3), pp. 51–63, 2005.
- [11] G. Lawton, 3D displays without glasses: Coming to a screen near you, *IEEE Computer*, vol. 44(1), pp. 17–19, 2011.
- [12] P. Leger, R. Deen and R. Bonitz, Remote image analysis for Mars Exploration Rover mobility and manipulation operations, *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, pp. 917–922, 2005.
- [13] C. Martin, LG Optimus 3D vs. HTC Evo 3D head to head review, *TheInquirer.net*, September 7, 2011.
- [14] Y. Moudden, P. Venault, A. Barnacka, D. Calvet, J. Glicenstein and M. Vivier, The level 2 trigger of the H.E.S.S. 28 meter Cherenkov Telescope, *IEEE Transactions on Nuclear Science*, vol. 58(4), pp. 1685–1691, 2011.
- [15] J. Mrovlje and D. Vrancic, Distance measuring based on stereoscopic pictures, *Proceedings of the Ninth International Ph.D. Workshop on Systems and Control*, vol. 2, pp. 1–6, 2008.

- [16] A. Rocha, W. Scheirer, T. Boult and S. Goldenstein, Vision of the unseen: Current trends and challenges in digital image and video forensics, *ACM Computing Surveys*, vol. 43(4), article 26, 2011.
- [17] J. Stamas, Sony unveils HDR-TD10 3D camcorder, *Camcorder-Info.com*, January 5, 2011.
- [18] M. Stockfisch, Prospective standards for in-home 3D entertainment products, *Digest of Technical Papers of the International Conference on Consumer Electronics*, pp. 133–134, 2010.
- [19] S. Volbracht, K. Shahrababaki, G. Domik and G. Fels, Perspective viewing, anaglyph stereo or shutter glass stereo? *Proceedings of the IEEE Symposium on Visual Languages*, 1996.
- [20] J. Wang, F. Shi, J. Zhang and Y. Liu, A new calibration model of camera lens distortion, *Pattern Recognition*, vol. 41(2), pp. 607–615, 2008.
- [21] S. Wang, J. Chen, Z. Dong and R. Ledley, SMIS – A real-time stereoscopic medical imaging system, *Proceedings of the Seventeenth IEEE Symposium on Computer-Based Medical Systems*, pp. 197–202, 2004.
- [22] C. Wheatstone, On some remarkable, and hitherto unobserved, phenomena of binocular vision, *Philosophical Transactions of the Royal Society of London*, vol. 128, pp. 371–394, 1838.
- [23] A. William, Stereoscopic visualization of scientific and medical content for education: Seeing in 3D, *Proceedings of the Fourth IEEE International Conference on eScience*, pp. 319–320, 2008.