**Article**

# Camouflage Breaking with Stereo-vision Assisted Imaging

Han Yao , Libang Chen , Jinyan Lin , Yikun Liu [*] , Jianying Zhou [*]

*Article*

# Camouflage Breaking with Stereo-Vision Assisted Imaging

**Han Yao [1,†], Libang Chen [2,†], Jinyan Lin [1,†], Yikun Liu [2,*] and Jianying Zhou [1,3*]**

[1]   School of Physics, Sun Yat-sen University, Guangzhou 510275, China; yaoh29@mail2.sysu.edu.cn (H.Y.); linjy1603@163.com (J.L.)

[2]   School of Physics and Astronomy, Sun Yat-Sen University (Zhuhai Campus), Zhuhai 519082, China; chenlb35@mail2.sysu.edu.cn (L.C.)

[3]   Guangzhou Midstereo Co., Guangzhou 510275, China

[*]   Correspondence: liuyk6@mail2.sysu.edu.cn (Y.L.); stszjy@mail.sysu.edu.cn (J.Z.)

[†]   These authors contributed equally to this work.

**Abstract:** Camouflage is a natural or artificially process to prevent an object from being detected while camouflage breaking is a countering process for the identification of the concealed object. We report that a perfectly camouflaged object in a two-dimensional scene can be retrieved and detected with stereo-vision assisted three-dimensional (3D) imaging perceived with stereopsis. The analysis is based on binocular energy model applied to general 3D settings. We show that a perfectly concealed object with random noise background can be retrieved with vision's stereoacuity to resolve the hidden structures. The theoretical analysis is further tested and demonstrated with distant natural images taken by a drone-camera, processed with a computer and displayed in an autostereoscopy. The recovered imaging is presented with removal of the background interference to demonstrate the general applicability for camouflage breaking with stereo imaging and sensing.

**Keywords:** camouflage breaking; binocular energy model; stereo-vision assisted 3D imaging

## 1. Introduction

Camouflage is a natural concealment capability of animals to hunt and to avoid being hunted [1,2]. The making and breaking of camouflage become important for achieving advantages for various confrontation circumstances [3,4]. Camouflage is usually achieved by combination of structures, colors and illuminations to hide the identity of the object from its background [5,6]. There has been consistent interest in the subject of camouflage, as the topic is not only scientifically interesting but technologically important. Among various technique for camouflage breaking, 3D convexity, machine learning and artificial intelligence have been widely studied [7,8]. It has been suggested that breaking camouflage is one of the major functions of stereopsis [9,10]. Recent development with 3D photographic and stereopsis technology [11,12] paves the way for stereo image reconstruction. However, the capability of stereo computer vision is yet not in any way comparable with human vision. It is then necessary to re-examine the effect of stereopsis on camouflage breaking and to analyze its impact on its real-world applications.

In this work, we show with vision's energy model [13–15] that 3D structures hidden deep in the background can be retrieved with binocular vision via stereoacuity. This theoretical analysis is applied to the extreme case where the object is completely submerged with background noises. The results show that the submerged 3D structures can be recovered if the recorded images are rightly perceived with human vision's stereopsis. A field experiment is carried out with natural images taken by a drone camera. These image data are processed with computer, displayed in an autostereoscopy [16,17] and perceived with vision. The removal of the background noises for both the random noise and the natural images clearly demonstrates the general applicability of the novel technique for camouflage breaking.

**2. Theoretical Analysis**

The theoretical analysis presented in this work is based on the vision's energy model proposed by Ohzawa *et al.* [13–15]. The image information is provided by the spatial array of binocular energy neurons that are identical with each other except for their receptive field locations. We use $L(x)$ and $R(x)$ to respectively represent the complex monocular linear responses of the left and right eye, at receptive position $x$. The spatial array of energy responses, $E(x)$, is expressed as:

$$E(x) = |L(x) + R(x)|^2$$
$$= |Re[L(x)] + Re[R(x)]|^2 + |Im[L(x)] + Im[R(x)]|^2 \tag{1}$$

While *Re* means the real part, *Im* means the imaginary part. The left and right eye responses are denoted as $L(x) = \rho_l^2 e^{i\phi_l(x)}$, $R(x) = \rho_r^2 e^{i\phi_r(x)}$, where $\rho_l^2$ and $\rho_r^2$ are the monocular energy containing amplitude information, $\phi_l(x)$ and $\phi_r(x)$ are the phase angle of the monocular response. With the amplitude and phase information, the above equation can be simplified as:

$$E(x) = \rho_l^2(x) + \rho_r^2(x) + 2\rho_l\rho_r\cos(\Delta\phi) \tag{2}$$

$\Delta\phi = \phi_l(x) - \phi_r(x)$ refers to interocular phase difference.

Notice that the equation shows that the images processed with stereo-vision contains a crossed term between the left and right eyes. This binocular interaction gives rise to the additional information that is not available with single eye vision or with their addition and it can be recovered only with simultaneous binocular vision.

To analyze the stereo-vision with the energy model, it is important to note how the lateral and longitudinal spatial resolutions are related to the stereo-vision. Stereoacuity is defined as human's vision ability to resolve the smallest depth difference measured with angular difference, described as $\Delta T$ [18,19]:

$$\Delta T = \frac{dZ \times a}{Z^2} \tag{3}$$

The parameters for stereoacuity are shown in Figure 1, where $a$ is the interocular separation of the observer, $Z$ represents the distance of the fixed peg from the eye and $dZ$ means the position difference.
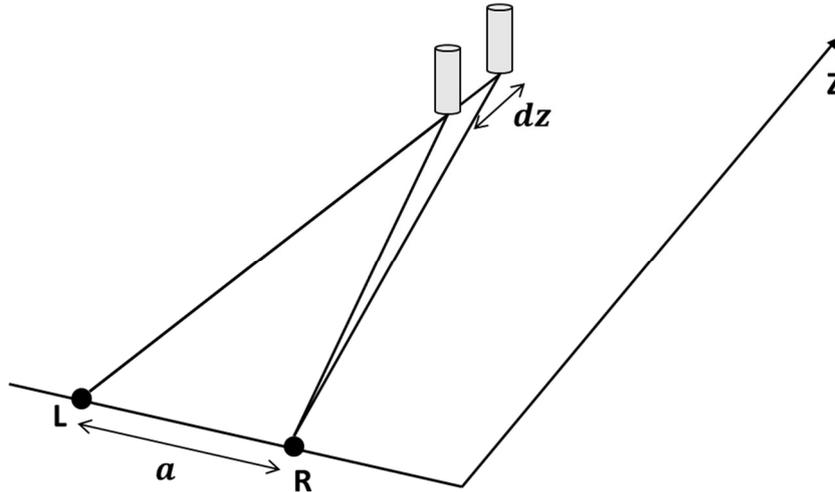


**Figure 1.** The architecture of stereoacuity measurement with the respect to observer's interocular separation.

For object taken with a camera and displayed on a stereopsis, images can be magnified with optics and stereopsis screen size. Considering total magnification factor and baseline difference, the stereoacuity can be approximated with Equation (4): [18,19]

$$\Delta T = \frac{dZ \times M \times N \times a}{Z^2} \tag{4}$$

where $M$ is magnification (ratio of display FOV and camera FOV), $N$ is the effective interocular separation (ratio of baseline and interocular separation).

An interesting but rarely discussed topic is how the stereoacuity is affecting the lateral resolution of the image. In general, the depth resolved measurement should enhance the lateral resolution in a similar way as with a confocal configuration while the rejection of the out of focus component is expected to increase the lateral resolution [20–22].

## 3. Results and Discussion

As the first step, we examine an extreme case where the object is a randomly sampled blocks of the letter "SYSU" of which each block in the left-hand and right-hand image consists of a random disparity. The object is completely swamped with random dot noises as shown in the Figure 2a and 2b. The Figure 2a and 2b show no useful information except for a noise background. On the other hand, when these left and right images are perceived with a stereopsis, meaningful signal clearly appears, as shown in the Figure 2c. The appearance of Figure 2c can be readily observed by delivering the left and right images into the left and right eyes, with which a pattern of "SYSU" is imaged and located at the center of the picture. The estimation of pixel disparity between Figure 2a and Figure 2b is necessary for the calculation of Figure 2c. Equation (2) could be regarded as the summation of two components, the constant item originated from monocular energy and the binocular crossed item which takes the disparity as variable.
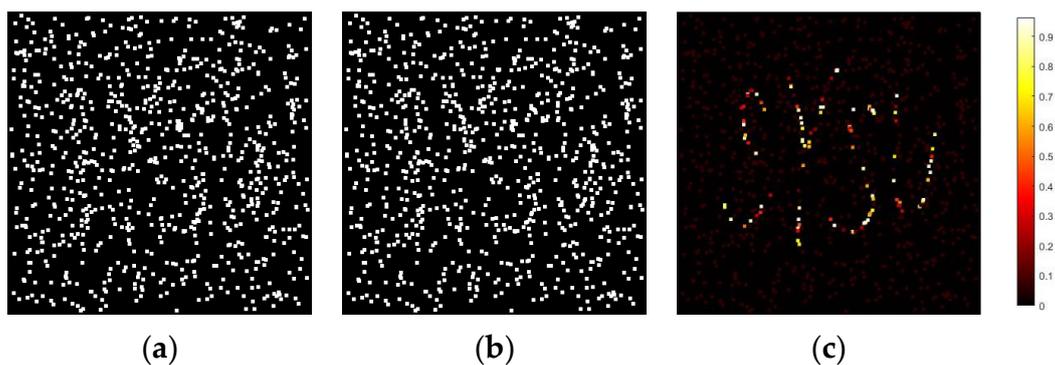


**Figure 2.** Random noise pattern viewed by the left eye (**a**), right eye (**b**). (**c**) shows the relative sense intensity when two eyes focus on the nonzero disparity layer.

Notice that the appearance of the pattern with stereopsis is a conventional technique for the screening of stereoacuity [23] and it is an implication of how binocular viewing can be applied to perceive the signal that is not possible with a single eye. With this result, one can see that the object can be concealed within a 2D imaging, either by the left or right eye, but it becomes visible with binocular viewing.

For the test of the camouflage breaking in natural scene, a field experiment is carried out to show the retrieval of object with stereoacuity assisted image sensing. The difficulty of 3D remote sensing with stereopsis arises from the required large baseline distance, which is studied in this work with a drone equipped with a single camera. For the purpose of executing stereoacuity functioning, an experimental architecture similar as Figure 1 is carried out and left- and right-images are selected from a video taken by the flying camera. The key parameters to be considered are the relative positions of the object with its background, and the baseline of the two-positions to provide the stereo-viewing so that the object can be resolved with stereoacuity measurement. Assuming, for example, an object located 100 meters away from the camera, and an object is 1 meter away from its background. It can be estimated that the baseline of the dual cameras must be approximately 57 cm apart. With the recorded navigating speed of the drone, it is straight forward to choose left and right images from the video format to meet the requirement of the baseline length.

To meet the sensing requirement, a drone (DJI Mavic2 pro: 28 mm equivalent lens on the 1-inch CMOS sensor) equipped with a camera is employed to take the video for the scene and the two

images from selected positions are employed as input for the left- and right-eye images. The flying altitude of the drone is 90 m. The dual images are transmitted to an autostereoscopy for 3D viewing at a viewing distance of 90 cm. The autostereoscopy is a directional backlight illuminated liquid crystal display with the feature size of 24 inch and single eye resolution at 1080 P (Type Midstereo 2468) [16,17]. With this autostereoscopy, the 3D scene is clearly visible. For the convenience of 3D perception, we convert the left and right images by coding it with blue and red, so that the image can be perceived with a 3D anaglyph.

The use of an autostereoscopy makes it possible to study the difference between 2D and 3D perception. For 2D mode, the left and right eyes are perceiving planar image. With 3D perception, the left and right images are delivered to the left and right eyes respectively. It becomes evident that 3D perception provides detailed object structures due to the perception of layered structures of the scene assisted with stereoacuity. With 2D mode, the information in different depths is intermixed, giving rise to the opportunity for camouflaging. With 3D, the concealed can be retrieved with binocular vision.

In order to further view the effect of stereopsis on the resolved 3D object (compared with Figure 3a), image is synthesized using disparity discriminated image pixels as input information. Detailed structures can be observed with stereo-vision concentrating on a unique disparity layer. By enhancing the image structure in one particular layer while removing structures from other layers, the lateral structures of the scene is substantially enhanced.
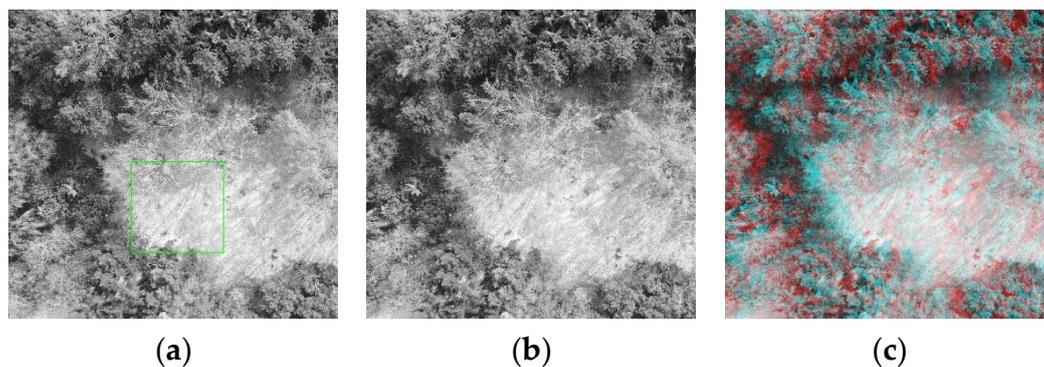


(**a**)          (**b**)          (**c**)

**Figure 3.** Left- (**a**) and right- (**b**) eye images and stereo anaglyph (**c**) synthesized by drone images taken at the altitude of 90 meters above the ground. The baseline distance between the left and right views is approximately 7 m. The tree leaves obliqued in 2D is made prominent with 3D.

Figure 4a shows the energy distribution, in which binocular vision focus on the ground layer, while Figure 4b illustrates the situation when two eyes are concentrating on the top tree layer, within which the tree branches and leaves are clearly visible. The pixel disparity between two visual angles, which is related to the interocular phase difference in Equation (2), is indispensable for the calculation of the energy distribution. The summation of the monocular energy from two eyes could be regarded as a constant offset.
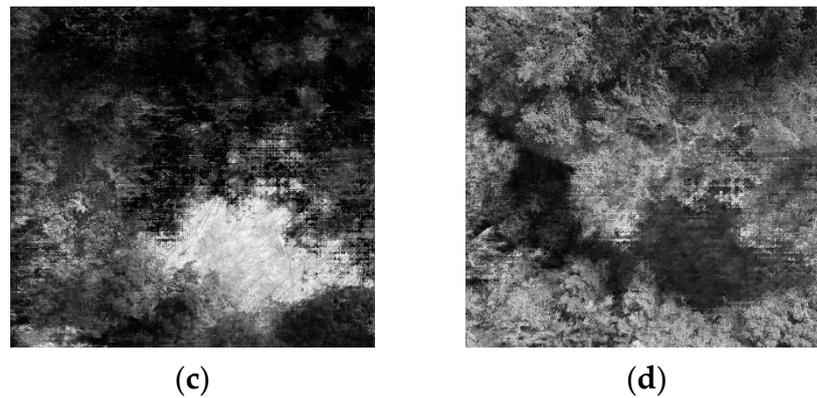


(**a**)          (**b**)

(c)    (d)
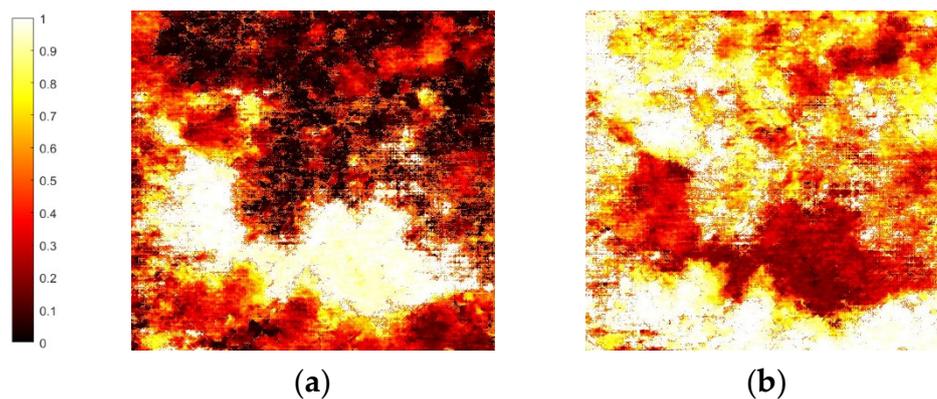
**Figure 4.** The energy distribution for the scene while two eyes focusing on the ground (**a**) and focusing on the tree (**b**). (**c**) and (**d**) are synthesized structures with weights respectively according to (**a**) and (**b**).

Figure 4c is obtained by dot multiplication of the energy pattern at a particular height of the tree with input of Figure 3a. Similarly, Figure 4d is resulted from multiplying the energy pattern at the height of the ground with a Figure 3a. Prominent difference is clearly visible in the comparison of these two cases mentioned above, showing substantially enhanced camouflage breaking capability. With binocular energy model, Figure 4c and Figure 4d show corresponding sectioning of the 3D scene for the different disparities, corresponding to pattern at different altitude. The sectioning images are in accord with the direct perception shown in Figure 3c. With the synergistic effect of both eyes, the focusing layer could be differentiated in depth and clearer in lateral resolution by rejecting structure information from different sections of the object. Compared with the corresponding region of the original image given in Figure 3a, neither ground or tree characteristics can be identified with the original image, which proves a promising way for camouflage breaking for stereo-vision and stereo-vision based optical imaging.

To better illustrate the advantages of layered discrimination with the energy model, we select a region in Figure 3a marked by a green square, which is amplified in Figure 5. It becomes clear that the image structures unclear from a bright background can be sectioned. The 2D image sections focusing to the ground and to the tree branches are clearly visible.
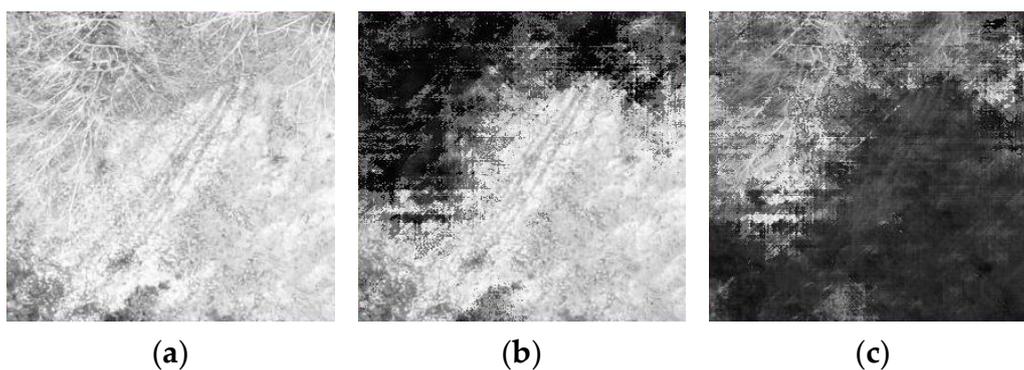


(a)    (b)    (c)

**Figure 5. (a)** is the perceived 2D image pattern and (**b**) and (**c**) show the sectioning image patterns focused to the ground and to the top of the trees.

## 4. Conclusion

In conclusion, based on the energy model of binocular vision, we analyze stereoacuity for image sensing, which is applied to camouflage breaking to the case where the object is completely obscured by the background noises. We employed the stereoacuity resolved sensing to a field experiment, and we restore the object which is concealed in 2D settings, hence demonstrating the enhanced capability for camouflage breaking even for complex background. It is shown that the present technique is

versatile and powerful on the basis of a well understanding of stereopsis and binocular interaction. We anticipate that the proposed and demonstrated technique has immediate applications for scientific research and technological applications.

**Author Contributions:** H.Y. carried out field measurement and image analysis; J.L. did theoretical analysis, energy model simulation, and image reconstruction; L.C. did theoretical analysis for energy model and field experimental measurement; Y.L. conducted the project coordination and data analysis; J.Z. proposed the project and wrote the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mondal A. Camouflage design, assessment and breaking techniques: a survey. *Multimedia Syst.* **2022**, *28*, 141-160.
2. Stevens M.; Merilaita S. Animal camouflage: current issues and new perspectives. *Philos. Trans. R. Soc. B* **2009**, *364*, 423-427.
3. Lamdouar H.; Xie W.; Zisserman A. The making and breaking of camouflage. *Proc. IEEE/CVF Int. Conf. Comput. Vis.* **2023**, 832-842.
4. Susan G.; John C.; Kevin R.; David A. Breaking camouflage: Binocular disparity reduces contrast masking in natural images. *J. Vis.* **2010**, *10*, 38-38.
5. Tankus A.; Yeshurun Y. Convexity-based visual camouflage breaking. *Comput. Vis. Image Underst.* **2001**, *82*, 208-237.
6. Stevens M.; Ruxton G. The key role of behaviour in animal camouflage. Biol. Rev. **2019**, *94*, 116-134.
7. Fan D.; Ji G.; Sun G.; Cheng M.; Shen J.; Shao L. Camouflaged object detection. *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* **2020**, 2777-2787.
8. Ji G.; Fan D.; Chou Y.; Dai D.; Liniger A.; Van G. Deep gradient learning for efficient camouflaged object detection. *Mach. Intell. Res.*, **2023**, *20*, 92-108.
9. Julesz B. *Foundations of cyclopean perception*. Publisher: U. Chicago Press., 1971.
10. Wardle S.; Cass J.; Brooks K.; David A. Breaking camouflage: Binocular disparity reduces contrast masking in natural images. *J. Vis.* **2010**, *10*, 38.
11. Brown M.; Burschka D.; Hager G.; Advances in computational stereo. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 993-1008.
12. O'Connor A.; Birch E.; Anderson S.; Draper H. The functional significance of stereopsis. *Invest. Ophthalmol. Vis. Sci.*, **2010**, *51*, 2019-2023.
13. Ohzawa I.; Deangelis G; Freeman R. Encoding of binocular disparity by complex cells in the cat's visual cortex. *J. Neurophysiol.* **1997**, *77*, 2879-2909.
14. Fleet D.; Wagner H.; Heeger D. Neural encoding of binocular disparity: energy models, position shifts and phase shifts. Vis. Res. **1996**, *36*, 1839-1857.
15. Hibbard P. Binocular energy responses to natural images. *Vis. Res.* **2008**, *48*, 1427-1439.
16. He Y.; Chen X.; Zhang G.; Fan Y.; Liu X.; Deng D.; Yan Z.; Liang H.; Zhou J. A directionally illuminated pixel-selective flickering-free autostereoscopic display. *Displays* **2024**, *82*, 102651.
17. Zhang A.; Chen X.; Wang J.; He Y.; Zhou J. Directionally Illuminated Autostereoscopy with Seamless Viewpoints for Multi-Viewers. *Micromachines*, **2024**, *15*, 403.
18. Jennings J.; Charman W. Depth resolution in stereoscopic systems. *Appl. Opt.* **1994,** *33*, 5192-5196.
19. Kytö M.; Nuutinen M.; Oittinen P. Method for measuring stereo camera depth accuracy based on stereoscopic vision. *Proc. SPIE* **2011**, *7864*, 168-176.
20. Webb R. Confocal optical microscopy. *Rep. Prog. Phys.* **1996**, *59*, 427.
21. Xie X.; Chen Y.; Yang K.; Zhou J. Harnessing the point-spread function for high-resolution far-field optical microscopy. *Phys. Rev. Lett.* **2014**, *113*, 263901.

22.   Yang L.; Xie X.; Wang S.; Zhou J. Minimized spot of annular radially polarized focusing beam. *Opt. Lett.* **2013**, *38*, 1331-1333.
23.   Wang Y.; Zhong J.; Cheng M.; Li J.; Ma K.; Hu X.; Li N.; Liang H.; Zhu Z.; Zhou J.; Yuan J. A novel clinical dynamic stereopsis assessment based on autostereoscopic display system. *Ann. Transl. Med.* **2022**, *10*.