

Article

Foreground Scattering Elimination by Inverse Lock-in-Like Spatial Modulation

Yueyu Lin ^{1,2,3}  and Sune Svanberg ^{1,2,3,*}

¹ National Center for International Research on Green Optoelectronics, South China Normal University, Guangzhou 510006, China; yueyu.lin@coer-scnu.org

² Guangdong Provincial Key Laboratory of Optical Information Materials and Technology & Center for Optical and Electromagnetic Research, South China Academy of Advanced Optoelectronics, South China Normal University, Guangzhou 510006, China

³ Department of Physics, Lund University, P.O. Box 118, SE-221 00 Lund, Sweden

* Correspondence: sune.svanberg@fysik.lth.se

Received: 9 July 2020; Accepted: 10 August 2020; Published: 13 August 2020



Abstract: We describe a simple approach to enhance vision, which is impaired by close range obscuring and/or scattering structures. Such structures may be found on a dirty windscreen of a car, or by tree branches blocking the vision of objects behind. The main idea is to spatially modulate the obscuration, either by periodically moving the detector/eye or by letting the obscuration modulate itself, such as branches swinging in the wind. The approach has similarities to electronic lock-in techniques, where the feature of interest is modulated to enable it to be isolated from the strong perturbing background, but now, we modulate the background instead to isolate the static feature of interest. Thus, the approach can be denoted as “inverse lock-in-like spatial modulation”. We also apply a new digital imaging processing technique based on a combination of the Interframe Difference and Gaussian Mixture models for digital separation between the objects of interest and the background, and make connections to the Gestalt vision psychology field.

Keywords: vision; scattering; modulation; lock-in techniques; background subtraction; Gestalt psychology

1. Introduction

Unimpaired vision is a very fundamental quality of life. Most visual defects can be corrected by glasses. Glare related to slant-angle specular reflections can be reduced by polarization filters. Good vision is also clearly related to adequate illumination. Normal human color vision requires object illumination by white light to activate our three types of color receptors, signaling particular colors after processing the color sensitive absorptions of the different image parts in the brain [1,2]. These receptors in the cones have low sensitivity and are, thus, inactive in low light level environments, where instead, the more sensitive rods are employed in black-and-white colorless vision. Thus, although the reflective “colors” are clearly physically equally present in low-light conditions, night vision devices employing, e.g., electro-optical image intensifiers, give black-and-white images only, actually black and green images, because of the emission characteristics of the most efficient electro-optical phosphor. However, night color vision can be obtained by separating different color bands before amplification, and then add the color coded enhanced images into a normal color image. This can be done by image processing, following, e.g., imaging color splitting by a split mirror Cassegrainian telescope, where three images are intensified simultaneously in a single unit [3]. Alternatively, rotating color-glass sectors are used in front of the image intensifier, which produces a magnified image on a white-phosphor screen, which is then observed through co-rotating color-glass sectors for flicker-less color image formation in the brain once the rotating frequency is high enough [4].

Vision can also be strongly influenced by the properties of the light transmission medium, normally the atmosphere. While there is no major molecular absorption in the spectral region of human vision (400–700 nm), scattering due to fog or smoke can certainly strongly impair vision. A remedy, frequently used in military applications, is to perform the observation at infra-red wavelengths, where Mie and Rayleigh scattering are less prominent due to the strong wavelength dependence of these processes [5,6].

A further origin of vision impairment is the scattering of a close lying, physically structured semi-transparent layer, making the contrast or isolation of remote objects very difficult. The present paper deals with this problem. We describe the phenomenon and provide a remedy based on vision physiology. We also make a connection to well-known digital imaging processing techniques used in background subtraction, and foreground removal by integral imaging.

2. The Problem and Its Solution

A common imaging situation yielding strongly impaired recognition is depicted in Figure 1. The objects of interest are at a distance but are viewed through a strongly scattering window, featuring structure. Light from the close lying object completely dominates over the light derived from the distant objects of interest. An everyday situation along the lines just described is viewing through a dirty or frosty windscreen while driving, making it difficult to see the road and the objects on or adjacent to the road. The common remedy is, of course, to use an efficient screen wiper, frequently combined with spraying the screen from outside with water or an anti-frost solution. Experience shows that this does not always make the situation satisfactory, especially not in situations with counter-illumination from the sun or from the headlights from approaching vehicles. The objects of interest remain hard to discern. We illustrate this situation in Figure 1.



Figure 1. Impaired vision sometimes encountered during nighttime driving. An animal, approaching from the left, is very hard to discern.

A further, somewhat different situation might be when the object is partly hidden by vegetation, i.e., leaves and branches, which may be moving in the wind. How could the object of interest behind be imaged clearly? Further, there may be raindrops or snowflakes falling in the line of sight to the object of interest, which impair vision. Such situations can be handled by well-known digital image processing techniques of background subtraction and integral imaging (see, e.g., [7–13]), but the human brain is also remarkably able to link together image elements, which are not accessible simultaneously, but only in a temporal sequence. We will discuss such situations in the present paper, while not covering cases when the object of interest is totally and permanently blocked from direct view, but where many interesting image processing approaches have recently been presented (see, e.g., [14], and references therein).

It is important to point out that we are also not considering situations where the scattering medium is more or less homogeneous, such as fog or biological tissue. Then the object of interest, frequently having sharp structures, is observed on a strong, homogenous background, making contrast minute or absent. For such situations, new and fascinating methods have recently been developed based on scattering compensation with spatial light modulators, etc. [15–18]. We are also not considering cases when the illumination of the object can be made structured, to ensure a spatial tagging, which enhances the object while reducing the non-structured background, e.g., in SLIPI (structured laser illumination planar imaging) applications [19,20]. A further way to enhance distinct distant objects over a close lying scattering medium (whether homogeneous or structured) is to simply apply far field focusing, resulting in distant objects of interest with sharp edges, which can be enhanced by digital processing [1,21,22].

The motivation to the present approach for improved vision comes from concepts of frequency and phase tagged (lock-in) detection, frequently used for retrieving weak signals residing on a strong background. The idea here is to identify the feature of interest by modulating some parameter related to the feature only, while not affecting the background. Examples of this approach are numerous in the field of spectroscopy, e.g., in wavelength modulation in diode-laser absorption spectroscopy, where the derivative of a weak but sharp absorption structure is observed by modulating the wavelength to an extent corresponding to a fraction of the feature line width (see, e.g., [23]). Another example is found in optical pumping or optical double resonance experiments [23], where the radiofrequency field inducing transitions between atomic sublevels is chopped on and off, and the minute changes in the detected light due to the action of the radio frequency are monitored selectively. The idea is to lift off and isolate the signal due to the feature of interest from the unrelated strong background. An imaging variety of such an approach could be to use modulated illumination for vision enhancement. Such illumination modulation could be taken, e.g., at nighttime driving, where the (LED) headlights of your car might be modulated in intensity (above the flicker frequency). Normal human vision would then be supported by a forward-looking video camera (possibly with a head-up display), which records synchronously and in phase with your own modulation, to enhance objects of interest to you, while direct light, or windscreen scattered light, from meeting car headlights, would be suppressed since it would be of a different modulation frequency and/or phase.

We will here focus on a different and more general approach for vision enhancement, where the enhancement occurs on the detection side only, rather than also taking the illumination side into account. One important application is when no artificial illumination occurs, e.g., in daylight applications, where the illumination obviously cannot be modulated, enabling the object of interest to be isolated from the disturbing background. The bearing principle in this paper is then to modulate instead the disturbing background, of which we can have control. It is achieved by a systematic spatial modulation by a periodic movement of the detector. If a total signal is observed and part of that signal is modulated, it can be lifted off, leaving the (frequently weak) residue isolated and clearly identifiable. Since the technique utilizes modulation, but for the opposite purpose as normally employed, the new approach could be denoted as “inverse lock-in-like spatial modulation” (ILLSPAM). A situation where spatial modulation could be implemented is the one shown in Figure 1.

3. Practical Implementation

The implementation of the proposed technique, using the physiology of vision and image perception, is the most straightforward application. The eye/brain system has an impressive power to “stabilize” an image, even if the detector is moving, and can easily discern the static parts of an image, while large parts of the “image” characterized by a periodic or uniform movement can be disregarded. Regarding vision psychology, there is a clear connection to the well-established *Gestalt* direction of vision apprehension (see, e.g., [24,25] and references therein). Here, the “figure–ground segregation” approach is taken, emphasizing the capability of the brain to figure out contexts, when considering the “whole” as a concept, in contrast to “the sum of the parts”; a holistic situation pertains. We, in the present treatment, more focus on the parallels to electronic signal enhancement, making the connection

to electronic lock-in isolation of the feature of interest. We put our approach in perspective to all the recent technological developments in vision enhancement in environments impaired by scattering.

An illustrative example, which is quite useful for drivers, is illustrated in Figure 2. If surprised with a sudden situation with foreground scattering, due to the headlights of a meeting car, as illustrated in Figure 1, or the impact of strong, counter sunlight, such as in Figure 2, vision is strongly impaired. For the latter case, the resulting image Figure 2c consists of the background scene of interest, as shown in Figure 2a, and the close-by foreground scene shown in Figure 2b added scattering dirt on the windscreen. The suggested method to isolate the two scenes for obtaining a better vision of the scene in Figure 2a is to move the head a few centimeters up and down periodically and concentrate on the nonmoving parts of the scene observed. This is illustrated in the movie, connected to Figure 2c, where the foreground scattering image, Figure 2b, is moved periodically up and down to simulate the car driving situation (Supplementary Materials, Video S1). We can clearly see how the physiological vision approach works very well to increase the apprehension of the road situation, including the pedestrian and the animal.

Likewise, the eye/brain can easily see and “lock” on a building, partly blocked by, e.g., close-distance tree branches, moving in the wind and exposing different parts of the building at different times, as further discussed below.

The simple physiological approach proposed above brings to mind a further, somewhat trivial trick, well known to myopic persons, who at nighttime from the bed would see a completely blurred distant LED-based clock, which becomes crystal clear, if observed through the small opening between the thumb and two fingers brought tightly together at their tips (central-ray imaging)!

Many digital image processing approaches of removing obscuring foreground or removal of uninteresting background have been developed (see, e.g., [7–13]), and we will put the new physiological approach into perspective by showing an illustrative example, where a standard smartphone is utilized. The approach can be readily applied to the case of active spatial modulation of a close range partially obstructing object (say, a dirty windscreen) when the camera or an attached optical arrangement is periodically spatially translated. Cases of passive, close or far range modulation can also be considered when, e.g., vegetation, rain, or snow in motion impairs the visualization of a distant object, e.g., a building.

We will now put the physiological vision approach in relation to digital image processing and will present a method to isolate the objects of interest. Figure 3 illustrates a laboratory set up, where a fixed smartphone camera was subjected to a transparent screen, 0.5 m from the camera, which was cluttered with an irregular pattern of scattering grey spots, and which were in periodic motion. An object of interest (the logos of the participating universities) was placed at a distance of 1 m, as seen in the background of Figure 3. Images were recorded with successive displacements of the disturbing cluttering screen.

Figure 4a shows an individual frame, where parts of the object of interest were blocked, while the right-hand part shows a processed image, which was mostly free of obstruction. A short video sequence is displayed in Video S2 in the Supplementary Materials, with individual frames occurring to the left, while the video sequence, processed as indicated above, is playing to the right in the figure, without obstruction. We note that the “modulation frequency”, i.e., the rate of foreground movement should be substantially higher than that of the scene recorded to achieve an optimal processed image impression for the case that the image parts of interest would be moving. Then our procedure was performed with the reconstructed image calculated as a “sliding average”, and the objects of interest appeared in a somewhat delayed film.

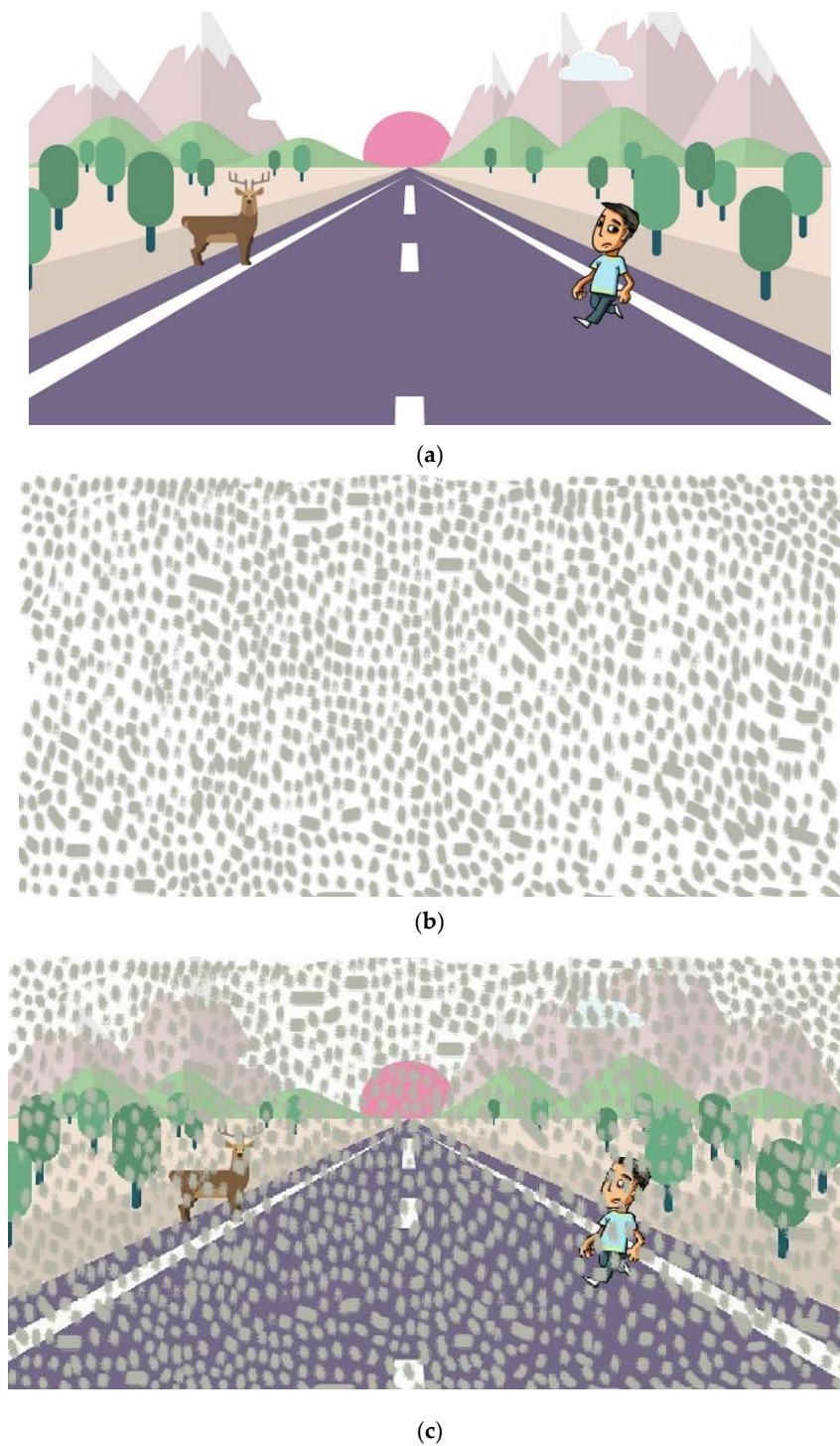


Figure 2. Illustration of how impaired vision sometimes encountered during car driving can be improved by a simple spatial modulation approach, utilizing the unprecedented image processing capability of the human brain. (a) The scene we want to observe without any image blurring; (b) Foreground image blurring, due to, e.g., dirt on the windscreen; (c) The resulting blurred image obtained as a result of the superposition of Figure 2a,b. Video S1 in the Supplementary Materials shows how Figure 2b is periodically translated up and down in front of Figure 2a, simulating a periodic movement of the observers head.

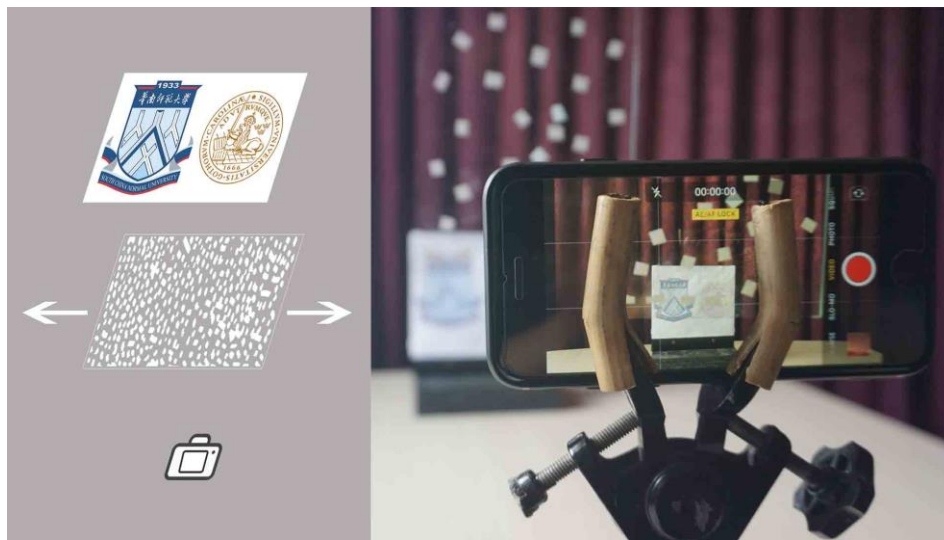


Figure 3. The laboratory setup with the object, the obscuring foreground screen and the smartphone video camera. (left) schematic setup, and (right) a photograph of the setup.



Figure 4. Individual frame (a) and a digitally processed image (b). Video S2 in the Supplementary Material shows a short sequence, with individual partly obscured frames to the left.

The approach we used in our demonstration, shown in Figure 4, is based on two already known methods. The Interframe Difference Method [11] is commonly used in motion detection, while it is not sensitive to targets, which are moving slowly. However, the Interframe Difference Method is simple and has good stability for light variations. The Gaussian Mixture Model (GMM) [12] builds the background by using M models, which come from the intensity of pixels. It also cannot detect targets with slow speed, because it builds the background by prehistory. GMM collects the intensity of a particular pixel value over a period of time for modeling. If the target moves slowly, it can wrongly be classified as a non-interesting (background) information, because the intensity of the pixel does not change much. At the same time, the Gaussian mixture model cannot deal with the problems arising from light changes because the background quality obtained by GMM often depends on the prehistory. We operated an interframe action based on the result obtained from GMM. It improved the sensitivity of GMM to light changes and improved the quality of background models as a feedback variable. By combining the two methods in our approach, certain advantages were obtained. We proceed as outlined below:

1. To increase the adaptability in updating of the training set, we defined a time T as the interval for the repeated processes of adding a new frame and discarding the oldest one. To prevent excessive computation time or possible motion, T should not be chosen too long. At time $t-1$, we then had the data set χ_{T-1} :

$$\chi_{T-1} = \{x^{(t-1)}, \dots, x^{(t-1-T)}\} \tag{1}$$

Here, $x^{(t-1)}$ is the current pixel value at time $t - 1$. The probability of observing the current pixel value $p(x)$ is:

$$p(x) = \sum_{m=1}^M \omega_m \cdot \eta(x, \mu_m, \sigma_m^2) \tag{2}$$

Here, the ω_m , values, which are positive and estimate the mixing weights, add up to one at time t . μ_m is the mean value of the i -th ($1 \leq i \leq m$) Gaussian in the mixture, and σ_m^2 is the estimate of the variance that describes the Gaussian models at time t . η is a Gaussian probability density function:

$$\eta(x, \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3}$$

- In every time period, T , foreground and background pixels were present at the same time. Considering the background target as always stable and similar, we ranked all the Gaussian models according to the estimated mixing weights ω_m . So we could approximate the background model from the first B largest models (i.e., those having the largest mixing weights). The probability of observing the current background pixel value $p(X|BG)$ should be:

$$p(X|BG) = \sum_{m=1}^B \omega_m \cdot \eta(x; \mu_m, \sigma_m^2) \tag{4}$$

- When a new frame comes (at time t), we did the following calculation for all the pixels in this frame. We calculated the distance between the pixel value x and the mean value of all Gaussian distributions in the mixed Gaussian model. If one of the distances was less than 2.5 times its standard deviation, we could say the pixel value x was defined as a background pixel belonging to the i -th ($1 \leq i \leq m$) Gaussian model. By doing this, we could figure out the background (static parts) and foreground (moving objects) in the frame. After calculation, the frame pixels could be divided into foreground and background pixels according to the following relation:

$$f_c(x, y) = \begin{cases} B(x, y), & \text{if } |f_c(x, y) - \mu_{m,t}| \leq 2.5\sigma_{m,t} \\ F(x, y), & \text{otherwise} \end{cases} \tag{5}$$

Here, $f_c(x, y)$, $B(x, y)$ and $F(x, y)$ are the current frame, background frame parts (static parts), and the current foreground frame parts (moving objects).

- After step 3, we built a temporary background image $BG(x, y)$ based on the background B_{n-1} corresponding to the previous frame. The pixels in B_{n-1} which locate in the area of $B(x, y)$ in function 5 were updated by using the pixel value which was in the same location $f_c(x, y)$. The other pixels in the area of $F(x, y)$ in function 5 remained the same.
- We converted the current frame $f_c(x, y)$ and temporary background image $BG(x, y)$ at time t to gray images $f_{cg}(x, y)$ and $BG_g(x, y)$. Then we did the subtraction with the two pixel values of the corresponding position. If the distance between the pixel in $f_{cg}(x, y)$ and the corresponding pixel in $BG_g(x, y)$ was smaller than an adaptive threshold H , this meant that the current pixel could be consider as a background pixel. The gray frame could be divided into gray foreground and gray background according to the following relation:

$$f_{cg}(x, y) = \begin{cases} B_g(x, y), & \text{if } |f_{cg}(x, y) - BG_g(x, y)| \leq H \\ F_g(x, y), & \text{otherwise} \end{cases} \tag{6}$$

- After step 5, we could reconstruct the new background B_n at time t based on the background B_{n-1} corresponding to the previous frame. The new background B_n could be obtained as follows:

Based on B_{n-1} , update the pixels which locate in the area of $B_g(x, y)$ in function 6 by using the pixel values which are in the same location in the current frame $f_c(x, y)$. The other pixels which locate in the area of $F_g(x, y)$ remain the same.

- When a new frame arrives, loop steps 1 through 6 to update the background.

Based on the result of the Gaussian mixture model (GMM), we performed a subtraction with the corresponding frame, which means that we combined these two methods. To evaluate the results of our approach, we used the CDnet2014 data [13] to test it. We used the “intermittentObjectMotion” and “thermal video” data files available at the site. We compared our method with the Gaussian Mixture Model (GMM) [12] and the Interframe Difference Method [11]. The data set also has an official test code that specifically tests the effectiveness of different methods. However, the test code only tests the moving parts to evaluate the results of background reconstruction. Now, instead of eliminating the moving parts of the image, as done in Figure 4, we focused on the moving parts. The results regarding the identification of the moving parts (the foreground) are shown in Figure 5.

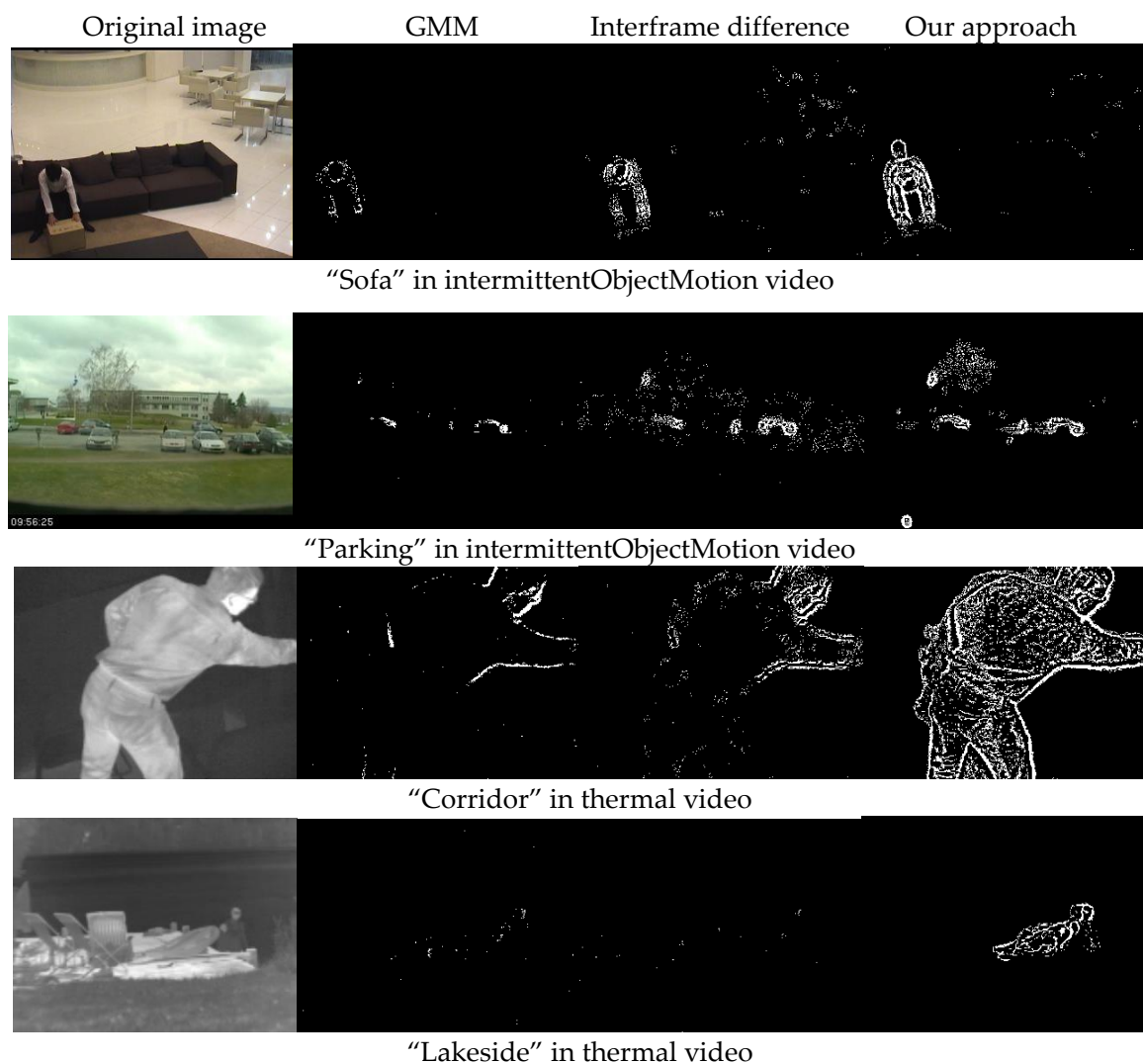


Figure 5. Comparison of benchmark videos between our approach and other approaches. From left to right for each scene, we present the original image, and the results for the foreground (the moving part) obtained with the Gaussian Mixture Model (GMM), the Interframe Difference Model and our new approach.

We note that a better image comprehension was obtained with our new method. However, we also applied more objective general foreground detection evaluation indices to compare the performance of our approach with two other approaches. The indices *Recall*, *Precision*, *PWC*, and *F-measure* were chosen, with the following definitions:

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$PWC = 100 * \frac{FN + FP}{TP + FN + FP + TN} \tag{9}$$

$$F - measure = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{10}$$

Here, *TP*, *TN*, *FP*, and *FN* denote true positive rate, true negative rate, false positive rate, and false negative rate regarding the pixel identification. After running the test code, the indices were generated automatically. Each type of sequence contained different kinds of categories, and the average indices were saved in a text file. The average results are shown in the form of a line chart in Figure 6.

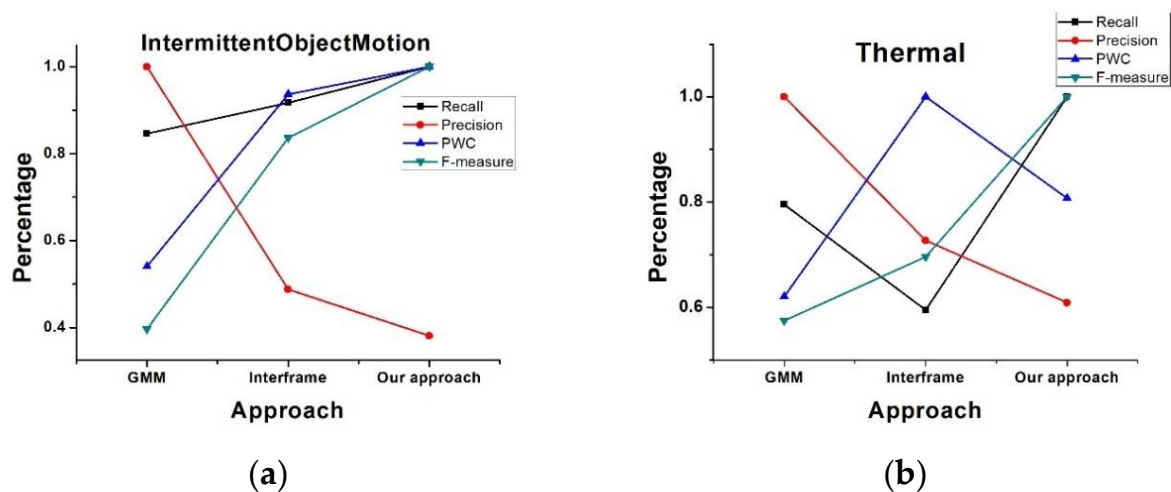


Figure 6. Comparison of evaluation indices of various approaches. (a) Performances of different approaches in intermittentObjectMotion video. (b) Performances of different approaches in thermal video.

From 6a, we can see the *Recall*, *PWC*, and *F-measure* curves were better than for the other two approaches applied to the intermittent ObjectMotion videos. Although *PWC* of our approach was not the best in the thermal videos (right), the *Recall* and *F-measure* values were still good. We can also see that *Precision* of our approach was the worst in both two types of videos, which was due to our sensitivity to light change. In a next step, we will focus on the ability to handle light changes to achieve higher *Precision* values.

4. Discussion

We note that the aspect primarily utilized in the present approach to foreground scattering elimination was that the object of interest does not move, or at least does not move in concord with the imposed modulation. We employed modulation (by, e.g., moving the eyes/head periodically slightly up and down) at a rate allowing the object of interest to be quasi static over a period when the foreground blurring pattern moved periodically. We considered the parallel to normal lock-in (frequency- and phase-tagged) detection. In this case, the modulation must be fast enough to move away from the strongest influence of the 1/f noise in the non-interesting, overwhelming signal contributions that

we wanted to discriminate against. Here, an increased spatial modulation frequency allowed us instead to mark the background clearly so that it could be discerned against the desired, weaker signal. We achieved the improved vision by using the image processing capability of the human brain, as shown in Figure 2, or by computerized image processing, as illustrated in Figure 4.

Instead of moving the detector, the foreground scattering object may instead be moved. Again, the physiological image processing of the brain could be used similarly, as illustrated in Figure 2. Clearly, here, digital image processing of the recordings of a fixed camera could also be used to eliminate the disturbing, partially obstructing close range perturbation, as illustrated in Figure 4. Incidentally, we note that a rotating round window is sometimes installed on the bridge of ships—but now for the purpose of screen wiping without moving the wiper, but rather the window. We note that even if there were no wipers, the present method would allow better vision of distant objects.

We introduced a new method, combining the established Interframe Difference and Gaussian Mixture Models to digitally separate moving and static parts of a video sequence, and demonstrated certain advantages, while there is still room for improvements, by appropriate handling of intensity variations.

We noticed that in the physiological imaging process, the brain over some time “remembers” the parts of the scene, which were non-obscured at a certain time, and is able to add to that later observed parts, which were subsequently uncovered, and finally, to synthesize a full image of the object, at any time partially obscured. This relates to the Gestalt direction of vision psychology [24,25]. This field, with origins back over 100 years [26], continues to attract much attention (see, e.g., [27,28], and analyzes the fascinating capability of human vision to ensure that “the perceptual whole is more than the sum of the parts”.

Customary imaging techniques used in video conferencing could apply, where only the moving parts of the frame (e.g., the mouth of a speaking person) are updated at video rates, while the “static” features, such as the conference room, remain at a relatively slow updating rate, thus massively saving transmission bandwidth. In the cases that we discuss, we instead eliminated the moving parts and focused on the “semi-static parts”, which constituted our object to be visualized. Again, the analog to the lock-in concept comes to mind; however, now ignoring the modulated part and focusing on the “static background”—making further meaning to the suggested acronym “inverse lock-in-like spatial modulation, ILLSPAM”. We note that the concept of modulating an accessible background instead of the feature of interest, when inaccessible for modulation, may have important applications also outside the field of vision.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2411-5150/4/3/37/s1>, Video S1: Figure 2 animation; Video S2: Figure 4 animation.

Author Contributions: All authors have read and agreed to the published version of the manuscript. Conceptualization, S.S.; Data curation, Y.L.; Funding acquisition, S.S.; Investigation, S.S.; Software, Y.L.; Supervision, S.S.; Visualization, Y.L.; Writing—original draft, S.S.

Funding: This work was partially funded by the SCNU Study Abroad Program for Elite Postgraduate Students, the Science and Technology Program of Guangzhou (No. 2019050001), and the Guangdong Provincial Key Laboratory of Optical Information Materials and Technology (No. 2017B030301007).

Acknowledgments: We gratefully acknowledge the support of Sailing He, Guofu Zhou, and Karl Åström. We are also very grateful to an anonymous reviewer, who brought to our attention the relation of our work to Gestalt psychology.

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

References

1. Watson, A.B. (Ed.) *Digital Images and Human Vision*; Bradford Books, MIT Press: Cambridge, MA, USA, 1993.
2. Li, Z.P. *Understanding Vision, Theory, Models and Data*; Oxford University Press: Oxford, UK, 2014.
3. Svanberg, S. Color image registration in low-light-level conditions. Swedish Patent SE 468 414, 1993.

4. Svanberg, S.; Johansson, J. Image registering in color at low light intensity. US Patent 5 483 379. European Patent 0584163, 1996.
5. Farmer, W.M. *The Atmospheric Filter*; JCD Publishing: Oviedo, FL, USA, 2001; Volumes 1 and 2.
6. Vollmer, M.; Möllmann, K.-P. *Infrared Thermal Imaging—Fundamentals, Research and Applications*; Wiley-VCH: Weinheim, Germany, 2010.
7. Szeliski, R. *Computer Vision-Algorithms and Applications*; Springer: Berlin/Heidelberg, Germany, 2010; p. 531. [[CrossRef](#)]
8. Toyama, K.; Krumm, J.; Brumitt, B.; Meyers, B. Wallflower: Principles and practice of background maintenance. In Proceedings of the Seventh International Conference on Computer Vision (ICCV'99), Kerkyra, Greece, 20–27 September 1999; pp. 255–261.
9. Hong, S.-H.; Jang, J.-S.; Javidi, B. Three-dimensional volumetric object reconstruction using computational integral imaging. *Opt. Express* **2004**, *12*, 483–491. [[CrossRef](#)] [[PubMed](#)]
10. Xiao, X.; Javidi, B.; Martinez-Corral, M.; Stern, A. Advances in three-dimensional integral imaging: Sensing, display, and applications. *Appl. Opt.* **2013**, *52*, 546–560. [[CrossRef](#)] [[PubMed](#)]
11. Lipton, A.J.; Fujiyoshi, H.; Patil, R.S. Moving target classification and tracking from real-time video. In Proceedings of the Fourth IEEE Workshop on Applications of Computer Vision, Princeton, NJ, USA, 19–21 October 1998; pp. 8–14.
12. Stauffer, C.; Grimson, W.E.L. Adaptive background mixture models for real-time tracking. In Proceedings of the Computer Vision and Pattern Recognition Conference, Fort Collins, CO, USA, 23–25 June 1999; pp. 246–252.
13. Goyette, N.; Jodoin, P.; Porikli, F.; Konrad, J.; Ishwar, P. Changedetection.net: A new change detection benchmark dataset. In Proceedings of the Computer Vision and Pattern Recognition Conference, Providence, RI, USA, 16–21 June, 2012; pp. 1–8.
14. O'Toole, M.; Lindell, D.B.; Wetzstein, G. Confocal non-line-of-sight imaging based on light-cone transformation. *Nature* **2018**, *555*, 338–341. [[CrossRef](#)] [[PubMed](#)]
15. Vellekoop, I.M.; Mosk, A.P. Phase control algorithms for focusing light through turbid media. *Opt. Commun.* **2008**, *281*, 3071–3080. [[CrossRef](#)]
16. Vellekoop, I.M.; Lagendijk, A.; Mosk, A.P. Exploiting disorder for perfect focusing. *Nat. Photonics* **2010**, *4*, 320–322. [[CrossRef](#)]
17. Vellekoop, I.M. Feed-back based wavefront sensing. *Opt. Express* **2015**, *23*, 12189. [[CrossRef](#)] [[PubMed](#)]
18. Yu, H.; Lee, J.P.; Lee, K.; Jang, J.; Lim, J.; Yang, W.; Jeong, Y.; Park, Y. In vivo deep tissue imaging using wavefront shaping optical coherence tomography. *J. Biomed. Opt.* **2016**, *21*, 101406. [[CrossRef](#)] [[PubMed](#)]
19. Kristensson, E.; Berrocal, E.; Wellander, R.; Richter, M.; Aldén, M.; Linne, M. Structured illumination for 3-D Mie imaging and 2-D attenuation measurements in optically dense sprays. *Proc. Comb. Institute* **2011**, *33*, 855–861. [[CrossRef](#)]
20. Kristensson, E.; Bood, J.; Aldén, M.; Nordström, E.; Zhu, J.J.; Huldt, S.; Bengtsson, P.E.; Nilsson, E. Berrocal, H.; Ehn, A. Stray light suppression in spectroscopy using periodic shadowing. *Opt. Express* **2014**, *22*, 7711–7721. [[CrossRef](#)] [[PubMed](#)]
21. Solomon, C.; Breckon, T. *Fundamentals of Digital Image Processing*; Wiley-Blackwell: Hoboken, NJ, USA, 2011.
22. Gonzales, R.C.; Woods, R.E. *Digital Image Processing*, 4th ed.; Pearson: London, UK, 2018.
23. Svanberg, S. *Atomic and Molecular Spectroscopy—Basic Aspects and Practical Applications*, 4th ed.; Springer: Berlin/Heidelberg, Germany, 2004; Chapters 7 and 10.
24. Wagemans, J.; Elder, J.H.; Kubovy, M.; Palmer, S.E.; Peterson, M.A.; Singh, M.; von der Heydt, R. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychol. Bull.* **2012**, *138*, 1172–1217. [[CrossRef](#)] [[PubMed](#)]
25. Wagemans, J.; Feldman, J.; Gepshtein, S.; Kimchi, R.; Pomerantz, J.R.; van der Helm, P.A.; van Leeuwen, C. A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychol. Bull.* **2012**, *138*, 1218–1252. [[CrossRef](#)] [[PubMed](#)]
26. Wertheimer, M. Experimentelle Studien über das Sehen von Bewegung [Experimental studies on the seeing of motion]. *Z. Psychol.* **1912**, *61*, 161–265.

27. Nie, Q.-Y.; Maurer, M.; Müller, H.J.; Conci, M. Inhibition drives configural superiority of illusory Gestalt: Combined behavioral and drift–diffusion model evidence. *Cognition* **2016**, *150*, 150–162. [[CrossRef](#)] [[PubMed](#)]
28. Costa, T.L.; Orsten-Hooge, K.; Rego, G.G.; Wagemans, J.; Pomerantz, J.R.; Boggio, P.S. Neural signatures of the configural superiority effect and fundamental emergent features in human vision. *Sci. Rep.* **2018**, *8*, 13954. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).