Active Multispectral Illumination and Image Fusion for Retinal Microsurgery

Raphael Sznitman, Seth Billings, Diego Rother, Daniel Mirota,
Yi Yang, Jim Handa, Peter Gehlbach, Jin U. Kang,
Gregory D. Hager, and Russell Taylor

Johns Hopkins University
3400 North Charles Street, Baltimore, MD, 21218, USA
{sznitman,sbillin3,yyang30,jthanda,pgelbach,jkang,hager,rht}@jhu.edu

Abstract. It has been shown that white light exposure during retinal microsurgeries is detrimental to patients. To address this problem, we present a novel device and image processing tool, which can be used to significantly reduce the amount of phototoxicity induced in the eye. Our device alternates between illuminating the retina using white and limited spectrum light, while a fully automated image processing algorithm produces a synthetic white light video by colorizing non-white light images. We show qualitatively and quantitatively that our system can provide reliable images using far less toxic light when compared to traditional systems. In addition, the method proposed in this paper may be adapted to other clinical applications in order to give surgeons more flexibility when visualizing areas of interest.

1 Introduction

Retinal microsurgery is one of the most demanding types of surgery. The difficulty stems from the microscopic dimensions of tissue planes and blood vessels in the eye, the delicate nature of the neurosensory retina and the poor recovery of retinal function after injury. Many micron-scale maneuvers are physically not possible for many retinal surgeons due to inability to visualize the tissue planes, tremor, or insufficient dexterity. To safely perform these maneuvers, microscopes are required to view the retina. A central issue for the surgeon is the compromise between adequate illumination of retinal structures, and the risk of iatrogenic phototoxicity either from the operating microscope or endoilluminators, which are fiber-optic light sources that are placed into the vitreous cavity to provide adequate illumination of the retina during delicate maneuvers.

Retinal phototoxicity from an operating microscope was first reported in 1983 in patients who had undergone cataract surgery with intraocular lens implantation [1]. Retinal phototoxicity is now a well recognized potential complication of any intraocular surgical procedure, and the frequency is reported to occur from 7% to 28% of patients undergoing cataract surgery [2, 3]. As a result, the International Commission on Non-Ionizing Radiation Protection (ICNIRP) now

provides safety guidelines for illumination of the fundus in both phakic and apha-kic subjects [4]. Blue wavelength and ultraviolet light induce the greatest degree of retinal injury. In fact, in [5] it was found that commercially available light sources for endoillumination exceeded the ICNIRP guidelines for retinal damage by visible light within 3 minutes, and in 9 of 10 sources, the safe exposure time was exceeded in less than 1 minute. In vitrectomy for macular hole repair, up to 7% of the patients have been reported to have experienced visually significant phototoxicity [6–8].

Phototoxicity can be either thermal or photochemical in nature from excessive ultraviolet (UV) or blue light toxicity. Ham et al. showed the action spectrum or relative risk of UV or blue light toxicity when the retina was exposed to var-ious wavelengths of light [9]. The action spectrum was then used to create a relative risk of phototoxicity associated with a given wavelength of light. The current state of the art endoillumination systems have taken advantage of these data. The new systems are significantly brighter and when compared to earlier systems, they appear to be safer. Previously, halogen light sources containing significant blue wavelength light have been replaced by xenon and mercury il-lumination systems, with reduced blue wavelengths. The risk, while reduced, is still significant for intraocular surgery. Given the advancing age of the popula-tion and increasing prevalence of retinal diseases, further improvements aimed at reducing iatrogenic retinal phototoxicity would be a welcome addition.

To address this issue, we present a novel visualization system that can be used to significantly reduce the emission of highly toxic wavelengths over existing sys-tems. While changing the spectral composition of the illumination toward longer wavelengths could help reduce phototoxicity, we have created a new device which cyclicly illuminates the retina using white and less damaging non-white light, allowing for maximal phototoxicity reduction. Consequently, images provided by this device are either fully colored or monochromatic.

To avoid visually straining any potential user (*i.e.* surgeon) this device is used together with an image recoloring scheme. Computer colorization schemes have existed since the 70's [10] and have since been further developed [11, 12]. In general, however, such systems rely on a user to pre-select regions of the image that correspond to specific colors, making them ill-suited for this application. More recently, a time series analysis was proposed to model the retinal image scene [13]. This methods however relies on having all visual cues (*e.g.* color and texture) available at all times to maintain an accurate retina model. To our knowledge, no previous work has focused on fusing images taken under varying spectrum illumination to form continuous and coherent image sequences.

Therefore, in order to facilitate the use of this device for surgical procedures, we have developed a fully automated algorithm for colorizing non-white light images, so that they appear fully colored. Our method is simple and requires little parameter tuning making it easy to use. We have experimentally shown that our method provides a quantitative and qualitative improvement in coloring accuracy over naive recoloring schemes.
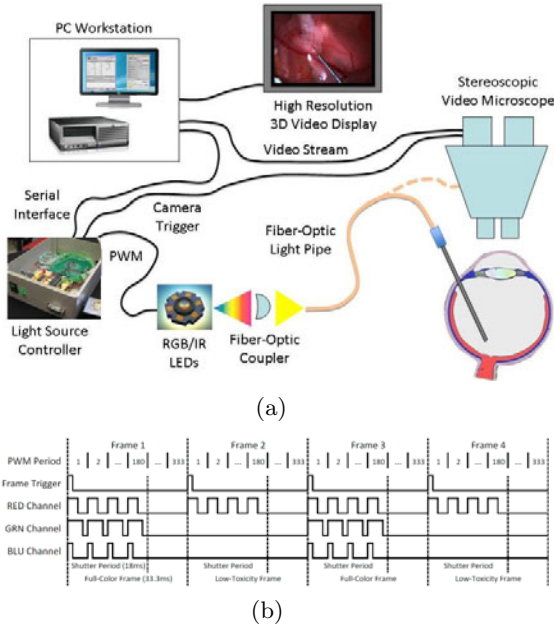
(a)

(b)

Fig. 1. (a) System overview and (b) example light source control signals

The remainder of this article is as follows: in section 2 we present the system overview and our device. In section 3 our colorization algorithm is presented. Section 4 details in depth experiments on our system, while we discuss some of the pit-falls of our system and future works in section 5.

2   Light Source

Figure 1(a) shows an overview of the system. Visualization of the retina occurs via a 3D LCD display, which accepts a video feed from high-resolution cameras mounted on a surgical microscope. Illumination is provided using high power LEDs in RGB configuration, and may be delivered either through a fiber-optic probe or through a surgical microscope. A control box, housing an embedded microcontroller and LED driver electronics, controls illumination activity via independent pulse-width modulation (PWM) of each LED channel. A serial command interface between the control box and a PC provides user-level control of the light source from a GUI application. The control box also sources a trigger signal to the cameras for synchronizing illumination activity with the shutter periods of the cameras.

The system reduces white light exposure by two means. Firstly, illumination is deactivated during the shutter closed period of each camera frame. For a typical shutter period of 18 milliseconds and a frame rate of 30 frames per second, the light exposure time is reduced by 50%. Secondly, rather than illuminating every

frame with white light, some frames are rendered under limited spectrum light having low toxic potential, such as red or infrared light. The result is a video feed interleaving full-color frames, when white light is active, with limited-spectrum (red) frames, when low toxicity light is active. Thus, the exposure period of highly toxic wavelengths is reduced by a further fraction equal to the repeat rate of white light images in the frame sequence. Figure 1(b) shows control signals from the light source controller for generating a sequence of alternating full color and monochrome images.

3   Image Recoloring

From the device described above, white and red light images are cyclicly produced at a fixed rate. Naturally, emitting fewer white light images allows for lower levels of phototoxicity for the patient. However, reducing the number of white light images increases the difficulty of the procedure for the surgeon. Hence, a method which restricts the number of white light images used, and still provides a *typical* view for the surgeon, is required. Ultimately, it is necessary to produce an accurate colored image of the retina at any given time, irrespective of which illumination was used.

To provide a coherent colored image sequence, we present two methods: a *naive* (section 3.1) and an *active scene rendering* (section 3.2) approach. Due to the lack of previous work on this particular topic, we treat the naive approach as a baseline algorithm. This algorithm is simple and only useful in cases with high fractions of white light. In section 4.1 we compare both methods on image sequences where ground truth is available, thus demonstrating improvements produced by non-naive methods.

At each discrete time step, t, we denote the type of illumination the device is triggering as L_t, where $L_t = 1$ when white light is used, and $L_t = 0$ for non-white light. Associated with each illumination, $I_t = \{R_t, G_t, B_t\}$ is the corresponding RGB image acquired. The rate at which white light illuminates the retina is then defined as $\phi = \frac{\sum_{i=1}^{t} L_i}{t}$.

In order to perform recoloring, it is necessary to correctly account for the color of the non-white illuminant. We define the color space of the acquired images as the usual RGB color space denoted by $S \subset R(3)$. Following [14], we define a separate color space $S' \subset R(3)$ such that the color of the non-white illuminant is (1,0,0). We relate S and S' by a linear transformation F of the form $F = sR$, where s is a scale factor and R is a rotation. Then for any RGB value $p \in S$, we can compute $p' \in S'$ as $p' = Fp$. The optimal F can be computed by first acquiring a non-white illuminated image, finding the largest principal component, x, and subsequently constructing two orthogonal components y and z as in [14]. R is constructed from these components. The scale s can then be computed by comparing a while light and non-white light image under the (color) rotation R.

Since our non-white illuminant is largely red, in the remainder of the paper we will continue to refer to the non-white image as the "red" image and the two

orthogonal components as green and blue with the understanding that these are, in general, not the raw color values from the camera.

We denote $\mathcal{I}_t$ as the final fully colored image rendered by our system. As the device sequentially provides us with images, we will maintain a color model for the retina, $\mathcal{M} = \{m_G, m_B\}$, where m_B and m_G are the green and blue color models (represented in the form of images), respectively. Such a color model will be maintained over time, and we thus denote $\mathcal{M}_t$ as the color model at time t. In order to have a color model at any given time, t, let I_1 be a white light image.

3.1   Naive Approach

Perhaps the simplest method to create and maintain a colored image model, $\mathcal{M}$, is to assume that images do not significantly change over time. In other words, a strong continuity in the appearance in color from I_t to $I_{t+\delta t}$ is assumed.

The corresponding algorithm is simple: if $L_t = 1$, then the model $\mathcal{M}_t$ is updated, $\mathcal{M}_t = \{G_t, B_t\}$ and $\mathcal{I}_t = I_t$. Otherwise, $L_t = 0$ and we let $\mathcal{I}_t = (R_t, m_G, m_B)$. Following such a procedures ensures that all $\mathcal{I}_t$ are fully colored images. Figures 2(a) and 2(b) show an example I_t and $\mathcal{I}_t$, respectively. Notice that continuity is violated, as the tool has been displaced since the last white-light image received, thus causing "ghosting" effects.
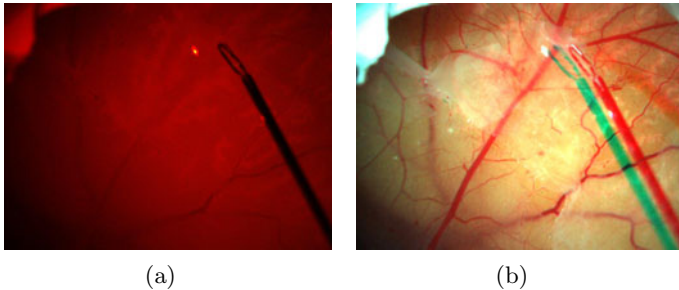

(a)   (b)

Fig. 2. (a) Example of a non-white light image from our device and (b) the image rendered by the naive algorithm. Notice that simply using the G and B channels from the last white frame does not generate good color images. This is particularly the case when there is motion in the scene.

3.2   Active Scene Rendering Approach

A natural extension of the naive approach is to infer the motion observed in the image sequence and correct the associated artifacts. We present our novel color fusing algorithm: Active Scene Rendering (ASR). Here the idea is to estimate the different forms of motion which appear in the scene and take this information into account when rendering the colored images.

Here, it is still assumed that a strong temporal correlation between adjacent images is present. Furthermore, it is stipulated that a transformation, T, from

image I_t to I_{t+1} can be inferred. Intuitively, T can be regarded as the motion, induced by the surgeon, which the eye undergoes during a procedure. Notice that this transformation only accounts for the eye motion and not the tool motion. Hence, to further reduce colorization errors (as those in figure 2(b)), we model the tool and its motion as well. The idea is to first detect the pose of the tool to obtain a 2D segmentation and then use this information to recolor the image correctly. We now describe how the estimation of the transformation T and the 2D tool segmentation are performed.

Image Stabilization. As previously mentioned, the surgeon is free to manipulate the eye. To compensate for this motion, a simple translation model for the motion of the retina is assumed. Although it has been shown that wide angle retinal image deformation is best modeled with a quadratic deformation [15], small motion can be approximated with pure translation when under high magnification. To estimate the translation we first extract SIFT features [16] (I_t is treated as a gray scale image for any value of L_t), find correspondences and then apply the robust ASKC method [17] to find the translation that best explains the correspondences. This permits us to find a transformation regardless of whether the tool is present in the image or not. Note that in order to present coherent image sequences, images are cropped by removing border regions.

Tool Detection. Given that the most consistent clue for the tool is its constant and known 3D shape, we use the framework proposed in [18] for simultaneous segmentation and pose estimation which exploits this information. This framework requires, as input, the 3D shape (represented as voxel occupancies) and color probability distribution (represented as a mixture of Gaussians) of the tool, and the color probability distribution for each background pixel (represented as a single Gaussian). The output of the framework is a segmentation of the tool in each frame, and also an estimate of the 3D pose of the tool in the 3D coordinate system of the camera, for each frame. The estimated 3D pose in one frame is used to initialize the segmentation and pose estimation in the following frame. Using this method guarantees finding the globally optimal 3D pose and segmentation in a computationally efficient manner.

The algorithm for ASR is similar to that of section 3.1. At $t = 1$, we let $\mathcal{I}_1 = I_1$ and set $\mathcal{M}_1 = \{G_1, B_1\}$. I_1 is then treated as the initial frame of reference, such that subsequent images are stabilized with regards to I_1. That is, for every new image I_t, we compute the transformation T_t from I_t to I_1. Then, using T_t, we translate I_t and compute a rectified image, $\tilde{I}_t$. When $L_t = 1$, we set $\mathcal{M}_t = \{\tilde{B}_t, \tilde{G}_t\}$ and $\mathcal{I}_t = \tilde{I}_t$.

If $L_t = 0$ (figure 3(a)), the 2D segmentation of the tool is determined (figure 3(b)). To do this, $\mathcal{M}_t$ and the known color model of the tool are used to initialize the detection process described above. Once the segmentation has been computed, this region is rendered using the tool color model. The rest of the image is rendered as $\mathcal{I}_t = (\tilde{R}_t, m_G, m_B)$ (figure 3(c)).
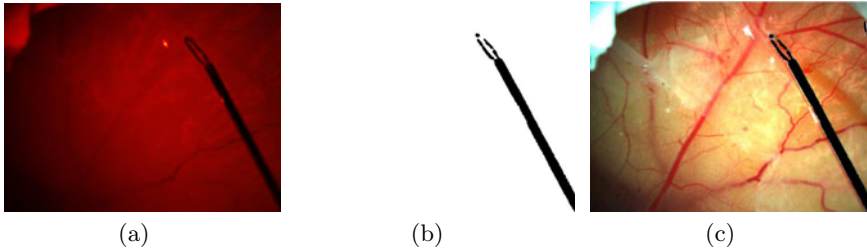
<div align="center">(a)   (b)   (c)</div>

Fig. 3. (a) Example non-white light image from our device, (b) segmented tool, and
(c) rendered image by ASR. Here, the tool is correctly estimated and image regions
are updated in a coherent manner.

4   Experimentation

We now show how our system performs on phantom image sequences. A quanti-
tative comparison of both methods is described in section 4.1, where it is shown
that ASR surpasses the naive approach in a setting where ground truth is known.
This is shown by measuring the error for different values of ϕ. Qualitative results
of our system on image sequences are then provided in section 4.2.

4.1   Validation with Ground Truth

To validate both approaches described in section 3 we recorded two image se-
quences of membrane peeling on embryonic eggs using only white light. Doing
so allows us to synthetically generate limited-spectrum images at any rate ϕ,
by using only the red band of white light images. Hence, we know that the
transformation F (see section 3) is known to be $F = I$. As detailed in [19, 20],
this phantom setup provides a similar environment to in-vivo settings. Image se-
quences consist of 500 images, acquired at 20 frames per second using the system
described in section 2.

Using this data, 5 image sequences are generated where $\phi = \{1/2, 1/4, 1/8,$
$1/16, 1/32\}$. For each sequence, both naive and ASR colorization approaches are
evaluated. Since the ground truth – the original recorded white light images
– is always available, an error can be computed for each frame generated by
either approach. In the following experiments the L2 (or mean squared error)
norm is chosen to measure the error between the ground truth and the rendered
images. In addition, we also compute the error using the Bounded Variation
(BV) norm, which has been used to quantify image quality during denoising
tasks [21]. This provides us with a measure of image quality, taking into account
both photometric and rectification errors.

Figure 4(a) shows the L2 norm error when varying ϕ for both methods. Figure
4(b) shows a similar result for the BV norm. In general, one can observe that
as ϕ decreases, the error rate increases. This is expected as the assumption
of similarity between frames, discussed in section 3.1, is increasingly violated.
Naturally, the naive approach suffers greatly when ϕ is small, as the true color
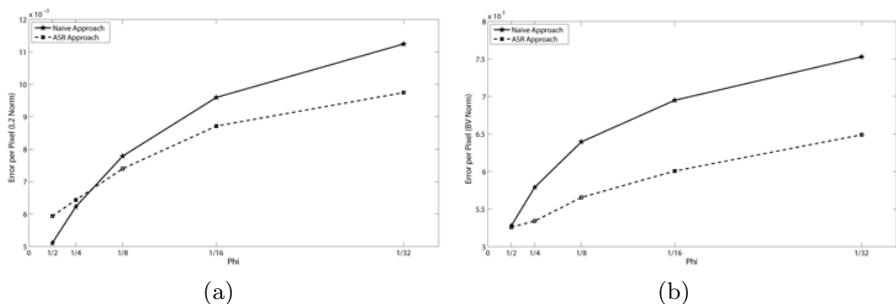
Fig. 4. Evaluation of the proposed methods when varying the fraction of available white-light images, ϕ. The average error per pixel ((a) L2 norm (b) BV norm) is computed on image sequences where ground truth is known. Both error metrics indicate that ASR (dashed line) provides accuracy gains over the naive approach (solid line).
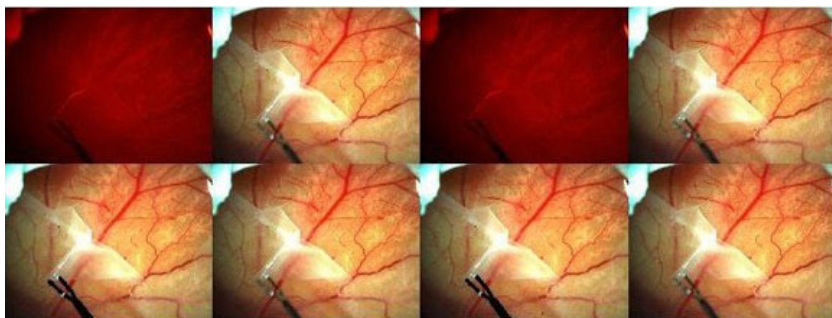

Fig. 5. (*top row*) Example image sequence from the device ($\phi = 1/2$) and corresponding rendered sequence using ASR (*bottom row*)

may differ greatly from the last observed color. ASR however suffers significantly less from small ϕ values, as it is able to maintain a more accurate color model.

4.2   Egg Peeling

Having observed that ASR provides a better way to model retinal-type scenes, we set up our system to record and display images for different values of ϕ. We record several image sequences in a similar setup as in section 4.1 and show the resulting recolored sequence. Note that the color mapping transformation F is assumed to have $R = I$, and a uniform scaling factor (determined empirically).

In figure 5 we show a typical image sequence ($\phi = 1/2$), of a chorioallatonic membrane peel from an 11 day old chicken embryo. The resulting recolored images rendered by our system are shown. The entire image sequence and recolored images can be viewed in a video included in the supplementary materials. The video shows four similar peeling sequences where each row corresponds to a different ϕ value (1/2,1/4,1/8,1/16). The first column shows the images provided by

the device, while the second and third columns show how the naive and model approach, respectively, render the sequence. Since the device is being used to obtain these image sequences, no ground truth is available for quantitative comparison.

Notice that in general the model approach renders a more coherent image than the naive approach. This is particularly true at smaller ϕ values, concurring with the results of section 4.1.

5   Discussion and Conclusion

In this paper we have presented a novel system that can be used to reduce toxic light exposure in retinal microsurgeries. The system consists of a new lighting device which reduces emission of highly toxic wavelengths. In addition we have developed a novel algorithm that can be used with this device in order to render a fully colored image sequence to the user, thus avoiding visual discomfort. We have shown qualitatively and quantitatively that our method provides superior rendering over naive approaches. Even at low ϕ rates (*e.g.* 1/8 or 1/16), we showed that maintaining high color fidelity is possible, allowing for low levels of phototoxicity.

However, most retinal surgeries involve changing the structure of the retina, and hence the color of the retina (as described in [13]). As seen in our image sequences, regions of the retina which are altered by the surgeon cannot be recolored correctly until a new white light image is provided. Hence a potential improvement of this method would involve a dynamic ϕ, which could change as a function of the activity in the image sequence.

Another natural future direction of this work would consist in understanding the effect of using such a system and phototoxicity levels observed in live tissue. Determining the exact relationship would allow for a quantitative evaluation in toxicity reduction; something which currently can not be determined.

Although phototoxicity reduction in retinal surgery provided the motivating focus for this paper, our technical approach to the problem is potentially more broadly applicable. We have developed methods for actively controlling the illumination spectrum in video microscopy and endoscopy and for fusing the resulting image sequences to form continuous and coherent image sequences. These methods are applicable in many clinical applications, including neurosurgery and cancer surgery. For example, changing the illumination spectrum can be used to improve tissue contrast or discrimination or the depth of penetration of light into tissue structures. The methods proposed in this paper may be adapted to such cases while still giving the surgeon more options on the actual visualization. Exploring these options will be another direction for future work.

Acknowledgements

References

1. McDonald, H., Irvine, A.: Light-induced maculopathy from the operating microscope in extracapsular cataract extraction and intraocular lens implantation. Ophthalmology 90, 945–951 (1983)
2. Khwarg, S., Linstone, F., Daniels, S., Isenberg, S., Hanscom, T., Geoghegan, M., Straatsma, B.: Incidence, risk factors, and morphology in operating microscope light retinopathy. American Journal of Ophthalmology 103, 255–263 (1987)
3. Byrnes, G., Antoszyk, A., Mazur, D., Kao, T., Miller, S.: Photic maculopathy after extracapsular cataract surgery a prospective study. Ophthalmology 99, 731–738 (1992)
4. International Commission on Non-Ionizing Radiation Protection: Guidelines on limits of exposure to broad-band incoherent optical radiation (0.38 to 3). Health Phys. 73, 539–554 (1997)
5. van den Biesen, R., Berenschot, T., Verdaasdonk, R., van Weelden, H., van Norren, D.: Endoillumination during vitrectomy and phototoxicity thresholds. British Journal of Ophthalmology 84, 1372–1375 (2000)
6. Poliner, L., Tornambe, P.: Retinal pigment epitheliopathy after macular hole surgery. Ophthalmology 99, 1671–1677 (1992)
7. Michels, M., Lewis, H., Abrams, G., Han, D., Mieler, W., Neitz, J.: Macular phototoxicity caused by fiberoptic endoillumination during pars plana vitrectomy. American Journal of Ophthalmol. 114, 287–292 (1992)
8. Banker, A., Freeman, W., Kim, J., Munguia, D., Azen, S.: Vision-threatening complications of surgery for full-thickness macular holes. Ophthalmology 104, 1442–1453 (1997)
9. Ham, W.J., Mueller, H., Ruffolo, J.J., Guerry, D., Guerry, R.: Action spectrum for retinal injury from near-ultraviolet radiation in the aphakic monkey. American Journal of Ophthalmology 93, 299–306 (1982)
10. Museum of Broadcast Communication: Encyclopedia of Television (online), `http://www.museum.tv/archives/etv/c/htmlc/colorization/colorization.htm`
11. Skora, D., Burinek, J., Zra, J.: Unsupervised colorization of black and white cartoons. In: Int. Symp. NPAR, Annecy, pp. 121–127 (2004)
12. Yatziv, L., Sapiro, G.: Fast image and video colorization using chrominance blending. IEEE Transactions on Image Processing 15, 1120–1129 (2006)
13. Sznitman, R., Lin, H., Manaswi, G., Hager, G.: Active background modeling: Actors on a stage. In: International Conference on Computer Vision, Workshop on Visual Surveillance (2009)
14. Mallick, S., Zickler, T., Belhumeur, P., Kriegman, D.: Specularity removal in images and videos: A PDE approach. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 550–563. Springer, Heidelberg (2006)
15. Stewart, C., Chia-Ling, T., Roysam, B.: The dual-bootstrap iterative closest point algorithm with application to retinal image registration. IEEE Transactions on Medical Imaging 22(11), 1379–1394 (2003)
16. Lowe, D.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 20, 91–110 (2003)
17. Wang, H., Mirota, D., Hager, G.: A generalized kernel consensus based robust estimator. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(1), 178–184 (2010)

18. Rother, D., Sapiro, G.: Seeing 3D objects in a single 2D image. In: International Conference on Computer Vision (2009)
19. Leng, T., Miller, J., Bilbao, K., Palanker, D., Huie, P., Blumenkranz, M.: The chick chorioallantoic membrane as a model tissue for surgical retinal research and simulation. Retina 24(3), 427–434 (2004)
20. Fleming, I., Balicki, M., Koo, J., Iordachita, I., Mitchell, B., Handa, J., Hager, G., Taylor, R.: Cooperative robot assistant for retinal microsurgery. In: International Conference on Medical Image Computing and Computer Assisted Intervention, vol. 11(2), pp. 543–550 (2008)
21. Chang, Q., Chern, I.: Acceleration methods for total variation-based image denoising. SIAM Journal of Applied Mathematics 25(3), 982–994 (2003)