Intra-Operative Medical Fluorescence Imaging Improvement

Rob Jones
Department of Electrical Engineering
Stanford University
Stanford, CA USA
rjones10@stanford.edu

***Abstract*—An approach to mitigating noise, the effects of vignetting and illumination non-uniformity, reflections, and distance in surgical fluorescence imaging is proposed and explored experimentally.**

I. Introduction

Dyes, such as Indocyanine Green (ICG), are used in several surgical applications, including highlighting regions of interest in anatomy when visible light is not ideal for doing so, visualizing blood flow, and locating pathogens. Visualization is typically accomplished by injecting the dye into the bloodstream or region of interest, illuminating the scene with a laser tuned to the absorption wavelength of the dye, and capturing the scene with a camera system that filters out the laser's wavelength but is sensitive to the dye's emission wavelength. The resulting fluorescence image is typically overlaid on a visible light image of the scene.

Fluorescence images are more difficult to use than typical visible spectrum images for a few reasons. Chief among these reasons is the physical difference between the two modes of imaging. Images in the visible spectrum consist almost entirely of surface reflections of the illuminant, whereas in fluorescence imaging in endoscopic surgical applications, the illuminant must travel through at least one layer of tissue and be absorbed by the dye that has concentrated in the blood or bile, which then emits photons at lower energy levels, which must travel back through the tissue before finally being captured by the camera. The deleterious effect of all of these stages of energy transfer, of course, is considerably magnified as distance from the subject increases due to the exponential relationship between incident power and distance. Consequently, fluorescent signals are even more sensitive to degradation in the optical path and irregularities in the pattern of the illuminant than are visible signals. Additionally, in the endoscopic surgical application, the subject being imaged is typically dominated by moist tissue, which means all forms of incident light can easily produce specular reflections. Since these reflections can slightly alter the wavelength of the reflected light, any reflection of the dye excitation light can actually cause it to move outside of the effective range of the imaging system's excitation light filter, making it no longer distinguishable from the dye's emission wavelength. Again, this not really a problem for the visible light portion of the image. These problems, and potentially others, contribute to the quality of surgical fluorescence images being very dependent on circumstances, which reduces their overall diagnostic and reference utility.

The goal of this project is to process a set of surgical fluorescence videos to make the fluorescence more uniform across the scene, more robust to the effects of distance, and sharper in its definition of anatomy.

II. Previous work

A. Noise

Regarding noise from the system, Qu, Zhang, and Jia suggest an efficient method for denoising images affected by salt and pepper noise, combining the benefits of adaptive filtering and weighted filtering [1]. The noise in the images in the current project are not necessarily subject to salt and pepper noise, but they do have a good deal of noise due partially to the low lighting conditions, so this approach still seems beneficial as a basis for denoising. Removing noise from the grayscale image will be important because it allows subsequent processing steps to produce more reliable results.

B. Vignetting and illumination non-uniformity

Zheng et al discuss a method of correcting for vignetting using radial and semi-circular tangential gradient distributions [2]. They assume based on observation of many natural images that positive and negative gradient distributions tend to be symmetric, so when they are calculated based on an estimated optical center and the results are not symmetric, a new estimated center can be found by minimizing the asymmetry, with a similar approach leading to an approximation of the vignetting in the optical system. In [3], Zheng works with a different group on a related premise – that intensity gradients in the image should be smooth except near the edges of objects. Again the approach is to apply estimated corrections to the illumination bias in different parts of the image until the image gradient properties match the assumption. The bias can then be used to correct for lighting non-uniformity. Both of these approaches are promising for the current application.

C. Specular reflections

Shafer proposes a dichromatic projection of color images [4] into a parallelogram in a plane defined by an achromatic

vector and a colored vector. The achromatic vector is essentially a grayscale representation, while the colored vector codes information about the body reflection color of the image subject. Using such an analysis in the current application can aid in finding reflections which may contribute to false detection of fluorescence, which would allow a processing algorithm to remove the potentially false fluorescence information from the image.

III. Algorithm Development

The images and videos used for this project were obtained at the author's place of work during the normal course of research and development activities. All images are from animal labs and therefore contain no patient information whatsoever. These images consist of a green color, indicating the presence of fluorescence (which usually occurs outside the visible spectrum), overlaid on a grayscale background, showing the scene as illuminated by visible light. Both components of the image are captured simultaneously by the same endoscopic camera system. The rigid endoscope uses a fiber optic transmission channel to deliver illumination, both in the visible range and at the excitation wavelength of the fluorescent dye, from an external light source to the inside of the subject. Reflected light is then transported from inside the subject through a series of lenses housed in the rigid endoscope and out to the camera's sensors. The resulting combined image is sampled regularly to construct a video signal, which is displayed on screens in the operating room to allow the surgeon to visualize the anatomy without performing open surgery on the subject. The video signal is also captured and stored digitally by an archival device.

The works mentioned above combined with working knowledge of the imaging systems used in endoscopic surgical fluorescence applications were implemented in various forms to process videos of surgical fluorescence in an attempt to improve their clinical usefulness. Fig. 1 shows an example of an image that could benefit from processing, and this same image will be used throughout this paper to demonstrate the effects of each processing step. The amount of noise is quite obvious, and it makes distinguishing features in the image difficult. The areas near the top of the screen show a noticeable decay in fluorescence intensity. Specular reflections can be seen in quite a few places in the image, and many of them exhibit a distinct green tinge, indicating likely false detection of fluorescence. Finally, the fluorescence signal itself, even where it is contiguous and distinct, lacks obvious edges and gradually bleeds into its surroundings. All of these characteristics are undesirable, and the goal of this project is to correct problems like these.

Ideally, the fluorescent and grayscale images could be processed independently at the time of acquisition. In this project, however, post-processing of the images was the chosen approach because real-time processing of each component would only be achievable through implementation of the image processing algorithms in an FPGA design within the camera system. The scope of such an effort far exceeds the amount of time allotted for this project, so only the combined image from the camera output, via the archival device, is available. The videos are MPEG4-encoded, which means they lack some of the original information and can include encoding artifacts, but the still images captured by the archival system are bitmaps that do not suffer the same problems. These bitmaps were therefore used more often during algorithm development than individual frames extracted from the videos. Once the algorithm was in place, it was used to post-process groups of video frames.

A. Circular masking and splitting

The videos in this application are recorded through a circular scope attached to a rectangular image sensor. As a result, part of the active image is actually just a black circular border. Some light does make its way into this region and some noise from the sensor occurs, so there is a bit of signal present. But there should not be, and algorithms like bias correction would seem to function better when the area that should be black is truly black. Therefore, circle detection was implemented in order to create a mask that could be used to enforce this border. Fig. 2 shows a detected circle imposed on the image, and Fig. 3 shows a mask developed from that circle.

The images also must be split into grayscale and fluorescence components for proper processing. This is accomplished by simply averaging the non-green channels of the image to find a grayscale representation, then subtracting the grayscale image from the green channel to leave only the fluorescence. The images resulting from applying this process to Fig. 1 are shown in Fig. 4 and Fig. 5. It should be noted that the averaging and subtracting process means that there is almost no apparent noise in the fluorescence image; denoising therefore will only be necessary for the grayscale image.

B. Denoising

Denoising was initially attempted by adapting the method proposed in [1]. One of the central features of this method is the so-called "ROAD" operator, which calculates absolute differences between all pixels in the window, at each pixel position, and aggregates a number of the smallest differences. This is a metric of how similar the pixel in question is to the rest of the frame. This metric is ultimately used to decide whether the pixel is really noise or just normal image data that should be left as-is. The window size for these "ROAD" calculations is also based on an iterative determination of the number of similar neighboring pixels – a mix of noise and non-noise data is required for the algorithm to function correctly. Before windowing, though, this method employs a guess about whether the pixel is even noise to begin with; in theory, this allows the algorithm to be more computationally efficient than other non-discriminating denoising approaches.

Due to the computational load of iterating over so many data points and their neighbors, especially in 1920x1080 resolution images (as is the case in this project), this method turned out to be quite slow, which made attempts to characterize and improve the algorithm prohibitively long, and no proper conclusion on the appropriate parameters was reached. The attempted implementation is therefore included in the code base, but it was not used in the final processing algorithm for this project. Instead, simple median filtering was

used, with an 11x11 kernel deemed appropriate for single-image processing and a 7x7 kernel chosen for many-frame video processing to save on computation time while sacrificing results to a certain extent. Fig. 6 shows the grayscale image after median filtering with the 11x11 kernel, where a drastic reduction in the noise level is apparent but the sharp features of the image do not appear to be greatly diminished.

C. Non-uniformity correction

The next step was to address the compounding issues of vignetting (light falloff radially from the optical center of the image due to the shape of the lenses used to acquire the image) and lighting non-uniformity (the differing intensity of incident light across the image due to the directionality and shape of the illuminant) in the fluorescence image. The visible image suffers from these effects, but as mentioned earlier, the fluorescence is degraded more noticeably since its overall intensity is lower. However, since the fluorescence image often consists of zero data in large regions, characterization of non-uniformities and degradations was carried out on the grayscale image first to attain knowledge of the system's behavior which could then be applied to correct the fluorescence image.

The vignetting correction algorithm produced promising results in some cases, but it seemed to struggle when image intensity was noticeably asymmetrically distributed. This makes sense since the operating principle of the algorithm is symmetric equalization of the positive and negative gradient distributions. Unfortunately, it must be assumed that the image intensity for most surgical video of the type being processed here will be decidedly asymmetric given the irregular shape of the anatomy being visualized and the relative differences in depth between one part of the image and another in such a confined space. This means the vignetting correction algorithm will see larger intensity gradients in some areas near the center of the image than would be expected in, for example, an outdoor image with natural lighting. As a result, it will incorrectly map the light falloff in the image and will weight the output improperly.

The bias correction algorithm, on the other hand, seemed to do a very good job of making the intensity across the image more even, bringing out details near the darker edges of the scene while dampening the brightest portions. Fig. 7 and Fig. 8 compare the effects of the two algorithms on the test image. It should be fairly obvious that the bias correction is positioned to much more effectively accomplish the stated goal of leveling the fluorescence intensity across the whole image to account for lighting non-uniformity and light falloff. Fig. 9 shows the fluorescence image resulting image after bias correction. Note especially the area near the top edge of image where the fluorescence of the central duct becomes much more obvious and the area 45 degrees clockwise is significantly amplified as well.

D. Specular Reflections

Shafer's method for finding specular reflections depends on the use of three channels of color data. In the images used for this project, there are miniscule differences between the grayscale data in each of the three color channels. But these differences are small enough that, as discussed above, a basic assumption is that the average of the red and blue channel data is sufficient to approximate the grayscale component of the green channel data. This implies that Shafer's method itself will not be perfectly applicable in this scenario. Instead, Shafer's insights provided a path forward for a different approach. The key finding was that specular reflections tend to be those parts of the image closest to the achromatic – grayscale – color axis.

Since specular reflections really are achromatic, the only way to differentiate them from other achromatic data is by comparing intensity. But an intensity threshold alone quickly proved not to be sufficient to separate reflections from useful data. Instead, it was also observed that reflections have defined edges. In light of this, an image gradient was calculated using the Sobel operator, and then both the gradient image and the intensity image were transformed into binary images by choosing thresholds based on their histograms. (This choice was intended to make the algorithm resilient to changes in overall image intensity and scene variation.) The binary images were then combined with an AND operation to locate the most defined edges of the brightest regions in the image – a good description of the borders of specular reflections. Finally, morphological image processing was applied to this combined binary image to close these regions and attempt to replicate the reflections from the original image. Fig. 10 shows the result for the test image, which corresponds very well with the original image data. The structural elements chosen for the closing and erosion operations were optimized for smaller reflection regions under the assumption that camera shuttering functionality would prevent very large regions of reflections from occurring.

E. Correction of reflections and background fluorescence

Once the reflection locations were obtained, the goal was to reduce their effect on misperception of fluorescence. A gamma-correction type of approach was deemed appropriate for this. For the grayscale image, the reflections are distracting since they are so bright. For the fluorescence image, for reasons explained earlier, it is likely that the excitation source also reflects in the same regions, which could result in false detection of fluorescence. The assumption is therefore that both the grayscale and fluorescence signals should be de-emphasized in these regions, even if it makes one or both of them a bit discontinuous. Using the binary mask of the reflection regions, pixel values for both images in these regions are subjected to exponentiation, reducing their intensity since their pixel values are in the interval [0, 1].

A similar approach was taken to minimize the bleed of fluorescence away from the boundaries of the true fluorescence in the image. As with the reflections, a histogram-and-threshold approach was used to isolate the brightest fluorescence values (which had to occur after reflection correction, which should illuminate the choice behind the ordering of these operations). As before, taking this approach means the algorithm is somewhat generalized and therefore robust to changes in illumination intensity and the scene. This is especially important for the fluorescence image since, as was mentioned previously, intensity depends greatly on distance

from the subject, so a goal of this project is to make that variation less obvious. Below the threshold, intensity values are weighted by their ratio with the threshold before exponentiation, resulting in a more drastic falloff characteristic than in the de-emphasis of the reflection correction. This is helpful in suppressing low-level fluorescence signals that gradually appear as the endoscope gets closer to tissue in the scene.

F. Reconstruction

After the above corrections are made, one task remains. The grayscale and fluorescence images are recombined into a single image by reversing the process that was used to split them from the original image. The combination of the reflection corrections and suppression of background fluorescence with all the previous steps, combined back into a single image, can be seen in Fig. 11. The image has seen significant noise reduction; reflections have been dampened and the green tinge has been removed from them; the intensity of fluorescence in the correct anatomy is consistent across the image, and it rapidly decays everywhere else; in short, the resulting image is more clinically desirable and therefore indicates that the algorithm is a good candidate for experimentation.

IV. Experimental results and Analysis

The algorithm was fit into a processing loop that would read in many frames from a video, process each frame, and then assemble a resulting processed video. To save on computing time, some expensive steps were only carried out on the first frame – notably, circle fitting and bias correction. This introduces inherent disadvantages, especially concerning bias correction, since a change in endoscope position relative to the scene can substantially affect the distribution of light. It is possible for the endoscope to move in relation to the camera and therefore invalidate the circular mask approximation, but this is far more unlikely. The results from processing the test videos, however, indicate that in most cases the assumption of constant bias correction was not too damaging.

The final algorithm was run for a set of test videos, four of which are uploaded to the course site for reference. They are labeled A, B, C, and D. The resulting videos had "_proc" appended to their filenames and are also uploaded.

Video A is a clip that shows a good deal of ambient fluorescence. It also features reflections with green components on both the surgical instrument and various parts of the anatomy. Variations in intensity correlated to changing distance from the subject and changing viewing angle are present as well. The processed version of Video A shows improvement in all of these aspects, and it does a much better job of highlighting only the tissue that really is fluorescing except for a couple of moments when the scope and the anatomy are moving. Artifacts from the bias correction are visible in the 6 o'clock to 9 o'clock regions of the image, however, especially towards the end of the video. The fluorescence intensity in this area is boosted enough that the ambient suppression portion of the algorithm no longer removes it.

Video B shows an extremely challenging situation in which the entire image is flooded with apparent fluorescence, largely because the endoscope is so close to the anatomy. The defined step changes in intensity visible in the video are the result of changes in the light source power and camera system gain levels in attempts by the user to manually improve the image. The approach is clearly only partially successful until the scope is pulled back from the scene, and then uncertainty remains about future visualization since the system's visualization settings have changed. The processed version of the video shows the algorithm is not able to completely remedy the situation, but it does de-emphasize fluorescence on the liver wall and create a defined border around the fluorescence of the ducts. It also maintains much more constant visualization qualities during the application of new system settings, which equivalently means that it is somewhat robust to distance from the subject – another major goal of the algorithm.

Video C shows a scene that should have a relatively constant medium level of fluorescence across the middle third of the image (the liver), and quite a few reflections are present as well. The processed version of the video handles these reflections generally as desired, and suppressed ambient fluorescence near the top and bottom of the image, but it mitigates much of the fluorescence that should still be present. A very noticeable combination of the MPEG encoding and the suppression algorithm's hard thresholding appears here as well, demonstrating some flaws in this approach.

Video D is similar to Video C, and the results after processing are also similar. The artifacts continue to be very noticeable here, but once again the algorithm does accomplish good recognition and suppression of both reflections and background fluorescence. These benefits come at the cost of removing useful information.

V. Conclusion

The algorithm developed for this project is a promising first step towards intelligent processing of surgical fluorescence video. Many parameters can and should be tweaked to improve performance, and the transition from threshold levels to exponential corrections should be smoothed. Additionally, the success of current results is only analyzable by surgeons or medical device professionals familiar with the procedures and technologies involved. Even though that analysis by experts is similar in nature to evaluating the success of an optical character recognition system, it would be worthwhile to try to pursue a numerical metric of success. If found, such a metric could be used to automate future algorithm development and improve ability to iterate and fine-tune the approach. This work will be carried forward further by the author.

Acknowledgment

Great thanks are due to Professor Gordon Wetzstein, my advisor for this project, for his careful reading of my verbose emails and vague ideas, and for helping direct this work remotely. Thanks are also due to the rest of the EE368 teaching staff for their patience and helpfulness throughout the course.

REFERENCES

[1] Xiujie Qu, Fu Zhang, and Huan Jia, "An Efficient Adaptive Denoising Algorithm for Remote Sensing Images," Mathematical Problems in Engineering, vol. 2013, Article ID 207461, 5 pages, 2013. doi:10.1155/2013/207461

[2] Yuanjie Zheng, Stephen Lin, Sing Bing Kang, Rui Xiao, James C. Gee, Chandra Kambhamettu. Single-Image Vignetting Correction from Gradient Distribution Symmetries. IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 35 Issue 6.

[3] Zheng Y, Grossman M, Awate SP, Gee JC. Automatic Correction of Intensity Nonuniformity from Sparseness of Gradient Distribution in Medical Images. Medical image computing and computer-assisted intervention : MICCAI . International Conference on Medical Image Computing and Computer-Assisted Intervention. 2009;12(0 2):852-859.

[4] Shafer, S. A. (1985), Using color to separate reflection components. Color Res. Appl., 10: 210–218. doi:10.1002/col.5080100409

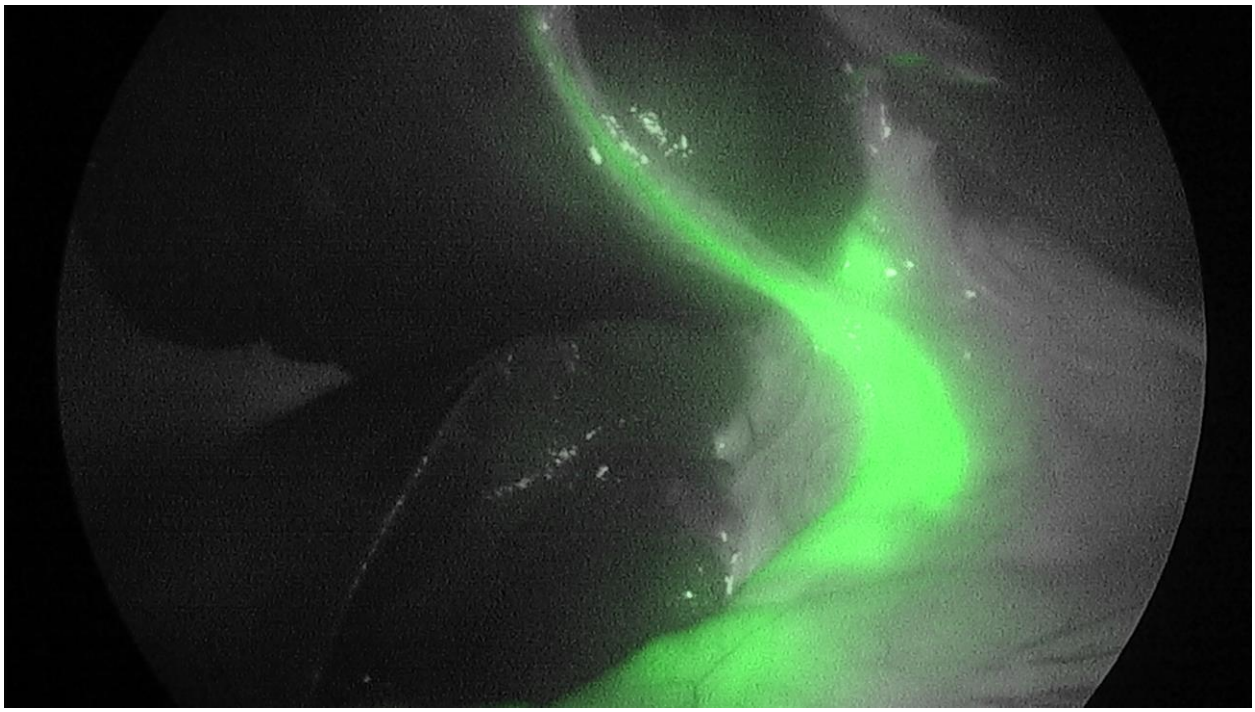

Fig. 1. Bitmap of surgical fluorescence image

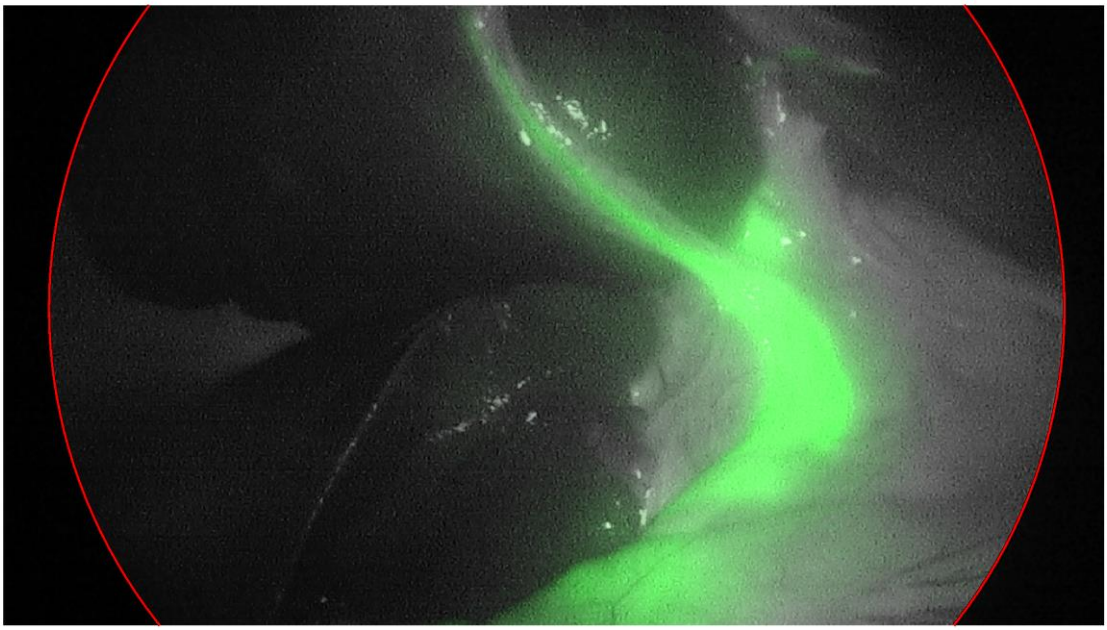

Fig. 2. Detected circle imposed on image

Fig. 3. Circular scope mask

Fig. 4. Grayscale image

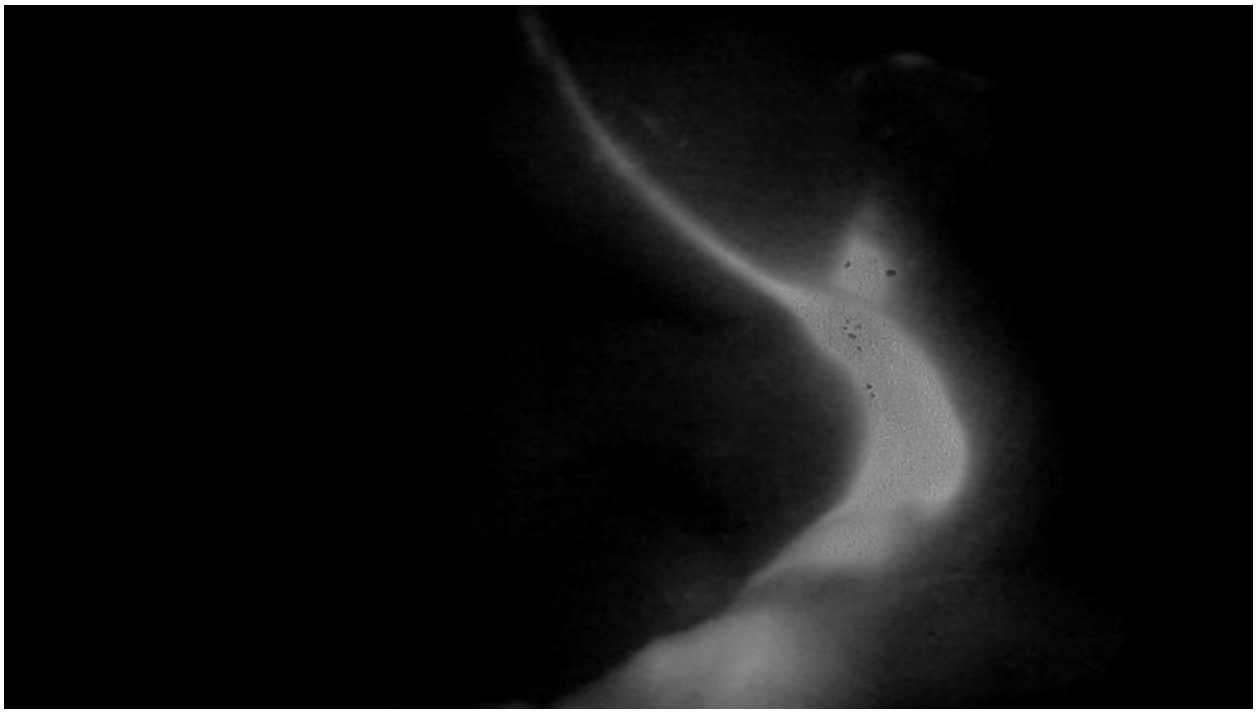

Fig. 5. Fluorescence image

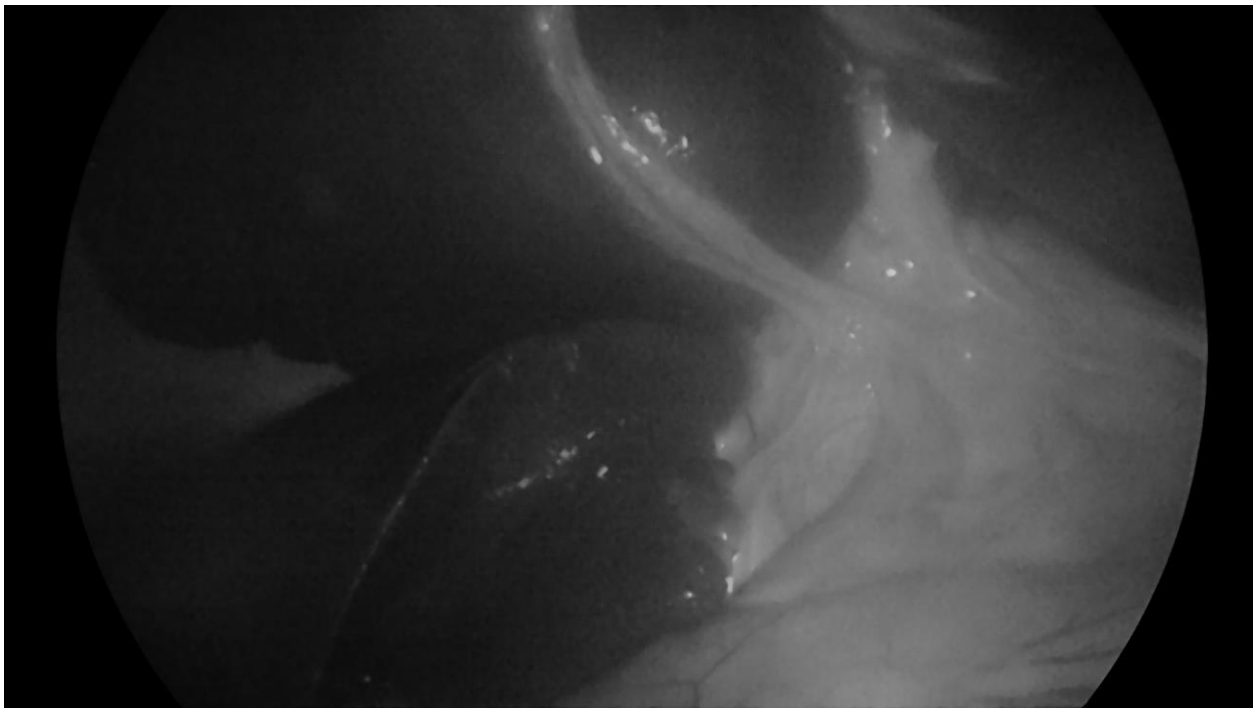

Fig. 6. Denoised grayscale image using an 11x11 median filter

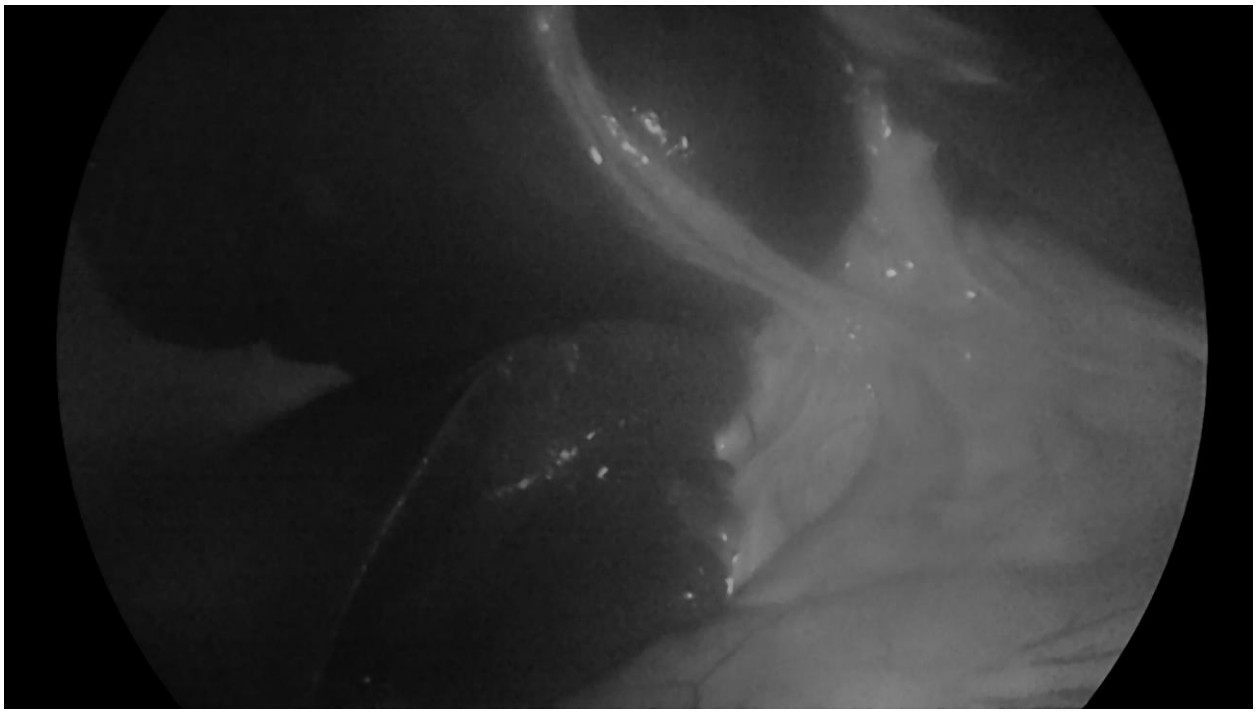

Fig. 7. Vignetting correction

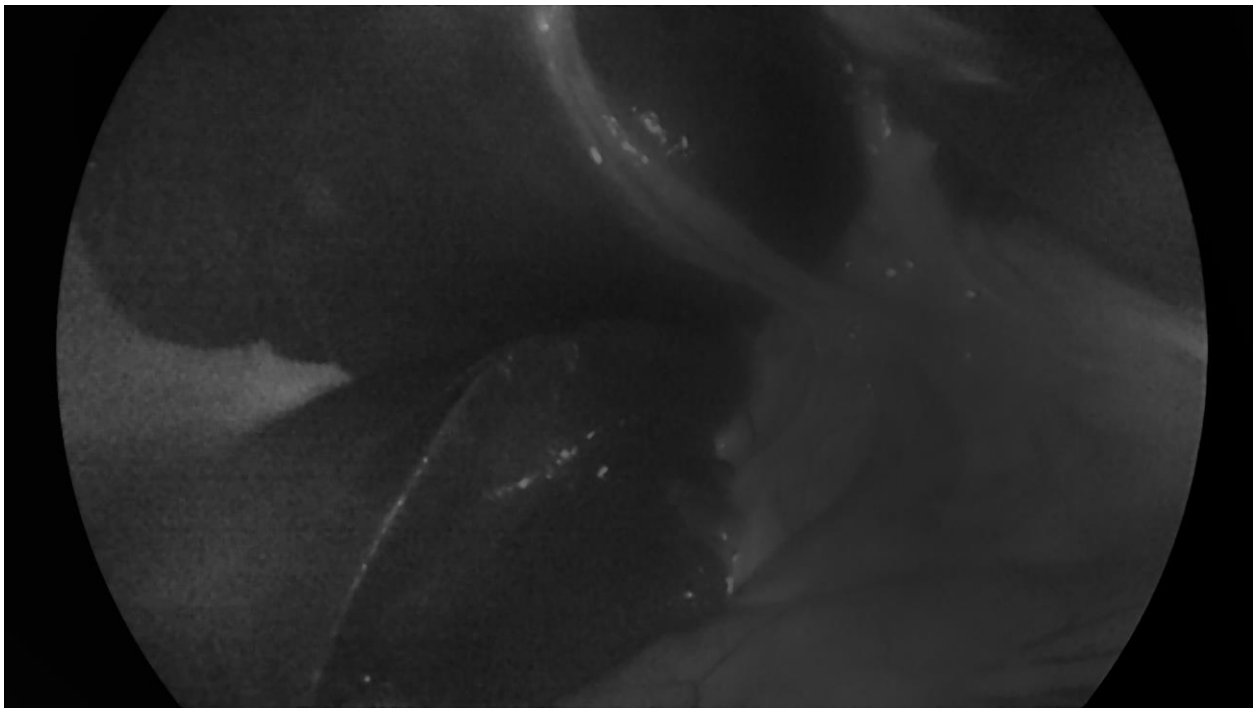

Fig. 8. Bias correction

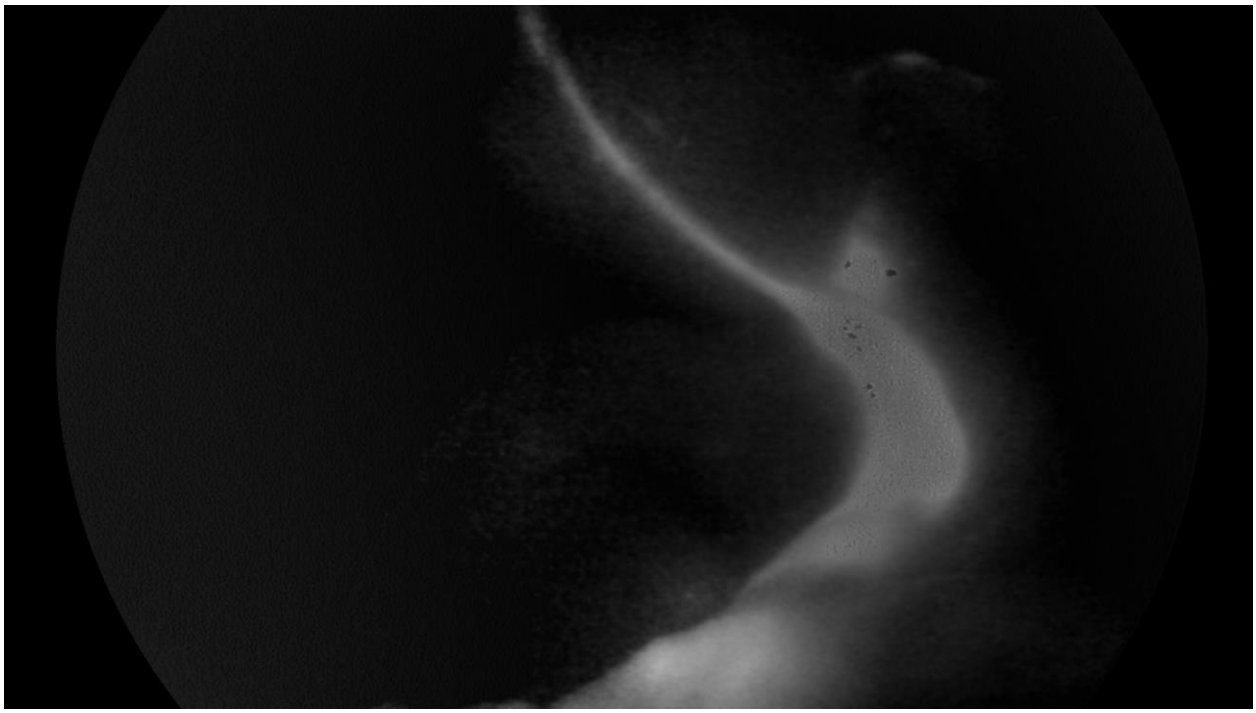

Fig. 9. Bias correction of fluorescence image

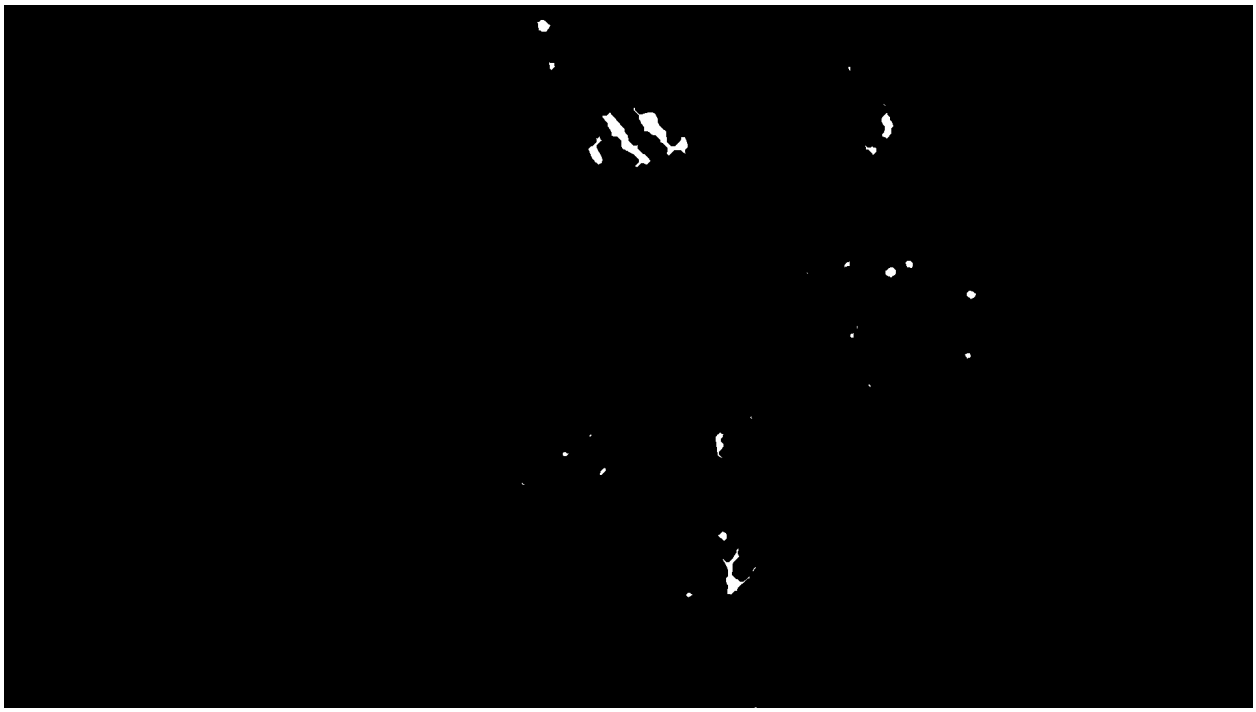

Fig. 10. Binary approximation of specular reflections in the grayscale image

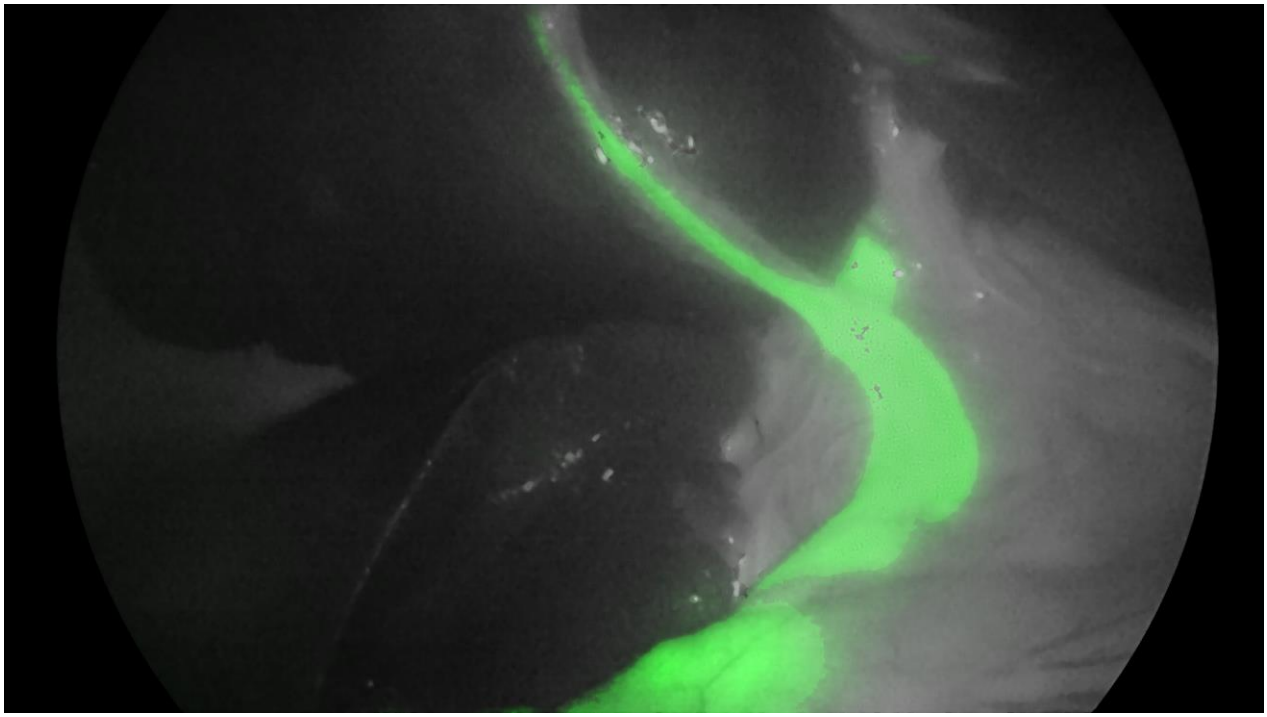

Fig. 11. Combined processed image