Chromaticity Based Smoke Removal in Endoscopic Images

Kevin Tchaka[a,b], Vijay M. Pawar[a], and Danail Stoyanov[a,b]

[a]Dept of Computer Science, University College of London
[b]Center for Medical Imaging Computing, University College of London, Gower Street WC1E 6BT, London, United Kingdom

ABSTRACT

In minimally invasive surgery, image quality is a critical pre-requisite to ensure a surgeons ability to perform a procedure. In endoscopic procedures, image quality can deteriorate for a number of reasons such as fogging due to the temperature gradient after intra-corporeal insertion, lack of focus and due to smoke generated when using electro-cautery to dissect tissues without bleeding. In this paper we investigate the use of vision processing techniques to remove surgical smoke and improve the clarity of the image. We model the image formation process by introducing a haze medium to account for the degradation of visibility. For simplicity and computational efficiency we use an adapted dark-channel prior method combined with histogram equalization to remove smoke artifacts to recover the radiance image and enhance the contrast and brightness of the final result. Our initial results on images from robotic assisted procedures are promising and show that the proposed approach may be used to enhance image quality during surgery without additional suction devices. In addition, the processing pipeline may be used as an important part of a robust surgical vision pipeline that can continue working in the presence of smoke.

Keywords: surgical vision, minimally invasive surgery, endoscopy, smoke removal

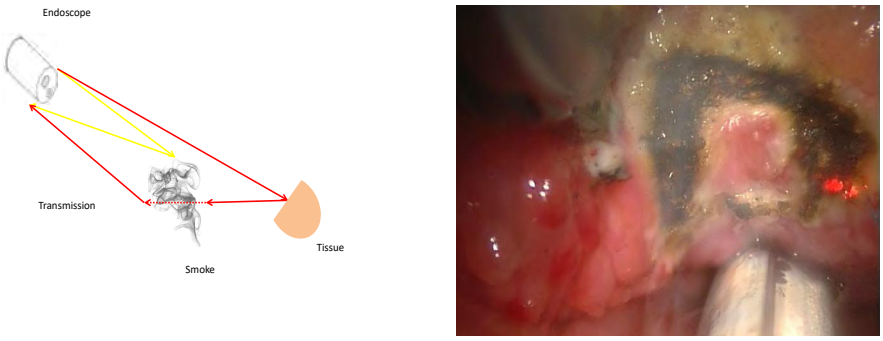
1. INTRODUCTION

In robotic assisted minimally invasive surgery (MIS), the visibility of the surgical site is of the utmost importance for the operating surgeon. When manipulating the surgical site it is common to use energy based surgical instruments in order to avoid bleeding due to vessel hemorrhage during tissue dissection. The use of these instruments results in the production of smoke.[1] This means that at times the procedure is paused to clear out smoke by using suction devices in order to recover clear visibility.[2] Removing the smoke from surgical video images without pausing the procedure using a software smoke removal solution would potentially simplify clinical work flow and shorten the duration of surgery. In addition to the potential workflow convenience of using software based smoke removal, the algorithms to achieve it would also form a component within computer assisted interventions using surgical vision where robustness in the presence of smoke could be improved.[3–5]

The removal of mediums that obscure image formation through computer vision has been studied for haze removal in outdoor scenes affected by bad weather.[6–10] Image dehazing methods usually rely on an optical model of light transmission through a particle-filled medium to estimate the image degradation and thus recover the haze-free image. Fattal et al.[6] assumed that transmission and surface shading are locally uncorrelated to first infer the transmission in the area affected by thin haze before propagating to denser haze areas via a Markov Random Field formulation and statistical smoothing. Although this method produces excellent results, it is computationally expensive, and not well suited to cases where the image is globally affected by an obscuring medium. Tan et al.[7] made the observations that haze free images generally present more contrast than hazy ones, and that the airlight variation over a local patch tends to be smooth. These allowed them to recover the radiance by maximizing contrast over a local window of the hazy image with a final algorithm that provided compelling visual results but also shifted regions within the image into a different color map. Zhang et al.[8] used chromaticity prior to relieve the computational load of the dehazing operation. Their algorithm extracted the transmission

Further author information: Kevin Tchaka:
E-mail: kevin.tchaka@gmail.com, Telephone: 44 787 945 8127

Medical Imaging 2017: Image Processing, edited by Martin A. Styner, Elsa D. Angelini, Proc. of SPIE
Vol. 10133, 101331M · © 2017 SPIE · CCC code: 1605-7422/17/$18 · doi: 10.1117/12.2254622



(a) (b)

Figure 1. Schematic illustration of image formation through a hazy medium. (a) schematic illustrating the formation of an image through a smoky medium. (b) Actual endoscopic image impacted by smoke from a cholecystectomy sequence.

iteratively with the assumption that large-scale chromaticity variations are due to transmission whereas small-scale luminance variations should be due to the scene albedo. the transmission map is finally refined by a non-linear edge preserving filter. This technique provides good results at a relatively low computational cost when compared to previous works. He and al.[9] proposed the dark-channel prior. This approach stems from the observation that most local patches in haze free images contain some pixels with very low intensity values in at least one color channel. Using this prior information the transmission can be extracted from the hazy image by a minimum operator. This method is simple and effective though it also causes a slight chromatic change in the image.

In this paper, we propose to extend existing work to remove smoke from stereo laparoscopic surgery, using image and video processing techniques in order to remove smoke from endoscopic surgery images. Considering smoke as a type of haze and excluding the assumptions pertaining to the presence of infinitely positioned objects in the image, the optical model used to explain image degradation in hazy medium haze still applies to endoscopic images. As a result, we expect that methods developed for outdoor image dehazing should be applicable to smoke removal for surgical images. Particularly, we investigate the results of applying the method described in He and al.[9] to surgical video in order to remove smoke globally before applying local contrast enhancement[11] to improve image quality.

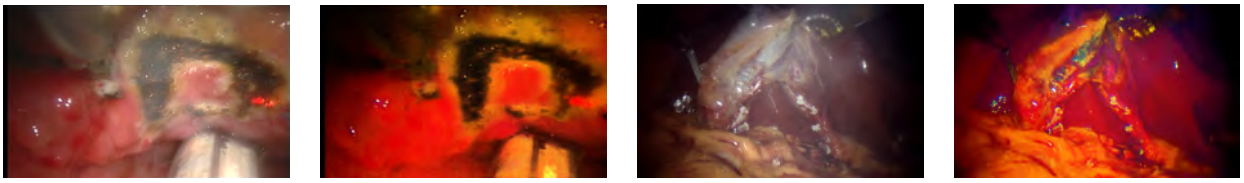
2. METHODS

2.1 Optical Model

The degradation of an image in the presence of smoke occurs due to the scattering of the light reflected from a surface in the atmosphere before reaching the camera. As such the light reaching the camera is composed of a sum of previously scattered rays plus those reflected from the smoke particles. These are commonly referred to as *transmission* and *airlight*. The formation of an image through a hazy medium is illustrated in **Fig. 1** (a). The light emitted from the endoscope is separated in two types of rays. The first, represented in yellow, bounces directly of the hazy medium while the second, represented in red, bounces of the tissues on the right-hand side before being scattered through the smoke. Both then reach the endoscope's camera being summed to form the image in (b).

Mathematically, image aberration due to smoke can be represented by the following assumptions. Along a very short distance along a light ray, let it be called $d\boldsymbol{r}$, $\mathbf{r}$ being an arc-length parametrization of the light ray, the fraction of light absorbed can be assumed to be linearly related to the medium extinction coefficient or *albedo* β. Integrating this process along an entire light ray we can express the transmission t over a surface s at a distance d from the surface to the sensor as:[12]

$$t = e^{-\int_0^d \beta(\boldsymbol{r}(s))ds} \tag{1}$$



before after before after

Figure 2. Two example stereo endoscopic images from robotic assisted surgery before and after smoke removal. The two images on the left side show transoral robotic surgery where a laser is used to debunk the base of tongue which creates smoke. The two on the right are from robotic assisted laparoscopic cholecystectomy where diathermy creates the smoke artifacts. The smoke is removed however the image color space is significantly shifted due to the illumination color. This can potentially be resolved through color consistency and illumination calibration.

In a pixel sense this is usually expressed as:

$$t(x) = e^{-\beta d(x)} \tag{2}$$

Where $d(x)$ is the scene depth at pixel x. Pixels are hereafter, considered individually but in practice, x should be the coordinate vector of an individual pixel. From these the optical model for the formation of hazy images can be expressed by the McCartney model[9] as follows:

$$I\left(x\right) = J\left(x\right)t\left(x\right) + A\left(1 - t\left(x\right)\right) \tag{3}$$

Where I represents the resulting hazy image, x is a single pixel, A is the airlight, t is the transmission and J is the smoke free image. $J\left(x\right)t\left(x\right)$ is commonly referred to as the attenuation module, and $A(1 - t\left(x\right))$ is the airlight module. Considering (3) the goal of smoke removal is to recover J by estimating the transmission and airlight. It should be noted that the model assumes that there is no other light source hidden within the haze and that the hazy medium is not a mix of different types of haze. Should these assumptions be violated methods relying on the model are susceptible of failure.

2.2 Dark-Channel Prior

The dark-channel prior (DCP) rule[9] is a reliable and effective approach for haze removal from single images. The method stems from the basic observation that most smoke-free images contain pixels with very low intensities in at least one color channel, due to shadows, colorful objects or dark objects. This is the dark-channel. On the other hand, smoke is mostly gray / whitish and hazy patches will mostly contain pixels with their three color channel at a pretty high value when compared with their smoke free counterpart. It is possible to roughly estimate the thickness of the smoke from the intensity of the dark-channel point as it becomes higher the thicker the smoke get. The dark-channel is mathematically expressed as:

$$J^{dark}\left(x\right) = \min_{c\in(r,g,b)}\left(\min_{y\in\Omega(x)} I^{c}\left(y\right)\right) \tag{4}$$

Where J^{dark} is the dark-channel, c is a color channel, $\Omega\left(x\right)$ represents a patch of the image centered on pixel x, and y represents a pixel belonging to the local patch Ω. Taking the 0.1% brightest patch of the dark-channel, the airlight is estimated as the color of the brightest of these pixels from the original image I. Then from the dark-channel prior theory and some manipulation of the McCartney model[9] the transmission can be obtained as:

$$t\left(x\right) = 1 - \omega\min_{c}\left(\min_{y\in\Omega(x)}\left(\frac{I^{c}\left(y\right)}{A^{c}}\right)\right) \tag{5}$$

Where $\omega \in]0\ 1]$ is a transmission constant introduced to preserve a minimum amount of haze. It is set to 0.95 in this paper as specified by He and al.[9] This constant is introduced because even in haze-free images, the air is not bereft of particles, plus the presence of haze is a fundamental cue for human to perceive depth. Thorough removal of all haze may make the image seem unnatural. Ω is a local patch of pixel considered. From this the scene radiance is recovered by a simple transformation on (3).

$$J(x) = \frac{I(x) - A}{\max(t(x), t_0))} + A \tag{6}$$

Where t_0 is a minimum transmission constant to avoid division by 0. In this paper we set $t_0 = 0.1$ as in He and al.[9] Both parameters value t_0 and ω were empirically found working values in He and al.[9] and reused exactly as is. The results are shown in **Fig. 2**. The left image of each pair is the original image where smoke is present and the right image is the resulting image after applying the global dark-channel method.[9] The smoke has been successfully removed, however the process darkens the global image.

2.3 Refined Dark-Channel Prior

This color change is due to the fact that some assumptions of the original DCP[9] do not hold in our case. According to He and al.[9] the dark-channel of an image was to relate directly the brightness of a pixel in the dark-channel to the haze density in the scene, allowing for accurate and simple estimation of the airlight and transmission. This was due to the assumption that infinitely distant objects would be present in the scene (e.g. sky region). the brightest pixels in the dark-channel would thus correspond to the regions of highest haze density in the input image. As it happens in our case there is no infinitely distant point, thus the dark-channel often ends up polluted by pixel brightness uncorrelated with the haze density of the region (e.g. surgical devices, reflection of endoscopic light on tissues, etc.). To address this issue it was important to find a way to refine our dark-channel and ensure that pixels impacted by haze would have their brightness maximized. To achieve this goal two methods were investigated.

First, considering that the intensity of smoky pixels would generally not be too close to white (1.0 if using floating values) we considered thresholding our dark-channel to eliminate outliers. Empirically, we found the value $T = 0.88$ to achieve our purposes. Although this method intuitively seems good, and provides acceptable results, its justification is shaky at best. Plus, the threshold value, though good in the cases tested cannot be ascertained and justified precisely. It might or not need readjustment in presence of a different case.

Second, we considered maximizing the brightness values of the smoke region expected to be in the range $[0.35; 0.65]$ while minimizing other values. Due to difference in lighting, depth and other factors, it is impossible to define a specific value for smoky pixels. However, sampling over multiple sequences and frames, it was observed that the brightness of smoky pixels in the dark-channel tended to be roughly situated in the $[0.35; 0.65]$ range. To account for the fact this range is but a rough estimate it was rather decided to put the emphasis on pixels close to 0.5 and decrease the emphasis as their distance to this value increase. When considering floating brightness value (e.g. $x \in [01]$) this can be easily achieved via the following formula:

$$J^{'dark}(x) = J^{dark}(x) \times \left(1.0 - J^{dark}(x)\right) \tag{7}$$

Where $J^{'dark}(x)$ is the refined dark-channel while $J^{dark}(x)$ is the original dark-channel computed using (4). The obtained refined dark-channel is then further normalized to the desired range.

$$J^{'dark}(x) = J^{dark}(x) \,/ \max_{y \in J^{'dark}} J^{'dark}(y) \tag{8}$$

Fig. 3 shows the two different refined dark-channel priors, compared to the original one and their associated recovered radiance. In the thresholded case, the surgical instruments as well as some patch of reflected light from the tissues were effectively removed but some patches of smoke were also wrongly considered as noise. When applying (7) the color space of the results was closer to the original, though the surgical instrument and the reflected dots of light on the tissue were not ignored as well as before. However the smoke region was in core

(a) (b) (c) (d) (e)

Figure 3. Comparison of the different dark-channel refining. (a) Original dark-channel computed with (4), (b) Dark-channel thresholded at 0.88, (c) Recovered radiance using the thresholded dark-channel, (d) Dark-channel refined with (7), (e) Recovered radiance using the dark-channel computed with (7).

well highlighted. Although in both cases the recovered radiances tended to be much closer to the original color space of the input image, refining the dark-channel using (7) seems to bring better resulting colors.

2.4 Image Contrast Adjustment

Using the DCP the result image normally has lower brightness than the original scene. Therefore it is useful to enhance the contrast in order to improve the recovered image quality. This enhancement brings about the increase of the contrast of brightness of dark regions, thus highlighting image edges and details. Many algorithms are available for this purpose such as Retinex, gamma correction and histogram equalization. Following in the footsteps of Ji et al.[13] in this paper we decided to apply histogram equalization due to its simplicity and low computation requirements.

Histogram equalization enhances the information in the image by widening the gray level interval and spreading out the most frequent intensity values. Since we were dealing with color images, the histogram equalization algorithm was applied to the value channel of the HSV conversion of the recovered radiance before converting back to RGB color space. The basic principle is to use the histogram of the image to compute a gray mapping table. Using this table each pixel value is calculated making the distribution uniform. The histogram equalization improves the image contrast, however it can also bring out the noise, thus an additional smoothing technique was used in addition. Median filtering is a simple and fast-smoothing technique that perfectly suits our purpose as it can easily suppress noise while preserving image detail.

3. RESULTS

The system was implemented using the OpenCV library[14] in Microsoft Visual Studio 2012. It took 347 milliseconds to process a 1440 x 288 image on a PC with a 2.40 GHz Intel Core i7-4700MQ processor. The results are shown in **Fig. 4**. The images left of each pair are the original frames where smoke is present and the images right are the resulting frames after going through our system. For this instance the results are obtained using the refined dark-channel presented in (7). Visually, it can be seen that the smoke has been partly removed while retaining the original color map. Interestingly, in the protastectomy (a), the smoke is mostly light and diffuse throughout the whole scene, whereas in case (b) it is more localized, but our algorithm processes both cases with acceptable results.

In order to try and quantify the results of the method a simple distortion measure was performed. The motivation was the following; since the scene is mostly static the moving tools used to perform the surgical operation stay mostly out of the field of view, one of the smoke-free frame at the beginning of the original sequence provided a ground truth clear radiance. Therefore, computing a mean-squared error (MSE) analysis against it, most of the distortion would be due to smoke impacted regions. We thus computed the MSE over both the original and processed sequence frames using the same ground truth image and compared the results to get an estimate of the smoke reduction. Running this analysis on the transoral laser surgery sequence yielded the following results.

Because the first frames of the sequence included a lot of movements from the robot's end-effectors which would corrupt the measurements, they were ignored and a suitable frame was selected as ground truth. The original sequence contained 532 frames of which 15 were ignored and the 16^{th} was selected as the ground truth

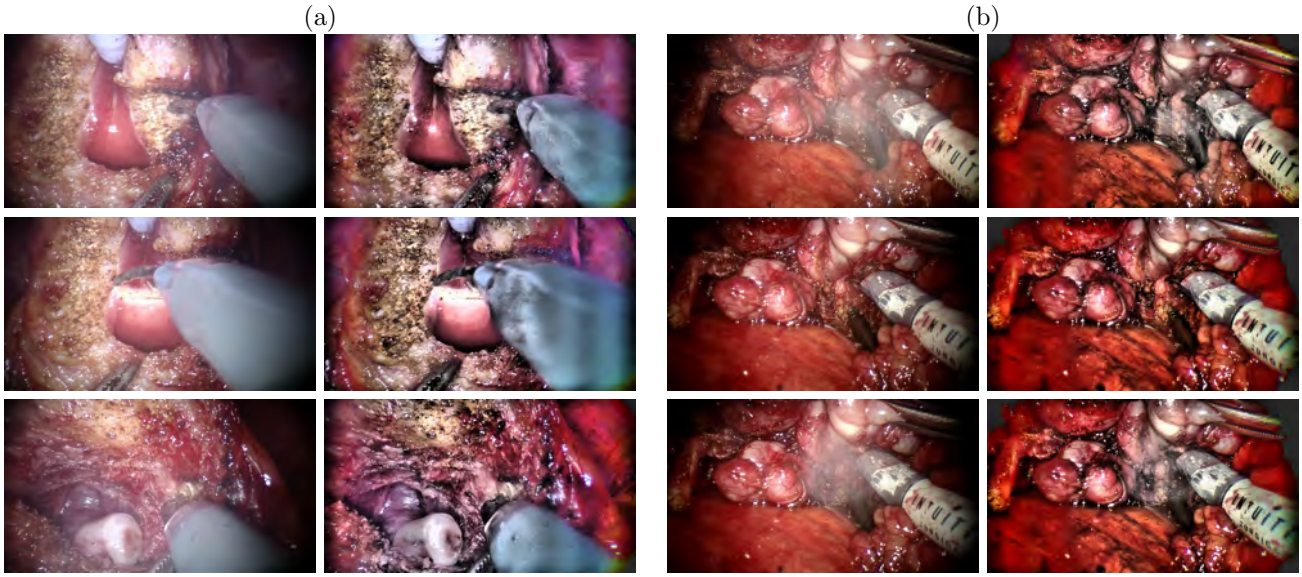


Figure 4. Result images from two prostatectomy operation. Each pair of image has the original frame on the left and the processed frame on the right.

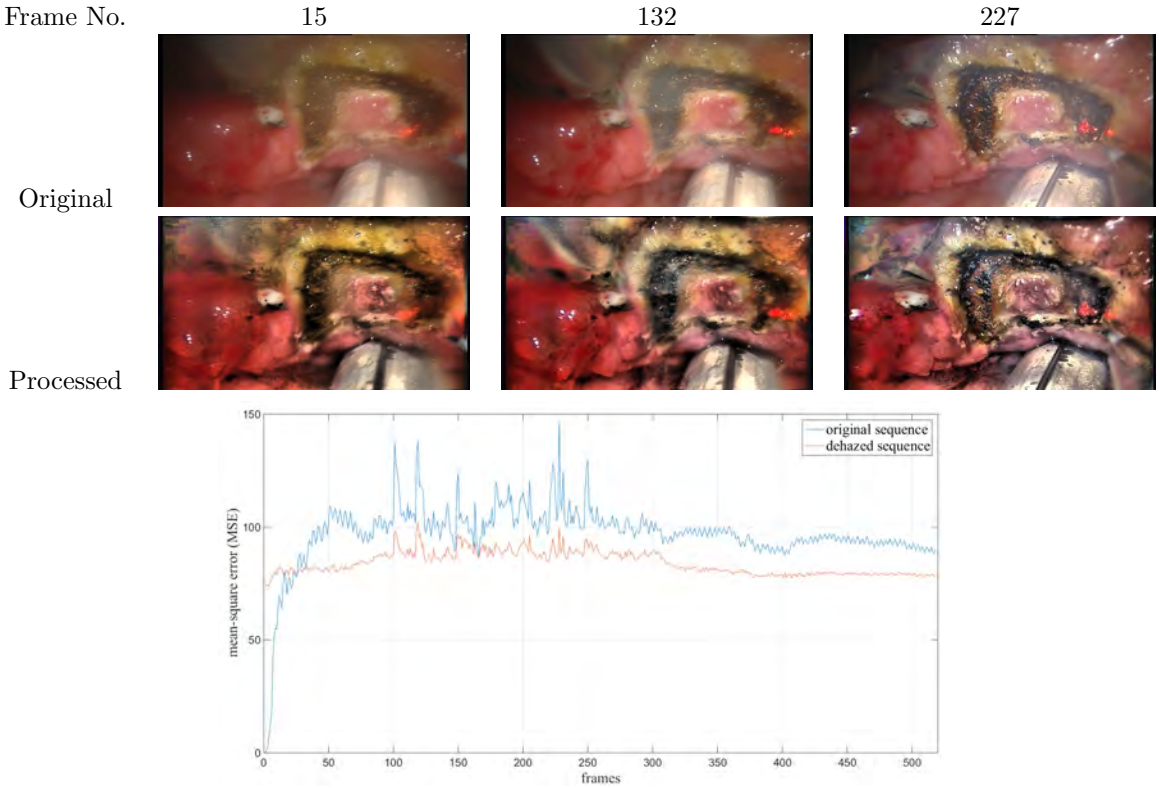


Figure 5. Comparison of the MSE values computed for both the original sequence (blue curve) and the dehazed one (red curve) over the different frames. The ground truth image used for the MSE estimation is the same in both cases and is the first static smoke-free frame of the original sequence. All error measures are made in comparison to this frame. Images on the top show the difference between the original sequence frames and the processed ones at specific frames.

Table 1. MSE comparison for some frames

Frame No	Original frame	DCP[9]	thresholded	Refined (7)
118	134.53	100.59	120.22	102.80
131	106.94	98.54	103.27	90.253
149	118.45	96.20	112.81	95.39
205	120.48	96.27	115.63	96.24
228	148.81	104.45	129.37	99.69
Mean	125.84	99.21	116.26	96.87
σ	16.15	3.44	9.59	4.73

case. Leaving over 516 frames for the computations. The graph presented in **Fig. 5** shows the evolution of the MSE over the frame for both the original and processed (dehazed in the graph's legend) frames.

The frame-by-frame MSE analysis ran over the sequence gives a mean error value of 97.03 while the same operation over the corresponding frames of the processed sequence returns a mean error of 83.98. The error is in average reduced by 13.04. Although the error measured cannot be completely imputed to the haze due to minutes modifications of the operated region and the movements of the laser itself. These have a minor impact on the measurement and the major pikes in **Fig. 5** do correspond to the most severely smoke-impacted frames. The dehazing process as well innately creates an increase in contrast which would impact the error measurement as can be seen over the very first frames of **Fig. 5** where a very wide distinction between the original and processed sequence is visible. Analysis of the standard deviation of both MSE sequences gives a clearer view of the impact of the method on the variation of the error. The original sequence MSE has a standard deviation value of $\sigma_{or} = 14.165$ while the processed sequence shows a standard deviation value of $\sigma_{proc} = 5.08$. From this it can be seen that the dehazing process quite clearly has a reduction impact on the smoke-induced distortion of the frames.

Some smoke impaired frame from the transoral sequence were selected to compare the performance using the original method,[9] the thresholded dark-channel (see **Fig. 3** images 2 and 3) and the refined dark-channel prior using (7). The MSE was computed with the same previous smoke-free frame as ground-truth.

Table I shows the MSE result of each approach, from it we can see that in spite of the color space shift, the original method[9] yields the lowest MSE result. For both refinement of the dark-channel method (threshold and using (7)) the results though higher than for He and al.[9] are still better than the original smoke impacted frame. This correlates with our previous intuition that the refinement have impacted slightly the reduction of the smoke. However, when considering both the final visual quality and the smoke reduction, the refined dark-channel using (7) yields the best compromise.

4. CONCLUSIONS

In order to improve the visibility of images and video during robotic assisted surgery it is important to remove aberrant artifacts, such as smoke. Smoke may be generated during instrument tissue interaction with electro-cautery instruments and it may obscure the field of view for the operating surgery. In this study, we investigated the process of image degradation in the case of smoke and proposed that algorithms normally used for outdoor images dehazing might be used in robotic assisted surgery to build a robust surgical vision pipeline. We used the dark-channel prior and image contrast enhancement to restore and enhance the clarity of the images. However the colors were not natural and looked more vivid than the true colors of the actual scene. A refining scheme for the dark-channel prior was implemented to cover this issue, results are quite promising but come with a reduction of the smoke density removal from the image. The actual relation between these two occurrences would benefit from further investigation and quantification of the smoke density removal attained.

5. ACKNOWLEDGMENTS

Danail Stoyanov receives funding from the EPSRC (EP/N013220/1, EP/N022750/1, EP/N027078/1, NS/A000027/1), TheWellcome Trust (WT101957, 201080/Z/16/Z) and the EU-Horizon2020 project EndoVESPA (H2020-ICT-2015-688592).

REFERENCES

[1] Ulmer, B. C., "The hazards of surgical smoke," {*AORN*} *Journal* **87**, 721 – 738 (April 2008).

[2] Ball, K. A., *Surgical smoke evacuation guidelines - compliance among perioperative nurses*, PHD dissertation, Virginia CommonWealth University, Richmond, Virginia (March 2009).

[3] Stoyanov, D., Mylonas, G., Lerotic, M., Chung, A., and Yang, G.-Z., "Intra-operative visualizations: Perceptual fidelity and human factors," *JOURNAL OF DISPLAY TECHNOLOGY* **4**, 491–501 (2008).

[4] Stoyanov, D., "Surgical vision," *Annals of Biomedical Engineering* **40**(2), 332–345 (2012).

[5] Maier-Hein, L., Groch, A., Bartoli, A., Bodenstedt, S., Boissonnat, G., Chang, P. L., Clancy, N. T., Elson, D. S., Haase, S., Heim, E., Hornegger, J., Jannin, P., Kenngott, H., Kilgus, T., Mller-Stich, B., Oladokun, D., Rhl, S., dos Santos, T. R., Schlemmer, H. P., Seitel, A., Speidel, S., Wagner, M., and Stoyanov, D., "Comparative validation of single-shot optical techniques for laparoscopic 3-d surface reconstruction," *IEEE Transactions on Medical Imaging* **33**, 1913–1930 (Oct 2014).

[6] Fattal, R., "Single image dehazing," *ACM Trans. Graph.* **27**, 72:1–72:9 (August 2008).

[7] Tan, R. T., "Visibility in bad weather from a single image," in [*Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*], 1–8 (June 2008).

[8] Zhang, J., Li, L., Yang, G., Zhang, Y., and Sun, J., "Local albedo-insensitive single image dehazing," *The Visual Computer* **26**(6), 761–768 (2010).

[9] He, K., Sun, J., and Tang, X., "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**, 2341–2353 (December 2011).

[10] Nayar, S. K. and Narasimhan, S. G., "Vision in bad weather," in [*Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*], **2**, 820–827 (1999).

[11] Gijsenij, A. and Gevers, T., "Color constancy using natural image statistics and scene semantics," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**, 687–698 (April 2011).

[12] Koschmieder, H., [*Theorie der horizontalen Sichtweite: Kontrast und Sichtweite*], no. vol. 2 in Beiträge zur Physik der freien Atmosphäre, Keim & Nemnich (1924).

[13] Ji, X., Cheng, J., Bai, J., Zhang, T., and Wang, M., "Real-time enhancement of the image clarity for traffic video monitoring systems in haze," in [*Image and Signal Processing (CISP), 2014 7th International Congress on*], 11–15 (October 2014).

[14] Bradski, G. *Dr. Dobb's Journal of Software Tools* (2000).