JOINT DESMOKING AND DENOISING OF LAPAROSCOPY IMAGES

Alankar Kotwal[1] *Riddhish Bhalodia*[1] *Suyash P. Awate*[2]

[1] Department of Electrical Engineering
Indian Institute of Technology (IIT) Bombay

[2] Department of Computer Science and Engineering
Indian Institute of Technology (IIT) Bombay

ABSTRACT

Laparoscopic images in minimally invasive surgery get corrupted by surgical smoke and noise. This degrades the quality of the surgery and the results of subsequent processing for, say, segmentation and tracking. Algorithms for desmoking and denoising laparoscopic images seem to be missing in the medical vision literature. This paper formulates the problem of *joint desmoking and denoising* of laparoscopic images as a *Bayesian inference* problem. It relies on a novel *probabilistic graphical model* of the images, which includes *novel prior models* on the uncorrupted color image as well as the transmission-map image that indicates color attenuation due to smoke. The results on simulated and real-world laparoscopic images, including clinical expert evaluation, shows the advantages of the proposed method over the state of the art.

Index Terms— Laparoscopy, desmoking, denoising.

1. INTRODUCTION AND RELATED WORK

Surgical smoke [1] and noise in laparoscopy images reduces the surgeon's visibility [2, 3] and degrades the results of subsequent processing for segmentation [4, 5], tracking [6], stereoscopy [7, 8], and augmented reality [9]. Algorithms for desmoking and denoising laparoscopic images seem to be missing in the literature. This paper aims to fill this void.

The problem of desmoking laparoscopic images has some similarity to the problem of dehazing landscape images, but unlike haze that is caused by fog and is determined by scene depth, smoke concentration is a local phenomenon and can vary independent of depth. An early method for dehazing [10] models the transmission map (a function of the scene depth) as a Gaussian Markov random field (MRF) that fails to preserve edges. Another work [11] models a spatially varying airlight image as a MRF using edge-preserving potential functions. However, both [10, 11] use *no* priors on color probability density functions (PDFs) in the uncorrupted image.

Later approaches for dehazing [12, 13] (i) propose a dark-channel prior on the colors in the uncorrupted image, (ii) use this prior to estimate the transmission map via soft matting or guided filtering, and (iii) use this map to estimate the hazeless image. However, this sequential estimation [10, 11, 12, 13] cannot exploit the information in the estimated hazeless image to improve the transmission-map estimate.

Later methods perform denoising in addition to dehazing. A recent work [14] couples the dark-channel prior with local Wiener filtering for real-time dehazing and denoising, but, due to time constraints, cannot employ edge-preserving prior models on the transmission map or the uncorrupted image. Some methods [15] sequentially denoise and dehaze via repeated acquisitions, but this is inapplicable to laparoscopy. A joint dehazing and denoising algorithm [16] uses the result of [12] as an initial solution, but subsequently performs kernel-regression based iterations without any constraints / priors on the transmission map or the uncorrupted image. Thus, [16] does *not* define a unified optimization problem over the image and has *no* guarantees for convergence or uniqueness of the solution. A recent work [17] uses a MRF to model albedo and depth, but, unlike us, does *not* use a prior on the color PDFs and uses a smoothness prior only on the chromaticity in the uncorrupted image; thus, their results [17] evidently depict unnatural colors and brightness variations.

We formulate the problem of *joint desmoking and denoising* of laparoscopic images as a *Bayesian inference* problem. It relies on a novel *probabilistic graphical model* including the uncorrupted image, transmission map, and the corrupted image. The proposed model includes novel *priors* to (i) preserve natural color and contrast in the uncorrupted color image, and (ii) enforce edge-preserving spatial regularity on the uncorrupted color image and the transmission-map image. Unlike [12] that estimates the airlight color relying on the dark channel prior that is sensitive to outliers, we propose a new estimation method, using a learned prior, which is robust to outliers. The results on simulated and real-world laparoscopic images, including clinical expert evaluation, show that the proposed method produces improved estimates of uncorrupted images, compared to the state of the art.

2. METHODS

This section details our novel probabilistic graphical modeling and inference. We denote random variables by uppercase letters and their specific instances by the lowercase letters.

We assume that the *unknown uncorrupted image* ap-

The first two authors contributed equally. We thank funding via the IIT Bombay Seed Grant 14IRCCSG010.

pears piecewise smooth, where each piece corresponds to an anatomical structure or a surgical instrument. We model the uncorrupted image, comprising I pixels, by an edge-preserving MRF $\mathbf{X} := \{X_i\}_{i=1}^{I}$, where, at each pixel i, the random variable X_i is vector valued (3-vector for the 3 color components). For the MRF, we define a neighborhood system $\mathcal{N} := \{\mathcal{N}_i\}_{i=1}^{I}$, where the set $\mathcal{N}_i$ gives the set of neighboring pixels for pixel i. We give the model details later.

Image Degradation Model. Laparoscopic images get degraded due to: (i) surgical smoke of gray color, (ii) diffuse-reflected light of blood red color, and (iii) random noise introduced via insufficient lighting and sensor errors.

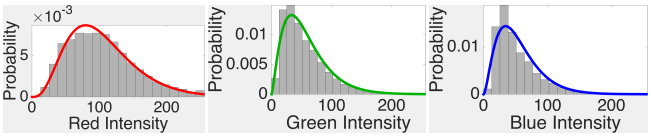
We model the unknown *corruption color* resulting from surgical smoke and diffuse reflection by random variable S that corrupts the color at pixel i via a convex weighted combination of the uncorrupted image color X_i and S. We call an image of one of these weights as the *transmission map*. We model the *unknown transmission map* by an edge-preserving MRF $\mathbf{T} := \{T_i\}_{i=1}^{I}$, where, $\forall i$, the random variable T_i takes real values in $(0, 1)$. For this MRF on $\mathbf{T}$, we define a neighborhood system $\mathcal{N}^T := \{\mathcal{N}_i^T\}_{i=1}^{I}$, where the set $\mathcal{N}_i^T$ gives the set of neighboring pixels for pixel i. We give the model details later. Given the uncorrupted image $\mathbf{X}$, corruption color S, and transmission map $\mathbf{T}$, the corrupted image color at pixel i is $T_iX_i + S(1-T_i)$. While T_iX_i models *direct attenuation* due to the loss in scene radiance, $S(1-T_i)$ models the contribution due to scattered and diffuse-reflected light. We assume the transmission map to be piecewise smooth, where its values vary with smoke concentration and scene depth.

We assume each pixel's color to be also corrupted by independent and identically distributed (i.i.d.) random noise. We approximate the overall per-pixel noise model to be additive zero-mean Gaussian with variance σ^2. We model the *observed corrupted image* using the random field $\mathbf{Y} := \{Y_i\}_{i=1}^{I}$. Given the uncorrupted image $\mathbf{X}$, corruption color S, and transmission map $\mathbf{T}$, the observed image color Y_i at pixel i, corrupted by smoke and noise, is $Y_i := T_iX_i + S(1-T_i) + \eta_i$, where $\eta_i \sim G(0, \sigma^2)$.

Problem Formulation. We formulate laparoscopic image desmoking and denoising as a maximum-a-posteriori (MAP) estimation problem. Given the corrupted image $\mathbf{y}$, we first find a robust estimate of corruption color s and then the MAP estimate of the uncorrupted image $\widehat{\mathbf{x}} := \arg\max_{\mathbf{x}} (\max_{\mathbf{t}} P(\mathbf{x}|\mathbf{y}, s, \mathbf{t}))$, where the posterior $P(\mathbf{x}|\mathbf{y}, s, \mathbf{t}) \propto P(\mathbf{y}|\mathbf{x}, s, \mathbf{t})P(\mathbf{x}, s, \mathbf{t})$. Here, $P(\mathbf{y}|\mathbf{x}, s, \mathbf{t})$ is the likelihood of observing the data $\mathbf{y}$. We design the prior $P(\mathbf{x}, s, \mathbf{t}) \propto P(\mathbf{x})P(\mathbf{t})$ to incorporate independent models on the uncorrupted image X and the transmission-map image T.

The i.i.d. additive Gaussian noise leads to the likelihood as $P(\mathbf{y}|\mathbf{x}, s, \mathbf{t}) := \prod_{i=1}^{I} G(y_i|t_ix_i + s(1-t_i), \sigma^2)$.

Prior Model on Uncorrupted Image X. We design a MRF prior on X through its (i) pixel color PDF, to preserve natural color and contrast, and (ii) joint PDFs of colors in neighborhoods, to enforce spatial regularity.



(a) (b) (c)

Fig. 1. Learning Prior PDFs on Color. The empirical histogram (bar plot) and fitted gamma PDF (solid curve) in uncorrupted laparoscopic images, for the color components **(a)** red: PDF Γ_R, **(b)** green: PDF Γ_G, and **(c)** blue: PDF Γ_B.

We use a set of high-quality (virtually, uncorrupted) images to learn the PDF of each color component. We propose to model the 3 empirical PDFs as gamma PDFs ($\Gamma_R(\cdot)$, $\Gamma_G(\cdot)$, $\Gamma_B(\cdot)$), which give a good fit (Figure 1). We design the prior on the color components to penalize the deviation between the gamma cumulative distribution functions (CDFs) and the color component CDFs ($F_R(\cdot)$, $F_G(\cdot)$, $F_B(\cdot)$) in the estimated uncorrupted image x. We penalize the Kolmogorov-Smirnov (KS) distance between the CDFs: $\text{KS}(F_\Gamma, F) := \text{KS}(F_{\Gamma_R}, F_R) + \text{KS}(F_{\Gamma_G}, F_G) + \text{KS}(F_{\Gamma_B}, F_B)$. We can model KS distance as a MRF potential function using a fully-connected pixel-neighborhood system $\mathcal{N}$ [18]. Gamma PDFs have been used for some grayscale medical images [19, 20].

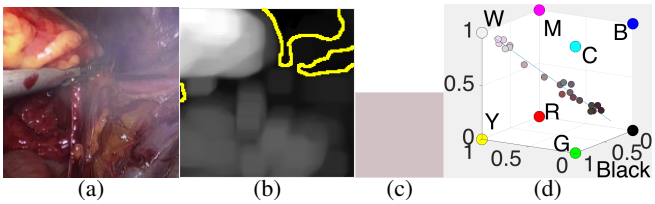
We enforce edge-preserving spatial regularity by penalizing the L_1 norm of the difference between the colors X_i and X_j at adjacent pixels i and j; all other 2-cliques have potential zero. This leads to a term, one for each adjacent-pixel clique $\{i, j\}$ in the neighborhood system $\mathcal{N}$, in the Gibbs energy as $\| X_i - X_j \|_1 = |X_{iR} - X_{jR}| + |X_{iG} - X_{jG}| + |X_{iB} - X_{jB}|$, where X_{iR} is the red color component in the color X_i at pixel i. We regularize the L_1 norm as the Huber function [18].

Thus, the prior model on the uncorrupted image X is $P(X) := (1/Z)\exp(-E(X))$, with partition function Z and Gibbs energy $E(X) := \alpha\text{KS}(F_\Gamma, F) + \sum_{i=1}^{I}\sum_{j\in\mathcal{N}_i} \beta \| X_i - X_j \|_1$, where $\alpha, \beta \in \mathbb{R}^+$ are free parameters.

Prior Model on Transmission-Map Image T. The transmission-map image is affected by the smoke concentrations and the depths in the scene. We propose a prior to model the transmission-map image as a piecewise-smooth image using the MRF $P(T) := (1/W)\exp(-F(T))$ with partition function W and Gibbs energy $F(T) := \sum_{i=1}^{I}\sum_{j\in\mathcal{N}_i^T} \gamma|T_i - T_j|$, with free parameter $\gamma \in \mathbb{R}^+$.

Robust Estimation of Corruption Color s. We propose a novel 2-stage algorithm to estimate the corruption color S, whose hue resembles smoke gray with a tinge of blood red.

Stage I. We take the corrupted image data y (Figure 2(a)) and automatically select regions with colors that are (i) homogeneous and (ii) close to a gray shade, i.e., similar RGB color components. We do this by constructing a grayscale image called the *smokiness map* (Figure 2(b)) as follows: (i) compute the standard deviation (SD) of the 3 color-component values at each pixel, (ii) set the intensity at pixel i to the 95-th percentile of these SD values in its 10×10 mm neighbor-



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

Fig. 2. **Estimating Corruption Color S. (a)** Corrupted image y. **(b)** Smokiness map (see text); bottom 1 percentile intensity region (yellow boundary). **(c)** Estimated color $\widehat{s}$ (shown as a constant-color image to aid visualization). **(d)** 1D subspace empirically learned for $\widehat{s}$ from training data.

hood. We select pixels (Figure 2(b)) having values in the bottom 1 percentile of the intensities in the smokiness map and estimate $\widehat{s}$ (Figure 2(c)) as the average color of the selected pixels. The neighborhood and percentiles (instead of max / min) ensures that $\widehat{s}$ is robust to outliers in the data y.

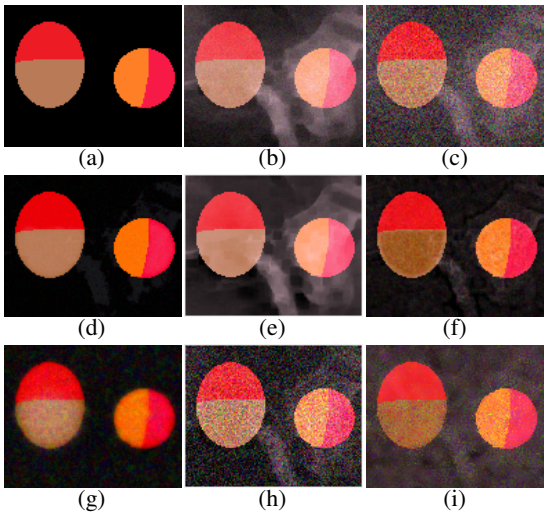
Stage II. We apply this estimation scheme on a *training* image set to learn an empirical distribution of $\widehat{s}$ (Figure 2(d)). We find $\widehat{s}$ to lie on a 1D subspace. For application on a corrupted laparoscopic image $\mathbf{y}$, we compute $\widehat{s}$ and then project it on this 1D subspace, for added robustness.

MAP Estimation of Images X, T. We find the MAP estimates of $\mathbf{t}, \mathbf{x}$ by using a gradient-ascent algorithm with adaptive step size to ensure that each update increases the log posterior. We find that a few iterations suffice for convergence. We replace the gradient of the KS distance by the difference between the current estimated x and its CDF-transformed image that matches its color-component histograms to the prior color-component histograms. We initialize the (i) transmission-map image as: $T_i^{\text{init}} \leftarrow 1, \forall i$ and (ii) uncorrupted image as: $X_i^{\text{init}} \leftarrow Y_i, \forall i$. We estimate the corruption color $\widehat{s}$ once before the MAP image estimation. We tune the free parameters α, β, γ empirically (just once for all experiments in the paper); Section 3 shows that the results are robust to the choice of values for the free parameters.

3. VALIDATION AND RESULTS

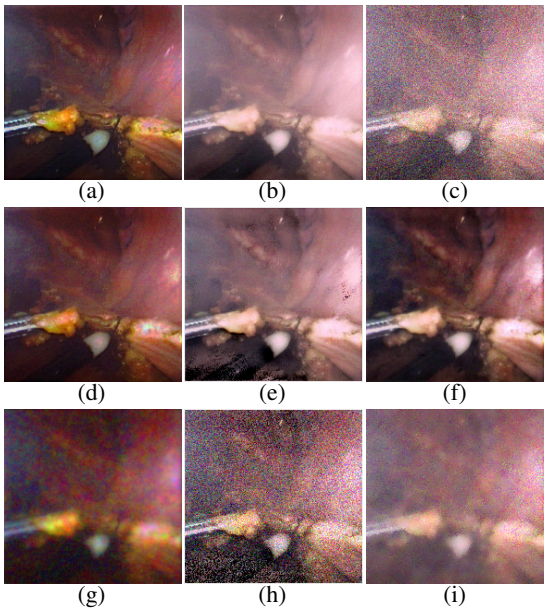
We perform quantitative and qualitative evaluation on (i) simulated data and (ii) 50 real-world laparoscopic images obtained from (`hamlyn.doc.ic.ac.uk/vision`) [21] and our clinical collaborators. Due to the absence of desmoking methods on laparoscopic / endoscopy images, we evaluate 2 other state-of-the-art algorithms that focus on dehazing and denoising of landscape images. To the best of our knowledge, ours is the first paper to propose a joint desmoking and denoising algorithm for laparoscopic images. For validating the methods, we use high-quality images, with very low smoke and noise, as ground truth. We introduce smoke in the images using transmission-map estimates obtained using [12].

Validation on Simulated Data. We simulate a phantom image, corrupt it with smoke and red diffuse-reflected light (using s) and noise. We evaluate the performance of (i) the proposed method, (ii) adaptive filtering [14], and (iii) a method performing denoising (we choose an edge-preserving



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 3. **Validation on Simulated Data. (a)** Simulated phantom (color component values $\in [0, 255]$). Smoky phantoms with **(b)** low noise ($\sigma = 3$) and **(c)** high noise ($\sigma = 30$). Results of processing image (b), using **(d)** our method, **(e)** adaptive filtering [14], and **(f)** denoising followed by dehazing [12]; a strategy considered in [15, 16]. **(g)–(i)** Results with the same 3 methods for processing image (c).



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 4. **Validation on High-Quality Laparoscopic Images. (a)** Ground-truth laparoscopic image (color component values $\in [0, 255]$). Smoky image with **(b)** low noise ($\sigma = 3$) and **(c)** high noise ($\sigma = 30$). Results of processing image (b), using **(d)** our method, **(e)** adaptive filtering [14], and **(f)** denoising followed by[12]; a strategy considered in [15, 16]. **(g)–(i)** Results with same 3 methods for processing image (c).

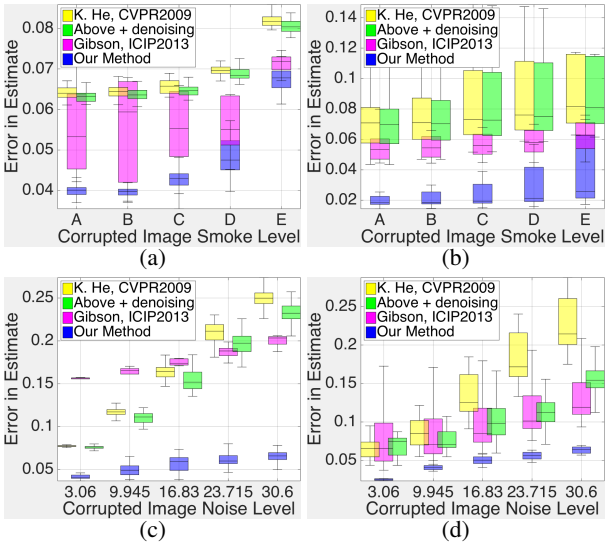


Fig. 5. **Validation on Simulated and High-Quality Laparoscopy Images.** For a fixed moderate noise level, *varying smoke levels* (low to high) with 50 simulated runs per level, box plots for RRMSE (defined in text) for the **(a)** phantom, **(b)** a set of 10 high-quality laparoscopic images. For a fixed moderate smoke level, *varying noise levels* (low to high) with 50 simulated runs per level, box plots for RRMSE for the **(c)** phantom, **(d)** a set of 10 high-quality laparoscopic images.

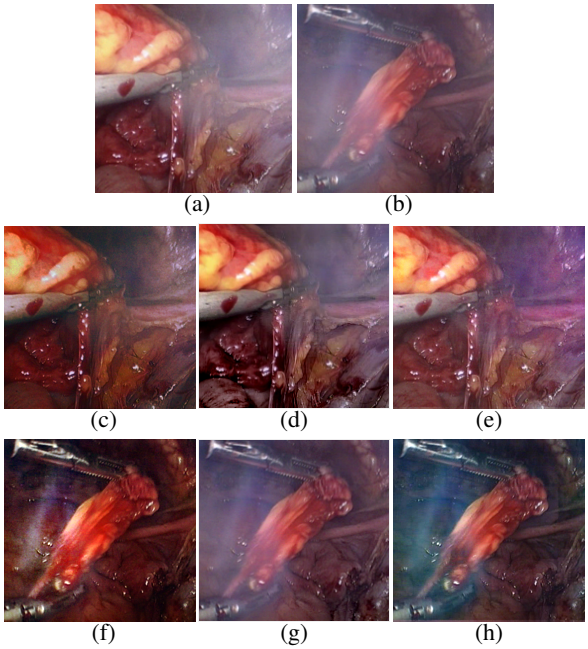


Fig. 6. **Results on Corrupted Laparoscopic Images. (a),(b)** Corrupted images. Results of processing the corrupted images, using **(c),(f)** our method, **(d),(g)** adaptive filtering [14], and **(e),(h)** denoising followed by dehazing [12]; a strategy considered in [15, 16].

bilateral filter [22]) followed by dehazing [12], a strategy considered in [15, 16]. We measure performance using the relative root mean square error (RRMSE): $\| \widehat{\mathbf{x}} - \mathbf{x} \|_2 \; / \; \| \mathbf{x} \|_2$ between the ground truth $\mathbf{x}$ and its estimate $\widehat{\mathbf{x}}$. Qualitative (Figure 3) and quantitative (Figure 5(a),(c)) evaluation shows that our method consistently performs better at preserving natural color and contrast, as well as smoke and noise reduction.

Validation on High-Quality Laparoscopic Images. We select 10 high-quality images with negligible smoke or noise, consider them as ground truth, corrupt them, and evaluate the performance of the 3 methods, qualitatively (Figure 4) and quantitatively (Figure 5(b),(d)). The proposed method leads to a significantly superior contrast, truer colors, and smaller RRMSE values for the estimated uncorrupted image.

Results on Corrupted Laparoscopic Images. We select laparoscopic images with a high amount of surgical smoke and noise. Figure 6 shows the estimated uncorrupted images from the 3 different methods, where the proposed method desmokes better. While the results in Figure 6(e),(h) introduce some artificial bluish hues and degrade image contrast, our improved performance stems from the proposed image-color priors and a reliable estimation of S.

Robustness to Parameter Tuning. The number of free parameters is (i) 5 in the sequential dehazing and denoising method, (ii) 3 (α, β, γ) in our method, and (iii) 1 in adaptive filtering. We use high-quality laparoscopic images corrupted with moderate smoke and low noise for evaluating the sensitivity of the results to parameter tuning by perturbing the parameter values around their empirically-tuned values (SD of 10%). The mean and SD of the resulting RRMSE values for (i) our method were 0.002 and 0.0007, which were far lower than those for (ii) the next best method, denoising followed by dehazing [12], i.e., 0.05 and 0.003.

Clinical Evaluation. We perform a blind evaluation by 4 experts who rated each processed image on a scale of 1 (poor) to 5 (excellent). The mode, median, mean, and SD of the ratings for (i) our method were 5, 5, 4.4, and 0.8, and (ii) the next best method, i.e., denoising followed by dehazing [12], were 3, 3, 3.1, and 0.8, respectively. Our method gives statistically significantly better ratings; p-value $< 10^{-15}$ for a Pearson's chi-square 2-sample test on the categorical distributions.

4. CONCLUSION

We formulated laparoscopic image desmoking and denoising as a novel Bayesian inference problem, using (i) a probabilistic graphical model, (ii) priors on color PDFs (to preserve natural color and contrast) and edge-preserving priors (to reduce noise) on the uncorrupted image and the transmission map, and (iii) a robust estimation scheme for corruption color. Qualitative and quantitative results on laparoscopic images, along with clinical evaluation, show superior performance of our method over the state of the art. Future work will explore better image models [23, 24] as priors.

5. REFERENCES

[1] W Barrett and S Garber, "Surgical smoke: A review of the literature. Is this just a lot of hot air?," *Surgical Endoscopy*, vol. 17, no. 6, pp. 979–87, 2003.

[2] R Silva, D Sehrt, W Molina, J Moss, S Park, and F Kim, "Significance of surgical plume obstruction during laparoscopy," *J. Soc. Laparoendoscopic Surgeons*, vol. 18, no. 3, pp. 269–71, 2014.

[3] K Weld, S Dryer, C Ames, K Cho, C Hogan, M Lee, P Biswas, and J Landman, "Analysis of surgical smoke produced by various energy-based instruments and effect on laparoscopic visibility," *J. Endourology*, vol. 21, no. 3, pp. 347–51, 2007.

[4] K Prokopetc, T Collins, and A Bartoli, "Automatic detection of the uterus and fallopian tube junctions in laparoscopic images," in *Info. Proc. Med. Imag. (IPMI)*, 2015, pp. 552–63.

[5] S Voros, E Orvain, P Cinquin, and J Long, "Automatic detection of instruments in laparoscopic images: a first step towards high level command of robotized endoscopic holders," in *IEEE Int. Conf. Biomed. Robotics Biomechatronics*, 2006, pp. 1107–12.

[6] R Wolf, J Duchateau, P Cinquin, and S Voros, "3D tracking of laparoscopic instruments using statistical and geometric modeling," in *Proc. Med. Imag. Comp. Comput. Assist. Interv. (MICCAI)*, 2011, pp. 203–210.

[7] L Maier-Hein, A Groch, Alberto Bartoli, S Bodenstedt, G Boissonnat, P-L Chang, NT Clancy, DS Elson, S Haase, E Heim, Joachim Hornegger, P Jannin, H Kenngott, T Kilgus, B Muller-Stich, D Oladokun, S Rohl, TR Dos Santos, H-P Schlemmer, A Seitel, S Speidel, Michael Wagner, and D Stoyanov, "Comparative validation of single-shot optical techniques for laparoscopic 3-D surface reconstruction," *IEEE Trans. Med. Imag. (TMI)*, vol. 33, no. 10, pp. 1913–1930, 2014.

[8] L Maier-Hein, P Mountney, A Bartoli, H Elhawary, D Elson, A Groch, A Kolb, M Rodrigues, J Sorger, S Speidel, and D Stoyanov, "Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery," *Med. Imag. Anal. (MedIA)*, vol. 17, no. 8, pp. 974–996, 2013.

[9] P Mountney, J Fallert, S Nicolau, L Soler, and P Mewes, "An augmented reality framework for soft tissue surgery," in *Int. Conf. Med. Imag. Comp. Comput. Assist. Interv. (MICCAI)*, 2014, pp. 423–431.

[10] R Fattal, "Single image dehazing," *ACM Trans. Graphics*, vol. 34, no. 1, 2008.

[11] R Tan, "Visibility in bad weather from a single image," in *Proc. IEEE Int. Conf. Comp. Vis. Pattern Recog. (CVPR)*, 2008, pp. 1–8.

[12] K He, J Sun, and X Tang, "Single image dehazing using dark channel prior," in *Proc. IEEE Int. Conf. Comp. Vis. Pattern Recog. (CVPR)*, 2009, pp. 1956–63.

[13] J Pang, O Au, and Z Guo, "Improved single image dehazing using guided filter," in *Asia Pacific Signal Info. Proc. Assoc. Annual Summit and Conf.*, 2011.

[14] K Gibson and T Nguyen, "Fast single image fog removal using the adaptive Wiener filter," in *Proc. IEEE Int. Conf. Imag. Proc. (ICIP)*, 2013, pp. 714–18.

[15] N Joshi and M Cohen, "Seeing Mt. Rainier: Lucky imaging for multi-image denoising, sharpening, and haze removal," in *Proc. IEEE Int. Conf. Comp. Photography (ICCP)*, 2010, pp. 1–8.

[16] E Matlin and P Milanfar, "Removal of haze and noise from a single image," in *SPIE Optics Photonics*, 2012.

[17] K Nishino, L Kratz, and S Lombardi, "Bayesian defogging," *Int. J. Comp. Vis. (IJCV)*, vol. 98, pp. 263–278, 2012.

[18] S. Z. Li, *Markov Random Field Modeling in Computer Vision*, Springer, 1995.

[19] Zhong Tao and Hemant D. Tagare, "Tunneling descent level set segmentation of ultrasound images," in *Info. Proc. Med. Imag. (IPMI)*, 2005, pp. 750–761.

[20] K Oshio, H Shinmoto, and R Mulkern, "Interpretation of diffusion MR imaging data using a gamma distribution model," *Magn. Reson. Med. (MRM)*, vol. 13, no. 3, pp. 191–195, 2014.

[21] B Huo and F Yin, "Probabilistic tracking of affine-invariant anisotropic regions," *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, vol. 35, no. 1, pp. 130–43, 2013.

[22] C Tomasi and R Manduchi, "Bilateral filtering for gray and color images," in *Proc. Int. Conf. Comp. Vision (ICCV)*, 1998, pp. 839–846.

[23] S P Awate and R T Whitaker, "Unsupervised, information-theoretic, adaptive image filtering for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, vol. 28, no. 3, pp. 364–376, March 2006.

[24] S P Awate and R T Whitaker, "Feature-preserving MRI denoising: A nonparametric empirical-Bayes approach," *IEEE Trans. Med. Imag. (TMI)*, vol. 26, no. 9, pp. 1242–1255, 2007.