Laparoscopy Image Enhancement

A Dissertation

Submitted in partial fulfillment of
the requirements for the degree of

Master of Technology
by

Ayush Baid

(12D100002)

Supervisors:

Prof. S. N. Merchant

and

Prof. S.P. Awate

Department of Electrical Engineering
Indian Institute of Technology Bombay

18 June 2017

Dedicated to my parents

Approval Sheet

This dissertation entitled "Laparoscopy Image Enhancement" by Ayush Baid is approved for the degree of Master of Technology.

_____

_____

_____

Examiners

_____

_____

_____

Supervisor (s)

_____

Chairman

Date: _____

Place: _____

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I declare that I have properly and accurately acknowledged all sources used in the production of this report. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Ayush Baid

Date: 18 June 2017

(12D100002)

Abstract

Laparoscopy images suffer from artifacts like surgical smoke, specular highlights, and noise. These artifacts hinder visibility, and degrade post processing (e.g. segmentation). There is a lack of literature on smoke removal in laparoscopy images, and to the best of our knowledge, no prior work on jointly tackling these three artifacts. We tackle these degradations as a novel *unified Bayesian inference problem*. We use probabilistic graphical models and employ sparse dictionary priors and novel image intensity distribution priors. We obtain maximum-apriori probability (MAP) estimate by *expectation maximization* and use variational Bayesian factorization to overcome the encountered intractabilities. Results on simulated and real-world laparoscopy images show that our joint optimization strategy outperforms the state-of-the-art.

Index terms — Laparoscopy, smoke removal, specular highlights removal, denoising, variational Bayes, EM, graphical models, distribution matching.

Table of Contents

List of Figures

Chapter 1

Introduction

Laparoscopy is a popular *minimally invasive surgery* technique in which operations are performed by inserting equipments through small incisions. Laparoscopy surgery has advantages like less pain and hemorrhaging, shorter recovery times over open surgical procedures. The key equipment is a *laparoscope*, an optical imaging instrument which relays the visuals on a screen. Another main equipment is a cold light source to illuminate the area of operation.

The closed nature of laparoscopy surgery presents some challenges. The images can get severely corrupted with specular highlights [3, 4], surgical smoke [5], and noise. Specular highlights result from strong reflection of the light source by body fluids like blood and mucus. Speckles interfere with post-processing like segmentation [6, 7] and tracking [8]. Electrical cauterization of a tissue generates surgical smoke, which hinders visibility for surgeons and robots alike. Noise is present in all optical imaging systems and a laparoscope is no exception.

Chapter 2 covers the related literature in these areas. Recent works in specular highlight removal in laparoscopy use isotropic diffusion and do not preserve texture and edges. There is lack of literature for smoke removal in laparoscopy images, and hence we look into haze removal in outdoor images. These methods do not employ any special properties of laparoscopy images, which can be imposed due to the smaller variation in this class of images.

In Chapter 3, we will introduce the variables in the system and the image formation model. We use probabilistic graphical models to design prior on the variables. The prior has components like sparse coding on a dictionary, and a novel probability distribution matching penalty. We formulate a unified Bayesian inference problem in Chapter 4, which is solved using the expectation-maximation (EM) algorithm. We introduce variational Bayesian approximations to overcome the analytical intractability in the optimization scheme.

We present validation metrics and results of our algorithm compared to the state-of-the-art in Chapter 5. We conclude and discuss some future directions in Chapter 6.

Chapter 2

Literature Survey

To the best of our knowledge, no existing work tackles smoke, specular highlights, and noise in a joint setting. We will cover these three and some related problem separately. First, we will look into specular highlights removal in laparoscopy images, which is mostly tackled as inpainting problem using some form of averaging. Inpainting is a process of filling in missing information, usually using true information in the surroundings. Then, we will cover dehazing, which is haze removal in outdoor images and bears similarity with desmoking in laparoscopy images. This will be followed with desmoking itself. We will not cover denoising as an independent domain.

2.1 Speckle removal in laparoscopy images

Arnold et al. [9] use a 2-step inpainting process. In the first step, they fill in the missing data with the centroid of available data in the local surroundings and perform strong smoothing using a Gaussian kernel. The smooth image output of the first step and the original image is combined using a weight mask in the second step. The weight mask has high weights near the speckles and decays non-linearly with distance. This results in a gradual transition between original image and the smooth median filtered image. The results however, are smooth and lack texture. This is expected because median filtering is not suitable to interpolate texture.

Isotropic color diffusion is used by Saint-Pierre et al. [4]. They use discrete convolutions with a kernel repeatedly until convergence is reached. The results, however, do not maintain edges and sharp texture. This is expected from an isotropic diffusion process. Stoyanov and Yang [3] use temporal non-rigid registration to obtain pixel values lost due to speckles. The location of speckles shift with time, and hence missing data in one frame may be present in other frames. Interpolation is performed using control points obtained after registration of frames captured at different instances. This method is non-temporal in nature, and also the averaging leads to over-smooth inpainting.

2.2   Dehazing in outdoor images

Outdoor images, particularly of landscapes are often plagued by haze. Haze can be natural (fog) or artificial due to pollution. Haze corrupts the color of image, and when present in large concentration, it can completely obscure the subjects. The effect of haze is modeled by a linear combination of object's radiance and haze color [10]. The following equation, ubiquitous in literature, captures the effect of haze

$$Y_i = T_i X_i + (1 - T_i) A \qquad (2.1)$$

where at pixel location i, Y_i is observed image pixel, $T_i \in [0, 1]$ is the haze transmission coefficient, X_i is radiance of the scene sans haze, and A is the airlight (considered constant for all pixels).

An important property which is exploited quite often is that the haze transmission coefficient $\{T_i\}$ is directly proportional to scene depth, and is hence spatially smooth. Fattal [11] uses Markov random field (MRF) to model the transmission map. Squared difference for four nearest-neighbors for each pixel location is penalized to enforce spatial regularity. Along with MRF for spatial smoothness, Tan [12] also increases image contrast by optimizing for the number of edges in the image. Both the methods do not utilize any information about the distribution of colors in the image.

He et al. [1] observe a statistical property that most local patches in outdoor haze-free images contain some pixel that has low intensity in at least one color channel. Infact, the lowest intensity in a local patch is called the *dark channel* and serves as an estimate for the transmission coefficient at that location. Soft matting is used to obtain a smooth final estimate of transmission map in [1]. Pang et al. [13] use adaptive patch size and replace the soft matting step with guided filtering. Gibson and Nguyen [2] calculate the dark channel and then apply adaptive wiener filter to smooth out the transmission map. Matlin and Milanfar [14] argue that the *dark channel* will be susceptible to outliers resulting from noise. They propose an iterative non-parametric kernel regression. The optimization is performed by alternating between minimization in terms of the transmission map and the uncorrupted image estimate to obtain simultaneous dehazing and denoising. While the dark channel principle is also applicable to laparoscopy images, we can derive more stringent properties to work with. For example, in laparoscopy images, the red channel usually contains the maximum value at each pixel out of the three color channels.

124 images are used by Joshi and Cohen [15] to generate a final image performing weighted averaging of transitionally aligned images. Mt. Rainer, the subject of interest in the paper, has large white glaciers which do not obey the local dark channel property. They also assume airlight constant for a scan-line and not for the whole image and compute the dark channel value per scan-line. This is then used to compute the transmission map and dehaze the image. This method is impractical for laparoscopy as it requires large number of images for a subject.

2.3   Desmoking in laparoscopy images

Kotwal et al. [16] use MRFs to model the image and smoke transmission map and perform joint optimization for smoke and noise removal. They penalize distance between the empirical distribution and template distribution of image intensities and combine it with Huber penalty in local patches to create a prior on the uncorrupted image. Spatial smoothness prior is used on transmission map.

Chapter 3

Formulation

We start this chapter with the concept of sparse coding, which will be used in our design of prior probability distribution on the uncorrupted image estimate $\mathbf{X}$. We then introduce a new colorspace and introduce a novel penalty function which matches two cumulative distribution functions (CDFs). We then design an MRF prior on the $\mathbf{X}$. This will be followed by the image formation model and then design of the remaining variables in the system.

Our notation scheme is to use upper case for random variables and lower case for their specific instance. Bold fonts denote a set of random variables over the whole image, and normal fonts denote an element from that set.

3.1   Sparse coding over a dictionary

Given a data vector W, a linear decomposition over the dictionary $\mathbf{D}$ in terms of coding S is expressed as

$$W \approx \mathbf{D}S \qquad (3.1)$$

In sparse coding [17], we want to represent an input vector using as few atoms of the dictionary as possible, i.e. S should have only a few entries with significant magnitude. Olshausen and Field [18] demonstrated that the dictionary atoms learnt via sparse coding bears similarity to processing in brain's visual cortex by learning features which are localized, oriented, and bandpass. A common way to enforce sparsity is to regularize the codes. A penalty function which uses regularization is of the form

$$J(\mathbf{D}, S) = ||W - \mathbf{D}S||_2^2 + \lambda \sum_j f(S_j) \tag{3.2}$$

where λ is called the regularization constant and controls the trade-off between sparsity and reconstruction accuracy, and f typically is a strictly increasing function of the absolute value of its argument and hence favor sparsity.

Sparse coding using L1 regularization, i.e. $f(S_j) = |S_j|$, is used to perform digit recognition in [19]. Non-negative sparse coding [20] constrains the dictionary and codes to be non-negative. The choice of sparsity function is $f(S_j) = |S_j| = S_j$. Mairal et al. [21] propose an online algorithm for dictionary learning in sparse coding framework using L1 regularization. They use least angle regression to solve the sparse coding problem for a given dictionary. Their algorithm and accompanying code provide a significant speedup for learning a dictionary on large datasets.

The L1 regularization problem can be approximated by an iterative reweighted least square optimization algorithm. The core of the iterative scheme is a weighted least square problem, which is described in the following text. Given a dictionary $\mathbf{D}$ and input vector W, the weighted least square optimization problem and its solution are

$$S^* = \arg \min_S ||W - \mathbf{D}S||_2^2 + \lambda ||\Gamma S||_2^2 \tag{3.3}$$

$$= (\mathbf{D}^\intercal \mathbf{D} + \lambda \Gamma^\intercal \Gamma)^{-1} \mathbf{D}^\intercal W \tag{3.4}$$

where Γ is the weight matrix. This core optimization scheme is used iteratively in iterative reweighted least squares algorithm, where in each step the weighted L2 regularized least square

problem is solved. It can be used to solve L1 regularized least square problem when the weights in each iteration are chosen to be inverse of the absolute value of the current estimate of the coding [22]. If S^m is the estimate before iteration m, then the next estimate will be

$$S^{m+1} = \arg\min_{S} ||W - \mathbf{D}S||_2^2 + \lambda \sum_j \frac{1}{|S_j^m|} (S_j)^2 \tag{3.5}$$

This iterative method of fixing the weights and solving a weighted L2 problem runs until convergence.

3.2   Image intensity distribution

Surgical smoke is generally gray in color. It perturbs the color in the image and hence the overall distribution of colors. We need a model distribution for pixel intensities of uncorrupted high quality laparoscopy images to compare the observed image with. A distribution in a 3 dimensional colorspace will be the best fit on the data. But it gives rise to high computational complexity, and also no generic distribution fits the laparoscopy dataset. So, we consider modeling a distributions along each of the three color channels independently.

The channels in the *RGB* space exhibit very high correlation and modeling the three channels independently is a poor choice in this colorspace. Hence, we use a data-adaptive decorrelated colorspace [23]. We first transform the data from *RGB* space to *LMS* space, as the latter is closely related to human perception. We then calculate the eigenvectors in the *LMS* space and use them as the basis vectors for the new space, which we call $l\alpha\beta$ space. The final transformation is

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0.3568 & 0.8413 & 0.5304 \\ 0.0760 & -0.2006 & 0.1239 \\ 0.2267 & 0.3574 & -0.6512 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3.6}$$

The cross-correlation between channels in *RGB* space is $[0.96, 0.99, 0.94]$ and that in $l\alpha\beta$ space is $[0.62, 0.63, 0.81]$. The new space has a clear advantage when we model the intensity
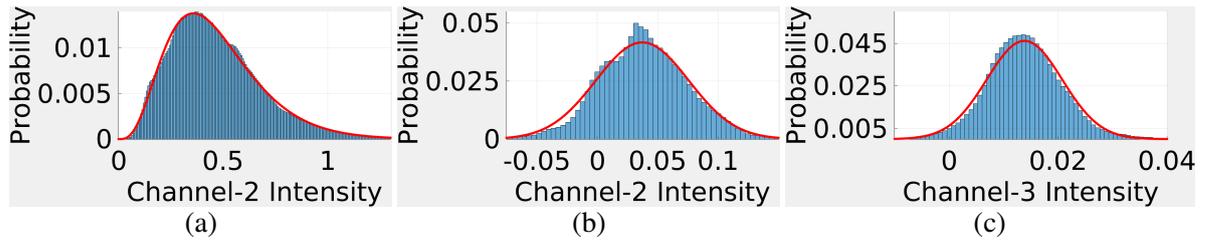
Figure 3.1:   **Learning Prior PDFs on Color.** Empirical histograms (bar plots) and fitted parametric PDFs (solid curves) in uncorrupted laparoscopy images, for 3 channel components: **(a)** gamma Γ_1, **(b)** Gaussian G_2, **(c)** Gaussian G_3.

distribution channel-wise and independent of other channels. We use a Gamma distribution to fit the l channel and Gaussian distributions to fit α and β channels respectively. The empirical distributions and the model fits are illustrated in Figure 3.1

3.3   Kernel density estimation

Kernel density estimation (KDE) is a non-parametric method of estimating the probability density function. Let $\mathbf{w} = (w_1, w_2, ..., w_n)$ be i.i.d. samples from a probability distribution. The kernel density estimate is

$$f^{\mathbf{w}}(b) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{b - w_i}{h}\right) \tag{3.7}$$

where $K(.)$ is a kernel (which takes non-negative values, integrate to one, and has a zero mean), and $h > 0$ is the bandwidth and controls the trade-off between bias and variance of the estimator. We will use a Gaussian kernel due to its mathematical property like differentiability. Due to our choice of Gaussian kernel, we can use the rule of thumb estimate [24] for bandwidth using the sample standard deviation $\hat{\sigma}$. The rule of thumb estimate is

$$h = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}} \tag{3.8}$$

We will now derive the CDF estimator $F^{\mathbf{w}}\left(\cdot\right)$ and its gradients using the Gaussian kernel.

$$f^{\mathbf{w}}(b) = \frac{1}{nh\sqrt{2\pi}} \sum_{i=1}^{n} \exp\left(-\frac{(b-w_i)^2}{2h^2}\right) \tag{3.9}$$

$$F^{\mathbf{w}}(b) = \int_{l=-\infty}^{b} f^{\mathbf{w}}(l)\,\mathrm{d}l \tag{3.10}$$

$$\frac{\partial f^{\mathbf{w}}(b)}{\partial w_i} = \frac{1}{nh\sqrt{2\pi}} \exp\left(-\frac{(b-w_i)^2}{2h^2}\right)\frac{(b-w_i)}{h^2} \tag{3.11}$$

$$\frac{\partial F^{\mathbf{w}}(b)}{\partial w_i} = \int_{l=-\infty}^{b} \frac{\partial f^{\mathbf{w}}(l)}{\partial w_i}\,\mathrm{d}l \tag{3.12}$$

We will formulate a *novel* penalty which compares an empirical CDF $F^{\mathbf{w}}$ with a template CDF F^{ref}. Our formulation is motivated by CDF matching. We map fixed set of points B under the CDF matching transform from empirical to template CDF, and penalize squared difference of the set of output points with the input. When the two CDFs match, the outputs will be the same as input points and hence the penalty will be zero. The penalty function and its gradient with respect to a point w_i in $\mathbf{w}$ are

$$J_{\mathrm{dist}}(\mathbf{w}) = \sum_{b \in B} \left\{b - H\left(b; F^{\mathbf{w}}, F^{\mathrm{ref}}\right)\right\}^2 \tag{3.13}$$

$$\frac{\partial J_{\mathrm{dist}}(\mathbf{w})}{\partial w_i} = 2\sum_{b \in B} \left\{b - H\left(b; F^{\mathbf{w}}, F^{\mathrm{ref}}\right)\right\}\frac{\partial H\left(b; F^{\mathbf{w}}, F^{\mathrm{ref}}\right)}{\partial w_i} \tag{3.14}$$

where $H\left(\cdot\right)$ evaluates the mapping of a point b under CDF matching transform between two distributions. The function will have gradients with respect to w_i via the empirical CDF. The function and its gradient are

$$H\left(b; F^{\mathbf{w}}, F^{\mathrm{ref}}\right) = \left(F^{\mathrm{ref}}\right)^{-1}\left(F^{\mathbf{w}}(b)\right) \tag{3.15}$$

$$\frac{\partial H\left(b; F^{\mathbf{w}}, F^{\mathrm{ref}}\right)}{\partial w_i} = \frac{\partial\left(\left(F^{\mathrm{ref}}\right)^{-1} F^{\mathbf{w}}(b)\right)}{\partial F^{\mathbf{w}}(b)}\frac{\partial F^{\mathbf{w}}(b)}{\partial w_i} \tag{3.16}$$

3.4   MRF model on image X

The uncorrupted image containing I pixels is modeled by an MRF $\mathbf{X} := \{X_i\}_{i=1}^{I}$, where $X_i \in [0,1]^3$ is a vector of *RGB* values at each pixel. We will define potential functions on cliques of this MRF and use it as prior probability on $\mathbf{X}$. We use the sparse coding prior to preserve texture and remove noise, and image intensity distribution prior to enforce color statistics. The neighborhood system $\mathcal{N}^X$ is a fully connected system.

Using the kernel density estimation based penalty function (Equation (3.13)), we define potential function on the clique containing all the pixels in the image. The expression for the same is

$$J_{\text{dist}}(\mathbf{X}) = \sum_{i=1}^{3} \sum_{b \in B_{c_i}} \left\{ b - H\left(b; F_{c_i}^{\mathbf{X}}, F_{c_i}^{\text{ref}}\right) \right\}^2 \tag{3.17}$$

where B_{c_i} is a fixed set of points in channel c_i. The colorspace of operation is $l\alpha\beta$ introduced in Section 3.2.

We will use L1 regularized sparse coding on full patches (i.e. patches fully contained in the image) of size $\acute{m} \times \acute{m}$ of $\mathbf{X}$ as the second penalty function. We define potentials on each square clique of size $m \times m$, where $m := 2\acute{m} - 1$. This translates to a sparse coding cost penalty on full patches $\{\bar{\mathbf{X}}_i\}_{i=1}^{I'}$, which is defined as

$$J_{\text{dict}}(\mathbf{X}, \mathbf{S}) = \sum_{i=1}^{I'} \|\bar{\mathbf{X}}_i - \mathbf{D}S_i\|_2^2 + \lambda\|S_i\|_1 \tag{3.18}$$

where $\mathbf{D} := \{D_j\}_{j=1}^{J}$ is a fixed dictionary, learnt from a training set. The value of λ is fixed by a heuristic mentioned in [21].

The prior probability on $\mathbf{X}$ is defined as

$$P(\mathbf{X}) = \frac{\int_{\mathbf{S}} \exp\left(-E\left(\mathbf{X}, \mathbf{S}\right)\right) d\mathbf{S}}{Z} \tag{3.19}$$

where $E\left(\mathbf{X}, \mathbf{S}\right)$ is the Gibbs energy function, given by

$$E\left(\mathbf{X}, \mathbf{S}\right) = \gamma_1 J_{\text{dict}}\left(\mathbf{X}, \mathbf{S}\right) + \gamma_2 J_{\text{dist}}\left(\mathbf{X}\right) \tag{3.20}$$

and Z is the normalization constant. $\gamma_1, \gamma_2 \in \mathbb{R}^+$ are free parameters.

3.5 Image formation

The artifacts are captured in a 3 step fashion. In this section, i denotes the pixel location. Equation (3.21) captures the effect of speckles using a binary speckle map $\mathbf{R}$, where $R_i = 1$ denotes the presence of speckle at pixel i. Equation (3.22) captures the smoke using transmission coefficient $T_i \in [0, 1]$. To capture noise, an i.i.d. zero mean Gaussian η_i with standard deviation σ is added at each pixel. $\mathbf{Y}$ is captured and available for processing. The equations are

$$Z_i = (1 - R_i)X_i + R_i K_{\text{spec}} \tag{3.21}$$

$$Y_i = T_i Z_i + (1 - T_i)K_{\text{smoke}} + \eta_i \tag{3.22}$$

where K_{spec} and K_{smoke} are speckle and smoke color respectively.

Over all pixels, $\mathbf{X}$ is an uncorrupted image, which we want to estimate. $\mathbf{Y}$ is the observed image, $\mathbf{T}$ is the smoke transmission map, and $\mathbf{R}$ is the speckle map. K_{spec} and K_{smoke} are considered constant across pixels.

3.6 Prior on transmission map T

The MRF $\mathbf{T} := \{T_i\}_{i=1}^{I}$ is defined with a neighborhood system $\mathcal{N}^T := \{\mathcal{N}_i^T\}$, where $\mathcal{N}_i^T$ is the set of pixels in 5×5 patch centered at pixel i. To enforce spatial smoothness, we will penalize deviations in a local neighborhood. The prior distribution is defined using potentials on all the

cliques as follows

$$P(\mathbf{T}) = \frac{1}{Z} \exp\left(-\gamma_3 \sum_{i=1}^{I} \sum_{j \in \mathcal{N}_i^T} (T_i - T_j)^2\right) \tag{3.23}$$

where the outer sum is performed over all pixels, $\gamma_3 \in \mathbb{R}^+$ is a free parameter, and Z is the normalization constant.

Chapter 4

Estimation

For our algorithm, we assume that we are provided with a speckle label map $\mathbf{r}$, speckle color K_{spec}, and smoke color K_{smoke}. We use maximum aposteriori probability (MAP) estimation scheme for image $\mathbf{X}$. Given the speckle map $\mathbf{r}$, we remove the pixel locations of $\mathbf{y}$ where specular highlights are present, as they convey no information about the underlying image. We call the new image with holes in place of specular highlights as $\mathring{\mathbf{y}}$. The MAP estimate is

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} P\left(\mathbf{x}|\mathring{\mathbf{y}}, \mathbf{r}\right) = \arg\max_{\mathbf{x}} P\left(\mathring{\mathbf{y}}|\mathbf{x}, \mathbf{r}\right) P\left(\mathbf{x}\right) \tag{4.1}$$

The first part of the final term in Equation (4.1) is the likelihood of observing the output $\mathring{\mathbf{y}}$. The second part is the prior probability of the image estimate $\mathbf{x}$.

4.1   Variational Bayesian expectation maximization

We will introduce dictionary coding $\mathbf{S}$ and smoke transmission map $\mathbf{T}$ as latent variables in the system. The posterior probability is

$$P\left(\mathring{\mathbf{y}}|\mathbf{x},\mathbf{r}\right)P\left(\mathbf{x}\right) = \int_{\mathbf{s},\mathbf{t}} P\left(\mathring{\mathbf{y}},\mathbf{x},\mathbf{s},\mathbf{t}|\mathbf{r}\right) \mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{t} \tag{4.2}$$

$$= \int_{\mathbf{s},\mathbf{t}} P\left(\mathring{\mathbf{y}}|\mathbf{x},\mathbf{s},\mathbf{t},\mathbf{r}\right)P\left(\mathbf{x},\mathbf{s}\right)P\left(\mathbf{t}\right) \mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{t} \tag{4.3}$$

$$= \int_{\mathbf{s},\mathbf{t}} P\left(\mathring{\mathbf{y}}|\mathbf{x},\mathbf{t},\mathbf{r}\right)P\left(\mathbf{x},\mathbf{s}\right)P\left(\mathbf{t}\right) \mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{t} \tag{4.4}$$

Equation (4.4) follows as $\mathbf{Y}$ is independent of $\mathbf{S}$, given $\mathbf{T}$ and $\mathbf{X}$ due to our image formation model. The final term in Equation (4.4) has three components. The first one is the likelihood of the output being observed. The second and third component are priors on $\mathbf{X}, \mathbf{S}$ and $\mathbf{T}$.

Due to our modeling of the noise as i.i.d. Gaussian, the log likelihood probability distribution turns out to be

$$\log P\left(\mathring{\mathbf{y}}|\mathbf{x},\mathbf{t},\mathbf{r}\right) = -\frac{1}{\sigma^2}\sum_{i|r_i=0}\left(y_i - t_i x_i - (1-t_i)K_{\text{smoke}}\right)^2 \tag{4.5}$$

We will now cover the expectation and maximization steps in the EM algorithm. Variational Bayesian factorization is introduced in the expectation step to overcome intractability.

4.1.1   Expectation step

Let $\mathbf{x}^m$ be the estimate of the uncorrupted image after m iterations. The Q function defined for the next iteration is

$$Q(\mathbf{x};\mathbf{x}^m) = \mathbb{E}_{P(\mathbf{S},\mathbf{T}|\mathring{\mathbf{y}},\mathbf{x}^m,\mathbf{r})}\left[\log P\left(\mathring{\mathbf{y}},\mathbf{x},\mathbf{S},\mathbf{T}|\mathbf{r}\right)\right]$$

$$= \mathbb{E}_{P(\mathbf{S},\mathbf{T}|\mathring{\mathbf{y}},\mathbf{x}^m,\mathbf{r})}\left[P\left(\mathring{\mathbf{y}},\mathbf{T}|\mathbf{x},\mathbf{r}\right)P\left(\mathbf{S}|\mathbf{x}\right)P\left(\mathbf{x}\right)\right] \tag{4.6}$$

The posterior probability over the latent variables can be written as

$$P\left(\mathbf{S}, \mathbf{T}|\mathbf{y}, \mathbf{x}^m, \mathbf{r}\right) = P\left(\mathbf{T}|\mathbf{y}, \mathbf{x}^m, \mathbf{r}\right) P\left(\mathbf{S}|\mathbf{x}^m\right) \tag{4.7}$$

The codes S_i for I' full patches in $\mathbf{X}$ are independent of each other. Hence

$$P\left(\mathbf{S}|\mathbf{x}^m\right) = \prod_{i=1}^{I'} P\left(S_i|\mathbf{x}^m\right) \tag{4.8}$$

The simplified expression for $Q\left(\cdot\right)$ is

$$Q(\mathbf{x}; \mathbf{x}^m) = \mathbb{E}_{P(\mathbf{T}|\mathring{\mathbf{y}}, \mathbf{x}^m, \mathbf{r})} \log\left[P\left(\mathring{\mathbf{y}}, \mathbf{T}|\mathbf{x}, \mathbf{r}\right)\right] + \sum_{i=1}^{I'} \mathbb{E}_{P(S_i|\mathbf{x}^m)}\left[\log P\left(S_i|\mathbf{x}\right)\right] + \log P\left(\mathbf{x}\right) \tag{4.9}$$

The expectations in Equation (4.9) are analytically intractable. We use variational factorization for the latent variables $\mathbf{S}$ and $\mathbf{T}$. The posterior probability for $\mathbf{T}$ is factorized over each pixel in the image, and the posterior probability for $\mathbf{S}$ is factorized over coefficient corresponding to each of the J atoms of dictionary $\mathbf{D}$.

$$P\left(\mathbf{T}|\mathring{\mathbf{y}}, \mathbf{x}^m, \mathbf{r}\right) \approx \prod_{i=1}^{I} G_i^\dagger\left(T_i|\mathring{\mathbf{y}}, \mathbf{x}^m, \mathbf{r}\right) \tag{4.10}$$

$$P\left(S_i|\mathbf{x}^m\right) \approx \prod_{j}^{J} G_{ij}(S_{ij}|\mathbf{x}^m) \tag{4.11}$$

$G_i^\dagger$ is the factor of $\mathbf{T}$ at pixel i, and G_{ij} is the factor corresponding to j^{th} coefficient of code vector S_i. $\{G_i^\dagger\}_{i=1}^I$ are truncated Gaussians with support $[0, 1]$, means $\{\mu_i^{\mathbf{T}}\}$, standard deviations $\{\sigma_i^{\mathbf{T}}\}$. $\{G_{ij}\}_{j=1}^J$ are Gaussian distributions with means $\{\mu_{ij}^{\mathbf{S}}\}$ and standard deviation $\{\sigma_{ij}^{\mathbf{S}}\}$. We have to optimize the factorization before solving for the Q function. Given all other factors, the optimum factor for $G_i^{\dagger*}$ and G_{ij}^* are derived as

$$\log G_i^{\dagger*}(T_i) = \mathbb{E}_{\prod_{k \neq i} G_k^{\dagger*}(T_k|\mathring{\mathbf{y}}, \mathbf{x}^m, \mathbf{r})}\left[\log P\left(\mathring{\mathbf{y}}, \mathbf{T}|\mathbf{x}^m, \mathbf{r}\right)\right] + c_1 \tag{4.12}$$

$$\log G_{ij}^*\left(S_{ij}\right) = \mathbb{E}_{\prod_{k \neq j} G_{ik}^*(S_{ik}|\mathbf{x}^m)}\left[\log P\left(S_i|\mathbf{x}^m\right)\right] + c_2 \tag{4.13}$$

Equation (4.12) and Equation (4.13) are solved iteratively for optimum factors $\{G_i^{\dagger *}\}$ and $\{G_{ij}^*\}$ until convergence. c_1 and c_2 are additive constants which will be absorbed in normalization of factor distributions. The parameters of optimum factors are derived in Appendix A and Appendix B. Once we have the optimum factorization, solving for the Q function becomes simple. We substitue $\mu_i^{\mathbf{T}}$ for T_i, $\left(\mu_i^{\mathbf{T}}\right)^2 + \left(\sigma_i^{\mathbf{T}}\right)^2$ for T_i^2 and $\mu_i^{\mathbf{T}}\mu_j^{\mathbf{T}}$ for T_iT_j, and $\mu_{ij}^{\mathbf{S}}$ for S_{ij} in

$$Q(\mathbf{x};\mathbf{x}^m) = \mathbb{E}_{\prod_{i=1}^{I} G_i^{\dagger *}(T_i|\mathbf{y},\mathbf{x}^m)}\left[\log P\left(\mathbf{y},\mathbf{T}|\mathbf{x},\mathbf{r}\right)\right] + \sum_i \mathbb{E}_{\prod_{j=1}^{J} G_{ij}^*(S_{ij}|\mathbf{x}^m)}\left[\log P\left(S_i|\mathbf{x}\right)\right] + \log P\left(\mathbf{x}\right)$$

$$(4.14)$$

4.1.2   M step

The final expression for $Q\left(\cdot\right)$ is

$$\begin{aligned}
Q\left(\mathbf{x};\mathbf{x}^m\right) = &- 0.5\gamma_1\bar{\mathbf{x}}_i^{\mathsf{T}}\left(\bar{\mathbf{x}}_i - \mathbf{D}\mu_i^{\mathbf{S}}\right) - \log\left(\gamma_2 J_{\text{dist}}\left(\mathbf{x}\right)\right) \\
&- \frac{1}{2\sigma^2}\sum_{\{i|r_i=0\}}\left(\left(\mu_i^{\mathbf{T}}\right)^2 + \left(\sigma_i^{\mathbf{T}}\right)^2\right)\left(x_i - K_{\text{smoke}}\right)^2 \\
&- 2\left(y_i - K_{\text{smoke}}\right)\mu_i^{\mathbf{T}}\left(x_i - K_{\text{smoke}}\right) + \delta'
\end{aligned} \qquad (4.15)$$

where δ' is independent of $\mathbf{x}$.

In this step, we will maximize $Q(\mathbf{x};\mathbf{x}^m)$ with respect to $\mathbf{x}$ to obtain a new estimate $\mathbf{x}^{m+1}$. The next estimate is given by

$$\mathbf{x}^{m+1} = \arg\max_{\mathbf{x}} Q\left(\mathbf{x};\mathbf{x}^{m+1}\right) \qquad (4.16)$$

This optimization problem is solved using projected gradient descent with adaptive step size.

Chapter 5

Results

In this chapter, we compare our methods with the existing literature. To the best of our knowledge, there is no prior work on joint removal of smoke, speckle, and noise in laparoscopy images. We combine multiple methods which solve the subproblems and use them for end-to-end comparison. We use anisotropic diffusion for inpainting which preserves texture better than isotropic diffusion used in laparoscopy imaging [4, 9]. The methods used for comparison are

1. Desmoking and denoising with adaptive wiener filtering by Gibson ICIP13 [2], followed by anisotropic diffusion inpainting.

2. Noise removal with edge preserving bilateral filtering, followed by smoke removal with He PAMI11 [1] and anisotropic diffusion inpainting.

We observe the qualitative aspect of output images such as texture, presence of noise, and accuracy and naturalness of colors in the output. We also use 4 quantitative measures, which are (1) relative root mean square difference (RRMSE); (2)sum of channel-wise mean structure similarity index measure (SSIM) across RGB channels; (3) sum of channel-wise mean quality index based on local variance (QILV) across RGB channels; (4) sum of channel-wise chi-squared distance between histograms along the 3 channels in the LMS colorspace. These metrics cover a wide range of features such as noise level, blur, structure, and color distribution. RRMSE, chi-squared distance should be low and QILV, SSIM should be high for the best method.

We present examples and quantitative metrics by processing synthetically corrupted simulated data and high quality laparoscopy data in the following sections.

5.1 Experiment details

Proposed as well as competing methods are tuned for best performance at 3 percent noise level. Our method has three free parameters (γ_1, γ_2 and γ_3) which are to be tuned. For synthetic corruption, we use 3 transmission coefficient maps and 20 specularity maps. We get 3 times 5 transmission coefficient maps using 5 scalar multiples or smoke levels. We add i.i.d. Gaussian noise of standard deviation ranging from 0 % to 7%.

5.2 Synthetic corruption on simulated data

We use this experiment to provide proof on concept of our method. The simulated data is designed to provide basis of evaluation in terms of image texture and color. We use 6 noise levels, 3 smoke maps, and 5 smoke levels. This results in 90 experiments.

An example of processing on synthetically corrupted phantom is Figure 5.1. Gibson ICIP13 [2] plus inpainting does a poor job at removing smoke. Bilateral filtering, He PAMI11 [1] plus inpainting does a decent job at desmoking, but the inpainting performance is poor. Our dictionary prior does a better job of filling in texture. Proposed method VBEM has the best removal of smoke, better texture and colors.

We perform quantitative benchmarking using 4 metrics. The results are presented in Figure 5.2. Our proposed method has better performance across all metrics. RRMSE and QILV metric in particular demonstrate robustness to high smoke concentration and high noise level.

## 5.3   Synthetic corruption on high quality laparoscopy images

We now perform validation on laparoscopy data. We take high quality laparoscopy images, corrupt them synthetically, and then process using different algorithms and compare the outputs with the ground truth. We run experiments on 6 image, where each image is corrupted with 90 different corruption combinations.

Figure 5.3, Figure 5.4, Figure 5.5, and Figure 5.6 show the results of processing on laparoscopy data. Figure 5.3 and Figure 5.4 demonstrate better noise removal while maintaining sharper texture.Figure 5.5 is an example of better speckle removal by our method. Figure 5.6 demonstrated higher contrast and better colors after smoke removal. These observations are similar to those for simulated data. Gibson ICIP13 [2] does not remove satisfactory levels of smoke. Bilateral filtering, He PAMI11 [1] plus inpainting does a decent job at smoke removal, but the result produces unnatural color in certain examples. Bilateral filter also leads to loss of edges and texture. The inpainting of specular highlights has clear seams in certain regions, and the filled regions lack texture. Proposed method VBEM has the best removal of smoke, better texture and colors. Quantitative evaluation in Figure 5.7 demonstrates that we perform better for each metric.

5.4   Real world laparoscopy images

We present some results of proposed method VBEM and other competing algorithms on real world laparoscopy images. Figure 5.8 demonstrates better texture in our output. Figure 5.9 and Figure 5.10 demonstrate better color accuracy after smoke removal. Finally, we demonstrate some results on speckle removal in Figure 5.11.

From the examples, we observed that our method generates natural colors after removing smoke, preserved texture while removing equal amounts of noise, and performs better filling of speckles.
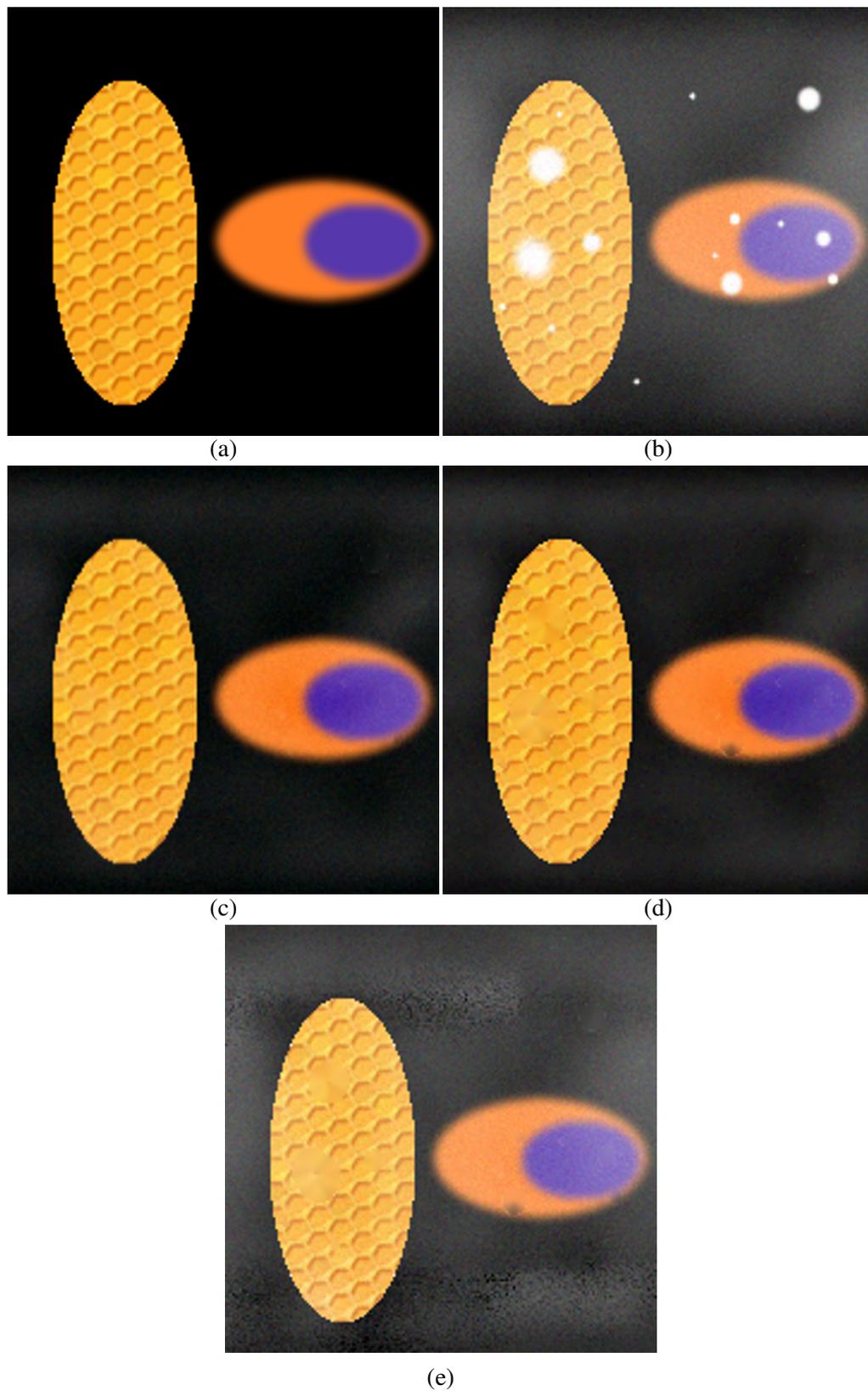
Figure 5.1:   **Validation on Simulated Data.**   **(a)** Phantom (color component values $\in [0, 255]$).
(b) Corrupted phantom with smoke, specularities, and low noise ($\sigma = 5$). Results of processing
image **(b)**, using: **(c)** *our method*; **(d)** bilateral filter denoising followed by dehazing [1] followed
by inpainting; **(e)** adaptive filtering [2] followed by inpainting.
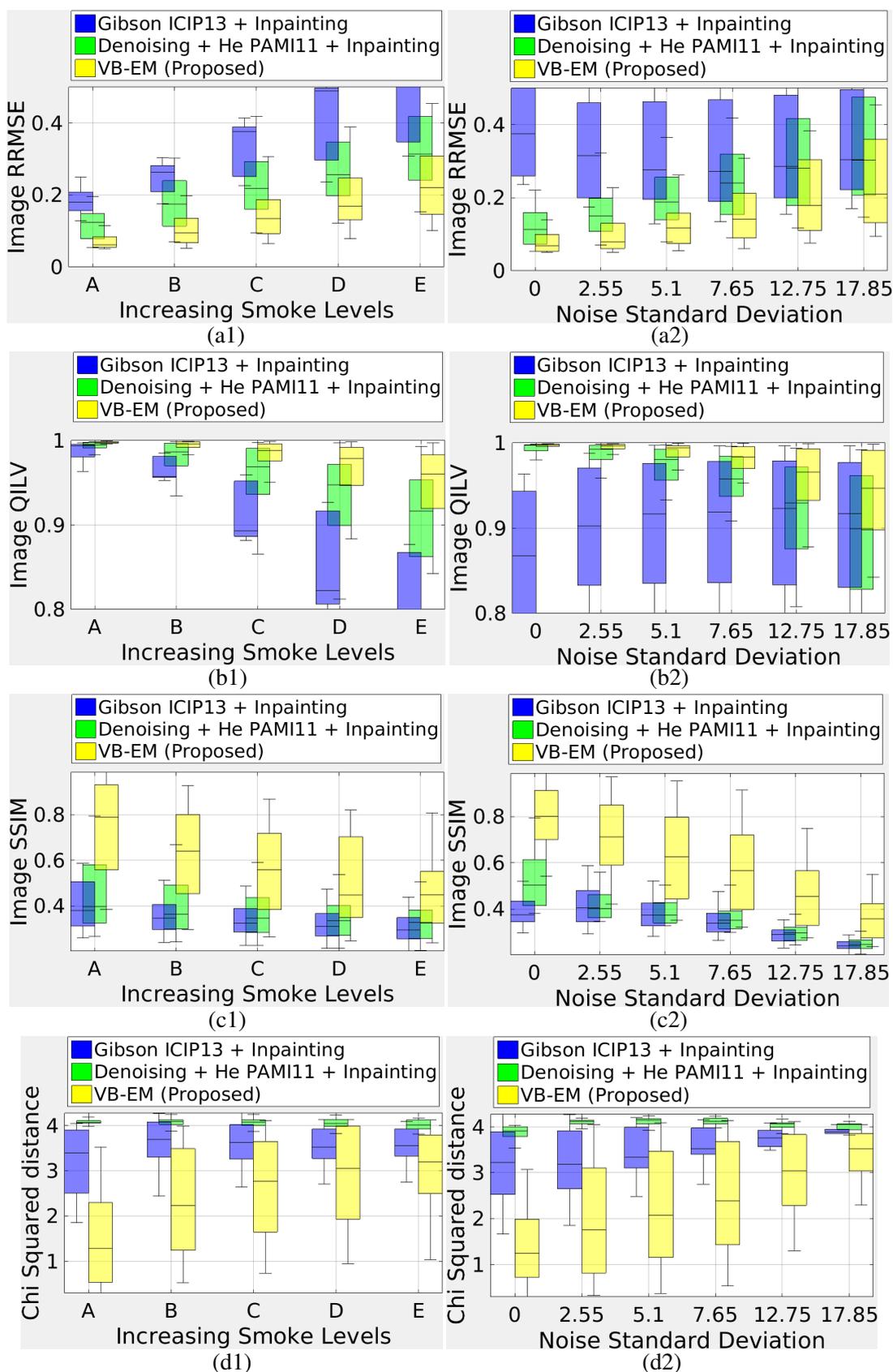
Figure 5.2: **Quantitative Validation on Simulated Data.** Boxplots for RRMSE **(a)**, QILV **(b)**, SSIM **(c)**, and chi-squared distance between histograms **(d)**. In column 1, results are grouped by smoke level and in column 2, grouped by noise level.
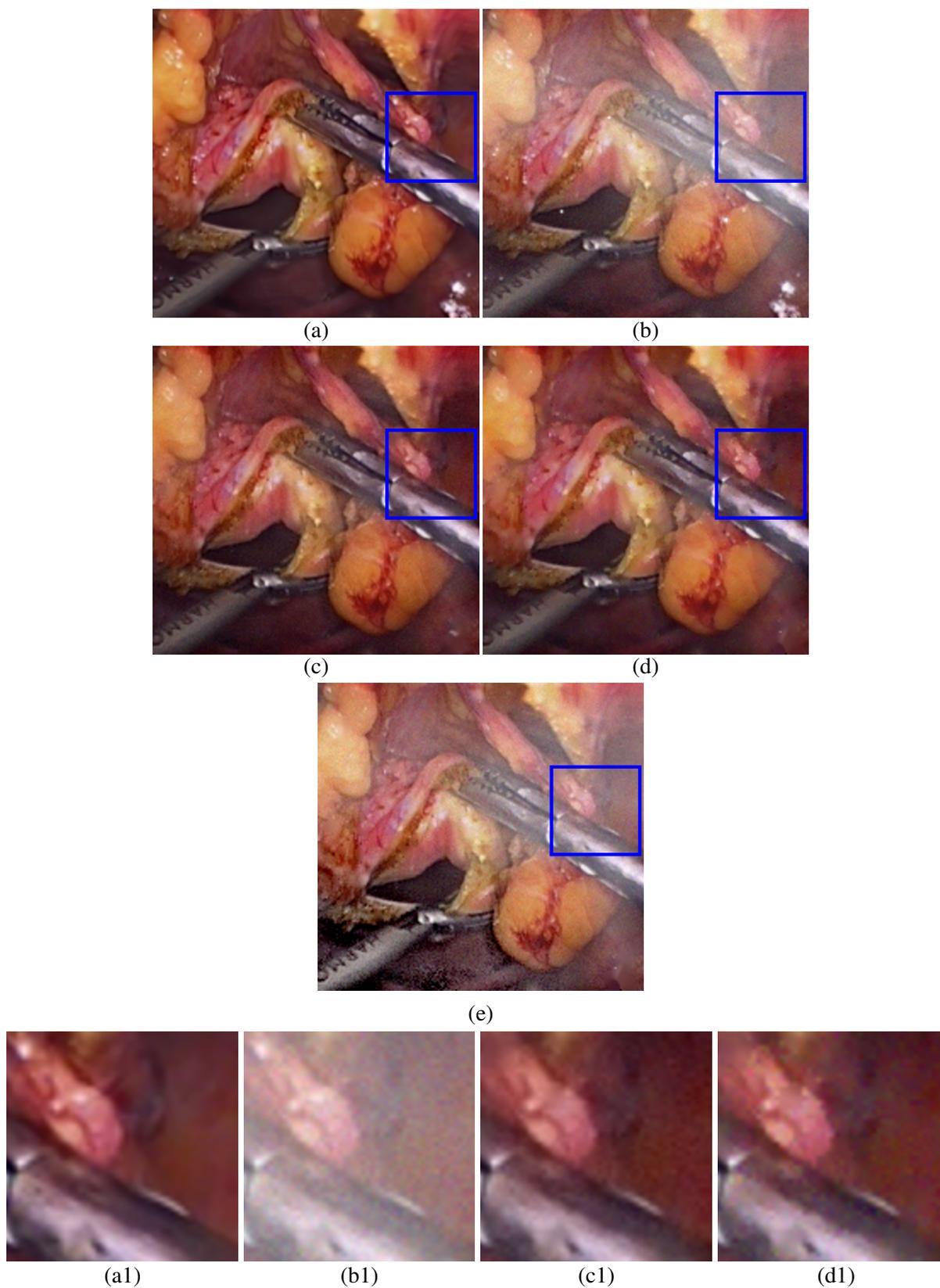
Figure 5.3: **Validation on Laparoscopy Data.** **(a)** High quality laparoscopy image (color component values $\in [0, 255]$). **(b)**: **(a)** corrupted synthetically with smoke, specularities, and moderate noise ($\sigma = 5$). Results of processing image **(b)** using: **(c)** *our method*; **(d)** bilateral filter denoising followed by dehazing [1] followed by inpainting; **(e)** adaptive filtering [2] followed by inpainting. Zoomed insets of **(a)**, **(b)**, **(c)**, **(d)** are **(a1)**, **(b1)**, **(c1)**, **(d1)**.
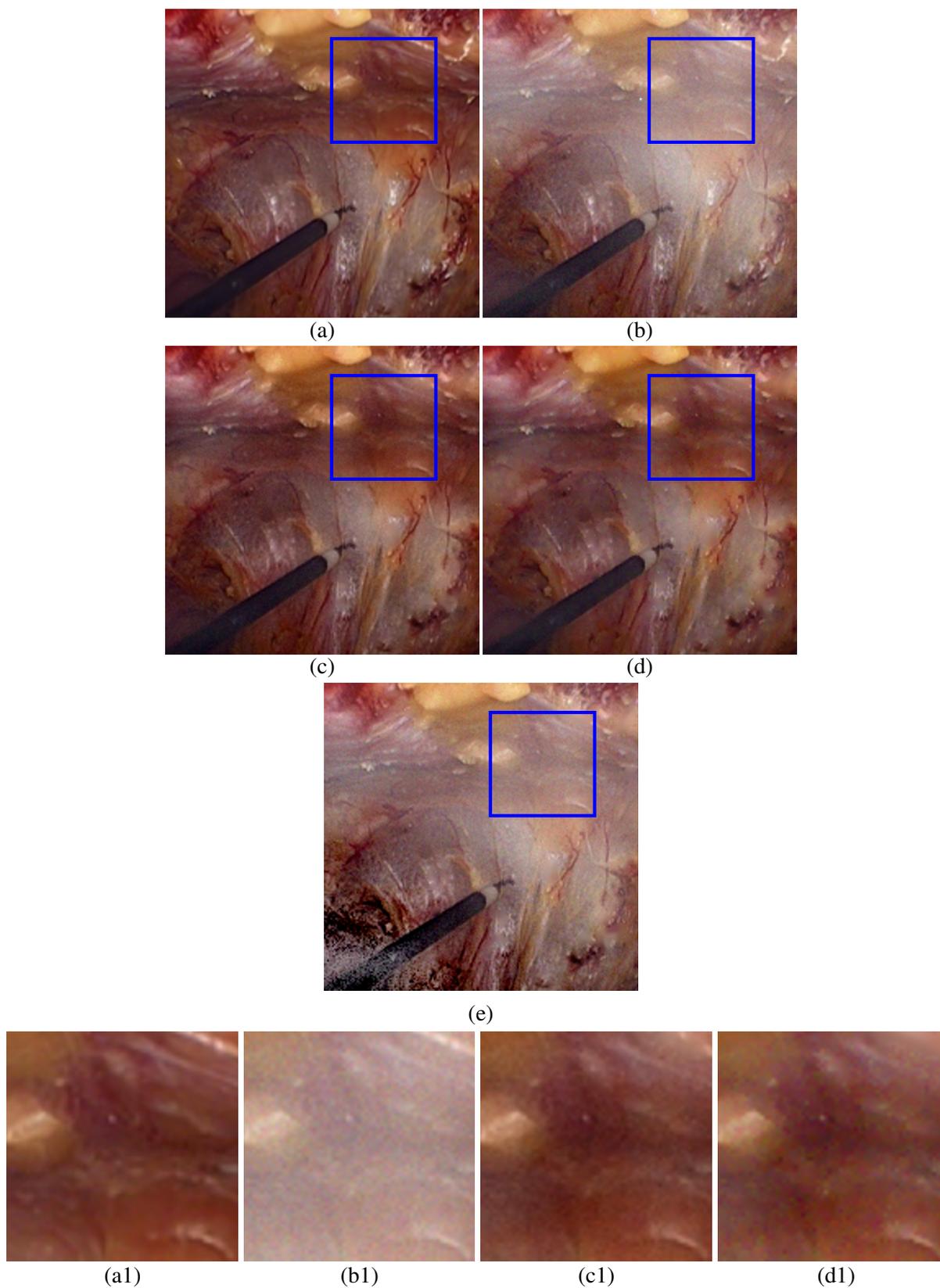
Figure 5.4: **Validation on Laparoscopy Data.** **(a)** High quality laparoscopy image (color component values $\in [0, 255]$). **(b)**: **(a)** corrupted synthetically with smoke, specularities, and moderate noise ($\sigma = 5$). Results of processing image **(b)** using: **(c)** *our method*; **(d)** bilateral filter denoising followed by dehazing [1] followed by inpainting; **(e)** adaptive filtering [2] followed by inpainting. Zoomed insets of **(a)**, **(b)**, **(c)**, **(d)** are **(a1)**, **(b1)**, **(c1)**, **(d1)**.
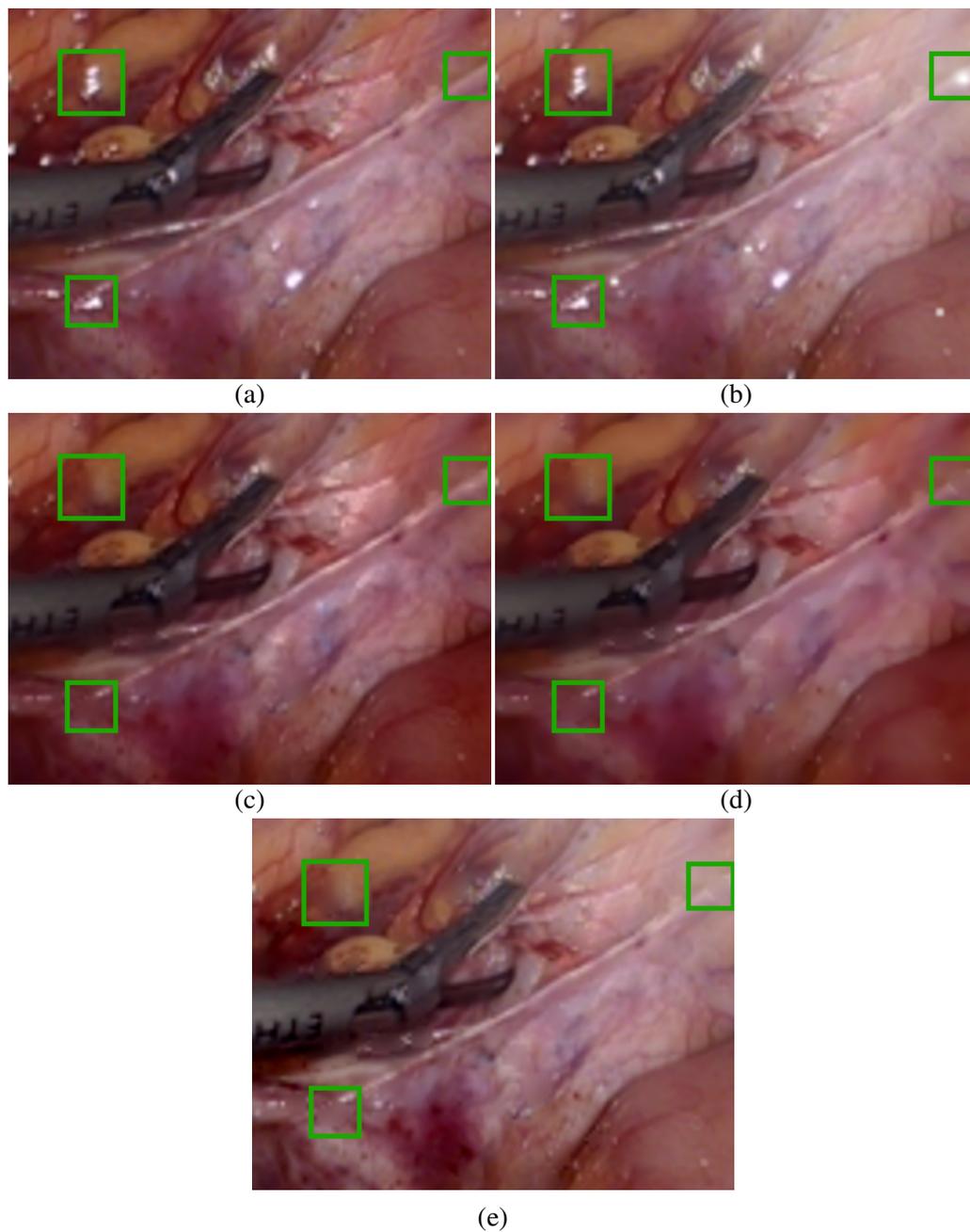
Figure 5.5:   **Validation on Laparoscopy Data.**   **(a)** High quality laparoscopy image (color component values $\in [0, 255]$). **(b)**: **(a)** corrupted synthetically with smoke and specularities. Results of processing image **(b)** using: **(c)** *our method*; **(d)** bilateral filter denoising followed by dehazing [1] followed by inpainting; **(e)** adaptive filtering [2] followed by inpainting.
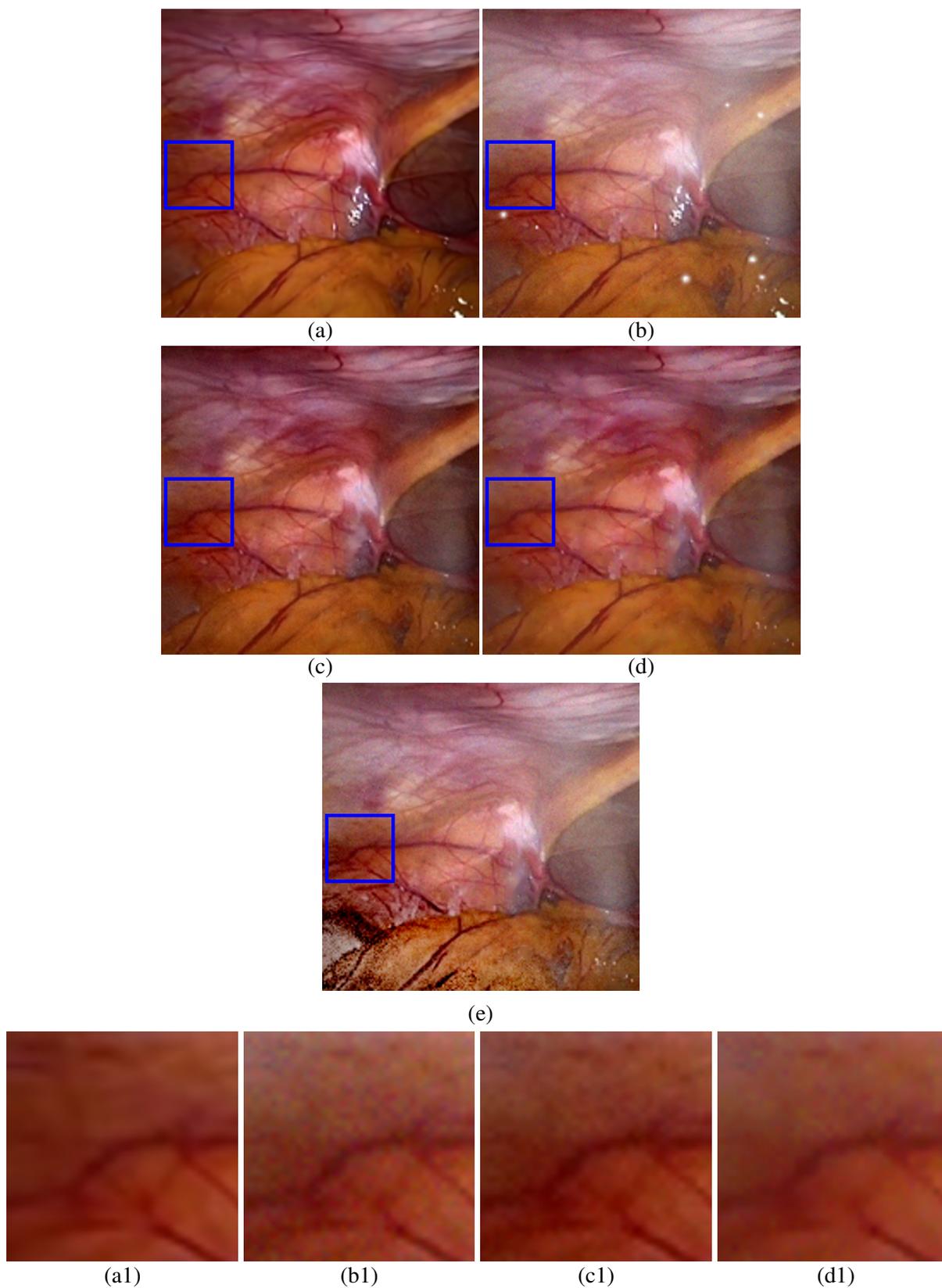
Figure 5.6:   **Validation on Laparoscopy Data.**   **(a)** High quality laparoscopy image (color component values $\in [0, 255]$). **(b)**: **(a)** corrupted synthetically with smoke, specularities, and moderate noise ($\sigma = 5$). Results of processing image **(b)** using: **(c)** *our method*; **(d)** bilateral filter denoising followed by dehazing [1] followed by inpainting; **(e)** adaptive filtering [2] followed by inpainting. Zoomed insets of **(a)**, **(b)**, **(c)**, **(d)** are **(a1)**, **(b1)**, **(c1)**, **(d1)**.
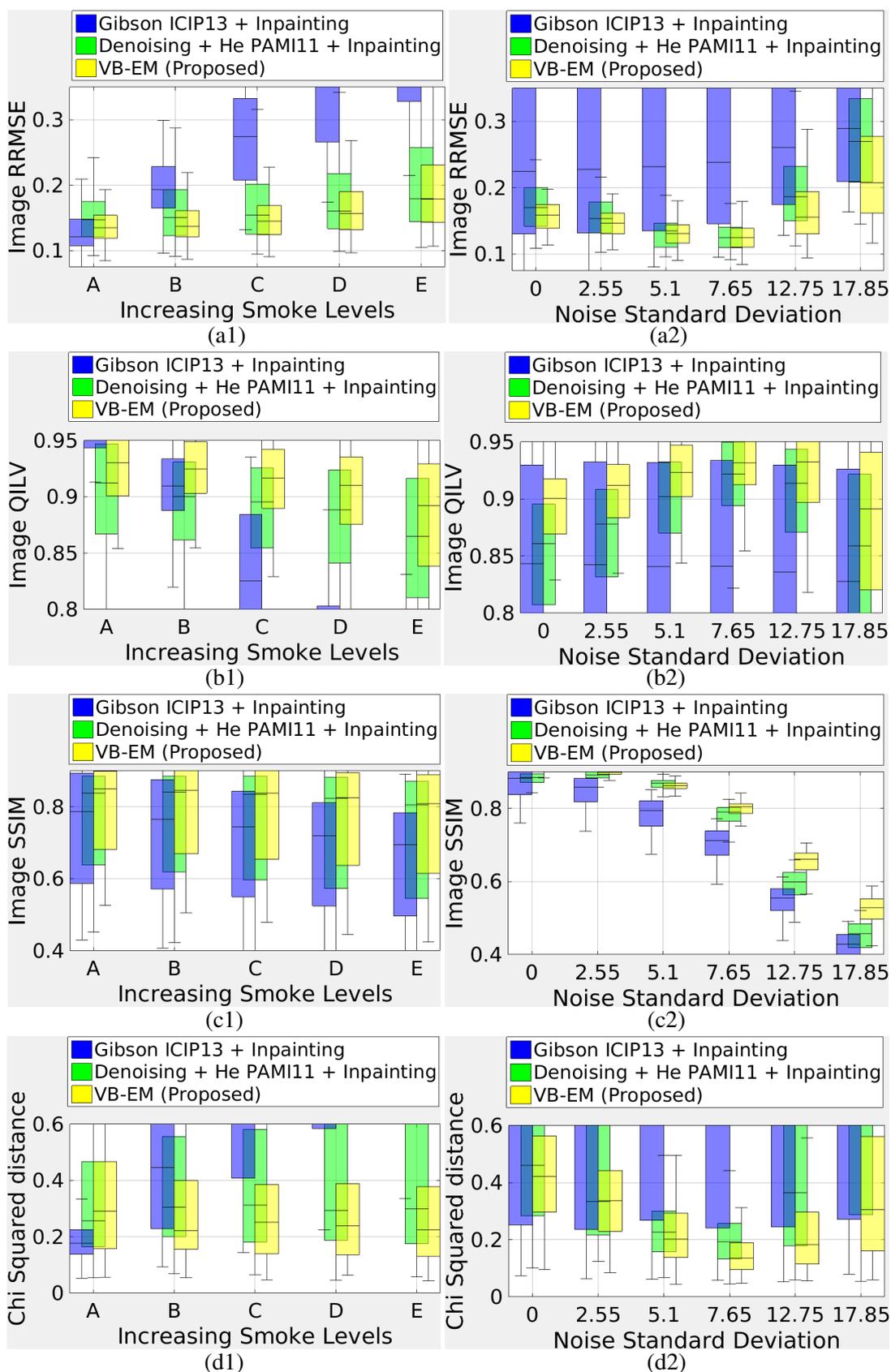
Figure 5.7: **Quantitative Validation on Laparoscopy Data.** Boxplots for RRMSE **(a)**, QILV **(b)**, SSIM **(c)**, and chi-squared distance between histograms **(d)**. In column 1, results are grouped by smoke level and in column 2, grouped by noise level.
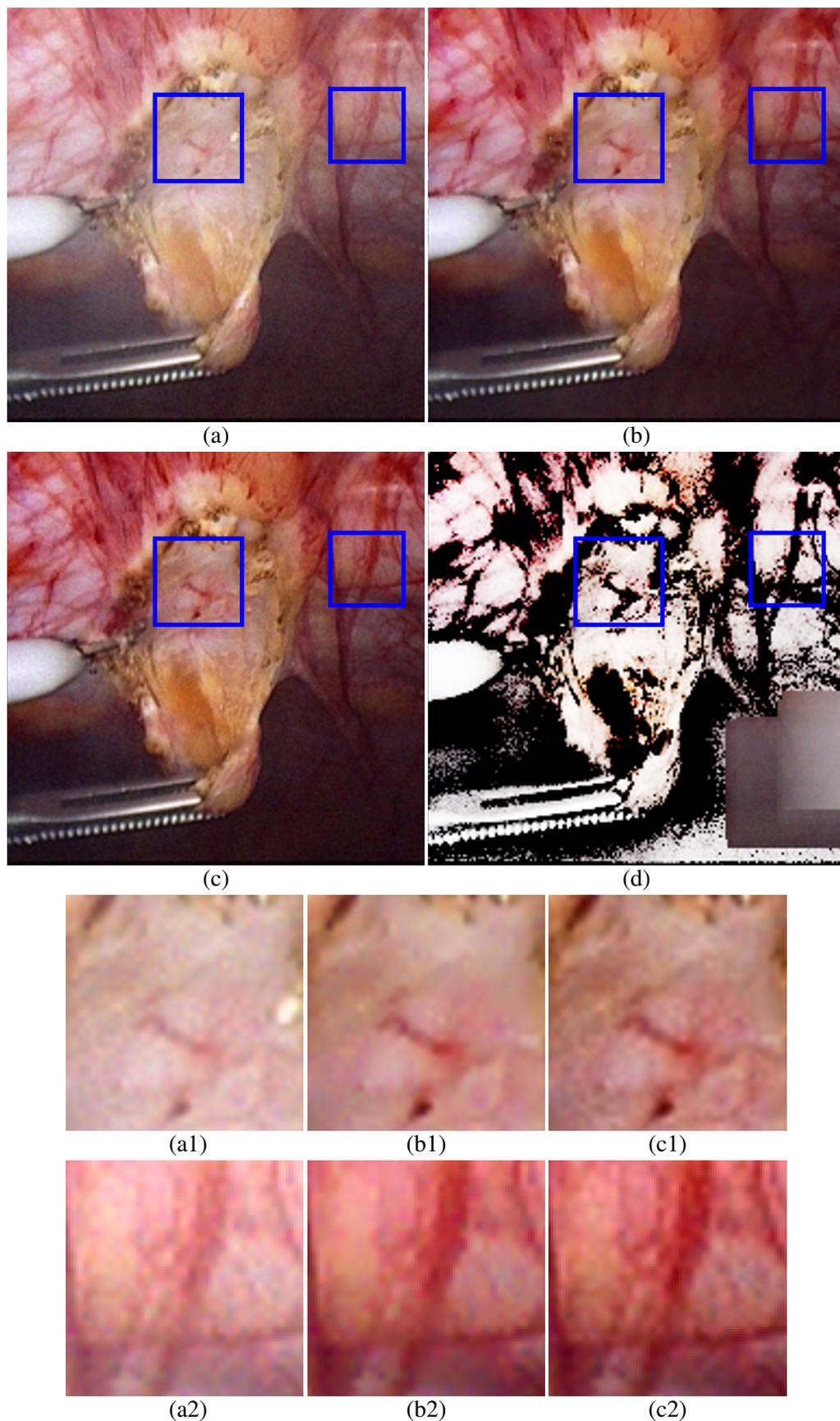
Figure 5.8:   **Results on Real World Laparoscopic Image.** **(a)** Observed image. Results of processing image **(a)** using: **(b)** denoising using bilateral filtering followed by dehazing [1] followed by inpainting; **(c)** *our method*; **(d)** adaptive filtering [2] followed by inpainting. Zoomed insets of **(a), (b), (c)** are in **(a1), (b1), (c1)** and **(a2), (b2), (c2)**.
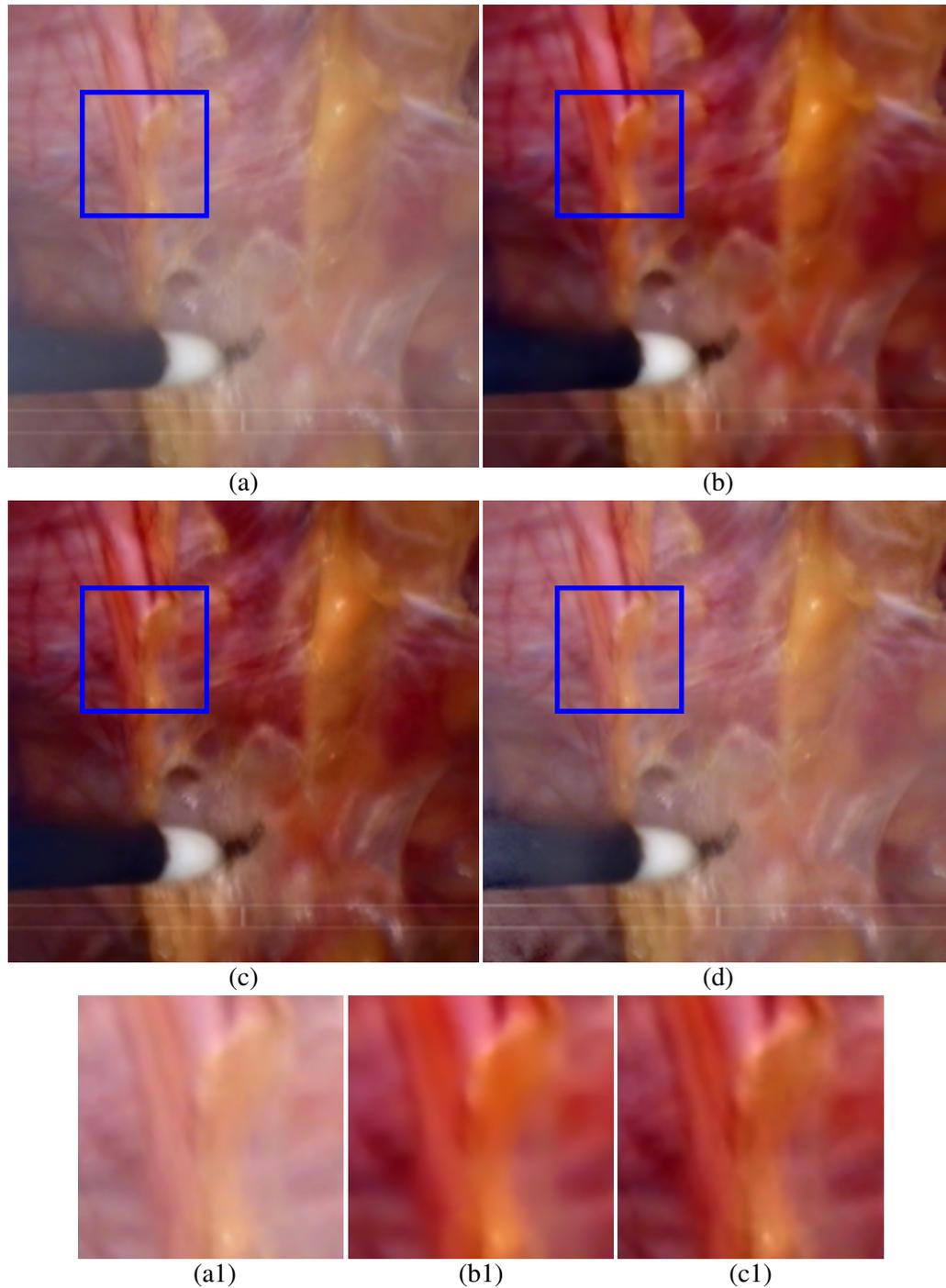
Figure 5.9:   **Results on Real World Laparoscopic Image.** **(a)** Observed image. Results of processing image **(a)** using: **(b)** denoising with bilateral filter followed by dehazing [1] followed by inpainting; **(c)** *our method*; **(d)** adaptive filtering [2] followed by inpainting. Zoomed insets of **(a), (b), (c)** are in **(a1), (b1), (c1)**.
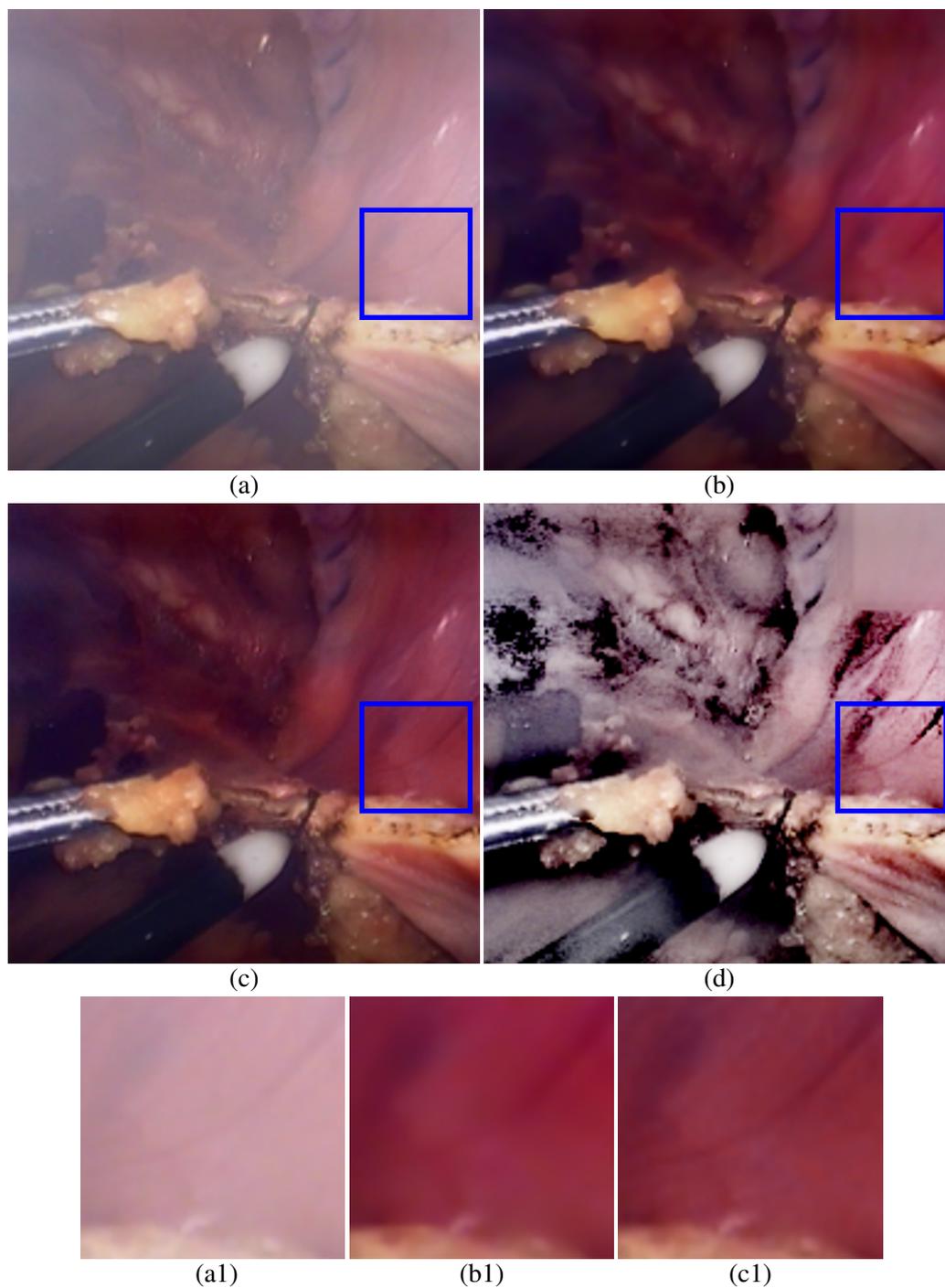
Figure 5.10: **Results on Real World Laparoscopic Image.** **(a)** Observed image. Results of processing image **(a)** using: **(b)** denoising with bilateral filter followed by dehazing [1] followed by inpainting; **(c)** *our method*; **(d)** adaptive filtering [2] followed by inpainting. Zoomed insets of **(a), (b), (c)** are in **(a1), (b1), (c1)**.
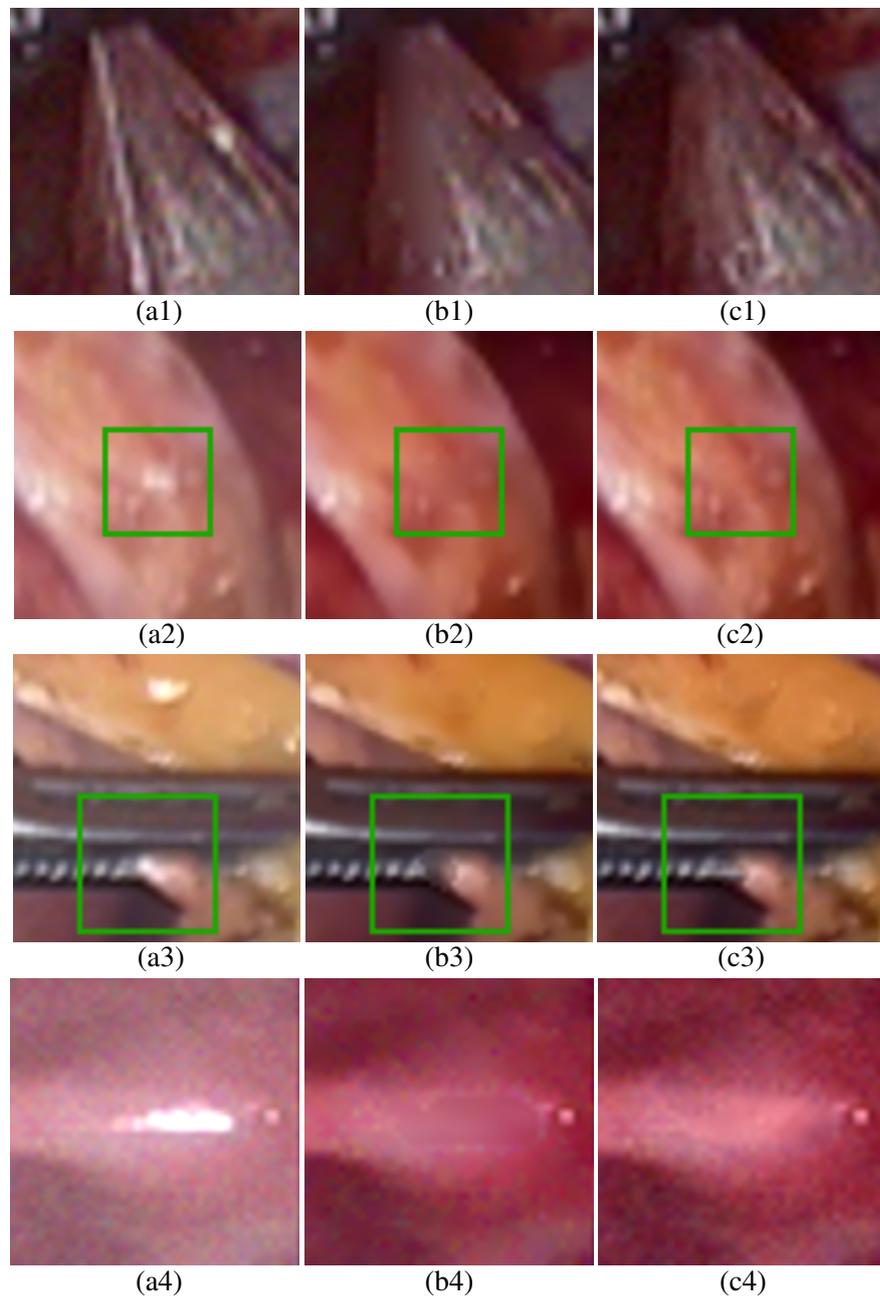
Figure 5.11:   **Results on Observed Laparoscopic Images. (a)** Observed Image. Results of processing image (a) using: **(b)** denoising with bilateral filter followed by dehazing [1] followed by inpainting; **(c)** *our method*.

Chapter 6

Conclusion and Future Work

Our proposed algorithm VBEM does a better job at removing smoke, specular highlight, and noise than the compared sequential methods. We produce natural colors after haze removal, preserve texture after removing noise, fill in better and sharp details in the holes created by specular highlight. These claims are supported by qualitative examples, as well as the metrics.

Our work can be extended by application on temporal sequences. The priors can be constructed in an online manner, using the temporal information. Our developments can be tried on different class of images and the performance can be evaluated.

Appendix A

Optimal factors for transmission map $\mathbf{T}$

We will now derive the parameters of the optimal factorization for $G_i^\dagger$.

$$\log P\left(\mathring{\mathbf{y}}, \mathbf{T} | \mathbf{x}^m, \mathbf{r}\right) = -\sum_{i:r_i=0}\left(\mathring{y}_i - T_i x_i^m - (1-T_i)K_{\text{smoke}}\right)^2$$
$$-\gamma_3 \sum_i \sum_{j \in \mathcal{N}_i^T}(T_i - T_j)^2 \tag{A.1}$$

To get optimal factor in Equation (4.12), we replace $\{T_k | k \neq i\}$ with current optimum means $\{\mu_k^\mathbf{T}\}$ and replace $\{T_k^2 | k \neq i\}$ with $\{\left(\mu_j^\mathbf{T}\right)^2 + \left(\sigma_j^\mathbf{T}\right)^2\}$ in Equation (A.1). This will result in a quadratic equation in T_i, from which the parameters $\left(\mu_i^\mathbf{T}, \sigma_i^\mathbf{T}\right)$ of truncated Gaussian factor $G_i^{\dagger*}$ can be solved [25]. The final expression for mean and standard deviation are is derived as

$$\mu_i^\mathbf{T} = \bar{\mu}_i + \bar{\sigma}_i \frac{\phi\left(\alpha_i\right) - \phi\left(\beta_i\right)}{\Phi\left(\beta_i\right) - \Phi\left(\alpha_i\right)} \tag{A.2}$$

$$\sigma_i^\mathbf{T} = \bar{\sigma}_i^2\left[1 + \frac{\alpha_i\phi\left(\alpha_i\right) - \beta_i\phi\left(\beta_i\right)}{\Phi\left(\beta_i\right) - \Phi\left(\alpha_i\right)} - \left(\frac{\phi\left(\alpha_i\right) - \phi\left(\beta_i\right)}{\Phi\left(\beta_i\right) - \Phi\left(\alpha_i\right)}\right)^2\right] \tag{A.3}$$

where ϕ and Φ are the PDF and the CDF of a standard normal distribution, and

$$\bar{\mu}_i = \frac{(1 - r_i)(y_i - K_{\text{smoke}})(x_i^m - K_{\text{smoke}}) + 2\gamma_3 \sum_{j \in \mathcal{N}_i^T} \mu_j}{(1 - r_i)(x_i^m - K_{\text{smoke}})^2 + 2\gamma_3},$$

$$\bar{\sigma}_i = \frac{1}{\sqrt{2}} \left((x_i^m - K_{\text{smoke}})^2 + 2\gamma_3 \right)^{-\frac{1}{2}},$$

$$\alpha_i = -\frac{\bar{\mu}_i}{\bar{\sigma}_i}, \text{and}$$

$$\beta_i = \frac{1 - \bar{\mu}_i}{\bar{\sigma}_i}.$$

Appendix B

Optimal factors for codes S

With a prelearnt dictionary $\mathbf{D} := \{D_j\}_{j=1}^J$, the probability distribution on the sparse code S_i for patch i of $\mathbf{x}$ is

$$\log P(S_{ij}|\mathbf{x}^m) \propto -\|\bar{\mathbf{x}}_i^m - \mathbf{D}S_i\|_2^2 - \lambda\|S_i\|_1 \tag{B.1}$$

Due to the absolute value operator in Equation (B.1), we are not able to fit any standard distribution for G_{ij}. We therefore replace L1 regularization with weighted L2 regularization. Iterative reweighted least squares (IRLS) can be used to approximate the solution of a sparse solution due to L1 regularization with an appropriate choice of weights [22]. Specifically, if the estimate after m iterations if S_i^m, then the next iteration gives the solution as

$$S_i^{m+1} = \arg\min_{S_i} \|\bar{\mathbf{x}}_i^m - \mathbf{D}S_i\|_2^2 + \sum_j \Gamma_{ij} S_{ij}^2 \tag{B.2}$$

where the weight $\Gamma_{ij} = 1/|S_{ij}^m|$.

We use this development to propose a strategy for factorization. At each iteration, we set the weights as the inverse of the absolute of the mean of the factors, followed by solving for optimum factors until convergence. When we have the weights $\mathbf{\Gamma}$ fixed, the new probability

distribution (instead of Equation (B.1)) is

$$\log P(S_{ij}|\mathbf{x}^m; \mathbf{\Gamma}) = -\|\bar{\mathbf{x}}_i^m - \mathbf{D}S_i\|_2^2 - \lambda\Gamma_{ij}S_{ij}^2 + c \tag{B.3}$$

where c is a constant.

We solve Equation (4.13) by substituting $\{S_{ik}|k \neq j\}$ with $\mu_{ik}^{\mathbf{S}}$ in Equation (B.1) to obtain a quadratic in S_{ij} sans an additive constant. We can then obtain the mean and sigma of the optimal Gaussian factor G_{ij} as

$$\mu_{ij}^{\mathbf{S}} = \frac{D_j^{\mathsf{T}}\left(\bar{\mathbf{x}}_i^m - \sum_{k \neq j} D_k \mu_{ik}^{\mathbf{S}}\right)}{D_j^{\mathsf{T}}D_j + \lambda\Gamma_{ij}} \tag{B.4}$$

$$\sigma_{ij}^{\mathbf{S}} = \frac{1}{\sqrt{D_j^{\mathsf{T}}D_j + \lambda\Gamma_{ij}}} \tag{B.5}$$

Bibliography

[1] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.

[2] K. B. Gibson and T. Q. Nguyen, "Fast single image fog removal using the adaptive wiener filter," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*. IEEE, 2013, pp. 714–718.

[3] D. Stoyanov and G. Z. Yang, "Removing specular reflection components for robotic assisted laparoscopic surgery," in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 3. IEEE, 2005, pp. III–632.

[4] C.-A. Saint-Pierre, J. Boisvert, G. Grimard, and F. Cheriet, "Detection and correction of specular reflections for automatic surgical tool segmentation in thoracoscopic images," *Machine Vision and Applications*, vol. 22, no. 1, pp. 171–180, 2011.

[5] W. L. Barrett and S. M. Garber, "Surgical smoke: a review of the literature," *Surgical endoscopy*, vol. 17, no. 6, pp. 979–987, 2003.

[6] K. Prokopetc, T. Collins, and A. Bartoli, "Automatic detection of the uterus and fallopian tube junctions in laparoscopic images," in *International Conference on Information Processing in Medical Imaging*. Springer, 2015, pp. 552–563.

[7] S. Voros, J.-A. Long, and P. Cinquin, "Automatic detection of instruments in laparoscopic

images: A first step towards high-level command of robotic endoscopic holders," *The International Journal of Robotics Research*, vol. 26, no. 11-12, pp. 1173–1190, 2007.

[8] R. Wolf, J. Duchateau, P. Cinquin, and S. Voros, "3d tracking of laparoscopic instruments using statistical and geometric modeling," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2011, pp. 203–210.

[9] M. Arnold, A. Ghosh, S. Ameling, and G. Lacey, "Automatic segmentation and inpainting of specular highlights for endoscopic imaging," *EURASIP Journal on Image and Video Processing*, vol. 2010, no. 1, p. 814319, 2010.

[10] H. Koschmieder, *Theorie der horizontalen sichtweite: kontrast und sichtweite*. Keim & Nemnich, 1925.

[11] R. Fattal, "Single image dehazing," *ACM transactions on graphics (TOG)*, vol. 27, no. 3, p. 72, 2008.

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

[13] J. Pang, O. C. Au, and Z. Guo, "Improved single image dehazing using guided filter," *Proc. APSIPA ASC*, pp. 1–4, 2011.

[14] E. Matlin and P. Milanfar, "Removal of haze and noise from a single image." in *Computational Imaging*, 2012, p. 82960T.

[15] N. Joshi and M. F. Cohen, "Seeing mt. rainier: Lucky imaging for multi-image denoising, sharpening, and haze removal," in *Computational Photography (ICCP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–8.

[16] A. Kotwal, R. Bhalodia, and S. P. Awate, "Joint desmoking and denoising of laparoscopy images," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 1050–1054.

[17] G. F. Harpur and R. W. Prager, "Development of low entropy coding in a recurrent network," *Network: computation in neural systems*, vol. 7, no. 2, pp. 277–284, 1996.

[18] B. A. Olshausen and D. J. Field, "Sparse coding of natural images produces localized, oriented, bandpass receptive fields," *Nature*, vol. 381, no. 60, p. 609, 1996.

[19] J. Mairal, J. Ponce, G. Sapiro, A. Zisserman, and F. R. Bach, "Supervised dictionary learning," in *Advances in neural information processing systems*, 2009, pp. 1033–1040.

[20] P. O. Hoyer, "Non-negative sparse coding," in *Neural Networks for Signal Processing, 2002. Proceedings of the 2002 12th IEEE Workshop on*. IEEE, 2002, pp. 557–565.

[21] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proceedings of the 26th annual international conference on machine learning*. ACM, 2009, pp. 689–696.

[22] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Acoustics, speech and signal processing, 2008. ICASSP 2008. IEEE international conference on*. IEEE, 2008, pp. 3869–3872.

[23] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer graphics and applications*, vol. 21, no. 5, pp. 34–41, 2001.

[24] B. W. Silverman, *Density estimation for statistics and data analysis*. CRC press, 1986, vol. 26.

[25] N. Johnson, S. Kotz, and N. Balakrishnan, "Continuous univariate probability distributions,(vol. 1)," 1994.

[26] A. Baid, A. Kotwal, R. Bhalodia, S. Merchant, and S. P. Awate, "Joint desmoking, specularity removal, and denoising of laparoscopy images via graphical models and bayesian inference," in *Biomedical Imaging (ISBI), 2017 IEEE 14th International Symposium on*. IEEE, 2017.

[27] A. Baid, S. Awate, and S. Merchant, "Joint desmoking, specularity removal, and denoising of laparoscopy images via graphical models and bayesian inference," *IEEE Transactions on Medical Imaging*, in prep.

List of Publications

1. A. Baid, A. Kotwal, R. Bhalodia, S. Merchant, and S. P. Awate, "Joint desmoking, specularity removal, and denoising of laparoscopy images via graphical models and bayesian inference," in *Biomedical Imaging (ISBI), 2017 IEEE 14th International Symposium on*. IEEE, 2017

2. A. Baid, S. Awate, and S. Merchant, "Joint desmoking, specularity removal, and denoising of laparoscopy images via graphical models and bayesian inference," *IEEE Transactions on Medical Imaging*, in prep

Acknowledgements

I express by deep gratitude towards Prof. Awate and Prof. Merchant for working with me. They were always very helpful along each step of the work. I would like to thank my classmates Alankar Kotwal, Riddhish Bhalodia, and Anand Pathak for helping in planning, problem formulation, and derivations. Alankar also helped me with debugging and tuning the code, and also actively participated in weekly discussion.

Ayush Baid

IIT Bombay

18 June 2017