

Unsupervised Learning of Surgical Smoke Removal from Simulation

L. Chen¹, W. Tang¹, and N. W. John²

¹*Department of Creative Technology, Bournemouth University*

²*School of Computer Science, University of Chester*

{chenl, wtang}@bournemouth.ac.uk, nigel.john@chester.ac.uk

INTRODUCTION

The surgical smoke produced during minimally invasive surgery can not only reduce the visibility of the surgeons, but also severely affect the performance of image processing algorithms used for image guided surgery such as image tracking, segmentation, detection and retrieval. Besides from physical smoke evacuation devices, many research works [3] [1] [4] [6] [7] address this issue by using vision-based methods to filter out the smoke and try to recover the clear images. More recently, end-to-end deep learning approaches [2] have been introduced to solve the de-hazing and de-smoking problems. However, it is extremely difficult to collect large amounts of data for the effective learning of the implicit de-smoking function, especially for surgical scenes. In this paper, we propose a computational framework for unsupervised learning of smoke removal from rendering smoke on laparoscopic video. Compared to conventional image processing approaches, our proposed framework is able to remove local smoke and recover more realistic tissue colour but will not affect the areas without smoke. Although trained on synthetic images, the experimental results show that our network is able to effectively remove smoke on laparoscopic images with real surgical smoke.

MATERIALS AND METHODS



Fig. 1: Overview of our framework for unsupervised learning of smoke removal

With the recent development of Convolutional Neural Networks (CNNs), machines can solve many computer vision problems when provided with large-scale ground-truth data for supervised learning. However, the acquisition of ground truth is an expensive, time-consuming and labour-intensive task. Especially for surgical scenes, the amount of data and the accuracy of data must be satisfied to ensure acceptable results for clinical use. To address this issue, we proposed an unsupervised framework for learning smoke removal. As can be seen from Figure 1, our frame-

work is composed of a render engine for synthesizing smoked images from laparoscopic videos and a CNN for the end-to-end training of de-smoking.



Fig. 2: The random smoke is rendered on the background of laparoscopic images to synthesize smoke image.

Smoke Synthesis

We use Blender¹ – an open source 3D creation software for the synthesis of smoke image for training. As shown in Figure 2, the real laparoscopic images from Hamlyn Centre Laparoscopic / Endoscopic Video Datasets² [8] are used as background images, and the smoke is rendered with randomly intensity, density and location on the background to simulate the real smoke. The variation of smoke render ensures that our network will not over-fitting to certain smoke intensity, density and location. With the help of powerful render engine, we are able to synthesize unlimited amount of realistic images with the presence of simulated surgical smoke. More importantly, this process is done automatically without any human intervention and cost.

End-to-end Learning of Smoke Removal

As smoke removal is a pixel-wise task, we adopt a fully convolutional encoder-decoder network to generate the same size de-smoked image with the input image. Since the smoke removal task needs to preserve most of the details from the input images, following the U-Net structure [5], skip connections are implemented for directly transferring high-level information to the bottom of the network to prevent the loss of high quality details.

¹<https://www.blender.org/>

²<http://hamlyn.doc.ic.ac.uk/vision/>

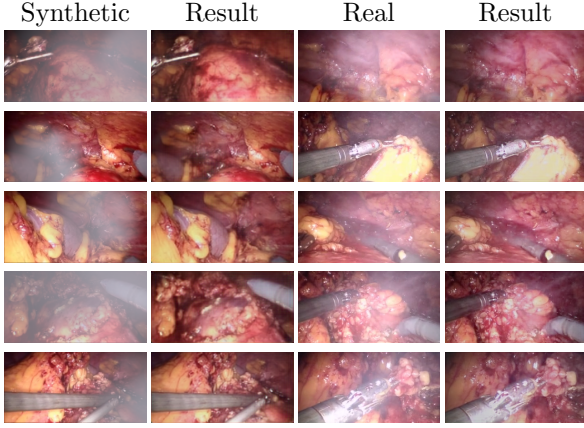


Fig. 3: The result for our smoke removal framework. Left: Synthetic smoke images and results; Right: Real surgical smoke images and results

The loss function for the network is the L1 error of original image and de-smoked image from synthesized smoke image.

$$L = \sum_{xy} |I_{original}(x, y) - I_{desmoked}(x, y)| \quad (1)$$

RESULTS

Our smoke removal network was implemented on Tensorflow platform and trained on a workstation with a single NVidia Titan X GPU. We used an *in-vivo* partial nephrectomy procedure dataset from Hamlyn Centre Laparoscopic / Endoscopic Video Datasets [8] that contains 20,000 stereo image pairs from 11 video sequences, in which 7,553 images without surgical smoke are manually picked as ground-truth dataset. We randomly rendered smoke on these 7,553 images to construct our training dataset. The images are scaled to 256x128 to accelerate training and inference. After training, our model only takes 12ms to remove smoke for a 256x128 image.

For testing, we run our trained network on laparoscopic images with real surgical smoke, the results are shown in Figure 3. The qualitative visual inspection of the results shows that our proposed framework can effectively remove real surgical smoke and recover correct tissue colours. The code used for data synthesis and training can be found at here ³.

DISCUSSION

In this work, we present a novel unsupervised deep learning framework for automatic surgical smoke removal during minimally invasive surgery. Our framework only needs laparoscopic images as input, and 3D render engine is used to randomly render smoke on these images to synthesize dataset for training. The preliminary results show that although our network is fully training on synthesize dataset, it has the ability to remove real surgical smoke during minimally invasive surgery at real-time speed. This solves

the problem that deep learning need large amounts of hand-labelled ground truth for training, especially for medical datasets that profession knowledge is needed for labelling data. In future work, it is interesting to see whether training networks from synthetic dataset can be extended to many other tasks such as laparoscope tracking, surgical instruments detection and tissue/organ segmentation, which will overcome the shortage of medical ground-truth data and greatly benefit the deep learning technology to be used in surgical scenes.

REFERENCES

- [1] A. Baid, A. Kotwal, R. Bhalodia, S. N. Merchant, and S. P. Awate. Joint desmoking, specularly removal, and denoising of laparoscopy images via graphical models and bayesian inference. In *IEEE ISBI*, apr 2017.
- [2] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. De-hazeNet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198, nov 2016.
- [3] A. Kotwal, R. Bhalodia, and S. P. Awate. Joint desmoking and denoising of laparoscopy images. In *IEEE ISBI*, apr 2016.
- [4] X. Luo, A. J. McLeod, S. E. Pautler, C. M. Schlachta, and T. M. Peters. Vision-based surgical field defogging. *IEEE Transactions on Medical Imaging*, 36(10):2021–2030, oct 2017.
- [5] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 2015.
- [6] K. Tchaka, V. M. Pawar, and D. Stoyanov. Chromaticity based smoke removal in endoscopic images. In *SPIE Medical Imaging*, feb 2017.
- [7] C. Wang, F. Alaya Cheikh, M. Kaaniche, and O. J. Elle. A Smoke Removal Method for Laparoscopic Images. *ArXiv e-prints*, Mar. 2018.
- [8] M. Ye, E. Johns, A. Handa, L. Zhang, P. Pratt, and G.-Z. Yang. Self-supervised siamese learning on stereo image pairs for depth estimation in robotic surgery. In *Hamlyn Symposium on Medical Robotics*, 2017.

³https://github.com/melights/cnn_desmoke