Intra-operative Image Enhancement and Registration for Image Guided Laparoscopic Liver Resection

PhD Candidate: Congcong Wang[1]

Supervisors

Prof. Faouzi Alaya Cheikh[1], Prof. Ole Jakob Elle[2,3], Prof. Azeddine Beghdadi[4]

[1]Norwegian University of Science and Technology, Norway

[2]Oslo University Hospital, Norway

[3]University of Oslo, Norway

[4]University Sorbonne Paris Nord , France

About the research project

- The Norwegian Research Council supported project IQ-MED: Image Quality enhancement in MEDical diagnosis, monitoring, and treatment.
  - WP1: Image quality enhancement
  - WP2: Capsule video endoscopy
  - **WP3: Video guided surgery**
  - WP4: Skin imaging
- Collaborations
  - Color lab of NTNU, IVS of Oslo University Hospital, L2TI of University Sorbonne Paris Nord
  - The EU-funded ITN-project HiPerNav (High-Performance soft-tissue Navigation)

Outline

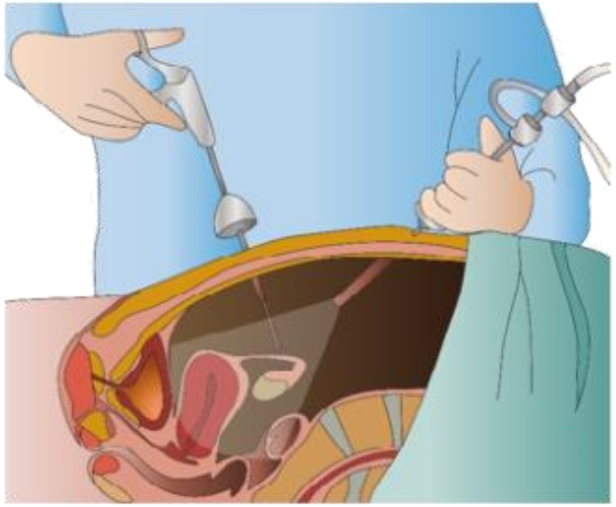
- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke removal
- Conclusion & Future perspectives

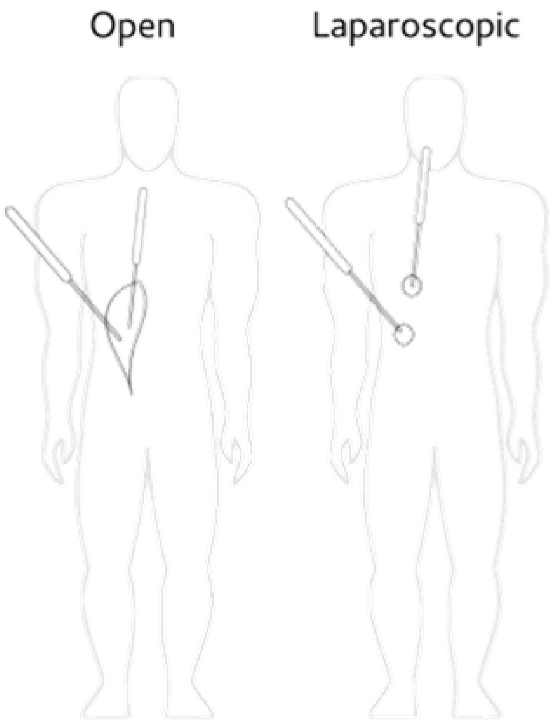
Outline

- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke removal
- Conclusion & Future perspectives

Laparoscopic Liver Resection

- Surgery is performed through small incisions.



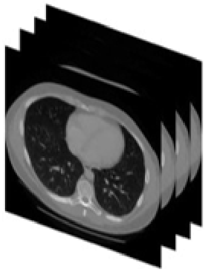


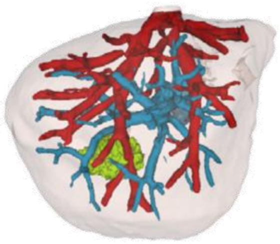
Advantages:

- Less hospitalization (shorter recovery)
- Less bleeding
- Better cosmetic output for the patient

Laparoscopic Liver Resection

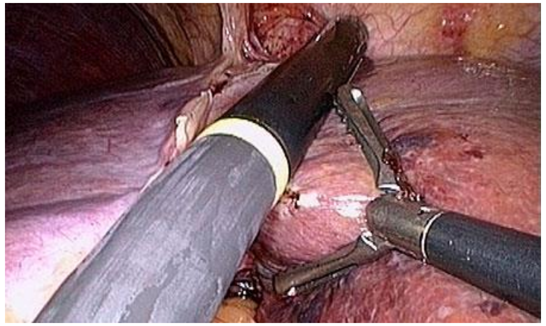
Pre-operative





R. Palomar, The Intervention Center, Oslo University Hospital

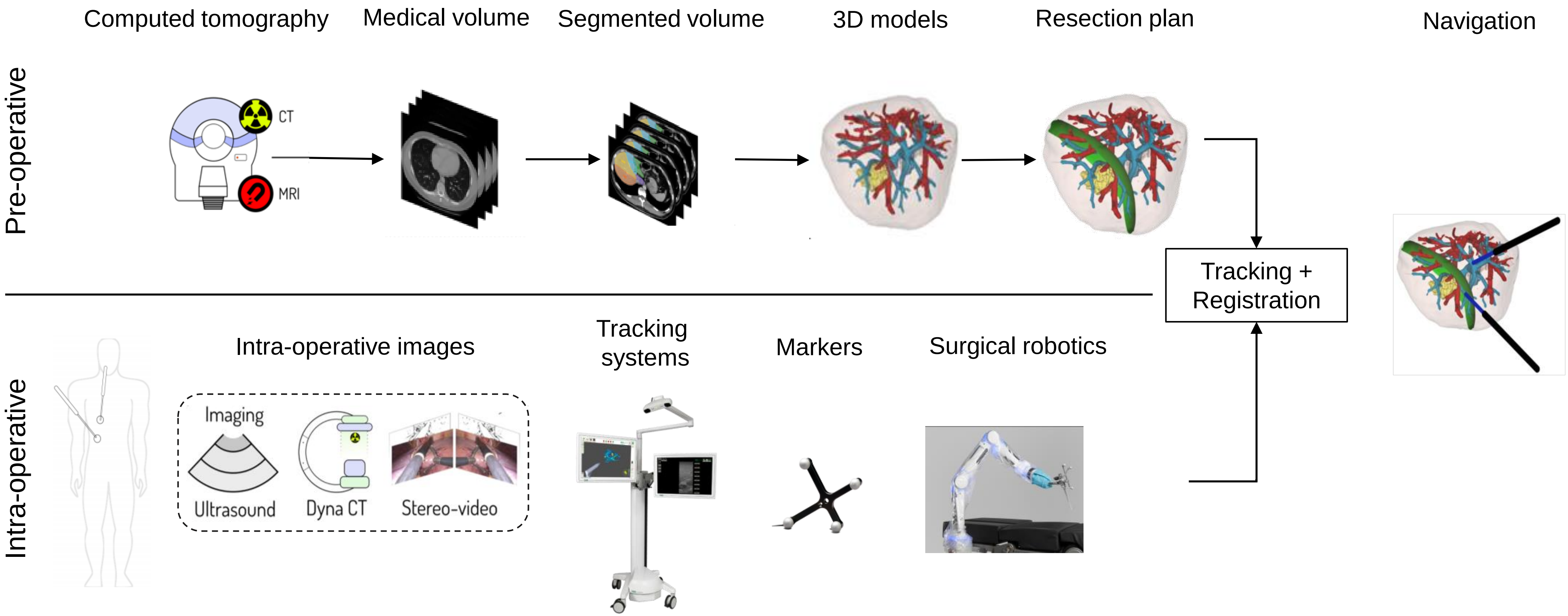
Intra-operative



Challenges:

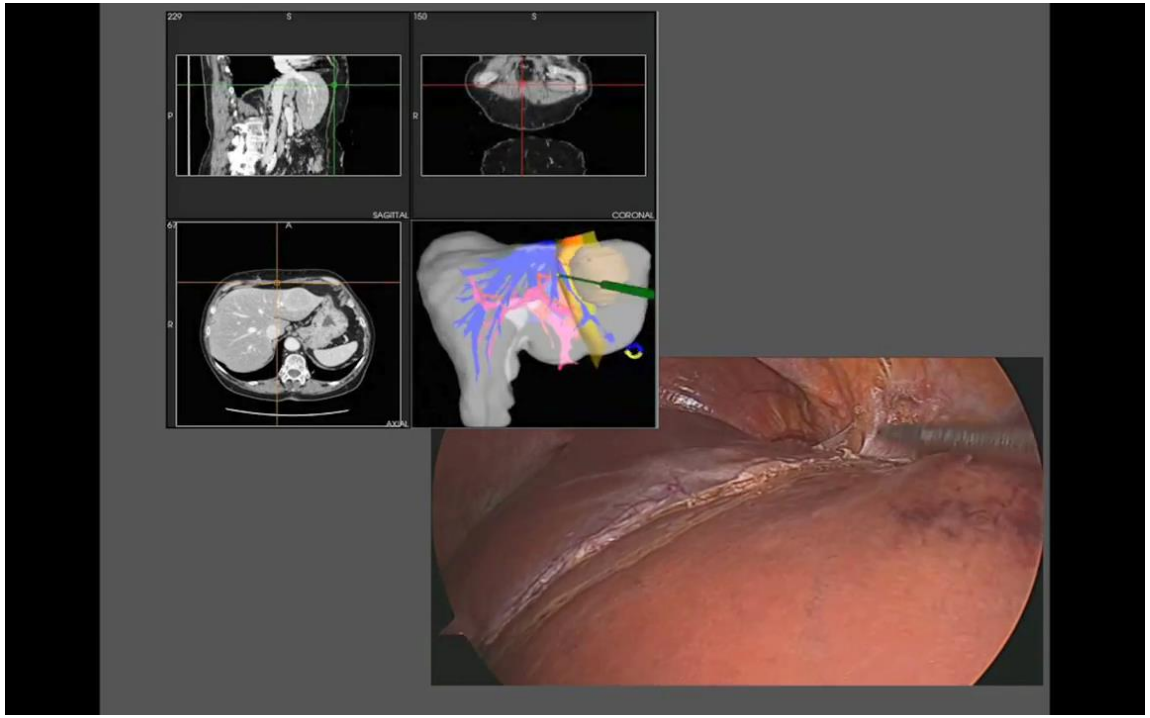
- Diagnostic images from MRI or CT presented somewhere else, not present in the OR in relation to the patient
- Liver is a soft tissue that moves and deforms during surgery
- The surface of the liver has few anatomical landmarks
- Difficult to do precise resections especially when it is close to big vessel that is not visible

Navigation for Laparoscopic Liver Resection

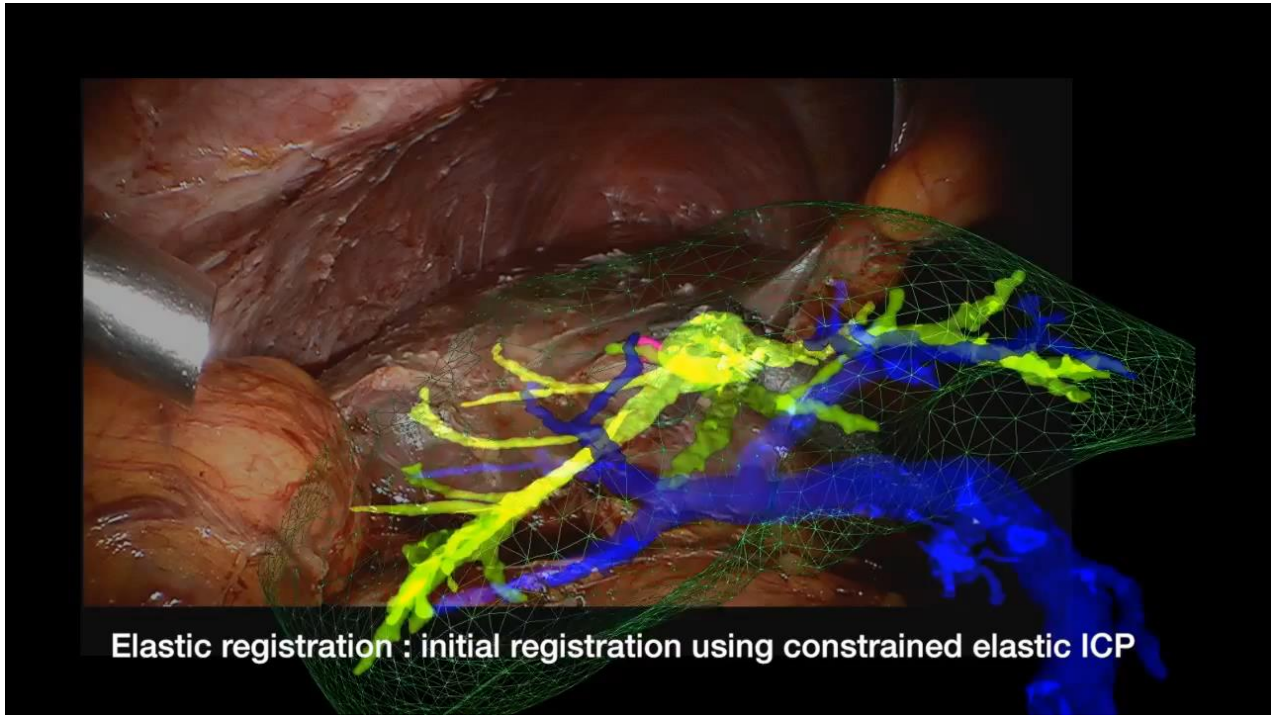


Images are mainly from R. Palomar, The Intervention Center, Oslo University Hospital

Examples of Intra-operative Navigation



Pathsurg

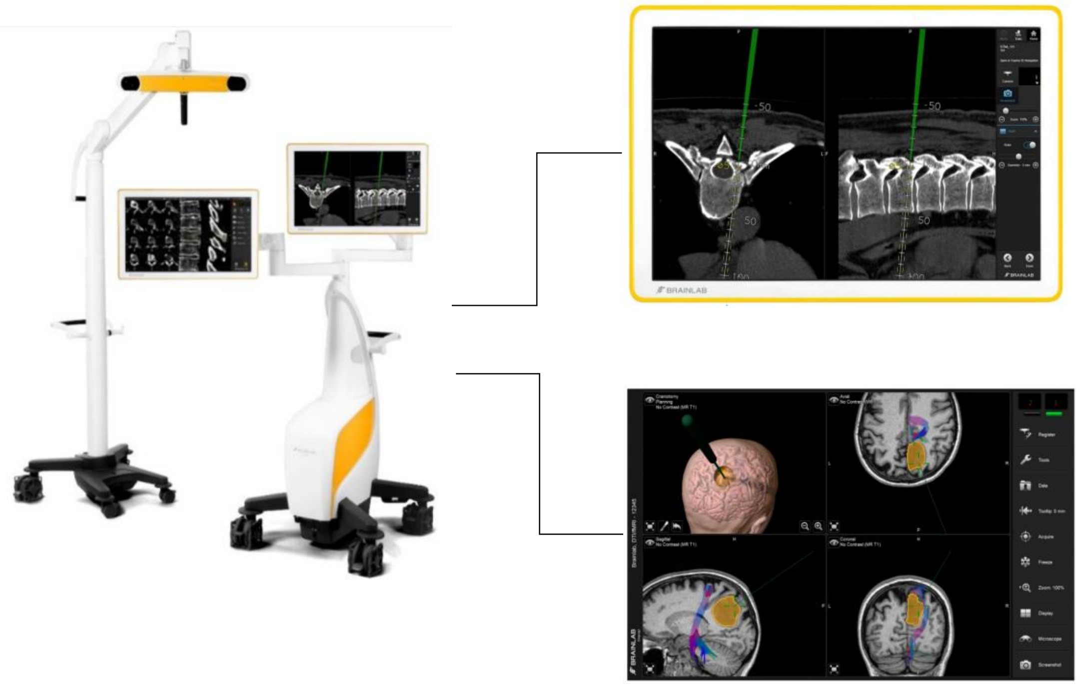


N. Haouchine et al.[8]

Navigation Systems

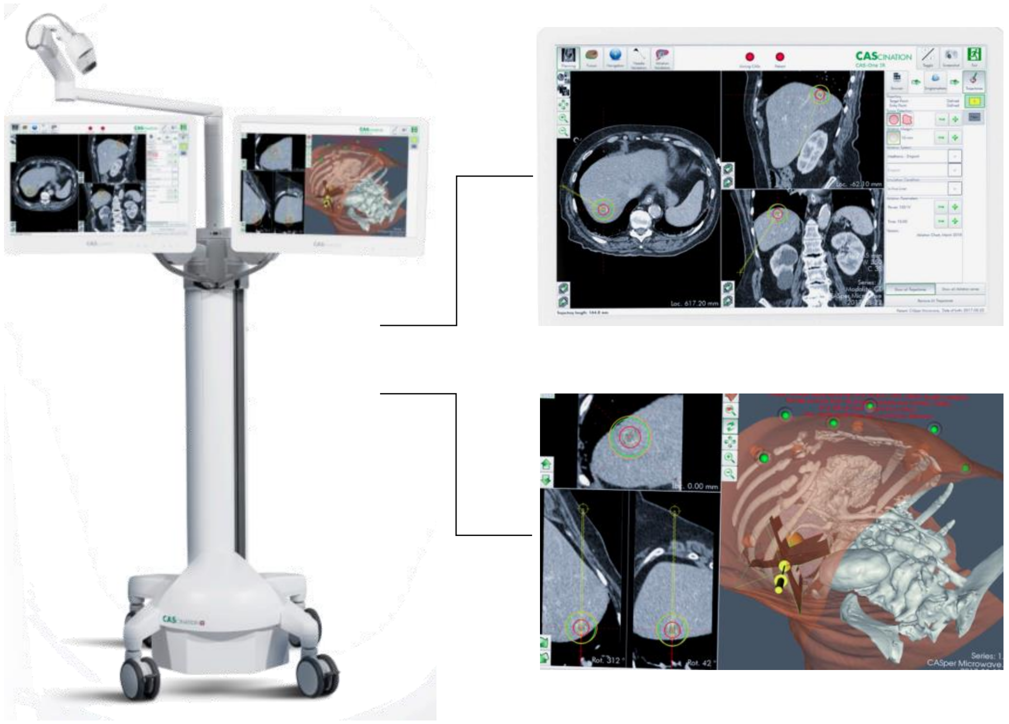
- Brainlab

Navigation in spine and brain surgery



- Cascination

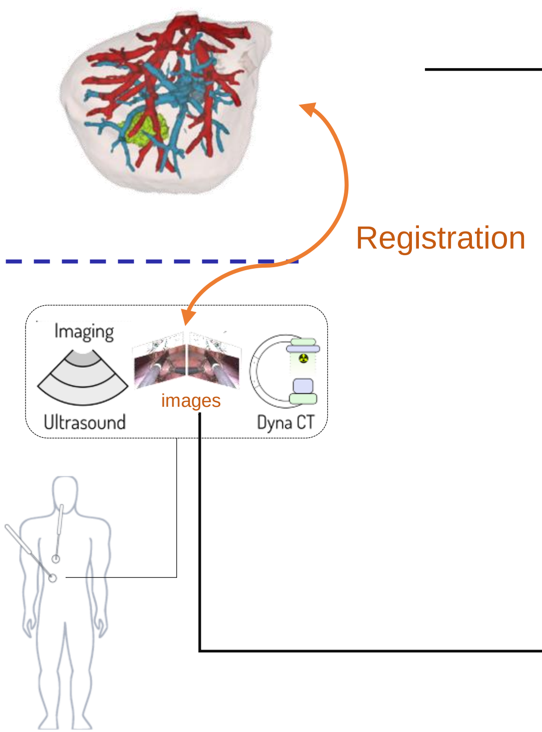
Navigation for interventional oncology



Research Objective - Stereo Vision Based Surface Registration

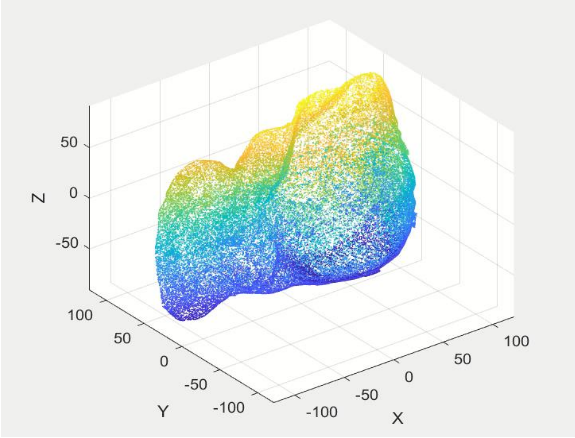
Pre-operative

3D modeling

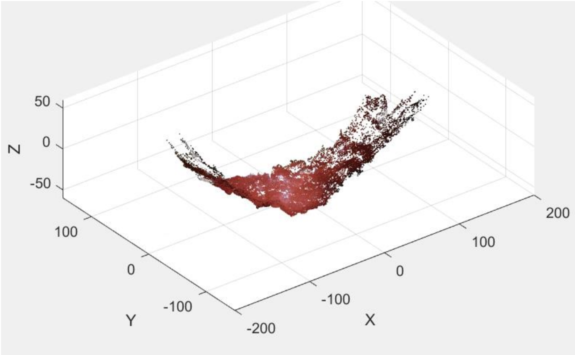


Intra-operative

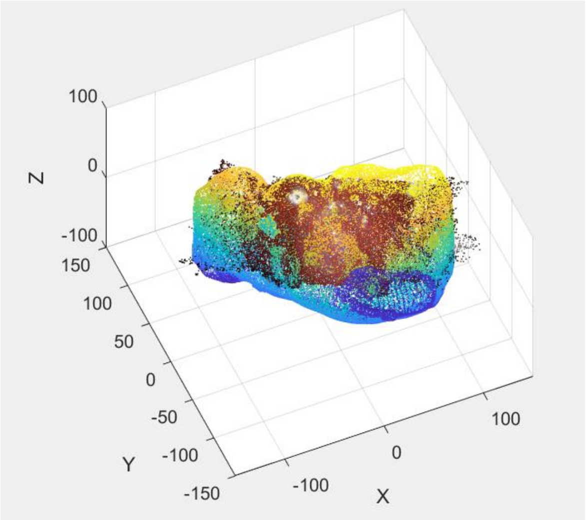
(1) Extract surface from 3D model



(2) Surface reconstruction from laparoscopic images



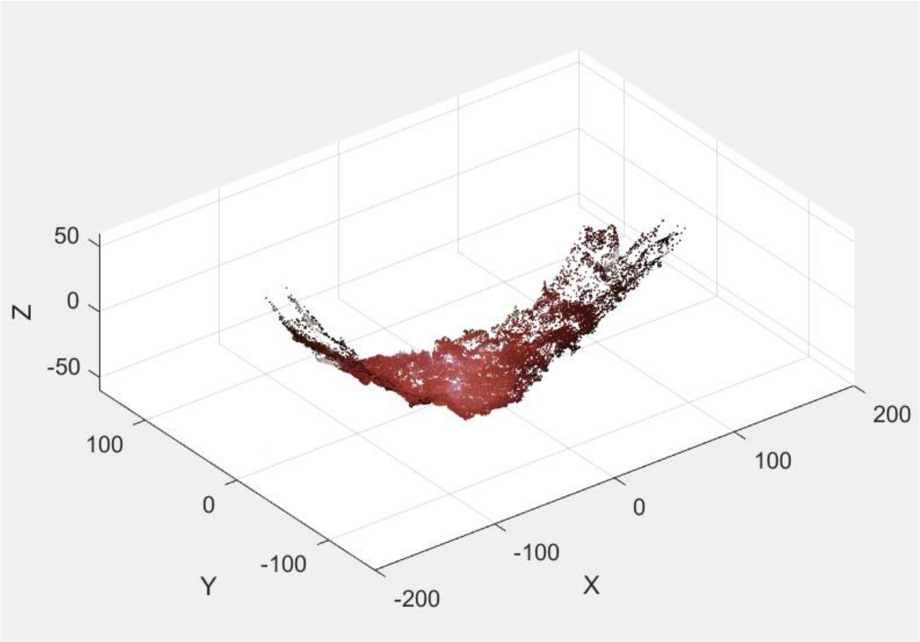
(3) Register the surfaces



Research Outline: Surface Reconsruction

- Goal: Surface reconstruction from stereo images

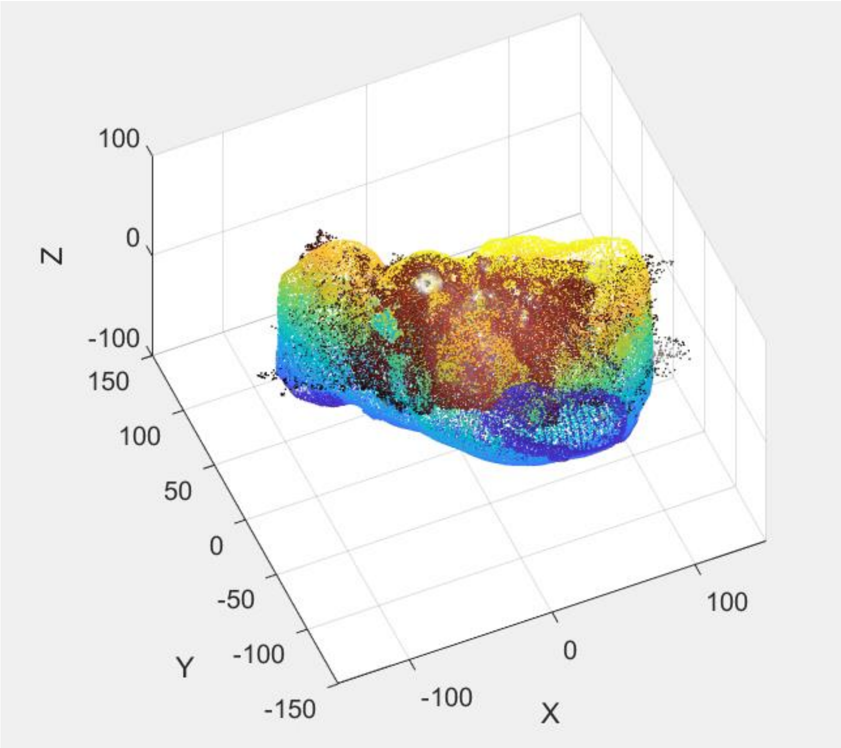
  Q1.1 How to reconstruct organ surface from stereo laparoscopic images? (**Paper F**)



Surface reconstruction (**Q1.1**, **Paper F**)

Research Outline: Surface Registration

- Goal: Surface based registration
  - **Q1.2** How to register the reconstructed organ surface with the pre-operative 3D volume? (**Paper G**)

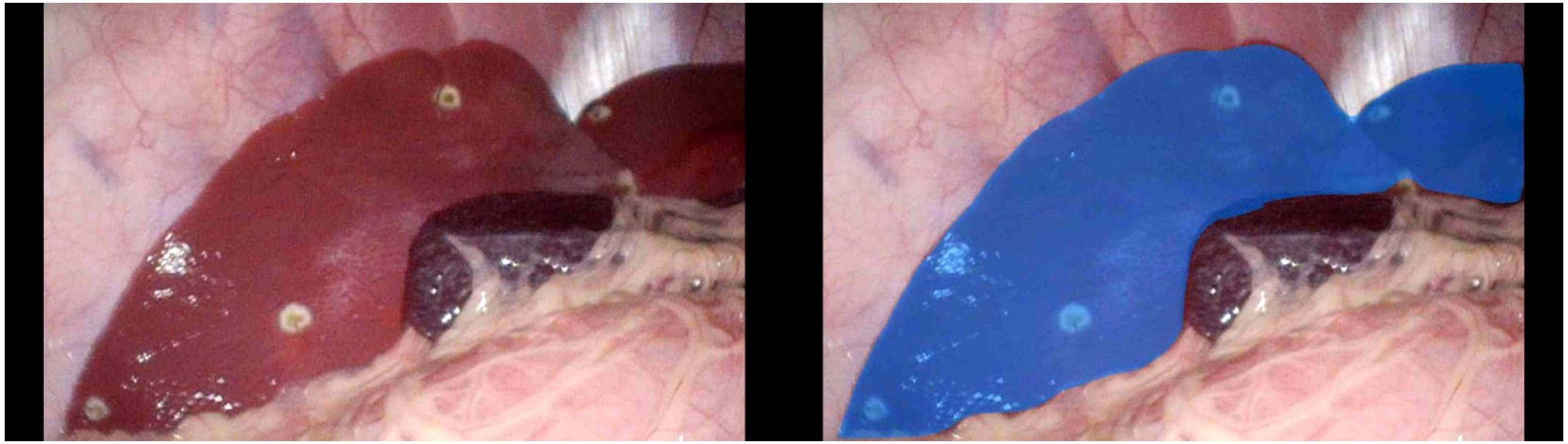


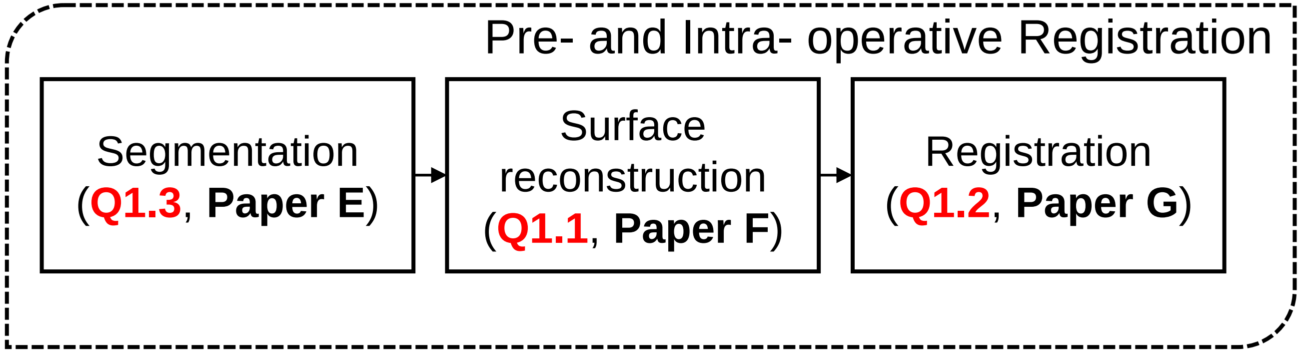
Surface reconstruction (**Q1.1**, **Paper F**) → Registration (**Q1.2**, **Paper G**)

Research Outline: Semantic Segmentation

- Goal: Semantic segmentation

  Q1.3 How to perform automatic semantic segmentation of the surgery scene especially the targeted organ? (**Paper E**)





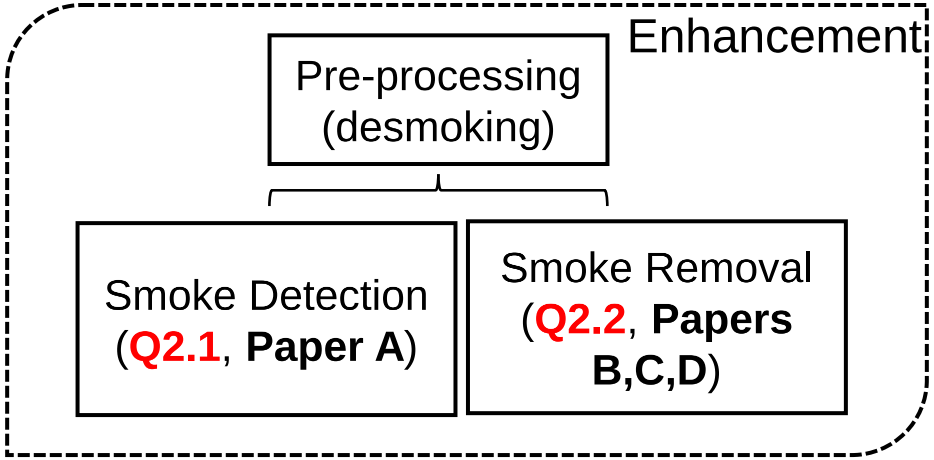
Research Outline

- Goal: Remove smoke in laparoscopic images using an image processing method.
  - For a better visualization of the surgical field
  - For a more robust performance of the following computer vision algorithms

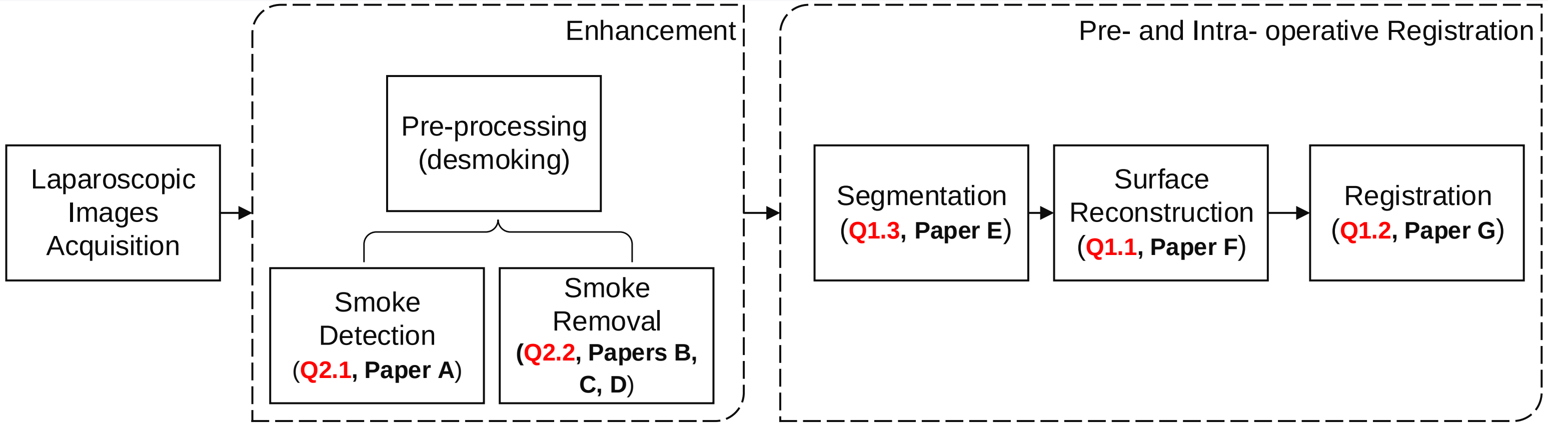
Smoke images Desmoked images

Q2.1 How to discriminate smoke and non-smoke frames? (**Paper A**)

Q2.2 How to remove the smoke without affecting the structural information and the visibility of features of interest so as to guarantee an acceptable surgical vision? (**Papers B, C, D**)



Research Outline: Publications



A. **Wang, C.***, Sharma, V.*, Fan, Y., Alaya Cheikh, F., Beghdadi, A., and Elle, O.J., and Stiefelhagen, R. "Can Image Enhancement be Beneficial to Find Smoke Images in Laparoscopic Surgery?". In 26th Color and Imaging Conference, pp. 163-168, 2018, Society for Imaging Science and Technology. (* denotes equal contribution.)

B. **Wang, C.**, Alaya Cheikh, F., Kaaniche, M., Beghdadi, A., and Elle, O. J. (2018). "Variational based smoke removal in laparoscopic images". Biomedical engineering online, 17(1), 139.

C. Bolkar, S., **Wang, C.**, Alaya Cheikh, F., and Yildirim, S. "Deep smoke removal from minimally invasive surgery videos". In 25th IEEE International Conference on Image Processing (ICIP), pp. 3403-3407. IEEE, 2018.

D. **Wang, C.***, Mohammed, A.K.*, Alaya Cheikh, F., Beghdadi, A., and Elle, O.J. "Multiscale deep desmoking for laparoscopic surgery". Medical Imaging 2019: Image Processing. Vol. 10949. International Society for Optics and Photonics, 2019. (* denotes equal contribution.)

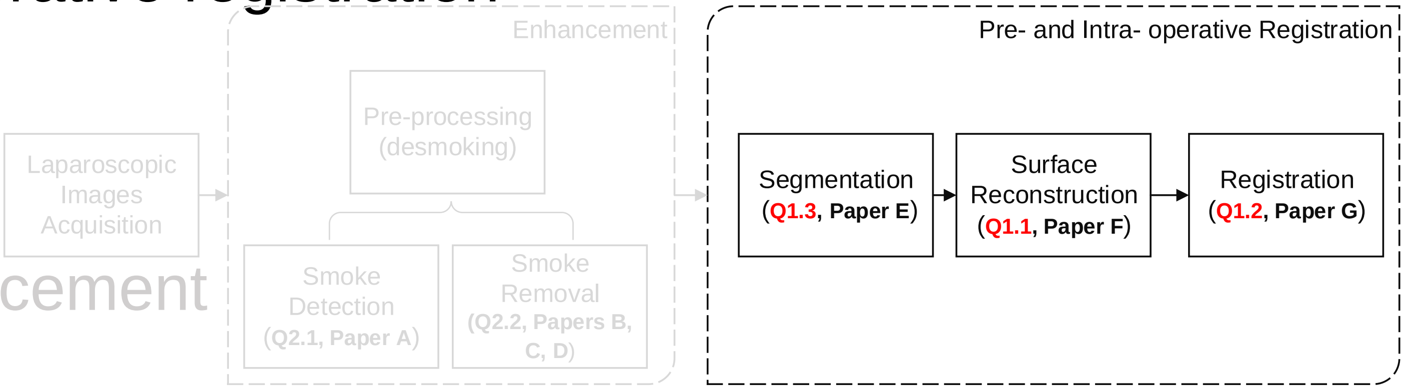
E. **Wang, C.**, Alaya Cheikh, F., Beghdadi, A. and Elle, O. J. "Adaptive context encoding module for semantic segmentation". In Electronic Imaging 2020: Image Processing: Algorithms and Systems. Society for Imaging Science and Technology, 2020.

F. **Wang, C.**, Alaya Cheikh, F., Kaaniche, M., and Elle, O. J. "Liver surface reconstruction for image guided surgery". Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling. Vol.10576. International Society for Optics and Photonics, 2018.

G. **Wang C.***,Teatini, A.*, Palomar R., Alaya Cheikh F., Beghdadi A., Edwin B., and Elle O.J. "Validation of stereo vision based liver surface reconstruction for image guided surgery." In 2018 Colour and Visual Computing Symposium (CVCS), pp. 1-6. IEEE, 2018. (* denotes equal contribution.)

Outline

- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke ramoval
- Conclusion & Future perspectives

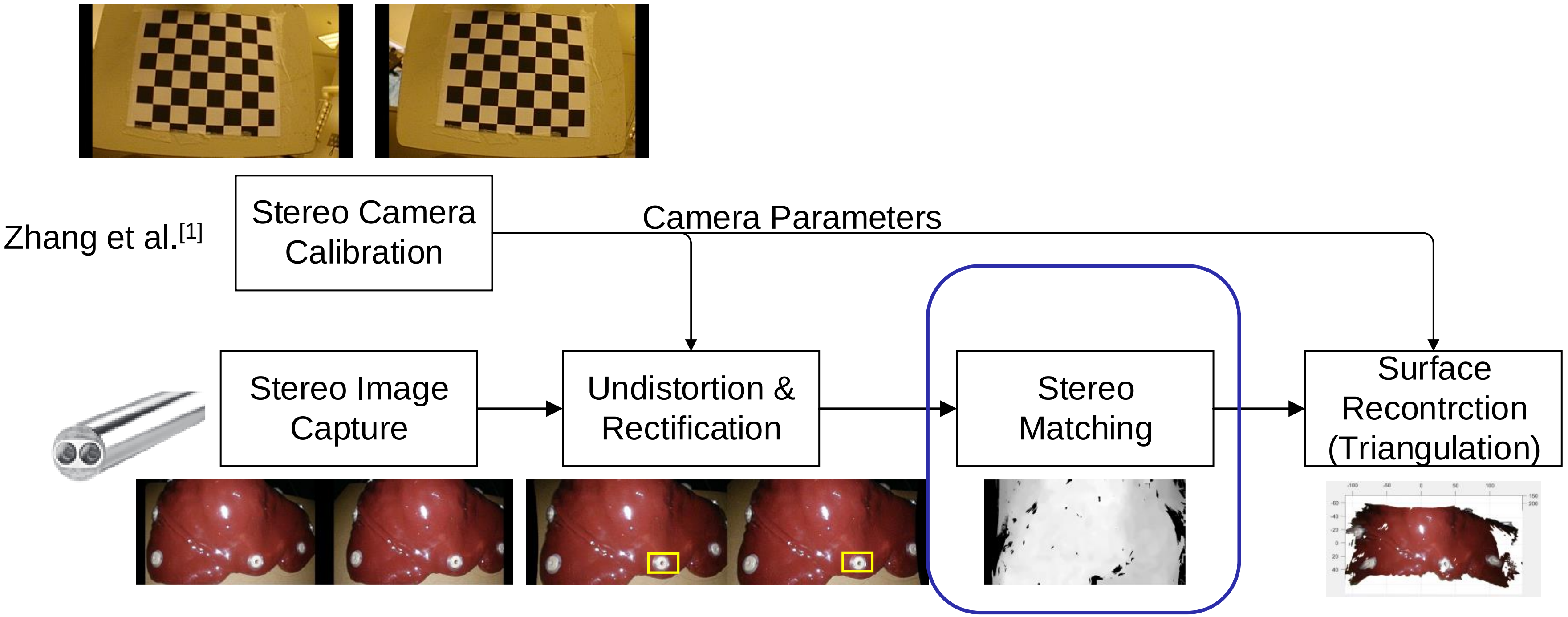


Outline

- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke ramoval
- Conclusion & Future perspectives

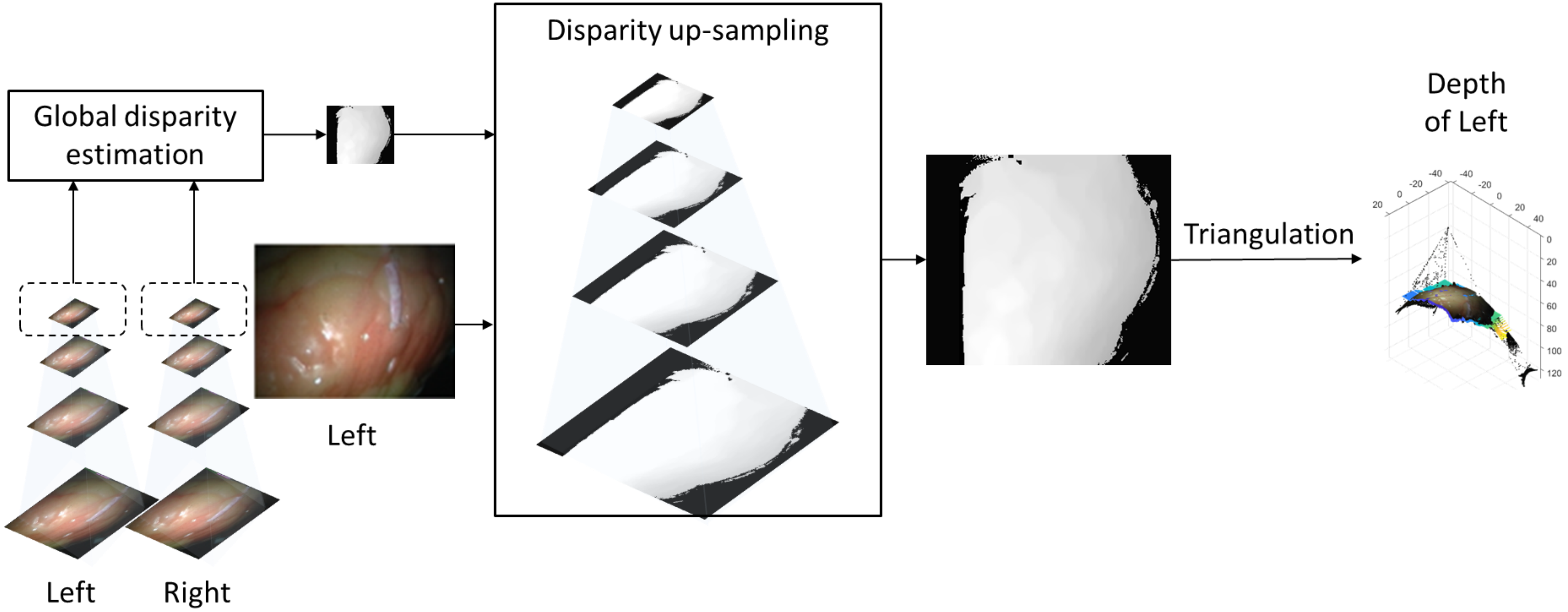
Liver Surface Reconstruction

- Depth estimation based on stereo matching
  - Workflow:



Liver Surface Reconstruction

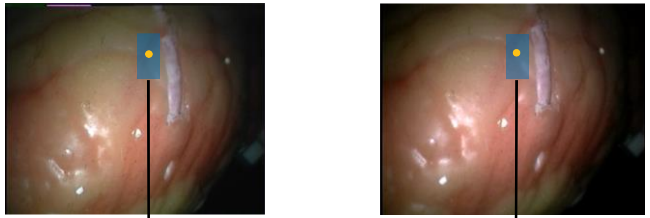
- Depth estimation based on stereo matching
  - Step 1：Global (variational) disparity estimation on down-sampled images
  - Step 2: Disparity map Up-sampling



Liver Surface Reconstruction

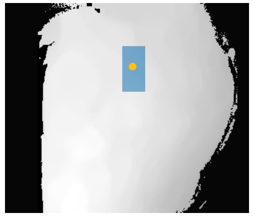
- Step1: Variational Disparity Estimation
  - Minimizing a global energy function over the entire image.
  - The gray values and gradient of the two corresponding pixels in the left and right images are the same.
    - Gray value diff.: $\left|\mathrm{I_l}(\mathrm{i},\mathrm{j}) - \mathrm{I_r}\left(\mathrm{i} - \mathrm{d_{i,j}},\mathrm{j}\right)\right|$
    - Gradient diff..: $\left|\nabla\mathrm{I_l}(\mathrm{i},\mathrm{j}) - \nabla\mathrm{I_r}\left(\mathrm{i} - \mathrm{d_{i,j}},\mathrm{j}\right)\right|$

    ➔ $\mathrm{E_{data}} = \Psi(\left|\mathrm{I_l}(\mathrm{i},\mathrm{j}) - \mathrm{I_r}\left(\mathrm{i} - \mathrm{d_{i,j}},\mathrm{j}\right)\right| + \left|\nabla\mathrm{I_l}(\mathrm{i},\mathrm{j}) - \nabla\mathrm{I_r}\left(\mathrm{i} - \mathrm{d_{i,j}},\mathrm{j}\right)\right|)$
  - The organ surface is smooth, so we assume that the disparity is smoothly distributed.
    - Pixel wise: $\mathrm{E}_{smooth} = \Psi(|\nabla d|)$
    - Non-local: $\mathrm{E}_{non\_local} = \sum_{(i,j)\in\Omega_d}\sum_{(i',j')\in N_{i,j}}|d_{i,j} - d_{i'j'}|$



Pixel p_l (i,j)

Pixel p_r (i-$d_{i,j}$,j)



Disparity map

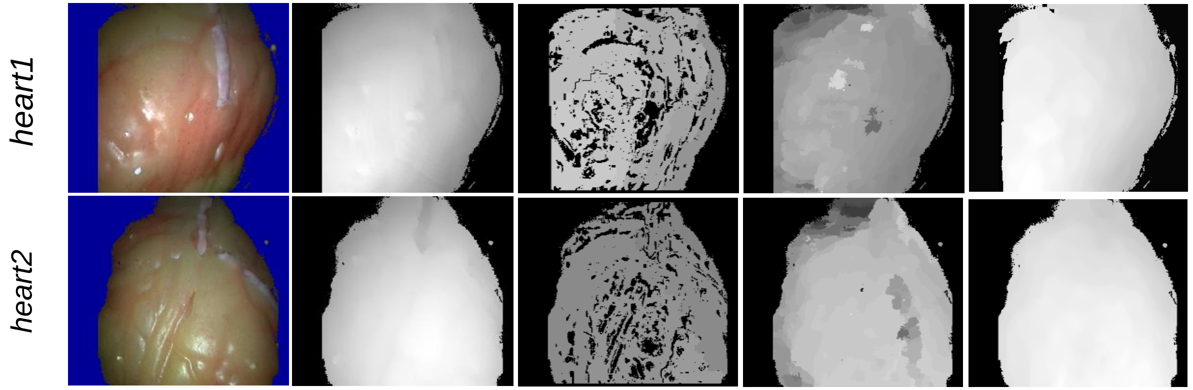
➔ $\min\limits_{\mathbf{d}} \mathbf{E}(\mathbf{d})$,

where $E(d) = \mathrm{E}_{data} + \lambda_s\mathrm{E}_{smooth} + \lambda_{nl}\mathrm{E}_{non\_local}$, $\Psi(x^2) = \sqrt{x^2 + \epsilon^2}$

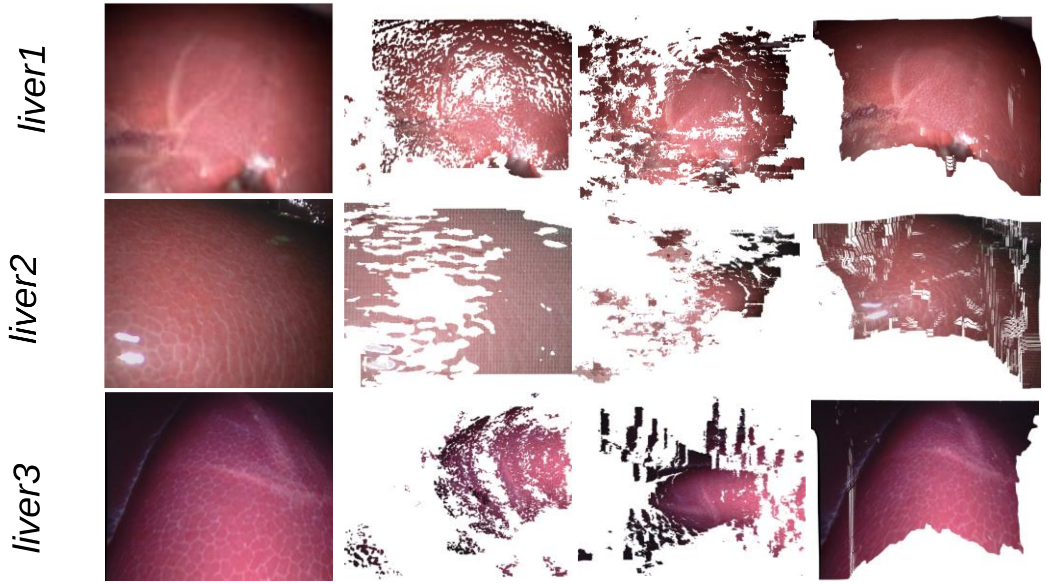
Liver Surface Reconstruction

Dataset

- Cardiac phantom datasets with ground truth: *heart1* and *heart2*.
- Porcine liver datasets without ground truth: *liver1*, *liver2* and *liver3*.



(a) Original image (b) Ground truth (c) QuasiDense [3] (d) FCVF[4] (e) Proposed

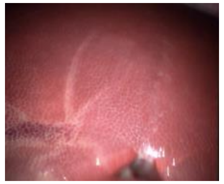


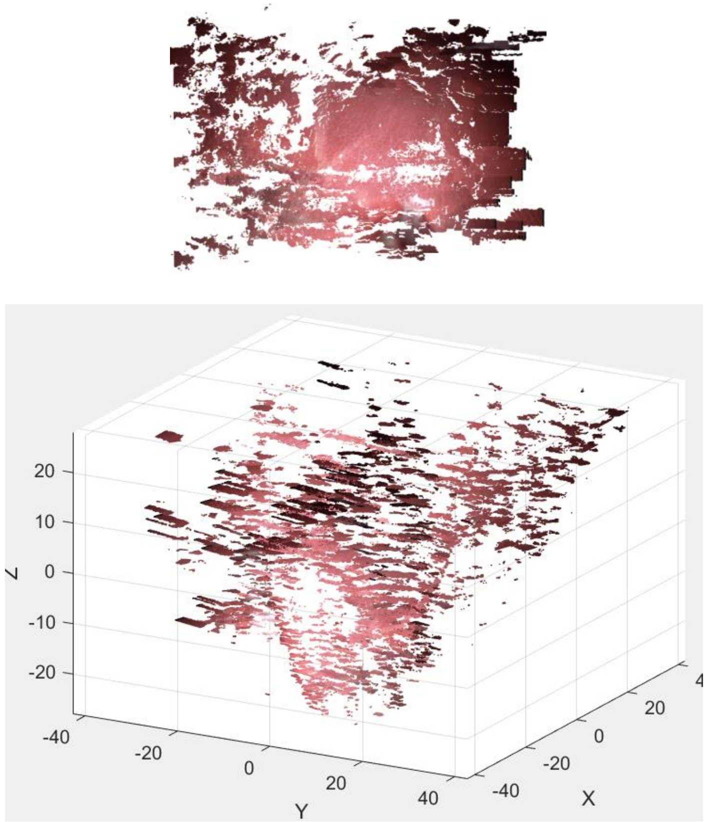
(a) Original image (b) QuasiDense [3] (c) FCVF[4] (d) Proposed

	heart1		*heart2*	
Methods	MAE(mm)	% Match	MAE(mm)	% Match
Proposed	**2.16 ± 0.65**	**97.25 ± 1.13**	**2.14 ± 0.83**	**99.96 ± 0.11**
QuasiDense[3]	2.33 ± 0.61	78.64 ± 2.00	2.26 ± 0.52	80.64 ± 1.87
FCVF[4]	4.43 ± 0.81	95.96 ± 1.57	4.21 ± 1.20	88.92 ± 2.38

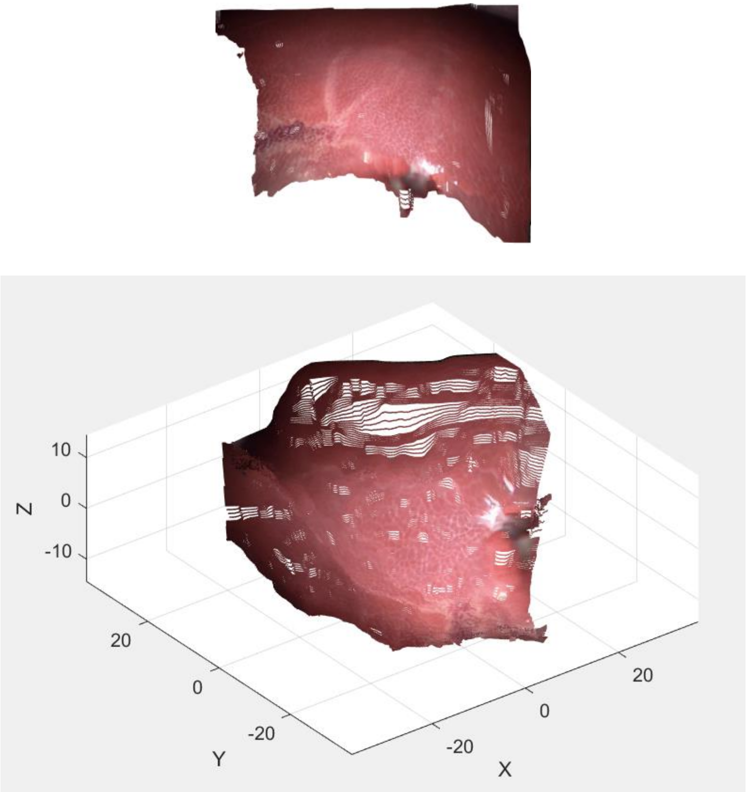
MAE (mm) and % Match for phantom datasets *heart1* and *heart2*.

Liver Surface Reconstruction





(c) FCVF[4]



(d) Proposed

Test Dataset 1 – Mean per-pixel error (mm)

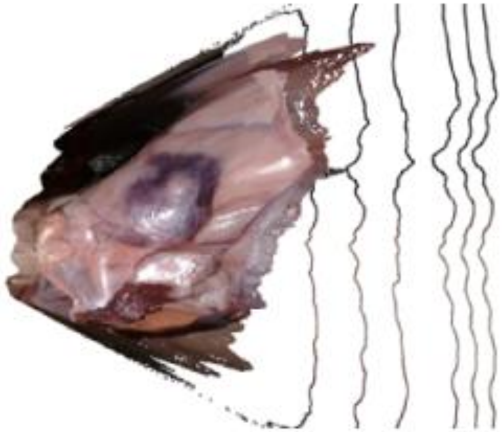
Team Name	Keyframe 1	Keyframe 2	Keyframe 3	Keyframe 4	Keyframe 5	Average
Congcong Wang	6.30	2.15	3.41	3.86	4.80	**4.104**
Jean-Claude Roshenthal	8.25	3.36	2.21	2.03	1.33	**3.436**
KeXue Fu	30.49	18.32	19.73	19.30	16.86	20.94
Trevor Zeffiro	7.91	2.97	1.71	2.52	2.91	**3.604**
Wenyao Xia	5.70	7.18	6.98	8.66	5.13	6.73
Zhu Zhanshi	14.64	7.77	7.03	7.36	11.22	9.604
Huoling Luo	29.68	16.36	13.71	22.42	15.43	19.52
Xiran Zhang	12.53	6.13	3.60	3.34	5.07	6.134
Xiaohong Li	34.42	20.66	17.84	27.92	13.00	22.768
Lalith Sharan	30.63	46.51	45.79	38.99	53.23	43.03

Slides from the challenge organizer- Intuitive

Test Dataset 1 – Keyframe 1


Ground Truth


Huoling Luo


Xiaohong Li


Congcong Wang


Trevor Zeffiro

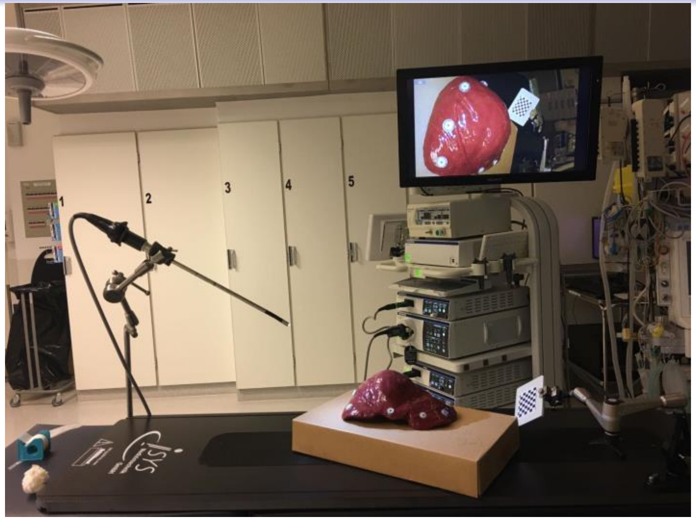

Jean-Claude Rosenthal

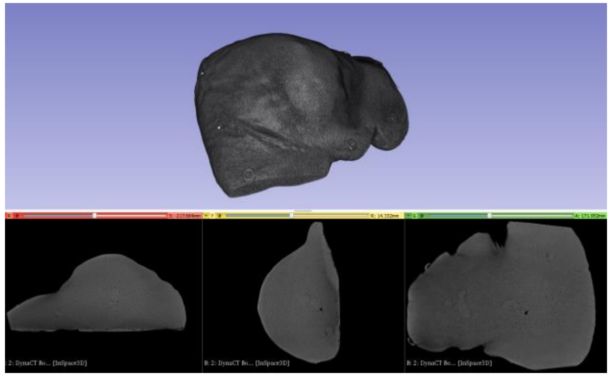
Outline

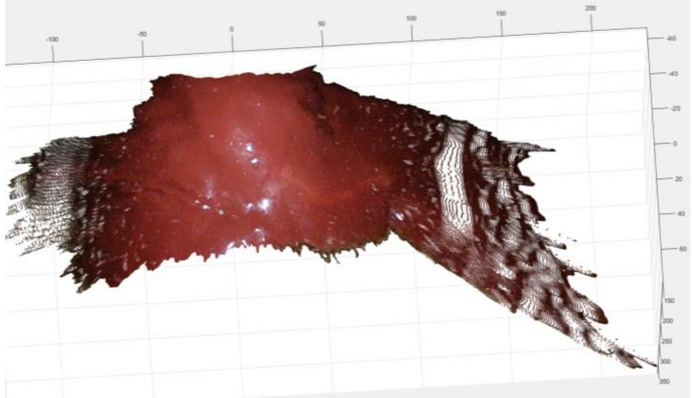
- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke ramoval
- Conclusion & Future perspectives

Surface Based Registration

Patient specific liver phantom

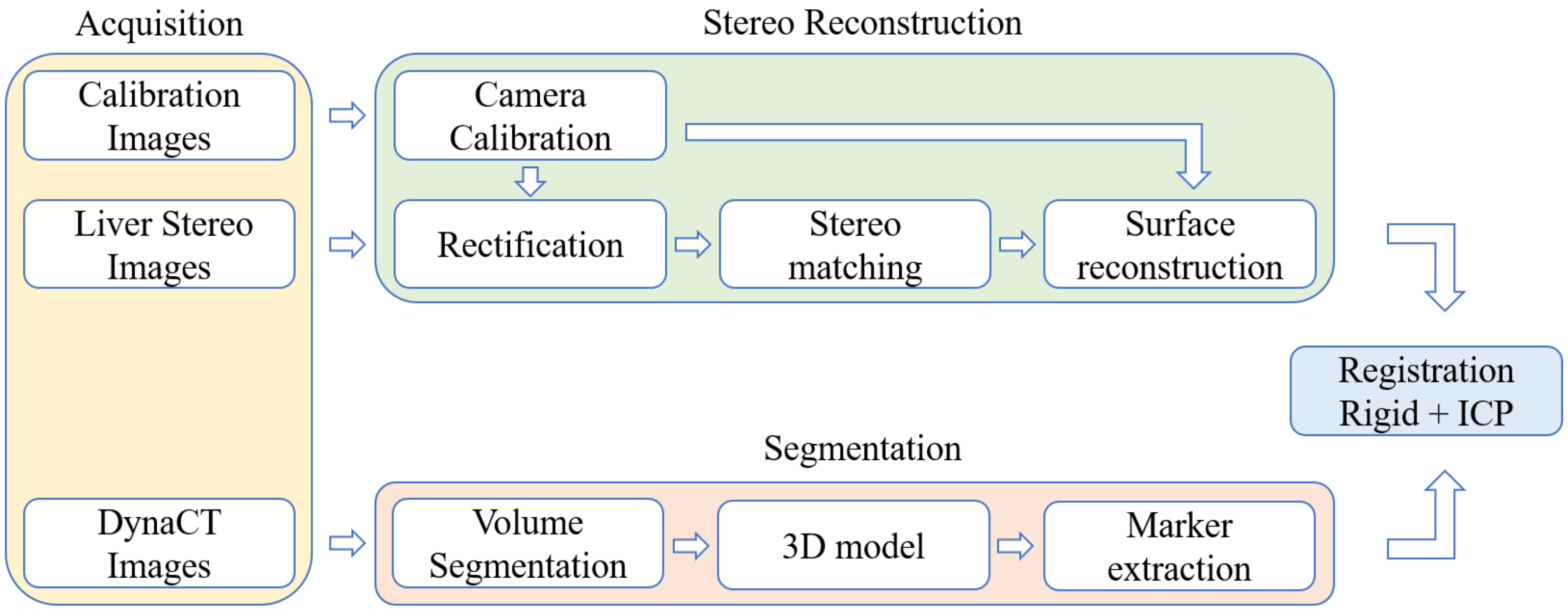






3D CT reconstruction ⟺ 3D stereo-reconstruction

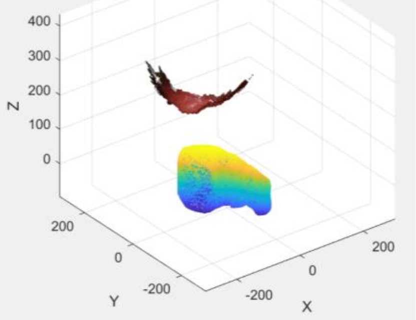
Surface Based Registration



- Registration via surface solely is not effective
- Markers to register CT and Stereo.
- Refine through Iterative Closest Points (ICP) [34] registration.

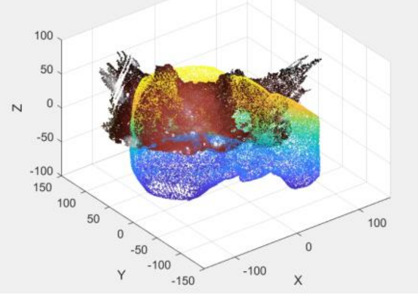
Before registration

(a)



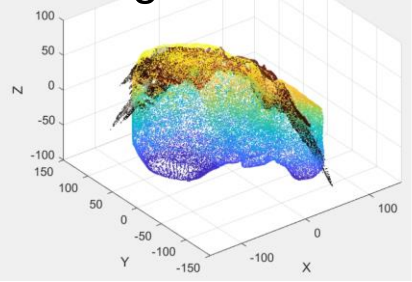
Surface registration by ICP

(b)



Rigid + ICP

(c)

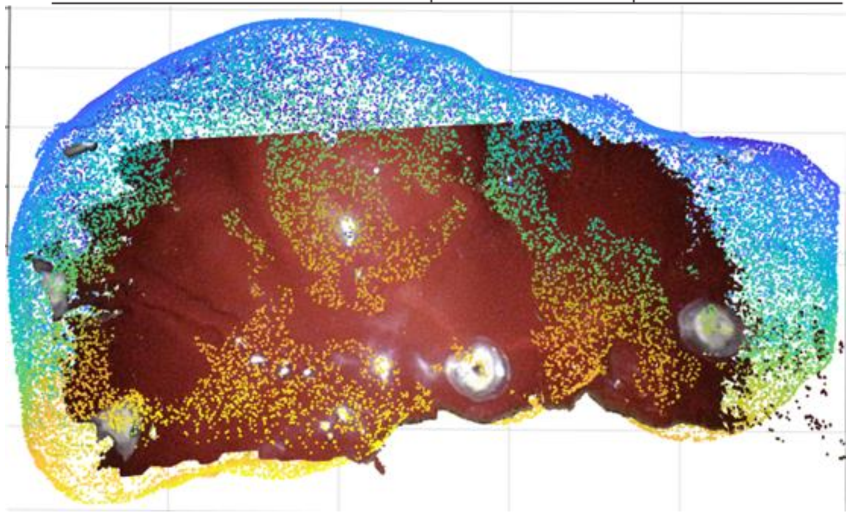


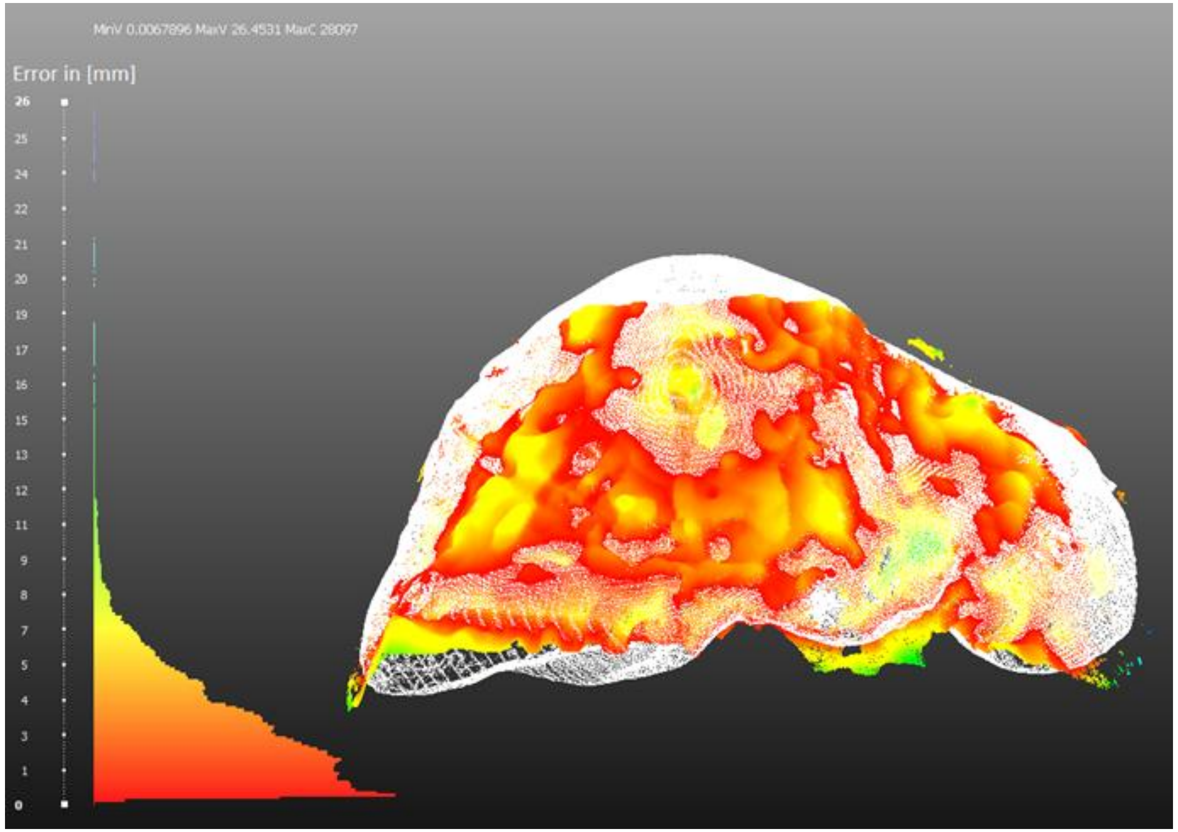
Surface Based Registration

- MAE and Hausdorff were tested on 2 datasets, (*Dataset1* and *Dataset2)* of 15 surface reconstructions for a total of 30 reconstructions.

TABLE 1. MAE AND HAUSDORFF IN [MM] IN TERMS OF MEAN μ, STANDARD DEVIATION σ AND MAXIMA FOR *Dataset1* AND *Dataset2*.

	Dataset1	*Dataset2*
MAE ($\mu \pm \sigma$)	4.6±1.0	4.4 ± 0.8
max_{MAE}	128.8	105.2
Hausdorff ($\mu \pm \sigma$)	3.7±0.8	3.6 ± 0.8
max_H	78.5	106.6





Discussion

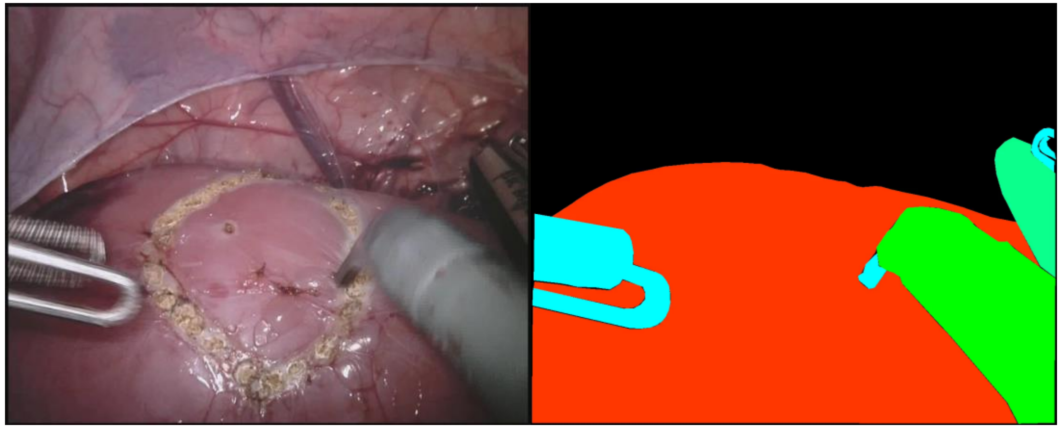
- Contributions
  - Simulation framework is successfully established
  - Experimentally validated and analyzed the surface reconstruction method of **Paper F**
- Discussion
  - Small patch surface is not effective for registration
    - Need initial registration (Marker based in **Paper G**)
    - Need larger surface, e.g. SLAM (simultaneous localization and mapping), tracked camera, etc.
  - Segment the images manually via Graph Cut method in **Paper G**
    - Automatic semantic segmentation is essential

Outline

- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke ramoval
- Conclusion & Future perspectives

Semantic Segmentation

- Semantic Segmentation
  - Understanding an image at pixel level
  - Partitioning an image into regions of meaningful objects
  - Comprehensive scene description: object category, location, etc.

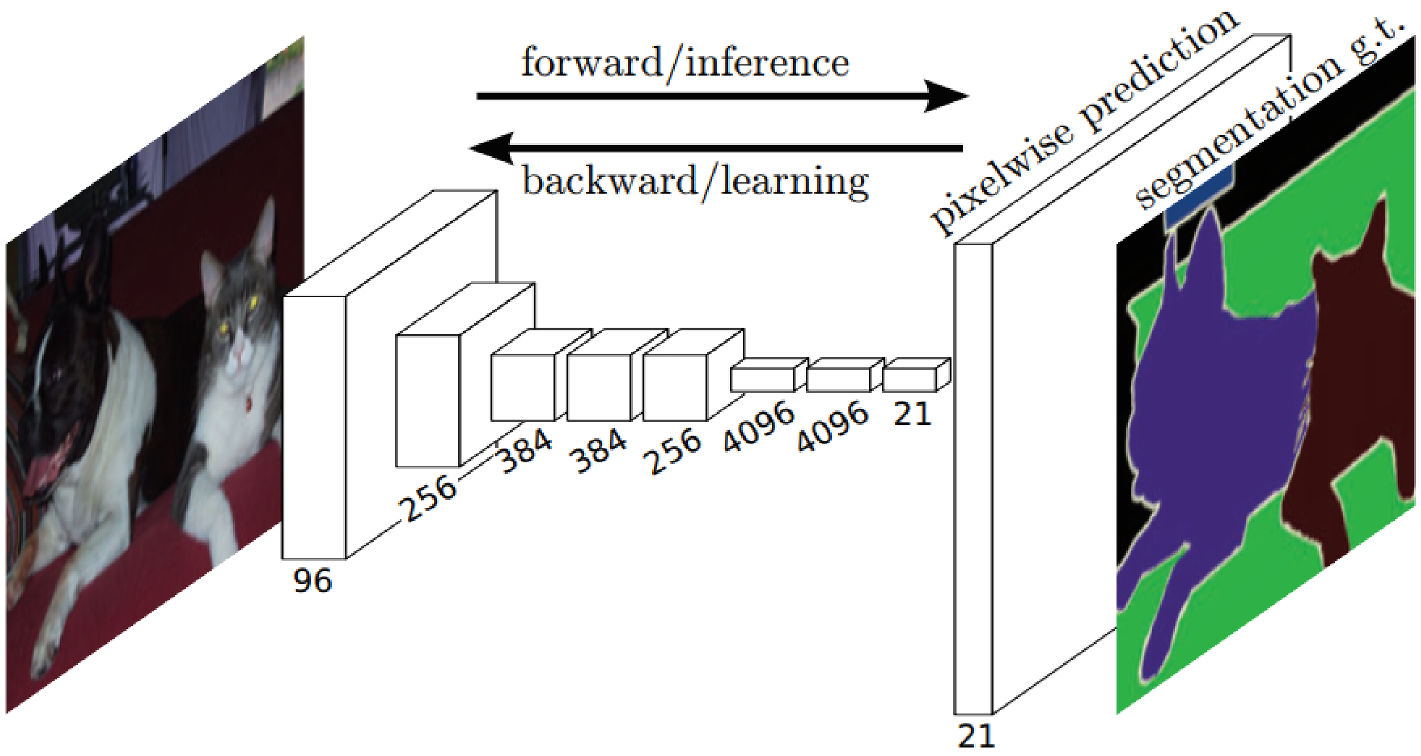


Allan et al.[24]

Zhang et al. [27]

Semantic Segmentation

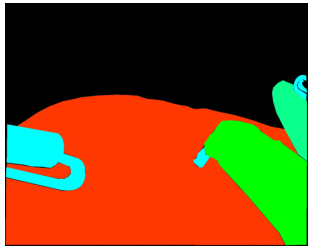
- Fully Convolutional Network (FCN)[5]



- Pre-trained CNN + Decoder
- Trained end-to-end
- → Meta algorithm for semantic segmentation

Semantic Segmentation

- What problem we are addressing?
  - The existence of objects at multiple scales.



Allan et al.[24]

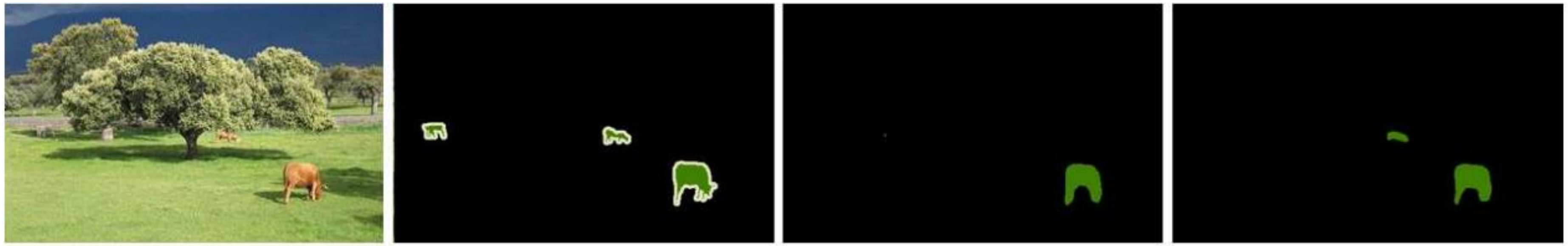


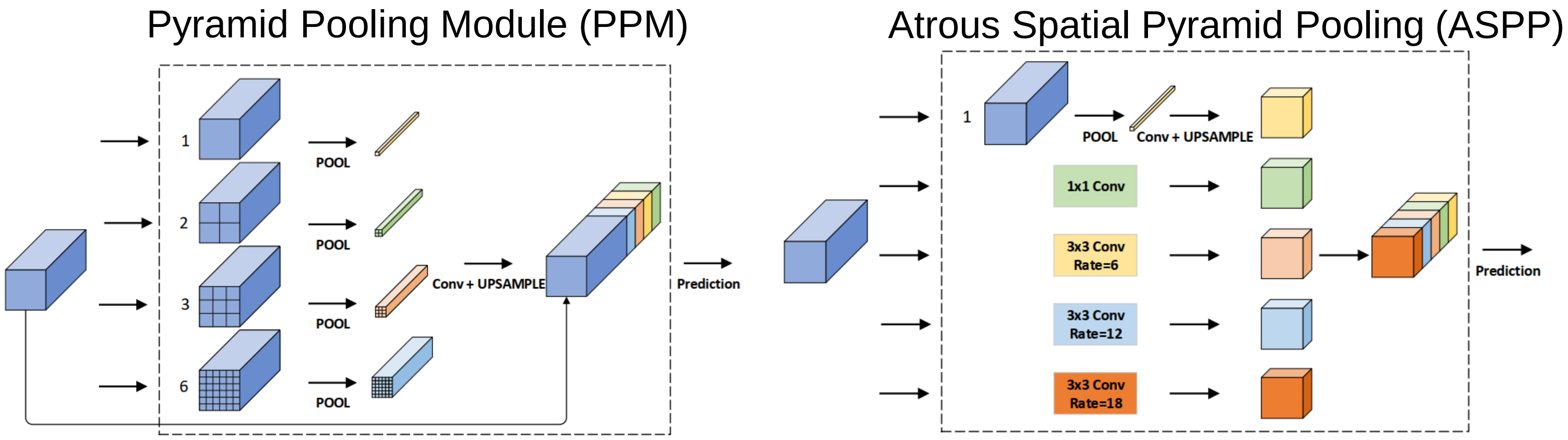
Image | Ground Truth | Result obtained by single scale | Result obtained by multi-scale

He et al. [6]

  - How to capture multiple scale information?

Semantic Segmentation

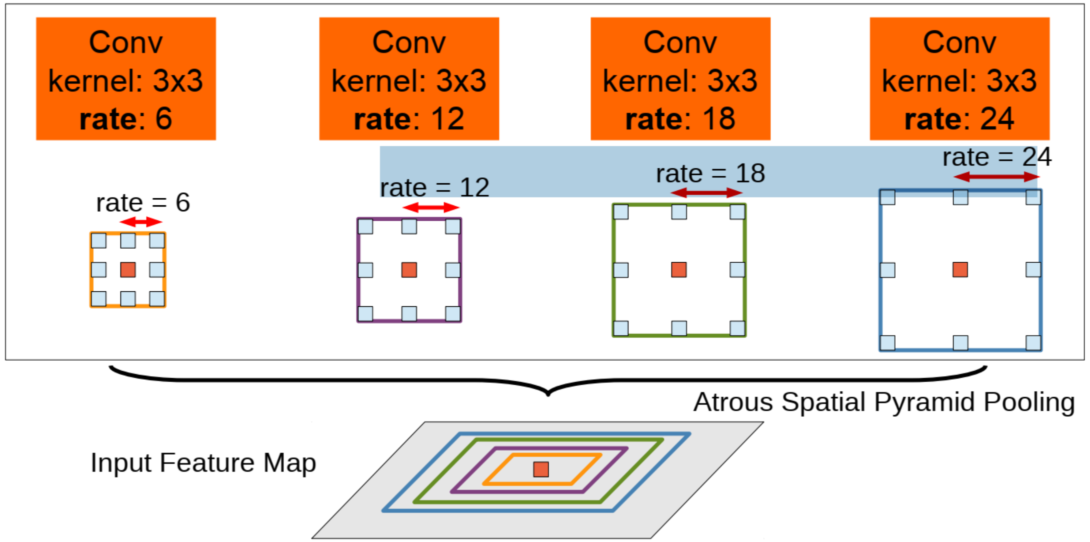
- Prior work



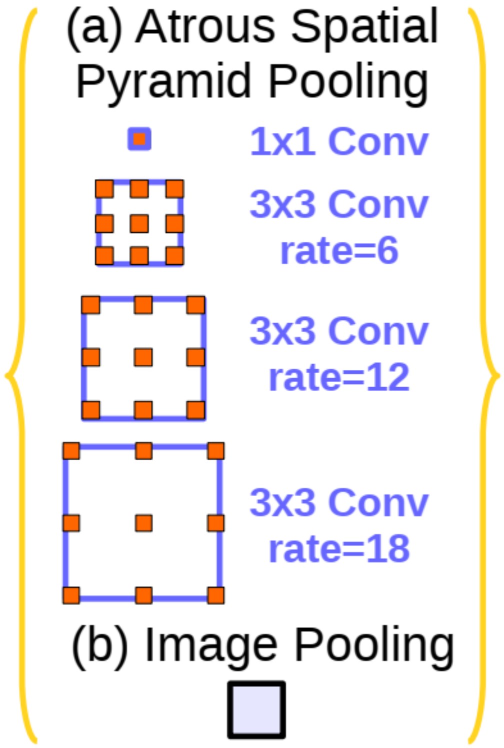
  - Two problems:
    - (1) The numbers of sub-region of PPM in PSPNet[33] and the atrous rates of ASPP module from Deeplabs[34] are selected empirically.
    - (2) PPM and ASPP both extract the context information by sampling from rigid rectangular regions which contain pixels from different object categories.
- How to solve the above two problem?

Semantic Segmentation

- Rethinking ASPP
  - Problem: Existence of objects at multiple scales
  - Solution: Different values of atrous rate → different sample locations of the convolution operation → multiple effective fields of view → capturing multi-scale context information

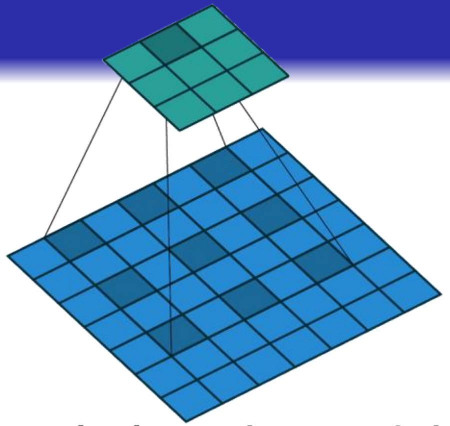


Chen et al. Deeplabv2

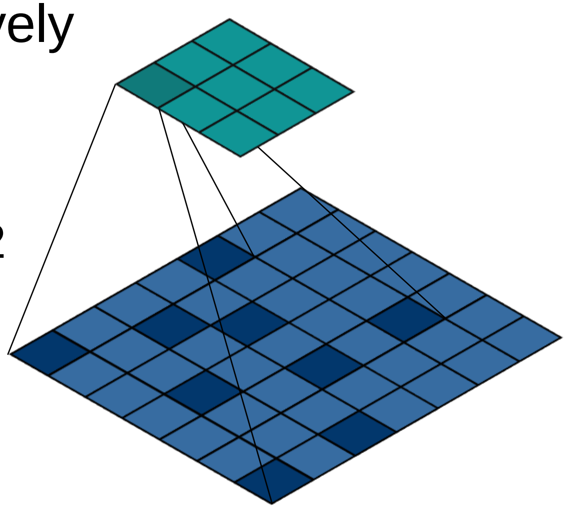


Chen et al. Deeplabv3

Semantic Segmentation



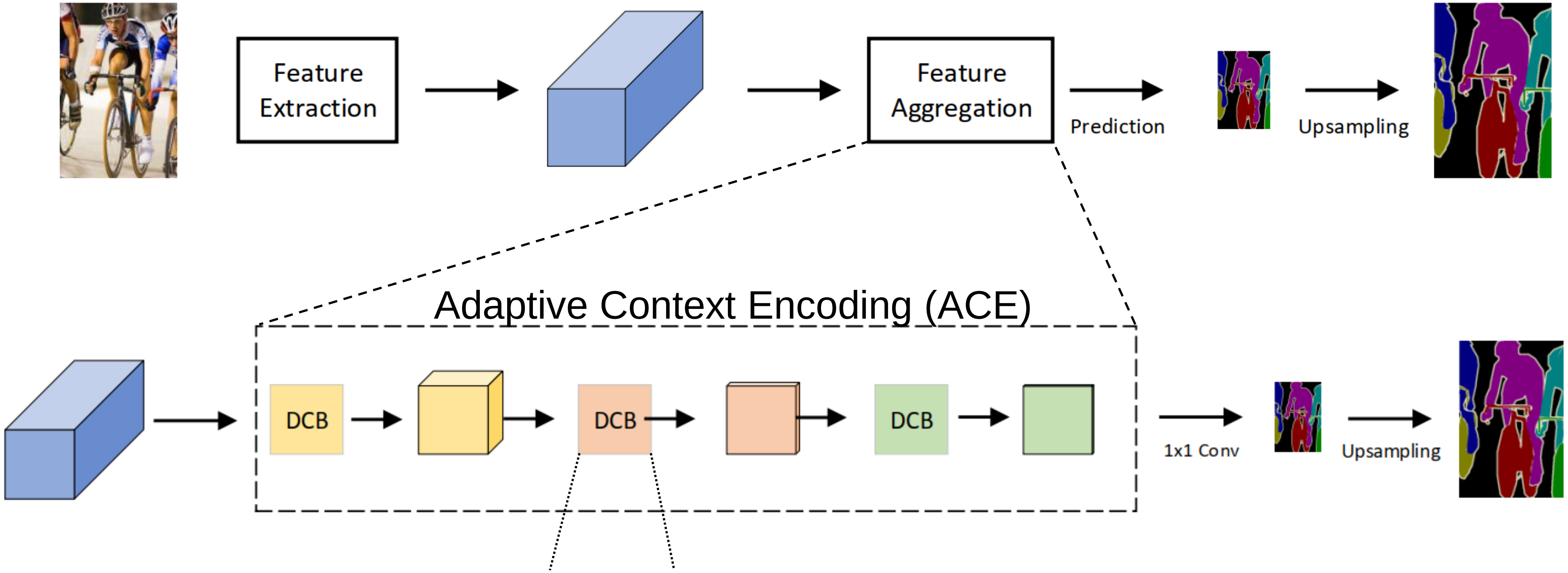
- Main idea
  - Problem: existence of objects at multiple scales
  - Solution of ASPP: Different values of atrous rate → different sample locations of the convolution operation → multiple effective fields of view → capturing multi-scale context information
  - Our idea: Learned sample locations of the convolution operation → learned effective fields of view → capturing multi-scale context information adaptively
  - Learn the sample locations
    - No need to decide the atrous rate manully – solving issue 1
    - No need to sample the pixel in a rectangular region– solving issue 2
    - Tool: Deformable convolution



Dai et al., Deformable ConvNets v1[10]
Zhou et al., Deformable ConvNets v2[11]

Semantic Segmentation

- Main idea
  - Network Architecture



Deformable convolution → BN → ReLU

Semantic Segmentation

- Comparison to PSP and Deeplab

  - Pascal-Context
    - 4998 training images
    - 5105 testing images
    - 59 object classes with background

  - ADE20k
    - 150 object classes
    - 20k images for training
    - 2k/3k images validation and testing

Batch Size	**Head**	**pixAcc%**	**mIoU%**
4	ASPP	75.42	43.62
	PPM	75.58	45.68
	Proposed	**77.68**	**48.07**
6	ASPP	77.19	46.53
	PPM	77.45	48.32
	Proposed	**78.35**	**49.36**
16	ASPP	78.68	49.04
	PPM	78.41	49.54
	Proposed	**78.85**	**50.35**

Batch Size	**Head**	**pixAcc%**	**mIoU%**
4	ASPP	78.11	37.11
	PPM	77.39	37.80
	Proposed	**78.62**	**38.51**

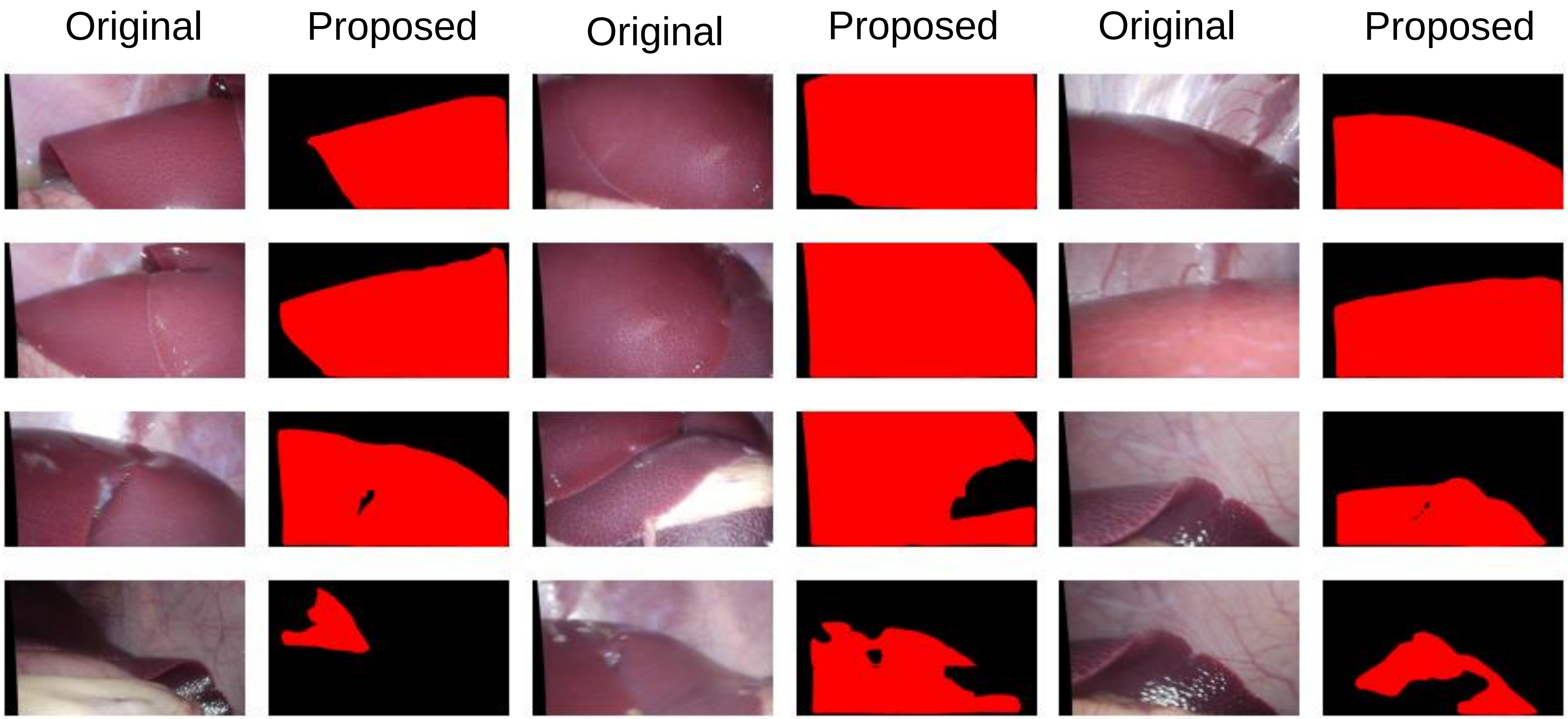
Semantic Segmentation

- Some visual results

Original Images | Ground Truth | Deeplabv3* | Proposed

Semantic Segmentation

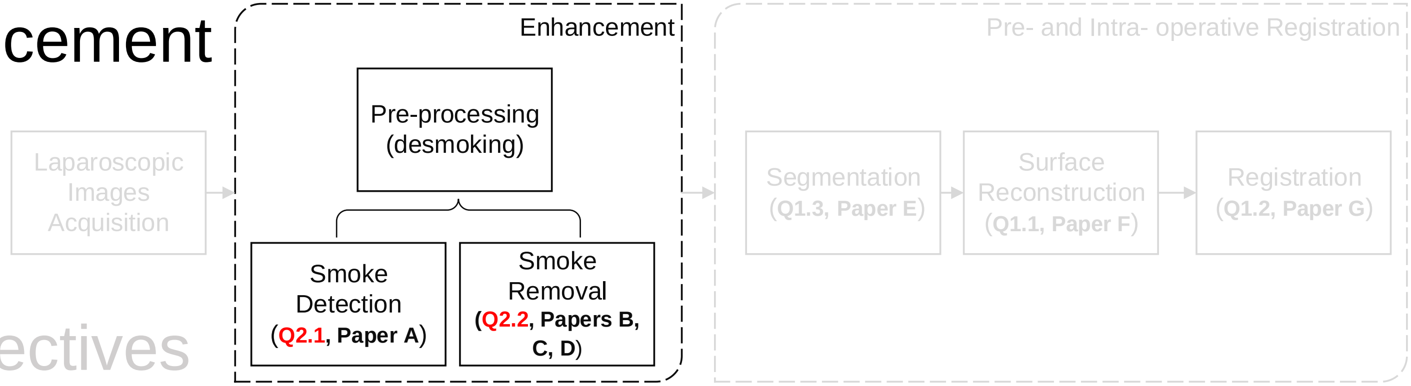
- Some visual results



Original images are from Oslo University Hospital

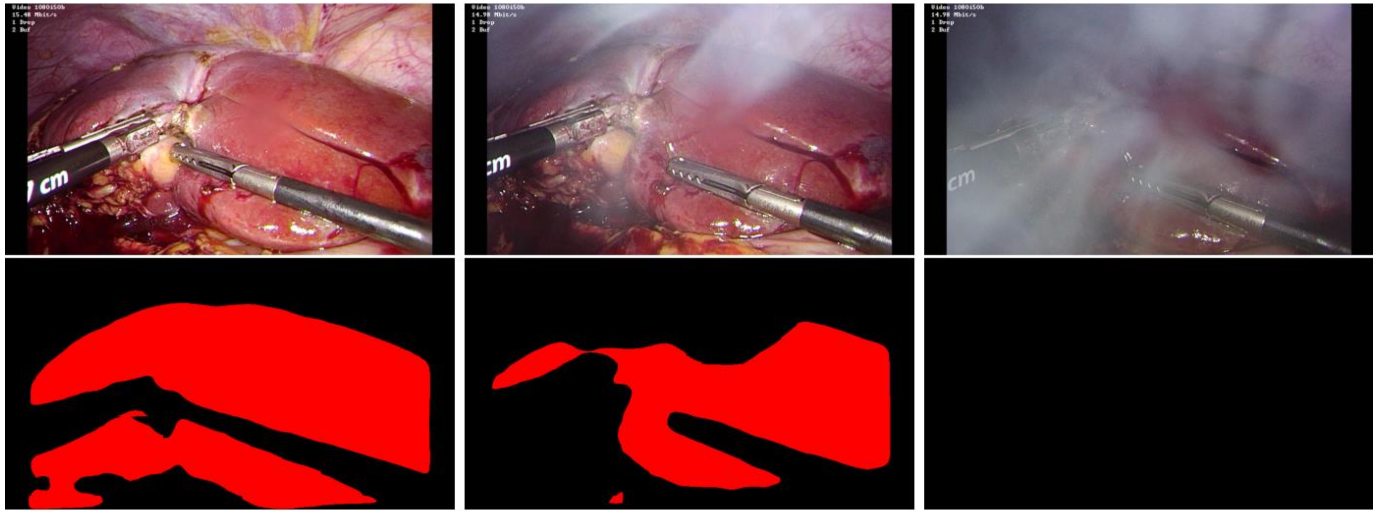
Outline

- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke removal
- Conclusion & Future perspectives

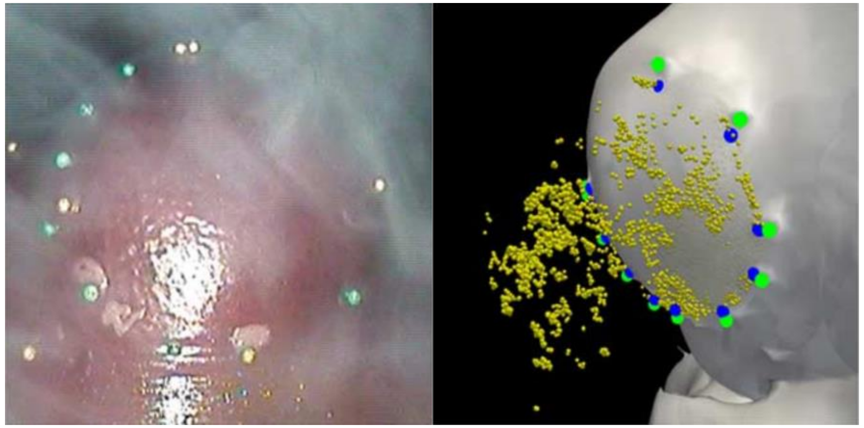


Why We Need Enhancement?

- Smoke degrades laparoscopic video quality.
  - Influences surgeon’s visibility
  - Influences the performance of computer vision based navigation systems



Original images are from Oslo University Hospital



Maier-Hein et al. [12]

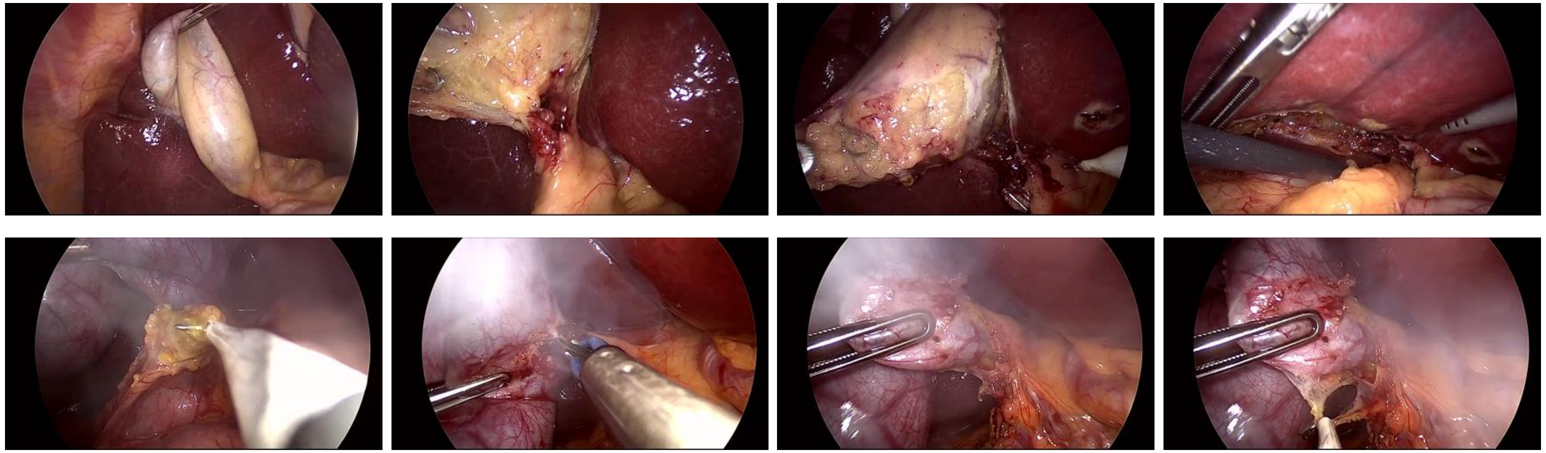
Smoke Detection

- Desmoking techniques
  - Computer vision algorithms
  - Smoke evacuation techniques

- When to start to remove smoke?
  - Smoke/non smoke images classification

0 - - non smoke

1 - - smoke

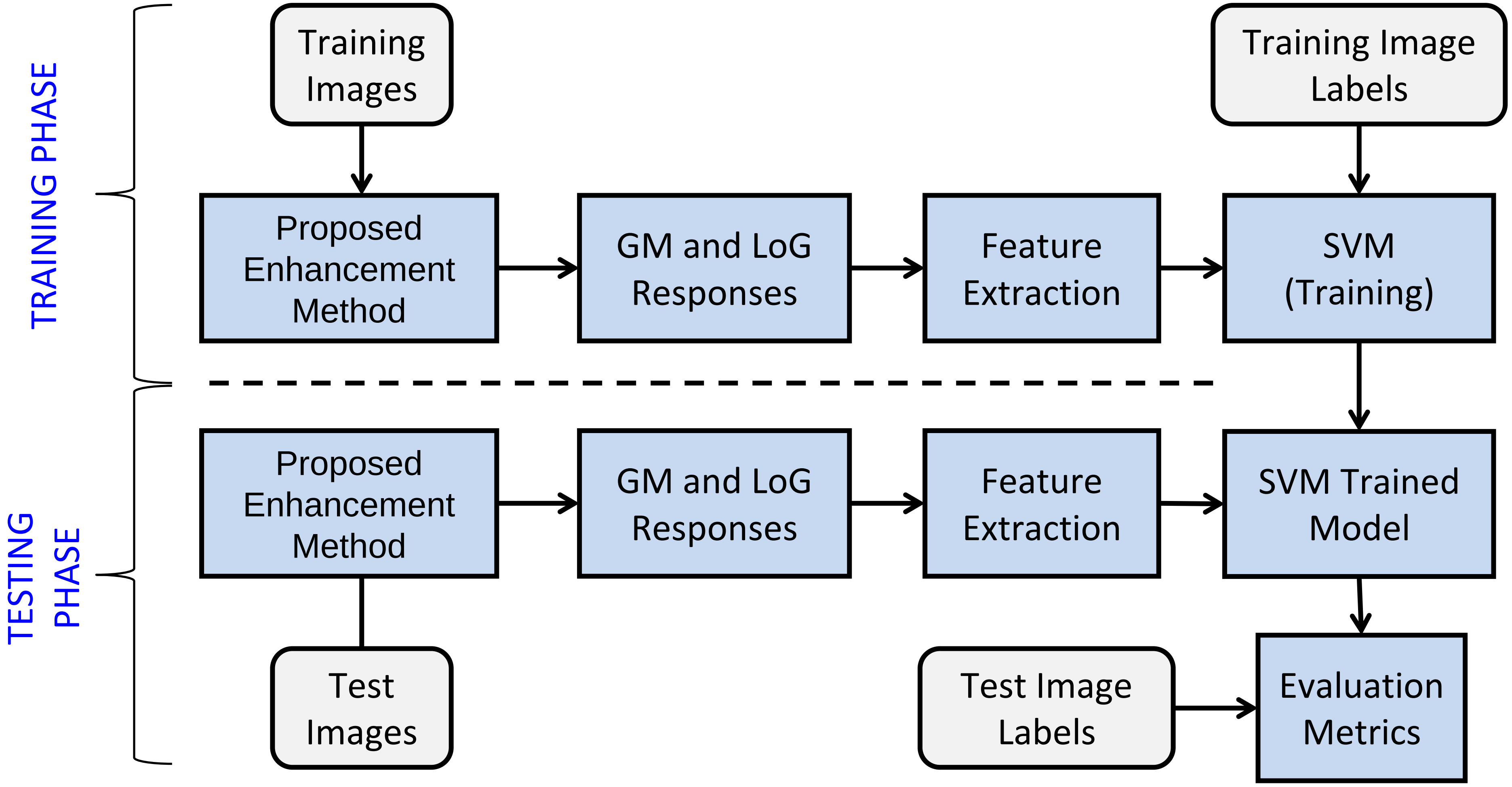


Outline

- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke removal
- Conclusion & Future perspectives

Smoke Detection

- Main idea



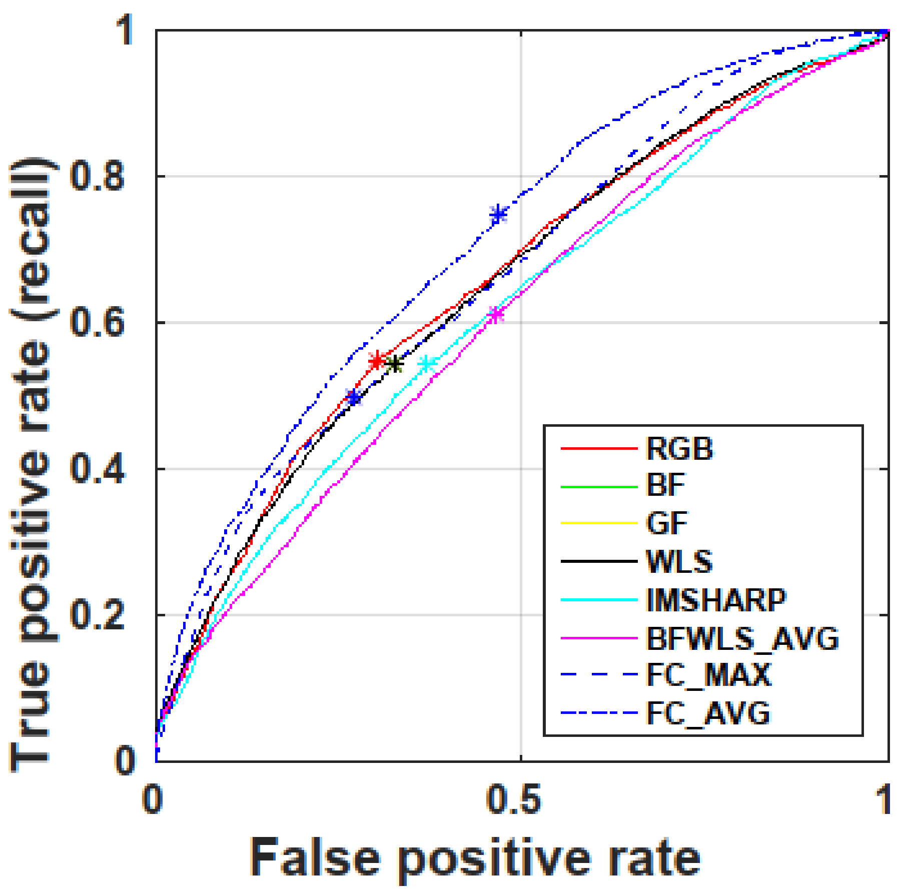
Smoke Detection

Method	Accuracy	F1-Score
RGB	0.60	0.60
IMSHARP	0.58	0.58
BF	0.60	0.59
GF	0.60	0.59
WLS	0.60	0.59
BFWLS_AVG	0.57	0.56
FC_MAX(Ours)	0.60	0.59
FC_AVG(Ours)	**0.64**	**0.64**

Tab.1 Comparison with the baseline RGB images and other enhancement methods

Method	Accuracy	F1-Score
SPA	0.63	0.58
SAN	0.63	0.59
FC_AVG(Ours)	**0.64**	**0.64**

Tab.2 Comparison with the saturation histogram based classification methodologies Saturation Analysis (SAN) and Saturation Peak Analysis (SPA)

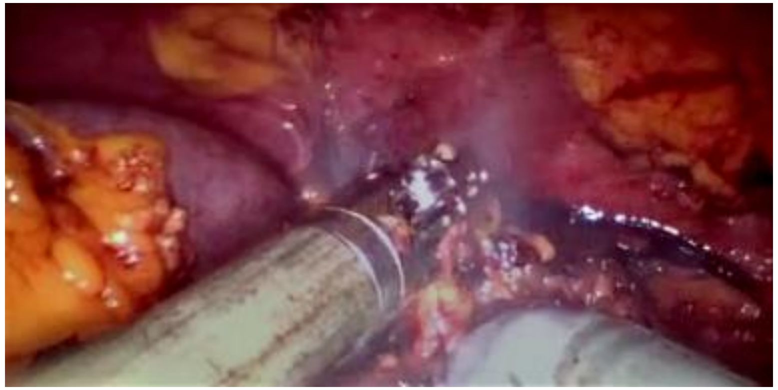


Surgical Smoke Removal

- Smoke removal methods
  - Mechanical solutions
  - Image processing based approaches

Smoke image

Desmoked image

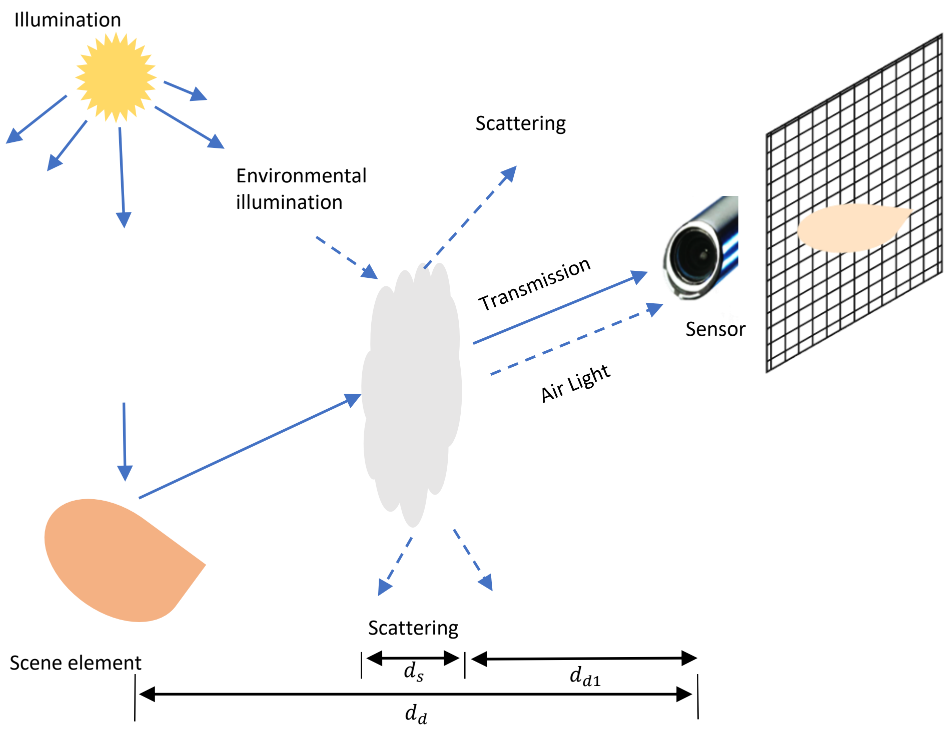


Outline

- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke removal
- Conclusion & Future perspectives

Surgical Smoke Removal

- Physical model of smoke image acquisition based on atmospheric scattering model



Smoke model

$$I = L + F = J_s t_s + A_s (1 - t_s)$$

Medium Transmission

Smoke Image

Radiance/smoke -free image

illumination

- *direct attenuation (Transmission)*

$$\mathbf{L} = g\frac{\eta\rho(\lambda)e^{-\beta(\lambda)d_s}}{d_d^2} = \mathbf{J}_s e^{-\beta(\lambda)d_s} = \mathbf{J}_s t_s$$

- *smoke veil (Air Light)*

$$\mathbf{F} = \int_{d_{d1}}^{d_{d1}+d_s} g\eta\beta(\lambda)e^{-\beta(\lambda)d} dd = ge^{-\beta(\lambda)d_{d1}}\eta(1 - e^{-\beta(\lambda)d_s})$$
$$= \mathbf{A}_s(1 - e^{-\beta(\lambda)d_s}) = \mathbf{A}_s(1 - t_s),$$

Surgical Smoke Removal

- Physical model of smoke image acquisition

Medium Transmission

Direct attenuation

$$\mathbf{I} = \mathbf{L} + \mathbf{F} = \boxed{\mathbf{J}_s t_s} + \boxed{\mathbf{A}_s(1 - t_s)}$$

Smoke veil

Smoke Image

smoke-free image

illumination

- A_s -- depends on the illumination property and the distance where smoke appears
- t_s -- the smoke thickness

hard ill-posed problem

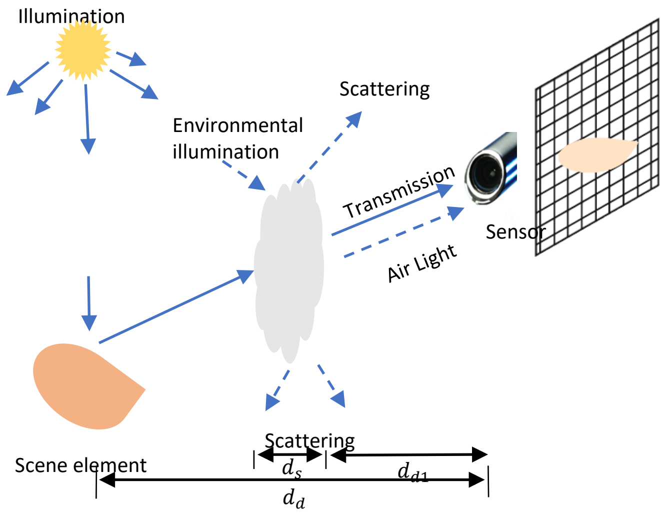
- (1) Estimate *smoke veil* **F**
- (2) Estimate J_s from *direct attenuation* **L**

Variational Smoke Removal

- (1) Estimate *smoke veil* F

$$\mathbf{I} = \mathbf{L} + \mathbf{F} = \mathbf{J}_s t_s + \boxed{\mathbf{A}_s(1 - t_s)}$$

Smoke veil **F**



- Mainly a function of the **properties of illumination** and **smoke thickness**.
  - Illumination is smoothly distributed
  - Smoke thickness may change when there is large depth jump of the scene

➔ Assumption 1 (strong) of the properties of *smoke veil*: Smoke veil is smoothly distributed except in regions exhibiting high scene depth changes

(a) F has a low contrast

- Scattering and transmission properties can assumed to be independent from wavelength.
  - Smoke veil value is added to the RGB channels equally

➔ Assumption 2: *smoke veil*'s RGB channels' intensity are equal

(b) F has low inter-channel differences

Variational Smoke Removal

- (1) Estimate *smoke veil* F $\mathbf{I} = \mathbf{L} + \mathbf{F} = \mathbf{J}_s t_s + \mathbf{A}_s(1 - t_s)$

$$E = \frac{\gamma}{2}\|\mathbf{F} - \mathbf{I}\|^2 + \|\mathbf{F}_{TV}\|_2, \quad \text{where } \|\mathbf{F}_{TV}\|_2 = \sum_i \sqrt{\boxed{\theta_x^2[\mathbf{D}_x\mathbf{F}]_i^2 + \theta_y^2[\mathbf{D}_y\mathbf{F}]_i^2} + \boxed{\theta_c^2[\mathbf{D}_c\mathbf{F}]_i^2}}$$

- Regularization term
  - (a) F has a low contrast: low derivative value of **F** with respect to variable *x* and *y*
  - (b) F has low inter-channel differences: low derivative value of **F** with respect to variable *c*

Original

Smoke veil **F**

Original

Smoke veil **F**

Variational Smoke Removal

- (2) Estimate J_s from *direct attenuation* L

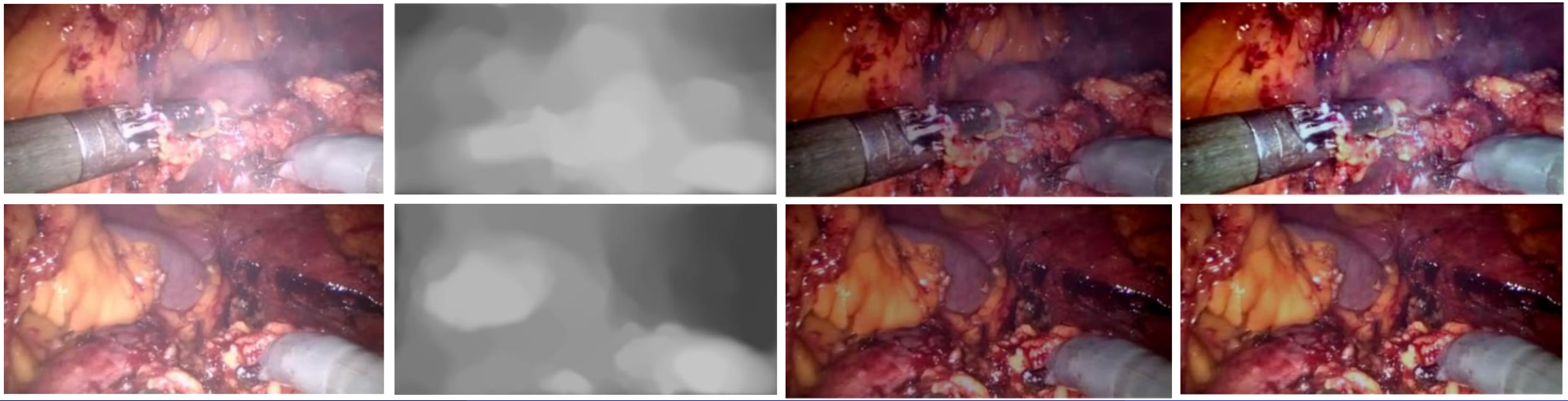
$$\mathbf{I} = \mathbf{L} + \mathbf{F} = \mathbf{J}_s t_s + \mathbf{A}_s(1 - t_s) \quad \Rightarrow \quad \mathbf{L}(x, y, c) = \mathbf{I}(x, y, c) - \alpha(c) \cdot \mathbf{F}(x, y, c)$$

$$\mathbf{L} = g\frac{\eta\rho(\lambda)e^{-\beta(\lambda)d_s}}{d_d^2} = \mathbf{J}_s e^{-\beta(\lambda)d_s} = \mathbf{J}_s t_s$$

J_s is attenuated exponentially with the thickness of smoke

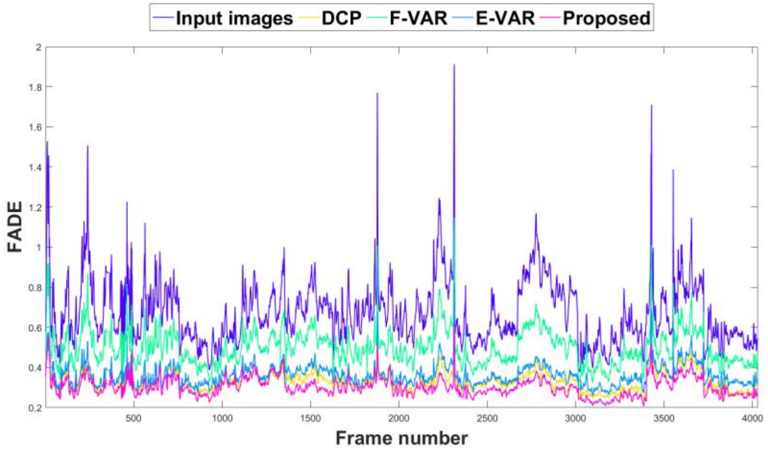
The depth range of surgery scene is limited → the variation range of t_s would be small (strong assumption)

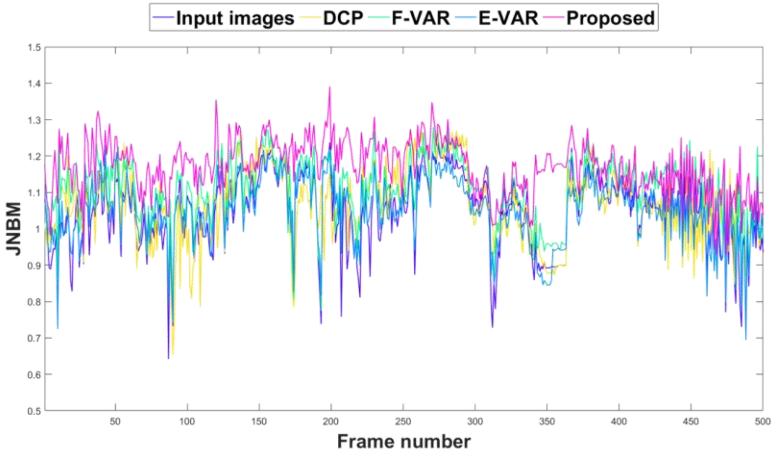
➔ Linear transformation of **L** to [0;255]

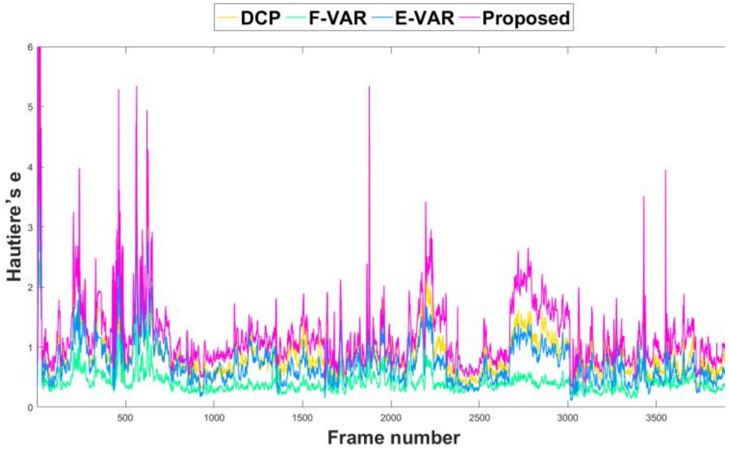


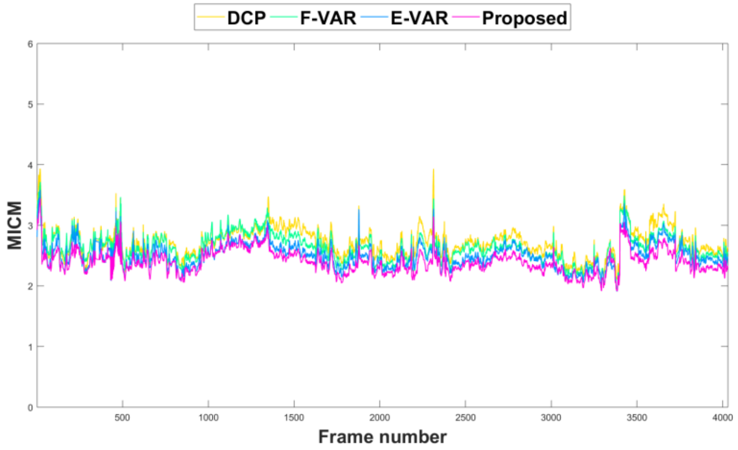
Variational Smoke Removal

	Dataset1				Dataset2			
	FADE	JNBM	RE	MICM	FADE	JNBM	RE	MICM
Input images	0.40	1.42	$N.A.$	2.62	0.67	1.03	$N.A.$	2.85
DCP	0.27	1.57	0.38	2.28	0.33	1.06	0.88	2.72
F-VAR	0.43	1.62	0.12	2.50	0.50	1.09	0.41	2.63
E-VAR	0.35	1.50	0.24	2.13	0.36	1.05	0.73	2.50
Proposed	**0.23**	**1.77**	**0.39**	**2.02**	**0.30**	**1.16**	**1.19**	**2.40**

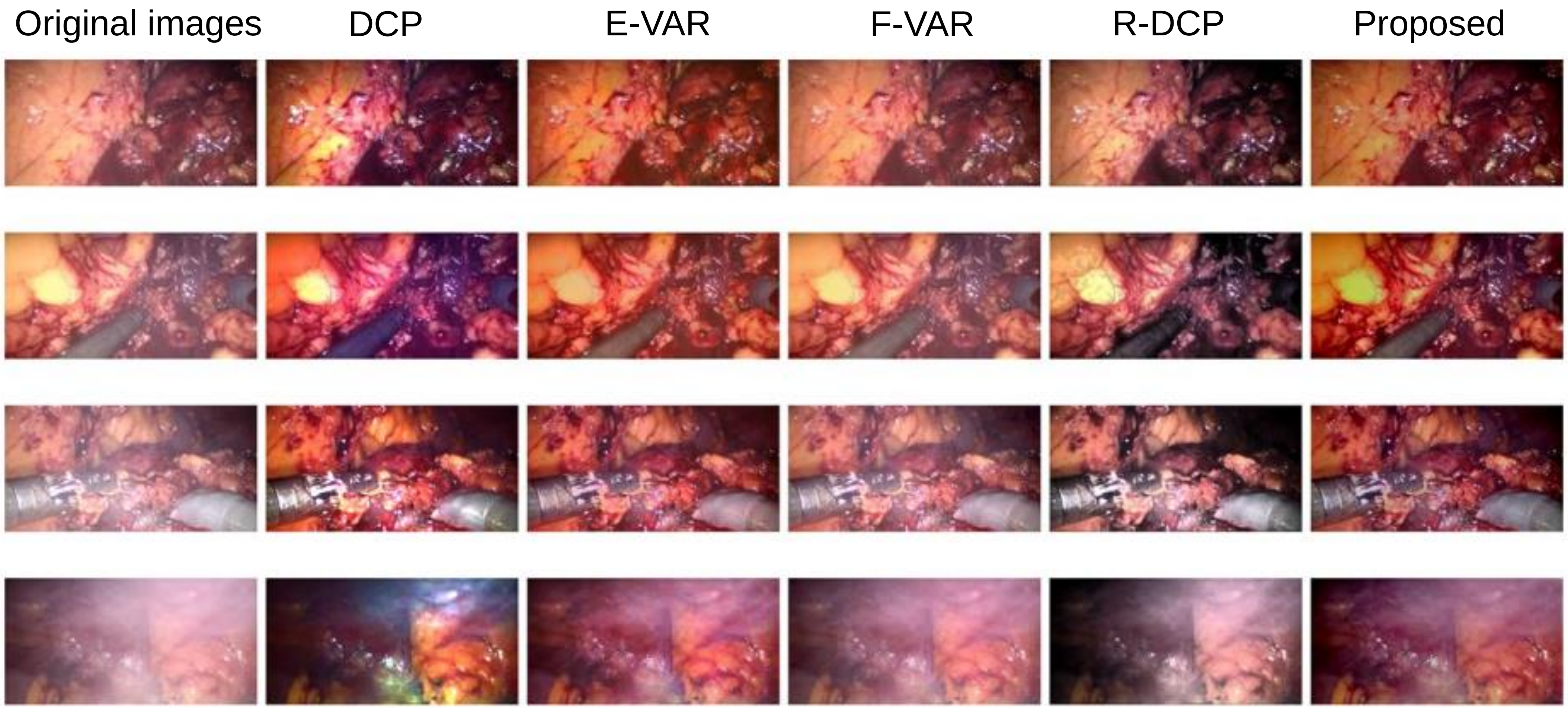








Variational Smoke Removal



Discussion

- Contributions
  - Physical model of smoke image formation is analyzed
  - A variational desmoking method is proposed
- Discussion
  - Computational speed:
    - GPU implementation of the variational method
    - Deep learning

Outline

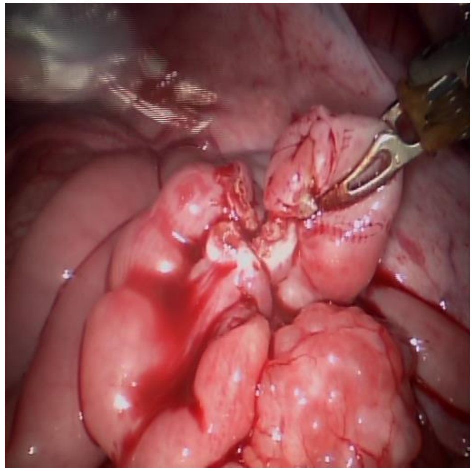
- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke removal
- Conclusion & Future perspectives

Deep Smoke Removal - Transfer Learning

- Dataset?
  - Synthetic Smoke Generation by Perlin noise
    - 19.600 images: 19500 images for training, 100 images for testing.

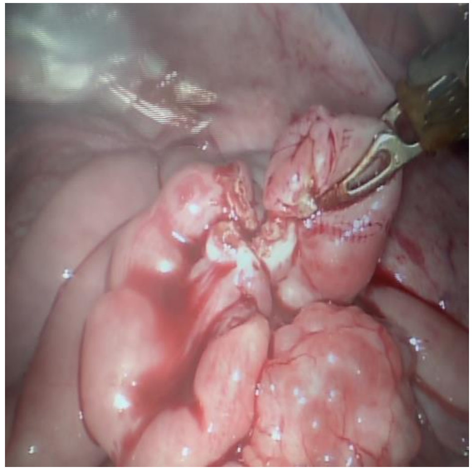
$$I_e^c(x) = I_g^c(x) + 0.8(I_s^c(x) - 1/N \sum_{i=1}^{N} I_s^c(i)),$$

Smoke free images



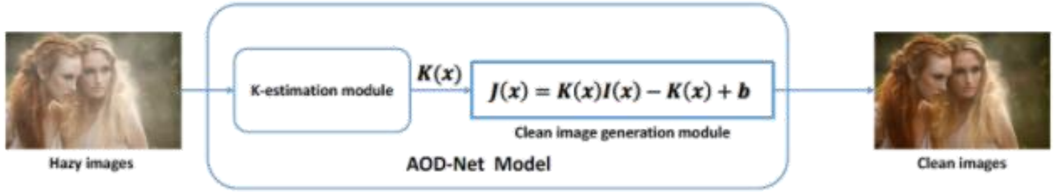
Generated Smoke

Image with smoke

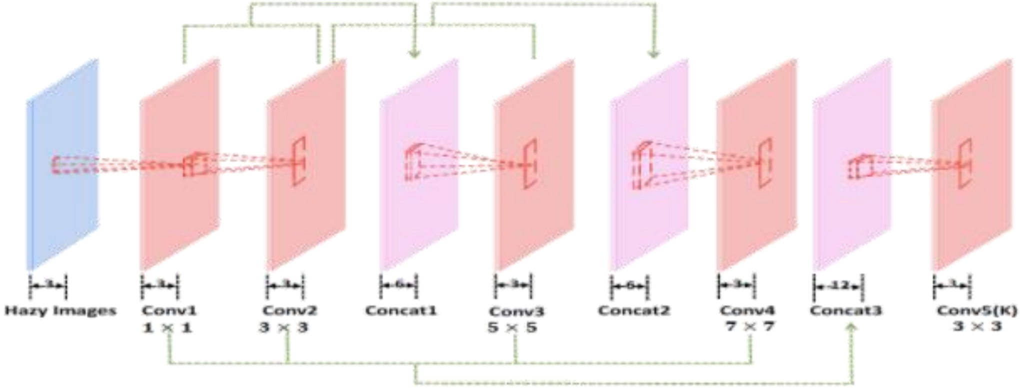


Deep Smoke Removal - Transfer Learning

- AOD-Net[22]: All-in-One Dehazing Network



(a) The diagram of AOD-Net



(b) K-estimation module of AOD-Net

$$I(x) = J(x)t(x) + A(1 - t(x))$$

$$\downarrow$$

$$J(x) = \frac{1}{t(x)} I(x) - A \frac{1}{t(x)} + A.$$

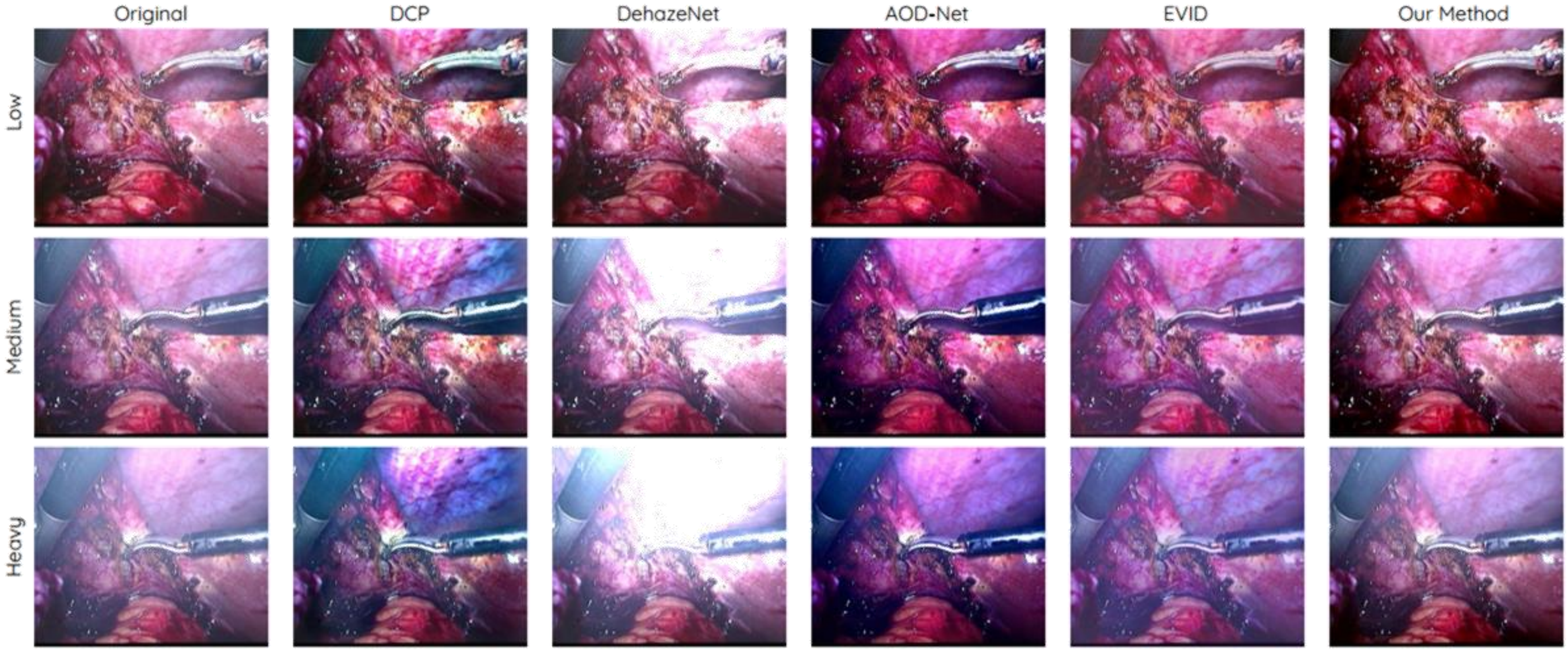
$$\downarrow$$

$$J(x) = K(x) I(x) - K(x) + b, \text{where}$$

$$K(x) = \frac{\frac{1}{t(x)}(I(x) - A) + (A - b)}{I(x) - 1}.$$

Deep Smoke Removal - Transfer Learning

- Experiments

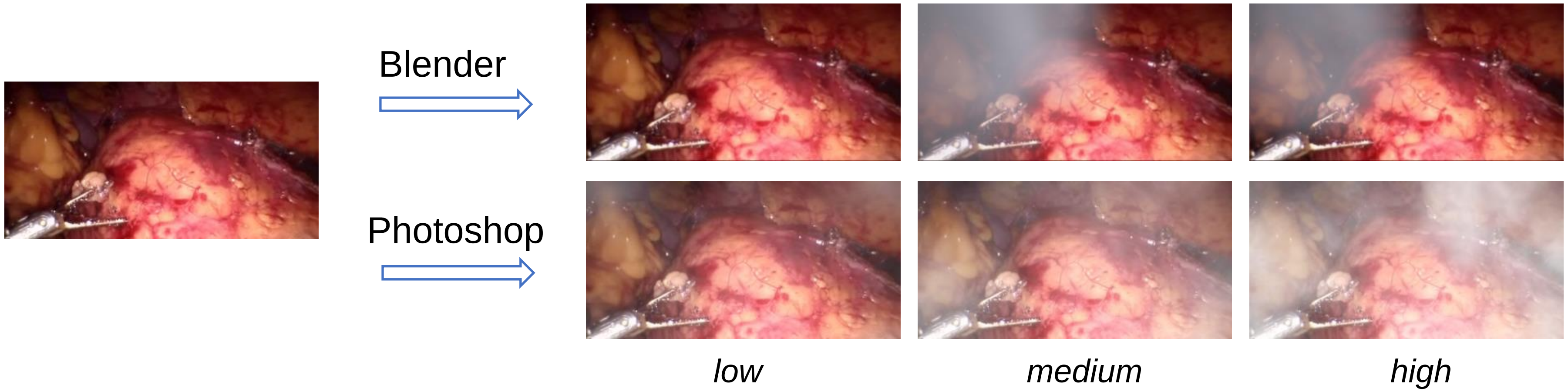


Discussion

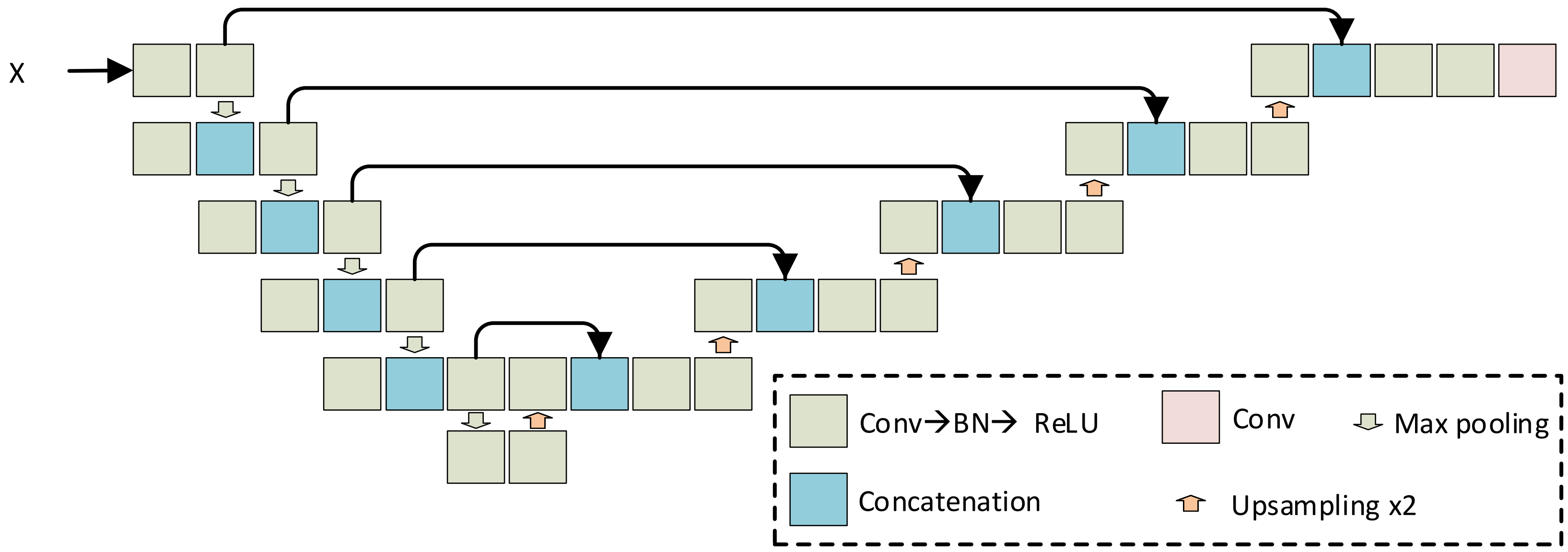
- Contributions
  - First known application of CNN based surgical smoke removal.
  - Employ synthetic smoke to generate training dataset
  - Proves the effectiveness and fast computational speed of applying deep learning to smoke removal purpose
- Discussion
  - Relies on physical model
  - Relies on pre-trained dehazing network

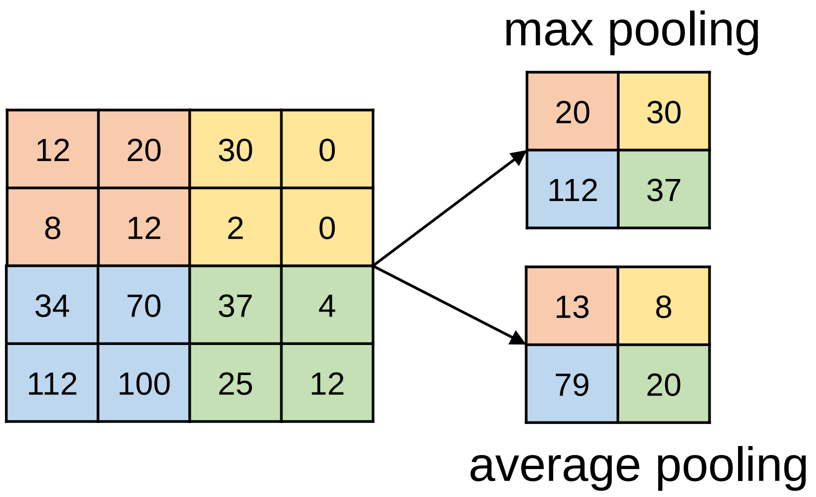
Deep Smoke Removal

- Dataset
  - Manually select 7553 smoke free images, synthesize smoke images by Blender and Adobe Photoshop with three smoke density: *low*, *medium* and *high*:



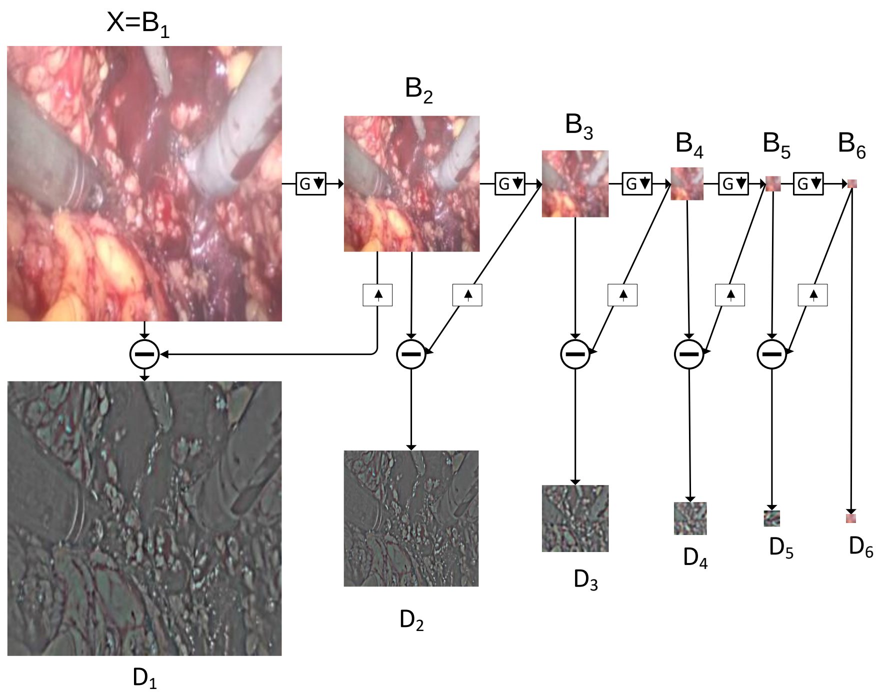
Deep Smoke Removal





Deep Smoke Removal

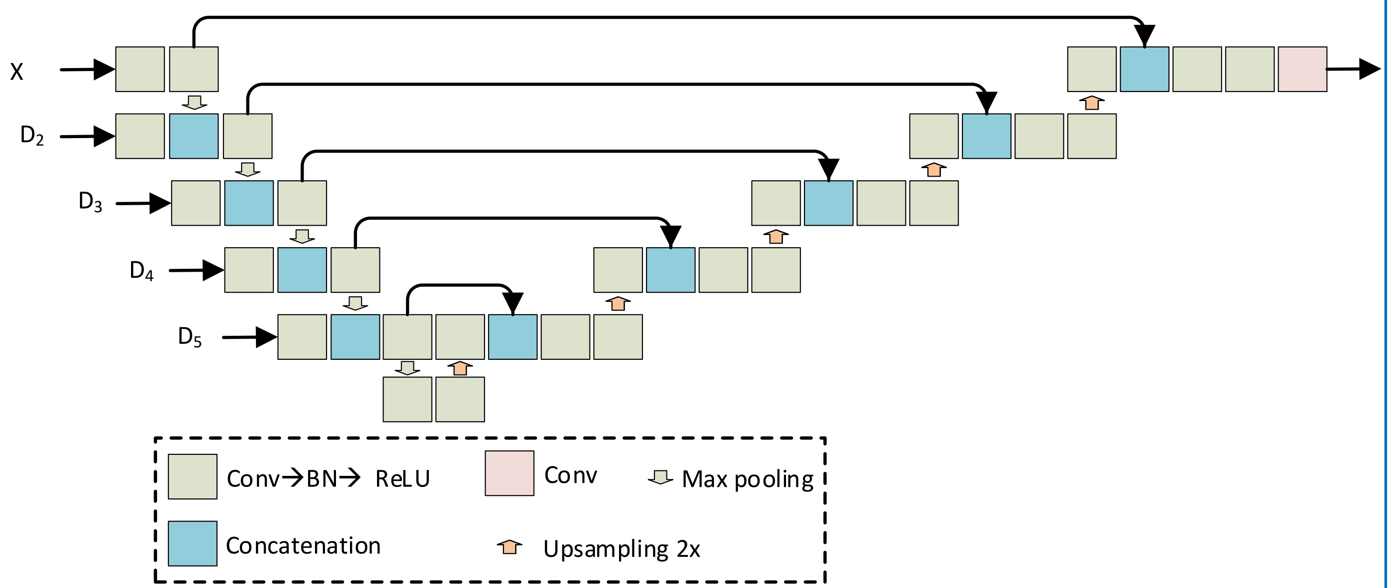
1. Image Pyramid Decomposition



G Smooth filtering followed by downsampling by 2x

Upsampling by 2x

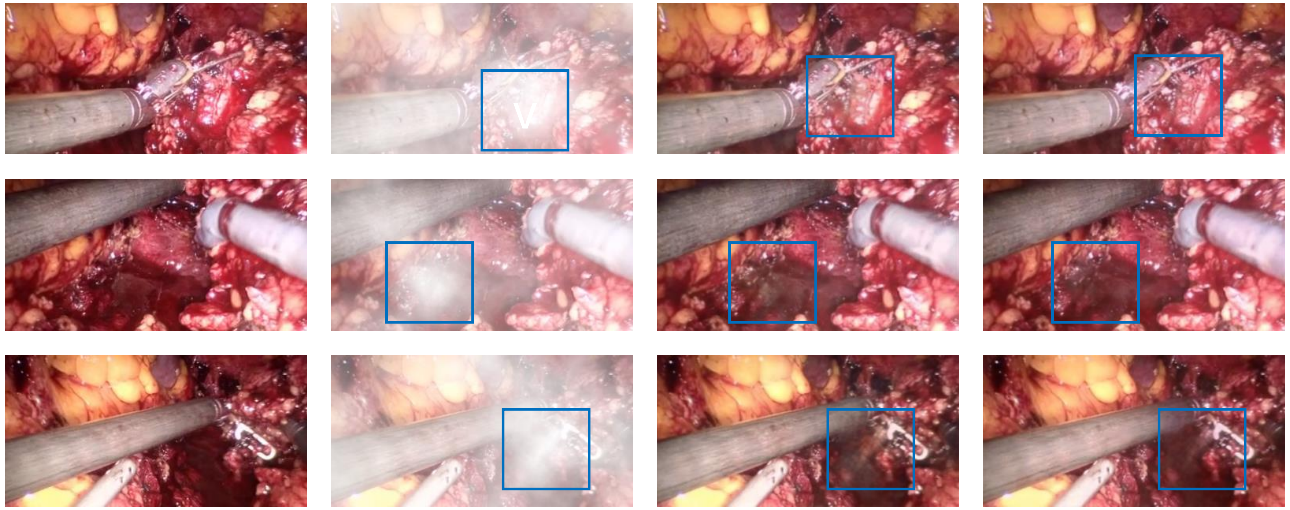
2. Network Structure



Deep Smoke Removal

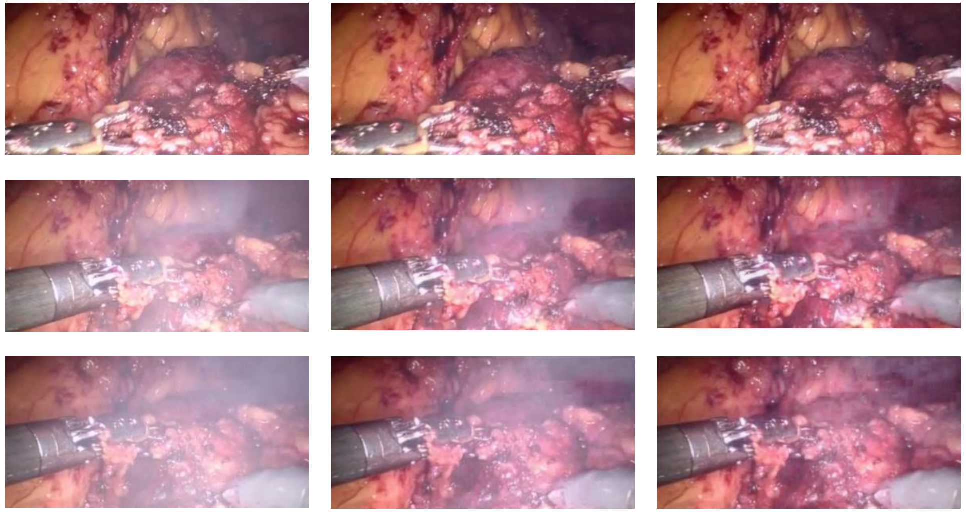
- Some Visual Results

Results of synthetic images



Orig. Smoke U-Net[23] Proposed

Results of real smoke images

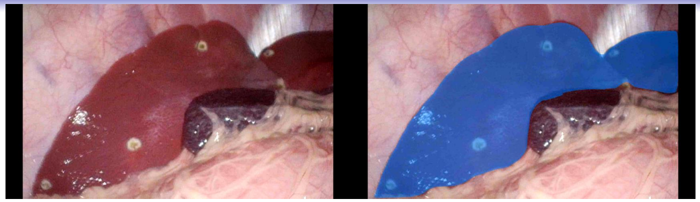


Orig. U-Net[23] Proposed

Outline

- Introduction
  - Background
  - Thesis outline
- Pre-operative and intra-operative registration
  - Surface reconstruction
  - Surface based registration
  - Semantic segmentation
- Laparoscopic image enhancement
  - Smoke detection
  - Variational smoke removal
  - Deep smoke removal
- Conclusion & Future perspectives

Conclusion



Laparoscopic Images Acquisition

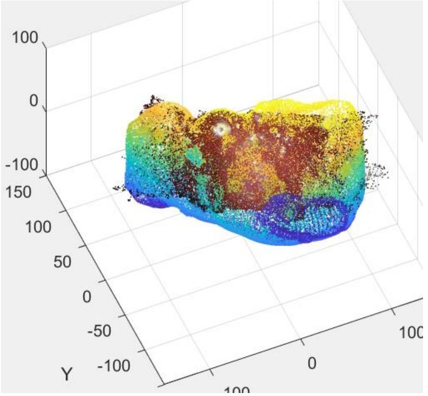
Enhancement

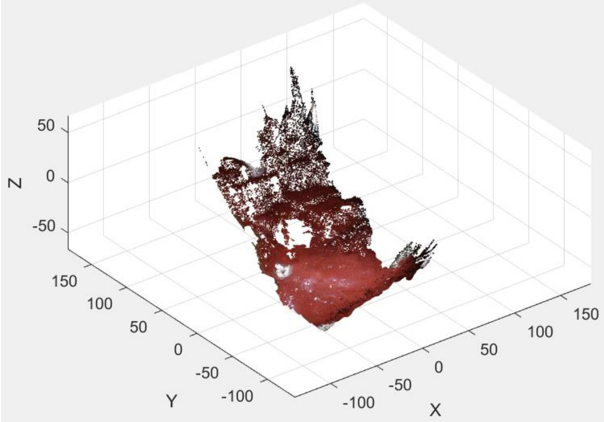
Pre-processing (desmoking)

- Smoke Detection (Q2.1, **Paper A**)
- Smoke Removal (Q2.2, **Papers B, C, D**)

Pre- and Intra- operative Registration

Segmentation (Q1.3, **Paper E**) → Surface Reconstruction (Q1.1, **Paper F**) → Registration (Q1.2, **Paper G**)





Conclusion

- Successfully implemented the stereo vision based surface registration workflow
- Propose proper algorithms for some steps of the workflow:
  - Smoke removal: classical image processing method, deep learning based methods
  - Automatic semantic segmentation
  - Stereo matching
  - Registration experimental setup

Future Considerations

- Smoke removal: classical image processing method, deep learning based methods
  - GPU implememtation of the variational method
  - **Validation**: image quality metrics, evaluation on the flowing tasks for surgical navigation
- Automatic semantic segmentation
  - Is deformable convolution operation the right tool?
  - Visual result indicates that the network is prone to output smooth result and easily ignore small detail. How to explain and improve it?

Future Considerations

- Stereo matching
  - Better energy function?
    - Occlusion, edge information, etc.
  - Parameters selection, computational speed
  - Deep learning especially unsupervised deep methods
- Registration
  - Surface stitching via deep learning based SLAM (Simultaneous localization and mapping)

Thank you

References

[1] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, 2000.

[2] D. Chan, H. Buisman, C. Theobalt, and S. Thrun, “A noise-aware filter for real-time depth upsampling,” in *ECCV Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications*. Springer, 2008.

[3] D. Stoyanov, M. V. Scarzanella, P. Pratt, and G.-Z. Yang, “Real-time stereo reconstruction in robotically assisted minimally invasive surgery,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*.
Springer, 2010, pp. 275–282.

[4] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, “Fast cost-volume filtering for visual correspondence and beyond,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504–511, 2013.

[5] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *IEEE conference on computer vision and pattern recognition*. IEEE, 2015, pp. 3431–3440.

[6] He, Junjun, Zhongying Deng, Lei Zhou, Yali Wang, and Yu Qiao. "Adaptive pyramid context network for semantic segmentation." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7519-7528. 2019.

[7] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, “Rethinking atrous convolution for semantic image segmentation,” arXiv preprint arXiv:1706.05587, 2017.

[8] N. Haouchine, S. Cotin, I. Peterlik, J. Dequidt, M. S. Lopez, E. Kerrien, and M.-O. Berger, “Impact of soft tissue heterogeneity on augmented reality for liver surgery,”
IEEE transactions on visualization and computer graphics, vol. 21, no. 5, pp. 584– 597, 2014.

[9] Song, Y., Totz, J., Thompson, S., Johnsen, S., Barratt, D., Schneider, C., Gurusamy, K., Davidson, B., Ourselin, S., Hawkes, D., et al. 2015. Locally rigid, vessel-based registration for laparoscopic liver surgery. International journal of computer assisted radiology and surgery, 10(12), 1951–1961.

[10] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, “Deformable convolutional networks,” in IEEE international conference on computer vision. IEEE,
2017, pp. 764–773.

[11] X. Zhu, H. Hu, S. Lin, and J. Dai, “Deformable convnets v2: More deformable,
better results,” in IEEE Conference on Computer Vision and Pattern Recognition.
IEEE, 2019, pp. 9308–9316.

[12] L. Maier-Hein, A. Groch, A. Bartoli, S. Bodenstedt, G. Boissonnat, P.-L. Chang, N. Clancy, D. S. Elson, S. Haase, E. Heim *et al.*, “Comparative validation of singleshot optical techniques for laparoscopic 3-d surface reconstruction,” *IEEE transactions on medical imaging*, vol. 33, no. 10, pp. 1913–1930, 2014.

[13] A. P. Twinanda, S. Shehata, D. Mutter, J. Marescaux, M. de Mathelin, and N. Padoy, “Endonet: A deep architecture for recognition tasks on laparoscopic videos,” TMI,
vol. 36, no. 1, pp. 86–97, 2017.

References

[14] L. K. Choi, J. You, and A. C. Bovik, “Referenceless prediction of perceptual fog density and perceptual image defogging,” IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 3888–3901, 2015.

[15] R. Ferzli and L. J. Karam, “A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb),” IEEE transactions on image processing, vol. 18, no. 4, pp. 717–728, 2009.

[16] M. A. Qureshi, A. Beghdadi, and M. Deriche, “Towards the design of a consistent image contrast enhancement evaluation measure,” Signal Processing: Image Communication, vol. 58, pp. 212–227, 2017.

[17] N. Hautière, J. P. Tarel, D. Aubert, and E. Dumont, “Blind contrast enhancement assessment by gradient ratioing at visible edges,” Image Analysis & Stereology, vol. 27, no. 2, pp. 87–95, 2011.

[18] 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.

[19] A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmío, “Fusion-based variational image dehazing,” IEEE Signal Processing Letters, vol. 24, no. 2, pp. 151–155, 2017.

[20] A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmío, “Enhanced variational image dehazing,” SIAM Journal on Imaging Sciences, vol. 8, no. 3, pp. 1519–1546, 2015.

[21] K. Tchaka, V. M. Pawar, and D. Stoyanov, “Chromaticity based smoke removal in endoscopic images,” in Medical Imaging 2017: Image Processing, vol. 10133. International Society for Optics and Photonics, 2017, p. 101331M.

[22] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “Aod-net: All-in-one dehazing network,” in IEEE International Conference on Computer Vision. IEEE, 2017, pp. 4770–4778.

[23] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.

[24] Allan, M., Kondo, S., Bodenstedt, S., Leger, S., Kadkhodamohammadi, R., Luengo, I., Fuentes, F., Flouty, E., Mohammed, A., Pedersen, M. and Kori, A. "2018 Robotic Scene Segmentation Challenge." *arXiv preprint arXiv:2001.11190,* 2020.

[25] Dumoulin, V., & Visin, F. (2016). “A guide to convolution arithmetic for deep learning”. arXiv preprint arXiv:1603.07285.

[26] Yu F, Koltun V. “Multi-scale context aggregation by dilated convolutions”. arXiv preprint arXiv:1511.07122. 2015 Nov 23.

References

[27] Zhang, H., Dana, K., Shi, J., Zhang, Z., Wang, X., Tyagi, A. and Agrawal, A., 2018. “Context encoding for semantic segmentation”. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 7151-7160, 2018.

[28] Zhou, Haoyin, and Jayender Jagadeesan. "Real-time dense reconstruction of tissue surface from stereo optical video." *IEEE transactions on medical imaging* (2019).

[29] Le, H.N., Nguyen, H., Wang, Z., Opfermann, J., Leonard, S., Krieger, A. and Kang, J.U., 2018. Demonstration of a laparoscopic structured-illumination three-dimensional imaging system for guiding reconstructive bowel anastomosis. Journal of biomedical optics, 23(5), p.056009.

[30] van den Berg, N.S., Buckle, T., KleinJan, G.H., van der Poel, H.G. and van Leeuwen, F.W., 2017. Multispectral fluorescence imaging during robot-assisted laparoscopic sentinel node biopsy: a first step towards a fluorescence-based anatomic roadmap. European urology, 72(1), pp.110-117.

[31] Teatini A., Brunet J.N., Nikolaev S., Wang C., Edwin B., Cotin S., Elle O.J., "Use of stereo-laparoscopic liver surface reconstruction to compensate for pneumoperitoneum deformation through biomechanical modeling", submited.

[32] P. J. Besl and N. D. McKay, “Method for registration of 3-d shapes,” in Sensor Fusion IV: Control Paradigms and Data Structures, vol. 1611. International Society for Optics and Photonics, 1992, pp.586–607.

[33] Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J., 2017. Pyramid scene parsing network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2881-2890).

[34] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A.L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, 40(4), pp.834-848.