Real-Time GPU Implementation of Correlation-Based Dense Matching of Stereo Endoscopic Images for Surface Reconstruction

Fahad Islam[1], Alborz Amir-Khalili[2], Jean-Marc Peyrat[1], Julien Abi-Nahed[1], Abdulla Al-Ansari[3], Rafeef Abugharbieh[2]

[1]Qatar Robotic Surgery Centre, Qatar Science & Technology Park, Doha, Qatar
[2]Biomedical Signal and Image Computing Lab, University of British Columbia, Vancouver, Canada
[3]Urology Department, Hamad Medical Corporation, Doha, Qatar

Introduction:

Robot-assisted surgeries benefit from the availability of a stereo endoscopic camera to provide the surgeon with a 3D view of the surgical scene. This intra-operative imaging modality also enables the reconstruction of 3D soft-tissue surface, which is a prerequisite for many clinical applications such as image-guidance with cross-modality registration, telestration, expansion of the surgical scene by mosaicing, and collision detection. The bottleneck in stereo surface reconstruction is the dense matching between left and right cameras as it is the most computationally expensive step. Nevertheless, being the most parallelizeable step too, we propose a GPU implementation to achieve real-time processing which is a requirement in any real life surgical setting.

Method:

The stereo rig is manually calibrated and resulting parameters used to rectify the pair of stereo (left and right camera) images. Rectification simplifies the dense matching problem by reducing the problem from being one of matching in 2D to that of 1D matching along parallel epipolar lines (Fig.1(a)). Dense matching is performed on the rectified images by identifying kernel patches of neighbouring pixels in the left and right images that have the highest correlation computed using the normalized cross correlation (NCC) metric. The distance between matching patches, also known as disparity, d, in the rectified image domain, is used to triangulate the location of the matched patch in 3-space. We implemented the correlation-based dense matching in two GPU versions (GPUstd and GPUopt) where GPUopt has been optimized to reduce the memory footprint, to eliminate redundant computations, and to modify the precision to maximize performance without significantly altering the results. We compare CPU and GPU computation times on stereo endoscopic images of real surgerical procedures. We consider different parameters: image sizes (480x270, 640x480, 960x540) and kernel patch sizes (5x5, 7x7, 15x15, 25x25, 31x31). The disparity search range is set to be 12% of the mage width.

Material:

We used clinical stereo endoscopic videos of robot-assisted partial nephrectomy surgeries captured at full HD 1080i resolution using a da Vinci S HD Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation Surgical System. Experiments were performed on a workstation with 128 GB RAM, an Intel Xeon Processor E5-2690, and an NVIDIA Tesla C2075.

Results:

Computation times are summarized in Fig.1(b)(c) and show substantial improvement when running on GPU compared to CPU (GPUstd is 14x faster and GPUopt is 56x faster). The CPU implementation hardly reaches real-time processing, whereas GPUopt can achieve real time for selected parameter setting, or in near realtime for others. The difference between the standard GPU implementation and the optimized one (4x faster) also shows that the quality of GPU implementation is important to reach full GPU capabilities. Our algorithm also performed better than a recent approach, GPU-OpenCV, reporting 125ms for stereo endoscopic images (compared to 87ms for our GPUopt.) on 640x480 images and 7x7 kernel patch size with similar GPU hardware (NVIDIA Tesla C2070).

Conclusion:

We presented a real-time GPU implementation of correlation-based dense matching of stereo endoscopic images. This work is the first but essential step towards the real-time implementation of the whole stereo surface reconstruction.

Acknowledgements:
This abstract was made possible by NPRP Grant from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

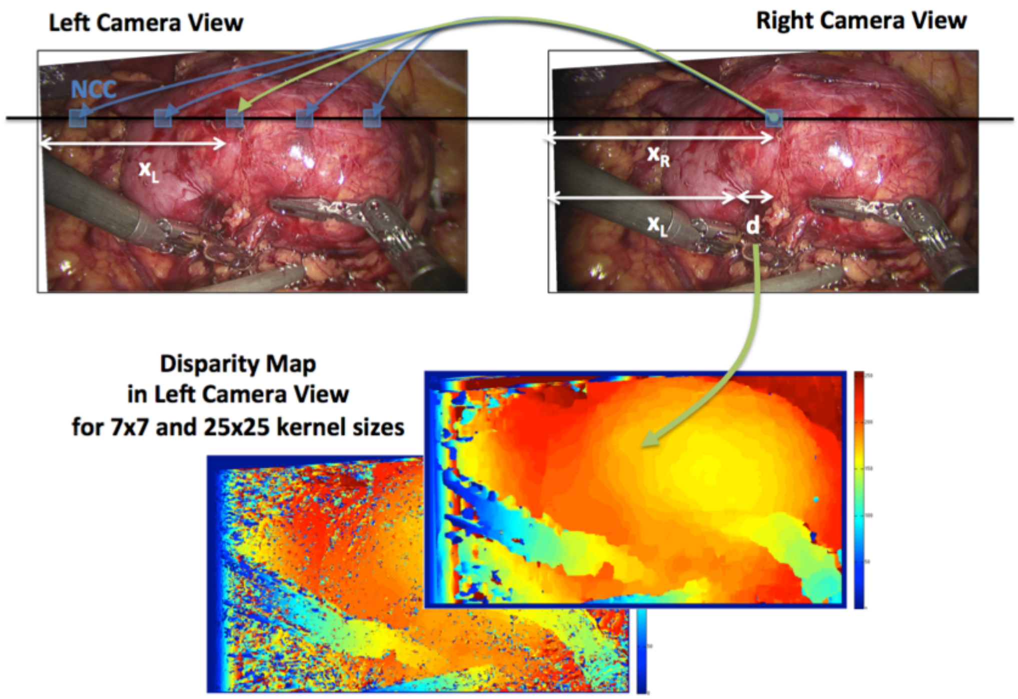

Fig.1(a) - Correlation-based dense matching between left and right camera view. Normalized Cross Corelation (NCC) ratio is computed between left and right images along an epipolar line (in black). The disparity is the distance between the two matching points along this epipolar line.

Image size	Disparity Search Range	Kernel size (pixel)	CPU Time (ms)	GPUstd Time (ms)	GPUopt Time (ms)
960 x 540	64	5 x 5	4,555	321	88
		7 x 7	8,627	604	164
		15 x 15	37,877	2,744	681
		25 x 25	101,900	7,666	1,803
		31 x 31	153,832	11,831	2,739
640 x 480	56	5 x 5	2,402	163	46
		7 x 7	4,477	312	87
		15 x 15	19,590	1,418	363
		25 x 25	54,602	3,954	947
		31 x 31	85,429	6,095	1,438
480 x 270	32	5 x 5	577	40	11
		7 x 7	1,092	77	20
		15 x 15	4,758	348	80
		25 x 25	12,870	970	205
		31 x 31	19,422	1,495	304

Fig.1(b) - Computation times (in ms) of CPU, GPU and GPU optimized implementations with respect to image and kernel sizes.

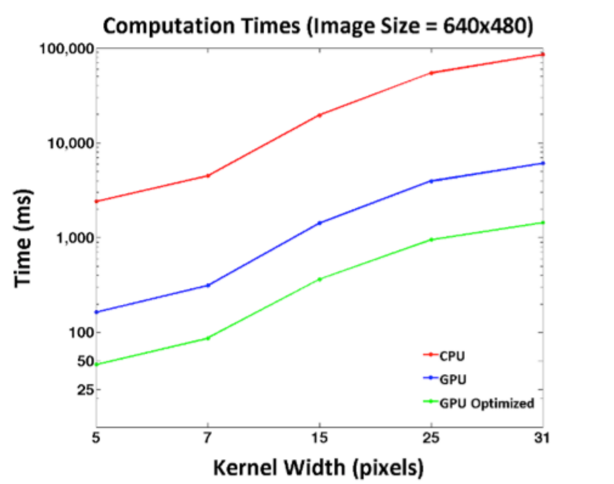

Fig.1(c) - Computation times (in ms) of CPU, GPU and GPU optimized implementations for an image size of 640x480 with respect to kernel sizes.