Real-Time Stereo Vision

Christian Richardt,
Douglas Orr &
Neil Dodgson

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Overview

We have built a real-time stereo vision application that estimates depth from stereo imagery such as existing stereo images and videos, as well as live streams from a stereo webcam.

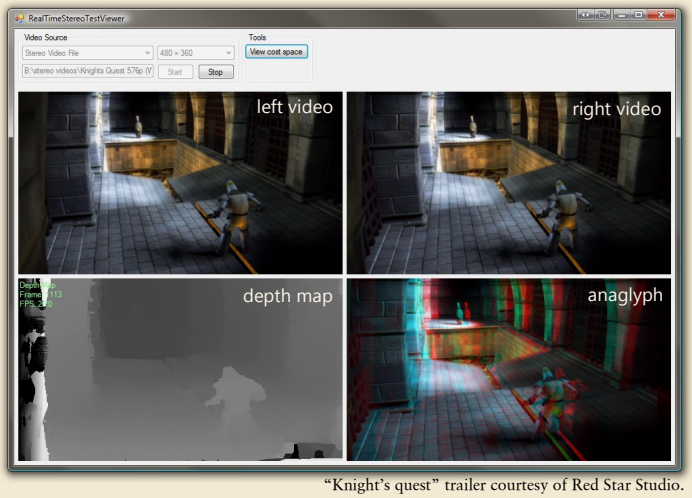
All core processing parts are implemented using *C for CUDA* and run on a Quadro FX 5800.

Motivation

Depth maps enable interesting and useful effects in video conferencing, such as video refocusing and background replacement. Stereo vision is also used in many automated systems for object detection and recognition. Both require real-time frame-rates for interactivity and a fast response.

Adaptive Support Weights

We developed an approximation to *Adaptive Support Weights* that is 200× faster than a naïve GPU implementation. Our implementation runs at 20 frames per second on a video of 480×270 pixels and 40 disparity levels (104 MDE/s).

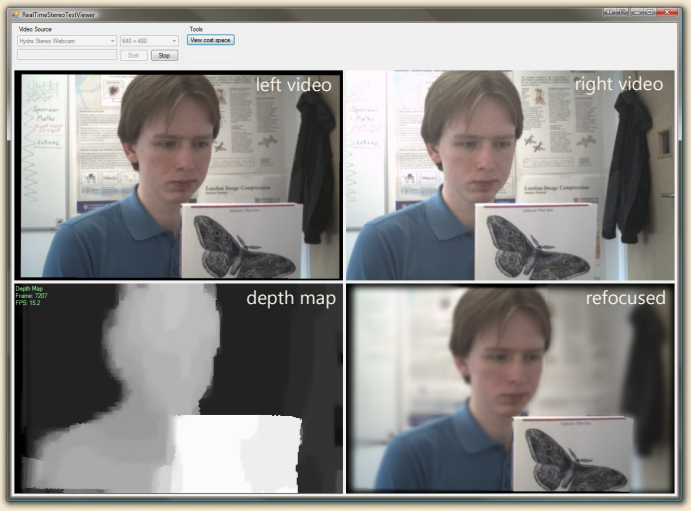


"Knight's quest" trailer courtesy of Red Star Studio.

Hierarchical Belief Propagation

Our implementation runs at 4.5 frames per second on full-resolution stereo webcam footage (752×480×64), which is about 60× faster than a comparable CPU implementation.

For the refocusing demo below, we down-sample the video to 320×205 with 30 disparities which runs at 18 frames per second, or 15 frames per second when including the effect rendering time.



Future Work

Depth estimation algorithms that work on stereo videos need to incorporate temporal evidence to resolve ambiguities and to reduce flickering for creating temporally more coherent depth maps.

Small Print: References & Acknowledgements

Kuk-Jin Yoon & In So Kweon. Adaptive Support-Weight Approach for Correspondence Search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):650–656, 2006.

Pedro F. Felzenszwalb & Daniel P. Huttenlocher. Efficient Belief Propagation for Early Vision. *International Journal of Computer Vision*, 70:41–54, 2007.

Christian Richardt is supported by an EPSRC Doctoral Training Studentship. Douglas Orr was supported as an undergraduate research intern by Presenccia (FP6-FET-27731).

We are grateful to NVIDIA for donating the Quadro FX 5800 through their CUDA Centre of Excellence at the University of Cambridge. We thank nVela Ltd for providing us with two Hydra stereo webcams which we used throughout the development of our application.

RAINBOW RESEARCH GROUP

William Gates Building
15 JJ Thomson Avenue
Cambridge CB3 0FD

www.cl.cam.ac.uk/~cr333