Introduction
Stereo vision is the extraction of 3D information from digital images. Stereo matching represents the process of determining object depth. Depth is inferred from appropriate projections of the same 3D points on 2 stereo images. Raw images are firstly rectified, so that vertical disparity is omitted (only horizontal search).

Theory/Method/Hypothesis
There are various algorithms for disparity estimation, usually divided into 4 steps:

- Cost computation – the similarity between pixels determined based on their gradients and color pixel values
- Cost aggregation – filtering of costs; box filter used, later only a summation of costs, to avoid multiplier utilization
- Disparity selection – choosing disparity with the smallest aggregated cost; winner-takes-all (WTA) approach
- Disparity refinement – post-processing: 1) left-right check – compares disparity maps for the left and right image and marks valid disparities, while filling invalid 2) segmented median filtering – weighted median filtering of disparities by segmentation based on grayscale left input image 3) guided image filtering – linear edge-preserving filtering that takes the most hardware resources; \( q = a \cdot I + b \) where \( I \) is a guidance image, and \( a \) and \( b \) are coefficients
- 4) median filtering – used for final pike removal.

All results are degraded because of the existence of occluded pixels.

Results
Several resolutions (450x375, VGA, SVGA) and disparity ranges (32, 64, 128) were tested. Comparable results obtained with biggest advantage in small number of DSPs and registers, as well as large MDE and highest possible fps (due to processing of 1 disparity on every clock cycle).

Conclusion/Perspectives
The proposed algorithm was successfully done in both software and hardware. The component can be used as a part of a larger system thanks to a convenient pipeline. Future improvements should include pre-processing (with rectification), better hardware implementation of the mean filter, use of off-chip memory to decrease BRAMs.

Acknowledgments
I would like to thank Prof. Siew Kei Lam and Prof. René Beuchat for guidance and supervision. Additional gratitude goes to all of my friends, family and my girlfriend.

References