Real-time Halfway Domain Reconstruction of Motion and Geometry

Lucas Thies[1] Michael Zollhöfer[2] Christian Richardt[2,3,4] Christian Theobalt[2] Günther Greiner[1]

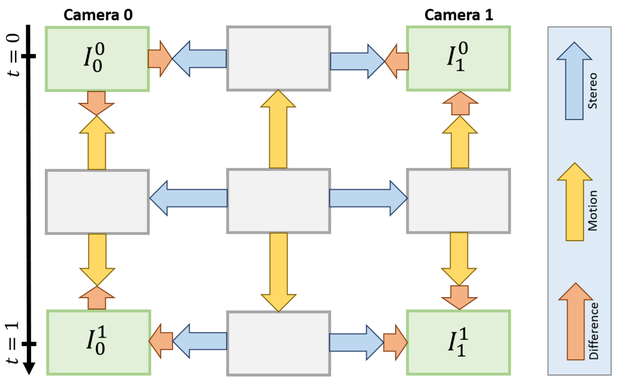
[1]University of Erlangen-Nuremberg [2]Max Plank Institute for Informatics [3]Intel Visual Computing Institute [4]Univeristy of Bath

ABSTRACT

We present a novel approach for real-time joint reconstruction of 3D scene motion and geometry from binocular stereo videos. Our approach is based on a novel variational halfway-domain scene flow formulation, which allows us to obtain highly accurate spatiotemporal reconstructions of shape and motion. We solve the underlying optimization problem at real-time frame rates using a novel data-parallel robust non-linear optimization strategy. Fast convergence and large displacement flows are achieved by employing a novel hierarchy that stores delta flows between hierarchy levels. High performance is obtained by the introduction of a coarser warp grid that decouples the number of unknowns from the input resolution of the images. We demonstrate our approach in a live setup that is based on two commodity webcams, as well as on publicly available video data. Our extensive experiments and evaluations show that our approach produces high-quality dense reconstructions of 3D geometry and scene flow at real-time frame rates, and compares favorably to the state of the art.

HALFWAY-DOMAIN

The four captured images define the corners of the scene flow geometry. The five in-between frames define intermediate states of warping between the captured images based on the flow data. We define the scene flow as a combination of three flows (stereo, motion and difference) relative to the halfway domain in the middle.



SCENE FLOW OBJECTIVE

Our approach is based on energy minimization. The objective function takes into account both spatial alignment of the inputs warped to the halfway-domain as well as the validity of the flows.

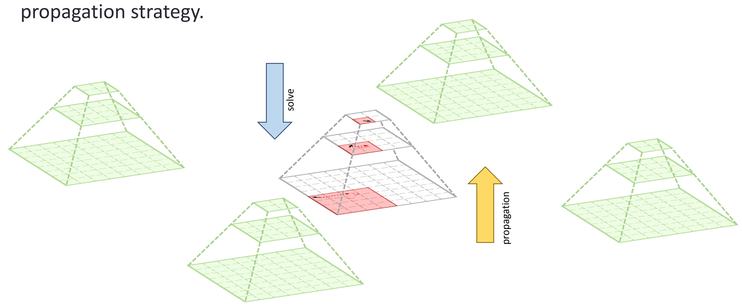
$$E(S) = E_{align}(S) + E_{reg}(S)$$

$$E_{align}(S) = \sum_N V(x_i)\Phi\left(I_1^0(x_i + s_i - m_i - d_i) - I_0^0(x_i - s_i - m_i + d_i)\right) + \dots$$

$$E_{reg}(S) = \sum_G \omega_1 E_{smooth}(S) + \omega_2 E_{epi}(S) + \omega_3 E_{mag}(S)$$

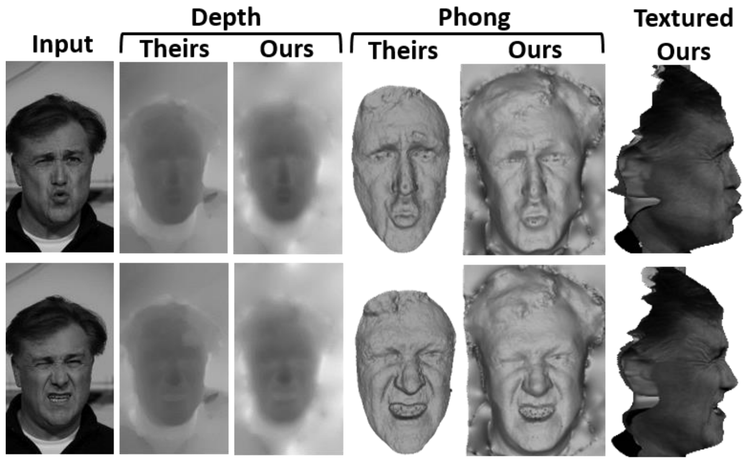
DATA-PARALLEL OPTIMIZATION

We devise a data-parallel hierarchical Gauss-Newton solver, following Zollhöfer et al. 2014, that exploits the computational power of modern graphics cards. Our hierarchy encodes flows based on deltas to the next coarser level. This enables us to handle large displacements and allows for fast convergence based on a temporal propagation strategy.



COMPARISON

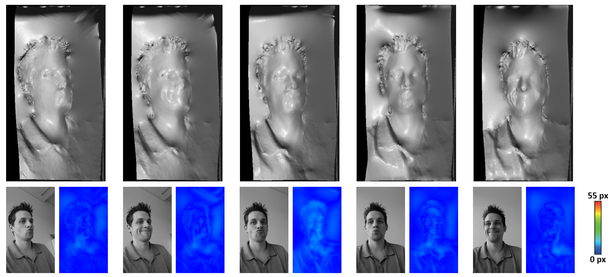
Valgaerts et al. 2010:
(1920 x 1088 @ 10Hz)



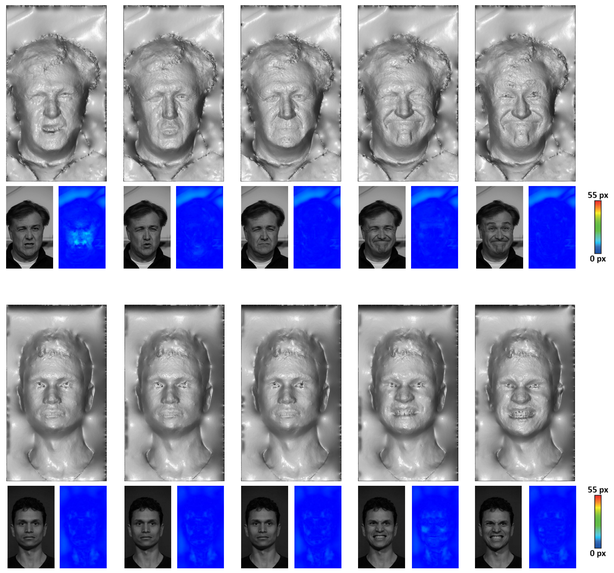
Blumenthal-Barby and Eisert 2014:
(4288 x 2848 @ 1,4Hz)

RESULTS

Live reconstruction results:
(1280 x 720 @ 30Hz)

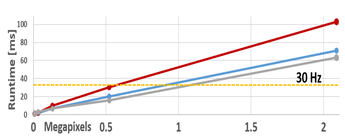


Results on the data of Valgaerts et al. 2010, 2012:
(1920 x 1088 @ 10Hz)



PERFORMANCE & CONVERGENCE

The runtime is linear in the image resolution and we obtain **real-time performance** for up to 0.9 MP.



Convergence behavior of our solver for the finest hierarchy level:

Error peaks correspond to the start of new non-linear Gauss-Newton steps.