Using Active Illumination for Accurate Variational Space-Time Stereo

Sergey Kosov, Thorsten Thormählen, Hans-Peter Seidel

Max-Planck-Institute for Computer Science
Saarbrücken, Germany

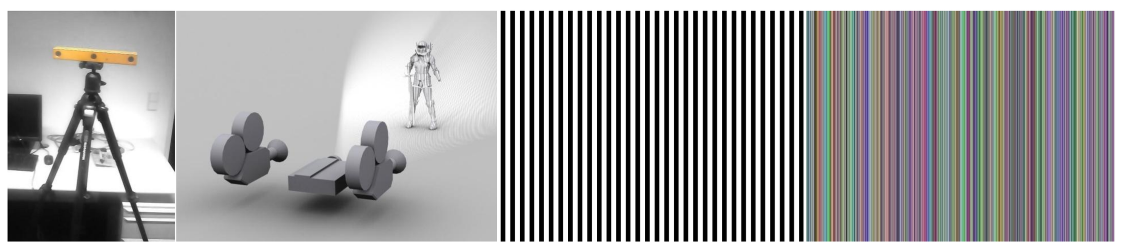
Algorithms for dense disparity map reconstruction are often a basic building block of more complicated systems for automatic 3D scene analysis, event detection, or object recognition. We present a combination of structured light and fast variational space - time stereo

1 First, we extend the 2D state-of-the-art variational framework to the 3D framework, which allows to consider the problem not only in space, but also in time. In combination with synchronized active illumination we make the solver independent from the scene texture. **2** The main disadvantage of the variational solvers is that they tend to become slow, especially as the number of equations is increased by adding information from different points in time. In our work we show that speed improving techniques like full multigrids (FMG) and multi-level adaptation technique (MLAT), can still be applied. This novel combination of algorithms runs in near sub-real-time and significantly improves the reconstruction accuracy. **3** In most stereo matching algorithms, the inherent ambiguity of image values in homogeneous image regions leads to a loss of accuracy in computation of dense dispatity maps. We argue that projected vertical color strip patterns are very good suited to be used in combination with the variational method. **4** The experiential results show the advantage of the novel variational space-time processing (STP) in combination with active illumination (AI) over the classical variational single frame processing (SFP).

Stereo Setup and Algorithm

- Point Grey Bumblebee® XB3 stereo camera
- Synchronized projector, casting a structured light pattern onto a scene
- Datasets with a resolution of 640x480x24 voxels
- Estimation of disparity map (variational approach, including regularization)

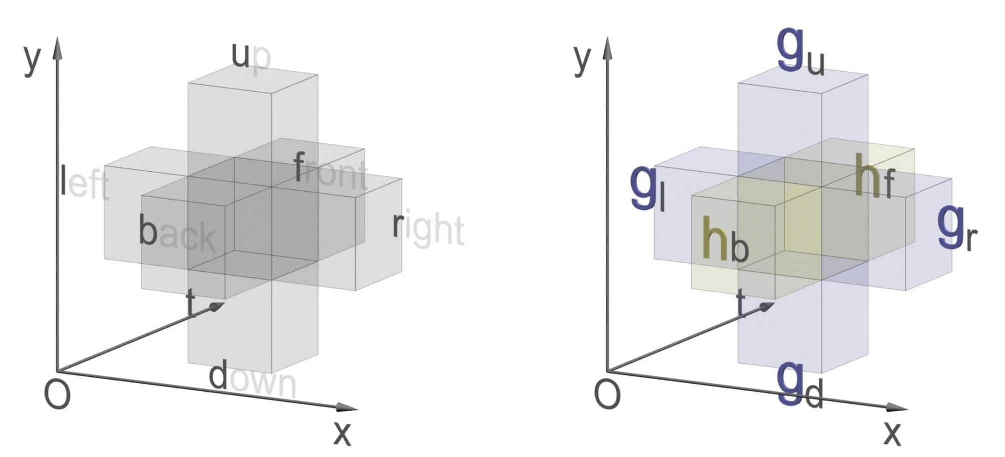
Our stereo setup and examples of stereo patterns:



1 Space-Time Variational Method

- Considering additionally the scene dynamics
- Processing the rectified *stereo sequences* instead of *stereo pairs*
- Classical 1D data-term
- Extended 3D smoothness-term
- Tichonov regularization for time dimention
- Charbonnier and Perrona-Malik regularization for space dimentions

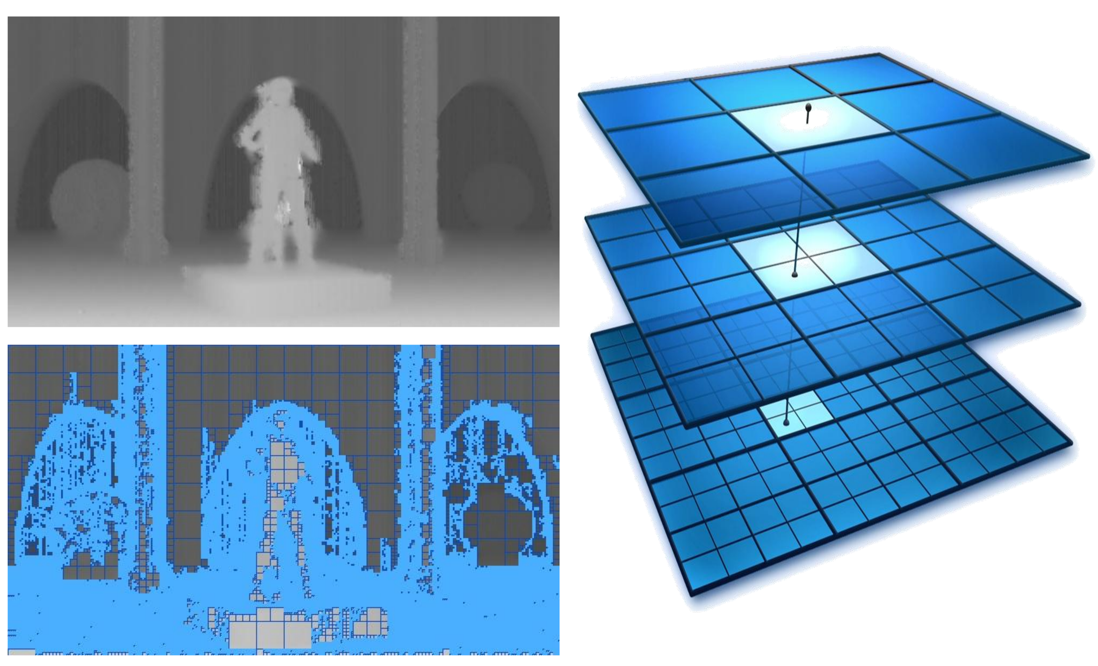
3D stencil for smoothness term discretization:



2 Space-Time FMG and MLAT

- The multigrid method implies the usage of a pyramid of scaled versions of the initial images.
- The maximal reasonable number of levels: $\#levels \le \log_2 \min\{X, Y, T\}$
- Usually the time dimension is much smaller than the space dimentions: $T \ll min\{X, Y\}$
- Use the full-multi-grid approach only in the spatial direction for each time slice independently
- Assume that the solution $u(x,y,t)$ is smooth with time: $u_t \to 0$
- Use MLAT approach for calculating the adapted grid once per currently processed space-time block

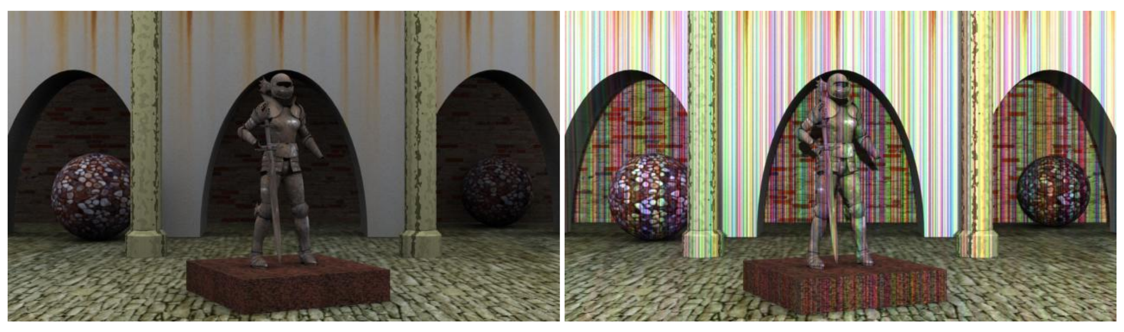
The finest MLAT grid example and multigrid:



3 Active Illumination

- Project intensity coded light in order to introduce the artificial texture onto the scene
- Define the best suited projection pattern for the scene and for the variational method
- Adapt the pattern in such a way, that interference by the scene is minimized
- Do not identify the pattern during the reconstruction process: 1) no need in exact positioning of the projector; 2) no need prior knowledge about the structure of the projected light
- Use random generated, high-contrast color pattern

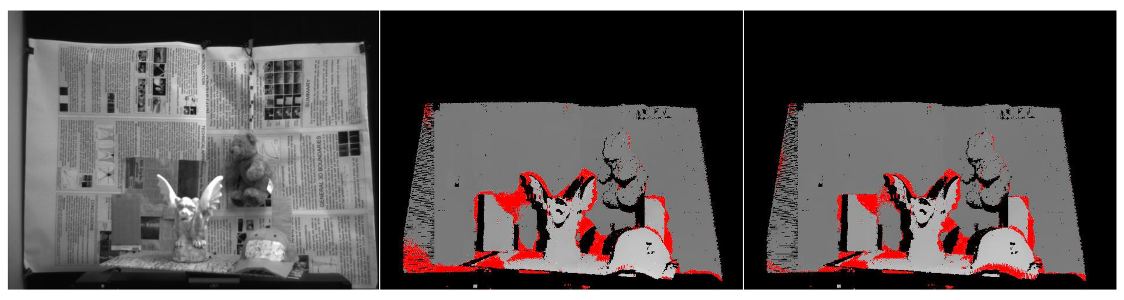
A scene without and with active illumination:



4 Experimental Results

- **Dynamic scene without active illumination**
- Improvement up to 1,21 times
- Average gain 1,3% improvement (SFP: 10,42% bad pixels; STP: 9,16% bad pixels)

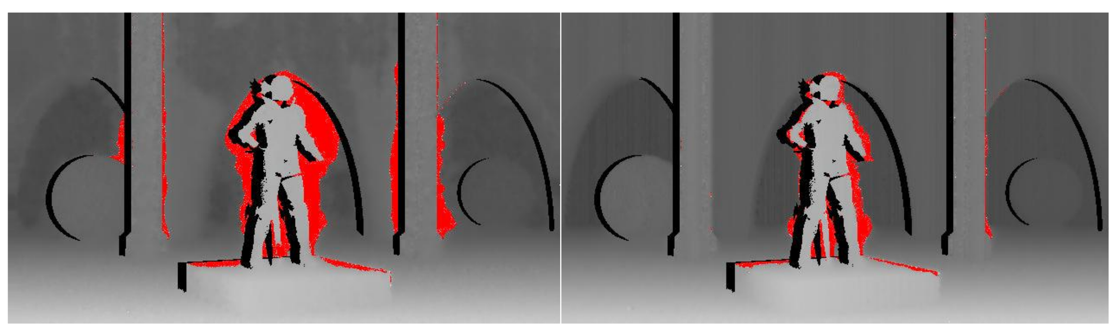
Left image of the 27-th stereo frame, solutions with bad-pixels maps for SFP (9,69%) and STP (7,99%) approaches :



- **Static scene with active illumination**
- Improvement up to 1,42 times
- Average gain 4% improvement (SFP: 14,56% bad pixels; STP: 10,54% bad pixels)

- **Dynamic scene with active illumination**
- Improvement more then 2 times
- Average gain 1,1% improvement (SFP: 3% bad pixels; STP: 1,9% bad pixels)

Solutions for the 7-th stereo frame with bad-pixels maps for SFP (4,46%) and STP (2,13%) approaches :



Comparison of the average percentage of bad pixels for the dynamic scene with AI scene (error threshold = 1 pixel). Improvement ratio are given in brackets:

	without AI		with AI	
	SFP	STP	SFP	STP
Variational method	5.9%	5.2% (1.1x)	3% (**2**x)	1.9% (**3.1**x)
Expansion method	2.9%	3.7% (0.8x)	2.4% (1.2x)	2.2% (1.3x)
Belief propagation	3.1%	2.9% (1.1x)	1.9% (1.6x)	1.4% (2.2x)
Swap method	10.3%	8.8% (1.2x)	7.2% (1.4x)	4% (2.6x)
Infection method	16.3%	16.2% (1.0x)	9.1% (1.8x)	8.9% (1.8x)
TRW method	3.6%	2.4% (**1.5**x)	2.4% (1.5x)	1.3% (2.8x)