Preface

He stood in the center of Heaven and looked about it, having decided to have four eyes today. He noticed that with less than two looking in any one direction, he couldn't see as well as he ought. He resolved to set someone to discover the reason for this.

Steven Brust, "To Reign in Hell".

Nowadays, Computer Vision has progressed far beyond the relatively simple question posed in the quote. Not only do we know that two views are necessary for stereo vision in order to obtain correspondences for triangulation, we can even design algorithms which mimic this process. And although the human visual system still performs far superior to machines when it comes to scene reconstruction, the latter progress fast and well. Vision algorithms become more and more robust and sophisticated, and modern techniques can integrate the information from a lot more than just two cameras and views to create a geometric model of a scene. Recently, researchers also began to exploit temporal coherence in video sequences.

The intention of this thesis is to progress one step further towards accurate and reliable scene reconstruction. The focus is on being able to render the model from novel viewpoints. Each chapter after the introduction is a revised and extended version of a refereed conference paper I published during my three years as a researcher at the MPI Informatik in Saarbrücken. Chronologically, the first one was the Light Field Rendering System in Chapter 12, which was presented at the *Vision, Modeling and Visualisation* 2002

in Erlangen [37]. In its first incarnation, it was based on an existing depth map reconstruction algorithm. Shortly after, we developed our own technique, which performs an additional background subtraction and adds temporal coherence for better playback quality. Presented at the *Computer Vision and Pattern Recognition* 2003 in Madison [42] and in a sketch at Siggraph 2003 in San Diego, this technique now forms the basis for Chapters 5 and Chapter 6.

My second line of research also started with rendering, this time free-viewpoint rendering of the existing visual hull geometry provided by Christian Theobalt using projective texturing. The result was the billboard rendering technique presented first at the *Visual Communications and Image Processing* 2003 in Lugano [43], and in an extended version at the *International Conference on Image Processing* 2003 in Barcelona [38]. It can be found in Chapter 13 of this thesis. Subsequently, I developed the necessary 3D surface reconstruction algorithms, which I now consider my major contribution and the core of this thesis. Analytical techniques based on weighted minimal surfaces seemed most appealing to me because of the concise mathematical approach. After I had the idea to represent spatio-temporal geometry as a hypersurface, I started to develop the necessary mathematical tools, which were presented at the *European Conference on Computer Vision* 2004 in Prague, now forming Chapter 8 of this thesis. The reconstruction technique based on the mathematical method, Chapter 9, was later accepted for publication for the *Computer Vision and Pattern Recognition* 2004 in Washington [39]. Reconstruction of refractive materials, the latest method based on the surface evolution and joint work with Ivo Ihrke, will be presented at the *International Conference on Computer Vision* 2005 in Beijing [50].

I am very thankful for the excellent working conditions I was provided with here at the MPI. In particular, my gratitude goes to my supervisor Marcus Magnor, who expertly guided and supported my research. Thanks also to my friends and fellow researchers for helpful scientific discussions and lots of fun in pursuit of less serious matters. Finally, a special thanks to my family for their unconditional love and support over the years. I dedicate this thesis to Susanne, who makes my life so much better in countless ways.

Saarbrücken, Germany, *Bastian Goldlücke*
October 12, 2005