Methods for 3D Geometry Processing in the Cultural Heritage Domain

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Dipl.-Math. Gerhard Heinrich Bendels

Bonn, Dezember 2006
Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms Universität Bonn

Dekan: Prof. Dr. Matthias Winiger

1. Referent: Prof. Dr. Reinhard Klein
2. Referent: Prof. Dr. Stefan Gumhold

Tag der Promotion: 30.05.2007
Erscheinungsjahr: 2007
Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert
Contents

Preface vii
Abstract ix
Zusammenfassung xi
Acknowledgements xiii

I Introduction 1

1 3D Geometry in the CH Domain 3

2 Basics on 3D Geometry Processing 7
  2.1 Digital Object Representations 8
    2.1.1 Point Sets 8
    2.1.2 Parameterised Surfaces 10
    2.1.3 Triangle Meshes 10
    2.1.4 Implicit Representations 11
    2.1.5 Other Representations 12
  2.2 Acquisition 12
  2.3 Registration 14
  2.4 Reconstruction 17
    2.4.1 Moving Least Squares 17
    2.4.2 Radial Basis Functions 19
    2.4.3 Multi-level Partition of Unity Implicits 22
    2.4.4 Contour Extraction 23

3 Contributions 25

iii
II Model Generation

4 Reconstruction from Scan Data
   4.1 Related Work ........................................ 32
      4.1.1 Pairwise Registration .......................... 32
      4.1.2 Multiview Registration ......................... 35
   4.2 Feature Detection and Matching .................... 36
      4.2.1 Feature Surface Elements ....................... 37
      4.2.2 Consistent Feature Matching .................... 39
   4.3 Pairwise Registration ................................ 41
   4.4 Multiview Registration .............................. 44
      4.4.1 Graph Setup ..................................... 45
      4.4.2 Graph Collapse .................................. 46
      4.4.3 Relaxation ....................................... 46
   4.5 Results and Conclusions ............................. 48

5 Reconstruction from Dense Image Sets ............. 51
   5.1 Motivation ............................................ 51
   5.2 Related Work ......................................... 54
   5.3 Overview ............................................... 57
   5.4 Multi-Camera Grid .................................... 57
   5.5 Geometry Acquisition ................................ 59
      5.5.1 Visual Hulls ..................................... 60
      5.5.2 Efficient Evaluation on the GPU ................. 61
   5.6 Bidirectional Texture Functions (BTF) ............ 62
      5.6.1 Compression ..................................... 63
      5.6.2 Parameterisation ................................. 63
      5.6.3 Incomplete BTF data for non-planar objects .... 64
   5.7 Results and Conclusions ............................. 66

III Surface Completion .................................. 71

6 Hole Detection ........................................ 73
   6.1 Related Work ......................................... 74
   6.2 Overview ............................................... 75
   6.3 Boundary Probability ................................ 76


<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.3.1</td>
<td>Parameterisation</td>
<td>136</td>
</tr>
<tr>
<td>8.3.2</td>
<td>Shape Functions</td>
<td>138</td>
</tr>
<tr>
<td>8.3.3</td>
<td>Separating Handles and Anchors</td>
<td>139</td>
</tr>
<tr>
<td>8.4</td>
<td>Mesh Forging Process</td>
<td>141</td>
</tr>
<tr>
<td>8.4.1</td>
<td>The Algorithm</td>
<td>141</td>
</tr>
<tr>
<td>8.4.2</td>
<td>Defining the Occluder Field</td>
<td>144</td>
</tr>
<tr>
<td>8.5</td>
<td>Adaptive Refinement</td>
<td>147</td>
</tr>
<tr>
<td>8.6</td>
<td>Editing Examples</td>
<td>149</td>
</tr>
<tr>
<td>9</td>
<td>Modelling for Surface Inpainting</td>
<td>155</td>
</tr>
<tr>
<td>9.1</td>
<td>Introduction</td>
<td>155</td>
</tr>
<tr>
<td>9.2</td>
<td>Related Work</td>
<td>156</td>
</tr>
<tr>
<td>9.3</td>
<td>Framework Overview</td>
<td>159</td>
</tr>
<tr>
<td>9.4</td>
<td>Template Insertion</td>
<td>160</td>
</tr>
<tr>
<td>9.4.1</td>
<td>Non-Rigid Alignment</td>
<td>161</td>
</tr>
<tr>
<td>9.4.2</td>
<td>Generalised Constraints</td>
<td>162</td>
</tr>
<tr>
<td>9.5</td>
<td>The Modelling Framework</td>
<td>164</td>
</tr>
<tr>
<td>9.5.1</td>
<td>Geodesic Distances for Point Clouds</td>
<td>167</td>
</tr>
<tr>
<td>9.5.2</td>
<td>Dynamic Point Insertion</td>
<td>168</td>
</tr>
<tr>
<td>9.6</td>
<td>Surface Completion</td>
<td>168</td>
</tr>
<tr>
<td>9.7</td>
<td>Results</td>
<td>171</td>
</tr>
</tbody>
</table>

Conclusions 173
Bibliography 177
Data Sources 201
Curriculum Vitae 203
Abstracting information from its original medium or carrier to make its properties visible, understandable, or accessible in another manner is one of the fundamental concepts of scientific research. The concrete form or method of abstraction varies in time, in reply to the field of research and the scientific question at hand. It is also naturally influenced by the technological progress, as new technology has always spawned new research methodologies and led to the formation and validation of new scientific hypotheses.

The one technological development that originated and motivated much of the research presented in this thesis is 3D Photography – the generation of digital copies of three-dimensional objects. While this technology has a vast impact in various fields of automation, manufacturing and reverse engineering, the focus of the present work is its use – and the methods necessitated by its use – in the context of cultural heritage applications.

In the cultural heritage sector, the use of virtual, yet faithful reconstructions of valuable artefacts can break ground for numerous fascinating applications that were not within possibility before, given the various severely limiting restrictions that characterise the handling of their physical counterparts. On the other hand, applying computer graphics technologies in this sector also poses several special demands, such as dependability of the data, as well as the differentiation between acquired, reconstructed and guessed data.

Before this background, the present thesis investigates methods for acquiring, reconstructing and recreating virtual representations of cultural heritage artefacts, focussing on fully automatic methods wherever possible and on intuitive and easy to use interaction paradigms where necessary to take into account that future users of the presented methods will most likely be experts in a field different from computer science.
This thesis presents methods for 3D geometry processing under the aspects of cultural heritage applications. After a short overview over the relevant basics in 3D geometry processing, the present thesis investigates the digital acquisition of 3D models. A particular challenge in this context are on the one hand difficult surface or material properties of the model to be captured. On the other hand, the fully automatic reconstruction of models even with suitable surface properties that can be captured with Laser-range scanners is not yet completely solved. This thesis presents two approaches to tackle these challenges. One exploits a thorough capture of the object’s appearance and a coarse reconstruction for a concise and realistic object representation even for objects with problematic surface properties like reflectivity and transparency. The other method concentrates on digitisation via Laser-range scanners and exploits 2D colour images that are typically recorded with the range images for a fully automatic registration technique.

After reconstruction, the captured models are often still incomplete, exhibit holes and/or regions of insufficient sampling. In addition to that, holes are often deliberately introduced into a registered model to remove some undesired or defective surface part. In order to produce a visually appealing model, for instance for visualisation purposes, for prototype or replica production, these holes have to be detected and filled. Although completion is a well-established research field in 2D image processing and many approaches do exist for image completion, surface completion in 3D is a fairly new field of research. This thesis presents a hierarchical completion approach that employs and extends successful exemplar-based 2D image processing approaches to 3D and fills-in detail-equipped surface patches into missing surface regions.
In order to identify and construct suitable surface patches, self-similarity and coherence properties of the surface context of the hole are exploited.

In addition to the reconstruction and repair, the present thesis also investigates methods for a modification of captured models via interactive modelling. In this context, modelling is regarded as a creative process, for instance for animation purposes. On the other hand, it is also demonstrated how this creative process can be used to introduce human expertise into the otherwise automatic completion process. This way, reconstructions are feasible even of objects where already the data source, the object itself, is incomplete due to corrosion, demolition, or decay.

Auch nach der Registrierung sind die erfassten Datensätze in vielen Fällen unvollständig, weisen Löcher oder nicht ausreichend abgetastete Regionen auf. Darüber hinaus werden in vielen Anwendungen auch, z.B. durch Entfernen unerwünschter Oberflächenregionen, Löcher gewollt hinzugefügt. Für eine optisch ansprechende Rekonstruktion, vor allem zu Visualisierungszwecken, im Bildungs- oder Unterhaltungssektor oder zur Prototyp- und Replik-Erzeugung müssen diese Löcher zunächst automatisch detektiert und anschließend geschlossen werden. Obwohl dies im zweidimensionalen Fall der Bildbearbeitung bereits ein gut untersuchtes Forschungsfeld darstellt und vielfältige Ansätze zur automatischen Bildvervollständigung existieren, ist die Lage im dreidimensionalen Fall anders, und die Übertragung von zweidimensionalen Ansätzen in den 3D stellt vielfach eine große Herausforderung dar, die bislang keine
zufriedenstellenden Lösungen erlaubt hat. Nichtsdestoweniger wird in dieser Arbeit ein hierarchisches Verfahren vorgestellt, das beispielbasierte Konzepte aus dem 2D aufgreift und Löcher in Oberflächen im 3D unter Ausnutzung von Selbstähnlichkeiten und Kohärenzeigenschaften des Oberflächenkontextes schließt. Um plausible Oberflächen zu erzeugen werden die Löcher dabei nicht nur glatt gefüllt, sondern auch feinere Details aus dem Kontext rekonstruiert.

Abschließend untersucht die vorliegende Arbeit noch die Modifikation der vervollständigten Objekte durch Freiformmodellierung. Dies wird dabei zum einen als kreativer Prozess z. B. zu Animationszwecken betrachtet. Zum anderen wird aber auch untersucht, wie dieser kreative Prozess benutzt werden kann, um etwaig vorhandenes Expertenwissen in die ansonsten automatische Vervollständigung mit einfließen zu lassen. Auf diese Weise werden auch Rekonstruktionen ermöglicht von Objekten, bei denen schon die Datenquelle, also das Objekt selbst z. B. durch Korrosion oder mutwillige Zerstörung unvollständig ist.
Acknowledgements

Many have contributed in many ways to the work presented in this thesis, and they should not go unmentioned.

First and foremost, I have to thank my supervisor, Prof. Dr. Reinhard Klein, who was an inexhaustible source of inspiration and guidance. With his enthusiasm for computer graphics and related fields he created in only a couple of years a working environment here at the university of Bonn as creative and fruitful as it is. For invaluable feedback I owe thanks also to Prof. Dr. Andreas Weber who was always available on short notice for advice in so many things including (but not restricted to) computer graphics, photography, and travelling.

My sincere appreciation is due to the various co-authors of the papers I published over the years, and I am also sincerely indebted to my current and former colleagues here in the computer graphics group in Bonn, especially Marcin Novotni, Michael Guthe, Ruwen Schnabel, Patrick Degener, and Simone von Neffe. It is obvious to me that my research was only possible in the collaborative and friendly atmosphere that they provided. Particular thanks go to my office mates Ralf Sarlette and Mirko Sattler; the exciting and comprehensive discussions on matters of graphics kind, politics, and the countless patent-pending ideas on how to get rich will not be forgotten. Likewise I owe thanks to Gero Müller and Jan Meseth for compassionately discussing the terrible defeats or glorious victories of our favourite football teams.

It is with deep gratefulness that I think of my family, without whose love and support this work would not have been possible. I would have loved my father, Heinrich Bendels, to live and see this thesis of mine.
Part I

Introduction
Making pieces of art or of particular historic importance available to an audience as wide as possible is a key interest of historians, archaeologists and museums’ curators. Ideally, from a didactic point of view, each artefact would be demonstrated within its historic and semantic context, maybe even giving the observer the opportunity to interact and participate, granting access to everyone interested. Aside from the inherent and unsolvable dimensionality problem that any object can be demonstrated physically in one state representing a certain temporal snapshot only – even though it may have undergone important changes over the course of time, these requirements conflict with another fundamental interest at the very heart of every historian: The preservation of the objects under his auspices.

Generation of digital three-dimensional copies of the historic artefacts can build a bridge between these discordant and contradictory requirements associated with the handling of valuable artefacts. It therefore comes to no surprise that 3D content generation has gained (and is gaining at an increasing speed) much attention from the cultural heritage community, even although 3D photography is still a sophisticated process that comes yet nowhere near the practicability and ease of use of traditional 2D photography. Nevertheless, the use of computer graphics methods opens up manifold opportunities and paves the way to novel solutions to long standing challenges associated with standard tasks in the cultural heritage domain.
Among the most imminent concerns of museums curators is the documentation and cataloguing of the vast amount of artefacts in the museums’ inventories. Not infrequently do in particular high ranked museums have a backlog of up to 50%, and only a fraction of the valuable artefacts in a museum’s possession can typically be displayed due to limited space and limited personnel for preparation and handling. Instant and random access to the full spectrum of the items hidden in storage rooms and depots is not even envisioned yet for many museums. Here, 3D photography has the potential to tackle traditional challenges and facilitate novel ways for dissemination and accessibility in a manner unknown with traditional photography.

In addition to this, the use of computer graphics methodologies in the cultural heritage domain can be expected to boost (and is doing so today) a number of applications that are feasible for the first time with the aid of captured 3D geometry:

- Quantitative / statistical analysis
- Restoration planning / documentation
- VR / AR applications
  - Virtual reassembly
  - virtual stress tests / check for plausibility of previous reconstructions
  - Virtual historic/spatial contexts
  - Virtual reconstruction in different states of evolvement
- Prototype generation for
  - Manufacturing of restoration parts
  - Mold generation for replica generation or merchandising
  - Mold generation for packaging and transport
  - Precisely formed supports / stands.

Although for most of the aforementioned applications a faithful digital reproduction of the object under consideration is required
and suffices, the endeavour to exploit the acquired data in education and entertainment applications also calls for methodologies to perform artistic and creative operations. To successfully enable experts in the field to incorporate their knowledge into the virtual restoration, make-up and presentation, however, these methodologies have to fulfill some specific requirements:

- The modelling interface must be as simple and intuitive as possible, as future users can be expected to be experts in a field different from computer science.

- Precise definition of the region of influence of an editing operation is obligatory, as in particular historians always need to be able to distinct modified from original data.

- The modelling technology cannot be restricted to the deformation of the object only; it is essential that also considerable material may be added or removed, including parts of other objects.

- Although detail preservation in this context is a virtue, it stands back behind the importance of the possibility to reconstruct or recreate fine surface detail in regions of added material to match the surrounding of the editing region.
CHAPTER 2

BASICS ON 3D GEOMETRY PROCESSING

Understood as a general scientific term, Modelling refers to the creation of a (typically simplified) representation of a system or phenomenon that retains specifically those properties of the original required by the application or scientific question at hand (cf. [Costello 1991]). In a less abstract sense, a model can also be an artificial instance of a physical object which is stripped from all its properties that are irrelevant in the respective context. As such, the prototypical clay miniature that designers create in the early stages of conceptualising a new car can serve as an example of a model, in this case capturing shape as the most important aspect. In the present thesis it is this last type that is referred to as models – virtual, digital descriptions of a physical object. In principle, such models can be generated in two fashions, either by ab initio-creation or by capturing the desired properties using

Figure 2.1: Model generation / Abstraction.
some kind of 3D photography, e.g. using Laser range scanners, structured light, tactile sensors, or volumetric methods like CT or \(\mu\)CT. The properties that are primarily required to be captured in the model are shape, appearance, and sometimes some aspects of the material the physical object is made of.

With respect to shape, a number of representations have been developed over the years in computer graphics and can be considered mainstream: Polygonal meshes, subdivision surfaces, level set surfaces, point sets, NURBS-surfaces, to name a few. All of these have their specific strengths and weaknesses, depending on the application they are used in – not all representations are equally suitable for all applications for all types of models, just as not all processing strategies are equally suitable for all representations.

Unlike in the automotive and manufacturing industries, where CAD-tools dominate the digital creation process, creating 3D models from scratch is generally not the method of choice in the cultural heritage sector, as it would antagonise many of the applications mentioned above, in particular with respect to documentation and analysis. Various projects do exist that aim at generating virtual models of buildings and architecture using CAAD-tools, and the grammar-based modelling of historic architecture is currently gaining increasing research attention (see [Havemann 2005]), but modelling in particular sculptures etc. is yet virtually unfeasible. The object representations predominantly used in this thesis are therefore those directly related to the data capturing process, i.e. point sets, implicit representations as distance fields, and triangle meshes. The following sections will give a short overview over some of the most popular surface representations.

2.1 Digital Object Representations

2.1.1 Point Sets

Among the more popular representations of a 2D-surface in 3D, the point set representation certainly constitutes the conceptually simplest. For a given 2-manifold surface \(\mathcal{S}\) in \(\mathbb{R}^3\), it consists of no
2.1. Digital Object Representations

Figure 2.2: Shape, its digital representation, the processing tools, and the application have a natural influence onto each other.

more than a finite set $\mathcal{P}$ of point samples of $\mathcal{S}$:

$$\mathcal{P} = \{ p_1, \ldots, p_N \in \mathbb{R}^3 \},$$

where $N \in \mathbb{N}$, and $p_i \in \mathcal{S}$ for all $i = 1, \ldots, N$. In its basic formulation no additional knowledge such as connectivity, spatial structure, etc. is required. Nevertheless, a point’s position is often paired with other attributes like colour and normal, and the resulting n-tuple is usually referred to as *surfel* ([Pfister et al. 2000]). For many applications, e.g. rendering and the transformation into other surface representations (see section 2.4), in particular the *surface normals* are required and either captured together with $\mathcal{P}$ from the surface $\mathcal{S}$ or derived from $\mathcal{P}$ using local surface analysis operators. Throughout this thesis, $\mathbf{n}(p)$ denotes the surface normal at some point $p \in \mathbb{R}^3$.

Representing surfaces through point sets has become increasingly popular in the past few years. One reason for this is the spreading of affordable 3D capture technology in form of laser range scanners outputting easily millions of point samples, and thereby producing a faithful sampling of the scanned surface. The other reason, that is equally important, is the development of powerful algorithms that prepare the ground for numerous applications to be employed directly on the point set itself without the need to perform a full-scale reconstruction first.
Figure 2.3: An ellipsoidal surface represented (from left to right) implicitly as level set \( \{ \mathbf{p} \in \mathbb{R}^3 \mid p_x^2 + 0.5p_y^2 + p_z^2 - 1 = 0 \} \), as triangle mesh, and as point set (overlaid over the triangle mesh for illustration purposes).

2.1.2 Parameterised Surfaces

Let \( I \subseteq \mathbb{R} \) be an interval, and let \( f : I \times I \rightarrow \mathbb{R}^3 \) be a continuous function. Then a parameterised surface \( S \) is defined as the set

\[
S = \{ \mathbf{p} \in \mathbb{R}^3 \mid \exists (u, v) \in I \times I : \mathbf{p} = f(u, v) \}.
\]

An example of this representation are the famous Non-uniform Rational B-Spline (NURBS) surfaces (see e.g. [Farin 1990]) that were used extensively (and are still today) in the engineering industries to describe the building components of cars. However, even in the automotive industries, where the desirable computational properties of NURBS surfaces are indispensable, their use is currently pushing limits, as the number of trimmed NURBS patches for a single car reach the order of millions. Constructing parameterised surfaces for captured physical models is non-trivial and computationally prohibitive.

2.1.3 Triangle Meshes

One special case of parameterised surfaces are triangle meshes. A triangle mesh is a piecewise linear surface that can be formulated as the pair \((\mathcal{V}, \mathcal{E})\), consisting of a set \(\mathcal{V}\) of vertices in \(\mathbb{R}^3\), representing the surface’s geometry and a set \(\mathcal{E} \subset \mathcal{V} \times \mathcal{V}\) of edges capturing the connectivity and the topology of the surface. The edges in \(\mathcal{E}\) form a planar graph whose faces are triangles in \(\mathbb{R}^3\).
2.1. Digital Object Representations

One reason for the popularity of this type of surface representation is the fact that nowadays’ graphics hardware is heavily tuned to rapidly handle large amounts of triangle mesh data, such that more than 100 million triangles can be rendered per second (according to NVIDIA, up to 181 million triangles per second on the Quadro FX 4500).

2.1.4 Implicit Representations

In contrast to parameterised surfaces which can be considered the image of a function, level set surfaces are defined implicitly as the kernel of a functional: Let $F : \mathbb{R}^3 \to \mathbb{R}$ be a continuous functional, then a level set surface $S$ is defined as the set

$$S = \{ p \in \mathbb{R}^3 \mid F(x,y,z) = 0 \}.$$

Algebraic surfaces, e.g. the ellipsoid $\{ p \in \mathbb{R}^3 \mid p_x^2 + 0.5p_y^2 + p_z^2 - 1 = 0 \}$, depicted in figure 2.3, are examples of this type of surface representation, where the surface points are roots of polynomials up to a given degree. Another, more abstract approach to implicit surfaces, is the zero-set of the distance field defined by the surface itself:

$$F(p) = \delta(p, S),$$

where $\delta$ is some appropriate (signed) distance measure between a point in $\mathbb{R}^3$ and the surface (or some surface approximation, see section 2.4).

By definition, implicit surfaces deliver an ubiquitous availability of inside/outside information by simply evaluating the respective functional $F$. In addition to that, implicit surfaces are powerful representations of topologically complex surfaces, that can easily handle topological changes during modelling. On the other hand, implicit surfaces do not enable any direct access to the surface itself. For an elaborate coverage of implicit surfaces see [Bloomenthal 1997].
2.1.5 Other Representations

All of the aforementioned surface representations are concerned with the geometry of the surface to be represented. In some applications and in particular for rendering, however, its appearance is more important. As a consequence, *image-based methods* that allow regarding the object from arbitrary view-points and under arbitrary lighting conditions have been developed. It has been demonstrated e.g. by Hawkins et al. [2001], that for demonstration and illustration purposes of an object itself, it can be represented by a dense set of photographic images without explicit geometry reconstruction. By basing the representation on photographs only, this approach is viable even for objects for which a faithful geometry capture is at least challenging, e.g. materials such as leather, feather, or fur. Nevertheless, numerous applications involve at least coarse geometry. Chapter 5 will discuss this hybrid type of representation where geometric fine detail is represented by images whereas the coarse geometry is represented explicitly as triangle mesh.

2.2 Acquisition

As of today, the technologies available to digitise three-dimensional physical objects can typically be categorised into three groups: *Laser-based, Computer Tomography* (CT or \(\mu\)CT),
2.2. Acquisition

and Light Field methods that are based on the exploitation of dense sets of photographic images.

<table>
<thead>
<tr>
<th></th>
<th>Surface</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact</td>
<td>Tactile Sensors</td>
<td>Histologies</td>
</tr>
<tr>
<td>Non-Contact</td>
<td>3D Photography</td>
<td>CT, µCT, MRT</td>
</tr>
</tbody>
</table>

Among the methods that use Laser-emitting devices, one can in turn distinguish three different range finding principles: Time-of-Flight, triangulation, and interferometric scanners. Due to the limited precision of time-of-flight scanners and the considerable cost of interferometric scanners, triangulation scanners are most widespread. They combine a simple setup (compared to interferometric scanners) resulting in mid-price devices with a reasonable precision of few microns under ideal conditions. For a given viewpoint, these non-contact capturing devices record a rectangular array of depth values (a so-called Range Image) by tracking contour lines projected onto the physical object (cf. figure 2.5). The depth values represent the distance between the capturing camera and the corresponding point on the object’s surface and therefore encode the point’s position in space relative to the scanning device. As a precise tracking of the global position of the scanner is generally unfeasible, each range image is given in its own coordinate system. This necessitates the transfer of each individual range image into one common, global coordinate system – a process that is usually referred to as registration. This next step in the acquisition pipeline after capturing the raw data will be described in the following section (section 2.3).

Computer tomography methods raster the space containing the object into a finite set of voxels and record density values for each voxel. Here, the full data is recorded in a common coordinate system such that no registration is necessary. Moreover, computer tomography allows a look on the inside of the object, which is particularly useful for objects with an intricate interior structure. On the other hand, the fact that only density values are recorded necessitates contour extraction algorithms to be applied before an actual surface can be generated. This step of the acquisition
pipeline is usually referred to as reconstruction and will be the topic of section 2.4.

Data acquisition using CT suffers from a limited resolution, while $\mu$CT delivers high resolution recordings at the cost of considerably confined measurement dimensions.

Acquiring the light field representation of an object involves the sampling of the space of all light directions and of all viewing directions. To this end, the object is lit from predefined lighting positions and under each such elementary illumination, a photograph is taken from equally predefined viewing positions. The acquisition setups presented so far ([Malzbender et al. 2001], [Hawkins et al. 2001], [Hawkins et al. 2005], [Koch 2006]) differ mainly in the sampling density with respect to light and viewing directions, and the extent to which the acquisition is parallelised. In particular the setup developed by Sarlette et al. described in [2006] delivers fast acquisition times of less than 20 minutes (three times as much for high dynamic range recordings) by massive parallelisation.

2.3 Registration

Range images captured with a laser range scanner are point clouds, each sampling a certain part of the digitised object and
given in its own local coordinate system. To derive a complete closed surface sampling, the range images have therefore to be transferred into a common global coordinate system. This registration step is traditionally performed by a manual pre-registration step that coarsely pairwise aligns the range images, followed by an iterative automatic fine alignment.

The goal of range image alignment is to find a coordinate transformation that maps points representing an identical position on the scanned surface to coinciding positions in space. Pairs of such points in the overlapping regions of two range images are called corresponding. Unfortunately, the correspondences between recorded range images are generally unknown and – although at least in part easily identifiable for a human observer – not accessible for an automated registration. It is worth noting that knowledge of the complete set of correspondences is equivalent to knowing the coordinate transformation between two range images. The concept of most successful approaches to (semi-)automatic registration is therefore an iterative procedure that alternately estimates correspondences and transformations (cf. figure 2.6).

**Iterative Closest Points**

Let $\mathcal{P} = \{p_1, \ldots, p_N\} \subset \mathbb{R}^3$ be a range image, and let $\mathcal{X} = \{x_1, \ldots, x_N\} \subset \mathbb{R}^3$ be a point set with which $\mathcal{P}$ is to be aligned. In the context of registration, $\mathcal{P}$ is also denoted as the *data*, while
Chapter 2. Basics on 3D Geometry Processing

$\mathcal{X}$ is called the model. Let $C_{\mathcal{X}} : \mathbb{R}^3 \to \mathcal{X}$ be the operator that finds for any point $p$ in the Euclidean space the closest point $x_p$ in the model $\mathcal{X}$:

$$x_p = C_{\mathcal{X}}(p) = \arg\min_{x \in \mathcal{X}} d(p, x)$$

according to some distance function $d$, typically the Euclidean distance

$$d(p, x) = \langle p - x, p - x \rangle^{1/2}.$$ 

Here, $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in $\mathbb{R}^3$.

The key component of the popular ICP method introduced by Besl et al. [1992] is to assume that the geometrical proximity of $p$ to $C_{\mathcal{X}}(p)$ indicates correspondence, i.e. semantic equivalence. ICP consequently interprets $\{(p_i, x_{p_i})\}_{i=1,\ldots,k} \subset \mathcal{P} \times \mathcal{X}$, with $k \leq N$, to be corresponding point pairs and tries to find a rotation $R$ and a translation $t$, such that

$$\frac{1}{k} \sum_{i=1}^{k} d(x_{p_i}, Rp_i + t)^2 \to \min_{R,t}!$$

After the minimising transformation $(R, t)$ is found, it is applied to $\mathcal{P}$ and the procedure is iterated, each time updating the correspondences using the closest-point-operator $C_{\mathcal{X}}$.

Initially, the above assumption is unjustified. The minimisation is therefore prone to lead to a local minimum that is arbitrary distant to the true solution. To counteract this, the iteration is initiated by a manual identification of corresponding point pairs, i.e. by selecting and declaring a sparse set of corresponding point pairs in the data and the model point set.

Although several extensions to the ICP algorithm have been developed over the years that help avoiding local minima and circumvent false correspondence computations, the fully automatic registration without manual pre-alignment is still a matter of current research. Chapter 4 will give a more detailed overview over the existing approaches and will also introduce a solution to this problem that exploits 2D-intensity images that are frequently recorded during scanning by most available off-the-shelf Laser range scanners.
2.4 Reconstruction

Given powerful and efficient algorithms to directly handle large point sets outputted from the previous stages in the content creation pipeline, many applications can deal with the point set representation itself and no further processing is required. For other applications however, a closed surface representation in form of a polyhedral surface (such as a triangle mesh) is still required or can at least be exploited for performance or efficiency reasons.

The goal of surface reconstruction as formulated by Hoppe et al. [1992] is therefore as follows:

Given a surface sampling $\mathcal{P}$ of points $p_i \in \mathbb{R}^3$, $i = 1, \ldots, N$ on or near a surface $S$, determine a surface approximation $\hat{S}$ to $S$.

This conversion from a point sampling into a surface approximation $\hat{S}$ is typically performed as a two-step process: Firstly, an implicit representation is constructed from the point set – for instance in form of the zero set of the surface’s distance field ([Hoppe et al. 1992], [Ohtake et al. 2003]), of a radial basis function fitted to the given surface data ([Carr et al. 2001]), or in form of the stationary set of a projection operator ([Levin 2003]). In a subsequent contouring step, this implicit representation is converted into a polygonal mesh by an iso-surface extraction algorithm.

This two-step reconstruction paradigm is also applied if faced with data from intermediate stages in industrial CAD processes which often suffers from topological inconsistencies, cracks in the tessellation etc. In this case the implicit representation is chosen to faithfully reproduce the surface properties, whereas the contour extraction can be tuned to generate surface meshes with the desired properties.

2.4.1 Moving Least Squares

The basic idea of the well-known moving least squares surface interpolation scheme by Levin [2003] is to define $\hat{S}$ as stationary
Figure 2.7: The moving least squares projection operator. Left: For a given point \( r \in \mathbb{R}^3 \) near the approximated surface \( S \), the minimising hyperplane \( H \) is found and a local coordinate system with origin at \( q \) is defined. Right: The projection operator is defined as a mapping of \( r \) onto the best fitting polynomial \( \pi \in \Pi_m^2(H) \).

A set of a projection operator \( P_m \) that projects points close to sufficiently sampled regions of the original surface \( S \) onto \( \hat{S} \): Let \( r \) be a point near the approximated surface \( S \); a hyperplane \( H = (a, D) \) with normal \( a \in \mathbb{R}^3 \), \( \|a\| = 1 \) and distance \( D \in \mathbb{R}^\geq 0 \) to \( r \) is defined s.t.

\[
\sum_{p_i \in P} (\langle a, p_i \rangle - D)^2 \theta(\|p_i - q\|) \rightarrow \min_{a,D}, \tag{2.1}
\]

where \( \theta \) is a non-negative weight function and \( q \) is \( r \)'s projection onto \( H \) (see figure 2.7, left). The resulting hyperplane \( H \) defines a local reference domain with its origin located at \( q \in H \).

In a second step, a polynomial \( \pi \in \Pi_m^2(H) \) of degree \( m \) is fitted to the residuals \( f_i = \|p_i - q_i\| \) such that

\[
\sum_{p_i \in P} (\pi(q_i) - f_i)^2 \theta(\|p_i - q\|) \rightarrow \min_{\pi}
\]

becomes minimal. It is worth noting that in the above equations the weight function is evaluated depending on the distance between the sample points \( p_i \) and the origin \( q \), i.e. the projection of \( r \) onto \( H \).

The MLS surface finally is defined to be the stationary set of the projection operator that maps a point \( r \) near \( S \) to the corresponding point on the polynomial:

\[ r \mapsto r' = q + \pi(0)a. \]
In many cases, normals are not known by measurement and need to be assigned to each point \( p \in \mathcal{P} \), and it is tempting to use the normal \( a(p) \) of the approximating hyperplane \( H \) as defined above. However, as stressed by Alexa and Adamson [2004], \( a(p) \) is not necessarily collinear with the surface normal, although it is frequently employed as such. For further details cf. the above publication, as well as [Adamson & Alexa 2003] and [Adamson & Alexa 2004].

As pointed out by Klein and Zachmann [2004], the weight function \( \theta \) in equation (2.1) depends on the Euclidean distance, which does not respect any topology potentially present in the data, and hence may declare points ”close” that are indeed, at least topologically, far away. Although on first sight this corresponds well to the fact that point sets do not explicitly store topology information, Klein and Zachmann correctly argue that proximity graphs can be used to approximate geodesic (and therefore surface-inherent) distances to overcome these limitations at the cost of only little storage overhead.

Although moving least squares surfaces in their original formulation are smooth by definition and therefore unable to reproduce sharp features, recent approaches have introduced piecewise smooth surface representations based on the MLS. See e.g. Fleishman et al. [2005], who used statistics methods to detect and preserve sharp features present in point sampled data in the MLS representation.

### 2.4.2 Radial Basis Functions

In contrast to the local approximation nature of the moving least squares scheme, where the approximating function (the approximant) is a low degree polynomial defined over a local domain available only in close vicinity to the approximated surface (the approximand), radial basis function (RBF) interpolation derives one implicit function whose zero set globally defines the approximating surface \( \hat{S} \).\(^1\) Its task is therefore to find a function \( \hat{s} \) such

\[\text{\footnote{In order to avoid confusion between interpolation and approximation, please note that the approximation} \( \hat{S} \) \text{ of the original surface} \ S \text{ is derived in this section via interpolation}}\]
that \( \hat{s}(p) = 0 \) for all \( p \in \mathcal{P} \).

In order to avoid trivial solutions like \( s = 0 \), boundary constraints are inserted that define non-zero values for off-surface points, one obvious choice being the signed distance of these points to the approximated surface (cf. [Carr et al. 2001]). The complete interpolation problem can then be stated as:

Given a surface sampling \( \mathcal{P} \) of points \( p_i \in \mathbb{R}^3, \ i = 1, \ldots, N \) on or near a surface \( S \), and further a set of off-surface points \( p_j \in \mathbb{R}^3, \ j = N + 1, \ldots, N + M \) find a function \( \hat{s} \) such that

\[
\hat{s}(p_i) = s_i \quad \text{for all} \quad i = 1, \ldots, N + M
\]

where \( s_i = 0 \) for \( i = 1, \ldots, N \). For \( i = N + 1, \ldots, N + M \), \( s_i \) denotes the off-surface points’ distance to the approximated surface.

In the above problem statement, the space from which the optimal function \( \hat{s} \) may stem is unspecified. In practice, the choice of an appropriate function space is influenced by additional smoothness assumptions for the approximand, which lead to optimality criteria for the approximant, and by computational considerations, which typically lead to finite dimensional function spaces.

Both conditions are fulfilled by setting the space of allowable interpolants to be

\[
\left\{ s(x) = p(x) + \sum_{i=1}^{N+M} \lambda_i \phi(||x - p_i||) \right\},
\]

where \( \phi : \mathbb{R}^2 \rightarrow \mathbb{R} \) are the so-called radial basis functions and \( p(x) \) is a low degree polynomial.

The simplest example for an RBF interpolation is the interpolation with finite linear combinations of translations of the radially symmetric function \( \phi(r) = r \):

\[
s(x) = p(x) + \sum \lambda_i ||x - p_i||,
\]

of the sample points in \( \mathcal{P} \). In this sense, \( \hat{S} \) is an approximant to \( S \) and an interpolant to \( \mathcal{P} \).
with a linear polynomial $p$ and the Euclidean norm $\| \cdot \|$. It was shown by Duchon [1977] that the smoothest interpolant in the space of Beppo-Levi distributions on $\mathbb{R}^3$ is guaranteed to have this particular form. (See e.g. [Carr et al. 2001] for further details.)

Other examples for frequently employed radial basis functions include

$$
\begin{align*}
\phi(r) &= r & \text{Biharmonic} \\
\phi(r) &= r^2 \log(r) & \text{Thin plate} \\
\phi(r) &= r^3 & \text{Triharmonic} \\
\phi(r) &= e^{-ar^2} & \text{Gaussian}
\end{align*}
$$

The specific choice of a radial basis function for a given interpolation problem strongly influences not only the resulting interpolant but also the computational effort to solve for the required coefficients $\lambda_i$ and those of $p$. In particular, radial basis functions with global support (for instance the triharmonic RBF $\phi(r) = r^3$) deliver fair surface approximations that are able even to cover large holes in the input sampling at the price of a dense matrix in the corresponding linear system (see below). RBFs with a more local support lead to sparse matrices that can be solved far more efficiently, but may generate surfaces with undesired properties.

Let $\pi_1, \ldots, \pi_l$ be the basis for the space of polynomials up to degree $l$ and let $c_1, \ldots, c_l$ be the corresponding coefficients of $p$ in this basis. The linear system for the desired interpolant can then be written as

$$
\begin{pmatrix}
\Phi & \Pi \\
\Pi^T & 0
\end{pmatrix}
\begin{pmatrix}
\lambda \\
c
\end{pmatrix}
= 
\begin{pmatrix}
s \\
0
\end{pmatrix},
$$

where $\Phi_{ij} = \phi(\|\mathbf{p}_i - \mathbf{p}_j\|)$ and $\Pi_{ij} = \pi_j(\mathbf{p}_i)$. This linear system also includes the orthogonality conditions that require

$$
\sum_{i=1}^{N} \lambda_i = \sum_{i=1}^{N} \lambda_i \mathbf{p}_i = 0.
$$

Regarding this linear system, it becomes obvious why RBF interpolation has traditionally been considered inappropriate for reconstruction purposes where point clouds easily reach the size of
millions of points. Even for reduced problem sizes in the context of surface modelling, solving the above linear equation remains computationally involved, although recent approaches have proven their feasibility even in real-time applications [Botsch & Kobbelt 2005].

The biggest advantage of RBF interpolation techniques is that they are able to derive smooth interpolants under only very mild conditions on the placement of the RBF centres – by construction, since the placement of the centres itself influences the function space. For function spaces that are defined without respect to the positions of the points to be interpolated, it is typically not difficult to construct examples that lead to singularities and non-invertible matrix representations.

For in-depth reading please refer to [Buhmann 2003], [Duchon 1977], [Carr et al. 2001] and [2003], and the references given therein.

2.4.3 Multi-level Partition of Unity Implicits

One of the most efficient approaches known to date to build a surface representation from large sets of sampled surface points are the so-called Multi-level Partition of Unity Implicits introduced by Ohtake et al. [2003]. This approach derives an implicit surface representation from point samples by computing local quadratic surface approximations in octree cells (the so-called local shape functions\(^2\)). The local surface approximations are then blended in order to generate a closed, smooth surface. The name MPU is due to the fact that the blending functions (the weights) sum to one at every input point.

More specifically, let \( \mathcal{P} \) as usual denote a set \( \{ \mathbf{p}_1, \ldots, \mathbf{p}_N \} \) of sample points representing a 2D surface \( S \) in \( \mathbb{R}^3 \) with associated surface normals \( \{ \mathbf{n}_1, \ldots, \mathbf{n}_N \} \). The goal is now to derive, in an adaptive manner a function \( f : \mathbb{R}^3 \to \mathbb{R} \) whose zero level set approximates the unknown underlying surface.

\(^2\)The name local shape functions is adopted here for reference reasons. This should not cause any confusion, however, with the notion of shape functions for modelling used in chapter 8.
To this end, the space surrounding the point set is subdivided adaptively and local surface approximations are constructed hierarchically in a top-down fashion. For each cell $\Omega$ of the space subdivision scheme, a quadratic functional is fitted. It is then checked, if the approximation error of this function with respect to the points in $\Omega$ exceeds a certain threshold. In this case, $\Omega$ is subdivided and the procedure is performed recursively on its child cells.

During fitting, the points contained in $\Omega$ are analysed to select the specific type of functionals to be used for approximation. In order to be able to also represent edge- and corner-shaped sharp features, the options include a heuristic to detect sharp features and the use of piecewise quadratic functions.

The main advantages of the MPU as implicit surface approximation is that it is error-adaptive and fairly fast.

### 2.4.4 Contour Extraction

The standard approach for contour extraction from implicit functions is the famous *Marching Cubes*-algorithm introduced by Lorensen and Cline [1987]. In its original formulation, it generates
a piecewise linear approximation of the implicit surface based only on inside / outside information that is specified at the nodes of a uniform grid. Exploiting symmetry relations, the set of possible inside / outside combinations at the eight grid nodes belonging to a grid cell can be reduced to 15 (see figure 2.8). Each grid cell is then processed and triangulated individually. If a consistent strategy is pursued to resolve ambiguous cases, the marching cubes algorithm is guaranteed to produce a closed 2-manifold mesh that separates nodes marked inside from those marked outside.

Over the years, numerous extensions and improvements of the marching cubes algorithm have been introduced, among others enabling the reconstruction of features inside octree cells ([Kobbelt et al. 2001]), and allowing for adaptive grids to be handled ([Bloomenthal 1988],[Shu et al. 1995]). Adaptivity was restricted by the fact, though, that adjacent grid cells were not allowed to differ by more than one octree level. Combining the dual surface nets approach by S.F.F.Gibson [1998] with ideas from [Kobbelt et al. 2001], this restriction was relieved by the dual contouring approach introduced by Ju et al. [2002]. Later, the dual surface nets approach was adapted to kd-trees by Greß and Klein [2003; 2004], further enhancing adaptivity and allowing even for thin solid structures to be faithfully reconstructed. For further reading, cf. [Greß & Klein 2004] and the references therein.
Motivated by the numerous and manifold applications in the cultural heritage sector that become viable – in some cases even for the first time – using computer graphics techniques, this thesis investigates methods for 3D geometry processing under consideration of the specific aspects of its use in this particular field. The work presented in this thesis summarises (and extends) work published in various papers as listed in section 9.7. The main contributions are:

- A selection of 3D model generation algorithms that enable users to efficiently produce digital copies of existing models in a fully- or semi-automatic fashion on the basis of various types of data sources (part II). Of particular relevance for the subsequent contributions is a fully automatic range image registration approach that exploits features contained in the 2D photographic images typically recorded together with the range images.

- A fully automatic hole detection approach for point sampled surfaces that paves the way for a context-based, hierarchical surface completion algorithm for point sampled surfaces that is able to reconstruct large scale as well as fine detail features in the hole region (part III).

- An interactive and semi-automatic modelling paradigm (part IV) that allows for intuitive and efficient free-form modification of 3D surface data and can seamlessly be integrated into the surface completion previously introduced in part III.
The methods and algorithms presented in this thesis are described and motivated with their applicability in the context of cultural heritage that offers a variety of applications and at the same time poses some specific demands. Nevertheless, most of the contributions of this thesis are applicable and relevant as such in many other contexts.

The thesis is organised in four parts. After the introductory part, part II (from page 27) deals with the generation of 3D models based on range images and from dense image sets. Focussing on the results from the registered point-based objects, the topic of part III (from page 71) is the repair and completion of the generated models, while the last part (from page 121) of this thesis considers modelling methods, that can also be used in conjunction with the automatic surface completion approach described in the preceding part.