ROBUST RANGE IMAGE REGISTRATION

USING GENETIC ALGORITHMS AND THE SURFACE INTERPENETRATION MEASURE

Luciano Silva • Olga R. P. Bellon • Kim L. Boyer

SERIES IN
MACHINE PERCEPTION & ARTIFICIAL INTELLIGENCE
Volume 60

World Scientific

Series in Machine Perception and Artificial Intelligence – Vol. 60

ROBUST RANGE IMAGE REGISTRATION

USING GENETIC ALGORITHMS AND THE SURFACE INTERPENETRATION MEASURE

Luciano Silva and Olga R P Bellon
Universidade Federal do Paraná, Brazil

Kim L Boyer
Ohio State University, USA

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ROBUST RANGE IMAGE REGISTRATION USING GENETIC ALGORITHMS AND THE SURFACE INTERPENETRATION MEASURE
Series in Machine Perception and Artificial Intelligence — Vol. 60

ISBN 981-256-108-0

Printed in Singapore by Mainland Press

ROBUST RANGE IMAGE REGISTRATION

USING GENETIC ALGORITHMS AND THE SURFACE INTERPENETRATION MEASURE

SERIES IN MACHINE PERCEPTION AND ARTIFICIAL INTELLIGENCE*

Editors: **H. Bunke** (Univ. Bern, Switzerland)
P. S. P. Wang (Northeastern Univ., USA)

Vol. 44: Multispectral Image Processing and Pattern Recognition
(Eds. *J. Shen, P. S. P. Wang and T. Zhang*)

Vol. 45: Hidden Markov Models: Applications in Computer Vision
(Eds. *H. Bunke and T. Caelli*)

Vol. 46: Syntactic Pattern Recognition for Seismic Oil Exploration
(*K. Y. Huang*)

Vol. 47: Hybrid Methods in Pattern Recognition
(Eds. *H. Bunke and A. Kandel*)

Vol. 48: Multimodal Interface for Human-Machine Communications
(Eds. *P. C. Yuen, Y. Y. Tang and P. S. P. Wang*)

Vol. 49: Neural Networks and Systolic Array Design
(Eds. *D. Zhang and S. K. Pal*)

Vol. 50: Empirical Evaluation Methods in Computer Vision
(Eds. *H. I. Christensen and P. J. Phillips*)

Vol. 51: Automatic Diatom Identification
(Eds. *H. du Buf and M. M. Bayer*)

Vol. 52: Advances in Image Processing and Understanding
A Festschrift for Thomas S. Huwang
(Eds. *A. C. Bovik, C. W. Chen and D. Goldgof*)

Vol. 53: Soft Computing Approach to Pattern Recognition and Image Processing
(Eds. *A. Ghosh and S. K. Pal*)

Vol. 54: Fundamentals of Robotics — Linking Perception to Action
(*M. Xie*)

Vol. 55: Web Document Analysis: Challenges and Opportunities
(Eds. *A. Antonacopoulos and J. Hu*)

Vol. 56: Artificial Intelligence Methods in Software Testing
(Eds. *M. Last, A. Kandel and H. Bunke*)

Vol. 57: Data Mining in Time Series Databases
(Eds. *M. Last, A. Kandel and H. Bunke*)

Vol. 58: Computational Web Intelligence: Intelligent Technology for
Web Applications
(Eds. *Y. Zhang, A. Kandel, T. Y. Lin and Y. Yao*)

Vol. 59: Fuzzy Neural Network Theory and Application
(*P. Liu and H. Li*)

Vol. 60: Robust Range Image Registration Using Genetic Algorithms
and the Surface Interpenetration Measure
(*L. Silva, O. R. P. Bellon & K. L. Boyer*)

*For the complete list of titles in this series, please write to the Publisher.

Luciano Silva and Olga Bellon would like to dedicate this work to each other. Kim Boyer would like to dedicate this work to his family: wife Ann, son Jeff, and daughter Sandy.

Preface

The book addresses the range image registration problem for automatic 3D model construction. We focus on obtaining highly precise alignments between different view pairs of the same object to avoid 3D model distortions. In contrast to most prior work, the view pairs may exhibit relatively little overlap and need not be prealigned. To this end, we define a novel effective evaluation metric for registration, the Surface Interpenetration Measure (SIM). This measure quantifies the interleaving of two surfaces as their alignment is refined, putting the qualitative evaluation of "splotchiness", often used in reference to renderings of the aligned surfaces, onto a solid mathematical footing. The SIM is shown to be superior to mean squared error (i.e. more sensitive to fine scale changes) in controlling the final stages of the alignment process.

We then combine the SIM with Genetic Algorithms (GAs) to develop a robust approach for range image registration. The results confirm that this technique achieves precise surface registration with no need for prealignment, as opposed to methods based on the Iterative Closest Point (ICP) algorithm, the most popular to date. We present thorough experimental results, including an extensive comparative study and propose enhanced GA-based approaches to improve the registration still further. Additionally, we develop a global multiview registration technique using our GA-based approach. The results show considerable promise in terms of accuracy for 3D modeling.

L. Silva, Olga R.P. Bellon, Kim L. Boyer

Contents

Chapter 1

Introduction

Building realistic 3D models from sensor data is still a challenging problem. In recent years the demand for reconstruction and modeling of objects is increasing and it is widely used in many research areas, including medical imaging, robotic vision and archaeology [Ikeuchi and Sato (2001)]. Most applications focus on developing techniques to construct precise 3D object models of physical objects, preserving as much information as possible.

Recently, projects in digital archaeology have presented new challenges and are gaining popularity in computer vision community such as the Digital Michelangelo project [Levoy *et al.* (2000)] and the Great Buddha project [Ikeuchi and Sato (2001)]. A primary objective of these efforts is the digital preservation of cultural heritage objects before degradation or damage caused by environmental factors, erosion, fire, flood, or human development. Some collaborative efforts have supported the repair and restoration of historic buildings, construction of virtual museums, teaching using 3D visualization, and the analysis of complex structures by their 3D models [Bernardini *et al.* (2002)]. Additionally, in 2003 important conferences have organized sections to discuss the issue, such as the IEEE/CVPR Workshop on Applications of Computer Vision in Archaeology (ACVA'03) in association with the IEEE Conference on Computer Vision and Pattern Recognition and the Special Session on Heritage Applications of 4th International Conference on 3-D Digital Imaging and Modeling (3DIM'03).

Since a physical object cannot be completely scanned with a single image, multiple scans from different views are required to supply the information needed to construct the 3D model. The most important issues in this process are to minimize the number of views to reduce error accumulation in the 3D model and because data acquisition is expensive [Ikeuchi and Sato (2001)]. Therefore, it is fundamental to adopt a proper and robust technique to align the views in a common coordinate frame, *i.e.* a multiview registration process, to avoid model distortion in subsequent surface reconstruction stage.

There are many methods to perform the registration of views to create 3D models, including calibrated pose measurement and manual registration and verification [Turk and Levoy (1994), Ikeuchi and Sato (2001)]. By using mechanical equipment (*e.g.* robot arm or controlled turntable) to obtain the absolute poses for the scanner views, the calibrated pose methods generally are limited to small objects [Blais and Levine (1995)]. In manual registration and verification, time is the main problem because the user must search for corresponding feature points in the view pair by hand.

Typically, a 3D model is built by the alignment and integration of multiple range views (range images are described in Section 1.1) of an object or scene [Sharp *et al.* (2002), Dorai *et al.* (1998), Blais and Levine (1995), Huber and Hebert (2003), Reed and Allen (1999)]. These two basic steps can be performed sequentially or simultaneously. Usually, sequential methods [Turk and Levoy (1994), Chen and Medioni (1992)] result in imprecise object models since the transformations errors accumulate and propagate from one iteration to another. However, if one can guarantee precise transformations, this method is more attractive and requires less computation resources (*i.e* memory) than others.

The simultaneous method generally is a more robust way to reach precise 3D models. In this category, a global registration between all pairs of views is performed, followed by their integration [Ikeuchi and Sato (2001), Dorai *et al.* (1998), Shum *et al.* (1997), Masuda (2002)]. In this process the accumulation error between the previously registered views is distributed among all alignments, avoiding model distortions while preserving the geometry. Stoddart and Hilton [Stoddart and Hilton (1996)] proposed one of the first global registration methods based on a physical equivalent model. A similar approach was proposed by Eggert *et al.* [Eggert *et al.* (1998)] based on a multi-resolution framework. Bergevin *et al.* [Bergevin *et al.* (1996)] minimized the registration error of all views simultaneously using a well-balanced network of views. Huber and Hebert [Huber and Hebert (2003)] proposed a similar approach to find the minimum spanning tree in a graph, which represents possible model hypotheses for a set of views. Although these methods have been successfully applied in a number of cases, they have some drawbacks such as the computational complexity of the algorithms, and the loss of small details of the object because of imprecise alignments or error accumulation.

In the multiview registration process, the precise alignment of two views is the fundamental stage. Therefore, it is important to adopt a robust registration technique to properly avoid incorrect alignments and further model distortion. It is also important to minimize the number of views to be aligned and, consequently, the registration method must be able to deal with low-overlapped views. To reduce the complexity of the problem, most methods proposed for multiview registration perform an initial registration stage between each pair of overlapped views before the global registration process [Ikeuchi and Sato (2001)].

The registration process for two views consists of finding the best geometric transformation that, when applied to one view, aligns it with the other in a common coordinate system. When registering partially overlapped views, one of the most important issues is to develop methods to deal with low-overlapped views that can guarantee a precise alignment [Rodrigues *et al.* (2002)]. Recently, some papers have directly addressed the problem of calculating the overlapping area between views and measuring the registration quality [Huber and Hebert (2003), Silva *et al.* (2003e), Silva *et al.* (2003d), Dalley and Flynn (2002)]. In this book we show experiments using our robust registration method to align low-overlapped views. Our method is robust because it can deal with non-Gaussian noise, outliers and low-overlapped views.

Since the object's views are effectively aligned one can obtain the entire model by merging the views using a variety of well-known approaches for 3D model representation [Rusinliewicz and Levoy (2001b)]. Usually, a triangular mesh is used and further simplifications are performed to reduce the number of points in the model while preserving the shape of the object [Levoy *et al.* (2000)]. In addition, texture mapping of the reconstructed 3D surfaces is a challenging problem in creating realistic 3D models [Ikeuchi and Sato (2001)].

1.1 Range images

In recent years, range scanners have been improved, allowing an increased number of applications in important areas, such as digital archaeology [Bernardini *et al.* (2002)], building reconstruction and restoration [Reed and Allen (1999)], and medicine [Sinha *et al.* (2003)].

There is a number of range scanner models with distinct methods of acquisition and varying accuracy [Besl (1989)]. Some can acquire different images by combining time-of-flight (range image) and amplitude (reflectance image) of laser beams. These images can be combined to improve image processing stages, such as segmentation [Silva *et al.* (2002)] or edge-based representations [Silva *et al.* (2001)].

Most up-to-date range scanners use a laser beam to precisely measure the distance from the sensor to points in the surface of the object or scene [Besl (1989)], typically in a regular grid pattern as shown in Figure 1.1. This grid can be defined as a *range image* in which each pixel corresponds to a range sample. By defining the resolution of the range image one can obtain the 3D coordinates for each sampled point in the surface of the object. The range image is also known as a $2\frac{1}{2}D$ representation because the 3D information relates only to the visible surface of the object as seen from a given view point.

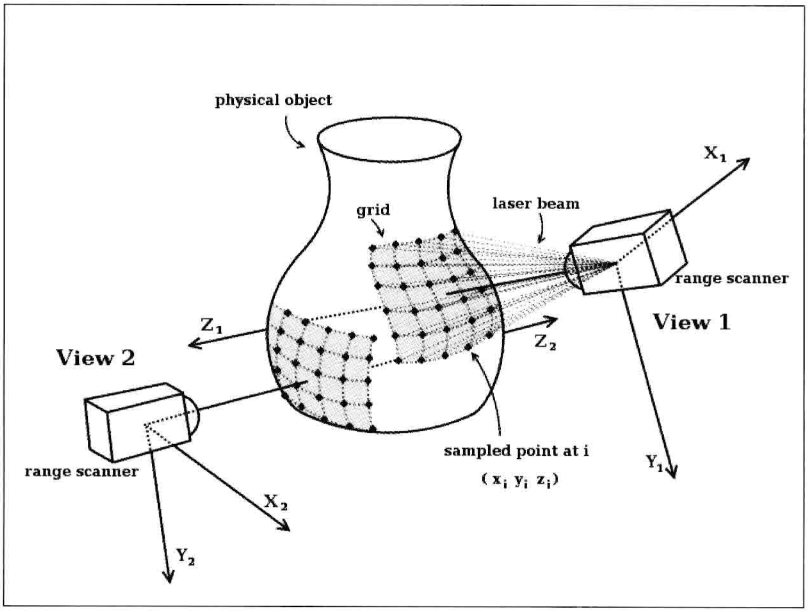

Fig. 1.1 Diagram of the range image acquisition process for two different views of the same object. The range images are composed by the sampled points of regular grids. Each sampled point has 3D information, (x_i, y_i, z_i), where i is a position on the grid.

More formally, a range image can be defined as a set of K discrete samples of a scalar function $j : \mathcal{I}^2 \rightarrow \mathcal{R}$, with $r_i = j(u_i)$, where $u_i \in \mathcal{I}^2$ is the index of the 2D grid (as shown in Figure 1.1), $r_i \in \mathcal{R}$ and $i = \{1, ..., K\}$.

A range image gives the distances between the image plane and points on the objects surfaces in the scene. By consulting a lookup table that indicates the relationship between the image coordinate system and the range scanner coordinate system, a range image can be further converted to range data. These are defined as a set of K discrete samples points of a vector function $h : \mathcal{I}^2 \rightarrow \mathcal{R}^3$, with $d_i = h(u_i)$, where $d_i \in \mathcal{R}^3$ and $i = \{1, ..., K\}$. Then, each sampled point has 3D coordinates (x_i, y_i, z_i).

Since that only part of the object can be seen from any given view point, multiple views are needed to obtain the entire 3D surface of a physical object (see Figure 1.1). Also, it is necessary that these views have some overlap to allow their registration. As can be seen in Figure 1.1 there is no overlap between views 1 and 2. Figure 1.2 shows an example of a range image in which it is possible to note unseen regions on the object surface when it is observed from differents points of view.

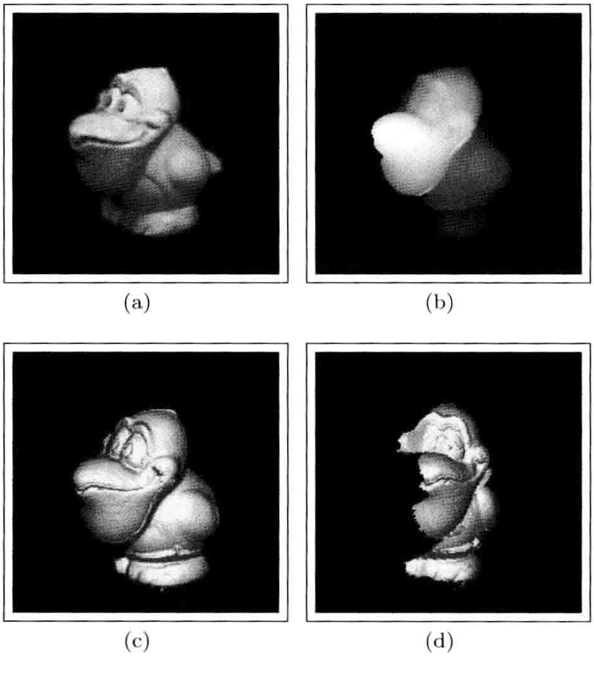

Fig. 1.2 An example of range image acquired with the Minolta Vivid 700 scanner: (a) the picture of the object; (b) the range image, in which the lighter pixels are closer to the sensor than the dark ones; (c) and (d) two rendered views of the range image of (a) observed from different points of view.

In this book we work with several different range image databases, some are available on the Internet and others were kindly supplied by research groups, as listed bellow:

- OSU range image database: Maintained by the Signal Analysis and Machine Perception Laboratory (SAMPL) at The Ohio State University - USA, coordinated by Prof. Kim L. Boyer. We used the range images from different small objects acquired with the Minolta Vivid 700 range scanner. Each object was imaged in 18 views acquired at 20 degree intervals using a computer-controlled turntable (`http://sampl.eng.ohio-state.edu`).
- IMAGO range image database: Maintained by the IMAGO group at the Universidade Federal do Paran - Brazil, coordinated by Prof. Olga R.P. Bellon. The images of small objects with high resolution and precision were acquired using the Roland MDX-15 scanner (`http://www.inf.ufpr.br/imago`).
- The Digital Michelangelo Project: Developed at Stanford University - USA, coordinated by Prof. Marc Levoy. This database has a number of range images from many statues created by Michelangelo in Italy. The images were acquired with high resolution using a special scanner fabricated for them by Cyberware (`http://graphics.stanford.edu/projects/mich`).
- The Great Temple in Petra: Supplied by Prof. Frederic Leymarie of the SHAPE lab at Brown University - USA. The images were acquired with a ShapeGrabber laser scanner in June 2002 at the site of the Great Temple in Petra, Jordan (`http://www.lems.brown.edu/vision/extra/SHAPE`).
- The Cathedral of Saint Pierre: Supplied by Prof. Peter Allen of the Robotics Group at Columbia University - USA. The images were acquired with a Cyrax 2400 scanner in 2001 for a project to model the Cathedral of Saint Pierre, Beauvais, France (`http://www1.cs.columbia.edu/~allen/BEAUVAIS`).
- The Thomas Hunter Building: Supplied by Prof. Ioannis Stamos of Hunter College - USA. The external views of the Thomas Hunter building in New York were acquired with a Cyrax 2400 scanner.
- Stuttgart range image database: This database contains a collection of synthetic range images with different view points taken from high-resolution polygonal models available on the web (`http://range.informatik.uni-stuttgart.de`).

We performed several experiments using these databases to compare our methods for range image registration with others and the results are reported in this book. Additionally, we performed some exhaustive evaluations using the OSU range image database to prove statistically the efficiency of our methods, as presented in the following chapters.

1.2 Applications

There are many applications for which it is desirable to generate 3D models of real objects, including object recognition, robot navigation and reverse engineering [Brown (1992)]. However, most do not demand a precise 3D model to achieve their objective.

In some applications, such as digital archaeology, it is essential to obtain precise 3D models since the objective is to preserve historic sites with as much information and detail as possible [Ikeuchi and Sato (2001)]. In this context, it is important to develop a robust range image registration method that is able to precisely align views of the object, generally by using a set of range images. In this book we address this problem and develop a robust method expected to be useful for many applications, including:

- Archaeology - Creating digital replicas of huge sculptures for heritage preservation. Thus, archaeologists can share their sites for in-depth investigation by other research groups. We can list some important contribution in this area, such as the Digital Michelangelo project [Levoy *et al.* (2000)], the Great Buddha project [Ikeuchi and Sato (2001)] and the Piet project [Bernardini *et al.* (2002)].
- Museology - Generating reproductions of memorable statues and sculptures from different places around the world to make them available in other museums. Furthermore, one can provide virtual visits to virtual museums to visualize the 3D models remotely.
- Architecture - Providing realistic environments from a set of views of buildings for structural analysis. Precise 3D models can aid building construction by providing information regarding different stages for visual inspection or automatic measurements. Also, for building reconstruction or restoration, it is important to plan and to test the proposed project in advance using 3D models to reduce potential the damage to historical sites.
- Military - Providing a robust registration method to deal with clut-

tered environments for many applications, such as target detection and object recognition, in which precision is a fundamental concern.

- Medicine - For critical processes, such as neurosurgery, the 3D models need to be reliable to avoid irreversible damage [Sinha *et al.* (2003)]. Additionally, the practice of dentistry requires accurate 3D representations for the study of the evolution of clinical treatments [Ahmed *et al.* (1997)].

1.3 Book outline

The remainder of this book is organized as follows: Chapter 1.3 presents the background of range image registration with related work and provides the strengths and weaknesses of each approach. Chapter 2.5 introduces a novel evaluation measure, called the Surface Interpenetration Measure (SIM), to assess the quality of registration results and to guide the registration process. We also present a number of experiments to show the stability of this measure. Chapter 3.5 presents the main concepts of Genetic Algorithms as applied to range image registration and shows experiments performed with different strategies. Chapter 4.10 describes a robust range image registration combining Genetic Algorithms (GAs) and the SIM. We also provide an in-depth discussion of the experiments with several comparison results. In Chapter 5.4 we describe experiments in multiview range image registration using our GA-based method with a critical analysis comparing it with other related approaches. Finally, in Chapter 6.5 we conclude by summarizing the contributions of this book and discussing directions for future work.