Bjarne Kjær Ersbøll
Kim Steenstrup Pedersen (Eds.)

LNCS 4522

Image Analysis

15th Scandinavian Conference, SCIA 2007
Aalborg, Denmark, June 2007
Proceedings

SCIA2007

IAPR

Springer

Bjarne Kjær Ersbøll
Kim Steenstrup Pedersen (Eds.)

Image Analysis

15th Scandinavian Conference, SCIA 2007
Aalborg, Denmark, June 10-14, 2007
Proceedings

Springer

Volume Editors

Bjarne Kjær Ersbøll
Technical University of Denmark
Department of Informatics and Mathematical Modelling
Richard Petersen Plads, DTU-Building 321, 2800 Lyngby, Denmark
E-mail: be@imm.dtu.dk

Kim Steenstrup Pedersen
University of Copenhagen, Department of Computer Science
Universitetsparken 1, 2100 Copenhagen, Denmark
E-mail: kimstp@diku.dk

Library of Congress Control Number: 2007928118

CR Subject Classification (1998): I.4, I.5, I.3

LNCS Sublibrary: SL 6 – Image Processing, Computer Vision, Pattern Recognition, and Graphics

ISSN 0302-9743
ISBN-10 3-540-73039-7 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-73039-2 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12075904 06/3180 5 4 3 2 1 0

Preface

The present volume contains the proceedings of the Scandinavian Conference on Image Analysis, SCIA 2007, held at Hotel Hvide Hus, Aalborg, Denmark, June 10–14, 2007.

Initiated in 1979 by Torleiv Orhaug in Sweden, SCIA 2007 represented the 15th in the biennial series of conferences. It is arranged in turn by the Scandinavian countries of Sweden, Finland, Denmark, and Norway, making it a regional conference. However, judging by the nationalities of contributing authors and participants, it is also an international conference.

Worldwide, there is no lack of conferences on image analysis. Indeed, hardly a day passes without an announcement of yet another conference. Therefore the pattern recognition societies of the Scandinavian countries take particular pride in being able to continue the SCIA tradition. SCIA has indeed matured over the many years it has been in existence, but in our opinion SCIA has maintained flexibility and has been able to adopt and incorporate necessary changes and adjustments over that time span. An important key to the success of SCIA is the constant and continuing high quality of the scientific content. Furthermore, the relaxed and friendly atmosphere of the conference itself is well known within the community. The objective to keep in mind must be to continue along those lines.

The number of submissions for this year's event was an impressive 228. Of these, 99 can be found in the present proceedings, leading to an acceptance rate of 43%. In order to optimize the outcome for the participants, the conference was organized as a single track event. Thirty-three papers were presented in the oral sessions and 66 were presented in the poster sessions. Each paper was reviewed by at least two peers. Acceptance was based solely on these reviews. The papers can roughly be grouped into the following topics: computer vision; 2D and 3D reconstruction; classification and segmentation; medical and biological applications; appearance and shape modeling; face detection, tracking and recognition; motion analysis; feature extraction and object recognition. Two tutorials preceded the conference.

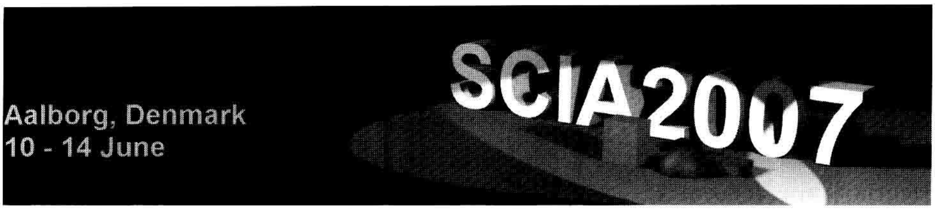
A conference is the result of careful planning and lots of work from numerous volunteers. It is important that we acknowledge these important contributions. We thank the invited speakers for enlightening us in their areas of research and the contributing scientists for their presentations. Furthermore, we thank the reviewers for the pleasant interaction during the review process and for the excellent work in helping to maintain the high quality of SCIA.

It is our sincere hope that the participants had an enjoyable and fruitful experience, both scientifically and socially, in Aalborg.

June 2007

Bjarne Kjær Ersbøll
Kim Steenstrup Pedersen

Organization



SCIA 2007 was organized jointly by Aalborg University, the Technical University of Denmark (DTU), the Department of Computer Science at University of Copenhagen, and the IT University of Copenhagen.

AALBORG UNIVERSITY

DTU

KØBENHAVNS UNIVERSITET

IT University of Copenhagen

Executive Committee

General Chair	Peter Johansen, University of Copenhagen, Denmark
Program Chairs	Bjarne Kjær Ersbøll, Technical University of Denmark
	Kim Steenstrup Pedersen, University of Copenhagen, Denmark
	Søren Ingvor Olsen, University of Copenhagen, Denmark
Local Chair	Claus Brøndgaard Madsen, Aalborg University, Denmark

Program Committee

Bjarne Kjær Ersbøll (Chairman)	Technical University of Denmark
Ingela Nyström	Uppsala University, Sweden
Ivar Austvoll	University of Stavanger, Norway
Janne Heikkilä	University of Oulu, Finland

Invited Speakers

Paul Debevec	University of Southern California, USA
Henrik Wann Jensen	University of California at San Diego, USA and Technical University of Denmark
Jitendra Malik	University of California at Berkeley, USA
Henrik I. Christensen	Georgia Institute of Technology, USA and Royal Institute of Technology, Sweden
Aapo Hyvärinen	University of Helsinki, Finland

Tutorials

Henrik Wann Jensen	University of California at San Diego, USA and Technical University of Denmark
Jitendra Malik	University of California at Berkeley, USA

Reviewers

Lukas Ahrenberg
Fritz Albregtsen
Hans Jørgen Andersen
Mats Andersson
Ulf Assarsson
Ivar Austvoll
Adrien Bartoli
Ewert Bengtsson
Magnus Borga
Gunilla Borgefors
Marleen de Bruijne
Anders Brun
Jakob Andreas Bærentzen
Stefan Carlsson
Jens Michael Carstensen
Niels Jørgen Christensen
Anne Cuzol
Anders Bjorholm Dahl
Ola Dahl
Sune Darkner
Jan-Olof Eklundh
Søren G. Erbou
Kenny Erleben
Bjarne Kjær Ersbøll
Søren Forchhammer
Robert Forchheimer
Vincent Gay-Bellile
Magnus Gedda
Mads Fogtmann Hansen
Lars Kai Hansen
Michael Sass Hansen
Markku Hauta-Kasari
Janne Heikkilä
Knud Henriksen
Anders Heyden
Jouni Hiltunen
Bjarke Jakobsen
Peter Johansen

Fredrik Kahl
Juho Kannala
Sune Høgild Keller
Hans Knutsson
Anders Wang Kristensen
Volker Krüger
Norbert Krüger
Iivari Kunttu
Heikki Kälviäinen
Bent Larsen
Rasmus Larsen
Reiner Lenz
Leena Lepistö
Joakim Lindblad
Marco Loog
Claus Brøndgaard Madsen
J. Birgitta Martinkauppi
Brian Mayoh
Thomas B. Moeslund
Allan Aasbjerg Nielsen
Henning Nielsen
Mads Nielsen
Michael Nielsen
Ingela Nyström
Erkki Oja
Matti Okkonen
Hildur Ólafsdóttir
Søren Ingvor Olsen
Niels Chr. Overgaard
Jussi Parkkinen
Kim Steenstrup Pedersen
Mathieu Perriollat
Johanna Pettersson
Daniel Pizarro
Esa Rahtu
Joakim Rydell
Pekka Sangi
Sigurjon A. Sigurdsson
Karl Sjöstrand
Arnold Skimminge
Örjan Smedby
Jon Sporring
Rasmus Stenholt
Björn Svensson
Stina Svensson
Jean-Philippe Tardif
Martin Vester-Christensen
Erik Vidholm
Ole Winther
Henrik Aanæs
Kalle Åström

Sponsoring Organizations

The International Association for Pattern Recognition

Danish Pattern Recognition Society (DSAGM).

Table of Contents

Computer Vision

2D and 3D Reconstruction

Motion Analysis

Feature Extraction and Object Recognition

Classification and Segmentation

Poster Session 1

Poster Session 2

Accurate Interpolation in Appearance-Based Pose Estimation

Erik Jonsson and Michael Felsberg*

Computer Vision Laboratory
Dept. of Electrical Engineering, Linköping University
erijo@isy.liu.se, mfe@isy.liu.se

Abstract. One problem in appearance-based pose estimation is the need for many training examples, i.e. images of the object in a large number of known poses. Some invariance can be obtained by considering translations, rotations and scale changes in the image plane, but the remaining degrees of freedom are often handled simply by sampling the pose space densely enough. This work presents a method for accurate interpolation between training views using local linear models. As a view representation local soft orientation histograms are used. The derivative of this representation with respect to the image plane transformations is computed, and a Gauss-Newton optimization is used to optimize all pose parameters simultaneously, resulting in an accurate estimate.

1 Introduction

Object recognition and pose estimation can be done in several ways. In the bag-of-features approach, local coordinate frames are constructed around points of interest [5], [9], and features from each local frame vote for a certain object and pose hypothesis. In the model-based approach [2], [11], a geometrical model is fitted to the observed image. This approach is often very accurate, but requires a good initial guess and a manually constructed 3D model. Global appearance-based methods extract features from the appearance of the entire object and match these to training views in memory. Ever since [10], [7], the most common approach seems to be using PCA.

In this paper, we use an appearance-based method using full object views, but avoid PCA due to the global nature of this representation. The main goal is to maximize the accuracy of the pose estimate by interpolating between a limited number of training views. The interpolation method is based on representing the views with *channel-coded orientation* [3], [4], and optimizing all pose parameters (including position, rotation and scale in the image plane) simultaneously using a Gauss-Newton method. The method requires an initial guess, which in a real system could be obtained using your favorite fast but inaccurate bag-of-features approach.

* This work has been supported by EC Grant IST-2003-004176 COSPAL. This paper does not represent the opinion of the European Community, and the European Community is not responsible for any use which may be made of its contents.

B.K. Ersbøll and K.S. Pedersen (Eds.): SCIA 2007, LNCS 4522, pp. 1–10, 2007.

The motivation for using full object views is two-fold. The first reason is that once we have formed an initial object hypothesis, it makes sense to use as much image information as possible in order to get an accurate estimate. The second reason is that using full views, we can focus on the interpolation and view representation, and ignore other aspects like how to choose interest points and construct local frames in a bag-of-features approach. This makes it easier to compare different view representations. Similar interpolation techniques as proposed here should however be possible to integrate also in a bag-of-features framework.

In contrast to model-based methods, our approach requires no knowledge of 3D geometry in the system, and is in no way specific to 3D pose estimation. The training set could consist of any parameterized image set, e.g. a robotic arm in different configurations etc.

2 Algorithm

2.1 Pose Estimation

The appearance of an object is determined by the object state $\mathbf{p} = [\theta, \phi, s, \alpha, x, y]$. The parameters s, α, x, y represent the scale, rotation and position of the object in the image plane and will be referred to as the *image parameters* $\mathbf{p}_{\text{img}}$. The two auxiliary angles θ and ϕ cover all pose variations not explained by rotation in the image plane and will be referred to as the *pose angles* $\mathbf{p}_{\text{pose}}$.

During training, we learn the appearance of the object given (θ, ϕ) using canonical image parameters. The result of the learning can be seen as a function $\mathbf{f}$ that maps the pose angles to a predicted feature vector:

$$\hat{\mathbf{c}} = \mathbf{f}(\theta, \phi) \ . \tag{1}$$

During operation of the system, we maintain a current hypothesis of the object state, and cut out an image patch around the current (x, y) with rotation α and size s. This can be formalized by a function

$$\mathbf{c} = \mathbf{g}(s, \alpha, x, y) \tag{2}$$

producing an observed feature vector from the image given certain image parameters. The pose estimation problem is now to find an object state $\mathbf{p}_*$ which minimizes the difference between the observed and predicted feature vectors:

$$\mathbf{p}_* = \arg\min_{\mathbf{p}} \|\mathbf{r}(\mathbf{p})\|^2 \tag{3}$$

where

$$\mathbf{r}(\mathbf{p}) = \mathbf{f}(\theta, \phi) - \mathbf{g}(s, \alpha, x, y) \ . \tag{4}$$

This can be solved using your favorite optimization method. We use a Gauss-Newton method, with a simple backtracking line search [8]. The update step direction $\mathbf{s}$ is given by

$$\mathbf{Js} = -\mathbf{r} \ , \tag{5}$$