Marginal Space Learning for Medical Image Analysis

Yefeng Zheng • Dorin Comaniciu

Marginal Space Learning for Medical Image Analysis

Efficient Detection and Segmentation of Anatomical Structures

Springer

Yefeng Zheng
Imaging and Computer Vision
Siemens Corporate Technology
Princeton, NJ, USA

Dorin Comaniciu
Imaging and Computer Vision
Siemens Corporate Technology
Princeton, NJ, USA

ISBN 978-1-4939-0599-7 ISBN 978-1-4939-0600-0 (eBook)
DOI 10.1007/978-1-4939-0600-0
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2014934133

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Muna, Allen, and Amy

– Y. Z.

To my family

– D. C.

Preface

Medical imaging is today an integrated part of the healthcare continuum, supporting early disease detection, diagnosis, therapy, monitoring, and follow-up. Images of the human body help in estimating the organ anatomy and function, reveal clues indicating the presence of disease, or help in guiding treatment and interventions. All these benefits are achieved by extracting and quantifying the medical image content, answering questions such as: "Which part of the 3D image represents the heart and what is the ejection fraction?", "What is the volume of the liver", "Which are the axillary lymph nodes with a diameter larger than 10 mm?", "Is the artificial heart valve being positioned at the right location, with the right angulation?"

With the continuous increase in the spatial and temporal resolution, the informational content of images increases, contributing to new clinical benefits. While most of the content extraction, quantification, and decision making are guided and validated by the clinicians, computer-based image systems benefit from efficient algorithms and exponential increase in computational power. Thus, they play an important role in analyzing the image data, performing tasks such as identifying the anatomy or measuring a certain body function.

Systems based on machine learning have recently opened new ways to extract and interpret the informational content of medical images. Such systems learn from data through a process called training, thus developing the capability to identify, classify, and label the image content.

Learning systems have been initially applied to nonmedical images for two-dimensional (2D) object detection problems such as face detection, pedestrian or vehicle detection in 2D images, and video sequences. In these methods, object detection or localization is formulated as a classification problem: whether an image block contains the target object or not. The robustness of the methods comes from the exhaustive search with the trained classifier during object detection on an input image. The object pose parameter space is first quantized into a set of discrete hypotheses covering the entire space. Each hypothesis is tested by a trained classifier to get a detection score and the hypotheses with the highest score are taken as the detection output. In a typical setting, only three pose parameters are estimated, the position (X and Y) and isotropic scale (S), resulting in a three-dimensional search space and a search problem of relatively low complexity.

On the other hand, most of the medical imaging data used in clinical practice are volumetric and three-dimensional (3D). Computed tomography, C-Arm X-Ray, magnetic resonance, ultrasound, and nuclear imaging create 3D representations of the human body. To accurately localize a 3D object, one needs to estimate nine pose parameters: three for position, three for orientation, and three for anisotropic scaling. However, a straightforward extension of a 2D object detection method to 3D is not practically possible due to the exponential increase in the computation needs attributed to exhaustive search. How do we solve this problem? What kind of learning strategy would help to perform efficient search in a nine-dimensional pose parameter space?

This book presents a generic learning-based method for efficient 3D object detection called Marginal Space Learning (MSL). Instead of exhaustively searching the original nine-dimensional pose parameter space, only low-dimensional marginal spaces are searched in MSL to improve the detection speed.

We split the estimation into three steps: position estimation, position-orientation estimation, and position-orientation-scale estimation. First, we train a position estimator that can tell us if a position hypothesis is a good estimate of the target object position in an input volume. After exhaustively searching for the position marginal space (three-dimensional), we preserve a small number of position candidates with the largest detection scores. Next, we perform joint position-orientation estimation with a trained classifier that answers if a position-orientation hypothesis is a good estimate. The orientation marginal space is exhaustively searched for each position candidate preserved after position estimation. Similarly, we only preserve a limited number of position-orientation candidates after this step. Finally, the scale parameters are searched in the constrained space in a similar way.

Since after each step we only preserve a small number of candidates, a large portion of search space with low posterior probability is pruned efficiently in the early steps. Complexity analysis shows that MSL can reduce the number of testing hypotheses by six orders of magnitude, compared to the exhaustive full space search. Since the learning and detection are performed in a sequence of marginal spaces, we call the method Marginal Space Learning (MSL).

As it will be shown in this book, the MSL has been applied to detect multiple 2D/3D anatomical structures in the major medical imaging modalities. Several key techniques have later been proposed to further improve its detection speed and accuracy: Constrained MSL to exploit the strong correlation existing among pose parameters in the same marginal spaces; Iterated MSL to detect multiple instances of the same object type in a volume; Hierarchical MSL to improve the robustness by performing learning/detection on a volume pyramid; Joint spatio-temporal MSL to detect the trajectory of a landmark in a volume sequence.

With these improvements, we can reliably detect a 3D anatomical structure with a speed of 0.1–0.5 s/volume on an ordinary personal computer (3.2 GHz duo-core processor and 3 GB memory) without the use of special hardware such as graphics processing units.

The MSL can also be applied to generate accurate shape initialization for the segmentation of a nonrigid anatomical structure. To further improve the initialization accuracy, the MSL has been extended to directly estimate the nonrigid deformation parameters in combination with a learning-based boundary detector that guides the boundary evolution.

Several practical anatomy segmentation systems have been built and evaluated at multiple clinical sites. Examples include four-chamber heart segmentation, liver segmentation, and aorta segmentation. At the time of publication they all outperformed the state of the art in both speed and accuracy.

This book is for students, engineers, and researchers with interest in medical image analysis. It can also be used as a reference or supplementary material for related graduate courses. Preliminary knowledge of machine learning and medical imaging is needed to understand the content of the book.

Princeton, NJ, USA

Yefeng Zheng
Dorin Comaniciu

Acknowledgements

We owe our gratitude to all colleagues who have made this work possible.

At the beginning, this was a joint project with a former colleague at Siemens Corporate Research, Adrian Barbu. During the summer of 2006 we were looking for an efficient and robust 3D detection method to automatically identify the heart chambers in cardiac computed tomography volumes. Based on our previous experience, we knew machine learning was the right direction; however, we lacked an efficient search method. After a series of systematic investigations, we jointly proposed a simple and elegant method called marginal space learning (MSL) as presented here. We really appreciated Adrian's enthusiasm during this development.

We would also like to thank many other colleagues at Siemens Corporate Research for applying, tuning, and improving the MSL on various 2D/3D object detection and segmentation problems in medical imaging. Bogdan Georgescu and Kevin Zhou contributed to the early development stages of this technology; Zhuowen Tu provided the implementation of the probabilistic boosting-tree and 3D Haar wavelet features; Haibin Ling helped in performing comparison experiments for Constrained MSL and Nonrigid MSL on liver detection; and Xiaoguang Lu and Gustavo Carneiro contributed to the comparison experiments of full space learning and MSL on 2D left ventricle detection in magnetic resonance images. Furthermore, several important improvements of MSL originated from our colleagues' work, including Hierarchical MSL from Michal Sofka and Jingdan Zhang, Iterated MSL from Michael Kelm, and Joint Spatio-Temporal MSL from Razvan Ionasec and Yang Wang.

We thank our colleagues from Siemens Healthcare, especially Michael Scheuering, Fernando Vega-Higuera, Dominik Bernhardt, and Michael Suehling from Computed Tomography; Matthias John, Jan Boese, and Martin Ostermeier from Angiographic and Interventional X-Ray Systems; Arne Littmann, Edgar Mueller, and Berthold Kiefer from Magnetic Resonance; and Helene Houle and Sara Good from Ultrasound. They provided us challenging medical imaging problems so that we had a chance to develop the MSL. They collected data all over the world for us and coordinated the clinical evaluation of the resulting systems. Most importantly, they asked us to continuously improve the detection/segmentation speed and robustness of the algorithms.

We are thankful to the Springer editors, Courtney Clark and Jennifer Evans, for their enthusiasm and support to the project.

Yefeng owes his deepest thanks to his wife, Muna Tang, for allowing him to work overtime, often during the weekends.

Dorin would like to thank his family for being so inspiring and supportive.

Contents

Acronyms

A2C	Apical two chamber view
A3C	Apical three chamber view
A4C	Apical four chamber view
AAM	Active appearance model
AF	Atrial fibrillation
ASM	Active shape model
AV	Aortic valve
CPU	Central processing unit
CT	Computed tomography
CTA	Computed tomography angiography
ED	End-diastole
EF	Ejection fraction
ES	End-systole
FSL	Full space learning
GPU	Graphics processing unit
LA	Left atrium
LAA	Left atrial appendage
LIPV	Left inferior pulmonary vein
LSPV	Left superior pulmonary vein
LV	Left ventricle
LVOT	Left ventricular outflow tract
MRF	Markov random field
MPR	Multi-planar reformatting or multi-planar reconstruction
MRI	Magnetic resonance imaging
MSL	Marginal space learning
MV	Mitral valve
PBT	Probabilistic boosting-tree
PCA	Principal component analysis
PDM	Point distribution model
PV	Pulmonary vein or pulmonary valve
RA	Right atrium

RIPV	Right inferior pulmonary vein
ROI	Region of interest
RSPV	Right superior pulmonary vein
RV	Right ventricle
RVOT	Right ventricular outflow tract
SAX	Short axis view
SNR	Signal-to-noise ratio
SVD	Singular value decomposition
SVM	Support vector machine
TAVI	Transcatheter aortic valve implantation
TAVR	Transcatheter aortic valve replacement
TEE	Transesophageal echocardiography
TPS	Thin-plate spline
TV	Tricuspid valve