(On the Spine)

SELF-LEARNING SHAPE RECOGNITION IN MEDICAL IMAGES

MA JINGTING

2019

NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

SELF-LEARNING SHAPE RECOGNITION IN MEDICAL IMAGES

MA JINGTING

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

2019

SELF-LEARNING SHAPE RECOGNITION IN MEDICAL IMAGES

MA JINGTING

MA JINGTING

School of Computer Science and Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Master of Doctor of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

Input Date Here	Input Signature Here
03.10.2019	
. . . . .	. . . . .
Date	Input Name Here
03.10.2019	Ma Jingting

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

Input Date Here	Input Supervisor Signature Here
03.10.2019	
. . . . .	. . . . .
Date	Input Supervisor Name Here
03.10.2019	Lin Feng

Authorship Attribution Statement

Please select one of the following; *delete as appropriate:

This thesis contains material from [3] paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Chapter 3 is published as Jingting Ma, Anqi Wang, Feng Lin, Stefan Wesarg, and Marius Erdt. “Nonlinear Statistical Shape Modeling for Ankle Bone Segmentation Using a Novel Kernelized Robust PCA.” In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI’17), Quebec City, Canada, 10-14 September, 2017, pp. 136-143.

The contributions of the co-authors are as follows:

- I proposed the method, did the experiments and wrote the drafts of the manuscript.
- Mrs. Wang assisted to conduct the experiment of MRI ankle bone segmentation.
- Prof. Lin revised the manuscript together with Dr. Wesarg.
- Prof. Erdt proposed the topic.

Partial contents in Chapter 4 is published as Jingting Ma, Feng Lin, Jonas Honsdorf, Katharina Lentzen, Stefan Wesarg, and Marius Erdt. “Weighted Robust PCA for Statistical Shape Modeling.” In Inter- national Conference on Medical Imaging and Virtual Reality (MIAR’16), Bern, Switzerland, August 24-26, 2016, pp. 343-353.

The contributions of the co-authors are as follows:

- I proposed the method, did the experiments and wrote the drafts of the manuscript.
- Prof. Lin revised the manuscript together with Dr. Wesarg.
- Mr. Honsdorf and Mrs. Lentzen implemented the Robust PCA and did partial experiments together.
- Prof. Erdt proposed the topic and revised the manuscript.

Chapter 5 is published as Jingting Ma, Feng Lin, Stefan Wesarg, and Marius Erdt. “A Novel Bayesian Model Incorporating Deep Neural Network and Statistical Shape Model for Pancreas Segmentation.” In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI’18), Granada, Spain, 16-20 September, 2018.

The contributions of the co-authors are as follows:

- I did all the experiments and wrote the drafts of the manuscript.
- Prof. Lin revised the manuscript together with Dr. Wesarg.
- Prof. Erdt suggested the topic and revised the manuscript.

Input Date Here	Input Signature Here
03.10.2019	
. . . . .	. . . . .
Date	Input Name Here
03.10.2019	Ma Jingting

Abstract

A massive amount of medical image data, e.g. from Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), is generated from hospitals every day. Biological structure segmentation is very useful to support surgery planning and treatments, as an ideal delineation of the outline of the target object can offer a precise location and quantitative analysis for further clinical diagnoses such as identification of tumorous tissues. However, the large dimension and complex patterns in medical image data make manual annotation extremely time-consuming and problematic. Accordingly, automatic biomedical image segmentation becomes a crucial pre-requisite in practice and has been a critical research issue over tens of years.

However, major challenges exist in medical image segmentation such as the low intensity contrast to surrounding tissues and complex geometry of shape. Moreover, limited amounts of labeled training data give rise to difficulties as well. Numerous approaches have been proposed to mitigate these challenges, from low-level image processing to supervised machine learning techniques. It is worth mentioning that statistical shape models (SSMs) based segmentation approaches have achieved remarkable success in a widespread of applications. SSMs are trained mostly using self-learning approaches to parameterize the significant variabilities of biological shapes, subsequently, the learned shape prior is adopted in image adaption to guide the shape fitting. Despite the success, SSMs-based segmentation approaches suffer from the limitation that the power of SSMs rises and falls with the quality of training data and geometrical complexity of the target shape. Furthermore, the existing image adaption may not be efficient in cases where the target object has a small and distorted structure. Therefore, this thesis aims to derive SSMs that are robust to training data corruption and are able to represent complex

patterns, and address the problem of the poor image adaption to realize the challenging object segmentation.

As training data is often corrupted by many factors like inherent noise/artifacts and non-ideal delineations in this thesis, many efforts have been devoted to developing SSMs that are robust to data corruption. First, early attempts proposing an imputation method and weighted Robust Principal Component Analysis (WRPCA) have been made to address arbitrary corruptions under the assumption of linear distribution. Nevertheless, deriving a quality model is still demanding as the shape variance of biological structures may not simply follow Gaussian distribution. To combat this, a kernelized RPCA is proposed to cope with outliers in a nonlinear distribution. The idea is performing the low-rank modeling on the kernel matrix to achieve nonlinear dimensionality reduction, and outlier recovery thereof.

To increase the generality and feasibility, this thesis, furthermore, presents a general nonlinear data compression technique, the Robust Kernel PCA (RKPCA), with the aim of constructing a low-rank nonlinear subspace free of outliers. In terms of evaluation, the proposed RKPCA delivers high performance on not only creating SSMs but also on outlier recovery. Experiments are conducted using two representative datasets, a set of 30 public CT kidneys and a set of 49 internal MRI ankle bones. Embedded into an existing segmentation framework, experimental results show that SSM built with the proposed RKPCA outperforms the state-of-the-art modeling techniques in terms of model quality and segmentation accuracy.

Since SSMs fail to adopt in cases where the target structure occupies a relatively small or distorted area, deep neural networks that remedy this shortcoming are considered thereof. However, redundant background contents in 3D volume may significantly influence the accuracy of deep deep neural networks. Aiming at challenging structures that occupy relatively small areas and have large variances, a novel unified segmentation framework is proposed that incorporates SSM on the top of deep neural network for detailed refinement. The motivation is aggregating both spatial and intensity based features from a limited amount of data. Globally optimized via Bayesian inference, the segmentation is driven by a dynamic weighted Gaussian Mixture Model integrating the probability

scores from the deep neural network and the shape prior from the SSM. Under a public NIH dataset of CT pancreas, the proposed segmentation framework achieves the best average Dice Similarity Coefficient compared to the-state-of-the-art approaches.

The majority of this work is based on public tools: the **Medical Imaging Interaction Toolkit (MITK)** for SSMs investigation and analysis and the public library **Keras** for deep neural networks development. All medical image datasets used in this thesis have been validated by clinical experts.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisors, Prof. Lin Feng and Prof. Marius Erdt, for their persistent guidance, patient and meticulous suggestions over the past few years. I appreciate all their contributions of time, ideas and funding to make my Ph.D. experience productive and stimulating. Also, I gratefully acknowledge the precious advice and encouragement of Prof. Dieter W. Fellner and Dr. Stefan Wesarg, who gave me the freedom I needed to go through difficulty.

I would like to thank my thesis advisory committee members, Prof. Wolfgang Meuller and Prof. Alexei Sourin, for their brilliant comments. Many thanks to my friends and colleagues working in Fraunhofer Singapore and Fraunhofer IGD, for their great help and accompany.

Last but not least, I would like to give a special gratitude to the joint Ph.D program and all relevant institutions. It is great honor for me to have the opportunity to study in Singapore and Germany. This valuable experience will be a treasure to my whole life.

Contents

Abstract . . . ii
Acknowledgments . . . v
List of Figures . . . x
List of Tables . . . xiv

1 Introduction **1**
1.1 Background . . . 1
1.2 Motivation and Hypothesis . . . 3
1.2.1 Data Contamination . . . 3
1.2.2 Complex Geometry . . . 5
1.2.3 Poor Image Adaption . . . 6
1.3 Summary of Contributions . . . 7
1.4 Thesis Organization . . . 9

2 Literature Review **11**
2.1 Medical Image Segmentation . . . 11
2.1.1 Overview of Application Areas . . . 11
2.1.1.1 Anatomical organs and substructures segmentation . . . 12
2.1.1.2 Pathology segmentation . . . 15
2.1.2 Overview of Major Approaches . . . 16
2.1.3 Discussion . . . 20
2.2 Statistical Shape Models . . . 21
2.2.1 Baseline of Statistical Shape Modeling . . . 21
2.2.2 SSM-based Segmentation Applications . . . 22
2.2.3 Nonlinear SSMs . . . 23

2.2.4 Discussion 24

3 Nonlinear Statistical Shape Modeling Using A Novel Kernelized Robust Principal Component Analysis 27

3.1 Introduction 29
3.1.1 Limitations of Robust Linear SSM 31
3.1.2 Weighted Robust PCA 33
3.1.3 KPCA for Statistical Shape Modeling 36
3.2 KRPCA for Statistical Shape Modeling 37
3.2.1 Kernelized RPCA 37
3.2.1.1 Solving K_X 39
3.2.1.2 Solving K_E 40
3.2.2 Applying KRPCA to statistical shape modeling 40
3.3 Evaluation 41
3.3.1 Dataset 42
3.3.2 Parameters 43
3.3.3 Ability of Outlier Recovery 44
3.3.4 Model Quality Evaluation 46
3.3.5 Application in Ankle Bone Segmentation 47
3.4 Discussion 48

4 A Novel Robust Kernel Principal Component Analysis for Nonlinear Statistical Shape Modeling from Erroneous Data 51

4.1 Introduction 52
4.2 Robust Kernel PCA 53
4.2.1 Objective formulation 55
4.2.2 Updating $E^{(t+1)}$ 55
4.2.3 Updating $K^{(t+1)}$ 56
4.2.4 Updating $X^{(t+1)}$ 57
4.2.5 Convergence Analysis 58
4.3 Evaluation 59
4.3.1 Datasets 60

4.3.2 Parameters . . . 61
4.3.3 Computational Complexity Analysis . . . 61
4.3.4 RKPCA for Outlier Recovery . . . 62
4.3.5 RKPCA for Missing Area Completion . . . 66
4.3.6 Model Evaluation . . . 68
4.3.7 Segmentation Evaluation . . . 73
4.4 Discussion . . . 75

5 A Novel Bayesian Model Incorporating Deep Neural Network and Statistical Shape Model for Pancreas Segmentation 78
5.1 Introduction . . . 79
5.1.1 Overview of Pancreas Segmentation . . . 79
5.1.2 Hypothesis . . . 81
5.2 DenseUNet for Model Initialization . . . 82
5.2.1 DenseUNet Architecture . . . 83
5.2.2 Model Initialization . . . 85
5.3 Bayesian Model . . . 86
5.3.1 Dynamic Weighted Gaussian Mixture Model . . . 88
5.3.2 Shape Prior . . . 89
5.4 Evaluation . . . 90
5.4.1 Dataset . . . 90
5.4.2 Parameter . . . 92
5.4.3 Segmentation Results . . . 92
5.4.4 Comparison to State-of-the-art . . . 93
5.5 Discussion . . . 94

6 Conclusion and Future Work 99
6.1 Conclusion . . . 99
6.2 Future Work . . . 102
6.2.1 Statistical Shape Modeling Without Correspondence . . . 102
6.2.2 Embedding RKPCA into Deep Neural Network . . . 103

References 105

List of Publications **122**

List of Figures

1.1 This figure illustrates the challenges in medical image segmentation, e.g. low-contrast of intensity and large variabilities of shape. 2

1.2 This figure illustrates examples of data corruption: (a) is a normal case of ankle bone shape; (b) shows a corrupted shape data with abnormal subpart (marked by a black square); (c) is an incomplete ankle shape with the missing areas (marked by a black square). 4

1.3 This figure illustrates a multi-object model of kidney pair, where (a) indicates the mean shape with normal variance, (b) and (c) show the exceptions with abnormal poses. 6

1.4 This figure illustrates an example of small organ (marked in red), i.e. pancreas, from the Axial, Sagittal and Coronal view respectively. 7

3.1 (a) Ankle bone structure (b) Corrupted ankle bone with anomalous overlap marked by a black square (c) Incomplete ankle bone with missing areas marked with black squares 28

3.2 Difference between reconstructions using WRPCA and RPCA. Note that the shading indicates the areas of mesh that are approximately ground truth and the red bits are the reconstruction of missing parts. 34

3.3 The figure illustrates the flow chart of modeling and back projection. (a) indicates the training mesh without landmark correspondence assigned; (b) indicates that the shape has been assigned with a groupwise correspondnece, where the color indiates the landmark ID; (c) indicates an SSM with the mean shape in the middle. The procedure from (b) to (c) means the model back-projection, and (d) is the reconstruction of the shape (b). It can be found that the abnormal overlap in (b) disappears in (d) with the correction by the model (c). 45
3.4 Generalization ability and Specificity for the models built with PCA, KPCA, KRLRR and KRPCA. Smaller value indicates better result. . . . 47
3.5 Figure shows the comparison of segmentation results of competing models in Sagittal, Coronal and Axial position respectively. 48

4.1 This figure illustrates the pipeline of our RKPCA and the procedure of model back projection. Based on a set of corrupted shapes, the training matrix D can be constructed. : when an unseen shape $\boldsymbol{Z}$ is back projected onto the model, a set of nonlinear principal components are propagated and further used to approximate the reconstruction $\hat{z}$ through the standard pre-image $f(z)$. 54
4.2 The figure plots a corrupted shape (a) with its reconstructions from competitive models (b) - (g) illustrated from the perspective of Coronal-Axial view (left) and 3D rendering (right), where the red shape indicates the ground truth. (a) is artificially generated with 20% of landmarks randomly removed. 64
4.3 The figure plots a corrupted shape (a) with its reconstructions from competitive models (b) - (g) illustrated from the perspective of Coronal-Axial view (left) and 3D rendering (right), where the red shape indicates the ground truth. (a) is artificially generated with 40% of landmarks randomly removed. 65
4.4 This figure plots the procedure of incomplete shapes generation: based on a ground truth shape (a), several pieces with varying areas are randomly selected (marked in red in (b)) and removed. An incomplete shape (c) with missing areas is generated thereof. 67

4.5 The figure plots the ground truth ankle bone shape (a) and the artificially generated incomplete shapes with 10%, 20% and 30% removing points in (b), (c) and (d) respectively. 67
4.6 The figure plots an incomplete ankle bone shape (a) with its reconstructions from competitive models (b) - (g) illustrated from the perspective of Coronal-Axial view (left) and 3D rendering (right), where the red shape indicates the ground truth. (a) is artificially generated with 30% of pieces randomly removed. 69
4.7 The variance for the RPCA and RKPCA model, where the shape in the middle is the mean and the left and right shapes are generated from $-3\sqrt{\Lambda_1}$ and $+3\sqrt{\Lambda_1}$ for RPCA model, and $-\sqrt{N_s\Lambda_1}$ and $+\sqrt{N_s\Lambda_1}$ for RKPCA model respectively, with other modes set to the 0. 70
4.8 The variance for the KPCA and RKPCA model, where the shape in the middle is the mean and the left and right shapes are generated from $-\sqrt{N_s\Lambda_1}$ and $+\sqrt{N_s\Lambda_1}$ for the first mode, with other modes set to 0. . . 70
4.9 Generalization ability G, specificity S and compactness C for the kidney models. 71
4.10 Generalization ability G, specificity S and compactness C for the ankle bone models. 72
4.11 Comparisons of kidney pair segmentation results. 75
4.12 Comparisons of ankle bone segmentation results. 76

5.1 This figure illustrates the whole segmentation pipeline: (a) implies the test image to be segmented, through the pre-trained deep neural network (b), a probability score (c) is obtained; (d) indicates that the shape model is initialized with the probability score; and (e) implies the proposed Bayesian model, where the color indicates the probability for each landmark; (f) is the final segmentation output. 82
5.2 Architecture of UNet proposed in [1]. 83
5.3 Architecture of our proposed DenseUNet. A convolutional with large kernel is adopted at the very beginning, followed by three 5-layer dense block. Each dense block is followed by a transition layer to achieve downsampling. 84

5.4 This figure illustrates the procedure of model initialization: (a) is the probability score map from deep neural network, where the color indicates the probability value; (b) is the ICP alignment of shape model and the probability score; (c) is the model initialization. 86
5.5 This figure illustrates the pipeline of segmentation approach: given the test image with probability map (a), the shape model is initialized to fit the detected region (b); considering the neighborhood region around each landmark (c), a Gaussian Mixture Model is trained (d) to guide shape adaption (e); afterwards, project the shape onto statistical shape model (f); we obtain the segmentation output (g) when the convergence is reached. 87
5.6 Figure shows the segmentation procedure of NIH Case #1 in the first row, Case #2 in the second row and Case #3 in the third row. Note that (a) is the primary segmentation result from deep neural network, (e) indicates the final result and the ground truth is marked in red. Specifically, the average DSC of Case #1 is raised from 68.90% in (a) to 82.17% in (e), the average DSC of Case #2 is raised from 71.52% to 84.83% and that of Case #3 is raised from 69.51% to 78.07%. 94

List of Tables

2.1 Major Application Areas 12
2.2 Approaches in medical image segmentation 16
2.3 Major challenges in statistical shape modeling 25

3.1 Overview of Matrix Factorization 30
3.2 Parameters and Descriptions 44
3.3 Back projection error MAD 45
3.4 Segmentation results for reference PCA model and competing models. Note that smaller Hausdorff distance indicates better result; for Dice coefficient and overlapping volume percentage, larger value represents better result. 50

4.1 Parameters and Descriptions 62
4.2 Reconstruction errors (mm) of kidney pairs for different methods and proportion of missing values. The star $*$ indicates a statistically significant difference between the corresponding results and our method at a significant level of 0.01. 64
4.3 The computation time (s) for reconstruction of kidney pairs with various proportion of missing values from 0% to 50% for different methods. 65

4.4 The computation time (s) for reconstruction of kidney pairs (first row) and ankle bones (second row) using RKPCA with various proportion of missing values from 0% to 50%, where m, n denote the scale of data matrix. 65

4.5 Reconstruction errors (mm) of ankle bones for different methods and proportion of missing areas. The star $*$ indicates a statistically significant difference between the corresponding results and our method at a significant level of 0.01. 68

4.6 Statistical analysis for Generalization ability of competitive kidney pair models and ours for the first 6 modes. Note that the star $*$ indicates that the corresponding value of ours method is statistically significant better than the marked result. 71

4.7 Statistical analysis for Specificity of competitive kidney pair models and ours for the first 6 modes. Note that the star $*$ indicates that the corresponding value of our method is statistically significant better than the marked result. 71

4.8 Statistical analysis for Generalization ability of competitive ankle bone models and ours for the first 8 modes. Note that the star $*$ indicates that the corresponding value of our method is statistically significant better than the marked result. 72

4.9 Statistical analysis for Specificity of competitive ankle bone models and ours for the first 8 modes. Note that the star $*$ indicates that the corresponding value of our method is statistically significant better than the marked result. 72

4.10 Kidney Segmentation Results 74

4.11 Ankle Bone Segmentation Results 74

5.1 Overview of Pancreas Segmentation 96

5.2 Pancreas segmentation results comparison. Note that the star $*$ indicates a statistically significant difference between the corresponding results and our method at a significant level of 0.01. 97

5.3 Pancreas segmentation results comparing with the state-of-the-art. '–' indicates the item is not presented. 98

Chapter 1

Introduction

1.1 Background

With the development of digital data acquisition, massive amounts of medical image data from imaging modalities like Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are generated from hospitals every day. Biological structure and organ segmentation, i.e. delineation of the outline of a target object, is greatly useful to support surgery planning and treatments as it allows for quantitative analysis of intensity and contour. However, the manual annotation by experts is extremely expensive, time-consuming and problematic due to the large dimension and ever-increasing quantity of 3D radiological scans. Therefore, it is impracticable to label a sufficiently large datasets for quantitative analysis, which makes automatic segmentation a crucial pre-requisite in the computer-aided detection pipeline.

Particular challenges exist in biomedical image segmentation. First, often the boundary of target object is ambiguous due to the low intensity contrast to its surrounding tissues, e.g. the boundary between CT pancreas and liver may often be invisible (cf. Fig. 1.1). Secondly, the shape variability of biological structures, e.g. the pancreas (cf. Fig. 1.1) is often quite large across the population. Moreover, with the rapid development of

pattern recognition and machine learning, the low-level image processing techniques are gragually replaced by the model-based approaches, e.g. shape models, appearance models and neural networks. However, it is usually hard to collect a sufficiently large dataset of labeled medical images processed by pathologists to train models for segmentation. Addressing these major challenges in biological structure segmentation is the focus of this thesis.

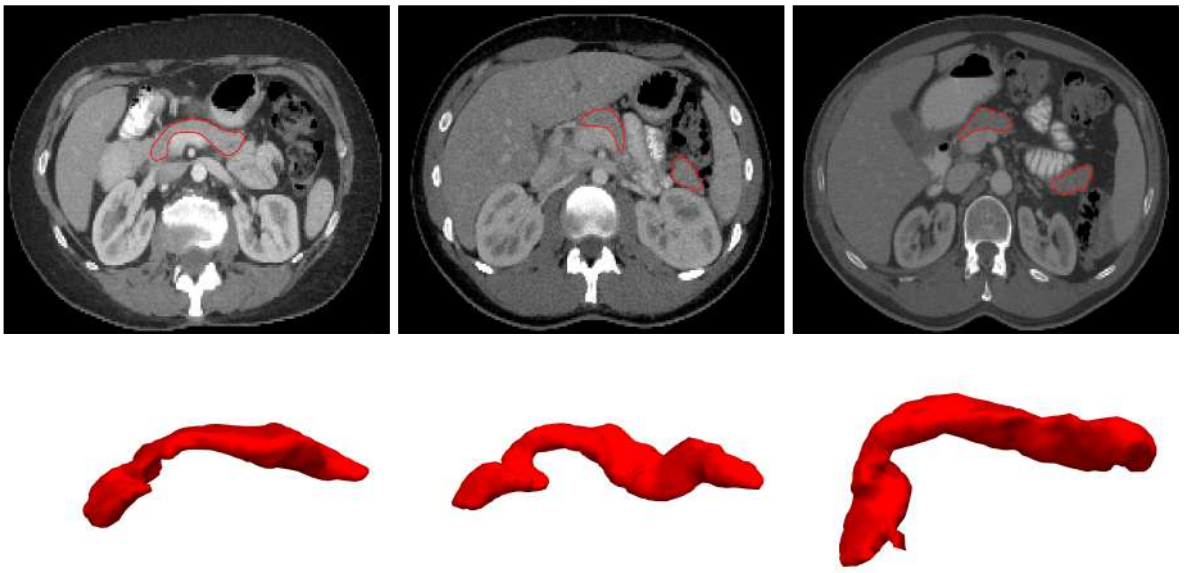


Figure 1.1: This figure illustrates the challenges in medical image segmentation, e.g. low-contrast of intensity and large variabilities of shape.

Automatic segmentation has been a critical issue for tens of years, and plenty of sophisticated approaches have been developed and achieved considerable results in various applications. As aforementioned, machine learning approaches, e.g. clusterings, classifiers and deformable shape models, have taken place of low-level pixel processing techniques over the past decade. Furthermore, deep neural networks have recently pervaded the whole field of medical imaging and become the preferred choice for researchers in segmentation according to a rapidly growing number of publications [2]. However despite its impressive success, particular challenges in medical imaging exist, e.g. small

size of training data and less intensity features, which inhibit deep neural networks from achieving satisfactory accuracies. Motivated by this, statistical shape models (SSMs) [3] are chosen to form the segmentation baseline because of the remarkable accuracy in various applications based on the shape prior. Furthermore, training SSMs takes less computational cost compared to the deep neural networks as SSMs are mostly trained in a self-learning manner. This thesis aims to develop novel statistical shape modeling techniques for 3D medical image segmentation.

1.2 Motivation and Hypothesis

It is known that a SSM-based segmentation framework usually consists of two subparts: shape modeling and image adaption. First, principal component analysis (PCA) is employed to extract a smaller number of principal components to represent the significant variabilities across the training shapes, forming a statistics model of shape. Subsequently, segmentation is realized by fitting the shape template to an unseen image until the optimal boundary is found, which is called image adaption in this thesis. At this stage, the image adaption is determined by the low-level intensity features and the high-level shape prior simultaneously. However, SSM-based segmentation approaches may not achieve satisfactory results when the shape model itself is not high quality due to the non-ideal training data and complex geometrical structure, and/or the image adaption is not capable to handle challenging cases, specifically, as outlined as follows.

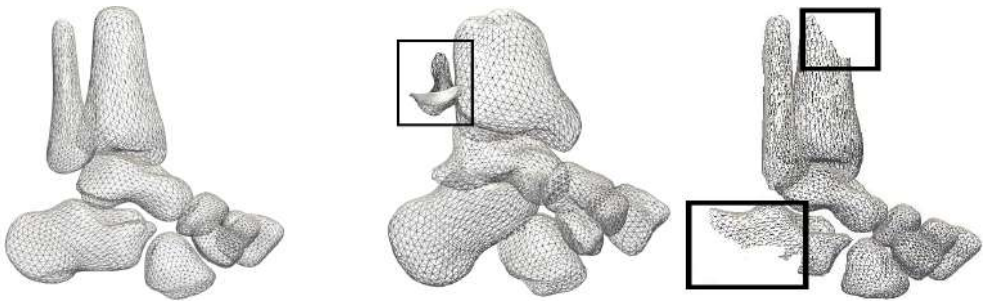
1.2.1 Data Contamination

The power of SSM-based segmentation rises and falls with the quality of the shape model. Unfortunately, the quality of training data is often severely deteriorated by factors like imaging artifacts, missing areas and label noise caused by non-ideal annotations even

by domain experts (cf. Fig. 1.2), which may significantly degrade the efficiency of SSMs.

Suppose we have a data matrix D constructed from the training samples, PCA aims to seek a low-dimensional subspce X where the dominant data lie on by minimizing $\|D - X\|_2$, i.e., the sum of l_2 norms of the residuals of the matrix decomposition [5]. However, it is brittle because any gross outlier will affect the optimization results and make the estimated subsapce X far from the true subspace X. Therefore, PCA is known to be sensitive to gross outliers.

This motivates us to develop robust matrix factorization techniques to learn the significant variabilities for statistical shape modeling, which are robust to arbitrarily distributed outliers (cf. Fig. 1.2-(b)) and/or missing data (cf. Fig. 1.2-(c)) in training samples. Subsequently, the proposed algorithms can be validated upon a set of training data containing varying degree of corruptions, by evaluating the ability of outlier recovery.



(a) (b) (c)

Figure 1.2: This figure illustrates examples of data corruption: (a) is a normal case of ankle bone shape; (b) shows a corrupted shape data with abnormal subpart (marked by a black square); (c) is an incomplete ankle shape with the missing areas (marked by a black square).

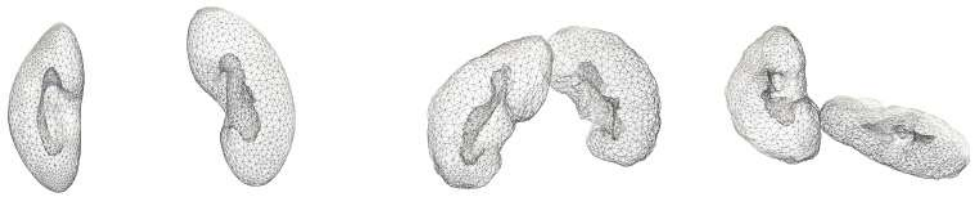
1.2.2 Complex Geometry

In addition to the data contamination, it is still demanding to derive a quality model for the biological structures with complex geometry, as the shape variance does not always follow a simple Gaussian distribution. Fig. 1.3 illustrates such an example of multimodally distributed object. Each single kidney, i.e. the left kidney and the right kidney, can be modelled by a unimodal distritbuion, whereas, the shape of combined left and right kidneys follows the multimodal distribution. Similar example can be found in Fig. 1.2-(a), where the ankle bone consists of 9 independent bones and follows a multimodal distribution as well. Besides, some nonlinear biological structures do not have a valid mean shape that cannot represent any existing vertebrae, e.g. the vertebrae in [90].

It is known that PCA model represents significant variances in a linear subspace that assumes to be a Gaussian distribution without bounded support. As a consequence, it lacks the ability to capture nonlinear patterns from such non-Gaussian distributions, which usually have bounded or semi-bounded domains. For example, the shape of kidney pair follows a non-Gaussian distribution (cf. Fig. 1.3) where the landmarks on the shape have bounded ranges. When the linear PCA model is applied on the kidney pair, invalid samples are generated from the model that have abnormal poses (cf Fig. 1.3-(b), (c)). This motivates us to consider nonlinear models for non-Gaussian distributions using the Kernel PCA (KPCA), which can be considered as the nonlinear version of PCA.

Even though the nonlinear patterns can be exploited by KPCA and combined multi-object models, data corruption in nonlinearity is rarely addressed in literature. As aforementioned, data contamination can be solved by robust matrix factorization. Accordingly, this thesis aims to project the linear robust matrix factorization onto a nonlinear space to remove outliers in complex geometry. Upon a set of training shapes with complex

geometrical structure, e.g. a multimodally distributed object, the proposed algorithms can be validated by artificially corrupting the training data.



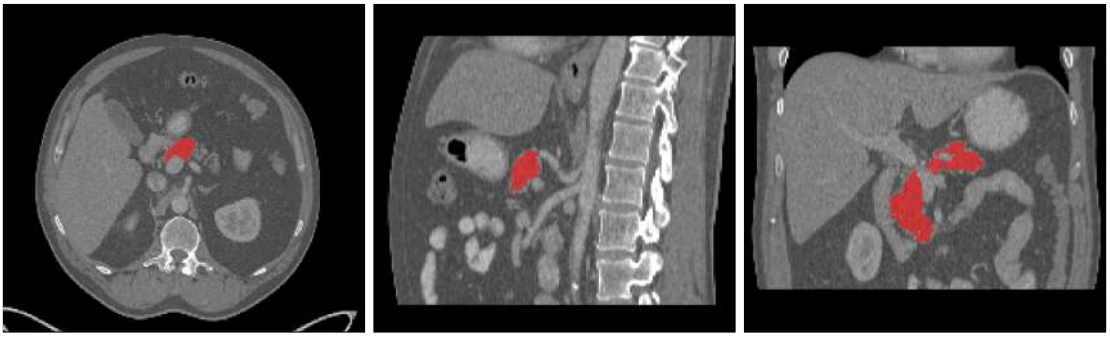
(a) (b) (c)

Figure 1.3: This figure illustrates a multi-object model of kidney pair, where (a) indicates the mean shape with normal variance, (b) and (c) show the exceptions with abnormal poses.

1.2.3 Poor Image Adaption

Although the SSM has a good quality, SSMs-based segmentation approaches may fail to achieve considerable results mainly because of the poor image adaption. On one hand, SSMs are often falsely initialized when the target object has a small, soft and highly deformed structure (cf. Fig. 1.4). This will affect the further shape fitting when the image adaption is not robust to the false model initialization. On the other hand, it is commonly seen that the target region has a blurry and invisible boundary due to the low contrast to its surroundings. Therefore, an efficient image adaption procedure is necessary to cope with such challenging cases.

Motivated by this, we aim to develop a robust image adaption procedure to achieve satisfied results in challenging segmentation cases. Sophisticated machine learning methods can be leveraged to extract discriminative intensity features not only to guide the landmark fitting, but also to provide a rough initialization to the SSM. In a word, this thesis aims to develop a novel unified SSM-based segmentation framework for challenging cases.



(a) Axial View (b) Sagittal View (c) Coronal View

Figure 1.4: This figure illustrates an example of small organ (marked in red), i.e. pancreas, from the Axial, Sagittal and Coronal view respectively.

1.3 Summary of Contributions

The major contributions are listed as follows.

- Focusing on self-learning shape modeling, this thesis presents a novel kernelized Robust Principal Component Analysis (KRPCA) for statistical shape modeling aiming at handling data corruption in nonlinear distributions, which is the major contribution in Chapter 3. Based on the well-known RPCA [5], an efficient tool of seeking a lower rank subspace free of outliers from the arbitrarily corrupted training data. The idea of the proposed kernelized RPCA is performing RPCA on the kernel space projected from training data. Since the kernel matrix represents the affinities between each pair of training data, a low-rank subspace recovery is performed to extract the primary clustering information and filter out unrelated ones. The outliers are discarded through dimensionality reduction in this kernel matrix. In this manner, the SSM created with the proposed KRPCA is robust to arbitrary corruption and capable to handle nonlinear data. Experiments are conducted to evaluate the model quality derived with the proposed approach and the state-of-the-art techniques based on a set of in-house MRI ankle bones, where each ankle bone involves 9 individual sub-parts. Evaluation results show that the

model built with the proposed KRPCA has a higher quality than other state-of-the-art methods.

- Extended from the KRPCA, a novel generic Robust Kernel Principal Component Analysis (RKPCA) is proposed for nonlinear shape modeling with a rigorous and complex mathematical derivation, which is the major contribution in Chapter 4. Quantitative evaluation is conducted using two different representative datasets: (1) a set of 30 public CT volumes of kidney pair and (2) a set of 49 MRI volumes of ankle bones. As explained, both the shapes of kidney pair and ankle bone are multimodally distributed with nonlinear features. Rich experiments are conducted to validate the proposed RKPCA and compare with the state-of-the-art methods. In order to evaluate the ability of recovering outliers, we artificially generate training shapes with various degree of corruption and compute the reconstruction errors, i.e. the distance between the reconstructions by the competitive methods and the corresponding ground truth data. Moreover, generalization ability, specificty and compactness are computed to measure the quality of SSMs created using competitive methods. The quality of models reflects the performance of the methods. Grounded with an existing image adaption, the built SSMs are employed in segmentation afterwards. In terms of the reconstruction error, model quality and segmentation accuracy, our proposed RKPCA shows a statistically better performance compared to the others.

- On the basis of the built SSMs, we focus on developing an image adaption approach to achieve the whole segmentation procedure aiming at challenging structures. Therefore, a novel Bayesian model is proposed that incorporates the SSMs with the output probability score from deep neural network for segmentation, which is the major contribution in Chapter 5. The motivation is aggregating both the spatial context exploited from the shape model and the intensity features from the

deep neural network. Moreover, the local shape statistics and global initialization clues together boost the segmentation accuracy. This full use of the data benefits the segmentation, especially when the datasets is limited. It is worth pointing out that there are no limitations to the choice of the deep neural network and the SSM. In this thesis, we improve the popular architecture UNet by replacing the stacked convolutional layers in downsampling path with the dense blocks from the DenseNet [6], in order to ameliorate the computational cost and strengthen the feature propogation. This neural network is utilized as a pre-prossing and called 'DenseUNet'. In evaluation, experiments are performed on 82 CT datasets of the challenging public NIH pancreas dataset, where the pancreas demonstrates very high inter-patient anatomical variability (cf. Fig. 1.1). The experimental results report 85.32 % of the mean Dice coefficient similarity that outperforms the state-of-the-art and approximately 12 % improvement from the predicted segment of deep neural network.

1.4 Thesis Organization

The organization of thesis is outlined as follows:

Chapter 2 provides a concise overview of medical image segmentation areas and approaches, followed by a detailed introduction of statistical shape models with practical applications. In Chapter 3, an overview of the matrix factorization techniques addressing the problem of data contamination is provided, followed by the pipeline of statistical shape modeling using the proposed Kernelized RPCA. Chapter 4 reviews the nonlinear compressive sensing methods and presents the detailed mathematical explanation of the proposed generic RKPCA, which is an extension from Kernelized RPCA and validated with rich experiments. In Chapter 5, the whole pipeline of the novel segmentation framework for challenging structures is provided, consisting of the DenseUNet and the Bayesian

model. Experiments are conducted upon a publicly available NIH CT pancreas, where segmentation accuracies are compared with the state-of-the-art approaches. Chapter 6 concludes the thesis and provides some research directions for future work as well.

Chapter 2

Literature Review

This chapter first conveys a concise overview of medical image segmentation, consisting of major application areas and approaches that are the most frequently used; secondly, introduces the pipeline of statistical shape modeling as well as its recent applications. Publications referred in this survey are mainly published at medical imaging related journals like *IEEE Transactions on Medical Imaging*, *Medical Image Analysis*, and conference proceeding of *MICCAI*. Note that the mentioned references are very representative, namely, either the most often cited or the latest.

2.1 Medical Image Segmentation

2.1.1 Overview of Application Areas

In the following, we briefly review the key applications and make copious reference to the recent research issues listed on **All Grand Challenges**[1]. Table 2.1) reports an overview of a widespread of medical image segmentation applications, which can be generally divided into two categories: anatomical organ/substructure segmentation and pathology segmentation. To be specific, anatomical organ/substructure segmentation

[1] http://www.grand-challenge.org

assumes the pixels belong to the region of interest have similar image properties like texture and intensity. Furthermore, the shape variability can be learned under certain criterion. Differently, pathological structures, e.g. tumor and nodule, are usually very small and have uncertain appearance.

Table 2.1: Major Application Areas

	Category	Remark
Anatomical Organ/Structure	Brain	brain ROI [7]; hippocampus[8]; brain tissue [9, 10]; AVP [11]
	Chest	lung structure [12]; airway [13]; thoracic vessel [14]
	Cardiac	left [15] /right [16] ventricle; whole heart [17]
	Eye	blood vessel [18]; retinopathy [19]
	Abdomen	liver [20]; kidney [21]; pancreas; spleen [22]; prostate [23, 24]
	Skeletal	knee cartilage [25]; vertebrae [26]; hip [26]; foot [27]; head and neck [28]; hand [29]
Pathology	Tumor	brain tumor [30]; prostate cancer [31]; breast cancer [32, 33]; skin lesion [34]; cervical cell [35]
	Gland	gland segmentation challenges [36, 37]
	Nodule	pulmonary nodules [38, 39]

2.1.1.1 Anatomical organs and substructures segmentation

Brain: A large number of works address segmentation of brain and its anatomical substructures, e.g. hippocampus and tissue. In [9], MRI brain is segmented into several anatomical regions using deep neural network; in [10], a novel Voxel-wise Residual Network (VoxResNet) is applied to segment key brain tissues. Addressing the injured brain, [7] presents a robust segmentation approach for MRI brain with gross abnormalities. Besides brain tissues, a novel approach based on locally normalized cross correlation incorporating an extension of conventional STAPLE algorithm is proposed and applied in hippocampus segmentation in [8]. In [11], Mansoor et al. present an automated segmentation framework for MRI Anterior Visual Pathway (AVP) using a deep neural network

incorporating a partitioned statistical shape model.

Chest: Thoracic analysis, e.g. detection and segmentation, from radiography and CT is commonly addressed in literature. In [12], lung boundary is detected based on a patient-specific atlas with a local graph cut refinement, where the proposed method is validated using both normal cases and lungs with various diseases. Addressing 3D airway tree structure segmentation, Xu et al. [13] propose a computational framework to quantify airways considering of the speculated shapes with branches and bifurcation points on airways. Challenge *VESSEL12* [14] is proposed with the aim of automatic CT thoracic vessel segmentation, as manual vessel segmentation is laborious and time-consuming.

Cardiac: Several challenges on cardiac segmentation are presented on the All Grand Challenges, e.g. *2009 Cardiac MR Left Ventricle Segmentation Challenge*, *2012 MRI RV Segmentation*, *2014 CETUS 3D Ultrasound Cardiac Segmentation* and *2017 Multi-modality Whole Heart Segmentation*. Upon the datasets provided by *2009 Cardiac MR Left Ventricle Segmentation Challenge*, [15] employs deep neural networks combined with deformable shape model in MRI left ventricle (LV) cardiac segmentation. In [16], an automatic MRI right ventricle (RV) segmentation framework is presented using random forest classifier, which gives probabilities of being RV and background for each pixels. In [17], a multi-modality MRI whole heart segmentation framework is presented using a novel multi-scale patch scheme, which offers different levels of structural information.

Eye: Retinopathy detection and segmentation is a crucial pre-requisite step in realistic applications, e.g. diabetic retinopathy screening. In [19], Zhang et al. propose an automatic exudate detection framework based on random forests, which can be used for denoising and detection reflections and imaging artifacts. In [18], a novel thresholding segmentation approach for blood vessel structure is proposed addressing large variations in illumination and thin vessels.

Abdomen: To date, a large number of abdominal organs segmentation approaches are proposed, and large anatomical structures such as liver, kidney and spleen have met considerable segmentation accuracies. In [20], a novel liver segmentation framework is proposed, which is robust to pathologies and validated upon two popular public databases *SILVER07* and *IRCADb*. In [21], Norajitra et al. propose a 3D SSM based segmentation approach incorporating with landmark-wise random regression forests for CT kidney segmentation. The work [22] focuses on MRI spleen segmentation addressing the challenge of large structural variation in 3D MRI scans. Within the scope of abdominal organs, pancreas segmentation is an extremely challenging task due to the ambiguous boundaries and large anatomical variations. A concise survey of pancreas segmentation is provided in Chapter 5. MR Prostate segmentation has raised extensively attention as well as it plays an important role in prostate cancer radiotherapy, on which challenges are proposed, e.g. *2012 Prostate MRI Segmentation (PROMISE12)*. In [23, 24], an automatic MRI prostate segmentation approach is proposed addressing the inhomogeneous and diverse appearance variance of prostate.

Skeleton: In addition to the soft tissues, skeletal anatomical segmentation is commonly addressed as well, e.g. vertebrae, knee cartilage, hand and ankle bones. Several challenges on vertebrae segmentation are proposed, e.g. *2014 Spine and Vertebrae Segmentation Challenge*, *2015 MICCAI Spine Workshop* and *xVertSegt Challenge 16*. Early works aim at segmenting each vertebrae, whereas, recent approaches are proposed to extract multi-vertebrae bones simultaneously. In [26], a novel shape and pose model is proposed for CT spin segmentation by exploiting the common statistics among the different anatomies. In [25], deformable model incorporating weighted shape priors is applied in MRI hip joint segmentation. Prasoon et al. propose [40] a novel system for knee cartilage segmentation that integrating three 2D deep neural networks, which are trained based on the slices extracted from Axial, Sagittal and Coronal planes respectively.

2.1.1.2 Pathology segmentation

Tumor: Tumor detection and segmentation has become an active research area recently, which has challenges like low-contrast to surroundings and irregular borders. The work [30] provides a report of comparable segmentation results addressing the challenge *BraTS* organized by MICCAI conference, which aims at segmentation of brain tumor in MRI scans. Furthermore, [31] presents an automated computed-aided detection (CAD) system for MRI prostate cancer detection and segmentation, achieved by multi-atlas based approach incorporating a classifier. Public challenges on breast cancer detection, e.g. *CAMELYON17 Challenge* , have raised attention in literature as breast cancer is one of the most prevalent cancers for female. In [32], Veta et al. provide an overview of breast cancer analysis followed by a discussion of techniques and applications. In [33], a novel statistical learning technique for tumor segmentation is proposed, which is based on Hidden Markov Models and applied in breast cancer diagnosis. In [34], a skin lesion segmentation is proposed by deep neural networks trained end-to-end, which is validated upon the public datasets provided by the public challenge *2016 ISBI Skin Lesion Analysis Towards Melanoma Detection*. In [35], Song et al. present an automatic segmentation of cervical abnormal cells based on robust shape priors in order to tackle the challenges of overlap.

Gland: Accurate glands segmentation from histology images is very useful for quantitative diagnosis by pathologists. In [36], Chen et al. propose a unified deep contour-aware network (DCAN) for gland segmentation. The proposed approach is robust to large pathology and achieves the best result in the public challenge *2015 MICCAI Gland Segmentation*. Addressing colorectal cancer segmentation, challenge *2015 Gland Segmentation in Colon Histology Images Challenge Contest* is proposed and raised much attention. In [37], a detailed introduction is provided, including the datasets, evaluation measurements and state-of-the-art techniques tackling colon cancer.

Nodule: Recently, a popular challenge *LUNA16* is proposed [38] with the aim of automatic detection of pulmonary nodules in thoracic CT scans. In [39], Want et al. propose a novel lung nodule segmentation framework based on convolutional neural networks addressing the difficulties like heterogeneity and low-contrast of characteristics to neighboring tissues.

To summarize, medical image segmentation application areas can be roughly categorized into two classes: anatomical organ/substructure segmentation, including soft tissues and hard tissues, and pathology segmentation, including tumor and nodule. Apparently, it can be found that the number of publications addressing anatomical structure segmentation is larger than that of pathology segmentation. Whereas, with the rapid development of deep neural networks, lesion detection and segmentation have raised more and more attention in recent years.

2.1.2 Overview of Major Approaches

In literature, a great deal of methods are presented for a widespread of medical image segmentation applications, which are categorized as follows.

Table 2.2: Approaches in medical image segmentation

Category	Techniques
Low-level pixel processing	edge detection [41]; region growing [42]
Clustering	k-means [43]; fuzzy c-means [44]
Classifier	SVM [45]; KNN [46]; random forests [47]
Graphical models	MRF [48]; CRF [49]; graph cuts [50]
Shape models	ASM [4]; SSM [3]
Deep neural networks	FCN [51]; UNet [1]; SegNet [52]
Others	level set [53]; atlas-based methods [54]

Low-level pixel processing: Initially, medical image segmentation is achieved using the low-level image processing techniques. For instance, a novel edge detection technique

is proposed in [41] based on mathematical morphology, e.g. dilation and erosion, to suppress noise and adapt to different edges. In addition to edge filters, region-based approaches like region growing [42] are commonly used in segmentation applications as well, which learn the homogeneity properties automatically from the seeds located in the target region where the learned homogeneity and characteristics are updated continuously in the segmentation.

Clustering: With the development of pattern recognition and machine learning, the low-level pixel processing is gradually replaced by the statistics analysis based on a set of training data. That is to say, discriminative features are learned automatically from a high-dimensional data space and used to determine the optimal boundary. Hereby, numerous clustering and classifier techniques are employed in segmentation. Under an unsupervised learning, clustering approaches, e.g. k-means and fuzzy c-means, compute the mean and covariance for each class, followed by clustering each pixel to the closest class iteratively. In [43], a novel approach incorporating k-means and an improved watershed is proposed for MRI segmentation, where the unsupervised k-means clustering is used to obtain a primary result followed by the watershed refinement. In [44], a modified fuzzy C-means (FCM) is developed for MRI segmentation.

Classifier: As opposed to clustering, classifiers partition the feature space into classes under supervised learning. In [45], Wang et al. integrate the landmark-based shape deformation and a regularization using a novel constrained support vector machine (SVM) to extract scalp contours from brain in cryosection images. Addressing the problem of high dimensional data and limited size of training set, a Bayesian model incorporated with appearance shape model is proposed in [55] to extract intra- and inter- variabilities for brain ROI segmentation. K-nearest neighbor classifier is adopted in [46] to classify the brain tissues in MRI. To address the challenges in segmentation, authors prefer to integrate multiple basic classifiers to achieve satisfactory. Furthermore, random forests

[47] deliver considerable efficiency in local intensity features classification in a wide range of applications. In [56], random forests are used to quantify the learned discriminative features in Cardiac RV, liver and kidney segmentation. In [57], a modified random forests is utilized to learn local features for kidney components segmentation. Wu et al. employ the random forest as a regressor to guide deformable model fitting for MRI brain ROI segmentation.

Graphical models: Graphical probabilistic models, e.g. Markov Random Field (MRF) [58] and conditional random field (CRF) [49], are commonly used in medical image segmentation, under the assumption that pixels belong to region of interest (ROI) have similar intensity and texture. As a statistical model, MRF can be easily incorporated in any segmentation framework due to its flexibility. In [48], a MRF based energy function incorporated with level set is proposed in order to enhance the robustness to noise for medical image segmentation. Graph cut, proposed in [50], efficiently forms the basis of segmentation constraint and can be easily incorporated with other machine learning methods with a balance strategy. In [59], Chen et al. integrate graph cuts with active appearance model (AAM) to segment liver, kidney and spleen on a public CT datasets. In [60], the basic graph cut is improved and reformulated to a deep cut and embedded into the deep neural network.

Shape models: From 1990s, the deformable shape models have achieved success in segmentation, such as, active shape model (ASM), active appearance model (AAM) [4], active contour model (ACM) and [61]. The motivation behind is fitting a shape template to an unseen image data, where the adaption is determined by low-level image intensity features and high-level shape prior simultaneously in a unified framework. The shape prior is represented by a statistical shape model [3, 62], which learns and quantizes the significant shape variabilities. The shape prior learned from SSMs is very valuable

to boost the segmentation performance. With long tradition, SSMs have been successfully employed in a various of applications, especially in anatomical organ/substructure segmentation.

Deep neural networks: Deep neural networks have become preferred in computer vision over the past five years. It is found that the deep neural networks have permeated almost the entire field, in particular, the number of publications grow rapidly since the year of 2015. Despite the impressive success in computer vision, it is known that the deep neural networks highly depend on the number of labeled training data. However, often the medical image data with labels are hard to collect, which inhibits many deep neural networks from achieving satisfactory results. Still, deep neural networks like UNet [1] as well as 3D UNet [63] have achieved considerable results in segmentation. It is worth mentioning that a survey of deep learning in medical imaging is provided in [2]. As aforementioned, many existing deep neural networks cannot achieve satisfactory results in the challenging cases of segmentation, due to the factors like small number of training data and limited computing memory. To combat this, usually a basic deep neural network is combined with other machine learning approaches in practice like graphical models MRFs, CRFs and SSMs to refine the segmentation output, which are applied on top of the probability score map produced. In [64, 65, 66], CRF is applied at the end for segmentation refinement. Similarly, MRF is often incorporated into segmentation as well. The MRFs parameter the spatial interactions between neighbor pixels, where the pixels belong to the same class are taken into account. Zhang et al. propose a novel hidden MRF in [58] under the EM inference for MRI brain segmentation. In [67, 68], MRFs are associated with deep neural networks. In addition to the graphical models, deep neural networks are combined with shape prior in [11, 69]. Other applications in medical image segmentation using deep neural networks can be found in [70, 71, 72, 73].

Level set: Level set [53] is another popular technique in segmentation due to its flexibility. Different from the active contour models, level set does not rely on landmark representation, as it leverages an implicit high dimensional function to represent the evolving contours as a zero level set, which can easily handle the topological changes. In [74], Farhangi et al. present a general segmentation framework for X-ray CT lung nodules detection, which is driven by a level-set evolution and an additional term representing the valid prior shapes. In [75], the level set evolution is reformulated and embedded into the deep neural network architecture.

Atlas-based methods: Atlas-guided method [54] is a powerful in medical image segmentation, where the atlas is generated from the target anatomy and used as a reference in segmentation. In [76], a multi-atlas based segmentation framework is proposed integrating an image similarity metric with a contour similarity prior for prostate segmentation. In [77], Wachinger et al. propose an automatic atlas-based segmentation framework under Bayesian inference for glands in head and neck.

2.1.3 Discussion

This section overviews the major segmentation application areas (cf. Table 2.1) and the most commonly used approaches tackling different segmentation problems (cf. Table 2.2). Conclusions can be drawn from the surveys that: (1) supervised classifiers and deep neural networks efficiently harness multi-hierarchical image intensity based features, whereas, deformable models are powerful in representing shape variabilities; (2) deep neural networks have achieved satisfactory results in challenging segmentation cases, however, they suffer from small size of training data and high computational cost; (3) traditional machine learning approaches, e.g. classifiers and shape models, are flexible in learning local features and do not highly depend on a large amount of training data; (4) image-based approaches are more efficient in pathology segmentation where the

target object has various shapes, and deformable models are more powerful in anatomical structure segmentation, especially in cases of low-contrast intensity to surrounding tissues.

2.2 Statistical Shape Models

Considering the major challenges and the application, this thesis focuses on SSM-based segmentation framework. This section will introduce the baseline of statistical shape modeling as well as its applications in medical image segmentation over the past five years.

2.2.1 Baseline of Statistical Shape Modeling

Given a set of N_s training shapes extracted from 3D volumes via Marching Cubes, a crucial pre-requisite step is the Procrustes alignment [78], the aim of which is to reduce unnecessary variabilities caused by poses such as orientation and location. Subsequently, usually a point-to-point groupwise correspondence is applied to the N_s meshes, in order to assign the areas with similar features among all training samples the same landmark index. Accordingly, each shape contains N_p landmarks that is represented as a column vector under the Point Distribution Model (PDM): $\boldsymbol{x} = \left(p_x^{(1)}, p_y^{(1)}, p_z^{(1)} \cdots p_x^{(N_p)}, p_x^{(N_p)}, p_x^{(N_p)}\right)$, where $\left(p_x^{(i)}, p_y^{(i)}, p_z^{(i)}\right)$ is the 3D Cartesian coordinate of the i^{th} point on the shape. A column stacked data matrix X can be constructed that $X \in \mathbb{R}^{3N_p \times N_s} = [\boldsymbol{x_1}, \cdots, \boldsymbol{x_{N_s}}]$.

PCA is applied to reduce the dimensionality as usually $N_p \geq N_s$. By performing eigendecomposition, the dominant first m eigenvalues $\lambda_{1\cdots m}$ and eigenvectors $\mathbf{N}$ are learned. In this manner, the high dimensional training data can be represented using a compressed low-dimension feature space, which contains dominant data variability. Namely, a mapping from each data point $\boldsymbol{x}$ in the input space and the feature space is constructed

as:

$$\boldsymbol{x} = \bar{X} + \mathbf{N}\boldsymbol{b}, \tag{2.1}$$

where $\boldsymbol{b} \in \mathbb{R}^m$ is the shape parameter in the feature space and $\bar{X}$ implies the mean shape vector in the training data set. Each component in $\boldsymbol{b}$ varies in the range $b_i \in [-3\sqrt{\lambda_i}, +3\sqrt{\lambda_i}]$. In terms of (2.1), each shape sample $\boldsymbol{x_i}$ in the training set is corresponded to a unique parameter vector $\boldsymbol{b_i}$, conversely, there exist only one reconstruction in the training set from the shape parameter in feature space. At this stage, the probability density function of the class of shape can be drawn from the learned shape feature space.

2.2.2 SSM-based Segmentation Applications

Over the past decade, SSMs have been very successful in segmentation tasks which range from organs to lesions, and from soft tissues to hard tissues. Early attempts using SSMs on CT **liver** segmentation can be found in [79, 80], where the segmentation is driven by an local image adaption with a global constraint learned from SSMs. In [81], a unified model integrating two appearance sub-models and an adaptive shape sub-model is proposed to segment pathological and healthy **lungs**. Albà et al. [82] present a generic SSM for abnormal **hearts** segmentation, where the abnormality in patient data is approximated and constrained by back propagation onto the SSM. In [83], a nonlinear SSM is derived to represent nonlinear variance of the **left ventricle** and used in MRI Cardiac segmentation. Associated with random regression forests, 3D SSM is applied in **kidney** segmentation in [21]. Besides the single subject, SSMs are employed in multi-object segmentation as well. In [84], a level-set-based SSM is proposed by conducting a branch-and-bound search over an eigen-shape space and used in **pancreas** and **spleen** segmentation. In [85, 86, 87], SSMs are employed in **multi abdominal organ** segmentation based on a limited amount of labeled training data. In [85], a general framework of multi-organ

segmentation is proposed that incorporates interrelations among multiple organs and models conditional shape and local priors without the prior intensity knowledge. Besides organs, SSM as well as its variations have proven to achieve fairly good results in small region recognition [88, 89]. In [89], 4D-SSM is derived for **lungs with large tumors** segmentation in 4D CT data.

In addition to the highly deformed soft tissues, segmentation for articulated joints of hard tissue has attracted extensive attention as well. For instance, **lumbar vertebrae** is an extremely challenging organ for segmentation due to its high complexity in geometry and large variability of individuals, numerous works address this challenging segmentation tasks in literature [90, 26, 91, 92]. In [91], a statistical shape decomposition and conditional model is proposed, the key technique of which is to reduce the complexity of subparts and model inner-relationships. In [92], inner-space among individuals are modeled in order to avoid the abnormal overlap, leading to the Statistical Interspace Model (SIM). In [26, 93], a statistical multi-shape model joint with a pose model is adopted in vertebrae and **wrist bone** segmentation respectively, note that the multi-object model captures the variabilities of shape and pose simultaneously. Similarly, Chen et al. [94] create a SSM incorporated the pose variation arising from articulations for wrist bone segmentation. In [29], a shape and appearance model is derived for X-ray **hand** segmentation, where the segmentation is optimized by a single global criterion under the Maximum a posteriori inference. In our publications [27, 95], robust SSMs are derived for MRI **ankle** bone segmentation which aims to cope with the data contamination and capture nonlinear patterns respectively.

2.2.3 Nonlinear SSMs

The majority of approaches deal with complex geometry by incorporating other machine learning techniques and/or utilizing integrated shape and pose models under the

assumption of a Gaussian-like distribution. In contrast to linear SSMs, the number of approaches exploiting nonlinear variabilities of shape data is relatively small. However, it is commonly seen that nonlinearity and linearity exist simultaneously in a probabilistic distribution of a class of shape, e.g. the inner space between sub-parts. This motivates an active area that generalizing the standard linear modeling techniques to nonlinear spaces. A quantity of modeling techniques introduce a multi-object combined model to model the nonlinear patterns, e.g. [26, 93]. A recent modeling approach is proposed in [96], which endows a differential representation of shape with a nonlinear Riemannian structure for identification of intra- and inter- population variability. Many biological structures, however, contain nonlinearities even though the non-linearity might not be always strong. In these cases, a linear model does not represent the anatomical structure well. On the other hand, Kernel PCA (KPCA) [97, 98] is often applied to capture nonlinear variance of a model instead of standard PCA, e.g. [90, 83]. The idea of KPCA is projecting the nonlinear input space onto an implicit high dimensional feature space, where linear variances can be learned from. Due to the implicitness of the feature space, the affinities of each pair of samples in the feature space is learned through a kernel trick, therefore, a kernel matrix is formed. Note that the kernel trick can be chosen from many different kinds depending on applications, where RBF is the most frequently used and defined as:

$$\kappa(x, y) = \exp\Big(-\frac{\|x - y\|^2}{2\sigma^2}\Big),$$

where σ decides the degree of smoothness.

2.2.4 Discussion

Summarizing from the widespread of SSM-based segmentation applications, it can be found that: (a) SSMs are mainly used in anatomical structure segmentation, from soft tissues to hard tissues, owing to its efficiency in representing shape variability; (b) SSMs

are extremely flexible in practice that it can be combined with clusterings and classifiers, and other segmentation frameworks, e.g. level set and deep neural networks; (c) SSMs do not highly depend on the amount of training data compared to the deep neural networks, namely, shape variabilities can be learned even from a small amount of training data; based on data compression, SSMs are trained with less computational cost and light weight.

Table 2.3: Major challenges in statistical shape modeling

Major Challenges	**Reference**
Nonlinearity	shape + pose model [26, 91, 93]
	KPCA [90, 83]
	Riemannian model [96]
Data Contamination	Probabilistic PCA [99]
	low-rank modeling [100, 101]

Trained by PCA, standard SSMs suffer from limitations that they lack the ability to represent nonlinear patterns, e.g. the inner space between subparts in spine, and are not robust to corrupted training data. Table 2.3 delivers a brief overview of literature works that address the major challenges. First, to represent nonlinear patterns, the work [26, 93] create a statistical model parameterizing the shape variance that incorporates an extra statistical model representing the pose and appearance of each subparts. Similarly, in [91], complex geometry of spine is modeled based on PCA SSMs of each single vertebrae with a conditional model for inter parts. Besides, the work [96] parameterizes nonlinear shape variance using the a differential shape representation projecting onto a non-Euclidean Riemannian space. Under a unified modeling scheme, KPCA is commonly used to learn nonlinear patterns instead of joint linear models, e.g. KPCA model is created in [90, 83] for vertebrae and left ventricle respectively. Secondly, data corruption, which is commonly seen in various applications, is also addressed in statistical shape modeling. In [99], a probabilistic PCA under EM optimization is used to handle corruptions in

training data. Under the same dataset, [100] employs the Robust PCA (RPCA) [5] that delivers a better performance in outlier recovery compared to the probabilistic PCA in [99]. The baseline of low-rank modeling presented in RPCA is applied to cope with pathological regions in segmentation in [101].

Despite the success of addressing major challenges, nevertheless, the data corruption in nonlinear distribution is rarely mentioned in literature that motivates us to fill this gap. From the aspect of SSM-based segmentation applications, it is apparently found that shape models are sensitive to initialization. To combat this issue, traditional machine learning techniques are triggered in segmentation to offer an initialization to shape model. Therefore, this thesis aims to create a robust nonlinear SSM and apply it to realize segmentation.

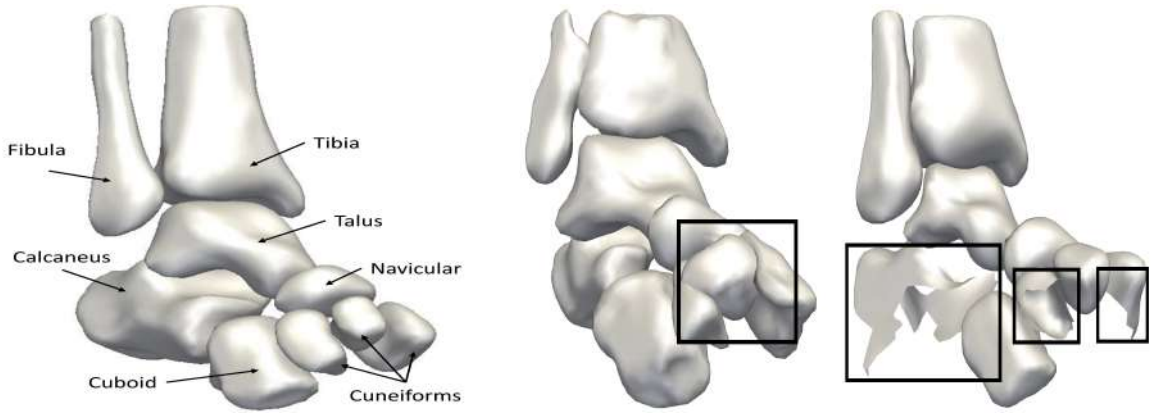
Chapter 3

Nonlinear Statistical Shape Modeling Using A Novel Kernelized Robust Principal Component Analysis

Statistical Shape Models have been employed in various applications of medical image segmentation. Using a set of shape training data, SSMs parameterize the significant variabilities of the anatomical structure of interest via PCA. The learned prior knowledge can then be used to effectively boost the performance of segmentations by projecting any distorted input shape to the SSM and constraining it to the most plausible modes of variance. In many practical applications, a high quality model significantly contributes to the final segmentation result. In spite of the considerable success, it is still challenging to create high quality models due to the following challenges:

- *Multimodal Shape Distributions:* The shape variance of many biological structures does not follow a simple Gaussian distribution. For example, the mean shape of a complete vertebrae model [90] is not meaningful as it does not represent any existing vertebra. The same applies to combined shapes with non-uniform structure (cf. Fig. 3.1-(a)), where linearity and nonlinearity exist simultaneously in a probabilistic

distribution. Nonlinear SSMs have been proposed, but existing methods cannot accurately project highly distorted input shapes to the nonlinear shape space in order to remove non-plausible distortion. A high level distortion in input shape, however, is likely to occur with most boundary detection methods.



(a) (b) (c)

Figure 3.1: (a) Ankle bone structure (b) Corrupted ankle bone with anomalous overlap marked by a black square (c) Incomplete ankle bone with missing areas marked with black squares

- *Data Corruption:* There are many factors influencing training data quality. Imaging artifacts, missing areas in images, non-visible organ boundaries, as well as inter- and intra-subject variance may lead to non-ideal delineations (cf. Fig. 3.1-(b), (c)) which in turn degrade the quality of SSMs using these delineations as training data. Furthermore, in practice, the amount of available ground truth data is usually limited since manual delineation is time-consuming and costly. Severe corruption in training data therefore has a significant impact on the whole SSM.

Addressing these challenges, a novel kernelized Robust PCA (KRPCA) with the aim to create SSMs that are not only robust to abnormal training data but also satisfied with nonlinear distribution, which is the major contribution in this chapter. It should be pointed out that partial content in this chapter comes from our publication "Nonlin-

ear Statistical Shape Modeling Using A Novel Kernelized Robust Principal Component Analysis" presented at MICCAI 2017.

The remaining of this chapter is organized as follows: first, a concise overview of matrix factorization techniques addressing the training data contamination is presented, followed by the introduction of statistical shape modeling using RPCA and our early contributions Weighted RPCA [27] as well as KPCA; subsequently, the proposed KRPCA is illustrated in detail; afterwards, experiments are conducted to validate the proposed method's ability of recovering outliers in training data, and evaluate the quality of model derived using competitive approaches as well as segmentation accuracy using these models. To be mentioned, evaluation results show that the model built with the proposed KRPCA has a higher quality than other state-of-the-art methods.

3.1 Introduction

As aforementioned, PCA is employed to extract the significant variabilities of the class of shape. However, PCA is extremely sensitive to gross corruption as it only assumes a small proportion of Gaussian noise. The problem of data contamination has raised intensively attention as it is commonly seen in various applications, hereby, a number of matrix factorization approaches have been proposed to address this issue (cf. Table 3.1).

Given a data matrix $D \in \mathbb{R}^{m \times n}$ constructed from training samples, the aim is to recover a low rank subspace free of outliers. Note that m decides the number of training samples and n decides the dimension of each sample. Compressive sensing is considered to reconstruct the high-dimensional training data based on a small number of orthonormal basis X, where $\|X\|_0$ denotes the number of non-zero entries in X, $\mathcal{A}$ is a given underdetermined linear operator and b is a given vector with smaller dimension. As opposed

Table 3.1: Overview of Matrix Factorization

Method	Objective	Subject to
Compressive Sensing [102]	$\min_X \|X\|_0$	$\mathcal{A}(X) = b$
Nuclear Norm Minimization [103]	$\min_X rank(X)$	$\mathcal{A}(X) = b$
Robust Non-negative Matrix Factorization [104]	$\min_{W,H,E} \|D - WH - E\|_F^2 + \lambda \|E\|_1$	$W \geq 0, H \geq 0$
Low Rank Representation [105]	$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1}$	$D = XZ + E$
Robust Principal Component Analysis [5]	$\min_{X,E} \|X\|_* + \lambda \|E\|_1$	$D = X + E$

to compressive sensing, the motivation behind nuclear norm minimization is seeking the best rank of X as nuclear norm is a convex relaxation of rank minimization. Such techniques show significant efficiency in matrix completion [106], defined as $\min_X \|X\|_*$, $X_{ij} = D_{ij},\ \forall (i,j) \in \Omega$. Non-negative Matrix Factorization (NMF) [107] is another commonly used tool for dimensionality reduction [108, 109], which aims to find two non-negative matrices W, H from the data matrix that is $D = WH$. Upon NMF, a robust NMF is proposed to split out the noise with an additional l_1 norm term. Based on the dictionary learning [110, 111], the Low Rank Representation (LRR) seeks the best rank of "dictionary" subset Z with a separated sparse component E containing outliers. To avoid the loss of generality, usually $X = D$ when the noise is small. LRR-based techniques are popular in sparse matrix recovery [112, 113]. By contrast, Robust PCA (RPCA) recovers a low-rank subspace X from training data straightforwardly, where the data sparsity is filtered out through minimizing the nonlinear l_1 norm. Note that λ is a positive trade-off to balance the two terms in LRR and RPCA. By efficiently addressing the sensitivity to outliers of PCA, RPCA shows fairly good performance in a wide range of applications, from video surveillance to signal processing.

Nevertheless, data corruption is rarely addressed in statistical shape modeling, even though it is one of the most common problems in computer vision. The probabilistic PCA (PPCA) is used to derive SSMs from lousy training data in [99], which is under

an Expectation Maximization optimization inference. Furthermore, RPCA is successfully applied in statistical shape modeling in [100] and delivers better performance than PPCA under the same evaluation criterion. Addressing the data contamination in linear distribution, many efforts have been devoted in this study. In our prior contribution [114], an imputation method is developed to cope with outliers in training shapes. The idea is to replace the identified outliers with the mean value derived from the population, where the outliers are determined in terms of data frequency in distribution. Based on RPCA, a Weighted RPCA (WRPCA) is proposed in [27], the motivation behind is leveraging some prior knowledge to separate outliers and preserve underlying data in the constructed low-rank subspace.

3.1.1 Limitations of Robust Linear SSM

Given a set of N_s training shapes, each of which is represented as a point-set $d_i = \left(p_x^{(1)}, p_y^{(1)}, p_z^{(1)}, \ldots, p_x^{(N_p)}, p_y^{(N_p)}, p_z^{(N_p)}\right)^T$ using the Point Model Distribution (PDM), where $(p_x^{(j)}, p_y^{(j)}, p_z^{(j)})$ indicates the 3D Cartesian coordinate of the j^{th} point on the shape. Afterwards, a column stacked matrix $D \in \mathbb{R}^{N_s \times 3N_p} = [d_1, \ldots, d_{N_s}]$ is constructed for training a shape model. Deriving a standard SSM, PCA performs eigen-decomposition on D and extract the first n_k modes of variance from eigenvectors $\{\boldsymbol{\alpha^i}\}_{i:1\ldots n_k}$ and eigenvalues $\Lambda_{i:1\ldots n_k}$ to describe the class of shape.

Taking advantage of the data sparsity, RPCA assumes that the data matrix D can be decomposed into a latent clean low-rank subspace X free of outliers and a sparse component E representing the data sparsity, i.e. outliers. The objective is formulated as

$$\min_{X,E} \|X\|_* + \lambda \|E\|_1, \quad \text{s.t. } D = X + E. \tag{3.1}$$

This can be seen as minimizing the nuclear norm $\|X\|_*$, a convex relaxation of minimizing $rank(X)$, and l_1 norm of E with a positive trade-off λ. In [115], Lin et al. present an

inexact augmented Lagrange multiplier (IALM) using an additional quadratic penalty in contrast to the standard Lagrange multiplier. We adapt IALM to solve (3.1) that delivers better performance than the dual approach, proximal gradient and the Exact ALM and formulate the Lagrangian:

$$\mathcal{L}(X, E, Y, \mu) = \min_{X,E} \|X\|_* + \lambda \|E\|_1 + \langle Y, D - X - E\rangle + \frac{\mu}{2} \|D - X - E\|_F^2, \tag{3.2}$$

where Y is multiplier and μ is a positive scalar that penalizes the violation of the linear constraint, the notation $\langle \cdot, \cdot \rangle$ denotes the standard trace inner product between two matrices of the same size and $\|\cdot\|_F$ is the induced Frobenius norm of a matrix. Under an iteration strategy, the sequence $\{X, E, Y, \mu\}$ are updated in turn, note that the order does not affect the optimization. Therefore, the Lagrangian (3.2) is separated into two sub-problems that are optimized:

$$E^{(t+1)} = \min_{E} \lambda \|E\|_1 + \frac{\mu^{(t)}}{2} \left\| D - X^{(t)} - E + \frac{1}{\mu^{(t)}} Y^{(t)} \right\|_F^2$$
$$X^{(t+1)} = \min_{X} \|X\|_* + \frac{\mu^{(t)}}{2} \left\| D - X - E^{(t+1)} + \frac{1}{\mu^{(t)}} Y^{(t)} \right\|_F^2$$

Note that (t) indicates the t^{th} iteration in optimization. To solve $E^{(t+1)}$, an element-wise soft-thresholding operator $\mathcal{S}_\tau[\cdot]$ [115] is introduced for sparsity identification in D, which is also used to introduce nonlinearity to the whole optimization procedure and defined as:

$$\mathcal{S}_\tau[X] = \min(X + \tau, 0) + \max(X - \tau, 0). \tag{3.3}$$

Hence, E can be obtained that $E^{(t+1)} = \mathcal{S}_{\lambda/\mu^{(t)}}\left[D - X^{(t)} + \frac{1}{\mu^{(t)}} Y^{(t)}\right]$ (cf. [115] for detailed explanation). With unconcerned variables $\{E^{(t+1)}, Y^{(t)}, \mu^{(t)}\}$ fixed, a Singular Value Shrinkage (SVS) operator $\mathcal{D}_\tau[\cdot]$ presented in [116] is adopted to update X that is defined as:

$$\mathcal{D}_\tau[X] = U \mathcal{S}_\tau[\Sigma] V^T, \quad X = U \Sigma V^T, \tag{3.4}$$

where $U\Sigma V^T$ is the singular value decomposition (SVD) of X. The motivation of $\mathcal{D}_\tau$ is shrinking the singular values with a pre-defined penalty τ to achieve dimensionality reduction of matrix X. Hence, X is updated that $X^{(t+1)} = \mathcal{D}_{1/\mu^{(t)}}\left[D - E^{(t+1)} + \frac{1}{\mu^{(t)}}Y^{(t)}\right]$. Afterwards, Y is updated with respect to X and E that $Y^{(t+1)} = Y^{(t)} + \mu^{(t)}\left(D - X^{(t+1)} - E^{(t+1)}\right)$. The procedure is repeated until the Lagrangian optimization converges that $\left\|D - X^{(t+1)} - E^{(t+1)}\right\|_F / \|D\|_F < \epsilon_R$, where ϵ_R is a pre-defined value for stop criterion. At this moment, the low-rank component $X_L = X^{(t+1)}$ in last iteration is considered as the expected clean subspace and further used for learning significant variabilities via PCA.

3.1.2 Weighted Robust PCA

In conventional RPCA, the low-rank matrix X is recovered automatically which leads to a loss of information that are non-outlier but not statistically significant in the datasets. Our assumption is that the prior knowledge can be obtained from the original clinical images to determine the possibility of being an outlier for each landmark. In a simple case, the probability of being an outlier for the parts that are not visible in the image is naturally 100%. However, our formulation allows all probabilities between 0 and 100% to describe areas with noisy or fuzzy boundaries. In order to account for these different probabilities and therefore to raise the accuracy of the low-rank matrix X, we propose a WRPCA method by introducing a weighting scheme in the solution of RPCA. Let $W = [w_1, ..., w_{N_s}]$ be the weighting matrix with the same dimensions as D; each column vector w_j is associated with the j^{th} training shape and matrix element w_{ij} denotes the weighting of the i^{th} landmark on the j^{th} shape, where w_{ij} is in the interval of $[0, 1]$. Fig. 3.2 shows the improvement of WRPCA in the construction of low-rank subspace by preserving landmarks with high probability.

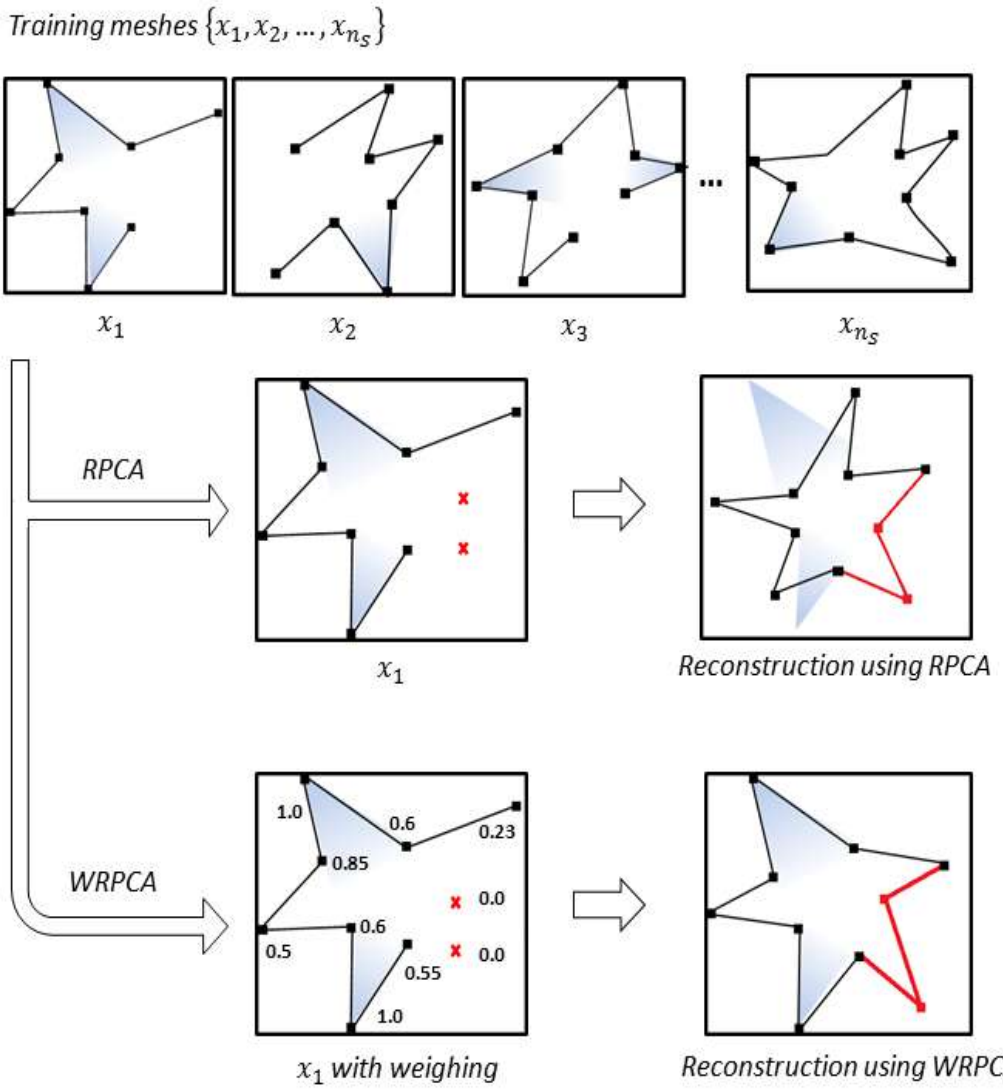


Figure 3.2: Difference between reconstructions using WRPCA and RPCA. Note that the shading indicates the areas of mesh that are approximately ground truth and the red bits are the reconstruction of missing parts.

Hereby, the objective function of WRPCA is formulated as:

$$\min_{X,E} \|X\|_* + \lambda \|W \circ E\|_1, \quad \text{s.t. } D = X + E, \tag{3.5}$$

where $\circ$ indicates the element-wise Hadamard product. Solved via IALM, the lagragian

is written as:

$$\mathcal{L}(X, E, Y, \mu) = \|X\|_* + \lambda \|W \circ E\|_1 + \langle Y, D - X - E \rangle + \frac{\mu}{2} \|D - X - E\|_F^2 , \tag{3.6}$$

An iteration strategy is adopted to minimize the Lagrangian; given a sequence $\{(X^{(t)}, E^{(t)}, Y^{(t)}, \mu^{(t)})\}$ of the t^{th} iteration, the Lagrange function is solved by the following subproblems to get the $(t+1)^{th}$ iteration:

$$X^{(t+1)} = \min_X \|X\|_* + \frac{\mu^{(t)}}{2} \left\| D - X - E^{(t)} + \frac{1}{\mu^{(t)}} Y^{(t)} \right\|_F^2 \tag{3.7a}$$

$$E^{(t+1)} = \min_E \lambda \|W \circ E\|_1 + \frac{\mu^{(t)}}{2} \left\| D - X^{(t+1)} - E + \frac{1}{\mu^{(t)}} Y^{(t)} \right\|_F^2 . \tag{3.7b}$$

Analogous to RPCA, $X^{(t+1)}$ can be efficiently solved via the singular value shrinkage operator $\mathcal{D}_\tau[X]$. Now, the weighting matrix W is introduced to turn the RPCA into a WRPCA formulation. $E^{(t+1)}$ is now computed as:

$$E^{(t+1)} = \mathcal{S}_{\lambda \frac{1}{\mu^{(t)}}}[(1 - W) \circ (D - X^{(t+1)} + \frac{1}{\mu^{(t)}} Y^{(t)})]. \tag{3.8}$$

The purpose of introducing the weighting matrix W is to influence the likelihood of being identified as an outlier for each landmark. Therefore, each entry $E_{ij}^{(t+1)}$ of $E^{(t+1)}$ in (3.7b) is computed as:

$$E_{ij}^{(t+1)} = \max((1 - w_{ij}) \cdot x'_{ij} - \frac{\lambda}{\mu^{(t)}}, 0) + \min((1 - w_{ij}) \cdot x'_{ij} + \frac{\lambda}{\mu^{(t)}}, 0), \tag{3.9}$$

where x'_{ij} is the matrix element in $X' = D - X^{(t+1)} + \frac{1}{\mu^{(t)}} Y^{(t)}$. That is if a weighting value w_{ij} of the landmark x_{ij} is closer to 1, $(1 - w_{ij})$ is closer to 0, i.e. the corresponding element in the sparse matrix $E_{ij}^{(t+1)}$ is more likely to be set to 0. A higher weighting value decreases the possibility to be identified as an outlier. Conversely, smaller weighting leads to a larger result of $E_{ij}^{(t+1)}$, indicating that the landmark x_{ij} is more likely to be an outlier. After separating the outliers and non-outliers in the t^{th} iteration, the multiplier $Y^{(t)}$ is updated in accordance with rules in RPCA. The process is repeated when the stop

criterion is satisfied, which is the same as that of RPCA. Subsequently, in order to build the SSM, PCA is applied to the output low-rank matrix X to capture the significant modes of variance of the input shapes.

Experimental results show a boost of performance of WRPCA compared to RPCA refer to our previous conference publication [27]. However, several limitations exist that WRPCA is apparently restricted to the prior knowledge, i.e., the probability of each landmark of being an outlier. Furthermore, WRPCA is not capable to handle nonlinear patterns that are often seen in realistic applications. This motivates us to put focus on the nonlinear SSMs in futher work.

3.1.3 KPCA for Statistical Shape Modeling

Linear techniques lack the capability to cope with nonlinear distribution. It is known that often the target shape has a complex geometry where linearity and nonlinearity exist simultaneously in one probabilistic distribution. Thereby, robustness to arbitrary outliers and fitness to nonlinear distribution are both desired properties for an ideal model. In literature, often linear matrix factorization approaches can be extended into a nonlinear distribution through nonlinear transformations, such as, kernel PCA (KPCA) [97, 98] is extended from PCA. This makes it feasible to derive a nonlinear SSM using KPCA.

In KPCA, a kernel trick κ is introduced to establish an implicit mapping Φ from the original shape space $D \in \mathcal{I}$ to a much higher dimensional feature space $D \rightarrow \Phi(D)$ where PCA is performed afterwards. Note that the popular Radial Basis Function (RBF) is adopted in this work as kernel trick. From the training data matrix D, a gram matrix $K = \mathcal{K}(D)$ is constructed with each element $K_{ij} = \kappa(d_i, d_j) = \langle \Phi(d_i), \Phi(d_j) \rangle$, where d_i denotes the i^{th} column vector in D. The gram matrix plays an important role as it represents the affinities between each pair of data point in the training set. Subsequently,

eigen-decomposition is performed on K to extract the first n_k dominant eigenvectors $\boldsymbol{\alpha}$ and eigenvalues Λ instead of carrying out the explicit eigen-decomposition on $\Phi(D)$. Each data point $d_i \in X$ is projected onto a set of nonlinear principal components $\boldsymbol{\beta_i} = \sum_{n=1}^{N_s} \boldsymbol{\alpha_n^{n_k}} \kappa(d_i, d_n)$ with dimension of n_k thereof. Mika et al. [97] presented a pre-image approach to inversely map the kernel space to the original data space by minimizing

$$\rho(z) = \left\| \Phi(z) - \mathbb{P}_n \Phi(D) \right\|^2, \tag{3.10}$$

where $\mathbb{P}_n \Phi(D)$ denotes projection from the feature space $\Phi(D)$ to the kernel space spanned by nonlinear principal components $\boldsymbol{\beta}$ and z is a shape vector approximated as the stationary point by setting the gradient in (3.10) zero:

$$\nabla_z \rho(z) = \sum_{i=1}^{N_s} \gamma_i \nabla_z \kappa(d_i, z), \quad \gamma_i = \sum_{k=1}^{n_k} \boldsymbol{\beta_k} \boldsymbol{\alpha_i^k}, \tag{3.11}$$

note that $\nabla_z \kappa(d_i, z)$ denotes the partial derivative of $\kappa(d_i, z)$ with respect to z.

3.2 KRPCA for Statistical Shape Modeling

3.2.1 Kernelized RPCA

Inspired by KPCA that performing PCA on the gram matrix, this thesis proposes a novel kernelized RPCA (KRPCA) to cope with data corruption in nonlinear distribution. The fundamental idea behind is to propagate a low-rank modeling as RPCA on the gram matrix to achieve dimensionality reduction, accordingly, reduce sparse outliers from the training data. Subsequently, the shape model is built based on the optimized gram matrix.

Analogous to RPCA in linear subspace, the implicit high-dimensional feature space $\Phi(D)$ is decomposed into two parts: $\Phi(D) = \Phi(X) + \Phi(E)$, where $\Phi(X)$ represents the feature space projection of a clean low-rank matrix $X \in \mathcal{I}$ and $\Phi(E)$ represents that of a

sparse matrix $E \in \mathcal{I}$. However, this complex system is infeasible to be directly solved due to the implicitness. The problem is reformulated thereof by considering the distributive property over matrix addition:

$$\Phi(D)^T\Phi(D) = \Phi(D)^T\Phi(X) + \Phi(D)^T\Phi(E)\ , \tag{3.12}$$

$K = \Phi(D)^T\Phi(D)$ be the constant kernel matrix where $K_{ij} = \Phi(d_i)^T\Phi(d_j) = \kappa(d_i, d_j)$. We define the matrix $\Phi(D)^T\Phi(X) = K_X$ that:

$$\begin{aligned}\Phi(D)^T\Phi(X) &= \begin{bmatrix}\Phi(d_1)^T\Phi(x_1) & \dots & \Phi(d_1)^T\Phi(x_n)\\ \vdots & \ddots & \vdots\\ \Phi(d_n)^T\Phi(x_1) & \dots & \Phi(d_n)^T\Phi(x_n)\end{bmatrix}\\ &= \begin{bmatrix}\kappa(d_1, x_1) & \dots & \kappa(d_1, x_n)\\ \vdots & \ddots & \vdots\\ \kappa(d_n, x_1) & \dots & \kappa(d_n, x_n)\end{bmatrix} = K_X,\end{aligned} \tag{3.13}$$

where $K_X \in \mathbb{R}^{n\times n}$. Similarly, $\Phi(D)^T\Phi(E) = K_E$ is defined and the decomposition problem in feature space is rewritten as:

$$\min_{K_X, K_E} \ \|K_X\|_* + \lambda \|K_E\|_1\ , \quad \text{s.t. } K = K_X + K_E. \tag{3.14}$$

In this manner, the low-rank model is applied on the kernel matrix K that determines the similarity of shapes in input space, in order to find the underlying clusters of similar shapes. The inexact augmented Lagrange multiplier (IALM) is applied to solve (3.14):

$$\mathcal{L}(K_X, K_E, Y, \mu) = \|K_X\|_* + \lambda \|K_E\|_1 + \langle Y, K - K_X - K_E\rangle + \frac{\mu}{2}\|K - K_X - K_E\|_F^2\ , \tag{3.15}$$

where Y decides the multiplier, μ is a positive parameter for adaptive penalization and λ is used to balance nuclear and l_1 norms. With an iteration strategy, $K_X^{(t+1)}$ and $K_E^{(t+1)}$ are obtained for the $(t+1)^{th}$ iteration.

3.2.1.1 Solving K_X

Fix the unconcerned variables $\{K_E, Y, \mu\}$, $K_X^{(t+1)}$ can be obtained by solving the sub-problem:

$$K_X^{(t+1)} = \min_{K_X} \|K_X\|_* + \frac{\mu^{(t)}}{2} \left\| K_X - \left(K - K_E^{(t)} + \frac{1}{\mu} Y^{(t)}\right) \right\|_F^2 , \tag{3.16}$$

the analytical solution to (3.16) is given below and the proof is provided in Lemma 3.1.

$$K_X^{(t+1)} = \mathcal{D}_{1/\mu^{(t)}} \left[syl\left(\mu^{(t)} I,\ \left((K_X^{(t)})^T K_X^{(t)}\right)^{-\frac{1}{2}},\ -\mu^{(t)}(K - K_E^{(t)} + \frac{1}{\mu^{(t)}} Y^{(t)})\right)\right]. \tag{3.17}$$

Lemma 3.1 *Let $F(X) = \|X\|_* + \theta \|X - H\|_F^2$, where θ and H are constant. The solution X^* can be given by deriving the subgradient of $F(X)$ and seeking its stationary point as $F(X)$ is convex. To reduce dimensionality, the Singular Value Shrinkage operator (3.4) $\mathcal{D}_\tau[X] = U diag(\Sigma_{ii} - \tau) V^T$ is leveraged to shrink the rank of X^*, where $U\Sigma V^T$ is the singular value decomposition of X.*

First we set the subgradient of $F(X)$ with respect to X zero and have:

$$\frac{\partial}{\partial X} F(X) = X(X^T X)^{-\frac{1}{2}} + 2\theta X - 2\theta H = 0, \tag{3.18}$$

obviously it is not feasible to get X^* and $(X^T X)^{-\frac{1}{2}}$ simultaneously, we consider $(X^T X)^{-\frac{1}{2}}$ as constant in computation of X^* and iteratively obtain X^* and $(X^T X)^{-\frac{1}{2}}$. Thus the problem of (3.18) is well known as Sylvester equation that $AX + XB + C = 0$. Here it is solved by the function syl from C++ Armadillo Library, that $\hat{X} = syl(A, B, C)$. By applying the shrinkage operator $\mathcal{D}_\tau$ to $\hat{X}$, we arrive at the solution:

$$X^* = \mathcal{D}_\tau[\ syl(\ 2\theta I, (X^T X)^{-\frac{1}{2}}, -2\theta H\ )\ ]\ . \tag{3.19}$$

3.2.1.2 Solving K_E

Fix the unconcerned variables $\{K_X, Y, \mu\}$, K_E^{t+1} is obtained by solving the subproblem:

$$K_E^{(t+1)} = \min_{K_E} \lambda \|K_E\|_1 + \frac{\mu^{(t)}}{2} \left\| K_E - \left(K - K_X^{(t+1)} + \frac{1}{\mu^{(t)}} Y^{(t)}\right) \right\|_F^2 , \tag{3.20}$$

the problem can be efficiently solved via the soft-thresholding operator $\mathcal{S}_\tau[X] = \max(X - \tau, 0) + \min(X + \tau, 0)$ in (3.3). Thus $K_E^{(t+1)}$ is defined as:

$$K_E^{(t+1)} = \mathcal{S}_{\lambda/\mu^{(t)}} \left[K - K_L^{(t+1)} + \frac{1}{\mu^{(t)}} Y^{(t)} \right]. \tag{3.21}$$

By updating $Y^{(t+1)} = Y^{(t)} + \mu^{(t)}(K - K_X^{(t+1)} - K_E^{(t+1)})$ and $\mu^{(t+1)} = \min(\rho\mu^{(t)}, \mu_{max})$, a new iteration is generated. The procedure converges to a point when $\left\| K_X^{(t+1)} - K_X^{(t)} \right\|_F \to 0$ and $\left\| K - K_X^{(t+1)} - K_E^{(t+1)} \right\|_2 < \epsilon$, where the optimal solution K_X^* to the optimization is reached.

3.2.2 Applying KRPCA to statistical shape modeling

To reduce dimensionality in feature space, K_X^* is leveraged to compute the first n_k eigenvectors $\boldsymbol{\alpha}$ and eigenvalues Λ by $K_X^* \boldsymbol{\alpha} = N_s \Lambda \boldsymbol{\alpha}$, a lower dimensional KPCA space is constructed thereof. For a sample $x \in \mathcal{I}$, nonlinear principal components with dimension of n_k in KPCA space $\boldsymbol{\beta}$ is computed by projecting x onto the selected eigenvectors that

$$\boldsymbol{\beta_i}(x) = \sum_{n=1}^{N_s} \boldsymbol{\alpha_n^i} \kappa(x_n, x), \quad i = 1 \ldots n_k \ . \tag{3.22}$$

To apply KPCA into SSM, Davies et al. gave a definition of probability density function $\mathcal{P}$ of KPCA model in [3]:

$$\mathcal{P}(x) \propto \sum_{i=1}^{n_k} \boldsymbol{\beta_i}(x)\boldsymbol{\beta_i}(x) = \sum_{a,b=1}^{N_s} \sum_{i=1}^{n_k} \boldsymbol{\alpha_a^i}\boldsymbol{\alpha_b^i} \kappa(x_a, x)\kappa(x_b, x), \tag{3.23}$$

representing the square of distance to the origin in KPCA space. As a result, the model captures nonlinear patterns by considering this "proximity to data" measure $\mathcal{P}$. Up to now, the KRPCA is built and the whole procedure is summarized in Algorithm 1.

Algorithm 1 Algorithm of Statistical Shape Modeling Using KRPCA

Input: Observation matrix $X \in \mathbb{R}^{3N_p \times N_s}$
Initialize: $Y = K_X = K_E = 0$, $K = \Phi(D)^T\Phi(D)$, $k = 0$

while not converge **do**
  Estimate $K_X^{(k+1)}$ by (3.17)
  Estimate $K_E^{(k+1)}$ by (3.21)
  Update $Y^{(k+1)}$, $\mu^{(k+1)}$
  $k \to k+1$
end while

Perform Eigen-Decomposition $K_X^* \boldsymbol{\alpha} = N_s \Lambda \boldsymbol{\alpha}$
Compute principal components in KPCA space via (3.22)
Get the PDF $\mathcal{P}$ for nonlinear statistical shape model

Output: the built KRPCA model

Usually it is trivial to back project parameter vector $\boldsymbol{\beta^i}(x)$ onto the input space, however, it is necessary to generate sample vectors in the input space for further evaluation. Thus we employ the reversed reconstruction approach in [117] to get the reconstructed shape $\hat{x} \in \mathcal{I}$ by:

$$\hat{x} = \frac{\sum_{i=1}^{N_s} \gamma_i \kappa(\hat{x}, x_i) x_i}{\sum_{i=1}^{N_s} \gamma_i \kappa(\hat{x}, x_i)} \ , \quad \gamma_i = \sum_{j=1}^{n_k} \boldsymbol{\beta_j} \boldsymbol{\alpha_i^j} \ . \tag{3.24}$$

3.3 Evaluation

The evaluation consists of three parts: (1) investigating the ability of outlier recovery of competitive methods **PCA**, **KPCA**, **KRLRR** [118] and our proposed KRPCA, namely, comparing the distance between the reconstructions and corresponding ground truth data; (2) evaluting the quality of SSMs built with competitive methods; (3) applying the

built models in ankle bone segmentation as the segmentation quality reflects the model efficiency.

3.3.1 Dataset

49 internal MRI ankle bone datasets are employed in evaluation, where each ankle bone consists of 9 discontinuous subparts with narrow inner space (cf. Fig. 3.1-(a)) to validate our approach on corrupted training data with a complex geometrical structure. The motivaton is that each bone follows normal distribution, however, the restricted inner space among independent bones brings nonlinearity to the whole. The volume size varies from $384 \times 384 \times 147$ to $512 \times 512 \times 139$ and the voxel size ranges from $0.46 \times 0.46 \times 0.50$ mm^3 to $0.49 \times 0.49 \times 0.50$ mm^3. Patients were 20 females and 29 males with age interval of 13 to 17 years. For resolution consistency, all the MRI ankle images as well as ground truth are rescaled to $512 \times 512 \times 128$. Afterwards, we obtain training shapes by extracting polygon meshes from the binary ground truth through Marching Cubes.

As it is known that a crucial step before modeling is creating a groupwise correspondence among all training shapes for statistics analysis. We employ the existing groupwise point-to-point correspondence approach presented in [119], in order to assign the same order of landmarks to areas with similar features (from Fig. 3.3-(a) to (b), the color represents the ID of landmark). It automatically computes a representation of training shapes as a vector of corresponding landmarks, with the same index representing the same feature over the whole training set. Specifically, the algorithm firstly computes an area-preserving parameterization $\mathcal{P}_{ref}$ of a single reference shape $\mathcal{M}_{ref}$ which is randomly chosen from the training set. Then, it computes an initial mapping of $\mathcal{M}_{ref}$ to the unit sphere and optimize $\mathcal{P}_{ref}$. Afterwards, this parameterization $\mathcal{P}_{ref}$ is propagated to all other shapes (cf. [119] for a detailed description). Generally, it generates an optimal

number of landmarks in terms of the boundary features, i.e. a large number of landmarks are generated for the boundary with complex variabilities and vice versa. As a consequence, the generated landmarks represent dominant boundary features. The correspondence is built with each subbone respectively, e.g. Fibula and Tibia. As a result, the whole ankle bone is merged with the bones with correspondence consisting of 5148 landmarks in total. Subsequently, the training shapes with correspondence are centred at the origin and aligned via rigid Procrustes alignment.

To make the ankle shapes arbitrarily corrupted, we create an SSM with 5 ground truth datasets via PCA and embed it into an existing segmentation framework [28] to re-segment the 49 datasets. Due to the limitation of this segmentation procedure, the segmented shapes may contain corrupted regions to different extent, e.g, Fig. 3.3-(a) shows an example of corrupted training shape where bones are abnormally overlapped. Hereafter, the ankle bone SSMs are all derived based on the generated 49 shapes $X = \{x_i : i = 1 \cdots N_s\}$ with arbitrary corruptions, while the ground truth datasets $Y = \{y_i : i = 1 \cdots N_s\}$ are used for evaluation.

3.3.2 Parameters

Even though a number of parameters and variables are leveraged in our technique, most of them are tuned related to the training data following certain rules. Table 3.2 shows the parameters and descriptions that are used in KRPCA. The trade-off λ, $Y^{(0)}$ (initialization for the multiplier Y), μ and the update rate η are all chosen according to the work [115], where $\|K\|_2$ decides the l_2 norm of matrix K and $\|K\|_\infty$ denotes the maximum absolute value of the matrix entries.

It remains to discuss the kernel width σ in RBF trick that has a vital effect on the underlying degree of model. Specifically, a large width reduces the difference between

Table 3.2: Parameters and Descriptions

Parameter	Description	Value
λ	trade-off in (3.14)	$\frac{1}{\sqrt{N_s}}$
$Y^{(0)}$	the initial Lagrange multiplier	$\frac{K}{\min\left(\|K\|_2, \lambda^{-1}\|K\|_\infty\right)}$
η	update rate of μ and μ_k	1.6
$\mu^{(0)}$	the initial value of μ	$1.25/\|K\|_2$
$\bar{\mu}$	maximum of μ	10^5
ϵ	terminate value	10^{-6}

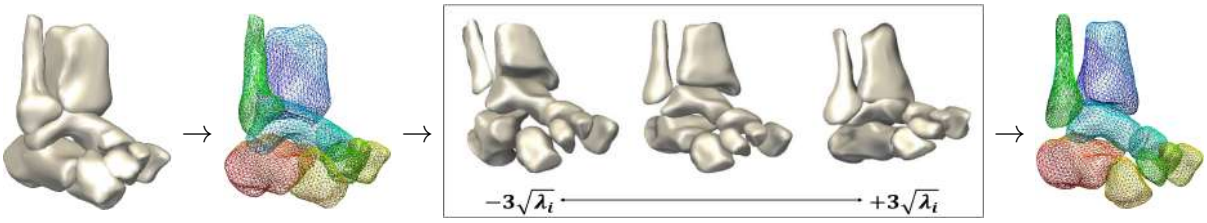
any two data points leading to a compact feature space, while a small width enlarges this difference leading to a model with large variabilities. Essentially, a small sigma preserves the features while a large sigma is a better choice for larger dimensionality reduction. To keep consistency with other works, the kernel width σ is computed by:

$$\sigma = \frac{1}{2N_s^2} \sum_{i=1}^{N_s} \sum_{j=1}^{N_s} \delta(x_i, x_j), \tag{3.25}$$

where $\delta(x_i, x_j) = \|x_i - x_j\|$ is the average symmetric distance (ASD) between the shape x_i and x_j.

3.3.3 Ability of Outlier Recovery

Fig. 3.3 illustrates the pipeline of modeling procedure as well as the model back projection. To evaluate the robustness to outliers of the competing models, the corrupted training shapes are projected back onto the model (from Fig. 3.3-(b) to (c)) to generate the corrected reconstructions (Fig. 3.3-(d)). We define the back projection error by computing the mean absolute distance MAD between the ground truth $Y = \{y_i : i = 1 \cdots N_s\}$ and reconstructions $\widehat{X} = \{\hat{x}_i : i = 1 \cdots N_s\}$ as $MAD = 1/N_s \sum_{i=1}^{N_s} \delta(\tilde{x}_i, \hat{x}_i)$. Table 3.3 shows the MAD results, where the row "Training Datasets" is MAD between training shapes and ground truth that indicates the degree of corruption of training shapes.



(a) (b) (c) (d)

Figure 3.3: The figure illustrates the flow chart of modeling and back projection. (a) indicates the training mesh without landmark correspondence assigned; (b) indicates that the shape has been assigned with a groupwise correspondnece, where the color indiates the landmark ID; (c) indicates an SSM with the mean shape in the middle. The procedure from (b) to (c) means the model back-projection, and (d) is the reconstruction of the shape (b). It can be found that the abnormal overlap in (b) disappears in (d) with the correction by the model (c).

Table 3.3: Back projection error MAD

	Mean (mm)	Min (mm)	Max (mm)
Training Datasets	7.77 ± 3.51	2.66	17.03
PCA	7.72 ± 3.45	2.72	16.87
KPCA	7.17 ± 3.35	2.41	16.78
KRLRR	9.20 ± 6.67	2.56	25.98
KRPCA (Ours)	**6.68 ± 3.27**	**2.39**	**16.10**

Conclusions can be drawn from Table 3.3. (1) As the training shapes are generated from the exsiting segmentation approach, the training shapes contain a low degree of corruption, i.e. the average distance between each training shape and its corresponding ground truth is only $7.77mm$ (cf. the first row in Table 3.3). (2) In terms of the back projection errors of PCA and KPCA, it can be found that PCA and KCPA nearly do not have the ability of outlier recovery. (3) Apparently the reconstructions from KRLRR model have more erroneous than the original training shapes as LRR-like approaches are often assumed a large number of training data. (4) Finally, our method reduces the erroneous in terms of its smallest back projection error.

3.3.4 Model Quality Evaluation

To evaluate the quality of SSMs, the most common measures to quantize the quality of SSMs are Generalization ability G, Specificity S and Compactness C provided in [3]. A model with good generalization is able to represent not only the trained shapes, but also deviations. In contrast to a general model, a specific model represents shapes close to the training data. Thus it does not cover large variances. In addition, a model is compact in case it is capable to describe the model's probability density function (PDF) by a small number of variances. Given a set of N_s training datasets $\{x_i : i = 1 \cdots N_s\}$, we denote a set of M shapes $\{y_A : A = 1 \cdots M\}$ randomly generated from the model's PDF, which approximately covers the whole PDF of the model when $M \to \infty$. G, S and C are defined as:

$$\begin{aligned} G_m &= \frac{1}{N_s} \sum_{i=1}^{N_s} \delta(x_i, \tilde{x_i}), \\ S_m &= \frac{1}{M} \sum_{A=1}^{M} \min_i \delta(y_A, x_i), \\ C_m &= \sum_{i=1}^{m} \Lambda_i / \sum_{i=1}^{n_k} \Lambda_i \times 100\%. \end{aligned} \tag{3.26}$$

where m is the number of modes used to generate the samples y_A from the model's PDF. Upon a leave-one-out mode that excluding each of training shapes x_i in turn and then project it onto the model derived with the remaining datasets, G is measured as the distance between x_i and its reconstruction $\hat{x}_i$ thereof. Lower value of G implies higher ability to validate the unseen data, namely, a better performance of generalization ability. On the other hand, S measures the distance from all generated samples to the closest training data. Smaller values of S indicate the model is more specific, i.e. a higher quality model. Compactness measures the relative variance learned by the dominant modes, i.e. higher ratio indicates better compactness.

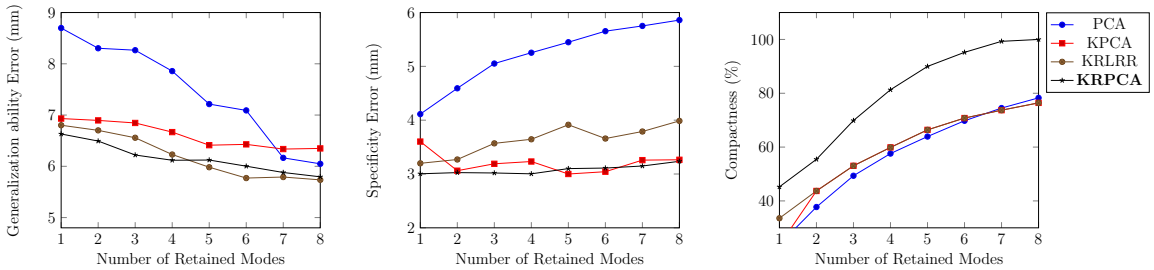


Figure 3.4: Generalization ability and Specificity for the models built with PCA, KPCA, KRLRR and KRPCA. Smaller value indicates better result.

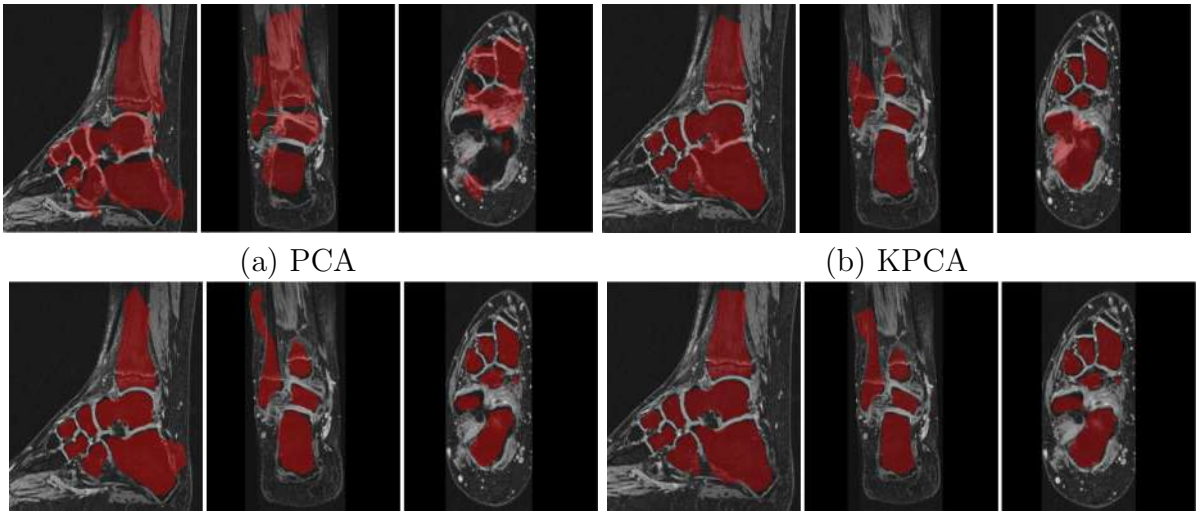
Fig. 3.4 reports the results for Generalization ability and Specificity for the first 8 modes of all competing models. It can be found that our model has the smallest generalization ability and specificity error for the dominant modes, and the largest compactness compared to the others. That is to say, our model has the best quality in terms of generalization ability, specificity and compactness.

3.3.5 Application in Ankle Bone Segmentation

Even though the shape of ankle has a zero mean, it is still challenging for the SSM to learn the exact variation of each bone. As a result, often the segmented bones are abnormally overlapped, which motivates us to create a nonlinear model. We apply the built models to an existing segmentation approach [28] based on conventional SSM. The accuracy of segmentation is measured by Hausdorff distance, Dice coefficient and overlap volume percentage (compare Table 3.4). For an intuitive view, Fig. 3.5 shows the comparison of segmentation results from the PCA model, KPCA model, KRLRR model and our KRPCA model.

In summary, our proposed KRPCA model has a higher quality in terms of all the measurements; KRLRR and KPCA both perform better than PCA model, although KRLRR has a larger back projection error and Hausdorff distance than PCA. However,

the large standard deviation explains the larger error, which means, KRLRR is not robust enough in conditions of a limited subset of training data. This is also the strong motivation for us to kernelize RPCA rather than LRR.



(a) PCA (b) KPCA

(c) KRLRR (d) **KRPCA**

Figure 3.5: Figure shows the comparison of segmentation results of competing models in Sagittal, Coronal and Axial position respectively.

3.4 Discussion

In this chapter, we aim to create SSMs what are robust to abnormalities in training data and satisfied with nonlinear distribution at the same time. Motivated by this, a novel kernelized RPCA approach is proposed by performing RPCA on the nonlinear kernel space obtained via KPCA. Evaluation results show that the model built with KRPCA has a better quality compared to the competed models in terms of generalization ability, specificity, compactness and the segmentation accuracies. Moverover, KRPCA shows a better performace on outlier recovery.

The proposed KRPCA shows the feasibility to handle the nonlinear sparsity through the kernel matrix compression, which offers a solid basis. However, it still remains to

evaluate the ability of gross outlier recovery, as the experimental results only show that the proposed KRPCA handles slight corruption in training data. Furthermore, it seeks the primary affinities across the training data and filters out the sparseo ones through performing RPCA on the gram matrix. It can be expected that, the proposed KRPCA may lose the benefits with smaller number of training datasets. Therefore, it is promising to artificially generate training datasets with large missing or corrupted areas in evaluation in the future work.

Table 3.4: Segmentation results for reference PCA model and competing models. Note that smaller Hausdorff distance indicates better result; for Dice coefficient and overlapping volume percentage, larger value represents better result.

	Hausdorff Distance (mm)			Dice Coefficient			Volume Overlap (%)		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
PCA	7.39±3.30	3.22	15.00	0.86±0.10	0.60	0.93	77.94±12.90	43.02	88.62
KPCA	8.13±3.59	3.64	14.41	0.87±0.07	0.76	0.93	82.38±10.99	61.42	87.43
KRLRR	10.32±5.12	3.22	20.99	0.90±0.03	0.87	0.95	82.38 ± 4.32	76.94	90.46
KRPCA	**7.28±2.93**	**2.36**	**11.05**	**0.91±0.02**	**0.88**	**0.96**	**83.67±4.31**	**79.22**	**92.29**

Chapter 4

A Novel Robust Kernel Principal Component Analysis for Nonlinear Statistical Shape Modeling from Erroneous Data

Statistical Shape Models (SSMs) have achieved considerable success in medical image segmentation. A high quality SSM is able to approximate the main plausible variances of a given anatomical structure to guide segmentation. However, it is technically challenging to derive such a quality model because: (1) the distribution of shape variance is often nonlinear or multi-modal which cannot be modeled by standard approaches assuming Gaussian distribution; (2) as the quality of annotations in training data usually varies, heavy corruption will degrade the quality of the model as a whole.

In this chapter, these challenges are addressed by introducing a generic SSM that is able to model nonlinear distribution and is robust to outliers in training data. Without losing generality and assuming a sparsity in nonlinear distribution, a novel Robust Kernel Principal Component Analysis (RKPCA) for statistical shape modeling is proposed with the aim of constructing a low-rank nonlinear subspace where outliers are discarded, which is the major contribution of this chapter. The proposed approach is validated on two

different datasets: a set of 30 public CT kidney pairs and a set of 49 MRI ankle bones volumes. Experimental results demonstrate a significantly better performance on outlier recovery and a higher quality of the proposed model as well as lower segmentation errors compared to the state-of-the-art techniques.

The remaining of this chapter is organized as follows: first of all, an overview of nonlinear statistical shape modeling approaches is provided, as well as the techniques of nonlinear SVM, LRR and RPCA. Secondly, a detailed explanation and derivation of our proposed RKPCA is illustrated in detail. Afterwards, rich experiments are conducted in order to evaluate the performance of RKPCA. It should be pointed out that partial content in this chapter comes from our publication "A Novel Robust Kernel Principal Component Analysis for Statistical Shape Modeling from Erroneous Data", which is under review of journal Medical Image Analysis.

4.1 Introduction

Among a variety of applications using SSMs, the number of approaches exploiting nonlinear variabilities of shape data is relatively small in contrast to linear models. The majority of approaches deal with complex geometry by incorporating other machine learning techniques and/or utilizing integrated shape and pose models under the assumption of a Gaussian-like distribution. In [90], KPCA is leveraged to capture nonlinear variance of SSMs for vertebrae; KPCA is also applied to capture nonlinear patterns of left ventricle in [83]. Without using PDM, a recent modeling approach is proposed in [96], which endows a differential representation of shape with a nonlinear Riemannian structure for identification of intra- and inter- population variability. Many biological structures, however, contain nonlinearities even though the non-linearity might not be always strong.

Based on the first attempt of performing RPCA on the gram matrix, we aim to improve its generality and presents a generic Robust Kernel Principal Component Analysis (RKPCA) technique for nonlinear subspace compression. As aforementioned, often the matrix factorization in linear space can be extended to solve nonlinear distribution through some operations, of which the kernel trick is the most popular way, e.g. KPCA [97, 98], kernel SVM [120] and kernel compressive sensing [121, 122]. However, such nonlinear techniques do not address the gross noise in training data. Efforts have been made to make KPCA robust for handling noise in the work [123]. Regarding the robust compressive sensing techniques like LRR and RPCA, which decompose the data matrix using a nonlinear thresholding, such extension ostensibly brings mathematical difficulty to identify and separate the nonlinear sparsity. A number of works on kernelizing LRR have been proposed [124, 118]. As discussed in the last chapter, LRR does not work well when the training database is relatively small, namely, the "dictionary" lacks generality. Motivated by this, this thesis aims to propose a generic nonlinear subspace recovery based on the motivation of matrix decomposition from RPCA.

Addressing these challenges, in this chapter, a generic SSM framework is proposed that can handle multimodal distribution and at the same time is robust to corrupted data. Another advantage of the proposed model in practice is that its robustness to outliers allows for a greater use of non-ideal training data generation methods for statistical shape modeling. For example, semi-automatic segmentation approaches could be used to a greater extent than currently possible to augment the training data base as the proposed SSM framework can still create high quality models in the presence of errors.

4.2 Robust Kernel PCA

Assuming a sparsity in nonlinear data probability density, we hypothesize that a latent low-rank nonlinear subspace can be recovered from corrupted training data and a compact

shape model is derived from the compressed clean subspace afterwards. This motivates us to perform RPCA on a nonlinear data distribution where the dominant patterns are captured by KPCA. In the previous conference work [95] presented at MICCAI 2017 (cf. 3), a kernelized RPCA is proposed for creating SSMs which outperforms linear models in coping with corruption in nonlinear data. For ease of presentation, the kernelized RPCA is detonated as **MICCAI-17**. In this chapter, the previous work is expanded addressing a novel nonlinear data decomposition technique with complete mathematical derivations and explanations.

In this section, a detailed explanation and derivation of RKPCA is illustrated. Fig. 4.1 plots the pipeline of RKPCA for modeling, as well as the model back-projection procedure.

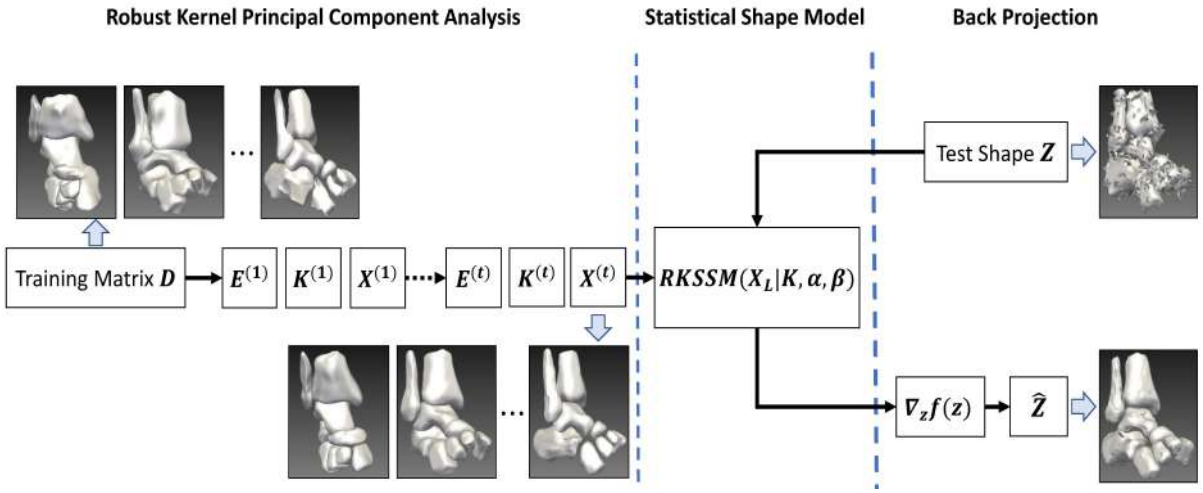


Figure 4.1: This figure illustrates the pipeline of our RKPCA and the procedure of model back projection. Based on a set of corrupted shapes, the training matrix D can be constructed. : when an unseen shape $\boldsymbol{Z}$ is back projected onto the model, a set of nonlinear principal components are propagated and further used to approximate the reconstruction $\hat{z}$ through the standard pre-image $f(z)$.

4.2.1 Objective formulation

To seek a low-rank component X that is free of outliers and represents primary nonlinear patterns, a straight forward solution is to decompose the high-dimensional feature space by minimizing $\|\Phi(X)\|_* + \lambda \|\Phi(E)\|_1$, which is in favor of the arguments in RPCA (cf. (3.1)). However, the implicitness of Φ makes it infeasible to determine the stationary point of $\|\Phi(E)\|_1$, as the l_1 norm is an element-wise nonlinear activation. To address this issue, a straightforward way would be minimizing $\|\Phi(X)\|_* + \lambda \|E\|_1$. However, since $\Phi(X)^T\Phi(X) = \mathcal{K}(X)$, $\|\mathcal{K}(X)\|_*$ can be considered to act as surrogate of $\|\Phi(X)\|_*$ because minimizing $\|\Phi(X)\|_*$ is equivalent to minimizing $tr\big((\Phi(X)^T\Phi(X))^{1/2}\big) = tr\big(\mathcal{K}(X)^{1/2}\big)$. Moreover, seeking the optimal $\|\mathcal{K}(X)\|_*$ alleviates computational complexity and its feasibility has been demonstrated in the previous work 3. A similar idea of minimizing the rank of the gram matrix is also presented in [125, 126].

Therefore, we aim to seek the best rank of $\mathcal{K}(X)$ instead of explicitly carrying out the unknown dimensional feature space $\Phi(X)$ and present the objective function:

$$\min_{X,E} \ \|\mathcal{K}(X)\|_* + \lambda \|E\|_1 , \quad \text{s.t. } X + E = D. \tag{4.1}$$

IALM is employed to form the objective Lagrangian:

$$\begin{aligned} \mathcal{L}(X, E, Y, \mu) = \min_{X,E} \|\mathcal{K}(X)\|_* + \lambda \|E\|_1 + \\ \langle Y, D - X - E\rangle + \frac{\mu}{2} \|D - X - E\|_F^2 , \end{aligned} \tag{4.2}$$

which is solved under an iteration scheme.

4.2.2 Updating $E^{(t+1)}$

With the unconcerned variables $\{X, Y, \mu\}$ fixed, the Lagrangian $\min_E \mathcal{L}(E, Y, \mu)$ with respect to the sparse component E is written as:

$$\min_E \lambda \|E\|_1 + \frac{\mu}{2} \left\| E - \left(D - X^{(t)} + \frac{1}{\mu} Y \right) \right\|_F^2 ,$$

and E is updated by introducing the soft-thresholding operator (3.3) that:

$$E^{(t+1)} = \mathcal{S}_{\frac{\lambda}{\mu^{(t)}}}\left[D - X^{(t)} + \frac{1}{\mu^{(t)}}Y^{(t)}\right], \tag{4.3}$$

where t indicates the iteration that starts from 0.

4.2.3 Updating $K^{(t+1)}$

We arrive at optimization to X based on the updated $E^{(t+1)}$, the Lagrangian $\mathcal{L}(X, Y, \mu)$ with respect to X is given by:

$$\min_X \|\mathcal{K}(X)\|_* + \frac{\mu}{2}\left\|X - \left(D - E^{(t+1)} + \frac{1}{\mu}Y\right)\right\|_F^2. \tag{4.4}$$

It is infeasible to directly apply the singular value shrinkage $\mathcal{D}_\tau[\cdot]$ to the Lagrangian, which brings difficulty to giving the optimal $X^{(t+1)}$. Hence, we decompose the optimization (4.4) into two subproblems with respect to X at each iteration. Specifically, we consider X a constant to seek the optimal rank of the gram matrix $\mathcal{K}(X)$. We subsequently derive $X^{(t+1)}$ with the obtained $\boldsymbol{K^{(t+1)}}$ by reaching the sub-stationary point:

$$\begin{cases} \boldsymbol{K^{(t+1)}} = \min\limits_{\mathcal{K}(X)} \|\mathcal{K}(X)\|_* \\ \boldsymbol{X^{(t+1)}} = \min\limits_X \mathcal{K}(X) + \frac{\mu}{2}\left\|X - \left(D - E^{(t+1)} + \frac{1}{\mu}Y\right)\right\|_F^2 \end{cases}$$

The Lagrangian (4.4) is separable because the nuclear norm and frobenius norm are both convex.

To achieve dimensionality reduction for outlier removal, at this stage, we apply the low-rank modeling that is introduced in RPCA to the gram matrix and the objective function is formulated as:

$$\min_K \left\|\boldsymbol{K^{(t)}}\right\|_*, \quad \text{s.t. } \boldsymbol{K^{(t)}} = \mathcal{K}\left(D - E^{(t+1)}\right). \tag{4.5}$$

To be mentioned, it is a special case of matrix decomposition that is called matrix completion in [116]. Associated with a Lagrange multiplier A, the low-rankness of gram

matrix is promoted iteratively by

$$\begin{cases} K^{new} = \mathcal{D}_{\mu_k}\left[\boldsymbol{K^{(t)}} + \frac{1}{\mu_k}A^{old}\right] \\ A^{new} = A^{old} + \mu_k\Big(\boldsymbol{K^{(t)}} - K^{new}\Big) \end{cases} \tag{4.6}$$

where $\{K^{new}, A^{new}\}$ are optimal solutions in each iteration and initialized as $K^0 = \mathcal{K}(D - E^{(t+1)})$ and $A^0 = 0$ respectively. It should be pointed out that the sought of low-rank $\boldsymbol{K^{(t)}}$ is an inner loop embedding in the whole RKPCA optimization, hereby, the extra superscript implying the iteration is dropped to avoid confusion in presentation. The procedure of low-rank modeling will converge to an accumulation point where $\left\|\boldsymbol{K^{(t)}} - K^{new}\right\|_F / \left\|\boldsymbol{K^{(t)}}\right\|_F \leq \epsilon_k$, a theoretical guarantee is provided in [106]. Until the convergence condition for the inner loop is satisfied, we simply take the final output and denote $\boldsymbol{K^{(t+1)}} = K^{new}$, which may not be symmetrical and full-rank. In this manner, $\|\mathcal{K}(X)\|_*$ is solved and the updated $\boldsymbol{K^{(t+1)}}$ is then applied to derive $X^{(t+1)}$.

4.2.4 Updating $X^{(t+1)}$

For the ease of presentation, we denote the constant $C = \mu/2\big(D - E^{(t+1)} + 1/\mu Y\big)$. Denote $\rho(X) = \mathcal{K}(X) + \mu/2\,\|X - C\|_F^2$, $X^{(t+1)}$ is optimized by deriving the subgradient of $\rho(X)$ and assumed to be the stationary point. The subgradient of $\rho(X)$ is defined as:

$$\frac{\partial}{\partial X}\rho(X) = \frac{\partial}{\partial X}\mathcal{K}(X) + \frac{\partial}{\partial X}\frac{\mu}{2}\|X - C\|_F^2\,. \tag{4.7}$$

The partial derivative of frobenius norm is easily obtained by:

$$\frac{\partial}{\partial X}\frac{\mu}{2}\|X - C\|_F^2 = \mu\big(X - C\big). \tag{4.8}$$

Subsequently, we seek the partial derivative of $\mathcal{K}(X)$ with respect to X. As opposed to the standard pre-image in (3.10) which also aims to approximate the projections from gram matrix to the input space, we leverage the value of low-rank $\boldsymbol{K^{(t+1)}}$ straightforwardly, implying the new affinities between data point, instead of using any principal

components derived from eigen-decomposition. Theoretically, computing the derivative of Matrix-by-Matrix can be accomplished by vectorizing the denominator matrix so that the Kronecker product will be leveraged to approximate the derivatives of Matrix-by-Scalar. However, this will raise the computational complexity on account of the large dimensionality in the derivative matrix. Moreover, the Kronecker product might cause high sparsity because each element K_{ij} in the gram matrix is only related to x_i and x_j but not to all the members. Addressing this issue, we present a novel solution to the partial derivative of $\partial_X \mathcal{K}(X)$ that is defined as:

$$\frac{\partial}{\partial X}\mathcal{K}(X) = \begin{bmatrix} \sum_{i=1}^{N_s} \frac{\partial \kappa(x_i, x_1)}{\partial X_{11}} & \cdots & \sum_{i=1}^{N_s} \frac{\partial \kappa(x_i, x_{N_s})}{\partial X_{1N_s}} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{N_s} \frac{\partial \kappa(x_i, x_1)}{\partial X_{3N_p1}} & \cdots & \sum_{i=1}^{N_s} \frac{\partial \kappa(x_i, x_{N_s})}{\partial X_{3N_pN_s}} \end{bmatrix} \tag{4.9}$$

More specifically, we sum up all the derivatives related to X_{ij} in order to avoid the sparsity caused by the Kronecker product.

Substituting (4.8) and (4.9) into (4.7) and set it to zero, thus we have:

$$\sum_{n=1}^{N_s} \boldsymbol{K}_{\boldsymbol{nj}}^{(\boldsymbol{t+1})} X_j^{(t)} - \sum_{n=1}^{N_s} \boldsymbol{K}_{\boldsymbol{nj}}^{(\boldsymbol{t+1})} + \mu X_j^{(t)} - \mu C_j = 0, \tag{4.10}$$

where j indicates each column in X and C, namely, $X^{(t+1)}$ is computed in column-by-column. Therefore, we come to the final solution:

$$X^{(t+1)} = \sum_{j=1}^{N_s} \frac{\mu C + \sum_{n=1}^{N_s} \boldsymbol{K}_{\boldsymbol{nj}}^{(\boldsymbol{t+1})} X_j}{\mu + \sum_{n=1}^{N_s} \boldsymbol{K}_{\boldsymbol{nj}}^{(\boldsymbol{t+1})}}. \tag{4.11}$$

4.2.5 Convergence Analysis

Based on the updated $X^{(t+1)}$ and $E^{(t+1)}$, we update the multiplier $Y^{(t+1)} = Y^{(t)} + \mu^{(t)}\big(D - X^{(t+1)} - E^{(t+1)}\big)$. Positive scalars μ and μ_k are updated exactly in the same way that

$\mu^{(t+1)} = \min\left(\mu^{(t)}(1+\eta), \bar{\mu}\right)$, and $\mu_k^{new} = \min\left(\mu_k^{old}(1+\eta), \bar{\mu_k}\right)$. Analogous to the inner loop (4.6), sequence $\{X, E, Y, \mu\}$ converges when $\left\|D - X^{(t+1)} - E^{(t+1)}\right\|_F / \|D\|_F \leq \epsilon$. In this way, a low-rank component X_L is recovered from the original data D where outliers are discarded and KPCA is performed to extract nonlinear principal components $\boldsymbol{\beta}$ and corresponding eigenvectors $\boldsymbol{\alpha}$ for statistical shape modeling.

We summarize the whole procedure of statistical shape modeling using the proposed RKPCA in Alg. 2.

Algorithm 2 Nonlinear statistical shape modeling via RKPCA

Input: training matrix D.
Initialize: $Y^{(0)}, \mu^{(0)}, \mu_k^{(0)}, t = 0$.

while not converged **do**
  Update $E^{(t+1)}$ via (4.3)
  Update $\boldsymbol{K^{(t)}} \leftarrow \mathcal{K}\left(D - E^{(t+1)}\right)$
  Initialize: $A^{(0)} = 0, K^{(0)} = \boldsymbol{K^{(t)}}, iter = 0$.

  while not converged **do**
    Update $\{K^{iter}, A^{iter}, \mu_k^{iter}\}$ via (4.6)
    $iter = iter + 1$
  end while

  Output: $\boldsymbol{K^{(t+1)}}$
  Update $X^{(t+1)}$ via (4.11)
  Update $Y^{(t+1)}, \mu^{(t+1)}$
  $t = t + 1$
end while

Perform KPCA on the recovered low-rank matrix X_L.
Output: $RKSSM(X_L|K, \boldsymbol{\alpha}, \boldsymbol{\beta})$.

4.3 Evaluation

Our method is evaluated on two representative datasets involving a set of kidney pairs and a set of ankle bones with arbitrary outliers. The results are compared with five closely per-

tinent approaches: conventional **PCA**, **RPCA** [5], **KPCA** [97], **NIPS-09** (the robust KPCA approach [127]) and our previous conference work KRPCA 3. To avoid confusion, we denote KRPCA as **MICCAI-17** for ease of presentation, as it is published at MICCAI 2017. Even though KPCA is not bounded by pre-image approaches, we choose the popular pre-image strategy presented by Mika et al. [97] in this work. In [127], Nguyen et al. present a Robust KPCA approach in order to make the KPCA pre-image robust to outliers. In contrast to the objective in (3.10), the reconstruction $\Phi(z)$ is required to be close to $\mathbb{P}_n\Phi(z)$ as well as the input sample with a balance trade-off constant ω, of which the objective is formed as $\min_z \|\Phi(x) - \Phi(z)\|_2^2 + \omega \|\Phi(z) - \mathbb{P}_n\Phi(z)\|_2^2$.

First, experiments are conducted to evaluate the ability of dealing with arbitrarily corrupted data from various degrees of competitive methods. Furthermore, we create SSMs for all these methods and evaluate the quality of the resulting models as well as the accuracy of segmentation using these SSMs.

4.3.1 Datasets

To investigate the ability for representing the nonlinear patterns of the competitive approaches, we create a kidney SSM with each pair consisting of the left and right kidney (cf. Fig. 4.7). Despite that each single kidney (left/right) follows a Gaussian distribution, the restricted inner space between two single kidneys brings nonlinearity to the whole, which is exactly the reason for choosing the data. 30 public CT datasets are collected from MICCAI 2015 Challenge Multi-Atlas [128], with volume sizes varying from $512 \times 512 \times 85$ to $512 \times 512 \times 198$, the slice thickness varying from 2.5 to 5.0 mm and in-plane resolution varying from 0.54 mm to 0.98 mm. Moreover, we utilize the 49 internal ankle bone datasets as aforementioned in chapter 3 consisting of 49 ground truth meshes and corrupted meshes with arbitrary outliers.

For resolution consistency, kidney images with ground truth are rescaled to iso-cubic volumes with dimension $256 \times 256 \times 256$ and the MRI ankle images as well as ground truth are rescaled to $512 \times 512 \times 128$. Then, we obtain training shapes by extracting polygon meshes from the binary ground truth through Marching Cubes. As it is known that a crucial step before modeling is creating a groupwise correspondence among all training shapes for statistics analysis. We employ the point-to-point correspondence approach presented in [119], in order to assign the same order of landmarks to areas with similar features. For the kidney datasets, we establish correspondence to the left and right kidneys respectively and merge both afterwards. As a result, each kidney pair contains 4000 landmarks.

4.3.2 Parameters

Generally speaking, the low-rank modeling techniques have compelling advantages of being light weight and trained unsupervised compared to the recent trend deep neural networks. Even though a number of parameters and variables are leveraged in our technique, most of them are tuned related to the training data following certain rules. Table 4.1 shows the parameters and descriptions that are used in RKPCA. It should be pointed out that the trade-off λ, initialization for multiplier Y, μ, μ_k, and the update rate η are all chosen according to the work [115], where $\|D\|_2$ decides the l_2 norm of matrix D and $\|D\|_\infty$ denotes the maximum absolute value of the matrix entries. The chosen of sigma follows the rule in (3.25).

4.3.3 Computational Complexity Analysis

Regarding the computational complexity of the proposed RKPCA, we denote $m = 3N_p$ and $n = N_s$ as the dimension of training matrix for ease of presentation. In each iteration, the main computational cost comes from (1) computation of gram matrix with $\mathcal{O}(mn^2)$

Table 4.1: Parameters and Descriptions

Parameter	Description	Value
λ	trade-off in (4.1)	$\frac{1}{\sqrt{\max(N_s, 3N_p)}}$
$Y^{(0)}$	the initial Lagrange multiplier	$\frac{D}{\min\left(\|D\|_2, \lambda^{-1}\|D\|_\infty\right)}$
η	update rate of μ and μ_k	1.6
$\mu^{(0)}$	the initial value of μ	$1.25/\|D\|_2$
$\bar{\mu}$	maximum of μ	10^5
ϵ	terminate value	10^{-6}
$\mu_k^{(0)}$	the initial value of μ_k	$1.25/\|\boldsymbol{K}\|_2$
$\bar{\mu_k}$	maximum of μ_k	10^5
ϵ_k	terminate value	10^{-5}

time complexity, (2) singular value shrinking the updated gram matrix from (4.6) that takes $\mathcal{O}(n^3)$ and (3) the partial derivative to update X from (4.7), which takes $\mathcal{O}(dmn)$. Overall, the time complexity of Algorithm 2 for one iteration is $\mathcal{O}(mn^2+n^3)$. In contrast to the competitive linear approaches, the main computational complexity of PCA is from the covariance matrix and eigen-decomposition, i.e. $\mathcal{O}(m^2n+n^3)$, which of RPCA comes from the SVD with $\mathcal{O}(\min(mn^2, m^2n))$. Regarding the other nonlinear modeling techniques, the computational complexity is contributed by the gram matrix computation as well as eigen-decomposition and/or SVD, and approximately measured as $\mathcal{O}(mn^2+n^3)$. That is to say, our RKPCA does not have much more computational cost when $n \ll m$.

4.3.4 RKPCA for Outlier Recovery

As aforementioned, outliers are often distributed arbitrarily across the training data, e.g. in noisy images, and our proposed RKPCA is supposed to construct approximately the same subspace from a corrupted training dataset as that from an uncorrupted one. Under this assumption, experiments are conducted to assess the ability of outlier recovery for competitive methods. We artificially make training datasets corrupted and

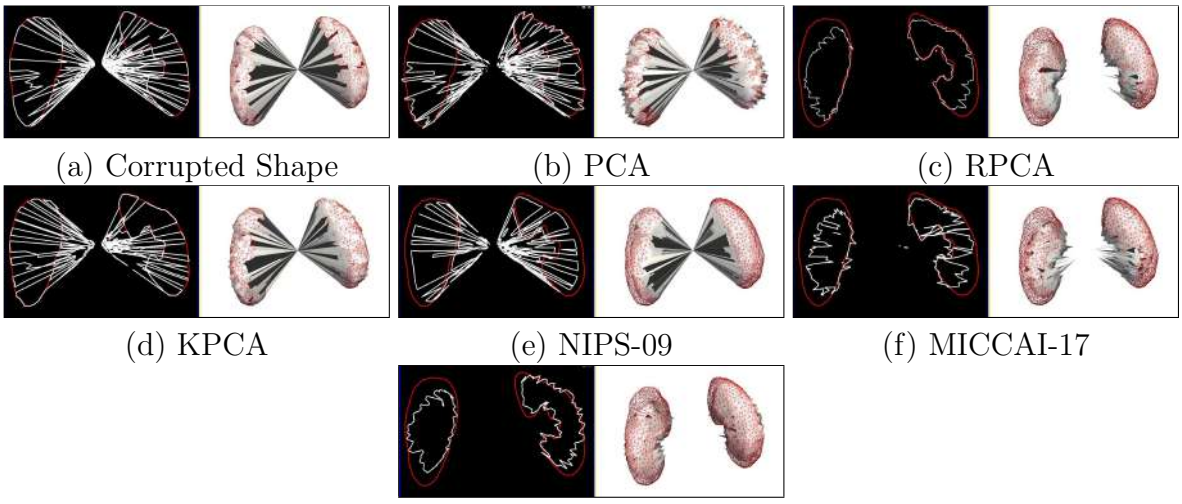
compare the distances between the reconstruction of corrupted training data (cf. Fig. 4.1, $\{X_L^{(i)}\}_{1\cdots N_s}$ denotes the reconstruction in our RKPCA) and its corresponding ground truth $\{Y^{(i)}\}_{1\cdots N_s}$, i.e. smaller distance indicate higher accuracy of outlier recovery. To be specific, a proportion of landmarks of all the training shapes are removed at once, which makes the training data matrix contain arbitrarily distributed missing entries. Through varying the amount of landmark removal, we have training datasets that are corrupted to varying degree thereof. In the experiments, (1) the removing entries are randomly selected per training shape and are different throughout the population; (2) all the missing entries are set to the same point, the location of which is not restricted, and in general it does not effect the outlier recovery results; (3) both datasets of kidney pair and ankle bone are employed in the outlier recovery evaluation. We illustrate the results for kidney pairs for more intuitive comparison.

Since the kidney pairs are pre-centralized, a proportional of points are randomly removed from the original datasets and these 3D coordinates are set to $(0, 0, 0)$ (cf. Fig. 4.2 and Fig. 4.3). Table 4.2 reports the symmetrical distance between the reconstructed shapes and corresponding ground truth with corruption degrees ranging from 0% to 50%. Each values in Table 4.2 is computed by $1/N_s \sum_{i=1}^{N_s} \delta(X_L^{(i)}, Y^{(i)})$. The kernel width σ is chosen as 473.33, 1062.33, 1327.75, 1484.77, 1579.65, 1631.16 in cases where the proportion of removed landmarks is 0%, 10%, 20%, 30%, 40% and 50% respectively. The paired t-test under the null hypothesis is used as the test statistic to decide whether the results are statistically significant. Fig. 4.2 and Fig. 4.3 intuitively illustrate the comparisons of ground truth and reconstructions from competitive methods, where the corruption degrees are set to 20% and 40% respectively.

Correspondingly, Table 4.3 reports the computation time in application (in second) on **Intel** Core i7 processor. With the increase of proportion of missing entries, the number of convergence iteration also increases in the nonlinear subspace compression in our RKPCA

Table 4.2: Reconstruction errors (mm) of kidney pairs for different methods and proportion of missing values. The star $*$ indicates a statistically significant difference between the corresponding results and our method at a significant level of 0.01.

	0%	10%	20%	30%	40%	50%
PCA	2.451 ± 0.380*	16.974 ± 4.947*	28.174 ± 6.572*	35.898 ± 6.691*	40.994 ± 7.265*	45.330 ± 8.084*
RPCA	3.302 ± 1.246*	5.797 ± 1.932	9.650 ± 5.182*	13.862 ± 2.188*	20.177 ± 2.665*	32.176 ± 3.603*
KPCA	5.523 ± 5.831*	17.938 ± 4.593*	24.404 ± 6.979*	32.273 ± 7.616*	39.718 ± 8.931*	44.962 ± 9.869*
NIPS-09	2.668 ± 12.735*	18.430 ± 4.002*	21.778 ± 5.004*	27.454 ± 7.308*	37.932 ± 3.950*	45.450 ± 8.709*
MICCAI-17	2.985 ± 1.463*	9.181 ± 2.576*	10.844 ± 2.648*	13.217 ± 3.168*	18.552 ± 2.970*	20.639 ± 2.903*
RKPCA	**2.050 ± 0.392**	**6.310 ± 1.330**	**8.702 ± 1.863**	**10.648 ± 2.339**	**14.032 ± 3.017**	**18.534 ± 2.692**



(a) Corrupted Shape (b) PCA (c) RPCA

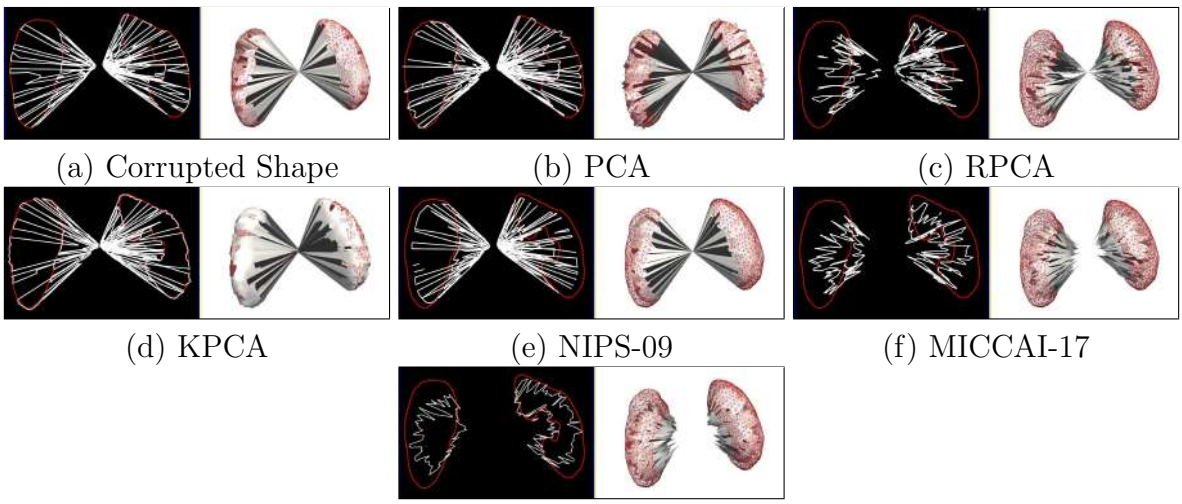
(d) KPCA (e) NIPS-09 (f) MICCAI-17

(g) RKPCA

Figure 4.2: The figure plots a corrupted shape (a) with its reconstructions from competitive models (b) - (g) illustrated from the perspective of Coronal-Axial view (left) and 3D rendering (right), where the red shape indicates the ground truth. (a) is artificially generated with 20% of landmarks randomly removed.

and the pre-image in other methods. Table 4.4 reports the computation time with respect to the scale of training data matrix $D \in \mathbb{R}^{m \times n}$, where the first row indicates the results for kidney pair reconstruction and the second row indicates the results for ankle bone reconstruction.

Conclusions can be drawn from Table 4.2. First of all, RPCA, MICCAI-17 and our RKPCA all have the ability to deal with arbitrary outliers. Particularly, RPCA outperforms in cases of small corruption, whereas, RKPCA shows significant efficiency



(a) Corrupted Shape (b) PCA (c) RPCA

(d) KPCA (e) NIPS-09 (f) MICCAI-17

(g) RKPCA

Figure 4.3: The figure plots a corrupted shape (a) with its reconstructions from competitive models (b) - (g) illustrated from the perspective of Coronal-Axial view (left) and 3D rendering (right), where the red shape indicates the ground truth. (a) is artificially generated with 40% of landmarks randomly removed.

Table 4.3: The computation time (s) for reconstruction of kidney pairs with various proportion of missing values from 0% to 50% for different methods.

Method	0%	10%	20%	30%	40%	50%
PCA	1.846	1.632	1.465	1.791	1.257	1.641
RPCA	2.936	2.927	2.497	2.378	2.759	2.134
KPCA	12.340	36.120	56.120	99.150	60.720	35.520
NIPS-09	13.510	14.312	15.043	15.248	16.309	16.151
MICCAI-17	3.800	7.181	16.520	30.900	45.900	26.920
RKPCA	**2.152**	**2.610**	**2.923**	**3.105**	**3.502**	**3.370**

Table 4.4: The computation time (s) for reconstruction of kidney pairs (first row) and ankle bones (second row) using RKPCA with various proportion of missing values from 0% to 50%, where m, n denote the scale of data matrix.

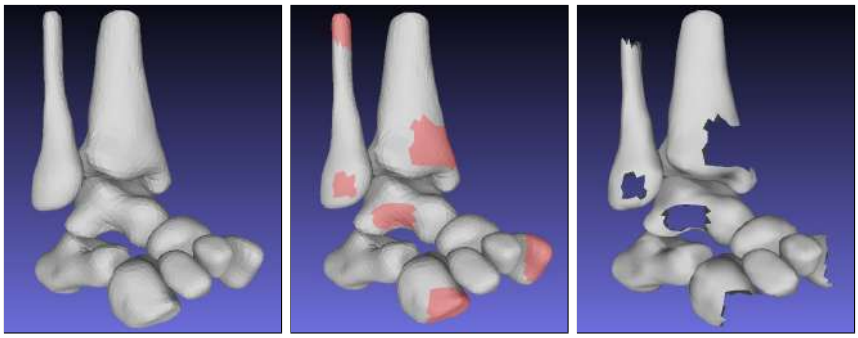
m	n	0%	10%	20%	30%	40%	50%
12000	30	2.152	2.610	2.923	3.105	3.502	3.370
15444	49	5.190	6.331	5.253	6.320	7.534	5.250

when the corruption degree becomes higher and the power of RPCA degrades (cf. Fig. 4.2-(c), (g) and Fig. 4.3-(c), (g)). This is mainly because RKPCA exploits the affinities of each pair of training data across the population and RPCA exploits the variance for every dimension individually. RPCA offers a finer reconstruction compared to RKPCA when the training data is slightly corrupted. However, when the corruption degree becomes higher, the number of reliable variables in each dimension is reduced, which makes it theoretically difficult to learn the dominant variability. In contrast, RKPCA is still capability to exploit the primary nonlinear patterns because the impact of outliers is alleviated during the kernelizing. Secondly, the reason for the reconstruction error with none missing values is that, all the modeling techniques are based on dimensionality reduction in accordance with compressive sensing where usually 95% variabilities are retained. In terms of the intuitive comparison illustrated in Fig. 4.2 and Fig. 4.3, where we find the competitive models eliminate corruptions to different extent but our RKPCA delivers the best reconstruction quality overall.

4.3.5 RKPCA for Missing Area Completion

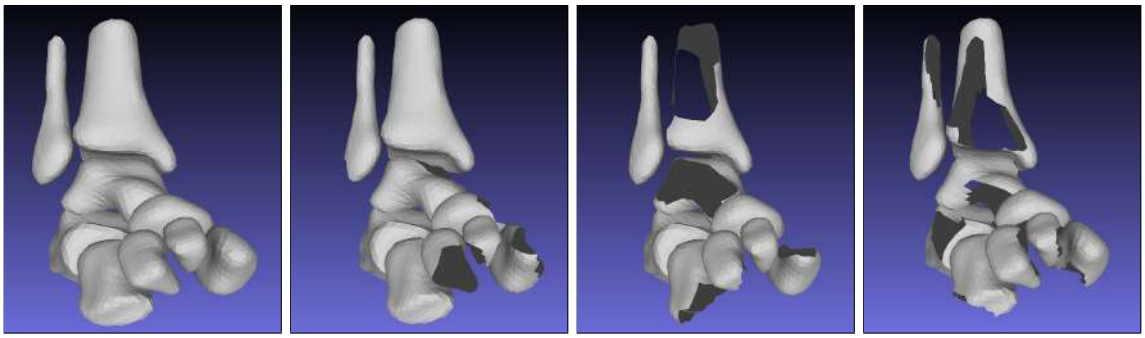
Besides the outliers, training shapes are often incomplete with missing parts in realistic applications. Therefore, experiments are conducted to assess the ability of missing areas completion. Assume that the missing areas are arbitrarily distributed across the whole shape, a subset of landmarks are randomly selected and removed together with their neighboring landmarks (cf. Fig. 4.4). In this manner, we artificially make all the training datasets incomplete and aim to assess the reconstruction accuracy, namely, closer to the corresponding ground truth data indicates higher accuracy. In this set of experiments, about 15 landmarks are randomly selected from each training shape and all the incomplete shapes are different across the population. Through varying the area of neighborhood around these selected landmarks, we thus generate training shapes with

different degree of incompleteness (cf. Fig. 4.5), i.e. the proportion of missing areas is equivalent to the radio between the sum of missing points and the total number of landmarks.



(a) (b) (c)

Figure 4.4: This figure plots the procedure of incomplete shapes generation: based on a ground truth shape (a), several pieces with varying areas are randomly selected (marked in red in (b)) and removed. An incomplete shape (c) with missing areas is generated thereof.



(a) (b) (c) (d)

Figure 4.5: The figure plots the ground truth ankle bone shape (a) and the artificially generated incomplete shapes with 10%, 20% and 30% removing points in (b), (c) and (d) respectively.

Table 4.5 reports the reconstruction errors for competitive methods, where the degree of incompleteness varies from 0% to 50%. Each values in Table 4.5 is computed by $1/N_s \sum_{i=1}^{N_s} \delta(X_L^{(i)}, Y^{(i)})$. For an intuitive view, Fig. 4.6 illustrates the reconstructions of incomplete training shapes with 30% missing areas compared to the ground truth

(marked in red). To be specific, the white pieces in Fig. 4.6-(a) denote the overlap of ground truth and incomplete shape, and the missing landmarks are close to the central position because the ankle bones are pre-centralized.

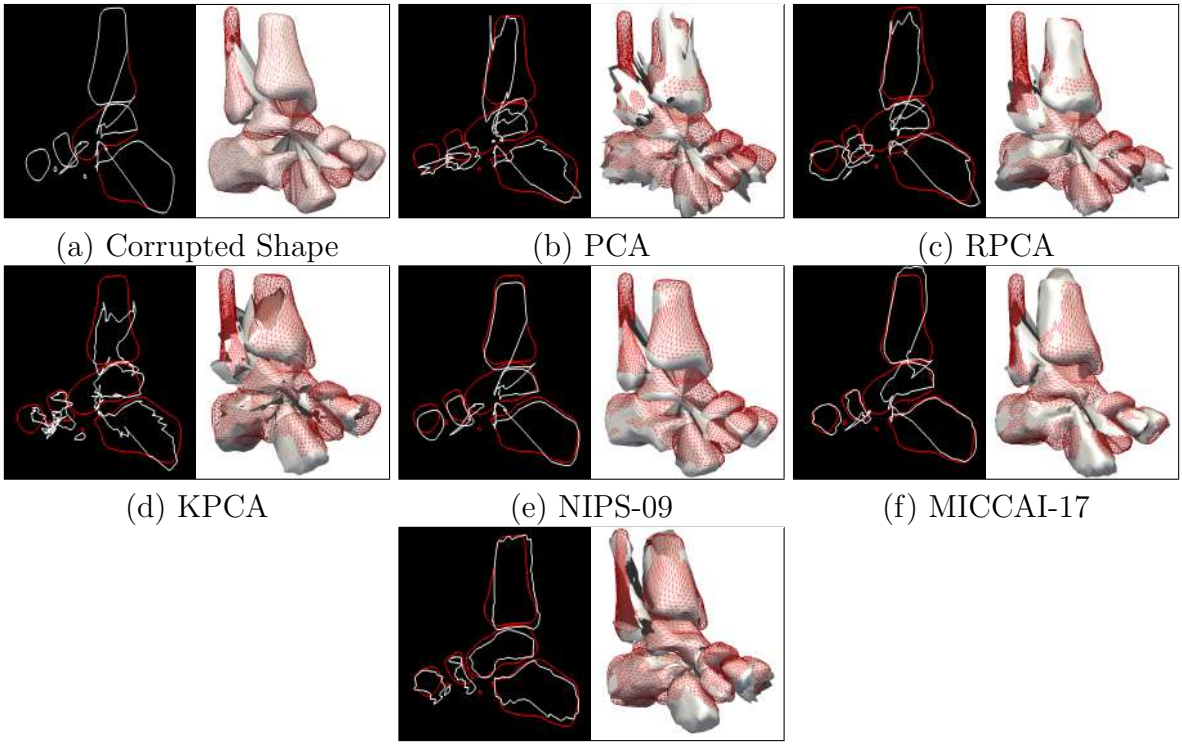
Table 4.5: Reconstruction errors (mm) of ankle bones for different methods and proportion of missing areas. The star $*$ indicates a statistically significant difference between the corresponding results and our method at a significant level of 0.01.

	0%	10%	20%	30%	40%	50%
PCA	1.567 ± 0.295	5.845 ± 0.357*	10.110 ± 1.055*	13.724 ± 1.804*	16.403 ± 2.266*	22.666 ± 3.646*
RPCA	1.811 ± 0.623*	2.740 ± 0.991	6.114 ± 1.285*	9.812 ± 2.547*	14.516 ± 3.132*	19.351 ± 2.126*
KPCA	5.709 ± 4.594*	7.023 ± 4.453*	10.496 ± 4.171*	14.290 ± 3.737*	17.018 ± 3.566*	20.296 ± 3.470*
NIPS-09	3.567 ± 1.891*	6.013 ± 2.345*	9.897 ± 1.487*	13.833 ± 2.157*	16.454 ± 2.754*	19.491 ± 2.460*
MICCAI-17	3.409 ± 0.871*	5.658 ± 1.806*	8.338 ± 2.567*	9.614 ± 2.036*	11.096 ± 3.780*	17.697 ± 2.960*
RKPCA	**1.663 ± 0.196**	**3.528 ± 1.502**	**5.981 ± 1.563**	**7.544 ± 2.268**	**9.762 ± 3.039**	**14.950 ± 2.953**

Conclusions can be drawn from Table 4.5 that RPCA, MICCAI-17 and our proposed RKPCA have much more capability of dealing with missing areas among the competitive methods. With the degree of incompleteness increases, nonlinear approach KPCA performs slightly better than PCA and RKPCA beats RPCA significantly.

4.3.6 Model Evaluation

For an intuitive understanding of the robustness to nonlinearity, we compare the shapes generated from the first mode of variance for models derived with RPCA and our RKPCA under the datasets of kidney pair in Fig. 4.7. It is apparently seen that the shapes generated from RPCA model (cf. Fig. 4.7-(a) left and right) make no sense due to the impossible deformation, where the left kidney is too close to the right kidney. Note that the variance for linear models varies in the range $\left[-3\sqrt{\Lambda_i}, +3\sqrt{\Lambda_i}\right]$, for nonlinear models which varies in the range $\left[-\sqrt{N_s\Lambda_i}, +\sqrt{N_s\Lambda_i}\right]$ as defined in [98], where $i = 1 \cdots n_k$. Similarly, the visualization of the first mode of variance for KPCA and RKPCA models shown in Fig. 4.8, which suggests the robustness to data corruption of the proposed



(a) Corrupted Shape (b) PCA (c) RPCA

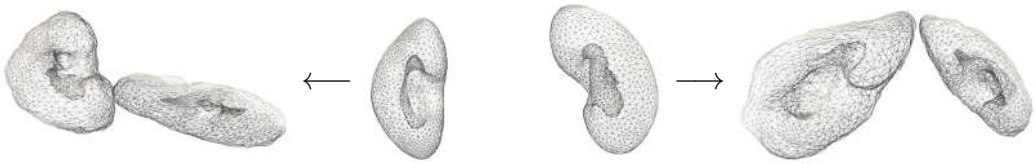
(d) KPCA (e) NIPS-09 (f) MICCAI-17

(g) RKPCA

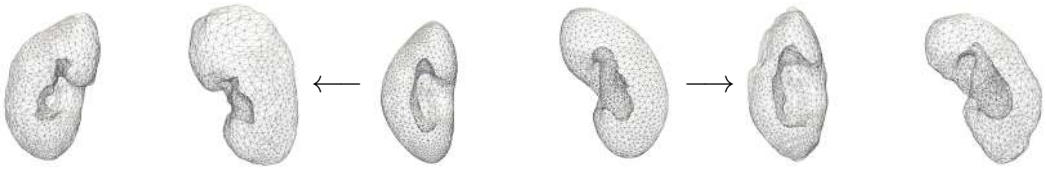
Figure 4.6: The figure plots an incomplete ankle bone shape (a) with its reconstructions from competitive models (b) - (g) illustrated from the perspective of Coronal-Axial view (left) and 3D rendering (right), where the red shape indicates the ground truth. (a) is artificially generated with 30% of pieces randomly removed.

RKPCA. Obviously it is found that the shapes from KPCA model still have erroneous areas, e.g. the overlap (cf. Fig. 4.8-(a) left) and anomalously deformed subparts (cf. Fig. 4.8-(a) right), by contrast, our RKPCA model efficiently eliminates the abnormalities in training data.

To evaluate the quality of SSMs derived with competitive methods, Generalization ability (leave-one-out mode) G, Specificity S and Compactness C are computed based on (3.26). We report the G, S and C for the kidney models and ankle bone models in Fig. 4.9 and Fig. 4.10 respectively. Under the paired t-test with the null hypothesis, the $p-$values are calculated with a significance level of $\alpha = 0.01$ for the generalization ability, specificity for kidney models (cf. Table 4.6 and Table 4.7) and ankle bone models



(a) RPCA model



(b) RKPCA model

Figure 4.7: The variance for the RPCA and RKPCA model, where the shape in the middle is the mean and the left and right shapes are generated from $-3\sqrt{\Lambda_1}$ and $+3\sqrt{\Lambda_1}$ for RPCA model, and $-\sqrt{N_s\Lambda_1}$ and $+\sqrt{N_s\Lambda_1}$ for RKPCA model respectively, with other modes set to the 0.

(a) KPCA model

(b) RKPCA model

Figure 4.8: The variance for the KPCA and RKPCA model, where the shape in the middle is the mean and the left and right shapes are generated from $-\sqrt{N_s\Lambda_1}$ and $+\sqrt{N_s\Lambda_1}$ for the first mode, with other modes set to 0.

(cf. Table 4.8 and Table 4.9) respectively.

To compute specificity, $M = 1000$ shapes in total are randomly drawn from the model

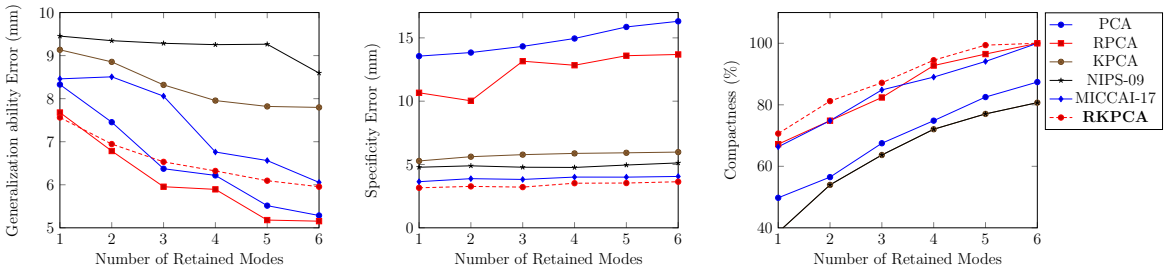


Figure 4.9: Generalization ability G, specificity S and compactness C for the kidney models.

Table 4.6: Statistical analysis for Generalization ability of competitive kidney pair models and ours for the first 6 modes. Note that the star $*$ indicates that the corresponding value of ours method is statistically significant better than the marked result.

	1	2	3	4	5	6
PCA	8.3278*	7.4536*	6.3727	6.2175	5.5150	5.2862
RPCA	7.6781	6.7840	5.9545	5.8933	5.1798	5.1543
KPCA	9.1343*	8.8541*	8.3192*	7.9553*	7.8195*	7.7964*
NIPS-09	9.4526*	9.3465*	9.2856*	9.2542*	9.2645*	8.5948*
MICCAI-17	8.4595*	8.5069*	8.0583*	6.7610	6.5624	6.0550
RKPCA	**7.5621**	**6.9456**	**6.5321**	**6.3210**	**6.0945**	**5.9545**

Table 4.7: Statistical analysis for Specificity of competitive kidney pair models and ours for the first 6 modes. Note that the star $*$ indicates that the corresponding value of our method is statistically significant better than the marked result.

	1	2	3	4	5	6
PCA	13.5585*	13.8404*	14.3276*	14.9500*	15.8621*	16.3088*
RPCA	10.6642*	10.0309*	13.1569*	12.8369*	13.5912*	13.6890*
KPCA	5.2807*	5.6164*	5.7778*	5.8765*	5.9236*	5.9804*
NIPS-09	4.7846	4.8956	4.7852	4.7674	4.9584	5.1250*
MICCAI-17	3.6432	3.8820	3.8208	4.0067	4.0023	4.0562
RKPCA	**3.1588**	**3.2711**	**3.2143**	**3.5226**	**3.5360**	**3.6391**

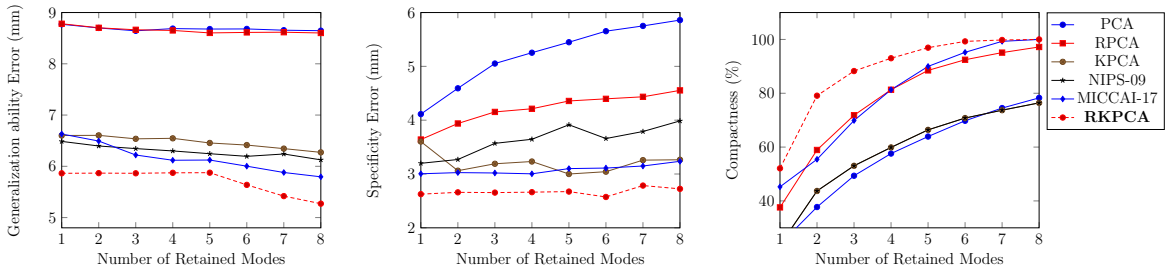


Figure 4.10: Generalization ability G, specificity S and compactness C for the ankle bone models.

Table 4.8: Statistical analysis for Generalization ability of competitive ankle bone models and ours for the first 8 modes. Note that the star $*$ indicates that the corresponding value of our method is statistically significant better than the marked result.

	1	2	3	4	5	6	7	8
PCA	8.7706*	8.6995*	8.6432*	8.6868*	8.6778*	8.6802*	8.6553*	8.6457*
RPCA	8.7813*	8.7024*	8.6634*	8.6509*	8.6024*	8.6124*	8.6160*	8.6012*
KPCA	6.6015*	6.6041*	6.5363*	6.5461*	6.4554*	6.4145*	6.3452*	6.2724*
NIPS-09	6.4859*	6.3964*	6.3458*	6.3001*	6.2459*	6.1945*	6.2403*	6.1260 *
MICCAI-17	6.6305*	6.4927*	6.2207*	6.1182	6.1225	6.0025	5.8795	5.7951*
RKPCA	**5.8636**	**5.8671**	**5.8640**	**5.8732**	**5.8757**	**5.6378**	**5.4189**	**5.2712**

Table 4.9: Statistical analysis for Specificity of competitive ankle bone models and ours for the first 8 modes. Note that the star $*$ indicates that the corresponding value of our method is statistically significant better than the marked result.

	1	2	3	4	5	6	7	8
PCA	4.1130*	4.5924*	5.0523*	5.2537*	5.4484*	5.6536*	5.7512*	5.8596 *
RPCA	3.6402*	3.9385*	4.1535*	4.2121*	4.3571*	4.3967*	4.4345*	4.5535*
KPCA	3.6025*	3.0616*	3.1899*	3.2320*	3.0000*	3.0428*	3.2578*	3.2645*
NIPS-09	3.1986*	3.2695*	3.5695*	3.6452*	3.9145*	3.6589*	3.7895*	3.9864*
MICCAI-17	3.0041*	3.0259*	3.0190*	3.0034*	3.1004*	3.1106*	3.1487*	3.2357*
RKPCA	**2.6257**	**2.6594**	**2.6546**	**2.6626**	**2.6734**	**2.5732**	**2.7852**	**2.7239**

PDF. As the models of the ankle bone are derived based on corrupted training data, we compute G and S using corresponding ground truth data $\{x_i\}$ and $\{y_A\}$ generated from model's PDF. Fig. 4.9 shows that our proposed RKPCA consistently boost the performance of model, in particular, the improvement of specificity suggests the considerable ability in drawing nonlinear PDF. On the other hand, Fig. 4.10 reports a higher quality of RKPCA model in coping with data corruption, especially in terms of generalization ability. It should be pointed out that **KPCA** and **NIPS-09** extracts nonlinear pattern variabilities in the same way, resulting in the equivalent eigenvalues $\Lambda_{i:1\cdots n_k}$.

4.3.7 Segmentation Evaluation

To assess the effect on segmentation accuracy, the derived SSMs are used in two existing segmentation approaches for CT kidney and MRI ankle bone respectively. Both approaches are based on a hierarchical ASM framework. Dice Similarity Coefficient DSC, Hausdorff Distance HD, and Volumetric Similarity VS are used as quality measures, whereas DSC and VS are listed in percent (larger values indicate better performance) and HD is given in millimeters (smaller values indicate better performance). We utilize the framework presented in [79] for kidney segmentation which is driven primarily by an automatic adaption to local features and a global model back propagation. In Table 4.10 we report the segmentation results using the competitive models. Our approach outperforms the other models in all measures (cf. Fig. 4.11).

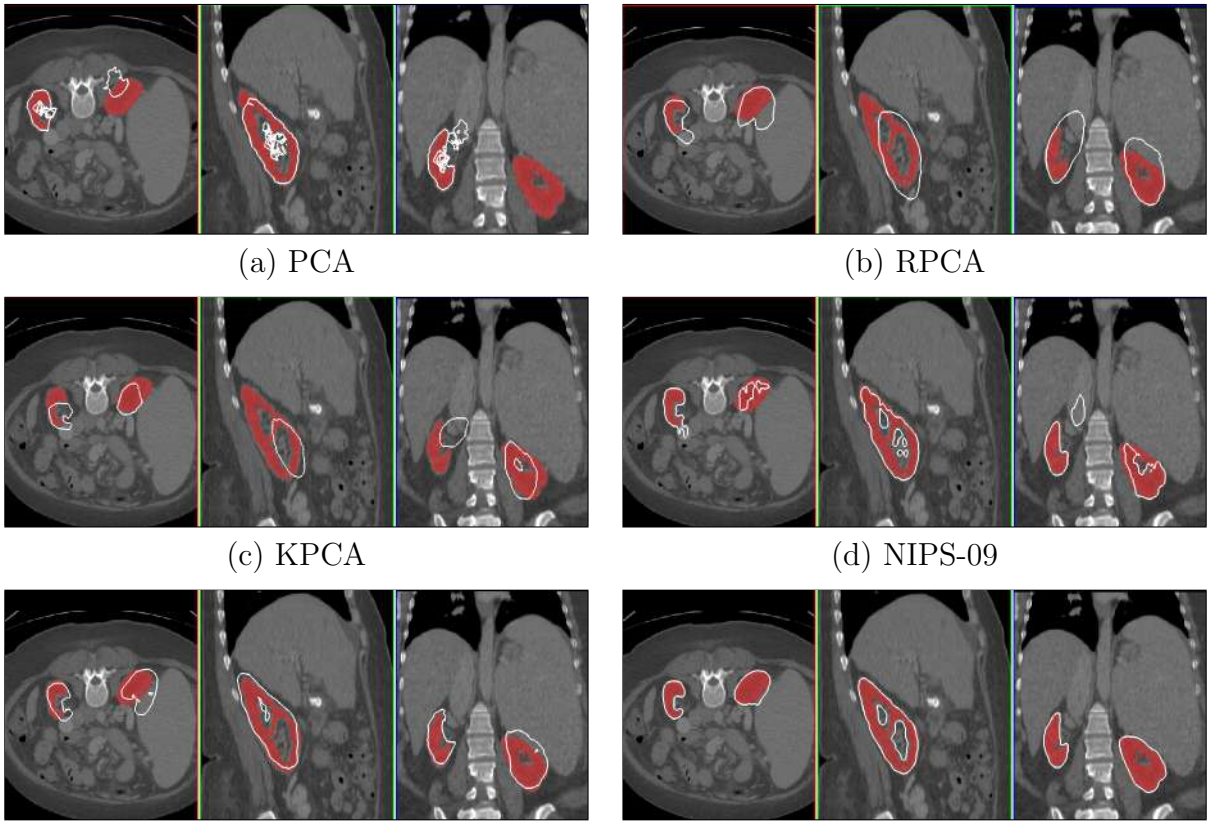
We utilize the same ankle bone segmentation approach presented in [28] as in our previous conference work [95] and report results in Table 4.11. Our proposed model delivers an improvement from the previous work MICCAI-17 as well as other competitive approaches. Fig. 4.12 illustrates a segmentation result: the narrow inner space accounts for the poor image adaption and back propagation, whereas our model effec-

Table 4.10: Kidney Segmentation Results

	DSC			HD (mm)			VS		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
PCA	0.761±0.144	0.303	0.961	15.063±7.536	4.898	38.131	0.932±0.080	0.595	0.991
RPCA	0.802±0.064	0.608	0.965	15.962±7.198	2.015	35.818	0.942±0.060	0.620	0.990
KPCA	0.816±0.093	0.559	0.901	14.982±8.825	0.434	39.824	0.923±0.062	0.688	0.993
NIPS-09	0.821±0.083	0.519	0.955	14.074±6.417	5.385	32.802	0.935 ±0.049	0.811	0.997
MICCAI-17	0.863±0.086	0.570	0.960	13.664±8.320	5.385	36.576	0.925±0.071	0.707	0.995
RKPCA	**0.905±0.016**	**0.876**	**0.945**	**7.845±1.676**	**5.099**	**11.180**	**0.970±0.018**	**0.917**	**0.994**

Table 4.11: Ankle Bone Segmentation Results

	DSC			HD (mm)			VS		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
PCA	0.827±0.110	0.596	0.931	8.029±2.454	2.735	16.572	0.852±0.099	0.717	0.953
RPCA	0.863±0.031	0.746	0.901	8.432±2.925	4.735	16.572	0.902±0.034	0.807	0.953
KPCA	0.843±0.038	0.746	0.901	9.018±2.959	4.735	16.572	0.892±0.059	0.797	0.953
NIPS-09	0.872±0.016	0.841	0.904	8.398±2.265	4.960	14.021	0.918±0.057	0.791	0.969
MICCAI-17	0.908±0.033	0.825	0.960	7.881±2.210	4.361	13.210	0.940±0.030	0.886	0.990
RKPCA	**0.930±0.025**	**0.880**	**0.972**	**6.115±2.360**	**2.280**	**10.753**	**0.980±0.035**	**0.894**	**0.997**



(a) PCA (b) RPCA

(c) KPCA (d) NIPS-09

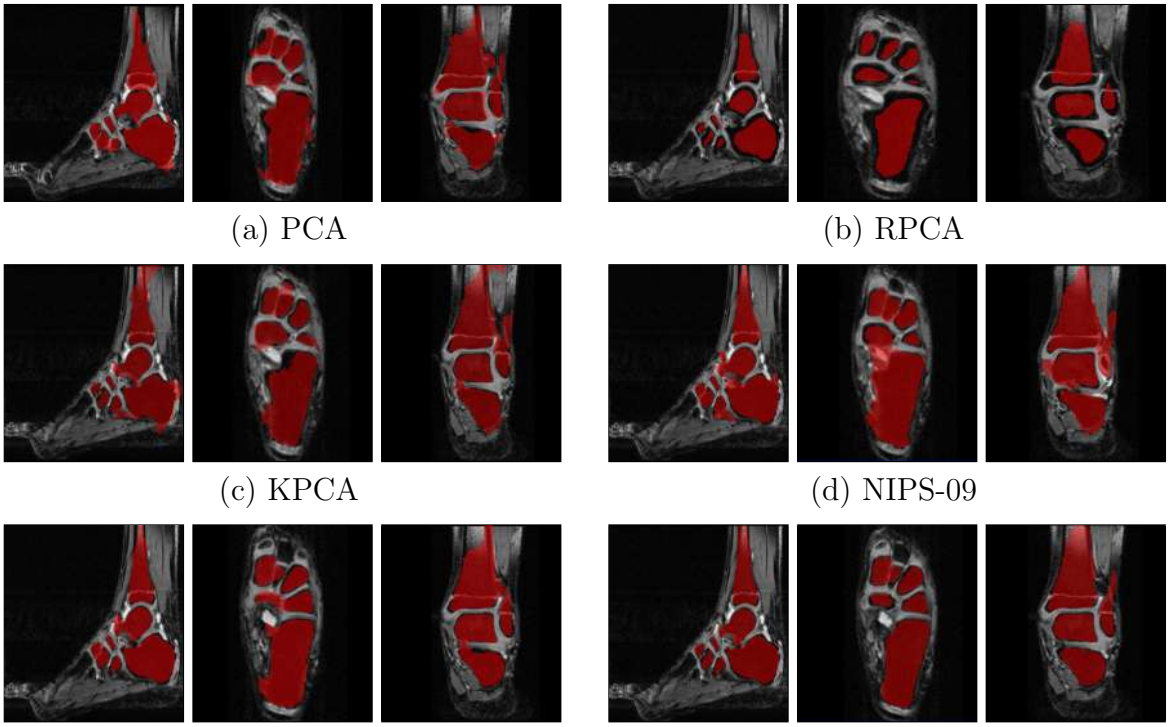
(e) MICCAI-17 (f) **RKPCA**

Figure 4.11: Comparisons of kidney pair segmentation results.

tively alleviates the overlap and abnormal deformation due to its ability of handling corruptions.

4.4 Discussion

In this chapter, we propose a novel Robust Kernel Principal Component Analysis approach for statistical shape modeling with the aim to derive an SSM that is robust to arbitrarily corrupted data and capable to represent linear and non-linear variabilities simultaneously. Experiments are conducted on 30 public CT kidney pair volumes to validate the models' capability of representing non-linear features. Furthermore, 49 MRI ankle bones are used to validate the models' capability of handling highly corrupted



(a) PCA (b) RPCA

(c) KPCA (d) NIPS-09

(e) MICCAI-17 (f) **RKPCA**

Figure 4.12: Comparisons of ankle bone segmentation results.

data. To investigate the ability for arbitrary outlier recovery, we artificially generate corrupted datasets by removing a certain proportion of landmarks and measure the distance between its reconstructions from competitive models and ground truth afterwards. Experimental results show a significant advantage of our RKPCA in coping with arbitrary outliers. Moreover, our proposed model outperforms the state of the art in all SSM relevant measures: generalization ability, specificity, and compactness. Furthermore, the application of our model in segmentation frameworks also demonstrates higher accuracy in comparison to state of the art SSMs.

In addition to the application in statistical shape modeling, our proposed RPCA can also be used in other compressive sensing areas, e.g. image and signal denoising, as it is capable to capture significant nonlinear patterns from any data matrix and it is robust to arbitrary outliers as well.

Several limitations do exist in practice, though. Generally, the training data needs to sufficiently cover the shape variances of the target object, i.e. as with other statistical shape modeling approaches, any abnormal shape not contained in the training data will not be modeled correctly. For example, for unhealthy organs with strong deformations or large tumors, RKPCA's robustness to corruptions may falsely try to correct these abnormalities. The same applies for significant articulated joint motions which are not present in the training data. Furthermore, if it comes to modeling of normally distributed training data, RKPCA does not provide major advantages other than speed in comparison to KPCA. Moreover, SSMs are generally sensitive to initialization and RKPCA is no exception. Especially for small and highly deformable structures this can be problematic and accurate pre-initialization methods are needed. In this work, though, as the ankle bone and kidney pair both occupy a relatively large region, practically, the initialization does not significantly degrade the segmentation accuracy. Still, this motivates us to combine our RKPCA with other approaches like deep neural networks to provide an accurate initialization method in future work.

Chapter 5

A Novel Bayesian Model Incorporating Deep Neural Network and Statistical Shape Model for Pancreas Segmentation

In recent years, deep neural networks have permeated the field of medical imaging, particularly, achieved mentionable success in medical image segmentation. However, poor contrast to surrounding tissues and high flexibility of anatomical structure of the interest object are still challenges. On the other hand, statistical shape model based approaches have demonstrated promising performance on exploiting complex shape variabilities but they are sensitive to localization and initialization. This motivates us to leverage the rich shape priors learned from statistical shape models to improve the segmentation of deep neural networks.

In this chapter, a novel Bayesian model is proposed that incorporates the segmentation results from both deep neural network and statistical shape model for segmentation. In evaluation, experiments are performed on 82 CT datasets of the challenging public NIH pancreas dataset. We report 85.32 % of the mean DSC that outperforms the state-of-the-art and approximately 12 % improvement from the predicted segment of deep neural

network. It should be pointed out that partial content in this chapter comes from our publication "A Novel Bayesian Model Incorporating Deep Neural Network and Statistical Shape Model for Pancreas Segmentation" presented at MICCAI 2018.

The chapter is organized as follows: first, an overview of pancreas segmentation approaches is introduced; the proposed DenseUNet as well as a concise introduction of UNet and DenseNet is provided in 5.2, followed by the shape model initialization; subsequently, the proposed Bayesian inference for segmentation is illustrated in 5.3, which is the major contribution of this work; afterwards, experiments for CT pancreas segmentation are introduced as well as a discussion.

5.1 Introduction

With the rapid development of Convolutional Neural Networks (CNNs) in semantic segmentation, deep neural networks like FCN [51], UNet [1], SegNet[52] and RCNN [129] have become a popular trend in medical image segmentation and achieved remarkable success in organ and lesion segmentation. However, difficulties still remain in some challenging subjects segmentation, such as CT pancreas, due to the relatively small region in the whole volume, highly complex anatomical structure and significantly ambiguous boundary. On the other hand, usually the amount of labeled medical image data is limited which inhibits the segmentation from achieving considerable accuracy. To tackle these challenges, we aim to propose a robust segmentation approach for pancreas, which is one of the most challenging organs.

5.1.1 Overview of Pancreas Segmentation

Numerous works focus on pancreas segmentation in literature, table 5.1 conveys a survey. Early works using traditional shape models are proposed in [130, 85]. Subsequently,

shape models are chosen to be incorporated with other methods like level set [84] and multi-atlas [131]. Since the year of 2015, the number of publications using deep neural networks rapidly increased, note that the public NIH datasets is provided in [132]. With the popularization of UNet, a large number of segmentation architectures are built on the basis of UNet [73, 133, 134].

As CNN based neural networks lack the ability of learning 3D spatial features, Cai et al. employ the recurrent neural networks (RNN) to learn the spatial information across adjacent slices in their two related publications [135, 136]. From another aspect, graphical models like Conditional Random Field (CRF) and Markov Random Field (MRF), shape models, image classifiers [137] [138] and Gaussian Mixture models [133], are often utilized on the top of probability score obtained from the deep neural network for a local refinement. The BRIEFnet proposed in [139] leverages a binary sparse convolutional in the first layer, followed by the dilated convolutional layers to learn the rich spatial features. In [140], Xia et al. fuse the segmentation based on 2D slices into a 3D Volumetric Fusion Net (VFN) in order to integrate 3D information.

In small object segmentation, background contents occupy a relatively large area and seriously disturb the salient feature extraction, which degrade the efficiency of segmentation frameworks. To combat this, a coarse-to-fine framework is designed in [141] where the coarse network is trained to obtain the rough segment and remove the background regions, afterwards the shrunken region is passed to the fine network for precise segmentation. Still under the coarse-to-fine framework, Yu et al. proposed the Recurrent Saliency Transformation Network in [142], where the estimated probability score in the last iteration is passed to the new iteration as spatial weights. A similar coarse-to-fine framework is also applied in [143] that deals with multi-scale inputs. On the other hand, in [134], a novel Attention Gate (AG) model based on the UNet architecture is proposed that efficiently suppress the unconcerned regions but highlight salient features. In [73, 144], a

dense dilated convolutional neural network is used for abdominal organs segmentation, which adopt the dense block from DenseNet.

Conclusions can be drawn from the survey: first, obviously it is found that the majority of pancreas segmentation approaches select deep neural networks and SSMs as a basis; an end-to-end trained deep neural network cannot guarantee the segmentation accuracy even though they are preferred, therefore, authors leverage another machine learning approaches, such as the graphical models CRF and MRF, and shape models, to refine the segmentation results.

5.1.2 Hypothesis

Considering of the ambiguities on boundary, it is well worth to leverage the 3D shape variabilities to distinguish the non-visible boundary, this motivates us to employ statistical shape models in segmentation framework. Through back projection onto the shape model, the corruptness on input shape is supposed to be corrected. Owing to the high variability of pancreas shape, the RKPCA (cf. Chapter 4 for detail) is employed to create a nonlinear SSM as it has compelling advantages in handling corrupted and highly deformable training data than conventional PCA models. However, the model based approaches are sensitive to initialization, thus a deep neural network plays an important role in providing a rough segmentation for shape model initialization. With this motivation, we integrate the segmentation from deep neural network and statistical shape model under a Bayesian inference. A novel optimization principle joint with image feature and shape prior is proposed to guide segmentation. Our approach is demonstrated to be promising and efficient in terms of evaluation. The whole pipeline of segmentation is illustrated in Fig. 5.1, where a DenseUNet architecture is presented as a primary training, followed by the Bayesian model to achieve segmentation.

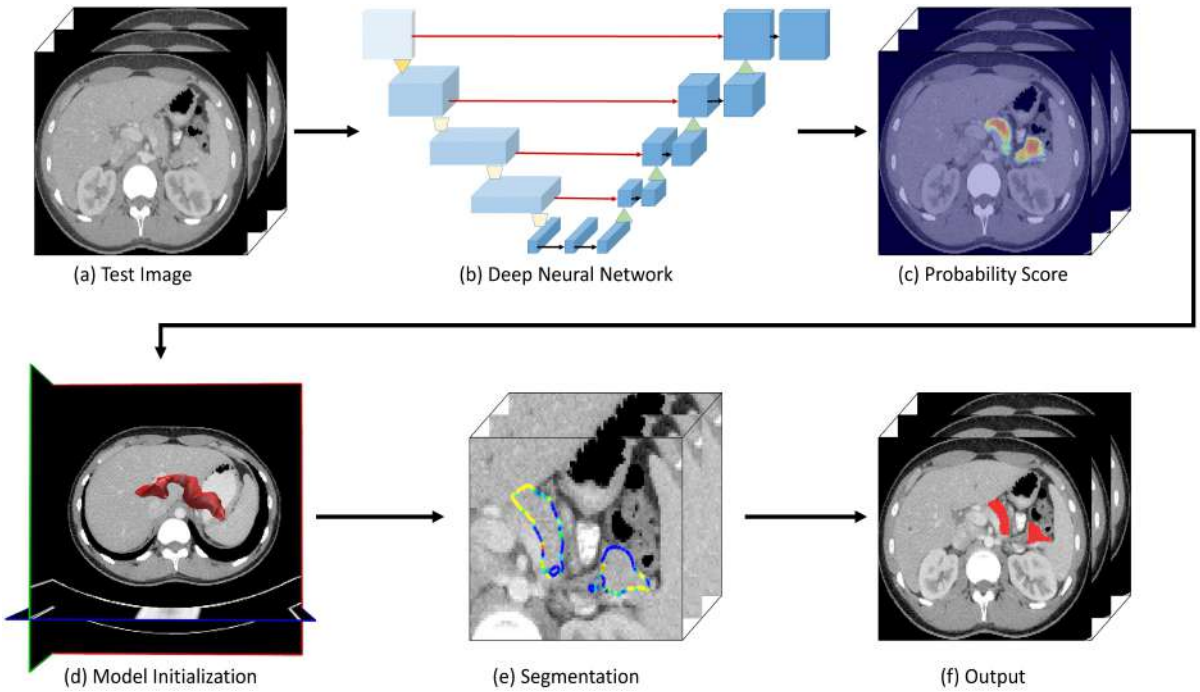


Figure 5.1: This figure illustrates the whole segmentation pipeline: (a) implies the test image to be segmented, through the pre-trained deep neural network (b), a probability score (c) is obtained; (d) indicates that the shape model is initialized with the probability score; and (e) implies the proposed Bayesian model, where the color indicates the probability for each landmark; (f) is the final segmentation output.

5.2 DenseUNet for Model Initialization

In this section, we elaborate our segmentation approach starting with the deep neural network architecture, followed by the Bayesian model. Given a set of 3D CT volumes $I = \{I_1, \ldots, I_N\}$ with corresponding ground truth mask $Y = \{Y_1, \ldots Y_N\}$ for training, shapes $S = \{S_1, \ldots, S_N\}$ can be extracted from the ground truth mask via Marching Cubes to train the robust kernel statistical shape model defined as $RKSSM(S|\mathrm{V}, \mathrm{K})$, where S implies the training dataset, V decides the eigenvectors in kernel space, K is the robust kernel matrix with elements $\mathrm{K}_{ij} = \kappa(S_i, S_j)$ and κ is the kernel trick.

5.2.1 DenseUNet Architecture

UNet, present at MICCAI 2015, has been the most popular neural network in medical image segmentation. Fig. 5.2 illustrates the architecture, where the hierarchical intensity based features are learned in the downsampling path by stacked 3×3 $Conv - BN - Relu$ layers, where BN indicates the batch normalization and $Relu$ indicates the nonlinear activation. Extended from FCN, UNet considers the full context of image and records the position of each pixel. This makes it possible to pass the entire image to UNet and obtain the probability score straightforwardly. More importantly, the key novelty of UNet is a skip connection from downsampling layer to its corresponding upsampling layer, which encourages the feature propagation and reuse.

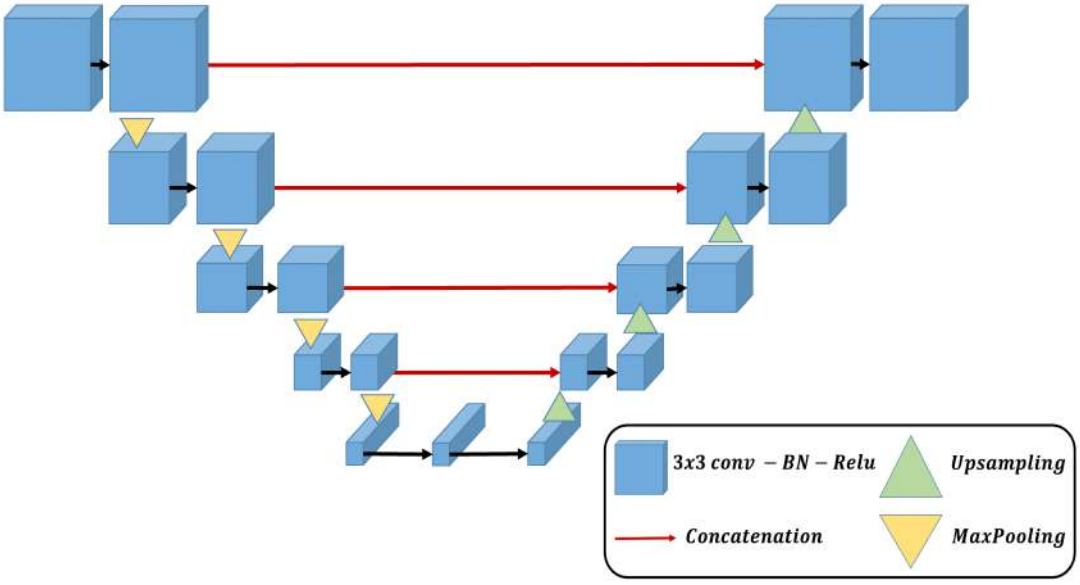


Figure 5.2: Architecture of UNet proposed in [1].

Despite the remarkable performance in localization, the UNet lacks the capability to learn local features. This is because usually medical image has a large dimension, which limits the batch size and number of feature maps in training phase. This motivate us to encourage the feature reuse and reduce number of parameters with the limited

GPU memory. DenseNet [6] shows advantages in narrowing the network width, reusing features and significantly alleviating the problem of gradient vanishing over the ResNet [145], which also aims to enhance the propagation of features. Therefore, we attempt to embed the DenseNet block in UNet architecture by simply replacing the stacked $Conv - BN - Relu$ and a following $MaxPooling$ operation at each downsampling step with a 5-layer dense block with a transition layer, meanwhile, keeping the upsampling path and concatenation unchanged. The architecture is illustrated in Fig. 5.3. In this work, the 5-layer dense block is used with the growth rate of 4.

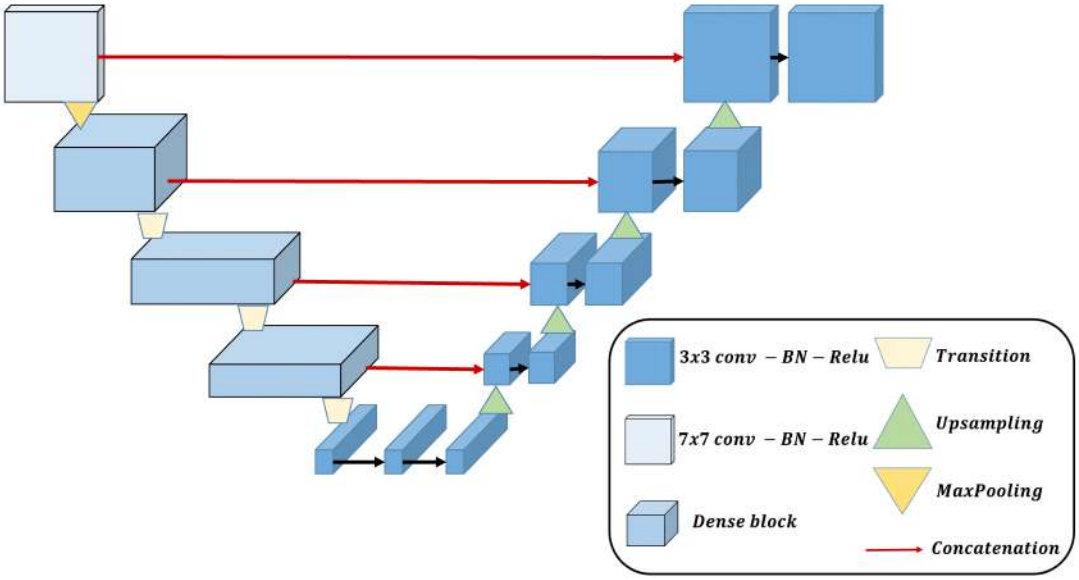


Figure 5.3: Architecture of our proposed DenseUNet. A convolutional with large kernel is adopted at the very beginning, followed by three 5-layer dense block. Each dense block is followed by a transition layer to achieve downsampling.

According to the most of segmentation related works, Dice coefficient loss with a smooth value is used and defined as:

$$l(Z, Y) = 1 - \frac{2 \times \sum_i z_i y_i + 0.1}{\sum_i (z_i^2 + \sum_i y_i^2) + 0.1},$$

where Z represents the predicted mask. Our DenseUNet is trained with 2D slices extracted from 3D training images from Axial view, Sagittal view and Coronal view respec-

tively, resulting in three predicted segment Z^A, Z^S and Z^C. Considering of the GPU memory consumption and computational cost, often the 2d slices extracted from 3D medical image volumes are used for training, e.g. the work [137, 140]. Due to the $ReLu$ activation in the output layer, the intensity range in predicted segment is in $[0, 255]$. To make use of the predicted segments in further Bayesian model, we generate probability maps $\Pi = \{\Pi_1, \ldots, \Pi_N\}$ by merging the three predicted segments and feeding into a sigmoid logistic function:

$$\Pi_i = \frac{1}{1 + \exp(-\frac{S_i^A + S_i^Z + S_i^C}{255})}, \tag{5.1}$$

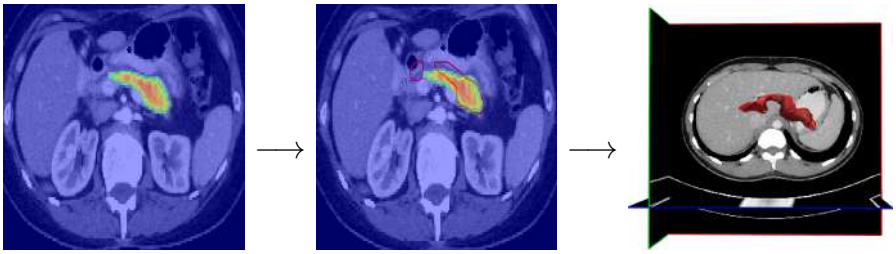
where Π_i indicates the probability map of the i^{th} image. Using the sigmoid function to compute probability map is because (1) this is a binary segmentation task with 2 classes in total, and (2) considering the uncertain accuracy of Dense U-Net, we make the probability for each pixel in range $[0.5, 1]$ that "1" indicates the pixel has a considerable probability of being ROI (Region of Interest) and "0.5" indicates the pixel is unsure to be ROI or NOI (Non of Interest). Apparently, the intersection region of Z^A, Z^S and Z^C is assigned higher probabilities, and uncertain or corrupted areas receive lower probabilities.

5.2.2 Model Initialization

A reliable initialization plays an important role in shape model based segmentation, as it not only locates the shape model but also offers the initial estimation for shape fitting. With the probability score map is obtained from the deep neural network, a necessary pre-requisite step is initialized the shape model with the probability map (cf. Fig. 5.4). First, the largest connected component extraction filter provided by the open source **Insight Segmentation and Registration Toolkit (ITK)** library[1] is applied to the

[1] https://www.itk.org/

probability, in order to extract the largest connected region that is assumed to be the most reliable. It is worth mentioning that this type of largest component thresholding is often used segmentations in literature. A 3D contour can be extracted from the new probability map after thresholding by Marching Cubes. Subsequently, the mean shape of the shape model is aligned to the contour extracted from probability map by the Iterative Closest Point Transform [146] provided by the public **Visualization Toolkit (VTK)** library[2].



(a) (b) (c)

Figure 5.4: This figure illustrates the procedure of model initialization: (a) is the probability score map from deep neural network, where the color indicates the probability value; (b) is the ICP alignment of shape model and the probability score; (c) is the model initialization.

5.3 Bayesian Model

The procedure of the proposed Bayesian model based segmentation refinement is depicted in Fig. 5.5. Let the shape model $RKSSM$ fed into Π for initialization (cf. Fig.5.5-(b)), we have an initial shape of segmentation $C = \{x_1, \ldots, x_{N_p}\}$, where landmark x_i represents the i^{th} pixel in the test image. Given the test image I, probability map Π and the shape model $RKSSM$, assume the optimal shape C can be derived using Bayes' rule as follows:

$$p(C|I, \Pi) \propto p(I, \Pi|C)\ p(C), \tag{5.2}$$

[2] https://www.vtk.org/

term $p(I, \Pi|C)$ is maximum likelihood estimation of C based on image and probability map and term $p(C)$ is considered as the prior distribution of the shape model. Shape C is guided towards the most probable mode by maximizing the posteriori in 5.2, which is equivalent to simply minimizing its negative logarithm leading to the energy function:

$$E(C) = -\log(p(I, \Pi|C)) - \log(p(C)), \tag{5.3}$$

the first term related to the intensity feature is solved via a Gaussian Mixture Model and the second term related to the shape prior is solved with the shape model. The optimal solution is reached by adapting the gradient descent to the energy. The overall procedure of segmentation algorithm is summarized in Alg. 3.

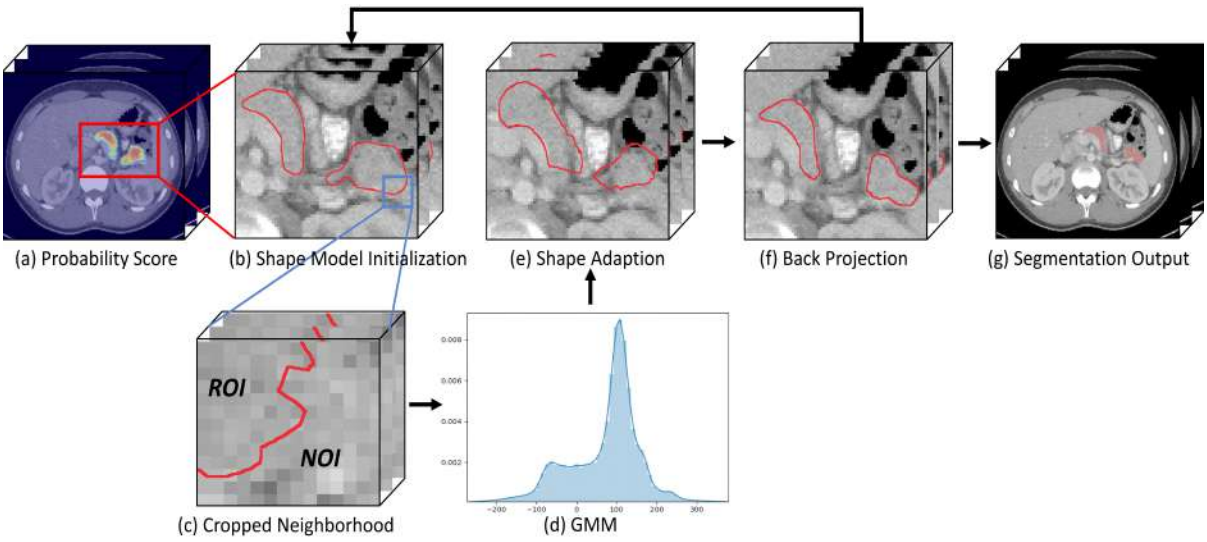


Figure 5.5: This figure illustrates the pipeline of segmentation approach: given the test image with probability map (a), the shape model is initialized to fit the detected region (b); considering the neighborhood region around each landmark (c), a Gaussian Mixture Model is trained (d) to guide shape adaption (e); afterwards, project the shape onto statistical shape model (f); we obtain the segmentation output (g) when the convergence is reached.

5.3.1 Dynamic Weighted Gaussian Mixture Model

To find the maximum likelihood of $p(I, \Pi|C)$, a dynamic weighted Gaussian Mixture Model $DWGMM$ is trained based on the image intensity, under the assumption that all the pixels are statistically independent from each other. In contrast to conventional GMM, the probability map Π is adopted as prior weights of different components in the model, and the probability density is updated with the shape fitting dynamically. Given a single Gaussian distribution:

$$\Psi(x|\mu, \Sigma) = \frac{1}{(2\pi)^{\frac{|x|}{2}}\sqrt{|\Sigma|}} \exp\big(-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)\big),$$

let $X = \{x_1, \ldots, x_{n_K}\}$ be a D-dimension image with n_K pixels, the probability density function (PDF) is defined as:

$$\mathcal{P}(X|\Pi, \Theta) = \prod_{i=1}^{n_K}\{\pi_i\Psi(x_i|\Theta_R) + (1-\pi_i)\Psi(x_i|\Theta_N)\}, \tag{5.4}$$

where the parameters $\{\Theta_R : \mu_R, \Sigma_R\}$ consists of mean value and standard deviation of image intensity, $\Psi(X|\Theta_N)$ is defined in the same way. This $DWGMM$ contains two independent components $\Psi(X|\Theta_R)$ and $\Psi(X|\Theta_N)$ representing ROI and NOI. As a result, the probability of pixel x_i being each component can be estimated from (5.4), we denote $w_R(x_i)$ and $w_N(x_i)$ as the probability of pixel x_i being ROI and NOI:

$$\begin{aligned} w_R(x_i) &= \frac{\pi_i\Psi(x_i|\Theta_R)}{\pi_i\Psi(x_i|\Theta_R) + (1-\pi_i)\Psi(x_i|\Theta_N)} \\ w_N(x_i) &= \frac{(1-\pi_i)\Psi(x_i|\Theta_N)}{\pi_i\Psi(x_i|\Theta_R) + (1-\pi_i)\Psi(x_i|\Theta_N)}. \end{aligned} \tag{5.5}$$

To reduce the non-related pixels' influence on $DWGMM$, only the neighborhood around each landmark is considered in training (cf. Fig. 5.5-(c)). Let $\Omega(x_i)$ donate the cubic neighborhood around the center x_i with radius r, thus each neighborhood contains $(2r+1)^3$ pixels. Let $\Omega^+(x_i)$ be the region inside the shape within $\Omega(x_i)$ and $\Omega^-(x_i) =$

$\Omega(x_i) - \Omega^+(x_i)$ be the outside region (cf. Fig. 5.5-(c)). Therefore, the parameters Θ_R, Θ_N are trained within $\int_{x_i \in C} \Omega^+(x_i)dx$ and $\int_{x_i \in C} \Omega^-(x_i)dx$ respectively. Similarly, the mean probability μ_{wR} and μ_{wN} of being ROI and NOI are obtained by only considering the pixels in region $\int_{i=1}^{N_p} \Omega(x_i)dx$. In this manner, more precise probabilities can be obtained by shrinking the region of neighborhood, leading to finer segmentation.

Theoretically, it would be ideal that the pixels inside shape C have the highest probability of being ROI and the pixels outside shape C have the highest probability of being NOI. Inspired by the popular Mumford-Shah function [61], a novel energy function term is formed:

$$\begin{aligned} -\log(p(I, \Pi|C)) = \int_{i=1}^{N_p} \int_{j \in \Omega(x_i)} & \big(w_R(x_j) - \mu_{wR}\big)^2 + \big(w_N(x_j) - \mu_{wN}\big)^2 \\ & + \big(w_R(x_j) - \mu_{wR}\big)\big(w_N(x_j) - \mu_{wN}\big)dx, \end{aligned} \tag{5.6}$$

at this stage, the landmarks are fitting to superior positions automatically in terms of the probability rules in 5.5. Since the pixels are statistically independent without global constraint, assume the landmark x_i will move along the outward curvature normal with direction $\overrightarrow{\boldsymbol{\jmath}}(x_i)$ to reach the optimal, we compute $\partial(p(I, \Pi|C))/\partial(C) = 0$ to obtain the movement direction $\overrightarrow{\boldsymbol{\jmath}}^*(x_i)$ for each landmark that:

$$\overrightarrow{\boldsymbol{\jmath}}^*(x_i) = \frac{(w_R(x_i) - \mu_{wR})^2 - (w_N(x_i) - \mu_{wN})^2}{(w_R(x_i) - \mu_{wR})(w_N(x_i) - \mu_{wN})}, \tag{5.7}$$

note that for pixels $x_j \in \Omega^+(x_i)$, $\overrightarrow{\boldsymbol{\jmath}}^*(x_j) < 0$, otherwise for pixels $x_j \in \Omega^-(x_i)$, $\overrightarrow{\boldsymbol{\jmath}}^*(x_j) > 0$. Namely, $\overrightarrow{\boldsymbol{\jmath}}^*(x_i) > 0$ indicates x_i moves along the normal to exterior and $\overrightarrow{\boldsymbol{\jmath}}^*(x_i) < 0$ indicates x_i moves along the inverse direction of outward normal to interior.

5.3.2 Shape Prior

Statistical shape models are demonstrated to have a strong ability in global shape constraint. In this work, we employ the proposed RKPCA method to train such a robust

kernel model $RKSSM(S|\mathrm{V},\mathrm{K})$. Specifically, the model statistics is leveraged to correct the erroneous modes and estimate the uncertain pieces (cf. Fig. 5.5-(e) to Fig. 5.5-(f)), which means we only focus on the model back-propagation at this stage. As aforementioned, often the standard pre-image of RKPCA fails to approximate a reconstruction in the shape space, furthermore, the shape to be projected onto the model at this stage already contains certain pieces that are supposed to be preserved. Therefore, the proposed constrained pre-image is used for model back-propagation. Assigning the initialization S_j for the shape C to be projected onto the model, we have $\kappa(C,S_j) = \max\{\kappa(C,S_i) : i = 1,\ldots,N\}$. Employing the shape model in Bayesian model, we consider the prior as:

$$-\log(p(C)) = \|\mathbb{P}_n\Phi(S) - \Phi(C)\|^2 + \omega\,\|\Phi(S_j) - \Phi(C)\|^2\,, \tag{5.8}$$

where ω is a positive trade-off to balance the supervised constraint and the standard pre-image. $\mathbb{P}_n\Phi(x)$ denotes the projection of $\Phi(x)$ onto the principal subspace of Φ. Afterwards, the shape projection is solved by taking gradient $\partial(-\log(p(C)))/\partial(C) = 0$ and the reconstructed shape vector is derived by:

$$\hat{C} = \frac{\sum_{i=1}^{N}\gamma_i\kappa(C,S_i)S_i + \omega S_j}{\sum_{i=1}^{N}\gamma_i\kappa(C,S_i) + \omega}, \quad \gamma_i = \sum_{k=1}^{N}\mathrm{V}_i^j\mathrm{K}_j\mathrm{V}_i^k. \tag{5.9}$$

The Algorithm of Bayesian model in segmentation is illustrated in Alg. 3.

5.4 Evaluation

5.4.1 Dataset

The proposed segmentation framework is validated upon a public NIH pancreas datasets [132], containing 82 abdominal contrast-enhanced 3D CT scans with the size ranges in

Algorithm 3 Algorithm of Segmentation with Bayesian Model

Input: a set of test images $I = \{I_1, \ldots, I_{N_s}\}$, the probability maps $\Pi = \{\Pi_1, \ldots, \Pi_{N_s}\}$, shape model $RKSSM$, radius $r = 3$
Feed shape model to the initial shape C extract from probability map

while neighborhood radius $r \geq 0$ **do**
  Train $\mathcal{P}(X|\Pi, \Theta)$ with current shape C in 5.4
  while not converged **do**
    Train $DWGMM$ in 5.6
    Shape Adaption in terms of 5.7 and obtain the new shape C^*
    if $\|C^* - C\|_2 \leq \epsilon$ **break**
  end while

  Update C by back projection onto $RKSSM$ in 5.9
  Shrink the neighborhood for fine tuning $r = r - 1$
end while

Output: the segment $\hat{Y}$ from the final shape $\hat{C}$

$512 \times 512 \times D$ ($D \in [146, 181]$), voxels resolution varies from $0.5mm$ to $1.0mm$. Note that CT pancreas is very representative as a challenging subject in segmentation.

Data augmentation is performed to artifically generate more datasets following the operations in [132], where affine transformation, i.e. rotation, scaling and translation, as well as non-rigid deformation are applied to each trainig sample. Rotation, scaling and translation are applied to the training samples simulatenously based on a randomly selected degree. Afterwards, the non-rigid deformation t is computed via fitting the TPS (thin-place-spline) to a grid of randomly transformed control points. As a result, 9 more samples are generated from every one training sample. Specifically, around $20,000$ images of 2d slice are used in each network training. For fair comparison, experiments are conducted upon 4 fold cross validation, i.e. the number of test datasets are 20, 20, 21, 21 respectively.

5.4.2 Parameter

On a *GeForce GTX* 1080*ti* GPU, it takes approximately 5 hours to process one fold cross validation. Learning rate is initially set as $lr = 0.01$ and decreased after every 100 epochs. Prior to statistical shape modeling, a point-to-point groupwise correspondence [119] is computed over all the 82 pancreas shapes, each pancreas shape contains 2500 landmarks thereof. The $RKSSM$ is trained following the pipeline as described in Chapter 4. The RBF kernel trick is employed and kernel width $\sigma = 250$ in accordance with the rule in (3.25). In the model back projection, the balance term is set as $\omega = 0.5$. We set $r = 3$ at the beginning in shape adaption with $DWGMM$. The convergence condition value for shape adaption is $\epsilon = 10^{-6}$.

We take the measures Dice Similarity Coefficient $DSC = 2(|Y_+ \cap \hat{Y}_+|)/(|Y_+| + |\hat{Y}_+|)$, Jaccard Index $JI = (|Y_+ \cap \hat{Y}_+|)/(|Y_+| \cup |\hat{Y}_+|)$, Hausdorff distance HD, Recall (true positive rate) and Precision. Note that Table 5.2 and Table 5.3 report the average results of each measure.

5.4.3 Segmentation Results

Table 5.2 reports the comparison of segmentation results from our own approaches and 2d UNet under the same datasets. To be mentioned, previously the SSMs derived using the proposed approach are embedded into an existing ASM segmentation framework, without the consideration of initialization. To evaluate the proposed Bayesian image adaption based on the dynamic weighted Gaussian Mixture Model, we embed the $RKSSM$ with ASM framework that is initialized in terms of the probability score from DenseUNet (cf. the row "ASM + DenseUNet" in Table 5.2).

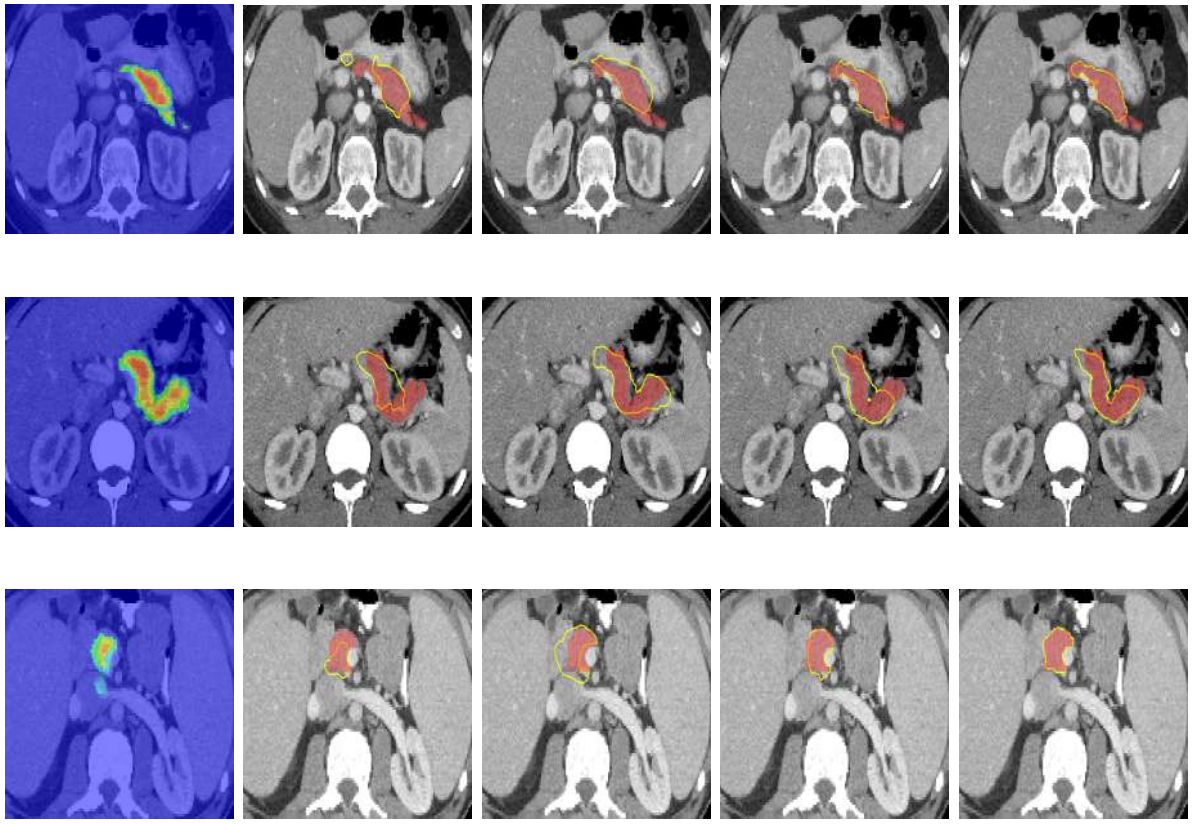
Under the paired t-test with the null hypothesis, the statistics analysis is performed on the segmentation results that the $p-$values are calculated with a significance level of

$\alpha = 0.01$. Conclusions can be drawn from the results in Table 5.2: 2D UNet achieves only 71.25% for the average DSC that is relatively low; our proposed DenseUNet boosts the performance slightly with 2% improvement in terms of the average DSC. Although the DenseUNet has a low accuracy, it is greatly useful in shape model initialization, namely, the ASM segmentation achieves 78.84% of the average DSC. It is worth to mention that with the same initialization, our proposed Bayesian model has significantly better performance than the conventional ASM.

For an intuitive view, the segmentation procedure of Bayesian model (cf. Fig. 5.6), where we compare the segmentation results for Case #1, Case #2, Case #3 at every stage with different cropped neighborhood radius. Specifically, Fig. 5.6-(a) illustrates the primary segmentation result from DenseUNet, (b) indicates the segmentation with radius $r = 3$, (c) indicates the segmentation with radius $r = 2$, (d) indicates the segmentation with radius $r = 1$ and (e) indicates the final result with radius $r = 0$. Apparently, the shape is closer to the ground truth (marked in red) with shrinking the cropped neighborhood radius, namely, the segmentation leads to precise gradually.

5.4.4 Comparison to State-of-the-art

We compare the segmentation results with related works using the same datasets in Table. 5.3. The introduction of the cited works is provided in Section 5.1. In terms of the segmentation results, we report the highest 85.32% average DSC with smallest deviation 4.19, and the DSC for the worse case reaches 71.04%. That is to say, our proposed method is robust to extremely challenging cases. We can also find an improvement of JI. More importantly, we can come to the conclusion that the proposed Bayesian model is efficient and robust in terms of the significant improvement (approximately 12% in DSC) from the neural network segmentation.



(a) (b) (c) (d) (e)

Figure 5.6: Figure shows the segmentation procedure of NIH Case #1 in the first row, Case #2 in the second row and Case #3 in the third row. Note that (a) is the primary segmentation result from deep neural network, (e) indicates the final result and the ground truth is marked in red. Specifically, the average DSC of Case #1 is raised from 68.90% in (a) to 82.17% in (e), the average DSC of Case #2 is raised from 71.52% to 84.83% and that of Case #3 is raised from 69.51% to 78.07%.

5.5 Discussion

Motivated by tackling difficulties in challenging organ segmentation, we integrate deep neural network and statistical shape model within a Bayesian model in this work. A novel optimization principle is proposed to guide segmentation. We conduct experiments on the public NIH pancreas datasets and report the average $DSC = 85.34\%$ that outperforms the state-of-the-art.

Despite the promising results, the proposed Bayesian segmentation pipeline has several limitations. First of all, it is not trained in an end-to-end manner, which makes it difficult to reproduce the results. Secondly, the proposed DenseUNet does not seem efficient in terms of the low accuracy. This is mainly because the ROI of pancreas only occupies a small area across the whole frame of image. Too much background information disturbs the feature extraction. Motivated by this, we aim to propose an end-to-end neural network that embeds the shape prior in the network architecture. The shape prior is designed to filter out background information in training.

Table 5.1: Overview of Pancreas Segmentation

Literature Work (Year)	Approach
Erdt et al. [130] (2011)	a pancreas tissue classifier + SSM
Okada et al. [85] (2015)	a conditional shape-location model + an organ correlation graph (OCG)
Roth et al. [132] (2015)	multi-level deep convolutional networks (ConvNets)
Saito et al. [84] (2016)	level-set-based SSM
Karasawa et al. [131] (2017)	a multi-atlas scheme based on vessel structure around pancreatic tissue
Zhou et al. [141] (2017)	FCN + A fixed-point model
Heinrich et al. [139] (2017)	BRIEFnet
Cai et al. [136] (2017)	a deep convolutional LSTM neural network
Gibson et al. [144] (2017)	Dense Dilated Convolutional network
Farag et al. [138] (2017)	an automated bottom-up approach that generates a hierarchical cascade of information
Zhu et al. [143] (2018)	3D Coarse-to-fine ResDSN
Guo et al. [133] (2018)	UNet + GMM
Roth et al. [137] (2018)	Holistically-nested convolutional networks (HNNs)
Gibson et al. [73] (2018)	a Dense V-network FCN
Yu et al. [142] (2018)	a Recurrent Saliency Transformation network
OKtay et al. [134] (2018)	UNet embedding an attention gate (AG) model
Xia et al. [140] (2018)	a 3D Volumetric Fusion Net (VFN)
Cai et al. [135] (2018)	a CNN-RNN model

Table 5.2: Pancreas segmentation results comparison. Note that the star $*$ indicates a statistically significant difference between the corresponding results and our method at a significant level of 0.01.

Method	DSC	JI	HD	Recall	Precision
Ours	**85.32±4.19**	**74.61 ± 6.19**	**7.11 ± 12.78**	**85.76 ± 4.91**	**86.33 ± 4.10**
ASM + DenseUNet	78.84 ± 5.10*	65.35 ± 6.68*	15.78 ± 11.53*	78.33 ± 5.13*	79.36 ± 7.01*
DenseUNet	73.39 ± 8.78*	58.67 ± 10.47*	18.729 ± 11.84*	73.52 ± 8.20*	75.36 ± 10.66*
UNet	71.25 ± 6.29*	55.60 ± 7.33*	19.38 ± 12.51*	72.08 ± 8.92*	71.53 ± 9.84*

Table 5.3: Pancreas segmentation results comparing with the state-of-the-art. '−' indicates the item is not presented.

Method	Mean DSC	Max DSC	Min DSC	Mean JI
Ours	$\mathbf{85.32 \pm 4.19}$	**91.47**	**71.04**	$\mathbf{74.61 \pm 6.19}$
Our DenseUNet	73.39 ± 8.78	86.50	45.60	58.67 ± 10.47
Xia et al. [140]	84.63 ± 5.07	91.57	61.58	–
Zhu et al. [143]	84.59 ± 4.86	91.45	69.92	–
Yu et al. [142]	84.50 ± 4.97	91.02	62.81	–
Cai et al. [135]	83.70 ± 5.10	91.00	59.00	72.30
Oktay et al. [134]	83.10 ± 3.80	–	–	–
Cai et al. [136]	82.40 ± 6.70	90.10	60.00	70.60 ± 9.00
Zhou et al. [141]	82.37 ± 5.68	90.85	62.43	–
Roth et al. [137]	81.27 ± 6.27	88.96	50.69	68.87
Roth et al. [132]	71.80 ± 4.10	86.90	25.00	–

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In summary, the research focus of this thesis is biological structure segmentation from 3D medical volumes. Considering of the low-contrast to surroundings, large shape variability and limited amount of labeled image data, statistical shape models (SSMs) are chosen as a fundamental approach owing to its considerable success in mitigating these challenges. SSMs extracts the dominant patterns of the target shape, afterwards the learned shape prior can be further used in segmentation associated with an image adaption. Specifically, SSMs parameterize the significant shape variability using a relatively small number of parameters, which makes SSMs-based approaches have compelling advantages of less computational cost in training phase and applicable flexibility in contrast to the recent popular trend deep neural networks.

Nevertheless, it is still demanding to derive a quality model due to the challenges that: (1) training shapes for deriving SSMs may contain arbitrary corruptions caused by factors like the artifacts, noise, non-ideal delineation and erroneous from mesh generation, such outliers degrade the quality of derived SSMs and the segmentation accuracy thereof; (2) it is common seen that the geometrical complexity in the target shape, where

linear and nonlinear patterns exist simultaneously. Furthermore, in practice, even a good quality SSM cannot guarantee the accuracy of segmentation without a power image adaption in cases of challenging situations, e.g. CT pancreas segmentation. In this thesis, many efforts have been made to developing techniques for statistical shape modeling and challenging subject segmentation mitigating the challenges as aforementioned.

Aiming at addressing the training data contamination, early attempts are made to investigate imputation method (cf. our publication [114]) where the high frequency data in the probabilistic distribution are considered as outliers and re-estimated by the proposed method. Owing to the great success in corrupted data recovery, the Robust PCA (RPCA) is leveraged for statistical shape modeling which forms a basis of outlier handling for further methodologies. Subsequently, a weighted RPCA (cf. our publication [27]) is proposed in order to improve the standard RPCA. The motivation behind is to keep the data with higher reliability in the underlying low-rank component in spite of its frequency in probabilistic distribution. Note that the degree of reliability is learned from prior knowledge.

The major contributions are summarized as follows:

- The successful attempts on coping with corruptions in linear distribution motivate us to address the data corruption in nonlinear distribution, as the nonlinear patterns are common seen in realistic scenarios, e.g. the narrow inner space between subparts in a multi-object model. Inspired by the Kernel PCA (KPCA), which maps nonlinear input data onto a high-dimensional implicit feature space and forms a kernel matrix by learning the affinities between each pair of data point, we perform the low-rank modeling on the kernel matrix to achieve dimensionality reduction and data compression thereof. Accordingly, a novel kernelized RPCA is proposed in our publication [95] to handle the data corruption in nonlinear distribution. The

proposed approach is validated upon a set of internal MRI ankle bone volumes, where the shape of ankle bone contains 9 independent sub-bones with narrow inner space. Experimental results show the feasibility of compressing the kernel matrix to achieve nonlinear corrupted data recovery.

- Based on the success of performing low-rank modeling on the kernel matrix, a novel Robust Kernel Principal Component Analysis (RKPCA) is proposed that expands from the motivation of the previous work. The motivation behind is to develop a generic nonlinear data compression technique, which has capability of handling heavily corrupted training data. In addition, a constrained pre-image is proposed to address the convergence problems in standard pre-image of KPCA, which is able to efficiently reconstruct severely corrupted or highly distorted shape data. Extensive experiments are conducted to validate the prosed RKPCA, in order to evaluate the ability of outlier recovery, the quality of derived SSMs as well as the segmentation accuracy using the derived models. Besides, statistical test is employed to prove the significantly better performance of the proposed RKPCA.

- As aforementioned, all the shape-model based approaches are sensitive to initialization, especially for the challenging cases where the subject occupies a small region and has large shape variability. In literature, often an image statistics classifier is leveraged to refine the segmentation. In this thesis, the deep neural networks are utilized as a crucial pre-training to offer an initialization for the shape model, as well as a prior probability score of related pixels. Therefore, a novel segmentation framework is proposed that integrate the shape prior and the probability score trained with deep neural network in a unified Bayesian model. Owing to the flexibility of SSMs, the segmentation is locally refined and outperforms the state-of-the-art in terms of the average DSC, upon the public NIH pancreas datasets.

Besides the algorithmic innovation, efficient implementation of the SSM and ConvNet tools ensures quality of the experimental results: the **Medical Imaging Interaction Toolkit (MITK)** is used for SSMs investigation and analysis; the public library **Keras** is leveraged for deep neural networks development. The proposed methods are evaluated upon a set of internal MRI ankle bone volumes, a set of public CT kidneys from **MICCAI 2015 Multi-atlas** and a set of public **NIH** CT pancreas. It is worth mentioning that all the medical image datasets used in this thesis are validated by clinical experts.

Despite the success in terms of the rich experimental results, the proposed techniques still have limitations that are discussed at the end of each chapter respectively. However, several general limitations should be argued here: first, the shape models are based on a groupwise point-to-point shape correspondence, however, the correspondence procedure itself is a time-consuming and difficult technical problem; second, even though our proposed segmentation approach has achieved the best performance compared to the state-of-the-art, the deep neural networks and shape models are trained individually, which gives rise to redundant computational cost and application time.

6.2 Future Work

With the foundation, two aspects of research directions for future work are presented as follows:

6.2.1 Statistical Shape Modeling Without Correspondence

Similar to the majority of shape modeling approaches, SSMs in this thesis are developed based on an existing point-to-point pre-correspondence. However, the correspondence computation itself is an extremely difficult research direction, which limits the SSMs'

flexibility in practice. Accordingly, it would be interesting to investigate a novel representation of shape for statistical shape modeling, with the aim of alleviating the burden of landmark correspondence. In [147], a novel shape representation called a Disjunctive Normal Shape Model (DNSM) is proposed, where the shapes are formed by the conjunctions of half-spaces. It shows considerable advantages of representing local statistics and does not require landmark correspondence. With the novel representation, the shape prior can be easily embedded into a wide range of machine learning approaches. In [29], a sparse image representation is utilized combining landmark position and appearance prior to create a SSM, instead of one-to-one correspondence. In a nutshell, what we expect is that reformulating the shape prior as a flexible energy function to embed in other frameworks. For instance, the level set is reformulated and combined with the deep neural networks in [75].

6.2.2 Embedding RKPCA into Deep Neural Network

Even though the deep neural networks become preferred recently, they still suffer from high computational cost and GPU memory consumption. By contrast, data compression techniques, e.g. compressive sensing and RPCA, have advantages of less computational cost and parameters in training phase over the deep neural networks. This encourages authors to reformulate such a low-rank modeling as a layer embedding into the deep neural network architecture. In [148], Yang et al. proposed an $ADMM - Net$ that is derived from the iterative procedures in Alternating Direction Method of Multipliers (ADMM) in order to achieve a fast MRI reconstruction. In [149], a novel convolutional neural network frame is proposed that consists of RPCA layers and the classical SVM as a statistics classifier for incomplete face recognition. The RPCA layers are used as filters to extract the low-rank components of the recovered images.

In this thesis, a novel RKPCA is proposed for statistical shape modeling which shows a satisfied ability of nonlinear data compression. It is worth to mention that it has potential to be used in compressive sensing areas. Therefore, it would be feasible to formulate the RKPCA as a layer in deep neural network in order to extract dominant nonlinear patterns and remove the sparsity for segmentation. Moreover, only a small number of parameters is required in training such a RKPCA layer, namely, it will not bring overmuch computational cost to the whole.

References

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2015, pp. 234–241.

[2] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[3] R. Davies, C. Twining, and C. Taylor, *Statistical models of shape: Optimisation and evaluation*. Springer Science & Business Media, 2008.

[4] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[5] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, p. 11, 2011.

[6] G. Huang, Z. Liu, K. Q. Weinberger, and L. van der Maaten, "Densely connected convolutional networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, no. 2, 2017, p. 3.

[7] C. Ledig, R. A. Heckemann, A. Hammers, J. C. Lopez, V. F. Newcombe, A. Makropoulos, J. Lötjönen, D. K. Menon, and D. Rueckert, "Robust whole-brain segmentation: application to traumatic brain injury," *Medical image analysis*, vol. 21, no. 1, pp. 40–58, 2015.

[8] M. J. Cardoso, K. Leung, M. Modat, S. Keihaninejad, D. Cash, J. Barnes, N. C. Fox, S. Ourselin, A. D. N. Initiative *et al.*, "Steps: Similarity and truth estimation for propagated segmentations and its application to hippocampal segmentation and brain parcelation," *Medical image analysis*, vol. 17, no. 6, pp. 671–684, 2013.

[9] A. de Brebisson and G. Montana, "Deep neural networks for anatomical brain segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 20–28.

[10] H. Chen, Q. Dou, L. Yu, J. Qin, and P.-A. Heng, "Voxresnet: Deep voxelwise residual networks for brain segmentation from 3d mr images," *NeuroImage*, 2017.

[11] A. Mansoor, J. J. Cerrolaza, R. Idrees, E. Biggs, M. A. Alsharid, R. A. Avery, and M. G. Linguraru, "Deep learning guided partitioned shape model for anterior visual pathway segmentation," *IEEE transactions on medical imaging*, vol. 35, no. 8, pp. 1856–1865, 2016.

[12] S. Candemir, S. Jaeger, K. Palaniappan, J. P. Musco, R. K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, and C. J. McDonald, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," *IEEE transactions on medical imaging*, vol. 33, no. 2, pp. 577–590, 2014.

[13] Z. Xu, U. Bagci, B. Foster, A. Mansoor, J. K. Udupa, and D. J. Mollura, "A hybrid method for airway segmentation and automated measurement of bronchial wall thickness on ct," *Medical image analysis*, vol. 24, no. 1, pp. 1–17, 2015.

[14] R. D. Rudyanto, S. Kerkstra, E. M. Van Rikxoort, C. Fetita, P.-Y. Brillet, C. Lefevre, W. Xue, X. Zhu, J. Liang, İ. Öksüz *et al.*, "Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the vessel12 study," *Medical image analysis*, vol. 18, no. 7, pp. 1217–1232, 2014.

[15] M. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri," *Medical image analysis*, vol. 30, pp. 108–119, 2016.

[16] D. Mahapatra, "Automatic cardiac segmentation using semantic information from random forests," *Journal of digital imaging*, vol. 27, no. 6, pp. 794–804, 2014.

[17] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of mri," *Medical image analysis*, vol. 31, pp. 77–87, 2016.

[18] M. U. Akram and S. A. Khan, "Multilayered thresholding-based blood vessel segmentation for screening of diabetic retinopathy," *Engineering with computers*, vol. 29, no. 2, pp. 165–173, 2013.

[19] X. Zhang, G. Thibault, E. Decencière, B. Marcotegui, B. Laÿ, R. Danno, G. Cazuguel, G. Quellec, M. Lamard, P. Massin *et al.*, "Exudate detection in color retinal images for mass screening of diabetic retinopathy," *Medical image analysis*, vol. 18, no. 7, pp. 1026–1043, 2014.

[20] G. Li, X. Chen, F. Shi, W. Zhu, J. Tian, and D. Xiang, "Automatic liver segmentation based on shape constraints and deformable graph cut in ct images," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5315–5329, 2015.

[21] T. Norajitra and K. H. Maier-Hein, "3d statistical shape models incorporating landmark-wise random regression forests for omni-directional landmark detection," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 155–168, 2017.

[22] Y. Huo, J. Liu, Z. Xu, R. L. Harrigan, A. Assad, R. G. Abramson, and B. A. Landman, "Robust multicontrast mri spleen segmentation for splenomegaly using multi-atlas segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 2, pp. 336–343, 2018.

[23] Y. Guo, Y. Gao, and D. Shen, "Deformable mr prostate segmentation via deep feature learning and sparse patch matching," *IEEE transactions on medical imaging*, vol. 35, no. 4, pp. 1077–1089, 2016.

[24] G. Litjens, R. Toth, W. van de Ven, C. Hoeks, S. Kerkstra, B. van Ginneken, G. Vincent, G. Guillard, N. Birbeck, J. Zhang *et al.*, "Evaluation of prostate segmentation algorithms for mri: the promise12 challenge," *Medical image analysis*, vol. 18, no. 2, pp. 359–373, 2014.

[25] S. S. Chandra, Y. Xia, C. Engstrom, S. Crozier, R. Schwarz, and J. Fripp, "Focused shape models for hip joint segmentation in 3d magnetic resonance images," *Medical image analysis*, vol. 18, no. 3, pp. 567–578, 2014.

[26] A. Rasoulian, R. Rohling, and P. Abolmaesumi, "Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+ pose model," *IEEE transactions on medical imaging*, vol. 32, no. 10, pp. 1890–1900, 2013.

[27] J. Ma, F. Lin, J. Honsdorf, K. Lentzen, S. Wesarg, and M. Erdt, "Weighted robust pca for statistical shape modeling," in *International Conference on Medical Imaging and Virtual Reality (MIAR'16), Bern, Switzerland*. Springer, 2016, pp. 343–353.

[28] S. Steger, F. Jung, and S. Wesarg, "Personalized articulated atlas with a dynamic adaptation strategy for bone segmentation in ct or ct/mr head and neck images," in *SPIE Medical Imaging: Image Processing*, vol. 9034. International Society for Optics and Photonics, 2014, p. 90341I.

[29] J. Krüger, J. Ehrhardt, and H. Handels, "Statistical appearance models based on probabilistic correspondences," *Medical Image Analysis*, vol. 37, pp. 146–159, 2017.

[30] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest *et al.*, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE transactions on medical imaging*, vol. 34, no. 10, p. 1993, 2015.

[31] G. Litjens, O. Debats, J. Barentsz, N. Karssemeijer, and H. Huisman, "Computer-aided detection of prostate cancer in mri," *IEEE transactions on medical imaging*, vol. 33, no. 5, pp. 1083–1092, 2014.

[32] M. Veta, J. P. Pluim, P. J. Van Diest, and M. A. Viergever, "Breast cancer histopathology image analysis: A review," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1400–1411, 2014.

[33] J. Jayender, E. Gombos, S. Chikarmane, D. Dabydeen, F. A. Jolesz, and K. G. Vosburgh, "Statistical learning algorithm for in situ and invasive breast carcinoma segmentation," *Computerized Medical Imaging and Graphics*, vol. 37, no. 4, pp. 281–292, 2013.

[34] Y. Yuan, M. Chao, and Y.-C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance," *IEEE Trans. Med. Imaging*, vol. 36, no. 9, pp. 1876–1886, 2017.

[35] Y. Song, E.-L. Tan, X. Jiang, J.-Z. Cheng, D. Ni, S. Chen, B. Lei, and T. Wang, "Accurate cervical cell segmentation from overlapping clumps in pap smear images," *IEEE transactions on medical imaging*, vol. 36, no. 1, pp. 288–300, 2017.

[36] H. Chen, X. Qi, L. Yu, and P.-A. Heng, "Dcan: deep contour-aware networks for accurate gland segmentation," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2016, pp. 2487–2496.

[37] K. Sirinukunwattana, J. P. Pluim, H. Chen, X. Qi, P.-A. Heng, Y. B. Guo, L. Y. Wang, B. J. Matuszewski, E. Bruni, U. Sanchez *et al.*, "Gland segmentation in colon histology images: The glas challenge contest," *Medical image analysis*, vol. 35, pp. 489–502, 2017.

[38] A. A. A. Setio, A. Traverso, T. De Bel, M. S. Berens, C. van den Bogaard, P. Cerello, H. Chen, Q. Dou, M. E. Fantacci, B. Geurts *et al.*, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge," *Medical image analysis*, vol. 42, pp. 1–13, 2017.

[39] S. Wang, M. Zhou, Z. Liu, Z. Liu, D. Gu, Y. Zang, D. Dong, O. Gevaert, and J. Tian, "Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation," *Medical image analysis*, vol. 40, pp. 172–183, 2017.

[40] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2013, pp. 246–253.

[41] Z. Yu-qian, G. Wei-hua, C. Zhen-cheng, T. Jing-tian, and L. Ling-Yun, "Medical images edge detection based on mathematical morphology," in *IEEE Annual International Conference on Engineering in Medicine and Biology Society*. IEEE, 2006, pp. 6492–6495.

[42] R. Pohle and K. D. Toennies, "Segmentation of medical images using adaptive region growing," in *SPIE Medical Imaging: Image Processing*, vol. 4322. International Society for Optics and Photonics, 2001, pp. 1337–1347.

[43] H. Ng, S. Ong, K. Foong, P. Goh, and W. Nowinski, "Medical image segmentation using k-means clustering and improved watershed algorithm," in *IEEE Southwest Symposium on Image Analysis and Interpretation*. IEEE, 2006, pp. 61–65.

[44] D.-Q. Zhang and S.-C. Chen, "A novel kernelized fuzzy c-means algorithm with application in medical image segmentation," *Artificial Intelligence in Medicine*, vol. 32, no. 1, pp. 37–50, 2004.

[45] S. Wang, W. Zhu, and Z.-P. Liang, "Shape deformation: Svm regression and application to medical image segmentation," in *IEEE International Conference on Computer Vision (ICCV)*, vol. 2. IEEE, 2001, pp. 209–216.

[46] H. A. Vrooman, C. A. Cocosco, F. van der Lijn, R. Stokking, M. A. Ikram, M. W. Vernooij, M. M. Breteler, and W. J. Niessen, "Multi-spectral brain tissue segmentation using automatically trained k-nearest-neighbor classification," *Neuroimage*, vol. 37, no. 1, pp. 71–81, 2007.

[47] A. Liaw, M. Wiener *et al.*, "Classification and regression by randomforest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.

[48] X. Yang, X. Gao, D. Tao, X. Li, and J. Li, "An efficient mrf embedded level set method for image segmentation," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 9–21, 2015.

[49] J. Lafferty, A. McCallum, and F. C. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," 2001.

[50] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in nd images," in *IEEE International Conference on Computer Vision (ICCV)*, vol. 1. IEEE, 2001, pp. 105–112.

[51] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *IEEE Conference on Computer Vision and Pattern recognition*, 2015, pp. 3431–3440.

[52] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[53] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *International Journal of Computer Vision*, vol. 72, no. 2, pp. 195–215, 2007.

[54] H. Park, P. H. Bland, and C. R. Meyer, "Construction of an abdominal probabilistic atlas and its application in segmentation," *IEEE Transactions on Medical Imaging*, vol. 22, no. 4, pp. 483–492, 2003.

[55] B. Patenaude, S. M. Smith, D. N. Kennedy, and M. Jenkinson, "A bayesian model of shape and appearance for subcortical brain segmentation," *Neuroimage*, vol. 56, no. 3, pp. 907–922, 2011.

[56] D. Mahapatra, "Analyzing training information from random forests for improved image segmentation," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1504–1512, 2014.

[57] C. Jin, F. Shi, D. Xiang, X. Jiang, B. Zhang, X. Wang, W. Zhu, E. Gao, and X. Chen, "3d fast automatic segmentation of kidney based on modified aam and random forest," *IEEE Transactions on Medical Imaging*, vol. 35, no. 6, pp. 1395–1407, 2016.

[58] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, 2001.

[59] X. Chen, J. K. Udupa, U. Bagci, Y. Zhuge, and J. Yao, "Medical image segmentation by combining graph cuts and oriented active appearance models," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2035–2046, 2012.

[60] M. Rajchl, M. C. Lee, O. Oktay, K. Kamnitsas, J. Passerat-Palmbach, W. Bai, M. Damodaram, M. A. Rutherford, J. V. Hajnal, B. Kainz *et al.*, "Deepcut: Object segmentation from bounding box annotations using convolutional neural networks," *IEEE Transactions on Medical Imaging*, vol. 36, no. 2, pp. 674–683, 2017.

[61] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[62] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3d medical image segmentation: a review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543–563, 2009.

[63] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: learning dense volumetric segmentation from sparse annotation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2016, pp. 424–432.

[64] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P. H. Torr, "Conditional random fields as recurrent neural networks," in *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1529–1537.

[65] P. F. Christ, M. E. A. Elshaer, F. Ettlinger, S. Tatavarty, M. Bickel, P. Bilic, M. Rempfler, M. Armbruster, F. Hofmann, M. DAnastasi *et al.*, "Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2016, pp. 415–423.

[66] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[67] Z. Liu, X. Li, P. Luo, C. C. Loy, and X. Tang, "Deep learning markov random field for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[68] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin, and P.-A. Heng, "3d deeply supervised network for automated segmentation of volumetric medical images," *Medical Image Analysis*, vol. 41, pp. 40–54, 2017.

[69] F. Milletari, A. Rothberg, J. Jia, and M. Sofka, "Integrating statistical prior knowledge into convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2017, pp. 161–168.

[70] D. Yang, D. Xu, S. K. Zhou, B. Georgescu, M. Chen, S. Grbic, D. Metaxas, and D. Comaniciu, "Automatic liver segmentation using an adversarial image-to-image network," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2017, pp. 507–515.

[71] R. Li, T. Zeng, H. Peng, and S. Ji, "Deep learning segmentation of optical microscopy images improves 3-d neuron reconstruction," *IEEE Transactions on Medical Imaging*, vol. 36, no. 7, pp. 1533–1541, 2017.

[72] S. Manivannan, W. Li, J. Zhang, E. Trucco, and S. J. McKenna, "Structure prediction for gland segmentation with hand-crafted and deep convolutional features," *IEEE Transactions on Medical Imaging*, vol. 37, no. 1, pp. 210–221, 2018.

[73] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. P. Pereira, M. J. Clarkson, and D. C. Barratt, "Automatic multi-organ segmentation on abdominal ct with dense v-networks," *IEEE Transactions on Medical Imaging*, 2018.

[74] M. M. Farhangi, H. Frigui, A. Seow, and A. A. Amini, "3-d active contour segmentation based on sparse linear combination of training shapes (scots)," *IEEE Transactions on Medical Imaging*, vol. 36, no. 11, pp. 2239–2249, 2017.

[75] N. Le, K. G. Quach, K. Luu, M. Savvides, and C. Zhu, "Reformulating level sets as deep recurrent neural network approach to semantic segmentation," *arXiv preprint arXiv:1704.03593*, 2017.

[76] S. Nouranian, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, "A multi-atlas-based segmentation framework for prostate brachytherapy," *IEEE transactions on medical imaging*, vol. 34, no. 4, pp. 950–961, 2015.

[77] C. Wachinger, K. Fritscher, G. Sharp, and P. Golland, "Contour-driven atlas-based segmentation," *IEEE transactions on medical imaging*, vol. 34, no. 12, pp. 2492–2505, 2015.

[78] C. Wang and S. Mahadevan, "Manifold alignment using procrustes analysis," in *International Conference on Machine Learning (ICML)*. ACM, 2008, pp. 1120–1127.

[79] M. Erdt, M. Kirschner, S. Steger, and S. Wesarg, "Fast automatic liver segmentation combining learned shape priors with observed shape deviation," in *IEEE International Symposium on Computer-Based Medical Systems*. IEEE, 2010, pp. 249–254.

[80] A. Saito, S. Yamamoto, S. Nawano, and A. Shimizu, "Automated liver segmentation from a postmortem ct scan based on a statistical shape model," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 2, pp. 205–221, 2017.

[81] A. Soliman, F. Khalifa, A. Elnakib, M. A. El-Ghar, N. Dunlap, B. Wang, G. Gimelfarb, R. Keynton, and A. El-Baz, "Accurate lungs segmentation on ct chest images by adaptive appearance-guided shape modeling," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 263–276, 2017.

[82] X. Albà, M. Pereañez, C. Hoogendoorn, A. J. Swift, J. M. Wild, A. F. Frangi, and K. Lekadir, "An algorithm for the segmentation of highly abnormal hearts using a generic statistical shape model," *IEEE Transactions on Medical Imaging*, vol. 35, no. 3, pp. 845–859, 2016.

[83] S. Faghih Roohi and R. Aghaeizadeh Zoroofi, "4d statistical shape modeling of the left ventricle in cardiac mr images," *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–17, 2013.

[84] A. Saito, S. Nawano, and A. Shimizu, "Joint optimization of segmentation and shape prior from level-set-based statistical shape model, and its application to the automated segmentation of abdominal organs," *Medical Image Analysis*, vol. 28, pp. 46–65, 2016.

[85] T. Okada, M. G. Linguraru, M. Hori, R. M. Summers, N. Tomiyama, and Y. Sato, "Abdominal multi-organ segmentation from ct images using conditional shape–location and unsupervised intensity priors," *Medical Image Analysis*, vol. 26, no. 1, pp. 1–18, 2015.

[86] J. E. Iglesias and M. R. Sabuncu, "Multi-atlas segmentation of biomedical images: a survey," *Medical Image Analysis*, vol. 24, no. 1, pp. 205–219, 2015.

[87] M. Wilms, H. Handels, and J. Ehrhardt, "Multi-resolution multi-object statistical shape models based on the locality assumption," *Medical Image Analysis*, vol. 38, pp. 17–29, 2017.

[88] S. Sun, C. Bauer, and R. Beichel, "Automated 3-d segmentation of lungs with lung cancer in ct data using a novel robust active shape model approach," *IEEE Transactions on Medical Imaging*, vol. 31, no. 2, pp. 449–460, 2012.

[89] M. Wilms, J. Ehrhardt, and H. Handels, "A 4d statistical shape model for automated segmentation of lungs with large tumors," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2012, pp. 347–354.

[90] M. Kirschner, M. Becker, and S. Wesarg, "3d active shape model segmentation with nonlinear shape priors," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2011, pp. 492–499.

[91] M. Pereañez, K. Lekadir, I. Castro-Mateos, J. M. Pozo, Á. Lazáry, and A. F. Frangi, "Accurate segmentation of vertebral bodies and processes using statistical shape decomposition and conditional models," *IEEE Transactions on Medical Imaging*, vol. 34, no. 8, pp. 1627–1639, 2015.

[92] I. Castro-Mateos, J. M. Pozo, M. Pereañez, K. Lekadir, A. Lazary, and A. F. Frangi, "Statistical interspace models (sims): Application to robust 3d spine segmentation," *IEEE Transactions on Medical Imaging*, vol. 34, no. 8, pp. 1663–1675, 2015.

[93] E. M. A. Anas, A. Rasoulian, A. Seitel, K. Darras, D. Wilson, P. S. John, D. Pichora, P. Mousavi, R. Rohling, and P. Abolmaesumi, "Automatic segmentation of wrist bones in ct using a statistical wrist shape + pose model," *IEEE Transactions on Medical Imaging*, vol. 35, no. 8, pp. 1789–1801, 2016.

[94] X. Chen, J. Graham, C. Hutchinson, and L. Muir, "Automatic generation of statistical pose and shape models for articulated joints," *IEEE Transactions on Medical Imaging*, vol. 33, no. 2, pp. 372–383, 2014.

[95] J. Ma, A. Wang, F. Lin, S. Wesarg, and M. Erdt, "Nonlinear statistical shape modeling for ankle bone segmentation using a novel kernelized robust pca," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2017, pp. 136–143.

[96] C. von Tycowicz, F. Ambellan, A. Mukhopadhyay, and S. Zachow, "An efficient riemannian statistical shape model using differential coordinates: With application to the classification of data from the osteoarthritis initiative," *Medical Image Analysis*, vol. 43, pp. 1–9, 2018.

[97] S. Mika, B. Schölkopf, A. J. Smola, K.-R. Müller, M. Scholz, and G. Rätsch, "Kernel pca and de-noising in feature spaces," in *Advances in Neural Information Processing Systems (NIPS)*, 1999, pp. 536–542.

[98] C. J. Twining and C. J. Taylor, "Kernel principal component analysis and the construction of non-linear active shape models." in *British Machine Vision Conference (BMVC)*, vol. 1, 2001, pp. 23–32.

[99] M. Lüthi, T. Albrecht, and T. Vetter, "Building shape models from lousy data," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2009, pp. 1–8.

[100] B. Gutierrez, D. Mateus, E. Shiban, B. Meyer, J. Lehmberg, and N. Navab, "A sparse approach to build shape models with routine clinical data," in *IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2014, pp. 258–261.

[101] C. Shi, Y. Cheng, J. Wang, Y. Wang, K. Mori, and S. Tamura, "Low-rank and sparse decomposition based shape model and probabilistic atlas for automatic pathological organ segmentation," *Medical Image Analysis*, vol. 38, pp. 30–49, 2017.

[102] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[103] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.

[104] L. Zhang, Z. Chen, M. Zheng, and X. He, "Robust non-negative matrix factorization," *Frontiers of Electrical and Electronic Engineering in China*, vol. 6, no. 2, pp. 192–200, 2011.

[105] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *International Conference on Machine Learning (ICML)*, 2010, pp. 663–670.

[106] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, p. 717, 2009.

[107] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Advances in Neural Information Processing Systems (NIPS)*, 2001, pp. 556–562.

[108] M. Jaggi, M. Sulovsk *et al.*, "A simple algorithm for nuclear norm regularized problems," in *International Conference on Machine Learning (ICML)*, 2010, pp. 471–478.

[109] Y. Xiao, Z. Li, T. Yang, and L. Zhang, "Svd-free convex-concave approaches for nuclear norm regularization," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 2017, pp. 3126–3132.

[110] J. Mairal, J. Ponce, G. Sapiro, A. Zisserman, and F. R. Bach, "Supervised dictionary learning," in *Advances in Neural Information Processing Systems (NIPS)*, 2009, pp. 1033–1040.

[111] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *International Conference on Machine Learning (ICML)*. ACM, 2009, pp. 689–696.

[112] Y. Xu, L. Duan, S. Lin, X. Chen, D. W. K. Wong, T. Y. Wong, and J. Liu, "Optic cup segmentation for glaucoma detection using low-rank superpixel representation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2014, pp. 788–795.

[113] C. Lu, Z. Lin, and S. Yan, "Smoothed low rank and sparse matrix recovery by iteratively reweighted least squares minimization," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 646–654, 2015.

[114] M. Jingting, K. Lentzen, J. Honsdorf, L. Feng, and M. Erdt, "Statistical shape modeling from gaussian distributed incomplete data for image segmentation," in *4th MICCAI Workshop on Clinical Image-based Procedures: Translational Research in Medical Imaging (CLIP'15)*. Springer, 2015, pp. 113–121.

[115] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *arXiv preprint arXiv:1009.5055*, 2010.

[116] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[117] J. T. Kwok and I. W. Tsang, "The pre-image problem in kernel methods," in *International Conference on Machine Learning (ICML)*, 2003, pp. 408–415.

[118] S. Xiao, M. Tan, D. Xu, and Z. Y. Dong, "Robust kernel low-rank representation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 11, pp. 2268–2281, 2016.

[119] M. Kirschner and S. Wesarg, "Construction of groupwise consistent shape parameterizations by propagation," in *SPIE Medical Imaging: Image Processing*, 2010.

[120] N. Kwak, "Nonlinear projection trick in kernel methods: An alternative to the kernel trick," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 12, pp. 2113–2119, 2013.

[121] H. Qi and S. Hughes, "Using the kernel trick in compressive sensing: Accurate signal recovery from fewer measurements," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2011, pp. 3940–3943.

[122] F. P. Anaraki and S. M. Hughes, "Kernel compressive sensing," in *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2013, pp. 494–498.

[123] J.-K. Im, D. W. Apley, and G. C. Runger, "Tangent hyperplane kernel principal component analysis for denoising," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 4, pp. 644–656, 2012.

[124] H. Nguyen, W. Yang, F. Shen, and C. Sun, "Kernel low-rank representation for face recognition," *Neurocomputing*, vol. 155, pp. 32–42, 2015.

[125] L. Zhang, T. Yang, J. Yi, R. Jin, and Z.-H. Zhou, "Stochastic optimization for kernel pca." in *AAAI*, 2016, pp. 2315–2322.

[126] L. Wu and Y. Wang, "Robust hashing for multi-view data: Jointly learning low-rank kernelized similarity consensus and hash functions," *Image and Vision Computing*, vol. 57, pp. 58–66, 2017.

[127] M. H. Nguyen and F. Torre, "Robust kernel principal component analysis," in *Advances in Neural Information Processing Systems (NIPS)*, 2009, pp. 1185–1192.

[128] B. Landman, Z. Xu, J. Eugenio Igelsias, M. Styner, T. Robin Langerak, and A. Klein (eds), "2015 miccai multi-atlas labeling beyond the cranial vault - workshop and challenge." doi:10.7303/syn3193805.

[129] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 142–158, 2016.

[130] M. Erdt, M. Kirschner, K. Drechsler, S. Wesarg, M. Hammon, and A. Cavallaro, "Automatic pancreas segmentation in contrast enhanced ct data using learned spatial anatomy and texture descriptors," in *IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2011, pp. 2076–2082.

[131] K. Karasawa, M. Oda, T. Kitasaka, K. Misawa, M. Fujiwara, C. Chu, G. Zheng, D. Rueckert, and K. Mori, "Multi-atlas pancreas segmentation: Atlas selection based on vessel structure," *Medical Image Analysis*, vol. 39, pp. 18–28, 2017.

[132] H. R. Roth, L. Lu, A. Farag, H.-C. Shin, J. Liu, E. B. Turkbey, and R. M. Summers, "Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2015, pp. 556–564.

[133] Z. Quo, L. Zhang, L. Lu, M. Bagheri, R. M. Summers, M. Sonka, and J. Yao, "Deep logismos: Deep learning graph-based 3d segmentation of pancreatic tumors on ct scans," in *IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2018, pp. 1230–1233.

[134] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.

[135] J. Cai, L. Lu, F. Xing, and L. Yang, "Pancreas segmentation in ct and mri images via domain specific network designing and recurrent neural contextual learning," *arXiv preprint arXiv:1803.11303*, 2018.

[136] J. Cai, L. Lu, Y. Xie, F. Xing, and L. Yang, "Pancreas segmentation in mri using graph-based decision fusion on convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2017, pp. 674–682.

[137] H. R. Roth, L. Lu, N. Lay, A. P. Harrison, A. Farag, A. Sohn, and R. M. Summers, "Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation," *Medical Image Analysis*, vol. 45, pp. 94–107, 2018.

[138] A. Farag, L. Lu, H. R. Roth, J. Liu, E. Turkbey, and R. M. Summers, "A bottom-up approach for pancreas segmentation using cascaded superpixels and (deep) image patch labeling," *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 386–399, 2017.

[139] M. P. Heinrich and O. Oktay, "Briefnet: deep pancreas segmentation using binary sparse convolutions," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2017, pp. 329–337.

[140] Y. Xia, L. Xie, F. Liu, Z. Zhu, E. K. Fishman, and A. L. Yuille, "Bridging the gap between 2d and 3d organ segmentation," *arXiv preprint arXiv:1804.00392*, 2018.

[141] Y. Zhou, L. Xie, W. Shen, Y. Wang, E. K. Fishman, and A. L. Yuille, "A fixed-point model for pancreas segmentation in abdominal ct scans," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2017, pp. 693–701.

[142] Q. Yu, L. Xie, Y. Wang, Y. Zhou, E. K. Fishman, and A. L. Yuille, "Recurrent saliency transformation network: Incorporating multi-stage visual cues for

small organ segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[143] Z. Zhu, Y. Xia, W. Shen, E. K. Fishman, and A. L. Yuille, "A 3d coarse-to-fine framework for automatic pancreas segmentation," in *International Conference on 3D Vision (3DV)*. IEEE, 2018.

[144] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. R. Davidson, S. P. Pereira, M. J. Clarkson, and D. C. Barratt, "Towards image-guided pancreas and biliary endoscopy: automatic multi-organ segmentation on abdominal ct with dense dilated networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2017, pp. 728–736.

[145] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[146] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor Fusion IV: Control Paradigms and Data Structures*, vol. 1611. International Society for Optics and Photonics, 1992, pp. 586–607.

[147] F. Mesadi, E. Erdil, M. Cetin, and T. Tasdizen, "Image segmentation using disjunctive normal bayesian shape and appearance models," *IEEE Transactions on Medical Imaging*, 2018.

[148] Y. Yang, J. Sun, H. Li, and Z. Xu, "Admm-net: A deep learning approach for compressive sensing mri," *arXiv preprint arXiv:1705.06869*, 2017.

[149] J. Zhao, Y. Lv, Z. Zhou, and F. Cao, "A novel deep learning algorithm for incomplete face recognition: Low-rank-recovery network," *Neural Networks*, vol. 94, pp. 115–124, 2017.

Publications

Journal Papers

(i) **Jingting Ma**, Anqi Wang, Feng Lin, Stefan Wesarg, and Marius Erdt. "A Novel Robust Kernel Principal Component Analysis for Statistical Shape Modeling from Erroneous Data." (Under review of Medical Image Analysis 2018).

Conference Papers

(i) **Jingting Ma**, Feng Lin, Stefan Wesarg, and Marius Erdt. "A Novel Bayesian Model Incorporating Deep Neural Network and Statistical Shape Model for Pancreas Segmentation." In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'18), Granada, Spain, 16-20 September, 2018.

(ii) **Jingting Ma**, Anqi Wang, Feng Lin, Stefan Wesarg, and Marius Erdt. "Nonlinear Statistical Shape Modeling for Ankle Bone Segmentation Using a Novel Kernelized Robust PCA." In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'17), Quebec City, Canada, 10-14 September, 2017, pp. 136-143.

(iii) **Jingting Ma**, Feng Lin, Jonas Honsdorf, Katharina Lentzen, Stefan Wesarg, and Marius Erdt. "Weighted Robust PCA for Statistical Shape Modeling." In International Conference on Medical Imaging and Virtual Reality (MIAR'16), Bern, Switzerland, August 24-26, 2016, pp. 343-353.

(iv) **Jingting Ma**, Katharina Lentzen, Jonas Honsdorf, Lin Feng, and Marius Erdt. "Statistical Shape Modeling from Gaussian Distributed Incomplete Data for Image Segmentation." 4th MICCAI Workshop on Clinical Image-based Procedures: Translational Research in Medical Imaging (CLIP'15), Munich, Germany, 5 October, 2015, pp. 113-121.