ABSTRACT

POLIANNIKOV, OLEG VICTOROVITCH. On Shape Description and Optimization for Object Classification. (Under the direction of Prof. Hamid Krim)

The purpose of this thesis has been to study several problems typical for shape analysis, computer vision and image processing in general. First, a problem of sampling planar shape and continuous curves is considered. It has been shown that samples of a discrete curve or a surface should contain the information about the object itself as well as the coordinate system used to produce a functional parameterization. The sampling algorithm has been designed as a two-cost optimization problem. Secondly, we have proposed a new algorithmic approach to implementing a simulated annealing type of optimization, which is based on the notion of scale. The scale is defined as a unit time interval used when converting a continuous-time evolution into a discrete one. It has been shown that by varying the scale as opposed to keeping it constant, one can achieve a better performance of an optimization algorithm, all in the absence of a complicated acceptance/rejection Markov chain mechanism. Finally, we have studied the problem of identification of symmetric planar shapes from a single view. It has been shown that the proposed notion of a skeleton of an arbitrary view of the shape is instrumental in constructing invariants enabling us to perform identification. It has further been demonstrated that for the case of noisy data one can describe the distribution of the skeleton points and thus construct an optimal skeleton estimation technique.

ON SHAPE DESCRIPTION AND OPTIMIZATION FOR OBJECT CLASSIFICATION

by

OLEG V. POLIANNIKOV

A dissertation submitted to the Graduate

Faculty of North Carolina State University

in partial fulfillment of the requirements for

the Degree of Doctor Of Philosophy

ELECTRICAL ENGINEERING

Raleigh

February 2003

APPROVED BY THE ADVISORY COMMITTEE:

Chair:	_____	Prof. Hamid Krim
Member:	_____	Prof. Jean-Pierre Fouque
Member:	_____	Prof. Brian L. Hughes
Member:	_____	Prof. J. Keith Townsend

To the memory of my grandfather, Prof. Mikhail N. Yakimenko (1931–2001)

PERSONAL BIOGRAPHY

Oleg V. Poliannikov received his M.S. in Mathematics from Moscow State University, Moscow, Russia in 1998. In 1999 he started his Ph.D. study in Electrical Engineering at North Carolina State University, Raleigh, NC. His research interests include random processes, shape theory and image processing.

ACKNOWLEDGMENTS

Working on this thesis has been many things, but first and foremost it has been four years of my life. Here I would like to acknowledge people who helped me not only survive through that time, but turn it into an exciting and unforgettable experience.

A lot of credit is certainly due to my advisor Prof. Hamid Krim. I would like to thank him for suggesting the problems to work on, offering me his ideas, making an often humanly impossible effort to keep me focused on the goals and finally paying my bills. I would further like to express my gratitude to the members of my committee: Prof. Brian L. Hughes, Prof. J. Keith Townsend and Prof. Jean-Pierre Fouque for their willingness to serve on the committee as well as constructive and friendly suggestions. I thank Prof. Arne Nilsson for having served on my committee during the initial phase of my program and Prof. H. Joel Trussell for having being a member of the committee as well as for the enormous amount of time that he had spent with me while I was working on the first part of this thesis.

I would like to thank the current as well as former students from our group whom I have enjoyed spending time with all these years. Those include Dr. Yun He, Dr. Bilge Karacali, Dr. Gozde Bozkurt Unal, Dr. Yufang Bao, Dr. A. Ben Hamza, Viraj Mehta, Aysegul Gunduz, Yang Wu, Sajjad Baloch, Christina Hammoc and many others. I'm grateful to all these people for providing me a great work environment as well simply being my good friends all this time.

Last but not the least, I want to thank all my family, including my wife here and my mother, sister and the rest of the relatives overseas. They have always been very supportive of me, and have been passionately waiting for a successful ending to this chapter of my life. For that my grandfather was one of those who took my life exceptionally personally and yet did not live to see me cross the final line, I dedicate this thesis to his memory.

TABLE OF CONTENTS

Page

LIST OF FIGURES vii

1 Preface 1

1.1 Problems Motivation, Formulation and Main Results 1
1.1.1 Discretization of Continuous Objects 2
1.1.2 Signal and Image Filtering Using Stochastic Optimization 6
1.1.3 Identification of an Object from a Single Image 9
1.2 Thesis Organization 13

2 Sampling Closed Planar Curves and Surfaces 14

2.1 Introduction 14
2.2 Planar Curves 16
2.3 Sampling Planar Curves 20
2.3.1 General Finite Sampling Problem 20
2.3.2 Sampling Closed Curves: Formulation 23
2.3.3 Curves Representable in Polar Coordinates 25
2.3.4 Sampling and Reconstructing Curves: Algorithms 30
2.4 General Applicability: Deviation From Model 35
2.4.1 Sampling By Optimized Approximation 36
2.4.2 Implementation and Results 37
2.5 Further Extension: Sampling of Surfaces 38
2.6 Chapter Summary 42

3 On a New Implementation of Simulated Annealing and Its Application to Signal and Image Filtering 43

3.1 Introduction 43
3.2 Simulated Annealing 48
3.2.1 Discrete Approximation of the Filtering Scheme 49
3.3 Adaptive temperature 51
3.4 Evolution of Several Perturbations of the Original Signal 60
3.5 Chapter Summary 61

Page

4 Identification of a Discrete Planar Symmetric Shape from a Single Noisy View . . . 63

4.1 Introduction . . . 63
4.2 Elements of Projective Geometry . . . 65
4.2.1 Basic Definitions . . . 65
4.2.2 Imaging in $3 - D$ Projective Space . . . 66
4.2.3 Views of Planar Shapes and Projective Transformations of Plane . . . 72
4.3 Identification of Shape from Single View . . . 75
4.3.1 Skeleton of Symmetric Planar Shape . . . 76
4.3.2 Reconstructing View of Shape from Skeleton . . . 81
4.4 Identification of Shape from Noisy Image . . . 84
4.4.1 Skeleton of Noisy Shape . . . 86
4.4.2 Detection of Shape by Best Linear Fit . . . 94

5 Conclusions . . . 102

5.1 Summary of Contributions . . . 102
5.1.1 Sampling of Planar Curves . . . 102
5.1.2 Implementation of a Simulated-Annealing Based Optimization . . . 103
5.1.3 Identification of a Planar Discrete Symmetric Shape from a Single View . . . 104
5.2 Future Work . . . 104

LIST OF REFERENCES . . . 107

APPENDICES

Appendix A: Proofs of Theorems from Chapter 2 . . . 112
Appendix B: On Implementation of Sampling Algorithms . . . 117

LIST OF FIGURES

Figure Page

2.1 Synthetic silhouette of a B-2 plane. 17

2.2 Samples with the same values but different locations correspond to different signals. . 24

2.3 Orientation of a pair $\{\boldsymbol{\tau}(t), \gamma_s(t)\}$ remains the same for all $t \in [a, b]$. 28

2.4 A curve and its admissible set. 29

2.5 Given a polar center a curve can be reconstructed from any samples as long as their number is sufficient. 31

2.6 A curve is sampled and then perfectly reconstructed from the samples *without* using any additional information. 34

2.7 A good approximation of a map of Germany may be obtained from as few as 15 samples. 38

2.8 A contour of a human kidney can be recovered from only 7 samples. 39

2.9 A contour of a human brain can be recovered from only 19 samples. 39

3.1 Example of a clean signal and the same signal corrupted by additive Gaussian noise. . 44

3.2 Many functionals exhibit multi-modal behavior, which means they have several local extrema. 45

3.3 When the local minimum is found via a deterministic descent it is possible to identify the exact location of the extremum much simpler than in the case of a stochastic diffusion. 56

3.4 A new implementation of simulated annealing for a signal filtering functional yields faster convergence to the global minimum. 58

3.5 Lower energy achieved naturally means better filtering result for a noisy signal. 58

Figure Page

3.6 A new implementation of simulated annealing for an image filtering functional yields faster convergence to the global minimum. 59

3.7 Higher energy level yields more visible noise while lower energy level yields cleaner image. 59

3.8 Evolving a perturbed version of the original signal may yield better minimization results. 61

4.1 Examples of discrete objects. 66

4.2 Pinhole camera. 67

4.3 Arbitrary coordinate frame. 71

4.4 Rotation parameters. 71

4.5 Frontal and skewed views of the same shape. 73

4.6 Projective transformations preserve collinearity of points. 74

4.7 Construction of skeleton. 77

4.8 Example of skeleton of view. 78

4.9 Skeleton signature. 80

4.10 Reconstruction of a view from a skeleton. 83

4.11 Different choice of $\mathbf{x}_1$ affects only the width of the underlying shape. 84

4.12 Noise-free and noisy skeletons. 87

4.13 Realizations of noisy skeletons. 89

4.14 Groups of correlated points in a skeleton. 91

4.15 Skeleton points correlation coefficients are functions of only 4 shape points, not the entire shape. 93

4.16 Finding the optimal skeleton. 98

Appendix
Figure Page

4.17 Shape of a wineglass is reconstructed from a noisy observation with known camera parameters. 98

4.18 Shape of a jar is reconstructed from a noisy observation with known camera parameters. 98

4.19 Estimating the skewed view of discrete shape corresponding to the wrench. 100

4.20 Estimating the skewed view of discrete shape corresponding to the airplane. 101

Chapter 1

Preface

1.1 Problems Motivation, Formulation and Main Results

Shape analysis is a common term which refers to a whole class of theories, methods and tools whose fundamental goal is to extract useful geometric information about an object from available data. A problem of such generality is by definition so broad and ambitious in nature and so vital to many applications, that it has attracted a lot of attention of people from different areas, such as differential geometry, computer vision, functional analysis, optimization, stochastic processes just to name a few.

This thesis includes work on different aspects of shape analysis. While each one of them may be given a separate mathematical formulation, one only starts to fully appreciate their elegance and importance by considering the global context in which these problems arise. For example, to recognize an airplane from its image taken by an optical camera one has to go through a multi-stage process which might include calibrating the camera, filtering the noise in the image, detecting the presence of the object of interest, extracting its shape, finding key features and finally running those features against a database , which includes all templates of airplanes of interest. While each

of the steps mentioned above constitutes in its own right a large and complex problem, which needs to be addressed, its exact formulation or the applicability of any particular tool or method are often carried out within the global context.

In the general description of the problems we have investigated in this thesis, we briefly review some of the related works, which have appeared in the literature. Each one of these problems constitutes a milestone within the big picture goal.

1.1.1 Discretization of Continuous Objects

Shapes of observed objects naturally arise in images and in other context. Any $3-D$ object, such as an airplane, tank, human kidney or brain is an example of shape. Mathematically we can think of the boundary of an object as being its shape. A lot of objects, such as airplanes or tools such as wrenches, are almost two-dimensional as their width to length ratio is rather small, which heads one to think it is reasonable to assume them as two dimensional shapes. Imaging a $3-D$ object naturally gives rise to another $2-D$ object (its image), with its boundary becoming another example of a $2-D$ shape.

There exist two fundamentally different paradigms for studying shapes. One is to view them as being continuous entities (e.g. [25]), and another one as discrete ([23]). Each approach has its own advantages and disadvantages, and consequently pros and cons. Continuous shapes are convenient for several reasons. Not the least of them is that the world we live in, is by its very nature continuous, and so are the shapes of objects that we encounter in our everyday lives. The boundary of an airplane, for example, is a continuous surface in $3-D$ space, and that of its image is a continuous $2-D$ contour. Furthermore, working with boundaries directly without any additional

pre-processing ensures no loss of geometric information and hence the higher quality of resulting methods. It is clear that these advantages may also be sources of major challenges, as a shape space of any significance, i.e. a family of shapes of interest, is very large and of infinite dimension. This further simplifies complicates the theoretical analysis and yields severe computational problems in the course of implementation.

Confronting these challenges has had mixed results, and promising new avenues have recently emerged. Specifically, a novel idea that has been gaining acceptance is that our interest in shape should often be limited to a set of a relatively small number of key features. For example, the estimation of the position of an airplane in the air or even to classify it up to a rather narrow class may only require the position of its head, tail and ends of the wings. Information about the entire surface would turn out to be redundant and any computations required to obtain it could be considered wasted resources. Reducing a shape to a handful of its key point, commonly referred to as landmarks, with no/minimal loss of information is hence not only promising but holds a great potential to significantly simplify any problem related to shape understanding. Furthermore, as it has been shown by Kendall, Mardia and others ([23, 24, 11, 17]), such a framework enables us to build a elegant and powerful theory of discrete shapes capable of encompassing both the deterministic as well as statistical settings.

While undoubtedly important and useful, the theory of discrete shapes suffers from one major shortcoming, which is often limiting its use in any fully automated shape analysis algorithm, namely sampling. The difficulty in summarizing a continuous shape by its discrete representation of a set of landmarks indeed remains an open problem. This challenge of unsupervised sampling

persists even if a priori information about the representability of a continuous shape in terms of a finite number of discrete points is present. In the first part of the thesis we address the problem of sampling of a continuous planar closed contour. This problem has been addressed in the literature, and perhaps the most common approach has been based on splines ([36, 29]). While splines achieve relatively good performance in discretizing a curve, they do not provide a one-to-one correspondence between a a continuous curve and its discrete counterpart.

In this work, we show that a choice of a coordinate system or a parameterization for the curve plays a crucial role in the construction of a sampling algorithm. Unlike the case of a standard $1-D$ or $2-D$ signals, curves and surfaces do not yield *unique* functional representations. The latter case entirely depends on a choice of a parameterization for the given geometric object. By providing a consistent way of associating to each curve or surface a unique functional representation, which in turn requires a pre-determined coordinate system, we avoid any ambiguity. Sampling of the resulting function would then have to account for the function itself as well as for the coordinate system, with respect to which it was obtained.

In Chapter 2, we formulate the problem of sampling a continuous non-self-intersecting closed planar curve as that of finding the class of such curves that are uniquely representable by a finite number of their samples. We show that the problem seems intractable in its generality and is only manageable if a particular class of parameterizations (parameterized family of coordinate systems) is selected a priori. The coordinate system we adopt in this work is that of polar coordinates. We present conditions under which a given curve can be represented with respect to some polar coordinate system, and develop a fast algorithm to achieve our objective. We further show that

fixing a coordinate system essentially turns a planar curve into a $1-D$ function. This in turn leads to be optimistic about a possible generalization of the existing theory of sampling $1-D$ functions to the case of curves.

The major obstacle to such generalization is the fact that while in the standard $1-D$ theory, samples of a function only carry the information about the function itself, they should, for the case of curves, provide knowledge about the particular coordinate system that the curve was represented in. For the polar coordinates, such knowledge includes the location of the polar center. Its position has to be recoverable from the samples of the curve, so that the coordinate system could be reconstructed first without any additional information, and then be used together with $1-D$ sampling theory to recover the functional form of the curve. Our sampling algorithm meets those requirements and is applicable to any curve representable in polar coordinates and satisfying additional smoothness constraints described in details in the chapter.

While sampling and perfect reconstruction from those samples is possible for a fairly large class of curves, many realistic planar shapes do not precisely meet the required constraints. In particular, often times curves tend to require additional smoothness to meet the algorithm's requirements. While there is no guarantee that a proposed sampling algorithm will work ideally in that case, and therefore the problem of sampling and perfect reconstruction may not be solvable, it is desirable that it would at least be robust in the sense that one could still sample the curve and reconstruct a close approximation from the samples. Indeed, we show that the sampling algorithm developed earlier for smooth curves is a particular class of a multi-cost Pareto optimization problem ([13, 31]).

This optimization problem is applicable to and can be solved ([16]) for a much broader class of curves, yielding sufficiently accurate approximations.

Finally, by using the spherical coordinates instead of polar ones, the results of this chapter are easily generalized to the case of $3 - D$ closed surfaces.

1.1.2 Signal and Image Filtering Using Stochastic Optimization

Many tasks in image processing in general and shape analysis in particular can be formulated as optimization problems. To an image, for instance, we can associate its cost that somehow measures the amount of noise in it. The stronger is the noise, the larger that cost would be. Filtering can then be defined as modifying the original image so as to reduce its cost. Image segmentation can also be expressed as a cost minimization problem, where now for a given image, possibly consisting of several pieces, to any curve we associate the cost which is inversely proportional to a similarity between image regions bounded by that curve, however that similarity may be defined. Assuming that homogeneous regions correspond to different objects on the image, it is clear that the closer the curve would resemble their boundaries, the smaller the cost would be.

Mathematically, a cost is defined as a functional $\mathscr{H}(f)$, where $f \in \mathscr{F}$ is a signal, image, curve etc, spanning some admissible space $\mathscr{F}$. The problem is then to find the optimal signal corresponding to the minimal cost, i.e. find $f_0 = \arg\min_{f \in \mathscr{F}} \mathscr{H}(f)$.

Digital signals, images and even more geometrically complex structures, such as curves and surfaces, are conveniently modelled as vectors from some possibly high but finite dimensional Euclidean space $\mathbb{R}^d$. A cost functional is then represented as a map $\mathscr{H} : \mathbb{R}^d \to \mathbb{R}$ from that space to the real numbers. Provided that the functional $\mathscr{H}$ has certain regularity properties, the

necessary condition for its minimizer is $\nabla \mathscr{H}(f_0) = 0$, where the symbol ∇ denotes the usual gradient. A standard approach for finding a point satisfying that condition is by using the gradient descent method, where we continuously descend in the direction of the gradient until the bottom is reached.

There are numerous problems with the gradient descent method, and the most challenging being that of a functional having more than one local minimum, i.e. several points satisfying the necessary condition, in which case the method is only capable of finding the nearest local minimum but not the global optimizer. A class of methods where the deterministic gradient-based search for an optimum is replaced by a stochastic procedure falls under the class of stochastic optimization techniques. An intuitive but rather simplistic and sometimes misleading way of thinking of stochastic optimization is in the absence of any guidance, one makes a random but hopefully educated guess, which in the long run averages to the proper direction.

A particular stochastic optimization technique that we will be interested in here is called simulated annealing. When applied to the functional of interest, the essence of this technique may be summarized as follows. We think of an evolving signal as a system, whose initial state coincides with the observed signal, and which is then being iteratively perturbed and modified in the pursuit of minimizing its cost. As was pointed out earlier, the gradient descent based method allows to decrease the cost until it reaches a local minimum, at which point the algorithm stops. To avoid such a situation, at each step of the evolution the system is perturbed by an independent random additive term, i.e. a small random change is added to each component of the discrete signal. The

energy of that change, as a function of time is sometimes referred to as the temperature schedule, a term which came from physics where the technique was first described and applied.

It has been shown that if the temperature monotonously and very slowly decays to zero, the system converges in probability to its optimal state. Despite its theoretical optimality, the practical and straightforward implementation of this technique has proven inefficient due to prohibitively slow convergence. The initially proposed temperature schedule was such that for any reasonable time period, the temperature level remains significant causing the system to keep changing its state without converging. It is important to note that a slow decay of the temperature is crucial for the convergence to take place, therefore implying that no faster decaying schedule is theoretically possible. In light of this phenomenon, research that followed up was primarily concentrated on modifying the algorithm while preserving the central idea of using stochastic perturbations to allow the system to escape local minima.

We propose to investigate the problem of coming up with a simulated annealing-based optimization technique, which would be

- suitable for a broad class of functionals;
- easy to implement;
- fast.

We argue that the speed of the cost decay may indeed be significantly improved by making the temperature schedule keep track of the time passed since the last local minimum was found and

forcing more aggressive search by increasing the temperature if the system fails to improve on the current state for a long period of time.

The key idea of the proposed algorithm is the notion of *scale of search*. Intuitively, if a functional is such that its local minima or at least principal modes are sufficiently far apart, the search should be conducted at a coarser scale, while if local minima are closely clustered together, the functional has to be explored at a much finer scale so as to minimize the chance of missing the mode containing the global minimum. While for any given functional it may be possible to find a corresponding optimal scale of search, for an arbitrary cost function the temperature schedule should automatically adapt to its behavior. We propose to start the search by exploring the functional at a finer scale first, which corresponds to a lower temperature, and then progressively increase the latter thus coarsening the evolution until a lower cost state is found.

The conducted experiments confirm that this intuitive argument is justified and performs well in practice. Furthermore, the absence of a complicated acceptance - rejection algorithm used by some other simulated annealing-based techniques is the key to the ease of implementations, which in turn enables us to apply the same algorithm to different functionals and achieve a satisfactory performance.

1.1.3 Identification of an Object from a Single Image

The problem of shape identification from an image is one of the most natural tasks in computer vision. Assume we observe an object in $3 - D$ space through a camera. Upon receiving an image from the camera, the goal is to characterize it relying on the data in the image. If we can extract some features of the object from its image, then it is possible to compare them to a database of

features of known objects. Provided that database is sufficiently complete, that would allow us to recognize the object.

We have argued earlier that a shape analysis is greatly simplified when shapes are represented as discrete entities in terms of their landmarks. In this part of the thesis, we will assume that an object is represented as a collection of discrete points $\{X_i\}$ in $3-D$ space, and its image is a $2-D$ discrete shape $\{x_i\}$. Obviously, an image $\{x_i\}$ is a function of the object and the camera. A recognition mechanism is then defined as any function of the image, $f\big(\{x_i\}\big)$, which remains invariant to the choice of camera. In addition, a good recognition mechanism should possess the following properties.

- The image of the function $f(\cdot)$ should have the smallest dimension possible.
- While it is typically not necessary for a function $f(\cdot)$ to be bijective, it is desirable that it would contain as much information about its argument as possible, so as to better discriminate between different objects.
- The recognition mechanism should be robust to noise in the data. Most recognition techniques rely on certain assumptions about the objects of interest. For example, a robot at a textile factory may know that buttons are round or at least have two round holes; when trying to identify an airplane we may assume it is symmetric etc. Imaging devices may introduce noise, which would make a circle look elongated, or a symmetric object skewed. It is important that such phenomena would receive proper attention when designing recognition mechanism.

Projective geometry provides a very elegant and convenient framework for the problem of recognition of a discrete shape ([15]). Let the set of points $X_i \in \mathbb{P}^3$ denote an object lying in the $3-D$ projective space, and $x_i \in \mathbb{P}^2$ be its corresponding image. The standard "pin-hole camera" model then assumes that the image and the object are related through the equation $x_i = \mathtt{P}X_i$, where $\mathtt{P}$ is a 3×4 matrix of a linear projective transformation. It can then be deduced that different images of the same object, x_i and x_i', corresponding to cameras $\mathtt{P}$ and $\mathtt{P}'$ respectively, are related through a linear transformation from $\mathbb{P}^2$ to itself (projective transformation), defined by some matrix $\mathtt{H}$.

The above suggests that to construct a recognition scheme, we should search for invariants of projective transformations ([9, 33, 37, 38, 39]). A variety of invariants have been proposed in an extensive research in the area. It has been proven that assumptions about the class of objects of interest are required for any recognition to proceed ([6]). The idea is to therefore impose constraints that would be sufficiently restrictive to make the recognition feasible, and sufficiently broad to encompass the class of shapes of interest in our applications.

Our interest in this work is in planar symmetric shapes. The primary potential application here is the recognition of airplanes, even though many other objects, such as tools fall into the same category. The goal is to then provide a classification mechanism for such an object based on its single image. The corner stone of our proposed recognition algorithm that we propose here is the notion of the skeleton of a view ([43]). A skeleton is a collection of points all lying on the same line, which is the image of the symmetry axis of the original object. As was discussed earlier, different views of the same objects are related to each other through projective transformations.

The same holds true for the skeletons of those views, i.e. the skeleton of the view $\{x_i\}$ is the projective image of the skeleton of another view $\{x'_i\}$ through some $1-D$ projective transformation. Furthermore, we formally show it that skeletons contain virtually as much geometric information about the original object, as the views. The problem of object recognition can therefore be reduced to classification of possible view skeletons.

Recall that a skeleton consists of several collinear points. The famous projective invariant of collinear points is the so-called cross-ratio. A cross-ratio is a function of four collinear points defined in terms of ratios of distances between them, which remains invariant to any projective transformation on those points. If a skeleton contains more than 4 different points, we can construct a vector of cross-ratios by taking different 4-tuples of points. That vector would obviously be projectively invariant. To summarize, given a view we can construct its skeleton which in turn gives rise to a vector of cross-ratios. The latters only depend on the initial object, and not on any particular image. Furthermore, this invariant captures almost all geometric information about the object, which makes it a good candidate for a recognition technique.

The second part of this work is devoted to the problem of robustness of the proposed recognition technique against noise in the data. As was discussed in the sampling part of the introduction, a discrete shape is always produced from its continuous original either by applying a discretization algorithm (See Chapter 2), or even by selecting the landmarks manually. In either case, the sampling process is inherently imperfect and yields discretization errors. A question that naturally arises is whether the previously proposed recognition algorithm is robust against this type of problems. This leads to the problem of statistical analysis of the proposed invariants.

While completely computing the multi-dimensional distribution of the signature vector is still an open problem ([34]), we show that under the assumption that the noise in data is independent and gaussian with zero-mean and identical variance, optimal identification is still possible, i.e. while it is certainly not feasible to do a perfect recognition in the noisy environment, for a given view, we can derive an optimal estimate of its skeleton, which would in turn allow us to compute the estimates of the cross-ratios. This statement is based on a proposition, which states that an independent gaussian noise in the data points leads to skeleton points becoming gaussian with the mean coinciding with the true skeleton and some covariance matrix Σ. The estimate of the true skeleton is then found by applying the generalized linear fit to the noisy skeleton points.

1.2 Thesis Organization

This thesis is organized as follows. Chapter 2 contains the discussion of sampling of planar shapes and $3-D$ curves along with the newly proposed sampling algorithm. In Chapter 3 we discuss the new implementation of a simulated annealing based stochastic optimization and its applications to signal and image filtering. Chapter 4 contains a development of the new technique of identification of a discrete planar symmetric shape from a single image. Finally Chapter 5 concludes this thesis by summarizing the work, providing the final remarks and discussing potential future research.

Chapter 2

Sampling Closed Planar Curves and Surfaces

2.1 Introduction

Sampling theory has a long history of development ([4, 55, 62]). The classical sampling theorem, which provides the basis for reconstructing bandlimited $1-D$ signals from discrete samples, was first proved by Cauchy, rediscovered by Whittaker ([48]) and Kotelnikov ([28]), and finally applied to problems of communications by Shannon ([49]), which ultimately resulted in the "digital revolution". The theorem is remarkable in that it allows converting an analog signal satisfying certain class constraints into a discrete sequence of numbers without any loss of information. Its mathematical statement is as follows.

Theorem 2.1 If a function $f : \mathbb{R} \to \mathbb{R}$ is bandlimited to $\omega_{\max} < \frac{\pi}{T}$, then it can be exactly reconstructed from its values at sampling points $t_n = nT$:

$$f(x) = \sum_{n=-\infty}^{+\infty} f(nT) sinc\left(\frac{x}{T} - n\right), \; x \in \mathbb{R}. \tag{2.1}$$

The formula (2.1) is nothing but a representation of a function with respect to a particular orthogonal basis. It was noted later that spaces other than that defined by basis of $sinc$-functions may be considered as models for "ideal" signals ([32]), i.e., those allowing perfect reconstruction from their discrete samples. Bases other than orthogonal (e.g. Riesz bases) can be efficiently used to

represent signals relative to them. Furthermore, it was shown that one might take advantage of stable representations of a signal with its coefficients with respect to systems with redundancy as opposed to orthogonal bases. That theory gave rise to the notion of frame and delivered many powerful sampling techniques ([2, 1]).

In this chapter we consider the problem of sampling a closed planar curve or a $3-D$ surface. Just like in the case of $1-D$ signals, the ability to convert a continuous contour to a collection of discrete points is extremely desirable and yields a number of useful applications. If the informative content of an image is limited to a single shape, then sampling the boundary of the object as opposed to the entire image may result in significant increase of compression efficiency.

On the other hand, when a shape is analyzed in the context of computer vision, then the importance of being able to find a discrete representation cannot be overemphasized. While there exists a number of methods dealing with continuous curves, the latter must often come from a very narrow class, e.g., quadrics ([56]). At the same time representing shapes as arrays of points is fundamentally simpler, and therefore a richer theory exists for such scenarios ([11, 50, 23]).

The main obstacle we have to tackle when sampling a curve is lack of its *a priori* given functional form ([41]). As will be discussed in detail later, what we normally observe is an image of a curve as opposed to the curve itself. That creates a problem of finding a parametrization that could be run through a sampling algorithm. Since a type of parametrization needs to be defined *a priori* ([42]), certain constraints have to be imposed on curves to guarantee its existence. We will show that once a parametrization of the curve is obtained, a sampling theorem based on sampling theory of $1-D$ signals can be applied ([44]).

This chapter is organized as follows. Our focus is mainly concentrated on $2-D$ curves. Section 2.2 provides a general background about planar curves. Its purpose is to fix the notation as well as provide the insight as to the fundamental differences between $1-D$ signals and planar curves from the point of view of sampling theory. In Section 2.3 we formulate the problem we address in this chapter, and describe in detail the solution that we propose, which includes class-constraints on curves, sampling and reconstruction algorithms and present examples of applying those algorithms to idea curves. Technical proofs are moved to Appendix A for the sake of clarity.

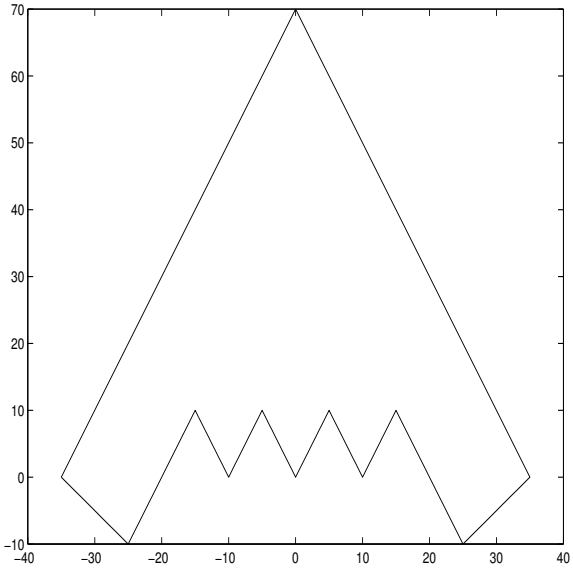
In practice, most if not all curves do not exactly satisfy any *a priori* given conditions. It follows then that no technique can provide a perfect sampling and reconstruction of realistic curves. In light of this, in Section 2.4 we propose an optimization technique which yields an optimal approximation to an original curve. We show that it is consistent with the sampling theorem for class-constrained curves in the sense that the latter becomes a particular case of the former.

Results obtained for $2-D$ curves can easily be generalized for $3-D$ surfaces. A brief summary of results for $3-D$ is presented in Section 2.5. Section 2.6 concludes the chapter by providing the final remarks.

2.2 Planar Curves

Curves arise in practice in many ways (see Figure 2.1). They normally represent a silhouette or a contour of some object in an image. From mathematical point of view, what we observe on a piece of paper or a computer screen is in fact an image of a curve. In general there exists no way to uniquely extract the functional form of a curve from its image, provided no *a priori* assumptions

Figure 2.1 Synthetic silhouette of a B-2 plane.



about the nature of the curve. Note that if we are given a graph of a $1-D$ signal, then a very specific coordinate system, namely Cartesian, is attached to it even if not explicitly mentioned. We will give the following definitions ([3, 19]).

Definition 2.2 A path (parameterized curve) is a continuous function $\gamma : [a, b] \subset \mathbb{R} \to \mathbb{R}^2$, i.e.,

$$\gamma(t) = \big(x(t), y(t)\big) \in \mathbb{R}^2,\ t \in [a, b],\ \gamma \in C\big([a, b]\big). \tag{2.2}$$

Here t, which varies in $[a, b]$, is the parameter of a curve.

Definition 2.3 The subset $\mathscr{C} \subset \mathbb{R}^2$ defined by

$$\mathscr{C} = Im(\gamma) \equiv \gamma\big([a, b]\big) \equiv \big\{\gamma(t) \mid t \in [a, b]\big\} \tag{2.3}$$

is called the image of a path $\gamma(\cdot)$.

The set $\mathscr{C}$ is what we really call a curve in everyday life. It is therefore intuitive that the set $\mathscr{C}$ should not depend on a choice of parameterization contrary to what is suggested by Definition 2.3.

The following construction assures that $\mathscr{C}$ is in fact independent from a parameterization so long as the latter belongs to an admissible class of paths.

Definition 2.4 Two paths $\gamma_1(t) : [a, b] \to \mathbb{R}^2$ and $\gamma_2(\tau) : [\alpha, \beta] \to \mathbb{R}^2$ are said to be equivalent (write $\gamma_1 \sim \gamma_2$), if there exists a change of variable $t = t(\tau)$, satisfying the conditions

1. $t : [\alpha, \beta] \to [a, b]$ is a bijection between the two intervals;
2. t and t^{-1} are continuously differentiable, i.e.,

$$t \in C^1\big([\alpha, \beta]\big),\ t^{-1} \in C^1\big([a, b]\big);$$

3. The two parameterizations have the same direction, i.e.,

$$\frac{d}{d\tau}t(\tau) > 0,\ \forall \tau \in [\alpha, \beta],$$

such that

$$\gamma_1\big(t(\tau)\big) = \gamma_2(\tau),\ \forall \tau \in [\alpha, \beta]. \tag{2.4}$$

It easily follows from Definition 2.4 that equivalent paths have the same image.

Theorem 2.5

$$\Big(\gamma_1 \sim \gamma_2\Big) \Rightarrow \Big(Im(\gamma_1) = Im(\gamma_2)\Big). \tag{2.5}$$

Furthermore, the relationship $\sim$ from Definition 2.4 is an equivalence relationship, which yields that all continuous paths are divided into disjoint equivalence classes. Those classes are called planar curves.

Definition 2.6 Let $\gamma_0 : [a, b] \to \mathbb{R}^2$ be a path. Then the equivalence class

$$C = \{\gamma \mid \gamma \sim \gamma_0\} \tag{2.6}$$

is called a curve with the representative γ_0.

Theorem 2.5 asserts that the image of a curve is uniquely defined.

Definition 2.7 Let C be a curve, and $\gamma_0 \in C$. Then by definition

$$Im(C) \equiv Im(\gamma_0). \tag{2.7}$$

Definition 2.8 A curve C is called closed if it contains at least one parametrization $\gamma_0 : [a, b] \to \mathbb{R}^2$, which is closed, i.e., $\gamma_0(a) = \gamma_0(b)$.

One easily establishes that *all* elements of a closed curve are closed parameterizations.

Definition 2.9 A closed curve C is called Jordan if it does not intersect itself except at the end points, i.e., if $\gamma_0 : [a, b] \to \mathbb{R}^2$ is its parametrization, and $t_1 < t_2$ then

$$\Big(\gamma_0(t_1) = \gamma_0(t_2)\Big) \Rightarrow \Big(t_1 = a,\ t_2 = b\Big). \tag{2.8}$$

In the remainder of the chapter, we will focus exclusively on Jordan curves.

A curve is defined as an object independent of a particular parametrization. Other geometric objects attributed to curves may be defined in a similar manner. A concept we find particularly useful is that of a unit tangent vector to a curve.

Definition 2.10 Let $\gamma : [a,b] \to \mathbb{R}^2,\ \gamma \in C^1\big([a,b]\big),\ \gamma(t) = \big(x(t), y(t)\big),\ t \in [a,b]$ be a path and $t_0 \in [a,b]$. Then the unit tangent vector to the path at the point $\gamma(t_0)$ is a vector $\boldsymbol{\tau}(t_0)$ defined as

$$\boldsymbol{\tau}(t_0) = \frac{\gamma'(t_0)}{\|\gamma'(t_0)\|} \in \mathbb{R}^2, \tag{2.9}$$

where

$$\gamma'(t_0) = \big(x'(t_0), y'(t_0)\big),\ \|\gamma'(t_0)\| = \sqrt{x'(t_0)^2 + y'(t_0)^2}. \tag{2.10}$$

As was noted earlier, it can be easily shown that the notion of a unit tangent vector to a curve is geometric, i.e., it does not depend on a particular parameterization of the curve. More precisely, the following theorem holds.

Theorem 2.11 Let $\gamma_1(t) : [a,b] \to \mathbb{R}^2$ and $\gamma_2(\tau) : [\alpha,\beta] \to \mathbb{R}^2$ be two equivalent paths, i.e. $\gamma_1 \sim \gamma_2$, such that $t = t(\tau)$ is a change of variable satisfying all the conditions listed in Definition 2.4. Let $\boldsymbol{\tau}_1(t), \boldsymbol{\tau}_2(\tau)$ be tangent vectors to paths γ_1 and γ_2 at points $\gamma_1\big(t(\tau)\big)$ and $\gamma_2(\tau)$ respectively. (Note that $\gamma_1\big(t(\tau)\big) = \gamma_2(\tau)$.) Then

$$\boldsymbol{\tau}_1\big(t(\tau)\big) = \boldsymbol{\tau}_2(\tau). \tag{2.11}$$

The proof of this theorem can be found in [19]. We will make use of this theorem in later sections of the chapter.

2.3 Sampling Planar Curves

2.3.1 General Finite Sampling Problem

Here we give a mathematical description of a finite sampling problem which is, generally speaking, suitable for any signals including $1-D$ function, images or curves on the plane. Our

description and solution of the problem will not only be rigorous and unambiguous, but it will also help us contrast and compare the relevance of traditional sampling theory of $1-D$ and $2-D$ signals to the problem we have at hand.

To proceed with the formulation of the finite sampling problem for an arbitrary signal, we state the following definition.

Definition 2.12 Let $f : X \to Y$ be a signal (function). Completely specifying the sampling problem of this signal entails imposing conditions on f, such that it is possible to find a triple $(N, \mathscr{X}, \mathscr{I})$, where $N \in \mathbb{N}$ is a number of samples, $\mathscr{X} \equiv \{x_1, \ldots, x_N\} \in X^N$ is a collection of N distinct samples, and $\mathscr{I}$ is an interpolation function (method of reconstruction): $\mathscr{I} : X^N \times Y^N \times X \to Y$, such that

$$\mathscr{I}\big(x_1, f(x_1), \ldots, x_N, f(x_N); x\big) = f(x),\ \forall x \in X. \tag{2.12}$$

The reader may find the above definition too abstract. However, it is extremely important to formally specify the problem we are going to be addressing. This definition also reveals the minimal information needed in order to formulate a sampling theorem. The following well-known theorem is a classical example of a solution to a $1-D$ finite sampling problem.

Theorem 2.13 Let $f : \mathbb{R} \to \mathbb{R}$ be a $1-D$ signal. Suppose also that it is periodic with a fundamental period 2π and bandlimited, i.e., there exists $N \in \mathbb{N}$ (for the sake of simplicity we will assume N even) and $c_{-\frac{N}{2}}, \ldots, c_{\frac{N}{2}} \in \mathbf{C}$, such that

$$f(t) = \sum_{n=-N/2}^{N/2} c_n e^{itn},\ \forall t \in \mathbb{R}. \tag{2.13}$$

Then if $t_1, \ldots, t_{N+1} \in \mathbb{R}$ is a collection of $N+1$ arbitrary but distinct points on the real line and $f(t_1), \ldots, f(t_{N+1})$ the corresponding sample values of $f(t)$, we can reconstruct the signal $f(t)$, and the interpolation function, $\mathscr{I}$, is defined by the right side of Equation 2.13, where the coefficients $\{c_n\}_{n=-\frac{N}{2},\ldots,\frac{N}{2}}$ are the unique solution of the following system of equations.

$$\begin{pmatrix} f(t_1) \\ \vdots \\ f(t_{N+1}) \end{pmatrix} = \begin{pmatrix} e^{i\left(-\frac{N}{2}\right)t_1} & \cdots & e^{i\frac{N}{2}t_1} \\ \vdots & \ddots & \vdots \\ e^{i\left(-\frac{N}{2}\right)t_{N+1}} & \cdots & e^{i\frac{N}{2}t_{N+1}} \end{pmatrix} \cdot \begin{pmatrix} c_{-\frac{N}{2}} \\ \vdots \\ c_{\frac{N}{2}} \end{pmatrix}. \tag{2.14}$$

The reader may refer to [32] among others for the proof of this theorem. It easily follows from the fact that any $N+1$ samples yield a linear system of equations with respect to $N+1$ signal Fourier coefficients, which is full-rank and thus has a unique solution. The Fourier coefficients, in turn, completely define a signal $f(t)$ (see, for example, [45]).

Definition 2.14 Let $f : X \to Y$ be a given signal. We will say that a point

$$y \in Im(f) \tag{2.15}$$

is an incomplete sample of f on X.

Definition 2.15 Let $f : X \to Y$ be a given signal. We will call a pair

$$\big(x, f(x)\big) \in X \times Y, \tag{2.16}$$

a complete sample of f on X.

Note, that when given an incomplete sample of a signal, we have its value at an unknown location. In a statement of a finite sampling problem a reconstruction of a signal is usually from

its complete samples. The following simple example demonstrates that the sample locations are critical.

Consider again a $1-D$ bandlimited signal $f(t)$, satisfying all of the conditions from Theorem 2.13 with $N = 3$. Then if we are given any 3 complete samples, Theorem 2.13 guarantees that we will be able to uniquely reconstruct $f(t)$. On the other hand, if we were given incomplete samples, then by varying the locations of those samples we would be able to construct an infinite number of signals, satisfying all the required conditions, e.g., if

$$f(t_1) = 1,\ f(t_2) = 0,\ f(t_3) = -1,$$

then the locations

$$t_1 = 0,\ t_2 = \frac{\pi}{2},\ t_3 = \pi$$

yield $f(t) = \cos t$, but for

$$t_1 = \frac{\pi}{2},\ t_2 = \pi,\ t_3 = \frac{3\pi}{2},$$

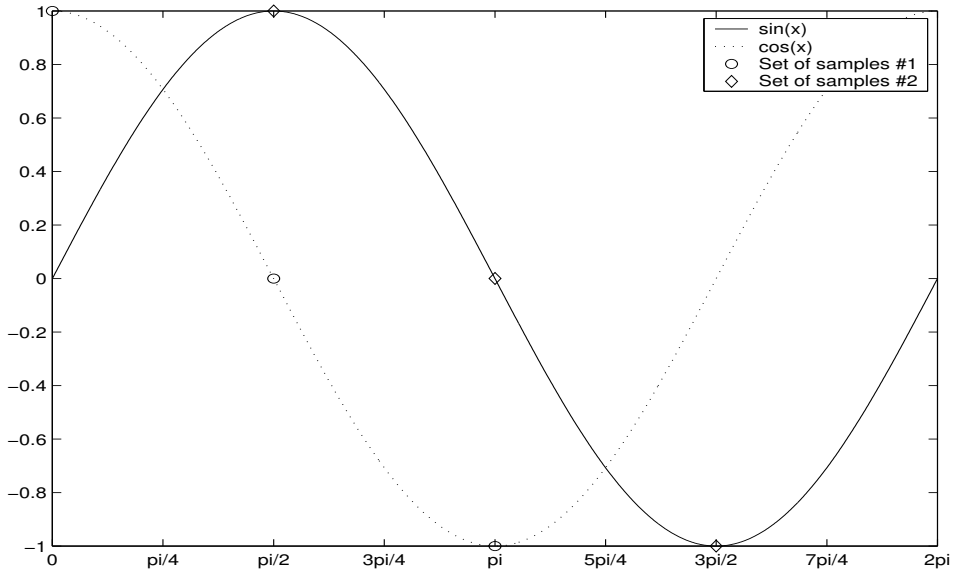
the formula for the signal would become $f(t) = \sin t$ (See Figure 2.2).

This very simple example underlines the difficulty we encounter when formulating a sampling problem for planar curves as we discuss next.

2.3.2 Sampling Closed Curves: Formulation

Recall, that a closed planar curve C is a class of equivalent (closed) parameterizations of the same image $\mathscr{C}$. We formulate the sampling problem for the curve C as follows. Find a parametrization $\gamma \in C$, whose finite sampling problem is well defined - seek a function $\gamma : [a, b] \to \mathscr{C} \subset \mathbb{R}^2$,

Figure 2.2 Samples with the same values but different locations correspond to different signals.



which satisfies some conditions so that a finite number of complete samples results yielding a reconstruction of $\gamma(\cdot)$.

To proceed, it is important to make the following observation. A complete sample for a curve has the form $\big(t;\gamma(t)\big) \equiv \big(t;x(t),y(t)\big)$. In practice, however, all we have is a set of point $\big\{(x_i,y_i)\big\}_{i=1}^{N}$, which do not constitute *complete* samples. In fact, a closer consideration reveals that these samples are not even *incomplete* as for arbitrary points on the plane $\big\{(x_i,y_i)\big\}_{i=1}^{N}$ there is no information about the parametrization at hand or its domain. The following two steps are in order to mitigate such a limitation.

- Make additional assumptions about the nature of a curve, namely about the *type* of coordinate system, this curve is represented in.

- Make the information of the coordinate system selection part of the sample specification so as to avoid any potential ambiguity presenting a closed curve and subsequentles.

2.3.3 Curves Representable in Polar Coordinates

As was

argued earlier, a choice of a coordinate system is important for extracting some meaningful information from curve samples. This may be demonstrated by a $1-D$ function, whose samples are as such in a cartesian coordinate system, and otherwise just a set of points in the plane. We glean the information projecting points onto the two axes. While this example may look trivial, the resulting procedure is instructive and indicative of the challenges we have to tackle.

To address the representation issues of closed (Jordan) curves, one natural reference system which surfaces is that of polar coordinates. Assuming that such a curve representation exists, its seemingly restrictive scope quickly pales next to its merit in suitability as well as adaptability to a wide array of settings. In the case of interest herein, we adopt a polar parametrization of a given curve C and next discuss the technical conditions which underly its existence.

Definition 2.16 Suppose we have a curve C. If for a given point $\gamma_0 \equiv (x_0, y_0)$ there exists a function $r(\theta) : [0, 2\pi] \to \mathbb{R}^2$, such that the curve C can be parameterized according to

$$\gamma(\theta) = \gamma_0 + \big(r(\theta)\cos(\theta), r(\theta)\sin(\theta)\big),\ \theta \in [0, 2\pi],\ \gamma(0) = \gamma(2\pi), \tag{2.17}$$

then we call γ_0 an admissible (polar) center of C.

While such a definition does not uniquely specify an admissible center of a curve, it allows us to address the existence issue, which in turn affords our constructing of curves whose class membership is easily carried out.

Given a curve C, and using the above definition, we may proceed to determine the admissibility of any point (x_0, y_0) by ensuring that every half-line originating at this point intersects C at one and only one point. To alleviate the computational load such a procedure may entail, we state the following theorem which forms the basis for a more efficient search of admissible points.

Theorem 2.17 Let C be a smooth curve, $(x_0, y_0) \in \mathbb{R}^2$, and ℓ - half-line starting from (x_0, y_0) and intersecting the image of C, set $\mathscr{C}$, in at least two points. Then there exists another half-line ℓ^t that also starts from (x_0, y_0) and is tangent to C.

The proof of this theorem can be found in the Appendix A.

It is important to note that the existence of a tangent half-line to a curve originating at a point (x_0, y_0) does not guarantee the non-admissibility of the latter as a center for the curve. It does, however, as further discussed below, rule all of those points out of the set from further consideration and hereby identify a so-called regular admissible set.

Definition 2.18 Let C be a curve and $(x_0, y_0) \in \mathbb{R}^2$ be a point. Then this point is called a regular admissible center if there exists no half-line ℓ going from (x_0, y_0) and tangent to C. A collection of all regular admissible centers will be called the admissible set of the curve C and denoted $\mathscr{A}_C$.

To simultaneously qualify the membership of any selected point (x_0, y_0) to $\mathscr{A}_C$ and avoid a computational explosion, the following describes an easily implementable procedure.

Theorem 2.19 Let C be a curve, and $\gamma : [a, b] \to \mathscr{C}$ be its parametrization. For any $t \in [a, b]$ define $\boldsymbol{\tau}(t)$ - a unit tangent vector to γ at a point $\gamma(t)$, and $\gamma_s(t)$ - a vector connecting the points $\gamma_0 \equiv (x_0, y_0)$ and $\gamma(t) \equiv \big(x(t), y(t)\big)$, i.e.,

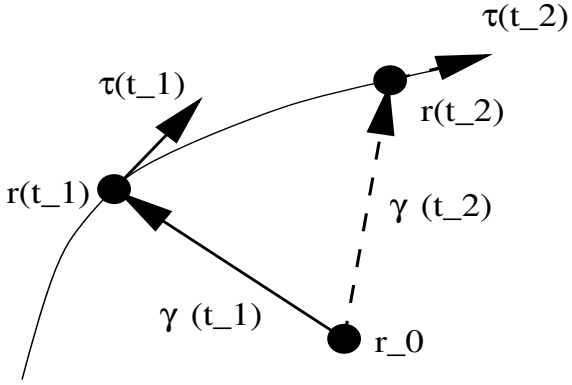
$$\gamma_s(t) = \big(x(t) - x_0, y(t) - y_0\big),\ t \in [a, b], \tag{2.18}$$

where $\gamma(t) \equiv \big(x(t), y(t)\big)$ (see Figure 2.3). Then $\gamma_0 \in \mathscr{A}_C$ if and only if the orientation of the pair of vectors $\big\{\boldsymbol{\tau}(t), \gamma_s(t)\big\}$ is constant (non-zero) for all $t \in [a, b]$. The latter condition means

$$sgn\left(\begin{vmatrix} \gamma_1(t) & \gamma_2(t) \\ \tau_1(t) & \tau_2(t) \end{vmatrix}\right) = const \neq 0,\ \forall t \in [a, b], \tag{2.19}$$

where $\gamma(t) \equiv \big(\gamma_1(t), \gamma_2(t)\big)^T$, $\tau(t) \equiv \big(\tau_1(t), \tau_2(t)\big)^T$.

Figure 2.3 Orientation of a pair $\{\boldsymbol{\tau}(t), \gamma_s(t)\}$ remains the same for all $t \in [a, b]$.



See Appendix A for the proof of this theorem. Appendix B contains more details regarding the numerical implementation of the test for admissibility based on Theorem 2.19.

In practice, a continuous curve is, of course, a polygon with a large number of vertices and ever smaller edges, and significantly simplifies the exhaustive search as called by Equation (2.19). In addition, as we show next, the structure of $\mathscr{A}_C$ may be used to advantage in avoiding to have to check the totality of points individually.

Theorem 2.20 Let C be a curve and $\mathscr{A}_C$ be its admissible set. Then $\mathscr{A}_C$ is convex, i.e.,

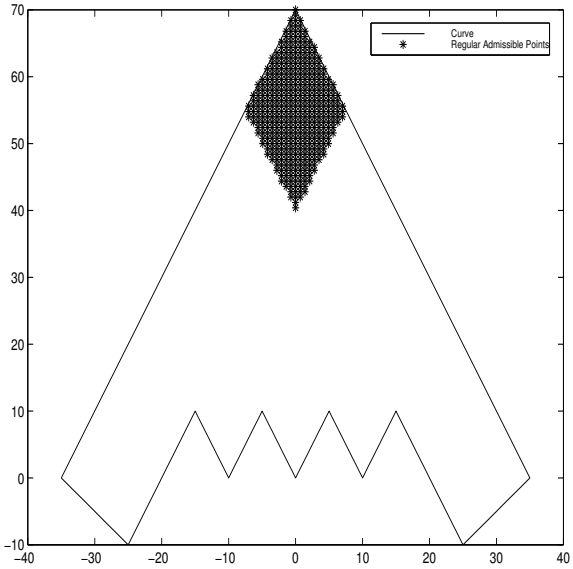
$$\Big(\gamma_1, \gamma_2 \in \mathscr{A}_C\Big) \Rightarrow \Big([\gamma_1, \gamma_2] \subset \mathscr{A}_C\Big), \tag{2.20}$$

where

$$[\gamma_1, \gamma_2] = \big\{\alpha\gamma_1 + (1-\alpha)\gamma_2 \mid \alpha \in [0, 1]\big\}. \tag{2.21}$$

The proof is presented in the Appendix A. A fast and simple approximation of $\mathscr{A}_C$ may, for instance, be carried out by a Monte Carlo method and illustrated in Figure 2.4.

Figure 2.4 A curve and its admissible set.



2.3.4 Sampling and Reconstructing Curves: Algorithms

Provided that a non-empty admissible set $\mathscr{A}_C$ exists, as spelled out above, our goal in this section is to describe practical algorithms enabling us to sample closed curve and reconstructing them. Specifically, when given a curve C, we have to, subject to any other additional constraints, identify and acquire its appropriate samples, which will also allow us to reconstruct it.

Suppose C is a curve and $\gamma_0 \in \mathscr{A}_C$. Then by definition there exists a polar parameterization $\big\{r(\theta),\ \theta \in [0, 2\pi]\big\}$ of the curve C. Since, $r(\theta)$ is a $1-D$ signal, we may therefore apply Theorem 2.13 to obtain the following result.

Theorem 2.21 Suppose that we have a curve C and $\gamma_0 \in \mathscr{A}_C$. Let $r(\theta)$ be a polar representation of C centered at γ_0. Then if $r(\theta)$ is bandlimited, and there exists $N \in \mathbb{N}$, such that

$$r(\theta) = \sum_{n=-N/2}^{N/2} c_n e^{i\theta n}, \tag{2.22}$$

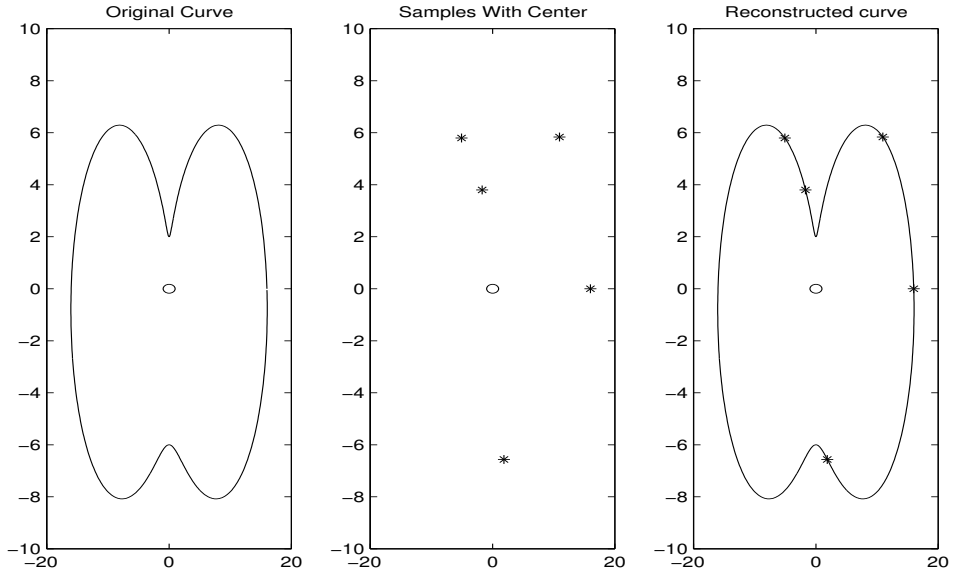
then $r(\theta)$ can be reconstructed from any $N+1$ samples (See Figure 2.5).

The reconstruction formula follows from Theorem 2.13. A few remarks regarding the last result are in order.

- Note that this result is not a finite sampling theorem as it heavily relies on a given point γ_0. Assuming such knowledge is quite unnatural, since in practice we are given only samples of a curve, and thus it is impossible to use this theorem.

- Providing the polar center of a curve explicitly is often undesirable because of noise that is present in virtually any real-world system. Suppose, for example, that the curve is to be

Figure 2.5 Given a polar center a curve can be reconstructed from any samples as long as their number is sufficient.



transmitted over a digital channel. It then would not greatly reduce the efficiency of transmission if we simply added γ_0 to transmitted samples. The downside of this approach is in that γ_0 becomes the bottle-neck of the system, since any error added to it, would significantly affect the reconstructed curve.

Our goal is now to come up with a sampling technique that would allow us to *extract* the polar coordinate system from samples in a reliable manner. Recall, that Theorem 2.13 states that a bandlimited signal can be reconstructed from any (possibly non-uniform) samples so long as their number matches the bandwidth of the signal. We show that it is possible to use this freedom to choose samples in such a way that they would contain the information about the coordinate system that was used to produce them.

We formulate the following proposition which will immediately provide us with a desirable sampling technique.

Proposition 2.1 Let C be a curve and $\gamma_0 \in \mathscr{A}_C$. Without loss of generality we will assume that $\gamma_0 = (0,0)$. Otherwise we can always shift the curve. Suppose that $\big\{r(\theta),\ \theta \in [0,2\pi]\big\}$ is a corresponding polar parametrization of C. Then if $r(\theta)$ is bandlimited, so that (2.22) holds, then we can always choose sampling points $\{\theta_i\}_{i=1}^{N+1}$ in such a way that

$$\frac{1}{N+1}\sum_{i=1}^{N+1}\big(r(\theta_i)\cos(\theta_i), r(\theta_i)\sin(\theta_i)\big) = (0,0). \tag{2.23}$$

The significance of the last statement lies in the implication of an existence of a set of sample points, whose orthocenter is at $(0,0)$, representing a parameterized curve so long as the latter is

bandlimited. This is tautologous, as is shown below, to saying that a unique reconstruction is possible with no additional assumptions. Thus we have the following sampling theorem.

Proposition 2.2 Let C be a curve. Assume $\mathscr{A}_C \neq \emptyset$, and $\gamma_0 \equiv (x_0, y_0) \in \mathscr{A}_C$. Suppose that $r(\theta)$ is a polar parametrization of C centered at γ_0. If $r(\theta)$ is bandlimited, so that (2.22) holds, then there exist distinct points $\theta_1, \ldots, \theta_{N+1} \in [0, 2\pi]$, such that the curve C is uniquely defined by the points

$$\big(r(\theta_i)\cos(\theta_i), r(\theta_i)\sin(\theta_i)\big) \in \mathbb{R}^2,\ i = 1, \ldots, N+1. \tag{2.24}$$

The reconstruction procedure easily follows from Equation (2.23) and Theorem 2.13.

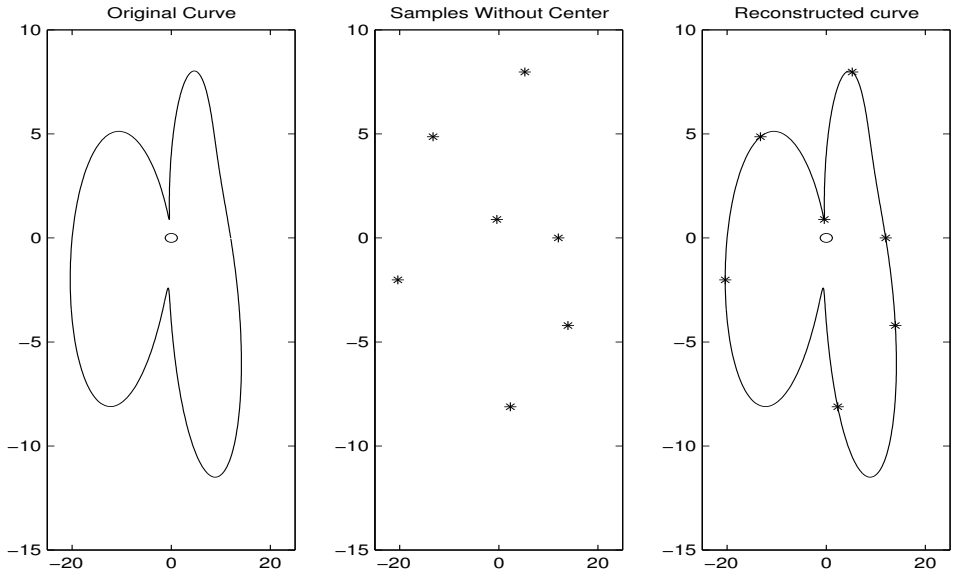
The above sampling procedure is sufficiently general to be applicable to a variety of curves. The sampling and unique reconstruction of a curve C is subject to the following conditions, which are, recall

- The admissible set from Definition 2.18 must be non-empty, i.e., $\mathscr{A}_C \neq \emptyset$.
- There must exist a point $\gamma_0 \in \mathscr{A}_C$, such that the corresponding polar parameterization of C is bandlimited.

The conditions we impose on a curve seem restrictive until we note that, unlike the $1 - D$ case, the proposed technique affords a reconstruction in spite of lack of complete samples and a fixed and unambiguous coordinate system.

The Figure 2.6 shows the result of applying the described technique to a curve, where we may also note that such an encoding of γ_0 is stable against random additive noise because of the smoothing that takes effect in the course of averaging in Equation (2.23).

Figure 2.6 A curve is sampled and then perfectly reconstructed from the samples *without* using any additional information.



2.4 General Applicability: Deviation From Model

As discussed in the beginning of this chapter, perfect reconstruction is impossible for arbitrary signals in general, and curves in particular. Any reconstruction algorithm requires additional information about the nature of a signal. In many cases a signal is constrained to lie in a certain functional space for the sampling theorem to be valid. The classical sampling theorem, for example, requires that a signal belong to the space of bandlimited functions.

Although class constraining, the above conditions, which a curve must satisfy to be a good candidate for sampling and exact reconstruction, are justifiable in light of the rather limited number of degrees of freedom imposed on an otherwise infinite dimensional space where our functions of interest live.

In the present context and for all practical purposes, where sampling of shapes is our main interest, a good approximation of curves will be perhaps no less important than an exact reconstruction. This in turn yields a number of interesting questions, such as:

- How close is the approximation to the original curve?
- Can we minimize the error of approximation?

An important point to note is that in the case of a non-bandlimited curve the bandlimited reconstruction may suffer from significant errors, if the proposed above procedure is directly applied. Recall, that if a curve is to be sampled and uniquely reconstructed, its sampling points are chosen so that the corresponding orthocenter coincides with the polar center. If we denote the distance between the orthocenter of samples and the polar center by $d(\theta_1, \ldots, \theta_{N+1})$, then our condition

may be written as

$$d(\theta_1, \ldots, \theta_{N+1}) = 0. \tag{2.25}$$

We claimed earlier that it is always possible to find $\{\theta_i\}_{i=1}^{N+1}$ for a bandlimited curve, such that (2.25) holds and where N matches the bandwidth. It is clearly seen that finite sampling and exact reconstruction are limited to bandlimited signals/curves. In the event that the curve is non-bandlimited, we may only seek to determine the samples, which will minimize $d(\theta_1, \ldots, \theta_{N+1})$ while optimizing the distance between the curve and its reconstructed approximation. Note that the previously described sampling procedure may be viewed as a particular case of this technique, which specializes to that above for the corresponding class of curves and for which the approximation error is zero, and the two criteria optimization described next reduces to one.

2.4.1 Sampling By Optimized Approximation

Let us assume that we have a curve C, such that $\mathscr{A}_C \neq \emptyset$, and $\gamma_0 \in \mathscr{A}_C$. Let $\left\{r(\theta), \theta \in [0, 2\pi]\right\}$ be a polar parameterization, which is not necessarily bandlimited. Let $N \in \mathbb{N}$ be a fixed integer number (for the sake of simplicity even). Towards formulating our generalized sampling define for each $(N+1)$-tuple $\theta_1, \ldots, \theta_{N+1} \in [0, 2\pi]$

$$d^2(\theta_1, \ldots, \theta_{N+1}) = \left(\sum_{i=1}^{N+1} r(\theta_i)\cos(\theta_i)\right)^2 + \left(\sum_{i=1}^{N+1} r(\theta_i)\sin(\theta_i)\right)^2, \tag{2.26}$$

and

$$l^2(\theta_1, \ldots, \theta_{N+1}) = \int_0^{2\pi} \left|r(\vartheta) - \hat{r}_{\theta_1,\ldots,\theta_{N+1}}(\vartheta)\right|^2 d\vartheta, \tag{2.27}$$

where $\hat{r}_{\theta_1,\ldots,\theta_{N+1}}(\vartheta)$ is the reconstruction achieved by the samples defined at $\theta_1, \ldots, \theta_{N+1}$. In an ideal situation we would like both d^2 and l^2 to be zero. It is clear, however, that the two functions

may achieve their minima at different points in an $(N+1)$-dimensional space. We search for a so-called Pareto solution of the two-criteria optimization problem ([13]), which may be viewed as a point, from which any deviation may not decrease in one component without increasing in the other. For more details refer to [31]. Define $F : [0, 2\pi]^{N+1} \to \mathbb{R}^2$ as follows

$$F(\theta_1, \ldots, \theta_{N+1}) = \left(d^2(\theta_1, \ldots, \theta_{N+1}), l^2(\theta_1, \ldots, \theta_{N+1})\right). \tag{2.28}$$

Definition 2.22 A point $(\theta_1^0, \ldots, \theta_{N+1}^0) \in \mathbb{R}^{N+1}$ is called the Pareto optimal point of F, if $\forall(\theta_1, \ldots, \theta_{N+1}) \in \mathbb{R}^{N+1}$ we have

$$\begin{cases} d^2(\theta_1, \ldots, \theta_{N+1}) & \leq & d^2\left(\theta_1^0, \ldots, \theta_{N+1}^0\right) \\ l^2(\theta_1, \ldots, \theta_{N+1}) & \leq & l^2\left(\theta_1^0, \ldots, \theta_{N+1}^0\right) \end{cases} \Rightarrow (\theta_1, \ldots, \theta_{N+1}) = (\theta_1^0, \ldots, \theta_{N+1}^0). \tag{2.29}$$

Consider the case when the function $r(\theta)$ is bandlimited, in the sense of Equation (2.22). There then exists according to Proposition 2.1 the corresponding number $N+1$ and $\theta_1^0, \ldots, \theta_{N+1}^0 \in [0, 2\pi]$, such that

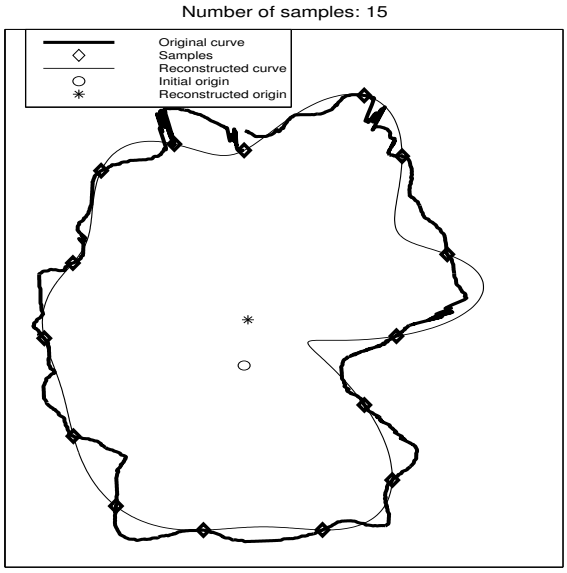
$$\begin{cases} d^2\left(\theta_1^0, \ldots, \theta_{N+1}^0\right) & = & 0, \\ l^2\left(\theta_1^0, \ldots, \theta_{N+1}^0\right) & = & 0. \end{cases} \tag{2.30}$$

Because of the non-negativity of the functions d^2 and l^2, we conclude that the point $\left(\theta_1^0, \ldots, \theta_{N+1}^0\right)$ is Pareto optimal, which also shows that our sampling technique described in the previous sections is a particular case of the multi-criteria optimization.

2.4.2 Implementation and Results

Multi-criteria optimization is quite well known in the mathematical community as well as engineering design, finance and other sciences. Several methods of finding a solution to a multi-criteria

Figure 2.7 A good approximation of a map of Germany may be obtained from as few as 15 samples.



optimization problem (Pareto optimal point) have been proposed. Here we present the results obtained by using the steepest-descent method, which is while generally slow but proven to achieve the global optimal point. For the sake of space we defer the description of this method to [16] and only present the simulation results.

Figure 2.7 contains a silhouette of Germany. One can see that a reasonable approximation can be achieved using only 15 samples. A similar example is presented in Figure 2.8, where a very good reconstruction of a human kidney is made only from 7 samples. Finally, the shape of the human brain can be accurately encoded into 19 samples (Figure 2.9).

2.5 Further Extension: Sampling of Surfaces

The results of the previous sections may be straightforwardly generalized to the case of surfaces in $\mathbb{R}^3$. Much like having a curve, when choosing a polar coordinate system makes it effectively a

Figure 2.8 A contour of a human kidney can be recovered from only 7 samples.

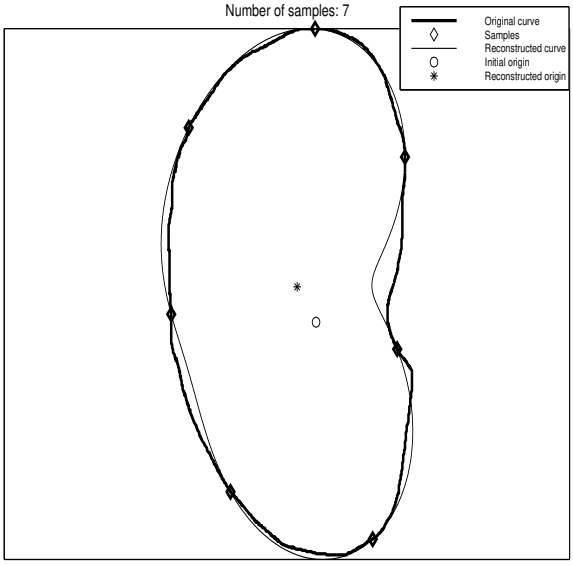
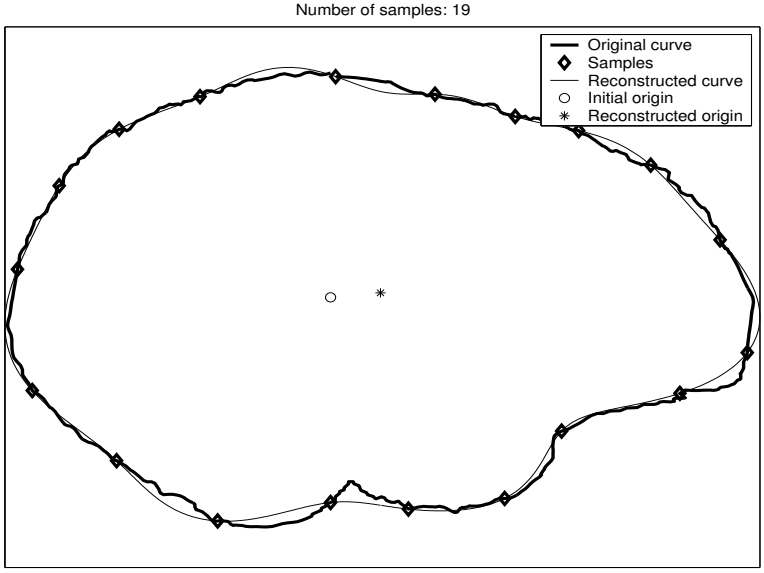


Figure 2.9 A contour of a human brain can be recovered from only 19 samples.



$1-D$ signal, a surface can be made into an image, by fixing spherical coordinates. Given that our interest focuses on $2-D$ curves, and that most of the results for surfaces are a simple generalization of the $2-D$ case, we forego any details and rather only state the results.

Definition 2.23 Suppose we have a surface S. If a point $\gamma_0 \equiv (x_0, y_0, z_0)$ is such that there exists a function $r(\theta, \phi) : [0, 2\pi]^2 \to \mathbb{R}^3$, such that the surface S can be parameterized according to

$$\gamma(\theta, \phi) = \gamma_0 + \Big(r(\theta, \phi)\cos(\theta)\cos(\phi), r(\theta, \phi)\cos(\theta)\sin(\phi), r(\theta, \phi)\sin(\theta)\Big), \tag{2.31}$$

where

$$\phi, \theta \in [0, 2\pi],\ \gamma(0, \phi) = \gamma(2\pi, \phi),\ \gamma(\theta, 0) = \gamma(\theta, 2\pi), \tag{2.32}$$

then we call γ_0 an admissible (spherical) center of S.

Theorem 2.24 Let S be a smooth surface, $(x_0, y_0, z_0) \in \mathbb{R}^3$, and ℓ - half-line starting from (x_0, y_0, z_0) and intersecting the image of S, set $\mathscr{S}$, in at least two points. Then there exists another half-line ℓ^t that also starts from (x_0, y_0, z_0) and is tangent to S.

Definition 2.25 Let S be a smooth surface and $(x_0, y_0, z_0) \in \mathbb{R}^3$ be a point. Then this point is called a regular admissible center if there exists no half-line ℓ going from (x_0, y_0, z_0) and tangent to S. A collection of all regular admissible centers will be called the admissible set of a surface S and denoted $\mathscr{A}_S$.

Theorem 2.26 Let S be a surface and $\mathscr{A}_S$ be its admissible set. Then $\mathscr{A}_S$ is convex, i.e.,

$$\Big(\gamma_1, \gamma_2 \in \mathscr{A}_C\Big) \Rightarrow \Big([\gamma_1, \gamma_2] \subset \mathscr{A}_C\Big), \tag{2.33}$$

where

$$[\gamma_1, \gamma_2] = \big\{\alpha\gamma_1 + (1-\alpha)\gamma_2 \mid \alpha \in [0,1]\big\}. \tag{2.34}$$

Theorem 2.27 Suppose that we have a curve S and $\gamma_0 \in \mathscr{A}_S$. Let $r(\theta, \phi)$ be a spherical representation of S centered at γ_0. Then if $r(\theta, \phi)$ is bandlimited, so that $\exists M, N \in \mathbb{N}$, such that

$$r(\theta, \phi) = \sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} c_{m,n} e^{i\theta n} \cdot e^{i\phi m}, \tag{2.35}$$

then $r(\theta, \phi)$ can be reconstructed from any samples so long as their number is sufficient.

Proposition 2.3 Let S be a curve and $\gamma_0 \in \mathscr{A}_S$. Without loss of generality we will assume that $\gamma_0 = (0,0,0)$. Suppose that $r(\theta, \phi),\ (\theta, \phi) \in [0, 2\pi]^2$ is a corresponding spherical parameterization of S. Then if $r(\theta, \phi)$ is bandlimited, then we can always choose sampling grid $\big\{(\theta_i, \phi_j)\big\}$ in such a way that

$$\sum_{i,j} \big(r(\theta_i, \phi_j)\cos(\theta_i)\cos(\phi_j), r(\theta_i, \phi_j)\cos(\theta_i)\sin(\phi_j), r(\theta_i, \phi_j)\sin(\theta_i)\big) = (0,0,0). \tag{2.36}$$

Proposition 2.4 Let S be a curve. Assume $\mathscr{A}_S \neq \emptyset$, and $\gamma_0 \equiv (x_0, y_0, z_0) \in \mathscr{A}_S$. Suppose that $r(\theta, \phi)$ is a spherical parameterization of S centered at γ_0. If $r(\theta, \phi)$ is bandlimited, then there exist distinct points (θ_i, ϕ_j) such that the curve S is uniquely defined by the corresponding samples of the surface.

2.6 Chapter Summary

Sampling of curves and surfaces is as important for a number of applications as sampling of $1-D$ or $2-D$ signals. It may help to compress certain images more efficiently, than other techniques that work with an entire image. Sampling also helps to provide us with shape landmarks, which are an essential tool for shape analysis and computer vision.

It was shown that a direct attempt to sample a curve encounters the difficulty, which is not topical for a signal or an image, namely that of finding a functional representation of the curve. Making sure that a curve will always have a parametrization of a prescribed type requires imposing constraints on the nature of the curve. It was demonstrated that for the case of polar parametrization such constraints could be found and efficient class identification numerical procedures could be implemented.

It was further discussed that unlike the $1-D$ case, samples of a curve are meaningless unless they carry information about a coordinate system in which the curve had been parameterized. We used the flexibility in choosing positions of sampling points when dealing with a $1-D$ bandlimited signal to incorporate the information about the polar center into the samples of a curve. The resulting procedure yielded a technique allowing us to reconstruct a curve from its samples without any additional information.

The approach we took for the case of ideal curves could be extended to curves outside that class. We proposed an optimization technique for sampling and approximating a curve for its samples, of which the previous sampling procedure arose as a particular case. That technique proved to provide us with very efficient discrete representations of a number of realistic contours.

Chapter 3

On a New Implementation of Simulated Annealing and Its Application to Signal and Image Filtering

3.1 Introduction

Many tasks in signal and image processing such as filtering, segmentation, pattern recognition and other applications may be posed in the context of an optimization problem. Given an initial or observed data f_0, to any signal f we associate a certain value $\mathscr{H}_{f_0}(f)$. The functional $f \rightarrow \mathscr{H}_{f_0}(f)$, which assigns a number to a signal, is generally referred to as the signal energy or cost. The energy (cost) of a signal is thus a function of the signal and the observed data. Our goal is then to minimize this cost, i.e. find a signal $\hat{f}$ with the lowest energy. Mathematically, we can write it as

$$\hat{f} = \arg\min_{f} \mathscr{H}_{f_0}(f). \tag{3.1}$$

The solution to the problem (3.1) then becomes the output of a system that is being designed.

The problem of finding a cost functional $\mathscr{H}$ suitable for a particular application is challenging. In applications related to signal or image filtering, the observed data consists of a noisy signal, and eliminating the noise component while maximally preserving the original signal (see Figure 3.1), is often the problem.

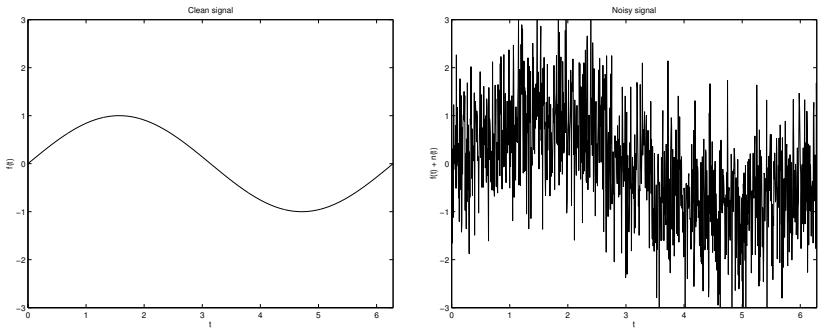


Figure 3.1 Example of a clean signal and the same signal corrupted by additive Gaussian noise.

A general approach to such a problem is to search for a signal which is close to the observation with all irregularities believed to be due to the noise factored out. A functional destined for filtering would therefore require two components. One would penalize large deviations from the observed data, and the other one would penalize signal abnormalities that are rather attributed to the noise than to the clean signal, i.e. impose some regularity on the signal.

The complex structure of the functional almost invariably results in the latter being multi-modal, i.e. having more than one local extremum (see Figure 3.2). A typical gradient descent method which is used to search for a local minimum thus becomes suboptimal, and additional effort is required to either improve on the extremum or in the best case converge to a global extremum.

Several approaches are available when searching for a global minimum of a multi-modal cost functional. Most of them fall into the category of stochastic optimization algorithms. Of particular interest is the algorithm referred to as simulated annealing together with its modifications. Simulated annealing is a technique that minimizes a given cost functional by combining a deterministic descent such as the gradient descent method, with stochastic perturbations which enable us to avoid attracting local extrema.

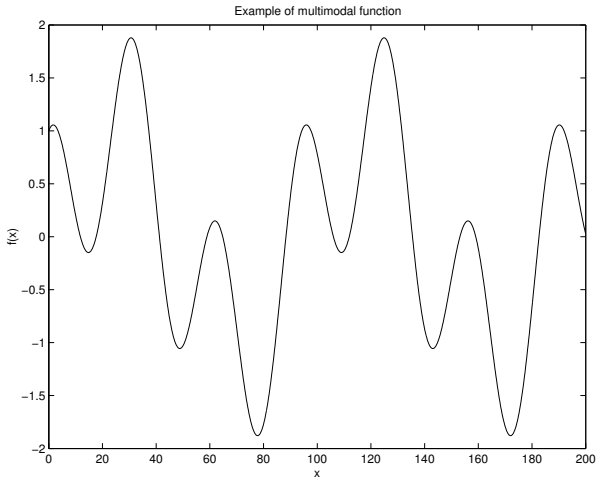


Figure 3.2 Many functionals exhibit multi-modal behavior, which means they have several local extrema.

Indisputable advantages of many simulated annealing based algorithms pertain to their ability accommodate complex cost functionals with large number of degrees of freedom in a straightforward and easy-to-implement manner. In addition, in many cases an algorithm from this class may be proven to be optimal in the sense of convergence to a global minimum.

On the other hand, straightforward implementations of simulated annealing suffer from serious shortcomings, perhaps the most important of which is their low rate of convergence. Many algorithms have been proposed in the literature in order to address this problem and potentially achieve a superior performance at a cost of greatly complicating the stochastic component of the algorithm. Our goal in this chapter is in this direction by proposing a modification of a classical simulated annealing scheme, to attain a faster descent while maintaining the simplicity of the original idea and its implementation.

The convergence of the simulated annealing algorithm is most often measured through the rate of decay of the temperature, which essentially controls the variance of the stochastic component. The logarithmic rate of the standard Boltzman technique in practice yields an almost constant temperature over a reasonable time interval, which is equivalent to an absence of convergence in practice. Even deeper insight may be obtained if we look at a discrete version of the algorithm from a particle viewpoint. The stochastic component forces the particle to "explore" the neighborhood of a particular local minimum at a certain scale, which can be heuristically understood as the size of a typical random jump. For a fixed temperature, the behavior of the functional (at a scale either too small or too large) will remain unexplored for a very long time, which often results

in an extremely large time period before the particle finds a better stable state, i.e. another local minimum with smaller energy.

A step in the direction of dealing with this problem was the well-known fast simulated annealing or the Cauchy machine proposed by Szu, whose idea was to make the tails of the distribution of jump sizes heavier, hence allowing occasional large jumps, which would help exploring areas far from the current position of the particle much faster. This approach unfortunately only partially addresses the scale problem described above. The technique uniformly fails at fine scales, as a particle may be moved by a large jump with no regards to whether a better local minimum is in proximity. It has also been reported that the Cauchy machine does not converge in theory, which limits its practical importance.

Our approach is based on systematically exploring all scales while maintaining the simplicity of implementation. In lieu of keeping the temperature constant, we propose to vary it non-monotonously by first exploring the immediate neighborhood of the current local minimum and then by proceeding to progressively increase its size. While much of the theoretical analysis is left as part of future work, experimental results show that a significant improvement in achieving the lowest possible energy state within a minimal time.

Most filtering algorithms are designed as discrete iterative procedures, where the observed data becomes the initial signal, to be evolved according to a particular minimization technique. In minimizing a given functional, initial states far away from the optimizer are usually worse than others. In those cases, achieving the global minimum may be very time consuming if at all attainable. We will demonstrate here that the evolution of several signals uniformly distributed in

the region where the global minimum is most likely to be found, results in further improvement of the search algorithm. Some initial states, due to their close proximity to the global minimum, faster converge to the optimal state.

This chapter is organized as follows. Section 3.2 provides the general background about stochastic optimization in general, and simulated annealing in particular. Section 3.3 is devoted to the description of the newly proposed algorithm based on using adaptive temperature to obtain a fast implementation of simulated annealing. Section 3.4 discusses the evolution of perturbations of the original signal as opposed to the original signal, which yields better minimization results. A summary of the chapter is given in Section 3.5.

3.2 Simulated Annealing

Consider a problem of finding the global minimum of $\mathscr{H}_{f_0}(f)$, where $f_0 \in \mathbb{R}^d$ for some integer $d \geq 1$, and $\mathscr{H}_{f_0}(\cdot) : \mathbb{R}^d \to \mathbb{R}$. In other words, find $\hat{f} \in \mathbb{R}^d$, such that

$$\hat{f} = \arg\min_{f \in \mathbb{R}^d} \mathscr{H}_{f_0}(f). \tag{3.2}$$

Here f, f_0 are multidimensional vectors, which represent discrete signals or images; f_0 being the observed data.

The problem (3.2) can be solved by introducing an interactive diffusion process $\{f(t)\}_{t\geq 0}$, such that $\forall d \geq 0,\ f(t) \in \mathbb{R}^d$. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a 1D function, which we will call the "temperature". Consider the following stochastic differential equation with the initial condition

$$df(t) = -\nabla_f \mathscr{H}_{f_0}(f)\, dt + \sigma(t)\, dw(t), \tag{3.3a}$$

$$f(0) = f_0, \tag{3.3b}$$

where ∇_f denotes the gradient $w(t)$ is a d-dimensional Brownian motion. The right hand side of the equation (3.3a) consists of two terms. The first one is deterministic and represents nothing but the famous gradient descent, while the second one is diffusive and called for to de-trap the process $f(t)$ from local minima of the functional $\mathscr{H}_{f_0}(\cdot)$.

The equation (3.3a) is by definition equivalent to the following integral representation for the process $f(t)$.

$$f_i(t) = f_{0,i} + \int_0^t \left[-\frac{\partial}{\partial f_i} \mathscr{H}_{f_0}\big(f(u)\big) \right] du + \sigma(t) \int_0^t dw_i(u),\ i = 1, \ldots, d, \tag{3.4}$$

where

$$f(t) \equiv \big(f_1(t), \ldots, f_d(t)\big)^T, f_0 \equiv \big(f_{1,0}, \ldots, f_{d,0}\big)^T, w(t) \equiv \big(w_1(t), \ldots, w_d(t)\big)^T.$$

Considering Equation (3.3a) with

$$\sigma(t) = \frac{\sigma_0}{\sqrt{\log(t+1)}}, \tag{3.5}$$

where σ_0 is a fixed constant, yields the convergence of the stochastic differential equation (3.3) to the solution $\hat{f}$, which is the global minimum of $\mathscr{H}_{f_0}(\cdot)$, i.e.

$$\mathscr{H}_{f_0}\left(\hat{f}\right) = \min_{f \in \mathbb{R}^d} \mathscr{H}_{f_0}(f). \tag{3.6}$$

3.2.1 Discrete Approximation of the Filtering Scheme

The implementation of the above scheme entails discretization of the above equations. Towards that end, we construct a discrete time process $\{g(n)\}_{n=1,2,\ldots}$, which will approximate the continuous diffusion process $\{f(t)\}_{t \geq 0}$.

Consider a fixed interval $(0, t)$ and let $\delta = \frac{1}{N}$, where N is a (large) integer number. Define

$$\tau_n = n\delta,\ n = 1, \ldots, N. \tag{3.7}$$

Then the process $\{g(n)\}$ may be constructed using the following iterative scheme,

$$g_i(n+1) = g_i(n) - \frac{\partial}{\partial g_i}\mathscr{H}_{f_0}\big(g(n)\big) \cdot \delta + \sigma(t) \cdot \big(w(\tau_{n+1}) - w(\tau_n)\big), \tag{3.8a}$$

$$g(0) = f_0. \tag{3.8b}$$

The diffusion (3.3) is approximated by the discrete process (3.8) as spelled out by the following theorem.

Theorem 3.1 (Descombes, Zhizhina) Under certain very mild assumptions on the functional $\mathscr{H}_{f_0}(\cdot)$, the approximation process (3.8) strongly converges to the diffusion (3.3) with order $\frac{1}{2}$. More precisely,

$$\forall t \geq 0,\ \max_{i=1,\ldots,d} \mathbb{E}\Big[\big|f_i(t) - g_i(n_t)\big|\Big] \leq C(t)\sqrt{\delta}. \tag{3.9}$$

Theorem (3.1) assures that for δ small enough, the behavior of the discrete approximation $\{g(n)\}$ is indistinguishable from its continuous counterpart $\{f(t)\}$. Thus we should expect from the discrete process, the same filtering properties that are attributed to the continuous diffusion.

The schemes described by Equations (3.3) and (3.8), are examples of a simulated annealing-based technique. That unfortunately means that they both suffer from the usual problems of such techniques. Observe from Equation (3.5) that the convergence of the diffusion process as well as

its discretized version is slow due to the slow decay of the diffusion coefficient. In practice, this tantamount to introducing a Gaussian component with virtually constant variance at each iteration of (3.8), which, of course, adversely affects the convergence of the algorithm within any reasonable computational time.

We next keep the difference equation (3.8) as a guide when constructing a framework with enhanced convergence properties.

3.3 Adaptive temperature

Consider a particular case of the optimization problem in Equation (3.2) when $d = 1$. Here we will use the following notation. Let $F : \mathbb{R} \to \mathbb{R}$ be a function defined on the real line and sufficiently smooth. Our goal is to find the global minimum $\hat{u}$ of $F(\cdot)$, i.e.

$$\hat{u} = \arg\min_{u \in \mathbb{R}} F(u). \tag{3.10}$$

The simulated annealing technique presented above yields the following stochastic differential equation to solve Equation (3.10).

$$du(t) = -F'\big(u(t)\big)\, dt + \sigma(t)\, dw(t), \tag{3.11a}$$

where $w(t)$ is a 1D Wiener process. The discrete approximation of equation (3.11a) can then be written as follows:

$$u_{n+1} = -F'(u_n) \cdot \delta + \sigma_n \cdot w_n, \tag{3.11b}$$

where $\{w_n\}_{n=1,2,\ldots}$ are independent identically distributed Gaussian random variables

$$w_n \sim \mathcal{N}(0, \delta),\ n = 1, 2, \ldots$$

and the temperatures $\sigma_n, n = 1, 2, \ldots$ are defined as

$$\sigma_n = \frac{1}{\sqrt{\log(n\delta + 1)}}.$$

It is apparent from Equation (3.11b) that a slow decay of the coefficients σ_n results in introducing Gaussian noise of virtually constant variance at each step of iteration. It is convenient to think of the time interval δ as determining the scale at which the search is being carried out. When δ is fixed and the temperature σ_n varies slowly, the search process takes place at essentially one fixed scale. If local minima are closely clustered in relation to a typical jump size of a diffusing particle, then many including those with lower energies will be overlooked for an extended period of time. On the other hand, if the distance between extrema is too large, then the diffusion energy is too small to allow a fast exploration of the functional and hence an efficient localization of new extrema, i.e. the particle is trapped in one mode for too long. The following proposition provides the exact relationship between the inter-local minima distance and the scale of the discrete diffusion.

Proposition 3.1 Let $F : \mathbb{R} \to \mathbb{R}$ be a 1D function, and $\alpha \in \mathbb{R}$ a fixed real constant. Define $G : \mathbb{R} \to \mathbb{R}$ by

$$\forall x \in \mathbb{R},\ G(x) = F(\alpha x). \tag{3.12}$$

Then the discrete procedure of minimization of the function G on the scale δ is equivalent to minimizing the function F at the scale $\alpha^2\delta$.

Proof. The minimization evolution for the function G is of the form

$$du(t) = -G'\big(u(t)\big)\, dt + \sigma(t)\, dw(t). \tag{3.13}$$

Let $\delta > 0$ be a time step. Then the discrete equation corresponding to the continuous evolution (3.13) is

$$u_{n+1} = u_n - G'(u_n) \cdot \delta + \sigma_n \cdot w_n, \tag{3.14}$$

where $\{w_n\}_{n\in\mathbb{N}}$ are independent random variables identically distributed according to $\mathcal{N}(0, \delta)$. Taking into account that $G'(x) = \alpha F'(\alpha x)$, we then have

$$u_{n+1} = u_n - \alpha F'(\alpha u_n) \cdot \delta + \sigma_n \cdot w_n \tag{3.15}$$

Multiply both sides by alpha and introduce a new variable

$$\forall n \in \mathbb{N},\ v_n = \alpha u_n. \tag{3.16}$$

Then

$$v_{n+1} = v_n - \alpha^2 F'(v_n) \cdot \delta + \alpha\sigma_n \cdot w_n. \tag{3.17}$$

Denote $\eta = \alpha^2\delta$ and $\omega_n = \alpha w_n$. The last equation can then be rewritten in the form

$$v_{n+1} = v_n - F'(v_n) \cdot \eta + \sigma_n \cdot \omega_n, \tag{3.18}$$

where ω_n are now independent Gaussian random variables with variance $\eta \equiv \alpha^2\delta$. Equation (3.18) is nothing but the discrete evolution for F with time step $\alpha^2\delta$. The proposition is therefore proven.

The significance of the above proposition lies in the shown fact that an optimal scale as defined by the size of the discrete time step δ, depends on the functional to be optimized. Different functionals should therefore require different δ's. To that end, we propose an algorithm, where

the discretization step would vary dynamically, thus automatically adapting to the behavior of the functional.

The algorithm described below has the following remarkable properties.

- It provides a reasonable solution to the scale problem outlined earlier by varying the scale of the discrete evolution when the solution cannot be found at a given scale.
- As a result of this dynamical approach, a particle tends to spend less time in local minima and is more efficient at localizing lower energy states.
- The algorithm is easy to implement, which allows us to easily use it for optimization problems over spaces of high dimensions.

Our algorithm is detailed next. For simplicity we will present its 1-D version. A multi-dimensional generalization is straightforward. Assume there is a fixed function $F : \mathbb{R} \to \mathbb{R}$. Let $u_0 \in \mathbb{R}$ be the initial state for the iteration.

1. Recall that a stochastic term was introduced in the original continuous differential equation and similarly its discrete counterpart so as to prevent the evolving particle from begin trapped in one local minimum. Until a local minimum is reached, the diffusion term is not needed. We therefore begin by setting up a deterministic gradient descent until the evolving point reaches the first local minimum.

$$u_{n+1} = u_n - F'(u_n) \cdot \delta, \ n = 0, 1, \ldots \tag{3.19}$$

2. Denote this local minimum by $u^{(1)}$.

3. Once the particle reaches a local minimum, i.e. $|F'(u_n)| < \epsilon$, where $\epsilon \ll 1$ is a predefined constant, a stochastic term is initiated. Define a sequence $\{\alpha_n\}_{n=1,2,\dots}$ starting from 1 and monotonously but slowly increasing.

4. Continue the particle evolution from point $u^{(1)}$, according to the stochastic difference equation

$$u_{n+1} = u_n - F'(u_n) \cdot \delta\alpha_n^2 + \alpha_n \sigma_n \cdot w_n, \tag{3.20}$$

where w_n are independent identically distributed random variables with Gaussian distribution $\mathscr{N}(0, \delta)$.

5. The stochastic diffusion stops once the particle leaves the neighborhood of a local minimum, i.e. when it finds a state $u^{(1,0)}$, such that

$$\left|F'\left(u^{(1,0)}\right)\right| > \epsilon.$$

6. The point $u^{(1,0)}$ is now regarded as a new initial point, and we go back to step 1.

A few remarks about the preceding algorithm are in order.

- Note again that each local minimum is found by applying a purely deterministic scheme. In particular, that allows to precisely pinpoint the location of the extremum as well as the time when it is achieved.

- Suppressing the stochastic term guarantees that once the point finds itself in the neighborhood of a local minimum, it will not be driven away from it at the next steps of the iteration but instead converge to the minimum.

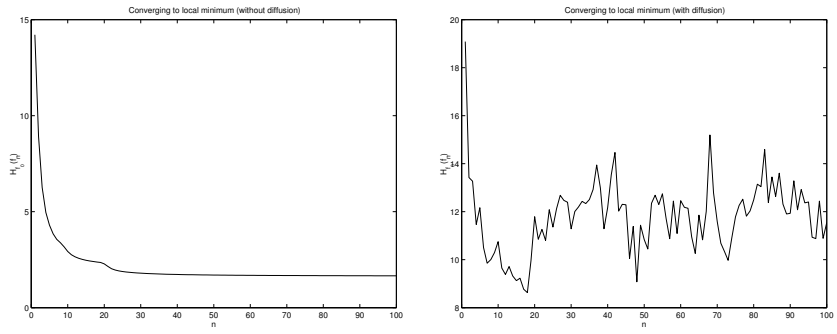


Figure 3.3 When the local minimum is found via a deterministic descent it is possible to identify the exact location of the extremum much simpler than in the case of a stochastic diffusion.

- Slowly increasing the sequence $\{\alpha_n\}$ in the Equation (3.20) is equivalent to varying the scale of the diffusion from finer to coarser. It therefore allows the particle to first explore the function in the immediate neighborhood. If no better extremum is found in the immediate vicinity of the current extremum, the search is performed progressively further and further away. Since a sufficient number of samples is required at each scale to ensure a high probability of locating the right point, the sequence must increase fairly slowly, e.g.

$$\alpha_{n_k} = \gamma n_k, \; k = 1, 2, \ldots$$

where $\gamma \ll 1$ and n_k is a time subsequence corresponding to step 4 of the above algorithm.

- While technically speaking the convergence to the global minimum is hard to guarantee, experiments we have conducted demonstrate that the algorithm is capable of finding better and better local minima much faster than the classical technique. In practice one can use heuristic stopping criteria such as the maximum number of iterations or the lowest energy, may be used to terminate the evolution in a finite time.

- A very significant feature of the proposed algorithm is its simplicity from a computational standpoint. It is therefore not surprising that its easy implementation makes it competitive with the Boltzmann machine.

Figures 3.4 and 3.6 illustrate typical scenarios that occur when one compares the Boltzmann technique to the newly proposed method. More aggressive and adapted search for a better state proves more fruitful and yields a better minimization result. The functionals chosen for this experiment is

$$\mathscr{H}_{f_0}(f) = \beta \sum_{i=1}^{d} (f_i - f_{0,i})^2 + (1-\beta) \sum_{i=2}^{d} \Delta_\theta (f_i - f_{i-1}), \tag{3.21}$$

where $\beta \in (0,1)$ is a fixed constant and

$$\Delta_\theta(x) = \begin{cases} x, & |x| \le \theta, \\ -\theta, & x < -\theta, \\ \theta, & x > \theta, \end{cases} \tag{3.22}$$

for some $\theta > 0$. This functional has a data fidelity term, which is a mean-square distance between an evolving signal and the original data, as well as a smoothness term, which progressively penalizes a high gradient attributing it to the noise, while putting smaller penalty on jumps that are likely to have come from the original clean signal.

The vector f represents the digitized signal or image being evolved, and the vector f_0 is a digital version of the original data.

Examples show that for this functional both in $1-D$ and in $2-D$ case, the proposed technique does indeed provide a superior performance for the optimization. Lower energy states are achieved much faster, corresponding to better filtering properties.

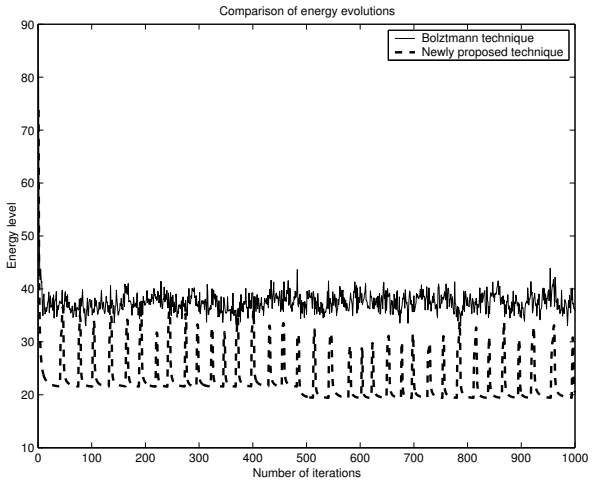


Figure 3.4 A new implementation of simulated annealing for a signal filtering functional yields faster convergence to the global minimum.

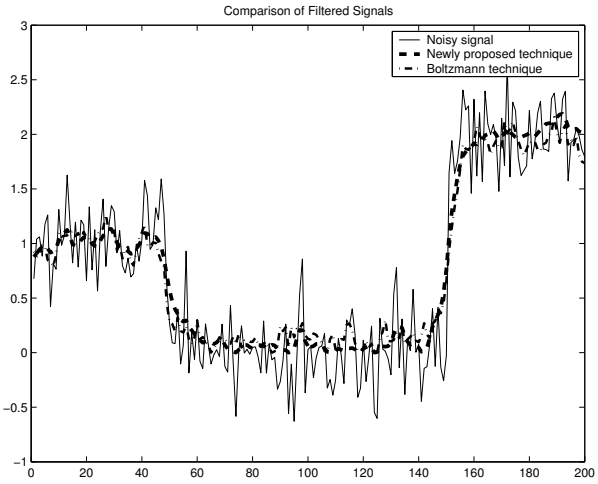


Figure 3.5 Lower energy achieved naturally means better filtering result for a noisy signal.

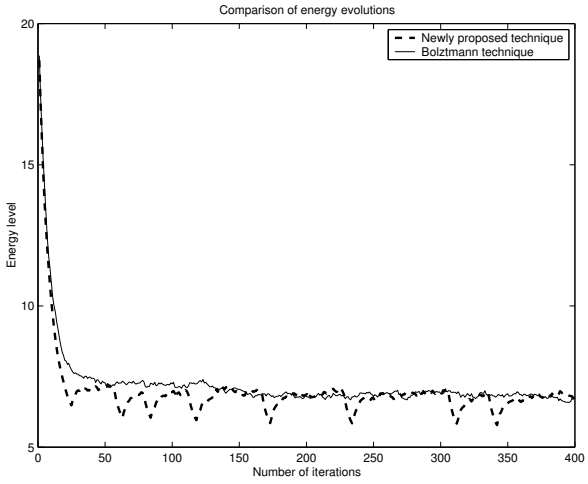


Figure 3.6 A new implementation of simulated annealing for an image filtering functional yields faster convergence to the global minimum.

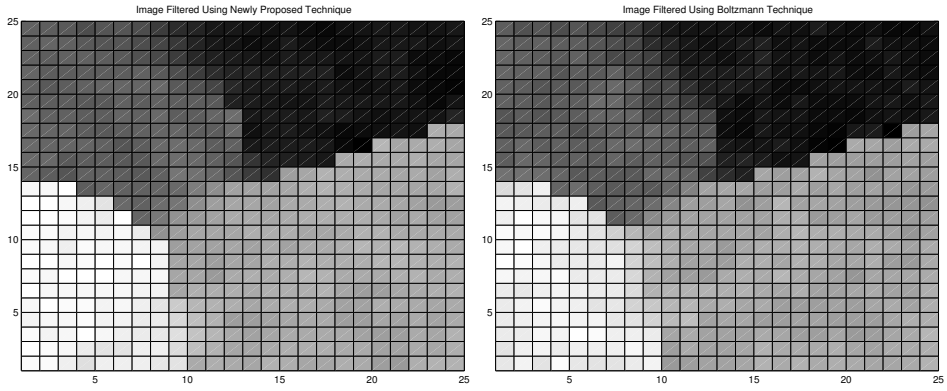


Figure 3.7 Higher energy level yields more visible noise while lower energy level yields cleaner image.

3.4 Evolution of Several Perturbations of the Original Signal

In the previous sections, a new simulated annealing based algorithm was developed and allowed us to achieve a better speed performance of an arbitrary functional energy decreasing towards its global minimum. The idea of adaptive temperature turns out to be quite intuitive as was illustrated in the $1-D$ case. Furthermore, it has been shown that it is applicable to a large class of functionals in multi-dimensional spaces.

The intuition of $1-D$ picture may be similarly transitioned and generalized to multidimensional discrete signals in search for the lowest possible energy.

Consider a simple example of a bi-modal function $F(u)$, which has two local minima, only one of which is global. It is clear that virtually any initial point located in the same cavity as the global minimum will yield a faster convergence to the global minimum than one positioned close to the local minimum. Thus we would like to start the evolution as close to the global minimum as possible. Because the exact form of the function $F(u)$ is not known a priori, to make a perfect guess is clearly an impossible task. The solution is then to uniformly throw several points in the interval where the global minimum is likely to be found and evolve them simultaneously, so that one realization would converge to the solution faster than others.

This simple idea can be exploited to solve minimization problems for high-dimensional functionals as well. Here, instead of a $1-D$ point u_0 the usual initial state is a signal $f_0 \in \mathbb{R}^d$. As it was argued earlier, the functionals used for filtering purposes (such as one defined by Equation 3.21) have a data fidelity term, its global minimum is unlikely to find too far from the original data. It is thus reasonable to produce the desired perturbations by adding random independent identically

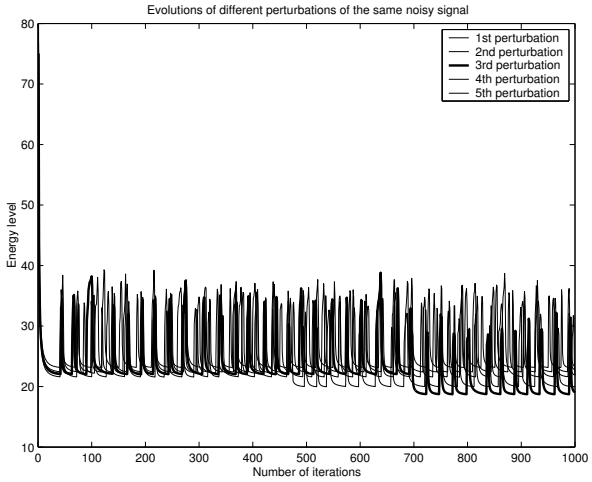


Figure 3.8 Evolving a perturbed version of the original signal may yield better minimization results.

distributed noise to each component of the digital vector, e.g.

$$f_i^j = f_{0,i} + n_i^j, \tag{3.23}$$

where n_i^j are independent uniformly distributed on the interval $[-n_{max}, n_{max}]$ random variables.

One or several of those perturbations is likely to descend to a global minimum faster than the original data. Figure 3.8 demonstrates an example of such a phenomenon occurring in practice.

3.5 Chapter Summary

In this chapter we have considered a problem of signal filtering using a stochastic optimization method. A particular technique that was employed was the celebrated simulated annealing, which combines the gradient descent method along with the diffusion that keeps the evolution from settling in local minima.

The major shortcoming of the classical implementation is its slow speed of convergence. We proposed a qualitatively new explanation for this phenomenon in terms of a particle that evolves over time searching for states with lower energy. It was argued that the classical technique along with some of its modifications restrict the search to certain scales that may not adapt well to the peculiarities of a functional of interest. This was confirmed by the superior performance of the newly proposed algorithm.

The new implementation of the simulated annealing based filtering diffusion is through using of non-monotonously varying (adaptive) adaptive temperature. By continuously increasing the scale, the evolution is capable of exploring the behavior of the functional being minimized at each scale, and successfully locate states with lower energy thus advancing the energy minimization process. The experimental data confirms that an intuitively promising idea is also a practical success.

Finally, we showed that one can further improve the search algorithm by evolving several perturbations of the original signal for a shorter time rather than just the original signal for a longer time.

The key observation behind that idea is that for an arbitrary functional, especially one with many modes, not all initial states are equivalent in terms of the resulting performance of the evolution. Starting the evolution far from the the location of the global minimum may result in trapping of the evolving signal in a wrong mode. That obstacle may be circumvented by randomly throwing several initial signals around the unknown global minimum. When several perturbations of the original signal are evolved, one realization may be more probable cadidate than the original signal and in fact converge to the global minimum much faster.

Chapter 4

Identification of a Discrete Planar Symmetric Shape from a Single Noisy View

4.1 Introduction

Automatic understanding of objects from their images is a fundamental problem of computer vision ([14]). In the vast majority of practical situations we do not have a hands-on access to an object we are interested in, but instead are bound to study it relying on its images obtained through various imaging devices. Because the world we live in is three-dimensional, and images are typically two-dimensional, the problem of extracting all possible information about the object from its images is fundamentally ill-posed.

Mathematical formulations of the problem of identifying an object very drastically depending on an exact practical setting and therefore assumptions that are considered natural ([35, 56, 57, 61, 47]). A lot of research has been devoted to studying object based on two and more images ([15, 7, 46]). It is clear that the more data is available about the object, the more information can potentially be extracted from it. On the other hand, in many cases a single image of an object is all that is known about the object and being able to fully utilize that knowledge becomes an important problem.

This chapter is devoted to the problem of identifying a symmetric planar discrete shape from a single image. By a discrete planar shape we will understand a collection of discrete planar points. A discrete shape is seemingly unnatural model because all objects that we observe in real life are continuous, and so are their images. However, one can make a strong argument that for all practical purposes, including recognition, a shape is often well defined by a finite number of its important points (landmarks) ([11, 23, 24, 17]). The continuous original then can either be reconstructed from those points, or the discrete shape itself can provide us with sufficient amount of information about the object that we seek.

It has been shown that if no additional assumptions have been imposed on a discrete object, no knowledge whatsoever can be learned from a single image, i.e. different objects can generate the same image and thus there is no way to tell any of them apart ([6]). The assumption of symmetry allows to avoid that difficulty, while at the same time keeping the class of shapes sufficiently broad and of practical interest in many applications.

In the first part of this chapter we show that virtually all geometric information about any shape satisfying the above conditions can be conveniently encoded in what we refer to as the skeleton of the shape and subsequently the signature of the skeleton ([43]). Skeleton of a shape has a very simple geometric structure, however it carries all information about the shape that we are interested in. What makes it a very useful object is that the skeleton of a shape can be computed from any image of an object, which gives us the ability to analyze the object regardless of the concrete image that we have.

The second part of the chapter is devoted to adapting methods developed in the first part to noisy shapes. As was noted in the earlier chapter, discretization can be performed either by a sampling algorithm or manually. In either case the accuracy of landmarks is a very relevant issue. Noise in samples of a shape yields that whatever properties discrete shapes of the class of interest have are, strictly speaking, no longer satisfied. It is only natural then that the developed methods appropriate for symmetric shapes stop working. As the shape skeleton is the keystone of our approach to study shapes, we study its behavior in a situation of a noisy image and derive the optimization procedure, which still allows us to recover the information about the object despite the noise.

4.2 Elements of Projective Geometry

4.2.1 Basic Definitions

Definition 4.1 Consider the Euclidean space $\mathbb{R}^{n+1}$. Define the equivalence relationship $\sim$ on $\mathbb{R}^{n+1}$ as follows:

$$\forall x, y \in \mathbb{R}^{n+1},\ (x \sim y) \Leftrightarrow (\exists \lambda \in \mathbb{R} :\ x = \lambda y) \tag{4.1}$$

The equivalence relationship defined above partitions the space $\mathbb{R}^{n+1}$ into disjoint equivalence classes. The space of these equivalence classes is called the n-dimensional projective space and is denoted $\mathbb{P}^n$.

Projective space may $\mathbb{P}^n$ be regarded as the extension of the Euclidean space $\mathbb{R}^n$. Indeed, consider a point $(x_1, \ldots, x_n)^T \in \mathbb{R}^n$. Its corresponding projective coordinates will then be $(x_1, \ldots, x_n, 1)^T$. If on the other hand, $(x_1, \ldots, x_n, x_{n+1})^T \in \mathbb{P}^n$ is a point in the projective space, then it is clear from the definition of the projective space that we can uniquely associate to this point its Euclidean

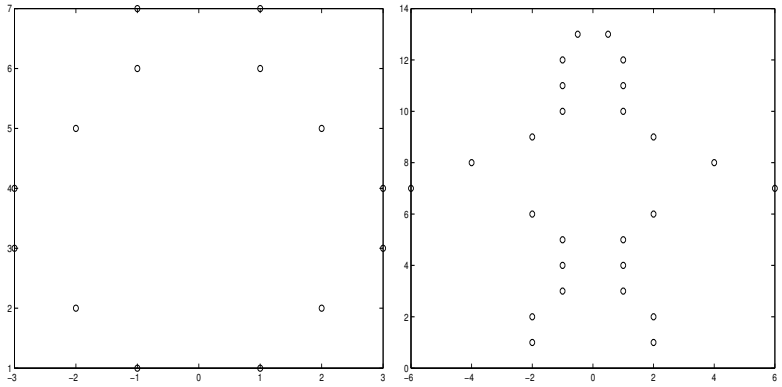


Figure 4.1 Examples of discrete objects.

coordinates $\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}\right) \in \mathbb{R}^n$ as long as $x_{n+1} \neq 0$. Projective points with the last coordinate equal to zero do not have Euclidean counterparts and may be regarded as "points at infinity". Indeed, points of the form $(Cx_1, \ldots, Cx_n, 1) \in \mathbb{P}^n$, where $C \to +\infty$, constitute a line in the projective space, which is equivalent to the Euclidean line $(Cx_1, \ldots, Cx_n) \in \mathbb{R}^n$. Unlike the Euclidean line, which is infinite, the projective line has an "end point".

$$\lim_{C \to +\infty} (Cx_1, \ldots, Cx_n, 1) = \lim_{C \to +\infty} \left(x_1, \ldots, x_n, \frac{1}{C}\right) = (x_1, \ldots, x_n, 0). \tag{4.2}$$

This equation justifies the term "point at infinity". Sometimes such a point $(x_1, \ldots, x_n, 0)$ is also referred to as the direction $(x_1, \ldots, x_n)$. For example, in the case of the $2 - D$ projective space the point $(1, 0, 0)$ corresponds to the horizontal direction, and $(0, 1, 0)$ corresponds to the vertical direction on the plane.

4.2.2 Imaging in $3 - D$ Projective Space

Since the objects surrounding us are $3 - D$, it is convenient to model them as collections of points lying in the $3 - D$ projective space $\mathbb{P}^3$.

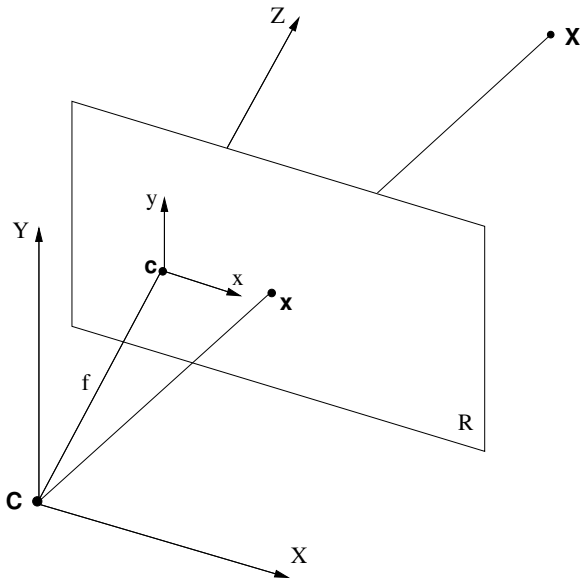


Figure 4.2 Pinhole camera.

An image of a given object in $\mathbb{P}^3$ is itself an object lying, however, in the $2-D$ projective space $\mathbb{P}^2$. Because our goal here is to infer information about the object from its single image, it is natural to start with understanding of the mathematical model of image formation.

A commonly used model for the camera is a so-called pinhole camera. A pinhole camera is defined by the retinal plane R, and the optical center C. For a given point $\mathbf{X}$ its image $\mathbf{x}$ through this camera is obtained by intersecting the line connecting the points C and $\mathbf{X}$ with the plane R.

The line that contains the optical center C and is perpendicular to the retinal plane R is called the optical or principal axis. It intersects the retinal plane at the point c. The distance between the points C and c is denoted by f and called the focal distance.

The relationship between the Euclidean coordinates of the world point $\mathbf{X}$ and its image $\mathbf{x}$ are then of the form

$$\begin{cases} x &= f\frac{X}{Z}, \\ y &= f\frac{Y}{Z} \end{cases} \tag{4.3}$$

or in the projective coordinates

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underbrace{\begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\mathtt{P}} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \tag{4.4}$$

We can see that in the projective coordinates the relationship between the coordinates of the object point and those of the image point is linear and of the form

$$\mathbf{x} = \mathtt{P}\mathbf{X}, \tag{4.5}$$

where $\mathbf{X}$ is the world point, $\mathtt{P}$ is the camera matrix, and $\mathbf{x}$ is the image of $\mathbf{X}$ through the camera $\mathtt{P}$.

4.2.2.1 Intrinsic Parameters of Camera

The relationship (4.4) assumes that the coordinate system in the retinal plane stems from the principal point c. This assumption may in general be violated. A more general form of the imaging mapping is then

$$\begin{cases} x &= f\frac{X}{Z} + c_x, \\ y &= f\frac{Y}{Z} + c_y \end{cases} \tag{4.6}$$

or in the projective coordinates

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} f & 0 & c_x & 0 \\ 0 & f & c_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \tag{4.7}$$

Furthermore, in real applications it is often unreasonable to assume that the two axes in the retinal plane have the same scale, i.e. pixels are not square (so-called CCD cameras).

To accommodate for that feature, we introduce the scaling coefficients α_x and α_y, which further generalize the form of the camera matrix to

$$\mathtt{P} = \begin{pmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{4.8}$$

where

$$\begin{aligned} f_x &= f\alpha_x, \\ f_y &= f\alpha_y. \end{aligned} \tag{4.9}$$

Introduce the matrix $\mathtt{K}$ as follows

$$\mathtt{K} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.10}$$

Then

$$\mathtt{P} = \mathtt{K}(\mathtt{I}_{3\times 3} \mid 0_{3\times 1}), \tag{4.11}$$

where $\mathtt{I}_{3\times3}$ is a 3×3 identity matrix. The matrix K is referred to as a camera calibration matrix and its entries f_x, f_y, c_x, c_y are called the intrinsic parameters of the camera. Each type of a physical camera has its own intrinsic parameters that can be carefully estimated in the lab conditions. For applications where we have access to the camera or its physical equivalent, we can safely assume the intrinsic parameters of the camera known.

4.2.2.2 Extrinsic Parameters of Camera

Recall that in the expression (4.11) the world point $\mathbf{X}$ has the projective coordinates $(X, Y, Z, 1)^T$, because the camera's optical center was assumed to be located at the origin of the Euclidean space and its principal axis was to coincide with the Z-axis. However, an arbitrary world coordinate system may be rotated and shifted with respect to the camera's Euclidean frame (see Figure 4.3) . Practically this is equivalent to saying that an object (or a camera) is shifted and rotated prior to imaging, i.e.

$$\mathbf{X}^0 = \mathtt{R}\left(\mathbf{X} - \mathbf{C}\right), \tag{4.12}$$

where $\mathtt{R}$ is the rotation matrix in $3-D$, and $\mathbf{C}$ is a vector representing the coordinates of the camera center in the actual world coordinate frame, $\mathbf{X}$ is the original object, and $\mathbf{X}^0$ is the transformed object to be imaged through the camera described above.

Combining (4.4) and (4.12), we get

$$\mathbf{x} = \underbrace{\mathtt{K}\mathtt{R}(\mathtt{I}_{3\times3} \mid -\mathbf{C})}_{\mathtt{P}}\,\mathbf{X}. \tag{4.13}$$

The rotation matrix $\mathtt{R}$ is a function of three parameters ϕ, ψ and α. The first two determine the Euclidean angles of the rotation axis, and the latter is the rotation angle (see Figure 4.4) .

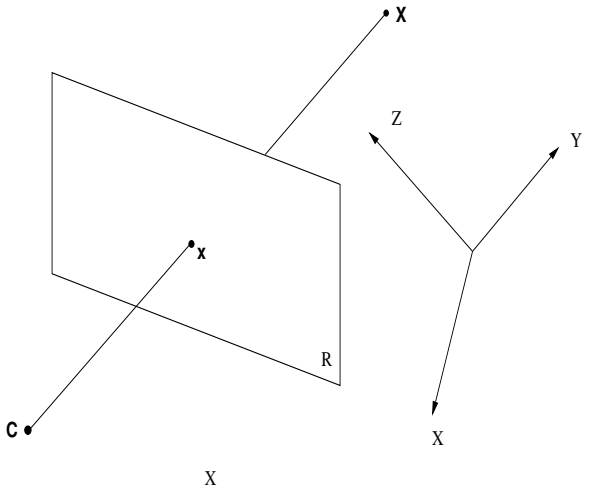


Figure 4.3 Arbitrary coordinate frame.

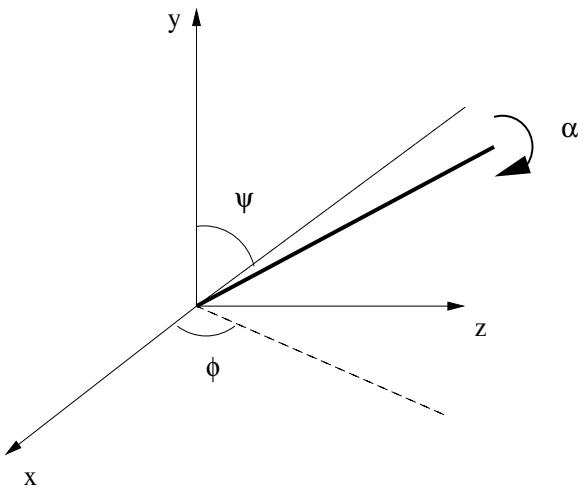


Figure 4.4 Rotation parameters.

The rotation parameters ϕ, ψ, α and the vector $\mathbf{C}$ are called the extrinsic parameters of the camera.

4.2.3 Views of Planar Shapes and Projective Transformations of Plane

Consider a planar point in the $3-D$ projective space $\mathbf{X}$, such that

$$\mathbf{X} = (X, Y, 0, 1). \tag{4.14}$$

Let $\mathtt{P}$ be an arbitrary camera matrix, and $\mathbf{x}$ be the image of $\mathbf{X}$ though that camera, i.e.

$$\mathbf{x} = \mathtt{P}\mathbf{X}. \tag{4.15}$$

Then

$$\mathbf{x} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{22} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{pmatrix}}_{\mathtt{H}} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}. \tag{4.16}$$

Definition 4.2 Introduce the notation $\mathbf{x}^0 = (X, Y, 1)^T$ and call it the frontal view of $\mathbf{X}$. A point $\mathbf{x}$ obtained by

$$\mathbf{x} = \mathtt{H}\mathbf{x}^0 \tag{4.17}$$

is called the projective transformation of $\mathbf{x}_0$.

A projective transformation is simply a linear transformation in projective coordinates. We have thus shown then that an arbitrary view of a planar shape can be viewed as the projective transformation with an arbitrary matrix $\mathtt{H}$ of its frontal view.

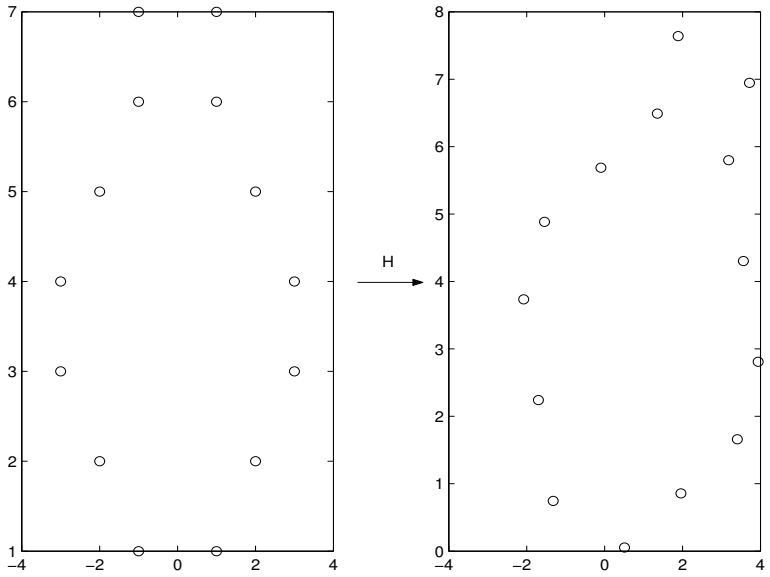


Figure 4.5 Frontal and skewed views of the same shape.

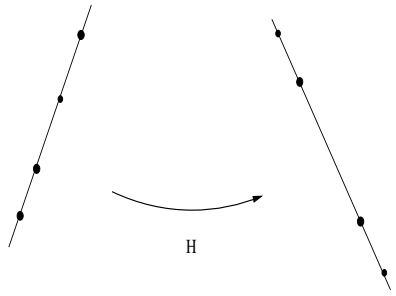


Figure 4.6 Projective transformations preserve collinearity of points.

4.2.3.1 Projective Transformations of Plane

Definition 4.3 A mapping $\mathscr{H} : \mathbb{P}^2 \rightarrow \mathbb{P}^2$ is called a projective transformation if it is of the form

$$\mathscr{H}(\mathbf{x}) = \mathtt{H}\mathbf{x}, \tag{4.18}$$

where $\mathtt{H}$ is a non-degenerate 3×3 matrix.

The following is the important property of the projective transformations.

Theorem 4.4 A map $\mathscr{H} : \mathbb{P}^2 \rightarrow \mathbb{P}^2$ is a projective transformation if and only if it preserves lines, i.e. if $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ belong to the same line in $\mathbb{P}^2$ then the points $\mathscr{H}(\mathbf{x}_1), \mathscr{H}(\mathbf{x}_2), \mathscr{H}(\mathbf{x}_3)$ all lie on the same line as well (see Figure 4.6).

It follows from this theorem that the projective transformations preserve collinearity of points - a property that will be made use of later.

4.3 Identification of Shape from Single View

Suppose we have a discrete symmetric planar object in $3-D$ projective space, by which we will understand a collection of points $\{\mathbf{X}_i\}_{i=1}^{2N}$, such that

$$\mathbf{X}_i \in \left\{\mathbf{X} \equiv (X, Y, Z, 1) \in \mathbb{P}^3 \mid Z = 0\right\},\ i = 1, \ldots 2N,$$

such that $\mathbf{X}_i$ is symmetric to $\mathbf{X}_{i+N}$ about the axis $\{Y = 0\}$, i.e.

$$X_i = -X_{i+N},\ i = 1, \ldots, N. \tag{4.19}$$

Assume further that the object is observed through a projective camera $\mathtt{P}$. A vector $\mathbf{x} \in \mathbb{P}^2$ obtained by

$$\mathbf{x} = \mathtt{P}\mathbf{X} \tag{4.20}$$

is then a view of the object $\mathbf{X}$ through the camera $\mathtt{P}$. Different camera matrices $\mathtt{P}$ generate different views of the object of interest. Exact parameters of a matrix $\mathtt{P}$ are functions of its extrinsic and intrinsic parameters of the physical camera modelled by $\mathtt{P}$.

Consider two different images, $\mathbf{x}_1$ and $\mathbf{x}_2$, of the object $\mathbf{X}$ produced by two different cameras $\mathtt{P}_1$ and $\mathtt{P}_2$. As was noted earlier, the two images are related through some projective transformation. More precisely, there exists a matrix $\mathtt{H}$, such that

$$\mathbf{x}_2 = \mathtt{H}\mathbf{x}_1.$$

In particular, all other views of the same object are projective transformations of $\mathbf{x}^0$.

4.3.1 Skeleton of Symmetric Planar Shape

Let $\{\mathbf{x}_i\}_{i=1}^{2N}$ be a fixed image of the object $\{\mathbf{X}\}_{i=1}^{2N}$, that is

$$\mathbf{x}_i = \mathtt{H}\mathbf{x}_i^0,\ i = 1, \dots 2N$$

for some fixed projective transformation (matrix) $\mathtt{H}$, where $\{\mathbf{x}_i^0\}_{i=1}^{2N}$ is a frontal view of the object. We now show that almost all projective geometry of $\{\mathbf{x}_i\}_{i=1}^{2N}$ can be conveniently captured by only $2N-3$ real parameters. Note that $\{\mathbf{x}_i\}_{i=1}^{2N}$ is in general determined by $4N$ real numbers ($2N$ $2-D$ points).

We start by introducing the so called skeleton of the shape (see Figure 4.7). ([43, 53]) Consider two pairs of adjacent symmetric points $\mathbf{x}_i, \mathbf{x}_{i+N}, \mathbf{x}_{i+1}, \mathbf{x}_{i+1+N}$. Using the notation $l_{\mathbf{x}_1,\mathbf{x}_2}$ for the line connecting two points $\mathbf{x}_1, \mathbf{x}_2$, we obtain two points

$$\begin{aligned} \mathbf{n}_i &= l_{\mathbf{x}_i,\mathbf{x}_{i+1+N}} \cap l_{\mathbf{x}_{i+1},\mathbf{x}_{i+N}}, \\ \mathbf{p}_i &= l_{\mathbf{x}_i,\mathbf{x}_{i+1}} \cap l_{\mathbf{x}_{i+N},\mathbf{x}_{i+1+N}}. \end{aligned} \tag{4.21}$$

The symmetry condition guarantees that the points $\mathbf{n}_i, \mathbf{p}_i$ belong to the imaged axis of symmetry, i.e. $\mathbf{n}_i, \mathbf{p}_i \in l_{\mathbf{n}_i,\mathbf{p}_i} \equiv l_{sym}^{\mathtt{H}} \equiv \{\mathtt{H}\mathbf{x} \mid \mathbf{x} \in l_{sym}\}$, where l_{sym} is the symmetry axis of the frontal view. Let

$$\mathbf{m}_i = l_{\mathbf{x}_i,\mathbf{x}_{i+N}} \cap l_{\mathbf{n}_i,\mathbf{p}_i}. \tag{4.22}$$

Obviously, $\mathbf{m}_i$ also belongs to the image of the symmetry axis $l_{sym}^{\mathtt{H}}$. Thus as the index i sweeps from 1 to $2N$, we get $2N-1$ collinear points

$$\{\mathbf{m}_1, \mathbf{n}_1, \mathbf{m}_2, \mathbf{n}_2, \dots, \mathbf{m}_{N-1}, \mathbf{n}_{N-1}, \mathbf{m}_N\} \in l_{sym}^{\mathtt{H}}.$$

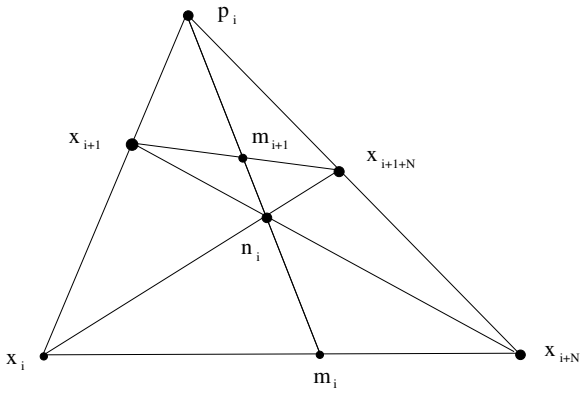


Figure 4.7 Construction of skeleton.

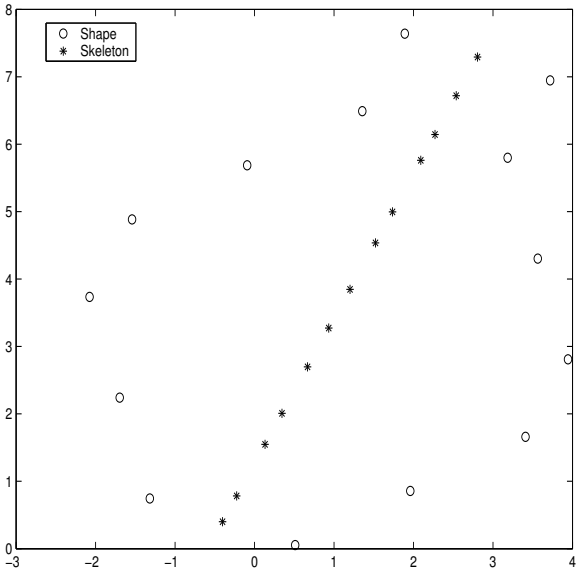


Figure 4.8 Example of skeleton of view.

Relabel these points as $\{\mathbf{s}_1, \ldots, \mathbf{s}_{2N-1}\}$ and call this collection the "skeleton" of the shape $\{\mathbf{x}_i\}_{i=1}^{2N}$. Our general interest is in the shape of our view, by which we will understand the geometry of the image of the original symmetrical shape. We will show that almost all such information is stored in the skeleton that was just defined. That would in turn yield that by studying the skeleton alone we can learn almost everything about the shape itself. This, as will be shown next, greatly simplifies the problem of analysis of planar symmetric shapes, making it more tractable and useful for practical purposes such as object recognition.

4.3.1.1 Skeleton's Signature

As was suggested in the previous sections, the main tool we will use in this chapter to analyze a projective image of a symmetric planar shape is the skeleton of the image. The skeleton is by definition a collection of points that lie on the same line, which is the projective image of the

symmetry axis of the fontal view of the original shape (see Figure 4.8) . The main geometric properties of a skeleton are the number of points and their positions with respect to each other. Because the former is trivially quantifiable, we will draw our attention to the latter. The relative position of the points on a line is completely determined by the corresponding inter-distances between all of them. However, in a situation of an arbitrary view of a given shape, Euclidean distances cannot serve as an an efficient geometrical measure, because they vary from one view to another.

Definition 4.5 Consider four points $\mathbf{s}_1, \ldots, \mathbf{s}_4$ lying on the same line. Their cross-ration is, by definition, the number

$$\sigma(\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4) = \frac{\|\overrightarrow{\mathbf{s}_1\mathbf{s}_3}\|}{\|\overrightarrow{\mathbf{s}_1\mathbf{s}_4}\|} \frac{\|\overrightarrow{\mathbf{s}_2\mathbf{s}_4}\|}{\|\overrightarrow{\mathbf{s}_2\mathbf{s}_3}\|}, \tag{4.23}$$

where by $\|\overrightarrow{\mathbf{s}_i\mathbf{s}_j}\|$ we denote the Euclidean length of the vector connecting the points $\mathbf{s}_i$ and $\mathbf{s}_j$.

The fundamental property of the cross-ratio of 4 collinear points is that it is invariant under any projective transformation ([33]). As a consequence of that, for any projective view of a symmetric planar shape, the cross-ratio of any fixed 4 points from the skeleton remains invariant, i.e. does not depend on a choice of a projective camera `P`. For a shape with sufficient number of points, there exist many ways to choose 4-tuples of skeleton points. We show that by choosing several cross-ratios, we can encode all geometric information about the skeleton.

Definition 4.6 Let $\mathbf{s}_1, \ldots, \mathbf{s}_M$, where $M \equiv 2N - 1 \geq 4$, be N collinear points (a skeleton of an projective image of a symmetric planar shape). We will define its signature as a vector of cross-ratios

$$\sigma \equiv \{\sigma_1, \ldots, \sigma_{M-3}\}, \tag{4.24}$$

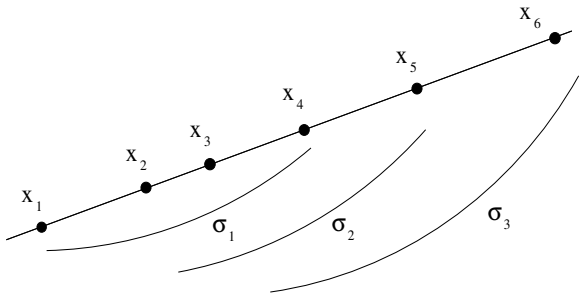


Figure 4.9 Skeleton signature.

where

$$\sigma_i = \sigma(\mathbf{s}_i, \mathbf{s}_{i+1}, \mathbf{s}_{i+2}, \mathbf{s}_{i+3}),\ i = 1, \ldots, M-3. \tag{4.25}$$

Since all individual cross-ratios are projective invariants and therefore do not depend on a choice of a projective camera, neither does the joint vector σ. Thus the signature defined above is a valid quantitative description of the geometry of the skeleton (see Figure 4.9).

The signature itself does not uniquely define a skeleton. It is obvious from the definition of cross-ratio, that any Euclidean transformation of a line which carries the skeleton points has no affect on the signature. Also, certain non-linear transformations have do not affect the signature as well. We can show, however, that the signature defines the skeleton up to a projective transformation, which is the maximum of what we can hope for, in the absence of additional information, such as an explicitly given camera matrix etc.

Suppose, we are given a signature vector σ, which consists of $M-3$ cross-ratios. The goal is to construct all possible skeletons, whose signature coincides with σ. It has already been mentioned that a Euclidean rotation of the line that contains all points in question does not affect the signature. It is therefore no loss of generality if we fix it to coincide with the ordinate axis. It is a simple

exercise to show that we can fix first three points s_1, s_2, s_3 at arbitrary locations, and then the rest can be computed by the following formula.

$$s_{i+3} = \frac{s_{i+2} - s_i}{(1-\sigma_i)\dfrac{s_{i+2} - s_i}{s_{i+1} - s_i} + \sigma_i} + s_i,\ i = 1, \ldots, N-3. \tag{4.26}$$

Thus, for a given signature the space of all skeletons matching that signature has 3 degrees of freedom. This proves that any two different skeletons with identical signatures are projective transforms of one another. Hence a signature indeed defines a skeleton up to a projective transform, and thus carries all the information about the projective geometry of the skeleton.

4.3.2 Reconstructing View of Shape from Skeleton

It was shown in the previous section that for a given planar symmetric shape, to each its projective image there corresponds its unique skeleton and thus a unique signature. This signature does not depend on a choice of a projective camera and thus truly is an attribute of the original shape. It was shown earlier that a signature defines a skeleton up to a projective transformation. We will show how it in fact almost defines a view of the original shape up to a projective transformation, which again will prove that a signature carries almost all projective geometry information about the original shape.

Suppose we are given a signature $\{\sigma_1, \ldots, \sigma_{M-3}\}$. Our goal is to reconstruct a view of an object up to a projective transformation. Towards that goal, we proceed as follows. We start by building up a skeleton that matches the above signature. Fix $\mathbf{s}_1 = (0, 1)$ and $\mathbf{s}_2 = (0, 2)$. Then according to the result in the previous subsection, we can find $\mathbf{s}_3, \ldots, \mathbf{s}_M$, such that the entire skeleton $\{\mathbf{s}_i\}_{i=1}^{M}$ matches the signature $\{\sigma_j\}_{j=1}^{M-3}$.

We now proceed to reconstructing a view of a symmetric planar shape that would match the above skeleton. Recall that for a frontal view of the shape $\{\mathbf{x}_i^0\}_{i=1}^{2N}$, the lines connecting that symmetric points are parallel (in the Euclidean geometry), which yields that in projective geometry they intersect at a vanishing point $\mathbf{x}_v^0 = (1, 0, 0)$. As was noted in the beginning of this chapter, a view being reconstructed is some projective transformation, $\mathtt{H}$, of the frontal view, i.e.

$$\mathbf{x}_i = \mathtt{H}\mathbf{x}_i^0.$$

The point $\mathbf{x}_v^0$ is then mapped to a new "vanishing point" $\mathbf{x}_v$. Since the projective transformation $\mathtt{H}$ can be arbitrary, so is $\mathbf{x}_v$. Denote by l_i^v the line that connects the points $\mathbf{s}_{2i-1}$ and $\mathbf{x}_v$. It follows then that the points $\mathbf{x}_i$ and $\mathbf{x}_{i+N}$ belong to the line l_i^v. Let us fix the point $\mathbf{x}_1 \in l_1^v$. We will show now that the rest of the shape can be reconstructed from the points already given.

To reconstruct the point $\mathbf{x}_{1+N}$, we do the following. Note that the points $\mathbf{x}_{1+N}, \mathbf{s}_1, \mathbf{x}_1, \mathbf{x}_v$ are by construction 4 collinear points. Thus their cross-ratio is a projective invariant and equals that of the corresponding points in the frontal view. The cross-ratio of any points of the form $-x, 0, x, +\infty$ is 2 as follows immediately from the definition of the cross-ratio. Coming back to the original view, we are given the three collinear points, and a known cross-ratio. As was shown previously, this information is sufficient to uniquely reconstruct the fourth point.

Now given the skeleton and the "foundation" of the shape, the rest of the points are reconstructed by inverting the skeleton construction procedure (see Figure 4.10).

It is important to note here that to reconstruct a view of a symmetric planar shape in addition to the skeleton we needed to choose two additional points: the vanishing point $\mathbf{x}_v$, and the exact location of the point $\mathbf{x}_1$. The vanishing point is a property of a camera and not an observed shape,

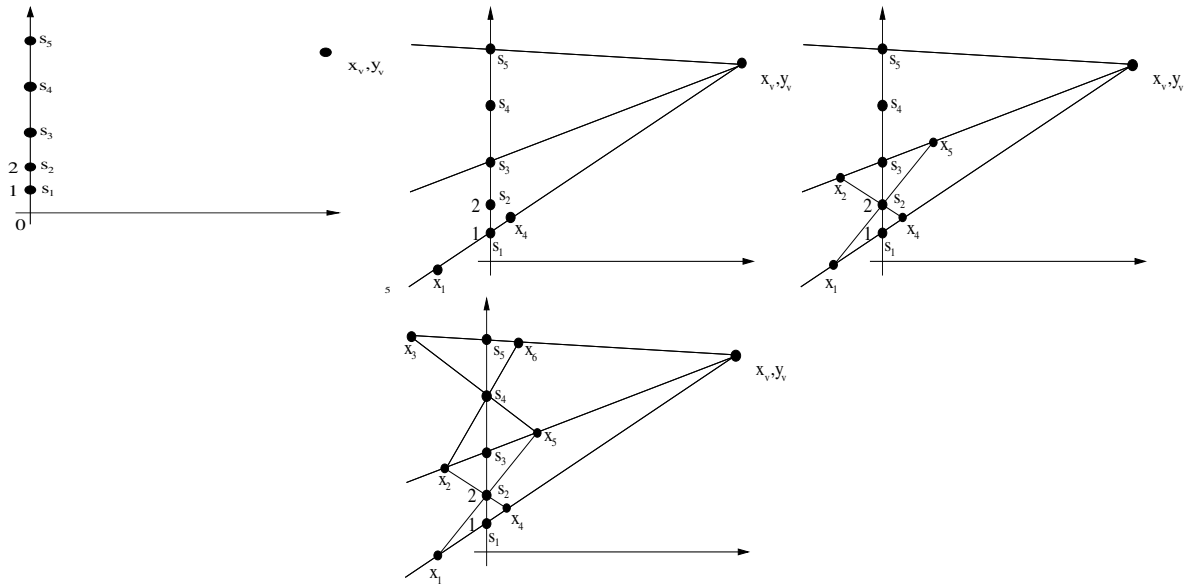


Figure 4.10 Reconstruction of a view from a skeleton.

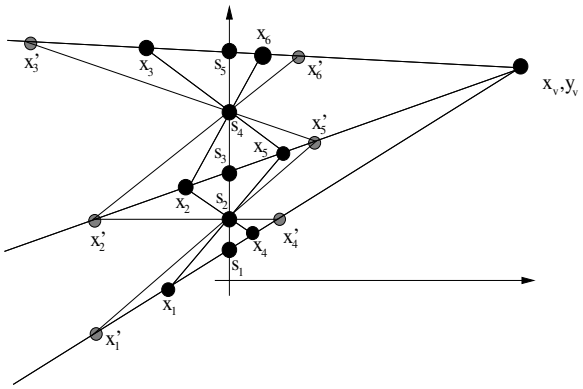


Figure 4.11 Different choice of $\mathbf{x}_1$ affects only the width of the underlying shape.

and thus by varying it we obtain different views of the same shape. However, different choice of $\mathbf{x}_1$ strictly speaking yields views of different shapes. As is easy to see the exact location of $\mathbf{x}_1$ controls the width of the underlying shape, i.e. all different shapes produced by different choices of $\mathbf{x}_1$ are copies of the same shape stretched in the horizontal direction (see Figure 4.11).

4.4 Identification of Shape from Noisy Image

It has been discussed in the previous chapters that the skeleton and its signature can play an important role in identification of a planar symmetric shape from an arbitrary view. More precisely, the following has been established.

- To any projective view of a planar symmetric shape, we can associate its skeleton.
- The skeleton is a collection of points lying on the same line, which is the projective image of the symmetry axis.
- Cross-ratios are effective in capturing the projective geometry of the skeletons and serve as a basis for the notion of signature of skeleton.

- The signature being a property of the original shape of interest and not a particular view can be used for the purposes of object identification.

While the above method to identify discrete shapes is theoretically sound, it suffers from an often unrealistic assumption that the observations that we have are noise-free. Shapes observed in practice are continuous. To obtain discrete representations of continuous contours, a sampling algorithm is used or the points (landmarks) are even selected manually. Either way of converting a continuous shape into a discrete counterpart yields unavoidable errors. A natural question to ask then would be whether the techniques discussed above are suitable to treat shape views in the presence of noise. Unfortunately the answer is no.

For a given noisy shape, it is possible to apply the same geometric construction which would yield the shape's "noisy skeleton points". However, as one can easily convince himself, those points in general are not going to be collinear. Furthermore, since the noise is added to the observation of the shape, not the shape itself, different *noisy* projective views of the same shape are strictly speaking not related through any projective transformation.

These two arguments warrant that an additional effort is to be made in order to be able to apply the previously designed concepts to the noisy environment. Our goal in this chapter therefore will be to study the noisy views of discrete planar symmetric shapes and suggest the estimating techniques, that would be capable of tackling these problems.

By a noisy view of a symmetric planar shape, we will understand the collection of Euclidean points $\{\mathbf{x}_i\}_{i=1}^{2N}$, such that

$$\mathbf{x}_i = (x_i, y_i),\ i = 1, \ldots, 2N,$$

and

$$(x_i, y_i) = \left(x_i^0, y_i^0\right) + (\Delta x_i, \Delta y_i),\ i = 1, \ldots, 2N, \tag{4.27}$$

where we will assume that $\Delta x_i, \Delta y_j$ are independent and identically distributed Gaussian random variables $\mathcal{N}(0, \varepsilon^2)$. The latter is a standard and well-accepted model for the noise. The points $\{(x_i^0, y_i^0)\}_{i=1}^{2N}$ in turn are a noise free observation of a shape through some arbitrary projective camera $\mathtt{P}$[1].

4.4.1 Skeleton of Noisy Shape

Observe that whether or not the observations contain noise, the same procedure of constructing a skeleton can be applied to yield $2N-1$ points. For a noisy view $\{(x_i, y_i)\}_{i=1}^{2N}$ there is certainly no guarantee that those points would lie on the same line. However, each individual point of a "noisy skeleton" $\{\mathbf{s}_i\}_{i=1}^{2N-1}$ will be an approximation of the corresponding point of the skeleton $\{\mathbf{s}_i^0\}_{i=1}^{2N-1}$ of the underlying clean view $\{(x_i^0, y_i^0)\}_{i=1}^{2N}$.

Our goal is to study the distribution of a noisy skeleton and then derive a technique for estimating the "clean skeleton". Let us fix a "clean view" $\{\mathbf{x}_i^0\}_{i=1}^{2N}$ of a symmetric planar shape and also assume that we know the projective camera $\mathtt{P}$ that was used to produce the view. Let us also define the "noisy view" according to (4.27). Denote the clean skeleton by $\{\mathbf{s}_i^0\}$ and the noisy skeleton by $\{\mathbf{s}_i\}$. Then the vector $(\mathbf{s}_1, \ldots, \mathbf{s}_{2N-1})^T$ is Gaussian $\mathcal{N}(\mathbf{s}^0, \Sigma)$, where

$$\mathbb{E}\mathbf{s}_i = \mathbf{s}_i^0, \tag{4.28}$$

and the covariance matrix Σ depends on a projection camera matrix $\mathtt{P}$ as well as on the shape itself.

[1] In this section of the chapter the superscript 0 has no relevance to the frontal view.

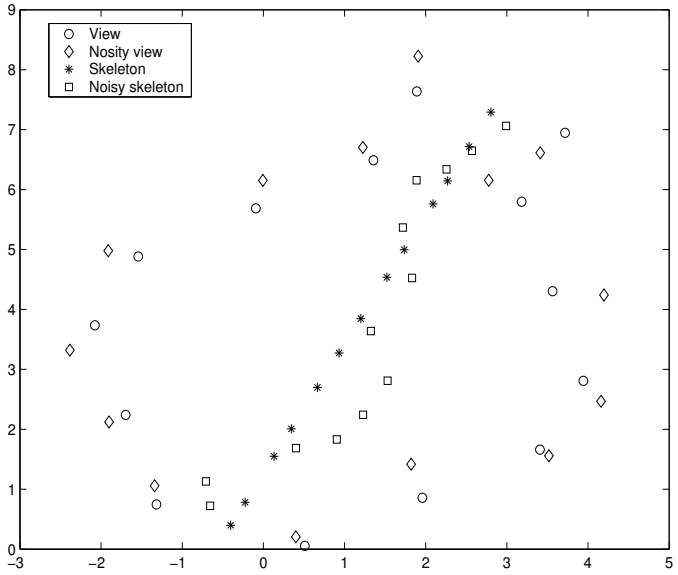


Figure 4.12 Noise-free and noisy skeletons.

While the formal complete proof of this fact remains to be an open problem because of non-trivial relationships tying together the shape points and the points of a skeleton, its validity can be demonstrated experimentally (see Figure 4.13).

4.4.1.1 Covariance Matrix of Skeleton Points

It has been established experimentally that for a given noisy view of a symmetric planar shape, its skeleton points are $2-D$ points with Gaussian distribution, whose means coincide with noise-free skeleton-points. While we offer no direct formula to compute a covariance matrix Σ, the following suggests a way to compute it numerically in terms of a function that does not depend on a particular choice of shape and thus can be a priori tabulated.

Suppose, that

$$\mathbf{s}_i = (s_i, t_i),\ \mathbb{E}s_i = s_i^0,\ \mathbb{E}t_i = t_i^0,\ i = 1, \ldots, 2N-1 \tag{4.29}$$

are the skeleton points. Define the vector

$$\tilde{\mathbf{s}} = (s_1, \ldots s_{2N-1}, t_1, \ldots, t_{2N-1})^T \tag{4.30}$$

The covariance matrix of interest then is

$$\Sigma = (\Sigma_{ij})_{i,j=1}^{4N-2},\ \Sigma_{ij} = cov(\tilde{\mathbf{s}}_i, \tilde{\mathbf{s}}_j). \tag{4.31}$$

The approximation of Σ certainly could be obtained via simulating a sufficiently large amount of data noise samples. Note, however, that separate simulations would be required for each view of each shape, which is costly computationally. Instead we propose the following way to compute the matrix Σ.

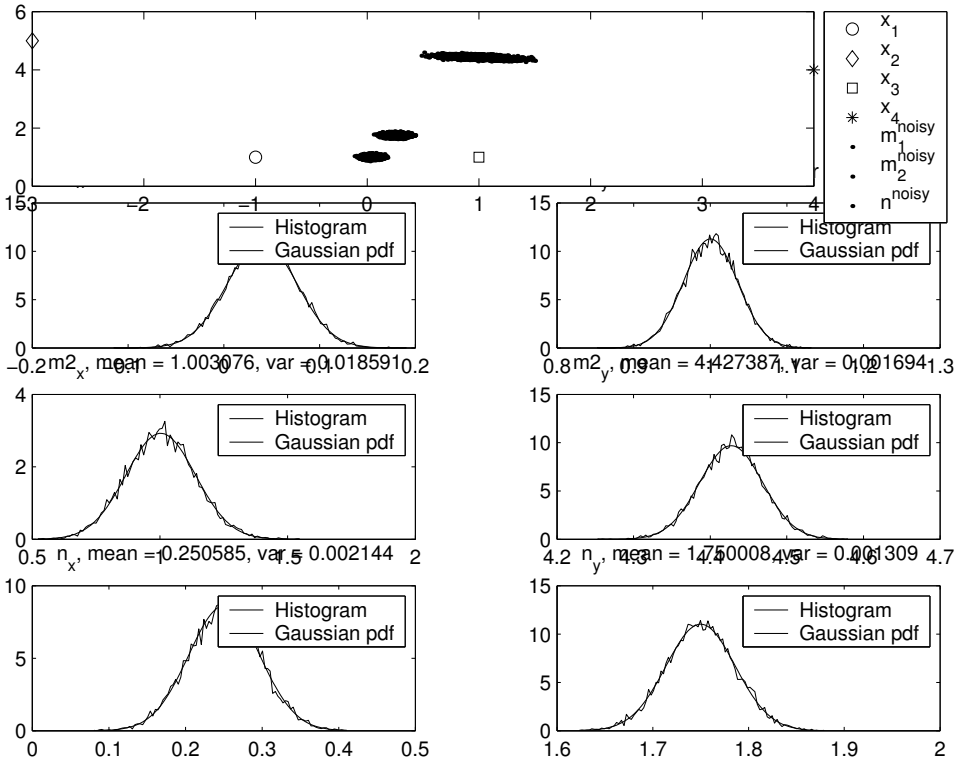


Figure 4.13 Realizations of noisy skeletons.

First, note that each skeleton point as defined previously is a local feature of a shape. It becomes apparent from the skeleton construction that not all skeleton points are correlated. More precisely,

$$\mathbb{E}[\Delta s_i \Delta s_{i+j}] = \mathbb{E}[\Delta s_i \Delta t_{i+j}] = \mathbb{E}[\Delta s_{i+j} \Delta t_i] = \mathbb{E}[\Delta t_i \Delta t_{i+j}] = 0, \tag{4.32}$$

where $j > 4, i = 2k - 1$, and

$$\mathbb{E}[\Delta s_i \Delta s_{i+j}] = \mathbb{E}[\Delta s_i \Delta t_{i+j}] = \mathbb{E}[\Delta s_{i+j} \Delta t_i] = \mathbb{E}[\Delta t_i \Delta t_{i+j}] = 0, \tag{4.33}$$

where $j > 3, i = 2k$ and

$$\Delta s_i = s_i - s_i^0,\ \Delta t_i = t_i - t_i^0. \tag{4.34}$$

This follows from the observation that only those skeleton points are correlated, which are constructed from the same data points. In particular, the values of $\Sigma_{1,i}, \Sigma_{2,i}, \Sigma_{2N,i}, \Sigma_{2N+1,i}$ can be computed if we know the locations of data points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_{1+N}, \mathbf{x}_{2+N}, \mathbf{x}_{3+N}$. Likewise, to compute the covariance coefficients corresponding to any two adjacent skeleton points $\mathbf{s}_{2k-1}, \mathbf{s}_{2k}$ all that is required to be known is the 6 data points whose noise affect the distribution of the said piece of the skeleton.

Define the function $\Lambda(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6)$, such that

$$\Lambda(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6) = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \lambda_{1,3} & \lambda_{1,4} \\ \underbrace{\lambda_{2,1}}_{=\lambda_{1,2}} & \lambda_{2,2} & \lambda_{2,3} & \underbrace{\lambda_{2,4}}_{=0} \end{pmatrix}, \tag{4.35}$$

where the function value is a matrix that consists of correlation coefficients of the skeleton points produced by the given data points. Then this function can be tabulated on a grid a priori, thus allowing us to inductively compute the full covariance matrix of the skeleton of any given view of a shape.

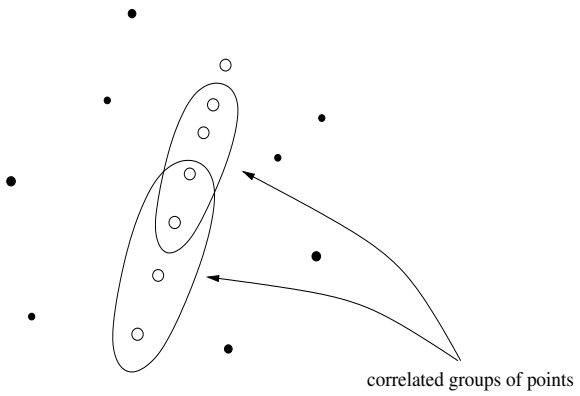


Figure 4.14 Groups of correlated points in a skeleton.

To reduce the amount of data that needs to be stored in order to compute Σ we note that actually it is not necessary to store the covariance coefficients of the skeleton points produced by *any* 6 data points, but only those that are defined up to rotation, scale and translation (see Figure 4.14). Indeed the translation of a shape does not change the covariance matrix, scaling of the shape results in the corresponding scaling of Λ, and finally if

$$\begin{pmatrix} s_i^2 \\ t_i^2 \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \cdot \begin{pmatrix} s_i^1 \\ t_i^1 \end{pmatrix}, \; i = 1, \ldots, N, \tag{4.36}$$

then

$$\begin{aligned}
\mathbb{E}\left[\Delta s_i^2 \Delta s_j^2\right] &= \cos^2\phi\, \mathbb{E}\left[\Delta s_i^1 \Delta s_j^1\right] + \sin\phi\cos\phi\, \left(\mathbb{E}\left[\Delta s_i^1 \Delta t_j^1\right]\right. \\
&\left. + \mathbb{E}\left[\Delta s_j^1 \Delta t_i^1\right]\right) + \sin^2\phi\, \mathbb{E}\left[\Delta t_i^1 \Delta t_j^1\right] \\
\mathbb{E}\left[\Delta s_i^2 \Delta t_j^2\right] &= \cos^2\phi\, \mathbb{E}\left[\Delta s_i^1 \Delta t_j^1\right] + \sin\phi\cos\phi\, \left(\mathbb{E}\left[\Delta t_i^1 \Delta t_j^1\right]\right. \\
&\left. - \mathbb{E}\left[\Delta s_i^1 \Delta s_j^1\right]\right) - \sin^2\phi\, \mathbb{E}\left[\Delta t_i^1 \Delta s_j^1\right] \\
\mathbb{E}\left[\Delta t_i^2 \Delta t_j^2\right] &= \cos^2\phi\, \mathbb{E}\left[\Delta t_i^1 \Delta t_j^1\right] - \sin\phi\cos\phi\, \left(\mathbb{E}\left[\Delta s_i^1 \Delta t_j^1\right]\right. \\
&\left. + \mathbb{E}\left[\Delta s_j^1 \Delta t_i^1\right]\right) + \sin^2\phi\, \mathbb{E}\left[\Delta s_i^1 \Delta s_j^1\right]
\end{aligned}$$

It follows then that the covariance matrix Σ is completely defined by the function of only 4 data points

$$\tilde{\Lambda}(\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_5, \mathbf{x}_6) = \Lambda\left(\mathbf{p}', \mathbf{x}_2, \mathbf{x}_3, \mathbf{p}'', \mathbf{x}_5, \mathbf{x}_6\right), \tag{4.37}$$

where

$$\mathbf{p}' = (-1, 1)^T, \; \mathbf{p}'' = (1, 1). \tag{4.38}$$

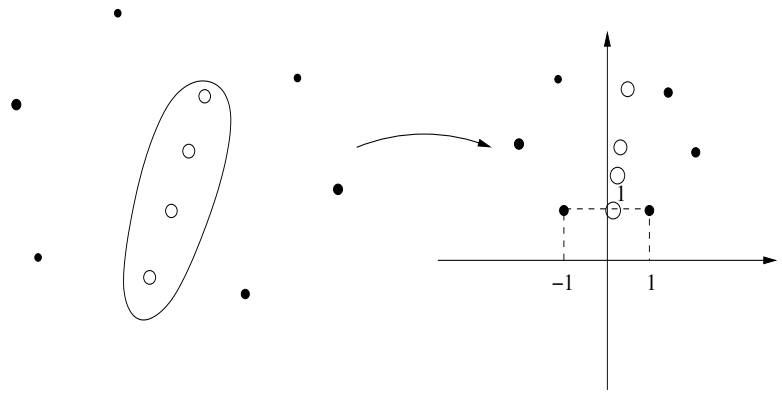


Figure 4.15 Skeleton points correlation coefficients are functions of only 4 shape points, not the entire shape.

4.4.2 Detection of Shape by Best Linear Fit

4.4.2.1 Detection of Shape from Known View

We know now that a noisy skeleton is a collection of correlated $2-D$ Gaussian random variables, whose means are collinear, i.e. lie on the same line, which is unknown. The problem of the best linear fit on the plane in the presence of correlated Gaussian noise is equivalent to the problem of non-linear fit in the presence of independent and identically distributed Gaussian noise. More precisely, we proceed as follows.

Suppose for now that the camera matrix P is known, and the problem is that of deciding which of the two given shapes yielded the noisy skeleton in question. We have the following hypothesis testing problem

$$\begin{aligned} H_0: &\quad (s_i,t_i) = (s_i^0,t_i^0) + (\Delta s_i^0, \Delta t_i^0),\ i=1,\ldots,2N-1, \\ H_1: &\quad (s_i,t_i) = (s_i^1,t_i^1) + (\Delta s_i^1, \Delta t_i^1),\ i=1,\ldots,2N-1. \end{aligned} \tag{4.39}$$

$\mathbf{s}_i = (s_i,t_i)\,,\ i=1,\ldots,2N-1$ are the observed skeleton points, and $\mathbf{s}_i^j = \left(s_i^j,t_i^j\right)\,,\ j=0,1$ are the points of the clean skeleton corresponding to each shape. The noise vector $\left(\ldots,\Delta s_i^j,\ldots,\Delta t_j^j,\ldots\right)$ is then Gaussian with zero mean and some known covariance matrix Σ^j. Both the means and the covariance matrix depend on an underlying shape, and in addition to that the covariance matrices Σ^j depend on a projective camera matrix P, which produced the shape view in study.

Let us introduce the following notation

$$\tilde{\mathbf{s}} = \begin{pmatrix} s_1 \\ \vdots \\ s_{2N-1} \\ t_1 \\ \vdots \\ t_{2N-1} \end{pmatrix}, \ \tilde{\mathbf{s}}^j = \begin{pmatrix} s_1^j \\ \vdots \\ s_{2N-1}^j \\ t_1^j \\ \vdots \\ t_{2N-1}^j \end{pmatrix}, \ j = 0, 1. \tag{4.40}$$

Then we have

$$\begin{aligned} H_0: \ \ & \mathbb{E}\left[\tilde{\mathbf{s}}\right] = \tilde{\mathbf{s}}^0, \ \mathbb{E}\left[\left(\tilde{\mathbf{s}} - \tilde{\mathbf{s}}^0\right)\left(\tilde{\mathbf{s}} - \tilde{\mathbf{s}}^0\right)^T\right] = \Sigma^0, \\ H_1: \ \ & \mathbb{E}\left[\tilde{\mathbf{s}}\right] = \tilde{\mathbf{s}}^1, \ \mathbb{E}\left[\left(\tilde{\mathbf{s}} - \tilde{\mathbf{s}}^1\right)\left(\tilde{\mathbf{s}} - \tilde{\mathbf{s}}^1\right)^T\right] = \Sigma^1. \end{aligned} \tag{4.41}$$

Let $A^j, \ j = 0, 1$ be whitening matrices, that is such that

$$\left(A^j\right)^T A^j = \Sigma^j, \ A^j \sim (4N-2) \times (4N-2). \tag{4.42}$$

Define the "whitened data" for each hypothesis

$$\bar{\mathbf{s}}_n^j = \left(\left(A^j\right)^T\right)^{-1} \tilde{\mathbf{s}}, \ \bar{\mathbf{s}}^j = \left(\left(A^j\right)^T\right)^{-1} \tilde{\mathbf{s}}^j. \tag{4.43}$$

We then deduce that $\bar{\mathbf{s}}_n^j$ is a Gaussian vector with

$$\mathbb{E}\left[\bar{\mathbf{s}}_n^j\right] = \bar{\mathbf{s}}^j. \tag{4.44}$$

and

$$\mathbb{E}\left[\left(\bar{\mathbf{s}}_n^j - \bar{\mathbf{s}}^j\right)\left(\bar{\mathbf{s}}_n^j - \bar{\mathbf{s}}^j\right)^T\right] = I. \tag{4.45}$$

Define the costs

$$C^j \equiv C\left(\tilde{\mathbf{s}}^j\right) = \left\|\bar{\mathbf{s}}_n^j - \bar{\mathbf{s}}^j\right\|^2. \tag{4.46}$$

The decision rule of choosing a hypothesis then becomes

$$C^0 \gtrless_{H_0}^{H_1} C^1. \tag{4.47}$$

The above rule is trivially generalized for the case when we have a library of N shapes, so that the detector is of the form

$$i_{opt} = \arg \min_{i=1,\ldots,N} C^i. \tag{4.48}$$

4.4.2.2 Estimation of a Shape from Known View using Skeleton

In many applications it is often the case that the set of possible shapes that may be observed is either undetermined or so large that running each individual shape against the matching function is impossible or impractical. The problem then becomes not of detecting but rather of estimating the shape whose noisy view we observe. Since by a shape we normally understand its frontal view, we formulate the problem as follows.

Let $\{\mathbf{x}_i\}$ be noisy data points defined according to Equation (4.27). Assume that the projective camera matrix $\mathtt{P}$ is given, so that the camera rotation parameters ϕ, ψ, α are known. The problem is then to estimate the frontal view $\left\{\mathbf{x}_i^f\right\}$ of a shape that matches the data.

Let $\{\mathbf{s}_i\}$ be a noisy skeleton computed form the data points $\{\mathbf{x}_i\}$. We propose an optimization technique, whose output is the estimate of the skeleton of the true shape corresponding to the observed data.

Denote by $l_{\phi,\psi,\alpha}$ the image of the the line $\{x = 0\}$ through the camera $\mathtt{P}(\phi, \psi, \alpha)$. Note that the skeleton points of the true shape must lie on $l_{\phi,\psi,\alpha}$. Define by the initial guess $\{\mathbf{s}_i^0\}$ the skeleton, obtained by projecting the noisy skeleton points on that line. The matching function (cost) of $\{\mathbf{s}_i^0\}$,

$C\left(\tilde{\mathbf{s}}^{0,f};\phi,\psi,\alpha\right)$, is defined very much analogously to the case of detection, except that now we have to take into account the fact that a skeleton of a frontal view defines a shape only up to the horizontal size.

Define by $\left\{\mathbf{s}_i^{0,f}=\left(s_i^{0,f},t_i^{0,f}\right)\right\}$ the skeleton of the frontal view corresponding to $\{\mathbf{s}_i^0\}$, and $\tilde{\mathbf{s}}^{0,f}=\left(s_1^{0,f},\ldots,s_{2N-1}^{0,f},t_1^{0,f},\ldots,t_{2N-1}^{0,f}\right)^T$. As was shown in the previous sections, any choice of $\mathbf{x}_1^0$ along with the skeleton $\{\mathbf{s}_i^{0,f}\}$, yields a reconstruction of the frontal view $\{\mathbf{x}_i^{0,f}\}$. If this reconstruction is a true shape then it follows that the skeleton of its rotated view should have the distribution $\mathscr{N}\left(\{\mathbf{s}_i^0\},\Sigma_{\{\mathbf{s}_i^0\},\mathbf{x}_1,\phi,\psi,\alpha}\right)$, which in turn allows us to define the matching cost

$$\mathscr{C}_{\mathbf{x}_1}(\tilde{\mathbf{s}},\phi,\psi,\alpha)=\left\|\bar{\mathbf{s}}-\bar{\mathbf{s}}^0\right\|^2, \tag{4.49}$$

where $\bar{\mathbf{s}}$ and $\bar{\mathbf{s}}^0$ are "whitened" vectors $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{s}}^0$. Minimizing over all possible $\mathbf{x}_1$, which do not depend on a skeleton, we get

$$C\left(\tilde{\mathbf{s}}^0;\phi,\psi,\alpha\right)=\min_{\mathbf{x}_1}\mathscr{C}_{\mathbf{x}_1}(\tilde{\mathbf{s}},\phi,\psi,\alpha). \tag{4.50}$$

The optimal skeleton is then found by minimizing the cost defined above, i.e. by numerically solving the problem:

$$\begin{cases} \left\{\mathbf{s}_i^{opt}\right\}=\arg\min\limits_{\{\mathbf{s}_i\}}C\left(\tilde{\mathbf{s}};\phi,\psi,\alpha\right) \\ \text{subject to:}\quad \{\mathbf{s}_i\}\in l_{\phi,\psi,\alpha},\ \forall i=1,\ldots,2N-1. \end{cases} \tag{4.51}$$

Figure 4.17 and Figure 4.18 demonstrate two example of estimating a shape through its skeleton from its single noisy observation.

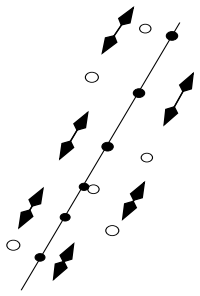


Figure 4.16 Finding the optimal skeleton.

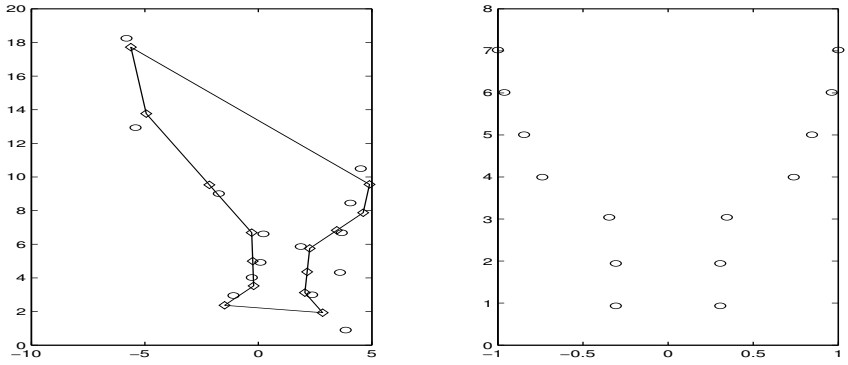


Figure 4.17 Shape of a wineglass is reconstructed from a noisy observation with known camera parameters.

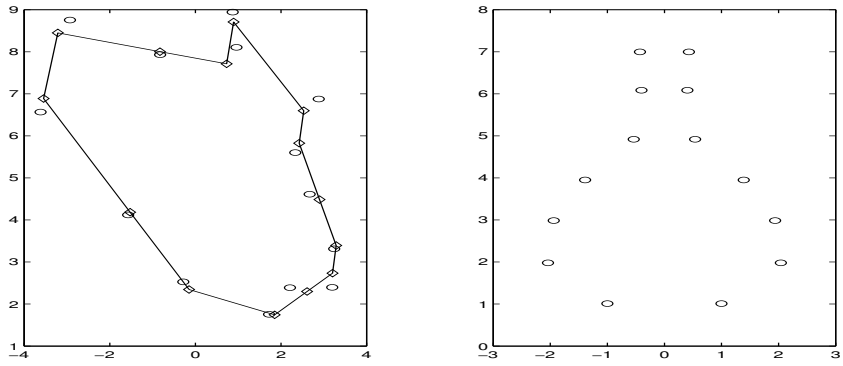


Figure 4.18 Shape of a jar is reconstructed from a noisy observation with known camera parameters.

4.4.2.3 Identification of Shape from Unknown View

The core of the skeleton detection algorithm is the generalized best linear fit in the presence of correlated $2-D$ Gaussian noise. It was pointed out that the correlation matrix of the vector representing the skeleton of a noisy skeleton $(\mathbf{s}_1, \ldots, \mathbf{s}_{2N-1})$ depends not only on a shape but also on a projective camera matrix P, that was used to produce that particular view of the shape, i.e. different camera matrices P yield different correlation matrices Σ.

Recall that a camera matrix depends on two groups of parameters: extrinsic and intrinsic. The intrinsic parameters model the inner structure of the camera, whereas the extrinsic parameters govern the location of the camera in the $3-D$ space with respect to the shape of interest.

From now on we will assume that all the intrinsic parameters of the camera are known. In practice that corresponds to the situation when the knowledge about what type of camera was used to produce the image is available as is often the case.

The extrinsic parameters that we are interested in are the rotation angles of the shape with respect to the camera along with the distance from the camera center to the object. The distance from the camera to the object in a practical situation when a clear view of the object available is relatively easy to estimate with modern equipment. Also, while strictly speaking by varying this distance we can change the projective geometry of a view, its affect in most situations is quite insignificant and thus can be neglected. Therefore we will assume that out of all camera parameters only the three rotation parameters ϕ, ψ and α are unknown.

Consider a noisy view $\{\mathbf{x}_i\}$ of a symmetric planar shape, where we assume that the noise is Gaussian with zero-mean and some variance δ^2. It was then established earlier that the conditional

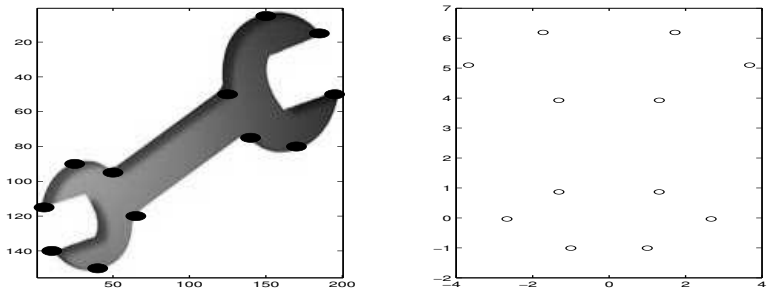


Figure 4.19 Estimating the skewed view of discrete shape corresponding to the wrench.

distribution of its skeleton given the camera matrix, i.e. the rotation parameters is Gaussian

$$p\left(\mathbf{s}_1,\ldots,\mathbf{s}_{2N-1}\right) \sim \mathscr{N}\left(\mathbf{s}_1^0,\ldots,\mathbf{s}_{2N-1}^0;\Sigma_{\mathbf{s};\phi,\psi,\alpha,\delta^2}\right). \tag{4.52}$$

Suppose that we know that the rotation parameters belong to some intervals:

$$\phi \in (\phi_{\min},\phi_{\max}),\ \psi \in (\psi_{\min},\psi_{\max}),\ \alpha \in (\alpha_{\min},\alpha_{\max}). \tag{4.53}$$

Then the shape estimation is obtained by optimizing the cost derived for a known view over all rotation angles from the admissible set defined by (4.53), i.e.

$$\{s_i\} = \arg \min_{\{\mathbf{s}_i\},\phi,\psi,\alpha} C(\tilde{\mathbf{s}},\phi,\psi,\alpha); \tag{4.54}$$

The examples on Figures 4.19 and 4.20 demonstrate reasonable performance for realistic shapes with the landmarks chosen manually with visible errors.

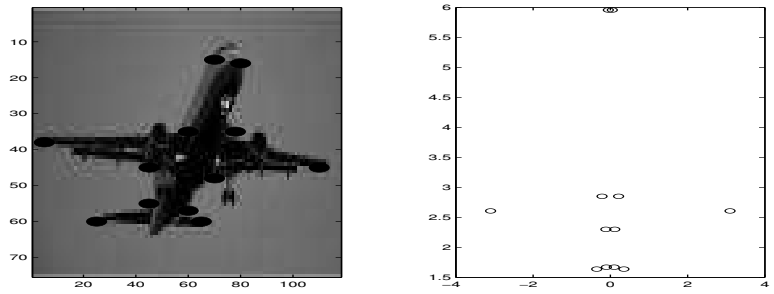


Figure 4.20 Estimating the skewed view of discrete shape corresponding to the airplane.

Chapter 5

Conclusions

In this chapter we summarize the contributions of the thesis and suggest future research avenues may follow in light of the results developed here.

5.1 Summary of Contributions

5.1.1 Sampling of Planar Curves

In this part of the thesis we proposed a new approach to sampling planar curves as well as $3-D$ surfaces. The essence of the proposed method is that in order to make use of the existing theory of sampling $1-D$ signals and $2-D$ images, we need to fix a parameterization or a coordinate system. Furthermore, it may be sufficient to know the family of coordinate systems.

A sampling algorithm is then obtained by choosing samples in such a way that they contain the information about the signal in a given coordinate system, as well as about the coordinates themselves. The reconstruction algorithm then consists of two steps:

1. Reconstruct a coordinate system;

2. Given the samples and the coordinate system, reconstruct a function.

We showed that it is possible to apply this method to polar coordinates and describe the class of curves, for which the resulting algorithm would yield perfect sampling and reconstruction.

Finally, we showed that the above sampling algorithm is a particular class of the multi-cost optimization technique, which is defined for a broad class of curves and yields good approximations even for curves which lie outside of the class, for which the sampling algorithm was derived .

5.1.2 Implementation of a Simulated-Annealing Based Optimization

Here we introduced a novel simulated-annealing based algorithm, which can be used to numerically optimize a functional with a good convergence rate. Optimization is in general performed as an evolution in the form of a stochastic differential equation. Our new algorithm is based on the notion of scale, by which we understand the unit time interval, which is used to convert a continuous diffusion into its discrete version.

We showed that different scales were optimal for different functionals, and in the absence of any knowledge about the latter, it is the best solution to instead vary and search for the best scale. Our strategy is to begin the evolution as a completely deterministic one, and then once a local minimum is found, to progressively increase the scale until there is a chance of finding a better energy state.

We showed the proposed technique to lead to superior performance, both in terms of the final energy level that can be achieved, as well as the time it takes for it to happen.

5.1.3 Identification of a Planar Discrete Symmetric Shape from a Single View

We considered a problem of identifying a discrete symmetric planar shape from its single view. We proposed the notion of a skeleton of a symmetric planar shape and made it the core of the recognition algorithm.

We showed that the skeleton was capable of capturing virtually all geometric information about the original object that was at all possible to extract. At the same time by introducing the signature of a skeleton, we obtained a handle that was unique to an object and thus was a perfect basis for a recognition algorithm.

We further considered the behavior of the skeleton in the presence of white Gaussian noise in the data points. We showed that such noise resulted in a correlated Gaussian noise in the skeleton points. A natural way to estimate the true skeleton therefore turned out to be generalized best linear fit.

5.2 Future Work

In this section we outline some of the work that would be natural new avenues in light of what has been done in the thesis.

- **Using spaces other than the one generated by sinc-functions for sampling curves.** The choice of $sinc$-functions is the most natural one in any sampling problem. It, however, comes at a significant cost. Due to their heavy-tails $sinc$-functions make any sampling algorithm very sensitive to noise. Significant errors at one point propagate far away from it, and corrupt the entire signal rather than just the local neighborhood.

One could potentially avoid this problem by considering other, more local orthogonal bases, such as wavelet.

- **Using coordinate systems other than simple polar coordinates.** Many shapes, such as airplanes and other natural and man-made objects cannot be well-represented in any polar coordinates. It would therefore be impossible to directly apply sampling techniques of this thesis to such objects. The proposed approach, however, is fairly general and could be taken in conjunction with other coordinate systems, perhaps application-driven, e.g. based on the mean-shape of objects of interest etc.

- **Sampling Skewed Views of Symmetric Shapes.** The purpose of the sampling algorithm is to convert a continuous shape into a discrete one, so that an identification algorithm such as the one presented in this thesis could be applied. There still remains a gap between the two steps due to the fact that the sampling algorithm proposed above does not guarantee the the resulting samples would correspond to symmetric points. To enforce that, one could add an additional symmetry constraint to the multi-cost optimization problem.

- **Theoretical analysis of the proposed simulated annealing based optimization technique.** While the proposed technique provides promising practical results, its theoretical analysis remains elusive. It is clear the the evolution is inherently non-Markovian and thus cannot be described in terms of a single stochastic differential equation. Perhaps some modification of it based on the same idea of scale could be more advantageous with regards to theoretical tractability.

- **Statistical analysis of the noisy skeleton's signature.** It was shown that a skeleton of a noisy shape consists of $2 - D$ Gaussian points, and the optimal estimation procedure was based on the generalized best linear fit. It would be of interest to study the distribution of a signature of noisy shape and derive an alternative estimation procedure based on that distribution.

LIST OF REFERENCES

[1] Akram Aldroubi and Karlheinz Gröchenig. Non-uniform sampling and reconstruction in shift-invariant spaces. *To appear in SIAM Reviews*, pages 1–47, 2001.

[2] Akram Aldroubi, Qiyu Sun, and Wai-Shing Tang. p-frames and shift invariant subspaces of l^p. *Journal of Fourier Analysis and Applications*, 7(1):1–22, 2001.

[3] A. T. Fomenko B.A. Dubrovin, S. P. Novikov and R. G. Burns. *Modern Geometry-Methods and Applications: Part Ii, the Geometry and Topology of Manifolds*, volume 104 of *Graduate Texts in Mathematics*. Springer Verlag, August 1995.

[4] John Benedetto and Paulo J.S.G. Ferreira. Introduction. [5], pages 1–28.

[5] John J. Benedetto and Paulo J. S. G. Ferreira, editors. *Modern Sampling Theory: Mathematics and Applications*. Applied and Numerical Harmonic Analysis. Birkhauser Boston, 2001.

[6] J. Brian Burns, Richard S. Weiss, and Edward M. Riseman. *The Non-Existence of General-Case View Invariants*, chapter 6. In Mundy and Zisserman [37], 1992.

[7] Stefan Carlsson and Daphna Weinshall. Dual computations of projective shape and camera positions from multiple images. *International Journal of Computer Vision*, 27(3):227–241, 1998.

[8] S. Chen and B. L. Luk. Adaptive simulated annealing for optimization in signal processing applications. *Signal Processing*, 79:117–128, 1999.

[9] Christopher Coelho, Aaron Heller, Joseph L. Mundy, David A. Forsyth, and Andrew Zisserman. *An Experimental Evaluation of Projective Invariants*, chapter 4. In Mundy and Zisserman [37], 1992.

[10] Xavier Descombes and Elena Zhizhina. Langevin algorithm in image restoration problems. to appear, 2002.

[11] Ian L. Dryden and Kanti V. Mardia. *Statistical Shape Analysis*. John Wiley & Son Ltd, 1998.

[12] Manuel Duque-Anton. Constructing efficient simulated annealing algorithms. *Discrete Applied Mathematics*, 77:139–159, 1996.

[13] Matthias Ehrgott. *Multicriteria Optimization*. Springer, 2000.

[14] Olivier Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, 1993.

[15] Olivier Faugeras, Quang-Tuan Luong, and Theo Papadopoulo. *The Geometry of Multiple Images: The Laws That Govern The Formation of Multiple Images of a Scene And Some of Their Applications*. MIT Press, 2001.

[16] Jörg Fliege and Benar Fux Svaiter. Steepest descent method for multicriteria optimization. *Mathematical Methods of Operations Research*, 51:479–494, 2000.

[17] Colin R. Goodall and Kanti V. Mardia. Projective shape analysis. *Journal of Computational and Graphical Statistics*, 1999.

[18] Richard I. Hartley. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(4):580–593, June 1997.

[19] David W. Henderson. *Differential Geometry: A Geometric Inroduction*. Prentice Hall, Inc, 1998.

[20] John E. Hopcroft, Daniel P. Huttenlocher, and Peter C. Wayner. *Affine Invariants for Model-Based Recognition*, chapter 18. In Mundy and Zisserman [37], 1992.

[21] Lester Ingber. Very fast simulated reannealing. *Journal of Mathematical Computational Modelling*, 12:967–973, 1989.

[22] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, second edition, 1991.

[23] D.G. Kendall. A survey of statistical theory of shape. *Statistical Science*, 4:87–120, 1989.

[24] D.G. Kendall, D. Barden, T.K. Carne, and H. Le. *Shape and Shape Theory*. John Wiley & Sons, Ltd, 1999.

[25] Eric Klassen and Anuj Srivastava. Analysis of planar shapes using geodesic lengths on a shape manifold. *submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence*, June 2002.

[26] Peter Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equation*. Applications of Mathematics, Stochastic Modelling and Applied Probability. Springer Verlag, third edition, January 1999.

[27] Peter E. Kloeden, Eckhard Platen, and Henri Schurz. *Numerical Solution of SDE Through Computer Experiments*. Springer, 1994.

[28] Vladimir A. Kotelnikov. On the transmission capacity of the ether and wire in electrocommunications. [5], pages 29–49.

[29] Jan Kybic, Thierry Blu, and Michael Unser. Generalized sampling: A variational approach. In *Proceedings of the 2001 Workshop on Sampling Theory and Applications, SampTA'01*, 2001.

[30] Thomas Liggett. *Interacting Particle Systems*. Springer Verlag, April 1985.

[31] Dinh The Luc. *Theory of Vector Optimization, Lecture Notes in Econ. Math. Systems, 319*. Springer-Verlag, 1989.

[32] Stéfane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press Ltd., 1998.

[33] Stephen J. Maybank. *Classification Based on the Cross Ratio*. Volume 825 of Mundy et al. [39], 1994. Second Joint European-US Workshop Ponta Delgada, Azores, Portugal, October 1993.

[34] Stephen J. Maybank. Stochastic properties of the cross ratio. *Pattern Recognition Letters*, 17(3):211–217, 1996.

[35] Stephen J. Maybank. Relation between 3d invariants and 2d invariants. *Image and Vision Computing*, 16(1):13–20, January 1998.

[36] Günter Meinardus and Guido Waltz. Best approximation by free knot splines. *BIT*, 41(1):158–178, 2001.

[37] Joseph L. Mundy and Andrew Zisserman, editors. *Geometric Invariance in Computer Vision*. The MIT Press, 1992.

[38] Joseph L. Mundy and Andrew Zisserman. *Introduction - Towards a New Framework for Vision*, pages 1–39. In [37], 1992.

[39] Joseph L. Mundy, Andrew Zisserman, and David Forsyth, editors. *Applications of Invariance in Computer Vision*, volume 825 of *Lecture Notes in Computer Science*. Springer-Verlag, 1994. Second Joint European-US Workshop Ponta Delgada, Azores, Portugal, October 1993.

[40] Bernt Øksendal. *Stochastic Differential Equations: Introduction with Applications*. Springer, fifth edition, 1998.

[41] Oleg V. Poliannikov and Hamid Krim. On sampling curves. In *Proceedings of the* 35^{th} *Asilomar Conference on Signals, Systems and Computers*, 2001.

[42] Oleg V. Poliannikov and Hamid Krim. Parameterization models for sampling planar curves. In *Proceedings of SPIE vol. 4667, SPIE Conference on Image Processing: Algorithms and Systems*, 2002.

[43] Oleg V. Poliannikov and Hamid Krim. Probabilistic analysis of noisy views of symmetric shapes. In *Proceedings of the* 36^{th} *Asilomar Conference on Signals, Systems and Computers*, 2002.

[44] Oleg V. Poliannikov and Hamid Krim. On sampling closed planar curves and surfaces. *The Journal of Sampling Theory in Image and Signal Processing (to appear)*, 2003.

[45] John G. Proakis and Dimitris G. Manolakis. *Digital Signal Processing: Principles, Algorithms, and Applications*. Prentice Hall, Inc., 1996.

[46] Long Quan. Two-way ambiguity in 2d projective reconstruction from three uncalibrated 1d images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):212–216, 2001.

[47] C. A. Rothwell, D. A. Forsyth, A. Zisserman, and J. L. Mundy. Extracting projective structure from single perspective views of 3d point sets. In *ICCV93*, pages 573–582, 1993.

[48] Claude E. Shannon. A mathematical theory of communications. *Bell Systems Technical Journal*, 27:379–423, 1948.

[49] Claude E. Shannon. Communication in the presense of noise. In *Proceedinds of the IRE*, volume 37, pages 10–21, January 1949.

[50] Christopher G. Small. *The Statistical Theory of Shape*. SV, 1996.

[51] H. Szu and R. Hartley. Fast simulated annealing. *Physics Letters A*, 122(3–4):157–162, 1987.

[52] Bill Lin Takayasu Sakurai and A. Richard Newton. Fast simulated diffusion: An optimization algorithm for multiminimum problems and its application to mosfet model parameter extraction. *IEEE Transactions on Computer Aided Design*, 11(2):228–234, February 1992.

[53] T. N. Tan. Monocular reconstruction of 3-d bilateral symmetrical objects. 1996.

[54] Constantino Tsallis and Daniel A. Stariolo. Generalized simulated annealing. *Physica A*, 233:395–406, 1996.

[55] Michael Unser. Sampling – 50 years after Shannon. *Proceedings of the IEEE*, 88(4):569–587, July 2000.

[56] Isaac Weiss. Model-based recognition of 3d curves from one view. *Journal of Mathematical Imaging and Vision*, 10(2):175–184, mar 1999.

[57] Isaac Weiss and Manjit Ray. Model-based recognition of 3d objects from single images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):116–128, feb 2001.

[58] Hermann Weyl. *The Classical Groups and Their Invariants*. Princeton University Press, Princeton, N.J., 2 edition, 1946.

[59] R. L. Yang. Convergence of the simulated annealing algorithm for continuous global optimization. *Journal of Optimization Theory and Applications*, 104(3):691–716, March 2000.

[60] Xin Yao. A new simulated annealing algorithm. *International Journal of Computer Mathematics*, 56:161–168, 1995.

[61] C. Yuan and H. Niemann. Neural networks for the recognition and pose estimation of 3d objects from a single 2d perspective view. *Image and Vision Computing*, 19(9–10):585–592, August 2001.

[62] Ahmed I. Zayed. *Advances in Shannon's Sampling Theory*. CRC Press (UK) LLC, 1993.

Appendix A: Proofs of Theorems from Chapter 2

A.1 Proof of Theorem 2.17

Let us suppose that there is no $\gamma(t)$, such that the half-line that starts at γ_0 and goes through $\gamma(t)$ is tangential to the curve C. By assumption there exist $t_1, t_2 \in [a, b],\ t_2 > t_1$, and the half-line ℓ starting from γ_0, such that it intersects $\mathscr{C}$ at the points $\gamma(t_1)$ and $\gamma(t_2)$.

- We first prove that we can choose t_1 and t_2 so that no point $\gamma(t)$ for $t \in (t_1, t_2)$ lies on ℓ. Suppose it is not true, i.e., for any chosen t_1, t_2 there exists $t_3 \in (t_1, t_2)$ such that $r(t_3) \in \ell$. Denote

$$t_1^{(1)} = t_1,\ t_2^{(1)} = t_3, \tag{A.1}$$

and consider the points $t_1^{(1)}$ and $t_2^{(1)}$. We have

$$\gamma\left(t_1^{(1)}\right), \gamma\left(t_2^{(1)}\right) \in \ell. \tag{A.2}$$

Then by assumption there must exist $t_3^{(1)} \in \left(t_1^{(1)}, t_2^{(1)}\right)$, such that $\gamma\left(t_3^{(1)}\right) \in \ell$. Denote

$$t_1^{(2)} = t_1^{(1)},\ t_2^{(2)} = t_3^{(1)}. \tag{A.3}$$

Proceeding in the same fashion, we see that we can construct two sequences: $\left\{t_1^{(n)}\right\}_{n\in\mathbb{N}}$, which is constant and $\left\{t_2^{(n)}\right\}_{n\in\mathbb{N}}$, which is strictly decreasing and bounded and thus convergent, i.e.,

$$\exists \lim_{n\to+\infty} t_2^{(n)} = t_1. \tag{A.4}$$

Because of smoothness of the curve $\gamma(t)$ we conclude, that ℓ is tangent to $\gamma(t)$ at the point t_1. This leads us to a contradiction, because we assumed that it was impossible. Thus we

have indeed proved that there exist two point t_1 and t_2, such that any point between $\gamma(t_1)$ and $\gamma(t_2)$ does not lie on ℓ.

- Let us now proceed as follows. Denote by $\phi(t)$ the angle between $\gamma(t) \equiv \gamma(t) - \gamma_0$ and $\tau(t)$. Suppose without loss of generality, that $\phi(t_1) \in (0, \pi)$. (It is clear that $\phi(t) \notin \{0, \pi\}$, and the case $\phi(t_1) \in (\pi, 2\pi)$ is absolutely analogous.) We will prove that at some point on $(t_1, t_2]$ the angle eventually becomes 0 or π, which means that the half-line connecting γ_0 and that point is tangential to C.

  Suppose, that $\exists \varepsilon > 0$ - arbitrarily small but fixed, such that $\forall t \in [t_1, t_2]$ $\phi(t) \in [\varepsilon, \pi - \varepsilon]$. It is clear then that the piece of the curve $\gamma(t), t \in [t_1, t_2]$ will completely lie in the sector. $\gamma(t_2) \in \ell$, and ℓ lies outside that sector. Since $\gamma(t)$ is continuous, we have a contradiction.

  The contradiction we have just come to means that $\forall n \in \mathbb{N}$ $\exists t_n \in [t_1, t_2]$ such that either $\pi - \frac{1}{2^n} < \phi(t_n) \leq \pi$ or $0 \leq \phi(t_n) < \frac{1}{2^n}$, which in turn means that there exists a subsequence $\{t_{n_k}\}$ such that $\{\phi(t_{n_k})\}$ converges to either 0 or π. Suppose, for example, that $\exists \{t_{n_k}\} \in [t_1, t_2]$, such that $\exists \lim\limits_{k \to +\infty} \phi(t_{n_k}) = 0$. Because $\{t_{n_k}\}$ is a bounded sequence, $\exists$ yet another subsequence $\left\{t_{n_{k_l}}\right\}$, such that $\exists \lim\limits_{l \to +\infty} t_{n_{k_l}} = s$ and $\exists \lim\limits_{l \to +\infty} \phi\left(t_{n_{k_l}}\right) = 0$. This proves that $\phi(s) = 0$, which means that the half-line going from γ_0 ro $\gamma(s)$ is tangential to C, which leads us to a contradiction with our assumptions again.

The assumption that there is no half-line connecting γ_0 and $\gamma(t)$ tangential to C is therefore wrong, which proves the theorem. $\square$

A.2 Proof of Theorem 2.19

Suppose first that there is no $\gamma(t)$ that is tangent to C, where $t \in [a, b]$. Denote the orientation of a pair of vectors $\{\gamma(t), \boldsymbol{\tau}(t)\}$ by $\mathscr{O}(t)$, i.e.,

$$\mathscr{O}(t) = sign \left(\begin{vmatrix} \gamma_1(t) & \gamma_2(t) \\ \tau_1(t) & \tau_2(t) \end{vmatrix} \right). \tag{A.5}$$

- *Forward part*:
  - Consider the case $\mathscr{O}(a) = 0$. Then $\gamma(0)$ and $\boldsymbol{\tau}(0)$ are collinear which is in contradiction with the assumption we just made.
  - Suppose then that $\mathscr{O}(a) \neq 0$. Without loss of generality assume $\mathscr{O}(a) = 1$. If $\exists t_0 \in [a, b]$, such that $\mathscr{O}(t_0) = -1$, then $\exists s \in [a, t_0]$, such that $\mathscr{O}(s) = 0$. Indeed, the determinant involved in the definition of the orientation is a continuous function of t, because of the smoothness of $\gamma(t)$. In addition, it is positive at a and negative at t_0. Then by Wierstrass theorem, there exists an intermediate point $s \in [a, t_0]$, where the continuous function equals 0.
- *Converse part*: Suppose that $\mathscr{O}(t) \equiv const \neq 0$. This means that that the determinant is either always positive or always negative, which indicates that $\gamma(t)$ and $\boldsymbol{\tau}(t)$ are never collinear. This proves the entire theorem. □

A.3 Proof of Theorem 2.20

Let C be a curve with the image $\mathscr{C}$. Let $\mathscr{A}_C$ be its admissible set and $\boldsymbol{\rho}^1, \boldsymbol{\rho}^2 \in \mathscr{A}_C$. Let $\alpha \in (0,1)$. We will show that the point $\boldsymbol{\rho} = \alpha\boldsymbol{\rho}^1 + (1-\alpha)\boldsymbol{\rho}^2$ belongs to $\mathscr{A}_C$. To proceed, per Theorem 2.19, we will have to prove the corresponding orientation function never changes its sign.

We have

$$\begin{vmatrix} \tau_1(t) & \tau_2(t) \\ \gamma_1(t) - \boldsymbol{\rho}_1 & \gamma_2(t) - \boldsymbol{\rho}_2 \end{vmatrix} = \begin{vmatrix} \tau_1(t) & \tau_2(t) \\ \gamma_1(t) - \alpha\boldsymbol{\rho}_1^1 - (1-\alpha)\boldsymbol{\rho}_1^2 & \gamma_2(t) - \alpha\boldsymbol{\rho}_2^1 - (1-\alpha)\boldsymbol{\rho}_2^2 \end{vmatrix} = \tag{A.6}$$

$$= \begin{vmatrix} \alpha\tau_1(t) + (1-\alpha)\tau_1(t) & \alpha\tau_2(t) + (1-\alpha)\tau_2(t) \\ \alpha\gamma_1(t) + (1-\alpha)\gamma_1(t) - \alpha\boldsymbol{\rho}_1^1 - (1-\alpha)\boldsymbol{\rho}_1^2 & \alpha\gamma_2(t) + (1-\alpha)\gamma_2(t) - \alpha\boldsymbol{\rho}_2^1 - (1-\alpha)\boldsymbol{\rho}_2^2 \end{vmatrix} = \tag{A.7}$$

$$= \alpha \begin{vmatrix} \tau_1(t) & \tau_2(t) \\ \gamma_1(t) - \boldsymbol{\rho}_1^1 & \gamma_2(t) - \boldsymbol{\rho}_2^1 \end{vmatrix} + (1-\alpha) \begin{vmatrix} \tau_1(t) & \tau_2(t) \\ \gamma_1(t) - \boldsymbol{\rho}_1^2 & \gamma_2(t) - \boldsymbol{\rho}_2^2 \end{vmatrix}. \tag{A.8}$$

Two options are to consider.

- Suppose $\mathscr{O}_{\boldsymbol{\rho}^1}(t) = \mathscr{O}_{\boldsymbol{\rho}^2}(t) = \mathscr{O}\ \forall t$. Then it follows from the above derivation that $\mathscr{O}_{\boldsymbol{\rho}}(t) = \mathscr{O}\ \forall t$. By assumptions, $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2 \in \mathscr{A}_C$, so that $\mathscr{O}$ is a constant not equal to zero. Which yields that $\boldsymbol{\rho} \in \mathscr{A}_C$.

- Alternatively, assume $\mathscr{O}_{\boldsymbol{\rho}^1}(t) \not\equiv \mathscr{O}_{\boldsymbol{\rho}^2}(t)$. Let us derive a contradiction with assumptions of the theorem. Without loss of generality we can assume that

$$\mathscr{O}_{\boldsymbol{\rho}^1}(t) = 1,\ \mathscr{O}_{\boldsymbol{\rho}^2}(t) = -1. \tag{A.9}$$

But there exist at least two points t_1, t_2, at which the two orientations must coincide. Indeed, the line, that contains the points $\boldsymbol{\rho}^1, \boldsymbol{\rho}^2$ intersects $\gamma(t)$ in two points. Denote one of them

$\gamma(t_1)$, and the other $\gamma(t_2)$. Then the vectors $\gamma(t_i) - \boldsymbol{\rho}^1$ and $\gamma(t_i) - \boldsymbol{\rho}^2$ are collinear for $i = 1, 2$, which yields that the orientations should be the same. Thus we get a contradiction with (A.9), which proves the theorem. $\square$

Appendix B: On Implementation of Sampling Algorithms

In this section, we discuss in details the implementation of theoretical results presented in the earlier sections of the paper. Very much like the vast majority of techniques presented in the continuous setting, the algorithms we have derived require the process of discretization.

In practice, we are given a binary image in the form

$$I(x,y) \in \{0,1\},\ x = 1,\dots,M,\ y = 1,\dots,N. \tag{B.1}$$

An image of a closed curve is then defined implicitly by points where the pixel value is 1, i.e.

$$\mathscr{C} = \{(x,y) \mid I(x,y) = 1\}. \tag{B.2}$$

Obviously, $\mathscr{C}$ is then a finite set and thus can be enumerated as

$$\mathscr{C} = \big\{(x_1,y_1),\dots,(x_P,y_P)\big\}. \tag{B.3}$$

B.1 Admissible Set

Let $\boldsymbol{r}_0 = (x_0,y_0),\ 1 \le x_0 \le M, 1 \le y_0 \le N$ be a given pixel in the image. Our goal is to determine if this point can serve as an admissible center of polar coordinates. Assume first that $\boldsymbol{r}_0$ is indeed an admissible polar center for $\mathscr{C}$. We will proceed as follows. Define a sequence

$$\mathscr{C}_{polar} = \big\{(\phi_1,r_1),\dots,(\phi_P,r_P)\big\},\ \phi_i \in [0,2\pi),\ r_i \ge 0,\ i = 1,\dots,P, \tag{B.4}$$

where

$$\forall i = 1,\dots,P,\ x_i = r_i\cos\phi_i,\ y_i = r_i\sin\phi_i. \tag{B.5}$$

Without loss of generality, we assume that

$$\phi_1 < \phi_2 < \cdots < \phi_{P-1} < \phi_P. \tag{B.6}$$

Otherwise we always can rearrange points. If for some distinct i and j, $\phi_i = \phi_j$, then we, of course, get a contradiction with the assumption that $\boldsymbol{r}_0$ is an admissible point, so in fact all inequalities in (B.6) are strict.

Note that any polar parametrization of the digital curve, if it exists, has the form

$$\big\{(\phi_1 + \Delta, r_1), \ldots, (\phi_P + \Delta, r_P)\big\},\ \Delta \in \mathbb{R}. \tag{B.7}$$

On the other hand, if (B.7) defines a polar parametrization for some $\Delta_0 \in \mathbb{R}$, then taking any other Δ yields a valid polar parametrization. We thus have shown that our construction provides us with a unique (up to an additive constant in angle) possible polar parametrization.

Recall, that we are still reasoning under the assumption that $\boldsymbol{r}_0 = (x_0, y_0)$ is an admissible polar center. Define vectors γ_i^s and τ_i by

$$\gamma_{i,1}^s = x_i - x_0,\ \gamma_{i,2}^s = y_i - y_0,\ i = 1, \ldots, P, \tag{B.8}$$

and

$$\tau_i = (x_i - x_0, y_i - y_0),\ i = 1, \ldots, P. \tag{B.9}$$

The vectors τ_i are discrete tangent vectors, and Theorem 2.19 assures then that the pairs $\big\{(\tau_i, \gamma_i^s)\big\}_{i=1,\ldots,P}$ should all have the same orientation.

Summarizing, the algorithm of determining if a particular point can be a polar center for a given curve works as follows.

1. Extract points belonging to the curve.

2. Convert them to a polar coordinate system centered at the point in question.

3. Convert the points' Cartesian coordinates to polar coordinates

4. Compute pairs of discrete vectors $\left\{(\gamma_i^s, \tau_i)\right\}$.

5. Accept the hypothesis that the point is admissible if the orientations of the pairs of vectors defined above does not change. Otherwise, conclude that the point is not in an admissible set.

Other algorithms used to produce numerical implementations are either described in the referenced literature or self-evident. We thus do not elaborate on them any further.