UNIVERSITY OF SOUTHAMPTON

On Analysing Deformable (Moving) Objects

by

Muayed S. Al-Huseiny

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the

Faculty of Physical and Applied Sciences

School of Electronics and Computer Science

February 2012

Contents

List of Figures

List of Tables

Nomenclature

$\mathbb{R}^2$	2D real domain
Ω	2D image domain (subset of $\mathbb{R}^2$)
Γ	evolving curve (active contour)
ϖ	an image region within Ω
u_0	input image formed of object and background
$F(\Gamma)$	segmentation functional
c_1	average of the image values inside the contour
c_2	average of image values outside the contour
μ	weight of the contour length
λ_1	weight of the force inside the contour
λ_2	weight of the force outside the contour
ϕ	signed distance function
$H(\cdot)$	Heaviside function
$\delta(\cdot)$	Dirac function
E_{cv}	Chan and Vese functional
$\vec{o}_1$	parameter vector of the distribution of the image foreground
$\vec{o}_2$	parameter vector of the distribution of the image background
$P(\cdot)$	probability density function
$p(\cdot)$	probability
$p_r(\cdot)$	prior probability
σ_1	variance of the distribution of the image foreground
σ_2	variance of the distribution of the image background
log	logarithm function
$\tilde{E}_{cv}$	statistical variant of E_{cv}

ϵ	regularisation parameter
h	space step
Δt	time step
ACWE	active contour model without edges
MRI	magnetic resonance imaging
PCA	principal component analysis
AR	autoregressive model
SDF	signed distance function
n	number of training shapes
Q	number of pixels of shape model training image
R	number of rows of a training image
C	number of columns of a training image
T	the set of training shapes
ϕ_i	i^{th} training SDF
$\bar{\mathbf{u}}$	mean shape
$\acute{\phi}$	mean offset distance function
$\vec{u}_i$	vectorised i^{th} training shape
$\mathbf{V}$	matrix of vectorised training shapes
Ψ	eigenvectors matrix of $\mathbf{V}\mathbf{V}^T$
υ_i	i^{th} eigenvalue of $\mathbf{V}\mathbf{V}^T$
$\grave{\Psi}$	eigenvectors matrix of $\mathbf{V}^T\mathbf{V}$
$\vec{\alpha}$	eigenmodes of the shapes
$cspline(\cdot)$	cubic spline interpolation
$Dist(\cdot,\cdot)$	number of error pixels in an estimated shape
ACF	autocorrelation function
NPP	normal probability plot
PPCC	probability plot correlation coefficients
lambda	distribution shape parameter
$Eigenmode^*$	vector of corresponding elements of $\vec{\alpha}$ of multiple subjects
M	number of training gait cycle
$\mathbf{S}$	matrix of shape vectors

PCA_{basis}	1^{st} PCA in the generative model
λ	eigenvalues of PCA_{basis}
$\text{PCA}_{feature}$	2^{nd} PCA in the generative model
$\bar{\alpha}$	mean eigenmode
$\mathbf{F}$	feature matrix
$\mathbf{K}$	finite dimensional kernel of $\mathbf{F}^T\mathbf{F}$
X	eigenvectors of $\mathbf{K}$
η	eigenvalues of $\mathbf{K}$
χ	eigenvectors of covariance matrix $\mathbf{F}\mathbf{F}^T$
$\vec{\beta}$	eigenmodes of the cycles in $\text{PCA}_{feature}$ (subject ID)
D	L^2 distance
Er	error function, the number of erroneous estimated shape elements
d	distance between gait cycles based on $\vec{\alpha}$
$\mathbf{V}^i$	i^{th} vectorised gait cycle
$\hat{d}$	distance between gait cycles based on $\vec{\beta}$
D_E	minimum Euclidean distance
$p(x, y)$	reference shape
$q(x, y)$	observed shape
B	boundary of the shape
I_B	interior of the shape
s	scaling factor
θ	rotation angle
T_x	translation along the x-axis
T_y	translation along the y-axis
E	distance term of the registration problem
$\mathbf{R}_\theta$	conventional 2D rotation matrix
r_x	x-axis element of the centroid of any shape, r
r_y	y-axis element of the centroid of any shape, r
ρ	radius element of the polar coordinates
ω	angle element of the polar coordinates
$\hat{\phi}$	centralised shape SDF ϕ

$\tilde{\phi}$	shape $\hat{\phi}$ in polar coordinates with ρ marginalised out
$\bar{\phi}$	normalised ϕ
$\tilde{\beta}$	scale factor as a function of rotation angle
ψ	Fourier transform of a polar coordinates shape
M_m	m^{th} radial moment
E_s	distance term E with respect to s only
$\acute{E}_s$	linearised E_s
$\hat{s}$	$\log s$
$\tilde{\zeta}$	scale factor as a function of translation
Θ	Fourier transform of a Cartesian coordinates shape
RBC	red blood cell
C3	3^{rd} cervical vertebra
DFT	discrete Fourier transform
MAP	maximum a posteriori
ϕ^{Pr}	prior shape SDF
E_{prior}	prior energy
E_{total}	total energy
γ_1	weight of the data energy
γ_2	weight of the prior shape energy
$\acute{\epsilon}$	quality of segmentation measure
SNR	signal to noise ratio
$\mathbf{S}_{train}$	classifier training set
M	total number of cycle in $\mathbf{S}_{train}$
c	the average number of frames in $\mathbf{S}_{train}$
$\mathbf{S}_{test}$	classifier testing set
r	the average number of cycles per subject in $\mathbf{S}_{test}$
MDA	multiple discriminant analysis
GEI	gait energy image
D_M	Mahalanobis distance
CCR	correct classification rate

Acknowledgements

I would like to express my gratitude to my supervisors, Dr. Sasan Mahmoodi and Professor Mark S. Nixon for their invaluable guidance and encouragement during this research. Sasan has taught me through every station of this journey, and Mark has always shown me the broader picture.

Thanks are also due to staff at University of Southampton, Tracey Cantlie, Amanda Goodacre, Sue Ganney, and John Wynn for their supportive attitude. I would like also to extend my thanks to all friends and colleagues in the ISIS research group, in particular to Ali Hassan, Mina Ibrahim, Amirthalingam Ramanan, and Tariq Hassan.

I should gratefully mention the support from the Government of Iraq via the Ministry of Higher Education and Scientific Research who granted me a full scholarship which made this project financially possible.

Last, but not least, I am indebted to my wife who shared an enormous part of the burden in order to make life easier for me.

Thank you all ...

To my dear family: father, mother, brothers and sisters and especially my wife and the little angels "Narjis" *and* "Ali" *for their love and support...*

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

by Muayed S. Al-Huseiny

Performing a high level vision is usually based on features extracted at low and intermediate levels of the process of perception of a visual scene.

Segmentation and matching are instrumental tasks in producing comparable features in applications such as medical imaging, mining and oil extraction, gaming consoles, face, ear and gait biometrics, and etc.

The ultimate goal of this study is to develop a fully functional prior aided segmentation framework to extract deformable shapes over a sequence of frames. This thesis acknowledges the demand by these applications for a robust and flexible approach which is particularly designed to extract deformable timely shape sequences. It is also recognised that existing methods are either too general, and thus inaccurate, or too specific, thereby limited in usability.

This thesis suggests a learning model for gait synthesis with the ability to extrapolate to novel data. It involves computing comparable features from multiple sources. We show that these features which we formulate as continuous functions can be modeled by linear PCA.

This thesis also proposes a new fast and robust shape registration algorithm to match shapes from different sources in the proposed framework. This algorithm is based on linear orthogonal transformations and shape moments. The registration parameters are computed directly by analysing the signed distance functions of

the shapes. This is in-line with the level sets based prior shape segmentation framework adopted here.

The segmentation is performed in a balanced framework between the data in the given images on one hand and the prior induced by the shape model and the registration algorithm proposed here on the other hand. This configuration ensures more control for the shape force over the overall shape geometry. Thus, favouring shapes familiar to the learned knowledge.

Chapter 1

Introduction

1.1 Context

Computer vision aims to mimic the basic principles of human visual perception. The process of vision involves the acquisition and analysis of a natural scene. Image analysis and computer vision relies on two main operations: feature extraction, and feature analysis. Feature extraction processes the captured image to extract a set of values or functions known as features, for example, by adopting image processing techniques. These features, are then analysed by using machine intelligence and learning algorithms to mediate understanding the sensed image.

This study is dedicated to techniques related to the extraction of visual features. Feature extraction can be performed in various ways. Segmentation, for instance, partitions an image domain into parts, each corresponding to a distinct object or region within the image. Detection can also be performed to find automatically the location of a known object within an image with minimal user intervention. Tracking on the other hand is the detection of an object in successive frames within temporally sequenced images such as videos. One other important operation is registration which aligns various shapes/images for comparison and further analysis.

To elaborate, segmentation seeks to define a boundary surrounding pixels with similar properties and features. This boundary represents the shape of a region of interest. An important issue is the segmentation of deformable shapes over time. The notion of shape deformation over time implies the continuous change of the region of interest. Therefore, instead of specifying one class of objects, there is a need to define a large set of classes in order to account for shape variation of the object over a period of time.

Researchers have also focussed on the use of human gait as a biometric. Recognition approaches are mostly founded on utilising sequences of the extracted shapes of walking subjects. The extraction is achieved either by manual segmentation or by background subtraction: the former is non automatic and the latter fails to concentrate on the moving subject alone, but extracts all moving objects. Also, very recently, some algorithms such as those enabling cars to avoid collision with pedestrians, or algorithms implemented in gaming consoles that help recognise human action have been proposed. The fact that these algorithms are built (in one way or another) on the assumption that the timely shape, location and pose of an objects at a scene are known, explains the need for reliably functional gait segmentation methodologies.

The fact that non-synthetic real-world images are often complex, cluttered, noisy and sometimes occluded makes the achievement of robust segmentation a serious challenge. Further, the segmentation of human gait is an even more challenging task due to the fact that scenes containing shapes of walking persons are prone to the kinds of distortions mentioned above, in addition to the presence of persistent occlusion and deformation.

Prior knowledge is known to provide better and more reliable results, due to its capability to compensate for missing or misleading information (Cremers et al. (2002), Rousson and Deriche (2003), Nixon and Aguado (2008)). This led the thinking towards the inclusion of prior knowledge in segmentation models. In the case of gait, this involves the prediction of shapes (by a prior shape model) and matching them to the detected data. However, the achievement of a reliable gait

(a)

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

FIGURE 1.1: (a) An of image of a walking subject, (b-e) A sequence of images of a walking subject

shape model is not straightforward due to the fact that the deformation of shapes is non-Gaussian (Cremers et al., 2006; Mahmoodi, 2009) (see Figure (1.2) and Section (4.3)), which exacerbates the difficulty. On the other hand, the periodicity and correlation of shapes over time can give clues on the shape and location in a dynamic fashion.

To explore the prospects of segmentation, consider for example, the scene in Figure (1.1(a)). It is formed of many objects (each with sub-objects), some of them are salient and others are less so, some objects are obstructed by other objects or by other parts of the same object. One approach to understand the scene is to segment all the objects in this image and then to identify those extracted objects. This (rather idealistic) approach assumes that perfect segmentation can be achieved, which is rarely the case due to the difficulties raised above.

Another approach is to decide beforehand what the scene should contain and then

to attempt to extract it. This (more realistic) procedure reduces the search space significantly and allows for incomplete or incorrect information in the data to be accommodated for.

In other words, the first approach phrases the problem as: given a detected (segmented) object, what is it? The issues here are, how reliable are the given features and how feasible is the search within all the modalities of objects and their possible variances? One more issue is the reliability of the segmentation of each individual object. The second approach states the problem as: given an identified object in an image, how to detect this object in this image. An issue here is, how likely is that such an object exists in the scene. This can be decided by a particular application. Another issue is that, whether the detection is feasibly and reliably achievable. Surely, the detection of a known object is likely to require less computational effort than the detection of all (probably unknown) objects. Hence, this thesis adopts the second approach.

By inspecting Figure 1.1(b-e), and assuming that the goal is to extract the walking person in these images. The segmentation of these images then according to the former approach is done by considering each image as a separate problem, and is solved by segmenting each object (perfectly). This is started over each time. The latter approach, however, can be taken by building knowledge gained by solving one problem. The solution is obtained by matching the most probable guess to the data. This solution is then used to predict a potential solution to the next problem (next frame).

Some authors suggest the use of kernel density (Cremers et al., 2004) to model the modes of deformation of gait shapes. The problem, however, with this approach is that the accuracy of estimation depends on the (empirically) chosen kernel bandwidth, also, the density is characterised by its high dimensionality. Furthermore, this method assumes that the data is independent and identically distributed (iid), whilst gait data actually expresses a high degree of correlation manifested by the time coherence of successive shapes.

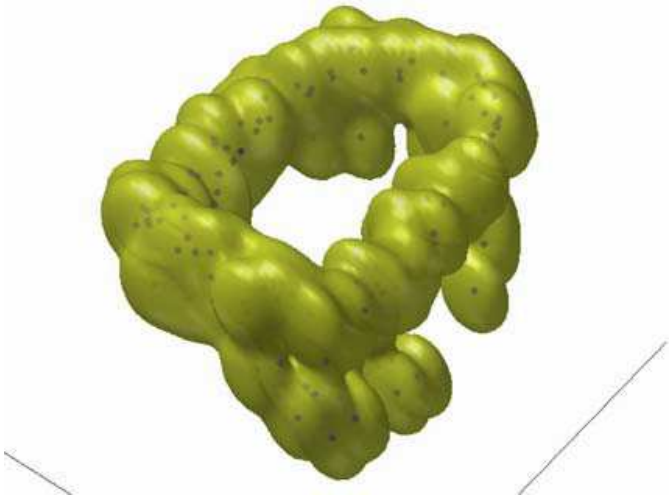


FIGURE 1.2: Density estimated for a 3D projection of 100 silhouettes (Courtesy of Cremers et al. (2006))

Another approach uses traditional linear PCA to reduce the dimensionality, while the temporally dependent deformation will be handled by some mechanism. Cremers (2006), as a leading example in this direction, have developed a gait segmentation model based on Autoregressive (AR) systems as the mechanism to synthesise new shapes. This model, though, has some issues: the algorithm memorises the current and previous frames' optimal segmentations, therefore, in the case when incorrect decision (segmentation) is made this would not be corrected at later stages of processing (Schmidt and Cremers, 2009). Such an approach therefore, results in removal of the temporal coherence of the shapes produced by the prior. The generation of shapes out of order with regard to their predecessor shapes would definitely degrade the quality of segmentation. Additionally, this algorithm can only be trained by using data from a single source, as no suggestion is made on how to synchronise and join data from different sources. Furthermore, the segmentation algorithm built on this model does not offer a scaling parameter between the template and data. This limits the usability of the algorithm to images of subjects with a similar size to that of the training set.

To better address gait segmentation, we propose a new approach to deal with this problem, the main aspects of the proposed framework are discussed in the next section.

1.2 Contributions

This thesis proposes a shape model which builds on the previous approach and overcomes some of the reported difficulties by using a better interpretation of the gait data. Also presented here is a powerful registration approach to match shapes to each other for the training of the shape model. The shape model and the registration technique are used to drive a region based segmentation model towards extracting successive shapes of human gait.

This work uses the segmentation model proposed by Chan and Vese (2001) to drive the data term of the segmentation energy. The shape and the data are matched by using the registration algorithm suggested in this thesis.

The proposed model dictates that entire gait cycles rather than individual shapes are modelled. Our model is based on using PCA, first to extract compatible individuals features, and then to model these features linearly. This is built on exploiting the quasi-periodic behaviour of the eigenmodes of a gait cycle which enables the automatic identification and extraction of a single period as a single cycle for each individual. These extracted cycles are then aligned and modeled together.

In this thesis, it is accepted that the deformation of shapes in a gait cycle along the time axis is non-Gaussian (Cremers, 2006). However, it is recognised that the deformation of corresponding shapes from different cycles can be approximated by a Gaussian distribution as explained in Section (4.3) (see Figure (1.3)). Then, the problem of non-Gaussian shape deformation is reduced into a simpler problem of Gaussian feature modelling. The non-Gaussianity of the deformation of the shapes is encapsulated in these features. This on one hand preserves the periodicity and time coherence of the data. On the other hand, it facilitates the automatic alignment of different gait cycles in order to model corresponding shapes together. Therefore, by using this configuration, we assumes an instantaneous Gaussian model at each time instant over the corresponding shape variations (Al-Huseiny et al., 2010).

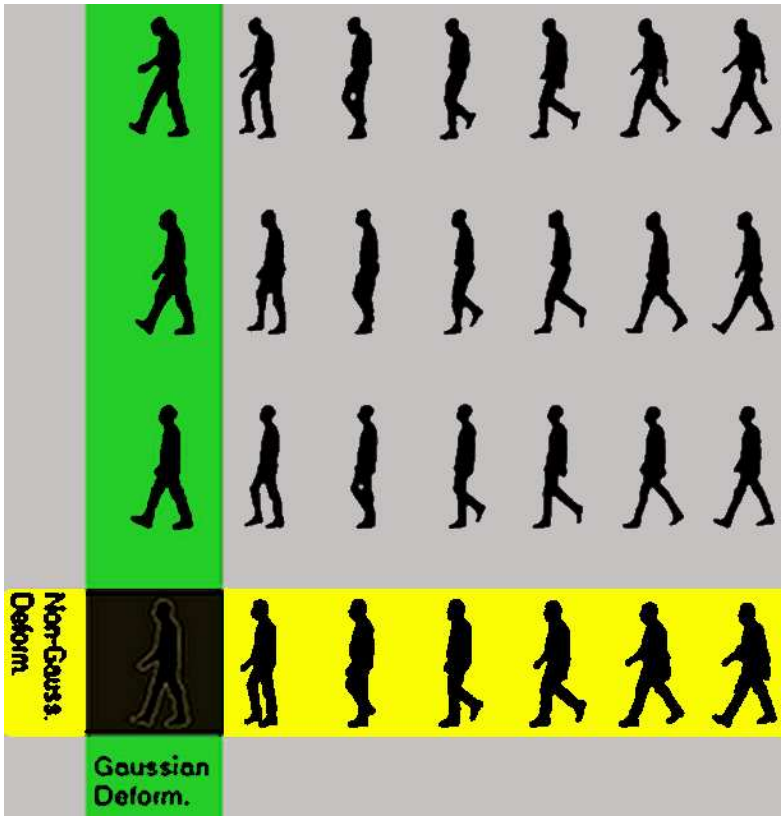


FIGURE 1.3: A set of different gait cycles synchronised chronologically. Yellow stripe: non Gaussian deformation of the shapes from the same cycle over time. Green stripe: Gaussian deformation of the corresponding shapes from different cycle

Another crucial component in this context is registration. In this thesis, a robust and non iterative shape registration algorithm is proposed. This technique employs the Fourier transform to factor out the phase shift between two signals (shapes to be registered). These shapes are mapped to appropriate domains in order to redefine the respective transformation parameters as a shifts. This framework is a novel adaptation and employment of these techniques using level sets.

The scale parameter, which describes the difference in size between the shapes, is computed by employing a new formulation to factorise the scale difference between two shapes from their corresponding radial moments. The definition of a signed distance function is also altered to suit the properties of the techniques used to compute the registration parameters.

The framework of the prior knowledge segmentation of gait shapes is achieved by the balancing between the analysis of image data and the estimated shape.

The proposed segmentation algorithm is formulated such that the shape term is the deciding force of the topology and geometry. The data driven term, on the other hand, acts locally around the evolving boundary to fine tune the result. This structure ensures robustness against distortions and relies on good shape estimates provided by the proposed shape model.

Publications

Papers based on this work include:

1. M. S. Al-Huseiny, S. Mahmoodi, and M. S. Nixon, Gait Sequence Synthesis and Reconstruction, in BMVC Workshop 2009 London, UK, 2009.

2. M. S. Al-Huseiny, S. Mahmoodi, and M. S. Nixon, Level Set Gait Analysis for Synthesis and Reconstruction, Proceedings of ISVC 2009, Las Vegas, USA, 2009.

3. M. S. Al-Huseiny, S. Mahmoodi, and M. S. Nixon, Gait Learning-Based Regenerative Model: a Level Set Approach, Proceedings of ICPR, Turkey 2010, Istanbul, 2010.

4. M. S. Al-Huseiny, S. Mahmoodi, and M. S. Nixon, Robust Rigid Shape Registration Method Using a Level Set Formulation, Proceedings of ISVC 2010, Las Vegas, USA.

5. M. S. Al-Huseiny, S. Mahmoodi, and M. S. Nixon, Robust and Fast Shape Registration Method, Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence in May 2011.

Chapter 2

Active Contour Segmentation

2.1 Summary

This chapter introduces the topic of image segmentation and discusses in particular the active contour without edges model. The subsequent sections describe the mathematical formulation of this model. Also, the assumptions made regarding the scenes included in the images are provided. The experimental results demonstrated in this chapter show the points of strength in this well established model, particularly, in segmenting bimodal images. The outcomes of applying this model to real-world image segmentation problems are also investigated.

2.2 Introduction

Segmentation is the process by which an image is divided into plausible (homogeneous) regions: an object and a background. These regions are characterised by visual consistency. Segmentation is an essential step for some high level vision tasks such as shape analysis, medical imaging, remote sensing, tracking, detection, and recognition.

One of the earliest segmentation models was introduced by Yuille (1991) who used a flexible template to find facial features specifically the eye. This approach requires the maximisation of a functional that gives a template which matches the edge points of the eye (Nixon and Aguado, 2008). This approach is limited in its application and is expensive to compute.

An important development in the field of segmentation has been achieved by introducing active contour or simply the snake by Kass et al. (1988). Snakes introduce spatial regularity into the pixel-wise edge detection operation (Chan and Shen, 2005). The active contour model is more flexible compared to the flexible template and is based on variational optimisation to ensure that the snake moves to a local minimum. However, detecting a change in topology requires special mechanisms (often heuristic). This approach is also non-intrinsic, since the energy depends on the parametrization of the curve and is not directly related to the objects geometry (Caselles et al., 1997).

The representation of the evolving contour received a new impetus when Osher and Sethian (1988) proposed the level set method to implicitly advance the contours. Level sets offer parameter free tool to characterise the shapes. Evolving the embedding function allows us to elegantly model topological changes of the boundary such as splitting and merging (Cremers, 2006). The fact that level sets are monotonically smooth functions with continuous derivatives makes them favorite for optimisation procedures.

Caselles et al. (1997) developed a level set based geometric active contour in which a geodesic term (derived form the image data) is included to give the contour its elasticity and smoothness. Although the geodesic active contours model implemented in the level set framework has many improvements over classic snakes, these improvements come at higher computational cost, which renders its utilization for time-critical applications problematic (Papandreou and Maragos, 2004).

The above mentioned active contour models have been developed considering that objects could be segmented by using edges gradients (Chan and Shen, 2005). This

notion is not always valid, in particular, in the case of low or even no gradient at some place of the edge the contour leaks and collapses.

Chan and Vese (2001) proposed a region based active contour as a special case of the Mumford-Shah functional (Mumford and Shah, 1989) to address the issue of weak and diffuse edges. As far as a pure segmentation task (using synthetic images) is concerned, this model may be one of the best to solve many of the problems associated with the active contours approach (such as initialisation, parameterisation, and contour leakage). This contour is implemented by using level set methods. This should facilitate the split and merge of the contour as it evolves implicitly (Chan and Vese, 2001).

The problem of snake evolution is solved via the employment of variational methods (Brechtken-Manderscheid, 1991). These optimisation approach define a quantity (energy) that reaches an optimum (maxima or minima) when the solution is reached. The process is accomplished iteratively by finding successive variations of the energy with respect to some parameters (the contour). Energy reduction is conducted numerically by applying the gradient descent technique.

2.3 Active Contour without Edges (ACWE)

This model is proposed by Chan and Vese (2001). It is a level set region based model, built on the Mumford-Shah functional (Mumford and Shah, 1989). This model assumes a bimodal image dominated by two Gaussian distributions representing the target and the background. Statistical properties are computed to drive the separation of the different regions. While the original model considers only the mean inside the regions, an extended version (Rousson and Deriche, 2003) includes regions' variances. By avoiding the use of edge gradients, this model achieves resistance to boundary leakage and contour collapse. This model hence has the ability to find objects when the boundary is weak or diffuse (Nixon and Aguado, 2008).

2.3.1 Model Description

Let Ω be the image domain (a bounded open subset of $\mathbb{R}^2$) and Γ be the evolving curve which is the boundary of the region $\varpi \subset \Omega$, where $\Gamma = \partial\varpi$. Also $inside(\Gamma)$ represents the region ϖ and $outside(\Gamma)$ represents the region $\Omega \backslash \overline{\varpi}$ (Chan and Vese, 2001).

Let u_0 be an image formed by two regions of approximately piecewise-constant intensities with values u_0^i and u_0^o. Assume further that the object to be detected is represented by the region with the value u_0^i and its boundary Γ_0. This leads to $u_0 \approx u_0^i$ inside the object i.e., inside Γ_0, and $u_0 \approx u_0^o$ outside the object i.e., outside Γ_0.

Now consider the following fitting term:

$$F(\Gamma) = \int_{inside(\Gamma)} (u_0(x,y) - c_1)^2 \, \mathrm{d}x\mathrm{d}y + \int_{outside(\Gamma)} (u_0(x,y) - c_2)^2 \, \mathrm{d}x\mathrm{d}y, \quad (2.1)$$

where the constants c_1 and c_2 (depending on Γ) are the averages of u_0 inside and outside Γ respectively (Chan and Shen, 2005).

In this case Γ_0 is the minimiser of the fitting term $F(\Gamma)$. These two terms are functionally equivalent to the external force of the classical snake models (Kass et al., 1988), with the distinction that the stopping criteria here is region based rather than edge based.

A length term is included in the snake functional for the purpose of spatial regularisation and to remove excessive contours, an area term can be added as well. These two terms resemble the contour internal force seen in the models of the classical snakes (Nixon and Aguado, 2008).

The model functional is then defined as:

$$\begin{aligned} F(c_1, c_2, \Gamma) = \mu \cdot |\Gamma| \quad &+ \quad \lambda_1 \cdot \int_{inside(\Gamma)} (u_0(x,y) - c_1)^2 \, \mathrm{d}x\mathrm{d}y \\ &+ \quad \lambda_2 \cdot \int_{outside(\Gamma)} (u_0(x,y) - c_2)^2 \, \mathrm{d}x\mathrm{d}y, \end{aligned} \quad (2.2)$$

where $\mu \geq 0$ and $\lambda_1, \lambda_2 > 0$, are fixed parameters.

In level set formulation, $\Gamma \subset \Omega$ is embedded as the zero level set of the function $\phi : \Omega \to \mathbb{R}$ such that:

$$\left\{ \begin{array}{ccc} \Gamma & = & \partial\varpi = \{(x,y) \in \Omega : \phi(x,y) = 0\}, \\ inside(\Gamma) & = & \varpi = (x,y) : \{(x,y) \in \Omega : \phi(x,y) > 0\}, \\ outside(\Gamma) & = & \Omega \backslash \overline{\varpi} = (x,y) : \{(x,y) \in \Omega : \phi(x,y) < 0\}. \end{array} \right\} \tag{2.3}$$

Hence, Γ is replaced by ϕ in the model functional, also the Heaviside function, H, and the Dirac function, δ, are employed. These are defined in Eq. (2.4) and Eq. (2.5) respectively,

$$H(\phi) = \left\{ \begin{array}{l} 1, \text{ if } \phi \geq 0 \\ 0, \text{ if } \phi < 0 \end{array} \right\}, \tag{2.4}$$

$$\delta_0(\phi) = \frac{\mathrm{d}}{\mathrm{d}\phi} H(\phi). \tag{2.5}$$

The energy functional in Eq.(2.2) is, accordingly, redefined in Eq. (2.6):

$$\begin{aligned} E_{cv}(c_1, c_2, \phi) = \quad & \mu \int_\Omega \delta_0 \left(\phi(x,y)\right) \left(\nabla\phi(x,y)\right) \mathrm{d}x\mathrm{d}y \\ + \quad & \lambda_1 \int_\Omega \left(u_0(x,y) - c_1\right)^2 H\left(\phi(x,y)\right) \mathrm{d}x\mathrm{d}y \\ + \quad & \lambda_2 \int_\Omega \left(u_0(x,y) - c_2\right)^2 \left(1 - H\left(\phi(x,y)\right)\right) \mathrm{d}x\mathrm{d}y \end{aligned} \tag{2.6}$$

Minimising $E_{cv}(c_1, c_2, \phi)$ with respect to c_1 and c_2 gives Eq. (2.7) and Eq. (2.8):

$$c_1(\phi) = \frac{\int_\Omega u_0(x,y) H\left(\phi(x,y)\right) \mathrm{d}x\mathrm{d}y}{\int_\Omega H\left(\phi(x,y)\right) \mathrm{d}x\mathrm{d}y}, \tag{2.7}$$

$$c_2(\phi) = \frac{\int_\Omega u_0(x,y) \left(1 - H\left(\phi(x,y)\right)\right) \mathrm{d}x\mathrm{d}y}{\int_\Omega \left(1 - H\left(\phi(x,y)\right)\right) \mathrm{d}x\mathrm{d}y}. \tag{2.8}$$

The Euler-Lagrange equation for ϕ is derived by fixing c_1 and c_2 and minimising $E_{cv}(c_1, c_2, \phi)$ with respect to ϕ (see Appendix A for further details), which gives turn to Eq. (2.9):

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon \left[\mu \, \mathrm{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 \right] \tag{2.9}$$

where t is the artificial time introduced to parameterise the contour evolution.

A generalised formulation of the Chan and Vese model is proposed by Rousson and Deriche (2003) to better fit the distributions of the regions in the image. This approach allows for the variance in addition to the mean to vary freely. This is unlike the original setting which assumes unity variance in the image regions. The extension suggested by Rousson and Deriche (2003) makes the model more robust against random noise added to the scene.

Let the image u_0 consist of two objects, target defined by parameters $\vec{o}_1$ and background defined by parameters $\vec{o}_2$, we can define the conditional probability density function $P(\vec{o}_i|u_0)$. The optimisation of the hypotheses regarding each object given the image and the prior probability p_r is given in Eq. (2.10) and Eq. (2.11) by using Bayes theorem:

$$P(\vec{o}_1|u_0) = \frac{P(u_0|\vec{o}_1) p_r(\vec{o}_1)}{p(u_0)}, \tag{2.10}$$

$$P(\vec{o}_2|u_0) = \frac{P(u_0|\vec{o}_2) p_r(\vec{o}_2)}{p(u_0)}, \tag{2.11}$$

where $p(u_0|\vec{o}_i)$ is the likelihood that a pixel (x, y) is distributed according to $\vec{o}_i$.

By assuming that all partitions are equally likely, i.e., $p_r(\vec{o}_1) = p_r(\vec{o}_2) = \frac{1}{2}$, and ignoring the normalisation factor, $p(u_0)$, the optimisation with respect to each hypothesis is achieved by computing the a posteriori probability:

$$p(\vec{o}_1|u_0) = p(u_0|\vec{o}_1), \tag{2.12}$$

$$p(\vec{o}_2|u_0) = p(u_0|\vec{o}_2). \tag{2.13}$$

The segmenting contour, therefore, is evolved by minimising the following energy (assuming that pixels in the regions are independent):

$$F(\Gamma) = \prod_{i=1}^{2} \left(\prod_{\Omega} p(\vec{o}_i|u_0) \right). \tag{2.14}$$

By employing a Gaussian distribution to represent the two regions and by using maximum likelihood, the segmentation is accomplished by minimising the energy in (2.15):

$$\begin{aligned} F(c_1, c_2, \sigma_1, \sigma_2, \Gamma) &= \mu.|\Gamma| \\ &- \lambda_1 \int_{inside(\Gamma)} \log \left(\frac{1}{\sqrt{2\pi\sigma_1}} e^{\frac{-(u_0-c_1)^2}{2\sigma_1^2}} \right) \mathrm{d}x\mathrm{d}y \\ &- \lambda_2 \int_{outside(\Gamma)} \log \left(\frac{1}{\sqrt{2\pi\sigma_2}} e^{\frac{-(u_0-c_1)^2}{2\sigma_2^2}} \right) \mathrm{d}x\mathrm{d}y, \end{aligned} \tag{2.15}$$

The functional in (2.15) is rewritten in Eq. (2.16) by using level sets representation:

$$\begin{aligned} \tilde{E}_{cv}(c_1, c_2, \sigma_1, \sigma_2, \phi) &= \mu \int_{\Omega} \delta\left(\phi(x,y)\right) \left|\nabla\phi(x,y)\right| \mathrm{d}x\mathrm{d}y \\ &+ \lambda_1 \int_{\Omega} \left(\frac{(u_0 - c_1)^2}{2\sigma_1^2} + \log(\sqrt{2\pi\sigma_1}) \right) H\left(\phi\right) \\ &+ \lambda_2 \int_{\Omega} \left(\frac{(u_0 - c_2)^2}{2\sigma_2^2} + \log(\sqrt{2\pi\sigma_2}) \right) \left(1 - H\left(\phi\right)\right) \end{aligned} \tag{2.16}$$

Similar to the procedure of minimising E_{cv} above, $\tilde{E}_{cv}$, is minimised with respect to c_1, c_2, σ_1 and σ_2 as in Eq. (2.7), (2.8), (2.17) and (2.18) respectively:

$$\sigma_1(\phi) = \frac{\int_{\Omega} (u_0 - c_1)^2 H(\phi)}{\int_{\Omega} H(\phi)}, \tag{2.17}$$

$$\sigma_2(\phi) = \frac{\int_{\Omega} (u_0 - c_2)^2 (1 - H(\phi))}{\int_{\Omega} (1 - H(\phi))}, \tag{2.18}$$

The solution to the segmentation problem can be found by evolving Eq. (2.19) with time (see Rousson and Deriche (2003) and the references therein for further

details):

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi) \left[\mu \, \mathrm{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \log \frac{\sigma_2}{\sigma_1} - \lambda_1 \left(\frac{(u_0 - c_1)^2}{2\sigma_1^2} \right) + \lambda_2 \left(\frac{(u_0 - c_2)^2}{2\sigma_2^2} \right) \right] \tag{2.19}$$

2.4 Methodology and Experimentation

In order to realise the energy functional in Eq. (2.9), regularised forms of H and δ_0, namely H_ϵ and δ_ϵ are employed, with $\epsilon \to 0$ as the regularisation parameter. These regularised functions are defined respectively in Eq.(2.20) and (2.21):

$$H_\epsilon(\phi) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan \left(\frac{\phi}{\epsilon} \right) \right), \tag{2.20}$$

$$\delta_\epsilon(\phi) = \frac{\partial}{\partial \phi} H_\epsilon(\phi) = \frac{\epsilon}{\pi \left(\epsilon^2 + \phi^2 \right)}. \tag{2.21}$$

Following Vese and Chan (2002), the energy functional of Eq. (2.9) is approximated for the purpose of numerical implementation into:

$$\begin{aligned} \phi_{i,j}^{n+1} = \frac{1}{C} [\, \phi_{i,j}^n \;\; + \;\; & m(\, c_3 \, \phi_{i+1,j}^n \; + \; c_4 \, \phi_{i-1,j}^n \; + \; c_5 \, \phi_{i,j+1}^n \; + \; c_6 \, \phi_{i,j-1}^n) \\ + \;\; & \triangle t \; \delta_\epsilon(\phi_{i,j}) \; (-\lambda_1 (u_0 - c_1)^2 \; + \; \lambda_2 (u_0 - c_2)^2 \;)], \end{aligned} \tag{2.22}$$

where $\triangle t$ is the time step parameter, c_1 and c_2 are computed as in Eq. (2.7) and (2.8). The computations of c_3, c_4, c_5, c_6, m, and C are listed in Appendix B:

The above numerical approximation is used in this chapter to test this model. In the following experiments the parameters are chosen as follows: $\lambda_1 = \lambda_2 = 1$, and $h = \epsilon = 1$. $\triangle t$ and μ are chosen differently in each experiment depending on the image.

Alternatively, Eq. (2.19) can be used in the segmentation. A numerical approximation of $\tilde{E}_{cv}$ is, thus, given in Eq. (2.23):

$$\begin{aligned}\phi_{i,j}^{n+1} = \frac{1}{C}[\phi_{i,j}^{n} &+ m(c_3\phi_{i+1,j}^{n} + c_4\phi_{i-1,j}^{n} + c_5\phi_{i,j+1}^{n} + c_6\phi_{i,j-1}^{n}) \\ &+ \triangle t\delta_{\epsilon}(\phi_{i,j}) - \lambda_1\left(\log\sigma_1 + \frac{(u_0 - c_1)^2}{2\sigma_1^2}\right) \\ &+ \lambda_2\left(\log\sigma_2 + \frac{(u_0 - c_2)^2}{2\sigma_2^2}\right)]. \end{aligned} \tag{2.23}$$

2.4.1 Synthetic Images

This experiment deals with the general (theoretical) problems of segmentation, including contour initialisation, curvature handling, multiple objects, and weak edge. By using the sort of bimodal images the assumption here is that the given scene is formed of two regions. The intensities in these regions form two Gaussian distributions with distinct parameters. The conclusion would be that the ACWE model is best suited to deal with such problems as it fits with the formulation of the model. The aim then of this experiment is to measure the compliance of the model with the above anticipations.

For all the three tasks of this experiment, three contour initialisations are used: inside (top rows), across (second rows) and outside (bottom rows) the object.

- The first task is to test the model behaviour when the object is acutely concave. Figure (2.1) shows that this model copes very well with acute corners. The figure also shows that with the choice of suitable value of μ the unnecessary contours are absorbed under the influence of the contour internal force. This is in particular attributed to the effect of elasticity term in removing the unnecessary contours in the rows (b) and (c).

- The next task is to explore the effect of multiple objects on the final segmentation by using this model. Figure (2.2) shows the benefit of using level set representation of the evolving contour to solve the persistent complex parameterisation problem associated with earlier active contours. In the

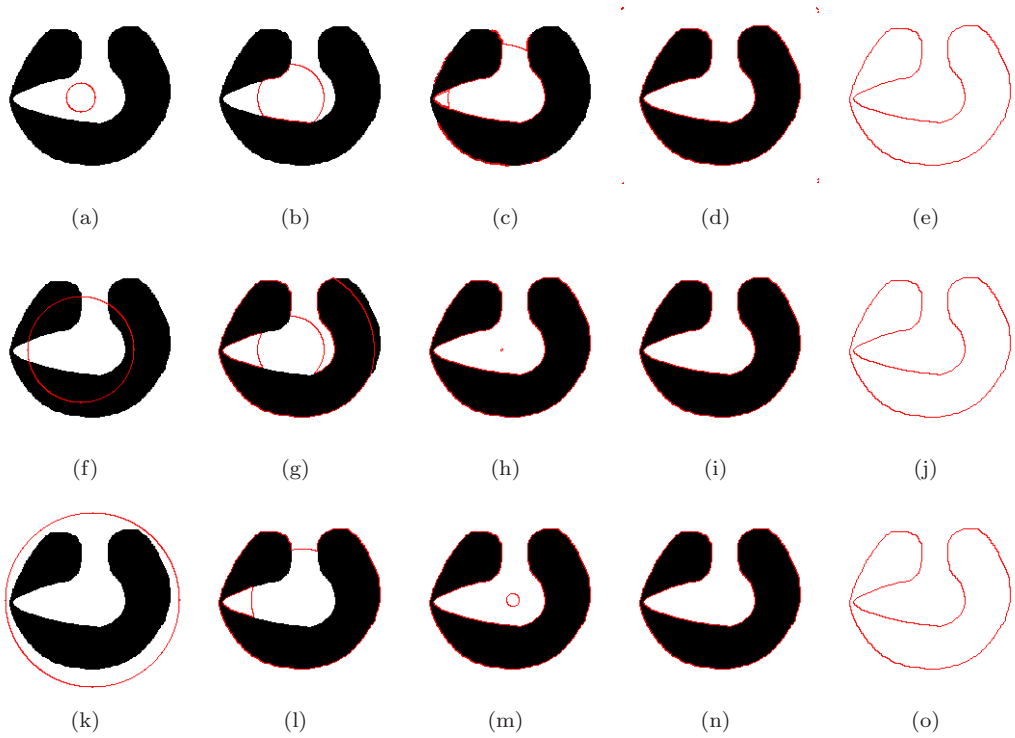


FIGURE 2.1: The segmentation of a shape with acute corners (image taken from Xie and Mirmehdi (2008)) by using ACWE. (a-e) inside initialisation: after 5 iterations with $\triangle t = 50$; (f-j) across initialisation: after 4 iterations with $\triangle t = 50$; (k-o) outside initialisation: after 4 iterations with $\triangle t = 50$

current context the contour merges and splits implicitly as the embedding level set function evolves.

In this example an image containing four discs is segmented. Very interesting to notice here is that the contour splits automatically and the resulting new contours act and evolve independently of each other.

- The last task is aimed to investigate the effect of diffuse or weak edges on this model. This is also one of the points that distinguish this model from the other approaches which are based on edge gradients in the sense that there are no weak regions as opposed to weak edges. The expectation then is that this model always stops at the likely edges of what constitutes a homogeneous (or quasi-homogeneous) region.

  Figure (2.3) shows that the contour stops at a sensible edge, avoiding leakage inside or outside the object and consequently contour collapse. This example demonstrates, clearly, the effectiveness of the region based stopping criterion

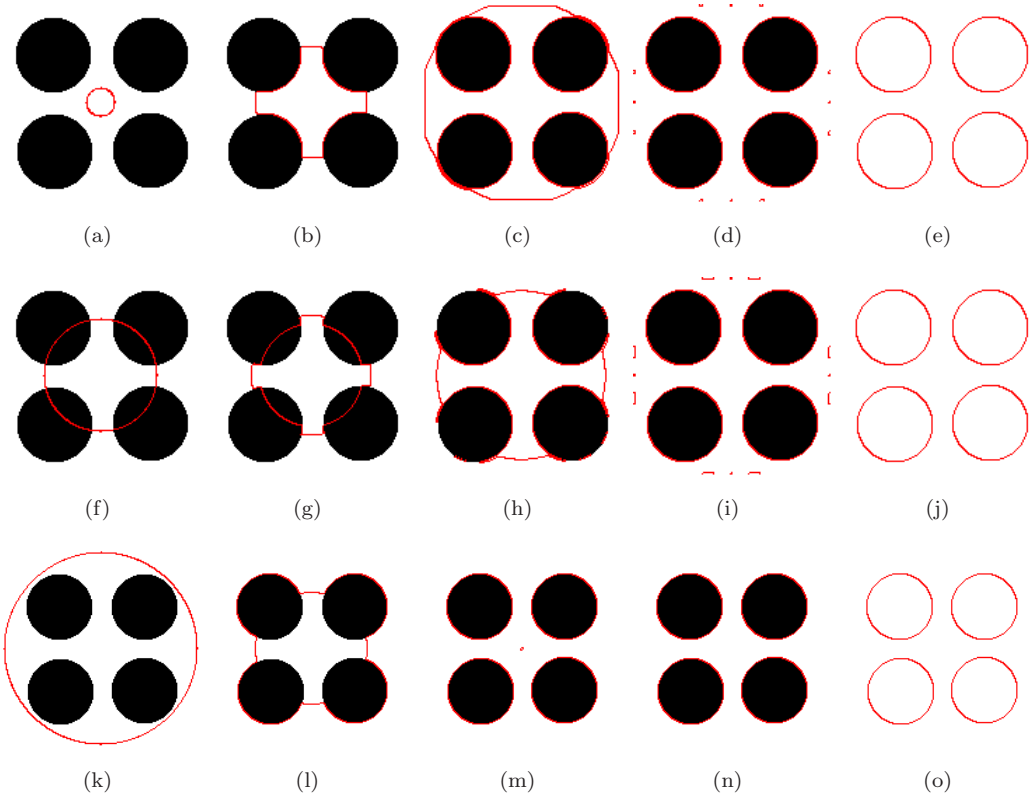


FIGURE 2.2: The segmentation of multiple objects image (image taken from Xie and Mirmehdi (2008)) by using ACWE. (a-e) inside initialisation: after 5 iterations with $\triangle t = 50$; (f-j) across initialisation: after 6 iterations with $\triangle t = 50$; (k-o) outside initialisation: after 6 iterations with $\triangle t = 50$

in holding the contour on the most likely edges. A direct ramification to that can be the deduction that this model with its less localised approach produces a higher level vision in terms of the meaningfulness of the segmented data in comparison to the more classical snakes.

2.4.2 Real World Images

The next set of experiments is conducted to address the sort of problems associated with the segmentation of real-world images, such as the lighting conditions, the presence of shadow, clutter, occlusion and etc.

- The first task is the segmentation of white matter in human brain MRI images. Figure (2.4) shows good results when there is a relatively clear separable regions in the scene. This is a prime example of when this model

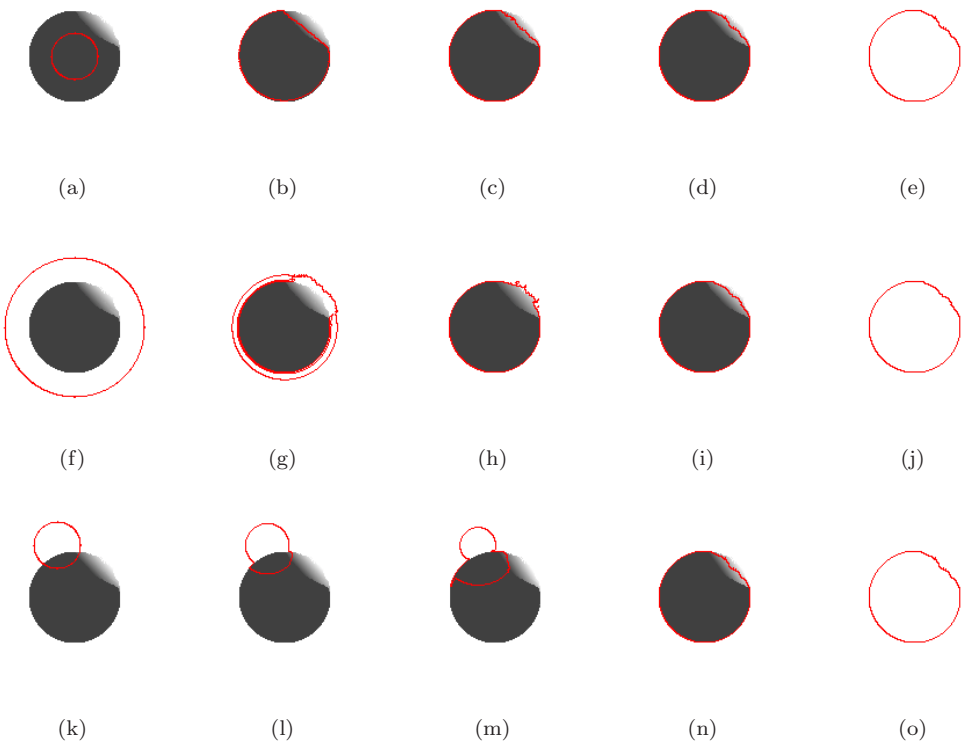
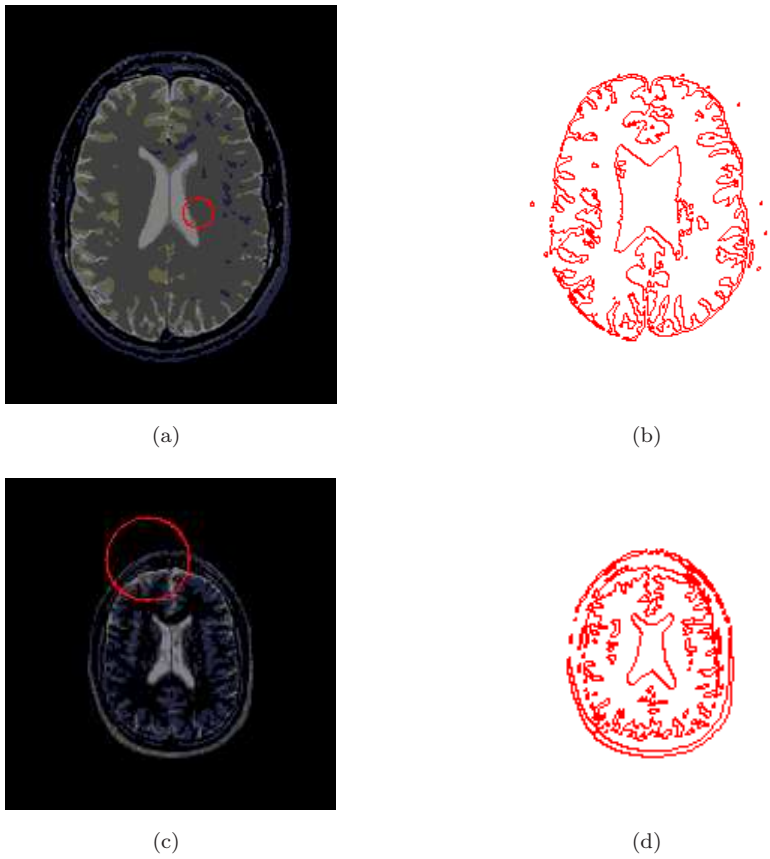


FIGURE 2.3: The segmentation of a shape with diffuse edge (image taken from Xie and Mirmehdi (2008)) by using ACWE. (a-e) inside initialisation: after 5 iterations with $\triangle t = 50$; (f-j) across initialisation: after 7 iterations with $\triangle t = 50$; (k-o) outside initialisation: after 6 iterations with $\triangle t = 50$

works as good with real-world images as it does with the toy examples seen before. For instance, the good handling of corners, the robustness on weak edges and the ease with which the contour changes its topology is repeated in this example.

- The second task is to segment the shape of a walking subject. This example shown by Figure (2.5) demonstrates the difference between the previous problems and the real-world challenges. In the previous examples the enclosing characteristic is that the images have two distinct regions, foreground (object) and background. In the current example the presence of shadow and clutter diminishes the bimodal assumption necessary to separate the image into a foreground and a background.

  Figure (2.5) shows poor outcomes caused directly by the significant regions of shadow and the cluttery patches in the background. These are the regions



(a) (b)

(c) (d)

FIGURE 2.4: The segmentation of white matter in two brain MRI images by using ACWE. (a) Initial contour appearing in the first MRI; (b) the final contour appearing in the first MRI; (c) the original image of the second MRI; (d) the final contour with the second MRI image

where in particular the results are unsatisfactory.

With the increase of scene complexity by incorporating, for example, some sort of noise or occlusion to the image, increases the confusion of the ACWE as how to extract the region of interest. This dilemma is one of the main motives behind incorporating prior knowledge into the process of segmentation. The segmentation of gait with the aid of the shape prior model is discussed in details in Chapter 6.

2.5 Conclusions

In this chapter, the problem of segmentation is introduced along with a well known image segmentation approach, the ACWE. The mathematical formulation of this

(a) (b) (c)

FIGURE 2.5: The segmentation of an image with walking subject in the presence of shadow and clutter. (a) The original image, (b) the initial contour, and (c) the final segmentation contour(s)

model shows that this approach is best suited to segment images with well structured (statistically) foreground and background, or so called, bimodal images.

This model is successfully tested with toy images in classical problems such as initialisation, curvature, object multiplicity, and in cases of weak edges. In all these instances the experimental results showed matching between the outcomes and the expectations. This gives an indication of the potential of this approach.

However, when it comes to segmenting real-world images, and due to the ambiguity caused by conditions like shadow, clutter, noise, and occlusion, this model sees its fundamental assumptions collapse and starts producing incorrect, less meaningful, or meaningless regions.

Chapter 3

Gait Feature Analysis

3.1 Summary

This chapter is devoted to the shape features used to capture the deformation of shapes in a sequence of gait images. The features used here are proposed previously to describe shapes in other problems. Same features are also proposed to model sequential gait shapes. The current framework is more accurate and carries more potential in terms of shape estimation. The experimental results presented here show that this approach generates more accurate reconstructions of the training data. The method is capable of compensating for missing data and can also be exploited in computing in-between frames which bears the potential to deal with other problems like synchronising multiple cameras. This approach is also applied successfully to identify pedestrian subjects based on their gait.

3.2 Introduction

Early work in shape modeling focused on creating a framework capable of generating shapes which are part of the class being modeled (Chen et al., 2002; Bresson et al., 2006). This generative approach received a lot of support from many computer vision applications due to their serious impact on recognition, segmentation,

tracking, and etc. Most successful of shape models, are approaches using statistical techniques. Frameworks built on assumptions that the data in the class being of Gaussian distribution constitutes the majority of such successful models (Cootes et al., 1995; Leventon et al., 2000; Tsai et al., 2003). The main attraction towards such assumptions is the simplicity and the abundance of tools available to analyse Gaussian data.

The modeling of human gait shapes is inherently more challenging, because the deformation of shapes is non-Gaussian (Cremers et al., 2006; Mahmoodi, 2009) and because gait shapes are self-occluding (Nixon and Carter, 2006). The shapes in gait are also periodic, and as such temporally coherent. That's to say shapes of a walking subject are not equally likely at all times (Cremers, 2006). Two main categories of statistical shape models have been suggested and employed so far in order to deal with these issues.

An interesting approach to model deformable moving objects is to use linear PCA, accompanied with some mechanism to synthesise new shapes. Meaning that the PCA will be used to reduce the dimensionality, while the temporally dependent deformation will be guided by this mechanism. In one early novel example, Cremers (2006) has developed a gait segmentation model based on Autoregressive (AR) systems as a mechanism to synthesise new shapes.

Motivated by the above, this chapter describes another approach to shape reconstruction which enjoys the simplicity of Gaussian models, and meanwhile leaves the non-linearity and time coherence in the data to be handled separately by using the interpolation cubic spline. This method is capable of reconstructing moving shapes in image sequences and is successfully used in the estimation accurate shapes in a human gait sequences.

So, in the rest of this chapter, Section (3.3) explains the computation of gait features, Section (3.4) deals with the interpolating cubic spline, the outcomes of this approach are presented and discussed in Section (3.5), and finally, conclusions are drawn in Section (3.6).

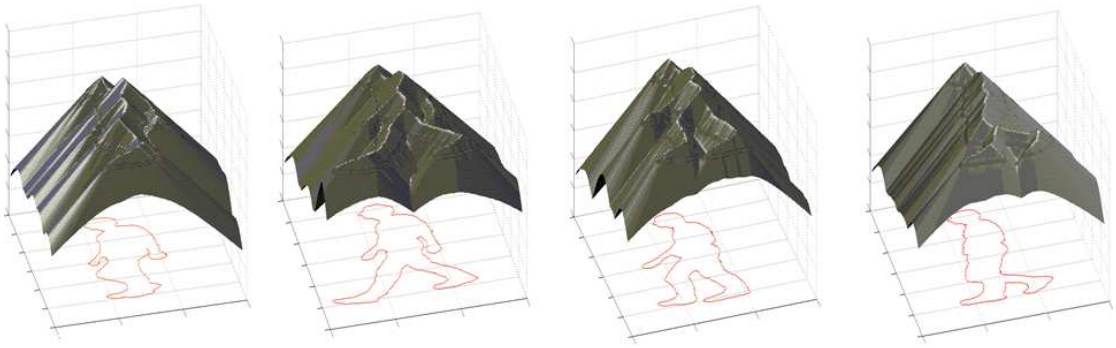


FIGURE 3.1: Example SDFs of four training shapes generated by using fast marching (Sethian, 1996). Also projected beneath them the corresponding contours that represent their zero level sets

3.3 Feature Computation

Following the lead by Leventon et al. (2000), the boundaries of n shapes with $Q = R \times C$ pixels each (R is the number of rows and C is the number of columns), constituting the training set, are embedded as the zero level set of n signed distance functions (SDFs), ϕ_i, by using fast marching (Sethian, 1996) (see Figure 3.1).

Before feature extraction, the shapes need to be aligned. Since the shapes are represented as SDFs (level sets), shape registration algorithm is proposed in Chapter 5 to perform the alignment of SDFs. Having aligned the SDFs, they are augmented into the set $T = \{\phi_1, \phi_2, \cdots, \phi_n\}$.

The mean shape (SDF), $\bar{\mathbf{u}}$ computed as $\bar{\mathbf{u}} = \frac{1}{n}\sum_{i=1}^{n}\phi_i$ is then subtracted from the SDF's ϕ_i to centralise these SDFs. The columns of each of the resulting mean-offset maps $\acute{\phi}$ are stacked lexicographically to form n column vectors, $\vec{u}_i$. These vectors collectively define the shape-variability $(Q \times n)$ matrix $\mathbf{V}$:

$$\mathbf{V} = [\vec{u}_1 \ \vec{u}_2 \ \cdots \ \vec{u}_n]. \tag{3.1}$$

The decomposition of the $Q \times Q$ covariance matrix $\mathbf{V}\mathbf{V}^T/n$ to its eigenvectors Ψ_i and eigenvalues υ_i as in Eq. (3.2) is numerically expensive.

$$\Psi \ \upsilon \ \Psi^T = \frac{1}{Q}\,\mathbf{V}\,\mathbf{V}^T. \tag{3.2}$$

In order to reduce the computational burden, the eigenvectors $\grave{\Psi}_i$ of an $n \times n$ ($n << Q$) kernel, $\frac{1}{n}\ \mathbf{V}^T\mathbf{V}$ are computed as in Eq. (3.3) by using SVD:

$$\grave{\Psi}\ \upsilon\ \grave{\Psi}^T = \frac{1}{n}\ \mathbf{V}^T\ \mathbf{V}. \tag{3.3}$$

The set of eigenvectors of the matrix $\frac{1}{Q}\ \mathbf{V}\mathbf{V}^T$ mentioned in Eq. (3.2) are then computed according to Eq. (3.4):

$$\Psi = \frac{1}{\sqrt{\upsilon}}(\mathbf{V} \times \grave{\Psi}). \tag{3.4}$$

A set of eigenmodes $\vec{\alpha}_i$ (see Figure 3.2(a)) is computed to quantify the contribution of each principal component to a certain shape:

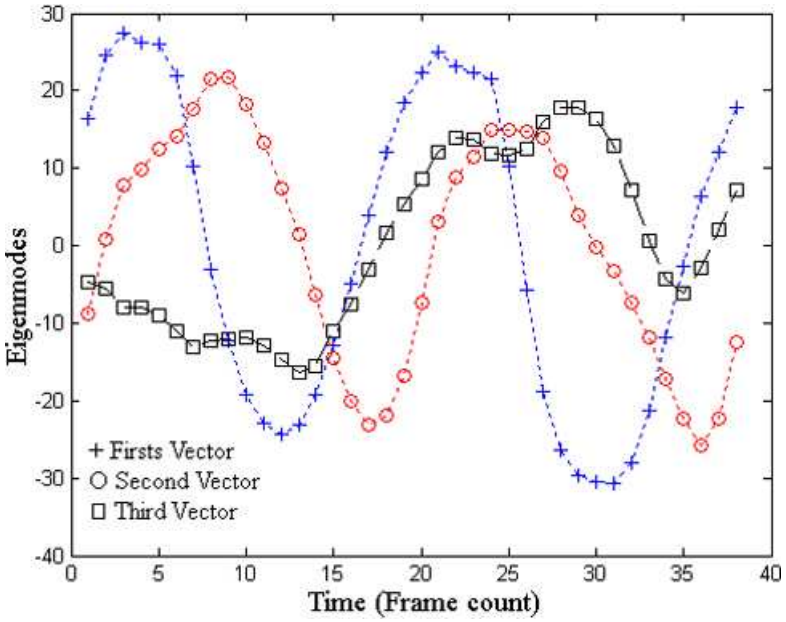
$$\vec{\alpha}_i = \Psi^T\ \vec{u}_i. \tag{3.5}$$

Accordingly an estimated valid shape vector $\hat{\mathbf{u}}_i$ similar to those of the training set $\vec{u}_i$, can be computed by using k ($k < n$) principal components according to Eq. (3.6):

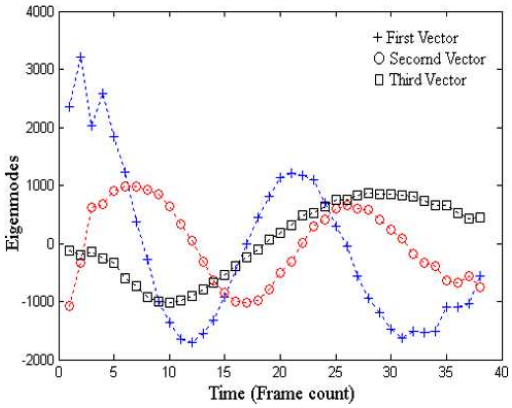
$$\hat{\mathbf{u}}_i = \Psi_k\ \vec{\alpha}_i \tag{3.6}$$

The accuracy of the shape estimate is, obviously, a function of k, and there is a trade off between the accuracy and the computational cost. For most applications the first few components are sufficient and the rest may be regarded redundant. The essence then of the ability of this approach to capture and reconstruct the pattern of the successive shapes of the training set is to properly model the eigenmodes, $\vec{\alpha}$.

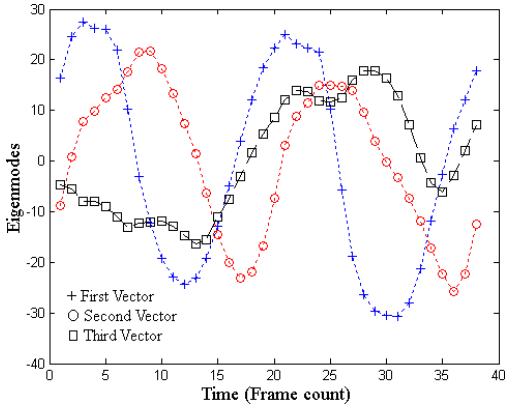
In their novel work, Cremers (2006) suggested the use of AR system over the set of eigenmodes $\vec{\alpha}$ to capture the temporal coherence of the consecutive shapes. The work presented here, however, proposes the use of the interpolating cubic spline to model the behaviour of $\vec{\alpha}$, due to its smoothness and simplicity. Although splines with other orders can be used to model the data, we expect a cubic spline would be



(a) The original eigenmode vectors



(b) The Eigenmodes modeled by AR



(c) The Eigenmodes modeled by the cubic spline

FIGURE 3.2: The first three eigenmode vectors of 38 training SDFs alongside with their synthesised counterparts. (a) Original, (b) generated by AR, and (c) generated by the cubic spline

suitable for modeling $\vec{\alpha}$ due to the non-linearity of these features. The outcomes of both approaches are compared, and further experiments are conducted to test the reliability of the cubic spline employed in this context.

3.4 Interpolating Cubic Spline

The interpolating cubic spline (Press et al., 2007) is a piecewise continuous curve. This curve passes through each of the values of a tabulated function ($y_i = y(x_i),\ i =$

$1 \cdots N$) it is supposed to model. There is a separate cubic spline polynomial for each interval, each with its own coefficients. For a single interval between x_i and x_{i+1} this polynomial is given by Eq. (3.7):

$$y = Ay_i + By_{i+1} + Cy_i'' + Dy_{i+1}'' \qquad \text{for } x \in [x_i,\ x_{i+1}] \tag{3.7}$$

where $A = \frac{x_{i+1}-x}{x_{i+1}-x_i}$, $B = 1 - A$, $C = \frac{1}{6}(A^3 - A)(x_{i+1} - x_i)^2$ and $D = \frac{1}{6}(B^3 - B)(x_{i+1} - x_i)^2$. y_i'' and y_{i+1}'' are the second derivatives of y_i and y_{i+1} respectively. These unknowns are computed by solving $N - 1$ linear equations according to the assumptions in Press et al. (2007). These equations are defined as follows:

$$\frac{x_i - x_{i-1}}{6}y_{i-1}'' + \frac{x_{i+1} - x_{i-1}}{3}y_i'' + \frac{x_{i+1} - x_i}{6}y_{i+1}'' = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} - \frac{y_i - y_{i-1}}{x_i - x_{i-1}}. \tag{3.8}$$

The polynomial in Eq. (3.7) constitutes the piecewise continuous smooth cubic spline which is used here to model the eigenmode vectors $\vec{\alpha}$ into time continuous signals:

$$\vec{\alpha}(t) = cspline(\vec{\alpha}). \tag{3.9}$$

3.5 Experiments and Discussions

The subsequent paragraphs utilise the proposed approach of applying cubic spline explained in Section (3.4) to model and reconstruct the shape eigenmodes mentioned in Section (3.3).

The experiments are aimed at achieving two goals: The first is a comparative application of AR system described in Cremers (2006) and the cubic spline based framework proposed here. The comparison concerns the modeling of the shape eigenmodes and subsequently using these modeled eigenmodes to reconstruct the training shapes.

The second goal includes further tests to verify the robustness of the proposed approach in perceiving the evolution of the implicit shapes. Namely, by examining

its ability to estimate the in-between shapes, by testing its adaptability to the lack of part of the training data, and finally by using the generated model parameters in the recognition of subjects.

A distance term *Dist* is introduced to assess the accuracy of the estimated shape by computing the total number of erroneous pixels in the estimated shape::

$$Dist(\phi_{original}, \phi_{estimated}) = \int_x \int_y |H(\phi_{original}) - H(\phi_{estimated})| \, \mathrm{d}x\mathrm{d}y, \qquad (3.10)$$

where H is the Heaviside function, $\phi_{original}$ is the original shape, defined by the training set, and $\phi_{estimated}$ is its estimated counterpart.

- The first experiment uses a set of shapes derived from a subject silhouette in each of the 38 frames of a video sequence recording their walk. The gait feature formulation (Section (3.3)) is then computed to calculate the eigenmode vectors $\vec{\alpha}_i$ of Eq. (3.5). The AR based model (Cremers, 2006) and the cubic spline based approach (Al-Huseiny et al., 2009) are both then used to model these eigenmodes (see Figure 3.2(b) and 3.2(c)). Next, both approaches are employed to construct an estimated gait sequence of shapes $\hat{\mathbf{u}}_i$ by using Eq. (3.6).

  Figure (3.3) shows *Dist* (Eq. (3.10)) for both approaches. This measure clearly suggests that for cubic spline there are on average 11.1 erroneous pixels per image of 12×10^4 pixels (0.009%) compared to 1802.3 erroneous pixels per the same image (1.5%) for AR.

  Figure (3.4) depicts these results, by showing some of the training sequence shapes alongside with their reconstructed counterparts.

  In the bottom row it is easy to observe a filtering effect imposed by the AR on the time series data. This filtering can be due to the fact that AR system expressed in Cremers (2006) and Neumaier and Schneider (2001) can be considered as a linear regressor (Oppenheim et al., 1999), while obviously, the eigenmodes, $\vec{\alpha}$, being modeled follow a non linear trend. This is reflected

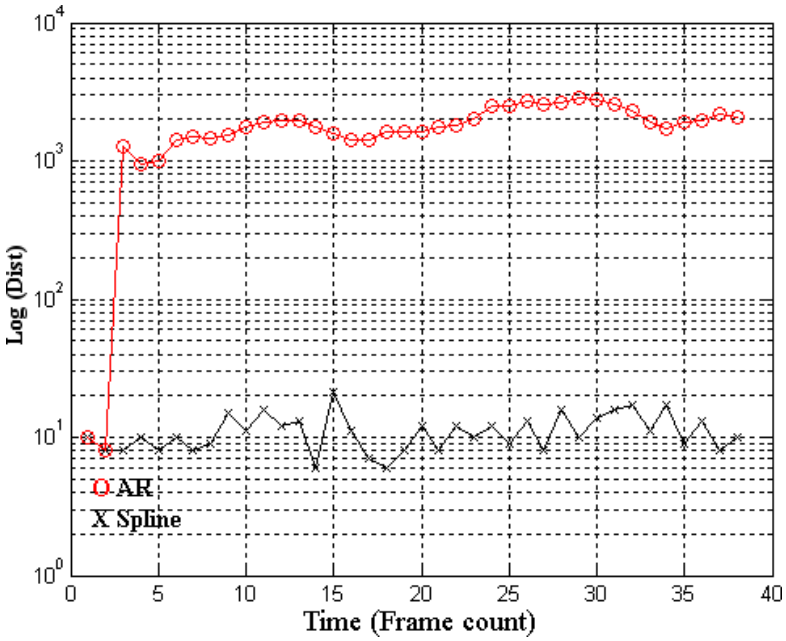


FIGURE 3.3: A semi-log plot of the error function *Dist* computed for the shapes generated by the AR based model (o-labeled) and by the cubic spline based approach (x-labeled)

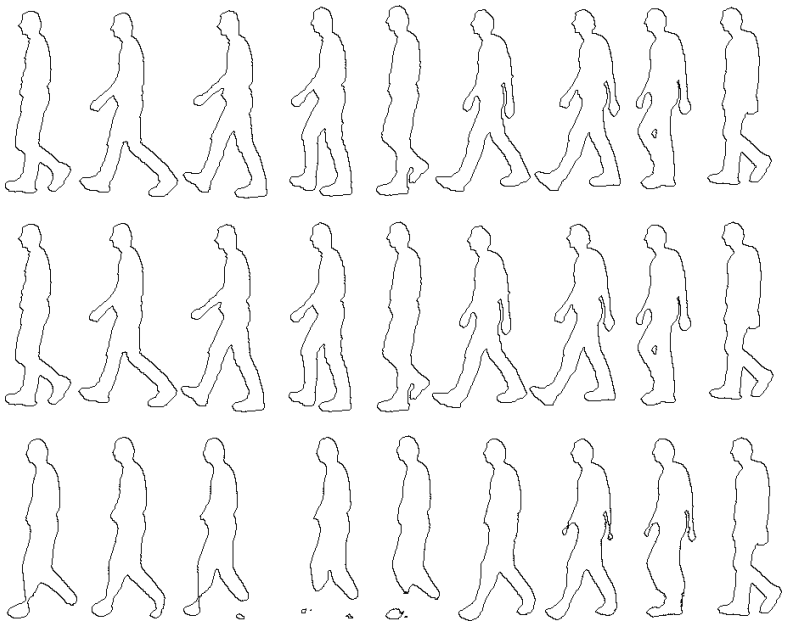


FIGURE 3.4: The estimation of the training shapes. Top row is a sample shapes of the training sequence with order (right to left):1, 5, 10, 13, 21, 26, 28, 33 and 38. Middle row is the same sequence reconstructed by cubic spline based regression. Bottom row is the sequence reconstructed by using AR based regression

on the reconstructed shapes in the form of smoothening or even erasing the thin parts of the shapes such as the hands. This also, in some cases, results in invalid shapes like having twisted hands or feet.

The middle row of Figure (3.4) shows the shapes reconstructed by the cubic spline based regressor which look very similar to training set. This is also consistent with the outcomes of the measure *Dist* seen in Figure (3.3).

One final point worth mentioning is that AR is (by nature) not self-starting, i.e. it depends on initial condition data that must be provided prior to the start of the reconstruction operation. When and if such data is not available or inaccurate the subsequent reconstruction becomes poor. Putting this in the context of segmentation, this means that the segmentation of the first few frames must rely on an alternative technique. Therefore, the overall segmentation depends on the accuracy of that alternative segmentation technique, this is of course in addition to the genuine shortcoming of the AR based regression demonstrated by the results above. Comparing this with the cubic spline based regression, which demonstrated excellent reconstruction results and is self starting data sequence reconstruction technique, the argument may be strongly held towards using this approach proposed here to model the eigenmodes $\vec{\alpha}$.

- In the second experiment, the proposed technique is tested for its capability to estimate unknown transitional shapes in between the successive training shapes, i.e. to up-sample the training data set.

  This has potential application in synchronising multiple cameras (Prismall et al., 2003). This is quite important because a large number of cameras placed in different locations, working with different sampling rates, produce unsynchronised footages. Up-sampling can numerically synchronise the captured frames by reconstructing in-between frames.

  Figure (3.5) exhibits the subtle movements in the contours of shapes estimated by the approach presented here by using two different sampling rates.

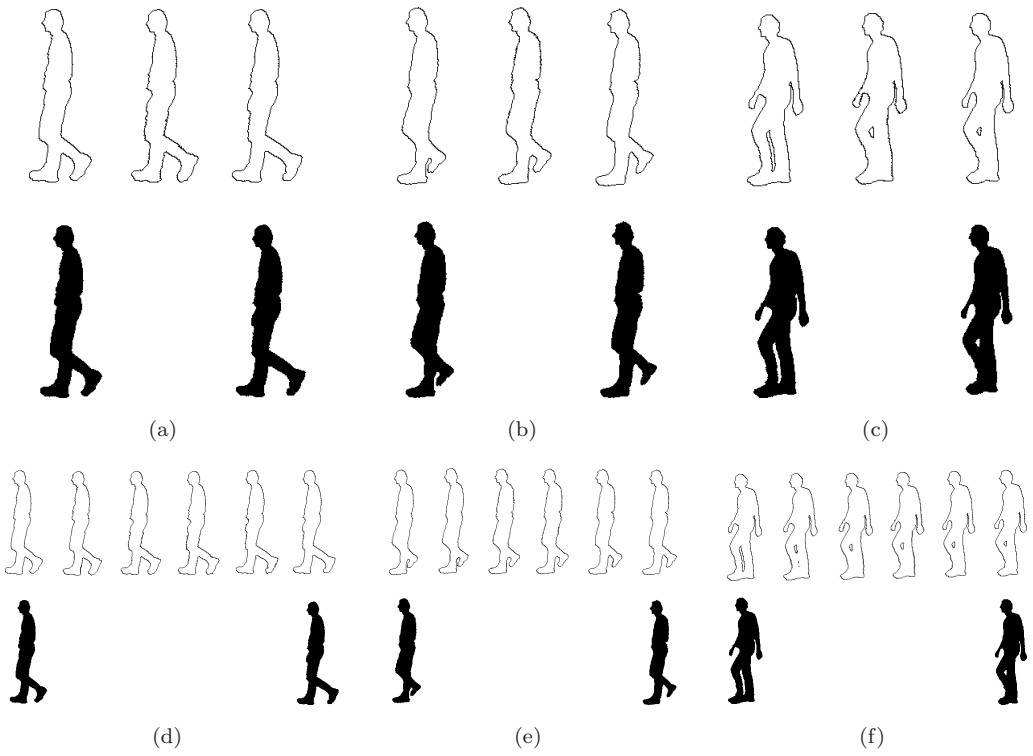
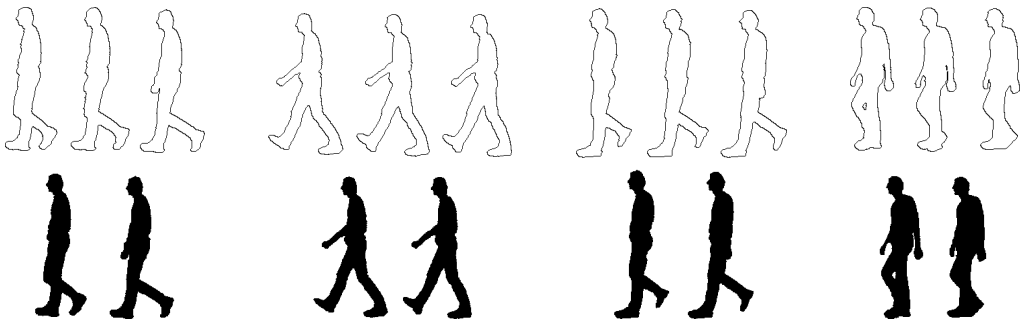


FIGURE 3.5: The up-sampling of the input gait cycle with two step sizes: (a) shapes 37-38 with step size 0.5, (b) shapes 20-21 with step size 0.5, (c) shapes 5-6 with step size 0.5, (d) shapes 37-38 with step size 0.2, (e) shapes 20-21 with step size 0.2, (f) shapes 5-6 with step size 0.2. In all cases, the bottom rows show couples of consecutive training sequence shapes, the top rows are their up-sampled counterparts

- In the third experiment the proposed approach is tested for its capability to compensate for missing data in the training set. This is achieved by applying the leave one out test, where each time one of the shapes of the training cycle is removed and the remaining shapes are used for training the algorithm. The algorithm is then used to estimate the whole gait sequence including the missing shape. This is repeated to all of the cycle shapes.

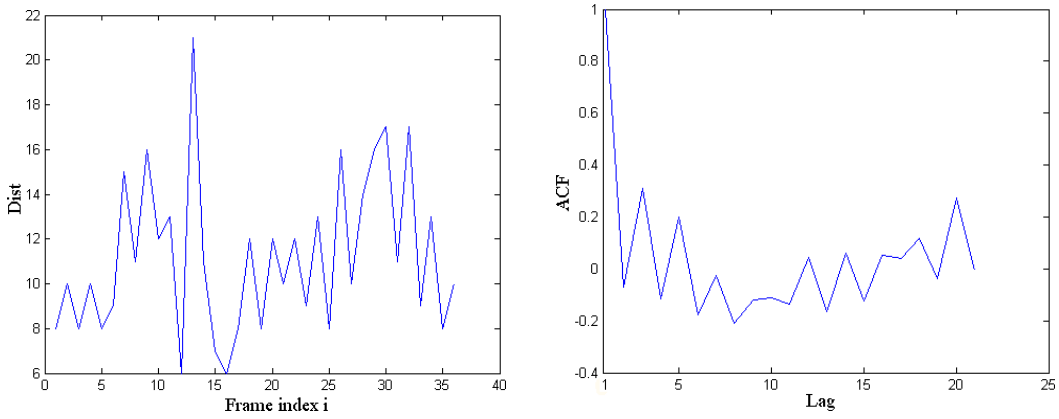
  Figure (3.6) shows examples of this test, in which the missing frames have been successfully reproduced.

  The error function $Dist$ of Eq. (3.10) is used to quantify this test. From Figure (3.7(a)) it can be seen that on average there are 11.5 erroneous pixels per image of 12×10^4 pixels (0.01%). The autocorrelation function (ACF) is also computed for $Dist$, which is shown in Figure (3.7(b)). The shape



(a) 36-38 (b) 29-31 (c) 17-19 (d) 3-5

FIGURE 3.6: Leave one out test. (a) estimating shape 37, (b) estimating shape 30, (c) estimating shape 18, and (d) estimating shape 4. In all cases the bottom rows show the two training sequence shapes before and after the removed one. The top rows are the contours of the reconstructed shapes including the missing one



(a) (b)

FIGURE 3.7: Error Analysis of the leave one out experiment for the estimated 38 shapes: (a) The error function, $Dist$. (b) The autocorrelation function ACF of the error function $Dist$

of ACF confirms that error in the estimated shapes is of random nature (random noise).

- In the fourth experiment, the proposed approach is used for the identification of walking subjects. This is based on the assumption that the way different persons walk induces different deformability coefficients, i.e., $\vec{\alpha}$.

  Hence, the gait cycles of four different subjects are used in producing four different sets of $\vec{\alpha}$ regarded as reference. For one of those subjects a new gait cycle different from the reference one used earlier is employed as a test cycle. The features $\vec{\alpha}$ corresponding to this test cycle are also computed.

TABLE 3.1: The identification of pedestrians by using d between the unknown and the reference cycles

	Distance d to the test cycle
1^{st} subject	28992
2^{nd} subject	25158
3^{rd} subject	26722
4^{th} ***subject***	*7868*

A distance measure $d = \sum_i \int_\tau |\vec{\alpha}_{1,i}(t) - \vec{\alpha}_{2,i}(t)|^2 \, \mathrm{d}t$ is used, where i is the index over the eigenmodes for each subject, τ is a single period of gait cycle, $\vec{\alpha}_{1,i}$ and $\vec{\alpha}_{2,i}$ are the i^{th} eigenmode vectors of the gait cycles of two subjects to whom the distance d is computed.

The empirical results show that this distance can be reasonably used to identify the the correct subject that has least distance to the test cycle. Table 3.1 shows that the cycle of the correct subject (the fourth) has the least distance to the unknown cycle.

3.6 Conclusions

This chapter introduces an interpolating cubic spline to better model walking subjects. This is done by modeling the time variation of shapes captured by the eigenmodes over a single gait cycle. This approach demonstrates improved performance and accuracy over the autoregressive system AR used in the literature for the same purpose. The technique presented here succeeded in capturing the key variability modes which led to success in the reconstruction of walking cycle shapes identical to the training set.

This method presented here is also used successfully in reconstructing the in-between frames which did not exist in the initial training set. Such result can be employed, for example, in numerically synchronising multiple cameras, an application which can prove vital with the expansion in the use monitoring cameras. The

method proposed here is also tested for its tolerance to missing parts of the training set for which the technique proves robust. Furthermore the present technique is employed in the identification of subjects by using the eigenmodes.

This approach, according to the presented results, provides a good answer in the case of modeling and regenerating the shapes of a single gait cycle. However, the expansion of the accuracy and generalisability of this approach by incorporating the capacity to produce entirely new cycles in addition to the current results of producing new shapes within a given cycle is further pursued in Chapter 4.

Chapter 4

Gait Generative Model

4.1 Summary

This Chapter proposes a learning method for gait synthesis from a sequence of shapes(frames) with the ability to extrapolate to novel data. It involves the application of PCA, first to reduce the data dimensionality to certain features, and second to model corresponding features derived from the training gait cycles as a Gaussian distribution. This approach transforms a non-Gaussian shape deformation problem into a Gaussian one by considering features of entire gait cycles as vectors in a Gaussian space. It is shown here that these features which we formulate as continuous functions can be modeled by PCA. This model is also used to in-between (generate intermediate unknown) shapes in the training cycle. Furthermore, the subsequent sections demonstrates that the derived features can be used rigorously in the identification of pedestrians.

4.2 Introduction

The construction of shape models has been an area of active research. Many applications share the need for good dependable shape models which should enjoy two

distinctive features: to be flexible enough to extrapolate beyond learning material, while being accurate enough to generate plausible estimates. Most prominent approaches in this context are the point distribution model (Cootes et al., 1995) which uses a Gaussian model for a set of corresponding points, later, the model by Leventon et al. (2000) which acts on signed distance functions (SDF's), a notion that avoids a strict point correspondence requirement. The common theme between these approaches is that the data they deal with, has in general a Gaussian distribution.

The issue with gait, as will be explained later in this chapter, is that the distribution of the human silhouette's deformation is non Gaussian (Cremers et al., 2006; Mahmoodi, 2009) (see Figures 1.2 and 1.3). Also a gait cycle is used as a sequence of shapes. Therefore a particular deformation has to be drawn over the whole sequence in order to produce a consistent gait cycle, i.e., the estimated cycle must show shapes belonging to the same subject at all times.

This chapter proposes a new method to model shape deformations of the entire gait sequences rather than individual shapes. Putting the problem in this form seems practical, since gait shapes are usually used in the context as complete cycles rather than particular shapes. This can be used to generate prior shapes in a prior shapes segmentation framework, and hence clean data is used here to train the model.

The proposed model is based on using PCA twice, first, to reduce the data dimensionality and extract the shapes' eigenmodes which are used to synchronise the different cycles. The second is used to model the deformation of these eigenmodes with a Gaussian distribution. We notice that the deformation of shapes in a gait cycle along the time axis is non Gaussian. However it is assumed here that the deformation of shapes over various gait cycles at a certain time is Gaussian (see Figure (1.3)). A set of Gaussianity tests is presented in Section 4.3 to support these assumptions.

So in the rest of this Chapter, the shape of the distribution of the features is discussed in Section 4.3, the structure of the proposed model is described in Section 4.4, the experimental outcomes are presented in Section 4.5. This Chapter concludes in Section 4.6.

4.3 Feature Distribution Analysis

It has been mentioned in several places throughout this thesis that the distribution of the gait data along the time axis is non-Gaussian. Furthermore, the proposed gait shape model detailed in the next section is based on the assumption that equivalent data at the same time instant from multiple subjects can be approximated by a Gaussian model. This section presents a set of tests to support these assumptions.

A set of 117 videos for walking subjects is used to compute the features, $\vec{\alpha}$ (see Section 3.3 for details). The $\vec{\alpha}$'s are then interpolated in order to align them and establish correspondence between features from various videos.

In this context we use graphical techniques such as histograms and normal probability plot (NPP) (Chambers et al., 1983), as well as, formal tests such as probability plot correlation coefficients (PPCC) (NIST/SEMATECH, 2011) to verify the shape of the distribution of the data.

Normal probability plot is a graphical technique for assessing whether or not a data set is approximately normally distributed. For data that are approximately Gaussian, NPP follows a linear order.

In PPCC, the test statistic is the correlation coefficient of the points that make up a NPP. This is expressed as a function of lambda, the shape parameter (the shape of the distribution). This test statistic is compared with tables of critical value (NIST/SEMATECH, 2011). If the test statistic is less than the tabulated value, the null hypothesis that the data came from a population with a normal distribution is rejected.

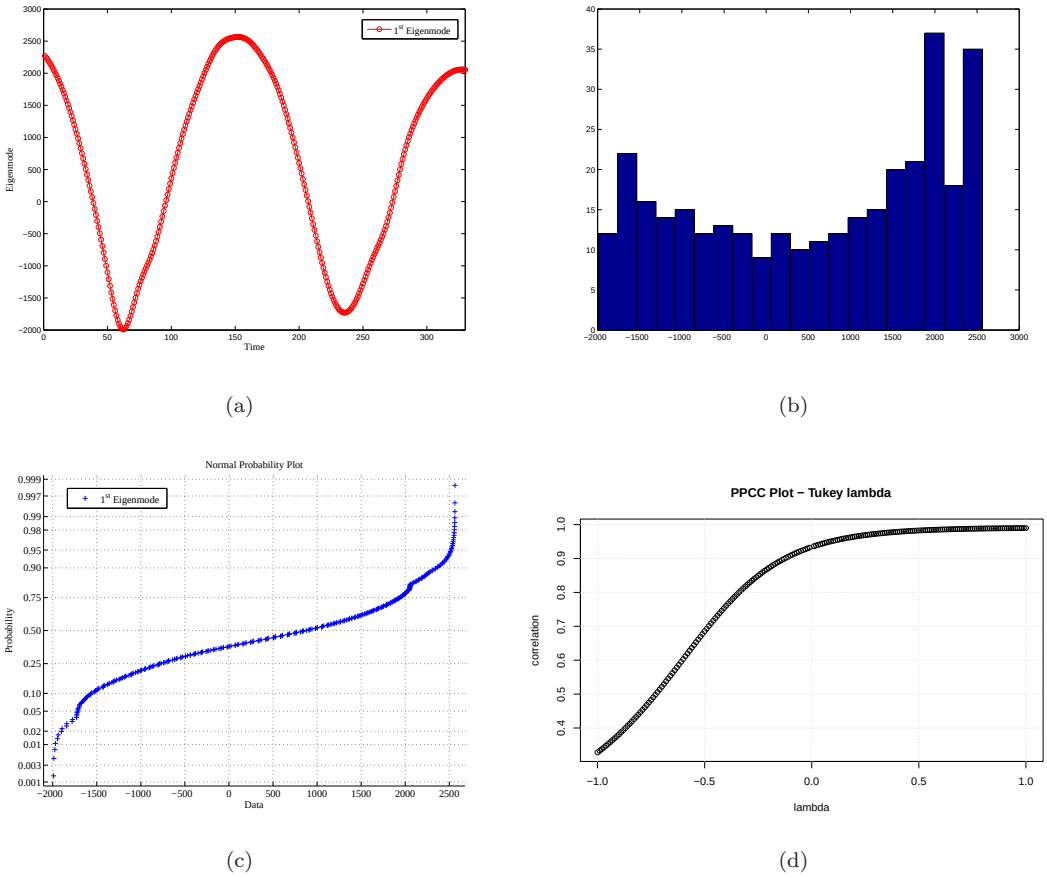


FIGURE 4.1: Gaussianity tests for $\vec{\alpha}_1$, (a) $\vec{\alpha}_1$, (b) histogram of $\vec{\alpha}_1$, (c) NPP, (d) PPCC plot as a function of lambda, PPCC=0.9898 at lambda=1 (Uniform)

4.3.1 Non-Gaussianity over time:

In order to assess the first assumption, that is, the distribution of gait data along time (i.e. shapes from a gait cycle of the same subject) is non Gaussian, the first ten eigenmodes $\vec{\alpha}_{1:10}$, which carry most of the information are analysed.

In Figure (4.1-a) it is noticed that $\vec{\alpha}_1$ does not follow a random pattern and there is a periodic behaviour in the data. This periodicity is a clear signal that the data cannot be approximated by a Gaussian distribution because of the randomness requirement of the central limit theorem. Also from the histogram of $\vec{\alpha}_1$ it is easy to tell that although there is a reasonable form of symmetry, the shape of the data is not normal. The NPP as shown in this figure is not linear. This is also confirmed by PPCC (computed by using the tool in Wessa (2011)) which peaks at lambda=1, this in turn means that this data can be approximated by a uniform

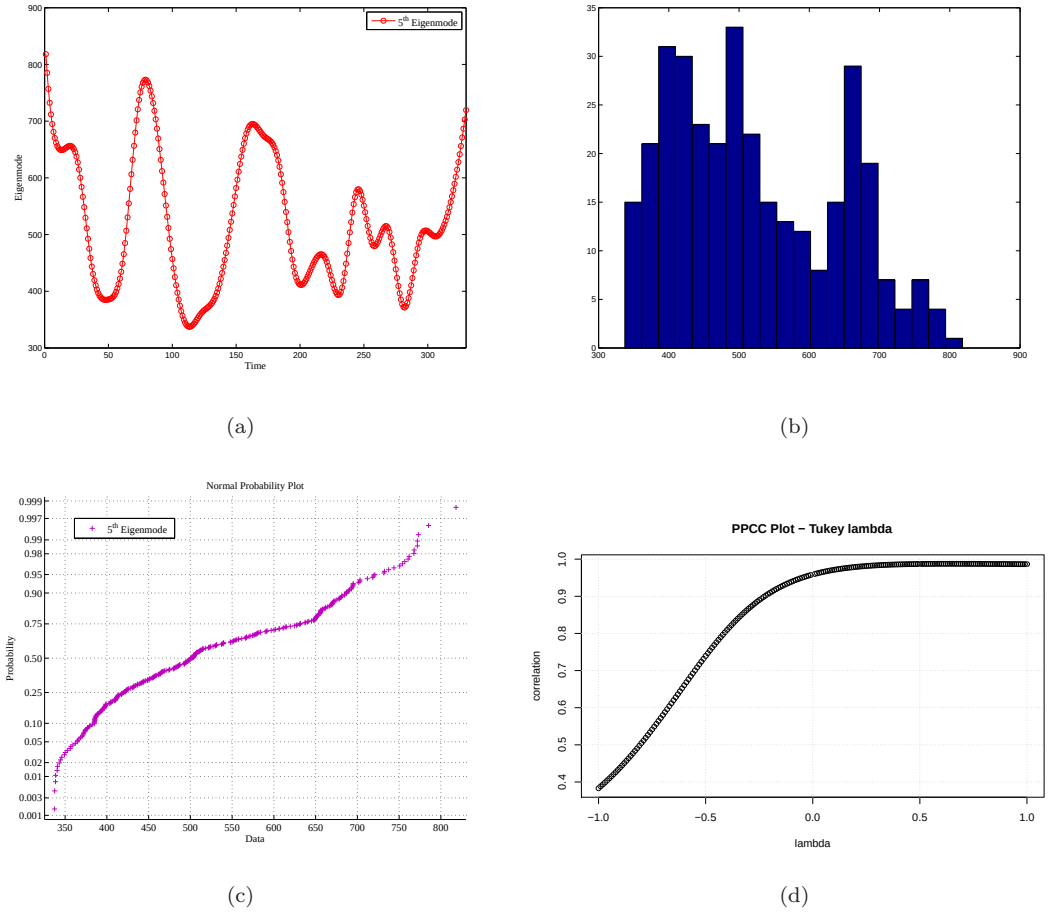


FIGURE 4.2: Gaussianity tests for $\vec{\alpha}_5$, (a) $\vec{\alpha}_5$, (b) histogram of $\vec{\alpha}_5$, (c) NPP, (d) PPCC, PPCC=0.9870 at lambda=0.5 (U-shaped)

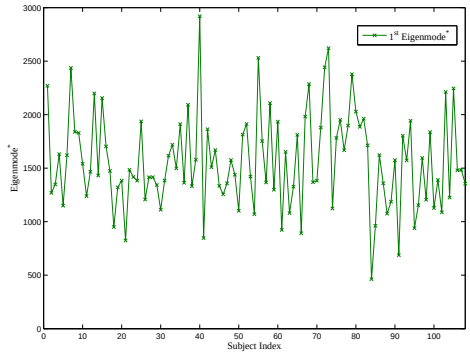
distribution (NIST/SEMATECH, 2011). Figure (4.2) shows similar results for the fifth eigenmode, $\vec{\alpha}_5$.

4.3.2 Gaussianity over Subjects:

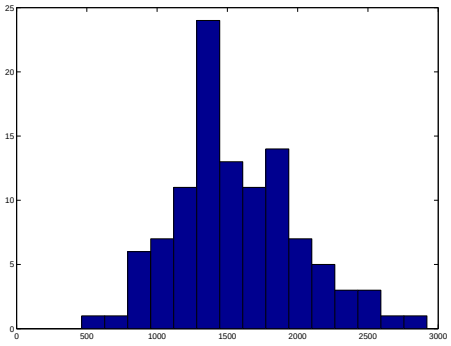
The second assumption which states a normal distribution over corresponding instances from multiple sources of gait features is tested here. Ten vectors representing the first instances of the first ten eigenmodes, $\vec{\alpha}_{1:10}$, over the 117 subjects are formed, these are called here $Eigenmodes^*_{1:10}$. $Eigenmodes^*_1$, for example, is formed of the first element of each of the 117 $\vec{\alpha}_1$'s, to constitute the variation modes of shapes from different sources at the same time instant (see Figure (4.3).

$$(Eigenmodes_1^*)^T \Leftarrow \begin{bmatrix} \boldsymbol{\alpha}_{1,\mathbf{1}} & \alpha_{1,2} & \alpha_{1,3} & \cdots & \alpha_{1,\tau} \\ \boldsymbol{\alpha}_{2,\mathbf{1}} & \alpha_{2,2} & \alpha_{2,3} & \cdots & \alpha_{1,\tau} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\alpha}_{117,\mathbf{1}} & \alpha_{117,2} & \alpha_{117,3} & \cdots & \alpha_{117,\tau} \end{bmatrix}$$

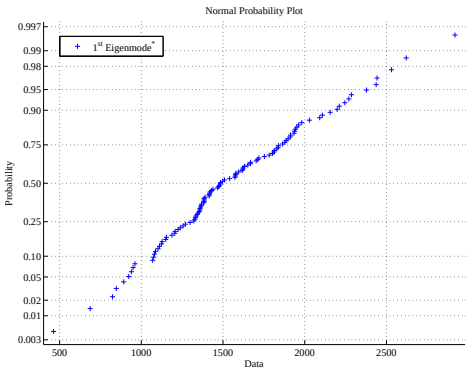
FIGURE 4.3: A matrix of the first eigenmodes, $\vec{\alpha}_1$, of 117 subjects over time τ, the first entry represents the subject index and the second entry represents the time index, i.e., each row is a single $\vec{\alpha}_1$ vector from a different subject. $Eigenmodes^*$'s are mapped from this matrix at the same time instant over all subjects, therefore, $Eigenmodes_1^*$ receives the bold α's



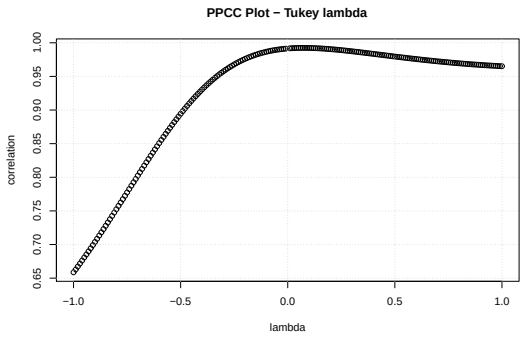
(a)



(b)



(c)



(d)

FIGURE 4.4: Gaussianity tests for $Eigenmode_1^*$, (a) $Eigenmode_1^*$, (b) histogram of $Eigenmode_1^*$, (c) NPP, (d) PPCC, PPCC=0.9918 at lambda=0.14 (Normal)

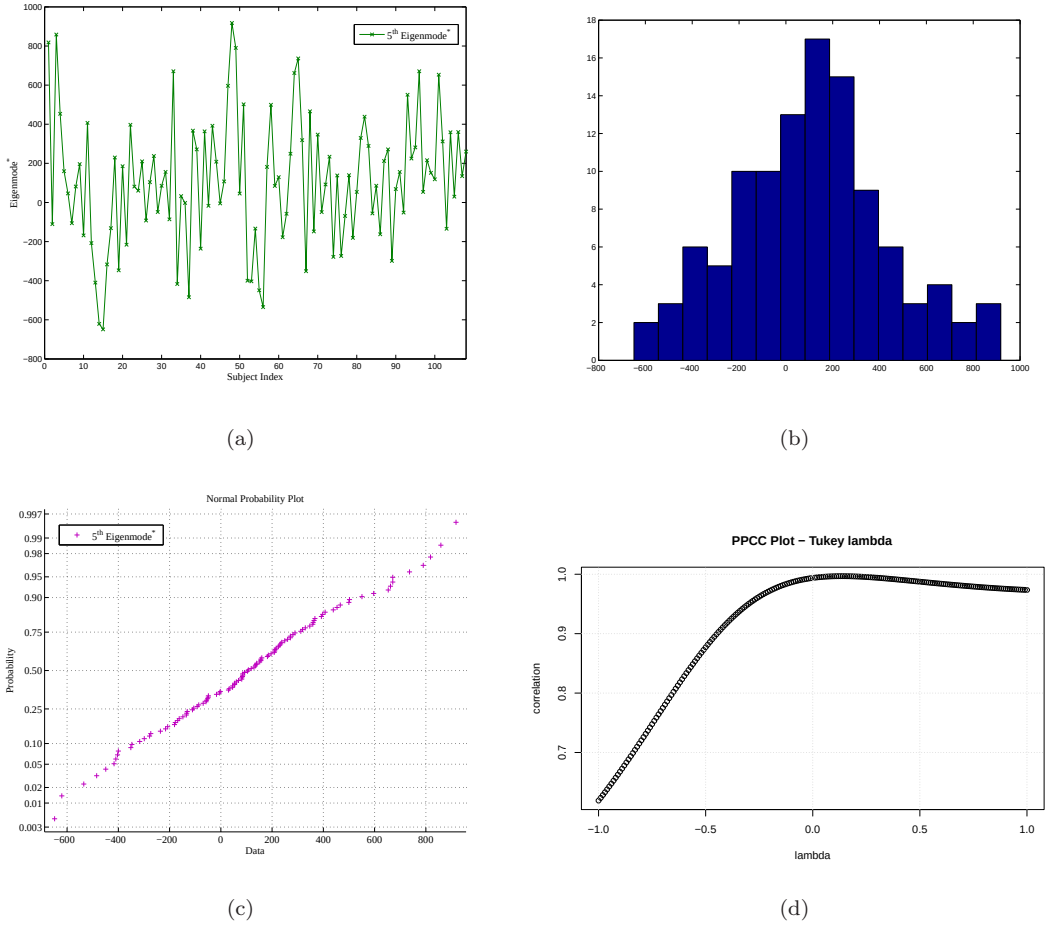


FIGURE 4.5: Gaussianity tests for $Eigenmode_5^*$, (a) $Eigenmode_5^*$, (b) histogram of $Eigenmode_5^*$, (c) NPP, (d) PPCC , PPCC=0.9967 at lambda=0.14 (Normal)

Figure (4.4) demonstrates the results of the above described tests for the $Eigenmode_1^*$. In this figure, it is clear that the data resembles random drawings, also the histogram shows reasonable symmetry with overall shape close to normal. The normal probability plot is very close to linearity. This plot indicates that the data represented by $Eigenmode^*$ is approximately Gaussian. Furthermore, PPCC peaks at 0.14 which verifies that a normal distribution is suitable for approximating the data.

From Figure (4.4-(d)), the maximum value of PPCC is 0.9918, and, from the tables in NIST/SEMATECH (2011), at the 5% significance level, the critical value is 0.9881. Since 0.9918 is greater than 0.9881, we cannot reject the null hypothesis that the data came from a population with a normal distribution. Figure (4.5) shows similar test results for $Eigenmode_5^*$.

These tests show that eigenmodes of the shapes of the gait cycles, $\vec{\alpha}$ (along time axis), do not obey a normal distribution. Also shown here that the corresponding shapes from different gait cycles (subjects) can be approximated by a Gaussian distribution.

Figure (4.6) summarises these findings by showing the NPPs for the first ten eigenmodes, $\vec{\alpha}_{1:10}$, and the first ten over-subject vectors, $Eigenmodes^{*}_{1:10}$. On the basis of these results, a gait shape model is proposed in the next section. This framework uses PCA to model corresponding data from multiple sources.

4.4 Gait Generative Model

In this section we describe the proposed framework to model a set of gait cycles. The boundaries of the training data consisting of M gait cycles with different number of shapes per cycle, are embedded as the zero level sets of N SDF's, ϕ_i by using fast marching (Sethian, 1996), where N is the total number of shapes in the training set. These SDFs are aligned to each other by using the algorithm presented in Chapter 5.

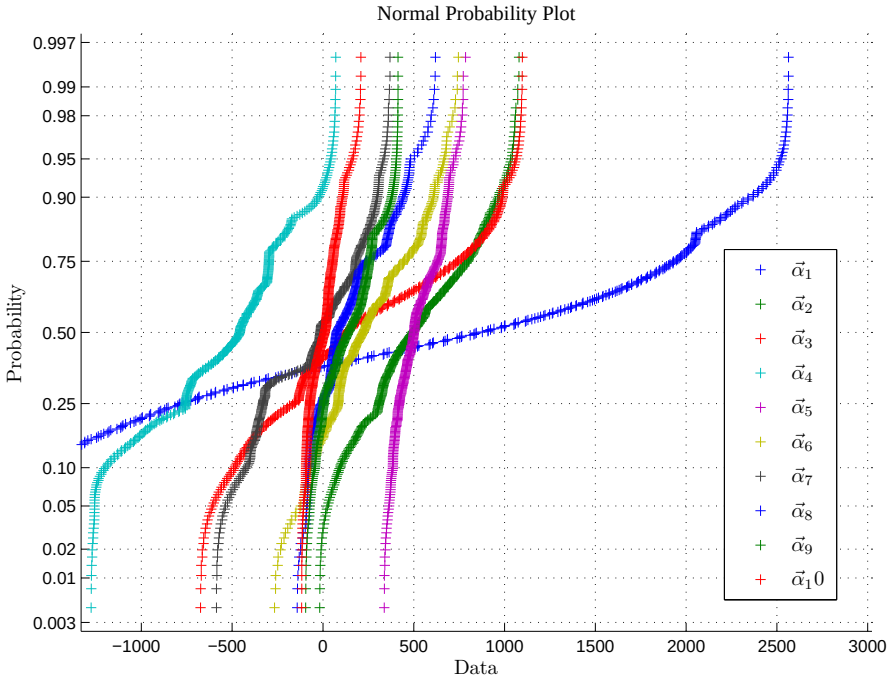
A mean shape is computed as $\bar{\mathbf{u}} = \frac{1}{N}\sum_{i=1}^{N}\phi_i$, this shape is then subtracted from the shapes to centralise the data, and the resulting distance maps $\acute{\phi}$ are then vectorised and augmented into the matrix $\mathbf{S}$ of shape vectors $\vec{u}_i$.

$$\mathbf{S}_{n\times N} \equiv \left[\vec{u}_1^1, \ldots, \vec{u}_l^1, \ldots, \ldots, \vec{u}_1^M, \ldots, \vec{u}_p^M\right], \tag{4.1}$$

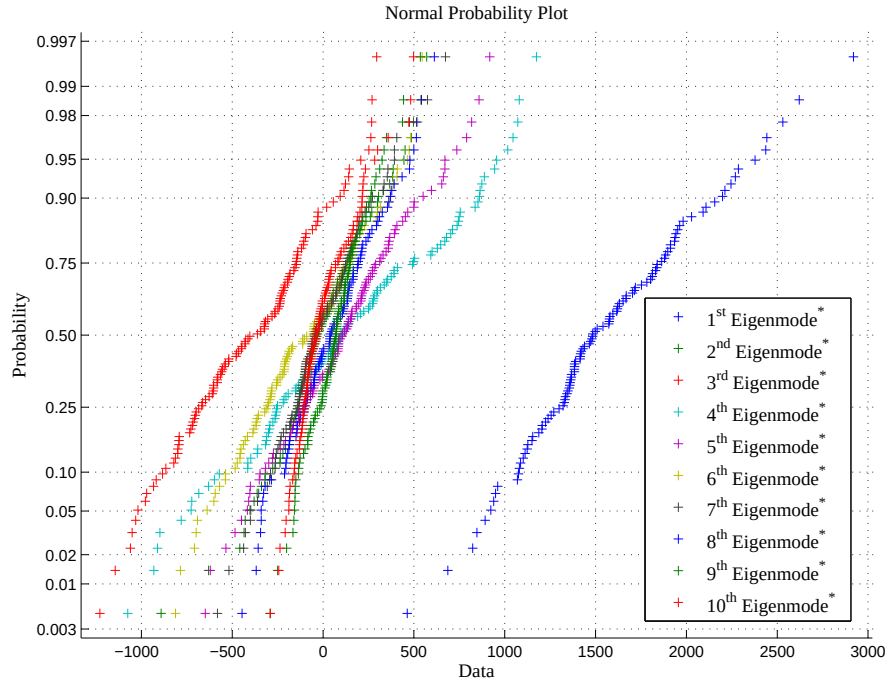
the superscripts refer to the particular gait cycle, and the subscripts refer to a shape within a cycle.

$\mathbf{S}$, is then subjected to the first principal component decomposition, PCA_{basis}, to generate a common eigenvector basis Ψ for the entire data set,

$$\Psi\ \lambda\ \Psi^T = \left(\mathbf{S}\ \mathbf{S}^T\right) \times 1/N. \tag{4.2}$$



(a)



(b)

FIGURE 4.6: Comparison of the normal probability plots, (a) NPPs for the first ten vectors (over-time), $\vec{\alpha}_{1:10}$, (b) NPPs for the first ten vectors (over-subject), $Eigenmode^*_{1:10}$

PCA_{basis} is crucial, because the projection of the shapes on the common Ψ produces eigenmodes $\vec{\alpha}$ corresponding to appropriate shapes,

$$\vec{\alpha}_i = \Psi^T \vec{u}_i. \tag{4.3}$$

Each cycle is now represented by a set of shapes' eigenmodes. The variations of these vectors (along the subject axis not the time axis), we argue, belong to a Gaussian distribution, and therefore, a second decomposition, $\text{PCA}_{feature}$, is applied to calculate the eigenvectors associated with these vectors.

The issue is that these eigenmode vectors, if considered as sampled periodic signals, are of different lengths and in different phases. A Hermite cubic spline (section 3.4) is therefore used to represent the underlying continuous function for each vector,

$$\alpha(t) = cspline(\vec{\alpha}), \tag{4.4}$$

these continuous shape eigenmodes $\alpha(t)$ are aligned and augmented as the columns of a zero-mean feature matrix $\mathbf{F}$,

$$\mathbf{F} \equiv [\alpha_1(t), \ldots, \alpha_M(t)]_{\infty \times M}. \tag{4.5}$$

The feature covariance matrix $\mathbf{C}_{\infty\times\infty} = \mathbf{F}\ \mathbf{F}^T$, has infinite dimensions and the decomposition of its eigenvector components is not numerically tractable. A finite dimensional kernel $\mathbf{K}$ is therefore defined as

$$\mathbf{K}_{M\times M} = \mathbf{F}^T\ \mathbf{F}. \tag{4.6}$$

Since the vectors of $\mathbf{F}$ are continuous functions, then the components of $\mathbf{K}$ can be computed by the following integration:

$$k_{i,j} = \int_\tau \alpha_i(t)\,\alpha_j(t)\,\mathrm{d}t. \tag{4.7}$$

Now, the eigenvector decomposition of $\mathbf{K}$ gives:

$$X\ \eta\ X^T = \mathbf{K}. \tag{4.8}$$

The eigenvectors $\chi(t)$, of the matrix $\mathbf{C}$ are found (Cootes et al., 1995),

$$\chi = \mathbf{F}\ X. \tag{4.9}$$

Equations (4.4-4.9) show that PCA can be applied to continuous functions as effectively as with discrete data (A numerical example is provided in Appendix D for further clarification). In other words, $\text{PCA}_{feature}$ places an instantaneous Gaussian model across the continuous functions at every time instant. The coefficients, $\vec{\beta}$, corresponding to the eigenmodes of a cycle, i, are computed as:

$$\vec{\beta}^i = \chi^T\ \left(\vec{\alpha}^i - \bar{\alpha}\right) \tag{4.10}$$

where $\bar{\alpha}$ is the mean set of eigenmodes computed as $\bar{\alpha} = \frac{1}{M}\sum_i^M \vec{\alpha}^i$. These computed values, $\vec{\beta}^i$, can only be used to recall the eigenmodes $\vec{\alpha}^i$ of the training set, a generalisation into estimating novel data, $\hat{\alpha}$, can be investigated by setting new values, $\hat{\beta}$, different from the computed $\vec{\beta}$'s. A set of eigenmodes $\hat{\alpha}(t)$ is then computed by using:

$$\hat{\alpha} = \chi\ \hat{\beta}\ +\ \bar{\alpha}. \tag{4.11}$$

These eigenmodes $\hat{\alpha}(t)$ are then translated into an estimated novel gait cycle of vectorised shapes by Eq. (4.12),

$$\hat{\mathbf{u}}_t = \Psi\ \hat{\alpha}(t), \tag{4.12}$$

where Ψ and $\bar{\mathbf{u}}$ are respectively the eigenvector basis and the mean shape of PCA_{basis}.

4.5 Evaluation

We have used the training cycles from the Southampton Gait database (Shutler et al., 2002). The images are initially segmented manually and their SDF's are then generated (Sethian, 1996). It is worth mentioning here that all images involved in the computation of this model are assumed to be generated under similar conditions (fixed lateral view).

4.5.1 Novel Data Generation

The main contribution of this work is that it consistently facilitates the manipulation of non Gaussian data in a linear fashion. That is to say, the generation of a novel cycle is achieved by adding a linear combination of the eigenvectors to the mean cycle of shapes. Hence, to test this idea, the model is trained by using 20 gait cycles of different subjects with a varying number of frames for each cycle. Then by using Eq. (4.11), three new cycles (shown in Figure 4.7) are generated by assigning 3 different values to the first element of $\hat{\beta}$ and setting the rest to zero. Each value set to $\hat{\beta}_1$ produces a different set of eigenmodes $\hat{\alpha}$, which in turn, by using Eq. (4.12) generate a new cycle of gait shapes.

For the sake of measurement, a simple L^2-norm distance D is computed between the set of eigenmodes $\vec{\alpha}^i$ for each of the training cycles and the mean set of eigenmodes $\bar{\alpha}$. This measure is intended to show the significance of the computed $\hat{\alpha}$'s away from $\bar{\alpha}$ compared to the distance of $\vec{\alpha}$'s of the training data from $\bar{\alpha}$ for the chosen $\hat{\beta}_1$.

The distances D corresponding to the chosen coefficients $\hat{\beta}_1$ (normalised by the 1^{st} eigenvalue) are shown in Figure 4.8. This indicates that changing one coefficient $\hat{\beta}_1$, produces a significantly new set of eigenmodes $\hat{\alpha}$, which gives rise to a novel gait cycle. Visually, it is easy to see from Figure 4.7 the effect of changing $\hat{\beta}_1$ in producing new sequences appearing in the rows (b-d).

FIGURE 4.7: The generation of novel gait cycles by using chosen values of $\hat{\beta}_1$ (normalised by the first eigenvalue). Row (a) is a cycle computed by using the mean eigenmodes $\bar{\alpha}$ by setting $\hat{\beta} = 0$. The rows (b-d) are novel cycles computed by using $\hat{\beta}_1 = 0.2291, \hat{\beta}_1 = -0.6872$, and $\hat{\beta}_1 = 1.3745$ respectively

Important to check is that all the shapes in a generated cycle are for the same subject in different walking positions. This is of course generated by altering one parameter. Including, and subsequently, altering more parameters adds more fidelity to the computed shapes.

4.5.2 Reconstruction: Leave One Out

The reconstruction by regenerating sequences of human walking subjects from the gait training data is assessed in this experiment. Leave one out test has proved to be a useful method to judge the reconstruction accuracy. Accordingly, for one of the training gait cycles, i, a shape, j, is removed each time, the remaining shapes of this cycle with the other 19 cycles together constituting the training set are used to train the model. We then reconstruct the i^{th} cycle by using Eq. (4.11) and

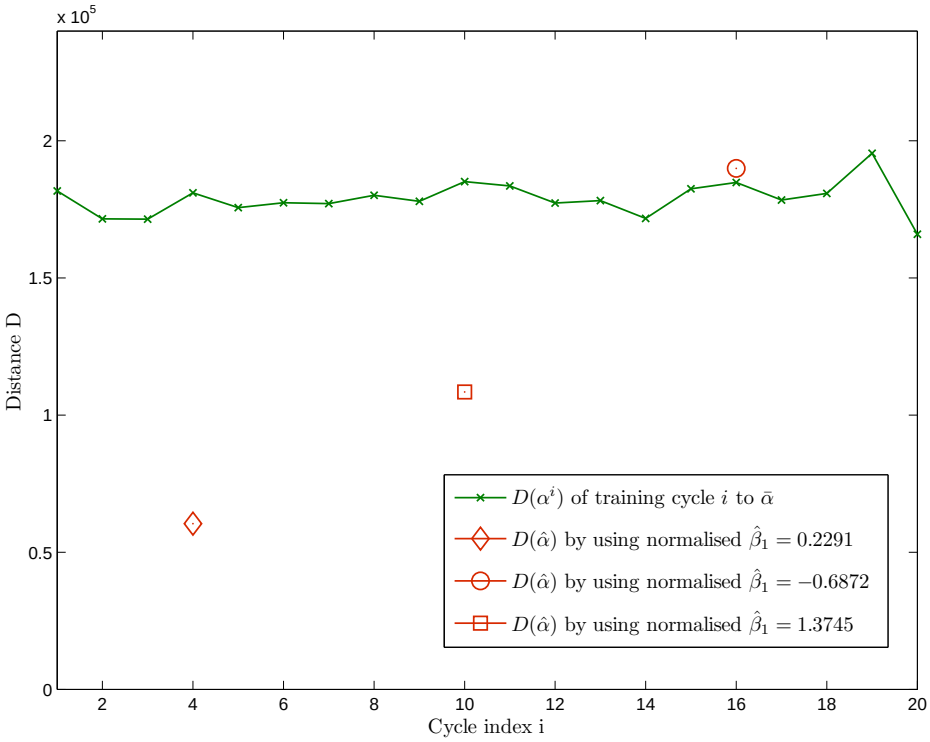


FIGURE 4.8: The distance D computed from the mean cycle for the eigenmodes $\vec{\alpha}^i$ of the training set, and the three computed sets of novel eigenmode $\hat{\alpha}$ generated by setting three new distinct values to $\hat{\beta}_1$

(4.12) by setting $\hat{\beta}$ to $\vec{\beta}^i$ corresponding to the cycle i. The reconstructed shapes are those used to train the model as well as the one removed.

This model proves robust against over-fitting and learns the underlying trend of the data, which in practice led to the generation of a good estimate to the missing shape. An error function Er between the removed shapes and their reconstructed estimates is computed as:

$$Er(j) = \sqrt{tr\left(\left(\phi_{rmv}(j) - \phi_{est}(j)\right)^T \left(\phi_{rmv}(j) - \phi_{est}(j)\right)\right)}, \tag{4.13}$$

where ϕ_{rmv} is the removed shape and ϕ_{est} is the estimated shape. This norm counts the number of erroneous elements in the estimate compared to the reference shape. Figure 4.9 compares the error function $Er1$ computed by using Al-Huseiny et al. (2009), and the error function $Er2$ computed by using this model to reconstruct the same cycle.

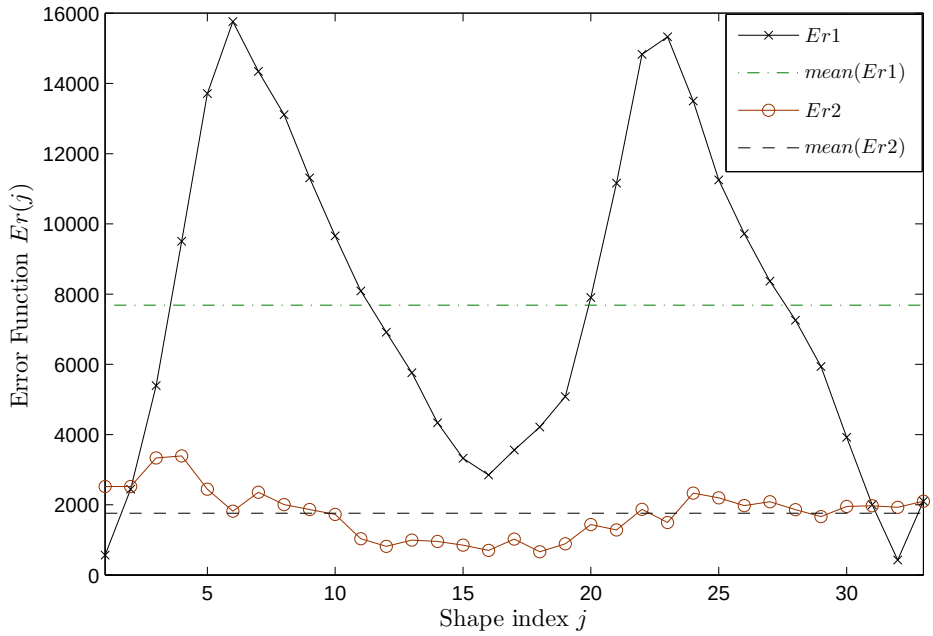


FIGURE 4.9: A comparison of the error function Er in the estimated shapes, $Er1$ computed for the shapes estimated by using (Al-Huseiny et al., 2009), $Er2$ computed for the shapes estimated by using the proposed model

It is noted that the average error is four times less by using the model proposed in this chapter. The reason for this is that in this model the missing shape is estimated by combining the effort of the instantaneous Gaussian contributed by $\text{PCA}_{feature}$ with the general trend of the data captured by the cubic spline, while in Chapter 3 (also in Al-Huseiny et al. (2009)), the estimation is based purely on the capturing of general trend of the data. Therefore, in the case of missing part of the training data, the instantaneous Gaussian model drives the production of a valid shape, i.e., a shape that belongs to the class of shapes at that specific time instant (specific minute action). The data (subject) specific parameters drives the generated shape closer to the look of the subject.

4.5.3 Pedestrian Identification

The shape eigenmodes generated by this model can be employed in the identification of pedestrians based on their gait by using the following theorem.

Theorem 1: In the model proposed here:

$$d = \left\|\mathbf{V}^1 - \mathbf{V}^2\right\|^2 = \sum_i \left\|\vec{\alpha}_i^1 - \vec{\alpha}_i^2\right\|^2, \tag{4.14}$$

where $\mathbf{V}^1$ and $\mathbf{V}^2$ are the vectorised SDF's of the gait cycles for two subjects and $\vec{\alpha}^1$ and $\vec{\alpha}^2$ are their corresponding eigenmodes.

Proof: Let $\mathbf{V}^1$ and $\mathbf{V}^2$ be the gait cycles for two subjects with R shapes each, such that:

$$\mathbf{V}^k = \bar{\mathbf{v}}^k + \sum_{i=1}^{R} \Psi_i^k \vec{\alpha}_i^k, \tag{4.15}$$

where $k = 1, 2$; then the distance, d, between the two cycles is computed by:

$$d = \|\mathbf{V}^1 - \mathbf{V}^2\|^2 \tag{4.16}$$

$$= \left\| \bar{\mathbf{v}}^1 + \sum_i \Psi_i^1 \vec{\alpha}_i^1 - \bar{\mathbf{v}}^2 - \sum_j \Psi_j^2 \vec{\alpha}_j^2 \right\|^2. \tag{4.17}$$

Since all the shapes in the proposed model are represented by using a common mean shape $\bar{\mathbf{u}}$ and eigenvectors Ψ, then $\bar{\mathbf{v}}^1 = \bar{\mathbf{v}}^2 = \bar{\mathbf{u}}$, and $\Psi^1 = \Psi^2 = \Psi$ and hence,

$$d = \left[\sum_i \Psi_i \vec{\alpha}_i^1 - \sum_j \Psi_j \vec{\alpha}_j^2\right]^T \left[\sum_i \Psi_i \vec{\alpha}_i^1 - \sum_j \Psi_j \vec{\alpha}_j^2\right] \tag{4.18}$$

and since $\Psi_i^T \Psi_j = \delta_{ij}$, where δ_{ij} is the *Kronecker* delta,

$$d = \sum_i |\vec{\alpha}_i^1|^2 - 2\vec{\alpha}_i^1 \vec{\alpha}_i^2 + |\vec{\alpha}_i^2|^2 = \sum_{i=1}^{R} \left\|\vec{\alpha}_i^1 - \vec{\alpha}_i^2\right\|^2 \quad \square \tag{4.19}$$

This distance consumes parameters generated by the PCA_{basis} only. Alternatively, theorem 2 shows that the full model provides simpler features for identification by using gait:

Theorem 2: In the model proposed here:

$$\hat{d} = \frac{1}{\tau}\int_{\tau} \left\|\mathbf{V}^1 - \mathbf{V}^2\right\|^2 \mathrm{d}t = \left\|\vec{\beta}^1 - \vec{\beta}^2\right\|^2, \tag{4.20}$$

where $\vec{\beta}^1$ and $\vec{\beta}^2$ are the cycle features for the first and the second subject respectively.

Proof: Let $\mathbf{V}^1$ and $\mathbf{V}^2$ be the gait cycles for two subjects over period τ, such that:

$$\mathbf{V}^k(t) = \bar{\mathbf{u}} + \Psi\ \alpha^k(t), \tag{4.21}$$

Eq. (4.21) can be written as:

$$\begin{aligned} \mathbf{V}^k(t) &= \bar{\mathbf{u}} + \Psi(\chi(t)\ \vec{\beta}^k + \bar{\alpha}(t)), \\ &= \bar{\mathbf{u}} + \Psi\ \chi(t)\ \vec{\beta}^k + \Psi\bar{\alpha}(t), \end{aligned} \tag{4.22}$$

Therefore, the distance, $\hat{d}$, is computed as:

$$\begin{aligned} \hat{d} &= \int_{\tau} \|\mathbf{V}^1 - \mathbf{V}^2\|^2 \mathrm{d}t \\ &= \int_{\tau} \left\|\bar{\mathbf{u}} + \Psi\ \chi(t)\vec{\beta}^1 + \Psi\bar{\alpha}(t) - \bar{\mathbf{u}} - \Psi\ \chi(t)\vec{\beta}^2 - \Psi\bar{\alpha}(t)\right\|^2 \mathrm{d}t, \\ &= \int_{\tau} \left\|(\Psi\chi(t))(\vec{\beta}^1 - \vec{\beta}^2)\right\|^2 \mathrm{d}t, \\ &= \int_{\tau} (\vec{\beta}^1 - \vec{\beta}^2)^T\ \chi(t)^T\ \Psi^T\ \Psi\ \chi(t)\ (\vec{\beta}^1 - \vec{\beta}^2)\mathrm{d}t, \\ &= \int_{\tau} (\vec{\beta}^1 - \vec{\beta}^2)^T\ (\vec{\beta}^1 - \vec{\beta}^2)\mathrm{d}t, \\ &= \tau \left\|\vec{\beta}^1 - \vec{\beta}^2\right\|^2, \end{aligned} \tag{4.23}$$

that is because $\Psi^T\Psi$ and $\chi^T\chi$ reduce to unit matrices (due to orthogonality).

The distance $\hat{d}$, then, becomes:

$$\hat{d} = \frac{1}{\tau}\int_{\tau} \left\|\mathbf{V}^1(t) - \mathbf{V}^2(t)\right\|^2 \mathrm{d}t = \left\|\vec{\beta}^1 - \vec{\beta}^2\right\|^2. \quad \square \tag{4.24}$$

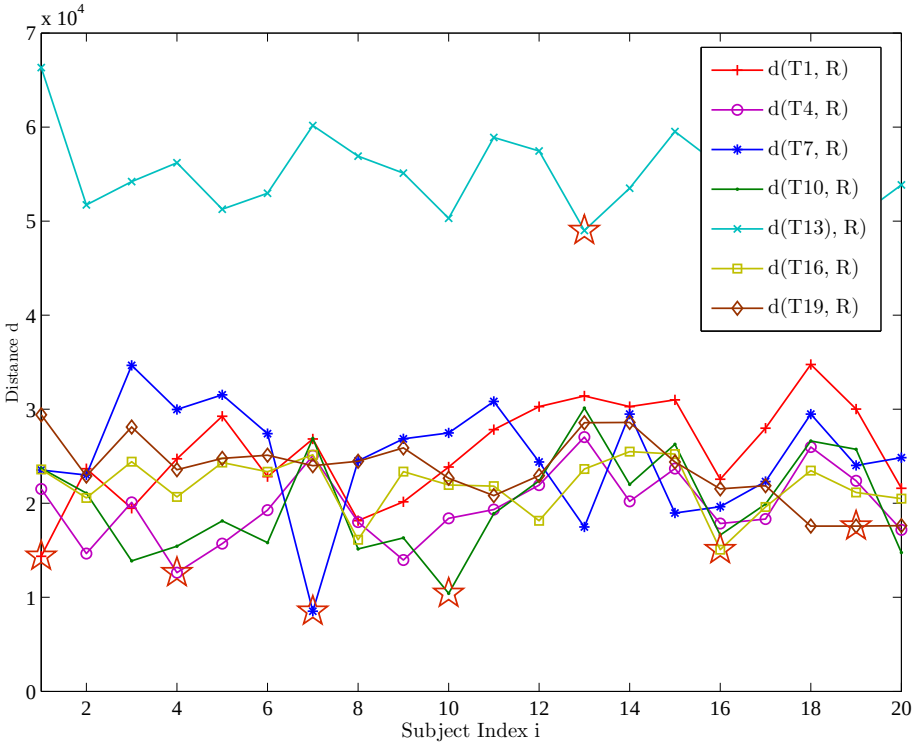


FIGURE 4.10: The identification of pedestrians by using their corresponding shape eigenmodes. The distance d measured between the $\vec{\alpha}$'s of the set of reference cycles R and those of the i^{th} test cycle, Ti

This theorem shows that the subject features, $\vec{\beta}$, can be used to identify individuals by their gait. A classification experiment based on these findings is given in section 7.2.

The distance, d, of Eq. (4.19) is used in the identification to measure the closest training cycle to an unknown cycle. So, for each of the subjects in the training set, a new unknown cycle (not in the training set) is used as the test cycle, the distance d is computed between the eigenmodes of these test cycles (referred to as T) and the eigenmodes of the known 20 reference cycles of the training set (referred to as R). In all of the 20 cases, d returns the lowest value for the correct cycle. Figure (4.10) shows the outcomes for a selected set of this experiment. The minimum value of d for each round is indicated by a star.

While this experiment shows good results in the identification of individuals by using the shape eigenmodes, $\vec{\alpha}$, the cycle parameters, $\vec{\beta}$, are more suitable (as seen

in Section 7.2) for such tasks due to the fact that $\vec{\beta}$ has much less dimensionality than $\vec{\alpha}$.

4.6 Conclusions

In this chapter a gait model is proposed. This model transforms the problem of non Gaussian shape deformation, into a Gaussian one by considering entire gait cycles as training data points (vectors). These points are modeled together by using PCA. This approach is applied to human gait analysis and synthesis, though it could be applied to similar problems elsewhere.

This approach preserves, and meanwhile exploits the time coherence of the shapes in the gait. This is important in applications such as tracking and prior shape based segmentation particularly in the presence of occlusion and with noisy images.

The interesting employment of PCA to deal with continuous functions makes it possible to align the gait cycles in the feature space. This made the case for accurate automatic shape correspondence.

The introduction of PCA over corresponding shapes results in producing entire novel cycles with valid shapes. The generation of new cycles is governed by changing the model coefficients which is a significant practical outcome of this model.

The proposed model is also used successfully in reconstructing the training data. Furthermore it proved more robust in the estimation of in-between shapes compared to the previous approaches. These results presented in this chapter demonstrate that the proposed method enjoys the flexibility of the statistical methods without over-fitting the known sample.

In addition to the above mentioned applications, the model proposed here, has showed and proved that the statistical shape eigenmodes can be used in pedestrian identification. This route can be extended into a gait biometric by applying this model to a large database (a preliminary example is given in section 7.2).

Unlike previous approaches there would be no need for computing complex shape descriptors such as the statistical moments in order to provide the common basis for measurement.

Chapter 5

Shape Registration

5.1 Summary

This chapter presents a fast algorithm for robust registration of shapes implicitly represented by signed distance functions (SDFs). The proposed algorithm aims to recover the transformation parameters (scaling, rotation, and translation) by minimising the dissimilarity between two shapes. To achieve a robust and fast algorithm, linear orthogonal transformations are employed to minimise the dissimilarity measures. The algorithm is based on phase correlation and statistical shape moments to compute the registration parameters individually. The algorithm proposed here is applied to various registration problems, to address issues such as registration of shapes with various topologies, registration of complex shapes containing various numbers of sub-shapes, applicability to databases, accuracy, convergence speed and stability. The outcomes are compared with other state-of-the-art shape registration algorithms and collectively suggest the advantages of the new technique.

5.2 Introduction

Shape registration can be viewed as the result of a point-wise transformation between an observation and a reference shape. It is a fundamental task used to match two or more shapes taken, for example, at different times, from different viewpoints, or from different scenes. Most image analysis systems which evaluate images from various sources require the registration or a closely related operation as an intermediate step (Brown, 1992)-(Zitova and Flusser, 2003). Shape registration is an essential requirement shared among many computer vision domains and applications, such as, pattern recognition, remote sensing, computer graphics, and medical image analysis to name a few.

Biomedical imaging, for instance, is a vital component of a large number of applications, which occur throughout the clinical track of events (Maintz and Viergever, 1998). Reliable matching and eventually integration of information from different sources demonstrate its importance from the fact that knowledge contained in an image is complemented by the knowledge contained in another image as a reference in the same medical setting.

The main motive to address the issue of shape registration (represented by SDFs) in this thesis is to build an integrated framework for prior guided gait segmentation. In such a framework, shapes produced by the model are usually normalised in terms of pose, size, and position of the shape. In order to be usable in a distance measure (within the segmentation framework), the shape needs to be aligned to the evolving contour. Existing approaches are either based on contours which later requires computing their corresponding SDFs to fit in the framework, or, are iterative and hence are slow and lack robustness. Furthermore, the construction of the shape model necessitate that the training shapes be registered to each other. As will be shown in this chapter, the proposed registration algorithm performs better than the existing approaches and integrates easily in the segmentation framework.

The construction of a registration method requires the harmonisation of the problem requirements with the choice of four components that form the registration

algorithm (Brown, 1992):

1. The feature space: this can be either the entire image (area-based or correlation like techniques) or selected features (feature based techniques) such as control points, line intersections, shape contours, and etc.

2. The search space: this has to do with the type of transformation the shapes are anticipated to have undergone such as rigid (translation and rotation), similarity (translation rotation and scale), affine (parallel lines remain parallel) and etc.

3. The search strategy: this represents the approach used to evolve from an initial state to the solution, examples are exhaustive search, gradient descent, and so on.

4. The similarity measure: this controls the quality of registration and triggers the convergence to a solution, such as correlation, norms, mutual information (Viola and Wells III, 1997) and many other similarity measures mentioned in the extensive literature that covers the registration subject.

The registration process encompasses the following steps: feature detection and matching, mapping function estimation, and finally image re-sampling and transformation (Zitova and Flusser, 2003).

In shape registration, the representation of the shape plays a crucial role in the registration process, and can significantly influence the overall performance of the registration algorithm. Active contours (Kass et al., 1988), Fourier descriptors (Zahn and Roskies, 1972) and active shape models (Cootes et al., 1995) are among the methods which employ explicit representations of shapes to describe arbitrary shapes. These representations are based on a significant parameterisation. Non-parametric shape representations such as the signed distance functions (SDFs) used in this thesis, are becoming a more popular choice, due to their implicit handling of important shape deformation, and the relatively simple extension to describe higher dimensions (Paragios et al., 2003)-(Paragios et al., 2002).

Contour-based registration methods, examples of which are found in Marques and Abrantes (1997), Markovsky and Mahmoodi (2009), Li et al. (1995), and Eugenio and Marques (2003), are among the techniques used widely in shape registration, due to their fast convergence. These techniques, however, rely merely on the contour points of the shape as the feature space to be matched during the registration. These techniques also require point correspondence for the boundary of the shapes. Consequently, contour-based methods fall short if the two shapes to be registered have Euler characteristic numbers other than 1 (i.e. solid shapes only), due to the ambiguity surrounding the process of establishing the point correspondence.

SDF-based shape registration techniques minimise the distance between the SDFs of two shapes by using the gradient descent algorithm. The methods in Paragios et al. (2003), Cremers et al. (2006), Vemuri et al. (2003), and El Munim and Farag (2007) are examples of this approach. These techniques are widely used with segmentation applications because it is relatively straightforward to embed such registration methods in functionals used for segmentation. The SDF-based shape registration methods (for example the seminal work by Paragios et al. (2003)) are in general capable of dealing with shapes with various Euler characteristic numbers (different topologies), although, with the increase in shape complexity the cost increases, as well as, the chances of falling into local minima. These methods however, have some drawbacks: i) low speed due to their iterative nature, ii) in some cases, the instability and convergence to local minima, which leads to a limitation in the range of transformation these techniques can handle, and finally iii) these methods are somewhat difficult to implement due to the need to tune the stopping parameters and time step for each transformation individually (see Cremers et al. (2006) for details).

This chapter presents a level set based shape registration algorithm, employing the shape boundary as the zero level set of an SDF. The search space is assumed over shapes with similarity transformation function. Our algorithm therefore estimates the translation, rotation and scale parameters to minimise a distance term. The algorithm proposed here employs orthogonal linear transformations and shape

moments to compute the parameters individually. While the extension of the algorithm suggested in this chapter to 3D registration problems is possible, it is beyond the scope of this thesis and may be considered in the future. The main advantages of this registration algorithm are as follows:

- It employs the Fourier shift theorem to compute the translation and rotation parameters, however, we must stress that unlike early deployments of this theorem, such as, in Casasent and Psaltis (1975) and De Castro and Morandi (1987) and many others which are area-based methods, this chapter presents an SDF-based method where the embedded boundaries of the shapes rather than the entire image (gray levels) information are used for registration and this is a fundamental difference.

- This chapter suggests a method to estimate the scale parameter based on the radial moments of shapes. It is demonstrated empirically that the terms used to compute rotation angle and scale are minimised in the correct scale and rotation angle.

- The evaluation of the approach proposed here, shows that this registration technique is robust, fast, and suitable for a wide range of registration problems. These problems address issues such as shapes with various topologies (i.e. shapes with different Euler characteristic numbers), and sustained performance over large sample volumes. The results presented here are compared with the state-of-the-art shape registration algorithms in the literature.

In the rest of this chapter, the registration problem is stated in Section 5.3. Section 5.4 then describes the proposed algorithm. The experimental results are presented in Section 5.5, and finally conclusions are drawn in Section 5.6.

5.3 The Statement of The Problem

Let $\phi_p(x, y) : \Omega \to \mathbb{R}$ and $\phi_q(x, y) : \Omega \to \mathbb{R}$ denote Lipschitz functions representing SDFs of shapes $p(x, y)$ and $q(x, y)$ where Ω is the bounded image domain. SDF

functions are generally defined as in (5.1),

$$\phi_B(x,y) = \left\{ \begin{array}{cc} 0, & (x,y) \in B, \\ +D_E((x,y),B), & (x,y) \in I_B, \\ -D_E((x,y),B), & (x,y) \in \Omega - (I_B + B), \end{array} \right\}, \tag{5.1}$$

where D_E represents the minimum Euclidean distance between the shape boundary B and each point in the domain Ω, and I_B is a subset of the domain Ω representing the interior of the shape (Paragios et al., 2002).

Parameters s, θ, T_x, and T_y representing scaling, rotation, and translations in x and y directions respectively are required to transform ϕ_q to minimise a distance term between ϕ_p and the transformed ϕ_q defined in (5.2),

$$E = \int\int \left| \phi_p(x,y) - \frac{1}{s}\phi_q\left(s\mathbf{R}_\theta\left(x + T_x, y + T_y\right)\right) \right|^2 \mathrm{d}x\mathrm{d}y \tag{5.2}$$

so that,

$$\left(\hat{\theta}, \hat{s}, \hat{T_x}, \hat{T_y}\right) = \underset{\boldsymbol{\theta},\, \boldsymbol{s},\, \boldsymbol{T_x},\, \boldsymbol{T_y}}{\operatorname{argmin}} \; E \tag{5.3}$$

where $\hat{\theta}$, $\hat{s}$, $\hat{T_x}$, $\hat{T_y}$ are respectively the estimated angle, scale, and translations. ϕ_p, ϕ_q, and the transformed SDF are in Ω, i.e. the aim is to compute transformation parameters such that the result is still in Ω. Also $\mathbf{R}_\theta$ is a conventional 2D rotation matrix,

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{5.4}$$

The minimisation of (5.2) leads to a set of non-linear equations with respect to the desired parameters as discussed in (Paragios et al., 2003). According to the authors in (Cremers et al., 2006), direct minimisation of these equations as utilised in (Paragios et al., 2003) is slow to converge, can fall into local minima and requires to continuously tune parameters for a smooth convergence, this conclusion is also verified in this chapter.

The objective of this chapter is, therefore, to propose a linear method minimising distance terms equivalent to (5.2). The method should also be robust against local minima, fast, and easy to implement.

5.4 Shape Registration

5.4.1 Rotation

Finding a closed form solution to minimise Eq. (5.2) with respect to the rotation angle θ is a difficult task if not impossible. Therefore, an alternative term is proposed here to compute the desirable optimal solution $\hat{\theta}$. It is numerically demonstrated here that the minimisation of the term proposed here is equivalent to the minimisation of (5.2). To this end, polar coordinate system is employed in order to find the desirable rotation angle. The notion that a rotation in Cartesian domain is a displacement in the angular component of the polar coordinate system (Casasent and Psaltis, 1975) is exploited here. Initially the shapes' SDFs become translation invariant, i.e.:

$$\hat{\phi}_p(x, y) = \phi_p(x - p_x,\ y - p_y), \tag{5.5}$$

$$\hat{\phi}_q(x, y) = \phi_q(x - q_x,\ y - q_y), \tag{5.6}$$

where (p_x, p_y) and (q_x, q_y) are respectively the centroids of shapes ϕ_p and ϕ_q. To this end, these centroids can be computed by using the Heaviside function, however, taking the modifications made in the remark on page (75) to the definition of SDFs into account, the centroids of SDFs can be directly computed.

A simple and efficient algorithm (Mukundan, 2009) with sub-pixel accuracy is used to map $\hat{\phi}_p(x, y)$ and $\hat{\phi}_q(x, y)$ to polar coordinates. We then calculate $\hat{\phi}_p(\rho, \omega)$ and $\hat{\phi}_q(\rho, \omega)$ such that: $x = \rho\cos\omega$, and $y = \rho\sin\omega$. These 2D centralised SDFs can be expressed as 1D signals by removing ρ without losing the rotation information. In order to vanish ρ in $\hat{\phi}_p(\rho, \omega)$ and $\hat{\phi}_q(\rho, \omega)$, we marginalise out ρ by integrating

$\hat{\phi}_p(\rho,\omega)$ and $\hat{\phi}_q(\rho,\omega)$ over ρ according to (5.7) and (5.8).

$$\tilde{\phi}_p(\omega) = \int_\rho \hat{\phi}_p(\rho,\omega)\mathrm{d}\rho, \tag{5.7}$$

$$\tilde{\phi}_q(\omega) = \int_\rho \hat{\phi}_q(\rho,\omega)\mathrm{d}\rho, \tag{5.8}$$

It will be demonstrated in this chapter that this step reduces the complexity and eventually increases the computational speed.

Let $\bar{\phi}$ denote a normalised instance of $\tilde{\phi}_p$, i.e.:

$$\bar{\phi}_p(\omega) = \frac{\tilde{\phi}_\omega}{\sqrt{\int_\omega |\tilde{\phi}_p(\omega)|^2 \mathrm{d}\omega}}. \tag{5.9}$$

Also, let us define $\tilde{\beta}$ as the optimal scale factor which is a function of the desirable rotation angle θ, i.e.:

$$\tilde{\beta}(\theta) = \int_\omega \left(\tilde{\phi}_q(\omega)\bar{\phi}_p(\omega+\theta)\right)\mathrm{d}\omega. \tag{5.10}$$

The desirable angle is estimated by minimising the dissimilarity term E_θ in (5.11),

$$\begin{aligned}
E_\theta &= \int_\omega \left|\tilde{\phi}_q - \tilde{\beta}\bar{\phi}_p\right|^2 \mathrm{d}\omega, \\
&= \int_\omega \left|\tilde{\phi}_q - \tilde{\beta}\bar{\phi}_p\right| . \left|\tilde{\phi}_q - \tilde{\beta}\bar{\phi}_p\right| \mathrm{d}\omega, \\
&= \int_\omega \left(\left(\tilde{\phi}_q\right).\tilde{\phi}_q - \tilde{\beta}\left(\tilde{\phi}_q\right).\bar{\phi}_p - \tilde{\beta}\left(\bar{\phi}_p\right).\tilde{\phi}_q + \tilde{\beta}^2\left(\bar{\phi}_p\right).\bar{\phi}_p\right)\mathrm{d}\omega, \\
&= \int_\omega \left|\tilde{\phi}_q\right|^2 \mathrm{d}\omega - 2\tilde{\beta}\left\langle\tilde{\phi}_q.\bar{\phi}_p\right\rangle + \tilde{\beta}^2 \int_\omega \left|\bar{\phi}_p\right|^2 \mathrm{d}\omega,
\end{aligned} \tag{5.11}$$

From (5.9), $\int_\omega \left|\bar{\phi}_p\right|^2 \mathrm{d}\omega = 1$; therefore:

$$\int_\omega \left|\tilde{\phi}_q - \tilde{\beta}\bar{\phi}_p\right|^2 \mathrm{d}\omega = \int_\omega \left|\tilde{\phi}_q\right|^2 \mathrm{d}\omega - \tilde{\beta}^2. \tag{5.12}$$

Hence the minimisation of (5.11) is achieved by maximising $\tilde{\beta}$. The optimal rotation angle, $\hat{\theta}$, is therefore calculated by maximising $\tilde{\beta}$ with respect to θ, or,

$$\hat{\theta} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ \tilde{\beta} \tag{5.13}$$

The maximum value for $\tilde{\beta}$ can be computed by using the Fourier transform. Let the Fourier transform of $\bar{\phi}_p$ and $\tilde{\phi}_q$ be $\bar{\psi}_p$ and $\tilde{\psi}_q$ respectively, such that,

$$\bar{\psi}_p(\zeta) = \int_{\omega} \bar{\phi}_p(\omega)\ e^{-i(\omega\zeta)2\pi} \mathrm{d}\omega, \tag{5.14}$$

$$\tilde{\psi}_q(\zeta) = \int_{\omega} \tilde{\phi}_q(\omega)\ e^{-i(\omega\zeta)2\pi} \mathrm{d}\omega. \tag{5.15}$$

Accordingly, by using Parseval's theorem, it can be written:

$$\tilde{\beta}(\theta) = \int_{\omega} \left(\tilde{\phi}_q(\omega)\bar{\phi}_p\left(\omega+\theta\right)\right) \mathrm{d}\omega = \int_{\zeta} \left(\tilde{\psi}_q(\zeta)\bar{\psi}_p^*(\zeta)e^{i(\zeta\theta)2\pi}\right) \mathrm{d}\zeta, \tag{5.16}$$

where $(*)$ denotes the complex conjugate. Hence, $\hat{\theta}$ is computed as:

$$\hat{\theta} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ \tilde{\beta} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \int_{\zeta} \left(\tilde{\psi}_q(\zeta)\ \bar{\psi}_p^*(\zeta)\ e^{i(\zeta\theta)2\pi}\right) \mathrm{d}\zeta. \tag{5.17}$$

It will be numerically demonstrated that the optimal (desirable) rotation angle $\hat{\theta}$ minimising term (5.11) is a minimiser of term (5.2). We conjecture here that terms (5.2) and (5.11) have the same minimiser. The numerical results presented in this chapter support such a conjecture. Two different shapes (see Figure 5.1(a) to 5.1(c)), one regarded as a reference shape and the other as an observed shape, are used to evaluate terms (5.2) and (5.11). The observed shape is rotated with angles between (-180 to 180) degrees. In Figure 5.1(d), terms (5.2) and (5.11) are plotted with respect to rotation angle. As can be seen from this figure, both terms are minimised at the same rotation angle corresponding to the correct desired angle.

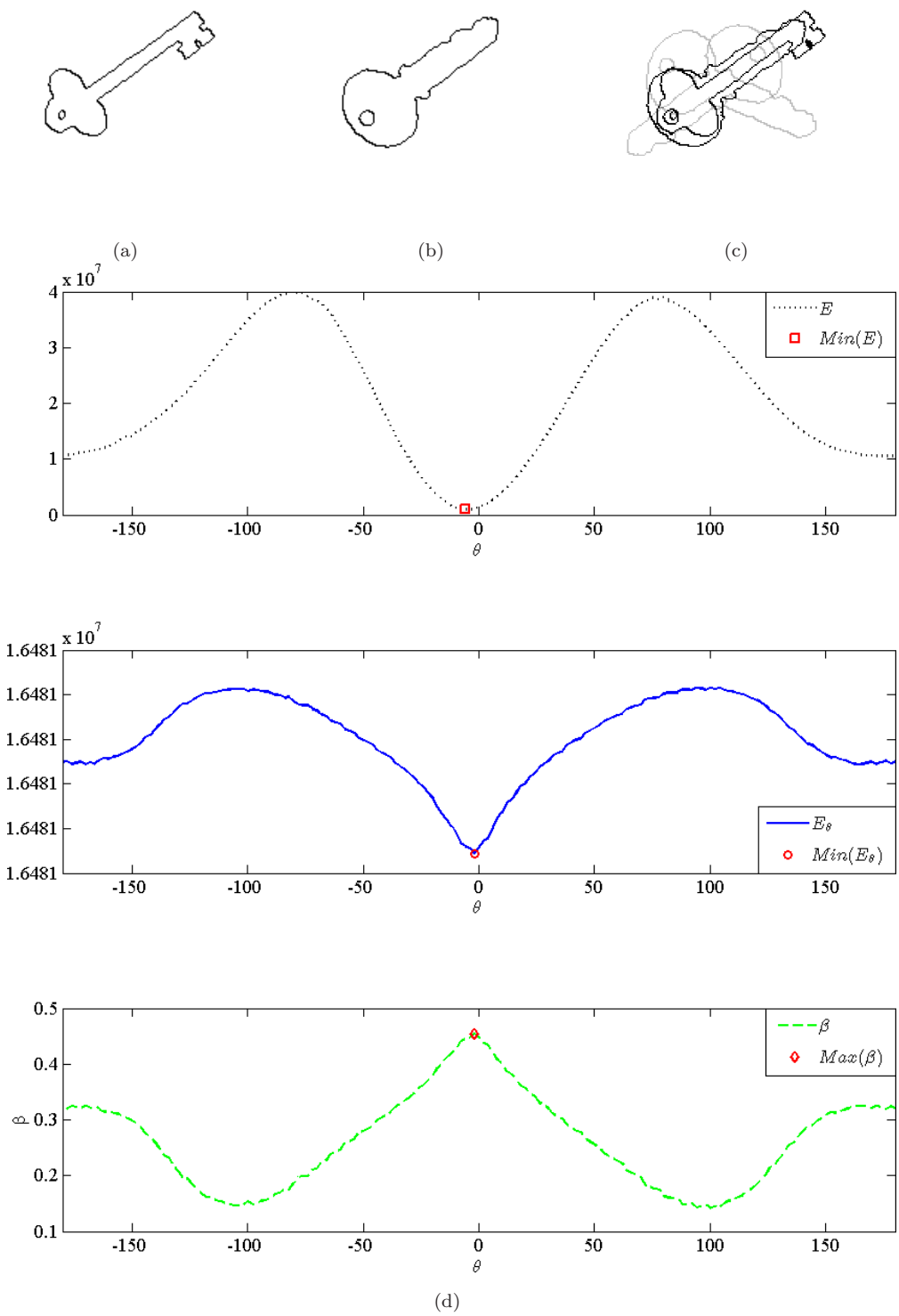


FIGURE 5.1: The equivalence of term (5.11) and term (5.2) with regard to the rotation angle. (a) reference shape-(b) observed shape, (c) both shapes with the observed shape being subjected to different rotations, (d) (top row) the error according to term (5.2) (d) (middle row) the error according to term (5.11), and (d) (bottom row) $\tilde{\beta}$ as calculated in (5.10); all graphs in (d) are plotted with respect to various rotation angle θ

5.4.2 Scale

Similar to the case in section 5.4.1, a closed form solution to minimise (5.2) with respect to the scale parameter, s, is difficult to achieve. Therefore an alternative term is suggested in this section to estimate s, and a numerical example is used to demonstrate that the minimisation of the term proposed here is equivalent to the minimisation of (5.2).

To calculate the scaling parameter between two shapes, the radial moments of SDFs are used to extract shapes' features. Paragios et al. (2003) have shown that if one shape is a scaled version of another, then the corresponding SDFs are proportional to the scale factor,

$$s\hat{\phi}_p(x, y) = \hat{\phi}_q(sx, sy), \tag{5.18}$$

where s is the scale parameter, and $\hat{\phi}_p$ and $\hat{\phi}_q$ are the translation invariant versions of ϕ_p and ϕ_q respectively as in (5.5) and (5.6).

The radial moments of the reference SDF $\hat{\phi}_p$ and the observed SDF $\hat{\phi}_q$ are computed as in (5.19) and (5.20),

$$M_m^{\hat{p}} = \int\int \left(\sqrt{x^2 + y^2}\right)^m \hat{\phi}_p(x, y)\mathrm{d}x\mathrm{d}y, \tag{5.19}$$

$$M_m^{\hat{q}} = \int\int \left(\sqrt{x^2 + y^2}\right)^m \hat{\phi}_q(x, y)\mathrm{d}x\mathrm{d}y, \tag{5.20}$$

where m represents the degree of the moment. By substituting (5.18) in (5.19), we arrive at (5.21):

$$M_m^{\hat{p}} = \frac{1}{s}\int\int \left(\sqrt{x^2 + y^2}\right)^m \hat{\phi}_q(sx, sy)\ \mathrm{d}x\mathrm{d}y \tag{5.21}$$

By using changes of variables, i.e., $X = sx$, and $Y = sy$, (5.21) can be written as,

$$\begin{aligned} M_m^{\hat{p}} &= \frac{1}{s}\int\int \frac{\left(\sqrt{X^2+Y^2}\right)^m}{s^m}\,\hat{\phi}_q(X,Y)\,\frac{\mathrm{d}X\mathrm{d}Y}{s^2}, \\ &= \frac{1}{s^{(m+3)}}\int\int \left(\sqrt{X^2+Y^2}\right)^m \hat{\phi}_q(X,Y)\,\mathrm{d}X\mathrm{d}Y, \\ &= \frac{1}{s^{(m+3)}} M_m^{\hat{q}}. \end{aligned} \tag{5.22}$$

Therefore, let E_s be the error term defined in (5.23):

$$E_s = \sum_m \left| M_m^{\hat{q}} - s^{m+3} M_m^{\hat{p}} \right|^2. \tag{5.23}$$

In order to find the desirable scale s, the above error term given in (5.23) needs to be minimised. However, this term is not linear with respect to variable s. By using a change of variable, the above non-linear least squares problem is, therefore, reduced to a linear one with respect to $\hat{s} = \log\ s$ by minimising the following error term:

$$\acute{E}_s = \sum_m \left| \log\left(\frac{M_m^{\hat{q}}}{M_m^{\hat{p}}}\right) - (m+3)\ \log s \right|^2. \tag{5.24}$$

Hence the optimal new parameter $\hat{s}$ is estimated by minimising $\acute{E}_s$:

$$\hat{s} = \underset{\boldsymbol{s}}{\operatorname{argmin}}\ \acute{E}_s. \tag{5.25}$$

It should be noted that the use of Chebyshev or Zernike moments leads to a non-linear least squares problem whose minimisation is more difficult and demanding than the current method proposed here. In the case of Chebyshev and Zernike moments, the non-linearity cannot be reduced to a linear problem by using a simple change of variables such as the one employed in (5.24).

Term (5.24) is used in this chapter to estimate the desired scale parameter s. In order to show that this term is equivalent to (5.2), two different shapes (see Figure (5.2(a)) to (5.2(c))), one regarded as reference shape and the other as observed shape, are used with scaling the observed shape by scaling factors between (0.5

-1.5). Terms (5.2), and (5.24) are computed for each scaling factor value. From Figure (5.2(d)), it is observed that terms (5.2) and (5.24) are minimised at the same value for the scaling factor.

5.4.3 Translation

By using the scaling and rotation parameters calculated in sections 5.4.1 and 5.4.2, term (5.2) is optimised to calculate the translation parameters T_x and T_y:

$$\phi_p = \phi_q(x - T_x, y - T_y). \tag{5.26}$$

By employing the same approach explained in section 5.4.1, the translation parameters are calculated. The only difference here is that since the translations are shifts in the Cartesian domain, there is no need to represents the SDFs in the polar coordinate system.

Let $\bar{\phi}_p$ denote a normalised version of ϕ_p, i.e.:

$$\bar{\phi}_p(x, y) = \frac{\phi_p(x, y)}{\sqrt{\int_{x,y} |\phi_p(x, y)|^2 \ \mathrm{d}x \ \mathrm{d}y}}. \tag{5.27}$$

Also, let us define $\tilde{\xi}$ as the optimal scale parameter which is a function of the translations T_x and T_y, such that, $\tilde{\xi}(T_x, T_y) = \int_{x,y} (\bar{\phi}_p(x, y) \ \phi_q(x + T_x, y + T_y)) \ \mathrm{d}x\mathrm{d}y$. The unknown desired translation parameters will be estimated by minimising the dissimilarity term in (5.28),

$$\begin{aligned}
\int_{x,y} \left|\phi_q - \tilde{\xi}\bar{\phi}_p\right|^2 \ \mathrm{d}x\mathrm{d}y &= \int_{x,y} \left(\left|\phi_q - \tilde{\xi}\bar{\phi}_p\right| . \left|\phi_q - \tilde{\xi}\bar{\phi}_p\right|\right) \ \mathrm{d}x\mathrm{d}y, \\
&= \int_{x,y} \left((\phi_q).\phi_q - \tilde{\xi}.(\phi_q).\bar{\phi}_p - \tilde{\xi}(\bar{\phi}_p).\phi_q + \tilde{\xi}^2.(\bar{\phi}_p).\bar{\phi}_p\right) \ \mathrm{d}x\mathrm{d}y, \\
&= \int_{x,y} |\phi_q|^2 \mathrm{d}x\mathrm{d}y - 2\tilde{\xi} \left\langle \phi_q.\bar{\phi}_p \right\rangle + \tilde{\xi}^2 \int_{x,y} |\bar{\phi}_p|^2 \mathrm{d}x\mathrm{d}y.
\end{aligned} \tag{5.28}$$

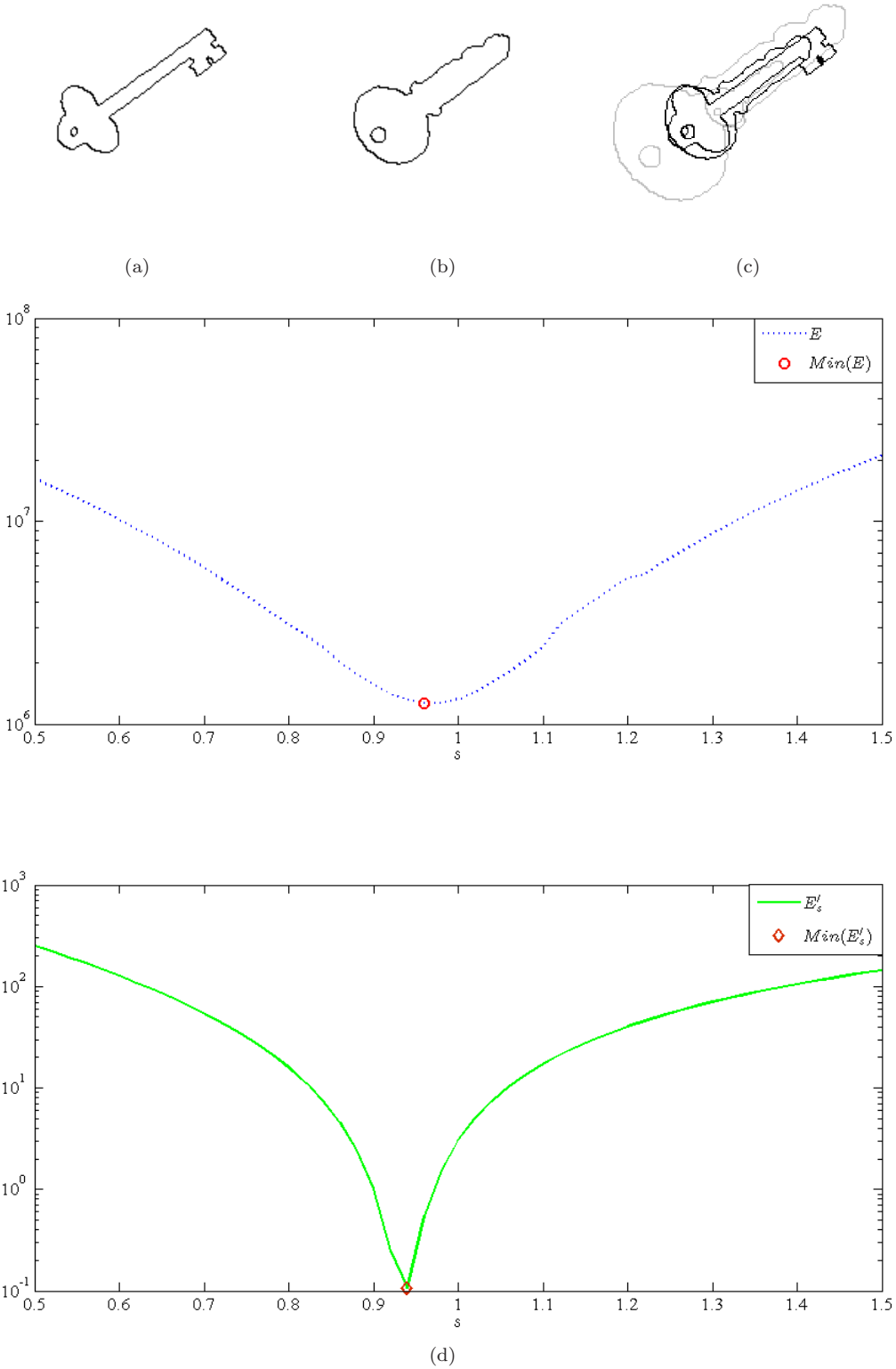


FIGURE 5.2: The equivalence of term (5.24) minimised here, and the statement of the problem as indicated in (5.2) with respect to various values for the scaling parameter, (a) reference shape (b) observed shape (c) both shapes with the observed shape being subjected to different scales (d) (top row) the error according to term (5.2), and (d) (bottom row) the error according to term (5.24)

From (5.27), $\int_{x,y} |\bar{\phi}_p|^2 \mathrm{d}x\mathrm{d}y = 1$; therefore,

$$\int_{x,y} |\phi_q - \tilde{\xi}\bar{\phi}_p|^2 \, \mathrm{d}x\mathrm{d}y = \int_{x,y} |\phi_q|^2 \, \mathrm{d}x\mathrm{d}y - \tilde{\xi}^2. \tag{5.29}$$

Accordingly the minimisation of (5.28) is achieved by maximising $\tilde{\xi}$:

$$\hat{T}_x, \hat{T}_y = \underset{\boldsymbol{T_x}, \boldsymbol{T_y}}{\operatorname{argmax}} \, \tilde{\xi} \tag{5.30}$$

Similar to what is presented in section 5.4.1, the maximum value for $\tilde{\xi}$ is computed in the frequency domain by using the Fourier transform, i.e.:

$$\begin{aligned} \left[\hat{T}_x, \hat{T}_y\right] &= \underset{\boldsymbol{T_x}, \boldsymbol{T_y}}{\operatorname{argmax}} \, \xi, \\ &= \underset{\boldsymbol{T_x}, \boldsymbol{T_y}}{\operatorname{argmax}} \int_{\omega_x, \omega_y} \left(\Theta_q(\omega_x, \omega_y)\bar{\Theta}^*_p(\omega_x, \omega_y) e^{i(T_x\omega_x + T_y\omega_y)2\pi}\right) \mathrm{d}\omega_x \mathrm{d}\omega_y, \end{aligned} \tag{5.31}$$

where $\hat{T}_x$, $\hat{T}_y$, $\Theta_q(\omega_x, \omega_y)$, $\bar{\Theta}_p(\omega_x, \omega_y)$, ω_x, and ω_y represent the estimated optimal translation parameters, the 2D Fourier transforms of ϕ_q and $\bar{\phi}_p$, and the spatial frequencies respectively.

Remark: Since images in practice are in the discrete domain, the Discrete Fourier transform (DFT) is employed instead of continuous Fourier transforms used in this section. It is therefore, required to modify the definition of SDFs to cope with the periodicity property introduced by DFT:

Let Ω be the image domain. This domain is partitioned by the shape perimeter into two regions, the shape interior I_B and the background, and let $\phi : I_B \to \mathbb{R}^+$ be a Lipschitz function that represents the distance transform for the boundary B of the shape, this is expressed in (5.32):

$$\phi_B(x, y) = \left\{ \begin{array}{cc} D_E\left((x, y), B\right), & (x, y) \in I_B, \\ 0, & (x, y) \in \Omega - (I_B + B) \end{array} \right\}, \tag{5.32}$$

where D_E,as before represents the minimum Euclidean distance between the shape boundary and each point in the domain. This modified SDF representation is induced by the periodicity requirements of the DFT used in Section 5.4. An algorithm of the technique proposed in this chapter is presented in Appendix C.

5.5 Results and Discussions

In this section, a set of examples is presented to demonstrate the performance of the shape registration method proposed here in comparison with other known registration methods in various shape registration problems. In all of the experiments presented in this chapter the moments up to the fifth degree are used to compute the scale parameter s by using the proposed method in Section 5.4.2. The higher the degree of the employed moments is, the more accurately s is estimated. However, this higher accuracy comes at the expense of higher numerical complexity.

Except for the first experiment, the results of the proposed technique are compared with two well-known shape registration methods. The first one is based on contours representing shapes, (e.g. see Markovsky and Mahmoodi (2009)). It will be called throughout this section the contour-based method. The other method used here for comparison employs SDFs to represent shapes similar to the technique proposed here; however, unlike the method proposed here, a gradient descent approach is employed to compute the optimal solutions (see Paragios et al. (2003) for more details). Throughout this section this method is termed as the SDF-based method.

In all but the first example a distance term $Dist$ is used to measure the quality of registration. This term is defined as:

$$Dist(\phi_{reference}, \phi_{registered}) = \int_x \int_y |H(\phi_{reference}) - H(\phi_{registered})| \, \mathrm{d}x\mathrm{d}y, \quad (5.33)$$

where H is the Heaviside function. In other words, this distance counts the number of un-overlapped pixels (area) in both shape SDFs. Also used as a measure

of quality is the cost in terms of run time (t) required to compute the optimal registration parameters.

For better demonstration, rather than showing the SDFs or the silhouettes of the shapes, the contours of the observed and reference shapes are shown in the figures throughout this section.

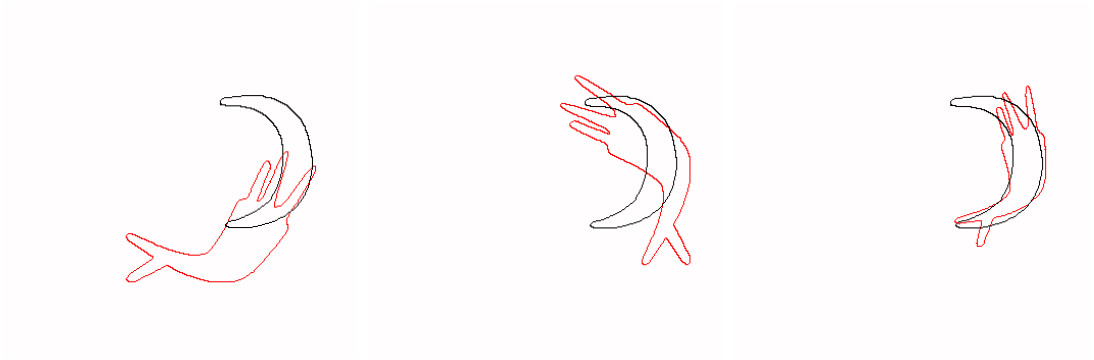
5.5.1 Comparison with Shapes' Centre of Mass Registration Methods

One simple but useful approach to find the position and the rotation angle is to compute the center of mass and the principal axes of two shapes and then to use the difference between these computed values respectively as the translation and rotation parameters (see Alpert et al. (1990), Dong and Boyer (1996), Chen et al. (2005) and many others). This method is widely used in medical image registration problems (Maintz and Viergever, 1998).

In the first example shown in Figure (5.3(a)), the shapes of two sickled red blood cells (RBCs) are registered. For the observed RBC, the center of mass is almost inside the shape while for the reference RBC, the center of mass is situated outside the shape. The use of the center of mass registration approach, shown in Figure (5.3(b)), results in shifting the observed shape incorrectly outside the reference shape. Figure (5.3(c)) also shows how the same problem is resolved by using the method proposed here. It is obvious that the observed shape is aligned to the reference shape despite the dissimilarity and the shift of mass between the two shapes.

5.5.2 Shapes with Different Topologies

In the example presented in this subsection, two shapes with different Euler characteristic numbers are used. These shapes obviously have different topologies. In



(a) (b) (c)

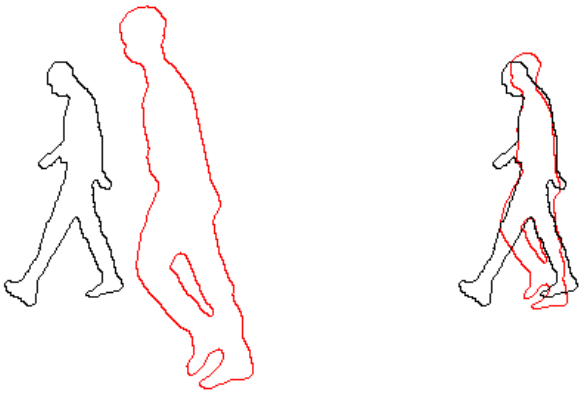
FIGURE 5.3: Two sickled RBCs, the observed shape (red) and the reference shape (black). (a) Initially, (b) the alignment by using centre of mass method, (d) by using the method proposed here

Figure (5.4(a)), the observed shape is an open '4' with Euler number one, and the reference shape is a closed '4' with Euler number zero.

As shown in Figure (5.4(b)), the two shapes are completely misaligned by using the contour-based method in Markovsky and Mahmoodi (2009). This shortcoming stems directly from the necessity to establish correct point correspondence in order for this method to work correctly. Such a requirement makes the algorithm sensitive to topological changes due to the ambiguity of establishing point correspondences in certain cases. Since in this case the shapes have different topologies, such correspondence may not be achievable. Figure (5.4(c)) is the result of applying the scheme in Paragios et al. (2003) to solve the registration problem. This figure shows a typical example of the local minima hurdle associated with this approach. In Figure (5.4(d)), the shapes in question are correctly registered by using the approach proposed in this chapter. Another example of the consistent performance of the proposed algorithm regarding the topological variance problem is showcased in Figure (5.5) where two gait shapes with different topologies are register

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

FIGURE 5.4: The registration of shapes with different topologies (size= 300×300). (a) initial shapes, (b) the result of attempting to register the shapes by using the contour based method in Markovsky and Mahmoodi (2009), (c) the registration of the shapes by using the SDF-based method in Paragios et al. (2003) (d) the two shapes superimposed optimally by using the approach proposed here



(a) (b)

FIGURE 5.5: The registration of two gait shapes with different topologies (size = 300×300). (a) initial shapes, (b) the two shapes aligned by using the algorithm presented here

5.5.3 Complex Shapes Containing Various Sub-Shapes

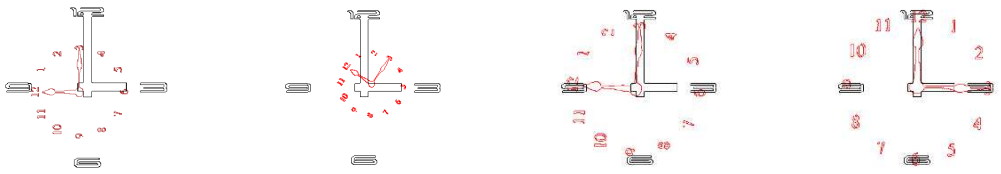
In the third example, two complex shapes, each having different number of components are registered. The employment of contour method in Markovsky and Mahmoodi (2009) to register such shapes, for example, by registering the individual objects in the observed shape to their counterparts in the reference shape may do partially. Wherein the objects with no counterparts remain unregistered and the overall shape integrity may be distorted. The other way is to register the shape components all together as a single entity. With this comes again the issue

TABLE 5.1: Comparison of the Results of the Registration Algorithms

	Dist (pixel) Original	*Dist* (pixel) in Markovsky and Mahmoodi (2009)	*Dist* (pixel) in Paragios et al. (2003)	*Dist* (pixel) in Proposed	t (Sec.) in Markovsky and Mahmoodi (2009)	t (Sec.) in Paragios et al. (2003)	t (Sec.) in Proposed
Number 4	3664	2296	2722	1078	1.09	28.25	2.83
Gait-Topology	13706	-	-	2305	-	-	3.02
Clock-face	5859	5162	6442	4775	3.13	29.18	9.361
Hands radiog.	8140	8303	8191	6194	3.15	1175.8	2.86
Similar hands	7249	668	916	171	1.34	3543.5	2.89
Gait-Similar	7420	314	499	220	1.74	859.72	2.89
Verteb.-x-ray	Av. = 2244	Av. = 345.2	Av. = 862.4	Av. = 339.2	Av. = 1.05	Av. = 470.5	Av. = 3.06
Gait-Sample	Av. = 3800	-	-	Av. = 1366	-	-	Av. = 3.39

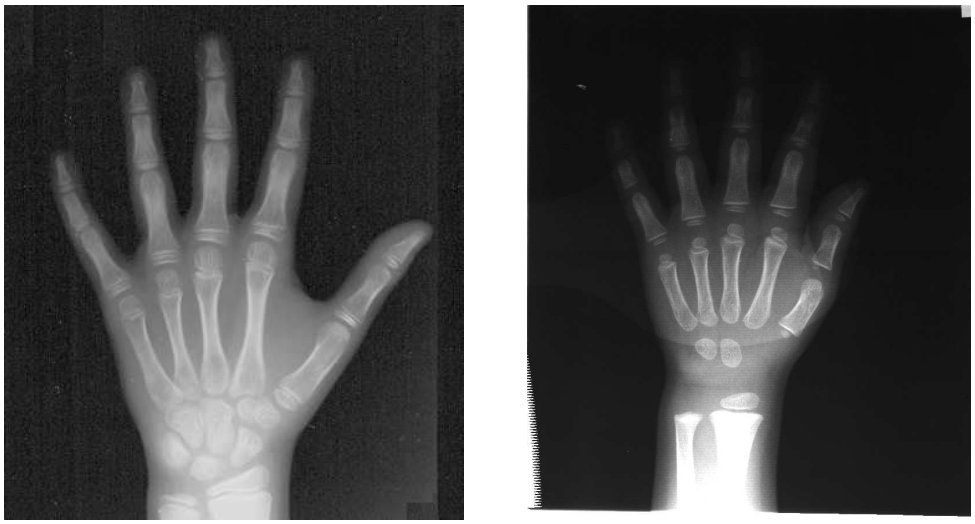
of point correspondence seen in the previous example. The algorithm in Paragios et al. (2003) on the other hand, is capable to certain extent of dealing with complex shapes. By increasing the shape complexity, the chance of falling into local minima increases as well. The time required to find the final parameters as demonstrated in Table (5.1) becomes longer.

In Figure (5.6), the observed shape is a clock face with conventional indicators, while the reference shape has compass point indicators. Similarly, in Figure (5.7), an x-ray of a child's hand with incomplete carpal bones is registered to an x-ray of an adult's hand with completely developed carpal bones. These two experiments demonstrate that the registration algorithm proposed here have the impetus to register two shapes, even if there is no direct (one to one) correspondence among components forming the complex shapes. This is in contrast to the methods in Markovsky and Mahmoodi (2009) and Paragios et al. (2003) which both produce poor results.

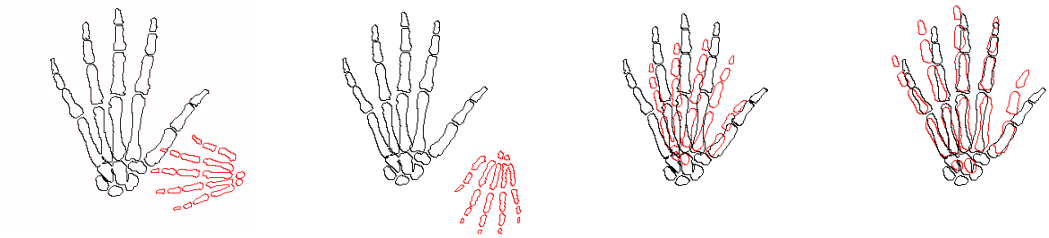


(a) (b) (c) (d)

FIGURE 5.6: The registration of clock-faces with different number of components (size= 480×480). (a) Initial shapes. (b) registration by using contour-based technique in Markovsky and Mahmoodi (2009). (c) registration by using the SDF-based algorithm proposed in Paragios et al. (2003), (d) registration by using the algorithm proposed here



(a) (b)



(c) (d) (e) (f)

FIGURE 5.7: The registration of two hand radiographies with various number of bones (size= 300×300), (a) original x-ray of a developed hand and (b) original x-ray of an underdeveloped hand with less bones (c) initial shapes before registration, (d) the registration by using the contour-based algorithm in Markovsky and Mahmoodi (2009), (e) the registration by using the SDF-based algorithm in Paragios et al. (2003). (f) the registration by using our registration algorithm

5.5.4 Similar Shapes with Artificial Transformations (Accuracy of Registration)

In the fourth example, similar shapes transformed by using various artificial transformations are registered. This is to ensure that maximum accuracy is achievable and quantifiable.

In Figure (5.8), two reference shapes (a hand and a walking subject shapes) are transformed by using two sets of parameters. Therefore, the observed shapes are replicas of the reference shapes. The algorithms in Markovsky and Mahmoodi (2009), Paragios et al. (2003), and the proposed algorithm are used to estimate the registration parameters. From this figure it is evident visually that the technique presented in this chapter delivers the highest matching of the two identical shapes. These results are also presented in Table 5.1.

5.5.5 Shapes' Database:

In the fifth experiment, a set of 86 x-ray images available from the U. S. National Library Of Medicine (2010) [1] are registered. These images are segmented to extract the shapes of the third cervical vertebra (C3). The corresponding SDF's are eventually computed by using Fast Marching algorithm (Sethian, 1996). The algorithm proposed here along with the algorithms in Paragios et al. (2003) and Markovsky and Mahmoodi (2009) are deployed to register these 86 shapes for the sake of performance comparison.

In Figure (5.9(a)), examples of this set are shown before registration. Figure (5.9(c)), (5.9(e)), and (5.9(g)) are the registered counterparts of the shapes shown in Figure (5.9(a)) by using respectively Markovsky and Mahmoodi (2009), Paragios et al. (2003), and the algorithm proposed here.

The quantitative assessment of this example is discussed in Figure (5.10). The figure shows that the average of the distance *Dist* (defined in 5.5) between the

[1] `http://archive.nlm.nih.gov/proj/ftp/ftp.php#available`

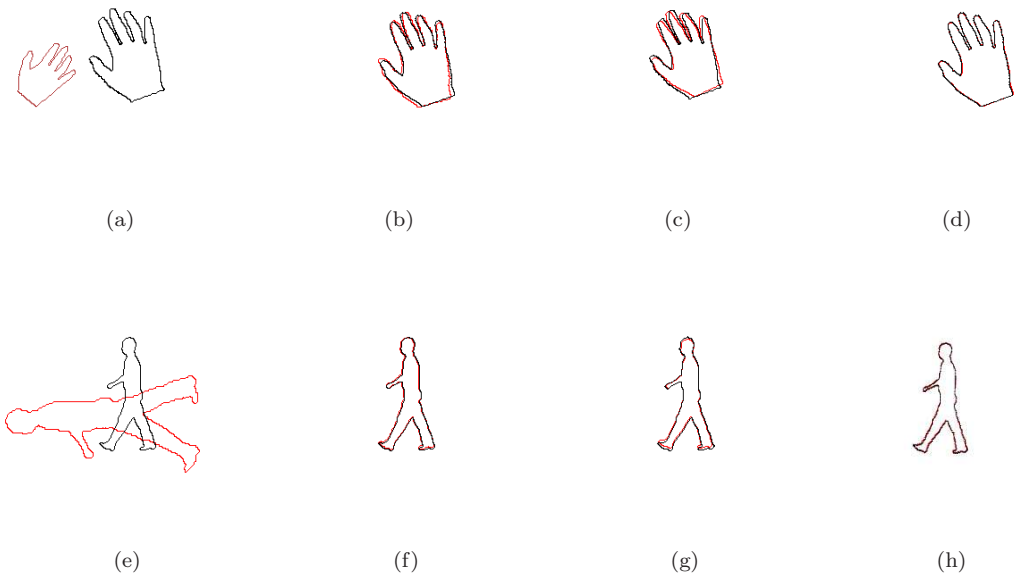


FIGURE 5.8: The registration of identical shapes with synthetic transformations (size= 300×300). (a) Initial shapes (hands example), the observed shape is generated by transforming the reference shape by using the parameters $\theta = -60, s = 0.7, T_x = -90, T_y = 20$. (b) The registration of the shapes in (a) by using the approach in Markovsky and Mahmoodi (2009), (c) the registration of the shapes in (a) by using the method in Paragios et al. (2003), (d) the registration of the shapes in (a) by using our method. (e) Initial shapes (gait exmple), the observed shape is generated by transforming the reference shape by using the parameters $\theta = 95, s = 1.7, T_x = -10, T_y = 20$, (f) the registration of the shapes in (e) by using Markovsky and Mahmoodi (2009), (g) the registration of the shapes in (e) by using the method in Paragios et al. (2003), (h) the registration of the shapes in (e) by using our method

reference and the observed shapes, for our method is reduced to 15.11% of that before registration. *Dist* is reduced to 15.38% for the algorithm in Markovsky and Mahmoodi (2009), and to 38.4% for the algorithm in Paragios et al. (2003). This example clearly demonstrates the robustness of our algorithm in registering large volumes of data while at the same time (as seen in Figure (5.10(b)) our method is characterised with a very competitive cost in terms of run time.

In another example, the shapes of walking subjects (in full-stride) are registered to a reference shape. In this example, Figure (5.11) shows that the average error, *Dist*, is reduced to the third after applying the algorithm proposed in this thesis to this sample.

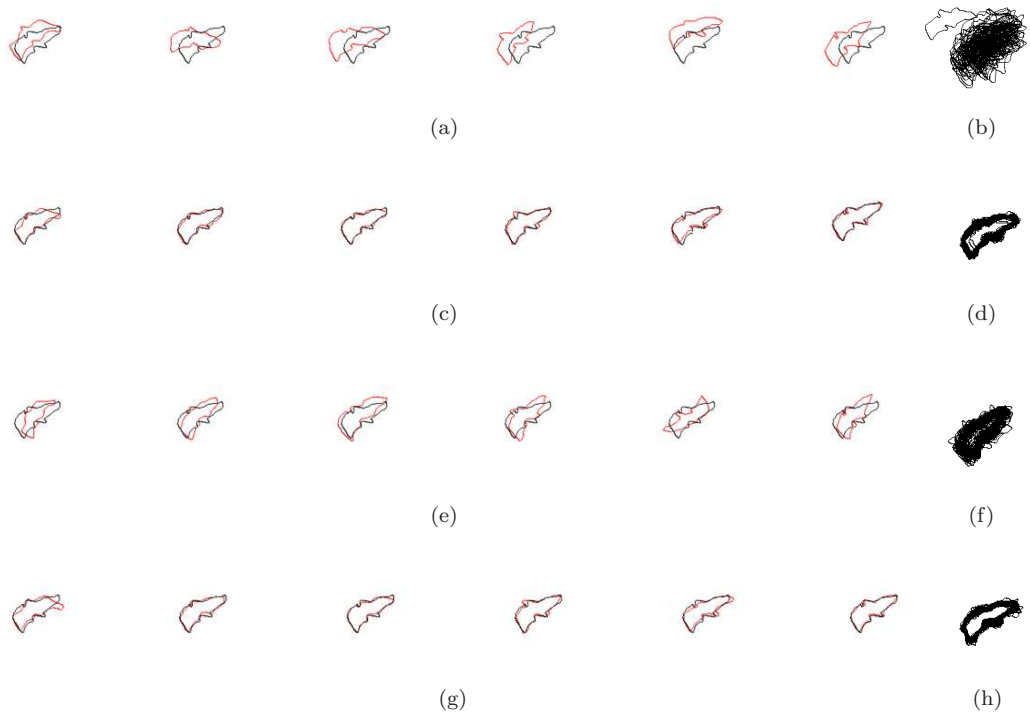
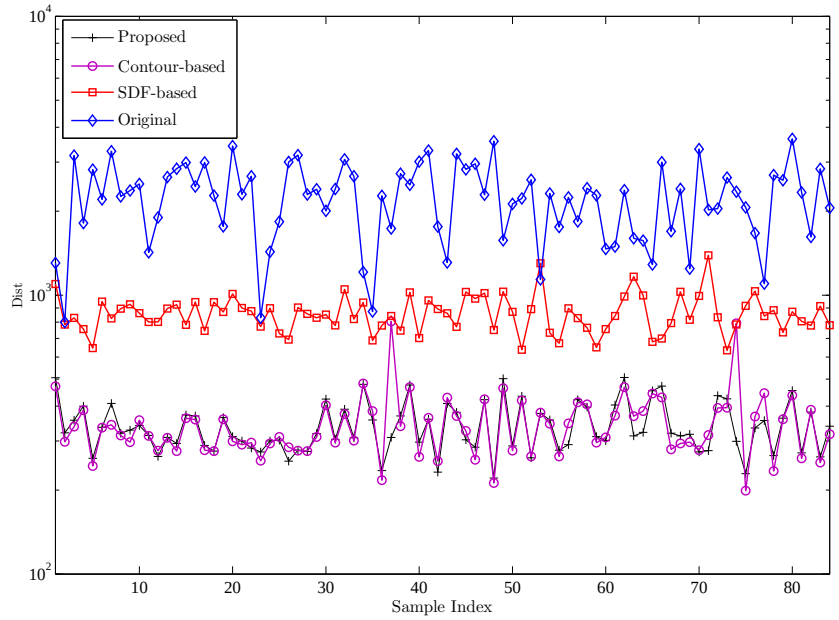


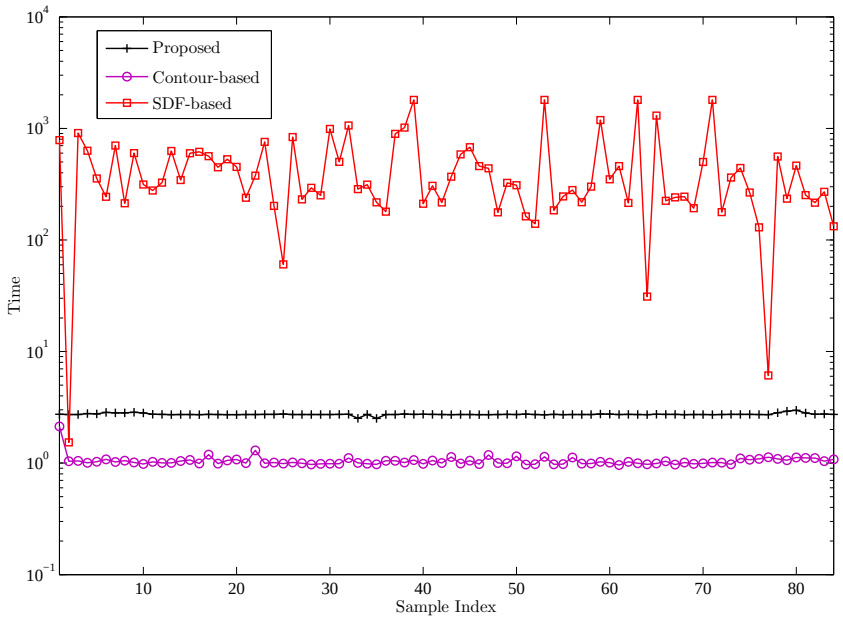
FIGURE 5.9: The registration of a medical image database, a set of 86 C3 vertebrae shapes (size= 300×300). (a) examples of the shapes before registration, (b) accumulated initial shapes, (c) the registration of the shapes by using the method in Markovsky and Mahmoodi (2009), (d) accumulated shapes after registration by using Markovsky and Mahmoodi (2009), (e) the registration of the shapes by using the algorithm in Paragios et al. (2003), (f) accumulated shapes after registration by using the method in Paragios et al. (2003), (g) the registration of the shapes by using the algorithm proposed here, (h) accumulated shapes registered by using our algorithm

The results summarised in Table (5.1), clearly suggest that in all of the experiments, the proposed algorithm have produced better registration performance compared to the other two approaches. It is shown here that the proposed algorithm is capable of registering shapes with various Euler characteristic numbers leading to different topologies and higher shape complexity. This is in contrast to the method in Markovsky and Mahmoodi (2009) which is limited to solid shapes (shapes with no holes).

The results also suggest that the algorithm proposed here is much faster than that proposed in Paragios et al. (2003) and is as fast as the algorithm presented in Markovsky and Mahmoodi (2009). The results presented in this section demonstrate a promising prospect for the shape registration algorithm proposed in this



(a)



(b)

FIGURE 5.10: : A quantitative assessment of the results demonstrated in Figure (5.9) by using *Dist* and the time t, (a) comparison of *Dist* produced by our algorithm and the algorithms in Markovsky and Mahmoodi (2009) and Paragios et al. (2003), (b) the time required for registration by using the three algorithms

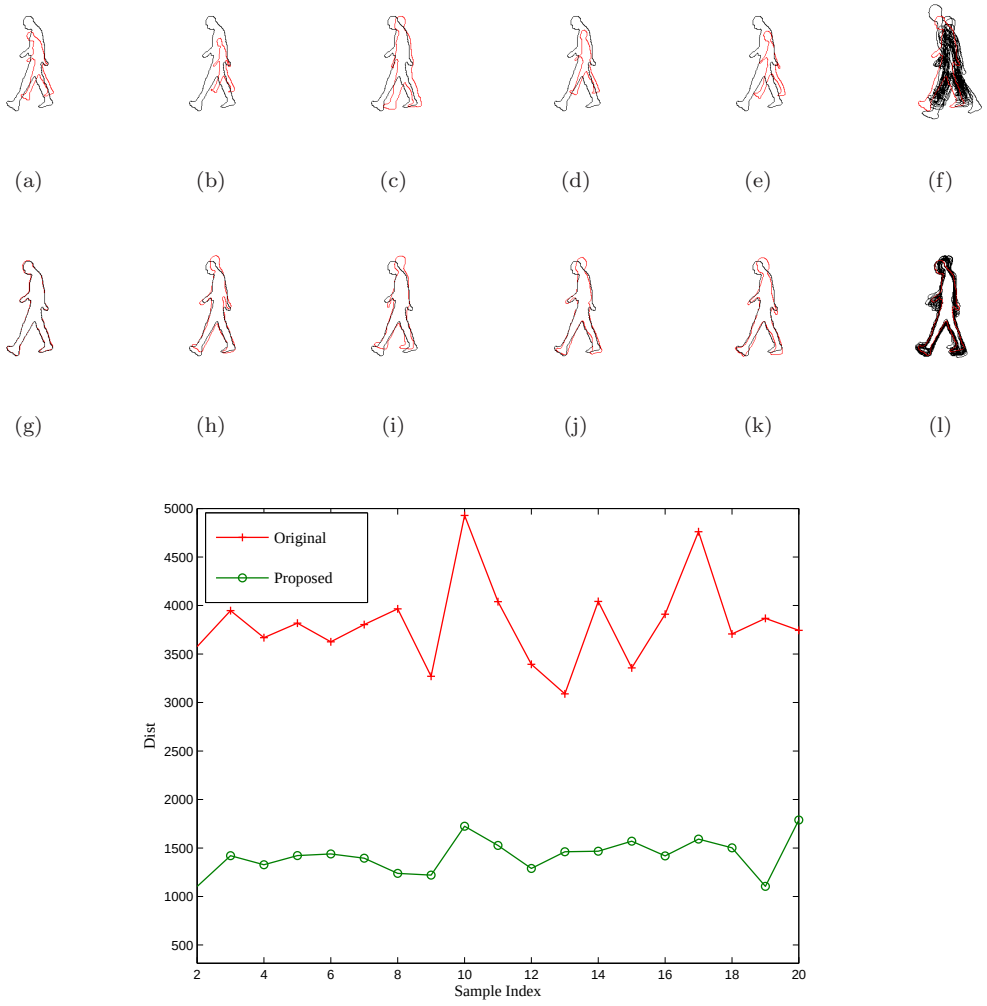


FIGURE 5.11: The registration of a sample of gait shapes, a set of 20 full-stride gait shapes (size= 300×300). (a-e) examples of the shapes before registration, (f) accumulated initial shapes, (g-k) the registration of the shapes by using the algorithm proposed here, (h) accumulated shapes registered by using our algorithm, (m) the error measure, *Dist*, computed for the shapes before and after registration

chapter and indicate that this algorithm is robust and fast and can perform registration for shapes with various complexities.

5.6 Conclusions

This chapter presents a shape registration algorithm which employs a modified signed distance function to represent the shapes. The proposed algorithm estimates the parameters by using closed form expressions. This algorithm takes the advantage of Parseval's theorem to estimate the rotation and the translation parameters by using Fourier transform. The algorithm also uses the radial moments to estimate the scale parameters.

This registration technique is tested successfully on a variety of problems including complex shapes and shapes with various topologies which cannot be registered by using contour based methods. The experimental results show that the proposed registration algorithm is fast, accurate, stable, and does not fall into local minima in contrast with other iterative techniques. Furthermore our technique is successfully exploited to register a large volume of medical shapes from human organs.

The results presented here show that our registration algorithm is overall more robust and accurate in comparison with the other algorithms investigated here. The algorithm proposed here, therefore, can be a solution to a wide range of registration problems encountered in image processing and computer vision especially medical problems.

As a future work, this algorithm can be generalised to register 3D shapes. The generalisation of this algorithm can be achieved for instance, by employing radial moments in a spherical coordinate system to estimate the scale parameter. The rotation parameter can be computed by using 2D DFT in the spherical coordinate system. Similarly, the translation parameters can be computed by employing a 3D DFT in the Cartesian coordinate system of the original shapes.

Chapter 6

Prior Aided Gait Segmentation

6.1 Summary

A major reason for the problems seen at the end of Chapter 2 upon dealing with real-world images, is that the segmentation model (as represented by the energy functional in Eq. (2.6) or (2.16)) does not specify the class which the shape being searched for falls into. Although this model establishes a good basis for the segmentation of an image by separating it into regions. The approach has its limitations, as noticed before. Some supervision can be included to guide the functional to a meaningful collection of regions constituting the target shape. Prior knowledge is known to improve segmentation/grouping significantly. This chapter discusses a proposed segmentation algorithm that employs the shape model defined in Chapter 4 and the registration algorithm presented in Chapter 5. The test results show accuracy, resilience and flexibility in segmenting real-world images in spite of various types of visual complications added to them.

6.2 Introduction

An important approach in the analysis of patterns is the employment of generative models. Predictions following from particular hypotheses are generated and

tested against the products of analysis of an image (Blake et al., 1998). An elegant combination of analysis and synthesis (Mumford, 1996) yields a productive partnership between the data and the learned knowledge.

One of the pilot uses of prior knowledge in image segmentation is to be found in the pioneering work by Cootes et al. (1995), in which, a Gaussian model is used to learn a set of training shapes represented by series of corresponding points. By adopting implicit shapes as the training set instead of only a set of markers, Leventon et al. (2000) developed an image segmentation approach which employs a geodesic active contour (Caselles et al., 1997) guided by statistical prior shape captured by PCA.

In the context of statistical shape learning, level sets allow the construction of shape dissimilarity measures defined on the embedding function which can handle shapes of varying topology (Cremers, 2006). Cremers et al. (2002) use level sets and Mumford-Shah functional with a non linear statistical shape model based on kernel density estimation.

In Leventon et al. (2000), the authors have developed active contours that use a shape model defined by a PCA. In their approach, the active contour evolves locally based on image gradients and curvature and globally towards the maximum a posteriori (MAP) probability of position and shape of the prior shape model.

By using the Chan and Vese (2001) segmentation model, Tsai et al. (2003) modified the Leventon et al. (2000) approach to develop prior shape segmentation for objects with linear deformations such as human organs in medical images. Optimisation is performed directly within the subspace of the first few eigenmodes.

The AR-based shape model (Cremers, 2006) mentioned earlier in this thesis is used to develop a prior shape gait segmentation algorithm. This algorithm addresses elegantly the time coherence between sequential shapes in a video. However, the algorithm in Cremers (2006) lacks the flexibility to learn from different sources because the transition from one video to another is meaningless in terms of time coherence. Also, this algorithm does not perform sufficient shape registration (it

does not include scale matching). Therefore the AR-based segmentation algorithm is limited to segmenting videos including shapes familiar (in terms of look and size) to those in the training set.

Motivated by the above mentioned and other approaches, a prior guided gait segmentation algorithm is proposed here. This algorithm exploits the proposed shape model which is capable of producing novel shapes while keeping the time coherence of the produced shapes. This model is engaged in a framework with the powerful region based segmentation addressed earlier. This combination achieves a balanced parting of roles between the prior and the active contour. Moreover, this framework, uses the registration algorithm proposed in this thesis to match the prior and the evolving contour to ensure.

This algorithm functions in the following way: initial few shapes are assumed know to start the analysis phase closer to the target subject features. The synthesis phase then produces a prior shape which is used by the evolving contour in conjunction with the image data to produce the next segmented shape. This shape is consequently used in the next analysis phase and so forth. Frame after frame the algorithm boosts its knowledge about the target and evolves toward more accurate guesses of the features of the unknown gait cycle.

This algorithm is evaluated by using real-world challenges. The outcomes are compared with the results of applying the algorithm suggested in the novel work by Cremers (2006). These results demonstrate more accuracy and reliability in the case of the proposed algorithm.

So, in the remaining part of this chapter The structure of the proposed segmentation algorithm is described in Section 6.3. The estimation of the prior shape is presented in Section 6.4. The proposed algorithm is tested and analysed in Section 6.5. Some conclusions are drawn Section 6.6.

6.3 Prior Shape Segmentation

Following the work in Cremers (2006), Cremers et al. (2006), Tsai et al. (2003), Bresson et al. (2006), and others, we merge prior shapes in the segmentation process. This is achieved here by incorporating the Chan and Vese model with the learned shape model discussed in Chapter 4 in a level set energy functional. Let ϕ be the evolving SDF, and let ϕ^{Pr} be the SDF of prior shape, then, the distance between these SDFs is defined by the energy in (6.1),

$$E_{prior}(\phi, \phi^{Pr}) = \int_{x,y} \left|\phi(x,y) - \phi^{Pr}(s\mathbf{R}_\theta(x+T_x, y+T_y))\right|^2 \mathrm{d}x\mathrm{d}y. \tag{6.1}$$

The total prior shape segmentation energy is then defined by (6.2),

$$E_{total}(\phi(x,y)) = \gamma_1 \tilde{E}_{cv}(c_1, c_2, \sigma_1, \sigma_2, \phi(x,y)) + \gamma_2 E_{prior}(\phi(x,y), s\mathbf{R}_\theta \phi^{Pr}(x+T_x, y+T_y). \tag{6.2}$$

The segmentation process is achieved by evolving the functional in (6.2) with time until equilibrium. The evolution of $\tilde{E}_{cv}$ (Eq. (2.16)) is detailed in Chapter 2.

Warp parameters $s, \theta, T_x,$ and T_y are optimised by solving the following equations:

$$\frac{\partial E_{prior}}{\partial \theta} = 0, \tag{6.3}$$

$$\frac{\partial E_{prior}}{\partial s} = 0, \tag{6.4}$$

$$\frac{\partial E_{prior}}{\partial T_x} = 0, \tag{6.5}$$

$$\frac{\partial E_{prior}}{\partial T_y} = 0. \tag{6.6}$$

The solution to Eq. (6.3), (6.4), (6.5) and (6.6) is achieved by using the algorithm proposed in Chapter 5.

The inference of ϕ^{Pr} to minimise E_{prior} is given in the next section.

6.4 Prior Shape Inference

The prediction of the prior is in fact a twofold problem: first, to estimate the particular subject (in the space of walking subjects captured by the model detailed in chapter 4), and second, to decide a particular shape (time instant in the cycle) that is most likely to have produced the given data.

In the proposed algorithm, an estimate to an unknown subject's identity is made by estimating the feature $\vec{\beta}_t$ from the distribution of walking subjects, followed by estimating a particular shape in the cycle by using $\alpha_t(i)$ at time instant i.

By recalling Section 4.4, in particular, Eq. (4.3), (4.10) and (4.12). Given the shapes segmented previously and the frame u_t. An update to the subject identity, $\vec{\beta}_t$, by using our prior shape model is computed according to Eq. (6.7):

$$\vec{\beta}_t = \chi^T \, (\vec{\alpha}_{t-1} - \bar{\alpha}), \tag{6.7}$$

where $\vec{\alpha}_{t-1}$ is the eigenmodes set computed by using the current shapes segmented up till $t-1$. $\vec{\beta}_t$, therefore, can be written as,

$$\vec{\beta}_t = \chi^T((\Psi^T \, (\phi_{1:t-1} - \bar{\phi})) - \bar{\alpha}), \tag{6.8}$$

where $\phi_{1:t-1}$ are (for simplicity[1]), the observed (segmented) shapes so far. This is the analysis phase.

The synthesis phase starts by computing the sequence of shapes representing the gait cycle of the estimated subject. Initially the eigenmodes, $\vec{\alpha}_t$, are computed:

$$\vec{\alpha}_t = \chi \, \vec{\beta}_t + \bar{\alpha}, \tag{6.9}$$

[1]This should in fact be the vectorised shapes $\vec{u}_{1:t-1}$ in stead of $\phi_{1:t-1}$.

The next prior shape is then computed:

$$\phi^{Pr} = \psi^T \, \vec{\alpha}_t(i) + \bar{\alpha}. \tag{6.10}$$

ϕ^{Pr} is plugged into energy (6.2) which evolves toward the next segmented shape ϕ_t. Shapes $\phi_{1:t}$ are then used to actuate the analysis-synthesis process for the next frame u_{t+1}. One final point is mentioned here, that is, the selection of a particular shape as the prior. In other words the computation of i in Eq. (6.10). This is done here by averaging the offsets of the known successive shapes. A detailed algorithmic procedure for the inference of the prior shape is given in Appendix E.

6.5 Evaluation and Discussion

In this section the proposed algorithm is implemented and tested by using data taken from the Southampton gait database (Shutler et al., 2002). The test data includes both indoor and outdoor samples. In order to put the proposed algorithm in context with the literature a set of experiments is initiated. In those experiments the results of this algorithm are compared to the results produced by the algorithm presented by Cremers (2006). This will be called here the AR-based algorithm. To show the potential of the tested algorithms, the indoor data is contaminated with different types of noise and with occlusion. The outdoor scenes normally offer high level of challenge and distraction to computer vision algorithms. Therefore, no noise or occlusion is introduced to those images.

As the algorithm presented in Cremers (2006) lacks the ability to accommodate for discrepancies in scale and rotation between the data and the prior due to the fact that no proper registration is performed (no scale is predicted); and also, due to the fact that the AR-based model can only be trained by using videos from a single source (subject), in the following experiments, the AR-based model is trained by using hand segmented data of the same subject in the scene being segmented. This is to ensure that the size of the shapes of the subject in the

training set is the same as those in the data to be segmented (which, obviously, favours the AR-based algorithm). Consequently, because of these limitations, the AR-based algorithm is compared to our algorithm only once in each category of experiments.

The following experiments help understand the way this algorithm works. These examples show that our shape model is producing shapes that are both plausible and increasingly similar to the subject in the scene. Such a guiding force indeed makes it easier for the evolving contour to split the region formed by an occlusion into an area which is inside and another area outside the contour. Although this seems to contradict the basics of the $\tilde{E}_{cv}$ force, it is perfectly compliant with the way smart (human, for example) visual system works. This means that the contour can be made to rely on the bi-modality in the scene only selectively. In the areas where bi-modality is the dominant force, $\tilde{E}_{cv}$ acts locally to fine tune segmented shape. The global geometry may therefore be driven by the prior. This flexible configuration is parameterised by γ_1 and γ_2. Such flexibility is of course required to make the algorithm adapt to the differing conditions in different scenes.

The quality of segmentation is compared to hand segmented images of the test sequence. For better demonstration, the figures display the given images, the images with the segmenting contour by the AR-based algorithm and the images with the segmenting contour by the algorithm proposed here.

6.5.1 Indoor cycles

In this example, a gait cycle captured inside lab is used. This video has few distracting factors, there is for instance, very limited clutter in the background. The scene is formed of chunks of quasi homogeneous regions. This however does not mean that this is an easy task for a segmentation algorithm. On contrary, there is still a problematic amount of shadow and difference in luminance levels. Furthermore, the target is formed mainly of two different regions.

In the first experiment, the scene is obstructed by a synthetic occlusion. From Figure (6.1(a)) it is noticeable that this algorithm is generating better outcomes compared to the AR-based algorithm. Clearly, the combination of the prior with the data term extracts the shape from the background. In the occluded area, the prior term takes full control of the segmentation. The AR-based algorithm, as explained in Chapter 4 depends on a prior model that seems insufficient. This can be concluded from the shapes produced which have some missing parts. Figure (6.1(b)) presents a quantitative comparison of the results of both approaches. In which the number of error pixel (in comparison with the ground truth) is used as measure. This graph backs the early observation about the robustness and accuracy of the proposed algorithm.

Also provided in Figures 6.2, 6.3, and 6.4, further examples of scenes containing occluded walking subjects in an indoor setting segmented by using this algorithm presented here.

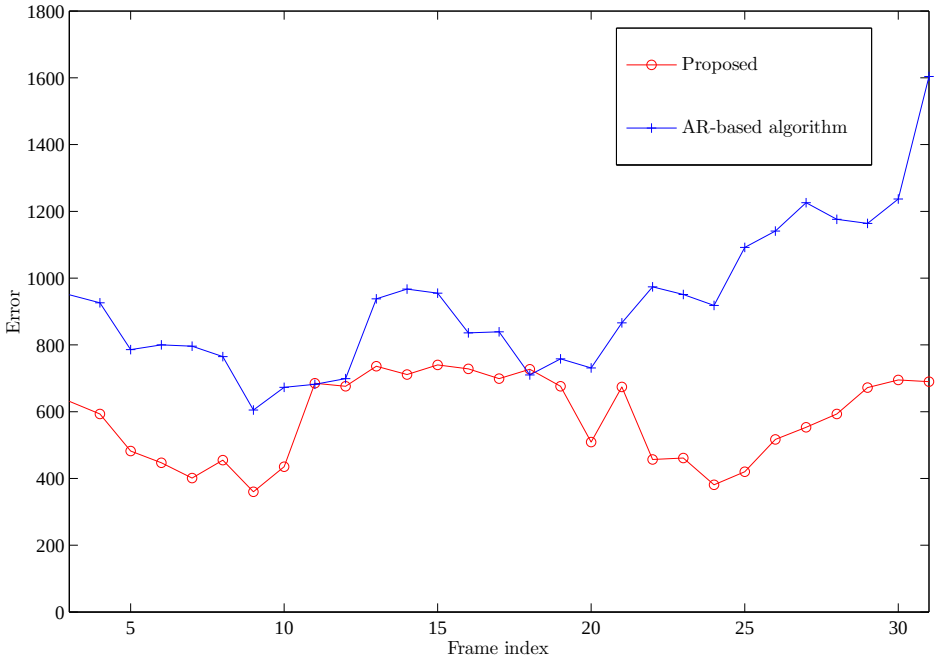
The second experiment is characterised by the inclusion of noise in addition to the occlusion. This is evidently a tough problem where the the region homogeneity is highly disturbed. Figure (6.5(b)) shows the scene glittered with Gaussian noise. This experiment is another example of the resilience of the algorithm proposed here. The example, nonetheless, can be regarded as a vote in favour of both compared algorithms as they managed to segment the target shapes. The graphs in Figure (6.5(a)) also supports this conclusion.

Further examples segmented by using this algorithm presented here are provided in Figures 6.6, 6.7, and 6.8, in which images of walking subjects are subjected to occlusion and Gaussian noise with SNR=1 in indoor settings.

6.5.2 Outdoor cycles

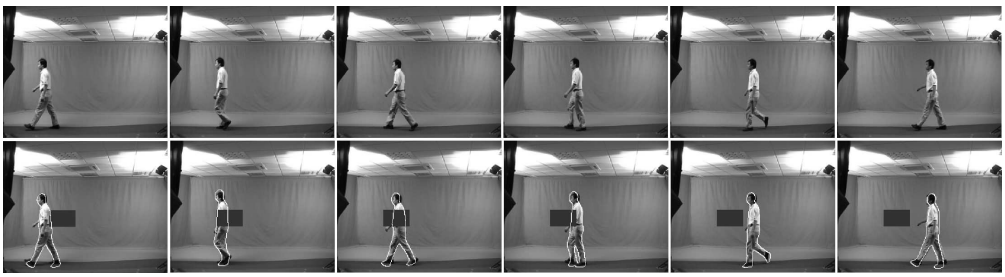
The experiment in this example is conducted on a sequence of images captured in an open place. The scene is rich in clutter and moving object. This is a typical real life visual challenge. In Figure (6.9(a)) it is shown that the AR-based

(a)

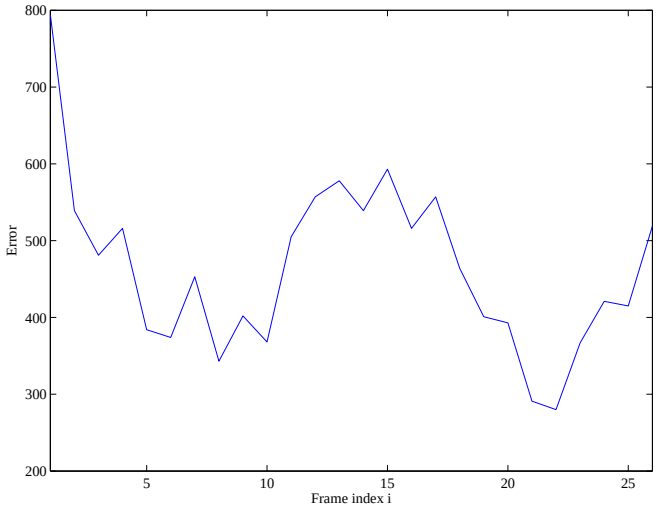


(b)

FIGURE 6.1: Indoor segmentation: occlusion, (a) the given sequence (top row), segmentation results by the AR-based algorithm (middle row), and the segmentation results by the algorithm proposed here (the bottom row) , (b) the number of error pixels in the segmented image sequence produced by each of the two algorithms



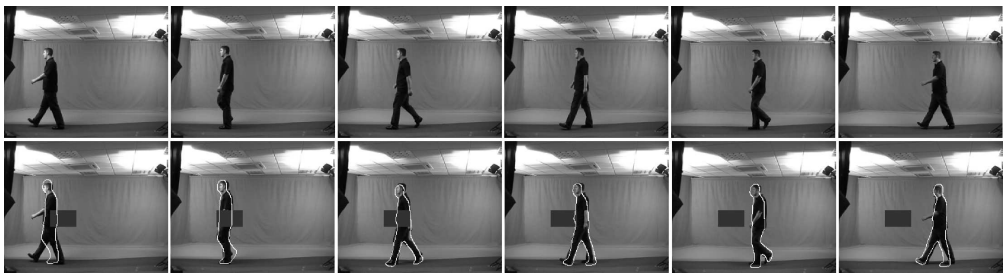
(a)



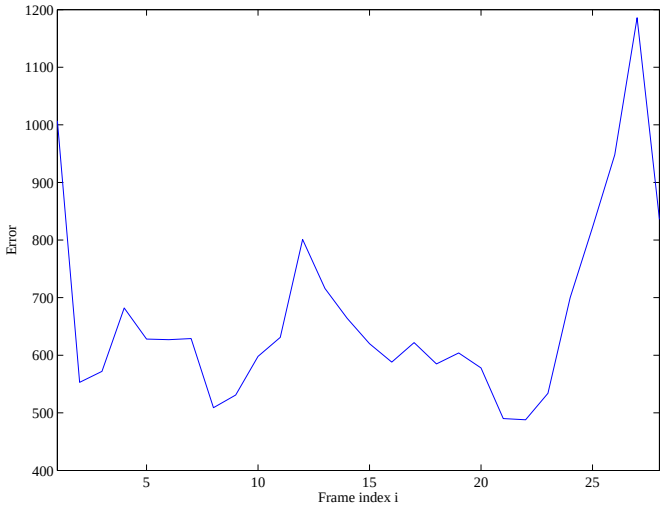
(b)

FIGURE 6.2: Indoor segmentation: occlusion, (a) the given sequence (top row), the segmentation results by using the algorithm proposed here (the bottom row) , (b) the number of error pixels in the segmented image sequence produced by this algorithms

algorithm starts to loose track of the target. This results in reduced performance and increase in the number of error pixels compared to the ground truth as seen in Figure (6.9(b)). Our algorithm on the other hand seems to hold well against the distractions presented by this scene. Interesting to mention that as compared with the indoor example, the error have dramatically increased here for both algorithms. However, as shown by the graphs in Figures (6.1(b), 6.5(b) and 6.9(b)), the AR-based algorithm has seen greater degradation in its segmentation output.



(a)



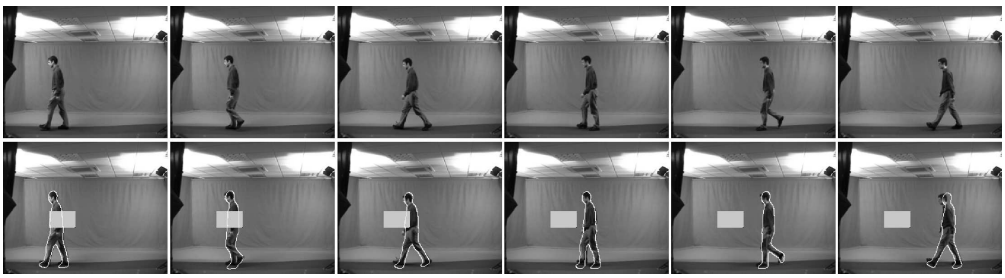
(b)

FIGURE 6.3: Indoor segmentation: occlusion, (a) the given sequence (top row), the segmentation results by using the algorithm proposed here (the bottom row) , (b) the number of error pixels in the segmented image sequence produced by this algorithms

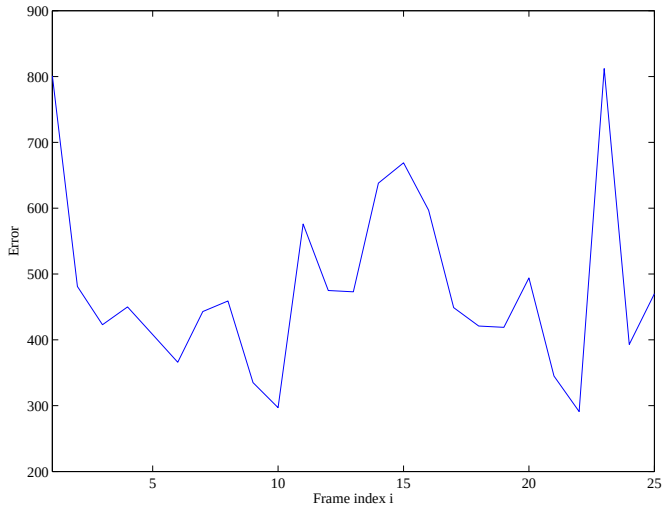
6.5.3 Noise analysis

The impact of various levels of noise on the quality of segmentation is investigated here. The indoor image sequence used in Section 6.5.1 is employed in this analysis. Increasing levels of Gaussian and uniform noise are added to the images of this sequence. The measure suggested by Cremers (2006) to quantify the quality of segmentation is used here as an indicator of the robustness. This measure is given in Eq. (6.11):

$$\acute{\epsilon} = \frac{\int_{x,y} \left(H(\phi) - H(\phi_0)\right)^2 \mathrm{d}x\mathrm{d}y}{\int_{x,y} H(\phi)\mathrm{d}x\mathrm{d}y + \int_{x,y} H(\phi_0)\mathrm{d}x\mathrm{d}y} \tag{6.11}$$



(a)



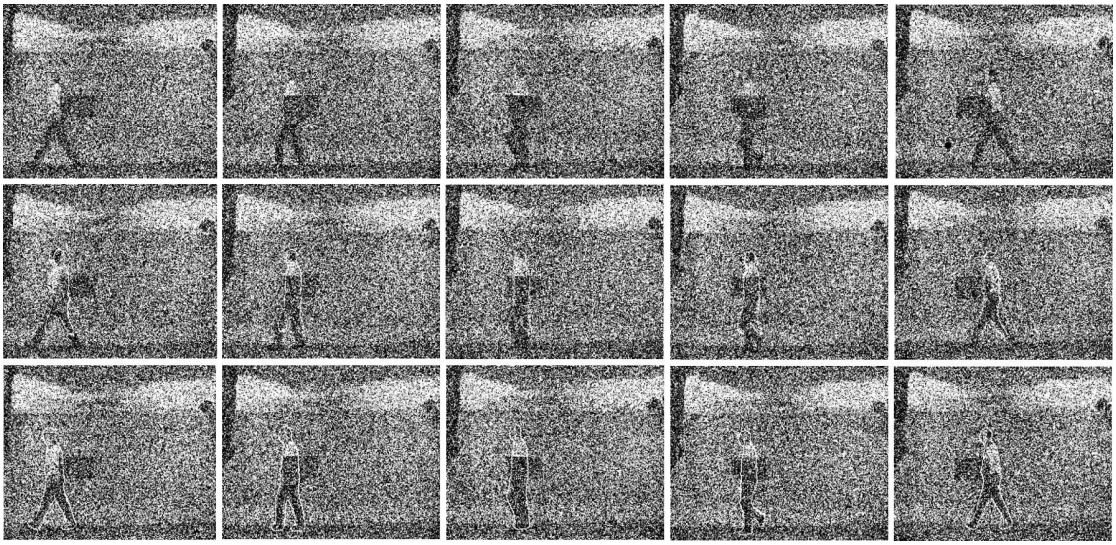
(b)

FIGURE 6.4: Indoor segmentation: occlusion, (a) the given sequence (top row), the segmentation results by using the algorithm proposed here (the bottom row) , (b) the number of error pixels in the segmented image sequence produced by this algorithms

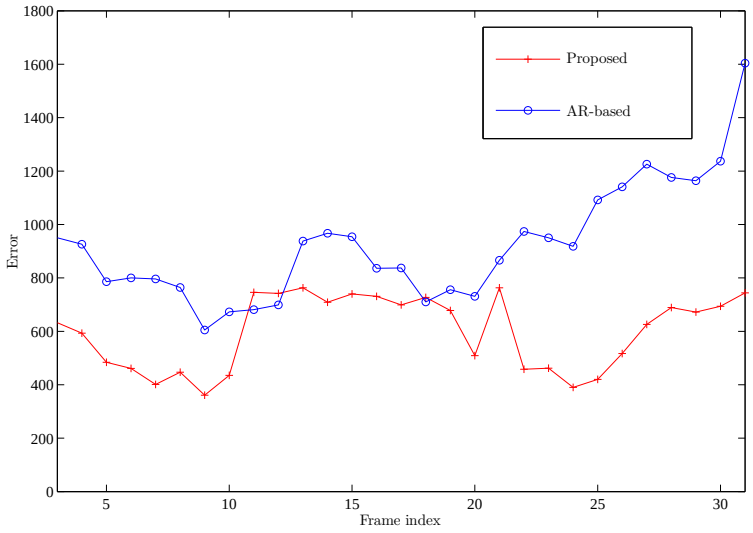
where H is again the Heaviside step function, ϕ_0 is the ground truth segmentation, and ϕ the shape segmented by the algorithm. This measure accounts for the relative area of the set symmetric difference and has a value between 0 and 1.

In Figure (6.10) examples of the shapes with various levels of noise are segmented by using this algorithm. It can be seen from this figure that in the high levels of noise most of the visual information is destroyed. Yet, it is still manageable by this algorithm. This is explained in better details in Figure (6.11).

Figure (6.11(a)) shows the segmentation error $\acute{\epsilon}$ curves resulting from different levels of noise. It shows in particular that overall the algorithm is steady. Figure (6.11(b)) demonstrates the average error $\acute{\epsilon}$ as a function of the noise level. As

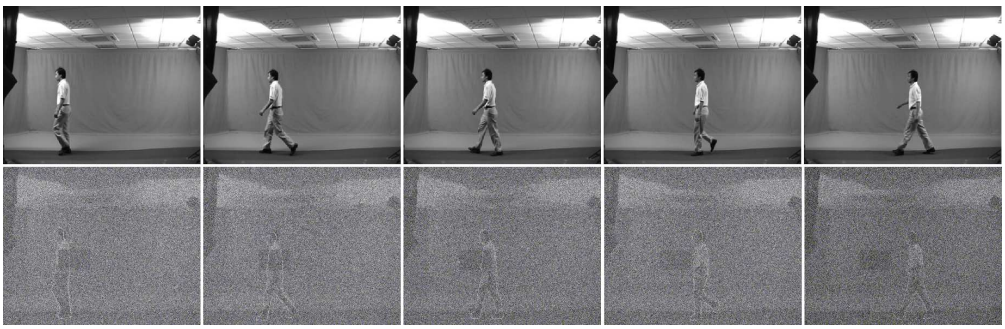


(a)

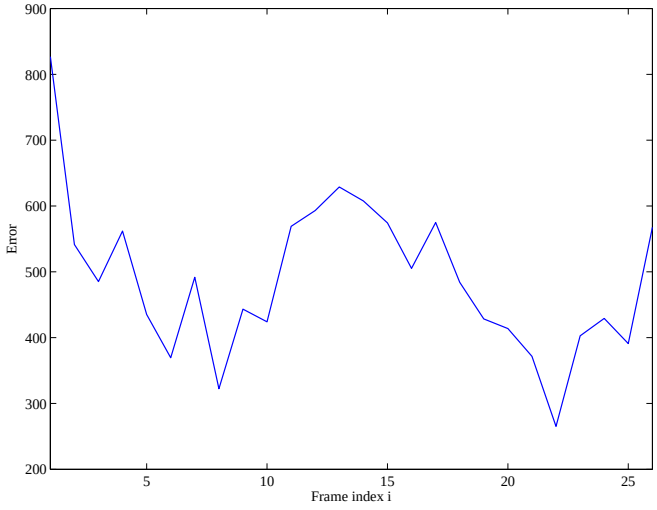


(b)

FIGURE 6.5: Indoor segmentation: occlusion and Gaussian noise with SNR=1, (a) the input sequence (top row), the segmentation results by using the AR-based algorithm (middle row), and the segmentation results by using this algorithm (bottom row), (b) the number of error pixels in the segmented images produced by the algorithms compared here



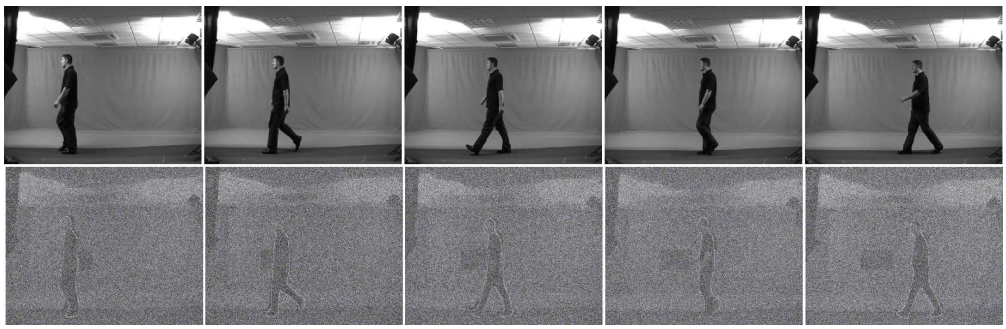
(a)



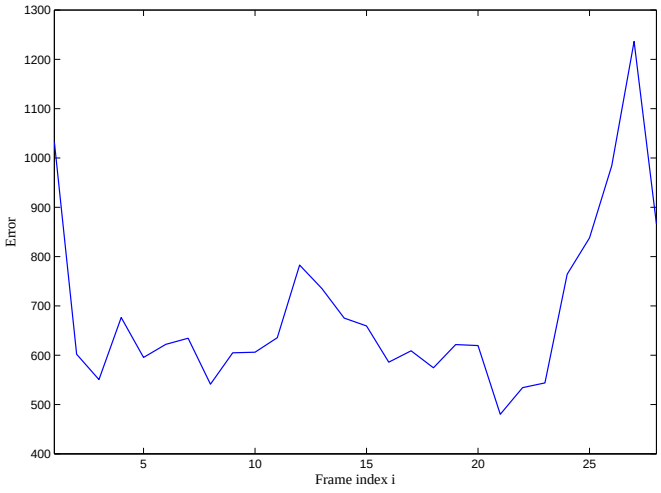
(b)

FIGURE 6.6: Indoor segmentation: occlusion and Gaussian noise with SNR=1, (a) the given sequence (top row), the segmentation results by using the algorithm proposed here (the bottom row) , (b) the number of error pixels in the segmented image sequence produced by this algorithms

is intuitive, the error increases with higher values of noise. This graph, however, shows that the error remains to some extent constant for noise levels below SNR=1. The residual error starts rising noticeably at values below SNR=0.8. The range of increase is, nevertheless, around 1 percent. This suggest clearly that this algorithm is very robust against noise.



(a)



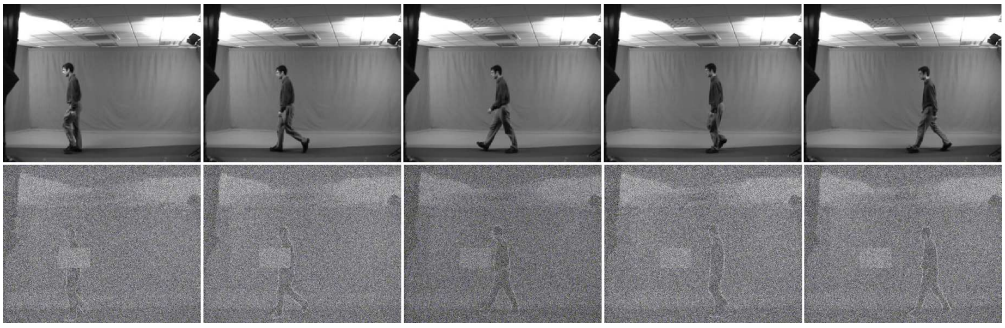
(b)

FIGURE 6.7: Indoor segmentation: occlusion and Gaussian noise with SNR=1, (a) the given sequence (top row), the segmentation results by using the algorithm proposed here (the bottom row) , (b) the number of error pixels in the segmented image sequence produced by this algorithms

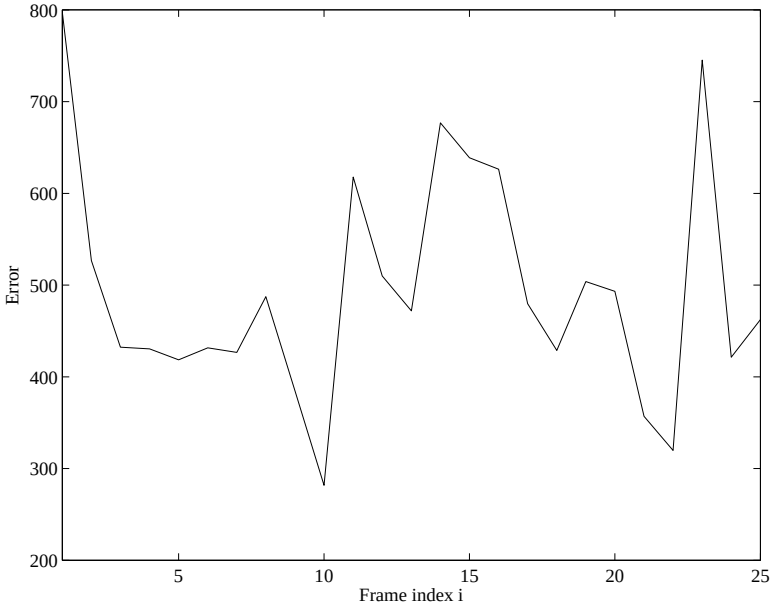
6.6 Conclusions

A gait segmentation algorithm with prior shape is presented in this chapter. This algorithm balances between classic active contour forces working on local level and prior shape governing the class of segmented shapes globally. The prior is derived by employing a framework of subsequent synthesis and analysis.

This new technique is evaluated by using a set of problems. These problems involve real-world image sequences of walking subjects. The images are further distorted by use of occlusion, noise, and clutter.



(a)



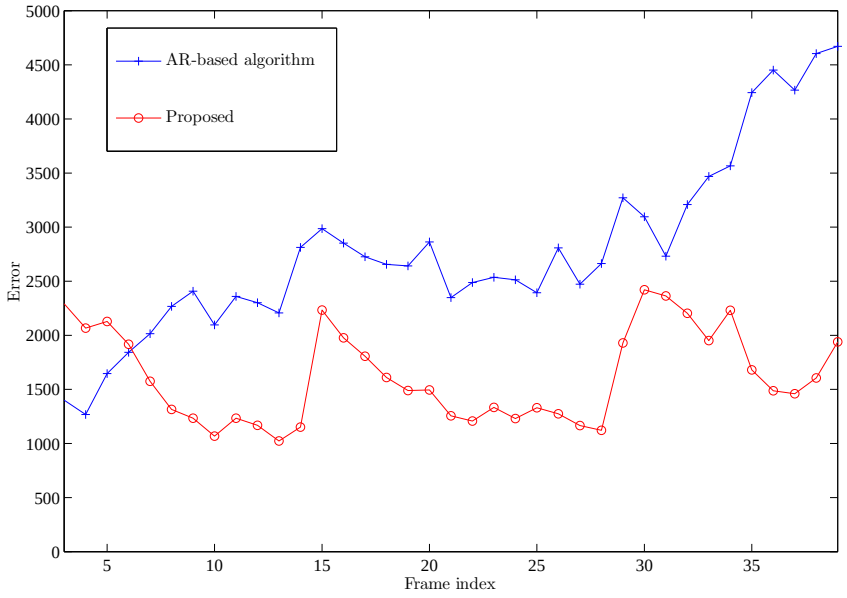
(b)

FIGURE 6.8: Indoor segmentation: occlusion and Gaussian noise with SNR=1, (a) the given sequence (top row), the segmentation results by using the algorithm proposed here (the bottom row) , (b) the number of error pixels in the segmented image sequence produced by this algorithms

The test results project the ability of this algorithm to perform good segmentation in all of the included distortions. Furthermore, the experiments demonstrate better results in comparison with previous approaches.

This algorithm can be further improved by including tracking methods such as Kalman filters. This is useful for instance in the extreme cases where the shape is occluded either partially or completely for a several frames. Such addition might increase resilience significantly.

(a)



(b)

FIGURE 6.9: Outdoor sequence segmentation example, (a) a sample of images before segmentation (top row), after segmentation by using AR-based algorithm (middle row), and by using this algorithm (bottom row), (b) the number of error pixels in the segmented images produced by these algorithms

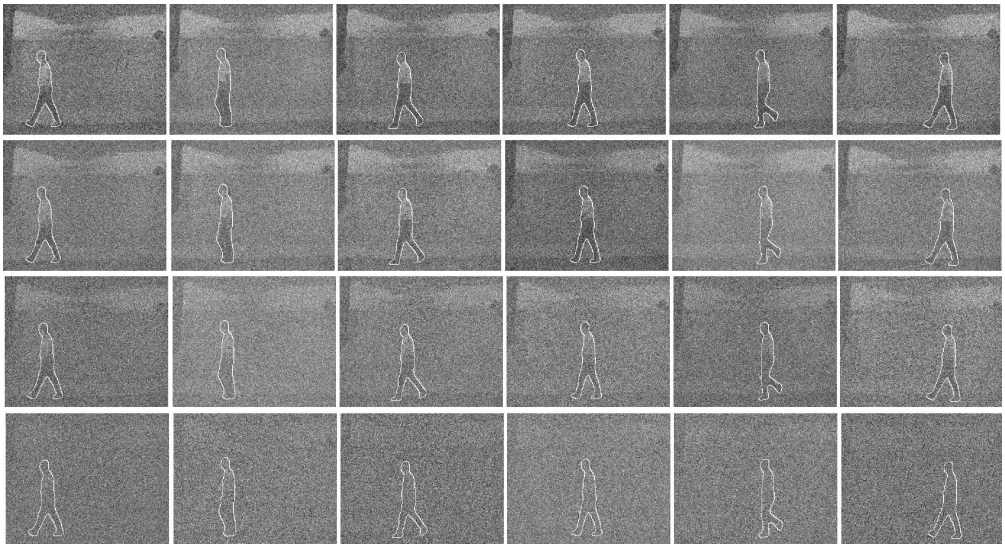
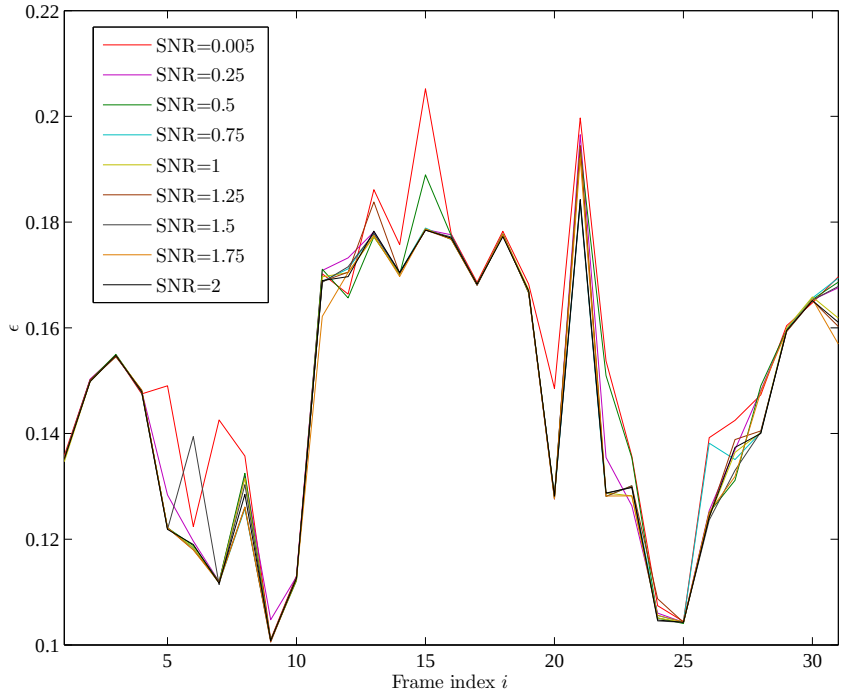
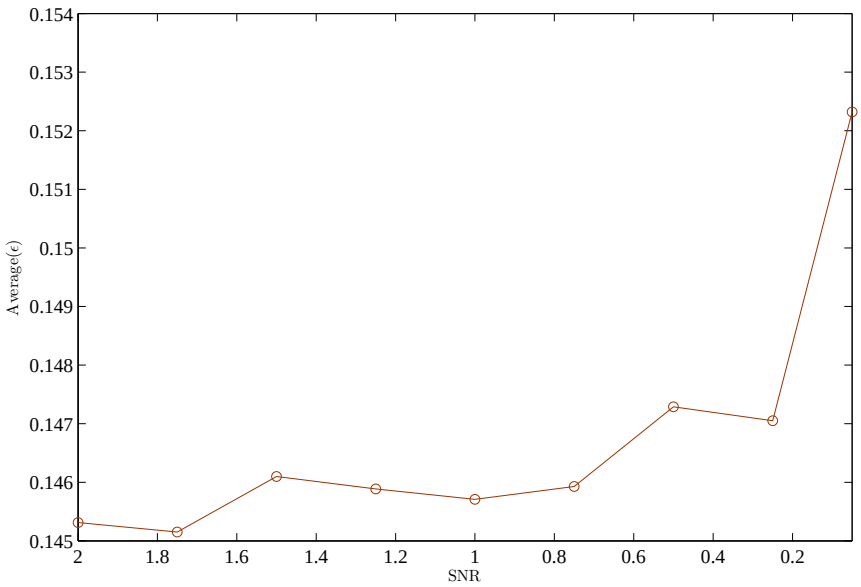


FIGURE 6.10: Some examples of image sequences contaminated with various levels of Gaussian noise: SNR=2 in the first row, SNR=1.25 in the second row, SNR=0.5 in the third row, and SNR=0.05 in the fourth row



(a)



(b)

FIGURE 6.11: Noise analysis, (a) the error ϵ in the segmented images as a function of different levels of noise, (b) a graph of the average values of ϵ plotted against the level of noise included in the images

Chapter 7

Conclusions and Further Work

7.1 Conclusions

This thesis communicates a set of algorithms aiming to advance the analysis and manipulation of temporally deforming shapes. In this thesis, shape features of entire cycles are packed as a single entity preserving the periodicity and time coherence. With the aid of an interpolating cubic spline, these features are matched in terms of length and phase. The correspondence of feature entities is automated.

The formed features are treated in a Gaussian framework by using PCA. Therefore, the manipulation of the model parameters leads to a holistic deformation pattern drawn over the entire sequence. The introduction of PCA over temporally matched shapes results in producing novel cycles with valid shapes. The generation of new cycles is governed by changing the model parameters which is a significant practical outcome of this model.

The generative model presented here is used successfully in reconstructing the in-between shapes which did not exist in the initial training set. Furthermore, this method is tested for its tolerance to missing parts of the training set for which the technique proves robust. One other application is to use the derived feature in the identification of pedestrian. A flavour of possible further investigation of this application can be found in the next section.

A shape registration algorithm is also contributed here. The proposed algorithm estimates the parameters by using closed form expressions. This algorithm employs Fourier transform properties to estimate the rotation and the translation parameters. The scale parameter is computed by using radial moments. This registration technique is tested by using a wide range of problems. The results show that this algorithm is fast (non-iterative), accurate, stable, and does not fall into local minima.

The final part of this thesis is devoted to the integration of the shape model with the ACWE segmentation functional. The algorithm suggested here to perform a prior shape informed segmentation also employs our shape registration technique. This algorithm is constructed such that the evolving contour takes its global geometrical structure from the prior. The data driven term brings the contour closer to the strong local forces in the data. This way the algorithm manages to overcome occlusion, ignore clutter and survive devastating levels of additive noise as demonstrated here.

The prior is estimated dynamically in an analysis-synthesis coupled iterations. Shapes known previously help identify the subject, which in turn help estimate the next prior. The prior guides the segmentation which subsequently provide an extra known shape for the next iteration.

There experimental results presented here show that the techniques presented in this work can help the progress in the field. Further research is needed to expand the findings of this report. The extension of the registration algorithm to 3D is foresee-ably possible as well as important. The inclusion of a tracking component in the segmentation algorithm can increase the reliability of the algorithm. Finally, there might be a good potential for the proposed gait features to be of use in the increasingly interesting gait biometrics area.

FIGURE 7.1: Heel strikes: the average image of the lower leg area of one of the videos (Arbab-Zavar et al., 2011)

7.2 Further Work

This section sheds some light on the potential use of the model proposed in Chapter 4 in gait biometrics. The experiment presented there shows on a very limited scale that the features derived are discriminative. This section, however, broadens the prospect by using a large database and including comparable classification techniques. This work was done in collaboration with B. Arbab-Zafar (Arbab-Zavar et al., 2011).

A sample from the Southampton Gait database (Shutler et al., 2002) is used. Videos were filmed for walking subjects laterally before a static camera at a rate of 25 frames per second in controlled conditions. We used a set of 1079 sequences from 116 subjects where the subject is moving from right to the left of the frame.

7.2.1 Single Cycle Detection

Gait cycles are detected automatically by detecting all the heel strike positions. Since the heels remain static while the rest of the body moves forward, it is expected that heels appear in the same position in the image for several frames. The successive frames are averaged and the heel strikes are detected by computing the highest intensity values beyond some threshold (see Figure 7.1). Then, the period of three successive frames where these positions first appear are considered a single gait cycle. An example of extracted gait cycles of a subject is shown in Figure 7.2.

FIGURE 7.2: Some extracted gait cycles of one of the recorded subject of the database (Arbab-Zavar et al., 2011)

7.2.2 Gait Feature Computation and Recognition

7.2.2.1 Gait Features

The sample is partitioned into two sets, $\mathbf{S}_{train} \equiv \{\mathbf{V}_1^1, \ldots, \mathbf{V}_1^M\}_{c \times M}$, where $\mathbf{V}_i^j$ is a vector of image sequence. The subscript, i, represent the cycle per subject and the superscript, j, represents the particular subject. M, is the total number of subjects and c is the average number of frames per all cycles. This set is used to train the regenerative model (i.e to compute $\bar{\mathbf{u}}$, Ψ, $\bar{\alpha}$, and χ) as detailed in chapter 4. The other (much bigger) set is $\mathbf{S}_{test} \equiv [\mathbf{V}_2^1, \ldots, \mathbf{V}_l^1, \ldots, \mathbf{V}_2^M, \ldots, \mathbf{V}_p^M]_{c \times r \times M}$, where r is the average number of cycles per subject. This set is used to generate the features $\vec{\beta}$ which are used in the classification:

$$\vec{\beta}_i^j = \chi^T(\vec{\alpha}_i^j - \bar{\alpha}), \quad \text{for } i = 2:r,\ j = 1:M. \tag{7.1}$$

where the eigenmode $\vec{\alpha}$ is computed as:

$$\vec{\alpha}_i^j(z) = \Psi^T(\mathbf{V}_i^j(z) - \bar{\mathbf{u}}) \tag{7.2}$$

where $(\mathbf{V}_i^j(z)$ is the z^{th} image of the sequence i of subject j.

7.2.2.2 Component and Discriminant Analysis

Due to the fact that gait image data (with high dimensionality) are problematic to classifiers. Researchers (Huang et al., 1999) combine PCA and Multiple Discriminant Analysis (MDA) to achieve the best data representation and class separability simultaneously.

Han and Bhanu (Han and Bhanu, 2006) proposed a simple spatio-temporal gait feature called Gait Energy Image (GEI). The authors employ PCA to reduce feature dimensionality while preserving the data, followed by MDA to maximise the ratio of between-class scatter (covariance) matrix and within-class scatter matrix.

By assuming Gaussian distributions for all the classes in the feature space with the same covariance matrix, Bayes optimal classifier becomes a minimum Mahalanobis distance classifier (Han and Bhanu, 2006),

$$D_M(\tilde{x}) = -\frac{1}{2}(\tilde{x} - \mu_i)^T \Sigma^{-1}(\tilde{x} - \mu_i) \tag{7.3}$$

where $\tilde{x}$ is a test feature, μ_i is the mean of the i^th class in the training sample, and Σ is the covariance matrix of the features.

The above described feature manipulation and classifier are employed for features computed by using the GEI and the model proposed here. However, since the features generated by our proposed gait model are in already low dimensionality (because of $\text{PCA}_{feature}$) PCA is of no use here. MDA is certainly useful and is employed in the same way as above.

7.2.3 Experiments

The data efficient leave-one-out cross validation is used here. Therefore, the classifier is trained by using one sample and tested by using the remaining samples in the dataset. Table 7.1 shows the correct classification rates (CCR) which is the number of test points classified correctly divided by the total number of points

TABLE 7.1: Leave-one-out cross-validation correct classification rates

Method	CCR (%)
GEI + PCA + MDA + Bayesian (no normalisation)	898/959 ($\approx$ 93.6%)
GEI + PCA + MDA + Bayesian (scale normalised)	947/959 ($\approx$ 98.7%
$\vec{\beta}$ + MDA + Bayesian (no normalisation)	928/959 ($\approx$ 96.8%)

in the data set. This experiment shows that the proposed features give better discriminant results than the GEI (Han and Bhanu, 2006) in the case of non normalised silhouettes.

Indeed, more scrutiny is required to involve data under different conditions such as scale normalisation, different viewpoints, noise and occlusion. Further, it is good to investigate the performance with randomly missing frames or lower temporal resolution, as reported earlier that this model can compensate for missing data. It is also interesting to study the use of other classification strategies.

Appendix A

Derivation of Euler Lagrange Equations

Let $F(c_1, c_2, \phi)$ be the functional of the ACWE model defined in Eq. (A.1):

$$\begin{aligned} F(c_1, c_2, \phi) = \quad & \mu \int_\Omega \delta(\phi(x,y)) \left|\vec{\nabla}\phi(x,y)\right| \mathrm{d}x\mathrm{d}y \\ + \quad & \lambda_1 \int_\Omega |u_0(x,y) - c_1|^2 H(\phi(x,y)) \, \mathrm{d}x\mathrm{d}y \\ + \quad & \lambda_2 \int_\Omega |u_0(x,y) - c_2|^2 (1 - H(\phi(x,y))) \, \mathrm{d}x\mathrm{d}y. \end{aligned} \tag{A.1}$$

The objective is to find ϕ that minimises F.

Therefore, the Lagrangian L of F is defined in Eq. (A.2):

$$\begin{aligned} L = \quad & \mu \left(\delta(\phi(x,y)) \left|\vec{\nabla}\phi(x,y)\right|\right) + \lambda_1 \left(|u_0(x,y) - c_1|^2 H(\phi(x,y))\right) \\ + \quad & \lambda_2 \left(|u_0(x,y) - c_2|^2 (1 - H(\phi(x,y)))\right) \end{aligned} \tag{A.2}$$

In what follows $\phi(x,y)$ and $u_0(x,y)$ will be respectively referred to as ϕ and as u_0 for short.

The function minimising F is the solution of the Euler-Lagrange Eq. (A.3):

$$\frac{\partial \phi}{\partial t} = \frac{\partial L}{\partial \phi} - \frac{\partial}{\partial x}\frac{\partial L}{\partial \acute{\phi}} - \frac{\partial}{\partial y}\frac{\partial L}{\partial \acute{\phi}} = 0. \tag{A.3}$$

Next, each term of this equation is computed individually. For $\frac{\partial L}{\partial \phi}$ this is given in Eq. (A.4):

$$\frac{\partial L}{\partial \phi} = \mu\ \delta_1(\phi)\left|\vec{\nabla}\phi\right| + \lambda_1(u_0(x,y) - c_1)^2\delta(\phi) + \lambda_2(u_0(x,y) - c_2)^2\delta(\phi) \tag{A.4}$$

For $\frac{\partial L}{\partial \acute{\phi}}$ this is given in Eq. (A.5)

$$\begin{aligned}\frac{\partial L}{\partial \acute{\phi}} &= \frac{\partial}{\partial \acute{\phi}}\left(\mu\ \delta(\phi)\left|\vec{\nabla}\phi\right|\right), \\ &= \mu\ \delta(\phi)\frac{\phi_x + \phi_y}{\sqrt{\phi_x^2 + \phi_y^2}}, \\ &= \mu\ \delta(\phi)\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}.\end{aligned} \tag{A.5}$$

From Eq. (A.5), $\frac{\partial}{\partial x}\frac{\partial L}{\partial \acute{\phi}}$ can be computed,

$$\frac{\partial}{\partial x}\frac{\partial L}{\partial \acute{\phi}} = \mu\ \left(\delta_1(\phi)\acute{\phi}_x\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|} + \delta(\phi)\frac{\partial}{\partial x}\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}\right). \tag{A.6}$$

Similarly $\frac{\partial}{\partial y}\frac{\partial L}{\partial \acute{\phi}}$ can be computed as in Eq. (A.7):

$$\frac{\partial}{\partial y}\frac{\partial L}{\partial \acute{\phi}} = \mu\ \left(\delta_1(\phi)\acute{\phi}_y\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|} + \delta(\phi)\frac{\partial}{\partial y}\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}\right). \tag{A.7}$$

Then we have,

$$\frac{\partial}{\partial x}\frac{\partial L}{\partial \acute{\phi}} + \frac{\partial}{\partial y}\frac{\partial L}{\partial \acute{\phi}} = \mu[\delta_1(\phi)\acute{\phi}_x\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|} + \delta_1(\phi)\acute{\phi}_y\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|} + \delta(\phi)\frac{\partial}{\partial x}\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|} + \delta(\phi)\frac{\partial}{\partial y}\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}]. \tag{A.8}$$

This is simplified in Eq. (A.9):

$$\frac{\partial}{\partial x}\frac{\partial L}{\partial \acute{\phi}} + \frac{\partial}{\partial y}\frac{\partial L}{\partial \acute{\phi}} = \mu\left(\delta_1(\phi)\frac{\vec{\nabla}\phi.\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|} + \delta(\phi)\ \nabla\left(\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}\right)\right). \tag{A.9}$$

This can be written as:

$$\frac{\partial}{\partial x}\frac{\partial L}{\partial \acute{\phi}} + \frac{\partial}{\partial y}\frac{\partial L}{\partial \acute{\phi}} = \mu\left(\delta_1(\phi)\frac{(\acute{\phi}_x)^2 + (\acute{\phi}_y)^2}{\left|\vec{\nabla}\phi\right|} + \delta(\phi)\ \nabla\left(\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}\right)\right), \tag{A.10}$$

or equivalently,

$$\frac{\partial}{\partial x}\frac{\partial L}{\partial \acute{\phi}} + \frac{\partial}{\partial y}\frac{\partial L}{\partial \acute{\phi}} = \mu\left(\delta_1(\phi)\frac{\left(\sqrt{(\acute{\phi}_x)^2 + (\acute{\phi}_y)^2}\right)^2}{\left|\vec{\nabla}\phi\right|} + \delta(\phi)\ \nabla\left(\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}\right)\right), \tag{A.11}$$

which gives:

$$\frac{\partial}{\partial x}\frac{\partial L}{\partial \acute{\phi}} + \frac{\partial}{\partial y}\frac{\partial L}{\partial \acute{\phi}} = \mu\left(\delta_1(\phi)\left|\vec{\nabla}\phi\right| + \delta(\phi)\ \nabla\left(\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}\right)\right). \tag{A.12}$$

Putting all parts together similar to Eq. (A.3), we have:

$$\begin{aligned}\frac{\partial\phi}{\partial t} &= \mu\ \delta_1(\phi)\left|\vec{\nabla}\phi\right| + \lambda_1(u_0 - c_1)^2\delta(\phi) - \lambda_2(u_0 - c_2)^2\delta(\phi) \\ &- \mu\ \delta_1(\phi)\left|\vec{\nabla}\phi\right| - \mu\ \delta(\phi)\nabla\left(\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|}\right).\end{aligned} \tag{A.13}$$

By deleting similar terms Eq. (A.13) gives rise to Eq. (A.14),

$$\frac{\partial \phi}{\partial t} = -\mu\, \delta(\phi) \nabla \cdot \left(\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|} \right) + \lambda_1 (u_0 - c_1)^2 \delta(\phi) - \lambda_2 (u_0 - c_2)^2 \delta(\phi). \qquad \text{(A.14)}$$

Rearranging,

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left(\mu\, \nabla \cdot \left(\frac{\vec{\nabla}\phi}{\left|\vec{\nabla}\phi\right|} \right) - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 \right) = 0. \qquad \text{(A.15)}$$

The PDE given in Eq. (A.15) is the minimiser of the functional $F(c_1, c_2, \phi)$. The solution to this equation is the evolving contour according to model proposed by Chan and Vese (2001).

Appendix B

Functional Parameters of ACWE

$$c_3 = \frac{1}{\sqrt{\left(\frac{\phi^n_{i+1,j}-\phi^n_{i,j}}{h}\right)^2 + \left(\frac{\phi^n_{i,j+1}-\phi^n_{i,j-1}}{2h}\right)^2}}. \tag{B.1}$$

$$c_4 = \frac{1}{\sqrt{\left(\frac{\phi^n_{i,j}-\phi^n_{i-1,j}}{h}\right)^2 + \left(\frac{\phi^n_{i-1,j+1}-\phi^n_{i-1,j-1}}{2h}\right)^2}}. \tag{B.2}$$

$$c_5 = \frac{1}{\sqrt{\left(\frac{\phi^n_{i+1,j}-\phi^n_{i-1,j}}{2h}\right)^2 + \left(\frac{\phi^n_{i,j+1}-\phi^n_{i,j}}{h}\right)^2}}. \tag{B.3}$$

$$c_6 = \frac{1}{\sqrt{\left(\frac{\phi^n_{i+1,j-1}-\phi^n_{i-1,j-1}}{2h}\right)^2 + \left(\frac{\phi^n_{i,j}-\phi^n_{i,j-1}}{h}\right)^2}}. \tag{B.4}$$

$$m = \frac{\triangle t}{h^2}\delta_\epsilon(\phi_{i,j})\mu. \tag{B.5}$$

$$C = 1 + m(c_3 + c_4 + c_5 + c_6). \tag{B.6}$$

where for all the above, h is the spatial step parameter.

Appendix C

Registration Algorithm

Algorithm1: Shape Registration Algorithm

Require: ϕ_p and ϕ_q {ϕ_p is the input reference shape, ϕ_q is the input observed shape}

$\hat{\theta} \Leftarrow$ the angle difference between ϕ_p and ϕ_q {by using Algorithm2}

Ensure: $\phi_{temp_1} \Leftarrow$ rotated ϕ_q by using $\hat{\theta}$ {Apply the computed rotation}

$\hat{s} \Leftarrow$ the scale difference between ϕ_p and ϕ_{temp_1} {by using Algorithm3}

Ensure: $\phi_{temp_2} \Leftarrow$ scaled ϕ_{temp_1} by using $\hat{s}$ {Apply the computed scale}

$\left[\hat{T}_x, \hat{T}_y\right] \Leftarrow$ the translation difference between ϕ_p and ϕ_{temp_2} {by using Algorithm3}

Ensure: $\acute{\phi}_q \Leftarrow$ tranlated ϕ_{temp_2} by using $[\hat{T}_x, \hat{T}_y]$ {Apply the computed translations}

return $\acute{\phi}_q$

%%%

Algorithm2: Rotation Angle Computation Algorithm

Require: ϕ_p and ϕ_q

$SampleAngle \Leftarrow$ the angular step size

Ensure: $\hat{\phi}_p \Leftarrow$ centralised ϕ_p {translation invariance by using Eq. (5.5)},

$\hat{\phi}_q \Leftarrow$ centralised ϕ_q {translation invariance by using Eq. (5.6)},

$\tilde{\phi}_p \Leftarrow$ map $\hat{\phi}_p$ to polar domain with angular step size $SampleAngle$,

$\tilde{\phi}_q \Leftarrow$ map $\hat{\phi}_q$ to polar domain with angular step size $SampleAngle$.

$\tilde{\phi}_p(\omega) \Leftarrow$ marginalise out ρ of $\hat{\phi}_p(\rho,\omega)$ {by using Eq. (5.7)}

$\tilde{\phi}_q(\omega) \Leftarrow$ marginalise out ρ of $\hat{\phi}_q(\rho,\omega)$ {by using Eq. (5.8)}

$\bar{\phi}_p \Leftarrow$ normalise $\tilde{\phi}_p$ {by using Eq. (5.9)}

$\bar{\psi}_p \Leftarrow$ DFT of $\bar{\phi}_p$ {according to Eq. (5.14)}

$\tilde{\psi}_q \Leftarrow$ DFT of $\tilde{\phi}_q$ {according to Eq. (5.15)}

$Phase \Leftarrow$ index of maximum of inverse DFT of $\bar{\psi}_p.\tilde{\psi}_q$ {according to Eq.(5.17)}

$\hat{\theta} \Leftarrow Phase \times SampleAngle \times 360$ {Convert the Phase shift to angle degrees}

return $\hat{\theta}$

%%%

Algorithm3: Scale Computation Algorithm

Require: ϕ_p and ϕ_q,

$Order \Leftarrow$ the degree of the shape momemnts,

$\hat{\phi}_p \Leftarrow$ centralise ϕ_p {translation invariance by using Eq. (5.5)},

$\hat{\phi}_q \Leftarrow$ centralise ϕ_q {translation invariance by using Eq. (5.6)}.

$MinTerm \Leftarrow$ derivative of Eq. (5.24) with respect to $\hat{s}$

for $i = 1 \rightarrow Order$ **do**

$M^i_{\hat{p}} \Leftarrow$ the i^{th} moment of $\hat{\phi}_p$

$M^i_{\hat{q}} \Leftarrow$ the i^{th} moment of $\hat{\phi}_q$

$\hat{s} \Leftarrow \hat{s} + MinTerm(M^i_{\hat{p}}, M^i_{\hat{q}}, i)$

end for

$\hat{s} \Leftarrow \exp \hat{s}$

return $\hat{s}$

%%%

Algorithm4: Translations Computation Algorithm

Require: ϕ_p and $\phi_q(x,y))$

$\bar{\phi}_p \Leftarrow$ normaliseϕ_p {by using Eq. (5.9)}

$\bar{\Theta}_p \Leftarrow$ DFT of $\bar{\phi}_p$ {according to Eq. (5.14)}

$\tilde{\Theta}_q \Leftarrow$ DFT of $\tilde{\phi}_q$ {according to Eq. (5.15)}

$[T_x\ T_y] \Leftarrow$ index of maximum of inverse DFT of $\bar{\Theta}_p.\tilde{\Theta}_q$ {accoring to Eq. (5.31)}

return $\hat{T}_x$ and $\hat{T}_y$

Appendix D

Application of PCA in Continuous Data

This is an example of the employment of PCA in modeling continuous functions. Let $f_i(x)$ be the set of functions defined in Eq. (D.1)

$$f_i(x) = \vartheta_i \, x^2, \tag{D.1}$$

where x and ϑ_i are defined in Eq. (D.2) and Eq. (D.3) respectively:

$$\forall x \quad : \quad x \in [-6, 6], \tag{D.2}$$

$$\vartheta_i \quad = \quad \{-2, -1, 0, 1, 3, 4, 5\}. \tag{D.3}$$

Figure (D.1) depicts the set of functions represented by $f_i(x)$. These functions are centralised by subtracting the mean $\bar{f}$ defined in Eq. (D.4),

$$\begin{aligned} \bar{f} &= \frac{1}{8} \sum_i \vartheta_i x^2, \\ \bar{f} &= 1.5x^2. \end{aligned} \tag{D.4}$$

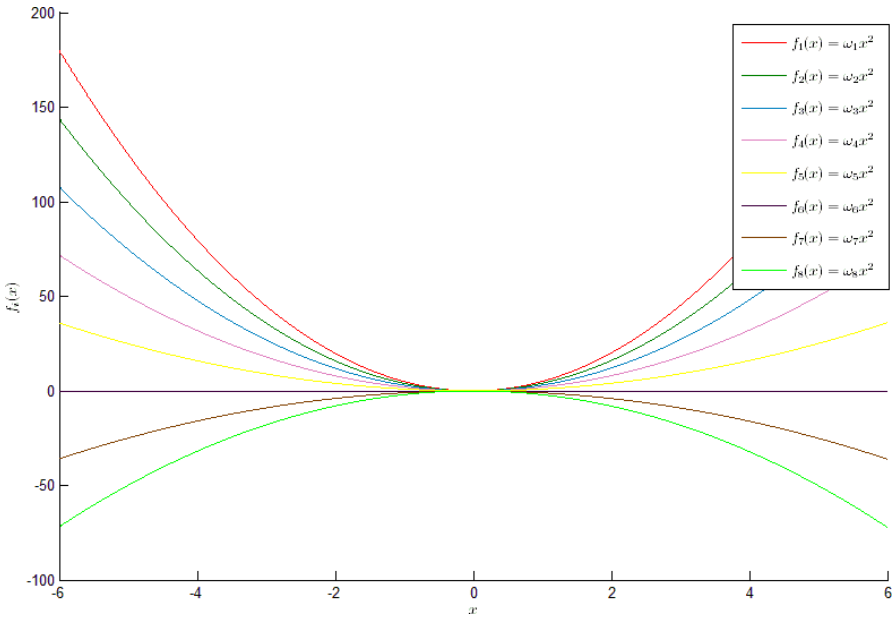


FIGURE D.1: A set of 8 functions f_i computed by using Eq. (D.1)

A matrix S which constitutes the set of functions f_i as its column elements is formed:

$$S(x) = [\ f_1\ f_2\ f_3\ f_4\ f_5\ f_6\ f_7\ f_8\ ]\,. \tag{D.5}$$

The covariance matrix SS^T has infinite dimensionality. Therefore, a kernel $K = S^T S$ (has dimensions 8×8) is formed instead. The components of K are computed as in Eq. (D.6):

$$\begin{aligned} k_{i,j} &= \vartheta_i \vartheta i \int_x x^4\ \mathrm{d}x, \\ &= \vartheta_i \vartheta i \left(\frac{x^5}{5} |_{x=-6}^{6} \right), \\ &= 3110.4\ \vartheta_i \vartheta_j. \end{aligned} \tag{D.6}$$

Hence, the set of eigenvalues ι and the corresponding eigenvectors o of K are computed according to Eq. (D.7),

$$o\ \iota\ o^T = K. \tag{D.7}$$

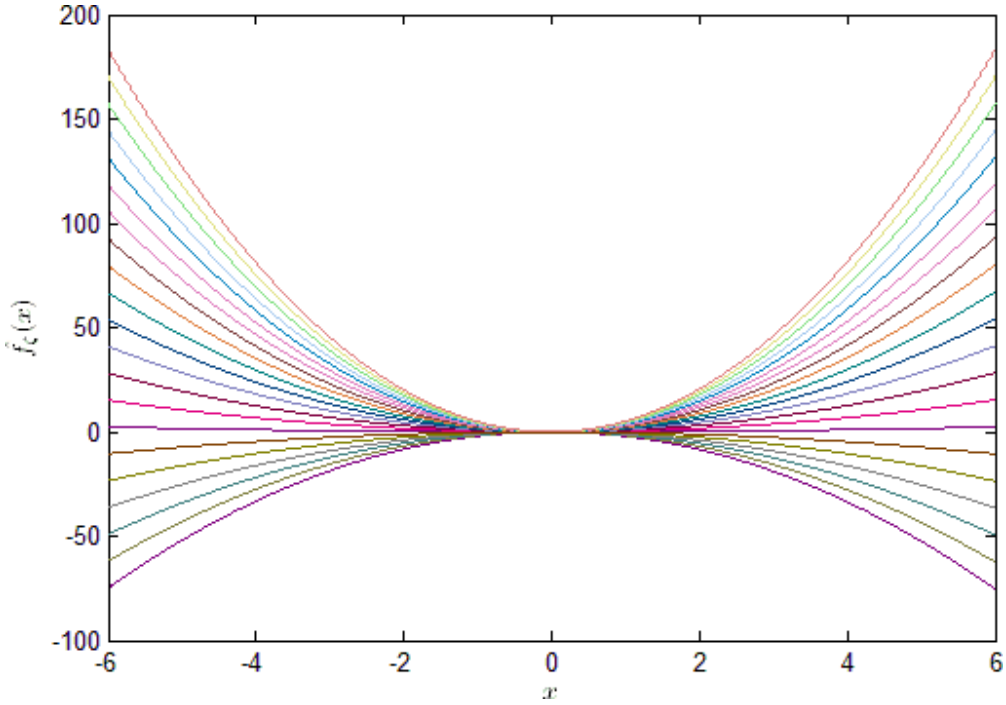


FIGURE D.2: A set of 21 functions $\hat{f}_\zeta$ generated by using Eq. (D.9) and setting $\zeta = \{-200, -180, -160 \cdots, 160, 180, 200\}$ respectively

Accordingly, the eigenvector basis of the covariance matrix of training set f_i is given in Eq. (D.8),

$$\chi = S\ o. \tag{D.8}$$

A novel data point (function of x), $\hat{f}$, belonging to the class of the training sample is generated (estimated) by using Eq. (D.9):

$$\hat{f}_\zeta = \zeta\ \chi + \bar{f}, \tag{D.9}$$

where ζ is some parameter (has an arbitrary value). Figure (D.2) and Figure (D.3) show some functions, $\hat{f}_\zeta(x)$, generated by setting various values to ζ.

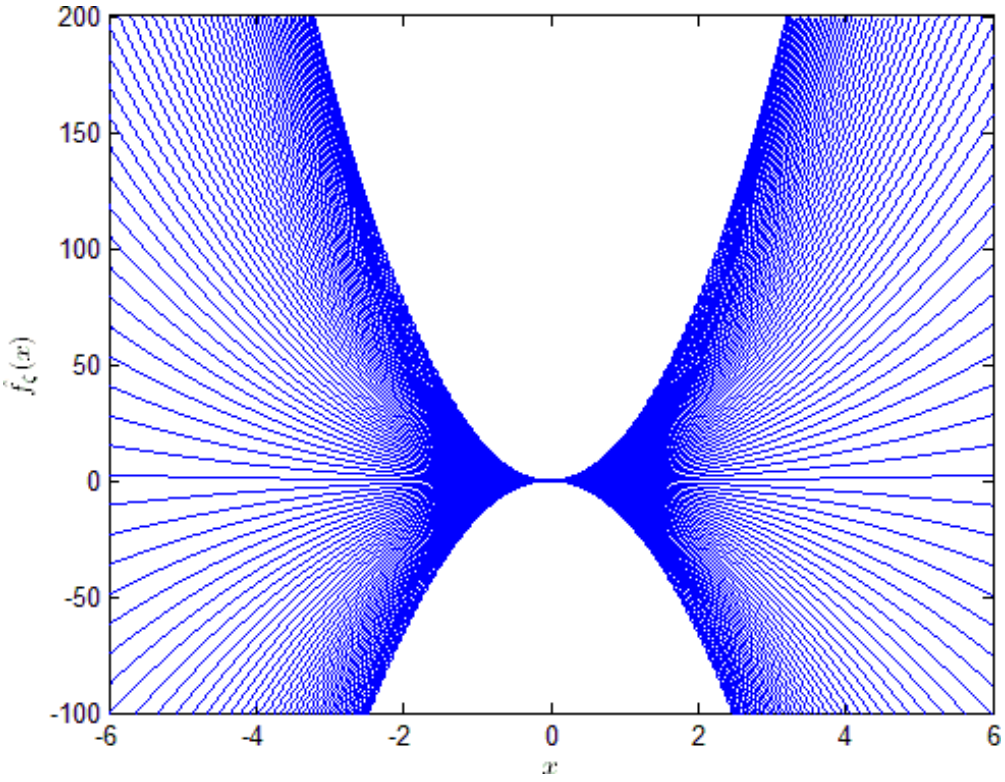


FIGURE D.3: A set of 101 functions $\hat{f}_\zeta$ generated by using Eq. (D.9). The values of ζ are chosen as: $\zeta = \{-1000, -980, -960 \cdots, 960, 980, 1000\}$ respectively

Appendix E

Algorithmic Prior Segmentation

$SequenceSegmentation(u_1, u_2, \ldots, u_n) \Longrightarrow [\phi_1, \phi_2, \ldots, \phi_n]$:

1. Initialise: $i = t-1$, $Shape_{off}$, $KnownShapes := [\phi_1, \ldots, \phi_{t-1}]$, ImageSequence:= $[u_j : j = (1, 2, \ldots n)]$, where $t-1$ is the number of known (given) shapes, and $Shape_{off}$ is the average offset of the indices of the given shapes,
2. Update identity of the subject (analysis), $\vec{\beta}$: $\beta etaUpdate(\vec{\beta}_{t-1}, [\phi_1, \ldots, \phi_{t-1}]) \Rightarrow \vec{\beta}_t$,
3. Set $i = i + Shape_{off}$,
4. Compute next prior shape (synthesis), ϕ^{Pr}:
   (a) compute $\vec{\alpha}_t := \chi\, \vec{\beta}_t + \bar{\alpha}$ (by using Eq. (4.11)),
   (b) compute $\phi^{Pr} = \Psi^T\, \vec{\alpha}_t(i) + \bar{\phi}$ (by using Eq. (6.9)),
5. Segment image u_t: $ImageSegment(u_t, \phi^{Pr}, \phi_{t-1}) \Rightarrow \phi_t$,
6. Add ϕ_t to the set of know shapes, $KnownShapes := [\phi_1, \ldots, \phi_{t-1}, \phi_t]$,
7. If $t \leq n$ go to step 2. Otherwise, Return $[\phi_1, \phi_2, \ldots, \phi_n]$.

$\beta etaUpdate(\vec{\beta}_{t-1}, [\phi_1, \phi_2, \ldots, \phi_{t-1}]) \Longrightarrow \vec{\beta}_t$:

1. Re-Compute a provisional entire gait cycle by using the previous subject identity $\vec{\beta}_{t-1}$:

   (a) $\hat{\alpha}_{t-1} := \chi\, \vec{\beta}_{t-1} + \bar{\alpha}$, (by using Eq. (4.11)),

   (b) $\hat{\phi}_{1:n} := \Psi\, \hat{\alpha}_{t-1} + \bar{\phi}$, (by using Eq. (6.9)),

2. Assign new sequence indices to the given shapes $[\phi_1, \ldots, \phi_{t-1}]$ by matching them to the closest shapes in the provisional cycle $[\hat{\phi}_1, \ldots, \hat{\phi}_n]$:
   $k = \underset{l}{\operatorname{argmin}} \sum_{l=1}^{n} \int_{x,y} |\phi_k - \hat{\phi}_l|^2 \mathrm{d}x\mathrm{d}y$,

3. Compute the eigenmodes corresponding to the (re-indexed) given shapes $\vec{\alpha}_{t-1}(k) := \Psi^T(\phi(k) - \bar{\phi})$ (by using Eq. (4.3)),

4. Fill the gaps in the newly constructed eigenmodes $\vec{\alpha}_{t-1}$ with zeros (thus, considering them as missing completely at random (Howell, 2009)),

5. Compute the updated subject identity, i.e., $\vec{\beta}_t := \chi^T(\vec{\alpha}_{t-1} - \bar{\alpha})$, (by using Eq. (6.7))

6. Return $\vec{\beta}_t$.

ImageSegment$(u_t, \phi^{Pr}, \phi_{t-1}) \Longrightarrow \phi_t$:

1. Initialise the evolving SDF $\phi := \phi_{t-1}$,

2. Iterate Eq. (6.2) until convergence,

3. Return ϕ_t.

Bibliography

Muayed S. Al-Huseiny, Sasan Mahmoodi, and Mark S. Nixon. Level set gait analysis for synthesis and reconstruction. In *Advances in Visual Computing*, volume 5876 of *Lecture Notes in Computer Science*, pages 377–386. Springer Berlin / Heidelberg, 2009.

Muayed S. Al-Huseiny, Sasan Mahmoodi, and Mark S. Nixon. Gait learning-based regenerative model: A level set approach. In *International Conference on Pattern Recognition*, pages 2644–2647, 2010.

N. M. Alpert, J. F. Bradshaw, D. Kennedy, and J. A. Correia. The principal axes transformation–a method for image registration. *Journal of Nuclear Medicine*, 31(10):1717–1722, 1990.

Banafshe Arbab-Zavar, Muayed S. Al-Huseiny, and Sasan Mahmoodi. Gait recognition based on the regenerative model. Technical report, University of Southampton, 2011.

A. Blake, B. Bascle, M. Isard, and J. MacCormick. Statistical models of visual shape and motion. *Proceedings of the Royal Society of London*, 356:1283–1302, 1998.

U. Brechtken-Manderscheid. *Introduction to the calculus of variations*. Chapman & Hall mathematics. Chapman & Hall, 1991. ISBN 9780412367007.

Xavier Bresson, Pierre Vandergheynst, and Jean-Philippe Thiran. A variational model for object segmentation using boundary information and shape prior driven by the mumford-shah functional. *International Journal of Computer Vision*, 68:145–162, June 2006.

Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24:325–376, December 1992.

D. Casasent and D. Psaltis. Position, rotation and scale invarient optical correlation. *Applied Optics*, 15:1795–1799, July 1975.

Vicent Caselles, Ron Kimmel, and Guillermo Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22:61–79, February 1997.

John Chambers, William Cleveland, Beat Kleiner, and Paul Tukey. *Graphical Methods for Data Analysis*. Wadsworth, 1983.

T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266 –277, feb 2001.

Tony F. Chan and Jianhong (Jackie) Shen. *Image Processing and Analysis*. Society for Industrial and Applied Mathematics, Philadephia, PA, 2005.

Weiqing Chen, Zongying Ou, and Weiwei Song. A coarse-to-refined approach of medical image registration based on combining mutual information and shape information. In *International Conference on Neural Networks and Brain, 2005. ICNN B '05.*, 2005.

Yunmei Chen, Hemant D. Tagare, Sheshadri Thiruvenkadam, Feng Huang, David Wilson, Kaundinya S. Gopinath, Richard W. Briggs, and Edward A. Geiser. Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50:315–328, December 2002.

T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models—their training and application. *Computer Vision and Image Understanding*, 61 (1):38–59, 1995.

Daniel Cremers. Dynamical statistical shape priors for level set-based tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(8):1262–1273, 2006.

Daniel Cremers, Timo Kohlberger, and Christoph Schnrr. Nonlinear shape statistics in Mumford Shah based segmentation. In *European Conference on Computer Vision*, 2002.

Daniel Cremers, Stanley J. Osher, and Stefano Soatto. Kernel Density Estimation and Intrinsic Alignment for Knowledge-Driven Segmentation: Teaching Level Sets to Walk. In *Lecture Notes on Computer Science*, pages 36–44. Springer, 2004.

Daniel Cremers, Stanley J. Osher, and Stefano Soatto. Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *International Journal of Computer Vision*, 69:335–351, 2006.

E. De Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:700–703, September 1987.

Lei Dong and Arthur L Boyer. A portal image alignment and patient setup verification procedure using moments and correlation techniques. *Physics in Medicine and Biology*, 41(4):697, 1996.

H.A. El Munim and A.A. Farag. Shape representation and registration using vector distance functions. In *IEEE Conference on Computer Vision and Pattern Recognition, 2007.*, June 2007.

F. Eugenio and F. Marques. Automatic satellite image georeferencing using a contour-matching approach. *IEEE Transactions on Geoscience and Remote Sensing*, 41(12):2869–2880, December 2003.

Ju Han and Bir Bhanu. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28:316–322, February 2006.

David C. Howell. Treatment of missing data [online resource retrieved 2009]. `http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/Missing.html`, July 2009.

P.S. Huang, C.J. Harris, and Mark S. Nixon. Recognising humans by gait via parametric canonical space. *Artificial Intelligence in Engineering*, 13(4):359 – 366, 1999.

M. Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

M. E. Leventon, W. E. L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 316–323, 2000.

Hui Li, B.S. Manjunath, and S.K. Mitra. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 4:320–334, 1995.

S. Mahmoodi. Shape-based active contours for fast video segmentation. *IEEE Signal Processing Letters*, 16(10):857–860, 2009.

J.B.Antoine Maintz and Max A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1 – 36, 1998.

I. Markovsky and S. Mahmoodi. Least-squares contour alignment. *IEEE Signal Processing Letters*, 16(1):41–44, January 2009.

Jorges S. Marques and Arnaldo J. Abrantes. Shape alignmentoptimal initial point and pose estimation. *Pattern Recognition Letters*, 18:49–53, January 1997.

R. Mukundan. A comparative analysis of radial-tchebichef moments and zernike moments. In *British Machine Vision Conference*, 2009.

D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.

David Mumford. Pattern theory: a unifying perspective. In David C. Knill and Whitman Richards, editors, *Perception as Bayesian inference*, pages 25–62. Cambridge University Press, New York, NY, USA, 1996.

Arnold Neumaier and Tapio Schneider. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software*, 27:27–57, March 2001.

NIST/SEMATECH. *NIST/SEMATECH e-Handbook of Statistical Methods*. NIST, USA, 2011.

Mark S. Nixon and Alberto Aguado. *Feature Extraction and Image Processing, Second Edition*. Academic Press, second edition, 2008.

Mark S. Nixon and J. N. Carter. Automatic recognition by gait. *Proceedings of The IEEE*, 94:2013–2024, 2006.

Alan V. Oppenheim, Ronald W. Schafer, and John R. Buck. *Discrete-time signal processing (2nd ed.)*. Prentice-Hall, Inc., 1999.

Stanley Osher and James A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-Jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.

G. Papandreou and P. Maragos. A fast multigrid implicit algorithm for the evolution of geodesic active contours. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages 689–694, 2004.

Nikos Paragios, Mikaël Rousson, and Visvanathan Ramesh. Matching distance functions: A shape-to-area variational approach for global-to-local registration. In *European Conference on Computer Vision*, pages 775–789, 2002.

Nikos Paragios, Mikaël Rousson, and Visvanathan Ramesh. Non-rigid registration using distance functions. *Computer Vision and Image Understanding*, 89(2–3): 142–165, 2003.

William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes with Source Code CD-ROM 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

S. Prismall, Mark S. Nixon, and J. Carter. Novel temporal views of moving objects for gait biometrics. In *Audio-and Video-Based Biometrie Person Authentication, 4th International Conference*, pages 725–733, 2003.

Mikaël Rousson and Rachid Deriche. Adaptive segmentation of vector valued images. In Stanley Osher and Nikos Paragios, editors, *Geometric Level Set Methods in Imaging, Vision, and Graphics*, chapter 11, pages 95–205. Springer New York, 2003.

Frank R. Schmidt and Daniel Cremers. A closed-form solution for image sequence segmentation with dynamical shape priors. In *Proceedings of the 31st DAGM Symposium on Pattern Recognition*, 2009.

J. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings on National Academy of Science USA*, 93(4):1591–1595, 1996.

J. Shutler, M. Grant, Mark S. Nixon, and J. Carter. On a large sequence-based human gait database. In *International Conference on Recent Advances Soft Computation*, pages 66–71, 2002.

Andy Tsai, Anthony Yezzi, William Wells, Clare Tempany, Dewey Tucker, Ayres Fan, W. Eric Grimson, and Alan Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2):137–154, 2003.

U. S. National Library Of Medicine. The second national health and nutrition examination survey (NHANES II), November 2010.

B. C. Vemuri, J. Ye, Y. Chen, and C. M. Leonard. Image registration via level-set motion: applications to atlas-based segmentation. *Medical Image Analisys*, 7 (1):1–20, March 2003.

Luminita A. Vese and Tony F. Chan. A multiphase level set framework for image segmentation using the mumford and shah model. *International Journal of Computer Vision*, 50:271–293, 2002.

Paul A. Viola and William M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24:137–154, 1997.

P. Wessa. Tukey lambda ppcc plot (v1.0.2) in free statistics software (v1.1.23-r7), 2011.

Xianghua Xie and M. Mirmehdi. Mac: Magnetostatic active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(4):632 –646, april 2008.

Alan L. Yuille. Deformable templates for face recognition. *Journal of Cognitive Neuroscience*, 3:59–70, January 1991.

Charles T. Zahn and Ralph Z. Roskies. Fourier descriptors for plane closed curves. *IEEE Transactions on Computer*, 21:269–281, March 1972.

Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977 – 1000, 2003.