IMAGE RECONSTRUCTION BY DOMINANT POINT DETECTION AND LEAST SQUARES RATIONAL QUADRATIC BÉZIER

LIM AI HUI

UNIVERSITI SAINS MALAYSIA

2019

IMAGE RECONSTRUCTION BY DOMINANT POINT DETECTION AND LEAST SQUARES RATIONAL QUADRATIC BÉZIER

by

LIM AI HUI

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

February 2019

ACKNOWLEDGEMENT

Doing the research to fulfill the requirement for the degree of PhD is one of the most challenging task in my life. I have struggled for years like other PhD students and tried my best to complete the thesis. The credit should not belongs to me alone but to many great persons who have sacrificed their precious time and effort in helping me along the way of my study.

Many thanks to Dr. Kong Voon Pang for giving me the excellent guidance in completing my thesis within limited duration. Thanks to him for his willingness to guide me in picking up certain related concept of mathematics and programming skill. Without his effort, I think the thesis is impossible to be done. I feel gratitute to him for the moral support and the motivating consultancy sessions which allowed me to see the bright side of my problem.

Not forget to thank the Bahagian Tajaan, KPM who provided financial support during the first three years of my study. I would not be able to start the study without the scholarship provided by KPM. Beside, thanks to the staffs in PPSM who gave good cooperation during my studies in USM. Thanks to Mr. Cheah Kah Wai who was willing to share his precious knowledge in LaTeX which has provides a convenient way to write my thesis.

Last but not least, a great gratitute credited to my beloved family and friends. Their support during my time struggling with the research is the key for me to continue without giving up. Thanks to my mother who takes excellent care of my little daughter during my study. Her real sense of commitment makes me feel worry-free during the hard time. Thank you.

TABLE OF CONTENTS

Acknowledgement ii

Table of Contents iii

List of Tables vi

List of Figures vii

List of Abbreviations xii

List of Symbols xiii

Abstrak xvii

Abstract xix

CHAPTER 1 – INTRODUCTION

1.1 Literature Review 2

1.2 Motivation 14

1.3 Problem Statements 15

1.4 Scope 16

1.5 Objectives 16

1.6 Methodology 17

1.7 Thesis Organization 18

CHAPTER 2 – PRELIMINARY WORK

2.1 Boundary Extraction 22

2.2 Region of Support 31

2.2.1 Eigenvalues of covariance matrices 32

2.2.2 Principal component analysis 34

2.3 Dominant Point Detection 36

2.3.1 *k*-Cosine 37

2.4 Curve Fitting 39

2.4.1 Continuity 41

CHAPTER 3 – REGION OF SUPPORT

3.1 Unit Bending Value 45

3.2 Marji & Siy Method 50

3.3 Eigen-Ratio 52

3.4 Eccentricity 55

3.5 Convex Hull Area-Length Ratio 58

3.6 Results 62

CHAPTER 4 – DOMINANT POINT DETECTION

4.1 Eigenvalues of Covariance Matrices in Dominant Point Detection 73

4.2 λ_C in Dominant Point Detection 75

4.3 λ_I in Dominant Point Detection 78

4.4 Comparison of λ_S, λ_R, λ_C and λ_I 80

4.5 Best Fit Vectors in k-Cosine Method 82

4.6 Improvement on Dominant Point Detection Methods 85

4.6.1 Additional dominant points by perpendicular distance error 86

4.6.2 Additional dominant points by ratio of significance 88

4.6.3 Suppress redundant dominant points 92

4.7 Performance Evaluation 92

4.8 Results 94

CHAPTER 5 – CURVE FITTING

5.1 Point of Inflection 99

5.2 Curve Fitting with Linear Segments 108

5.3 Curve Fitting with Circular Arc 109

5.3.1 A drawback of circular arc fitting 113

5.4 Curve Fitting with Rational Quadratic Bézier Curve 116
5.4.1 Parametrization 117
5.4.2 Least-square rational quadratic Bézier curve 118
5.4.3 G^1 continuity least-square curve 125
5.5 Implementation 126
5.6 Result 130

CHAPTER 6 – RESULT AND DISCUSSION

6.1 Results 135
6.2 Discussion 150

CHAPTER 7 – CONCLUSION

7.1 Summary and Conclusion 153
7.2 Limitation 155
7.3 Recommendations for Future Work 156

REFERENCES 158

APPENDICES

LIST OF PUBLICATIONS

LIST OF TABLES

		Page
Table 3.1	Size of support region for a corner point marked with "□".	63
Table 3.2	Size of support region for a non-corner point marked with "△".	63
Table 3.3	Size of support region for a non-corner point marked with "◯".	63
Table 4.1	λ_S, λ_R, λ_C and λ_I for semi-ellipses P, Q, R, S and T given in Figure 4.6.	82
Table 4.2	Mean error and compression ratio of rectangular shape.	96
Table 4.3	Mean error and compression ratio of curved shape.	96
Table 5.1	Data A.	113
Table 5.2	Data B_1.	121
Table 5.3	Data B_2.	121
Table 5.4	Mean error of "cat" data.	132
Table 6.1	Quantitative analysis for dominant point detection of sample data.	137
Table 6.2	Quantitative analysis for curve-fitting of sample data.	138
Table A.1	Coordinates boundary data of sample image "alphabet T", "gear", "calligraphy da" and "calligraphy niu".	163
Table B.1	Coordinates boundary data of sample image "bat", "cat", "lumbar" and "key".	171

LIST OF FIGURES

		Page
Figure 2.1	Reconstruction of image boundary.	22
Figure 2.2	Sample images used in the research.	25
Figure 2.3	Slide of CT scan (source: Laboratory of Human Anatomy and Embryology, University of Brussels (ULB), Belgium).	26
Figure 2.4	Boundary data of (a) “alphabet T” (b) “gear”.	27
Figure 2.5	Boundary data of (a) “calligraphy da” (b) “calligraphy niu”.	28
Figure 2.6	Boundary data of (a) “bat” (b) “cat”.	29
Figure 2.7	Boundary data of (a) “lumbar” (b) “key”.	30
Figure 2.8	Curvature measurement with k-cosine.	38
Figure 3.1	Two examples for vectors $\vec{v}_i$, $\vec{u}_i$ and $\vec{q}_i$.	47
Figure 3.2	Huge difference between the $\lvert\vec{v}_i\rvert$ and $\lvert\vec{u}_i\rvert$.	48
Figure 3.3	Corner points detected on hypocycloid by bending value.	49
Figure 3.4	Corner points detected on hypocycloid by unit bending value.	50
Figure 3.5	Data in an ellipse and the eigenvectors.	52
Figure 3.6	Eccentricity of an ellipse.	56
Figure 3.7	Convex hull (shaded region) of a set of points.	60
Figure 3.8	Monotonic curve(left) and non-monotonic curve(right).	60
Figure 3.9	Example of computing monotonicity value.	61
Figure 3.10	Two test data sets (a) rectangular shape (b) curved shape.	62
Figure 3.11	Endpoints of support region for corner point “□” by (a) Unit bending value (b) Eigen-ratio (c) Eccentricity with eigenvalues (d) Convex hull area-length ratio (e) Marji & Siy method.	65
Figure 3.12	Endpoints of support region for non-corner point “△” by (a) Unit bending value (b) Eigen-ratio (c) Eccentricity with eigenvalues (d) Convex hull area-length ratio (e) Marji & Siy method.	66

Figure 3.13 Endpoints of support region for non-corner point "◯" by (a) Unit bending value (b) Eigen-ratio (c) Eccentricity with eigenvalues (d) Convex hull area-length ratio (e) Marji & Siy method. 67

Figure 3.14 Endpoints of support region for corner point "□" by (a) Unit bending value (b) Eigen-ratio (c) Eccentricity with eigenvalues (d) Convex hull area-length ratio (e) Marji & Siy method. 68

Figure 3.15 Endpoints of support region for non-corner point "△" by (a) Unit bending value (b) Eigen-ratio (c) Eccentricity with eigenvalues (d) Convex hull area-length ratio (e) Marji & Siy method. 69

Figure 3.16 Endpoints of support region for non-corner point "◯" by (a) Unit bending value (b) Eigen-ratio (c) Eccentricity with eigenvalues (d) Convex hull area-length ratio (e) Marji & Siy method. 70

Figure 4.1 (a) Angle θ formed by $p_{i-\ell}$, p_i and p_{i+r}. (b) Eigenvalues (i) λ_L (ii) λ_S (iii) λ_R against angle θ (in degree). 75

Figure 4.2 The graphs of λ_{sum} with different lengths of arms. (i) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 1$ (ii) $\|p_{i-\ell}p_i\| = 4$; $\|p_ip_{i+r}\| = 4$ (iii) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 2$ (iv) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 3$ 76

Figure 4.3 The graphs of λ_C with different lengths of arms. (i) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 1$ (ii) $\|p_{i-\ell}p_i\| = 4$; $\|p_ip_{i+r}\| = 4$ (iii) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 2$ (iv) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 3$ 77

Figure 4.4 The graphs of λ_I with different lengths of arms. (i) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 1$ (ii) $\|p_{i-\ell}p_i\| = 4$; $\|p_ip_{i+r}\| = 4$ (iii) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 2$ (iv) $\|p_{i-\ell}p_i\| = 1$; $\|p_ip_{i+r}\| = 3$ 80

Figure 4.5 First point, middle point and last point formed right angle in semicircle. 81

Figure 4.6 Semi-ellipses with same size of support but different lengths of minor axis. 81

Figure 4.7 Dominant points of rectangular shape by methods of (a) k-cosine (b) k-cosine with best-fit vectors. 85

Figure 4.8 Dominant points of curved shape by methods of (a) k-cosine (b) k-cosine with best-fit vectors. 85

Figure 4.9 Detected dominant points on hypocycloid data. 88

Figure 4.10 Detected dominant points on epicycloid data. 89

Figure 4.11	Area of the rectangle formed by chord length of the segment and the maximum perpendicular distance, $A_s = d_{seg} \times d_{max}$.	90
Figure 4.12	Rectangle of whole image and rectangle of a segment.	90
Figure 4.13	Detected dominant points of hypocycloid and its additional dominant points by r_p.	91
Figure 4.14	Detected dominant points of epicycloid and its additional dominant points by r_p.	91
Figure 4.15	Dominant points of rectangular shape by (a) λ_C (b) λ_I (c) cosine of best-fit vectors.	95
Figure 4.16	Dominant points of curved shape by (a) visual approach (b) λ_C (c) λ_I (d) cosine of best-fit vectors.	96
Figure 5.1	Two type of "S" shaped curve segments.	100
Figure 5.2	Two type of "S" shaped segments with their end point best-fit vectors.	101
Figure 5.3	θ_{PQ} formed by $\vec{v_P}$ and $\vec{v_Q}$.	102
Figure 5.4	Particular segment with wiggly data.	104
Figure 5.5	Area of polygons formed by candidates of point of inflection.	104
Figure 5.6	Area of polygons formed by candidates of point of inflection.	105
Figure 5.7	Sample of "S" curve.	107
Figure 5.8	Maximum perpendicular distance d_{max} and length d_{seg}.	108
Figure 5.9	A semicircle is produced to fit set Data A.	114
Figure 5.10	Various circular arcs to set Data A.	115
Figure 5.11	Error function J in (5.4) against parameter t in Equation (5.10).	115
Figure 5.12	Experimental Data B_1.	122
Figure 5.13	(a) Graph ξ against w (b) Graph $\frac{d\xi}{dw}$ against w for Data B_1.	123
Figure 5.14	Experimental Data B_2.	123
Figure 5.15	(a) Graph ξ against w (b) Graph $\frac{d\xi}{dw}$ against w for Data B_2.	124
Figure 5.16	Flowchart of curve-fitting algorithm for a segment.	129

Figure 5.17 "Cat" data is fitted by (a) linear segments. (b) linear segments and circular arcs. (c) linear segments, circular arcs and rational quadratic Bézier curves. 133

Figure 5.18 "Cat" data with G^1 continuity. 134

Figure 6.1 Alphabet T (a) binary image (b) 237 boundary data points (c) 8 dominant points. 136

Figure 6.2 Gear (a) binary image (b) 247 boundary data points (c) 48 dominant points. 137

Figure 6.3 Calligraphy "da" (a) binary image (b) 362 boundary data points (c) 39 dominant points (d) 47 dominant points. 138

Figure 6.4 Calligraphy "niu" (a) binary image (b) 221 boundary data points (c) 34 dominant points (d) 40 dominant points. 139

Figure 6.5 Bat (a) binary image (b) 327 boundary data points (c) 34 dominant points (d) 49 dominant points. 140

Figure 6.6 Cat (a) binary image (b) 235 boundary data points (c) 28 dominant points (d) 44 dominant points. 141

Figure 6.7 Lumbar (a) binary image (b) 246 boundary data points (c) 38 dominant points (d) 43 dominant points. 142

Figure 6.8 Key (a) binary image (b) 189 boundary data points (c) 24 dominant points (c) 30 dominant points. 142

Figure 6.9 Curve fitting of alphabet T data. 143

Figure 6.10 Curve fitting of gear data. 143

Figure 6.11 Curve fitting of calligraphy "da" data (a) with 39 dominant points (b) with 47 dominant points. 144

Figure 6.12 Curve fitting of calligraphy "niu" data (a) with 34 dominant points (b) with 40 dominant points. 145

Figure 6.13 Curve fitting of bat data (a) with 34 dominant points (b) with 49 dominant points. 146

Figure 6.14 Curve fitting of cat data (a) with 28 dominant points (b) with 44 dominant points. 147

Figure 6.15 Curve fitting of lumbar data (a) with 38 dominant points (b) with 43 dominant points. 148

Figure 6.16 Curve fitting of key data (a) with 24 dominant points (b) with 30 dominant points. 149

Figure 6.17 Different scales used for (a) calligraphy “da” (b) key data. 152

LIST OF ABBREVIATIONS

ROS	region of support
CNC	computer numerical control
PNG	portable network graphics
DP	dominant point
CT	computerized tomography
DICOM	digital imaging and communications in medicine
PCA	principal component analysis
2D	two dimensions
3D	three dimensions

LIST OF SYMBOLS

e_p	perceptual error
η	type weight
L	segment length
G^0	zero-order geometric continuity
G^1	first-order geometric continuity
C^0	zero-order parametric continuity
C^1	first-order parametric continuity
W	boundary points
p_i	point of interest in Cartesian system
n	number of boundary points
m	number of dominant points
$S(p_i)$	set of points in supporting region of point p_i
ℓ	size of left arm of supporting region
r	size of right arm of supporting region
$L(p_i)$	set of points of left arm in supporting region
$R(p_i)$	set of points of right arm in supporting region
H	sequence of points in Cartesian system

C	covariance matrix
λ_L	largest eigenvalue of covariance matrix
λ_S	smallest eigenvalue of covariance matrix
$\vec{v_1}$	eigenvector associates with λ_L
$\vec{v_2}$	eigenvector associates with λ_S
D	set of dominant points
c_i	k-cosine value of p_i
θ_i	angle formed at p_i included by two vectors connecting the point p_i to a preceding point $p_{i-\ell}$ and a succeeding point p_{i+r}
d_i	perpendicular distance from a point p_i to a linear line
d_p	local maximum perpendicular distance
r_p	ratio of significance
ε	mean error of dominant point detection
$\bar{S}_j$	signed perpendicular distance at point p_j
Y	absolute difference of signed area
J	error function of circular arc fitting
R	radius of approximating circular arc
C_i	control points of rational quadratic Bézier curves
$B_i(t)$	quadratic Bernstein polynomials with parameter t

lim	limit
$\vec{v_i}$	backward vector formed by the point p_i and a preceding point p_{i-k}
$\vec{u_i}$	forward vector formed by the point p_i and a succeeding point p_{i+k}
b_{ik}	bending value
ψ_i	i^{th} dominant point
L_k	length of line segment that joins the point of interest p_i to p_{i-k} (or p_{i+k}), $k > 0$
e_k	the sum of squared perpendicular distance from each data point between p_i and p_{i-k} (or p_{i+k}) to the line segment joining p_i to p_{i-k} (or p_{i+k})
F	$F = L_k - e_k$ in Marji & Siy's method
μ	ratio of $\frac{\lambda_S}{\lambda_L}$
E	eccentricity
δ_{pq}	central moment of order pq
c	length from the center of an ellipse to the focus
g	length of a line connecting the end point of minor axis to the focus of the ellipse
A_{ch}	area of convex hull
L_e	length of two end points of a points set
M	monotonicity value
ζ	ratio $\frac{A_{ch}}{L_e M}$

$\bar{D}$ normalized maximum deviation

θ_b angle between best-fit vector and the vector of chord line

$\beta(t)$ rational quadratic Bézier curve

w_i weight of the curve

G set of data points in segment PQ

f error function of rational quadratic Bézier curve fitting

ξ derivative $\frac{df}{dw}$

T_i^- left tangent of a point

T_i^+ right tangent of a point

ϖ Mean error of curve fitting

PEMBINAAN SEMULA IMEJ DENGAN PENGESANAN TITIK DOMINAN DAN KUASA DUA TERKECIL BÉZIER KUADRATIK NISBAH

ABSTRAK

Teknik-teknik pengesanan titik dominan yang berasaskan sempadan dan pemasangan lengkung adalah penting dalam analisis imej digital. Mereka menyediakan maklumat utama bagi imej untuk diproses dan dianalisis dengan lebih lanjut dalam bilangan data yang lebih rendah, dengan itu mempercepatkan proses analisis. Ini berkaitan dengan proses penghuraian bentuk, pengenalpastian corak dan pemahaman imej. Data sempadan bagi imej binari diekstrak dan kemudian diwakili dalam koordinat bentuk planar. Data ini kemudian digunakan untuk proses penentuan rantau sokongan, pengesanan titik dominan dan akhirnya pemadanan lengkung. Rantau sokongan memainkan peranan penting dalam proses mencari titik dominan yang tepat bagi satu set data sempadan. Rantau sokongan bagi suatu titik adalah subset bagi data sempadan yang menanggung tahap kelinearan yang maksimum. Teknik-teknik yang terdiri daripada pendekatan geometri dan statistik dikemukakan untuk mencari rantau sokongan yang sesuai. Dua kaedah yang sedia ada dalam mencari rantau sokongan diperbaiki untuk meningkatkan prestasi mereka. Kaedah-kaedah ini dinamakan sebagai nilai membengkok unit dan kaedah Marji & Siy. Selain itu, tiga kaedah baru dicadangkan dalam kerja ini untuk menentukan rantau sokongan, iaitu eigen-nisbah, eksentrik dengan nilai eigen dan nisbah luas hul cembung-panjang. Pengukuran kelengkungan adalah kaedah biasa untuk mengekstrak titik dominan. Kaedah geometri dan statis-

tik digunakan untuk menganggar kelengkungan dan seterusnya mencari titik dominan. Titik dengan kelengkungan yang tinggi akan dipilih sebagai titik dominan untuk mewakili bentuk data. Di sini, nilai eigen bagi kovarian matriks digunakan untuk menganggar ketirusan bagi titik data. Dua ukuran λ_C dan λ_I dirumuskan untuk menganggar kelengkungan. Selain itu, kaedah kosinus dengan vektor terbaik dikemukakan sebagai peningkatan bagi kaedah k-kosinus. Titik dominan kemudian dihubungkan oleh lengkung untuk membina semula imej. Selalunya, segmen garisan digunakan untuk menyambungkan titik dominan kerana ini adalah satu cara termudah bagi perwakilan segmen. Walau bagaimanapun, penggunaan lengkok bulat kuasa dua terkecil dikatakan menghasilkan penghampiran yang lebih baik pada tahap kerumitan pengiraan yang lebih tinggi. Lengkung Beziér kuadratik nisbah juga diguna pakai untuk memadankan data sempadan untuk penghampiran yang lebih memuaskan dengan keselanjaran geometri sehingga tahap 1 iaitu G^1. Vektor terbaik yang ditaksirkan oleh kaedah analisis komponen utama digunakan untuk menganggar arah tangen pada titik dominan. Ini dapat membantu dalam mencari titik kawalan untuk membina lengkung Beziér kuadratik nisbah. Pemberat optimum untuk fungsi Beziér kuadratik nisbah adalah dianggar dengan pendekatan kuasa dua terkecil. Beberapa contoh geometri turut dipersembahkan untuk menunjukkan keupayaan bagi kaedah-kaedah yang dicadangkan.

IMAGE RECONSTRUCTION BY DOMINANT POINT DETECTION AND LEAST SQUARES RATIONAL QUADRATIC BÉZIER

ABSTRACT

The techniques of boundary-based dominant point detection and curve fitting are important in the analysis of digital images. They provide key information of images for further processing and analyzing in lower amount of data, hence speed up the process of analysis. This relates to the processes of shape description, pattern recognition and image understanding. Boundary data of binary images are extracted and then represented in coordinates of planar form. These data are then used for the processes of support region determination, dominant point detection and lastly, curve fitting. Region of support plays a vital role in the process of locating precise dominant points of a set of boundary data. Support region of a point is a subset of boundary data points which bear the maximum level of linearity. Techniques consist of geometrical and statistical approaches are presented to find proper support region. Two existing methods in finding support region are improved to enhance their performance. The methods are named as unit bending value and Marji & Siy method. Besides, three new methods are proposed in this work to determine support region, namely eigen-ratio, eccentricity with eigenvalues and convex hull area-length ratio. Curvature measurement is a common method to extract dominant points. Geometrical and statistical methods are applied to estimate curvatures and hence locating the dominant points. Data points with high curvature will be selected as dominant points to represent the shape of the

data. Here, eigenvalues of covariance matrix are used to estimate the sharpness of data points. Two measures λ_C and λ_I are formulated to estimate curvature. Besides, method of cosine of best-fit vectors is presented as an improvement of k-cosine method. Detected dominant points are then connected by curves to reconstruct the image. Very often, line segments are used to connect the dominant points as this is the easiest way for segment representation. However, using least-square circular arcs is claimed to produce better approximation at a higher level of computational complexity. Rational quadratic Bézier curve is also adopted to fit the boundary data to produce better approximation with geometric continuity up to order 1, i.e., G^1. The best-fit vectors defined by principal component analysis (PCA) method is employed to estimate tangent direction at the dominant points. This helps in finding control points for constructing rational quadratic Bézier curve. An optimum weight for rational quadratic Bézier function is approximated by least square approach. Some geometrical examples are illustrated to show the capability of the proposed methods.

CHAPTER 1

INTRODUCTION

In some particular area of animation or digital art, data transfer speed is more important than image details, hence a simplified model is sufficient (Yahya, 2009). Data storages become an important concern in this digital era. Images usually consist of high pixels can be drastically reduced its size by detecting significant points in order to optimize the storage. For instance, simple objects such as a diamond, triangle or even a star can be represented by a few main vertices rather than a large number of points (Stojmenovic and Nayak, 2007). In this case, the main vertices are taken as significant points. Simple images representation not only promotes effective storage but also speed up the process of image data transferring.

Suppose a set of ordered data points consists of points that described the outline of the image which is often known as boundary points. Significant points are points extracted from boundary points that mainly described the shape of the image. It is also known as dominant points or corner points in many research papers. Dominant or corner points usually possess local maximum curvature as compared to their neighboring points. Dominant point detection or corner detection has been a popular research area in pattern recognition and image processing of scholars all over the world. It is a process to extract significant points from a set of data that able to represent the main information carried by the data set. Dominant points serve to simplify the analysis of images by drastically reducing the amount of data to be processed (Subri et al., 2006).

Region of support refers to a set of points that consists of a point under consideration known as interest point and a number of preceding and succeeding points of the interest point. Region of support plays an important role to extract dominant points effectively. A proper region of support provides a reasonable number of preceding and succeeding points of the interest point to be considered. The interest point is tested its local property such as curvature based on its corresponding region of support in order to distinguish it as dominant point or vice versa.

Data points between two consecutive dominant points form a segment where the two end points of the segment consist of the said dominant points. Reconstruction of the shape of images can be done by applying approximated curves to the segments formed. Connecting those dominant points with only line segments to produce polygon representation is insufficient to preserve the shape of the image. Therefore, curve fitting which is able to construct curvy segments that well represent the outline of the image becomes a challenge to researchers. The fitting process is targeted to be simple, fast and of course, precision wise is the main concern.

1.1 Literature Review

Dominant point detection may be grouped into two main categories: boundary-based and gray-level approaches (Sun et al., 2007). Here, boundary-based dominant point detection is studied and discussed. The significant information brought by a shape or pattern mainly represented by its boundary corners. Therefore, an efficient boundary-based dominant point detection method is important to detect the prominent points which are able to represent the original shape of the image visually and analyti-

cally with minimum error incurred. The purpose of corner detection is to detect some potential feature points which can sufficiently describe the shape of an object (Wang et al., 1995).

Ever since Attneave (1954) discovered that the information of an image was concentrated along contours (i.e., the region where the colours changed tremendously) and at the points where its direction changes most rapidly, many researches started to focus in finding dominant points which identified as the points with local maximum curvature. His experiment has shown that a drawing of a "sleeping cat" still able to be identified even though the drawing was made by the points with maximum curvature from the contour. The curvature measures the rate at which the curve bends away from the tangent line at a particular point (Bradley and Smith, 1995). The common procedures for corner detection techniques are estimating the curvature of each point on the boundary, then identify the points which have local maximum curvature as the corner points.

Detected dominant points from the boundary of an image are hoped to have a reasonable number of points and able to represent the original image information. Therefore, shape-preserving also take into account while detecting dominant points. Juttler (1997), Constatini and Pelosi (2004, 2001) and Kong (2006) have done the studies in shape preserving approximation. Shape preservation of planar and space data has played an important role in curve fitting over a large amount of data. Applying shape preserving conditions may prevent undesirable inflection or oscillations of the fitting curve. In Constatini and Pelosi (2004), the shape of a ordered data set is defined by the shape of the piecewise linear function where it is the best least squares approximation

to the data. This is applied while connecting dominant points with piecewise linear segments so that the reconstructed polyline is well representing the shape of original data.

In Tsai et al. (1999), an approach that integrated the statistical and geometrical properties of the data was implemented to extract points with high curvature value. The small eigenvalue of the covariance matrix of a subset of boundary points was adopted in detecting corner points. The points with small eigenvalues which exceed a given predetermined threshold would be taken as corner points. Zhu et al. (2010) also presented an auto-corner detection based on eigenvalues of the covariance matrices of boundary points over multi-region of support. The algorithm computes the eigenvalues product of the covariance matrices of each data points at various sizes of supporting region. Yahya (2009) claimed that the smallest eigenvalue of covariance matrices proposed in Tsai et al. (1999) for corner detection is not only affected by sharpness of curves but also by the size of supporting region or the number of data points of the curve. Furthermore, Yahya showed that it is not suitable to detect corner for equidistant data points of circular arcs of various radii and curved corners of different sharpness and sizes. Therefore, Yahya (2009) proposed the ratio of eigenvalues of covariance matrix in detecting dominant points. Tuzel et al. (2006) claimed that covariance matrices are good as supporting region descriptors. A single covariance matrix extracted from a region is usually enough to match the region in different views and poses. They assumed that the covariance of a distribution is enough to discriminate it from other distributions. Covariance does not carry any information regarding the order and the number of points. This means that the covariance is invariant although certain scaling and rotation are applied to the data. Covariance descriptor is no longer

rotationally invariant if the information such as the orientation of the points are taken into account. During covariance computation, majority of noise of samples are filtered out by an average filter (Tuzel et al., 2006).

Sun (2008) and Wu (2003b) used k-cosine value as the curvature measure to detect dominant points where k basically indicates the number of size of a supporting region. According to Wu (2003b), corner detection by k-cosine method is proposed by Rosenfeld and Johnston (1973), and then been improved by Rosenfeld and Weszka (1975). In Sun (2008), the boundary of an object is first represented into curvature measured by k-cosine. Then, by modifying the corner detection error, Sun proposed a suitable region of support and curvature threshold for robust corner detection. Wang et al. (1995) used another measurement named bending value to assess the degree of possibility of a point being a corner. The bending value is defined by considering the projections of vectors formed in the region of support. The authors utilized the concept that the directions of the forward and backward vectors of a non-corner point will cancel each other. Wu (2003a) computed average of bending values in a particular segment known as smoothing bending value to extract dominant point. Wu suggested that the points with local maximum smoothing bending value can be located as the dominant points on the curve.

Tsai (1996) and Subri et al. (2006) proposed a neural network approach to detect corners. In Tsai (1996), curvature was measured by using neural network to identify the included angles at boundary points. Subri et al. (2006) presented a neural network classifier to detect the corners of chain code series. The Freeman chain code was used to derive the classifier as training, testing and validation set. Steps of developing the

classifier can be found in Subri et al. (2006). The results show that there are strength and drawbacks of using neural network in corner detection of chain code series. The strength of applying neural network is it makes corner detector more sensitive in detecting a corner. On the other hand, the drawback of the method is it can be classified as tedious, and trial and error process (Subri et al., 2006). They claimed that the method is tedious because it involves training samples that have to pass through three stages while the trial and error process will sometime lead to insignificant result. In Wu and Wang (1993), the curvature based polygonal approximation method was proposed. The method combined the corner detection and polygonal approximation techniques to detect the dominant points. Polygonal approximation was performed by using detected corner points as initial points. This detection method consists of three procedures, i.e. first, extract the break points that do not lie on a straight line, then, detect the potential corners; lastly, perform polygonal approximation by partitioning the data between two consecutive potential corners.

Region of support (ROS) plays vital role in dominant point detection. Many researches have found their own way to determine region of support before the curvature is measured. Generally, region of support can be categorized into symmetry and non-symmetry region. Symmetric support region refers to the same number of preceding and succeeding points of an interested point whereas non-symmetric support region of an interested point has different number of preceding and succeeding points. Many studies discussed on the symmetric region of support. Meanwhile, some papers such as Guru and Dinesh (2004) mentioned that the left arm and right arm of a point are not necessarily to be the same length and it is more reasonable for them to be asymmetric. Note that the sequence of points which come before the interest point is referred

to as left arm of region of support whereas the right arm formed by the sequence of points which come after the interest point irrespective of the direction of traversal of the boundary data. The authors determined the region of support of each boundary point based on the points local geometric properties. They proposed to choose the maximum sequence of points which hold approximately zero value in smallest eigenvalue of the covariance matrix as the respective arms on both sides of an interest point. Once the region of support is determined, they estimated the curvature at that point and also computed how many times a point has been the end point of the region of support of other points and denoted it as limit value. Results from the experiments told that the limit value, the size of region of support and the curvature of a true corner point are relatively larger than the respective values of its neighbor points.

According to Teh and Chin (1989), precise determination of the region of support is more important than curvature estimation. They determined symmetric region of support by first computing the length of the chord connecting a preceding and a succeeding point of the interest point and the perpendicular distance of the interest point to the chord. They stressed that the precision of determining region of support would lead to effectiveness in dominant point detection. The procedure in Teh and Chin (1989) can be summarized as follows - first determines the region of support for each point based on its local properties, then measure of significance (e.g., curvature) is computed for each data point and finally dominant points are detected by a process known as non-maxima suppression. However, Wu (2003a) showed that Teh-Chin's algorithm (1989) tends to find small region of support hence causing too many dummy points to be detected. On the other hand, relevant information may loss due to an extremely large region of support (Sun et al., 2007). Wu (2003b) modified the Teh-Chin's method such

that the region of support for each point depending on the region of support for the previous point. That is, the determination of the region of support for each point can be done dynamically by using the information of region of support found for the previous point. The author claimed that Teh-Chin's method detects too many insignificant dominant points due to the method prone to give small region of support. Therefore, Wu (2003b) proposed their strategy to overcome this problem.

Rosenberg (1972) aimed to extract fragments of a blob which are perceptively meaningful. He attributed each point by a set of neighboring points and finding the significant set. In his study, he introduced "dom points" for every point on the contour where this "dom points" consists of some consecutive points along the contour. Once "dom points" were determined, some measures of significance to estimate curvature would be performed. Davis (1977) pointed out that computing measure of significance over inappropriate region of support may result in true dominant point to be discarded. In Tsai et al. (1999), fixed region of support was proposed. However, fixed region of support is unreasonable as Langridge (1972) pointed out that each boundary point in a contour should have their own geometric characteristic. Therefore, every boundary point should own their individual region of support. Guru and Dinesh (2004) commented the results given by Tsai et al. (1999) is less precise in detecting true corner due to the static region of support.

In Zhu et al. (2010), the method of multi region of support was used. They claimed that the corner detector will be sensitive to noise if the single region of support is set to be too small, however, inaccuracy of corners detection arises if the size of support region is too big. Stojmenovic et al. (2006) had did a handy survey research on mea-

suring the linearity for unordered point sets. In the following year, Stojmenovic and Nayak (2007) continued their research on measuring linearity of ordered point sets. Their researches mainly motivated by the determination of "region of interest" which is very useful in image processing. According to Stojmenovic et al. (2006), linear points are often representing a "region of interest" in an image. In this context, region of interest having the same meaning as region of support. Stojmenovic et al. (2006) and Stojmenovic and Nayak (2007) have set four perspectives in analyzing their linearity algorithms. First, the linearity value is a number in the range [0, 1]. A highly linear shape will give the value of 1 and then gradually decreasing to highly non-linear shape which takes the value of 0. Second, linearity measure should be invariant under affine transformation, i.e., translation invariant, scale invariant and rotation invariant. Thirdly, the algorithm has to be resistant to protrusions in the data set because small protrusions could be due to noise. Lastly, the algorithm should be simple and able to calculate linearity value in short time.

There are several methods in literature to extract the dominant points from the boundary of image data. The efficiency and accuracy of the methods are compared analytically from the aspect of visual inspection in previous years (Rosin, 2003). In recent years, quantitative assessment gradually takes the place of comparative method as mentioned in Rosin (2003). The author developed several measures to assess the stability of polygonal approximation algorithms under two classes of variations, namely perturbations of the data and changes in the algorithms scale parameters.

Generally, there are two approaches of curve fitting which commonly used by researchers; interpolation and approximation. Curve generated in interpolation method

will pass through every data point whereas curve approximation is to construct a curve which is close to the data points depending on error criteria set by researchers. In image processing, polygonal approximation is crucial to constitute the preserved feature of the images in reduced storage space. It is a process of creating polyline to reconstruct the images. Common approach in polygonal approximation is to formulate the approximation as an optimization problem (Pei and Horng, 1996). Nevertheless, application of polygonal approximation has its limitation as many smooth contours are not suitable to be approximated by using polygons. The performance of approximation is said to be enhanced if higher order polynomial curves are used for approximation. However, higher order polynomial curves require higher computational complexity. In machinery industry, the tool path for Computer Numerical Control (CNC) machinery only use the piecewise circular arcs and straight line segments because of physical limitation. Designing shapes in CNC machinery using tangential circular arc segments are better and easier than the polygons consisting of line segments.

In Pei and Horng (1996), circular arcs are employed to fit the digital boundaries with lower computational complexity compared to splines curve fitting. They have provided a set of existing circular arcs to be chosen automatically to fit polygon-like boundaries. The algorithm only required to construct a few circular arcs to approximate image boundaries. The circular arc was obtained by determining the center of circle such that the arc gives the smallest sum of squared differences of radii to the concerned boundary data. The results are compared to Teh-Chin's algorithm of constructing polygonal approximations for digital images by evaluating the number of dominant points and integral square error. Pei and Horng (1996) claimed that satisfactory results were obtained after the algorithm has been tested on a number of digital

images.

Horng and Li (2001) has improved the method of Pei and Horng by fitting the digital curves with combination of line segments and circular arcs. They claimed that curve fitting with combination of line segments and circular arcs can yield better results than using circular arcs alone. The model of estimating the center of circular arc in Pei and Horng (1996) was employed in Horng and Li (2001). The authors proposed a dynamic programming to locate optimal segmentation points which minimizing the total perceptual error. The problem of digital curves fitting with line segments and circular arcs is then formulated in mathematical form as an optimization problem and solve it. In other words, an optimal analytic curve is designed to minimize the perceptual error of fitting. Two criteria for perceptual significance described by Ichoku et al. (1996) were considered in Horng and Li (2001). The first criterion is the priority will be given to straight line segments to fit the segments. Circular arcs are considered to fit the data only if they considerably simplify the representation. In second criterion, longer curve segments are preferred regardless the choice of line segments or circular arcs are used to fit the data. This can be seen from the perceptual error e_p of an analytic curve segment which is given by

$$e_p = \eta \left[e + \frac{c}{L} \right]$$

where η is the type weight, c is a constant, L is the segment length and e represents the integral square error. The value of type weight η controls the level of preference for different types of analytic curves (line segment or circular arc). Lower value of η assigned for line segment compared to circular arc. Thus, this coincides with the first

criterion, i.e. the line segments are preferred. The authors formulate perceptual error by modified integral square error by adding in the term $\frac{c}{L}$ to prevent short segments. The term $\frac{c}{L}$ provides an additional error for short segments, this is to achieve second criterion, i.e. the longer segments are preferred. Horng and Li (2001) claimed that the method is able to construct excellent representations of the boundaries of the real images.

Ichoku et al. (1996) presented a dynamic focusing algorithm for the segmentation of planar curves into straight lines and circular arcs. They strongly agreed that any boundaries could made up of line segments and circular arcs. Many digital boundaries are actually can be more reasonable to be approximated by the combination of straight line segments and circular arcs than using line segments alone. The algorithm recursively attempted to fit a candidate segment by using first a line segment and then a circular arc. The fitting process is started with the candidate segment of original digital boundary. If the fit is unsatisfied, then the candidate segment is decreased by a number of pixels and the process of fitting is performed recursively until the desired fit is obtained. The segmentation result is affected by the choice of the decreased candidate segment pixels, the threshold set and the starting direction. The fitting was based on three quality factors. One of the objective measure is the deviation of the data from the fitting curves. The authors proposed to use the "normalized maximum square deviation" which is given by dividing the square of maximum deviation by the length of line or the radius of the circular arc. Lastly, to make sure the deviations be evenly distributed along the fitting curves, a regression line is built. The slope of the regression line is the measure of validating the line or the arc fit. Furthermore, a pre-test for arc fitting was implemented to save time. If the length of the line joining two end points of

a segment is less than half the radius of the arc, then the segment can possibly suitable fitted with straight line instead of an arc. This is because the minor arc subtended by a chord joining two end points of a segment whose length is less than half the radius does not have significant concavity or convexity (Ichoku et al., 1996). The experimental results showed that the dynamic focusing algorithm performed good in the segmentation of noisy and smoothly winding curves.

Ahn and Kim (1998) has their concern on the curvature of a rational quadratic Bézier curve to be monotone in [0, 1], to have a unique local minimum, a unique local maximum, or both extremum in $(0, 1)$. They have found necessary and sufficient conditions to achieve the curvature mentioned above. Besides, they have contributed in finding the path of tool by presenting simple algorithm to obtain the offset curve which to be regular and to have the same tangent direction with the given rational quadratic Bézier curve at the first and the last control points.

Yahya et al. (2005) has used rational Bézier cubic curves to fit the Arabic letter "qaf" which is cursive in character and having varying curves and cusps. The representation of the letter is formed by the steps including contour extraction and corner detection. The last step is curve fitting by using rational Bézier cubic curves. In order to obtain G^1 curves which approximated to digitized data points, weights of the rational cubic were adjusted.

Various curve schemes have been explored in the literature of curve designing. Curve flexibility and complexity are two main criterion that make a curve scheme popular. For example, quadratic Bézier curves are important geometric entities in many

applications. It is often used as the generatrix curves of the surface of radars and to approximate circular arcs, which unable to be represented by polynomials in an exact way (Gu et al., 2009). Gu et al. (2009) constructs a composite curve consists of two quadratic Bézier curves to satisfy the endpoint constraints with both positions and directions of tangent vectors. According to the authors, two segments of quadratic Bézier curves are enough to satisfy the endpoint conditions of a segment.

1.2 Motivation

There is a wide use of dominant point detection (or corner detection) in engineering, medical sciences, digital art, motion analysis and other computer vision applications. Transforming digital images into numerical information for further processing and analysis is one of the important computer vision applications. This is due to long processing time and huge storage are required in reconstruction of images. Low computational complexity cum effective algorithm may help in facilitating the industrial growth such as medical field that involves image processing. This motivates the work to create a model which is accurate and consistent of original image as it is used to produce effective manufacturing, shorten time and energy consumption for subsequent designing and analysis, effective face recognition and perhaps to facilitate medical procedures. Dominant point detection refers to the detection of significant points for subsequent processing. Its ability in reducing data storage is one of the attractive feature provided by dominant point detection. Existing dominant point detection methods in the literature are worth for further investigating as they have their strengths and weaknesses. Discovering new method may provides more choices for the researchers, application developers and users. Furthermore, there are rooms to improve the existing

methods in order to obtain more accurate and desired results.

On the other hand, image data involves huge data storage and longer time to process. This may be a burden during data transmission. The set of detected dominant points significantly reduce storage size and hence increase data transfer rate from a data source to its subsequent analysis and practical use. On the other hand, dominant points of an image shall be able to preserve the main feature of the original image. Therefore, techniques of constructing curves to approximate the image boundary with reasonable level of accuracy where the approximated curves pass through the dominant points are one of the important challenges in the research. The techniques have to be low computation and complexity when approximate to the shape of the image.

1.3 Problem Statements

Irregular shape images are referred to jagged boundary data. The existing methods for detecting dominant point are not suitable for the data set of irregular shape and worth for further improvement. The same situation happened for region of support. It needs to be improved further. The existing techniques which using statistical analysis to detect the sharp corner are suitable to measure the curvature only in certain circumstance (Tsai et al., 1999; Yahya et al., 2005). The methods failed in detecting the points effectively if curvature exceeds certain angle. However, the eigenvalues and the corresponding eigenvectors of covariance matrix possess characteristics that are useful in dominant point detection.

Besides, methods of curvature measurement in detecting dominant points which applied in existing method can be enhanced. Simple and easy curve fitting techniques

are eagerly needed to preserve the original shape of the image boundary. Therefore, an appropriate fitting method should be developed to fit the image data with least error and able to produce smooth resulting curve.

1.4 Scope

The main focus of this study is detecting the dominant point from boundary data of digital images and hence applying low computational complexity curve fitting by using least square approximation to reconstruct the outline of the images. The simple digital images in binary data which consist of black and white colours are considered in this study. Hence, the data are 2D which can be presented by Cartesian plane system. Besides that, the study is limited to close data points where the outline of digital image is close contour or unbroken contour.

1.5 Objectives

The aim of this research is to survey and hence improve some existing techniques in the literature as well as develop new methods in determining region of support, detecting dominant points of image boundary data and approximating the data by smooth curve with least error. The objectives are listed as follow:

- To improve the existing methods as well as develop new methods to determine support region for every data point.
- To improve the existing methods as well as develop new methods of dominant point detection.

- To detect significant points in order to preserve the inherent shape of image boundary.
- To construct visually smooth curves with low computational complexity, least error approximation and yet able to represent the inherent feature of image data.

1.6 Methodology

The image reconstruction will go through a few steps. The steps include image data extraction from binary image, support region determination, dominant point detection and curve fitting by least square objective function. Binary images are digital images that has two possible values for each pixel. It can be stored in memory as a bitmap. Due to the small size of the image files, binary images with black and white are employed in this research.

In determining support region, improvement of existing methods are as follow.

(i) Unit bending value

(ii) Marji & Siy method

Besides, some new methods are suggested in this research to determine support region. They are named as

(i) Eigen-ratio

(ii) Eccentricity with eigenvalues

(iii) Convex hull area-length ratio

For detecting dominant point, statistical and geometrical methods used in the literature are considered. Two statistical estimates using eigenvalues are proposed in this work. They are λ_C and λ_I. On the other hand, an improvement of method of k-cosine is introduced. Best fitting vectors are studied and then employed in the k-cosine measure. That is k-cosine method with best fitting vectors. The boundary data are partitioned into several segments at the extracted dominant points. After that, the curves to be constructed are required to pass through dominant points and close to other data points. Thus, it is an approximation problem with objective. Piecewise linear segment, circular arc or rational quadratic Bézier curve are employed to fit the boundary data with least error. The geometric continuity G^1 is pursued to obtain smooth join at the dominant points.

1.7 Thesis Organization

The thesis is organized into seven chapters. Chapter 1 brings out the motivation, problem statements, objectives of the research and methodology. Reviews of various techniques in literature are also discussed in this chapter.

Chapter 2 briefly describes the preliminary concepts used in the research. This includes the process of boundary extraction. Besides, some mathematics concepts that will be applied in later chapters are discussed too. This includes the procedure started from the support region determination to curve-fitting of segments. An overview of this research is stated here.

Chapter 3 fully discusses the process of finding region of support. In this chapter, certain methods published by other authors are discussed. Then, adapted methods in

determining region of support of data points are elaborated in detail. This includes the underlying concepts and their corresponding algorithms. Two simple data sets in rectangular and curved shapes are employed to show the performance of proposed methods.

Chapter 4 deliberates the new estimation methods in extracting dominant points. Here, concepts of mathematics that will be applied are discussed in Chapter 2. Then, new estimation methods in dominant point detection are discussed. Detected dominant points can be improved by either increasing more significant points or suppressing the existing redundant dominant points. This is discussed in this chapter too. At the end of the chapter, the performance of the estimations are evaluated quantitatively and displayed using same rectangular and curved shapes data as given in Chapter 3.

Curves fitting to each segment formed by dominant points is described in Chapter 5. Curve-fitting using line segments and least-square circular arcs are presented. To improve the resulting curves, rational quadratic Bézier curves with least-square error are adopted. Then, a full algorithm of curve-fitting using combination of line segments, least-square circular arcs and rational quadratic Bézier curves is presented. The curve-fitting method is presented by using a “cat" image. The boundary data of the image are fitted with three approaches: (i) only line segments; (ii) combination of line segments and least-square circular arcs are used to fit the boundary data; then (iii) combination of line segments, least-square circular arcs and rational quadratic Bézier curves. Their mean errors approximation for respective methods are shown.

An overall discussion regarding the methods used in this study is done in Chapter

6. Results of proposed methods on sample images will be displayed. The extracted boundary data points, detected dominant points and resulting curves are illustrated graphically. Besides, their numerical results also shown to compare the strengths and weaknesses of the proposed methods.

The thesis ends with the conclusion in Chapter 7. This chapter gives a summary of the research and concludes with a discussion on limitation of the research. Besides, some possible research extensions are suggested in this chapter too.

CHAPTER 2

PRELIMINARY WORK

In this chapter, a brief discussion regarding the core of this research is given, i.e. region of support determination, dominant point detection and curve fitting. Figure 2.1 illustrates the flow of the research. It shows the research outline which gives an overall picture of the work. First, binary images are used as input in the research algorithm. The boundaries of the images are then extracted to form boundary data points. Next, support region of every data point is determined. Based on the support region, curvature of the data point is estimated and hence dominant points can be detected. Then, image reconstruction is carried out by fitting least square curve to the boundary points. The output are expected to be geometrical curve which approximates the image outline or informative data of dominant points and shape parameters. Each step in Figure 2.1 will be briefly discussed in the following sections. Image data that will be used as test images are displayed in Section 2.1. Besides, some mathematics concepts are discussed in this chapter as they are important as prerequisite knowledge and foundation to develop this research. They include eigenvalues of covariance matrices and principal component analysis (PCA) which are used to determine region of support and to detect dominant points. Concept about k-cosine method (Sun, 2008) for dominant point detection is briefly discussed too.

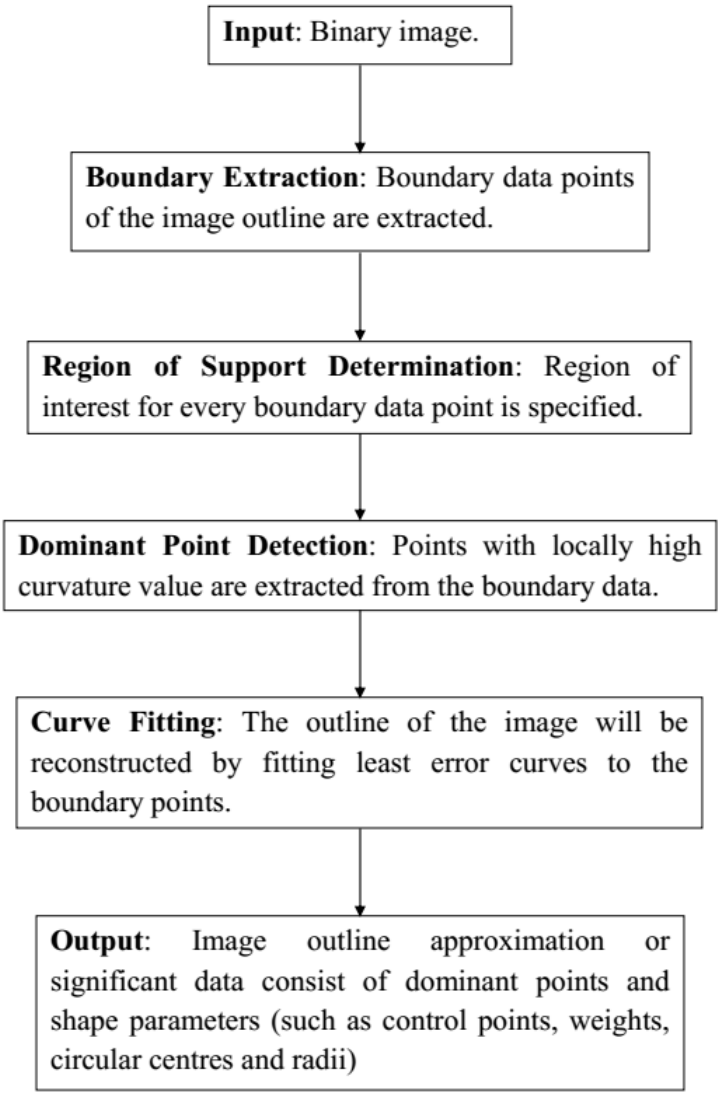


Figure 2.1: Reconstruction of image boundary.

2.1 Boundary Extraction

A digital image is a numeric representation of a two-dimensional image. Here, binary images that has only two possible values for each pixel are used in order to be easily processed by computer. Digitized images of binary form in Figure 2.2 are considered as a tool to visualize this research performance. The research is focusing on the boundary-based description and representation, thus the outline of an image known as contour line which carries the information of the shape of the image is considered.

Boundary extraction of a digital image is a process of extracting the contour line of the image and hence perform the related pixels in planar points. This yields ordered boundary points of the image. Figures 2.4-2.7 are the boundary points of the images given in Figure 2.2. The coordinates of boundary points are shown in Appendix A and B. The boundary points were obtained by the procedures as described below.

First of all, the binary images were saved as Portable Network Graphics (PNG) format as it is lossless image compression format that most widely used on the internet. Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. Furthermore, images in PNG format can be easily obtained from the internet. The contour lines and boundary points of the images will then be extracted by using Matlab built-in function named "bwtraceboundary()". The resulting contours have one pixel thick. To save time during samples testing, the contours with reasonable number of data points are needed. The number of points on contours can be reduced without disturbing the shape of the contours. This will speed up the processes during region of support determination, dominant point detection and curve fitting. Therefore, the contours of images are smoothed out/ filtered by diminishing systematically number of points on the boundaries. The boundary points filtered from the image are required to be almost equidistant from each other (for example, four pixel apart). Yahya (2009) extracted every third pixel and eliminated away the intermediate two. We should note that using larger spatial filter to get wider spaced boundary points may overlook the important corners that significantly describe the shape of images while smaller spatial filter will lead to huge number of data points that requires bigger size of storage and longer time

of processing. The retaining boundary points are denoted as

$$W = \{p_i | i = 1, 2, \ldots, n\} \tag{2.1}$$

where n is the number of boundary points. The position of a pixel in an digital image can be expressed by Cartesian planar system as $p_i = (x_i, y_i)$.

Figure 2.2 displays sample images in PNG format which being used in this research. They are named as follow, from left to right “alphabet T”, “gear”, “calligraphy da”, “calligraphy niu”, “bat”, “cat”, “lumbar” and “key” respectively. Note that image “lumbar” is obtained from a computerized tomography (CT) scan slide of lumbar as shown in Figure 2.3. The slide in format Digital Imaging and Communications in Medicine (DICOM) is read by Matlab and converted to binary image shown in Figure 2.2.

The respective boundary data of sample images in Figure 2.2 are extracted and filtered with 4 pixels apart to obtain boundary data as shown in Figures 2.4, 2.5, 2.6 and 2.7. The number of boundary points for every image is different since different sizes of images being considered. For instance, Figure 2.5(a) and Figure 2.6(a) has larger number of boundary points compared to other boundary data. Thus, they seen to have closer boundary points after the size of figures have been scaled to fit the pages. The boundary data in this work are closed boundary. Data obtained from open boundary of image can be included in this study.