

# CONTENTS

<b>ABSTRACT (in Chinese)</b> .....	<b>i</b>
<b>ABSTRACT (in English)</b> .....	<b>ii</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>iii</b>
<b>CONTENTS</b> .....	<b>iv</b>
<b>LIST OF FIGURES</b> .....	<b>vii</b>
<b>LIST OF TABLES</b> .....	<b>xi</b>

<b>Chapter 1 Introduction</b> .....	<b>1</b>
1.1 Motivation.....	1
1.2 Survey of Related Studies .....	3
1.3 Overview of Proposed Method .....	4
1.3.1 Definitions of Terms .....	4
1.3.2 Assumptions.....	6
1.3.3 Brief Descriptions of Proposed Method .....	6
1.4 Contributions .....	8
1.5 Thesis Organization .....	9
<b>Chapter 2 Overview of Proposed Method for Talking Cartoon Face Generation</b> .....	<b>10</b>
2.1 System Organization.....	10
2.2 Uses of Facial Features, Control Points, and Key Frames.....	12
2.2.1 Facial Feature Regions.....	12
2.2.2 Facial Feature Points.....	13
2.2.3 Control Points .....	13
2.2.4 Key Frames .....	14
2.3 Talking Cartoon Face Generation from Single Images.....	14
2.3.1 Basic Idea.....	15
2.3.2 Generation Process.....	15
2.4 Talking Cartoon Face Generation from Image Sequences .....	17
2.4.1 Basic Idea.....	17
2.4.2 Generation Process.....	18
<b>Chapter 3 Extraction of Facial Feature Regions</b> .....	<b>19</b>
3.1 Introduction.....	19
3.2 Review of Eye-pair Detection Method .....	20
3.3 Extraction of Background, Hair, and Face Regions.....	21

3.3.1	Proposed Hierarchical Bi-level Thresholding.....	22
3.3.2	1 <sup>st</sup> Level Bi-level Thresholding in Intensity Channel .....	23
3.3.3	Optimal Eye-Pair Detection.....	24
3.3.4	2 <sup>nd</sup> Level Bi-level Thresholding in Hue Channel .....	26
3.3.5	Region Refinements.....	27
3.4	Extraction of Facial Feature Regions.....	29
3.4.1	Eyeball Detection for Position Estimation.....	30
3.4.2	Knowledge-based Edge Detection by Local Thresholding.....	31
3.4.3	Knowledge-based Facial Feature Extraction .....	32
3.5	Detection of Cheek Boundaries and Ear Regions.....	36
3.6	Experimental Results .....	37
<b>Chapter 4</b>	<b>Extraction of Facial Feature Points.....</b>	<b>40</b>
4.1	Introduction.....	40
4.2	Face Model Construction.....	41
4.2.1	Review of MPEG-4.....	41
4.2.2	Review of Two Curve Drawing Algorithm.....	42
4.2.2.1.	Corner-cutting Subdivision .....	43
4.2.2.2.	Cubic Bezier Curve Approximation.....	45
4.2.3	Proposed Facial Feature Points.....	47
4.3	Extraction of Facial Feature Points.....	49
4.3.1	Extraction of Eyebrows, Eyes, Nose, Mouth, and Jaw Points.....	49
4.3.2	Extraction of Face, Ear, and Hair Points.....	54
4.4	Cartoon Face Deformation.....	57
4.4.1	Proposed Facial Animation Parameter Units .....	57
4.4.2	Assignment of Control Points.....	58
4.5	Experimental Results .....	59
<b>Chapter 5</b>	<b>Talking Cartoon Face Generation from Single Images..</b>	<b>62</b>
5.1	Introduction.....	62
5.2	Definition of Basic Mouth Shapes.....	63
5.2.1	Phonetic Transcription .....	64
5.2.2	Extraction of Basic Mouth Shapes.....	65
5.2.3	Basic Mouth Shapes.....	67
5.3	Analysis of Time Intervals between Mouth Shapes of Syllables.....	71
5.4	Talking Cartoon Faces Generation by Synthesizing Moving Lips .....	75
5.5	Talking Cartoon Faces Generation by Synthesizing Emotions.....	77
5.6	Experimental Results .....	77
<b>Chapter 6</b>	<b>Talking Cartoon Face Generation from Image Sequences</b>	<b>79</b>
	.....	<b>79</b>

6.1	Introduction.....	79
6.2	Tracking for Eyeball Regions .....	80
6.3	Tracking Mouth Regions .....	81
6.4	Extraction of Facial Feature Points.....	82
6.5	Experimental Results .....	83
<b>Chapter 7</b>	<b>Cartoon Generator and Experimental Results .....</b>	<b>86</b>
7.1	Introduction.....	86
7.2	Overview of SVG .....	86
7.3	SVG Animation.....	88
7.3.1	From the View of Spatial Domain .....	88
7.3.2	From the View of Temporal Domain .....	91
7.4	An Application to Electronic Books .....	93
7.5	Experimental Results .....	94
<b>Chapter 8</b>	<b>Conclusions and Suggestions for Future Works .....</b>	<b>98</b>
8.1	Conclusions.....	98
8.2	Suggestions for Future Works.....	99
<b>References</b>	<b>.....</b>	<b>101</b>



# LIST OF FIGURES

Fig. 1.1	A flowchart of proposed system. ....	6
Fig. 1.2	Environment setup of this study.....	7
Fig. 2.1	System organization and generation stages. ....	11
Fig. 2.2	Facial feature regions used in the proposed system. (a) Background, hair, and face regions. (b) Eyebrows, eyes, nose, and mouth regions. ....	12
Fig. 2.3	Facial feature points used in the proposed system.....	13
Fig. 2.4	Feature points of mouth. ....	13
Fig. 2.5	A basic concept of key frame use. ....	14
Fig. 2.6	Stages of talking cartoon face generation from single images.....	16
Fig. 2.7	Stages of talking cartoon face generation from sequential images.....	18
Fig. 3.1	An overall illustration result of extraction of facial feature regions.....	19
Fig. 3.2	Stages of the fast frontal face detection method by Chan <i>et al.</i> [10].....	20
Fig. 3.3	The square region proposed in Chan <i>et al.</i> [10].....	20
Fig. 3.4	HSI color space.....	21
Fig. 3.5	Flowchart of hierarchical bi-level thresholding method.....	22
Fig. 3.6	An experimental result of bi-level thresholding in intensity channel. (a) Neutral facial image in intensity channel. (b) Binary image $B_1$ . ....	24
Fig. 3.7	An experimental result of optimal eye-pair detection. (a) Binary image $B_1$ . (b) Probable eye-pairs generated by Chan's method. (c) An optimal eye-pair...	25
Fig. 3.8	An experimental result of bi-level thresholding in hue channel. (a) Binary image $B_1$ . (b) Neutral facial image in hue channel. (c) Binary image $B_2$ .	27
Fig. 3.9	Four sets of regions and two eye regions. (a) Dark regions $D$ (black) and light regions $L$ (white) of $B_1$ . (b) Face regions $F$ (black) and background regions (white) of $B_2$ . (c) Eye regions (red rectangles).....	28
Fig. 3.10	An experimental result of region refinements for face regions and hair regions. (a) Face region (black). (b) Bi-level image $B_3$ . (c) Hair region (black). ....	29
Fig. 3.11	A result of eyeball detection. (a) Eyeballs with a distance $d$ between them. (b) A $2d \times 2d$ square region.....	30
Fig. 3.12	Sobel operators. ....	31
Fig. 3.13	A result of knowledge-based edge detection by local thresholding. (a) Sobel edge image $S$ . (b) Binary edge image $B_{\text{edge}}$ with three horizontal divisions shown on it. ....	32
Fig. 3.14	Search range definition for knowledge-based facial feature extraction. ...	34

Fig. 3.15	A result of extraction of facial feature regions. (a) Speculated positions and ranges of facial features (yellow rectangle). (b) Final extraction results of facial feature regions. ....	36
Fig. 3.16	A result of detection of cheek boundaries. (a) Projection information. (b) Cheek boundaries shown on a neutral facial image.....	37
Fig. 3.17	An example of experimental results. (a) A given neutral facial image. (b) A result of optimal eye-pair detection. (c) A result of face regions. (d) A result of hair regions. (e) A result of eyeball detection and $2d \times 2d$ rectangle range. (e) A result of facial feature extraction. ....	38
Fig. 3.18	Another example of experimental results. (a) A given neutral facial image. (b) A result of optimal eye-pair detection. (c) A result of face regions. (d) A result of hair regions. (e) A result of eyeball detection and $2d \times 2d$ rectangle range. (e) A result of facial feature extraction. ....	39
Fig. 4.1	84 feature points in MPEG-4. ....	41
Fig. 4.2	FAPUs in MPEG-4. ....	42
Fig. 4.3	An illustration of corner-cutting subdivision.....	43
Fig. 4.4	An illustration of corner-cutting algorithm.....	44
Fig. 4.5	Cubic Bezier curve.....	46
Fig. 4.6	An experimental result of dragging control points with the same direction. (a) Iteration 1 with error distance of 19.75 pixels. (b) Iteration 2 after moving 19 pixels down warding with error distance of 5.75 pixels. (c) Recursively doing the approximation until the error distance smaller than 1 pixel. ....	46
Fig. 4.7	An experimental result of control points with diverse directions. (a) A normal jaw curve. (b) A sharper jaw curve. (c) A flatter jaw curve.....	47
Fig. 4.8	72 feature points in the proposed system.....	48
Fig. 4.9	An illustration of extraction of eyebrow points. (a) Left eyebrow points in the proposed face model. (b) Detection of the left eyebrow points in left eyebrow region. ....	49
Fig. 4.10	An illustration of extraction of eye points. (a) Left eye points in the proposed face model. (b) Detection of the left eye points in left eye region. ....	51
Fig. 4.11	Left eye points in the proposed face model. ....	52
Fig. 4.12	An illustration of extraction of mouth points. (a) Mouth points in the proposed face model. (b) Detection of the mouth points in mouth region...	53
Fig. 4.13	An illustration of extraction of face points. (a) Face points in the proposed face model. (b) Detection of the face points in face region.....	55
Fig. 4.14	An illustration of extraction of hair points. (a) Hair points in the proposed face model. (b) Detection of the hair points in hair region.....	56
Fig. 4.15	Facial animation parameter units in the proposed system. ....	57

Fig. 4.16 An illustration of eye deformations. ....	58
Fig. 4.17 An illustration of mouth deformations.....	58
Fig. 4.18 An example of experimental results. (a) A result of detection of facial feature points. (b) A result of neutral cartoon face rendered by SVG. (c) A result of a cartoon face with relocating some facial points. (d) A result of a cartoon face with relocating the control points of eyes and mouth.....	59
Fig. 4.19 The second example of experimental results. (a) A result of detection of facial feature points. (b) A result of neutral cartoon face rendered by SVG. (c) A result of a cartoon face with relocating some facial points. (d) A result of a cartoon face with relocating the control points of eyes and mouth. ....	60
Fig. 4.20 The third example of experimental results. (a) A result of detection of facial feature points. (b) A result of neutral cartoon face rendered by SVG. (c) A result of a cartoon face with relocating some facial points. (d) A result of a cartoon face with relocating the control points of eyeballs (shifting up) and mouth. ....	61
Fig. 5.1 Basic components for definition of basic mouth shapes. (a) Control points (solid circle) of mouth. (b) Facial animation parameter units of mouth and nose. ....	67
Fig. 5.2 Basic mouth shapes of Mandarin initials rendered by SVG. (a) Basic mouth shape <i>m</i> . (b) Basic mouth shape <i>f</i> . (c) Basic mouth shape <i>h</i> .....	68
Fig. 5.3 Basic mouth shapes of Mandarin finals rendered by SVG. (a) Basic mouth shape <i>a</i> . (b) Basic mouth shape <i>i</i> . (c) Basic mouth shape <i>u</i> . (d) Basic mouth shape <i>e</i> . (e) Basic mouth shape <i>o</i> . (f) Basic mouth shape <i>n</i> .....	70
Fig. 5.4 The waveform of the syllable “ㄚ” and the start time of the second phoneme. ....	71
Fig. 5.5 A statistic result of the start time of the second basic mouth shape of the syllables with 2 basic mouth shapes. ....	72
Fig. 5.6 An illustration of time intervals of two basic mouths.....	72
Fig. 5.7 A statistic result of the syllables with 3 basic mouth shapes. (a) The start time of the second basic mouth shape. (b) The time interval between the second and the third basic mouth shape. ....	73
Fig. 5.8 An illustration of time intervals of three basic mouths.....	74
Fig. 5.9 A statistic result of the syllables with 4 basic mouth shapes. (a) The start time of the second basic mouth shape. (b) The time interval between the second and the third basic mouth shape. (c) The time interval between the third and the forth basic mouth shape.....	74
Fig. 5.10 An illustration of time intervals of four basic mouths. ....	75
Fig. 5.11 An overall illustration of talking cartoon face generation. ....	76

Fig. 5.12	Some experimental results of emotions rendered by SVG. (a) Neutral cartoon face. (b) Smile. (c) Sad. (d) Angry. (e) Surprise. ....	77
Fig. 5.13	A resulting sequence of the talking cartoon face for speaking “您好” with a surprise emotion and a randomly generated eye-blinking. ....	78
Fig. 6.1	A flowchart of talking cartoon face generation from image sequence. ....	79
Fig. 6.2	A flowchart of tracking for eyeball regions in image sequences. ....	81
Fig. 6.3	A flowchart of tracking for Mouth regions in image sequences. ....	82
Fig. 6.4	An example of experimental results. (a) A given neutral facial image (the first facial image of the video sequence). (b) A result of automatic neutral cartoon faces generation. ....	83
Fig. 6.5	An example of tracking results. (a) A tracking result of one of the image sequence. (b) Related cartoon face of (a). (c) A tracking result of one of the image sequence. (d) Related cartoon face of (c). (e) A tracking result of one of the image sequence. (f) Related cartoon face of (e). ....	84
Fig. 6.6	Another experimental result. (a) A given neutral facial image (the first facial image of the video sequence). (b) A result of automatic neutral cartoon faces generation. (c)(e)(g) Tracking results of the image sequence. (d)(f)(h) Related cartoon faces. ....	85
Fig. 7.1	A result of an SVG source code. ....	87
Fig. 7.2	An illustration of spatial domain. ....	88
Fig. 7.3	A proposed face model with 84 facial feature points. ....	89
Fig. 7.4	An illustration of the polyline shape. ....	89
Fig. 7.5	An illustration of the circle shape. ....	90
Fig. 7.6	An illustration of eye drawing. ....	90
Fig. 7.7	An illustration of the cubic Bezier curve. ....	90
Fig. 7.8	A result of different types of cartoon face. (a) Normal type. (b) Different types of eyebrows and nose. (c) Cute type. ....	91
Fig. 7.9	A result of adding a layer for clothing with cute-type faces. ....	91
Fig. 7.10	An illustration of temporal domain. ....	92
Fig. 7.11	Cute type cartoon face with head shaking animation. ....	93
Fig. 7.12	An application of proposed methods to electronic books. ....	93
Fig. 7.13	The neutral facial image used for the experiment. ....	94
Fig. 7.14	A video sequence used for the experiment. ....	95
Fig. 7.15	A resulting sequence of the talking cartoon face for “您好” with a random generated eye-blinking. ....	96
Fig. 7.16	Another experimental results of talking cartoon face generation from image sequences, where the speaker in the video clip is speaking “夕陽”. ....	97

# LIST OF TABLES

Table. 5.1	An illustration of the basic elements of Taiwan Tongyoung Romanization.	64
Table. 5.2	Classification of initials according to the manner of articulation proposed in Yeh[16]	65
Table. 5.3	Three basic mouth shapes of Mandarin initials.	66
Table. 5.4	A set of combinations with 7 basic mouth shapes of Mandarin finals.	66

