

DECLARATION

This is to certify that

i) The thesis comprises my original work towards the degree of Master of Engineering in Computer Engineering at Institute of Science and Technology for Advance Research and Studies (ISTAR) and has not been submitted elsewhere for a degree.

ii) Due acknowledgement has been made in the text to all other material used.

Signature of Student

ACKNOWLEDGEMENT

I would like to acknowledge the contribution of certain distinguished people, without their support and guidance this research work would not have been completed.

I take this opportunity to express my sincere thanks and deep sense of gratitude to my project guide *Prof. N. M. Patel Sir*, for his guidance and moral support during the course of preparation of this project report. I really thank him from the rock bottom of my heart for always being there with his extreme knowledge and kind nature. I would also like to thanks all Staff Member of Department for timely help and encouragement of fulfillment of my thesis work.

I am grateful to my college Shri G H Patel College of Engineering and Technology, Vallabh Vidyanagar, for providing me all required stuff and good working environment. I take this opportunity to thank *Prof. Maulika Patel* , HOD, Department of computer engineering, GCET and all my colleagues for their all time support and help in each and every aspect during my ME studies.

I feel a deep sense of gratitude for my family members who formed part of my vision and taught me the good things that really matter in life. I would not have been able to finish my work without their help and love. Finally, I am thankful to almighty god without whose blessings, I would not have achieved anything in life.

Acknowledgements and thanks are also extended to all the authors whose articles have been referred to for the completion of this report.

Namrata Dave
(08CP918)

ABSTRACT

As the computer hardware and other technology become advanced, computer animation has also advanced. As technology has advanced in area of facial animation, need of technology arise which generates realistic facial animation. To generate realistic face animations it is necessary to generate the movement of lips. The goal of this work is to implement a system to analyse an audio signal containing speech, and produce a classification of lip shape categories i.e. viseme in order to synchronize the lips of a computer generated face with the speech. The thesis describes the work to derive a method that maps speech to lip movements. The method is implemented in Matlab on the Windows platform. The program reads speech from pre-recorded audio files that are transformed into sequence of phonemes using phoneme extraction procedure. We have also identified lip parameters for different viseme classes. Neural network is used to map sequence of phonemes to the corresponding viseme class.

Keywords: phonemes, visemes, neural network

CONTENTS

Certificate	ii
Declaration	iii
Acknowledgement	iv
Abstract	v
List of Figures	viii
List of Tables	viii
1. Introduction	01
1.1 Introduction	01
1.2 Thesis Outline	02
	06
2. Literature review	08
2.1	08
2.2	10
2.2.1	11
2.2.2	12
2.2.3	13
2.2.4	14
2.2.5	16
2.2.6	17
2.2.7	17
2.3	19
3. Our Approach	22
3.1	22
3.1.1	22
3.1.2	23

			3.1.3	25	
			3.1.4	25	
			3.1.5	27	
			3.1.6	27	
			3.1.7	28	
			3.1.8	28	
			3.1.9	29	
		3.2			29
		3.3			30
4. Implementation Details						32
		4.1	Solution Overview			32
		4.2				33
			4.2.1		34	
			4.2.2		35	
			4.2.3		35	
		4.3	Algorithm			36
		4.4				38
			4.4.1		38	
			4.4.2		42	
5. Results and Discussion						44
		5.1				44
		5.2				45
		5.3				46
6. Conclusions and Future Work						48
		6.1	Conclusion			48
		6.2	Future Work			49
References						50
Appendices						52
		Appendix A	-		52	
		Appendix B	-		56	
		Appendix C	-		58	
		Appendix D	-		59	

LIST OF FIGURES

I.	24
II.	24
III.	30
IV.	33
V.	45
VI.	46
VIII.	46

LIST OF TABLES

I.	21
II.	43