Visual Speech Recognition: Lip Segmentation and Mapping

Alan Wee-Chung Liew
Griffith University, Australia

Shilin Wang
Shanghai Jiaotong University, China
Table of Contents

Foreword ............................................................................................................................................. xvii
Preface ................................................................................................................................................ xvi

Section I
Introduction and Survey

Chapter I
Audio-Visual and Visual-Only Speech and Speaker Recognition: Issues about Theory, System Design, and Implementation ................................................................. 1

Derek J. Shiell, Northwestern University, USA
Louis H. Terry, Northwestern University, USA
Petar S. Aleksic, Google Inc., USA
Aggelos K. Katsaggelos, Northwestern University, USA

Chapter II
Lip Feature Extraction and Feature Evaluation in the Context of Speech and Speaker Recognition ......................................................................................................................... 39

Petar S. Aleksic, Google Inc., USA
Aggelos K. Katsaggelos, Northwestern University, USA

Chapter III
Lip Modelling and Segmentation ..................................................................................................... 70

A. Caplier, GIPSA-lab/DIS, France
S. Stillittano, GIPSA-lab/DIS, France
C. Bouvier, GIPSA-lab/DIS, France
P. Y. Coulon, GIPSA-lab/DIS, France

Chapter IV
Visual Speech and Gesture Coding Using the MPEG-4 Face and Body Animation Standard........ 128

Eric Petajan, VectorMAX Corporation, USA
Section II
Lip Modeling, Segmentation, and Feature Extraction

Chapter V
Lip Region Segmentation with Complex Background ................................................................. 150
Shilin Wang, Shanghai Jiaotong University, China
Alan Wee-Chung Liew, Griffith University, Australia
Wing Hong Lau, City University of Hong Kong, Hong Kong
Shu Hung Leung, City University of Hong Kong, Hong Kong

Chapter VI
Lip Contour Extraction from Video Sequences under Natural Lighting Conditions.................. 172
Marc Lievin, Avid Technology Inc., Canada
Patrice Delmas, The University of Auckland, New Zealand
Jason James, The University of Auckland, New Zealand
Georgy Gimelfarb, The University of Auckland, New Zealand

Chapter VII
3D Lip Shape SPH Based Evolution Using Prior 2D Dynamic Lip Features Extraction and Static 3D Lip Measurements ........................................................................................................ 213
Alfonso Gastelum, The University of Auckland, New Zealand, & Image Analysis Visualization Laboratory, CCADET-UNAM, Mexico
Patrice Delmas, The University of Auckland, New Zealand
Jorge Marquez, Image Analysis Visualization Laboratory, CCADET-UNAM, Mexico
Alexander Woodward, The University of Auckland, New Zealand
Jason James, The University of Auckland, New Zealand
Marc Lievin, Avid Technology Inc., Canada
Georgy Gimelfarb, The University of Auckland, New Zealand

Chapter VIII
Shafiq ur Rehman, Umeå University, Sweden
Li Liu, Umeå University, Sweden
Haibo Li, Umeå University, Sweden

Section III
Visual Speech Recognition

Chapter IX
Visual Speech Processing and Recognition .................................................................................. 261
Constantine Kotropoulos, Aristotle University of Thessaloniki, Greece
Ioannis Pitas, Aristotle University of Thessaloniki, Greece
Chapter X
Visual Speech Recognition Across Multiple Views .................................................. 294
Patrick Lucey, Queensland University of Technology, Australia
Gerasimos Potamianos, IBM T. J. Watson Research Center, USA
Sridha Sridharan, Queensland University of Technology, Australia

Chapter XI
Hidden Markov Model Based Visemes Recognition, Part I: AdaBoost Approach .......... 326
Say Wei Foo, Nanyang Technological University, Singapore
Liang Dong, National University of Singapore, Singapore

Chapter XII
Hidden Markov Model Based Visemes Recognition, Part II: Discriminative Approaches ... 356
Say Wei Foo, Nanyang Technological University, Singapore
Liang Dong, National University of Singapore, Singapore

Chapter XIII
Motion Features for Visual Speech Recognition ..................................................... 388
Wai Chee Yau, RMIT University, Australia
Dinesh Kant Kumar, RMIT University, Australia
Hans Weghorn, BA University of Cooperative Education Stuttgart, Germany

Chapter XIV
Recognizing Prosody from the Lips: Is It Possible to Extract Prosodic Focus from Lip Features? .......................................................... 416
Marion Dohen, GIPSA-lab, France
Hélène Levenbruck, GIPSA-lab, France
Harold Hill, ATR Cognitive Information Science Labs, Japan, & University of Wollongong, Australia

Chapter XV
Visual Speech Perception, Optical Phonetics, and Synthetic Speech ......................... 439
Lynne E. Bernstein, House Ear Institute, Los Angeles, USA
Jintao Jiang, House Ear Institute, Los Angeles, USA

Section IV
Visual Speaker Recognition

Chapter XVI
Multimodal Speaker Identification Using Discriminative Lip Motion Features ................. 463
H. Ertan Çetingül, John Hopkins University, USA
Engin Erzin, Koç University, Turkey
Yücel Yemez, Koç University, Turkey
A. Murat Tekalp, Koç University, Turkey
Chapter XVII
Lip Motion Features for Biometric Person Recognition ................................................................. 495
    Maycel Isaac Faraj, Halmstad University, Sweden
    Josef Bigun, Halmstad University, Sweden

About the Contributors .................................................................................................................. 533

Index............................................................................................................................................ 543