Wenyi Zhao
Shaogang Gong
Xiaoou Tang (Eds.)

LNCS 3723

Analysis and Modelling of Faces and Gestures

Second International Workshop, AMFG 2005
Beijing, China, October 2005
Proceedings

Springer

Wenyi Zhao Shaogang Gong
Xiaoou Tang (Eds.)

Analysis and Modelling of Faces and Gestures

Second International Workshop, AMFG 2005
Beijing, China, October 16, 2005
Proceedings

Springer

Volume Editors

Wenyi Zhao
Vision Technologies Lab
Sarnoff Corporation
Princeton, NJ 08873, USA
E-mail: wzhao@sarnoff.com

Shaogang Gong
Queen Mary, University of London
Department of Computer Science
London E1 4NS UK
E-mail: sgg@dcs.qmul.ac.uk

Xiaoou Tang
Microsoft Research Asia
Beijing Sigma Center
10080 Beijing, China
E-mail: xitang@microsoft.com

Library of Congress Control Number: 2005933260

CR Subject Classification (1998): I.4, I.5, I.3.5, I.2.10, I.2.6, F.2.2

ISSN	0302-9743
ISBN-10	3-540-29229-2 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-29229-6 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11564386 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 3723

Lecture Notes in Computer Science

For information about Vols. 1–3645

please contact your bookseller or Springer

Vol. 3750: J. Duncan, G. Gerig (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005, Part II. XL, 1018 pages. 2005.

Vol. 3749: J. Duncan, G. Gerig (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005, Part I. XXXIX, 942 pages. 2005.

Vol. 3739: W. Fan, Z. Wu, J. Yang (Eds.), Advances in Web-Age Information Management. XXII, 930 pages. 2005.

Vol. 3738: V.R. Syrotiuk, E. Chávez (Eds.), Ad-Hoc, Mobile, and Wireless Networks. XI, 360 pages. 2005.

Vol. 3735: A. Hoffmann, H. Motoda, T. Scheffer (Eds.), Discovery Science. XVI, 400 pages. 2005. (Subseries LNAI).

Vol. 3734: S. Jain, H.U. Simon, E. Tomita (Eds.), Algorithmic Learning Theory. XII, 490 pages. 2005. (Subseries LNAI).

Vol. 3731: F. Wang (Ed.), Formal Techniques for Networked and Distributed Systems - FORTE 2005. XII, 558 pages. 2005.

Vol. 3728: V. Paliouras, J. Vounckx, D. Verkest (Eds.), Integrated Circuit and System Design. XV, 753 pages. 2005.

Vol. 3726: L.T. Yang, O.F. Rana, B. Di Martino, J. Dongarra (Eds.), High Performance Computing and Communcations. XXVI, 1116 pages. 2005.

Vol. 3725: D. Borrione, W. Paul (Eds.), Correct Hardware Design and Verification Methods. XII, 412 pages. 2005.

Vol. 3724: P. Fraigniaud (Ed.), Distributed Computing. XIV, 520 pages. 2005.

Vol. 3723: W. Zhao, S. Gong, X. Tang (Eds.), Analysis and Modelling of Faces and Gestures. XI, 423 pages. 2005.

Vol. 3722: D. Van Hung, M. Wirsing (Eds.), Theoretical Aspects of Computing – ICTAC 2005. XIV, 614 pages. 2005.

Vol. 3721: A. Jorge, L. Torgo, P. Brazdil, R. Camacho, J. Gama (Eds.), Knowledge Discovery in Databases: PKDD 2005. XXIII, 719 pages. 2005. (Subseries LNAI).

Vol. 3720: J. Gama, R. Camacho, P. Brazdil, A. Jorge, L. Torgo (Eds.), Machine Learning: ECML 2005. XXIII, 769 pages. 2005. (Subseries LNAI).

Vol. 3719: M. Hobbs, A.M. Goscinski, W. Zhou (Eds.), Algorithms and Architectures for Parallel Processing. XI, 448 pages. 2005.

Vol. 3718: V.G. Ganzha, E.W. Mayr, E.V. Vorozhtsov (Eds.), Computer Algebra in Scientific Computing. XII, 502 pages. 2005.

Vol. 3717: B. Gramlich (Ed.), Frontiers of Combining Systems. X, 321 pages. 2005. (Subseries LNAI).

Vol. 3715: E. Dawson, S. Vaudenay (Eds.), Progress in Cryptology – Mycrypt 2005. XI, 329 pages. 2005.

Vol. 3714: H. Obbink, K. Pohl (Eds.), Software Product Lines. XIII, 235 pages. 2005.

Vol. 3713: L. Briand, C. Williams (Eds.), Model Driven Engineering Languages and Systems. XV, 722 pages. 2005.

Vol. 3712: R. Reussner, J. Mayer, J.A. Stafford, S. Overhage, S. Becker, P.J. Schroeder (Eds.), Quality of Software Architectures and Software Quality. XIII, 289 pages. 2005.

Vol. 3711: F. Kishino, Y. Kitamura, H. Kato, N. Nagata (Eds.), Entertainment Computing - ICEC 2005. XXIV, 540 pages. 2005.

Vol. 3710: M. Barni, I. Cox, T. Kalker, H.J. Kim (Eds.), Digital Watermarking. XII, 485 pages. 2005.

Vol. 3708: J. Blanc-Talon, W. Philips, D. Popescu, P. Scheunders (Eds.), Advanced Concepts for Intelligent Vision Systems. XXII, 725 pages. 2005.

Vol. 3707: D.A. Peled, Y.-K. Tsay (Eds.), Automated Technology for Verification and Analysis. XII, 506 pages. 2005.

Vol. 3706: H. Fuks, S. Lukosch, A.C. Salgado (Eds.), Groupware: Design, Implementation, and Use. XII, 378 pages. 2005.

Vol. 3703: F. Fages, S. Soliman (Eds.), Principles and Practice of Semantic Web Reasoning. VIII, 163 pages. 2005.

Vol. 3702: B. Beckert (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. XIII, 343 pages. 2005. (Subseries LNAI).

Vol. 3701: M. Coppo, E. Lodi, G. M. Pinna (Eds.), Theoretical Computer Science. XI, 411 pages. 2005.

Vol. 3699: C.S. Calude, M.J. Dinneen, G. Păun, M. J. Pérez-Jiménez, G. Rozenberg (Eds.), Unconventional Computation. XI, 267 pages. 2005.

Vol. 3698: U. Furbach (Ed.), KI 2005: Advances in Artificial Intelligence. XIII, 409 pages. 2005. (Subseries LNAI).

Vol. 3697: W. Duch, J. Kacprzyk, E. Oja, S. Zadrożny (Eds.), Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005, Part II. XXXII, 1045 pages. 2005.

Vol. 3696: W. Duch, J. Kacprzyk, E. Oja, S. Zadrożny (Eds.), Artificial Neural Networks: Biological Inspirations – ICANN 2005, Part I. XXXI, 703 pages. 2005.

Vol. 3695: M.R. Berthold, R. Glen, K. Diederichs, O. Kohlbacher, I. Fischer (Eds.), Computational Life Sciences. XI, 277 pages. 2005. (Subseries LNBI).

Vol. 3694: M. Malek, E. Nett, N. Suri (Eds.), Service Availability. VIII, 213 pages. 2005.

Vol. 3693: A.G. Cohn, D.M. Mark (Eds.), Spatial Information Theory. XII, 493 pages. 2005.

Vol. 3692: R. Casadio, G. Myers (Eds.), Algorithms in Bioinformatics. X, 436 pages. 2005. (Subseries LNBI).

Vol. 3691: A. Gagalowicz, W. Philips (Eds.), Computer Analysis of Images and Patterns. XIX, 865 pages. 2005.

Vol. 3690: M. Pěchouček, P. Petta, L.Z. Varga (Eds.), Multi-Agent Systems and Applications IV. XVII, 667 pages. 2005. (Subseries LNAI).

Vol. 3688: R. Winther, B.A. Gan, G. Dahll (Eds.), Computer Safety, Reliability, and Security. XI, 405 pages. 2005.

Vol. 3687: S. Singh, M. Singh, C. Apte, P. Perner (Eds.), Pattern Recognition and Image Analysis, Part II. XXV, 809 pages. 2005.

Vol. 3686: S. Singh, M. Singh, C. Apte, P. Perner (Eds.), Pattern Recognition and Data Mining, Part I. XXVI, 689 pages. 2005.

Vol. 3685: V. Gorodetsky, I. Kotenko, V. Skormin (Eds.), Computer Network Security. XIV, 480 pages. 2005.

Vol. 3684: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part IV. LXXIX, 933 pages. 2005. (Subseries LNAI).

Vol. 3683: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part III. LXXX, 1397 pages. 2005. (Subseries LNAI).

Vol. 3682: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. LXXIX, 1371 pages. 2005. (Subseries LNAI).

Vol. 3681: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LXXX, 1319 pages. 2005. (Subseries LNAI).

Vol. 3679: S.d.C. di Vimercati, P. Syverson, D. Gollmann (Eds.), Computer Security – ESORICS 2005. XI, 509 pages. 2005.

Vol. 3678: A. McLysaght, D.H. Huson (Eds.), Comparative Genomics. VIII, 167 pages. 2005. (Subseries LNBI).

Vol. 3677: J. Dittmann, S. Katzenbeisser, A. Uhl (Eds.), Communications and Multimedia Security. XIII, 360 pages. 2005.

Vol. 3676: R. Glück, M. Lowry (Eds.), Generative Programming and Component Engineering. XI, 448 pages. 2005.

Vol. 3675: Y. Luo (Ed.), Cooperative Design, Visualization, and Engineering. XI, 264 pages. 2005.

Vol. 3674: W. Jonker, M. Petković (Eds.), Secure Data Management. X, 241 pages. 2005.

Vol. 3673: S. Bandini, S. Manzoni (Eds.), AI*IA 2005: Advances in Artificial Intelligence. XIV, 614 pages. 2005. (Subseries LNAI).

Vol. 3672: C. Hankin, I. Siveroni (Eds.), Static Analysis. X, 369 pages. 2005.

Vol. 3671: S. Bressan, S. Ceri, E. Hunt, Z.G. Ives, Z. Bellahsène, M. Rys, R. Unland (Eds.), Database and XML Technologies. X, 239 pages. 2005.

Vol. 3670: M. Bravetti, L. Kloul, G. Zavattaro (Eds.), Formal Techniques for Computer Systems and Business Processes. XIII, 349 pages. 2005.

Vol. 3669: G.S. Brodal, S. Leonardi (Eds.), Algorithms – ESA 2005. XVIII, 901 pages. 2005.

Vol. 3668: M. Gabbrielli, G. Gupta (Eds.), Logic Programming. XIV, 454 pages. 2005.

Vol. 3666: B.D. Martino, D. Kranzlmüller, J. Dongarra (Eds.), Recent Advances in Parallel Virtual Machine and Message Passing Interface. XVII, 546 pages. 2005.

Vol. 3665: K. S. Candan, A. Celentano (Eds.), Advances in Multimedia Information Systems. X, 221 pages. 2005.

Vol. 3664: C. Türker, M. Agosti, H.-J. Schek (Eds.), Peer-to-Peer, Grid, and Service-Orientation in Digital Library Architectures. X, 261 pages. 2005.

Vol. 3663: W.G. Kropatsch, R. Sablatnig, A. Hanbury (Eds.), Pattern Recognition. XIV, 512 pages. 2005.

Vol. 3662: C. Baral, G. Greco, N. Leone, G. Terracina (Eds.), Logic Programming and Nonmonotonic Reasoning. XIII, 454 pages. 2005. (Subseries LNAI).

Vol. 3661: T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, P. Olivier, T. Rist (Eds.), Intelligent Virtual Agents. XIII, 506 pages. 2005. (Subseries LNAI).

Vol. 3660: M. Beigl, S. Intille, J. Rekimoto, H. Tokuda (Eds.), UbiComp 2005: Ubiquitous Computing. XVII, 394 pages. 2005.

Vol. 3659: J.R. Rao, B. Sunar (Eds.), Cryptographic Hardware and Embedded Systems – CHES 2005. XIV, 458 pages. 2005.

Vol. 3658: V. Matoušek, P. Mautner, T. Pavelka (Eds.), Text, Speech and Dialogue. XV, 460 pages. 2005. (Subseries LNAI).

Vol. 3657: F.S. de Boer, M.M. Bonsangue, S. Graf, W.-P. de Roever (Eds.), Formal Methods for Components and Objects. VIII, 325 pages. 2005.

Vol. 3656: M. Kamel, A. Campilho (Eds.), Image Analysis and Recognition. XXIV, 1279 pages. 2005.

Vol. 3655: A. Aldini, R. Gorrieri, F. Martinelli (Eds.), Foundations of Security Analysis and Design III. VII, 273 pages. 2005.

Vol. 3654: S. Jajodia, D. Wijesekera (Eds.), Data and Applications Security XIX. X, 353 pages. 2005.

Vol. 3653: M. Abadi, L. de Alfaro (Eds.), CONCUR 2005 – Concurrency Theory. XIV, 578 pages. 2005.

Vol. 3652: A. Rauber, S. Christodoulakis, A M. Tjoa (Eds.), Research and Advanced Technology for Digital Libraries. XVIII, 545 pages. 2005.

Vol. 3651: R. Dale, K.-F. Wong, J. Su, O.Y. Kwong (Eds.), Natural Language Processing – IJCNLP 2005. XXI, 1031 pages. 2005. (Subseries LNAI).

Vol. 3650: J. Zhou, J. Lopez, R.H. Deng, F. Bao (Eds.), Information Security. XII, 516 pages. 2005.

Vol. 3649: W.M. P. van der Aalst, B. Benatallah, F. Casati, F. Curbera (Eds.), Business Process Management. XII, 472 pages. 2005.

Vol. 3648: J.C. Cunha, P.D. Medeiros (Eds.), Euro-Par 2005 Parallel Processing. XXXVI, 1299 pages. 2005.

Vol. 3646: A. F. Famili, J.N. Kok, J.M. Peña, A. Siebes, A. Feelders (Eds.), Advances in Intelligent Data Analysis VI. XIV, 522 pages. 2005.

Preface

During the last 30 years, face recognition and related problems such as face detection/tracking and facial expression recognition have attracted researchers from both the engineering and psychology communities. In addition, extensive research has been carried out to study hand and body gestures. The understanding of how humans perceive these important cues has significant scientific value and extensive applications. For example, human-computer interaction, visual surveillance, and smart video indexing are active application areas. Aiming towards putting such amazing perception capability onto computer systems, researchers have made substantial progress. However, technological challenges still exist in many aspects.

Following a format similar to the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG) 2003, this one-day workshop (AMFG 2005) provided a focused international forum to bring together well-known researchers and research groups to review the status of recognition, analysis and modeling of faces and gestures, to discuss the challenges that we are facing, and to explore future directions. Overall, 30 papers were selected from 90 submitted manuscripts. The topics of these papers range from feature representation, robust recognition, learning, and 3D modeling to psychology. In addition, two invited talks were given, by Prof. Kanade and Dr. Phillips. The technical program was organized into four oral sessions and two poster sessions.

This workshop would not have been possible without the timely reviews provided by the members of the Technical Program Committee under a tight schedule.

October 2005

Wenyi Zhao
Shaogang Gong
Xiaoou Tang

Conference Organization

Workshop Chairs

Wenyi Zhao, Sarnoff Corporation, USA
Shaogang Gong, Queen Mary, University of London, UK
Andrew Senior, IBM T.J. Watson Research Center, USA
Xiaoou Tang, Microsoft Research Asia, China

Advisory Committee

Rama Chellappa, University of Maryland, USA
Thomas S. Huang, University of Illinois at Urbana Champaign, USA
Anil Jain, Michigan State University, USA

Local Arrangement Co-chairs

Shuicheng Yan, The Chinese University of Hong Kong, Hong Kong, China
Qingshan Liu, Chinese Academy of Sciences, China

Program Committee

Richard Bowden	Surrey, UK
Kevin Bowyer	Notre Dame, USA
Rama Chellappa	Maryland, USA
Jeff Cohn	Pittsburgh, USA
Robert Collins	Penn State, USA
Tim Cootes	Manchester, UK
David Hogg	Leeds, UK
Anil Jain	Michigan State, USA
David Jacobs	Maryland, USA
Ron Kimmel	Technion, Israel
Josef Kittler	Surrey, UK
David Kriegman	UCSD, USA
Stan Z. Li	NLPR, China
Chengjun Liu	NJIT, USA
Zhichen Liu	Microsoft Research, USA
Gerard Medioni	USC, USA

Baback Moghaddam	MERL, USA
Mark Nixon	Southampton, UK
Jonathon Phillips	NIST, USA
Long Quan	UST, HK
Pawan Sinha	MIT, USA
Matthew Turk	UCSB, USA
David Zhang	HKPU, HK

Table of Contents

Oral Sessions

Representation

Recognition

Learning

3D Modeling

Poster Sessions

Session I

Session II

Facial Expression Analysis

Takeo Kanade

U.A. and Helen Whitaker University Professor, Robotics Institute,
Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract. Computer analysis of human face images includes detection of faces, identification of people, and understanding of expression. Among these three tasks, facial expression has been the least studied, and most of the past work on facial expression tried to recognize a small set of emotions, such as joy, disgust, and surprise. This practice may follow from the work of Darwin, who proposed that emotions have corresponding prototypic facial expressions. In everyday life, however, such prototypic expressions occur relatively infrequently; instead, emotion is communicated more often by subtle changes in one or a few discrete features. FACS-code Action Units, defined by Ekman, are one such representation accepted in the psychology community.

In collaboration with psychologists, we have been developing a system for automatically recognizing facial action units. This talk will present the current version of the system. The system uses a 3D Active Appearance Model to align a face image and transform it to a person-specific canonical coordinate frame. This transformation can remove appearance changes due to changes of head pose and relative illumination direction. In this transformed image frame, we perform detailed analysis of both facial motion and facial appearance changes, results of which are fed to an action-unit recogniser.

W. Zhao, S. Gong, and X. Tang (Eds.): AMFG 2005, LNCS 3723, p. 1, 2005.

Modeling Micro-patterns for Feature Extraction

Qiong Yang[1], Dian Gong[1, 2], and Xiaoou Tang[1]

[1] Microsoft Research Asia, Beijing Sigma Center,
100080 Beijing, China
{qyang, xitang}@microsoft.com
[2] Department of Electronic Engineering, Tsinghua University,
100084 Beijing, China
gongd@wmc.ee.tsinghua.edu.cn

Abstract. Currently, most of the feature extraction methods based on micro-patterns are application oriented. The micro-patterns are intuitively user-designed based on experience. Few works have built models of micro-patterns for feature extraction. In this paper, we propose a model-based feature extraction approach, which uses micro-structure modeling to design adaptive micro-patterns. We first model the micro-structure of the image by Markov random field. Then we give the generalized definition of micro-pattern based on the model. After that, we define the fitness function and compute the fitness index to encode the image's local fitness to micro-patterns. Theoretical analysis and experimental results show that the new algorithm is both flexible and effective in extracting good features.

1 Introduction

Feature extraction is one of the most important issues in pattern recognition. In previous studies, people observed that the spatial context in images plays an important role in many vision tasks, such as character recognition, object detection and recognition. So they design micro-patterns to describe the spatial context of the image, such as edge, line, spot, blob, corner, and more complex patterns. Furthermore, it is observed that the regional characteristics of micro-patterns are more robust to shift and scale, so a number of features are developed to calculate the regional characteristics of micro-patterns. These features include:

a) *Orientation Histogram.* This kind of features designs the micro-pattern as directional line or edge, and calculates the histogram of each direction in the region. It has been used as an informative tool for various vision tasks. Sun and Si [1] used orientation histograms to find the symmetry axis in an image. Freeman and Roth [2] developed a method for hand gesture recognition based on the global orientation histogram of the image. Lowe [3] developed a scale-invariant feature from local orientation histograms for object recognition. Levi and Weiss [4] used local edge orientation histograms (EOH) as features to improve performance in object detection as well as face detection. Another example is that Four Directional Line Element (FDLE) [5] has been successfully used for character recognition.

b) *Filter Banks.* In this kind of features, a bank of filters is designed to extract the micro-structural features, and the regional characteristics are computed from the filter

W. Zhao, S. Gong, and X. Tang (Eds.): AMFG 2005, LNCS 3723, pp. 2–16, 2005.

response. Goudail et al. [6] designed a series of local autocorrelation filters for face recognition, and the filter response is summed over the global image to form the feature. Wang et al. [7] used the histogram of regulated outputs of Gabor filters for Chinese character recognition in low-quality images.

c) *Local Binary Pattern.* This feature is designed for texture analysis [8], face detection and face recognition [9]. The image is first divided into small regions, from which Local Binary Pattern (LBP) histograms are extracted and concatenated into a single feature histogram to efficiently represent the image.

In all these features, the micro-patterns are intuitively user-designed based on experience, and they are application oriented. The micro-patterns fit for one task might be unfit for another. For example, FDLE [5] is successful in character recognition, but might not achieve the same success in face recognition, since face image is much more complex than the character image so that it cannot be simply represented by directional lines. Another problem is that in some cases, it is difficult for the user to intuitively determine whether the micro-pattern is appropriate unless he refers to the experimental result. A similar problem exists for Gabor features. Although in many papers Gabor has been used to recognize a general object as well as face [10,11,12], the parameters are mainly adjusted by experimental results, and it costs a lot of time and efforts to find the appropriate parameters.

In this paper, we propose a model-based feature extraction approach, which uses Markov random field (MRF) to model the micro-structure of the image and design adaptive micro-patterns for feature extraction. The key idea is motivated by several observations:

First of all, image structure modeling can help us find good features in at least three aspects: 1) Modeling could provide sound theoretical foundations and guide us on how to design suitable micro-patterns. 2) Through modeling, the feature extraction method could be more general, and also more applicable to various applications. 3) Modeling will alleviate the efforts in adjusting parameters. Therefore, we introduce image structure modeling in the stage of feature extraction.

Secondly, Markov field [13,14,15,17,18,19,20,21] provides a flexible mechanism for modeling spatial dependence. If we study the spatial dependence in a local region of the image, it will model the micro-patterns, with different spatial dependency corresponding to different micro-patterns. It is also convenient for representing unobserved complex patterns of images, especially the location of discontinuities between regions homogeneous in tone, texture or depth. Therefore it is possible for using Markov field to model the micro-patterns, not limited to the simple ones, but also the complex patterns. Moreover, the parameters of the model can be statistically learned from samples, instead of intuitively user-designed. Thereby it is more adaptive to the local characteristics of images. Different micro-patterns will be designed for different kinds of images, different attributes of images, and even at different sites of an image, so features will be more flexible, and also more applicable to various applications.

Based on the above observations, we use MRF to extract block-level micro-structural features. We first divide the image into sub-blocks and use MRF to model the micro-patterns in each sub-block. Based on that, we compute the local fitness sequence to describe the image's local fitness to micro-patterns. Then, we extract the modified FFT (fast Fourier transform) feature of the local fitness sequence in each sub-block. Finally, we concatenate these features from all sub-blocks into a long feature vector. The new feature presents a description of the image on three levels: the Markov field model reflects the spatial correlation of neighborhood in a pixel-level; the local fitness sequence in each sub-block reflects the image's regional fitness to micro-patterns in a block level; and the features from all sub-blocks are concatenated to build a global description of the image. In this way, both the local textures and the global shape of the image are simultaneously encoded.

2 Feature Extraction from Micro-structure Modeling

2.1 Markov Random Field Model

Let I represent a $H \times W$ image with S as its collection of all sites, and let $X_s = x_s$ represent some attribute of the image I at site $s \in S$. The attribute may be grayscale intensity, Gabor attribute or other features. Also, we denote the attributes of all other sites in S excluding site s by $X_{-s} = x_{-s}$. The spatial distribution of attributes on S, $\boldsymbol{X} = \boldsymbol{x} = \{x_s, s \in S\}$, will be modeled as a Markov random field (MRF).

Let N_s denote the neighbors of site s, and the r-th order neighborhood is defined to be $N_s^{(r)} = \{t \mid dist(s,t) \le r, t \in S\}$, where $dist(s,t)$ is the distance between site s and site t. The 1-st and 2-nd order neighborhood structure are displayed in Fig. 1. Because of the local property (i.e. Markovianity: $p(X_s = x_s | X_{-s} = x_{-s}) = p(X_s = x_s | X_{N_s} = x_{N_s})$), the Markov model is equivalent to the Gibbs random field, so we use the energy function to calculate the probability as follows

$$p(X_s | X_{-s}) = p(X_s | X_{N_s}) = \frac{1}{\mathcal{T}} \exp\{-E_{\theta_s}(X_s, X_{N_s})\}, \tag{1}$$

where $E_{\theta_s}(X_s, X_{N_s})$ is the energy function at site s which is the sum of energies/potentials of the cliques containing site s, and $\mathcal{T} = \sum_{X_s} \exp\{-E_{\theta_s}(X_s, X_{N_s})\}$ is the partition function. Here, θ_s is the parameter set for site s, so we rewrite $p(X_s | X_{N_s})$ into $p_{\theta_s}(X_s | X_{N_s})$.

For a pair-wise MRF model, there is $E_{\theta_s}(X_s, X_{N_s}) = H_s(X_s) + \sum_{t \in N_s} J_{st}(X_s, X_t)$, where $H_s(X_s)$ is the "field" at site s, and $J_{st}(X_s, X_t)$ is the "interaction" between site s and site t. Furthermore, if $H_s(X_s) = 0$ and $J_{st}(X_s, X_t) = \frac{1}{(\sigma_{st})^2}(X_s - X_t)^2$, then we get the smooth model and there is $E_{\theta_s}(X_s, X_{N_s}) = \sum_{t \in N_s} \frac{1}{(\sigma_{st})^2}(X_s - X_t)^2$, $\theta_s = \{\sigma_{st}, t \in N_s\}$. If $H_s(X_s) = \alpha_s X_s$, $J_{st}(X_s, X_t) = \beta_{st} X_s X_t$ and $X_s \in \{+1, -1\}, s \in S$, then we get the Ising model and there is $E_{\theta_s}(X_s, X_{N_s}) = \alpha_s X_s + \sum_{t \in N_s} \beta_{st} X_s X_t$, $\theta_s = \{\alpha_s, \beta_{st}, t \in N_s\}$. For simplicity, we write θ_s as θ.

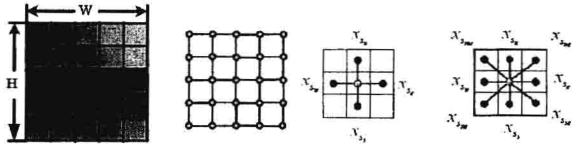

Fig. 1. (a) The image I with the size of H×W. (b) The site set S of the image I. (c) The 1-st order neighborhood structure. (d) The 2-nd order neighborhood structure. Here, H=W=5.

2.2 Feature Extraction

In this section, we will discuss how we extract features based on Markov random field model. Firstly, we propose a generalized definition of micro-pattern, and then we design a fitness function to extract the image's local fitness to micro-patterns.

2.2.1 Generalized Definition of Micro-patterns

Assume that Ω denotes the *micro-pattern*, and $\Omega_\theta(\gamma)$ is defined to be all the pairs of (x_s, x_{N_s}) that satisfy the constraint $g_\theta(x_s, x_{N_s}) = \gamma$ with given θ, i.e. $\{(x_s, x_{N_s}) : g_\theta(x_s, x_{N_s}) = \gamma\}$. Here, θ is the parameter set.

$\Omega_\theta(\gamma)$ has the following properties:

1. Given θ, $\{\Omega_\theta(\gamma), \gamma \in \mathcal{R}\}$ describes a series of micro-patterns where $\mathcal{R}$ is the value set of γ.
2. When γ is discrete, $\Omega_\theta(\gamma)$ is characterized by its probability $P(\Omega = \Omega_\theta(\gamma))$; when γ is a continuous variable, $\Omega_\theta(\gamma)$ is characterized by the probability density function $p(\Omega_\theta(\gamma))$.

In this paper, since we use MRF model, we define $g_\theta(x_s, x_{N_s}) = E_\theta(X_s = x_s, X_{N_s} = x_{N_s})$, therefore

$$\Omega_\theta(\gamma) = \{(x_s, x_{N_s}) : E_\theta(X_s = x_s, X_{N_s} = x_{N_s}) = \gamma\} \tag{2}$$

That is, (x_s, x_{N_s}) in the same level of energy belong to the same micro-pattern.

a) When we use the smooth model, i.e. $E_\theta(X_s, X_{N_s}) = \sum_{t \in N_s} \frac{1}{(\sigma_{st})^2}(X_s - X_t)^2$, then

$$\Omega_\theta(\gamma) = \left\{(x_s, x_{N_s}) : \sum_{t \in N_s} \frac{1}{(\sigma_{st})^2}(x_s - x_t)^2 = \gamma\right\} \tag{3}$$

In this sense, Fig. 2(a) and Fig. 2(b) are deemed to be same, while Fig 2(c) and Fig. 2(d) are deemed to be different micro-patterns.

b) When we use the Ising model, i.e. $H_\theta(X_s | X_{N_s}) = \alpha_s X_s + \sum_{t \in N_s} \beta_{st} X_s X_t$ (with 1-st neighborhood), where $X_s \in \{+1, -1\}, \forall s \in S$ and $\theta = \{\alpha_s, \beta_{st}, t \in N_s\}$ is as shown in Fig.2(e), there is