

## DOCTORAL THESIS

### Dependency modeling for information fusion with applications in visual recognition

Ma, Jinhua

*Date of Award:*  
2013

[Link to publication](#)

#### General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

# Dependency Modeling for Information Fusion with Applications in Visual Recognition

MA Jinhua

A thesis submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Principal Supervisor: Prof. YUEN Pong Chi

Hong Kong Baptist University

August 2013

# Abstract

While many pattern recognition algorithms have been developed in the last forty years, classifying images/videos in practical applications still faces the challenges of self/mutual occlusions, clustered backgrounds, illumination variations, etc. In order to improve the recognition performance, many systems are designed by fusing multiple complementary features for various classification tasks. This thesis addresses the independent assumption issue in the fusion process and proposes two novel frameworks for dependency modeling.

Under some mild assumptions, the first approach uses a linear combination of posterior probabilities to model the feature dependency. Based on the linear combination property, this thesis proposes a Linear Classifier Dependency Modeling (LCDM) method for classifier level fusion. Under the linear dependency modeling framework, this thesis shows that more information about the class label is available in feature level, so LCDM is generalized to feature level and the Linear Feature Dependency Model (LFDM) is proposed.

Since it is almost impossible to verify whether the assumptions in existing methods are valid in practice applications, fusion method with less demanding assumption should give better performance. In the second approach, this thesis develops an Analytic Dependency Model (ADM) for score level fusion without the assumptions in existing fusion algorithms. With the proposed ADM, this thesis gives an equivalent condition to the independent assumption from probabilistic properties of marginal distributions. Since the ADM may contain infinite number of undeter-

mined coefficients, this thesis further proposes the Reduced Analytic Dependency Model (RADM) based on the convergent properties of analytic functions.

While the proposed fusion methods overcome some limitations in existing approaches, the fusion performance can be further improved by combining more discriminative features. Among feature extraction algorithms, supervised manifold learning has been successfully applied to many image classification problems. However, for video applications, existing manifold learning methods do not take full advantage of the global constraint of temporal labels. To overcome this limitation, this thesis proposes a new Supervised Spatio-Temporal Neighborhood Topology Learning (SSTNTL) method for video classification.

The proposed methods have been extensively evaluated on publicly available databases such as PASCAL VOC 2007, Columbia Consumer Video, Hollywood Human Action, etc., and convincing experimental results have been achieved. In short, the major contributions of this thesis are summarized as follows.

- A linear dependency modeling framework is developed for classifier level and feature level fusion.
- A Reduced Analytic Dependency Model (RADM) is derived for score level fusion with less demanding assumption.
- A Supervised Spatio-Temporal Neighborhood Topology Learning (SSTNTL) method is proposed for video classification.

# Table of Contents

<b>Declaration</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Table of Contents</b>	<b>v</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>xi</b>
<b>List of Symbols</b>	<b>xiv</b>
<b>List of Abbreviations</b>	<b>xv</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.2 Motivations of This Project . . . . .	3
1.3 Contributions of This Thesis . . . . .	5
1.4 Thesis Overview . . . . .	7
<b>Chapter 2 Related Works and Databases</b>	<b>8</b>
2.1 Probabilistic Fusion Methods . . . . .	8
2.1.1 Fusion Models with Independent Assumption . . . . .	8
2.1.2 Fusion Models without Independent Assumption . . . . .	9

2.2	Non-Probabilistic Fusion Methods . . . . .	10
2.3	Databases for Evaluation . . . . .	12
2.3.1	Digit Database . . . . .	12
2.3.2	Oxford 17 Flower Database . . . . .	12
2.3.3	CMU PIE Face Database . . . . .	13
2.3.4	FERET Face Database . . . . .	14
2.3.5	Weizmann Human Action Database . . . . .	14
2.3.6	KTH Human Action Database . . . . .	16
2.3.7	UCF Sports Database . . . . .	18
2.3.8	Hollywood Human Action Database . . . . .	18
2.3.9	Hand Gesture Action Database . . . . .	19
2.3.10	VOC 2007 Object Categorization Database . . . . .	20
2.3.11	Columbia Consumer Video Database . . . . .	21

**Chapter 3 Manifold learning for Spatio-Temporal Feature Representation 22**

3.1	Introduction . . . . .	22
3.2	Review on Manifold Learning for Video Analysis . . . . .	24
3.3	Revisiting Locality Preserving Projection and Its Supervised Version	25
3.4	Supervised Spatio-Temporal Neighborhood Topology Learning . . . . .	26
3.4.1	Topological Analysis for Action Recognition . . . . .	27
3.4.2	Supervised Spatial Neighborhood Topology Construction . . .	30
3.4.3	Temporal Pose Correspondence Neighborhood Topology Con- struction . . . . .	31
3.4.4	Supervised Spatial and Temporal Pose Correspondence Neigh- borhood Topology Learning . . . . .	35
3.5	Experiments . . . . .	37
3.5.1	Settings and Classifier . . . . .	37
3.5.2	Results on Weizmann Human Action Database . . . . .	38

3.5.3	Results on KTH Human Action Database . . . . .	41
3.5.4	Results on UCF Sports Database . . . . .	44
3.5.5	Results on HOHA Database . . . . .	45
3.5.6	Results on Hand Gesture Action Database . . . . .	46
3.5.7	Comparing SSTNTL with and without TPC Neighbors . . . . .	51
3.6	Summary . . . . .	51
<b>Chapter 4 Linear Dependency Modeling</b>		<b>53</b>
4.1	Introduction . . . . .	53
4.2	Linear Dependency Modeling . . . . .	55
4.2.1	Linear Dependency Modeling in Classifier Level . . . . .	55
4.2.2	Linear Dependency Modeling in Feature Level . . . . .	58
4.2.3	Learning Optimal Linear Dependency Model . . . . .	60
4.2.4	Sensitivity to Density Estimation Error . . . . .	63
4.2.5	Remarks . . . . .	65
4.3	Experiments . . . . .	69
4.3.1	Results on Synthetic Data . . . . .	69
4.3.2	Results on Oxford 17 Flower Database . . . . .	70
4.3.3	Results on Digit Database . . . . .	71
4.3.4	Results on Human Action Databases . . . . .	73
4.3.5	Analysis of Dependency and Learning Results . . . . .	78
4.4	Summary . . . . .	81
<b>Chapter 5 Reduced Analytic Dependency Modeling</b>		<b>82</b>
5.1	Introduction . . . . .	82
5.2	Reduced Analytic Dependency Modeling . . . . .	84
5.2.1	Analytic Dependency Modeling . . . . .	85
5.2.2	Reduced Model . . . . .	88
5.2.3	Model Learning . . . . .	89
5.3	Experiments . . . . .	96

5.3.1	Results on Digit Database . . . . .	96
5.3.2	Results on Oxford 17 Flower Database . . . . .	97
5.3.3	Results on Face Databases . . . . .	98
5.3.4	Results on Human Action Databases . . . . .	101
5.3.5	Results on VOC 2007 Database . . . . .	102
5.3.6	Results on Columbia Consumer Video Database . . . . .	104
5.3.7	Comparing RADM with and without Marginal Distribution Constraint . . . . .	104
5.3.8	Fusion with SSTNTL . . . . .	106
5.4	Summary . . . . .	109
<b>Chapter 6 Conclusions and Future Works</b>		<b>110</b>
<b>Appendices</b>		<b>113</b>
<b>Bibliography</b>		<b>118</b>
<b>Curriculum Vitae</b>		<b>131</b>