香港浸會大學
HONG KONG BAPTIST UNIVERSITY

HKBU Scholars

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

Download date: 21 Jun, 2024

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