Journal of Fiber Bioengineering and Informatics 8:1 (2015) 117–124
doi:10.3993/jfbi03201511

Unsupervised Spectral Regression Learning for Pyramid HOG

Qiang Li, Zhongli Peng, Xiaomei Lin*

Changchun University of Technology, Changchun, Jilin 130012, China

Abstract

Applying the original raw data to machine learning will bring in a poor performance, because so many features are not necessary and redundant. Extracting a small number of good features will be an important issue, and it can be solved by using dimensionality reduction techniques. However, the popular dimensionality reduction method will suffer from the eigen-decomposition of dense matrix problem which is expensive in memory and time. We adopt unsupervised (unlabeled) spectral regression method for dimensionality reduction, which well avoids the problem of dense matrix eigen-decomposition problem and can be applied on large scale data sets. Histograms of Oriented Gradients (HOG) are robust features which not only well characterize the local shape and appearance but also show a certain degree of local optical and geometry invariance. In order to characterize the local shape and appearance better, we extract a three-tier pyramid HOG descriptor vector for one sample. Then we adopt the unsupervised spectral regression method for dimensionality reduction on these descriptor vectors. Our algorithm can be applied in the library entrance guard system of university and other research fields. Several experiments on well-known face databases have shown good performance and good invariance against illumination, occlusion and local deformation, etc.

Keywords: Dimensionality Reduction; Eigen-decomposition of Dense Matrix; Three-tier Pyramid HOG

1 Introduction

The original raw date in reality is often high-dimensional, which is expensive in memory and time without performing the dimensionality reduction operation. Spectral methods, a powerful tool for dimensionality reduction, aim to acquire the information resided in the eigenvectors of a neighborhood graph matrix of data sets to reveal the low dimensional structure in high dimensional data. However, the problem of eigen-decomposition of dense matrices will be brought in, which is expensive in memory and time. The unsupervised (unlabeled) spectral regression method well avoids this problem and will be well applied on large scale data sets [1].

How to extract robust features is an important issue in the field of machine learning, computer vision, pattern recognition, etc. The Histograms of Oriented Gradients (HOG) performs

*Corresponding author.
Email address: 27453582@qq.com (Xiaomei Lin).

quite well in characterising the local shape and appearance of object by capturing the gradient structure. What's more, HOG will show a certain degree of invariance against the local optical and geometrical deformation [2]. HOG will perform a local normalization operation on the block. The final result will be affected by the size of cells. So, in order to characterize the local shape and appearance better, we construct a three-tier pyramid for each sample, and extract the three-tier pyramid HOG features for each sample. Finally, we will carry out an unsupervised spectral regression learning for the three-tier pyramid HOG.

The rest of this paper is organized as follows: In Section 2, we introduce the unsupervised Spectral Regression Learning. Pyramid Histograms of Oriented Gradients (HOG) is reviewed in Section 3. In Section 4, we propose our Unsupervised Spectral Regression Learning for Pyramid HOG. Then, experiments and results is presented in Section 5, and followed by the conclusion in Section 6.

2 Unsupervised Spectral Regression Learning

The spectral embedding algorithms aim to seek a linear function which minimizes the object function for spectral embedding learning. These algorithms will suffer from the high computational time cost and expensive memory requirements, when it comes to the eigen-decomposition of dense matrices. So, these algorithms won't be suited to the large scale data sets. And, the data in reality is always unlabeled. We adopt the unsupervised spectral regression for dimensionality reduction, which avoids the eigen-decomposition of dense matrices. The rough steps of the unsupervised spectral regression learning are as follows [1].

1. Solve the generalized eigenvalue problem of the following function to get y, the eigenvectors in accordance with the maximum eigenvalues.

$$\hat{y} = \underset{y_T D y = 1}{argmax}\, y_T S y = argmax \frac{y_T S y}{y_T D y} \Rightarrow S y = \lambda D y \tag{1}$$

2. Use the technique to solve least square problem [3], such as LSQR to find the projection matrix a which satisfies $X^T w = y$ [4], and more importantly these efficient iterative algorithms can well handle large scale least square problems.

$$w = \underset{w}{argmax} \left(\sum_{i=1}^{N} (w^T x_i - y_i)^2 + \alpha \|w\|^2 \right) \tag{2}$$

$\alpha\|w\|^2$ the a penalty term, which can solve the ill posed minimization problem of the least square when the number of samples is smaller than the number of features. This is the regularized least square called ridge regression [5].

3 Pyramid Histograms of Oriented Gradients (HOG)

Histograms of Oriented Gradients (HOG) can well characterize the local shape and appearance of the objects. What's more, they show a certain degree of local optical and geometry invariance. But, we know that the size of the "cell" impacts an important role in fully playing the performance