

Image Classification Using Group Sparse Multiview Patch Alignment Framework Method

Ashok Kakad¹, Pandhrinath Ghonge²

¹Department of Computer Engineering, Yadavrao Tasgaonkar Institute of Engineering and Technology, Karjat, India

²Department of Electronics Engineering, Yadavrao Tasgaonkar Institute of Engineering and Technology, Karjat, India

Abstract: We cannot classify the images using Single feature. Multiview learning aims to unify different kinds of features to produce an efficient representation. This technique redefines part optimization in the patch alignment framework (PAF) and develops a group sparse multiview patch alignment framework (GSM-PAF). The new part optimization considers not only the complementary properties of different views, but also views consistency. In particular, view consistency models the correlations between all possible combinations of any two kinds of view. In contrast to conventional dimensionality reduction algorithms that perform feature extraction and feature selection independently, GSM-PAF enjoys joint feature extraction and feature selection which leads to the simultaneous selection of relevant features and learning transformation, and thus makes the algorithm more discriminative.

Keywords: GSM-PAF, Multiview learning, feature selection and feature extraction

1. Introduction

We are never to be classifying the images from image data set by using single feature. In this project, we extract various kinds of features for each image, and then generate different views based on the features. These views are regarded as multiple views (Multiview) of an image. Each view is assumed to have a particular physical meaning and statistical property.

Recently, a large number of methods of learning from multiview data (Multiview Learning) by considering the diversity of different views have been proposed. These views may be obtained from different visual views, multiple sources or different subsets. In this project, we focus on image classification and each image has multiple features. In the high-dimensional space of each view's representation, it is hard to distinguish images of different classes (e.g. features that do not carry discriminative information weaken the capability of a trained model to separate samples from different classes) and the trained model has the so-called "curse of dimensionality" problem (e.g. all training examples tend to pile up at the boundaries and become support vectors in support vector machine training, ending up with poor generalization).

Dimensionality reduction based on spectral analysis is the process of transform measurements from a high-dimensional space to a low-dimensional subspace through the spectral analysis on specially constructed matrices. It aims to reveal the intrinsic structure of the distribution of measurements in the original high-dimensional space and plays an important role in data mining, computer vision, and machine learning to deal with "curse of dimensionality" for various applications, e.g., biometrics, multimedia information retrieval, document clustering, and data visualization. Representative spectral analysis based dimensionality reduction algorithms can be classified into two groups: i) conventional linear dimensionality reduction algorithms and ii) manifold learning based algorithms.

Many problems in statistical pattern recognition begin with the pre-processing of multidimensional signals, such as images of faces or spectrograms of speech. Often, the goal of pre-processing is some form of dimensionality reduction: to compress the signals in size and to discover compact representations of their variability. Two popular forms of dimensionality reduction are the methods of principal component analysis (PCA) and multidimensional scaling (MDS). Both PCA and MDS are eigenvector methods designed to model linear variability's in high dimensional data. In PCA, one computes the linear projections of greatest variance from the top eigenvectors of the data covariance matrix. In classical (or metric) MDS, one computes the low dimensional embedding that best preserves pairwise distances between data points. If these distances correspond to Euclidean distances, the results of metric MDS are equivalent to PCA. Both methods are simple to implement, and their optimizations do not involve local minima. These virtues account for the widespread use of PCA and MDS, despite their inherent limitations as linear methods.

Recently, we introduced an eigenvector method—called locally linear embedding (LLE)—for the problem of nonlinear dimensionality reduction. This problem is illustrated by the nonlinear manifold in Figure 1. In this example, the dimensionality reduction by LLE succeeds in identifying the underlying structure of the manifold, while projections of the data by PCA or metric MDS map faraway data points to nearby points in the plane. Like PCA and MDS, our algorithm is simple to implement, and its optimizations do not involve local minima. At the same time, however, it is capable of generating highly nonlinear embedding's. Note that mixture models for local dimensionality reduction, which cluster the data and perform PCA within each cluster, do not address the problem considered here namely, how to map high dimensional data into a single global coordinate system of lower dimensionality.

Volume 3 Issue 12, December 2014

www.ijsr.net

Licensed Under Creative Commons Attribution CC BY

2. Literature Review

Recently, a large number of methods of learning from multiview data (multiview learning) by considering the diversity of different views have been proposed. These views may be obtained from different visual views, multiple sources or different subsets. For example, any object can be captured from a range of visual views Z.-G. Fan [2], C. M. Christoudias [3] e.g., (frontal and profile views of faces); a person can be identified by face, palmprint or iris with information obtained from multiple sources; or an image can be represented by its color or shape features, which can be treated as multiple features of the image. Multiple view learning was introduced by Blum [15] in semi-supervised learning. They proposed a co-training algorithm to use both labeled and unlabeled examples to train a classifier from two representations. The co-training algorithm trains one classifier on each view of the labelled examples and then iteratively allows each classifier to label the unlabeled examples it predicts with the highest confidence. Given independence between the classifiers, newly labelled examples from one classifier may give the other classifier new information to improve the model.

Much effort has been expended on multiview learning C. Xu [9], such as dimensionality reduction (feature extraction and feature selection, V. Bolón-Canedo [10]), classification, T. T. Nguyen [11] and clustering. Feature extraction algorithms, such as manifold learning, Steven P. Brumby [13] and subspace learning T. Li [5] learn to obtain low-dimensional representations of the high-dimensional examples. Most of the existing multiview feature extraction algorithms share at least one of the following two problems: the out-of-sample problem (because they cannot extract the feature representation for test images directly but have to re-compute the embedding), and the over-fitting problem (because they linearly encode all the features without discarding redundant features), which are inappropriate for practical applications.

A Patch alignment framework termed “patch alignment” to unify spectral analysis based dimensionality reduction algorithms. This framework consists of two stages: part optimization and whole alignment. For part optimization, different algorithms have different optimization criteria over patches, each of which is built by one measurement associated with its related ones. For whole alignment, all part optimizations are integrated to form the final global coordinate for all independent patches based on the alignment trick. This framework discovers that: i) algorithms are intrinsically different in the patch optimization stage; and ii) all algorithms share an almost identical whole alignment stage. As an application of this framework, we also develop a new dimensionality reduction algorithm, termed Discriminative Locality Alignment (DLA), by imposing discriminative information in the part optimization stage. Benefits of DLA are threefold: i) because it takes into account the locality of measurements, it can deal with the nonlinearity of the measurement distribution; ii) because the neighbour measurements of different classes are considered, it well preserves discriminability of classes; and iii) because

it obviates the need to compute the inverse of a matrix, it avoids the small sample size problems.

3. Problem Definition

Most of the existing multiview feature extraction algorithms share at least one of the following two problems:

a) Out-of-sample problem

The out-of-sample problem (because they cannot extract the feature representation for test images directly but have to re-compute the embedding), and

b) Over-fitting problem

The over-fitting problem (because they linearly encode all the features without discarding redundant features), which are inappropriate for practical applications.

To alleviate the problem illustrated above, we use a new linear multiview feature extraction method based on PAF (patch alignment framework).

4. Proposed Work

GSM-PAF is extraction method based on PAF. PAF is proposed as a framework for dimensionality reduction. PAF unifies popular dimensionality reduction algorithms, e.g., PCA (Principal Component Analysis). GSM-PAF is a flexible and can be either un-supervised or supervised. GSM-PAF (Group sparse multi-view patch alignment framework) is a new linear multiview feature extraction method based on PAF (patch alignment framework). PAF is proposed as a framework for dimensionality reduction. PAF unifies popular dimensionality reduction algorithms, e.g., PCA (principal component analysis) .GSM -PAF is a flexible and can be either un-supervised or supervised. In a GSM-PAF a framework of joint feature extraction and feature selection for multiview learning. GSM-PAF is a further yet comprehensive development of patch alignment framework which unifies many dimensionality reduction algorithms. This framework consists of two stages: part optimization and whole alignment. For part optimization, different algorithms have different optimization criteria over patches, each of which is built by one example associated with its related ones, for e.g. PCA. For whole alignment, all part optimizations are integrated to form the final global coordinate for all independent patches based on the alignment trick.

GSM-PAF first builds a patch for a sample of a view. Based on the patches from different views, part optimization can be performed to obtain the optimal low-dimensional representation for each view. All low-dimensional representations from different patches are subsequently unified as a whole by global coordinate alignment. Joint feature extraction and feature selection are then performed based on $l_{2,1}$ -norm. Finally, the solution of GSM-PAF is derived by using the alternating optimization. The GSM-PAF shown in below figure

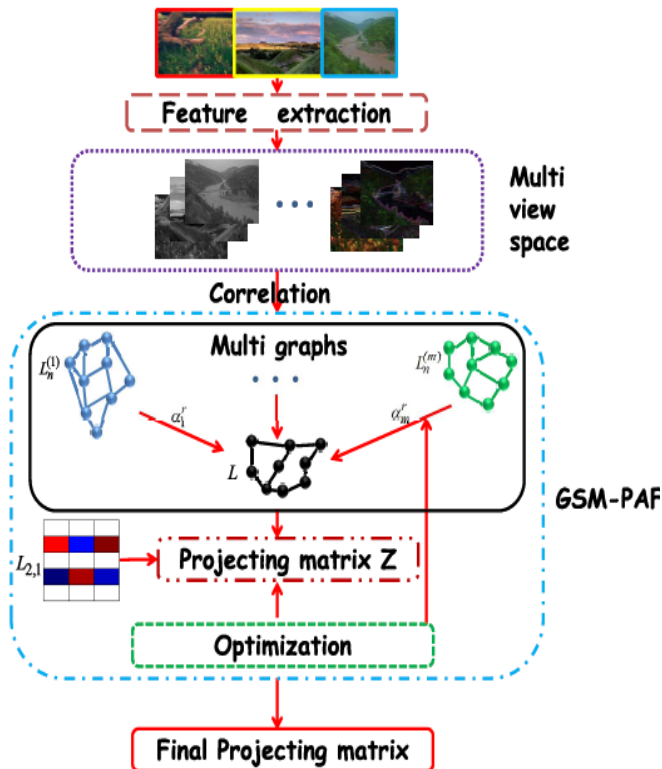


Figure 2: System Diagram for Image Classification

Let X represent data sets and x represent data vectors. Based on these notations, GSM-PAF can be described as follows, according to our previous patch alignment framework [38], [39]. Given a multiview feature data set X with n examples and m views (each “view” has a particular meaning and statistical property, e.g. texture, color, and shape), i.e.,

$$X = \left[\left(X^{(1)} \right)^T \dots \left(X^{(m)} \right)^T \right]^T,$$

$$A = \sum_{j=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{h=1}^m \left(\text{tr} \left(S_j V(S_l)^T \right) - \varepsilon \left(\text{tr} \left(S_j V(S_j)^T \right) + \text{tr} \left(S_l V(S_l)^T \right) \right) \right)$$

Where in $X(i) \in R^{D_i \times n}$ is the matrix of D_i –dimension feature vectors for the i th view representation and $X \in R^{D \times n}$ $D = \sum_{i=1}^m D_i$ contains all examples represented by all views. GSM-PAF aims to find a linear transformation matrix U which can project the high dimensional data X into a low dimensional embedding $Y \in R^{d \times n}$ ($d < D$), i.e., $Y = U^T X$, meanwhile retaining the neighborhood structures and the correlation between any two views in the multiple feature spaces. That is to say, all examples of all classes are mapped by a common matrix U .

GSM-PAG mainly consists of Part optimization and Whole alignment they are as follows:

A) Part optimization

For part optimization, different algorithms have different optimization criteria over patches, each of which is built by

one example associated with its related ones. For instance, the part optimization of PCA is

$$\frac{1}{n^2} \begin{bmatrix} -(n-1)^2 & -(n-1)e_{n-1}^T \\ -(n-1)e_{n-1} & e_{n-1}e_{n-1}^T \end{bmatrix},$$

where n is the number of examples and $e_{n-1} = [1, \dots, 1]^T \in R^{n-1}$.

B) Whole alignment

For whole alignment, all part optimizations are integrated to form the final global coordinate for all independent patches based on the alignment trick. As defined, Y is used to denote the low dimensional embedding of X . By summing up all the samples based on equation, the whole alignment is obtained as :

$$\begin{aligned} \arg \min_{Y, \alpha} & \left(\sum_{j=1}^n \sum_{i=1}^m \alpha_i \text{tr} \left(Y_j^{(i)} L_j^{(i)} \left(Y_j^{(i)} \right)^T \right) \right) - \beta \text{tr} \left(Y A Y^T \right) \\ & = \arg \min_{Y, \alpha} \left(\sum_{i=1}^m \alpha_i \text{tr} \left(Y L_n^{(i)} Y^T \right) \right) - \beta \text{tr} \left(Y A Y^T \right) \end{aligned}$$

where $L(i)$ n is a normalized graph Laplacian matrix by performing a normalization on Laplacian matrix $L(i)$. The Laplacian matrix is defined as $L(i) = D(i) - W(i)$, where $W(i) \in R^{n \times n}$ is the similarity matrix of the i th view and has non-zero weights on x 's k -nearest neighbors, zero on others. $D(i)$ is a diagonal matrix and its entries are the column sum of W . The matrix A in equation (8) is defined as follows:

where S_j is the selection matrix. The optimization of equation (8) cannot apply to test samples directly because of the out-of-sample problem. In this subsection, we obtain a linear transformation from equation which can be applied to the test samples directly.

In the proposed work, we propose GSM-PAF for better classification of Image. We first project the Multi view data into a low dimensional space and then perform classification using conventional classifier.

Here, first we take the Image Dataset as an Input then by using GSM-PAF algorithm we have to extract feature of each Image from Image Dataset. Then generate the Graph depend on the features of the image. After that for better classification of Image, we use the KNN algorithm or SVM algorithm.

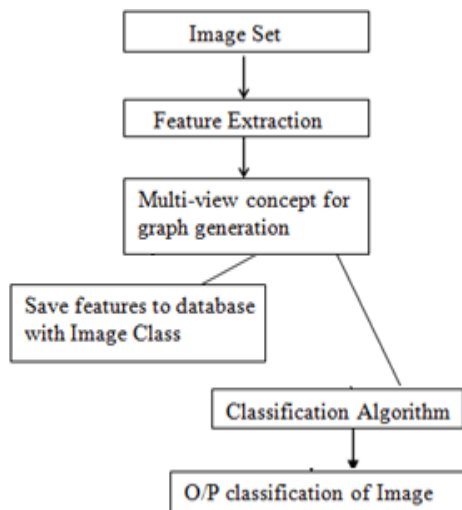


Figure 3: Flow graph for Image Classification using GSM-PAF

5. Advantages and Disadvantages

Advantages

Image classification is one of the important and complex processes in Image processing. Classification is an information processing task in which images are categorized into multiple groups. In supervised classification trained database is needed and also required human annotation. In supervised classification the operator can detect error and remedy them. In unsupervised classification human annotation is not needed and it is more computers automated. In unsupervised classification time taken is less and minimizes human errors.

Disadvantages

The image classification is more complex and difficult to classify if it contain blurry and noisy content. In supervised classification training data can be time consuming and costly and it is prone to human error. In unsupervised classification method are maximally-separable cluster in spectral space may not match our perception of the important classes on the landscape and also limited control over the menu of classes.

6. Conclusion

Using GSM-APF algorithm, we proposed a framework of joint feature extraction and feature selection for multiview learning. To speed up the process of image classification we use GSM-PAF algorithm. The contributions of this paper are twofold. First, we consider not only the independent information of each view and the complementary properties of different views, but also view consistency in linear multiview feature extraction. Second, we simultaneously perform feature extraction and feature selection for multiview learning based on the $l_2, 1$ -norm of the projection matrix. We have to categories the online data set in multiple groups.

References

- [1] Jie Gui, Dacheng Tao, "Sparse Multiview Patch Alignment Framework With View Consistency for Image Classification" IEEE transactions on image processing, vol. 23, no. 7, July 2014
- [2] Z.-G. Fan and B.-L. Lu, "Fast recognition of multi-view faces with feature selection," 10th IEEE ICCV, Oct. 2005
- [3] C. M. Christoudias, R. Urtasun, and T. Darrell, "Unsupervised feature selection via distributed coding for multi-view object recognition," in Proc. CCVPR, Jun. 2008,
- [4] T. Zhang, D. Tao, X. Li, and J. Yang, "Patch alignment for dimensionality reduction," IEEE Trans. Knowledge Data, vol. 21, Sep. 2009
- [5] Z.-Y. Zhang, T. Li, and C. Ding, "Non-negative tri-factor tensor decomposition with applications," Knowledge Information System., vol. 34, no. 2, pp. 243–265, 2013.
- [6] E. Kokiopoulou, D. Kressner, and P. Frossard, "Optimal image alignment with random projections of manifolds: Algorithm and geometric analysis," IEEE Transaction Image Process., vol. 20, no. 6, pp. 1543–1557, Jun. 2011.
- [7] C. Xu, D. Tao, and C. Xu, "A survey on multi-view learning," ArXiv: 5634, 2013.
- [8] V. Bolón-Canedo, N. Sánchez-Marño, and A. Alonso-Betanzos, "A review of feature selection methods on synthetic data," Knowledge Information System., vol. 34, 2013.
- [9] T. T. Nguyen, K. Chang, and S. C. Hui, "Supervised term weighting centroid-based classifiers for text categorization," Knowledge Information System, vol. 35, no. 1, 2013.
- [10] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," Science, vol. 290, no. 5500, pp. 2323–2326, 2000
- [11] X. He and P. Niyogi, "Locality preserving projections," in Proceeding Neural Information Processing System, 2004.
- [12] Blum and T. Mitchell, "Combining labeled and unlabeled data with co training," in Conference Computing Learn Theory, Jul. 1998,