Engineering Applications of Artificial Intelligence 29 (2014) 201–210

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Segmentation of endocardium in ultrasound images based on sparse representation over learned redundant dictionaries

Roberto Rosas-Romero [a,*], Hemant D. Tagare [b]

[a] Universidad de las Américas-Puebla, Puebla 72810, Mexico
[b] Yale University, New Haven, CT, USA

ARTICLE INFO

Article history:
Received 28 May 2013
Received in revised form
2 August 2013
Accepted 2 September 2013
Available online 18 October 2013

Keywords:
Echocardiographic image segmentation
Sparse representation
Dictionary learning

ABSTRACT

This paper considers the problem of segmenting the *endocardium* in 2-D short-axis echocardiographic images from rats by using the sparse representation of feature vectors over learned dictionaries during classification. We highlight important aspects of the application of the theory of sparse representation and dictionary learning to the problem of ultrasound image segmentation. Experiments were conducted following two directions for the generation of dictionaries for myocardium and blood pool regions; by manual extraction of image patches to build untrained dictionaries and by patch extraction followed by training of dictionaries. The results obtained from different learned dictionaries are compared. During classification of an image patch, instead of using features of the patch alone, features of neighboring patches are combined.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Advancements in ultrasound transducer design, improvements in image resolution, noninvasive nature, usefulness for medical diagnosis, portability as well as economy make segmentation of ultrasound images a motivating challenge; however, ultrasound data presents characteristics (such as speckle, non-homogeneities, shadowing, low contrast, and signal dropout) that make its segmentation a difficult task. Noble and Boukerroui (2006) provide an extensive review of different ultrasound segmentation methods.

We consider the problem of segmenting the endocardium on short-axis echocardiographic images by dividing an image in *patches* (Kumar and Hebert, 2006), and classifying an image patch as belonging to one of two classes or regions, *myocardium tissue* or *blood pool.* This problem is approached using the promising research field of *dictionary learning*, early addressed by Olshausen and Field (1997), with the motivation that there are no reports of echocardiographic image segmentation using the framework of dictionary learning. This research field focuses on the development of algorithms to learn dictionaries with elements, called *atoms*, so that a signal of interest can be decomposed as a linear combination of a few atoms. The sparse representation of a signal possesses an implicit *discriminative* nature by choosing the subset of atoms that give the best compact reconstruction of the signal, and discarding other representations.

Modeling of signals by means of their sparse representation is a natural and fundamental concept so that it becomes a very useful tool for different applications. The solution to the problem of dictionary learning for sparse representation of images has proven to be successful in many other applications such as image denoising (Aharon et al., 2006; Elad and Aharon, 2006), compression (Marcellin et al., 2000), super-resolution (Yang et al., 2010), handwritten digit classification (Mairal et al., 2008c, 2012), face recognition (Wright et al., 2009), texture segmentation and classification (Mairal et al., 2008a; Wright et al., 2010), object detection (Agarwal and Roth, 2002), color restoration (Mairal et al., 2008b).

There are different research directions that could be taken to push the frontiers of knowledge in this research field (Elad, 2012). An opposite alternative to the previously mentioned synthesis-based sparse representation model for signals is the analysis-based model where an analysis dictionary is learned so that this dictionary multiplies a signal to provide the corresponding sparse code (Rubinstein and Faktor Elad, 2012). Scale-Up of an image from a down-scaled noisy version with preservation of edges and small details has been accomplished by using sparse representation models and regularization (Zeyde et al., 2012). It has been shown that dictionary learning outperforms off-the-shelf fixed dictionaries for the case of denoising of astronomical images (Beckouche et al., 2013). An image has been separated into its texture and piecewise smooth components by modeling these components as sparse combination of atoms from dictionaries (Starck et al., 2005). Algorithms for multi-scale dictionary learning combine characteristics from multi-scale representation models (wavelets) and single-scale dictionaries to sparsely represent

* Corresponding author. Tel.: +52 222 229 26 77.
E-mail addresses: Roberto.rosas@udlap.mx (R. Rosas-Romero), hemant.tagare@yale.edu (H.D. Tagare).

0952-1976/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.engappai.2013.09.008

signals (Ophir et al., 2011). Atoms of learned dictionaries have been modeled as sparse representations over fixed based dictionaries (double sparsity) (Rubinstein et al., 2010). The *Minimum Length Description Principle* is incorporated to the dictionary learning framework in Ramírez and Sapiro (2012) to optimize not only the structure of atoms but also the number of atoms and to automatically tune the sparsity constraint.

Most computer vision problems, addressed with the framework of *sparse representation*, only use gray level information as features; however, spatial information, statistical parameters and other structures, such as contours, should be added to gray level information as entries in *feature vectors* (this has been done with penalized logistic models (Yue and Tagare, 2009)) to make dictionaries more suitable to handle nonhomogeneous spatial artifacts, such as contrast and noise, which are characteristic in ultrasound images. For the case of echocardiographic images, the contrast across endocardium as well as the brightness of myocardium vary spatially due to the interaction between the propagation of ultrasound waves and the geometry of myocardium tissue (Yue and Tagare, 2009; Holland et al., 1999).

There are three categories of dictionary learning algorithms. In this application, the algorithms follow the direction of the *clustering methods* (Tosic and Frossard, 2011). We use dictionaries which are constructed in two ways, by (1) manual extraction of image patches to build untrained dictionaries or (2) by patch extraction followed by dictionary training which consists in minimizing an energy function through optimization methods like *MOD* (Egan et al., 2000) or *K-SVD* (Aharon et al., 2006). In Wright et al. (2009), the application of untrained dictionaries in classification tasks has been successful for face recognition (a method called *Sparse Representation Classification—SRC*) where (1) the whole image under low resolution is used as an atom, (2) multiple atoms correspond to different instances of the same face under different conditions of illumination and/or occlusion, and (3) one single dictionary is built by concatenating face dictionaries from different subjects. This method is useless for our application since this method was designed to work on atoms which are generative models of entire images and not for the modeling of isolated image patches.

During estimation of the sparse representation of a signal, there is a search for a linear combination of atoms to approximate the signal of interest. One way of estimating the sparse representation of a signal consists in using *greedy algorithms* like *Matching Pursuit* (*MP*) (Mallat and Zhang, 1993) and *Orthogonal Matching Pursuit* (*OMP*) (Pati et al., 1993) which consist in minimizing the l^0 norm of the sparse code of the signal under the constraint of an undetermined system of linear equations. Motivated by the fact that atoms from untrained dictionaries do not evolve, we propose the straight use of these manually extracted atoms to reconstruct a signal by (1) looking for the nearest atoms to the signal and then (2) using the sub-set of the nearest atoms to estimate the best linear approximation to the signal. We called this approach *Matching Pursuit over the L Nearest Atoms* (*MPLNA*). Since atoms are l^2 normalized and might be extracted from overlapping image patches, image segmentation based on sparse representation can handle patch classification without previous registration.

During image segmentation, patch classification is accomplished by extraction of a feature vector from the patch, followed by a transformation of the feature vector into a residual feature vector, and finally a label is assigned to the patch by searching for the smallest residual feature inside the transformed vector (Aharon et al., 2006; Mairal et al., 2008c, 2012, 2008a; Wright et al., 2009, 2010; Rodríguez and Sapiro, 2008). Residual features are residuals of the approximation of a feature vector over different classes. Instead of looking for the smallest residual to classify a patch, we propose to linearly combine residuals from multiple patches in a neighborhood and use these combinations as features for classification. This approach accounts for filtering and for spatial interaction of neighboring patches allowing the reduction of isolated misclassified patches or clusters of misclassified patches.

In Section 2, an overview of the framework on sparse representation and dictionary learning is presented, highlighting the theory suitable to the current application. Section 3 discusses how sparse representation is applied to the task of segmentation of the endocardium in ultrasound images as well as contributions in our application. Section 4 provides the experimental results obtained by following different strategies of dictionary learning and matching pursuit as well as comparisons. The conclusions are presented in Section 5.

2. An overview on sparse representation and dictionary learning

Sparse representation of signals has received considerable attention as a tool to solve different problems in Computer Vision. An extensive survey of the challenges, motivation, approaches and applications of the main algorithms in the field of dictionary learning for sparse representation is presented by Tosic and Frossard (2011) and Elad (2010).

A *dictionary* is a collection of k elements stacked as column vectors in a matrix $\boldsymbol{D}=[\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_k] \in \boldsymbol{R}^{n \times k}$. Each element, called *atom*, is an n-dimensional vector $\boldsymbol{d}_i \in \boldsymbol{R}^n$ in the *hyper-sphere* $\|\boldsymbol{d}_i\|_2 = 1$. The set of atoms $\{\boldsymbol{d}_i\}_{i=1}^k$ serves as an over-complete basis ($k > n$) for reconstruction of a signal $\boldsymbol{x} \in \boldsymbol{R}^n$.

2.1. Sparse coding

Given a *dictionary* $\boldsymbol{D} \in R^{n \times k}$, a signal $\boldsymbol{x} \in \boldsymbol{R}^n$ can be approximated by a linear combination of its atoms, $x = \sum_{i=1}^{k} \alpha_i \boldsymbol{d}_i = \boldsymbol{D}\boldsymbol{\alpha}$, where $\boldsymbol{\alpha} \in \boldsymbol{R}^k$ is the *sparse code* of the signal. Since $\boldsymbol{D}$ is over-complete ($k \gg n$), there are many solutions to this undetermined system of equations. Thus, a constraint is imposed so that the objective is to find a *sparse code* $\boldsymbol{\alpha}$ with the smallest number of non-zero coefficients

$$\boldsymbol{\alpha} = \arg\min_{\alpha} \|\boldsymbol{\alpha}\|_0 \ni \boldsymbol{x} = \boldsymbol{D}\boldsymbol{\alpha} \tag{1}$$

By allowing certain degree of noise, a bounded error ε is imposed in the reconstruction of $\boldsymbol{x}$,

$$\boldsymbol{\alpha} = arg\min_{\alpha} \|\boldsymbol{\alpha}\|_0 \ni \|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{\alpha}\|_2^2 < \varepsilon \tag{2}$$

where $\|\boldsymbol{\alpha}\|_0$ is the l^0 *norm* that counts the number of nonzero components in $\boldsymbol{\alpha}$. The problem posed in (1) and (2) can be solved with *greedy algorithms* like *Matching Pursuit* (*MP*) (Mallat and Zhang, 1993) and *Orthogonal Matching Pursuit* (*OMP*) (Pati et al., 1993). In the sparse coding literature, the l^0 *norm* has been replaced by the l^1 *norm* and the *sparse coding* becomes a *convex problem* which is solved by different algorithms (Hale et al., 2007; Malioutov Willsky; Gorodnitsky and Rao, 1997).

2.2. Reconstructive dictionary learning

The goal of *dictionary learning* is to compute the *reconstructive dictionary* $\boldsymbol{D}=[\boldsymbol{d}_1\, \boldsymbol{d}_2 \ldots \boldsymbol{d}_k] \in \boldsymbol{R}^{n \times k}$ that provides the optimum sparse reconstruction for a given set of m training signals, $\{\boldsymbol{x}_i\}_{i=1}^m \in \boldsymbol{R}^n$, called *observations* or *training samples*, stacked as column vectors in a matrix $\boldsymbol{X}=[\boldsymbol{x}_1\, \boldsymbol{x}_2 \ldots \boldsymbol{x}_m] \in \boldsymbol{R}^{n \times m}$. During *dictionary learning*, the set of dictionary atoms $\{\boldsymbol{d}_i\}_{i=1}^k \in \boldsymbol{R}^n$ and the set of sparse codes $\{\boldsymbol{\alpha}_i\}_{i=1}^m \in \boldsymbol{R}^k$, stacked as column vectors in a matrix $\boldsymbol{A}=[\alpha_1\, \boldsymbol{\alpha}_2 \ldots \boldsymbol{\alpha}_m] \in \boldsymbol{R}^{k \times m}$, are simultaneously estimated

Download English Version:

https://daneshyari.com/en/article/380664

Download Persian Version:

https://daneshyari.com/article/380664

Daneshyari.com