Information Sciences 339 (2016) 175–188

Contents lists available at ScienceDirect

Information Sciences

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

Fuzzy density weight-based support vector regression for image denoising

Yun Zhang[a], Shuqiong Xu[a], Kairui Chen[a], Zhi Liu[a,*], C.L. Philip Chen[b]

[a] Faculty of Automation, Guangdong University of Technology, Guangzhou, Guangdong, China
[b] Faculty of Science and Technology, University of Macau, Macau, China

ARTICLE INFO

Article history:
Received 12 April 2014
Revised 18 November 2015
Accepted 1 January 2016
Available online 7 January 2016

Keywords:
Fuzzy density weight
Least squares support vector regression
Image denoising

ABSTRACT

Support vector machine (SVM) is a popular machine learning technique and its variant least squares support vector regression (LS-SVR) is effective for image denoising. However, conventional LS-SVR does not fully consider the sampling distribution of noisy images, which may degrade the performance of the algorithm. In this paper, we propose a new fuzzy density weight SVR (FDW-SVR) denoising algorithm, which assigns fuzzy priority to each sample according to its density weight. FDW is designed to estimate the joint probability density function via the fuzzy theory based on the pixel density and neighborhood density. Extensive experimental results show that FDW-SVR is superior to those state-of-the-art denoising techniques in light of both subjective and objective evaluations.

1. Introduction

Image denoising, as the image preprocessing, is an important task in image processing. In the past three decades, a heap of denoising methods have been developed by researchers [3,6,7,9,16,25,26,28,29,50,53,59]. The simplest method for noise removal is Gaussian filtering, which is to solve an isotropic heat diffusion equation, a second-order linear partial differential equation. Compared to Gaussian filtering, anisotropic diffusion [9,18,26,37,52] smoothes out noise without losing edges. However, it tends to over blur the image and sharpen the boundary, which cause the loss of many details. Besides, the wavelet-based methods [6,16,17,20,23,24,38,59] are popular and dominant in denoising, but hard to remove the ringing artifacts of wavelet reconstruction. In other words, wavelet-based methods tend to bring additional edges or structures to the image. Modeling methods such as Bayesian inference [3,22,27] and Gaussian Conditional Markov Random Field [14,34,36,46,48] can be applied to image denoising via learning parameters of the model from noisy images. However, these methods bring in the same drawbacks as most denoising algorithms do: over smoothing and edge sharpening.

SVR is a generalization of Vapnik's SVM method to estimate real-value functions based on the principle of structural risk minimization [8,45]. In the past decade, SVR has been extensively developed to deal with various kinds of problems, such as pattern recognition [4,13,19,39,44,56,60], medical diagnosis [40], image processing [11,15,31, 47,51,55], forecasting problems [10], etc. Li [28, 29] performed the SVR for image denoising under the assumption that different images satisfied local similarity condition [2,7]. In [33], Liu proposed a denoising algorithm based on wavelet transformation and LS-SVR. Based on LS-SVR, denoising operators used in wavelet domain were obtained and employed to image denoising according to the principle of wavelet denoising. The author in [12] proved that the LS-SVR with translation invariant kernel is linear

* Corresponding author. Tel.: +86 15918713538.
E-mail address: 9075846@qq.com, lz@gdut.edu.cn (Z. Liu).

http://dx.doi.org/10.1016/j.ins.2016.01.007
0020-0255/

time-invariant and found that common radial basis function kernel-based LS-SVR is actually low-pass and linear phase filter for image denoising. However, all the above-mentioned LS-SVR does not fully consider the sampling distribution of noisy images, which motives us to improve the LS-SVR algorithm from this perspective.

Fuzzy sets and systems are powerful tools to improve the robust performance of image processing [1,5,21,25,30,32,58] and provide solutions to a broad range of problems, such as intelligent control, pattern classification, reasoning, planning and computer vision. In [42], Pedrycz provided a superb introduction for fuzzy-neural networks and fuzzy-genetic systems and bridged the gap that has developed between theory and practice. He also presented the unified principles of granular computing along with its comprehensive algorithmic framework and design practices [41]. Based on the fuzzy weighted method [57], Lin [32] provided a method to separate the image from noises. He introduced two factors: the confident factor and the trashy factor, to build a relationship between the probability density function and the heuristic function, and automatically generate memberships of training data points by a heuristic strategy. In [35], Liu proposed a novel three-domain fuzzy support vector regression (3DFSVR), where three-domain fuzzy kernel function provided a solution to process uncertainties and input–output data information simultaneously. 3DFSVR is another way to improve traditional 2DFSVR by adding the fuzzy domain. H. Huang [21] proposed a fuzzy SVM algorithm to deal with the outliers and noise problems. It should be noted that the distribution of density plays an extremely important role in the process of image denoising. Considering the input-output image data-set, it is difficult to obtain the joint probability density distribution function (PDF) due to the strong coupling of the input neighborhood and output pixel as well as the disturbance of noises.

To address this problem, FDW-SVR method is proposed to optimize the SVR by estimating joint PDF of noisy images. By shifting a $r \times r$ neighborhood window over the noisy image, we can obtain an input set U_i of the FDW-SVR. Then, different weights are assigned to the measures of error. The contributions of this work are summarized as follows:

(1) A fuzzy approach is presented to estimate the joint density distribution of image data-set. It is rather difficult to obtain the joint PDF by traditional methods, due to the probabilistic independence of input-output image data-set $\{U_i, y_i\}_{i=1}^N$.
(2) As a novel denoising algorithm, FWD-SVR is developed such that the priority is assigned to each sample according to its density weight. This improves the performance of LS-SVR as the sampling distribution of noisy images is taken into account.

In the next section, the theoretical basis of LS-SVR and the analysis to the influence of sampling distribution on the model are presented. In Section 3 , all the details of the proposed FDW-SVR denoising algorithm are given. In Section 4 , the experimental results are discussed to verify the performance of FDW-SVR. The conclusion of this work is given in Section 5 .

2. Least squares support vector regression

2.1. LS-SVR algorithm

The training set is defined as $\{U_i, y_i\}_{i=1}^N$, where N is the number of samples, $U_i \in \mathbb{R}^d$ is the ith input and $y_i \in \mathbb{R}$ is the ith output. $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$ represent the real $d \times 1$ vector and the real $d \times d$ matrix, respectively. The norm of a vector $U = [u_1, \ldots, u_d]^T \in \mathbb{R}^d$ is defined as $\|U\| = \sqrt{U^TU}$. Nonlinear mapping $\psi : \mathbb{R}^d \to \mathbb{R}^{dh}$ maps samples from input space $\mathbb{R}^d$ to high dimension feature space $\mathbb{R}^{dh}$. d is the dimension of input space and d_h the dimension of feature space. The LS-SVR is represented as a quadratic programming problem with equality constrains and the goal is to minimize

$$\min J(\omega, b, e_i) = \frac{1}{2}\|\omega\|^2 + \frac{1}{2}\gamma \sum_{i=1}^{N} e_i{}^2 \tag{1}$$

$$s.t.\, y_i = \omega^T \psi(U_i) + b + e_i, i = 1, \ldots, N \tag{2}$$

where J is the objective function, $\omega \in \mathbb{R}^{dh}$ is the weight vector, ω^T is the transpose of ω, $b \in \mathbb{R}$ is the model bias, γ is the regularization parameter and e_i is the fitting error. ω and b are the regression parameters to be optimized.

2.2. The influence of sampling distribution on LS-SVR

The input–output image data-set $\{U_i, y_i\}_{i=1}^N$ is composed of input neighborhood U_i and output pixel y_i. The input U_i denotes the vector whose elements are the features of pixels in the neighborhood of noisy image. y_i denotes the central pixel of the neighborhood in the original image [54]. The experiential risk is

$$\begin{aligned} Q_{emp} &= N \sum_{i=1}^{N} e_i^2 f(U_i, y_i) \partial_{U\times y} \\ &= N \partial_{U\times y} \sum_{i=1}^{N} f(U_i, y_i) e_i^2 \end{aligned} \tag{3}$$