Neurocomputing 154 (2015) 284–295

Contents lists available at ScienceDirect

Neurocomputing

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

Learning local Gaussian process regression for image super-resolution

Jianmin Li [a], Yanyun Qu [a,*], Cuihua Li [a], Yuan Xie [b], Yang Wu [c], Jianping Fan [d]

[a] Computer Science Department, Xiamen University, P.R.China
[b] The State Key Lab of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, P.R. China
[c] Academic Center for Computing and Media Studies, Kyoto University, Japan
[d] School of Information Science and Technology, Northwest University, P.R.China

ARTICLE INFO

Article history:
Received 16 June 2014
Received in revised form
24 November 2014
Accepted 29 November 2014
Available online 9 December 2014

Keywords:
Super resolution
Data-driven Gaussian process regression
Prototype-based Gaussian process regression
Covariance matrix
Low resolution
High resolution

ABSTRACT

Learning based super-resolution (SR) methods, which predict the high-resolution pixel values but not directly provide an estimation of uncertainty, are typically non-probabilistic and have limited generalization ability. Gaussian processes can provide a framework for deriving regression techniques with explicit uncertainty models, but Gaussian Process Regression (GPR) has a significant drawback in being time consuming. The computational complexity of GPR is cubic in the number of training examples, which is prohibitively expensive for a large-scale training set. In this article, we have proposed learning local GPR for image SR. Two algorithms are developed to support local GPR for super resolution. A data-driven GPR based super-resolution algorithm is first developed to learn a local GPR model for every LR patch on an input oriented training dataset with moderate size. In order to further improve the running speed, a prototype based GPR algorithm is developed for super resolution. The proposed algorithm is about one-order faster than the data-driven GPR solution because it makes models for the prototypes of image patches rather than for each image patch. Thus, the local regression efforts are greatly reduced to just finding the nearest prototype for each LR image patch and applying its corresponding pre-computed projective matrix for super-resolution prediction. Our algorithms have greater robustness and usability as they provide a formularized way to automatically learn the hyper-parameters introduced for optimizing the covariance function, while most of the state-of-the-art super-resolution methods could only utilize these parameters in a cross-verification way. Moreover, our algorithms offer confidence values at the test points which benefit the pixels' post-processing. Our algorithms are evaluated on popular datasets that are widely used in the super-resolution literature, and the experimental results have demonstrated that the efficiency and effectiveness of our proposed algorithms are comparative with several state-of-the-arts super-resolution methods.

1. Introduction

Image super-resolution (SR) is widely used in many practical applications such as satellite and medical imaging, where the analysis or diagnosis from low-resolution (LR) images can be difficult. Therefore, SR has become a hot topic in the field of computer vision. The goal of image SR is to generate a high-resolution (HR) image from one or multiple low-resolution (LR) input images. The SR problem is ill-posed because a low-resolution image can be generated by many different high-resolution images under different transformations. Up to now, a large number of learning based SR approaches [11–19] have been developed and they can yield promising results. However, the state-of-the-art learning based super-resolution methods are typically non-probabilistic and can predict the high-resolution pixel values but not directly provide an estimate of uncertainty. Thus, they are inadequate when the uncertainty is required. Moreover, the parameters used in these methods are specified a priori. In this article, we present an explicit probabilistic model for SR. We introduce Gaussian processes regression (GPR) for SR and our probabilistic formulation provides a principled way to learn hyper-parameters.

The computer vision community has paid little attention to Gaussian processes (GP) due to the fact that the Gaussian processes conventionally limit the amount of training data, because the computational complexity of GP is $O(N^3)$, which is cubic in the number of training examples and it is prohibitively expensive when the training dataset is large scale. In this article, we focus on how to make GPR work when the training dataset is large-scale. We propose an approach which learns local GPR models for the SR problem. We make a local GPR model for an LR image patch rather than make a global GPR model. It is worth noting that each query LR patch has its special local model. Two SR algorithms are developed to support the local GPR solution. In the first algorithm, a training

* Corresponding author.

http://dx.doi.org/10.1016/j.neucom.2014.11.064
0925-2312/

collection is searched for an LR image patch from the training dataset, and then a GPR model is made depending on the training collection. Because each image patch has its special GPR model, the local GPR model is data-driven, and we name it data-driven Gaussian Process Regression (DDGPR). In the second algorithm, we first find prototypes for image patches and make GPR models for the prototypes of the image patches. We name it prototype-based GPR (PGPR). DDGPR focuses on the practicability of GPR and the super-resolution accuracy. And PGPR focuses on low running time.

The three main contributions of this article are: 1) an explicit probabilistic model is made for super-resolution based on Gaussian Process regression which could not only predict the high-resolution values but also give their confidence values; 2) the data-driven GPR based SR is developed, which is a local model and achieves the promising SR performance; 3) the prototype-based GPR scheme is developed, which reduces the computational complexity and is one-order faster than DDGPR. We have shown that the proposed approach can achieve state-of-the-art super-resolution results on the benchmark image set with superior scaling ability.

This article extends our previous work [23] by proposing a novel algorithm named PGPR. PGPR is different from DDGPR which is proposed in [23] in the two major ways: 1) in DDGPR, the local GPR models are made for image patches, while in PGPR, the local GPR models are made for prototypes of image patches; 2) when an HR image is generated from its LR image, we need to build an image training dataset and learn a GPR model for each image patch in DDGPR, while in PGPR we just anchor an image patch to its nearest prototype and look up the corresponding prototype model which is pre-computed and stored as a reference model. The two differences lead to much fewer GPR models and significantly less computational time of the HR image reconstruction for PGPR than for DDGPR. Thus, PGPR is about one-order faster than DDGPR without scarifying much accuracy. However, DDGPR can achieve more accurate results than PGPR. Moreover, we also add additional empirical results and making an in-depth analysis of our approach's performance.

The remainder of this article is organized as follows. We introduce the related work in Section 2, and we give a brief overview of Gaussian Process regression in Section 3, and then we describe the local Gaussian Process regression algorithms for image super resolution in Section 4. The experimental results are shown in Section 5. In the last section, we give our conclusions.

2. Related work

Generally, the state-of-the-art SR techniques can be categorized into three classes: the interpolation based methods [1–6], the reconstruction based methods [7–10] and the learning based methods [11–19].

The interpolation-based methods are simple and fast, but the quality of the interpolated super-resolution image is very limited, because such methods cannot recover high frequency details. The reconstruction based methods apply various smoothness priors and impose the constraint that makes the HR image reproduce the original LR image when properly downsampled. Therefore, the performance of these methods relies on the prior information and the compatibility with the given image. Moreover, it degrades rapidly with the increase of the magnification factor, or the decrease of the size of the input image.

Learning based SR methods are typically example-based and can be characterized as non-probabilistic estimation. They have yielded promising results in recent years. Freeman's work [11] is a milestone for the example-based super-resolution methods. Image super-resolution is transformed into the problem of estimating high-frequency details by interpolating an image into the desired scale. Then, the estimations of high-frequency patches are performed based on the nearest neighbor (NN) patches of the corresponding input low-frequency patches, and the compatibility of output patches is resolved by using a Markov network. Recently, approaches based on NN have achieved advanced progress [24,25]. Michaeli et al. [24] proposed a nonparametric blind SR method. In their method, the SR is treated as an image filtering problem. Hypothetically, an LR image is obtained by blurring and subsampling an HR image. Thus, an SR image can be obtained by using the optimal kernel to filter the LR image. The optimal blur kernel depends on the NN search and can be approximated iteratively. Zhu et al. [25] proposed a deformable patch based SR approach. The deformable similarity is developed and used to find the NN patches. It is well known that the nearest neighbor based estimation suffers from overfitting, that is, the estimation function can explain the training data perfectly but it cannot be generalized to unknown data. This is prominent for image super-resolution. Thus, it is reasonable to improve the NN-based methods by adopting learning algorithms with regularization capability to avoid overfitting.

Many methods made attempts to regularize the SR estimation. Chang et al. [12] made an assumption that the appearances of similar patches in low and high-resolution images may have similar local geometrics and form the manifolds. A high resolution image patch is the linear combination of its HR nearest neighbors and the combination weights are corresponding to those of low-resolution patches. But it is not clear whether the assumption of the manifold structure of similar LR or HR patches is satisfied.

In recent years, sparse regularization has become a very popular tool in super-resolution [15],[18] which regularizes the reconstruction coefficients. Yang et al. [18] first cast the super-resolution problem into a sparse representation problem. Sparse representation based SR methods usually learn a dictionary, and an input image patch can be represented by the dictionary. These methods assume that the representation coefficients are sparse. There are two disadvantages of sparse representation based SR: 1) a trivial solution makes the dictionary too large; 2) there is ambiguity between the HR and LR patches.

It is also rather straightforward to regularize the regressor itself. Ni et al. [16] utilized support vector regression (SVR) to solve the super-resolution problem in the frequency domain and pose the super-resolution as a kernel learning problem. Kim et al. [19] posed the estimation of high frequency details as a kernel ridge regression problem in which he regularized the function with a L_2 norm. The downside of these methods is that they assumed the type of fitting function and the used hyper-parameters are specified a priori. As we know, the regression problem is sensitive to the assumption function. If the assumption function doesn't match the distribution of the data, the correct solution cannot be achieved. The advantages of GP based SR are: 1) the regression function is a latent variable which is decided by the data rather than by the user; 2) the hyper-parameter used in the covariance kernel will be optimized rather than specified a priori.

The method most closely related to our work is reported in [17]. He et al. [17] found a local neighborhood at the position of a query LR patch and implemented the GPR on the local region. This operation is like a filter. An SR image is yielded from a single LR image without any external training set, and each HR image patch is the output of GP regression on its corresponding LR nearest neighbors which contain dozens of LR image patches. Our work is different from He's work [17] because we build a GPR model for SR based on external training data which can be large scale and we focus on low running time.

GP has been recently introduced in computer vision and achieved promising results. Ashish [26] implemented GP on active object categorization, Urtasun [27,28] used GP to model human motion, and William [29] utilized GP for stereo segmentation. However, few literatures discussed the GP framework for SR. In this article, we extend a GPR based approach to solve the SR problem.