

Multi-task Gaussian Process Regression-based Image Super Resolution

Xinwei Jiang
xinwei.jiang@ia.ac.cn
Jie Yang
yangjie@nlpr.ia.ac.cn
Lei Ma
lei.ma@ia.ac.cn
Yiping Yang
yiping.yang@ia.ac.cn

Institute of Automation
Chinese Academy of Sciences
Beijing, China

Image super resolution (SR) aims at recovering the missing high frequency details from single image or multiple images. Existing SR methods can be divided into three categories: interpolation-based, reconstruction-based and example learning-based. Our paper focuses on the third category. Example learning-based SR methods [6] utilize the LR-HR image pair to infer the missing high-frequency details in the LR image and achieve state-of-the-art performance. Recently, in the field of example learning-based SR, more and more researchers resort to learn the LR-HR relationship directly, i.e. $y = f(x)$, where x is the input LR image feature, y is the targeted HR image and f is the mapping function that transforms the LR feature into HR image. Instead of commonly used parametric models, non-parametric methods [3], especially gaussian process regression (GPR)-related methods [2, 4, 5] begin to emerge in the SR field. However, previous GPR-based SR methods simply learn all the GPR models independently and ignore the correlation between them. On the other hand, each pixel prediction can be treated as a task, so that inferring a HR patch can be regarded as a multi-task problem. In this paper, we focus on the multi-task gaussian process (MTGP) regression and apply it to super-resolution problem. We first give a brief overview of MTGP proposed in [1]. Then we study how SR problem corresponds to MTGP and propose the multi-task gaussian process super-resolution (MTGPSR) framework.

MTGP tries to solve the following problem: Given N distinct inputs x_1, \dots, x_N we define the complete set of responses for M tasks as $\mathbf{y} = (y_{11}, \dots, y_{N1}, \dots, y_{12}, \dots, y_{N2}, \dots, y_{1M}, \dots, y_{NM})^T$, where y_{ij} is the response for the j^{th} task on the i^{th} input x_i . We also denote the $N \times M$ matrix Y such that $\mathbf{y} = \text{vec}Y$. Given a set of observations \mathbf{y}_0 , which is a subset of \mathbf{y} , we wish to predict the unobserved values of \mathbf{y}_u of some input points for some tasks. MTGP wishes to learn M related latent functions $\{f_l\}$ by placing a GP prior over $\{f_l\}$ and directly induce correlations between tasks. Assuming that the GPs have zero mean we define

$$\langle f_l(x) f_k(x') \rangle = K_{lk}^f(x, x') \quad (1)$$

$$y_{ij} \sim N(f_j(x_i), \sigma_j^2) \quad (2)$$

where K^f is a positive semi-definite (PSD) matrix that specifies the inter-task similarities, k^x is a covariance function over inputs, and σ_j^2 is the noise variance for the j^{th} task. The key property of multi-task gaussian process model is the introduction of inter-task correlation matrix K^f , so that observations of one task can affect the predictions on another task.

Inference in the MTGP model can be carried out by using the standard GP formulation for the predictive mean and variance. The predictive mean on a test point x_* for task j is obtained by

$$f_j(x_*) = (k_j^f \otimes k_*^x)^T \Sigma^{-1} \mathbf{y}, \text{ where } \Sigma = K^f \otimes K^x + D \otimes I \quad (3)$$

where \otimes denotes the Kronecker product, k_j^f is the j^{th} column of K^f , k_*^x represents the vector consisting of covariances between the test point x_* and all the training points, K^x stands for the covariance matrix obtained by computing covariances between all the training points pairs, D is an $M \times M$ diagonal matrix with σ_j^2 in the $(j, j)^{th}$ position, and Σ is an $MN \times MN$ matrix.

In the learning stage, the MTGP hyper-parameters are optimized as follows: Let f be the vector of function values corresponding to \mathbf{y} , and similarity for F for Y . Further, let $y_{\cdot j}$ denote the vector $(y_{1j}, \dots, y_{Nj})^T$ and similarly for $f_{\cdot j}$. Given the missing data, which in this case is f , the

complete-data log-likelihood is

$$L_{comp} = -\frac{N}{2} \log |K^f| - \frac{M}{2} \log |K^x| - \frac{1}{2} \text{tr}[(K^f)^{-1} F^T (K^x)^{-1} F] - \frac{N}{2} \sum_{j=1}^M \log \sigma_j^2 - \frac{1}{2} \text{tr}[(Y - F) D^{-1} (Y - F)^T] - \frac{MN}{2} \log 2\pi \quad (4)$$

from which we have following updates:

$$\hat{l}_x = \arg \min_{l_x} (N \log |\langle F^T (K^x(l_x))^{-1} F \rangle| + M \log |K^x(l_x)|) \quad (5)$$

$$\hat{K}^f = N^{-1} \langle F^T (K^x(l_x))^{-1} F \rangle \quad (6)$$

$$\hat{\sigma}_j^2 = N^{-1} \langle (y_{\cdot j} - f_{\cdot j})^T (y_{\cdot j} - f_{\cdot j}) \rangle \quad (7)$$

where the expectations $\langle \cdot \rangle$ are taken with respect to $p(f|y_0, l_x, K^f)$, and $\hat{\cdot}$ denotes the updated parameters.

Processing pipeline of our method is as follows: in the training stage, we first construct HR/LR patch pairs by downsampling and then conduct K-means clustering on LR patch dataset. For each cluster, we learn one MTGP model to fit the training data. In the predicting stage, given an LR image, we first overlap sample the image getting test patch dataset and classify the data using the K-NN algorithm based on the cluster centers obtained from training stage. Next, each HR patch subspace corresponding to the LR patch subspace are recovered through learned MTGP regression. Finally, all the predicted patches are reconstructed into a HR image using average weighting scheme. Details of algorithm and experiment are presented in the paper.

In this paper we propose a super resolution framework based on the multi-task gaussian process regression and study how the models are optimized and inferred effectively. The proposed MTGPSR framework makes the pixel prediction correlation into consideration based on the image local structure. Experimental results show that the proposed algorithm achieves the comparative performance and makes the super-resolved image more accurate and natural.

- [1] Edwin V Bonilla, Kian M. Chai, and Christopher Williams. Multi-task gaussian process prediction. In *Advances in Neural Information Processing Systems 20*, pages 153–160. 2008.
- [2] He He and Wan-Chi Siu. Single image super-resolution using gaussian process regression. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 449–456, 2011.
- [3] Li He, Hairong Qi, and Russell Zaretzki. Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 345–352, 2013.
- [4] Younghee Kwon, Kwang In Kim, Jin Hyung Kim, and Christian Theobalt. Efficient learning-based image enhancement: Application to super-resolution and compression artifact removal. In *Proceedings of the British Machine Vision Conference (BMVC)*, pages 1–12, 2012.
- [5] Jianmin Li, Yanyun Qu, Cuihua Li, Yuan Xie, Yang Wu, and Jianping Fan. Learning local gaussian process regression for image super-resolution. *Neurocomputing*, 154(0):284–295, 2015.
- [6] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19(11):2861–2873, 2010.