Information Sciences 462 (2018) 315–330

Contents lists available at ScienceDirect

Information Sciences

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

# Piecewise linear regression-based single image super-resolution via Hadamard transform

Jingjing Luo[a,1], Xianfang Sun[b,1], Man Lung Yiu[c], Longcun Jin[a,*], Xinyi Peng[a]

[a] School of Software Engineering, South China University of Technology, Guangzhou 510006, China
[b] School of Computer Science and Informatics, Cardiff University, Cardiff, CF10 3AT, UK
[c] Department of Computing, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

ARTICLE INFO

*Article history:*
Received 3 February 2018
Revised 10 June 2018
Accepted 12 June 2018
Available online 14 June 2018

*MSC:*
00-01
99-00

*Keywords:*
Single image super-resolution
Hadamard transform
Decision tree

ABSTRACT

Image super-resolution (SR) has extensive applications in surveillance systems, satellite imaging, medical imaging, and ultra-high definition display devices. However, state-of-the-art methods for SR still incur considerable running times. In this paper, we thus propose a method based on the Hadamard pattern and tree search structure to significantly reduce the running time. In this approach, low-resolution (LR) and high-resolution (HR) training patch pairs are classified into different classes based on the Hadamard patterns generated from the LR training patches. The mapping relationship between the LR space and the HR space for each class is then learned and used for SR. Experimental results show that the proposed method can achieve an accuracy comparable to those of state-of-the-art methods with a much faster running speed. The dataset, pretrained models and source code can be accessed at the URL in the footnote[2].



## 1. Introduction

Image super-resolution (SR) is the process of recovering a visually pleasing high-resolution (HR) image from a low-resolution (LR) image. SR has many real applications, such as face recognition [15,27,42], visual question answering [47], visual speaker identification and authentication [20], object understanding [24], activity recognition [28], surveillance systems, satellite imaging, medical imaging, and ultra-high definition display devices. Most existing methods use certain prior information to address the SR problem, especially learned priors. The interpolation-based methods [19,23], reconstruction-based methods [6,35], and learning-based methods [7–10,12–14,21,22,25,26,29–31,34,36–39,41,43,44,48–50] are three classical types of methods for single-image SR. These SR methods are intended to solve natural image SR problems. In addition, a type of SR method exists that only deals with face images; it is called face SR (or face hallucination) [16]. In this paper, the former is our concern.

Learning-based SR methods divide the input LR image into patches and predict their corresponding HR patches using the mapping models that are learned from a dataset of LR-HR patch pairs. These LR-HR patch pairs are cropped from a database composed of LR-HR image pairs. Many learning algorithms have been proposed to learn the mapping models,

* Corresponding author.
*E-mail address:* lcjin@scut.edu.cn (L. Jin).
[1] Equal contribution.
[2] https://github.com/youyouyimu/PLRBSISRvHT

https://doi.org/10.1016/j.ins.2018.06.030

including dictionary learning [10,25,26,29,36,38,39,44,48,50], regression [30,31,38,39,43], decision tree [13,41], random forest [12,14,34] and convolutional neural network (CNN) [8,9,21,22,37].

The advantages and limitations of the above methods are summarized below. Most dictionary learning methods are sparse coding (SC)-based, whereby the sparse prior can well regularize the ill-posed SR problem. However, constructing sparse dictionaries requires expensive computation. Regression-based methods can solve the SR problem by several piecewise linear regression models or a global regression model. Both the decision-tree-based methods and random-forest-based methods are an ensemble of piecewise linear regression models. However, the complex tree structure and a large number of trees (forest) can decrease the retrieving speed of the regression models. The CNN-based methods train a global non-linear regression model to more accurately describe the mapping relationship between the LR space and HR space. The global non-linear regression model consists of a large number of parameters, whose computing processes involve heavy computational loads.

More recently, regression-based methods have achieved great improvements in SR. Linear regression models [18] have higher prediction speeds than non-linear regression models. These methods [8,9,21,22,30,31,37–39,43] learn the relationship between the LR space and HR space, and they use it to solve the SR problem. Timofte et al. [38,39] assumed that the mapping relationship between the LR space and HR space is locally linear and that many linear regressors are therefore learned and anchored to the feature space as a piecewise linearization. The methods [30,31] divide the feature space into many subspaces based on antipodally invariant metrics and learn a linear regressor for each subspace. In the papers [8,9,21,22], the mapping from the LR space to the HR space is described as a deep CNN that takes an LR image as the input and outputs an HR image, which is an end-to-end mapping (i.e., a global non-linear regressor). The method in [22] uses a generative adversarial network (GAN) for the image SR, in which a perceptual loss function [17] is adopted.

SR methods based on CNN require a significant amount of training time, making them not suitable to certain application scenarios. Some SR methods that are based on SC [38,39] and simple functions [43] use the gradients of the LR image patches and normalized image patches, respectively, to represent image features. However, this approach adds to the computational complexity of the training phase. To address the above limitations, we propose a Hadamard pattern-based SR method using the decision tree [13] for the single-image SR.

A Hadamard matrix [3] is a square matrix that consists of +1 and −1. It is symmetric with respect to the leading diagonal. Its rows (columns) are orthogonal to one another. The Hadamard transform uses a Hadamard matrix as its operator. It is thus useful in signal-image processing [32], including signal/image coding/decoding [1,33] and compressive sensing [45]. In some cases, it is also useful in image filtering and pattern recognition. In our paper, we perform Hadamard transform to code the LR training image patches. The coding results (Hadamard patterns) are bases of classifying training data.

The contributions of this paper are summarized in three aspects.

(1) We propose implementation of the Hadamard transform on LR image patches and use the obtained Hadamard patterns to represent image features. The Hadamard transform is fast because it requires only addition and subtraction without multiplication or division. This property makes feature extraction efficient.

(2) We employ a variant of the decision tree-the ternary decision tree-to conduct fast classification and regression.

(3) The experimental results show that the proposed method can achieve comparable accuracy with a much shorter running time compared to state-of-the-art methods.

The remainder of this paper is organized as follows. Section 2 defines the problem and briefly describes the related work. Section 3 presents our solution to the single-image SR problem. Section 4 analyzes the experimental results, and Section 5 concludes this paper.

## 2. Related work

### 2.1. Problem statement

Single-image SR is intended to reconstruct an HR image with high definition and fidelity from an LR image that has unsatisfactory resolution. It can be formulated as

$$\hat{X} = \uparrow Y, \ s.t. \ \hat{X} \approx X, \tag{1}$$

where $Y$ is the LR input image, $\hat{X}$ is the upscaled output image, $X$ is the original HR image, and $\uparrow$ is an upsampling operator. In the training phase, $X$ is known. This formula implies that the upsampling mapping models are trained to describe the relationship between the LR space and the HR space as accurately as possible. In the testing phase, $X$ is unknown. This formula implies that the learned mapping models generate an HR output from an LR input. The predicted HR output and the HR image generated from the same imaging model are as similar as possible.

In the literature, the following transformation is usually used to describe the real imaging process of LR images:

$$Y = \downarrow BX + n, \tag{2}$$

where $\downarrow$ is the downsampling operator, $B$ denotes the blurring operator, and $n$ is the additive noise. Most SR methods solve this problem at a patch level.