https://doi.org/10.1016/j.image.2018.12.001

Image Super-resolution via Feature-augmented Random Forest

Hailiang Li[1], Kin-Man Lam[1], Miaohui Wang[2]

[1]Department of Electronic and Information Engineering, The Hong Kong Polytechnic University
[2]College of Information Engineering, Shenzhen University, Guangdong, China

harley.li@connect.polyu.hk, enkmlam@polyu.edu.hk, mhwang@szu.edu.cn

Abstract — Recent random-forest (RF)-based image super-resolution approaches inherit some properties from dictionary-learning-based algorithms, but the effectiveness of the features working in RF is overlooked in the literature. In this paper, we present a novel feature-augmented random forest (FARF) method for image super-resolution, where the conventional gradient-based features are proposed to augment the features used in RF, and different feature recipes are formulated on different processing stages in an RF. The advantages of our method are that, firstly, the dictionary-learning-based features are enhanced by adding gradient magnitudes, based on the observation that the non-linear gradient magnitudes are highly discriminative. Secondly, generalized locality-sensitive hashing (LSH) is used to replace principal component analysis (PCA) for feature dimensionality reduction in constructing the trees, but the original high-dimensional features are employed, instead of the compressed LSH features, for the leaf-nodes' regressors. With the use of the original higher dimensional features, the regressors can achieve better learning performances. Finally, we present a generalized weighted ridge regression (GWRR) model for the leaf-nodes' regressors. Experiment results on several public benchmark datasets show that our FARF method can achieve an average gain of about 0.3 dB, compared to traditional RF-based methods. Furthermore, a fine-tuned FARF model can compare to, or (in many cases) outperform, some recent state-of-the-art deep-learning-based algorithms.

Keywords— Random forest; gradient magnitude filter; clustering and regression; image super-resolution; weighted ridge regression.

1. INTRODUCTION

In the past few years, random forest (RF) [3, 14], a machine-learning tool, working via an ensemble of decision trees, has been employed for efficient classification and regression problems, and applied to a large variety of computer-vision applications, such as object recognition [27], face alignment [15, 18, 21], data clustering [17], single image super-resolution (SISR) [8, 19], and so on. The RF-based SISR approach can be considered as a clustering/classification-based method, as shown in Fig. 1. However, the clustering and regression problems in RF require different discriminative features, which have not been systematically studied in the literature.

Feature engineering is a research hotspot on the image-restoration problems. Pioneer work in [45] used a simple high-pass filter, which is simply subtracting the output of a low-pass filter from the input image. Meanwhile, most algorithms [1, 2, 4, 5, 8] follow the approach in [28], which concatenates the first and second-order gradients to form the features, as an inexpensive solution to approximating high-pass filtering. Since RF can be used as a dictionary-learning-based tool, it inherits many properties from conventional dictionary-learning-based algorithms for feature extraction. However, the discriminative ability of those gradient-based features for random forest has been overlooked in the literature. We found,

from experiments, that augmented features based on two gradient-magnitude filters can achieve more than 0.1dB quality improvement in RF-based SISR, with the same parameter setting.

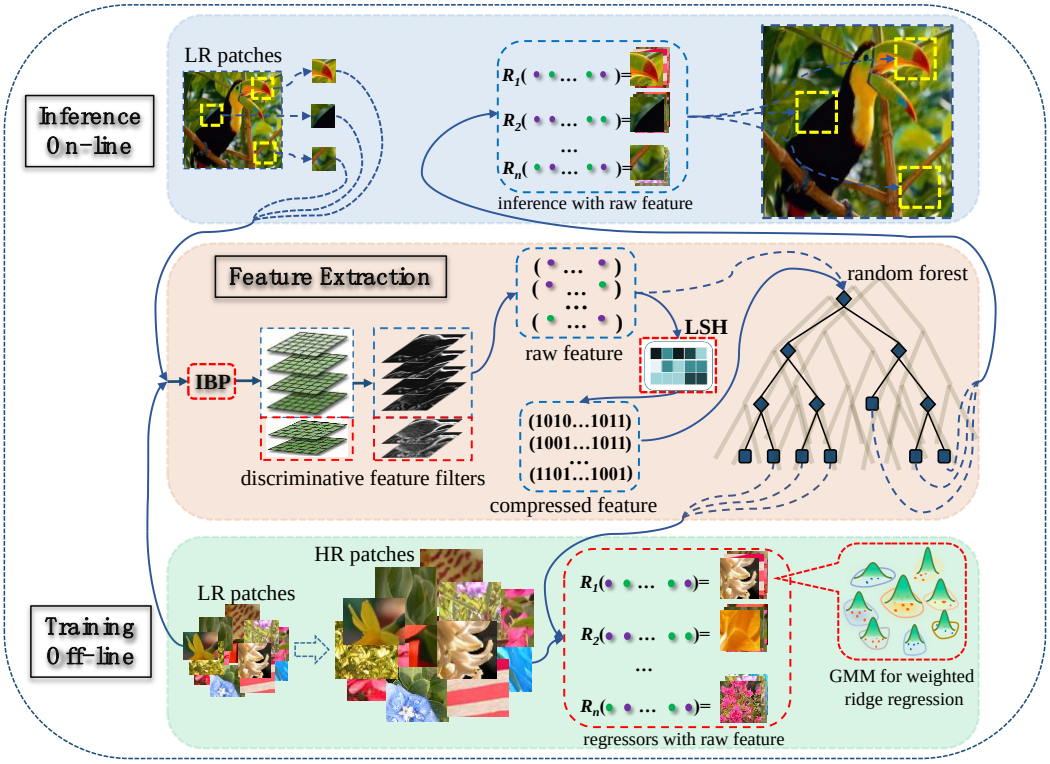


Fig. 1: An overview of the proposed FARF framework for image super-resolution.

In the FARF-based image SR scheme, firstly, the initial coarse estimation is generated by using iterative back projection (IBP) instead of bicubic interpolation. More discriminative features are extracted by using the first and second-order gradients, as well as their magnitudes. Then, the conventional PCA is replaced by the generalized LSH for dimensionality reduction, and the compressed features are used for clustering at the split nodes in an RF. Finally, the respective regressors at the leaf-nodes are learned, by using the original high dimensional features with the GWRR models.

In most dictionary-learning-based algorithms, principal component analysis (PCA) is used for dimensionality reduction before classification and regression. The impact of using PCA has also been paid less attention in the literature. PCA projection may damage the structure of features, which are originally discriminative for clustering at the split nodes and regression at the leaf nodes. Motivated by the content-based image retrieval (CBIR) techniques in [13, 46], where the coarse-level search uses compressed features, while the fine-level search uses augmented features, we propose a similar approach for building more efficient trees, with their leaf nodes having better regression performances. To achieve this, our method uses the original features, rather than the compressed features generated by PCA as worked in [1, 2, 4, 5, 8, 28], so that more accurate regression and higher image quality improvement can be achieved. Moreover, unsupervised locality-sensitive hashing (LSH), instead of PCA, is employed for feature dimensionality reduction, which can reduce the damage on the feature structure after compression. The compressed features are used for clustering at the split nodes. Therefore, the resulting forest can improve the quality of reconstructed images. For the regression problems at the leaf nodes, in addition to using the augmented features without compression, we propose a generalized weighted ridge regression (GWRR) model as an extension of the work in [1]. The GWRR models are generated based on data distributions at the leaf nodes.

The main contribution of our method is on feature augmentation, so we call our method feature-augmented random forest (FARF). The pipeline of our FARF method, which includes feature extraction, the training stage, and inference stages for SISR, is shown in Fig. 1.

Having introduced the main idea of our paper, the remainder of this paper is organized as follows. In Section 2, we review the related works on SISR, particularly the RF-based approaches and our insights. In Section 3, we introduce the proposed FARF method, including the discriminative feature augmented by the gradient-magnitude filters, the generalized weighted ridge regression (GWRR) model, and the fine-tuned FARF. In Section 4, we evaluate our FARF scheme on public datasets, and conclusions are given in Section 5.

2. Image Super-Resolution via Random Forest

2.1 Conventional Patch-based Image Super-Resolution

Image SR attempts to achieve an impressive HR quality image from one or a set of LR images via artistic skills, which has been an active research topic for decades in the image restoration area. Generalized SR includes interpolation algorithms, such as the classic bicubic interpolation, and other edge-preserving algorithms [23, 41, 42, 43, 44].

The traditional super-resolution algorithms are based on pixel operations. Intuitively, operating on a "big pixel", i.e. a patch [24], is more effective. Since patch-based algorithms can preserve the local texture structure of an image, various methods based on image patches, such as non-local means [23], self-similarity [31], manifold learning [29], block-matching and 3D filtering (BM3D) [25], sparse representation [28], etc. have been proposed.

The neighbor-embedding (NE) methods [29, 30] are the milestone for patch-based dictionary learning methods. NE learns the mapping between low and high-resolution patches, with the use of manifold learning. Based on locally linear embedding (LLE), an LR patch can be represented as a linear combination of its nearest neighbors in a learned dictionary, and its HR counterpart can be approximated as a linear combination of the corresponding HR patches of its LR neighbors, with the same coefficients. Although the NE method is simple and sounds practical, a problem with the method is how to build an effective patch dictionary.

An approach to reducing the dictionary size is to learn a relatively smaller dictionary with discrete cosine transform (DCT) or wavelet fixed basis. However, the adaptiveness to data is sacrificed. In 2010, Yang *et al*. [28] proposed a sparse prior for dictionary learning. Using sparse coding, image representation can work with a relatively smaller dictionary, while keep the adaptiveness by learning the basis from data directly. This approach opens the era for sparse coding in the image inverse problems. Although the l^0-norm of α (the sparse coefficients) is an ideal regularization term for the sparse constraint, this strong constraint leads to an NP-hard problem in solving the coefficients α. Yang *et al*. [28] relaxed the l^0-norm to l^1-norm, so as to achieve a feasible solution.

Meanwhile, the effectiveness of sparsity is challenged [1, 5] by researchers, as to whether sparsity or collaborative representation really helps in image classification and restoration. As a natural solution to that, Timofte *et al*. proposed an anchored neighborhood regression (ANR) [2] framework, where there is no sparse constraint in the model. ANR replaces the sparse-decomposition optimization (l^1-norm) with a ridge regression (i.e. l^2-norm), where the coefficients can be computed offline and each coefficient can be stored as an atom (anchor)'s projection matrix in the dictionary. The offline learning is computationally intensive, but the online or prediction stage is very efficient. This approach has subsequently led to several variants. Timofte *et al*. later extended the ANR approach to the A+ [5]. In A+ [5], the coupled dictionaries are trained from a large pool of training samples (in the order of millions)

rather than only from the anchoring atoms, which greatly improves the image quality. After that, more extensions based on ANR and A+ have emerged [1, 33, 35, 36].

2.2 Random-Forest-based Image Super-Resolution

Previously, in the above-mentioned dictionary-learning methods, the complexity of finding those similar patches by comparing an input patch with all the dictionary items has been overlooked. Recently, algorithms using random forest [2, 5, 8] have achieved state-of-the-art performances, in terms of both accuracy and efficiency for classification and regression tasks. This is mainly due to the use of ensemble learning and sublinear search based on binary decision trees. Schulter *et al*. [8] adopted random forest and the clustering-regression scheme to learn regressors from the patches in leaf nodes for SISR. With the same number of regressors, the RF-based algorithm can outperform or achieve comparable performance with A+ and its variants, in terms of accuracy but with less computational complexity.

RF-based image super-resolution, following a recent emerging stream [5, 31] on single-image SR, formulates the SR problem as a clustering-regression problem. These emerging approaches attempt to reconstruct an HR image from patches with the aid of an external database. These methods first decompose an image into patches, then classify the patches into different clusters, and later regressors are trained for all the respective clusters, which generate mappings from the features of an input LR patch to those of the corresponding HR patch. In the inference stage, an LR image follows the same procedures, such that it is divided into patches and features are extracted from each patch. Then, the patches are classified into different clusters using *K*-NN [8, 19] or RF [2, 5, 8], and their super-resolved HR patches are computed through regression in the leaf nodes (see Fig. 1). This kind of clustering-regression-based random forest [2, 5, 8] methods has achieved state-of-the-art performance in SISR, both in terms of accuracy and efficiency.

2.3 Deep Learning based Image Super-Resolution

In recent years, deep learning has achieved promising performances on image super-resolution [37, 38, 39, 40]. In [37, 38], milestone works on image super-resolution based on deep learning were presented, where a convolutional neural network (SRCNN) was proposed to learn an end-to-end mapping between LR and HR images for image super-resolution. Wang et al. [7] extended the SRCNN network to a sparse coding-based network by combining the domain knowledge. Later, a scheme with very deep networks for SISR was proposed in [39], where the convergence rate of the deep network is improved by using residual learning and extremely high learning rates. In addition, Ledig *et al*. [40] introduced a generative adversarial network (GAN)-based image super-resolution model (SRGAN), where the image perceptual loss function is reformulated as the combination of content loss and adversarial loss. Although deep-learning-based approaches have achieved promising progress on SISR, the heavy computational requirement is still a large burden even though the implementation is accelerated by GPU. This may limit them from those applications without powerful GPUs, such as smart mobile terminals.

3. Feature-Augmented Random Forest

Classification and regression can be regarded as probability problems from the statistical theory. Historical frequentist probability is the probability obtained from the relative frequency in a large number of trials. In contrast, the Bayesian probability is an interpretation of the concept of probability, in which probability is interpreted as an expectation taking the knowledge and personal belief into account. From the Bayesian theory, the posterior probability of a random event is a conditional probability, which can be calculated if the relevant evidence or context is considered. Therefore, the posterior probability is the probability $p(\theta|x)$ of the parameters θ given the evidence x. We denote the probability distribution function of the prior for parameters θ as $p(\theta)$, and the likelihood as $p(x|\theta)$, which is the probability

of x given θ. Then, based on the Bayesian rule, the posterior probability can be defined as follows:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}. \quad (1)$$

The posterior probability can be written in a memorable form as:

$$Posterior\ probability \propto Likelihood \times Prior\ probability.$$

Based on the Bayesian framework, the likelihood term and the prior term are both required to be determined in order to solve the inverse problems, and the extracted features are normally worked as a prior or a likelihood, particularly on some image-restoration problems. From this point of view, most research works, from classic feature extractors to deep-learning neural networks, are essentially done under the Bayesian inference framework.

Since SISR is a well-known ill-posed problem, researchers have put their efforts into the priors of the problem with skills from mathematics, computer vision and machine learning. One of the obvious and most studied priors is the edge prior, which can be found in many pioneering works: new edge-directed interpolation (NEDI) [41], soft-decision adaptive interpolation (SAI) [42], directional filtering and data-fusion (DFDF) [43], modified edge-directed interpolation (MEDI) [44], and so on. The edge prior is effective on image processing, and the first and second-order gradients were studied and employed by Yang *et al*. [28] in a pioneering dictionary-learning-based algorithm. However, the effect of edge-based features has not been investigated in depth.

3.1 Augmented Features via Gradient Magnitude Filters

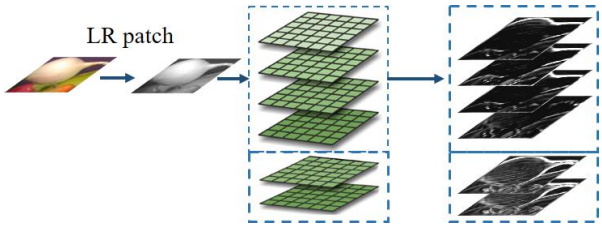


Fig. 2: Features extracted from an LR image through the first and second-order gradient filters in horizontal and vertical directions (the upper four), as well as the gradient magnitude filters (the below two), are concatenated to form augmented features with more discriminative properties.

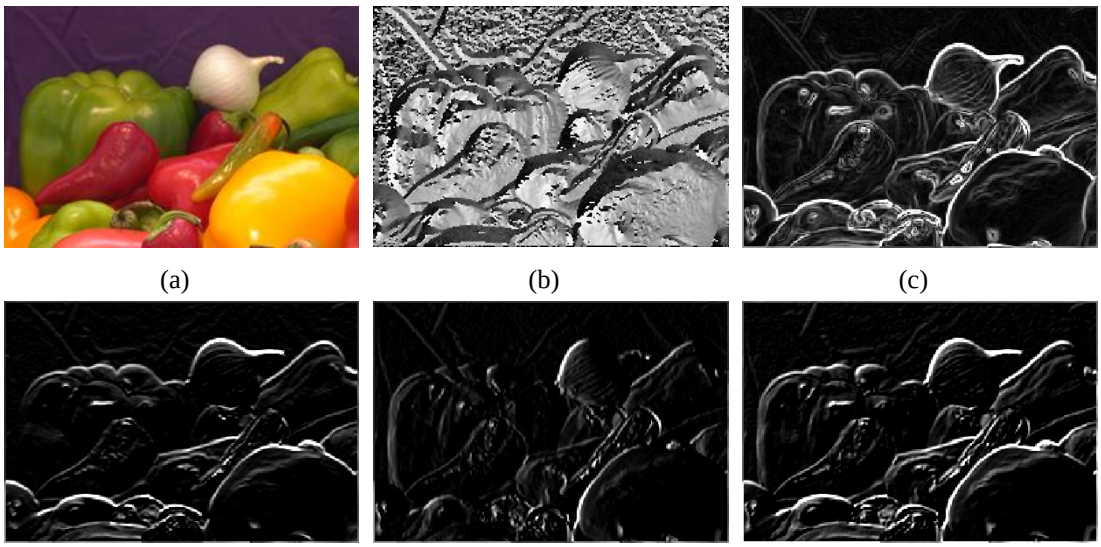
For the clustering and classification problems, feature engineering is a critical research point, and in some cases, the chosen features may dominate the performance. A feature filter F, whose coefficients are computed to fit the most relevant parts in the LR image patches, is employed, and the generated features can achieve more accurate predictions for reconstructing their corresponding HR image patches, as shown in Fig. 2.

Normally, it is unstable to directly use pixel intensities as features, which are susceptible to the environmental lighting variations and camera noise. Instead, the differences between the neighboring pixels' intensity values, which are computationally efficient, and are immune to lighting changes and noise, are examined. This type of feature can be implemented efficiently through convolutional filters. Typically, the feature filter F can be chosen as a high-pass filter, while in [2, 4, 5, 28], the first and second-order gradient operators are used to generate an up-sampled version from a low-resolution image, then four patches are extracted from the gradient maps at same location, and finally the patches are concatenated to form feature vectors. The four 1-D filters used to extract the gradients are described in Eqn. (2),

$$\left.\begin{matrix} F_1 = [-1, 0, 1], & F_2 = {F_1}^T \\ F_3 = [1, 0, -2, 0, 1], & F_4 = {F_3}^T \end{matrix}\right\}, \tag{2}$$

where F_1 and F_2 are first-order gradient filters, while F_3 and F_4 are second-order gradient filters.

These features can work well on dictionary-learning-based methods, because when searching a matched patch in a dictionary, the distance is calculated based on the whole feature vectors with the Euclidean distance. However, when training a split node in a decision tree of an RF, only one or a few of the features are chosen as candidate features for comparison. Therefore, more discriminative features are required for RF, when compared with dictionary-learning-based methods.



(a) (b) (c)

(d) (e) (f)

Fig. 3: Visualization of the features from a natural image: (a) original color image, (b) image *gradient orientation*, (c) image *gradient magnitude*; (d) *horizontal gradient* $\partial I/\partial x$, (e) *vertical gradient* $\partial I/\partial y$, (f) the sum: $(\partial I/\partial x + \partial I/\partial y)$.

The first and second-order *gradients* of an image can provide the directions of edges in a perceptual manner, as shown in Fig. 3, which can be calculated as follows:

$$\nabla I = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right]^T, \tag{3}$$

where $\partial I/\partial x$ and $\partial I/\partial y$ are the *gradients* in the *x*-axis and *y*-axis directions, respectively, at a given pixel. Meanwhile, the *gradient magnitude* image can provide the edge strength, as described in Eqn. (4).

$$\|\nabla I\| = \sqrt{(\frac{\partial I}{\partial x})^2 + (\frac{\partial I}{\partial y})^2}. \tag{4}$$

With the natural image shown in Fig. 3, it can be observed that the *gradient magnitude* image has more detailed textures than the *gradient* images ($\partial I/\partial x$ and $\partial I/\partial y$), as well as the sum of the horizontal *gradient* and vertical *gradient* image, i.e. $\partial I/\partial x + \partial I/\partial y$, perceptually. An explanation for this phenomenon is that non-linear features are usually more discriminative. Thus, in our work, all the first and second-order *gradients,* and *gradient magnitude* are employed, and are concatenated to form more discriminative, augmented features.

On the other hand, the image *orientation* (gradient angle) is defined by the following formulation,

$$\angle \nabla I = \tan^{-1}(\partial y/\partial x), \tag{5}$$

where $\tan^{-1}(.)$ is the gradient orientation, with a value between -90° and 90°. As shown in Eqn. (5), when the value of ∂x is equal to 0 or close to 0, the value of $\angle\nabla$ becomes infinitely large and unstable, i.e., different ∂y will result in approximately the same $\angle\nabla$ value.

Based on this analysis, we only use the two *gradient magnitude* filters derived from the four gradient filters [28], shown in (2), to generate the augmented features. Experiments validate that the use of the augmented features can improve the conventional RF algorithm [8] to achieve a performance gain of more than 0.1dB, which is a remarkable improvement, with the same setting and parameters.

3.2 Fine-grained Features for Regression

The inference stage of the RF-based image super-resolution process is similar to the content-based image retrieval (CBIR) framework, as shown in Fig. 4. The general approximated nearest neighbor (ANN) search framework [13, 46] is an efficient strategy for large-scale image retrieval, which mainly consists of 4 parts: (1) extracting compact features (e.g., the locality-sensitive Hashing (LSH) [10] features) for a query image; (2) coarse-level search using Hamming distance to measure the similarity between binary compact Hash features, then narrow the search scope into a smaller candidate group; (3) fine-level search by using Euclidean distance to measure the similarity between their corresponding feature vectors; and (4) finding the object in the smaller candidate group that is the nearest one to the query image.

Locality sensitive hashing (LSH) [6, 47] is a series of algorithms which are used for handling largescale data processing with high dimensionality. The main idea behind LSH is to formulate a family of functions which can map (hash) high dimensional features into buckets so that similar features can fall into the same bucket with a high probability.

In the inference stage of conventional RF-based SISR, PCA projection is worked as a hash-like function to reduce the feature dimension for decreasing the search range, which can speed up the searching as the coarse-level search in a CBIR framework. However, the impact of using PCA on feature dimensionality reduction has been overlooked in previous works [1, 2, 4, 5, 8, 28]. In our algorithm, the LSH based scheme from [32] is employed to further improve feature discrimination. This LSH based scheme [32] transforms the original feature space into a compactly pre-clustered feature space, via a trained rotation matrix. The additional load of the rotation in the inference stage is ignorable, and the feature dimension will keep the same as PCA does [1, 2, 4, 5, 8, 28], i.e., for an enlargement with scale factor 3, the original dimension is 486 (for the 9×9 size image patch with 6 filters), and the final dimension is about 50 after 90% dimensionality reduction with PCA or LSH.

Inspired by the fine-level search using augmented features in CBIR frameworks, the high-dimensional features used in the leaf nodes of an RF can further improve the prediction accuracy in the regression step, which has not been studied previously. Consequently, we use the original features, rather than the PCA or LSH compressed features, to perform ridge regression in the leaf nodes. Experimental results show that this new RF scheme, by using this augmented feature, can greatly improve the quality of super-resolved images. Another explanation for this is that the regression problems can benefit more from the higher dimensional features than classification problems.

Based on the observation that the original edge-like features are used for the final regressors in the leaf nodes and the compressed features (either produced by PCA or LSH) are used for clustering in the split nodes, a new clustering-regression-based SISR approach can be designed as shown in Fig. 4. In this new scheme, the original-compressed coupled feature sets are worked for different purposes at different stages, i.e., the original edge features are used for regression in the leaf nodes, and the compressed features derived from the LSH-like functions are employed for node splitting (clustering) in the training stage,

and node searching in the inference stage in the split nodes. On training the split nodes, the entropy is replaced by variance as the works in [8, 32].

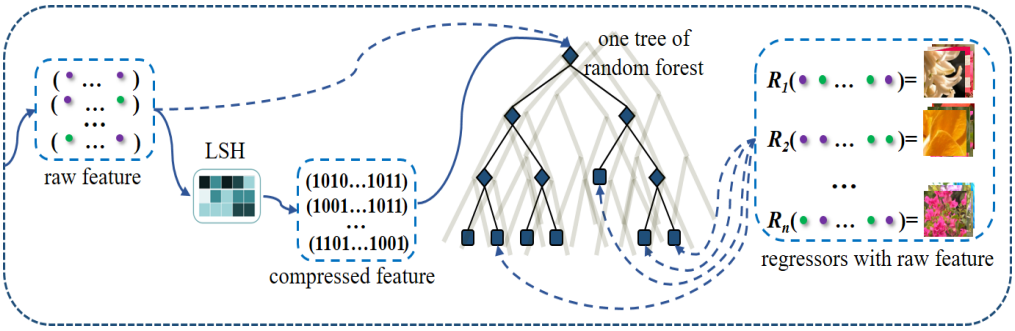


Fig. 4: Augmented features for regressors and the LSH compressed features for searching in the trees of a random forest.

In the new scheme, we unify the research of LSH-based SISR and image retrieval (CBIR) [46, 13]. In brief, the new achievement on unsupervised LSH can be evaluated not only in CBIR systems, but also in the clustering-regression RF-based SISR methods. Moreover, as evidence from [32], using proper unsupervised LSH models, e.g., iterative quantization (ITQ) [12], for feature-dimension reduction, instead of PCA, can reduce the damage on the image structure. This can further improve the super-resolved image quality. Different from [32] using an ITQ-like algorithm to rotate the original features into a new feature space, with the use of the proposed original-compressed coupled feature sets, any unsupervised LSH generated features can directly be employed.

3.3 Generalized Weighted Ridge Regression Model

In this sub-section, we further analyze the ridge regression employed in the RF leaf nodes. The anchored neighborhood regression (ANR) [2] model relaxes the l^1-norm to the l^2-norm, with least-squares minimization as the following equation,

$$\min_{\alpha} \|F\mathrm{D}_l\alpha - Fy\|_2^2 + \lambda\|\alpha\|_2, \tag{6}$$

where α is the coefficients, and F is a feature-extraction operator on the LR patches, which aims to extract discriminative features from LR patches, rather than using the raw pixel intensity. In our formulation, D_l and D_h represent the low and high-resolution coupled dictionaries trained jointly from LR and HR patch samples, respectively. Based on the ridge regression [16] theory, this l^2-norm constrained least square regression regularized problem has a closed-form solution, according to the Tikhonov regularization theory, as follows:

$$\alpha = (D_l^T D_l + \lambda I)^{-1} D_l^T Fy. \tag{7}$$

With the assumption in [28], where HR patches and their counterpart LR patches share the same reconstructed coefficient α, i.e. $x = \mathrm{D}_h\alpha$, from Eqn. (7) we have

$$x = \mathrm{D}_h(D_l^T D_l + \lambda I)^{-1} D_l^T Fy. \tag{8}$$

If we define P_G as a pre-calculated projection matrix, as follows,

$$P_G = \mathrm{D}_h(D_l^T D_l + \lambda I)^{-1} D_l^T, \tag{9}$$

then the HR patches can be reconstructed with $x = P_G Fy$.

Having studied the model in Eqn. (8), the authors in [1] argued that different weights should be given to different atoms when reconstructing an HR patch so as to emphasize the similarity to the anchor atom.

Based on this idea, authors in [1] proposed a weighted collaborative representation (WCR) model by generalizing the normal collaborative representation (CR) model in ANR, as follows:

$$\min_{\alpha} \|FD_l\alpha - Fy\|_2^2 + \|\lambda_{WCR}\alpha\|_2, \tag{10}$$

where λ_{WCR} is a diagonal weight matrix, whose non-zero diagonal entries are proportional to the similarities between the atoms and the anchor atom.

Same as the ANR model, a new closed-form solution can be computed offline through the following equation,

$$\alpha^* = (D_l^T D_l + \lambda_{WCR})^{-1} D_l^T Fy, \tag{11}$$

and the new projection matrix can be derived as

$$P_G^* = \mathrm{D}_h(D_l^T D_l + \lambda_{WCR})^{-1} D_l^T. \tag{12}$$

The WCR model further improves the ANR/A+ model in terms of image quality, while keeping the same level of computation. In [9], the local geometry prior of the data sub-space is used. However, all the weighted ridge regression models [1, 9] are constructed based on an existing dictionary, e.g., Zeyde *et al*. [4] used K-SVD to train a sparse-coding-based dictionary with 1,024 items.

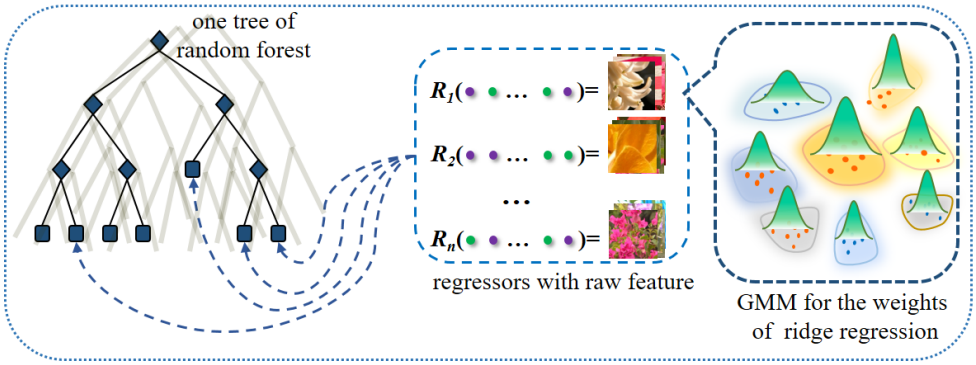


Fig. 5: Gaussian mixture model (GMM) is used to generate the weights for weighted ridge regression, and the weight of each entry lies on its belonging cluster's size and its centrality in the belonging cluster.

When training the regressors in an RF, there is no existing anchor point in the clustered groups of the leaf nodes, similar to the previous models [1, 9]. A solution to the mentioned problem is inspired by the work on image classification using locality-constrained linear coding (LLC) [34], where the Gaussian mixture model (GMM) is used to describe the locality-constrained affine subspace coding (LASC) [22]. We employ GMM to construct the data distribution in the sub-space for each leaf node, which derives the weights of all the entries in the ridge regression models. Through the derived weights, we can obtain a generalized weighted ridge regression (GWRR) model for ridge regression. The new projection matrix is given as follows:

$$P_G^* = \mathrm{D}_h(D_l^T D_l + \lambda_{GWRR})^{-1} D_l^T, \tag{13}$$

where λ_{GWRR} is a diagonal weight matrix, and diagonal entries are the weights of samples in the leaf-node. Each sample's weight is related to its belonging cluster's size and its local centrality in its belonging cluster, as illustrated on the right of Fig. 5. In other words, the cluster's weight is proportional to its sample amount, and the local weight of a sample is depending on how close it is to the cluster center. Obviously, a query input, falling into a bigger cluster and closer to the center of the belonging cluster, achieves a larger weight. In a rough form, the diagonal weight matrix λ_{GWRR} is given as follows:

$$\lambda_{GWRR} = diag\{[w_1; w_2; \dots; w_i; \dots]\}, w_i \propto C_i^k \times d_i^k, k = (1, \dots, K) \tag{14}$$

where w_i is the weight of the i^{th} entry, m is the number of samples in the leaf nodes, C_i^k is the k^{th} cluster's weight for the i^{th} entry, d_i^k is the i^{th} entry's local weight in the k^{th} cluster, which is approximated as the inverse of the distance to the center of the belonging cluster, and K is the number of clusters generated by the GMM model for a leaf node, e.g., K=32 is set in our experiments.

Experimental results in Table-1 show that the proposed GWRR model can achieve the same level of performance as WCR [1] and obtain 0.2dB gain than the ANR [2] model. Note that when the number of samples in a leaf node becomes bigger, the performance of the GWRR model will achieve less advantages than the normal regression model, because the higher weights will be averaged by a large number of other samples. Theoretically, the regression of a leaf node can benefit from the GWRR model, particularly when there are a few samples falling into the leaf nodes.

Images	*baboon*	*baby*	*bird*	*butterfly*	*foreman*	*head*	*lenna*	*woman*	***Average***
ANR	23.52	35.06	34.44	25.74	32.92	33.54	32.92	30.17	31.04
WCR	23.55	35.09	34.75	26.18	33.51	33.61	33.16	30.42	**31.28**
GWRR	23.54	35.09	34.74	26.13	33.46	33.58	33.12	30.38	31.25

Table-1: Performances of ANR [2], WCR [1], and the proposed GWRR, in terms of PSNR (dB) with an upscale factor (×3) on some public standard test images.

3.4 Fine-Tuning with Iterative Back Projection for Initial Coarse Estimation

In a broad sense, SISR is a low-level computer vision task, which attempts to restore an HR image $\mathcal{X}$ from a single input LR image $\mathcal{Y}$. A mathematical model for image degradation can be formulated as follows:

$$\mathcal{Y} = (\mathcal{X} * \mathcal{B}) \downarrow s, \tag{15}$$

where $\mathcal{B}$ is a low-pass (blur) filter and $\downarrow s$ denotes the down-sampling operator with a factor of s. Based on a given LR image $\mathcal{Y}$, how to achieve an approximated HR image $\hat{\mathcal{X}}$ is a classic inverse problem, which requires priors based on the Bayesian theory.

Irani and Peleg [26] firstly proposed an iterative back projection (IBP) method for SR reconstruction. IBP is an effective way to obtain an HR image when comparing it with other SR methods. In the IBP method, the reconstruction error of an estimated LR image $\hat{\mathcal{Y}}$ is the difference between the input LR $\mathcal{Y}$ and the synthesized image $\hat{\mathcal{Y}}$ generated from the estimated HR image $\hat{\mathcal{X}}$ as follows:

$$e(\hat{\mathcal{Y}}) = \mathcal{Y} - \hat{\mathcal{Y}} = \mathcal{Y} - (\hat{\mathcal{X}} * \mathcal{B}) \downarrow s. \tag{16}$$

IBP can efficiently obtain the HR image by minimizing the reconstruction error defined in Eqn. (16). For the IBP approach on SISR, the updating procedure can be summarized as the following two steps, performed iteratively:

- Compute the reconstruction error $e(\hat{\mathcal{X}})$ with the following equation:
$$e(\hat{\mathcal{X}}) = e(\hat{\mathcal{Y}}) \uparrow \mathrm{s} * p, \tag{17}$$
where $\uparrow$ is the up-sampling operator and p is a constant back-projection kernel to approximate the inverse operation of the low-pass filter $\mathcal{B}$.
- Update the estimated HR image $\hat{\mathcal{X}}$ by back-projecting errors as follows:
$$\hat{\mathcal{X}}^{t+1} = \hat{\mathcal{X}}^t + e(\hat{\mathcal{X}}^t), \tag{18}$$
where $\hat{\mathcal{X}}^t$ is the estimated HR image at the t-th iteration.

Most learning-based algorithms [1, 2, 4, 5] follow the milestone work in [28], which uses the coarse estimation firstly obtained via bicubic interpolation. As we know, the classic IBP algorithm is an efficient way to obtain high-quality up-scaled images, but it will inevitably produce artifacts (such as ringing,

jaggy effects, and noise) at the output, because the kernel operator p in Eqn. (17) is hard to estimate accurately. That is the reason why algorithms with IBP need an additional denoising process [11, 23, 26]. However, the sparse-constraint-based approach [28] does not have this denoising capability.

As the l^2-norm constraint-based ridge regression has the denoising effect, due to its averaging-like process, this means that the ridge regression-based RF scheme has the denoise capability intrinsically. Based on this observation, we obtain the coarse estimation of an HR image $\widehat{x}$ by applying IBP to the corresponding input LR image y. Experimental results in Table-2 and Table-3 validate that using IBP, instead of bicubic, to obtain the initial coarse estimation can help the RF-based SR method obtain a remarkable improvement.

3.5 Fine-Tuning with Proper Trees in Random Forest

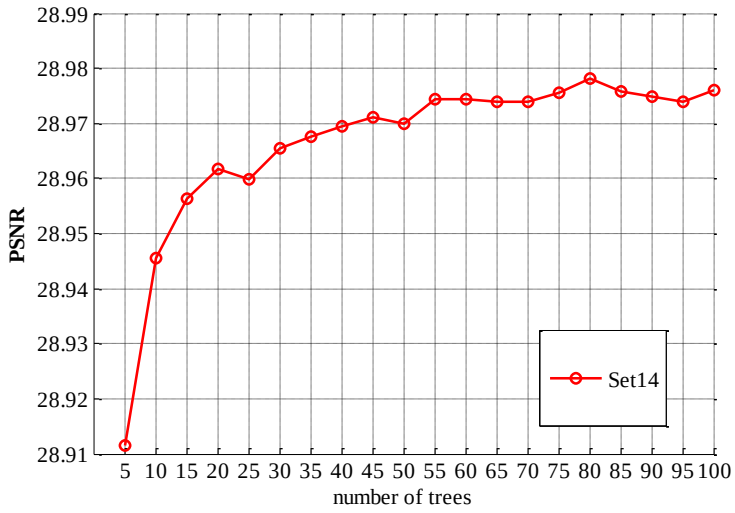


Fig. 6: The image super-resolution quality, in terms of PSNR, with different numbers of trees in a random forest for super-resolution (3x) experiments on Set14. The number of trees = 45 gives a better trade-off between efficiency and complexity.

As the number of trees is an important parameter in RF-based approaches, we plot the performance with respect to the number of trees. As shown in Fig. 6, the performance of the RF-based image super-resolution method increases as expected, but the increment becomes relatively smaller after a certain number of trees are used. The experimental results in Fig. 6 were obtained on the Set14 dataset, and 2 million samples from the dataset were used for all training stages. It shows that using 45 trees is an optimal number, as a tradeoff between performance and computational cost. Therefore, we set the number of trees for the proposed FARF method at 45, and our method with this number is denoted as FARF*. The performances of our methods, and other methods, are tabulated in Table-2 and Table-3. We also compare our methods with a recently proposed deep-learning-based algorithm, SRCNN algorithm [37, 38] and SCN algorithm [7], and our methods outperform them in some cases.

3.6 Algorithm Workflow

The training and inference stages of the proposed FARF algorithm are described in Algorithm 1 and Algorithm 2, respectively. To help the readers understand our paper, the source code of our algorithm will be available at: https://github.com/HarleyHK/FARF, for reference.

Algorithm 1: Training Stage of FARF-based Image Super-Resolution:

Input: $\{\mathcal{Y}^i, x^i\}_{i=1}^N$: training LR-HR patch pairs;

Output: the trained random forest $\mathcal{T}$ with regressors $\mathcal{R} = (\mathcal{R}_1, \dots)$, the LSH model: $\mathcal{M}_{LSH}$;

1: Upscale the input LR patches by using IBP for the initial coarse estimations; ⇒ {Eqns. (17, 18)}

2: Obtain the features calculated from LR image by using the first and second-order gradient filters, and the gradient magnitudes on the up-scaled coarse versions; ⇒ {Eqns. (3, 4)}

3: Conduct LSH on the raw features to obtain compressed features, at the same time obtain the trained LSH projection model $\mathcal{M}_{LSH}$;

4: Train a random forest with the compressed features via the LSH model $\mathcal{M}_{LSH}$;

5: Train the weighted ridge regressors $\mathcal{R}$ by the GWRR models in leaf nodes; ⇒ {Eqn. (13)}

6: Save the random forest $\mathcal{T}$ with ridge regressors $\mathcal{R}$, and the trained LSH model $\mathcal{M}_{LSH}$.

Algorithm 2: Inference Stage of FARF-based Image Super-Resolution:

Input: testing LR image $\mathcal{Y}$, the trained random forest $\mathcal{T}$ with ridge regressors $\mathcal{R} = (\mathcal{R}_1, \dots)$, the trained LSH model $\mathcal{M}_{LSH}$;

Output: super-solved image $\widehat{\mathcal{X}}$;

1: Upscale patches from LR $\mathcal{Y}$ to form an initial coarse estimation by IBP; ⇒ {Eqn. (17, 18)}

2: Compute the discriminative features for all the patches; ⇒ {Eqns. (3, 4)}

3: Compute the compressed feature with the LSH model $\mathcal{M}_{LSH}$;

4: For each patch, using the compressed feature to search in the leaf nodes to obtain its corresponding regressor from the trained random forest $\mathcal{T}$;

5: Compute the super-resolved image $\widehat{\mathcal{X}}$ through all the super-solved patches by weighted ridge regressors $\mathcal{R}$ in leaf nodes. ⇒ {Eqn. (10)}

4. EXPERIMENTS

In this section, we evaluate our algorithm on standard super-resolution benchmarks Set 5, Set14 and B100 [20], and compare it with some state-of-the-art methods. They are *bicubic* interpolation, adjusted anchored neighborhood regression (A+) [5], standard RF [8], alternating regression forests (ARF) [8], the convolutional neural-network-based image super-resolution (SRCNN) [37, 38], the naive Bayes super-resolution forest (NBSRF) [19] and the sparse coding based network (SCN) [7], as listed in Table-2 and Table-3. We set the same parameters for all the RF-based algorithms, i.e., the number of trees in an RF is 10, and the maximum depth of each tree is 15. We use the same set of training images (91 images) for all the algorithms, as previous works [2, 4, 5, 8] do. RF+ means a normal RF-based algorithm added with the two *gradient magnitudes* augmented features, and RF# is the normal RF-based algorithm, where the original raw features, instead of using the PCA compressed features, are used to learn the regressors in leaf nodes. FARF$^-$ denotes trimmed version of our proposed feature-augmented RF scheme, which combines RF+ and RF# by adding the gradient magnitude features and using the original raw features for regression. FARF means the normal version or our proposed algorithm, which is based on FARF$^-$, the superior, unsupervised LSH projection [32], instead of PCA, is employed for dimensionality reduction, and the generalized weighted ridge regression (GWRR) model is used for the leaf-nodes'

regressors. FARF* is a further refined version of FARF, by performing further fine-tuning: (1) employing IBP, instead of the traditional *bicubic* interpolation algorithm, to obtain the initial coarse estimation in the inference stage, and (2) setting the proper number of trees (e.g., 45) for training an RF.

dataset	#	*bicubic*	A+	RF	ARF	RF^{+}	RF$^{\#}$
Set5	×2	33.66/0.00	36.55/0.51	36.52/0.03	36.65/0.82	36.67/0.04	36.63/0.05
	×3	30.39/0.00	32.59/0.33	32.44/0.04	32.53/0.62	32.56/0.05	32.53/0.05
	×4	28.42/0.00	30.29/0.23	30.10/0.05	30.17/0.71	30.18/0.06	30.22/0.06
Set14	×2	30.23/0.00	32.28/1.11	32.26/0.05	32.33/1.43	32.37/0.06	32.32/0.07
	×3	27.54/0.00	29.13/0.66	29.04/0.07	29.10/1.12	29.17/0.08	29.11/0.09
	×4	26.00/0.00	27.33/0.47	27.22/0.08	27.28/0.91	27.31/0.09	27.29/0.09
B100	×2	29.32/0.00	30.78/0.00	31.13/0.06	31.21/1.52	31.22/0.07	31.23/0.07
	×3	27.15/0.00	28.18/0.00	28.21/0.07	28.26/1.23	28.27/0.08	28.26/0.08
	×4	25.92/0.00	26.77/0.00	26.74/0.07	26.77/1.02	26.78/0.09	26.79/0.09

dataset	#	FARF^{-}	FARF	FARF*	SRCNN	NBSRF	SCN
Set5	×2	36.68/0.06	36.78/1.03	**36.81**/3.02	36.66/3.21	36.76/0.04	36.58/5.12
	×3	32.62/0.06	32.73/1.12	**32.78**/3.13	32.75/3.36	32.75/0.05	32.68/5.31
	×4	30.27/0.07	30.39/1.20	30.45/3.18	**30.48**/3.11	30.44/0.06	30.41/5.36
Set14	×2	32.37/0.08	32.41/1.15	**32.45**/3.23	32.42/6.52	**32.45**/0.06	32.35/6.02
	×3	29.17/1.02	29.23/1.24	**29.29**/3.27	29.28/6.41	29.25/0.08	29.16/6.13
	×4	27.36/1.10	27.45/1.31	27.48/3.35	**27.49**/6.38	27.42/0.09	27.39/6.21
B100	×2	31.34/0.92	31.35/1.22	**31.38**/3.28	31.36/8.24	31.36/0.09	**31.38**/6.16
	×3	28.30/1.05	28.35/1.25	28.38/3.32	**28.41**/8.37	28.39/1.02	28.33/6.30
	×4	26.83/1.20	26.88/1.41	**26.91**/3.46	26.90/8.42	26.89/1.13	26.88/6.51

Table-2: Results of the proposed method compared with state-of-the-art works on 3 datasets in terms of PSNR (dB) and runtime (seconds) using three different magnification factors (#) (×2, ×3, ×4).

Table-2 summarizes the performances of our proposed algorithms on the 3 datasets, in terms of the average peak signal to noise ratio (PSNR) scores and runtime (seconds), with different magnification factors (×2, ×3, ×4), and different patch sizes (6×6, 9×9, 12×12). Table-3 gives more details of the results on some images from the Set5 dataset, with magnification factor ×3. As the results have shown based on the 3 datasets, our proposed algorithm FARF outperforms A+ and ARF for all the magnification factors.

Set5(×3)	*bicubic*	Zeyde	A+	RF	ARF	FARF^{-}	FARF	FARF*	SRCNN	NBSRF	SCN
baby	33.91	35.13	35.23	35.25	35.15	35.20	35.34	**35.37**	35.25	35.28	35.25
bird	32.58	34.62	35.53	35.23	35.31	35.39	35.53	**35.54**	35.47	35.49	35.42
butterfly	24.04	25.93	27.13	27.00	27.39	27.65	27.68	27.82	**27.95**	27.87	27.78
head	32.88	33.61	33.82	33.73	33.73	33.75	33.84	**33.85**	33.71	33.78	33.65
woman	28.56	30.32	31.24	30.98	31.08	31.11	31.27	31.34	**31.37**	31.33	31.28
average	30.39	31.92	32.59	32.44	32.53	32.62	32.73	**32.78**	32.75	32.75	32.68

Table-3: Results of the proposed method compared with state-of-the-arts methods on 3 datasets in terms of PSNR (dB) with magnification factors (×3) on dataset Set5.

The objective quality metric, PSNR, in Table-2 also shows that the fine-tuned FARF, i.e. FARF*, can further improve the image quality, which is comparable to recently proposed state-of-the-art deep-learning-based algorithms, such as SRCNN [37, 38] and SCN [7]. Comparing our proposed FARF algorithm to other methods, the improved visual quality of our results is obvious, as shown in Fig. 7. This shows that our method can produce more details, particularly on some texture-rich regions.

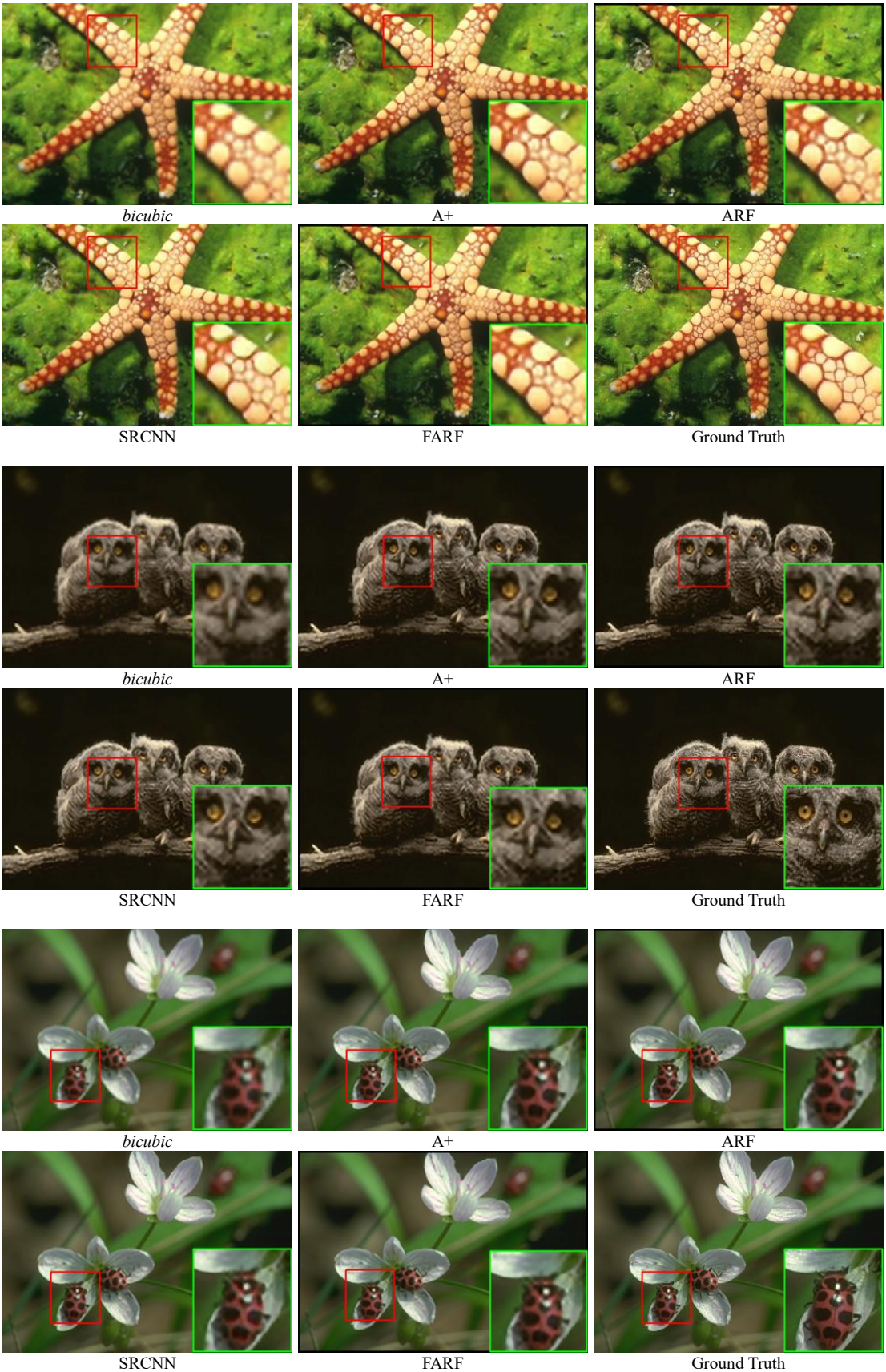


Fig. 7: Super-resolved (×3) images from B100, *bicubic*, A+ (ACCV-2014) [5], ARF (CVPR-2015) [8], SRCNN (PAMI-2016) [38], our proposed FARF algorithm, and ground truth. The results show that our FARF algorithm can produce more details and its performance is comparable to a recent state-of-the-art deep-learning method [38].

5. CONCLUSIONS

This paper presents a feature-augmented random forest (FARF) scheme for the single-image super-resolution (SISR) task by augmenting features and redesigning the inner structure of a random forest (RF), with different feature recipes used at different stages, where the compressed features are used for clustering in the split nodes and the original features are used for regression in the leaf nodes. The contributions of this paper are threefold: (1) the more discriminative gradient magnitude-based augmented features are proposed for clustering in split nodes and regression in leaf nodes; (2) By replacing principal component analysis (PCA) with a generalized unsupervised locality-sensitive hashing (LSH) model for dimensionality reduction, we lay out an original-compressed coupled feature set for tackling the clustering-regression tasks, and unify SISR and content-based image retrieval for LSH performance evaluation; and (3) we have extended the weighted collaborative representation model to a generalized weighted ridge regression model for ridge regression. The proposed FARF scheme can achieve highly competitive quality results, e.g., obtaining about a 0.3dB gain in PSNR, on average, when compared to the conventional RF-based super-resolution approaches. Furthermore, a fine-tuned version of our proposed FARF approach is provided, whose performance is comparable to recent state-of-the-art deep-learning-based algorithms.

References

[1] H. Li and K.-M. Lam, "Fast super-resolution based on weighted collaborative representation," in IEEE the 19th International Conference on Digital Signal Processing (DSP),2014, pp. 914-918.

[2] R. Timofte, V. De Smet, and L. Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2013, pp. 1920-1927.

[3] Y. Amit and D. Geman, "Shape quantization and recognition with randomized trees," Neural computation, vol. 9, no. 7, pp. 1545-1588, 1997.

[4] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in International conference on curves and surfaces, 2010, pp. 711-730: Springer.

[5] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in Asian Conference on Computer Vision (ACCV), 2014, pp. 111-126: Springer.

[6] P. Indyk and R. Motwani, "Approximate nearest neighbors: towards removing the curse of dimensionality," in Proceedings of the thirtieth annual ACM symposium on Theory of computing, 1998, pp. 604-613: ACM.

[7] Z. Wang, D. Liu, J. Yang, W. Han, and T. Huang. "Deep networks for image super-resolution with sparse prior," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015.

[8] S. Schulter, C. Leistner, and H. Bischof, "Fast and accurate image upscaling with super-resolution forests," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3791-3799.

[9] J. Jiang, X. Ma, C. Chen, T. Lu, Z. Wang, and J. Ma, "Single Image Super-Resolution via Locally Regularized Anchored Neighborhood Regression and Nonlocal Means," IEEE Transactions on Multimedia, vol. 19, no. 1, pp. 15-26, 2017.

[10] A. Gionis, P. Indyk, R. Motwani, et al. Similarity search in high dimensions via hashing. In VLDB, volume 99, pages 518–529, 1999.

[11] W. Dong, L. Zhang, G. Shi, and X. Wu, "Nonlocal back projection for adaptive image enlargement", IEEE International Conference on Image Processing (ICIP), pp. 349-352, November 2009.

[12] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin, "Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 12, pp. 2916-2929, 2013.

[13] H. Liu, R. Wang, S. Shan, and X. Chen, "Deep supervised hashing for fast image retrieval," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2064-2072.

[14] L. Breiman, "Random forests," Machine learning, vol. 45, no. 1, pp. 5-32, 2001.

[15] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1867-1874.

[16] A. N. Tikhonov, V. I. A. k. Arsenin, and F. John, "Solutions of ill-posed problems." Winston Washington, DC, 1977.

[17] F. Moosmann, B. Triggs, and F. Jurie, "Fast discriminative visual codebooks using randomized clustering forests," in Conference on Neural Information Processing Systems (NIPS), 2006, vol. 2, p. 4.

[18] H. Li, K.-M. Lam, M.-Y. Chiu, K. Wu, and Z. Lei, "Cascaded face alignment via intimacy definition feature," Journal of Electronic Imaging (JEI), vol. 26, no. 5, p. 053024, 2017.

[19] J. Salvador and E. Pérez-Pellitero, "Naive bayes super-resolution forest," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015, pp. 325-333.

[20] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in Proceedings of Eighth IEEE International Conference on Computer Vision (ICCV), 2001, vol. 2, pp. 416-423.

[21] S. Ren, X. Cao, Y. Wei, and J. Sun, "Face alignment at 3000 fps via regressing local binary features," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1685-1692.

[22] P. Li, X. Lu, and Q. Wang, "From dictionary of visual words to subspaces: locality-constrained affine subspace coding," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 2348-2357.

[23] H. Li and K.-M. Lam, "Guided iterative back-projection scheme for single-image super-resolution," in IEEE Conference on Global High Tech Congress on Electronics (GHTCE), 2013, pp. 175-180.

[24] I. Ram, M. Elad, and I. Cohen, "Image processing using smooth ordering of its patches," IEEE Trans. on Image Processing (TIP), vol. 22, no. 7, pp. 2764-2774, 2013.

[25] A. Danielyan, V. Katkovnik, and K. Egiazarian, "BM3D frames and variational image deblurring," IEEE Trans. on Image Processing (TIP), vol. 21, no. 4, pp. 1715-1728, 2012.

[26] M. Irani and S. Peleg. "Improving resolution by image registration, "CVGIP: Graphical Models and Image Processing, 53(3), 1991

[27] J. Gall and V. Lempitsky, "Class-specific hough forests for object detection, " in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[28] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," IEEE Transactions on Image Processing (TIP), vol. 19, no. 11, pp. 2861-2873, 2010.

[29] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-Resolution through neighbor embedding," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 275-282, 2004.

[30] M. Bevilacqua, A. Roumy, C. Guillemot, and M. L. Alberi-Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in the British Machine Vision Conference (BMVC), 2012.

[31] C.-Y. Yang and M.-H. Yang, "Fast direct super-resolution by simple functions," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 561-568, 2013.

[32] H. Li, K.-M. Lam, and D. Li, "Joint Maximum Purity Forest with Application to Image Super-Resolution," Journal of Electronic imaging (JEI) 27(4), 043005 (2018), doi: 10.1117/1.JEI.27.4.043005.

[33] Y. Zhang, Y. Zhang, J. Zhang, H. Wang, X. Wang, and Q. Dai, "Adaptive local nonparametric regression for fast single image super-resolution," in Visual Communications and Image Processing (VCIP), 2015, 2015, pp. 1-4: IEEE.

[34] J. Wang, J. Yang, K. Yu, F. Lv, T. S. Huang, and Y. Gong. "Locality-constrained linear coding for image classification," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.

[35] Y. Zhang, K. Gu, Y. Zhang, J. Zhang, and Q. Dai, "Image super-resolution based on dictionary learning and anchored neighborhood regression with mutual incoherence," in IEEE International Conference on Image Processing (ICIP), 2015, Québec City, QC, Canada, Sep. 2015, pp. 591–595.

[36] Y. Zhang, Y. Zhang, J. Zhang, and Q. Dai, "CCR: Clustering and Collaborative Representation for Fast Single Image Super-Resolution," IEEE Transactions on Multimedia (TMM), vol. 18, no. 3, pp. 405-417, 2016.

[37] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in European Conference on Computer Vision (ECCV), 2014, pp. 184-199: Springer.

[38] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," IEEE transactions on pattern analysis and machine intelligence (TPAMI), vol. 38, no. 2, pp. 295-307, 2016.

[39] J. Kim, J. Kwon Lee, and K. Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1646-1654.

[40] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2017.

[41] X. Li and M. T. Orchard, "New edge-directed interpolation," IEEE Transactions on Image Processing (TIP), vol. 10, no. 10, pp. 1521-1527, 2001.

[42] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation," IEEE Transactions on Image Processing (TIP), vol. 17, no. 6, pp. 887-896, 2008.

[43] L. Zhang and X. Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion," IEEE Transactions on Image Processing (TIP), vol. 15, no. 8, pp. 2226-2238, 2006.

[44] W.-S. Tam, C.-W. Kok, and W.-C. Siu, "Modified edge-directed interpolation for images," Journal of Electronic imaging (JEI), vol. 19, no. 1, pp. 013011, 2010.

[45] W. Freeman, T. Jones, and E. Pasztor, "Example-based super-resolution, " IEEE Computer Graphics and Applications (CG&A), vol. 22, no. 2, 2002.

[46] K. Lin, H.-F. Yang, J.-H. Hsiao, and C.-S. Chen. "Deep learning of binary hash codes for fast image retrieval," In Computer Vision and Pattern Recognition Workshops (CVPRW), pages 27–35, 2015.

[47] P. Indyk, R. Motwani, P. Raghavan, and S. Vempala, "Locality-preserving hashing in multidimensional spaces," in Proceedings of the twenty-ninth annual ACM symposium on Theory of computing, 1997, pp. 618-625: ACM.