Fast super-resolution algorithm using rotation-invariant ELBP classifier and hierarchical pattern matching

Dong Yoon Choi, Byung Cheol Song *

Department of Electronic Engineering, Inha University, 100 Inha-Ro, Nam-Gu, Incheon 22212, Republic of Korea

ABSTRACT

This paper proposes a fast super-resolution (SR) algorithm using content-adaptive two-dimensional (2D) finite impulse response (FIR) filters based on a rotation-invariant classifier. The proposed algorithm consists of a learning stage and an inference stage. In the learning stage, we cluster a sufficient number of low-resolution (LR) and high-resolution (HR) patch pairs into a specific number of groups using the rotation-invariant classifier, and choose a specific number of dominant clusters. Then, we compute the optimal 2D FIR filter(s) to synthesize a high-quality HR patch from an LR patch per cluster, and finally store the patch-adaptive 2D FIR filters in a dictionary. Also, we present a smart hierarchical addressing method for effective dictionary exploration in the inference stage. In the inference stage, the ELBP of each input LR patch is extracted in the same way as the learning stage, and the best matched FIR filter(s) to the input LR patch is found from the dictionary by the hierarchical addressing. Finally, we synthesize the HR patch by using the optimal 2D FIR filter. The experimental results show that the proposed algorithm produces better HR images than the existing SR methods, while providing fast running time.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

As digital displays become larger and larger in size and higher and higher in resolution, the need for full HD (FHD) and ultra HD (UHD) television (TV) is rapidly increasing [1]. With this trend, a sufficient amount of FHD/UHD content is required, but that content is hard to find, in practice. So, high-performance up-scaling to convert low-resolution (LR) legacy images into high-resolution (HR) images should be developed as an essential technique for FHD/UHD devices.

This up-scaling is a research issue that has been dealt with for a long time. Nearest neighbor interpolation, bi-linear interpolation, and bi-cubic interpolation [2] are the most popular methods for up-scaling. However, the above-mentioned linear interpolation algorithms cause a blur phenomenon due to low definition and various artifacts. In order to remove or mitigate such phenomena, slightly better methods such as the Lanczos filter based on an ideal sinc function and edge-directed interpolation [3] (taking into account edge direction) have been presented. These algorithms are easily implementable because of their simple structure, but they are not acceptable in terms of subjective visual quality,

because they cannot reconstruct the high-frequency (HF) signals lost in edge or texture areas, and may cause jagging artifacts.

In order to overcome the limitations of the above-mentioned interpolation methods, so-called super-resolution (SR) algorithms have been developed as an advanced up-scaling approach. In general, SR can be categorized into single-image SR and multiple-image SR. Typical multiple-image SR algorithms [4–6] require extensive computation in the registration step to estimate sub-pixel motion while positioning LR images on the HR grid. However, the accuracy of this registration is still insufficient to justify the computational complexity. Consequently, we do not know of any multiple-image SR algorithms being successfully implemented and embedded in cutting-edge digital devices.

On the other hand, single-image SR algorithms are usually block-based or patch-based. Such patch-based SR methods are divided into example-based methods [7–13,22] and content-adaptive filter-based SR algorithms [14–18,21,23]. First, example-based SR algorithms usually store an abundant number of LR-HR patch pairs in an off-line dictionary, and predict HF signals lost in an arbitrary input LR image by exploring the dictionary. As a classic example of external example-based methods [7–9], Freeman et al. employed Markov-model–based belief propagation to estimate the lost HF signals [7]. Freeman et al.'s algorithm is meaningful in that it is the first framework of the conventional example-based SRs. However, the algorithm is computationally heavy because it requires as many examples as possible to achieve

* Corresponding author.
  E-mail address: bcsong@inha.ac.kr (B. Cheol Song).

acceptable performance. In order to reduce computational burden, Chang et al. proposed a SR algorithm using nearest neighbor embedding where several nearest neighbor patches close to the input LR patch are explored, and the lost HF signals are estimated by finding the correlation between the input LR and its neighbors via Gram matrix [8]. Yang et al. generated a compact dictionary for efficient example-based SR by adopting sparse signal representation [9]. However, the above-mentioned external example-based SR algorithms still require a huge dictionary to achieve acceptable reconstruction quality.

Approaches to overcome such a drawback are called internal example-based methods [10,11]. Freedman and Fattal [11] proposed a self-example SR algorithm with an advantage that a pre-trained dictionary is not required. Inherent self-examples are explored from an input LR image by examining self-similarity between the LR image and its up-scaled images. Consequently, they provide outstanding visual quality for edge areas, because edges usually have good self-similarity. On the other hand, the conventional self-example–based SR algorithms do not provide acceptable visual quality for some texture areas [11]. So, joint methods which choose and combine merits of external example-based SR and internal example-based SR have been developed [12,13]. Wang et al. computed adaptive weighting by minimizing two loss functions which were derived from sparse coding based external examples and epitomic matching based internal examples, respectively, and synthesized the lost HF signals [13].

As another single-image SR approach, content-adaptive filter-based SR defines mapping function, i.e., finite impulse response (FIR) filter optimized for every possible LR patch in the learning stage, and reconstructs HR image by using the best matched FIR filter to each LR patch in the inference stage. Because of its merit of high processing speed, content-adaptive filter-based SR has received much attention recently [14–17,21,23]. Yang et al. proposed a typical filter-based SR where a number of LR-HR patch pairs are extracted from various natural images. LR-HR patch pairs are partitioned by LR-patch-based clustering, and mapping function(s) per cluster for transforming an LR patch to a HR patch is trained in the off-line learning stage [14]. Those mapping functions stored in a dictionary are used for on-line HF synthesis. Timofte et al. presented the anchored neighborhood regression (ANR) approach, which used ridge regression to learn exemplar neighborhoods off-line, and uses these neighborhoods to pre-compute projections to map LR patches onto the HR domain [15,16]. As a result, they achieved fast execution while retaining the qualitative performance of recent state-of-the-art methods. Dong et al. first employed a famous deep learning method, i.e., convolutional neural network (CNN) to solve this SR problem. Each convolution layer was modelled as a step of patch-based SR, and convolutional filters to reconstruct HR patches from LR patches were learned by using many training data [17]. The above-mentioned SR algorithms provide better visual quality than conventional SR algorithms, but they still suffer from some artifacts such as ringing near edges.

Kondo and Kawaguchi [18] proposed a content-adaptive SR algorithm using a pre-determined number of FIR filters which are trained during an offline learning step. In the learning step, LR-HR patch pairs are grouped into a specific number of clusters by using a simple classifier called adaptive dynamic range coding (ADRC), and the best 2D FIR filter per cluster is computed. In the on-the-fly inference step, the ADRC of each LR block is computed, and the 2D FIR filter corresponding to the ADRC is selected in the dictionary. Using the selected 2D FIR filters, an HR image is synthesized on a patch basis. The merit of such an approach is its low computational cost because the matching process is not required for the inference step. However, Kondo and Kawaguchi's method may suffer from halo artifacts during reconstruction because the classification performance of ADRC is very limited.

On the other hand, ADRC is a sort of local binary pattern (LBP). LBP, which was first introduced to image texture description for texture analysis, has recently been used for face description by adopting the region division and concatenation histogram strategy [19,20]. Instead of the original LBP, extended LBP (ELBP) not only extracts the relative gray value difference between the central pixel and its neighbors, as provided by LBP, but also focuses on their absolute differences, which are also critical to describe local shapes [20]. Since ELBP is usually superior to ADRC, we can achieve acceptable visual quality as well as high speed if we employ it as a classification tool for SR.

In this paper, we present a rotation-invariant ELBP classifier, and propose a fast learning-based super-resolution algorithm using the ELBP classifier and hierarchical pattern matching. This paper has several contribution points. First, we present the ELBP for the rotation-invariant classification. The benefit of the rotation-invariant ELBP is that similar ELBPs with respect to rotation can be grouped together and can share a common 2D FIR filter; hence, the memory size of the dictionary is reduced. Second, we dramatically reduce the total number of clusters by using a statistical characteristic of ELBP patterns. Note that the ELBP may exponentially increase the number of clusters in a dictionary as its bit-length increases. Fortunately, since most of the ELBPs do not exist or have a negligible frequency of occurrence in practice, we can choose a much smaller number of dominant clusters. Third, the matching process in the inference stage is unnecessary. This is because the ELBP itself can be the address of the corresponding 2D FIR filter in the dictionary. This makes a real-time SR process possible. Finally, we present a hierarchical addressing method which can rapidly find the right address of an input LR patch in the dictionary with minimum memory cost.

The proposed algorithm consists of the learning stage and the inference stage. First, a sufficient number of LR-HR patch pairs are extracted from the training-purpose LR-HR image pairs, and the rotation-invariant ELBP for each LR patch is derived. Next, LR-HR patch pairs having the same ELBP are grouped into a single cluster, and a specific number of dominant clusters are chosen. Then, the best 2D FIR filter for each selected cluster is computed by using the LR-HR patch pairs in the cluster. All the 2D FIR filters are hierarchically indexed according to their ELBPs and stored in a dictionary. In the inference stage, the rotation-invariant ELBP for each input LR patch is first computed. Second, a new address for the ELBP is assigned by the proposed addressing method, and the best-matched 2D FIR filter to the input LR patch is found at the address in the dictionary. Also, the rotation angle corresponding to the ELBP is computed, and the selected 2D FIR filter may be reversely rotated for the computed angle if necessary. Finally, the HR patch is synthesized using the rotated 2D FIR filter. The experimental results show that the proposed algorithm outperforms the previous SR algorithms, while having significantly faster processing speed than the previous ones.

This paper is organized as follows. LBP is briefly reviewed in Section 2, and the proposed algorithm is described in detail in Section 3. Section 4 presents the intensive experimental results. Finally, Section 5 concludes this paper.

2. Review of local binary pattern

LBP is briefly described in this section, because it is a basic classifier for the proposed SR algorithm. LBP, first proposed by Ojala [19], is a typical pattern classifier having light complexity and is widely used for various applications requiring local coding. LBP compares a center pixel X_C with its neighbor pixels, as seen in Fig. 1, and is obtained by binary decision, as expressed in Eq. (1).