J. Vis. Commun. Image R. 31 (2015) 305–311

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Short Communication

Learning-based image interpolation via robust k-NN searching for coherent AR parameters estimation ☆

Kwok-Wai Hung *, Wan-Chi Siu

Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong

ARTICLE INFO

Article history:
Received 9 September 2014
Accepted 5 July 2015
Available online 13 July 2015

Keywords:
Image interpolation
Autoregressive model
Learning-based algorithms
Super-resolution
Soft-decision estimation
K-means clustering
Weighted least squares
Robust estimation

ABSTRACT

Image interpolation is to convert a low-resolution (LR) image into a high-resolution (HR) image through mathematical modeling. An accurate model usually leads to a better reconstruction quality, and the autoregressive (AR) model is a widely adopted model for image interpolation. Although a large amount of works have been done on AR models for image interpolation, there are plenty of rooms for improvements. In this work, we propose a robust and precise k-nearest neighbors (k-NN) searching scheme to form an accurate AR model of the local statistic. We make use of both LR and HR information obtained from a large amount of training data, in order to form a coherent soft-decision estimation of both AR parameters and high-resolution pixels. Experimental results show that the proposed learning-based AR interpolation algorithm has a very competitive performance compared with the state-of-the-art image interpolation algorithms in terms of PSNR and SSIM values.

1. Introduction

IMAGE interpolation is a classic problem in various image processing applications [1–14]. This work proposes to use the autoregressive (AR) model, which has been widely adopted in the literature [4–13]. However, in previous approaches AR parameters were learnt online from the same input low-resolution (LR) image [4–12] or were pre-computed from a training data set through an offline process [13]. On the contrary, the proposed AR parameter estimation is done online in a much more precise way by searching the k-nearest neighbors (k-NNs) from a pre-stored training set, whereas the adaptive k-NN criterion [14] is adopted to adaptively select the number of k-NNs according to the amount of relevant training data, which can make the k-NN searching scheme robust to insufficient matches (outliers) and preserve the details of relevant matches. Using the LR and HR information from k-NNs, a coherent soft-decision process [5] is applied to estimate both AR parameters and high-resolution pixels. To our best knowledge, the proposed algorithm is the first to make use of the precise k-NN searching to form a coherent soft-decision estimation of both AR model parameters and HR pixels during the online estimation in the image processing tasks in the literature.

The rest discussions of the paper are as follows: Section 2 gives the background of the AR-based image interpolation [4–9], Section 3 explains the overall scheme, including the precise and robust k-NN searching, the coherent soft-decision estimation and the refinement process, Section 4 shows our experimental results and Section 5 concludes the paper.

2. Background

Let us briefly review the concept of image interpolation using local autoregressive models [4–9] and soft-decision estimations [5,6,8]. Fig. 1 shows a local region of a high-resolution (HR) image to be interpolated. Let us define the low-resolution (LR) patch as $X_S = \{x_i | i \in S\}$ and the HR patch as $Y_S = \{y_i | i \in S\}$, where S is the common support of both patches. Let us denote the diagonal neighbors of LR pixels as $\{y_{i(t)} | i(t) \in S\}$ and the diagonal neighbors of HR pixels as $\{x_{i(t)} | i(t) \in S\}$, as indicated by green arrows in Fig. 1. Any LR and HR pixel and its diagonal neighbors are related by the common model parameters, $\{A_{(t)}\}$,

$$x_i \approx \sum_t A_{(t)} y_{i(t)} \tag{1}$$

$$y_i \approx \sum_t A_{(t)} x_{i(t)} \tag{2}$$

where $\{y_{i(t)}\}$ are diagonal neighbors of x_i and $\{x_{i(t)}\}$ are diagonal neighbors of y_i, and t is the index of the model. "Geometric duality"

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author at: DE503, EIE, PolyU, Hung Hom, Hong Kong.
E-mail address: kwokwai.hung@connect.polyu.hk (K.-W. Hung).

http://dx.doi.org/10.1016/j.jvcir.2015.07.006
1047-3203/

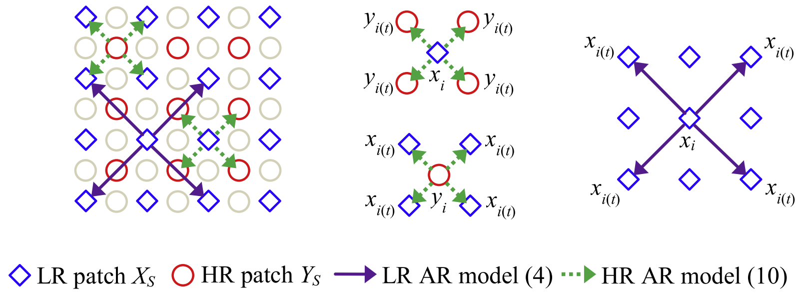

Fig. 1. Examples of LR pixels and HR pixels and corresponding AR models.

property [4] shows that the edge statistics modeled by the AR system (1) and (2) are generally scale-invariant. Thus, the diagonal neighbors of LR pixel, $\{x_{i(t)}\}$, can be replaced by nearest LR neighbors, $\{x_{i(t)}|i(t) \in S\}$, as indicated by purple arrows in Fig. 1,

$$x_i \approx \sum_t A_{(t)} x_{i(t)} \tag{3}$$

where $\{x_{i(t)}\}$ are approximated diagonal neighbors of x_i. Using a number of approximating samples, the weighted least squares (WLS) can be applied to estimate the AR model parameters [6–9],

$$\{\hat{A}_{(t)}\} = \arg\min_{\{A_{(t)}\}} \sum_{i,i(t)\in S} W_i \left(x_i - \sum_t A_{(t)} x_{i(t)} \right)^2 \tag{4}$$

where $W = \{W_i\}$ are weights of the residuals, which can be modeled by an exponential form [6–9], such as $W_i = \exp(-E[x_i - \hat{y}_i]^2)$, where exp(.) is the exponential function and $\hat{y}_i$ is an approximated value of the HR pixel using simple interpolation.

By substituting the estimated model parameters into (2), the HR pixel can be interpolated by $y_i \approx \Sigma_t A_{(t)} x_{i(t)}$. Instead, the AR models in (1) and (2) can combine together to form an unified soft-decision optimization [5]

$$\{\hat{y}_i\} = \arg\min_{\{y_i\}} \left\{ \sum_{i,i(t)\in S} \left(y_i - \sum_t A_{(t)} x_{i(t)} \right)^2 + \sum_{i,i(t)\in S} \left(x_i - \sum_t A_{(t)} y_{i(t)} \right)^2 \right\} \tag{5}$$

Eq. (5) simultaneously solves for all HR pixels, in order to constrain the statistical consistency within a local region.

3. Proposed learning-based AR model parameter estimation

3.1. Overview of the proposed algorithm

In this paper, a learning-based scheme for estimating the AR model parameters using a precise k-NN searching and a coherent soft-decision estimation is proposed. The proposed approach estimates the AR model parameters by searching the k-nearest LR–HR training pairs from a pre-stored training set formed by 25 standard images, where the training set was clustered into 32 subsets by k-means clustering to speed up the searching time. Both LR and HR information from the k-nearest training pairs are used in the proposed approach. This is essentially different from previous AR model parameters estimation approaches which rely on the information from the same LR input image [4–12] or a pre-computed set of AR parameters [13].

Fig. 2 shows an overview of the proposed algorithm.

Algorithm 1: Learning-based image AR interpolation scheme

Input: LR patches $\{X_S\}$, training data $\{X_S^n\}$ and $\{Y_S^n\}$
Output: Estimated HR patches $\{Y_S\}$
(1) Initialization: For each X_S
(a) Use normalized correlation coefficient $\mathrm{NCC}(X_S, X_S^n)$ to find the k-NNs, $\{X_S^n\}_{n=1}^k$ and $\{Y_S^n\}_{n=1}^k$, by (6) and (7).
(b) Use the k-NNs to find the AR model parameters, $\{A_{(t)}\}$, by (10).
(c) Use AR model parameters to estimate the HR patch, Y_S, by (5).
(2) Refinement: For each $Z_S = \{X_S \cup \{Y_S\}\}$
Repeat 1(a)–(c) using normalized correlation coefficient $\mathrm{NCC}(Z_S, Z_S^n)$ to find the k-NNs, by (13) and (14).

3.2. Precise search for k-nearest LR–HR training pairs

Let us denote the training data set for the LR patch X_S as $\{X_S^n = \{x_i^n | i \in S\}\}$ and the training data set for the HR patch Y_S as $\{Y_S^n = \{y_i^n | i \in S\}\}$, where n is the patch index. Let us also illustrate the process for interpolating the HR patch $Y_S = \{y_i | i \in S\}$ in Fig. 1 during the online estimation. Initially, the input LR patch X_S is used to determine which subset to be used by measuring the Euclidean distance with all centroids of subsets.

The normalized correlation coefficient (NCC) [15] is then adopted as the searching criterion because it is a second-order statistics measurement, which is scale invariant and robust to small noises [15]. To search for the k-nearest LR–HR training pairs, the NCC of the LR patch X_S and all LR patches in the training set $\{X_S^n\}$ are measured,

$$\mathrm{NCC}(X_S, X_S^n) = \sum_{i\in S} (x_i - Mean(X_S))\left(x_i^n - Mean(X_S^n)\right) / \sigma_X \sigma_X^n \quad \text{for} \quad \forall n \tag{6}$$

where $Mean(.)$ is the mean operation, and σ_X and σ_X^n are standard derivations of patch X_S and patch X_S^n respectively. The measured NCC values are then sorted in a descending order, such that $\mathrm{NCC}(X_S, X_S^n) \geqslant \mathrm{NCC}\left(X_S, X_S^{n+1}\right)$ for $\forall n$. Then, k-nearest LR training patches $\{X_S^n\}_{n=1}^k$ which have the k-largest NCC values are chosen by the adaptive k criterion [14],

$$\hat{k} = \arg\min_k k \quad \text{s.t.} \sum_{n=[1,k]} \mathrm{NCC}(X_S, X_S^n) > T_1 \tag{7}$$

where T_1 is a threshold to control the number of k-nearest neighbors depending on the sum of largest correlations of training neighbors. Specifically, the largest NCC value, $\mathrm{NCC}\left(X_S, X_S^1\right)$, is summed