Video Super-Resolution by Adaptive Kernel Regression

Mohammad Moinul Islam, Vijayan K. Asari,
Mohammed Nazrul Islam, and Mohammad A. Karim

Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, VA
{misla001,vasari,mislam,mkarim}@odu.edu

Abstract. A novel approach for super-resolution using kernel regression technique is presented in this paper. The new algorithm uses several low resolution video frames to estimate unknown pixels in a high resolution frame using kernel regression employing adaptive Gaussian kernel. Experiments conducted on several video streams to evaluate the effect of the proposed algorithm showed improved performance when compared with other state of the art techniques. This resolution enhancement technique is simple and easy to implement and it can be used as a software alternative to obtain high quality and high resolution video streams from low resolution versions.

1 Introduction

Super-resolution is a process of image resolution enhancement by which low quality, low resolution (LR) images are used to generate a high quality, high resolution (HR) image [1]. There are numerous applications of super-resolution in the areas of image processing and computer vision such as long range target detection, recognition and tracking. It has many applications in consumer products such as cell phone, webcam, high-definition television (HDTV), closed circuit television (CCTV) etc.

There have been several techniques developed and presented in the literature for image super-resolution [2-7]. These can be classified into two categories: single image super-resolution and super-resolution from several frames. In the first case, there is no additional information available to enhance the resolution. So, the algorithms are based on processes of either image smoothing or interpolation. Interpolation techniques give better performance than the smoothing technique but both methods smooth edges and cause blurring problems [2]. Single image super-resolution based on machine learning techniques uses patches from training images and a tree based approximate nearest neighbor search is performed to find the desired HR patch from the LR patch [3]. Li *et al* [4] used a support vector regression (SVR) to find the mapping function between a low resolution patch and the central pixel of a HR patch. These methods are computationally expensive and require huge memory for handling the training data. The classical super-resolution restoration from several noisy images was proposed by Elad and Feuer [5], where the mathematical model of obtaining a super-resolution image from several LR images is described as:

$$Y_k = D_k C_k F_k X + E_k \quad \text{for} \quad 1 \le k \le N \tag{1}$$

G. Bebis et al. (Eds.): ISVC 2009, Part II, LNCS 5876, pp. 799–806, 2009.

In the above equation, $\{Y_k\}_{k=1}^{N}$ represents N images of size $M \times M$ each of which is a different representation of the desired HR image X. Matrix F_k represents the geometric warp performed on the image X, C_k is a linear space variant blur matrix and D_k is a matrix representing the decimation operator in Y_k. E_k is the additive zero mean Gaussian noise in the k-th LR frame.

Another approach to super-resolution is iterative back projection (IBP) similar to the projections in computer aided tomography (CAT) [6]. This method starts with an initial guess of the HR image and simulated to generate a set of LR images which are compared with the observed image to update the HR image. A modified IBP algorithm was presented in [7] based on elliptical weighted area (EWA) filter in modeling the spatially-variant point spread functions (PSF). The rest of the paper is organized as follows: In section 2, an overview of kernel regression technique is presented. Section 3 describes the proposed algorithm. Simulation results and performance analysis are discussed in section 4. Finally, concluding remarks are presented in section 5.

2 Kernel Regression Technique

Kernel regression analysis is a nonparametric regression method to estimate the value of an unknown function $f(x)$ at any given point based on the observations. For two dimensional cases, the regression model is:

$$Y_i = f(x_i) + \varepsilon_i, \quad i = 1,2,\ldots,N, \quad x_i = [x_{1i}, x_{2i}]^T \tag{2}$$

where $\{(x_i),\ i = 1,2,\ldots,N\}$ are the design points (pixel position), $\{Y_i, i = 1,2,\ldots,N\}$ are observations (pixel values) of the response variable Y, f is a regression function and $\{\varepsilon_i, i = 1,2,\ldots,N\}$ are independent identically distributed (i.i.d.) random errors, and N is the number of samples (number of frames). The generalization of kernel estimate $\hat{f}(x)$ is given by the following minimization problem [8].

$$\min_{p,q_1..q_l} \sum_{i=1}^{N} \left[Y_i - \left\{p + q_1(x_i - x) + \ldots + q_l(x_i - x)^l\right\}\right]^2 K\left(\frac{x_i - x}{h}\right) \tag{3}$$

where $K(.)$ is the kernel function with bandwidth h and l is a positive integer which determines the order of the kernel estimator. The above equation is solved to determine the unknown regression coefficients $p, q_1, q_2 \ldots, q_l$. In this paper, we used Gaussian kernel in the above regression equation. Solving equation (3) for $l = 0$ gives zero-order local constant kernel estimator also known as Nadaraya-Watson kernel estimator given as [8]:

$$f_{NW}(x) = \frac{\sum_{i=1}^{N} Y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^{N} K\left(\frac{x_i - x}{h}\right)} \tag{4}$$

Higher order kernel estimators can also be applied but with increasing complexity. One major problem of the ordinary kernel is that it is not dependent on the image characteristics. So, we used data dependent kernel which depends not only on the pixel positions but also on the intensity values of the samples. A general form of adaptive kernel is Bilateral kernel and for zero order case, above estimator is modified as [9],

$$f_{NW}(x)=\frac{\sum_{i=1}^{N} Y_i K\left(\frac{x_i - x}{h}\right) K\left(\frac{Y_i - Y}{h_r}\right)}{\sum_{i=1}^{N} K\left(\frac{x_i - x}{h}\right) K\left(\frac{Y_i - Y}{h_r}\right)} \tag{5}$$

where h_r is the intensity dependent smoothing scalar.

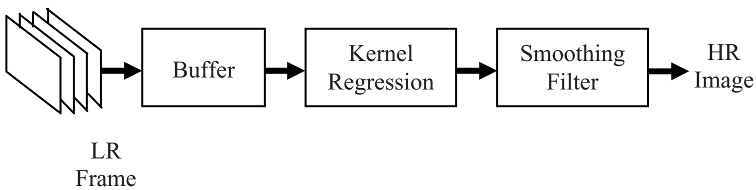

Fig. 1. Block diagram of kernel regression based video super-resolution

3 Proposed Algorithm

Figure 1 shows the basic block diagram of the proposed technique. As shown in the figure, low resolution frames are input to a buffer which stores *N* number of LR frames. Once the *N*-th frame is received kernel regression method is applied to construct an HR image. Given *N* low resolution video frames of size $m \times n$, the problem of super-resolution is to estimate high resolution frame of size $rm \times rn$, where *r* is the resolution enhancement factor. For simplicity, we assume $r= 2$. This means, a 2×2 LR grid is zoomed into a 4×4 HR grid. Hence for each 2×2 LR block, 12 additional pixel values need to be determined. In order to get super-resolution image with acceptable registration error we use four LR frames. Each LR frame can be considered as a downsampled, degraded version of the desired HR image. These LR frames contribute new information to interpolate sub-pixel values if there are relative motions from frame to frame [1]. Then the regression output is smoothed out with a smoothing filter. In Figure 2, four 2 × 2 LR data blocks are mapped into a single 4×4 HR grid. Subscripts in LR frames indicate frame numbers whereas in HR frame it indicates reconstructed pixel. Different symbols correspond to information from different 2×2 data blocks. As there are motions between successive blocks, pixels are geometrically shifted in the HR blocks which contribute sub-pixel information to construct a HR image.

The following assumptions are made regarding the video sequences:

1) Frame rates are not too small otherwise some motion information may be lost.
2) Distance between objects and camera are such that small change in camera motion doesn't cause significant displacement of the object regions.

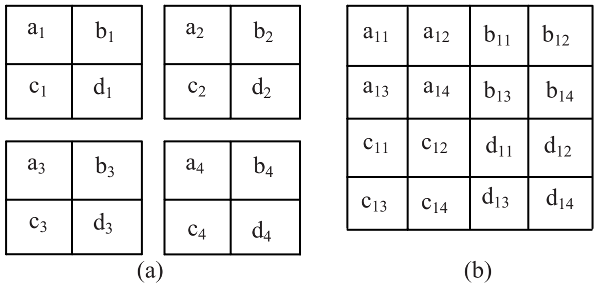

Fig. 2. (a) Four consecutive LR frames of size 2×2, (b) A 4×4 HR frame is constructed from the four 2×2 LR frames

Figure 3 depicts description of the proposed algorithm for resolution enhancement for a factor of 2. For each pixel in LR frame four pixels are generated by placing the kernel function at the desired LR pixel position and taking samples from four neighborhood positions. Small perturbation of the camera is allowed in order to estimate sub-pixel values. Thus pixel a_{11} is found by applying the kernel at pixel location a_1 with four observations in temporal direction of LR frames a_1, a_2, a_3 and a_4. Similarly, a_{12} is obtained by applying the kernel at pixel location a_1 but with four observations b_1, b_2, b_3 and b_4 of LR frames and so on.

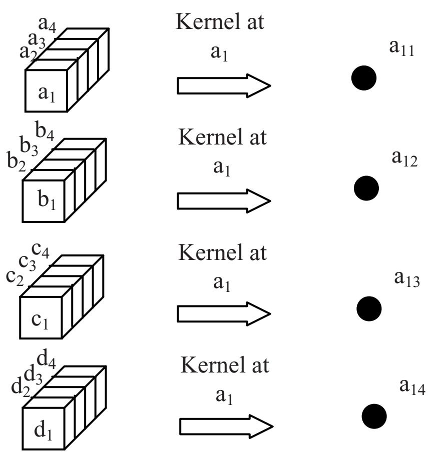

Fig. 3. Detailed diagram showing the construction of the above HR frame as shown in Fig. 2. For each position of the kernel, four pixels are generated taking four samples from the LR frames.

4 Simulation Results

The proposed algorithm is tested with several LR video sequences and their PSNR is measured as the performance index. Four consecutive LR frames are used to construct a HR frame with 2×2 scaling using our kernel regression method. Figure 4 shows the results of the super-resolution reconstruction of grayscale sequences by bi-cubic interpolation, iterative back projection (IBP), robust super-resolution and the proposed kernel regression method. Simulation results for IBP and robust super-resolution method are obtained by using the software available at [10]. It can be seen that the results obtained by IBP and the robust super-resolution method are noisy and the interpolated images are blurry, and the high resolution image created by our method is clear. Figure 5 shows the super-resolution recreation of a high resolution image with another set of gray scale images containing more textures. It can be seen that the proposed kernel regression method performs better than the other methods. Figure 6 shows the reconstruction of a color image.

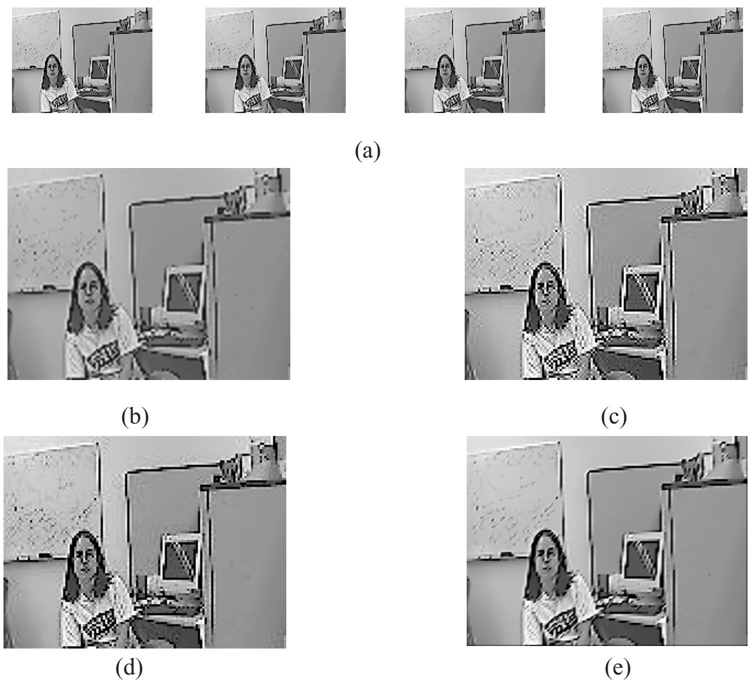

Fig. 4. (a) Four consecutive LR frames from Emily video [11], (b) HR image with a magnification factor 2 using bi-cubic interpolation, (c) Iterative back projection (IBP) (d) Robust super-resolution, and (e) Proposed kernel regression method

In order to measure the performance of the proposed technique, image frames are down-sampled and then reconstructed using the proposed algorithm to the original size. Since the original HR image is available, the restoration quality is measured by computing the PSNR as in [12].

$$\text{PSNR}= 10\log_{10}\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}255^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i,j)-\hat{f}(i,j)\right)^2} \tag{6}$$

where f is the original HR image and $\hat{f}$ is the reconstructed image. An image with higher PSNR means better reconstruction but it doesn't always represent the true quality of the image [12]. The PSNR results of the test videos are shown in Table 1. In all the cases, kernel regression method shows a higher PSNR value which supports the effectiveness of the algorithm. The algorithm performs well if the above mentioned assumptions hold true i.e. there is only a small motion from frame to frame. The main advantage of this algorithm is that it is simple to implement and hence computationally efficient.

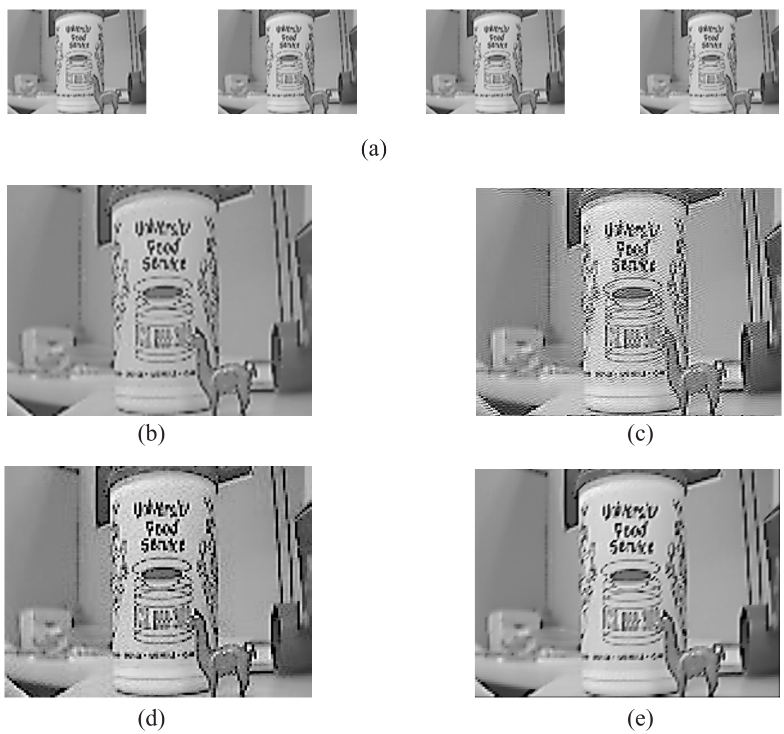

Fig. 5. (a) Four consecutive LR frames from Alpaca video [11], (b) HR image with a magnification factor 2 using bi-cubic interpolation, (c) Iterative back projection (IBP), (d) Robust super-resolution, and (e) Proposed kernel regression method

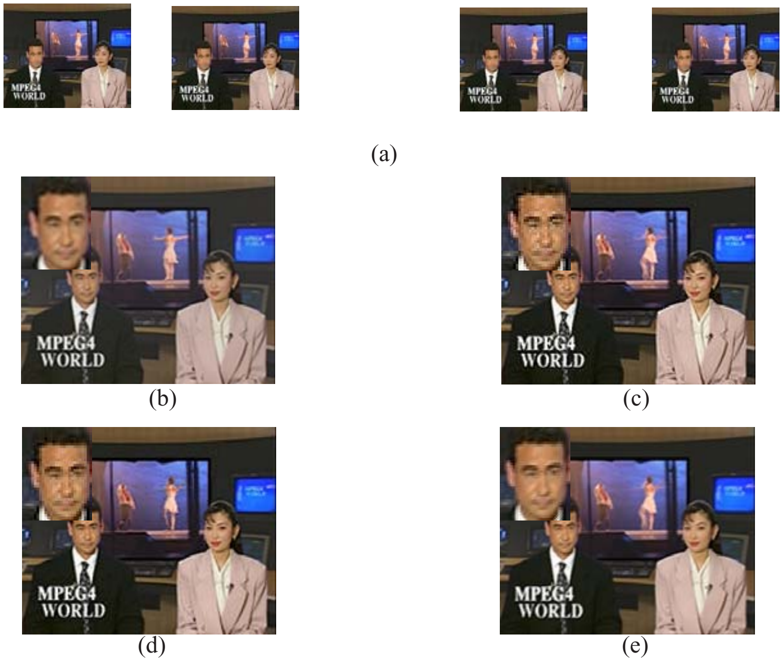

Fig. 6. (a) Four consecutive LR frames from news video [13], (b) HR image with a magnification factor 2 using bi-cubic interpolation, (c) Iterative back projection (IBP), (d) Robust super-resolution, and (e) Proposed kernel regression method

Table 1. Performance Comparison of the proposed super-resolution method

Image	Bi-cubic interpolation	IBP	Robust super-resolution	Proposed method
Emily	16.51	16.23	16.67	16.7
Alpaca	18.07	18.19	19.30	19.64
News	22.93	22.65	22.8	23.05

5 Conclusions

A new algorithm for image super-resolution has been presented in this paper. It uses a kernel regression technique in video frames to construct a high resolution frame. The main advantage of this method is that it doesn't need exhaustive motion estimations and hence it is simple in computations. However, this method is not suitable if there are significant motions between successive frames. Simulation results showed that the proposed algorithm performs comparably to other state of the art techniques.

References

[1] Nguyen, N., Milanfar, P.: A Computationally Efficient Superresolution Image Reconstruction Algorithm. IEEE Transactions on Image Processing 10, 573–583 (2001)

[2] Chang, H., Yeung, D.Y., Xiong, Y.: Super-resolution through Neighbor Embedding. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 275–282 (2004)

[3] Freeman, W.T., Jones, T.R., Pasztor, E.C.: Example-based Superresolution. IEEE Computer Graphics and Applications 22, 56–65 (2002)

[4] Li, D., Simske, S., Mersereau, R.M.: Single Image Super-resolution Based on Support Vector Regression. In: IEEE Proceedings of IJCNN, pp. 2898–2901 (2007)

[5] Elad, M., Feuer, A.: Restoration of a Single Superresolution Image from Several Blurred, Noisy and Undersampled Measured Images. IEEE Transactions on Image Processing 6, 1646–1658 (1997)

[6] Irani, M., Peleg, S.: Improving Resolution by Image Registration. CVGIP: Graphical Models and Image Processing 53, 231–239 (1991)

[7] Jiang, Z., Wong, T.T., Bao, H.: Practical Super-resolution from Dynamic Video Sequences. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 549–554 (2003)

[8] Qiu, P.: Image Processing and Jump Regression Analysis. Wiley Series in Probability and Statistics. Wiley, New Jersey (2005)

[9] Takeda, H., Farisu, S., Milanfar, P.: Kernel Regression for Image Processing and Reconstruction, vol. 16, pp. 349–366 (2007)

[10] Super-resolution software: `http://lcavwww.epfl.ch/software/superresolution/index.html`

[11] Super-resolution dataset: `http://www.soe.ucsc.edu/~milanfar/software/sr-datasets.html`

[12] Li, D., Mersereau, R.M., Simske, S.: Blind Image Deconvolution through Support Vector Regression. IEEE Transactions on Neural Networks 18, 931–935 (2007)

[13] Super-resolution dataset: `http://media.xiph.org/video/derf/`