# A Novel Interpolation Based Super-Resolution Of The Cropped Scene From A Video

A. John, Sakthivel Punniakodi, Ramesh Murukesan
*Dept., Information Technology, Christ College of Engineering and Technology*
*Puducherry, India.*

## Abstract

*Super resolution (SR) image reconstruction is the process of combing several low resolution images into a single high resolution image. The videos of the image change frame to frame. This paper is based on interpolation super-resolution method. An algorithm for enhancing the resolution of the scene through Segmentation of the video and cropping the required part of the scene, super-resolution using Interpolation, Regression, and Post-processing, is applied to the effective Super-resolution image output. Further object tracking and identification use the results of this work. We worked in traffic surveillance videos.*

***Keywords: super Resolution – Sequence of Image -Interpolation Method***

## 1. Introduction

Super Resolution of Image first appeared in the early 1980s, with one of the first papers in the signal processing community, the paper by Tsai and Huang [1]. Since then the topic has been active, with some of the early good results appearing in the 1990s. The last five or so years however have witnessed an enormous resurgence in SR activity. Any given set of source low resolution (LR) images only captures a finite amount of information from a scene; the goal of SR is to extract the independent information from each image in that set and combine the information into a single high resolution (HR) image. The LR images can come from a variety of sources: they can be taken from different frames of a video sequence, different still images taken from a single camera that has undergone translation or rotation, or multiple cameras capturing a single scene. The only requirement is that each LR image must contain some information that is unique to that image. This means that when these LR images are mapped onto a common reference plane their samples must be subpixel shifted from samples of other images – otherwise the images would contain only redundant information and SR reconstruction would not be possible. Various types of super resolution methods are available [2], but in this paper we applied interpolation based super-resolution method. We propose a novel super-resolution of the cropped scene from a video.



### 1.1. Converting video into frames

The first process of our work is to "Converting video into frames". Converting video into frames is the process of reading each and every frame in a sequence of image and saving it. Consider a sample video of "avi" format. Let us consider a sample video be "A" (A indicate the sample video of .avi format). The image sequence is mathematically mentioned below,

$$A = \sum_{i=0}^{n} \alpha_i$$

Where $\alpha$ is the Frames present per second in a sequence of images. i =0, 1,2,… n denotes the seconds of video. The number of frames per second is mathematically denoted below,

$$\alpha = \sum_{j=0}^{n} f_j$$

Where $\alpha$ denotes number of frames per second, f is the Frames present in sequence of image and j =0,

1,2,… n denotes the possession of the frame in sequence of image. (The available number of frames per second is 24fps / 25fps / 30fps)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 3, March - 2013

In this paper we consider a sequence of image with 24 fps (fps means frames per second). As our paper fully deals with traffic surveillance we consider traffic video as input. Reading each frame is expressed mathematically as below,

$$\sum f = f_0, f_1, f_2, \ldots f_n$$

Before converting we have to find number of frames in a sequence of image (video), which is denoted as "n". We start reading each frame until the end of frame in a sequence and save each frame in a directory. Some of the frames are displayed below,

Figure 1.1 Video to Frame (Frame-1, Frame- 38, Frame-64)

## 1.2. Cropping the number plate

Cropping is defined as removal of unwanted parts or cutting the required part of an image. In our paper we are cropping the required part of an image (frame / scene) which we gathered from a video (sequence of image). Cropping process is mathematically expressed as follows,

$$A = I\,(X_1, Y_1, X_2, Y_2)$$

Where I am the input frame to crop the required part to apply super resolution $X_1$ , $Y_1$, $X_2$, and $Y_2$ denotes the axis of the rectangle, which have to select and cropped, and A denote the cropped image from the frame.

In our paper as we concentrate on traffic surveillance (which means number plate detection), we crop the number plate present in a frame as shown below,

Figure 1.2.1 Input Frame, Selecting desired area

Figure 1.2.2 Cropped image of Selected area

## 1.3. Super-Resolution

Super resolution is a current researching technique, which is defined as generating high resolution of an image from a low resolution image or a set of low resolution image [3,13]. In this paper we use single image super resolution. Single image SR (super-resolution) is the task of constructing an HR (high resolution) enlargement of a single LR (low-resolution) image, (which we cropped from a frame). Our proposed method performs four main sub steps as mentioned below,

1. Interpolation.
2. Generation of a set of candidate images. (Regression)
3. Combining candidate images to produce a single image.
4. Post-processing.



### 1.3.1. Interpolation

Interpolation is defined as a process of providing specified values at specified points. There are four different Interpolations namely,

I. Nearest Neighbor Interpolation.
II. Linear Interpolation.
III. Cubic Interpolation.
IV. Spline Interpolation.

In our paper we perform Spline Interpolation. We select Spline Interpolation because it is more sophisticated and produces the smoother edges. Converting a cropped image into the desired scale.

**Algorithm to find the Spline Interpolation [2]**

Let us consider a third order polynomial p (x) for which we produce

$$P(x_1) = y_1$$
$$P(x_2) = y_2$$
$$P'(x_1) = k_1$$
$$P'(x_2) = k_2$$

We can write it in symmetrical form as,

$$P = (1\text{-}t)\, y_1 + t y_2 + t\,(1\text{-}t)\,(a\,(1\text{-}t) + bt)$$

Where,

$t = \dfrac{x - x_1}{x_2 - x_1}$ And

a = $k_1$ ($x_2 - x_1$) – ($y_2 - y_1$)
b = - $k_2$ ($x_2 - x_1$) + ($y_2 - y_1$)

Double Differentiate P we get as follows,

$$P'' = 2\frac{b - 2a + (a-b)3t}{(x_2 - x_1)^2}$$

Double differentiation with respect to $x_1$ and $x_2$ as given below,

$$P''(x_1) = 2\frac{b - 2a}{(x_2 - x_1)^2}$$

$$P''(x_2) = 2\frac{a - 2b}{(x_2 - x_1)^2}$$

In spline interpolation, left of the leftmost "knot" and the rightmost "knot" thus, it form of a straight line with $q'' = 0$, because ruler can move freely.

$$P''(x_n) = -2\{\, 3(y_n - y_{n-1}) - (2k_n + k_{n-1})(x_n - x_{n-1}) / (x_n - x_{n-1})^2 \,\} = 0$$

Following graph show the example of Spline Interpolation clearly.

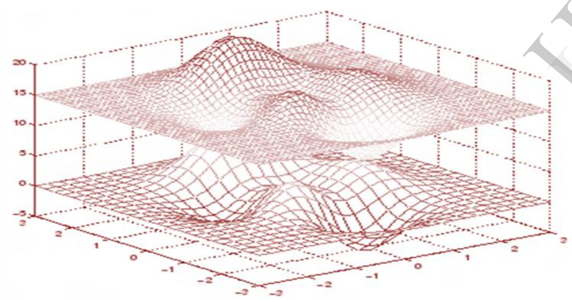

Figure 1.3.1 Graphical representation of interpolation

Applying Spline Interpolation in our cropped image is shown below,

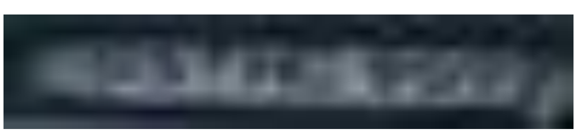

Figure1.3.2 Cropped image

Figure1.3.3 Interpolation image

## 1.4. Generation of set of candidate images.

A set of candidate images is generated based on patch-wise regression. To reduce the time complexity we utilize kernel ridge regression. By combining gradient descent and kernel matching pursuits we found a spare basis [4,6,7].

### 1.4.1. Gradient descent

We use Gradient descent because it's a best optimization algorithm. It provides plenty of data available everywhere and extract information efficiently. The column vector is denoted as,

$$\vec{\theta}_t = (\theta_t(1), \theta_t(2)...., \theta_t(n))^T$$

where T denotes transpose .

Gradient-descent general mathematical expression is given bellow,

$$\vec{\theta}_{t+1} = \vec{\theta}_t - \frac{1}{2}\alpha \nabla_{\vec{\theta}_t} \left[ V^{\pi}(s_t) - V_t(s_t) \right]^2$$

Where, $V_t(s_t)$ is a smooth differentiable function of $\vec{\theta}_t$ for all $s \in S$ .

Let us generate candidate images using the interpolation image as shown below,

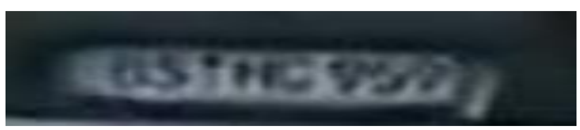

Figure 1.4.1. Interpolation Resultant Image

Figure 1.4.2. Regression results

## 1.5. Combining candidate images to produce a single image

We produce a single image by convex combinations of candidate images by noting that the KRR (kernel ridge regression) (spares) corresponds to the Map estimated with the GP (sparse) prior [8], GP (Gaussian process) is defined as a learning technique, seeks to predict the value of unknown function of any valid input when we provided with of a set of input/examples. By using GP we produce combiner and regresses simultaneously so that error measure is

minimized [9]. We produced set of linear repressors' (candidate images) which is trained such that for the each location (x, y) by applying GP we receive a patch of output image as Z (Nl (x, y) , :) and also produce estimation difference as given below,
$(\{d_1(x,y), \ldots d_n(x,y)\})$

Where d is the difference between the various x and y values present between the set of images (candidate images), the final estimation of the pixel we obtain as a convex combination of candidates which is formulated below,

$$Y(x, y) = \{\textstyle\sum_{i=1}^{N} \{W_i(x, y)\, Z(x, y, i)\}\}$$

Where,

$$W_i(x, y) = \{\exp(M) / \textstyle\sum_{j=1,\ldots,N} \exp(M)\}$$

$$M = -\left|d_i(x, y)\right| \frac{1}{\sigma_C}$$

We re-moving hyper parameters which are chosen based on the error rate of SR for a set of images. To combine two image values the difference is considered as given below,

$$D([x_i,y_i], [x_j,y_j]) = \sqrt{(\|x_i\, x_j\|^2 + (\sigma_x / \sigma_y)\, \|y_i - y_j\|^2)}$$

Where $\sigma_x$ and $\sigma_y$ are the variances of the distances between the pairs of training data points in x and y respectively [10].

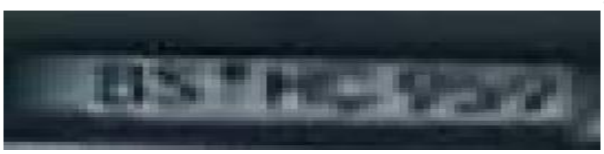

Figure 1.5.1. Combining candidate

**1.6. Post-processing**

Post processing is carried based on Based on Image Prior. Image prior is flexible high-order MRF (Based on Image Prior). Here we use a modification of the NIP (natural image prior) framework which was proposed by tappen et al[12].

$$P(\{x\}\{y\}) = \frac{1}{C} \prod_{(j,i \in Ns(j))} \exp\left[-\left(\frac{\left|\hat{x}_j - \hat{x}_i\right|}{\sigma_N}\right)^{\alpha}\right] . \prod_{j} \exp\left[-\left(\frac{\hat{x}_j - y_j}{\sigma_R}\right)^{2}\right]$$

Where,

$\{y\}$ -Observed variables corresponding to the pixel values of the y.

$\{x\}$ -Latent variable.

$N_s(j)$ - j's pixel location, 8-connected neighbours.

$C$ is normalization constant.

$\prod_{(j,i \in Ns(j))} \exp\left[-\left(\frac{|\hat{x}_j - \hat{x}_i|}{\sigma_N}\right)^{\alpha}\right]$ to prevent the output image flowing far away from input(regression based SR) result of y.

The factor graph representation is given below,

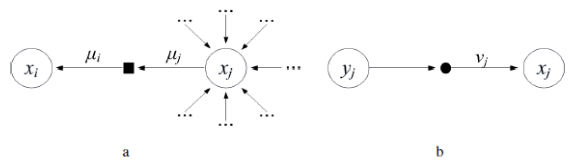


Figure 1.6.1. a. NIP term, and b. Deviation penalty term.

The input is blurred, and removed the very high spatial-frequency components from it. The major edges are found by the Thresholding the each pixel based on the Laplacian and range of pixel values present in local patches [13]. Applying the post processing in our sample we get the following output.

Figure 1.6.2. Final output

## Conclusion

The Super-Resolution of cropped scene from a video is performed. The super resolution is carried by using interpolation, regression (where we generate a set of candidate images), then we combined candidate images produced in regression and finally we perform post- processing to get a Super - Resolute image as output.. Future direction of this paper can be work object detection and traffic surveillances in multi direction and multi-angle.

## Reference

[1] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration,"in *Advances in Computer Vision and Image Proscessing*. Greenwich, CT: JAI Press, 1984, vol. 1, pp. 317–339.

[2] Jing Tian, Kai-Kuang Ma "A survey on super-resolution imaging " Signal, Image and Video Processing, Springer Link September 2011, Volume 5, Issue 3, pp 329-342.

[3]E.C.Pasztor, W.T.Freeman, & O.T.Carmichael, "Learning low-level vision; International Journal of Computer Vision, 2000.

[4]Research on an image magnifying algm based on cubic spline interpolation. Mei Wang, Yanhua Che, Yongling chu, ShaoChun Li, EMEIT, 2011 International Conference .

[5]Carin, Dobeck,G.J., Kernet Matching, Stack, J.R., Xuejun Liao, L. Neural Networks, IEEE Transaction. Application of Adaptive-Kernel Matching Pursuit to Estimate Mixtyre-Pixel Proportion. Wang Xiaoqin, Wu Bo, Huang Bo. Image & Graphics, 2007.

[6]The Kernel-match Pursuit for the Large-datasets. Vlad Popovici, Samy Bengio Jean–Philippe Thiran. EPFL. Signal Processing Institute.

[7]Y. Kwon and K. I. Kim, "Single image super resolution using sparse regression & natural image prior", *IEEE . Pattern Analysis & Machine Intelligence* 2010.

[8]M. F. Tappen, W. T. Freeman, amd B. C. Russel, "Exploiting sparse-derivative prior for the super resolution & image-demosaicing". *IEEE Workshop on Statistical & Computational Theories of Vision*-2003.

[9]Z. Ghahramani & E. Snelson, Sparse Gaussian process using pseudo-inputs. In Advances in Neural Info. Processing Sys., Cambridge, MA, 2006.

[10]R. A. Jacobs, I. Jordan ,Hierarchical mixture of experts & the em algorithm. Neural Computation,1994.

[11]B. C. Russel, W. T. Freeman, & M. F. Tappen. Exploiting sparse derivative prior for the super-resolution and the image demosaicing, IEEE - 2003.

[12]Kwang In Kim and Younghee Kwon Example based Learning Super-Resolution for Single-Image.

[13]Deqing Sun and Wai-Kuen Cham, Post processing of Low Bit Rate Block D.C.T. Coded Images Based on the Fields of Experts Prior, IEEE Transactions on Image Processing, Nov 2007.

[14]Y. Gong, S. Dai, Y. Wu, and M. Han. Bilateral back projection for a single image super-resolution. IEEE International Conference on Multimedia and Expo, 2007.