

# IMPROVING & RESTORING OF IMAGE USING SUPER-RESOLUTION TECHNIQUE

**Shwetambari Shinde**

*Student(M-tech)/C.S.E.  
CSVТУ  
Nagpur, 440009, India*

*shweta\_shinde9388@yahoo.co.in*

**Meeta Dewangan**

*Lecturer/C.S.E.  
CSVТУ  
Bhilai, 491001, India*

*meetadewangan@gmail.com*

---

## Abstract

Image Super-Resolution is a multithreaded Image for super-resolution. Super-resolution is an image fusion and reconstruction problem, where an improved resolution image is obtained from several geometrically warped, low resolution images. The high resolution image is not only an image that has more pixels (like in the case of interpolation), but it also has more visible details. A non-iterative method of image super-resolution based on weighted median filtering with Gaussian weights is proposed. Visual tests and basic edges metrics were used to examine the method. It was shown that the weighted median filtering reduces the errors caused by inaccurate motion vectors. Image re sampling is one of the most important problems in image processing. Many image re sampling algorithms use a priori information. As an example, image self-similarity at different resolutions is used in NEDI algorithm. Nevertheless this approach improves the image visual quality only if the a priori information is true.

**Keywords:** low resolution, Gaussian weight, median filter, super-resolution.

---

## 1. INTRODUCTION

Good techniques for super-resolution are especially useful where physical limitations exist preventing higher resolution images from being obtained. For example, in astronomy where public presentation of images is of significant importance, super-resolution techniques have been suggested. Whenever dynamic image enlargement is needed, such as on some web pages, super-resolution techniques can be utilised.

This paper focuses on the issue of how to increase the resolution of a single image using only prior information about images in general, and not relying on a special training set or the use of multiple images. The methods for achieving super-resolution are as varied as the applications. They range from simple use of Gaussian or preferably median filtering, to supervised learning methods based on learning image patches corresponding to low resolution regions from training data, and effectively sewing these patches together in a consistent manner. What method is appropriate depends on how easy it is to get suitable training data, how fast the method needs to be and so on. There is a demand for methods which are reasonably fast, which are generic in that they do not rely on having suitable training data, but which do better than standard linear filters or interpolation methods. This paper describes an approach to resolution doubling which achieves this. The method is structurally related to one layer of the dynamic tree model. We use a dynamic tree-like architecture to model the high resolution data. Approximate conditioning on the low resolution image is achieved through a mean field approach. Except that it uses real valued variables.

## 2. Related work

In the super-resolution literature there are many papers which do not deal with the simple case of reconstruction based on a single image. Many authors are interested in reconstruction based on multiple slightly perturbed subsamples from an image. This is useful for photographic scanners for example. In a similar manner other authors utilise the information from a number of frames in a temporal sequence. In other situations highly substantial prior information is given, such as the ground truth for a part of the image. Sometimes restrictions on the type of processing might be made in order to keep calculations in real time or deal with sequential transmission.

Simple approaches to resolution enhancement have been around for some time. Gaussian and Wiener filters (and a host of other linear filters) have been used for smoothing the blockiness created by the low resolution image. Median filters tend to fare better, producing less blurry images. Interpolation methods such as cubic-spline interpolation tend to be the most common image enhancement approach.

One important paper which deals specifically with the problem tackled here is by Freeman, Jones and Pasztor. They follow a supervised approach, learning a Low to high resolution patch model (or rather storing examples of such maps), and utilising a Markov random field for combining them and loopy propagation for inference. Later work simplifies and improves on this approach. Earlier Work tackling the same problem includes that of Schultz and Stevenson, which Performed an MAP estimation using a Gibbs prior.

There are two primary difficulties with smoothing (e.g. Gaussian, Wiener, Median filters) or interpolation (bicubic, cubic spline) methods. First smoothing is indiscriminate. It occurs both within the gradual change in colour of the sky, say, as well as across the horizon, producing blurring problems. Second these approaches are inconsistent: sub sampling the super-resolution

image will not return the original low-resolution one. Hence we need a model which maintains consistency but also tries to ensure that smoothing does not occur across region boundaries (except as much as is needed for anti-aliasing).

## 2.1 The model

Here the high-resolution image is described by a series of very small patches with varying shapes. Pixel values within these patches can vary, but will have a common mean value. Pixel values across patches are independent. Apriori exactly where these patches should be is uncertain, and so the pixel to patch mapping is allowed to be a dynamic one. The model is best represented by a belief network. It consists of three layers. The lowest layer consists of the visible low-resolution pixels. The intermediate layer is a high-resolution image (4 X 4 the size of the low-resolution image). The top layer is a latent layer which is a little more than 2 X2 the size of the low resolution image. The latent variables are positioned at the corners, centres and edge centres of

the pixels of the low resolution image. The values of the pixel colour of the high resolution nodes are each a single sample from a Gaussian mixture (in colour space), where each mixture centre is given by the pixel colour of a particular parent latent.

Another way of viewing this is that a further indicator variable can be introduced which selects which mixture is responsible for a given high-resolution node. We say a high resolution node chooses to connect to the parent that is responsible for it, with a connection probability given by the corresponding mixing coefficient. These connection probabilities can be specified in terms of positions.

The motivation for this model comes from the possibility of explaining away. In linear filtering methods each high-resolution node is determined by a fixed relationship to its neighbouring low-resolution nodes. Here if one of the latent variables provides an explanation for a high-resolution node which fits well with its neighbours to form the low-resolution data, then the posterior responsibility of the other latent nodes for that high-resolution pixel is reduced, and they are free to be used to model other nearby pixels. The high-resolution pixels corresponding to a visible node can be separated into two (or more) independent regions, corresponding to pixels on different sides of an edge (or edges). A different latent variable is responsible for each region. In other words each mixture component effectively corresponds to a small image patch which can vary in size depending on what pixels it is responsible for  $\epsilon$ .

Let  $v_j \in L$  denote a latent variable at site  $j$  in the latent space  $L$ . Let  $x_i \in S$  denote the value of pixel  $i$  in high resolution image space  $S$ , and let  $y_k$  denote the value of the visible pixel  $k$ . Each of these is a 3-vector representing colour. Let  $V$  denote the ordered set of all  $v_j$ . Likewise  $X$  denotes the ordered set of all  $x_i$  and  $Y$  the set of all  $y_i$ . In all the work described here a transformed colorspace of (gray, red-green, blue-yellow) is used. In other words the data is a linear transformation on the RGB colour values using the matrix

$$\begin{bmatrix} 0.66 & 1 & 0.5 \\ 0.66 & -1 & 0.5 \\ 0.66 & 0 & -1 \end{bmatrix}$$

The remaining component is the connectivity (i.e. the indicator for the responsibility) between the high-resolution nodes and the nodes in the latent layer. Let  $z_{ij}$  denote this connectivity with  $z_{ij}$  an indicator variable taking value 1 when  $v_j$  is a parent of  $x_i$  in the belief network. Every high resolution pixel has one and only one parent in the latent layer. Let  $Z$  denote the ordered set of all  $z_{ij}$ .

## 2.2 Distributions

A uniform distribution over the range of pixel values is presumed for the latent variables. The high resolution pixels are given by Gaussian distributions centred on the pixel values of the parental latent variable. This Gaussian is presumed

independent in each pixel component. Finally the low resolution pixels are given by the average of the sixteen high resolution pixels covering the site of the low resolution pixel. This pixel value can also be subject to some additional Gaussian noise if necessary (zero noise is assumed in this paper). It is presumed that each high resolution pixel is allowed to choose its parent from the set of latent variables in an independent manner. A pixel has a higher probability of choosing a nearby parent than a far away one. For this we use a Gaussian integral form so that:

$$P(Z) = \prod_{ij} P_{ij} \quad \text{where } P_{ij} \propto \int_{B_r} dr \exp(-(rj - r)^2 / 2\sum), \quad (1)$$

where  $r$  is a position in the high resolution picture space,  $rj$  is the position of the  $j$ th latent variable in the high resolution image space (where these are located at the corners of every second pixel in each direction as described above). The integral is over  $B_i$  defined as the region in image space corresponding to pixel  $x_i$ .  $\sum$  gives the width (squared) over which the probability decays. The larger  $\sum$  the more possible parents with non-negligible probability. The equations for the other distributions are given here. First we have

$$P(X | Z, V) = \frac{1}{(2\pi\Omega_m)^{\frac{1}{2}}} \exp(-Z_{ij}(x_i^m - v_j^m)^2 / 2\Omega_m) \quad (2)$$

Where  $m$  is a variance which determines how much each pixel must be like its latent parent. Here the indicator  $z_{ij}$  ensures the only contribution for each  $i$  comes from the parent  $j$  of  $i$ . Second with  $P_a(k)$  denoting the set of all the  $d = 16$  high resolution pixels which go to make up the low resolution pixel  $y_k$ . In this work we let the variance  $\square \rightarrow 0$ .  $\square$  determines the additive Gaussian noise which is in the low resolution image. Last,  $P(V)$  is simply uniform over the whole of the possible values of  $V$ . Hence  $P(V) = 1/C$  for  $C$  the volume of  $V$  space being considered.

$$P(Y | X) = \prod_{km} \frac{1}{(2\pi\square)^{1/2}} \exp(-(y_m^k - \frac{1}{d} \sum_{i \in P_a(k)} x_i^m)^2 / 2\square) \quad (3)$$

### 2.3 Setting parameters

The prior variance parameters need to be set. The variance  $\square$  corresponds to the additive noise. If this is not known to be zero, then it will vary from image to image, and needs to be found for each image. This can be done using variational maximum likelihood, where  $\square$  is set to maximise the variational log likelihood.  $\sum$  is presumed to be independent of the images presented, and is set by hand by visualising changes on a test set. The  $m$  might depend on the intensity levels in the image: very dark images will need a smaller value of 1 for example. However for simplicity  $\Omega^m = \Omega$  is treated as global and set by hand. Because the primary criterion for optimal parameters is subjective, this is the most sensible approach, and is reasonable when there are only two parameters to determine. To optimise automatically based on the variational log likelihood is possible but does not produce as good results due to the complicated nature of a true prior or error-measure for images. For example, a highly elaborate texture offset by one pixel will give a large mean square error, but look almost identical, whereas a blurred version of the texture would give a smaller mean square error, but look much worse.

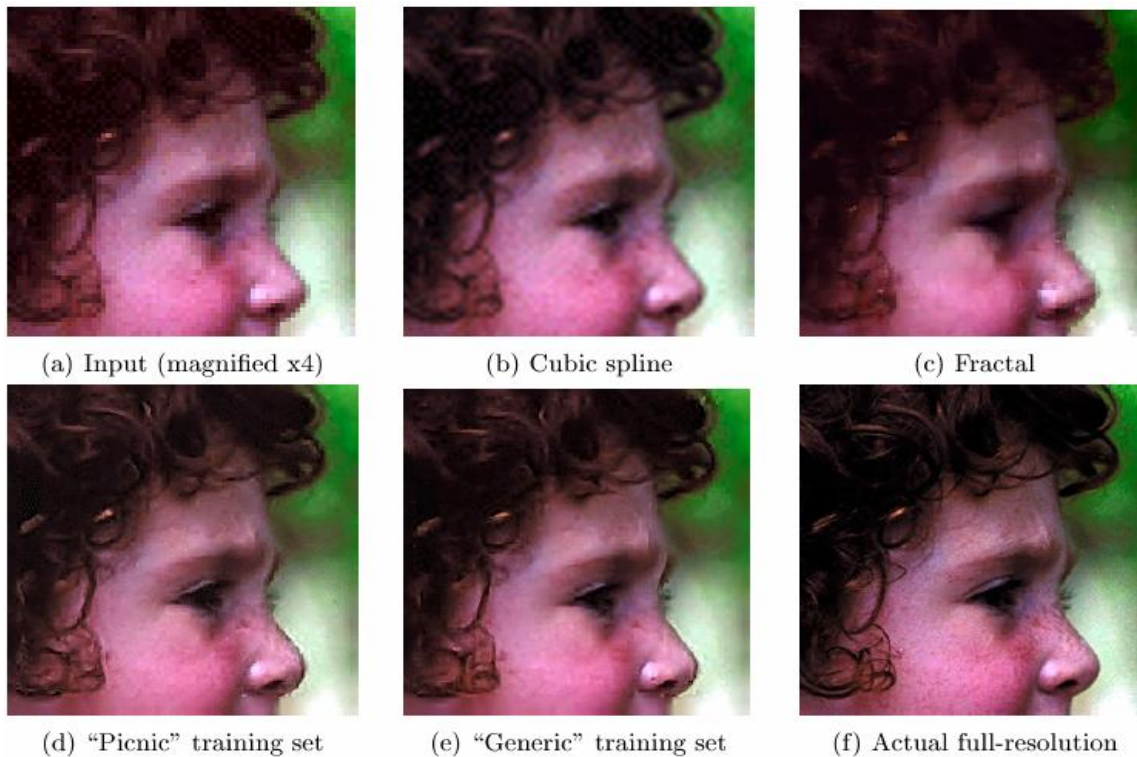
## 3. Implementation

The basic implementation involves setting the parameters, running the mean field optimisation and then looking at the result. The final result is a down sampled version of the  $4 \times 4$  image to  $2 \times 2$  sizes: the larger image is used to get reasonable Anti-aliasing.

To initialise the mean field optimisation,  $X$  is set equal to the bi-cubic interpolated Image with added Gaussian noise. The  $Q(Z)$  is initialised to  $P(Z)$ . Although in the examples here we used 25 optimisations  $Q(Z)$ , each of which involves 10 cycles through the mean field equations for  $Q(X)$  and  $Q(V)$ , it is possible to get reasonable results with only three  $Q(Z)$  optimisation cycles each doing 2 iterations through the mean field equations. In the runs shown here,  $\square$  is set to zero, the variance is set to 0.008, and  $\sum$  is set to 3.3.

## 4. Demonstrations and assessment

The method described in this paper is compared with a number of simple filtering and interpolation methods, and also with the methods of Freeman et al. The image from Freeman's website is used for comparison with that work (figure 1). Full colour comparisons for these and other images can be found at <http://www.anc.ed.ac.uk/~amos/superresolution.html>. First two linear filtering approaches are considered, the Wiener filter and a Gaussian filter.



**FIGURE 1:** Comparison

with approach of Freeman et al.

The third method is a median filter. Bi-cubic interpolation is also given. Quantitative assessment of the quality of super-resolution results is always something of a difficulty because the basic criterion is human subjectivity. Even so we compare the results of this approach with standard filtering methods using a root mean squared pixel error on a set of 8, 128 by 96 colour images, giving 0:0486, 0:0467, 0:0510 and 0:0452 for the original low resolution image, bicubic interpolation, the median filter and dynamic structure super-resolution respectively. Unfortunately the unavailability of code prevents representative calculations for the Freeman et al approach. Dynamic structure resolution requires approximately 30 - 60 ops per  $2 \times 2$  high resolution pixel per optimisation cycle, compared with, say, 16 ops for a linear filter, so it is more costly. Trials have been done working directly with  $2 \times 2$  grids rather than with  $4 \times 4$  and then averaging up. This is much faster and the results, though not quite as good, were still an improvement on the simpler methods.

Qualitatively, the results for super-resolution are significantly better than most standard filtering approaches. The texture is better represented because it maintains consistency, and the edges are sharper, although there is still some significant difference from the true image. The method of Freeman et al is perhaps comparable at this resolution, although it should be noted that their result has been down sampled here to half the size of their enhanced image. Their method can produce  $4 \times 4$  the resolution of the original, and so this does not accurately represent the full power of their technique. Furthermore this image is representative of early results from their work. However their approach does require learning large numbers of patches from a training set. Fundamentally the dynamic structure super-resolution approach does a good job at resolution doubling without the need for representative training data. The edges are not blurred and much of the blockiness is removed. Super-resolution provides a technique for resolution enhancement, and provides an interesting starting model which is different from the Markov random field approaches. Future directions could incorporate hierarchical frequency information at each node rather than just a single value.

## 5. REFERENCES

### Book

- [1] R. C. Gonzalez and P. Wintz, Digital Image Processing. New York: Addison-Wesley, 2010.
- [2] W. K. Pratt, Digital Image Processing. New York: Wiley, 2010.
- [3] A. K. Jain, Fundamentals in Digital Image Processing. Englewood Cliffs, NJ: Prentice-Hall, 1989.

### Article in a Journal

- [1] Andrew M. Maiden, Martin J. Humphry, Fucui Zhang, and John M. Rodenburg, "Super resolution imaging," J. Opt. Soc. Am. A 28, 604-612 (2011)

- [2] S.C. Park, M.K. Park, and M.G. KANG, "Super-Resolution Image Reconstruction: A Technical Overview", IEEE Signal Processing Magazine, Vol. 20, pp. 21-36, May 2003
- [3] W.T. Freeman, T.R. Jones, and E.C. Pasztor, "Example-Based Super-Resolution", IEEE Computer Graphics and Applications, Vol. 22, pp. 56-65, 2002.
- [4] N. J. Adams. Dynamic Trees: A Hierarchical Probabilistic Approach to Image Modelling. PhD thesis, Division of Informatics, University of Edinburgh, 5 Forrest Hill, Edinburgh, EH1 2QL, UK, 2001.
- [5] M. Elad and A. Feuer. Super-resolution reconstruction of image sequences. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(9):817-834, 1999.