Multiple Regression based Graphical Modeling for Images

Pavan S., Sridhar G., and Sridhar V.

Abstract—Super resolution is one of the commonly referred inference problems in computer vision. In the case of images, this problem is generally addressed using a graphical model framework wherein each node represents a portion of the image and the edges between the nodes represent the statistical dependencies. However, the large dimensionality of images along with the large number of possible states for a node makes the inference problem computationally intractable. In this paper, we propose a representation wherein each node can be represented as a combination of multiple regression functions. The proposed approach achieves a tradeoff between the computational complexity and inference accuracy by varying the number of regression functions for a node.

Keywords—Belief propagation, Graphical model, Regression, Super resolution.

I. INTRODUCTION

Graphical models are being extensively explored for inference applications in the case of images. A graphical model is used to infer the underlying scene given the observed image data. One of the commonly used techniques for solving approximate inference vision problems is belief propagation (BP). However, the large dimensionality of images tends to pose severe challenges for a tractable application of graphical models to inference problems in the case of images.

Each pixel or a block of pixels in the image can be considered as a node in the graph. Each node in the graph can take on a set of discrete states. As an example, in the case of a grayscale image, each node can take on any of the 256 distinct discrete states. The amount of computation required to estimate the corresponding scene for each observed node using BP is of the order of $N^2$, where $N$ is the number of discrete states for a node. A reduction in the number of states per node would result in a decrease in the accuracy of the estimate. Therefore, there is a need for an approach that can vary the number of states associated with a node in a flexible manner to achieve a tradeoff between computational complexity and inference accuracy. One of the applications of such a flexible computational approach is super resolution. In this paper, a novel super resolution algorithm (also referred to as image interpolation) based on multiple regressors is proposed for deriving a high-resolution (HR) image from a given low-resolution (LR) image.

II. RELATED WORK

Super resolution algorithms can be classified into two groups: (i) Functional interpolation and (ii) Learning based methods. Functional interpolation methods such as nearest neighbor, bilinear, bicubic, and cubic-spline, approximate the unknown pixel values from a known set of local functions [1], [7]. The major drawback of these methods is that the resulting images often have blurred edges. Even though, the resulting image can be sharpened by using image sharpening techniques [8], [9], this results in haloing artifacts. The performance can be improved by deconvolving the blurred image [14]. However, the deconvolution method only enhances the features that are present in the LR image.

The other class of methods requires some prior information that is obtained using either a training data set or some unrelated scenes having the desired artistic style. Freeman et al [2] have proposed a technique to learn the compatibility functions required to estimate the transformation between the observed image and the unobserved scene from a training data set. Chang et al [7] have suggested the use of local linear embedding which is a manifold learning scheme that infers the HR image from multiple nearest neighbors in the training set. Rosales et al [5] have proposed a method to render the image in a particular artistic style and having the desired resolution. Tappen et al [3] have suggested an approach to reduce the computational complexity by assigning a single state to a node using a regression function. The state assigned to a node is represented by a regression function. The set of possible regression functions is learned from a training data set. The use of a single regression function for each node tends to impose some limitations on the accuracy level that can be achieved with this method. The estimation accuracy also depends on the number of regression functions present in the set and the precise matching of the regression functions obtained using the data set with the desired transformation.

In this paper, we propose an approach wherein each node in the graphical model is represented using multiple regression
functions. This technique estimates latent image pixel values with better accuracy than the above-mentioned approaches. The approach is also flexible to achieve a tradeoff between the computational complexity and estimation accuracy.

III. SUPER_RESOLUTION ALGORITHM

The purpose of a super-resolution algorithm is to estimate the latent HR scene from the observed image data. Both the scene and image are divided into spatially non-overlapping blocks and represent them using a factor graph [13] as shown in Fig. 1. Each block is represented as a node in the network. The links connecting the nodes represent the statistical dependencies as illustrated in Fig. 1.

![Factor graph for super-resolution](image)

**A. Intensity based regression function**

Each node in the graph can take on a maximum of 256 distinct states. For large networks, as in the case of images, this type of representation would yield the most accurate estimate but is computationally not tractable. To ameliorate the computational efficiency, each node can be assigned a single state (using a regression function) as in [3]. This system would be computationally efficient but a good estimate of the accuracy would require the use of a large set of regression functions. Thus, there is a need for a tradeoff between the computational performance and accuracy.

To attain this objective, it is proposed to segment the intensity space into a finite number of clusters (C). Each of these intensity clusters is represented by a regression function (R). Each node in a latent HR image incorporates multiple regression functions that can be combined to obtain a sub-sampled LR image. Based on this, a model for the generation of an HR image is proposed. Specifically, it is assumed that an LR image is obtained from the HR image by convolving the intensity clusters by a low-pass filter (K) and weighting them based on their frequency of occurrence.

![Factor graph for super-resolution](image)

**B. Compatibility function model**

Let x and y represent the scene and image blocks respectively. Let the number of scene and image blocks be (M x M) where (M x M) is the size of the observed image. The joint probability distribution between the HR and LR nodes is modeled as a graphical model using the two compatibility functions \( \psi(x_i, x_j) \) and \( \phi(x_i, y_j) \) [12].

The reconstruction error between the HR and LR image \( \phi(x_i, y_j) \) is modeled as a Gaussian distribution.

\[
\phi(x_i, y_j) = \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{l=1}^{K} (C_l * K_j - y_j)^2\right)\right)
\]

where \( \sigma \) is the standard deviation parameter.

The other compatibility function \( \psi(x_i, x_j) \) models the structure of the HR image to be inferred. In the case of natural images, the image derivative can be modeled by the Laplacian distribution [6].

\[
\psi(x_i, x_j) = \exp\left(-\frac{1}{\sigma^2}\left(\sum_{l=1}^{K} (C_l * K_j - (C_j * K_j))\right)^2\right)
\]

where \( x_i \) and \( x_j \) represent the neighboring blocks in the HR image, \( \sigma \) is the standard deviation, and \( \alpha \) is the exponent parameter.

**C. Regression based latent image estimation**

Each node in a latent image can be represented by multiple regression functions. A new set of compatibility functions is defined for each regression function. The max-product BP algorithm is applied to each node in the graph. The BP algorithm has been widely used to perform inference in Bayesian networks [10]. It has been recently applied to graphs with cycles under the name of loopy belief propagation [4].

In the proposed method, the BP algorithm is used to find the marginal probabilities (or beliefs) of the node by passing messages between the nodes present in the graph. The set of regression functions, which yield the marginal probabilities greater than some threshold \( T \) are selected.

To obtain an estimate of a node present in the latent image, the multiple regression functions are linearly combined based on (i) their marginal probabilities (Pr), (ii) the similarity (D) of a regression function with the corresponding observed value, and (iii) the Euclidean distance between the coordinates of the pixel to be inferred and the origin. Let \( v_k \) (\( \forall k=1:b^2 \)) be a pixel co-ordinates in \( x_i \) and \( s_k \) is the corresponding pixel intensity value.

\[
s_k = y_i \times fn(Pr, D, E)
\]

where \( fn \) is a linear function of \( Pr \), D and E, illustrated in more detail in Fig. 2.
Let \( I \) be the observed image of size \( M \times M \). Let \( x_n \) and \( y_n \) \((\forall n=1:\{(M \times M)\})\) represent the unobserved HR nodes and observed LR nodes respectively.

**Intensity_based_regression(y)**

\[
\begin{align*}
\text{Segment the intensity space into a finite set of clusters} & \{C_1,…C_p\}, \forall \leq p \leq 256 \\
\text{Convolve each cluster } C_i \text{ with a lowpass anti-aliasing Gaussian kernel } K_i \\
\text{Divide the image into non-overlapping blocks } (B_i) & \forall i=1:\{(M \times M)\} \text{ of size } (b \times b) \\
\text{For each block } B_i, \text{ calculate the frequency of occurrence } & \{g_j\} \text{ of the pixels belonging to the various clusters } \{C_j\}, \forall j=1:p \\
\text{The LR image pixel value is obtained from the relation:} & \sum_{j=1}^{p} C_j \ast K_i \ast w_j \\
\text{where } w=f(r), \text{ where } f \text{ is some pre-defined function} & \\
\text{Compatibility_function_model(y, C, K, w)} \\
\text{Design the compatibility functions} & (\phi(x_i, y_i), \psi(x_i, x_j)) \text{ based on the model in equations (1) and (2) (refer to the text)} \\
\text{Latent_image_estimation( } \phi, \psi, y \}
\end{align*}
\]

**V. CONCLUSION**

In this paper the super resolution inference problem is mapped onto a graphical model framework. A tractable application of this model in the case of images is possible with each node in the graph being represented by multiple regression functions defined in the intensity space. The number of regression functions for a node can be varied to obtain a superior computational performance or an accurate estimate, that is a tradeoff between the computational complexity and inference accuracy can be achieved.

**REFERENCES**

FIG. 3. Comparison of 64x64 images. Fig 3(a)-3(d) Results of experiments conducted on the face image, Fig 3(f)-3(j) Results of experiments conducted on the peppers image, Fig 3(k)-3(o) Results of experiments conducted on the cat image, Fig 3(p)-3(t) Results of experiments conducted on the Lena image.

<table>
<thead>
<tr>
<th></th>
<th>Nearest</th>
<th>Bilinear</th>
<th>Bicubic</th>
<th>Proposed approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td>26.2070</td>
<td>27.5188</td>
<td>25.7917</td>
<td>4.7864</td>
</tr>
<tr>
<td>Pepper</td>
<td>23.2178</td>
<td>25.0369</td>
<td>24.2993</td>
<td>6.7939</td>
</tr>
<tr>
<td>Cat</td>
<td>21.5129</td>
<td>20.7490</td>
<td>20.3472</td>
<td>8.0535</td>
</tr>
</tbody>
</table>