World Academy of Science, Engineering and Technology
International Journal of Aerospace and Mechanical Engineering
Vol:1, No:2, 2007

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.

Pavan S. is with the Applied Research Group, Satyam Computer Services Limited, #14, Langford Avenue, Lalbagh Road, Bangalore, India -560 025. (e-mail: Pavan_S@satyam.com).

Sridhar G. is with the Applied Research Group, Satyam Computer Services Limited, #14, Langford Avenue, Lalbagh Road, Bangalore, India -560 025. (phone: +91-80-22231696; fax: +91- 80 –2227 1882;e-mail: Sridhar_Gangadharpalli@satyam.com).

Sridhar V. is with the Applied Research Group, Satyam Computer Services Limited, #14, Langford Avenue, Lalbagh Road, Bangalore, India -560 025. (e-mail: Sridhar@satyam.com).

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], [11]. 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.

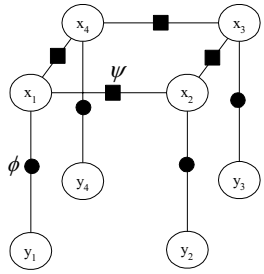

FIG.1.Factor graph for super-resolution

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 decimation 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.

B. Compatibility function model

Let x and y represent the scene and image blocks respectively. Let the number of scene and image blocks be (MxM) where (MxM) 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 $\phi(x_i, y_i)$ [12].

$$P(x_1, \ldots, x_{(MxM)}, y_1, \ldots, y_{(MxM)}) = \prod_{i,j} \psi(x_i, x_j) \prod_i \phi(x_i, y_i) \quad (1)$$

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

$$\phi(x_i, y_i) = \exp\left[-\frac{1}{2} \left\{ \frac{\left(\sum_{l=1}^{bxb} (C_l * K_l \times w_l) - y_i \right)}{\sigma} \right\}^2 \right] \quad (2)$$

where σ 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].

$$\psi(x_i, x_j) = \prod_{i,j} \exp\left[\frac{\left| \left((C_i * K_i) - (C_j * K_j) \right) \right| \alpha}{\sigma} \right] \quad (3)$$

where x_i and x_j represent the neighboring blocks in the HR image, σ is the standard deviation, and α 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 co-ordinates 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.

$$s_k = y_i \times fn(\Pr, D, E) \quad (4)$$

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 MxM.Let x_n and y_n ($\forall$ n=1: (MxM)) represent the unobserved HR nodes and observed LR nodes respectively.

Intensity_based_regression(y)

{

Segment the intensity space into a finite set of clusters $\{C_1 \ldots C_i \ldots C_p\}$ $\forall\ 1 \le p \le 256$

Convolve each cluster C_i with a lowpass anti-aliasing Gaussian kernel K

Divide the image into non-overlapping blocks (B_i $\forall$ i=1: (MxM)) of size (bxb)

For each block B, calculate the frequency of occurrence (rj) of the pixels belonging to the various clusters (C_j) $\forall$ j=1: p

The LR image pixel value is obtained from the relation:

$$y_i = \sum_{l=1}^{bxb} C_l * K_l \times w_l$$

where w=f(r), where f is some pre-defined function

}

Compatibility_function_model(y, C, K, w)

{

Design the compatibility functions ($\phi(x_i, y_i)$, $\psi(x_i, x_j)$) based on the model in equations (1) and (2) (refer to the text)

}

Latent_image_estimation (ϕ, ψ, y)

{

Calculate a regression function (R_i) for each cluster C_i for $1 \le i \le p$

For each y_i, calculate the new set of compatibility functions using the regression functions

Use the max-product belief propagation algorithm to choose a set of regression functions which yield probabilities (Pr_j, $\forall$ j=1: d) greater than some threshold T

The multiple regression functions are linearly combined based on probability (Pr), similarity with observed value, and euclidean distance to obtain the latent LR image.

FIG.2.Super Resolution Algorithm

IV. RESULTS AND DISCUSSION

The HR image is considered to be composed of non-overlapping 4x4 blocks. Each x_i represents the value of a 4x4 block. The HR image is the 4 times scaled version of an LR image. To evaluate the proposed approach, some test images of size 64x64 are decimated by a factor 4. The intensity space is divided into four equal sized clusters, that is, each cluster contains 64 intensity levels. The decimation is performed by applying a low-pass Gaussian kernel to each intensity cluster and combining them based on the frequency of occurrence of the clusters in the node. Using the proposed algorithm the LR image is super resolved back to the original dimensions. For the proposed algorithm the experiments are conducted for σ=0.01, α=0.7. For the latent image estimation step, the marginal probabilities which are greater than T = 0.2 are selected. Smoothing along the edges of a block reduces the artifacts resulting from the block-based approach. The proposed approach is compared with nearest neighbor, bilinear, and bicubic super resolution algorithms.

A comparison of the outputs of the different super-resolution algorithms is shown in Fig. 3 for four different images. The proposed approach produces visibly sharper images than the other algorithms. The proposed approach also outperforms the others in terms of the mean square error between each of the true HR images and the output of the different algorithms as listed in Table I.

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

[1] William K.Pratt, *Digital Image Processing,* 3rd ed., John Wiley and sons, 2003, pp. 393-397.

[2] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," in *International Journal of Computer Vision*, 40(1): 25-47, 2000.

[3] Marshall F. Tappen, Bryan C. Russell, and W .T. Freeman, "Efficient graphical models for processing images", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.

[4] Jonathan S. Yedidia, W.T.Freeman, and Yair Weiss, "Understanding belief propagation and its generalizations, *Technical Report TR2001-22, MERL,* 2001".

[5] R. Rosales, K. Achan, and B. Frey, "Unsupervised image translation," in *Ninth International Conference on Computer Vision (ICCV)*, 2003.

[6] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7):674-694, July 1989.

[7] Hong Chang, Dit-Yan Yeung, and Yimin Xiong,"Super-Resolution through neighbor embedding," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.

[8] H. Greenspan, C. Anderson, and S. Akber, "Image enhancement by nonlinear extrapolation in frequency space," *IEEE Trans. on Image Processing,* 9(6), 2000.

[9] B. Morse, and D. Schwartzwald,"Image magnification using level set reconstruction," *Proc. International Conf. Computer Vision (ICCV),* pages 333-341, 2001.

[10] J.Pearl,"*Probabilistic reasoning in intelligent systems: networks of plausible inference*," Morgan Kaufmann, 1988.

[11] H. H. Hou, and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," in *IEEE Trans.Acoust.Speech Signal Processing,* ASSP-26(6):508-517, 1978.

[12] D. Geiger, and F. Girosi, "Parallel and deterministic algorithms from MRF's: Surface reconstruction," in *IEEE Pattern Analysis and Machine Intelligence,* 13(5), 401-412, 1991.

[13] F. R. Kschischang, B. J.Frey, and H. A. Loeliger,"Factor graphs and the sum-product algorithm," in *IEEE Transactions on Information Theory*, 42(2): 498-519, 2001.

[14] M. Belge, M. Kilmer, and E. Miller,"Wavelet domain image restoration with adaptive edge-preserving regularity," in *IEEE Transactions on Image Processing*, 9(4): 597-608, 2000.

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 I.
MEAN SQUARE ERROR FOR THE DIFFERENT INTERPOLATION ALGORITHMS

	Nearest	Bilinear	Bicubic	Proposed approach
Face	26.2070	27.5188	25.7917	4.7864
Pepper	23.2178	25.0369	24.2993	6.7939
Cat	21.5129	20.7490	20.3472	8.0535
Lena	21.1587	21.6597	21.2371	9.9609