ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.6

Super-Resolution using Gaussian Mixture Model

Deepali .M. Bongulwar
Assistant Professor
Department of Electronics Engineering
RVS College of Engineering, Jamshedpur, India

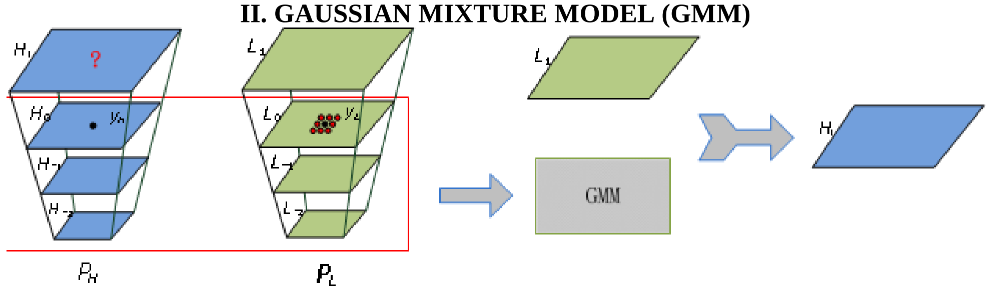
Abstract:
In recent years, super-resolution techniques in the field of computer vision have been studied in earnest owing to the potential applicability of such technology in a variety of fields. In this paper we present a method for single image super-resolution (SR). Given an input low resolution image, we create a pyramid pair: the ground truth pyramid and the interpolated pyramid. This method aims to model the relationship between pixel value in ground truth pyramid and its corresponding 8- neighborhood vector in interpolated pyramid using Gaussian Mixture Model (GMM). Each pixel in final high-resolution image is predicted by its corresponding 8- neighborhood vector through the trained GMM. Our algorithm only utilizes the information of input image rather than the external image database. Our proposed algorithm achieves much better results than the state of the art algorithms in terms of both objective measurement and visual perception.

Keywords: GMM, super-resolution.

I. INTRODUCTION

The resolution of the digital camera installed in cellular phones has increased dramatically in recent years. On the other hand, due to price competition, the need to reduce the cost of the image sensor has been a serious problem. For this reason, the technology for high-resolution digital image processing has been attracting much attention. If images are enlarged by using either linear interpolation or bicubic interpolation (a popular expansion processing technique), the resolution of the images decreases because their edge information is lost. Therefore, a method that assures the high resolution of the expanded images and adds an appropriate high-frequency component to the image is required. A considerable amount of research has been carried out on single-image super-resolution techniques in the field of computer vision. The goal of image super-resolution (SR) is to obtain the corresponding high-resolution (HR) image from the given low-resolution (LR) image. SR method can be broadly categorized into three classes: 1. Interpolation based method; 2. Reconstruction-based method; 3. Example-based method. The Interpolation-based methods [11] are fast and easy to implement but the results are blurred. The Reconstruction-based method considers image SR as a process of solving the inverse problem, which obtains the final HR image by introducing a series of prior knowledge to constraint the inverse problem and compress the solution space [6,10]. But this approach needs much preliminary work to obtain the prior knowledge. Example-based method generates the HR image by searching for the similar image patches from training LR image database and using their corresponding HR image patches [1,3,4,8,9]. It can produce high contrast details in HR image with powerful database searching method. But the training image database needs to be selected carefully. In this paper, we create a single image super resolution method that makes two significant advances: Firstly, we propose a novel SR framework and construct the image pyramid pair [1] that only uses the information of the input LR image without the external image database. Secondly, we consider the relationship between pixel in ground truth pyramid and its surrounding 8-Neighborhood pixels of the corresponding pixel in interpolated pyramid. These pixels are used to train a Gaussian Mixture Model (GMM). Pixels in the final HR image are estimated via the training GMM. In the following parts of this paper, we give a brief introduction of GMM in Section II. In section III, we describe the detailed process of our algorithm: Pyramids construction, GMM model training and prediction for SR. Section IV displays the results of our algorithm and the comparison with others.

II. GAUSSIAN MIXTURE MODEL (GMM)



(a) Pyramid pair (b) GMM (c) HR image

Figure.1. The flowchart of our algorithm framework. (a) The pyramid pair: ground truth pyramid PH and interpolation pyramid PL. (b) Training a GMM. (c) Predict the final HR image H1

Gaussian Mixture Model (GMM) is a parametric probability density function that represented as a weighted sum of Gaussian densities. GMM can approximate an arbitrary probability density function accurately [5]. For a D− dimension random variable Z , we can rewrite its probability density function (PDF) by GMM as,

$$P(z)=\sum_{i=1}^{M} \alpha_i G(Z,\mu_i,\Sigma_i) \quad (1)$$

Where M is the number of component Gaussian densities, α_i denotes weight and $\sum_{i=1}^{M} \alpha_i=1$,

$G(Z,\mu_i,\Sigma_i)$ represents the i^{th} component which is a D-dimension Gaussian function.
We denote the parameter set of the GMM as,

$$\Theta=\{\alpha_1,\alpha_2,\ldots\alpha_M,\mu_1,\mu_2,\ldots\mu_M,\Sigma_1,\Sigma_2,\ldots\Sigma_M\} \quad (2)$$

In this paper, we use EM algorithm[5] to estimate the parameter set θ,EM algorithm alternates the parameters by implementing the E-step and M-step.

For the training set $Z=\{Z_1, Z_2,\ldots\ldots\ldots Z_N\}$, the estimation of the new parameters in terms of the old parameters is as follows,

$$p(j \mid Z_i, \theta^{old})=\frac{\alpha_j^{old}\, G(Z,\mu_j^{old}\, \Sigma_j^{old})}{\sum_{i=1}^{M} \alpha_s^{old}\, G(Z_s,\mu_s^{old},\Sigma_s^{old})} \quad (3)$$

$$\alpha_j^{new} = \frac{1}{N}\sum_{s=1}^{M} p(j \mid Z_i, \theta^{old}) \quad (4)$$

$$\mu_j^{new} = \frac{\sum_{i=1}^{N} Z_i\, p(j|Z_i\theta^{old})}{\sum_{i=1}^{N} p(j|Z_i,\theta^{old})} \quad (5)$$

$$\Sigma_j^{new} = \frac{\sum_{i=1}^{N} p(j|Z_i,\theta^{old})(Z_i-\mu_j^{new})(Z_i-\mu_j^{new})^T}{\sum_{j=1}^{N} p(j|Z_i,\theta^{old})} \quad (6)$$

Note that the equation (3) performs the E-step and the equations (4) - (6) perform the M-step. The algorithm uses the new derived parameters as the prediction for the next iteration.

III. GMM FOR SUPER-RESOLUTION

Figure 1 shows the framework of our algorithm, our algorithm mainly consists of three steps: Pyramid Pair Construction, GMM Training and Prediction for SR. In Section A, we utilize the information about the input low-resolution image to construct a pyramid pair. In Section B, the pyramid pair is used to produce the training data set and train a GMM. In Section C it provides a way to estimate the final high-resolution image using GMM.

A. Pyramid Pair Construction

Motivated by [3], we construct a pyramid pair: the ground truth pyramid P_H and the interpolated pyramid P_L. Fig. 1(a) shows the structure of pyramid pair P_H and P_L. P_H consists of the image set $\{H_i\}$,i= -2,...,1. H_0 is the input low-resolution image and H_1 is the final high-resolution image to be estimated. The image H_i is generated from $H_{i+1}$by the following operation,

$$H_i= (H_{i+1}*B)\downarrow_s ,$$

where ∗ denotes the convolution operation, B is a blur matrix, ↓ represents a down sampling operation and sis the scale reduction factor between H_i and H_{i+1} . The pyramid P_L is made up of the image set $\{L_i\}$, i=-2…1.Image L_i is generated from H_{i-1} using bicubic interpolation by the scale factor of s. It’s obvious that the pyramid P_H corresponds to pyramid P_L . Since H_1 is the final high-resolution image, we could estimate H_1 by the relationship between H_i and L_i . Compared with prior example based methods, we do not utilize the external image database but only use the input LR image H_0.

B. GMM Training

The pixels with the similar edge tend to have similar neighborhoods whose intensities change fast in the direction perpendicular to that of the edge and pixels in a smooth region seems to have relatively invariant intensities within the neighborhood [4]. The similarity of two pixels can be reflected as the similarity of their corresponding surrounding 8−Neighborhood pixels. In our method we consider the relationship between pixel in pyramid P_H and the surrounding 8−Neighborhood pixels X of the corresponding pixel in P_L . Figure 1(a) shows the structure of a training dataZ . The black point y_H corresponds to y_L , The red points in L_0 represents the 8−neighborhood pixels X for y_L . Each training data consists of y_H and X , so Z can be rewritten as,

$$Z=[Y, X] \quad (8)$$

For each image pair H_i and L_i=−2, −1,0 . We get the training data for each pixel in H_i . Now we define training data as random variables $Z=\{Z_1, Z_2 \ldots\ldots Z_N\}$ where N is the number of training data. We can train a GMM p(Y ,X) for Z.

C. Prediction for SR

We consider the parameter set Θ and rewrite

$\mu_i=[\mu_{ix}, \mu_{iy}], \Sigma_i = \begin{bmatrix} \Sigma_{ixx} & \Sigma_{ixy} \\ \Sigma_{iyx} & \Sigma_{yy} \end{bmatrix}$,we can get

$$p(Y,X) = \sum_{i=1}^{M} \alpha_i G(Z_i, \mu_i, \Sigma_i)$$
$$=\sum_{i=1}^{M} \alpha_i\, G(Y|X, \mu_{iy/x}, \Sigma_{iy/x})\, G(X, \mu_{ix}, \Sigma_{ix}) \quad (9)$$

Where

$$\mu_{iy/x} = \mu_{iy} - \Sigma_{iyx}\, \Sigma_{ixx}^{-1}(\mu_{ix} - X)$$
$$\Sigma_{iy/x}=\Sigma_{iyy} - \Sigma_{iyx}\Sigma_{ixx}^{-1}\Sigma_{ixy} \quad (10)$$

If p(Y\X) is known, the optimal estimation of the pixels Y in H_1 can be given as

$$\hat{Y} = E(Y|X) \quad (11)$$

Under the Minimum Mean Square Error(M.M.S.E.)[2]
Combining the equation (9), we can easily get the p(Y\X) by Bayesian Theorem. That is

$$P(Y\backslash X) = \frac{p(Y,X)}{p(X)} = \sum_{i=1}^{M} \beta_i\, G(Y|X, \mu_{iy/x}\Sigma_{iy/x}) \quad (12)$$

Where

$$\beta_i = \frac{\alpha_i G(Y\backslash X, \mu_{ix},\ \Sigma_{ix})}{\sum_{j=1}^{M} \alpha_j G(X,\mu_{jx}\,\Sigma_{jx})} \quad (13)$$

The optimal for casting$\hat{Y}$ can be represented as the following form [10]:

$$\hat{Y} = E(Y|X) = \sum_{i=1}^{M} \beta_i \mu_{iy/x} \quad (14)$$

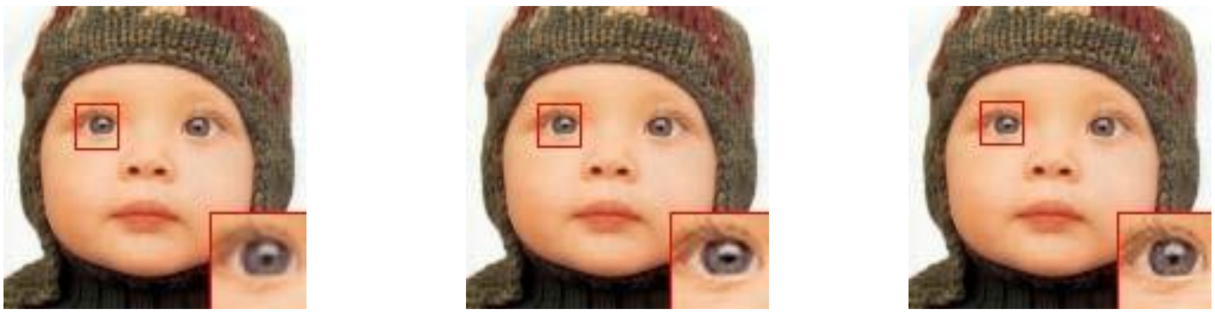
Where the $\mu_{iy/x}$ and β_i are the same as in equation (10) and (13). For each pixel y in the final high-resolution image H_1, it can be predicted by equation (14) and the corresponding vector X .

IV. EXPERIMENTS

Our method has two parameters to determine. The first parameter is the number of component Gaussian densities M in GMM. We set M to 5 for all our experiments. The second parameter is the level K of pyramid. We set K to 4 for preventing the smaller size of the lowest level image. The PSNR and SSIM values of three different test images were displayed in Table I (the scaling factor is 2). To compare our method with other SR methods, we consider bicubic interpolation method, and Sparse-Representation [7]. In Table I, the first row in each method is the PSNR and the second row is the SSIM. We can see that our method obtained higher PSNR and SSIM value than the other two algorithms.

Table.I. PSNR and SSIM values of three images with different algorithms

	tiger	pepper
Bicubic	25.82	26.91
	0.882	0.922
Sparse-Representation[7]	27.70	28.50
	0.916	0.925
Proposed Method	**29.20**	**30.12**
	0.925	**0.945**



(a) bicubic **(b) Sparse-Representation** **(c) Proposed**

Figure 2. Super-resolution (4×) comparison of bicubic interpolation, sparse-Representation [7] and proposed method

Figure 2 shows the results of using different super-resolution algorithms to a tiger image for 4X magnification. We compare our method with the bicubic interpolation and Sparse-Representation [7]. It's clear that our result is better especially in textual parts than the other two algorithms. Our approach produced lessnoise and artifacts.

V. CONCLUSIONS

We proposed an algorithm for single image SR in this paper. We construct a pyramid pair: the ground truth image pyramid and the interpolation image pyramid using only the input LR image without external image database. The pixels in ground truth pyramid and the corresponding 8−neighbor pixels are used to train GMM. The final HR image is predicted through the training GMM. Compared with interpolation-based method and other learning-based method, our approach achieved better result.

VI .REFERENCES

[1]. (2009) D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in IEEE ICCV.

[2]. Guoqiang Yu and Jianming Hu, (June 9-11, 2003) Short-term Traffic Flow Forecasting Based on Markov Chain Model, Intelligent Vehicles Symposium, Proceedings, Page(s): 208-212, IEEE

[3]. (2004) H. Chang, D. Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In CVPR, volume 1, pages 275–282 .

[4]. (2011) He and Wan-Chi Siu, Single Image Super-Resolution Using Gaussian Process Regression. In CVPR, pages 449–456.

[5]. Jeff A. Bilmes. (,Apr. 1998)A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. TR07021, International Computer Science Institute, Berkerley CA.

[6]. (2008) J. Sun, Z. Xu, and H. Y. Shum. Image super-resolution using gradient profile prior. In CVPR..

[7]. (2010) J. Yang, J. Wright, T. Huang and Y.Ma. Image super resolution via sparse representation. IEEE Transactions on Image Processing.

[8].(2008) K.I. Kim, Y. Kwon, "Example-Based Learning for Single-Image Super Resolution," Lecture Notes in Computer Science, Pattern Recognition 30th DAGM Symposium, Proceedings, pp.456–465.

[9]. (2011) M.-C. Yang, C.-H Wang, and Y.-C. F. Wang, "Learning context-aware sparse representation for single image super-resolution," in IEEE ICIP.

[10]. (2008) Qi Shan, Zhaorong Li, JiayaJia, Chi-Keung Tang. Fast Image/Video Upsampling. ACM Trans-actions on Graphics (SIGGRA-H ASIA 2008), Vol. 27.

[11]. (2010). W. S. Tam, C. W. Kok, and W. C. Siu. A modified edge directed interpolation for images. Journal of Electronic Imaging, 19(1), 013011:1–20.