APPROXIMATE ALPHA-STABLE MARKOV RANDOM FIELDS FOR VIDEO SUPER-RESOLUTION

Jin Chen, Jose Nunez-Yanez , and Alin Achim

Department of Electrical and Electronic Engineering, University of Bristol, UK
Email: eexjc@bris.ac.uk, J.L.Nunez-Yanez@bristol.ac.uk, Alin.Achim@bristol.ac.uk

ABSTRACT

In the paper, we present a Bayesian super resolution method that uses an approximation of symmetric alpha-stable ($S\alpha S$) Markov Random Fields as prior. The approximated $S\alpha S$ prior is employed to perform a maximum a posteriori (MAP) estimation for the high-resolution (HR) image reconstruction process. Compared with other state-of-the-art prior models, the proposed prior can better capture the heavy tails of the distribution of the HR image. Thus, the edges of the reconstructed HR image are preserved better in our method. Since the corresponding energy function is non-convex, the iterated conditional modes (ICM) method is used to solve the MAP estimation. Results indicate a significant improvement over other super resolution algorithms.

1. INTRODUCTION

Digital Image/Video Super-Resolution(SR) techniques are widely desirable in many fields, such as medical imaging, military information acquisition and consumer electronics. Physically increasing the pixel density of the charge-coupled device (CCD) arrays can be very expensive [1], therefore people are more in favour of using SR algorithms to increase the resolution of the observed data.

The first SR method proposed by Tsai and Huang [2] used a frequency domain approach based on the shifting property of the continuous Fourier transform. Some assumptions were made in their experiments, such as, *e.g.*, the purely translational motion of frames is known, and there is no blur. Later research was extended to consider the observation noise and the spatial blurring, and many spatial based SR approaches have been proposed to overcome the drawbacks of the frequency domain method. Spatially reconstructing high resolution images from their low resolution versions is an ill posed problem, which is solved by combining multiple successive sub-pixel shifted low resolution frames. Due to the presence of sub-pixel displacement, it is possible to obtain high frequency content beyond the Nyquist limit of the sampling equipment.

In last decade, SR methods using the Bayesian frameworks have become central to the design of novel SR algorithms. Since Bayesian methods include explicit prior constraints on the solution, an important idea is to utilize a constraint, as close to the real distribution of the original scene as possible. Among the Bayesian SR methods, Gaussian Markov Random Fields (GMRF) [3] is one of the mostly used priors. Actually, in image reconstruction, the choice of prior is of crucial importance for edge preservation. Farsiu et al. proposed the bilateral total variation (BTV) prior; this method is declared to have a good noise suppression ability, however the results are always blurred. In [4], the authors have adopted the Generalized Gaussian as prior, due to the heavy tail property it possesses, and this method can offer a good enhancement of visual quality of the super-resolved image.

In this paper, we adopt a Gaussian Scale Mixture approximation for α-stable prior, under the Markov Random Fields (MRFs) framework. The proposed approach enables the use of a Bayesian probabilistic image processing framework. The remaining paper is organized as follow: Section 2 briefly reviews Bayesian super-resolution together with a description of some standard priors. Section 3 gives some necessary preliminaries on the α-stable model, as well as details on how to approximate the α-stable using the *GSM* model. Both qualitative and quantitative results for super-resolved video frames are presented in Section 5. Finally, we conclude the paper in Section 6.

2. BAYESIAN SUPER-RESOLUTION

Suppose we are given F low resolution images and each image contains arbitrary motion relative to a reference frame. Every low-resolution (LR) image is represented as vectors, denoted as $Y_1...Y_f$. The aim of the SR method is to infer the high-resolution frame X from the observed low-resolution images. In this Section we study Super-Resolution as an inverse problem and address related regularization issues.

Fig. 1 shows diagrammatically the structure of our SR algorithm. The proposed method estimates the parameters for the $S\alpha S$ prior and then approximates it with a *GSM* model. This step improves the quality of the reconstructed image, by helping preserve edges better. There are many motion estimation methods proposed available, such as Block Matching, Pyramidal Lucas-Kanade and Combined Local Global (CLG)

ISSN 2076-1465

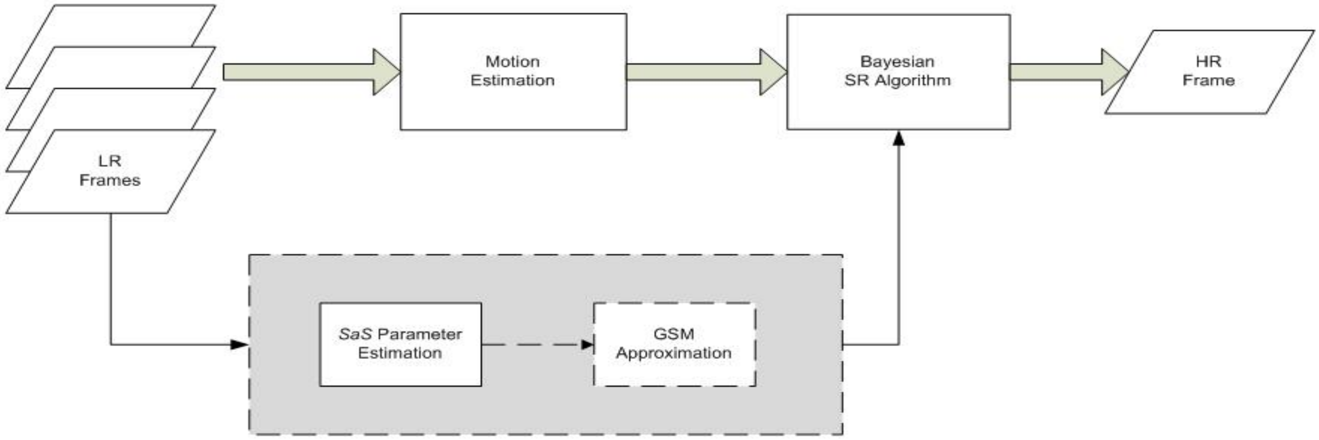


Fig. 1. Block diagram of the SR process

optical flow. Among those methods, we adopted the CLG optical flow as it was shown to outperform the other two methods [5] [4].

There are two fundamental terms included in a typical Bayesian SR framework, namely the conditional probability (likelihood) $P(y|X)$ and the prior $P(X)$. If we use X to represent the original HR image and $\widehat{X}$ corresponds to the estimated HR image, then Bayes SR can be derived starting from Bayes theory:

$$P(X|y,d) = \frac{P(X,d)P(y|X,d)}{P(y)} \quad (1)$$

where d is the motion vector and y are the observed images. $P(y)$ is called the evidence, and can be neglected as it is a constant value [4]. Therefore, the high-resolution image can be obtained by maximising $P(X|d,y)$, which constitutes the well-known MAP estimator:

$$\begin{aligned}\widehat{X} &= \arg\max_X P(X,d)P(y|X,d) \\ &= \arg\min_X \left(-logP(X,d) - logP(y/X,d)\right) \quad (2)\end{aligned}$$

Hence, we will have to specify the prior $P(X,d)$ to incorporate with the likelihood part $P(y/X,d)$.

Most state-of-the-art Bayesian SR approaches are based on a generative model for the observed low resolution images, comprising a prior over the high resolution image together with an observation model. The prior is also known as a regularization term. Among previously proposed priors, we consider two of them in the rest of the paper. Farsiu *et al.* proposed the BTV prior in [6]:

$$P(X) = \frac{1}{\alpha}exp\{\beta \sum_{n,m=-l}^{l} \gamma^{|n|+|m|}||X - S_x^n S_y^m X||_1\} \quad (3)$$

where α, β and γ are parameters of the prior, S_x^n S_y^m are shifting matrices that shift the image horizontally and vertically by n and m pixels respectively [6] [4]. This model can penalise the energy function by comparing the estimated HR image to versions of itself shifted by an integer number of pixels in various directions. In [4], authors is motivated by recent progresses on natural image statistics. The gradient magnitudes generally obey a heavy tailed distribution [7], hence GGMRF prior was used and described as:

$$P(X) = \frac{1}{\alpha}exp\{\beta \sum_{s,r\epsilon\ \Omega} ||g_{s,r}||_q\} \quad (4)$$

The parameter β determines the influence of the prior, while $||\cdot||_q$ refers to the L_q norm with q being constrained between 1 and 2. $g_{s,r}$ refers to the first order neighbourhood of pixels, while s and r indicate the location of these two pixels.

3. THE SYMMETRIC ALPHA-STABLE PRIOR

Recently, work on non-Gaussian modelling for image processing gained increasing interest in the research community. In [8], for image restoration, the authors employed an α-stable distribution to better capture the heavy-tailed nature of the data as well as the inter-scale dependencies of wavelet coefficients. The algorithm successfully removes noise from digital images, while preserving the visual quality of the image very well. α-stable distribution does not always have a closed-form expression, which makes it difficult to use in a Bayesian framework. To overcome this problem, Kuruoglu *et.al* have proposed an analytical approximation for α-stable probability density functions in [9], by using Gaussian Scaled Mixtures (GSM). Here, we model images with approximated *SαS* Markov Random Fields (MRFs).

A general stable distribution is determined by four parameters: shape parameter α is in $(0,2]$, also known as characteristic exponent. This is the most import parameter of stable distribution; the smaller the characteristic exponent α is, the heavier the tails of the $S\alpha S$ density. The skewness parameter β is in $[-1, 1]$ and measures the asymmetry of the distribution (which in our case is set to zero). Scale parameter c and location parameter μ indicate the width and the location of the distribution respectively. The characteristic function is used to define the general $S\alpha S$ PDF:

$$\varphi(\omega) = exp(j\mu\omega - c|\omega|^\alpha) \tag{5}$$

The $S\alpha S$ model is suitable for describing variables with heavier tails then what is assumed by exponential families. This is because the $S\alpha S$ distribution follows an algebraic rate of decay, whereas the decaying occurring in the GGD is exponential.

Nolan [10] describes a consistent maximum likelihood method to estimate parameters of $S\alpha S$, which gives reliable estimates and provides the most tight confidence intervals. Hence, we first estimate the parameters of the image gradient distribution by using Nolan's method [10]. The method in [9] is based on the corollary of mixing property of α-stable random variables. Specifically, [11] states that any $S\alpha S$ random variables can be written as the product of a Gaussian random variable and a positive stable random variables. In [12] this theorem is mathematically given as:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{Z|V}(z|v) f_V(v) J(z,v) dv \tag{6}$$

where $f_Z(z)$ is the $S\alpha S$ distribution we want to approximate, $f_{Z|V}(z|v)$ follows a Gaussian distribution with $f_V(v)$ being its weighting parameter. Herein $f_V(v)$ is a positive stable random variable, also referred to as the mixing function [9] [11]. The last term $J(z,v)$ represents the Jacobian of Z with respect to V. The analytic expression proposed in [9] for the *SaS* approximation is thus:

$$P_{\alpha,\gamma,\beta=0,\mu}(z) = \frac{\sum_{i=1}^{N} \frac{1}{\sqrt{2\pi vi}} exp(-\frac{(z-\mu)^2}{2vi^2}) f_V(vi)}{\sum_{i=1}^{N} f_V(vi)} \tag{7}$$

where the denominator is the normalization constant.

4. *SAS*-MRF SUPER-RESOLUTION ALGORITHM

In this section with provide details on our proposed $S\alpha S$ MRF SR method. In accordance with the Bayesian model described in Sec. 2, the likelihood term can expressed as:

$$P(Y|X) \propto \sum_{f=1}^{F} \|H_f X - Y_f\|^p \tag{8}$$

Parameter p refers to L_p norm. Note that when $p = 2$ the likelihood term corresponds to a Gaussian. H_f is a sparse matrix which is used to represent the blurring and downsampling process. The estimated motions of LR frames, against its reference frame, are used in H_f. This is in order to use additional information from other close frames in the super-resolution process. After including the regularization term, a HR image solution can be obtained by minimizing the cost function:

$$\widehat{X} = \arg\underset{X}{\text{Min}} \sum_{f=1}^{F} \|H_f X - Y_f\|^p + \beta\Psi(X) \tag{9}$$

In equation (9), function Ψ represents the prior/regularization term. In our case, this takes the form of a Gaussian Scale Mixture $S\alpha S$ approximating a $S\alpha S$ density:

$$\Psi(X) = \sum_{i=t}^{N} \frac{\omega_t}{\sqrt{2\pi}\delta_t} exp(-\frac{(\phi(X))^2}{2{\delta_t}^2}) \tag{10}$$

where $\phi(X)$ represents the adjacent pixel gradient in two directions (horizontal and vertical), ω_i is the weighting/scaling parameter for each Gaussian component. These parameters are estimated as explained in previous section.

Finally, the proposed alpha-stable MRF super resolution method is implemented through the following equation:

$$\widehat{X} = \arg\underset{X}{\text{Min}} \sum_{f=1}^{F} \|H_f X - Y_f\|_p^p + \beta \sum_{i=1}^{z^2 MN} (W_N \cdot G_N(\phi_i(X))) \tag{11}$$

where β is a tuning parameter for the prior term. W_N is a ten elements scaling vector used to form the approximated $S\alpha S$ and G_N represents the corresponding ten Gaussian distributions. Both W_N and G_N are obtained using equation (7). Because the cost function is not convex, gradient descent optimization is not suitable for finding a global minimum of the objective function. Hence we have to resort to other, non convex methods. Metropolis algorithm and Gibbs sampler are two often used random search methods. In order to help the algorithm jump out from a local minimum into a global minimum, same strategies employed in these method include: not always pursuing a descented energy value, as occasionally an energy increase is allowed.

It is well known that when a random search method is employed for MAP estimation, the computational cost becomes enormous. The Iterated Conditional Mode (ICM) algorithm presented in [13] is a computationally acceptable alternative in solving MAP super-resolution problems. ICM is employed in our method as part of a succession of steps including:

1. Estimate the $S\alpha S$ distribution's parameters from the LR frame. Find the weighting parameter of the Gaussian scale mixture by using the method described in [9], and constructing the analytic expression of the $S\alpha S$ distribution.

2. Initial guess the HR frame from LR frame by using cubic-interpolation, and estimate the sub-pixel motion of LR frames against the reference frame.

3. For i from 1 to z^2MN, update the value of X_i that minimises:

$$\widehat{X} = \arg\min_X \Bigg((Y - \boldsymbol{H}X)^2 + \beta \sum_i^{z^2MN} (ln(\sum_{t=1}^{N} \frac{\omega_t}{\sqrt{2\pi\delta_t}} exp(-\frac{(\phi_i(X))^2}{2{\delta_t}^2}))) \Bigg) \quad (12)$$

4. Repeat stage (3) *iter* times. We note that in most of the experiments, 5 or 6 iterations were sufficient to produce good results.

Throughout the experiments, the parameter β is fixed at 0.03 suggesting robustness of the algorithm over various images. In the first step of approximating the mixtures, using ten Gaussian terms is enough to provide a good $S\alpha S$ approximation. In equation 12, X and Y are both rearranged in lexicographic order, matrix *H* represents the blurring and decimation operator.

5. EXPERIMENTAL RESULTS

We assess the performance of our proposed SR algorithm by calculating the peak signal-to-noise ratio (PSNR), defined as

$$PSNR = 10log_{10}\frac{NP}{\|a-b\|} \quad (13)$$

where a and b are the estimated and original HR images, and the pixel values are normalized to be in [0,1]. Apart from measuring the PSNR of super resolved images, the structural similarity (SSIM) [14] was also used to measure the reconstruction quality. Compared to the more traditional PSNR and mean squared error (MSE), SSIM has been proven to be more consistent with human eye perception. The SSIM is given by

$$SSIM(a,b) = \frac{(2\mu_a\mu_b + c_1)(2\delta_{ab} + c_2)}{(\mu_a^2 + \mu_b^2 + c_1)(\delta_a^2 + \delta_b^2 + c_2)} \quad (14)$$

where μ_a and μ_b represent the mean values, δ_a and δ_b are their variances, and δ_{ab} is the covariance. The SSIM measure should be close to unity for an optimal effect of SR reconstruction. Our SSIM results and the corresponding PSNR values are summarized in Table. 1.

For comparison, we considered a number of methods including: the Gaussian Markov Random Field Super Resolution [3], (*GMRF*); the robust SR method based on bilateral TV priors [6] (*BTV*); Generalized Gaussian MRF Super Resolution [4](*GGMRF*).

In the experiments real life video data was used. Video sequence 'Lab' is shot by a handheld camera. The resolution of it is normalized to 180×144. To simulate the effect of camera point spread function, the HR images were convolved with a symmetric Gaussian low-pass filter of size 3×3 with standard deviation equal to one. The camera blurred images were downsampled by a factor of 2 in each direction. 20 adjacent low resolution frames were used to obtain a single super resolved frame.

The results are shown in Fig.2. The enhancement of the resolution of proposed method over the original image can be clearly noticed. Moreover, the proposed SR method exhibit clear visual quality beyond other methods. As what we are expecting, the objects edge are more clear after the SR process. Table. 1 indicates that our method outperforms other methods in terms of both PSNR and SSIM.

6. CONCLUSIONS

In this paper, we presented a novel method for the problem of increasing the spatial resolution of video frames. The Bayesian super-resolution technique is studied, together with four different prior models: approximated $S\alpha S$, GGMRF, GMRF and BTV. Besides, some basics of the alpha stable distribution are also introduced in paper. Simulation results have shown the superiority of the proposed algorithm with respect to three the recently proposed techniques. In future, we plan to extend our framework to more general super resolution applications characterised by more complex motion models.

7. REFERENCES

[1] Xuelong Li, Yanting Hu, Xinbo Gao, Dacheng Tao, and Beijia Ning, "A multi-frame image super-resolution method," *Signal Processing*, vol. 90, no. 2, pp. 405 – 414, 2010.

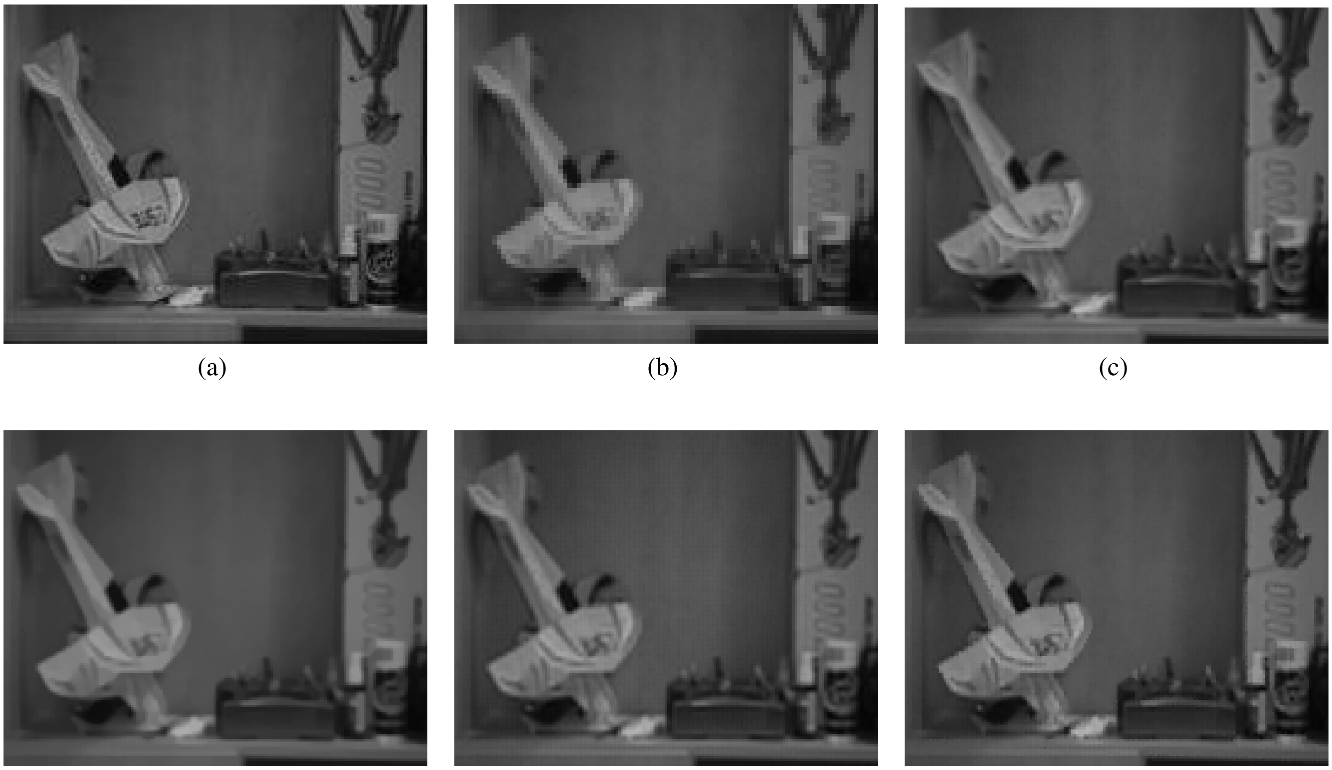
[2] RY Tsai and T.S. Huang, "Multiframe image restoration and registration," *Advances in computer vision and image processing*, vol. 1, no. 2, pp. 317–339, 1984.

[3] P. Cheeseman, R. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," *fundamental theories of physics*, vol. 62, pp. 293–308, 1996.

[4] J. Chen, J. Nunez-Yanez, and A. Achim, "Video super-resolution using generalized gaussian markov random fields," *Signal Processing Letters, IEEE*, vol. 19, no. 2, pp. 63 –66, feb. 2012.

[5] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/kanade meets horn/schunck: Combining local and global optic flow methods," *International Journal of Computer Vision*, vol. 61, no. 3, pp. 211–231, 2005.

[6] Michael Elad Sina Farsiu, M. Dirk Robinson and Peyman Milanfar., "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, pp. 1327–1344, october 2004.



(a) (b) (c) (d) (e) (f)

Fig. 2. LR and super-resolved frame (a)original (b)LR (c)GMRF (d)BTV (e)GGMRF (f)Proposed

Table 1. SR Results Lab

Case	LR	GMRF	BTV	GGMRF	Proposed
PSNR(in dB)	29.4628	33.7012	34.1664	34.8348	**35.0992**
SSIM	0.8673	0.9359	0.9373	0.9438	**0.9506**

[7] Jinggang Huang and D. Mumford, "Statistics of natural images and models," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, 1999, vol. 1, pp. 2 vol. (xxiii+637+663).

[8] A. Achim and E.E. Kuruoglu, "Image denoising using bivariate alpha-stable distributions in the complex wavelet domain," *Signal Processing Letters, IEEE*, vol. 12, no. 1, pp. 17 –20, Jan. 2005.

[9] EE Kuruoglu, C. Molina, and WJ Fitzgerald, "Approximation of α-stable probability densities using finite mixtures of gaussians," in *Proc. EUSIPCO98 Eur. Signal Processing Conf*, pp. 989–992.

[10] J.P. Nolan, "Numerical calculation of stable densities and distribution functions," *Communications in Statistics-Stochastic Models*, vol. 13, no. 4, pp. 759–774, 1997.

[11] G. Samorodnitsky, M.S. Taqqu, and RW Linde, "Stable non-gaussian random processes: stochastic models with infinite variance," *Bulletin of the London Mathematical Society*, vol. 28, no. 134, pp. 554–555, 1996.

[12] D.F. Andrews and C.L. Mallows, "Scale mixtures of normal distributions," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 99–102, 1974.

[13] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 259–302, 1986.

[14] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.