SUPER-RESOLUTION IMAGE RECONSTRUCTION USING THE ICM ALGORITHM

Martins, A. L. D.[*] *and Homem, M. R. P.*[†] *and Mascarenhas, N. D. A.*

Universidade Federal de São Carlos, Departamento de Computação
Via Washington Luís, Km 235, CP 676, CEP: 13.565-905, São Carlos, SP, Brazil
ana_martins@dc.ufscar.br; murillo_rodrigo@dc.ufscar.br; nelson@dc.ufscar.br

ABSTRACT

Super-resolution image reconstruction is a powerful methodology for resolution enhancement from a set of blurred and noisy low-resolution images. Following a Bayesian framework, we propose a procedure for super-resolution image reconstruction based on Markov random fields (MRF), where a Potts-Strauss model is assumed for the *a priori* probability density function of the actual image. The first step is given by aligning all the low-resolution observations over a high-resolution grid and then improving the resolution through the Iterated Conditional Modes (ICM) algorithm. The method was analyzed considering a number of simulated low-resolution and globally translated observations and the results demonstrate the effectiveness of the algorithm in reconstructing the desirable high-resolution image.

Index Terms— Markov random fields, Iterated Conditional Modes, super-resolution image reconstruction.

1. INTRODUCTION

In this work we consider the problem of resolution enhancement from a set of blurred and noisy low-resolution (LR) images that are also corrupted by aliasing. Indeed, it can be shown that if the set of LR images are corrupted by aliasing and also have different sub-pixel shifts from each other, the different information contained in each of them can be exploited to obtain a high-resolution image [1]. This problem is commonly referred to *high-resolution* (HR) or *super-resolution* (SR) *image reconstruction* [2, 3].

Several algorithms were proposed in the last years for SR image reconstruction [1, 3]. In general, as a preliminary classification, they can be divided in spatial or frequency domain approaches. It is well known that SR reconstruction algorithms through frequency domain are simpler and have more intuitive SR mechanism than that derived in the spatial domain [2]. Tsai and Huang [4] were the first to restore a high-resolution image from a sequence of low-resolution, undersampled, discrete frames with same displacement between each other. They used a frequency domain approach based on the shifting property of the continuous Fourier transform to restore the high-resolution image. In that work, each low-resolution image was seen as the same signal, but shifted by different quantities, i.e., the motion was considered purely translational. Thus, they found more signal frequency components, increasing the resolution. A remarkable point is that they did not consider blur and noise on their work. Later, other works like [5], extended that frequency domain approach to include blurred and noisy low-resolution images. Since [4], several resolution enhancement approaches have been developed, most of then using a spatial domain context. In fact, spatial domain methods are able to work with more general observations models such as spatially varying blurring [1]. Another advantage provided by the spatial approaches is, for instance, the capability for *a priori* constraints inclusion. The most important results were acquired by Irani and Peleg [6], which turned that work a reference on resolution enhancement problems. They used an iterative spatial domain approach similar to the back-projection algorithms employed on computed tomography reconstruction algorithms.

Following a Bayesian framework, we propose an alternative algorithm for SR image reconstruction based on a Markov random field (MRF), where a Potts-Strauss model is assumed for the *a priori* probability density function of the actual image. Given a set of LR images from the same scene (sometimes referred to as *observations*), under the assumption that there exist sub-pixel displacements from each other, we intend to determine these relative displacements among the under-sampled observations and then reconstruct an image on a HR grid using the Iterated Conditional Modes (ICM) algorithm.

2. THE PROPOSED METHOD

2.1. IMAGE FORMATION MODEL

Consider the high-resolution version $f[i, j]$, $0 \le i, j < M$, of a continuous signal $f : \Re^2 \to \Re$ that represents the scene of interest. Following a lexicographic ordering of $f[i, j]$, a model for the low-resolution version $d[k, l]$, $0 \le k, l < N$,

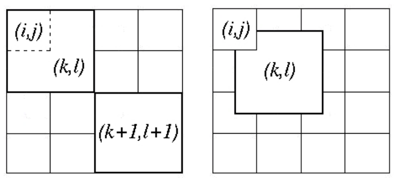
[*]The work of Ana L. D. Martins was supported by a CAPES scholarship, Brazil.

[†]The work of Murillo R. P. Homem was supported by FAPESP, Brazil, grant numbers 04/01632-1 and 2002/07153-2.

$N < M$, of $f[i, j]$, can be given by

$$\mathbf{d} = \mathbf{D}\mathbf{f}, \tag{1}$$

where $\mathbf{d} \in \Re^{N\times N}$ is the vector with components given by $d_n = \sum_m \delta_{n,m} \cdot f_m$, f_m are components of the vector $\mathbf{f} \in \Re^{M\times M}$ and $\delta_{n,m}$ are elements of the down-sampling operator $\mathbf{D}$ of size $M^2 \times N^2$. In this sense, the LR pixels are defined as a weighted sum of the related HR pixels. Figure 1 illustrates these possibilities. Therefore, the $\mathbf{D}$ operator under-samples



(a) (b)

Fig. 1. (a) The HR pixel (i,j) is completely within the LR pixel (k,l) and completely outside the pixel (k+1,l+1); (b) the HR pixel (i,j) is partially within the LR pixel (k,l).

the vector signal $\mathbf{f}$ yielding the observation vector $\mathbf{d}$.

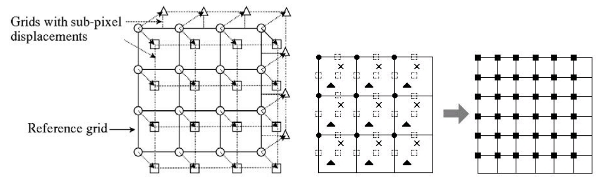
In a more realistic approach, digital images are often blurred by the optical system during the acquisition procedure [1] and also corrupted by noise. In that case, frequently, the blurring process is regarded as a linear, space-invariant operator and then a blurred vector image $\mathbf{b}$ is given by $\mathbf{b} = \mathbf{H}\mathbf{f}$, where $\mathbf{H}$ is the $M^2 \times M^2$ block-circulant matrix that gives the blurring degradation effects with elements given by samples of the point spread function (PSF) of the optical system. Hence, after the blurring, a LR version of a HR image $\mathbf{f}$, may be modelled as

$$\mathbf{d} = \mathbf{D}\mathbf{H}\mathbf{f} + \mathbf{n}, \tag{2}$$

where $\mathbf{n}$ stands for the noise in the observations, following an additive model.

2.2. IMAGE REGISTRATION

The first step in the proposed algorithm is to find an estimate of the actual image based on a sub-pixel image registration procedure. Indeed, SR image reconstruction is proved to be possible if multiple LR images of the same scene can be obtained [7], where the images are necessarily shifted with sub-pixel precision from each other. Therefore, a first estimate of the image can be obtained by align all the LR observations on a HR grid. Figure 2(a) illustrates the case for three images with different sub-pixel displacements from each other, where the images are aligned following a reference frame. Moreover, since the displacements can be different between each observation and the reference image, a non-uniform interpolation approach can be used to find out an estimate of the HR image. Figure 2(b) illustrates the last case. For the purpose of this work, in the image registration step, we have used



(a) (b)

Fig. 2. (a) Images with sub-pixel displacements; (b) interpolation on a high-resolution grid.

the procedure discussed in [8] that is a variant of the method proposed by Irani and Peleg [6].

2.3. BAYESIAN FORMULATION FOR SR

It is well known that HR image reconstruction is an ill-posed problem. Thus, some kind of regularization is need to reach a good approximation of the actual image. In fact, we need to find an estimate $\hat{\mathbf{f}}$ of the HR image given a set of LR observations $\mathbf{d_t}$, $t = 1, \cdots, T$, each of them modeled by equation (2). A Bayesian formulation of the problem provides a flexible and convenient way of using the *a priori* information to achieve a good approximation of the original scene. In this sense, the maximum *a posteriori* probability (MAP) solution choose the estimate that maximizes the conditional probability density of $\mathbf{f}$ given all the observations, that is,

$$\hat{\mathbf{f}} = arg\, max_{\mathbf{f}}\{p(\mathbf{f} \mid \mathbf{g})\}, \tag{3}$$

where $\mathbf{g}$ is constructed with all the LR observations $\mathbf{d}$.

2.4. THE ICM ALGORITHM

The Iterated Conditional Modes algorithm was proposed by Besag [9] as a computationally feasible alternative in computing the maximum *a posteriori* probability (MAP) for the actual image given the observations. Indeed, it is known that MAP algorithms make enormous computational demands due to the inherent difficulty in computing the MAP estimate. Further, close related to Markov random fields (MRF), the ICM algorithm is not only computationally undemanding but also ignores the large-scale deficiencies of the *a priori* probability for the true image [9]. It is an iteractive procedure and it is easily shown that for each iteration, the MAP estimate never decreases and eventual convergence is assured. The method is based on the equation (4) for the *a posteriori* probability of the value of the pixel i, given the observations $\mathbf{g}$ and the current values of all pixels in the neighborhood of the pixel i.

$$p(f_i \mid \mathbf{g}, \hat{\mathbf{f}}_{S\setminus i}) \sim p(g_i \mid f_i) \cdot p(f_i \mid \hat{\mathbf{f}}_{\partial_i}) \tag{4}$$

In the above equation $S \setminus i$ represents the set of *all neighbors* of the pixel i and ∂_i a small set of neighbors of the same pixel, defined by a neighborhood system. The usual neighborhood system in image analysis defines the first-order neighbors of a pixel as the four pixels sharing a side with the given pixel. Second-order neighbors are the four pixels sharing a corner. Higher order neighbors are defined in an analogous manner [10]. In this sense, $\hat{\mathbf{f}}_{S\setminus i}$ is the vector of all current values of the image excluding the pixel i and $\hat{\mathbf{f}}_{\partial_i}$ is a vector of some neighbors of f_i, following a neighborhood system.

Although it is proposed inside a Bayesian framework, the ICM is a deterministic algorithm and it is given by

1. Choose a MRF model for the true values of f_i;

2. Initialize $\hat{\mathbf{f}}$ by choosing f_i as the intensity $\hat{f}_i$ that maximizes $p(g_i \mid f_i)$ for each i;

3. For i from 0 to $M^2 - 1$, update $\hat{f}_i$ by the value of f_i that maximizes

$$p(g_i \mid f_i) \cdot p(f_i \mid \hat{\mathbf{f}}_{\partial_i})$$

4. Repeat item (3) τ_{iter} times.

2.5. THE POTTS-STRAUSS MODEL

The choice of a realistic image model is a critical step on the estimation process. Since MRF allows the introduction of context information presented on the observations, in this work, we assume a Potts-Strauss model for the *a priori* probability density function of the actual image. The Potts-Strauss model can be defined by the set of all the conditional distributions given by

$$p(f_i \mid \hat{\mathbf{f}}_{\partial_i}) \sim e^{\beta \cdot \sharp\{t \in \partial_i / f_t = f_i\}}, \tag{5}$$

where $\beta \in \mathfrak{R}$ is often referred to as the attraction or repulsion parameter if it is positive or negative, respectively [11].

Moreover, in the ICM algorithm, we also need to know $p(g_i \mid f_i)$, where in our case the vector $\mathbf{g}$ is constructed from all the observation vectors $\mathbf{d}_t$. We assume that $p(g_i \mid f_i)$ may be given by

$$p(g_i \mid f_i) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(g_i - m_i)^2}{2\sigma^2}}, \tag{6}$$

where $m_i = \frac{1}{C}(\sum_{j\in\partial_i} f_j) + \frac{1}{C} f_i$ and $C = \sharp\partial_i + 1$.

It is important to note that image models based on a MAP-MRF formulation usually implies in a uniform smoothness of the image. This oversmoothness do not respect discontinuities, where abrupt changes occur. According to [12], how to apply the smoothness constraint while preserving discontinuities has been an active research area. Since [13] introduced line fields to preserve edges, discontinuity-adaptive methods control the interaction between neighbors in such a way that when a discontinuity is detected, the degree of interaction is adjusted not to smooth the area.

3. RESULTS

An evaluation of the proposed method has been conducted by processing a set of simulated LR images. Figure 3(a) shows the 512x512 pixels image that was considerated the actual HR image. The image was convolved with a 3x3 uniform rectangular kernel to simulate the blur due to the image process acquisition. Then, 16 LR images were obtained by using the image formation model presented in Section 2.1. In order to simulate sub-pixel displacements, each of the 16 down-sampling operators under-sampling the HR blurred image by 4 in each direction, 16 times, each time starting from a different pixel within the first 4x4 block [14]. Later, each LR image of 128x128 pixels was contaminated by additive and independent Gaussian noise at 40 dB. Figure 3(b) shows the first LR image. It was considerated as the reference image for the sub-pixel image registration procedure. For comparison proposes, Figure 3(c) presents the result of the bilinear interpolation of the reference image and Figure 3(d) shows the result from the registration of all the LR observations on a grid of 512x512 pixels.

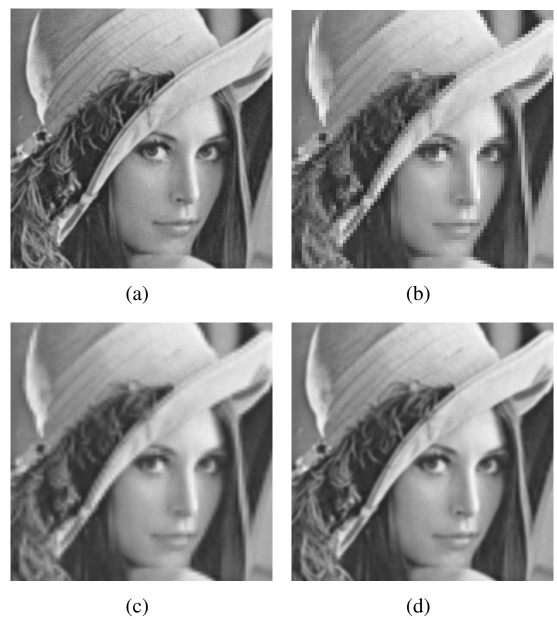


Fig. 3. (a) Actual image; (b) reference LR observation; (c) bilinear interpolation of the reference image; (d) first estimate of the HR image based on the sub-pixel image registration procedure.

As one can see, the registration procedure is able to give better results when compared with the interpolated image. Although in this simulation we have knowledge of the actual displacements between each LR image and the reference image, we have estimate the displacement values. We note that the proposed method for sub-pixel registration has demonstrated to be very accurate in all conducted experiments.

In figure 4(a) we present the HR estimate reconstructed using the proposed algorithm without a discontinuity-adaptive method and in figure 4(b) the result using a discontinuity-adaptive procedure. In both the simulations, the algorithm was initialized with the image in figure 3(d) and the β parameter in equation (5) was found following the procedure proposed in [11]. Also, in this experiment we do not take into account the blur from the optical system in the restoration process. From the presented results, we can see that the algorithm was able to improve the quality of the initial estimate. We also note that in the most of the experiments, the algorithm had a fast convergence rate, where 5 or 6 iterations were sufficient to produce good results.

(a) (b)

Fig. 4. (a) HR estimate without a discontinuity-adaptive procedure; (b) HR estimate with a discontinuity-adaptive procedure.

4. CONCLUDING REMARKS

We have presented an efficient algorithm for SR image reconstruction based on a Markov random field where we used the Iterated Conditional Modes algorithm for computing the maximum *a posteriori* conditional probability. Indeed, the results demonstrate that the algorithm can be extremely efficient in a SR reconstruction framework where the method has demonstrated good performance both in visual accuracy and computational cost. We also note that, although we do not address the image debluring procedure in this work, it can be easily incorporated into the proposed algorithm. In future works, we intend to make additional experiments in order to verify the accuracy of the proposed method when compared with the Irani-Peleg algorithm and also considering different levels for the signal to noise ratio in the observations. We also intent to test the algorithm with other models for the *a priori* probability density function of the actual image.

5. REFERENCES

[1] S. Borman and R. Stevenson, "Spatial resolution enhancement of low-resolution image sequences. A comprehensive review with directions for future research," *Technical Report, University of Notre Dame, Notre Dame*, 1998.

[2] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, pp. 21–36, 2003.

[3] D. Lim, "Achieving accurate image registration as the basis for super-resolution," *Master Thesis, The University of Western Australia*, 2003.

[4] R. Y. Tsai and T. S. Huang, "Multi-frame image restoration and registration," *Advances in Computer Vision and Image Processing*, pp. 317–339, 1984.

[5] S. P. Kim and W. Y. Su, "Recursive high resolution reconstruction of blurred multiframe images," *IEEE Trans. on Image Processing*, vol. 2, pp. 534–539, 1993.

[6] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graph. Models Image Process.*, vol. 53, pp. 231–239, 1991.

[7] S. Chaudhuri, "Super-resolution imaging," *Norwell, MA: Kluwer*, 2001.

[8] M. R. P. Homem, A. L. D. Martins, and N. D. A. Mascarenhas, "High-resolution image reconstruction with the ICM algorithm," *WVC2006 - II Workshop de Visão Computacional*, pp. 123–128, October 2006.

[9] J. Besag, "Statistical analysis of dirty pictures," *Journal of the Royal Statistical Society*, vol. 48, pp. 259–302, 1986.

[10] R. C. Dubes and A. K. Jain, "Random field models in image analysis," *Journal of Applied Statistics*, vol. 16, pp. 131–164, 1989.

[11] N. D. A. Mascarenhas and A. C. Frery, "SAR image filtering with the ICM algorithm," *International Geoscience and Remote Sensing Symposium*, pp. 2185–2187, August 1994.

[12] S. Z. Li, "On discontinuity-adaptive smoothness priors in computer vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17(6), pp. 576–586, June 1995.

[13] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6 (6), pp. 721–741, November 1984.

[14] H. Stark and Y. Yang, "Vector space projections, a numerical approach to signal and image processing, neural nets, and optics," *A Wiley-Interscience Publication*, 1998.