High-Resolution Image Reconstruction with the ICM Algorithm

Homem, M. R. P.; Martins, A. L. D.; Mascarenhas, N. D. A.
Departamento de Computação, Universidade Federal de São Carlos
Via Washington Luís, Km 235, CP 676, CEP: 13.565-905, São Carlos, SP, Brazil
murillo@if.sc.usp.br; analu_bcc@yahoo.com.br; nelson@dc.ufscar.br

Abstract

In this work, we consider the problem of resolution enhancement from a set of blurred and noisy low-resolution images with different sub-pixel displacements from each other. We propose a new procedure for high-resolution image reconstruction based on both the Iterated Conditional Modes (ICM) algorithm and a new approach for sub-pixel image registration. The first step is given by aligning all the low-resolution observations over a high-resolution grid and then improve the resolution through the ICM algorithm, where a Potts-Strauss model is assumed for the \emph{a priori} probability density function of the actual image. The method was analyzed considering a number of globally translated and low-resolution observations and the results demonstrate the power of the algorithm in reconstructing a high-resolution image.

1 Introduction

The number of pixels per unit area in a digital image is one of the most significant factors in determining the quality of a two-dimensional image in various applications such as in remote sensing, medical imaging, computer vision or video applications, just to point out a few. For instance, in computed tomography (CT), an image of higher resolution than that provided by the CT scanner can be desirable since it can reveal more detailed information about the scene. Indeed, small structure details can often supply additional information for the image analysis and precise diagnoses. Commonly, in digital image processing, the data are acquired by using sensor arrays that consist of many individual detectors such as in charged-couple devices (CCD) cameras. Then, the images are obtained directly from the data observations, or indirectly, as in CT scanning where the final results are images reconstructed from the recorded projections data. In this sense, a first approach to improve the resolution of an image is increasing the number of detectors and also reducing the size of them. However, it has some practical limitations and subsequently the size of each detector cannot be made arbitrarily small [1]. Therefore, a resolution enhancement approach using signal-processing techniques can be a useful choice toward increasing the spatial resolution in a digital image.

In this work we consider the problem of resolution enhancement from a set (or sequence) of blurred and noisy low-resolution (LR) images (sometimes referred to as \emph{observations}) that are also corrupted by aliasing [2]. The problem of achieving a high-resolution (HR) image from a set of observed LR ones is referred to high-resolution or super-resolution (SR) image reconstruction in the image processing literature [1, 2, 3]. It can be shown that if the 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 HR image [1, 2]. Following [1], the methodology for HR image reconstruction can be seen as a three-stage process. Initially, the LR images must be registered with sub-pixel accuracy in relation to a reference image that, usually, is one of the available LR observations. Following that, a HR grid from the set of aligned LR images are generated and, eventually, the deblurring process is performed to restore the actual image. We propose both an alternative algorithm for HR image reconstruction based on the Iterated Conditional Modes (ICM) algorithm and an efficient way to registrate translated images with sub-pixel accuracy. As can be seen, we are concerned with the first and second step of the general framework for HR [1], i.e., given a set of LR 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. Notwithstanding, although we do not address the image deblurring procedure in this work, we note that it can be easily incorporated into the proposed algorithm.

In Section 2, we describe the problem of the HR image reconstruction. Particularly, in Section 2.1, we show a straightforward way to improve the resolution through the image registration from a set of observations and in Section 2.2 we present a LR image formation model from a HR one. Both the sections will be important for the subsequent ones. In Section 3 we present the proposed methodology where we describe the method for sub-pixel registration in Section 3.1. Then, we briefly describe the principles of the ICM algorithm in Section 3.2 and in Section 3.3 the algorithm for HR image reconstruction is presented. Eventually, results and discussions are presented in Section 4. Conclusions of the work are presented in Section 5.

2 SR Image Reconstruction

Several algorithms have been proposed in the image processing literature in the last years for HR image reconstruction [2, 3]. In general, as a preliminary classification, they can be divided in spatial or frequency domain approaches. It is well known that HR reconstruction algorithms through frequency domain are simpler and have more intuitive HR mechanism than that derived in the spatial domain [1]. 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 (CFT) 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. Latter, other works like [5], extended that frequency domain approach to include blurred and noisy low-resolution images. Since [4], several resolution enhancements approaches appeared, most of then using a spatial domain context. Indeed, spatial domain methods are able to work with more general observations models such as spatially varying blurring [2]. Another advantage provided by the spatial approaches is, for instance, the capability for a priori constraints inclusion. The most important results were acquired in [6], which turned this work a reference on resolution enhancement problems. Irani and Peleg [6] used an iterative spatial domain approach similar to the back-projections algorithms employed on CT reconstruction algorithms. In that work, the authors have considered translations and rotations between the low-resolution images.

2.1 Image Registration

The 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. If the LR images are shifted by integer units, then each image contains the same information, and thus there is no new information that can be used to reconstruct an HR image [1]. On the other hand, if the displacements are not integer units and are known, or can be well estimated, then the LR images can be combined to construct a HR one. Figure 1(a) illustrates the case for three images with different sub-pixel displacements from each other, where the images are aligned following a reference frame.

![Figure 1](image.png)

Figure 1: (a) Images with sub-pixel displacements; (b) interpolation on a HR grid.

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 a first estimate of the HR image. Figure 1(b) [1] illustrates that case, where a non-uniform interpolation is used to produce an improved resolution image. However, although this procedure can be quite useful, it does not consider any information of the LR image formation model. Indeed, more accurate algorithms take into account a model
for the image acquisition process. Thus, in the next section, we present a model for the LR observations.

2.2 LR Image Formation Model

Consider a continuous signal $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ that represents the scene of interest and suppose that it is imaged onto a sensor array, which consists of a set of $N \times N$ small detectors. We note that for the purposes of this work we are not concerned with the nature of the individual detectors and, for simplicity, we only consider a square array. Also, consider that each detector in the sensor array covers only a small portion of the scene whereas the sensor array covers the entire scene as a whole. Thus, considering that the $n$-th detector has a spatial response characteristic described by $\delta_n(x,y)$, its output is given by \[ d_n = \int_{-\infty}^{+\infty} f(x,y) \cdot \delta_n(x,y) \, dx \, dy, \] where $0 \leq n < N^2$. Now, consider the discrete versions $f[i,j]$ and $\delta_{kl}[i,j]$ of $f(x,y)$ and $\delta_n(x,y)$, respectively, defined over a high-resolution rectangular grid of size $M \times M$ pixels, where $M > N$, $n = k \cdot N + l$ and $0 \leq k, l < N$. Then, the integration over the continuous spatial variables $(x,y)$ defined in equation (1) for the $n$-th detector can be approximated by summations on the high-resolution grid yielding

\[ d[k,l] = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} f[i,j] \cdot \delta_{kl}[i,j]. \] (2)

In this sense, the discrete image $d[k,l]$, formed by the output of the $N \times N$ detectors, represents the low-resolution version of the high-resolution image $f[i,j]$, $0 \leq i, j < M$. Considering that the discrete spatial response function $\delta_{kl}[i,j]$ assumes a uniform, unity response of the detector over its response region, it can be defined as \[ \delta_{kl}[i,j] = \begin{cases} 0 & \text{if pixel } (i,j) \text{ is outside the pixel } (k,l) \\ 1 & \text{if pixel } (i,j) \text{ is within the pixel } (k,l) \\ r_{k,l} & \text{if pixel } (i,j) \text{ is partially within } (k,l) \end{cases} \] (3)

where $0 < r_{k,l} < 1$ represents a pixel-contribution parameter. It can be seen as the fraction of overlap between the pixel at $(i,j)$ position on the high-resolution grid and the detector at $(k,l)$ position on the low-resolution grid. In this sense, the LR pixels are defined as a weighted sum of the related HR pixels. Figure 2 illustrates these possibilities.

![Figure 2](image)

Figure 2: (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 $(k,l)$.

Following a lexicographic ordering of $f[i,j]$ and $d[k,l]$, equation (2) can be written in vector-matrix notation as

\[ d = Df, \] (4)

where $d \in \mathbb{R}^{N \times N}$ is the vector with components given by $d_n = \sum_{m} \delta_{n,m} \cdot f_m$ and $f_m$ are components of the vector $f \in \mathbb{R}^{M \times M}$. Moreover, $\delta_{n,m}$ are elements of the down-sampling operator $D$ of size $M^2 \times N^2$, formed with elements of $\delta_{kl}[i,j]$. In this sense, the $D$ operator under-samples the vector signal $f$ yielding the observation vector $d$. We also say that $D$ under-samples the image by a scale factor $s$ in each direction, where $s$ depends on $M$ and $N$. In a more realistic approach, digital images are often blurred by the optical system during the acquisition procedure \cite{2} 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 $b$ is given by $b = Hf$, where $H$ 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 $f$, may be modelled as

\[ d = Dhf + \eta, \] (5)

where $\eta$ stands for the noise in the observations, following an additive model.

3 The Methodology

3.1 Sub-pixel Registration

As discussed in Section 2.1, the LR image registration with sub-pixel precision cannot be neglected in
the HR image reconstruction framework [3]. However, we have noted that this stage is often neglected in most of the proposed algorithms in the literature, where for practical proposes they assume to have knowledge about the relative displacements among the under-sampled observations. Therefore, in this section we present a straightforward algorithm for LR image registration with sub-pixel accuracy. It is a variant of the algorithm proposed in [8] and it has been proved to be very useful for our proposes. The method only considers a set of globally translated and low-resolution observations, i.e., we do not consider any other transformations [9] among the images. We observe that this assumption has been proved not to be so restrictive since we are considering very slight displacements between the images.

Consider two continuous signals \(f, g : \mathbb{R}^2 \to \mathbb{R}\) that represent a reference image and a slighited shifted version of the reference image, respectively. Then, we need to find \(x_0\) and \(y_0\) that minimize a similarity function given by

\[
s(x_0, y_0) = \sum_{x \in X} \sum_{y \in Y} [f(x, y) - g(x - x_0, y - y_0)]^2, \tag{6}
\]

where \(X\) and \(Y\) are finite sets of points. Under the assumption that \(f(x, y)\) and \(g(x, y)\) are analytic functions, or are at least \(C^1\) class, and expanding \(g(x, y)\) to the first term of its Taylor series, it is easy to show that \(x_0^*\) and \(y_0^*\) that minimize equation (6) is given by

\[
x_0^* = \frac{\sum_x \sum_y [(f - g)(x, y) - y_0^* g_y(x, y)] g_x(x, y)}{\sum_x \sum_y g_x^2(x, y)} \tag{7}
\]

and

\[
y_0^* = \frac{\sum_x \sum_y [(f - g)(x, y) - x_0^* g_x(x, y)] g_y(x, y)}{\sum_x \sum_y g_y^2(x, y)} \tag{8}
\]

where \(g_x(x, y)\) and \(g_y(x, y)\) are the first derivatives of \(g(x, y)\) in relation to \(x\) and \(y\), respectively.

### 3.2 The ICM Algorithm

The Iterated Conditional Modes algorithm was proposed by Besag [10] 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 [10]. It is an iterative 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 (9) for the \(a posteriori\) probability of the value of the pixel \(i\), given the observations \(g\) and the current values of all pixels in the neighborhood of the pixel \(i\).

\[
p(f_i \mid g, \hat{f}_{S \setminus i}) \sim p(g_i \mid f_i) \cdot p(f_i \mid \hat{f}_\partial) \tag{9}
\]

In the above equation \(S \setminus i\) represents the set of all neighbors of the pixel \(i\) and \(\partial_i\) 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 [11]. In this sense, \(\hat{f}_{S \setminus i}\) is the vector of all current values of the image excluding the pixel \(i\) and \(\hat{f}_\partial\) 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 \(\hat{f}\) 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

\[
p(g_i \mid f_i) \cdot p(f_i \mid \hat{f}_\partial)
\]

4. Repeat item (3) \(\tau_{iter}\) times.

### 3.3 The Proposed Algorithm

Now, we consider the problem of reconstructing an image on a HR grid of size \(M \times M\) pixels from a sequence of \(T\) blurred and noisy under-sampled LR images of size \(N \times N\) pixels with different sub-pixel displacements from each other. The goal is to find an estimate of the HR blurred image \(b\) given all LR observations \(d_i\). The proposed algorithm consist of
constructing an initial HR image from all the observations following the registration procedure discussed in Section 2.1 and then improve the resolution through the ICM algorithm. As mentioned above, we assume a global translate model for the observations and then, given a set of T observations, we use equations (7) and (8) to estimate the relative displacements between each observation and a reference image. For practical proposes, we assume that the reference image is the first LR observation. Then, a first estimate of the HR image is given by aligning all the low-resolution observations over a high-resolution grid. Following that, we use the ICM algorithm to improve the resolution of this first estimate, where we assume a Potts-Strauss model \[12\] for the actual value of the HR image. We note from the previous section, that the ICM algorithm is initialized by the maximum-likelihood estimate of \( \hat{f} \). However, we propose to use the estimate of the HR image from the registration stage as the initial image to the algorithm. Now, the Potts-Strauss model can be defined by the set of all the conditional distributions given by

\[
p(f_i | \hat{f}_{\partial_i}) \sim e^{\beta \xi \{t \in \partial_i / f_i = f_t\}},
\]

where \( \beta \in \mathbb{R} \) is often referred to as the attraction or repulsion parameter if it is positive or negative, respectively \[12\]. In the ICM algorithm, we also need to know \( p(g_i | f_i) \), where in our case the vector \( g \) is constructed from all the observation vectors \( d_t \). In this work, we assume that \( p(g_i | f_i) \) may be given by

\[
p(g_i | f_i) = \frac{1}{\sigma \sqrt{2\pi}} \cdot e^{- \frac{(g_i - m_i)^2}{2\sigma^2}},
\]

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

4 Results and Discussions

An evaluation of the proposed method has been conducted by processing a set of simulated LR images. Figure 3(a) shows the image that was considerated the actual HR image. It has 512x512 pixels and was downloaded from the USC-SIPI Image Database \[13\]. In order to simulate the blur due to the image process acquisition, that image was convolved with a 3x3 uniform rectangular kernel. Then, 16 LR images are obtained by using the image formation model presented in Section 2.2. 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. Latter, each 128x128 LR image was contaminated by additive and independent Gaussian noise at 40 dB. For comparison proposes, Figure 3(b) show four LR images from the set of 16 simulated images. In this experiment, we consider the image in the upper-left corner of the Figure 3(b) as the reference image. 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 the all LR observations on a 512x512 grid.

Figure 4: HR reconstruction using the ICM algorithm.

As one can see, the registration procedure is able to give better results when compared with the interpolated image. Although we have knowledge of
the actual displacements between each LR image and the reference image, we have used equations (7) and (8) to estimate the displacement values. We note that the proposed method for sub-pixel registration has demonstrated to be very accurate in all conducted experiments. Figure 4 shows the result for the HR reconstruction using the ICM algorithm. In this simulation, the algorithm was initialized with the image in Figure 3(d) and the β parameter in equation (10) was found following the procedure proposed in [12]. From the result, 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 for producing good results.

5 Concluding Remarks

We have presented an efficient algorithm for HR image reconstruction based on the Iterated Conditional Modes algorithm and also using a procedure for sub-pixel image registration. The results demonstrate that the algorithm can be extremely efficient in a HR reconstruction framework. Indeed, the method has demonstrated good performance both in visual accuracy and computational cost. In future works, we intend to make more 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 SNR ratio in the observations.

6 Acknowledgments

The work of Murillo R. P. Homem was supported by FAPESP, Brazil, grant numbers 04/01632-1 and 2002/07153-2. The work of Ana L. D. Martins was supported by a CAPES scholarship. Ana L. D. Martins also would like to thank Paulo E. Cruvinel for useful discussions.

References