PRACTICAL AND ROBUST SUPER RESOLUTION USING ANISOTROPIC DIFFUSION FOR UNDER-DETERMINED CASES

[1]*Hiêp Quang Luong*, [2]*Stefaan Lippens and* [2]*Wilfried Philips*

[1]hiep.luong@telin.ugent.be
[1,2]Ghent University, Dept. TELIN, St.-Pietersnieuwstraat 41, Ghent, Belgium

ABSTRACT

In this paper, we propose a fast super resolution (SR) scheme based on anisotropic diffusion. SR is the reconstruction of one or more high resolution images from a set of low resolution images. We focus our attention on a practical architecture of the SR algorithm. Realistic super resolution has to deal with different kinds of noise, blur and also the problem of under-determination, when very few low resolution images are available. These problems are tackled by our proposed method. Results show an improvement in visual quality: important image features are preserved while noise and blur are removed.

1. INTRODUCTION

In most image analyse applications, high resolution (HR) images are often required. However video sensors with low resolution (LR) are used nowadays, e.g. webcams and surveillance applications.
There is a limitation on increasing the spatial resolution by reducing the pixel size of the video camera. If the pixel size decreases, the number of incoming photons per pixel also decreases. In this case, shot noise becomes more important which degrades the image quality severely. The high cost for high precision optics and image sensors is also an important concern in many commercial applications regarding HR imaging.
When image resolution can not be improved at the sensor because of technological limitations or because of high costs, we need SR image reconstruction [1]. SR is a image processing technique to obtain a HR image from multiple observed LR images (e.g. from a video sequence).
Figure 1 shows us the relationship between the ideal HR image and the acquired LR images. In order to reconstruct the HR image, subpixel shifts between the LR images must exist. In this paper we take a closer look to the under-determined case where some pixels on the HR grid will have no observation after proper registration and projection of the LR images onto the HR grid. Most real data corresponds to an under-determined situation.
In the following paragraphs we will discuss the used SR image reconstruction scheme and its components. We also discuss some results and conclude our work.

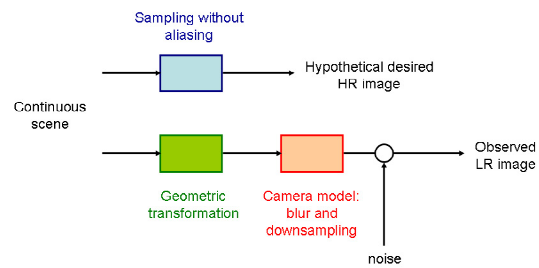


Figure 1: *Observation model of the image acquisition.*

2. THE PROPOSED METHOD

Since very fast resolution enhancement methods are desirable for practical applications, our SR scheme is similar to the fast method proposed in [2] and is based on our camera model as shown in figure 1. Our method consists of three consecutive steps (see also figure 2):

1. Subpixel registration of the LR images to a reference image I_{ref}, which is chosen from the LR images.
2. Noniterative data fusion and interpolation.
3. Iterative deblurring and denoising.

Unlike [2], we treat the interpolation not simultaneously with the iterative deblurring. The iterative restoration step will converge much faster, because the interpolation will provide a better initialisation image. In the rest of this paper we will treat each R,G,B-channel of colour images separately.

2.1. Subpixel Registration

To simplify the registration problem, we assume that we are only dealing with pure translational motions of the LR images. This assumption will enable us to develop a very fast algorithm. Better but slower results could be achieved using more accurate motion models.

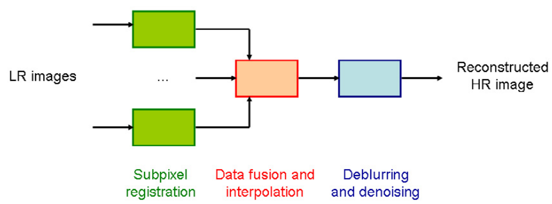


Figure 2: *Block diagram representation of our SR method.*

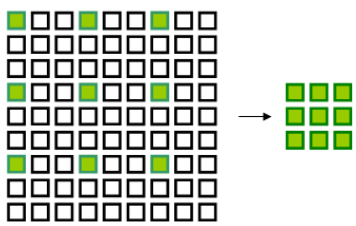


Figure 3: *The decimation operator maps an $rM \times rN$ image to an $M \times N$ image ($r = 3$).*

Common ways to achieve subpixel registration in the spatial domain is to interpolate either the image data or the correlation data [3, 4]. In order to save computation time we only resample the reference image I_{ref} on a higher resolution. In this way we represent the downsampling operator in the camera model as a simple decimation operator (see figure 3). We then find the registration parameters $\widehat{\mathbf{y}}_i$ by minimizing this energy function:

$$\widehat{\mathbf{y}}_i = \arg\min_{\mathbf{y}} \frac{1}{W} \sum_{\mathbf{x}\in\Omega} |I_{\mathrm{LR},i}(\mathbf{x}) - I_{\mathrm{ref}}(\mathbf{x}+\mathbf{y})| \qquad (1)$$

with W the total number of overlapping pixels between the ith LR image $I_{\mathrm{LR},i}$ and the resampled reference image I_{ref}. The set of overlapping pixels is denoted as Ω. As a simplification of the optimization problem, we use the HR grid as the discrete search space for $\widehat{\mathbf{y}}_i$.

Most existing techniques use linear interpolation to resample I_{ref}. However these interpolation methods typically suffer from blurring, staircasing and ringing. These artifacts not only degrade the visual quality, but also affect the registration accuracy. Therefore we adopt a nonlinear interpolation technique based on level curve mapping, which suffers less from these artefacts [5].

To speed up the computation time, we use a hierarchical registration method. First, the integer pixel shifts are calculated between the LR images and the original I_{ref}. Next the registration parameters are refined to subpixel shifts using the described method.

2.2. Data Fusion And Interpolation

In this step we determine a pixel value for every pixel of the HR grid. In the previous registration step we already obtained zero (because of under determination) or more observations for these pixels.

Several observations are available. Starting from the maximum likelihood principle, it can be shown that minimizing the norm of the residuals is equivalent to median estimation [2]. A residual is the difference between an observed pixel value and the predicted pixel value. The median is very robust to outliers, such as noise and errors due to misregistration. For this reason we adopt the median estimate for every pixel with at least one observation.

No observation is available. The unknown pixel values are initialised with the interpolated reference image I_{ref}. We do not need additional computations since this image is already constructed for the registration (see § 2.1). These pixel values are then corrected with a weighted average of neighbouring pixels for which we do have an observation. This is a kind of fast interpolation which prevents the loss of known pixel values in the image denoising.

2.3. Robust Deblurring And Denoising

We assume that the blur in the camera model is characterized by a shift-invariant *point spread function* (PSF). The inverse problem becomes highly unstable in the presence of noise. This can be solved by imposing some prior knowledge about the image. Typically we will try to force spatial smoothness in the desired HR solution. This is usually implemented as a penalty factor in the generalized minimization cost function:

$$\widehat{I}(\mathbf{x}) = \arg\min_{I(\mathbf{x})} \left[\rho_{\mathrm{R}}(I(\mathbf{x})) + \lambda\rho_{\mathrm{D}}(H * I(\mathbf{x}) - I(\mathbf{x},0))\right] \qquad (2)$$

where H denotes the PSF-kernel (typically Gaussian blur) and λ is the regularization parameter between the two terms, respectively called the regularization term ρ_{R} and the data fidelity term ρ_{D}. Image $I(\mathbf{x},0)$ is the image obtained in § 2.2.

The minimization problem of equation 2 could be transformed to the following partial differential equation (PDE) which produces iteratively diffused images $I(\mathbf{x},t)$ starting from the inialisation image $I(\mathbf{x},0)$:

$$\frac{\partial I(\mathbf{x},t)}{\partial t} = \rho'_{\mathrm{R}}(I(\mathbf{x},t)) + \lambda\rho'_{\mathrm{D}}(H * I(\mathbf{x},t) - I(\mathbf{x},0)) \qquad (3)$$

where $\rho'(I) = \frac{\partial}{\partial I}\rho(I)$ for both terms.

We use so-called edge-stopping functions in the regularization term, because it suppresses the noise better while retaining important edge information [6]. For this reason, we will use the anisotropic diffusion formulation for the regularization term:

$$\rho'_{\mathrm{R}}(I(\mathbf{x},t)) = \nabla \cdot [g(|\nabla I(\mathbf{x},t)|)\nabla I(\mathbf{x},t)] \tag{4}$$

where $g(x)$ is the edge-stopping function. In this paper we use the *Lorentzian* edge-stopping function (g_{Lor}):

$$g_{\mathrm{Lor}}(x) = \frac{1}{1+\frac{x^2}{k^2}} \tag{5}$$

where the parameter k is called the contrast parameter, which controls the shape of the diffusivity function [6].
For the data fidelity term, we will use the L_1-norm function due to its robust property according to [2]:

$$\rho_{\mathrm{D}}(\ldots) = |H * I(\mathbf{x},t) - I_0(\mathbf{x})| \tag{6}$$

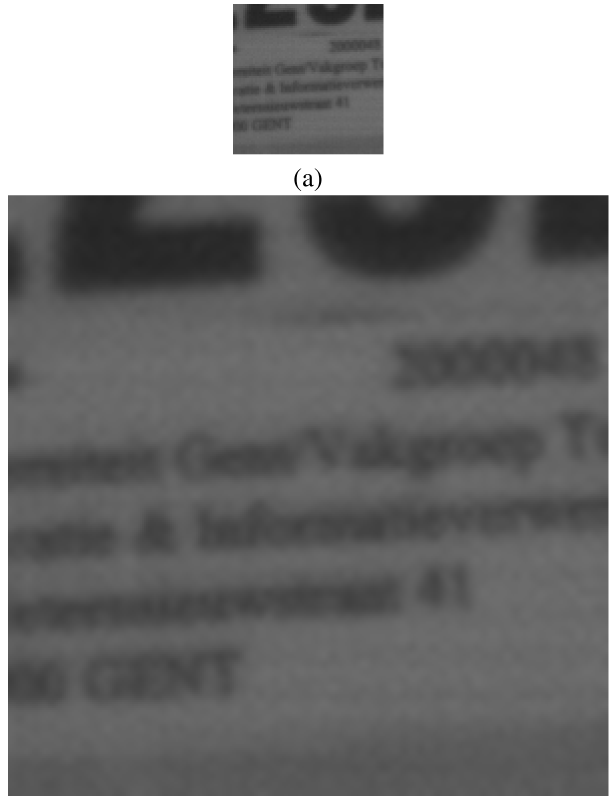
Finally the PDE of equation 3 is iteratively applied to update the blurred and noisy image.

3. RESULTS

As a realistic experiment, we have grabbed 30 LR images of 75×75 pixels with the *Philips Inca Smartcam* in rather poor lightning conditions and we have enlarged the images 4 times in each dimension. The following parameters are chosen: $\lambda = 30$, $\sigma_b = 3.0$ (this is the standard deviation of the Gaussian blur kernel). The parameter selection was based on trial and error, i.e. to produce the visually most appealing results. To restrict the computation time, we have limited the number of iterations for deblurring and denoising to 50. The computation times are measured on an AMD Athlon$^{\mathrm{TM}}$ 64 processor 3000+.
Figure 4 shows the result of cubic B-spline interpolation. Obviously, this non-SR technique is not able to make most characters readable. Figure 5 illustrates the *iterative back-projection method* (IBP), an SR technique proposed in [4]. Some characters are already more readable, but the result is still very blurry.
Beside the use of (4) as regularization term, we also plugged the *bilateral total variation* (BTV) prior as proposed in [2] in our SR scheme for comparison. This BTV regularization term is given by

$$\rho_{\mathrm{R,BTV}}(I(\mathbf{x})) = \underbrace{\sum_{l=-P}^{P}\sum_{m=-P}^{P}}_{l+m\geq 0} \alpha^{|l|+|m|}|I(\mathbf{x})-I(\mathbf{x}+(l,m))|$$

where the scalar weight α ($0 < \alpha < 1$) is a weighting factor for geometric distances and P is the window size (typically $P = 2$ or $P = 3$) [2].



(a)

(b)

Figure 4: *(a) One original LR input image and (b) standard non-SR cubic B-spline interpolation.*

Figure 6 shows the results of our proposed SR scheme with the BTV prior and the proposed Lorentzian edge-stopping function respectively. The result produced with the BTV prior is better denoised in comparison with the Lorentzian edge-stopping function. On the other hand, the use of the Lorentzian function preserves important edges and important features of the text. Moreover, from our experiments we clearly observe that our anisotropic diffusion based SR technique (9 s) is significantly faster than the IBP SR reconstruction (233 s) and the use of the BTV prior (18 s).

4. CONCLUSION

We presented a very fast and straightforward super resolution scheme based on anisotropic diffusion that gives very promising visual results. The architecture of the hierarchical subpixel registration and the robust data fusion was focused on speed. The choice of edge-stopping functions in the anisotropic diffusion offers denoised images with sharp edges and preservation of important features. Future work includes automatic parameter estimation, better (general) PSF estimation and noise estimation, the use of more complex motion models, etc.

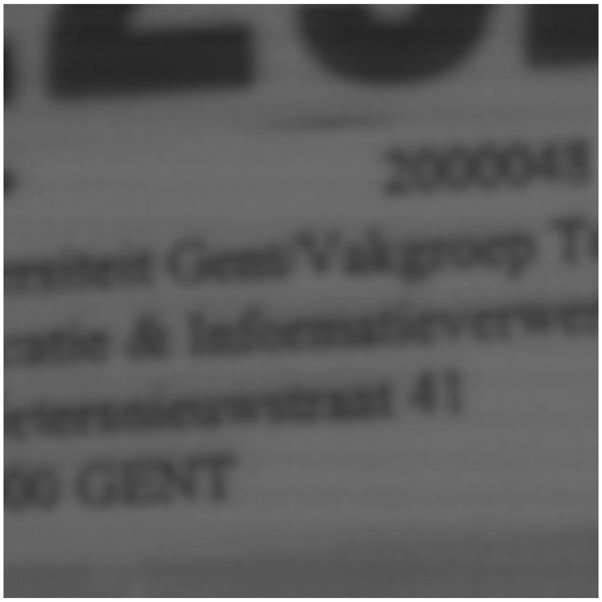


Figure 5: *SR result with 30 LR images and the iterative backprojection method [4] (computation time: 233 s).*

5. REFERENCES

[1] S.C. Park, M.K. Park, and M.G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.

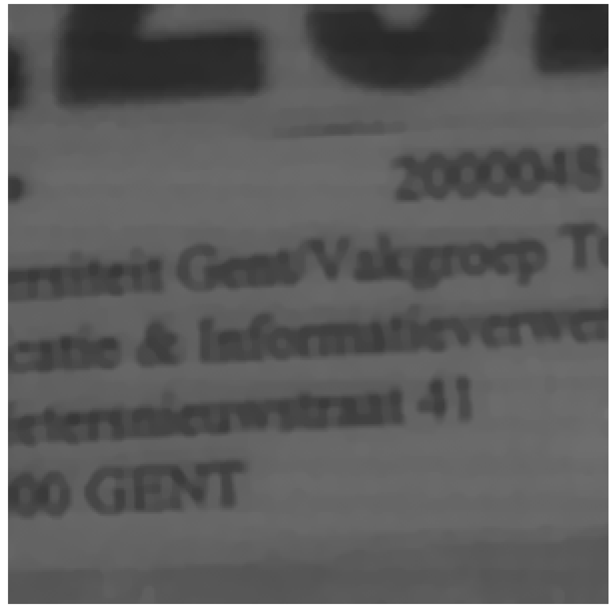
[2] S. Farsiu, M.D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions On Image Processing*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.

[3] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.

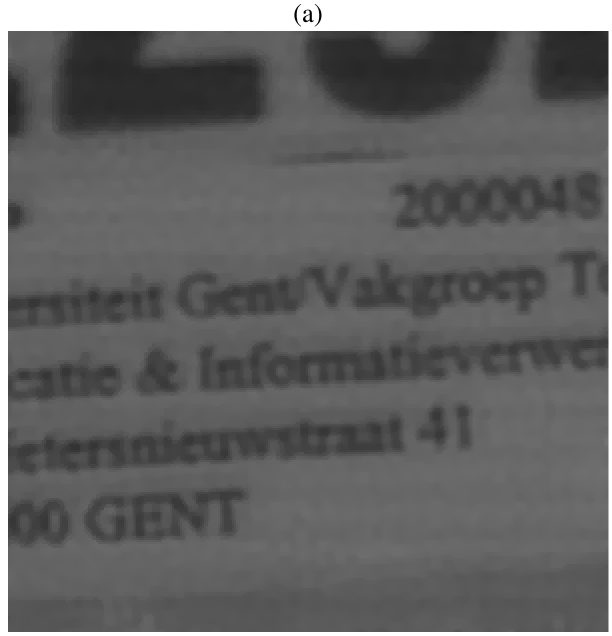
[4] M. Irani and S. Peleg, "Improving resolution by image registration," *Graphical Models And Image Processing*, vol. 53, no. 3, pp. 231–239, May 1991.

[5] H.Q. Luong, P. De Smet, and W. Philips, "Image interpolation using constrained adaptive contrast enhancement techniques," in *Proc. of IEEE International Conference on Image Processing 2005*, 2005, pp. 998–1001.

[6] A. Pizurica, I. Vanhamel, H. Sahli, W. Philips, and A. Katartzis, "A bayesian approach to nonlinear diffusion based on a laplacian prior for ideal image gradient," in *IEEE Workshop on Statistical Signal Processing SSP 2005*, July 2005.



(a)



(b)

Figure 6: *SR results with 30 LR images and the proposed scheme using (a) the BTV prior ($\alpha = 0.5$ and $P = 2$) (computation time: 18 s) and (b) the proposed Lorentzian edge-stopping function ($k = 1$) (computation time: 9 s).*