NON-LOCAL IMAGE INTERPOLATION

Hiêp Quang Luong, Alessandro Ledda and Wilfried Philips

Ghent University, Department of Telecommunications and Information Processing
St.-Pietersnieuwstraat 41, 9000 Ghent, Belgium

ABSTRACT

In this paper we present a novel method for interpolating images and we introduce the concept of non-local interpolation. Unlike other conventional interpolation methods, the estimation of the unknown pixel values is not only based on its local surrounding neighbourhood, but on the whole image (non-locally). In particularly, we exploit the repetitive character of the image. A great advantage of our proposed approach is that we have more information at our disposal, which leads to better estimates of the unknown pixel values. Results show the effectiveness of non-local interpolation and its superiority at very large magnifications to other interpolation methods.

Index Terms— Interpolation, image resolution

1. INTRODUCTION

In this paper we describe a new approach to image interpolation. Many interpolation methods already have been proposed in the literature, but all suffer from one or more artefacts. Linear interpolation methods deal with aliasing (e.g. jagged edges in the up scaling process), blurring and/or ringing effects [1]. Non-linear or adaptive interpolation methods incorporate a priori knowledge about images. Dependent on this knowledge, the interpolation methods could be classified in different categories. The edge-directed based techniques follow a philosophy that no interpolation across the edges in the image is allowed or that interpolation has to be performed along the edges. This rule is employed for instance in the AQua-2 method [4]. The restoration-based techniques tackle unwanted interpolation artefacts. Examples are methods based on isophote smoothing, level curve mapping and mathematical morphology [2, 8]. Some other adaptive techniques exploit the self-similarity property of an image, e.g. iterated function systems [5, 9]. Another class of adaptive interpolation methods is the example-based approach, which maps blocks of the low-resolution image into predefined interpolated blocks [6]. Adaptive methods still suffer from artefacts: their results often look segmented, yield important visual degradation in fine textured areas or random pixels are created in smooth areas [2].

When we use very big enlargements (i.e. linear magnification factors of 8 and more), then all these artefacts become more visible and hence more annoying.

2. REPETITIVE STRUCTURES

Fractal-based interpolation methods suppose that many things in nature possess fractalness, i.e. scale invariance [5]. This means that parts of the image repeat themselves on an ever-diminishing scale, hence the term self-similarity. This self-similarity property is exploited for image compression and interpolation by mapping the similar parts at different scales. Due to the recursive application of these mappings at the decoder stage, the notion of *iterated function systems* (IFS) is introduced.

Unlike IFS, we exploit the similarity of small patches in the same scale, i.e. spatially. In order to avoid confusion, we will use the term repetitivity. Another class of upscaling methods which also takes advantage of repetitivity, is called super resolution (SR) reconstruction. SR is a signal processing technique that obtains a high resolution image from multiple noisy and blurred low resolution images (i.e. from a video sequence) [7]. It is well known that SR produces superior results to interpolation methods. Typical SR reconstruction schemes consist of the following steps:

- Subpixel registration of the resolution images;
- Data fusion and interpolation;
- Restoration (i.e. deblurring and denoising).

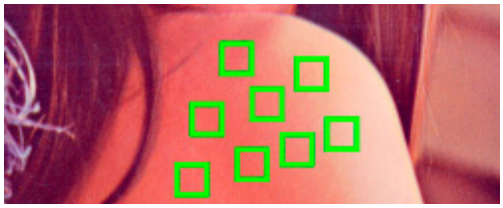
It is often assumed that true motion is needed for SR, however many registration methods do not yield true motion: their results are optimal to some proposed cost criterion, which are not necessarily equal to true motion. With this in mind, we can hypothetically assume that repetitive structures could serve as multiple noisy observations of the same structure (after proper registration). Results of our experiments in § 4 will confirm that this hypothesis holds for real situations. The concept of repetitive structures has already succesfully been used for image denoising [3]. Besides repetitivity in texture, we can also find this recurrent property in other parts of the image, some examples are illustrated in figure 1.

Our method is found perfectly suitable to some applications: text images (multiple repeated characters regardless of

their font), gigantic satellite images (long roads and a lot of texture provide a huge amount of training data).

(a) Repetition in different objects.

(b) Repetition along edges.



(c) Repetition in uniform areas.

Fig. 1. Examples of repetitive structures in images.

3. PROPOSED RECONSTRUCTION SCHEME

We propose a simple interpolation method which exploits this repetitive behaviour. Our scheme is quite straightforward and consists of three consecutive steps (similar to the SR scheme).

3.1. Matching and registration of repetitive structures

For the sake of simplicity, we define small squared windows (B with n^2 pixels) as basic structure elements. Two criterions are used in our algorithm to find matching windows across the whole image, namely the *zero-mean normalized cross correlation* (CC) and the *mean absolute differences* (MAD):

$$E_{CC} = \frac{\sum_{\mathbf{x}\in\Omega}(B(m(\mathbf{x})) - \overline{B})(B_{ref}(\mathbf{x}) - \overline{B_{ref}})}{\sqrt{\sum_{\mathbf{x}\in\Omega}(B(m(\mathbf{x})) - \overline{B})^2 \sum_{\mathbf{x}\in\Omega}(B_{ref}(\mathbf{x}) - \overline{B_{ref}})^2}} \quad (1)$$

$$E_{MAD} = \frac{1}{n^2}\sum_{\mathbf{x}\in\Omega}|B(m(\mathbf{x})) - B_{ref}(\mathbf{x})| \quad (2)$$

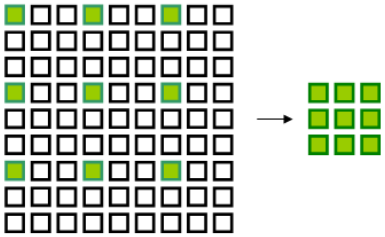


Fig. 2. The decimation operator maps a $3M \times 3N$ image to an $M \times N$ image.

where Ω contains all the pixels of the reference window B_{ref}; $\overline{B}$ and $\overline{B_{ref}}$ are denoted as the mean values of respectively B and B_{ref}. The transformation of the coordinates is characterized by the mapping function m. To simplify the registration problem and particularly to save computation time, we assume that we are only dealing with pure translation motions of B. The main motive to use these two criterions is because they are somewhat complementary: CC emphasizes the similarity of the structural or geometrical content of the windows, while MAD underlines the similarity of the luminance (and colour) information. A matched window is found if the two measures E_{CC} and E_{MAD} satisfy to the respective thresholds τ_{CC} and τ_{MAD}, more specifically: $E_{CC} > \tau_{CC}$ and $E_{MAD} < \tau_{MAD}$. Our implementation uses an exhaustive search in order to find the matching windows, but more intelligent (pattern-based) search algorithms could reduce the computation time enormously.

Common ways to achieve subpixel registration in the spatial domain, is to interpolate either the image data or the correlation data. In order to save computation time we only resample the reference window B_{ref} on a higher resolution. In this way we represent the downsampling operator in the camera model as a simple decimation operator (see figure 2). We estimate the subpixel shifts using the criterion of equation 2. After the registration, the pixel values of B are mapped onto the high resolution (HR) grid. Most existing techniques use linear methods to upscale B_{ref}. However, these interpolation methods typically suffer from blurring, staircasing and/or ringing. These artefacts not only degrade the visual quality but also affect the registration accuracy. That is why we adopt a fast non-linear restoration-based interpolation based on level curve mapping [2].

3.2. Data fusion

Starting from the maximum likelihood principle, it can be shown that minimizing the norm of the residuals is equivalent to median estimation, which is very robust to outliers, such as noise and errors due to misregistration [7]. For this reason we adopt the median estimate for each pixel in the HR grid for which we have at least one observation.

For the other unknown pixels, we simply initialize them with the values of the interpolated B_{ref}s: no additional calculations are needed since these interpolations are already constructed for the registration step. As in traditional interpolation, the original pixel values must not change [2], thus we map these pixel values onto the HR grid. In a nutshell, the HR grid consists of three classes: the original pixels (OR), the unknown pixels (UN) and the fused pixels (FU). The last mentioned class will provide the extra information which gives us better interpolation results compared to conventional upscaling techniques.

3.3. Robust denoising

The obtained HR image $I(\mathbf{x}, 0)$ after data fusion is still noisy. Diffusion-based denoising methods using partial differential equations (PDE) are very popular nowadays: one can suppress the noise while retaining important edge information by imposing some prior knowledge about the image in the regularization. The following PDE iteratively produces a family of diffused images $I(\mathbf{x}, t)$ starting from the input image $I(\mathbf{x}, 0)$:

$$\frac{\partial I(\mathbf{x}, t)}{\partial t} = \rho_1'(I(\mathbf{x}, t)) + \lambda \rho_2'(I(\mathbf{x}, t) - I(\mathbf{x}, 0)) \quad (3)$$

where λ is the regularization parameter between the two terms, respectively called the regularization term and the data fidelity term. One of the most successful edge-preserving regularization terms proposed for image denoising is the total variation (TV):

$$\rho_1(I(\mathbf{x}, t)) = |\nabla I(\mathbf{x}, t)| \quad (4)$$

According to [7], the following data fidelity term is very robust to outliers:

$$\rho_2(I(\mathbf{x}, t) - I(\mathbf{x}, 0)) = |I(\mathbf{x}, t) - I(\mathbf{x}, 0)| \quad (5)$$

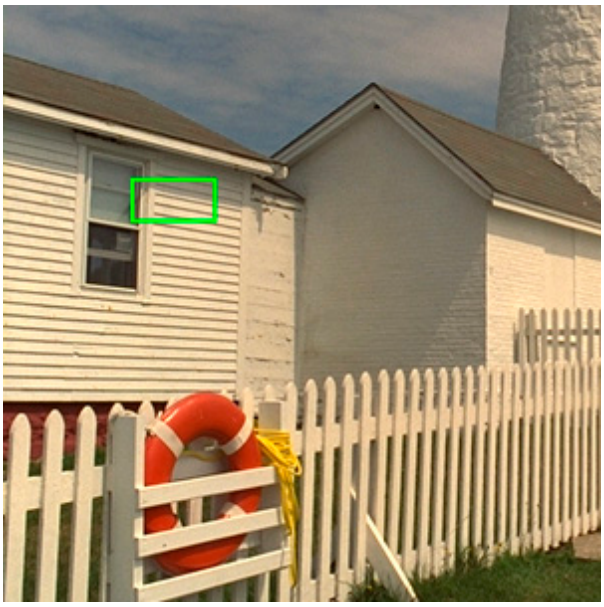
These ρ-functions are very easy to implement and are very computationally efficient. Other robust ρ-functions could also be employed. We now adapt the PDE in equation 3 locally to the several classes of pixels on the HR grid:

- Class OR: since these pixels are known and noise-free, no regularization term has to be applied. This means that these pixels depend only on the data fidelity term. Since the initialization starts with the same pixel value (see § 3.2), no changes will be made.
- Class UN: these pixels are most likely noise and depend only on the regularization term ($\lambda = 0$).
- Class FU: these pixels contain noise and relevant information. Equation 3 is applied here with λ proportional to the number of available observations.

4. RESULTS

In figure 3 we show a part of the original image with a $8\times$ nearest neighbour interpolation of the region of interest. As basic structure elements we use 5×5 windows and we have enlarged the image with a linear magnification factor of 8. For the matching step we have used the following threshold parameters: $\tau_{CC} = 0.9$ and $\tau_{MAD} = 9.0$. For the denoising we have applied the PDE within 100 iterations and with $\lambda = \min(\frac{\alpha}{20}, 1)$, where α denotes the number of observations. With these parameters we obtained the following partition of the different classes: 1.5625% (OR), 30.9091% (FU) and 67.5284% (UN).

Figure 4 shows our result compared to a linear interpolation (cubic B-splines), an edge-directed method (AQua-2 [4]) and an IFS method (obtained from commercial software [9]). Significant improvements in visual quality can be noticed in our method: there is a very good reconstruction of the edges and our result contains less annoying artefacts.



(a) A part of the original image.

(b) Nearest neighbour of the region of interest of (a).

Fig. 3. Parts of the original image.

5. CONCLUSION

In this paper we have presented a novel method for interpolating images based on the repetitive character of the image. Exploiting repetitivity brings more information at our disposal, which leads to better estimates of the unknown pixel values. Results show the effectiveness of non-local interpolation and its superiority for very large magnifications to other interpolation methods: edges are reconstructed well and artefacts are heavily reduced.

6. ACKNOWLEDGEMENT

The authors would like to thank D. Muresan for providing us his software (*Visere* with the AQua-2 algorithm) [4].

7. REFERENCES

[1] T. Lehmann, C. Gönner and K. Spitzer, "Survey: Interpolations Methods In Medical Image Processing," in *IEEE Trans. on Medical Imaging*, vol. 18, no. 11, pp. 1049-1075, 1999.

[2] H.Q. Luong, P. De Smet and W. Philips, "Image Interpolation Using Constrained Adaptive Contrast Enhancement Techniques," in *Proc. of IEEE International Conference of Image Processing*, vol. 2, pp. 998-1001, 2005.

[3] A. Buades, B. Coll and J. Morel, "Image Denoising By Non-Local Averaging," in *Proc. of IEEE International Conference of Acoustics, Speech end Signal Processing*, vol. 2, pp. 25-28, 2005.

[4] D. Muresan, "Fast Edge Directed Polynomial Interpolation," in *Proc. of IEEE International Conference of Image Processing*, vol. 2, pp. 990-993, 2005.

[5] H. Honda, M. Haseyama and H. Kitajima, "Fractal Interpolation For Natural Images," in *Proc. of IEEE International Conference of Image Processing*, vol. 3, pp. 657-661, 1999.

[6] W.T. Freeman, T.R. Jones and E.C. Pasztor, "Example-Based Super-Resolution," in *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56-65, 2002.

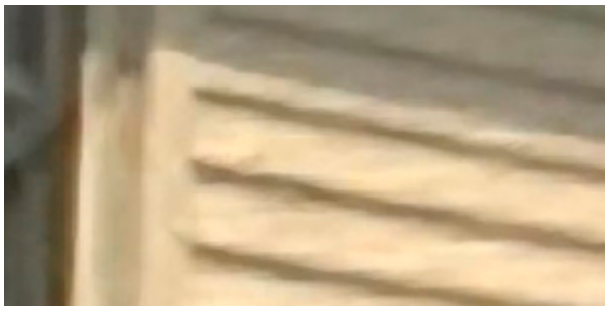
[7] S. Farsiu, M.D. Robinson, M. Elad and P. Milanfar, "Fast and Robust Multiframe Super Resolution," in *IEEE Trans. on Image Processing*, vol. 13, no. 10, pp. 1327-1344, 2004.

[8] A. Ledda, H.Q. Luong, V. De Witte, W. Philips and E.E. Kerre, "Image Interpolation Using Mathematical Morphology," in *Sixth FirW PhD Symposium*, Ghent University, Belgium, 2005.

[9] Genuine Fractals 4, *http://www.ononesoftware.com*

(a) Cubic B-spline.

(b) AQua-2 [4].



(c) IFS [9].

(d) Our proposed method.

Fig. 4. Results of several interpolation methods.