

# High Quality Image Magnification using Cross-Scale Self-Similarity

André Goossen<sup>1</sup>, Arne Ehlers<sup>1</sup>, Thomas Pralow<sup>2</sup>, Rolf-Rainer Grigat<sup>1</sup>

<sup>1</sup>Vision Systems, Hamburg University of Technology, D-21079 Hamburg

<sup>2</sup>General X-Ray, Philips Medical Systems, D-22335 Hamburg

andre.goossen@tu-harburg.de

**Abstract.** In medical imaging there is a frequent need to magnify a certain region of interest (ROI) of an image. However many modalities suffer from severe noise and traditional upscaling methods produce poor enlargement results. We present an approach based on cross-scale similarity of an image extendable to sequences using time domain information without explicit motion compensation. It combines noise suppression and upscaling and hence yields superior image quality compared to standard zooming algorithms. We evaluated our algorithm using artificial and clinical fluoroscopy acquisitions.

## 1 Introduction

Image magnification is the process of computing a high-resolved image from an observed low-resolution image. In the field of medical X-ray image processing the observation is provided by the X-ray detector. Because of the limited resolution of image sensors and the influence of noise during image acquisition, especially in X-ray fluoroscopy, the observation is a degraded version of the captured scene. Using traditional upscaling techniques [1, 2] such as cubic interpolation, Lanczos, etc. that incorporate neighboring pixel to interpolate the high-resolved image tend to smooth structures and do not perform well in the presence of noise. The recently proposed non-local means filter [3] besides denoising shows potential for image magnification [4].

## 2 Methods and materials

In this section we first briefly introduce the problem of image zooming that has to be solved and derive the technique of incorporating self-similarity from the general non-local means filtering approach [3].

### 2.1 Image zooming problem

For the image zooming problem we assume a simple degradation model in which the relation between the observation and the high-resolution image is given by

$$u = \mathcal{H}v + \eta \tag{1}$$

with  $u$  denoting the low resolution observation and  $v$  is the high-resolved image, that should be reconstructed.  $\eta$  is additive white Gaussian noise with zero mean and variance  $\sigma^2$ . Besides we assume that the resolution of the image  $v$  is  $z$ -times higher than the resolution of the observation with  $z \in \mathbb{N}$  being a positive integer.

The operator  $\mathcal{H}$  is a composition of blurring and downsampling and has to be inverted when solving the ill-posed problem of reconstructing  $v$  from  $u$ .

### 2.2 Non-local means filtering

Let  $u : \Omega \rightarrow \mathbb{R}$  be an image with  $\Omega \subset \mathbb{R}^2$  and let  $\mathcal{N}_u^r(\mathbf{x}) \subset \Omega$  denote a quadratic neighborhood of size  $R \times R$ ,  $R = 2r + 1$  within  $u$  around the center  $\mathbf{x} \in \Omega$ . Then the weighting function

$$w(\mathbf{x}, \mathbf{y}) = e^{-\frac{1}{h^2} \|\mathcal{N}_u^r(\mathbf{x}) - \mathcal{N}_u^r(\mathbf{y})\|^2} \tag{2}$$

maps similarity of neighborhoods around  $\mathbf{x}$  and  $\mathbf{y}$  to the interval  $[0, 1]$ . The parameter  $h$  controls the steepness of this mapping function. Exploiting this measure it is possible to average information according to its similarity:

$$v(\mathbf{x}) = \frac{1}{W(\mathbf{x})} \int_{\Omega} w(\mathbf{x}, \mathbf{y}) \cdot u(\mathbf{y}) d\mathbf{y}, \quad \mathbf{x}, \mathbf{y} \in \Omega . \tag{3}$$

The similarity weights are normalized by  $W(\mathbf{x}) = \int_{\Omega} w(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ . The result  $v(\mathbf{x})$  contains a filtered version of  $u(\mathbf{x})$  based on similarity of neighborhoods in  $u$ .

With  $N$  denoting the number of pixels in each dimension, i.e.  $|\Omega| = N^2$ , this method has a complexity of  $O(N^4 \cdot R^2)$ . The algorithm can be simplified by reducing the search space for similar neighborhoods to  $\mathcal{N}_u^s(\mathbf{x})$  of size  $S \times S$ ,  $S = 2s + 1$  which corresponds to updating the weights  $w(\mathbf{x}, \mathbf{y})$  to

$$\tilde{w}(\mathbf{x}, \mathbf{y}) = w(\mathbf{x}, \mathbf{y}) \cdot \sigma(\|\mathbf{x} - \mathbf{y}\|_{\infty}), \quad \sigma(\|\mathbf{x} - \mathbf{y}\|_{\infty}) = \begin{cases} 1 & \|\mathbf{x} - \mathbf{y}\|_{\infty} < s \\ 0 & \text{else} \end{cases} . \tag{4}$$

### 2.3 Image zooming using self-similarity

The algorithm introduced in this section allows to up-scale a single low-resolution image by a factor  $z \in \mathbb{N}$ . A method similar to the Non-local means filter is used which works on neighborhoods of different scales exploiting cross-scale similarity [4] within an image.

Let  $u$  denote the image to zoom and  $v$  denote the result of the magnification. The initial estimation  $v'$  is computed by upscaling  $u$  with a Lanczos interpolation. Then for a certain position  $\mathbf{x}$  we compare the neighboring pixels on a  $z$ -spaced grid, i.e.  $\mathbf{y} = \mathbf{x} + (m \cdot z, n \cdot z)^T$  with  $n, m \in \mathbb{Z}$ . Hence we derive the weights by

$$w(\mathbf{x}, \mathbf{y}) = e^{-\frac{1}{h^2} \|\mathcal{N}_{v'}^r(\mathbf{x}) - \mathcal{N}_{v'}^r(\mathbf{y})\|^2} . \tag{5}$$

Now instead of replacing the pixel at  $\mathbf{x}$  with samples from  $v'$  we extract the information from the corresponding pixel of the low resolution image  $u$ . Inverting

the decimation, we insert the averaged pixel only at  $\mathbf{x}$  avoiding the blurring effect of interpolating methods. Hence we derive the high-resolved magnified image by

$$v(\mathbf{x}) = \frac{1}{W(\mathbf{x})} \sum_{\Omega} w(\mathbf{x}, \mathbf{y}) \cdot u\left(\frac{1}{z}\mathbf{y}\right) d\mathbf{y}, \quad \mathbf{x}, \mathbf{y} \in \Omega. \quad (6)$$

This can be enhanced by not only using the image itself but also the recent frames within the sequence of acquisitions. In [5] we propose a temporal extension of the NLM filter based on a multi-scale motion estimation

$$v_{\bar{t}}[\mathbf{x}] = \frac{1}{W[\mathbf{x}]} \sum_{t=\bar{t}}^{\bar{t}-\tau} \sum_{\Omega} w_t[\mathbf{x}, \mathbf{y}] \cdot u_t\left[\frac{1}{z}\mathbf{y}\right], \quad W[\mathbf{x}] = \sum_{t=\bar{t}}^{\bar{t}-\tau} \sum_{\Omega} w_t[\mathbf{x}, \mathbf{y}] \quad (7)$$

with  $t \in \mathbb{N}$  denoting the index of a certain frame within the sequence history and  $\tau$  determining the number of frames to consider for filtering the current frame at  $t = \bar{t}$ .

For the motion estimation we decompose the image into a Gaussian pyramid representation and compute the sum of squared difference (SSD) for each level  $i$ . Thus the estimated position of the block center within  $u_t$  for a given block  $\mathcal{N}_{u_{\bar{t}}}^b(\mathbf{x})$  of size  $B \times B$ ,  $B = 2b + 1$  is calculated by

$$\mathbf{y}_i = \operatorname{argmax}_{\mathbf{y} \in \Omega} \left\| \mathcal{N}_{\mathcal{G}_i(u_{\bar{t}})}^b(\mathbf{x}) - \mathcal{N}_{\mathcal{G}_i(u_t)}^b(\mathbf{y}) \right\|, \quad i = 0, 1, \dots, n \quad (8)$$

$$\|\mathbf{y}_i - \mathbf{y}_{i+1}\| < B, \quad \mathbf{y}_{n+1} = \mathbf{0} \quad (9)$$

with  $\mathcal{G}_i(\cdot)$  denoting the Gaussian level  $i$  and  $\mathcal{G}_0(u) = u$ .

The restriction of higher level estimations  $\mathbf{y}_i$  to the vicinity of the lower level estimation introduces a natural regularization compared to matching only full resolution blocks and prevents outliers. However the proposed approach does not depend on accurate motion detection. A rough estimation is sufficient to select the correct neighborhood and even outliers do not spoil the algorithm as the contained neighborhoods will have very low weights and hence will not influence the reconstructed value.

To achieve best results  $v'$  can be assigned the result  $v$  and the method may be iterated  $l$  times. Experiments have shown that  $l = 2$  iterations are usually sufficient and further loops do not yield significant enhancements. Incorporating the simplification shown in Equation (4) the total complexity of the proposed method is  $O(N^2 \cdot S^2 \cdot R^2 \cdot \tau \cdot l)$ .

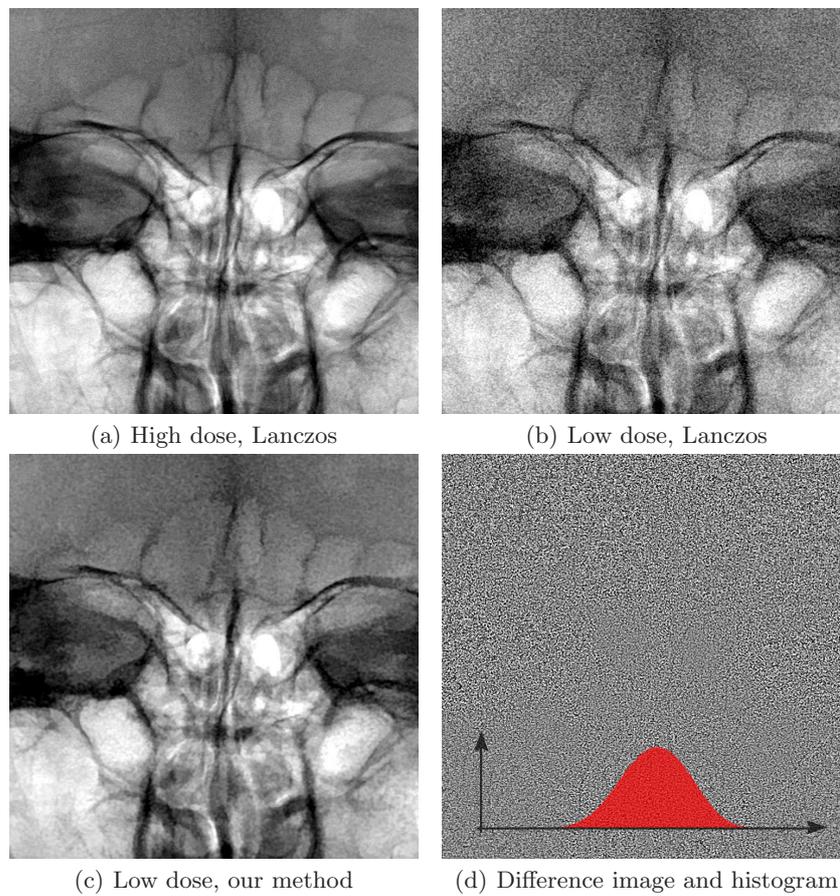
## 2.4 Test set

To evaluate the proposed method we used clinical fluoroscopy sequences containing continuous navigation frames acquired at low X-ray doses and single exposures at higher doses of the same anatomy. Thus we were able to compare impression of sharpness and noise with traditional upsampling as well as the high dose image serving as ground truth. In addition we processed a sequence containing an artificial rotating phantom and evaluated the noise suppression capability and resolution of our method.

### 3 Results

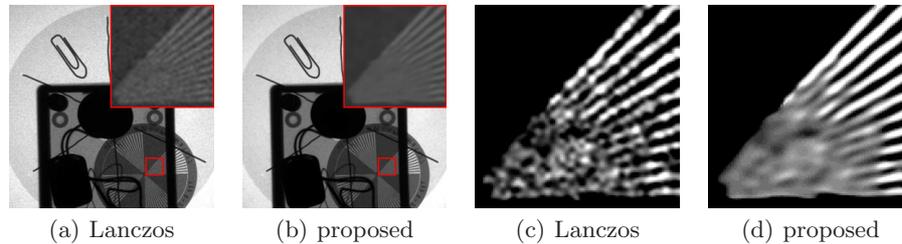
We visualize the results for two test sequences. The first one is a frame taken from a clinical angiography. Fig. 1 shows a part of the cranium magnified by traditional Lanczos interpolation compared to the proposed method.

As one can see, Lanczos interpolation significantly suffers from the inherent noise in low dose frames. The proposed method achieves a noise impression that is comparable to a high dose exposure and preserves sharpness but is not able to reconstruct all details contained in the high dose exposure from multiple low dose frames. The difference image between both methods contains only noise and the histogram shows an almost ideal Gaussian distribution.



**Fig. 1.** Magnifying a cranial angiography: (a) Single frame acquired with high X-ray dose enlarged with Lanczos interpolation, (b) same anatomy acquired with low dose enlarged with Lanczos filter and (c) from  $\tau = 10$  frames with the proposed magnification method, (d) contains the difference between Lanczos filtered and proposed image.

**Fig. 2.** Comparison of magnification techniques: The part marked by the red rectangle is upsampled for enhanced visibility. The complete detail view is magnified using (a) Lanczos interpolation and (b) the proposed method. (c) and (d) contain contrast enhanced detail views respectively.



The second sequence (Fig. 2) has been chosen to evaluate the spatial resolution. The Siemens star contained in this sequence is corrupted by noise and Lanczos as well as other interpolation methods even amplify this noise upon magnification. The proposed method though slightly degrading contrast achieves superior magnification and even reconstructs the corrupted pattern. Noise is heavily reduced and lines are better recognizable.

## 4 Discussion

In direct comparison to simple interpolation the proposed magnification shows superior quality. Evaluation of difference images indicates that no relevant structure is lost due to the filtering operation. Hence it is possible to achieve enhanced magnification of noisy image data without tampering the diagnostic information. As a rule of thumb the X-ray dose of a single high dose exposure corresponds to 30 seconds of navigation at low dose. As this technique might render some of the exposures unnecessary, we reduce the overall dose for patients and staff.

The high complexity yet prevents the algorithm from being executed interactively, however we have successfully implemented a real-time similarity based denoising [5] and as computational power grows it will be possible to do the same for the proposed method.

## References

1. Keys R. Cubic convolution interpolation for digital image processing. *IEEE Trans ASSP*. 1981;29(6):1153–60.
2. Turkowski K. Filters for common resampling tasks. *Apple Computer*; 1990.
3. Buades A, Coll B, Morel JM. A non-local algorithm for image denoising. *Proc IEEE CVPR*. 2005;2:60–5.
4. Ebrahimi M, Vrscay ER. Examining the role of rcale in the context of the non-local-means filter. *LNCS*. 2008;5112:170–81.
5. Gooßen A, Pralow T, Grigat RR. Medical X-ray image enhancement by intra-image and inter-image similarity. *Proc SPIE*. 2009.