

TOMOAND: A package for Anisotropic Nonlinear Diffusion in Electron Tomography

J. J. Fernandez ^{1,2,3}, S. Li ^{1,4}, V. Lucic ⁵

¹ MRC Laboratory of Molecular Biology, Hills Road, Cambridge, UK.

² Dept. Computer Architecture, University of Almeria, Almeria, Spain.

³ National Centre for Biotechnology (CNB-CSIC), Madrid, Spain. (present address)

⁴ Dept. Biochemistry and Biophysics, UCSF, USA. (present address)

⁵ Dept. Structural Biology, MPI Biochemistry, Martinsried, Germany.

September 2011

Contact: jfernandez.software@gmail.com

[Web site](#)

The package TOMOAND contains programs for denoising 3D tomograms according to the approach described in *J.Struct.Biol.* 144:151-162 (2003), *Comp.Sci.Eng.* 7(5):54-61 (2005) and *Lec.Not.Comp.Sci.* 4788:60-69 (2007). The program tomoand implements a 3D anisotropic nonlinear diffusion method that combines the EED (Edge-Enhancing Diffusion) and CED (Coherence-Enhancing Diffusion) diffusion modes simultaneously. The processing of the voxels as CED or EED is based on a threshold computed from an area of the input volume that contains only noise.

Globally, the program analyzes each voxel in the volume to define if it is part of (1) an isotropic feature or noise or (2) a plane-like or line-like feature. For the former (1), EED is applied, i.e. structure-preserving filtering is applied to smooth the volume, but preserving edges. For the latter (2), CED is applied, i.e. the idea is to enhance the structures: in lines, smoothing is applied along the lines; For planes, smoothing is applied across the plane. The present version contains some new improvements, such as a strategy to further smooth the background, the automatic tuning of the parameters related to EED and CED, and the stopping criterion based on *Noise Estimate Variance*.

The program tomoand has been optimized to reduce memory requirements and processing time. Multithreading techniques are used in the most expensive routines to take advantage of current multicore computers.

Table of Contents

1. Installation
2. Description
3. Parameters
4. The output report
5. Advices and comments
6. History
7. References

1 Installation

- Uncompress the file `TOMOAND.zip`.
- You will find these directories:
 - `bin`, with the executable (`tomoand`).
 - `com`, with an example of script for tomoand (`tomoand.csh`).
 - `doc`, with the documentation PDF (`tomoand.pdf`).
- Set up your `PATH` environment variable to have direct access to the executables

Tested machines

So far, the package has been built and tested on the following platforms:

Linux machines based on x86 processors (32 bits and 64 bits)
Intel-based Macs (64 bits) under OSX (10.6 and onwards)

In case of problems in these or other platforms, contact me and I'll try my best to sort them out.

2 Description

The method for denoising tomograms implemented in `tomoand` is based on Anisotropic Non-linear Diffusion, and combines two diffusion modes for structure preserving noise reduction:

- **EED** (Edge Enhancing Diffusion) mode. It denoises while preserving and enhancing edges.
- **CED** (Coherence Enhancing Diffusion) mode. It enhances curvilinear and planar structures.

The program proceeds as follows:

- The program uses a subvolume of 10x10x10 voxels that contains only noise (hereinafter referred as Noise subvolume) to monitor the denoising process. The user has to input the coordinates of that subvolume. This Noise subvolume is a key piece of information for the denoising process. So, it should be carefully selected from the tomogram.

The program carries out an analysis of the noise area, essentially consisting of computing the statistics: mean, variance, gradient, coherence, etc. All these metrics will be used to supervise the process and compute the values for some parameters.

Based on the statistics in the noise area, a threshold is computed. This threshold will be used to process a voxel as EED (if its local structure indicates that is likely to be noise or part of a homogeneous area) or CED (if its local structure indicates that is likely to be a linear or planar structure).

- The program then analyzes each voxel based on the gradient magnitude and the local structure. If local structure measured around the voxel is likely to be noise or part of a homogeneous area, EED diffusion mode is applied. Otherwise, CED diffusion mode is applied as presented in the article in JSB'03:
 - if the structures are considered linear, they are applied a smoothing along the preferred direction.
 - if the structures are considered as planes, they are applied a 2D smoothing along the orthogonal directions that define the plane.
- This program also includes a strategy to further smooth the background. It uses the noise subvolume to monitor the average density level in the background region. During the denoising process, all the voxels whose density are poorer than the average background density are smoothed out with linear diffusion (i.e. Gaussian filtering) at the current iteration. This strategy was introduced in the article CSE'05.
- The diffusion process runs up to a maximum number of iterations set up by the user. However, the diffusion may stop earlier according to an automatic stopping criterion called *Noise Estimate Variance (NEV)*, if set up by the user. This stopping criterion was introduced in the article CSE'05. Again, the Noise subvolume has an important role here.

- This version of the program has an enriched output throughout its execution. This allows the user to check what's exactly going on in both the whole volume and the noise subvolume.

In particular, there is a lot of information about the noise subvolume: statistics on the density values, and also on the gradient, and on the averaged coherence. The information on the gradient and the averaged coherence can provide clues about the values for the parameters K (for EED) and C (for CED), respectively.

- The program may also help the user to automatically and dynamically tune the parameters C and K . The automatic values are estimated from the noise subvolume, based on its statistics. This tuning is carried out at each iteration. As a consequence, the automatic values of C and K are not fixed, but they change dynamically through the process depending on the current status of the noise subvolume. This strategy was introduced in the article LNCS'07.

A detailed description of the procedures implemented in the program can be found in the following articles:

[An improved algorithm for anisotropic nonlinear diffusion for denoising cryotomograms.](#)

JJ Fernandez and S Li

Journal of Structural Biology 144:152-161, 2003.

<http://dx.doi.org/10.1016/j.jsb.2003.09.010>

[Anisotropic nonlinear filtering of cellular structures in cryoelectron tomography.](#)

JJ Fernandez and S Li

Computing in Science and Engineering 7(5):54-61, 2005.

<http://dx.doi.org/10.1109/MCSE.2005.89>

[Three-dimensional anisotropic noise reduction with automated parameter tuning. Application to electron cryotomography.](#)

JJ Fernandez, S Li and V Lucic

Lecture Notes in Computer Science 4788:60-69, 2007.

http://dx.doi.org/10.1007/978-3-540-75271-4_7

Please, cite (any of) these articles if you use TOMOAND in your works.

3 Parameters

(The reader is referred to the articles JSB 144:151-162 (2003), CSE 7(5):54-61 (2005) and LNCS'07 for a better comprehension of some of these parameters).

The set of parameters of the program are: (optional parameters are indicated with [])

- Name of the input volume
- Name of the output volume
- Number of iterations and [Use of Stopping Criterion]
- C constant for CED
- K constant for EED
- Parameter of CED/EED Balance (known as “pr”).
- Proportion of CED along 2nd eigenvector (known as “prced”).
- Proportion of Smoothing based on Grey level (known as “prgr”).
- Coordinates of the Noise Area (X Y Z)
- So and S: sigmas for initial gaussian filtering and for averaging structure tensor, respectively.
- ht and [N. Threads]
- [Diffusion mode] (this only for advanced users)

Description of the parameters:

- *Names of the input and output tomograms.*

The format of the input tomogram is MRC, any mode (0 - unsigned char; 1 - short; 2 - float), as typically used with standard packages like IMOD. The format of the output tomogram is MRC using the same mode as the input tomogram. The input map's header is preserved and transferred to the output map. The program automatically detects the endian of the numbers and applies the proper conversions.

- *Number of Iterations and [Use of Stopping Criterion]*

The number of iterations represents the maximum number of denoising cycles. The number of iterations depend on the time interval (ht) that has been chosen. For the standard value of ht=0.1, a number of iterations around [20,60] yields fairly good results. More iterations provide further smoothing/enhancing. If too many iterations are used, then some interesting structural features may turn out to be blurred or artefactually deformed.

The “Use of Stopping Criterion” is set up with a single character: 'Y' (stands for YES) or 'N' (stands for NO). This parameter is optional, and by default is taken 'N'. If “Use of Stopping Criterion” is 'Y', at the end of each iteration the denoising process checks out whether the NEV criterion is fulfilled. If the stopping conditions are reached, the denoising stops. The stopping criterion only works for tomograms with mode 2 (float).

- *C constant for CED.*

(see papers in JSB'03, CSE'05 and LNCS'07 for further details on this parameter).

It influences the smoothing along the preferred directions for CED voxels. Note that a value of 0 makes the strongest CED diffusion. The value of *C* depends on the density range in the volume. So, it is necessary to try several values in order to find out the optimal one. The report that the program shows on console with the statistics (in particular, the averaged coherence) in the Noise volume is useful to tune this parameter.

In the present version, a new feature has been included. If *C* is set up to -1, then this parameter is automatically and dynamically tuned at each iteration based on the averaged coherence in the Noise subvolume.

- *K constant for EED.*

(see papers in JSB'03, CSE'05 and LNCS'07 for further details on this parameter).

This parameter represents a threshold on the gradient magnitude. For a given voxel, if the magnitude of the gradient is greater than the threshold *K*, then it is considered as part of an edge. In such a case, smoothing across the edge is avoided. Therefore, edges and structure boundaries are preserved and enhanced. Otherwise (i.e., the gradient is below the threshold) the voxel is considered as part of an isotropic structure. Hence a smoothing more similar to isotropic gaussian filtering is applied.

The value of *K* depends on the density range in the volume. So, it is necessary to try several values in order to find out the optimal one. The output report that the program shows on console with the statistics (in particular, the gradient) in the Noise volume is useful to tune this parameter.

In the present version, a new feature has been included. If *K* is set up to -1, then the program automatically and dynamically tunes this parameter throughout the diffusion process based on the gradient in the Noise subvolume.

- *CED/EED Balance Parameter (pr).*

This parameter influences the threshold on which the switch between CED and EED is based. This is a factor in [0.0,1.0] that is applied over the maximum difference between the first and the third eigenvalues ($\mu_1 - \mu_3$) found in the area containing noise. The threshold that is used to switch between CED and EED modes is then set up as:

$$threshold = pr * \max_{noise}(\mu_1 - \mu_3).$$

Those voxels whose local structure given by $\mu_1 - \mu_3$ is greater than that threshold will be processed with CED. Otherwise, EED will be applied. Low values of *pr* involves less EED denoising and more CED enhancement. High values of *pr* involves more EED denoising and less CED enhancement.

In general, cleaner results are obtained if a *pr* parameter with higher value is applied. For example, *pr*=0.5 or *pr*=1.0 yield cleaner backgrounds. But if enhancement is preferred over smoothing, this parameter should have lower values, say *pr*=0.25 or *pr*=0.1. Note that a value of *pr*=0.0 makes the program apply only CED.

- *Proportion of CED along 2nd eigenvector (prced).*

This influences the diffusion along the 2nd eigenvector for CED. It is specially useful for structures that, at a local scale, can be considered as plane-like. *prced* is a factor in [0.0,1.0] that is applied over the CED diffusion formula along the 2nd eigendirection:

$$\lambda_2 = \alpha - (1 - \alpha) * prced * \exp\left(\frac{-C}{(\mu_1 - \mu_2)^2}\right)$$

Higher values of this parameter allows better enhancement along the second eigendirection. However, it should be taken into account that very high values may create artifacts. Advisable values are in [0.25,0.75].

- *Proportion of Smoothing based on Grey level (prgr).*

This parameter influences the amount of background smoothing as described in CSE'05. It must have a value in the interval [0.0,1.0]. The program, at each iteration, computes the mean grey level in the Noise subvolume. Then a threshold is computed according to the following formula:

$$(\text{NoiseMeanGrey} - \text{min}) * prgr + \text{min}$$

where min is the minimum density level in the tomogram at that iteration.

Then all the voxels in the tomogram with a density lower than that threshold are processed with linear diffusion (i.e. Gaussian filtering) in that iteration.

Note that a value of *prgr*=0 turns off background smoothing and makes the program behave as in the previous version, described in JSB'03. If *prgr*=1, all the voxels in the tomogram with a density lower than the mean grey level in the Noise subvolume will be applied linear diffusion. Intermediate values of *prgr* in the interval [0.0,1.0] makes the background smoothing not so strong.

The description above applies only for those tomograms where the density is “white” (i.e. higher density is higher value in the tomogram). If the input tomogram has the density in “black” (i.e. higher density is represented with lower value in the tomogram) you can get the same effect with only input the parameter with negative sign “-”. In such a case, the threshold is computed according to the following formula:

$$\text{max} + (\text{max} - \text{NoiseMeanGrey}) * prgr$$

where max is the maximum density level in the tomogram at that iteration.

Now the program will filter those voxels with density higher than that threshold by means of linear diffusion. If *prgr*=-1, all the voxels in the tomogram with a density higher than the mean grey level in the Noise subvolume will be applied linear diffusion. Intermediate values of *prgr* in the interval [-1.0,0.0] makes the background smoothing not so strong.

- *Coordinates of the Noise Area (X Y Z).*

These coordinates define the first voxel of a 10x10x10 area containing only noise (i.e. background). This area is considered for estimating the characteristics of the noise in the volume at each iteration, and it is continuously supervised as described in JSB'03, CSE'05 and LNCS'07, and throughout this manual. In setting up this parameter, take into account that (1 1 1) are the coordinates of the first voxel stored in the volume.

- *Initial sigma (So) and sigma for averaging struct.tensor (S).*

The initial sigma (S_o) refers to the scale of the noise and is used to apply an initial gaussian filtering at the beginning of each iteration. It is intended to regularize the computation of the gradients, in order to make them less sensitive to noise. It allows removing of fine-grain noise. Essentially, every feature with dimensions smaller than this sigma is considered noise and, hence, strongly smoothed. Be careful with this parameter. If you indicate a high value, the structures in the denoised volume may be blurred. A value in the range [0.25,1.0] is advisable.

The parameter sigma S refers to the scale of the features to be enhanced in the tomogram. It is used for averaging the structure tensor (this averaging is only applied for CED) with the aim of estimating the local structure at the scale S . The value of this parameter depends on the size of the features that the user may want to enhance. As an advice, a value around 2.0 may be useful.

- *ht and [N. Threads].*

ht is the time interval. The maximum recommended value is 0.1. Greater values may produce instabilities in the solution.

The parameter 'N. Threads' is optional. If no value is introduced, the program automatically creates as many threads as cores available in the computer, which will collaborate in the denoising process. In other words, by default the program intends to fully exploit the parallel capabilities of multicore (dual, quad, hexa, etc) computers. This parameter 'N. Threads' allows the user to specify a particular number of threads instead of using the default value. In that case, the user should bear in mind that the maximum number of threads should not be higher than the number of cores in the computer (i.e. 4 for a computer with a quad-core; 2 for a dual-core machine; etc).

- *[Diffusion mode] (this only for advanced users).*

This parameter is optional and is intended only for advanced users. This parameter can take 3 values:

0: stands for the standard hybrid approach based on EED and CED, as described in JSB'03, CSE'05 and LNCS'07. For EED, the structure tensor is used, for CED the averaged structure tensor (component-wise averaged with sigma S) is used. The CED/EED switch is based on the non-averaged structure tensor. The diffusion associated to EED then runs along v_2 and v_3 (i.e., the 2nd and 3rd eigenvectors) since μ_2 and μ_3 are both 0. This is the approach taken by default.

1: This value enables the use of the averaged structure tensor for both EED and CED. The CED/EED switch is still based on the non-averaged structure tensor. With this strategy, the diffusion associated to EED then can be set up according to the local structure (isotropic,plane,line), given as a function of μ_2 and μ_3 , in a similar way to that applied with CED. In this approach, the magnitude of the gradient used for the EED formula comes from the non-averaged gradient.

2: This value enables the use of the averaged structure tensor for EED, CED and the CED/EED switch. This strategy is similar to that just mentioned (1), except that now the CED/EED switch is also based on the averaged structure tensor. The setting up of the parameter pr is then a little more tricky. In this approach, the magnitude of the gradient used for the EED formula also comes from the non-averaged gradient.

On execution, the program asks the user for all these parameters, interactively. However, it is recommended the use of scripts to edit the parameters and launch the program. An example of script is included in the ZIP file: `tomoand.csh`.

Examples of parameter sets and scripts

Some examples of parameter sets and scripts are reproduced here:

Example 1

The use of stopping criterion is ON. The standard hybrid CED/EED approach is selected. As many threads as cores available in the computer will collaborate in the denoising process (by default if 'N. Threads' is left empty). The remaining parameters are chosen as described above.

```
tomoand << eof
Input_Tomogram.mrc # Input file
Output_Tomogram.mrc # Output file
20 Y # No Iterations and [Use of Stopping Criterion]
4 # C Constant for CED
2 # K Constant for EED
0.5 # CED/EED Balance Parameter
1.0 # Proportion of CED along 2nd eigenvector.
1.0 # Proportion of Smoothing based on Grey level
190 96 30 # Coordinates of the Noise Area ( X Y Z )
1.0 2.0 # Initial sigma and sigma for averaging struc.Tensor
0.1 # ht and [N. Threads]
0 # [Diffusion Mode]: Standard Hybrid (0); Averaged EED(1|2)
eof
```

Example 2

The use of stopping criterion is OFF (selected by default). The standard hybrid CED/EED approach is selected by default. The parameter C is left to be automatically tuned by the own program based on the statistics of the Noise subvolume. As many threads as cores are used (default). The remaining parameters are chosen as described above.

```
tomoand << eof
Input_Tomogram.mrc # Input file
Output_Tomogram.mrc # Output file
15 # No Iterations and [Use of Stopping Criterion]
-1 # C Constant for CED
2 # K Constant for EED
0.5 # CED/EED Balance Parameter
1.0 # Proportion of CED along 2nd eigenvector.
1.0 # Proportion of Smoothing based on Grey level
190 96 30 # Coordinates of the Noise Area ( X Y Z )
1.0 2.0 # Initial sigma and sigma for averaging struc.Tensor
0.1 # ht and [N. Threads]
eof
```

Example 3

The use of stopping criterion is OFF (selected by default). The diffusion approach based on the use of the averaged structure tensor for EED, CED and CED/EED switch is ON. The “Proportion of Smoothing based on Grey level” is set up to 1.0, and the negative sign - indicates that the density in the tomogram is black. The parameters C and K are left to be automatically tuned by the own program based on the statistics of the Noise subvolume. The N. threads (3) is introduced manually. The remaining parameters are chosen as described above.

```
tomoand << eof
Input_Tomogram.mrc # Input file
Output_Tomogram.mrc # Output file
25 # No Iterations and [Use of Stopping Criterion]
-1 # C Constant for CED
-1 # K Constant for EED
0.5 # CED/EED Balance Parameter
1.0 # Proportion of CED along 2nd eigenvector.
-1.0 # Proportion of Smoothing based on Grey level
190 96 30 # Coordinates of the Noise Area ( X Y Z )
1.0 2.0 # Initial sigma and sigma for averaging struc.Tensor
0.1 3 # ht and [N. Threads]
2 # [Diffusion Mode]: Standard Hybrid (0); Averaged EED(1|2)
eof
```

Example 4

This is an example where all the automatic features of the program are ON: the use of stopping criterion is ON and the parameters C and K are left to be automatically tuned by the own program based on the statistics of the Noise subvolume. The maximum number of threads (as many as cores) is used. The standard hybrid CED/EED approach is selected by default. The remaining parameters are chosen as described above.

```
tomoand << eof
Input_Tomogram.mrc # Input file
Output_Tomogram.mrc # Output file
20 Y # No Iterations and [Use of Stopping Criterion]
-1 # C Constant for CED
-1 # K Constant for EED
1.0 # CED/EED Balance Parameter
1.0 # Proportion of CED along 2nd eigenvector.
1.0 # Proportion of Smoothing based on Grey level
190 96 30 # Coordinates of the Noise Area ( X Y Z )
1.0 2.0 # Initial sigma and sigma for averaging struc.Tensor
0.1 # ht and [N. Threads]
eof
```

4 The output report

The tomoand program has an enriched output with respect to the previous versions. The information is provided throughout the denoising process, iteration by iteration. It helps the user to supervise what's going on, and how the denoising is affecting the input tomogram. It also provides clues for tuning the parameters.

Concerning the tuning of the parameters K and C , two are the most important items: the gradient and the averaged coherence in the noise subvolume:

- Gradient:

For every voxel in the noise subvolume, the gradient magnitude $|\nabla I|$ is computed. The noise subvolume has $10 \times 10 \times 10$ voxels. As a consequence, $10 \times 10 \times 10$ measures of the gradient in the noise subvolume is obtained. The program shows the mean (μ_g) and the sigma (i.e. standard deviation, σ_g) of that gradient magnitude population:

Gradient in Noise subvol. (mean,sigma): $\mu_g \quad \sigma_g$

The parameter K acts as a contrast threshold so that structures with $|\nabla I| > K$ are regarded as edges, otherwise as the interior of a region. Consequently, the information on the gradient is helpful to set up K .

The mean μ_g is actually used when automatic tuning of K is used. In turn, when manual tuning is used, this information provides clues about a value for K . Probably, a proper K should be a value in the range $[\mu_g, \mu_g + 2 * \sigma_g]$. Note that μ_g and σ_g decrease as iterations evolve. The user should choose the information on μ_g and σ_g provided in the early iterations to set up the parameter K .

- Averaged coherence:

The averaged coherence is defined as $(\mu_1 - \mu_3)^2$, where μ_1 and μ_3 are the eigenvalues of the averaged structure tensor.

For every voxel in the noise subvolume, the coherence $(\mu_1 - \mu_3)^2$ is computed. As a consequence, $10 \times 10 \times 10$ measures of the coherence in the noise subvolume is obtained. The program shows the mean (μ_c) and the sigma (i.e. standard deviation, σ_c) of that coherence population:

Av. Coherence in Noise subvol. (mean,sigma): $\mu_c \quad \sigma_c$

This parameter C acts as a threshold so that structures with $(\mu_1 - \mu_3)^2 > C$ are regarded as lines (something similar happens with planes). Consequently, the information on the averaged coherence is helpful to set up C .

The mean μ_c is actually used when automatic tuning of C is used. In turn, when manual tuning is used, this information provides clues about a value for C . Probably, a proper C should be a value in the range $[\mu_c, \mu_c + 2 * \sigma_c]$. Note that μ_c and σ_c decrease as iterations evolve. The user should choose the information on μ_c and σ_c provided in the early iterations to set up the parameter C .

Example of the output report

The following is a piece of the output report that the program shows during the execution. This piece corresponds to the information for the fifth iteration of the denoising process:

```
=====
iteration number:                5 / 20
diffusion time:                  0.500000

+ Analysis of noise subvolume:
-----
CED/EED switch threshold        :    65.582
Density level-based threshold   :   130.748
Gradient in Noise subvol.      (mean,sigma):    2.454    1.246
Av. Coherence in Noise subvol. (mean,sigma):    4.261    7.079
-----

+ Diffusion:
-----
AND Parameters:
      K: 2.453870
      C: 4.261443
CED voxels/total voxels        : 1084525/11132000
-----

Statistics at iter 5:           (mean,  sigma,  min,  max)
* Stats.whole_volume:         136.250  12.493  57.743  259.890
* Stats.noise_subvolume:     134.341   4.727  124.103  155.774
-----

NEV criterion:
  Noise Estimate Variance:      688.573 > 648.611 ... stop
=====
```

The index of the current iteration (5) out of the maximum (20), and the corresponding diffusion time are shown. In this case, ht was 0.1. Then, the analysis of the noise subvolume is carried out, and statistics are computed and shown: first, the threshold used as CED/EED switch; second, the threshold on the density if background smoothing is ON; third, the statistics on the gradient in the noise subvolume (mean and sigma of the $10 \times 10 \times 10 = 1000$ gradient measures). The information of the gradient can be used for proper manual tuning of the parameter K . If automatic K is ON, the program will use the mean gradient as the value for K . Finally, the statistics on the coherence from the averaged structure tensor in the noise subvolume are shown (mean and sigma). This information can be used to manually tune the parameter C . If automatic C is ON, the program will use the mean Av. Coherence as the value for C .

In the diffusion process, the values for K and C are shown if their automatic tuning is ON. Then, the number of voxels processed as CED is reported. After the diffusion, the density statistics of the whole tomogram and the noise subvolume are printed on output.

Finally, if the use of stopping criterion is ON, the noise estimate variance is shown (688.573 in the example) and the initial noise variance (646.611) are shown. If the former is greater, then the diffusion stops.

5 Advices and comments

- *Automatic tuning of the parameters C and K .*

Setting up the parameters C and K sometimes may be painful. In the present version, it has been included the possibility of automatic tuning those parameters by setting them up to a value of -1. This automatic tuning is carried out by the program iteration-by-iteration, and based on the Noise Area specified by the user. When setting up the parameters manually, the user has to take into account that those parameters will be fixed for the whole denoising process. On the contrary, automatic tuning make the parameters dynamically change, based on the statistics in the Noise area. At the beginning, at the early iterations, C and K have greater values. As iterations evolve, C and K get smaller values according to the remaining noise in the tomogram.

This automatic tuning seems to work fine and it is encouraged to use it. Note that it is possible to set up automatic tuning for one of those parameters, while the other one can be tuned manually.

- *Manual tuning of the parameters C and K .*

Manual Tuning of the values for C and K fixes those parameters for the whole denoising process. The user has to take into account that both parameters depend somewhat on the scale and the standard deviation of the densities in the input tomogram. As a consequence, every tomogram may require a proper tuning of those parameters.

The programs shows a report on the statistics of the noise area and the whole tomogram during the denoising process, as shown above. This allows the user to check what's exactly going on in both the whole volume and the noise subvolume. This information is key to properly tune the parameters.

In the output report, there is a lot of information about the noise subvolume: statistics on the density values, and also on the gradient, and on the averaged coherence. The information on the gradient and the averaged coherence can provide clues about the values for the parameters K (for EED) and C (for CED), respectively.

- *Preliminary recipes for tuning of the parameters C and K .*

In the earliest versions of the package, there was limited information about the statistics of the noise area shown in the output report of the program. Hence, it was not so easy to find out the best parameter set. At those dates, there was a recipe that helped. The idea was transforming the density range of the input tomogram to a given, known, range. This is the sequence of steps to follow:

- Once the tomogram is reconstructed, the density is rescaled to be in the range $[0.0,255.0]$. In order to do this, the mean μ and standard deviation σ of the original reconstructed tomogram have to be computed. Then a cutoff over the densities is applied so that the density in the tomogram is in the range $[\mu - 4 * \sigma, \mu + 4 * \sigma]$ or $[\mu - 5 * \sigma, \mu + 5 * \sigma]$. Subsequently, the densities are rescaled to $[0.0,255.0]$.
- For these range of density, the values of K in the range $[1,4]$ and values of C in the range $[1,8]$ usually work fine and produce good denoising/enhancement effects.

Another rule of thumb that was found to be useful without the need of rescaling the densities of the tomogram was to set up the parameters as: $K = (\text{sigma}/25)^4$ and $C = (\text{sigma}/25)^4$.

These recipes still can be useful. However, the automatic tuning of the parameters or the manual tuning based on the statistics of the noise area may be the best way to proceed.

- *Original denoising approach by Frangakis and Hegerl, 2001.*

This original approach by Frangakis and Hegerl described in JSB 135:239-250, 2001 can be obtained with this program by using the following parameter settings:

- * No use of Stopping criterion.
- * Null Proportion of CED along 2nd eigenvector, i.e. $prcd=0$
- * Null Proportion of Smoothing based on Grey Level, i.e. $prgr=0$
- * Coordinates of the Noise area pointing to the first voxel:, i.e. 1 1 1
- * Null Initial sigma, i.e. $So=0$
- * sigma for averaging struct.tensor, $S=0.1$
- * $ht=0.1$.
- * Diffusion Mode=0 to turn on the standard hybrid CED/EED approach.

- *Differences in EED with the previous version of the package.*

In the current version of the package, the diffusion associated to EED runs along v2 and v3 (i.e., the 2nd and 3rd eigenvectors) when the standard hybrid approach is used. The fact that the non-averaged structure tensor is used for EED produces null μ_2 and μ_3 . Under those circumstances, the diffusion must run along v2 and v3. However, the previous version of the package (May 2004) only applied diffusion along v3, as mentioned in JSB'03 and CSE'05.

- *Erratum in the JSB'03 paper concerning the CED/EED switch threshold.*

In the paper, it is said "The threshold is computed from the mean value of $(\mu_1 - \mu_3)$ in a subvolume of the image containing only noise.". The correct sentence is "The threshold is computed from the maximum value of $(\mu_1 - \mu_3)$ in a subvolume of the image containing only noise."

- *Bugs in the previous version of the package.*

There was a *bug* in the previous version of the package (May 2004) in the formula for EED. Due to a mistake, the formula implemented was:

$$1 - \exp(-3.31488/((\text{gradient}/K)^{32}))$$

Now the correct formula is implemented in the present version:

$$1 - \exp(-3.31488/((\text{gradient}/K)^8))$$

6 History

V. May 2004

- Very early version of the program as described in our article JSB 144:151-162 (2003).

V. April 2008

- The original package as described in our articles JSB 144:151-162 (2003), CSE 7(5):54-61 (2005) and LNCS'07.

V. September 2011

- The program reads and writes datasets in MRC format, any mode (0, 1, 2). It writes the output tomogram using the same mode as the input.
- The Stencil 5x5x5 is no longer available. The program now only uses the Stencil 3x3x3, using a maximum recommended time interval of $ht=0.1$.
- The program automatically detects the number of cores and by default, the program fully exploits the computer power in the computer by creating as many threads as cores.

Acknowledgements

We thank Tony Crowther for help, support and fruitful discussions during the entire work. The very early versions of this package were inspired on the code that Drs. A. Frangakis and R. Hegerl provided us in 2002 (JSB 135:239-250, 2001). Some routines related to I/O used in this package are inspired on Bsoft (Heymann, JSB 133:156–169, 2001). The development of this package has been possible thanks to grants from the Spanish Ministry of Science, the MRC, EMBO, HFSP and EU through the network of excellence 3DEM-NoE.

7 References

- of special interest.

References describing the methods implemented in the TOMOAND package

- J.J. Fernandez, S. Li. [An improved algorithm for anisotropic nonlinear diffusion for denoising cryotomograms](#). *J. Struct. Biol.*, vol. 144, pp. 152–161, 2003.
- J.J. Fernandez, S. Li. [Anisotropic nonlinear filtering of cellular structures in cryoelectron tomography](#). *Computing in Science and Engineering*, vol. 7, issue 5, pp. 54–61, 2005.
- J.J. Fernandez, S. Li, V. Lucic. [Three-dimensional anisotropic noise reduction with automated parameter tuning. Application to electron cryotomography](#). *Lecture Notes in Computer Science*, vol. 4788, pp. 60–69, 2007.

References on denoising in electron tomography with anisotropic/isotropic nonlinear diffusion

- A.S. Frangakis, R. Hegerl. [Noise reduction in electron tomographic reconstructions using nonlinear anisotropic diffusion](#) *J. Struct. Biol.*, vol. 135, pp. 239–250, 2001.
- A.S. Frangakis, A. Stoschek, R. Hegerl. [Wavelet transform filtering and nonlinear anisotropic diffusion assessed for signal reconstruction performance on multidimensional biomedical data](#). *IEEE Trans. BioMed. Engineering*, vol. 48, pp. 213–222, 2001.
- R. Narasimha et al., and S. Subramaniam. [Evaluation of denoising algorithms for biological electron tomography](#). *J. Struct. Biol.*, vol. 164, pp. 7–17, 2008.
- J.J. Fernandez. [TOMOBFLOW: Feature-preserving noise filtering for electron tomography](#). *BMC Bioinformatics* 2009, 10:178.

References on denoising in electron tomography with other methods

A. Stoschek , R. Hegerl. Denoising of Electron Tomographic Reconstructions Using Multiscale Transformations. *J. Struct. Biol.* vol. 120, pp. 257–265, 1997.

W. Jiang, M.L. Baker, Q. Wu, C. Bajaj, W. Chiu. Applications of a bilateral denoising filter in biological electron microscopy. *J. Struct. Biol.* vol. 144, pp. 114–122, 2003.

R.S. Pantelic, R. Rothnagel, C.Y. Huang, D. Muller, D. Woolford, M.J. Landsberg, A. McDowall, B.Pailthorpe, P.R. Young, J. Banks, B. Hankamera, G. Ericksson. The discriminative bilateral filter: An enhanced denoising filter for electron microscopy data. *J. Struct. Biol.* vol. 155, pp. 395–408, 2006.

P. van der Heide, X.-P. Xu, B.J. Marsh, D. Hanein, N.Volkman. Efficient automatic noise reduction of electron tomographic reconstructions based on iterative median filtering. *J. Struct. Biol.* vol. 158, pp. 196–204, 2007.

References on denoising with nonlinear diffusion

- D. Barash, "A fundamental relationship between bilateral filtering, adaptive smoothing and the nonlinear diffusion equation," *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 24, pp. 844–847, 2002.
- D. Barash and D. Comaniciu, "A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift," *Image and Vision Computing*, vol. 22, pp. 73–81, 2004.
- B. Jahne, H. Scharr, and S. Korkel, *Handbook of Computer Vision and Applications, Vol. 2: Signal Processing and Pattern Recognition (B. Jahne, Ed.)*. San Diego: Academic Press, 1999, ch. Principles of Filter Design, pp. 125–152.
- P. Mrazek and M. Navara, "Selection of optimal stopping time for nonlinear diffusion filtering," *Intl. J. Computer Vision*, vol. 52, pp. 189–203, 2003.
- P. Perona, J. Malik, "Scale-space and edge detection using anisotropic diffusion", *IEEE Trans. Patt. Anal. Mach. Intel.* vol. 12(7), pp. 629-639, 1990.
- H. Scharr and D. Uttenweiler, "3D anisotropic diffusion filtering for enhancing noisy actin filament fluorescence images," *Lecture Notes in Computer Science: Pattern Recognition*, vol. 2191, pp. 69–75, 2001.
- D. Uttenweiler, C. Weber, B. Jahne, R. H. Fink, and H. Scharr, "Spatiotemporal anisotropic diffusion filtering to improve snrs and object restoration in fluorescence microscopic image sequences." *J. Biomed. Opt.*, vol. 8, pp. 40–47, 2003.
- J. Weickert, *Anisotropic Diffusion in Image Processing*. Teubner, 1998.
- J. Weickert, "Coherence-enhancing diffusion filtering," *Int. J. Computer Vision*, vol. 31, pp. 111–127, 1999.
- J. Weickert, "Coherence-enhancing diffusion of colour images," *Image and Vision Computing*, vol. 17, pp. 201–212, 1999.
- J. Weickert and H. Scharr, "A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance," *J. Visual Comm. Imag. Repres.*, vol. 13, pp. 103–118, 2002.