A Quantitative Theory of the Non-Local Means Algorithm

N.A. Thacker[a,*], P.A. Bromiley[a] and J.V. Manjon[b]

[a] Imaging Science and Biomedical Engineering, University of Manchester, U.K.
[b] IBIME Group, ITACA Institute, Universidad Politécnica de Valencia, Valencia, Spain.

Noise filtering is a common pre-processing step in image analysis, and is particularly effective in improving the subjective appearance of images. Simple approaches, such as Gaussian filtering, use the signal averaging principle in order to take advantage of the spatial redundancy in the data. However, such techniques make no attempt to confirm that the data being averaged is drawn from the same intensity distribution. Therefore, in locations containing image structure such as edges, where spatially adjacent pixels are drawn from different intensity distributions, simple noise filtering techniques will degrade the image. Typical effects are the blurring of edges or even the removal of small structures. Such techniques are therefore unsuitable for quantitative or clinical use. Edge-preserving filters that average data in the direction orthogonal to the local image gradient, such as anisotropic diffusion [1, 2], have been proposed to avoid this problem. However, the lack of an explicit test for statistical equivalence of the data being averaged is still present, and so such techniques may still erase small features i.e. features whose spatial extent is of the same magnitude as the range used in the local gradient calculation and smoothing kernel. Wavelet-based filters have also been applied to MRI de-noising [3], but tend to introduce characteristic artefacts that can be problematic for clinicians.

The Non-Local Means (NLM) algorithm has recently been proposed to remedy this characteristic problem in noise filtering techniques. NLM explicitly incorporates a test of statistical equivalence to select the intensities to be averaged. The local context of each pixel, i.e. the surrounding patch of intensities, is compared to other image patches in the surrounding data. A similarity measure is applied to detect statistically equivalent patches. The central pixels from the similar patches are then averaged to produce the modified, noise-filtered intensity for the original pixel. The effectiveness of NLM has been demonstrated empirically on clinical data [4]. However, only the basic mathematical description of the technique has been proposed: no investigation of its statistical foundations has been attempted. Without this theoretical foundation it is impossible to identify the implicit assumptions made in NLM, and therefore impossible to confirm whether it is suitable for use on any particular data set. We cannot therefore use it with confidence in clinical or quantitative tasks.

In this paper we analyse the statistical foundations of the NLM approach by constructing an equivalent algorithm based on conventional statistical approaches, using a χ^2 test as the basis of the similarity measure and an analogy with Expectation-Maximisation [5] as the basis of the intensity averaging process. Both the original specification of NLM and the statistically based equivalent are applied to simulated data and real MR images. The aim here is not to evaluate the modified version of NLM through an algorithmic “shoot-out”: rather, the results of each algorithm are compared and, through studying the differences, we identify the characteristic behaviours of the NLM algorithm. In particular, we show that the intensity averaging process in the original algorithm does not incorporate a strict enough definition of statistical similarity to prevent contamination with non-equivalent data. Averaging over non-equivalent data will potentially introduce an image structure dependent bias into the results. Furthermore, we observe that the number of pixels averaged over is spatially varying, dependent on the amount of local equivalent data, in all versions of NLM. Therefore, the noise distribution in the filtered images will also vary spatially, an effect that would have to be taken into account in subsequent quantitative analysis. The two effects are related: the stricter the criterion for equivalence, the more pixels will be left unmodified. We conclude that a more sophisticated treatment of the data is required in order to extract all available information for the averaging process, and suggest avenues for pursuing this approach further.

References

1. G. Gerig, O. Kubler, R. Kikinis, and F. A Jolesz. Nonlinear Anisotropic. Filtering of MRI Data. IEEE Trans. Med Imag., vol. 11, pp. 221-232, 1992.
2. A. Samsonov and C. Johnson. Noise-Adaptive Nonlinear Diffusion Filtering of MR Images With Spatially Varying Noise Levels. Magnetic Resonance in Medicine 52:798-806, 2004.
3. R. D. Nowak, Wavelet-based Rician noise removal for magnetic resonance imaging, IEEE Transactions on Image Processing, 8:10, 1408 -1419, 1999.
4. J.V. Manjon, M. Robles, N.A. Thacker, Multispectral MRI De-noising Using Non-local Means. Proc. MIUA 2007, 41-46, Aberystwyth, 2007.
5. A.P. Dempster, N.M. Laird & D.B. Rubin, ”Maximum Likelihood from Incomplete Data via EM Algorithm”, *Journal of the Royal Society*, **39**, pp.1-38, 1977.

*E-mail: neil.thacker@man.ac.uk