Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Diffusion equations with negentropy applied to denoise mammographic images

P Mayo, F Rodenas, D Ginestar, G Verdú and R Miró

Abstract— **Mammography is a radiographic technique used for the detection of breast lesions. The analysis of the digital image normally requires a previous application of filters as a preprocessing step to reduce the noise level of the image, while preserving important details to carry out a suitable diagnostic. In the literature, there are a large amount of denoising techniques applied to different medical images.**

In this work we have studied the performance of a diffusive filter with a stopping condition based on the statistical concept of negentropy, applied to denoise mammographic images. The negentropy has been succesfully prove with other denoising methods as independent component analysis by the authors in [1]. We have evaluated the final image quality obtained, measuring a root mean squared error between the noise-free initial image and the final restored image and compared the results obtained by this diffusive filter with those obtained by an adaptative non-linear Wiener filter.

I. Introduction

The low contrast of the small tumours to the background, which is sometimes close to the noise, makes that small breast cancer lesions can hardly be seen in the mammography [2]. So a preprocessing stage for the reduction of the noise present in the original image, is important to carry out the analysis of the medical images. The noise reduction improves the visual evaluation of the image and the detection of important features of the medical image.

Different methods have been proposed for image restoration from noisy images [1], [3] and, in particular, they have been successfully employed to denoise medical digital images. An important class of image denoising methods is based on nonlinear diffusion equations [4], [5]. These methods usually appear associated to a variational problem for minimizing an energy functional. It is solved by means of a gradient like method, ending up with a time dependent non linear diffusion equation. In this work we will evaluate the possibility of using the negentropy of the image to develop an automatic stopping criterion for the diffusive filter for the mammographic image restoration process.

II. Image Restoration Method

A. Nonlinear Diffusive Filter

The usual model considered for noisy images is the following, let us assume that f is the sum of an ideal noise-free image $\tilde{f}$ and a noise signal n:

$$f = \tilde{f} + n \; . \tag{1}$$

where f is the observed image.

We assume that the image $\tilde{f}$ and the noise n are uncorrelated and the noise is gaussian and it has zero mean value. The initial point of the diffusion process is the observed image $f = u(0)$ and the iterative filtering produces a family of images $u(t)$ $(t = 0, 1, 2...)$. These images are filtered versions of the original f.

The nonlinear diffusive filter used in this paper is based on the dynamic equation (see [6] for more details) with a constraint:

$$\frac{\partial u}{\partial t} = \vec{\nabla}\left(\frac{\vec{\nabla} u}{\sqrt{\beta^2 + \left\|\vec{\nabla} u\right\|^2}}\right) + \epsilon \nabla^2 u + \mu(f-u) \tag{2}$$

$$\int_\Omega (u-f)^2 \, d\vec{x} = \sigma^2 \int_\Omega d\vec{x} \, , \tag{3}$$

where β, and ϵ are constant. The ϵ is a scale parameter based on the pixel size of the image. The σ^2 is a estimation of the variance of the noise present in the image and Ω is a convex region of R^2 constituting the support space of the surface u, representing the image.

To discretize the equation we have to consider that digital images define a structured grid given by the different pixels constituting the image. The corresponding diffusive filter is obtained discretizing the diffusion operator in each spatial node and for the time discretization we use an additive operator splitting (AOS) [4], [6]. The values of the parameters for the filter used in this work are: $\epsilon = 0.04$, $\beta = 1$, and time step $\Delta t = 0.1$.

B. Stopping time selection by the negentropy criterion

The differential entropy H of a random variable X with probability density function $p(x)$ is defined as:

$$H(X) = -\int p(x) \, \log p(x) \, dx \; . \tag{4}$$

Entropy could be used as a measure of nongaussianity, since a well-known result of information theory is that a gaussian variable has the largest entropy among all random variables of equal variance.

The negentropy function, is defined in terms of the differential entropy H by:

$$J(X) = H(X_G) - H(X) \, , \tag{5}$$

where X_G is a Gaussian random variable of the same covariance matrix as X. $J(X)$ is always non-negative and it is zero if and only if X is a gaussian random variable.

P Mayo and G Verdú are with Department of Chemical and Nuclear Engineering, Polytechnic University of Valencia, Camino de Vera s/n 46022 Valencia (Spain) gverdu@iqn.upv.es

F Rodenas and D Ginestar are with the Applied Mathematical Department of the Polytechnic University of Valencia (Spain)

Unfortunately, the estimation of differential entropy or negentropy, using the definition is quite difficult because it requires the estimation of the density of X. This is overtaken by using simple and robust approximations of entropy. Useful approximations of entropy have been proposed by A. Hyvrinen in the context of independent component analysis [1], [7]. These approximations are based on the estimation of the maximum entropy compatible with the measurements of the random variable X.

In the simplest case, the approximation of the negentropy function becomes:

$$J(X) \propto [\, E\{G(x)\} - E\{G(x_G)\}\,]^2 \,, \tag{6}$$

where X and X_G are normalized to zero mean and unit variance, and G is a nonquadratic function. The more common choices for G that have proved very useful [7], [8] are the following

$$\begin{aligned} G_1(x) &= \frac{1}{a}\log\cosh(ax) \quad (1 \le a \le 2) && (7)\\ G_2(x) &= -\exp(-x^2/2)\,. && (8) \end{aligned}$$

For the filter, one should choose the stopping time T such that the restored image $u(T)$ is as near to the free-noise image $\tilde{f}$ as possible, this means that the euclidean distance $\|u(T) - \tilde{f}\|$ is as small as possible. Let us call such T that minimizes the distance $\|u(T) - \tilde{f}\|$ the optimal stopping time.

The proposed criterion in this work is based on the evolution of the negentropy of the series $u(t)$ $(t = 0, 1, 2...)$. In this work, we show experimentally that it is possible to choose a stopping time from the variation of the negentropy of the series $(u(t)$ which is very close to the optimal stopping time minimizing the distance $\|u(t) - \tilde{f}\|$.

We introduce the relative change of the negentropy by:

$$\frac{J(u(t+\Delta t)) - J(u(t))}{J(u(t+\Delta t)) \cdot \Delta t}\,, \quad t = 0, 1, 2... \tag{9}$$

We observe that the minimum of the distance, which determines the optimal stopping time, corresponds with a small value of the relative variation of the negentropy (9). The negentropy criterion to select the stopping time in the diffusion process is evaluated and validated experimentally since we have not an analytical demonstration of the existence of the minimum of the negentropy. Then, we define the stopping time T for the diffusion filtering as the time T that satisfies

$$\frac{J(u(T+\Delta t)) - J(u(T))}{J(u(T+\Delta t)) \cdot \Delta t} = \varepsilon \tag{10}$$

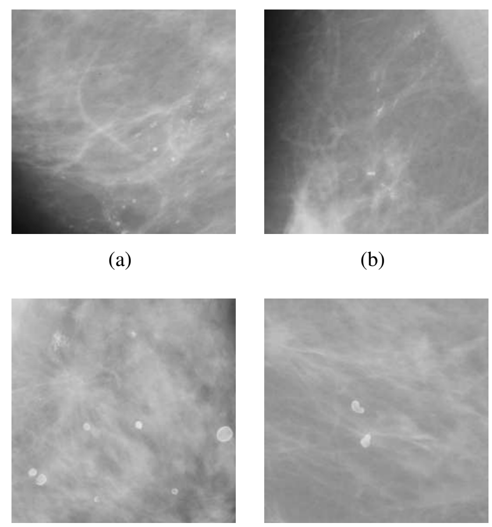
where ε is a constant whose value is chosen in such a way that the stopping time obtained will be as close as possible to the optimal stopping time. Its value is 0.005 and it is independent of the class of statistics of the mammographic image considered.

The main advantage of negentropy based method is that the diffusion stopping time obtained is estimated without any additional knowledge of the ideal free-noise image and using quite mild assumptions about noise.

III. Denoising experiments with mammographic images

Next, we test with mammographic images the denoising method described above. This method consists of applying to the image a diffusive filter a number of times determined by the beginning of an stable interval of the negentropy of filtered images $u(t)$.

The mammographic images considered in the experiments have been selected from the MIAS MiniMammographic Database, provided by the Mammographic Image Analysis Society [9]. The mammograms are digitised at 200 micron pixel edge, resulting that the images are 1024×1024 pixel resolution. The criteria for image selection was that the test set must be representative of the mammographic images, so we choose four images containing all kinds of typical mammogram regions as spiculated masses, distortion in breast architecture, asymmetries and clustered microcalcifications. (Fig. 1).



(a) (b) (c) (d)

Fig. 1. Test images used in the experiments from the MIAS Minimammographic Database.

To assess the applicability of the denoising method, we assume the test images to be noise-free and corrupt them by adding different levels of gaussian noise of zero mean. Thus, in these experimental cases the variance of the noise, which is needed as an input for the diffusion filter, is known. However, when the noise variance is unknown, a very robust estimation of this parameter may be performed by using Donoho's method [10] based on the wavelet transform of the image.

The estimation of the the standard deviation is given by

$$\sigma \approx \frac{\text{median}|Y_{ij}|}{0.6745}\,, \tag{11}$$

where Y_{ij} are the diagonal detail coefficients of the appropriate wavelet transformed.

Some statistical values of the original images such as the mean, the standard deviation and the kurtosis are given in

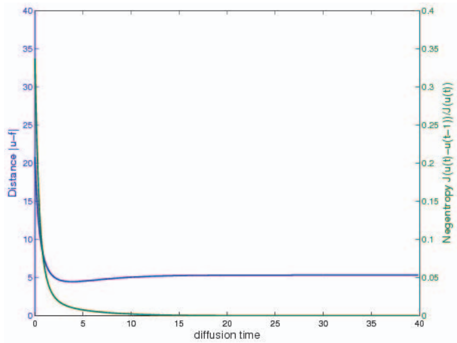


Fig. 2. The distance $\|u(t) - \tilde{f}\|$ and the relative negentropy evolution with the diffusion parameter for the mammographic image.

Table 1.

TABLE I
STATISTICAL VALUES FOR THE TEST IMAGES.

	Mean	Standard deviation	Kurtosis
Image (a)	150.06	30.84	4.639
Image (b)	132.82	35.08	0.961
Image (c)	165.25	17.04	2.135
Image (d)	160.59	10.93	0.259

Denoting the original image by $\tilde{f}$, for each time step of the filter, we compute the Euclidean distance

$$d(t) = \|u(t) - \tilde{f}\| \; . \tag{12}$$

The distance between the original noise-free image and the restored one is the root of the mean squared error (RMSE) between the two images. We use this RMSE to evaluate the final image quality obtained.

Let us denote by d_{OP} the optimum distance, *i.e.*, the minimum value of $d(t)$, and by d_{NE} the distance between the noise-free image $\tilde{f}$ and the restored image computed with the negentropy criterion.

The negentropy $J(u(t))$ is estimated using the approximations (6) described above. These approximations were obtained under the assumption that the probability density is not very far from Gaussian distribution, this holds in our case. For practical computation of the negentropy the signal $u(t)$ must be normalized making it zero-mean and of unit variance. In Fig. 2, the evolution of the distance $\|u(t) - \tilde{f}\|$ and the relative negentropy of $u(t)$ for a mammographic image is shown. We observe that the minimum of the distance, which determines the optimal stopping time, corresponds with a small value of the relative variation of the negentropy (9).

Each test image has been corrupted with four levels of gaussian noise of zero mean. The standard deviation of the added noise ranks from 0.1 to 0.4 times the standard deviation of the corresponding original image. In Tables 2-6, we show the results for the real noise variance added to the image, $\sigma^2 re$, and the estimated variance by the Donoho's formula, $\sigma^2 es$, and the distances (RMSE) between the original noise-free image and the restored one in three cases, namely, the optimal, the restored using the negentropy criterion and the restored using an adaptative Wiener filter.

TABLE II
STOPPING TIMES AND DISTANCES FOR THE TEST IMAGE (A).

f.noise	$\sigma^2 re$	$\sigma^2 es$	d_{NE}	d_{OP}	d_{WEI}
0.1	3.10	3.41	3.28	3.07	3.97
0.2	6.24	6.46	5.82	5.79	7.59
0.3	9.32	9.40	8.02	7.99	13.73
0.4	12.44	12.39	10.36	10.12	23.84

TABLE III
STOPPING TIMES AND DISTANCES FOR THE TEST IMAGE (B).

f.noise	$\sigma^2 re$	$\sigma^2 es$	d_{NE}	d_{OP}	d_{WEI}
0.1	3.50	3.72	3.23	3.23	4.01
0.2	7.02	7.14	5.93	5.69	8.60
0.3	10.54	10.90	8.64	7.65	17.01
0.4	13.98	14.01	10.62	9.67	31.53

TABLE IV
STOPPING TIMES AND DISTANCES FOR THE TEST IMAGE (C).

f.noise	$\sigma^2 re$	$\sigma^2 es$	d_{NE}	d_{OP}	d_{WEI}
0.1	1.71	2.18	2.01	1.67	3.23
0.2	3.41	3.68	3.95	3.49	4.32
0.3	5.09	5.29	5.71	5.04	6.24
0.4	6.82	6.97	6.96	6.46	8.88

TABLE V
STOPPING TIMES AND DISTANCES FOR THE TEST IMAGE (D).

f.noise	$\sigma^2 re$	$\sigma^2 es$	d_{NE}	d_{OP}	d_{WEI}
0.1	1.09	1.60	1.43	0.81	2.07
0.2	2.18	2.46	2.21	1.86	2.55
0.3	3.29	3.54	3.09	2.62	3.25
0.4	4.37	4.49	3.98	3.33	4.31

As derived from these tables and from the observation of the above-mentioned images, the use of the proposed method suggest that the negentropy criterion provides a good estimates of the stopping time in the diffusion process for moderate noise levels. In addition we observe that the distances (or RMSE) between the original noise-free image and the restored image with our method are very near to the optimal. The other standard technique Wiener filter provides bigger distances to the optimum for high levels of noise.

The values obtained for the estimation of noise added for the images are very similar for the real, so it suggest that the wavelet estimation could be useful to denoise real noisy mammographic images with unknown noise level.

We compare in Figure 3 the visual quality of noisy versions of the images at Figure 1 and the resulting denoised images obtained applying the diffusive filter with negentropy. We can see there that the final image obtained of the restored image is quite similar to that of the original noise-free image, so it is a good check for our develop method to denoise mammographic images.

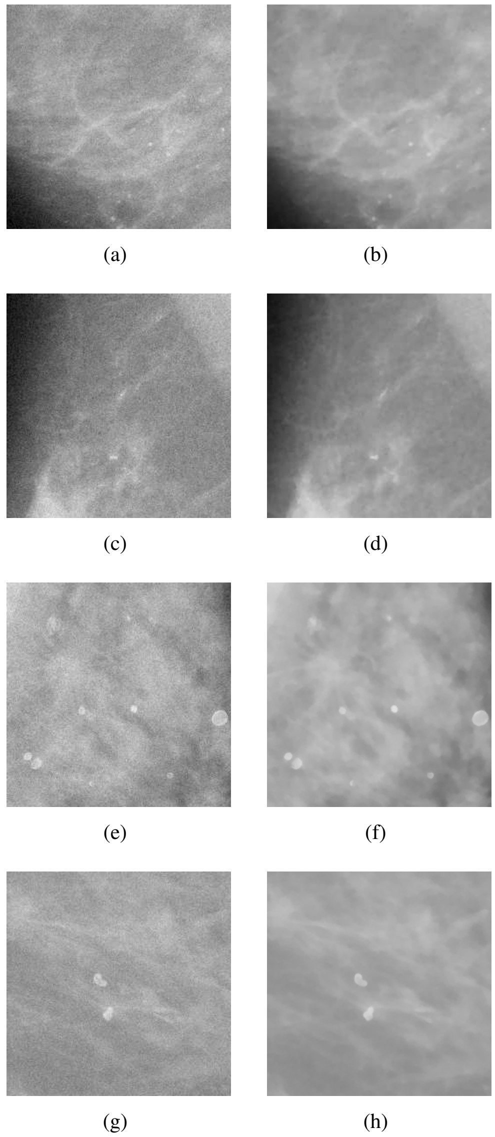


Fig. 3. Image restoration process. First column test images with noise added; second column filtered images stopping the diffusive process with the negentropy criterion.

IV. Discussion and conclusions

In this paper we have focused in the restoration of mammographic digital images by using a non-linear diffusive filter combined with the estimation of the stopping time at the beginning of stable intervals of the negentropy. The main conclusion is that the proposed method may be used with guarantee to reduce the noise of mammographies, since the distances between the original noise-free image and the restored image with our method are very near to the optimal.

The reduction of the noise in medical images, in particular in mammographies, is a very important stage of any computerized assisted diagnosis system. We have showed that our method works as well as other well established methods like Wiener classical method. In addition, the Donoho's wavelet variance estimation is very similar to the real variance of the noise added. This estimation would be very useful to denoise real noisy mammographies, although the real noise it is not known in these images, Donoho's estimation could approximate it in an objective way.

Acknowledgment

The authors would like to thank the Polytechnic University of Valencia project of title: Validacin de un sistema de diagnstico asistido por ordenador para mamografas. Aplicacin al programa de cncer de mama de la Comunidad Valenciana.

References

[1] A. Hyvrinen 1999 *Sparse code shrinkage: denoising of nongaussian data by maximum likelihood estimation* Neural Comput 11: 1739-1768
[2] J. Dengler, S. Behrens, J. F. Desaga 1993 *Segmentation of microcalcifications in mammograms* IEEE Trans Med Image 12: 634-642
[3] F. Catt, P. Lions, J. Morel, T. Coll 1992 *Image selective smoothing and edge detection by nonlinear diffusion* SIAM Num Anal 29: 182-193,
[4] J. Weickert, B.M. ter Haar Romeny, M.A. Viergever 1998 *Efficient and reliable schemes for nonlinear diffusion filtering* IEEE Trans Image Processing 7-3: 398-410
[5] A. Marquina, S. Osher 2000 *Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal* SIAM J Sci Comput 22-2: 387-405
[6] J. Weickert 2001 *Efficient image segmentation using partial differential equations and morphology* Pattern Recognition 34-9: 1813-1824
[7] A. Hyvrinen and E. Oja 2000 *Independent Component Analysis:algorithms and applications*Neural Networks 13: 411-430
[8] A. Hyvrinen 1998 *New approximations of differential entropy for independent component analysis and projection pursuit* Adv in Neural Information Processing Systems, 10: 273-279
[9] The Mammographic Image Analysis Society. The MiniMammography Database can be found at address: `www.wiau.man.ac.uk/services/MIAS/MIASmini.html`.
[10] D. Donoho, I. Johnstone, G. Kerkyacharian, D. Picard 1995 *Wavelet shrinkage: asymptopia?* Journal of the Royal Statistical Society 57: 301-369