Nonlinear Unsharp Masking for the Enhancement of Document Images

Stefano Chiandussi, Giovanni Ramponi
D.E.E.I., University of Trieste
via A. Valerio, 10, 34127 Trieste, Italy
Tel: +39 40 6767147; fax: +39 40 6763460
e-mail: ramponi@imagets.univ.trieste.it

ABSTRACT

A novel operator for the enhancement of the quality of document images is presented in this paper. This operator, which is a quadratic one, is based on the Unsharp Masking (UM) technique, but it is able to limit noise amplification because every pixel of the processed image depends upon a large portion of the input image; in the same time a good response on details is obtained. A formal description of the operator's response to noise is also presented.

1 INTRODUCTION

Many types of documents, ranging from terrain and urban maps to technical drawings or printed and handwritten text, can be digitally acquired by some electronic equipment and converted to *document images*, i.e. gray–level images the contents of which are completely characterized by a binary (black/white) field. Once acquired as a gray–level image and binarized, a document image can be further processed for tasks such as automatic character recognition. Document images vastly differ in quality: consequently, their acquisition may produce gray–level images which are difficult to effectively binarize. The performance of a character recognition algorithm or the readability of a binarized document by a human user are strongly influenced by the quality of the binarization process. In this paper, we present a technique that performs the enhancement and binarization of a digitized gray-scale document image. Experimental results demonstrating the quality of the binarized output images are also presented.

2 DESCRIPTION OF THE OPERATOR

Figure 1 displays the block diagram of the proposed gray–level enhancement technique. This method is based on a nonlinear version of the Unsharp Masking (UM) method. In conventional UM methods, a highpass-filtered version of the input image is added to the image itself so as to improve the contrast in the detail areas such as character edges. In a general image processing environment, the performance of the conventional linear UM technique is limited by the noise amplification caused by the highpass filter. Different linear and nonlinear approaches have been proposed in the past to reduce the sensitivity of UM techniques; in our work, we employ an operator which belongs to the class of the polynomial filters [1]. The structure of the UM method has been maintained, but: (i) a Laplacian–of–Gaussian (LoG) bandpass filter substitutes for the highpass filter in the lower branch, as suggested e.g. in [2]; (ii) most importantly, the signal which feeds the LoG filter is processed by a quadratic operator which introduces a significant noise smoothing in background areas. The overall structure of the algorithm corresponds to that of a homogeneous quadratic filter. Recent works have shown that quadratic filters are capable of edge enhancement in images, with limited noise amplification [3].

The lower branch of the block diagram contains the main computational operators of the algorithm. The input image is first scaled according to

$$u(m,n) = \frac{x(m,n) - a}{b - a} \;, \qquad (1)$$

where a is the average value of the luminance in the background of the processed image and b is the average luminance of the text/drawing. However the choice of these two parameters is not critical. After scaling, the background areas of the input image approximate a zero–mean signal, while the characters have mean one.

The scaled data are fed to a Gaussian filter, which may be characterized as a local mean value estimator. Its output, $v(m,n)$, is multiplied by the value of the present input sample $u(m,n)$. The resulting data are processed by a LoG filter, which amplifies their medium–frequency component. A fraction of this signal is added to the original image.

When compared to the conventional UM algorithm, the scheme described above exhibits small noise sensitivity due to the bandpass response of the LoG filter and to the local mean estimate which makes the operator capable of reducing the variance of the background areas.

The heart of the processing is formed by the product of the two signals $u(m,n)$ and $v(m,n)$: in this way,

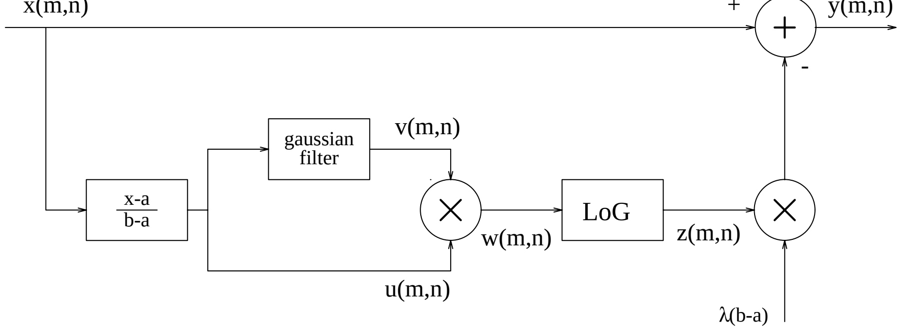

Figure 1: Block diagram of the preprocessing technique.

a correlation is evaluated between the reference pixel $u(m,n)$ and the elements in its neighborhood; each contribution is weighted with a coefficient which is inversely proportional to the distance of the pixel from the center of the processing mask. As it will be demonstrated in the next sections, this procedure has two effects: in the background areas, where each deviation from zero may be considered as due to noise, the distribution becomes more concentrated, hence reducing the noise variance. To verify this, we observe that a noisy background with Gaussian distribution of mean a and variance $(a-b)^2\sigma^2$ implies that $v(m,n)$ is Gaussian too, because $v(m,n)$ is a linear combination of the input samples, even if it is no longer an uncorrelated noise. Under these conditions about noise, it can be easily demonstrated that, in a background area:

$$Prob(|w = u\,v| \leq \alpha) > Prob(|u| \leq \alpha), \tag{2}$$

for $\alpha \in [0,1]$. We must remind indeed that after scaling the background has mean zero and the characters have mean one. Inequality (2) indicates that $w(m,n)$ is more concentrated around zero than $u(m,n)$ is.

In detail areas, on the contrary, the correlation is higher: if the reference pixel is located on the edge of a character or inside a relatively small character, the signal which is input to the LoG filter is much larger. Finally, it should be observed that the inner part of large characters is left almost unchanged by the operator: locally, the signal has mean one (which is blocked by the LoG filter) and a small high–frequency content (which is further attenuated by the Gaussian filter).

3 NOISE AMPLIFICATION

In this section we analyze the noise amplification of the filter. To this purpose, we evaluate the variance of the output, under the hypothesis that the input is a white ergodic noise (noisy background with Gaussian distribution of mean a and variance $(a-b)^2\sigma^2$: this hypothesis implies that $\{u(m,n)\}$ has zero mean and σ^2 variance). First of all, we have to calculate the average of the output:

$$E[y] = E[x - \beta z] = E[x] - \beta E[z] \tag{3}$$

where $\beta \doteq \lambda(b-a)$.

Because of the presence of the LoG (linear) filter, $E[z] = 0$, so we can conclude that the output mean coincides with the input one.

Now we can calculate the output noise variance:

$$\begin{aligned} var[y] &= E[(x-\beta z)^2] - E^2[x - \beta z] = \\ &= var[x] - 2\beta E[xz] + \beta^2 E[z^2], \end{aligned} \tag{4}$$

where:

$$\begin{aligned} E[xz] &= E[((b-a)u + a)z] = \\ &= (b-a)E[uz] + E[az] = \\ &= (b-a)E[uz]. \end{aligned} \tag{5}$$

It is easy to calculate the expression $E[uz]$ because it is a polynomial of the third degree in the (scaled) input samples. For the assumptions made, we know that:

$$E[u_{ij}^3] = 0, \quad E[u_{ij}^2 u_{kl}] = 0, \quad E[u_{ij} u_{kl} u_{mn}] = 0 \tag{6}$$

for any couple of indexes (i,j), (k,l) and (m,n). So, every term needed for the calculation of (5) is the product of three not always different samples: we are in the case of condition (6). For this reason, we can conclude that $E[xz] = 0$.

To complete the evaluation of expression (4), we need $E[z^2]$.

$$\begin{aligned} E[z^2] &= E\left[\left(\sum_{i,j} h(i,j) w(m-i, n-j)\right)^2\right] = \\ &= \sum_{i_1,j_1} \sum_{i_2,j_2} h(i_1,j_1) h(i_2,j_2) \cdot \\ &\quad \cdot E[w(m-i_1, n-j_1) w(m-i_2, n-j_2)]. \end{aligned} \tag{7}$$

where $h(i,j)$ is the impulse response of the LoG. Writing in full $E[w(m-i_1, n-j_1)w(m-i_2, n-j_2)]$ (as the output of the Gaussian filter), it becomes clear that we need the expression of $E[u(m,n)u(\mu,\nu)u(m-i_1, n-j_1)u(\mu-i_2, \nu-j_2)]$. The last one is different from zero only if:

$$(i)\; i_1 = i_2 = j_1 = j_2 = 0 \qquad (ii) \begin{cases} m - i_1 = \mu \\ n - j_1 = \nu \\ \mu - i_2 = m \\ \nu - j_2 = n \end{cases} \tag{8}$$

Thanks to these considerations, we can write:

$$E[w(m-i_1,n-j_1)w(m-i_2,n-j_2)] = \\ = \sigma^4 \left[g^2 + g^2(i_1-i_2,j_1-j_2)\right] . \tag{9}$$

where $g(i,j)$ is the impulse response of the Gaussian filter and $g \equiv g(0,0)$. Because of conditions (8), $E[w(m,n)w(m-i,n-j)]$ is zero only if the distance between the couple of pixels is too high; the image border, of course, is settled by the extension of the Gaussian filter window and it must be taken into account.

Now we have calculated every component of expression (4), so that it is possible to write the variance of the output noise; substituting (9) in (7) and (7) in (4) we obtain:

$$\begin{aligned} var[y] &= var[x] + \beta^2 \sum_{i_1,j_1}\sum_{i_2,j_2} h(i_1,j_1)h(i_2,j_2)\cdot \\ &\quad \cdot\sigma^4[g^2 + g^2(i_1-i_2,j_1-j_2)] = \\ &= var[x](1+\lambda^2\sigma^2 \sum_{i_1,j_1}\sum_{i_2,j_2} h(i_1,j_1)\cdot \\ &\quad \cdot h(i_2,j_2)[g^2+g^2(i_1-i_2,j_1-j_2)]. \end{aligned} \tag{10}$$

Among LoG properties, there is:

$$\sum_{i,j} h(i,j) = 0. \tag{11}$$

Substituting (11) in (10) we have:

$$\begin{aligned} var[y] &= var[x](1+\lambda^2\sigma^2 \sum_{i_1,j_1}\sum_{i_2,j_2} h(i_1,j_1)\cdot \\ &\quad \cdot h(i_2,j_2)g^2(i_1-i_2,j_1-j_2). \end{aligned} \tag{12}$$

which is the solution of our problem.

From this formula it results that the output variance is a function of the square of the input variance. On the other side, this effect is very limited due to the small values of the $g(i,j)$ and $h(i,j)$ coefficients and of the parameter λ. Experimental results have also proved that the proposed filter has the lowest noise amplification among general purpose filter such as [1], [2], [3] or dedicated filters as [4].

4 PARAMETERS DETERMINATION

In the description of the operator, we mentioned the problem of finding the a and b parameters. Of course, they can be set manually, but there is another possibility too. First of all, we obtain a binarized version of the original by manually or automatically choosing the threshold, so that we have a rough information about which pixels belong to the background (white pixels in the binarized image) and which belong to the characters (black pixel). Then we divide the input image in blocks (for ease we used non-overlapping ones), for every block we estimate the mean of the pixels which belong to the background or to the characters, obtaining a good evaluation of a and b for that block. Now it is possible to calculate (1) for every block, and process the obtained scaled image. Once we have the enhanced image, it is convenient to binarize the image obtained and apply again the described procedure: this time we have a better information about white and black pixels and, of course, a better result.

5 EXPERIMENTAL RESULTS

The results of processing a mail address test image is shown in the next page. Fig.2-a shows the original data; in Fig.2-c we can see the processed image, using the manually and globally set a and b values. In Fig.2-e we can see the result of the described filter with the two-step parameters setting. In the second column (Fig.2-b-d-f) we find the binarized versions of the corresponding images in the first one. A simple fixed threshold binarization has been used.

It should be observed that a significant improvement has been obtained (Fig.2-c) both in the dark area at the top of the image and in the readability of the printed address. Noise amplification on background is visible especially between "Louisiana Department" and "LA DEPT", and on the dark background: this produces some undesired black pixels in the binarized image.

In Fig.2-e, thanks to the local detection of parameters there is a significantly lower noise in the backgrounds, especially in the dark one (in Fig.2-c the parameters are tuned for a good result on the bright one): this improvement removes almost all the black spots in the binarized image. Moreover at the same time the characters are a bit more connected.

References

[1] G. Ramponi, "A Simple Cubic Operator for Sharpening an Image," *Proc. 1995 IEEE Workshop on Nonlinear Signal and Image Processing*, June 20-22, 1995, Neos Marmaras, Greece.

[2] F. Neycenssac, "Contrast Enhancement Using the Laplacian of a Gaussian Filter," *CVGIP – Graphical Models and Image Processing*, Vol. 55, No. 6, Nov. 1993, pp. 447-463.

[3] S.K. Mitra, H. Li, I. Li and T-H. Yu, "A New Class of Nonlinear Filters for Image Enhancement," *Proc. ICASSP '91*, pp. 2525-2528, 1991.

[4] Giovanni Ramponi and Paolo Fontanot, "Enhancing document images with a quadratic filter", *Signal Processing*, 33 (1993), pp 23-34.

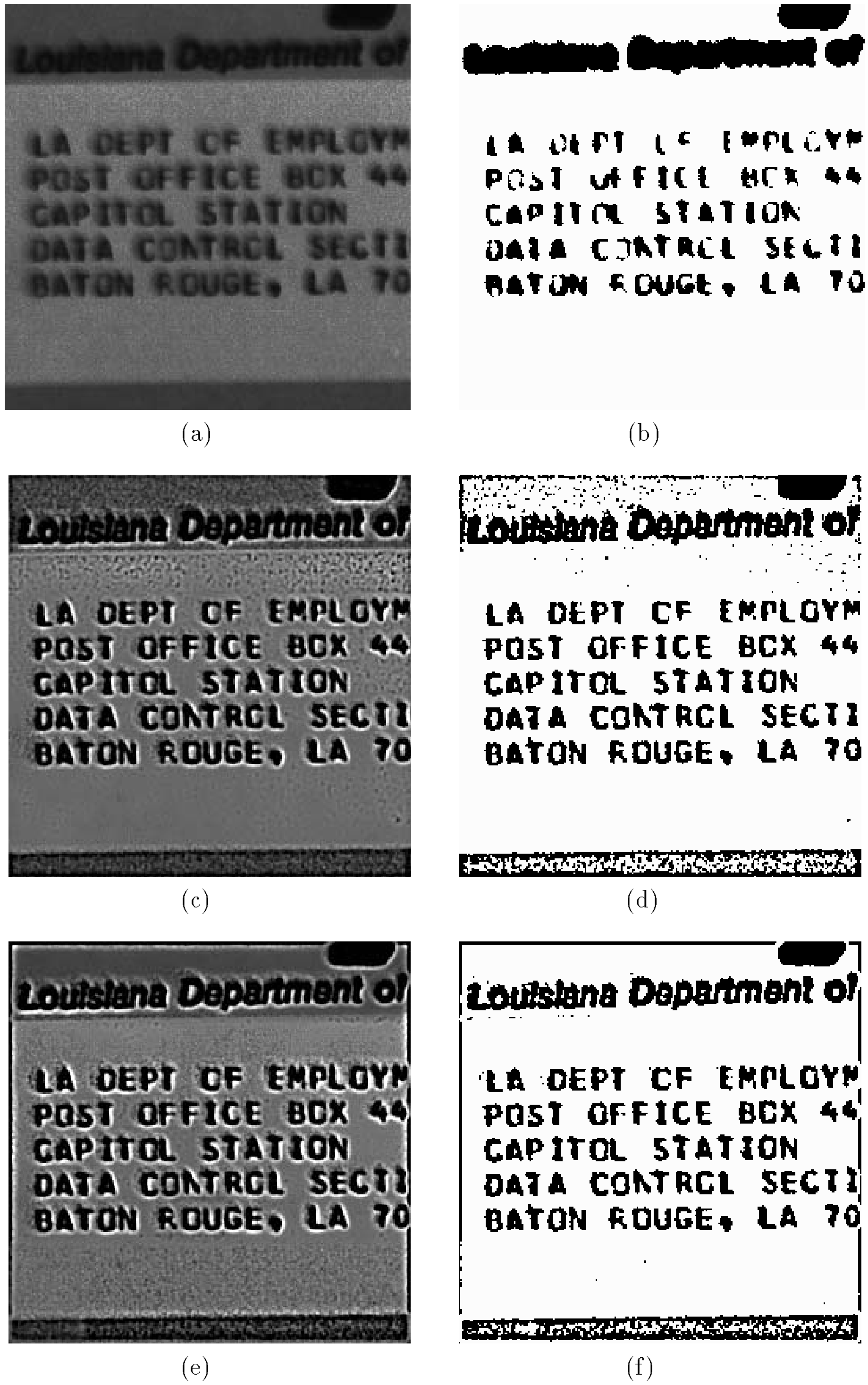

Figure 2: Original (a) and processed images (see text).