Enhancement of scanned documents in Besov spaces using wavelet domain representations

Kathrin Berkner[1]
Ricoh Innovations, Inc., 2882 Sand Hill Road, Suite 115, Menlo Park, CA 94025

ABSTRACT

After scanning, a document is typically blurred, and some noise is introduced. Therefore, the enhancement process of a scanned document requires a denoising and deblurring step. Typically, these steps are performed using techniques originated in the Fourier-domain. It has been shown in many image processing applications such as compression and denoising that wavelet-domain processing outperforms Fourier-domain processing. One main reason for the success of wavelets is that wavelets adapt automatically to smooth and non-smooth parts in an image due to the link between wavelets and sophisticated smoothness spaces, the Besov spaces. Recently smoothing and sharpening of an image – interpreted as an increasing and decreasing of smoothness of an image – has been derived using Besov space properties. The goal of this paper is to use wavelet-based denoising and sharpening in Besov spaces in combination with characterization of lines and halftone patterns in the wavelet domain to build a complete wavelet-based enhancement system. It is shown that characteristics of a scanned document and the enhancement steps necessary for a digital copier application are well-suited to be modeled in terms of wavelet bases and Besov spaces. The modeling results leads to a very simple algorithmic implementation of a technique that qualitatively outperforms traditional Fourier-based techniques.

Keywords: Wavelets, wavelet transform, Besov spaces, deblurring, scanning, halftone, moiré.

1. INTRODUCTION

Digital imaging has become a common part of day-to-day life as consumers have easier access to digital cameras, scanners, printers, etc. Especially in home or industrial offices working without printers and digital copiers is hard to imagine. A digital copier consist in part of a scanning unit, an image processing unit and a printing unit and is mainly used for copying documents. Without processing a document between the input and output unit the visual quality of the copied document is greatly reduced. Scanning introduces a blurring of the original document as well as some scanner noise. The preparation for printing has to consider toner characteristics and eliminate possible frequency shifts (moiré artifacts) due to different scanner and printer resolution as well as frequencies contained in the document, e.g. in halftone patterns.

Typical image processing steps to take care of all those issues are deblurring and denoising. Deblurring will create sharp text, whereas denoising will reduce the scanner noise and suppress some halftone frequencies in order to avoid moiré artifacts in the printed image. These image processing steps are typically performed via linear filtering of the digitized document, i.e via multiplication in the Fourier-domain. Common techniques include unsharp masking for sharpening of text and image contours, bandpass filtering for removal of halftone frequencies and lowpass or Wiener filtering for removal of scanner noise. Those techniques have a well-developed theoretical background and a long history of use in applications[1,10]. However, an important disadvantage of those techniques is that the filters are designed in the Fourier-domain and Fourier coefficients capture frequency information of the entire image, not local frequency information. Another more theoretical disadvantage of Fourier-domain processing is that approximation errors can be measures only in the function space L^2 which does not incorporate any information of the smoothness of the image (differentiability). That means in the L^2 space it is not possible to capture or measure the difference between a sharp and a smooth edge.

During the last decade the use of wavelet bases functions have become very popular in signal and image processing. Wavelets overcome the two disadvantages of Fourier-based methods mentioned above. First, wavelet coefficients contain

---

1) Email: berkner@rii.ricoh.com

not global, but local frequency information of the image. Secondly, wavelets form bases in a variety of function spaces that take into account the smoothness of a function. Those function spaces, called *Besov spaces*, can be interpreted as generalized L^p spaces that also consider the differentiability of a function. In the following the main facts about wavelets and wavelet transforms are summarized briefly.

Wavelet coefficients are computed similar to Fourier coefficients from inner products of a function with basis functions – the wavelets. Typically, a wavelet system used in the dyadic wavelet transform consists of two functions, the scaling function φ and the wavelet function ψ. Those two functions are linked to each other via specific scaling equations. The representation of a function *f* in a wavelet basis is given by

$$f = \sum_k c_{Jk}\varphi_{Jk} + \sum_{j \geq J}\sum_k d_{jk}\psi_{jk} , \tag{1}$$

where φ_{jk} and ψ_{jk} are dilates and translates of φ and ψ, namely $\varphi_{jk}(t) = 2^{j/2}\varphi(2^j t\text{-}k)$ and $\psi_{jk}(t) = 2^{j/2}\psi(2^j t\text{-}k)$. The *scaling coefficients* c_{Jk} are computed as $c_{Jk} = \langle f,\varphi_{Jk}\rangle$ and the *wavelet coefficients* d_{jk} as $d_{jk} = \langle f,\psi_{jk}\rangle$. The scaling coefficients can be interpreted as representing a low resolution version of the image, whereas the wavelet coefficients capture the detail information at various resolution levels j. The functions φ_{Jk} and ψ_{jk} $(j \geq J)$ form an orthonormal bases in the L^2 space. The wavelets used in this paper belong to the family of orthogonal Daubechies wavelets. Those wavelets have compact support, have D vanishing moments and a Hoelder regularity σ.[5]

If a function or signal is given in the form of its wavelet coefficients, modifying individual coefficients only affects a region of the entire image. This characteristic in combination with the multiresolution nature of the decomposition has lead to enormous success in various fields of image processing over the last years. Especially in denoising and compression wavelet-based techniques have become state-of-the-art and outperform classical Fourier-based techniques in most applications.[4,7] Theoretical results involving modeling of functions in Besov spaces explain the performance of wavelet-based techniques.[6]

Naturally the question arises whether other image processing steps such as deblurring could benefit from the use of wavelet bases. Several methods are known that combine a classical Fourier-based deblurring step with wavelet-based denoising [11]. Those methods require a switching of image processing domains. It would be of greater benefit if the entire processing for deblurring and denoising could be performed in one wavelet domain. A theoretical definition of sharpening and smoothing in Besov spaces using wavelet-based representations has been derived and compared to classical Fourier-based and the hybrid Fourier-wavelet-based methods.[2] Based on those results it is very easy to combine deblurring with denoising in one wavelet domain. Since an enhancement system for a scanned document has to apply different degrees denoising and deblurring to different regions of the document, the locality of wavelet coefficients provides another advantage.

In this paper the use of wavelet-domain representations for digital copier enhancement processing is studied. Four different types of images most commonly contained in compound documents are targeted: text, halftone patterns, background and photographic images. After introducing Besov spaces and their relations with wavelet bases, their use in denoising and deblurring applications is explained. The important step for the digital copier application is the modeling of original and scanned documents as well as the enhancement processing steps in terms of wavelet bases and Besov spaces. A theoretically motivated model is presented in this paper. Suppression of halftone frequencies is shown to benefit greatly from the ability to modify individual localized coefficients after a characterization of halftone-versus-line patterns is performed in the wavelet domain. Experimental results are presented along with conclusions at the end of this paper.

2. ENHANCEMENT OF SCANNED COMPOUND DOCUMENTS

The enhancement of a scanned document combines three image processing tasks: deblurring, denoising, and moiré reduction. The most obvious one is perhaps deblurring since the blurring introduced by the scanner is easily recognizable, especially around text characters.

Scanner blurring is typically modeled as a convolution with a point-spread function which is equivalent to a multiplication in the Fourier-domain. The theoretically exact way to invert the blurring is to divide the blurring filter in the Fourier-domain of the scanned document. However, it is well known that this theoretical approach is not applicable most of the time due to several reasons. Noise is typically introduced during the scanning process which makes inversion of the blurring by division in the Fourier-domain an ill-posed problem. The other reason is that if the blurring filter has zeros in its frequency response, an inversion by division does not exist. In both cases, typically, a regularized inverse filtering is applied. This filter contains a regularization parameter that controls the trade-off between noise suppression and filter inversion.[1,10] Another point that makes filter inversion in the Fourier-domain problematic is that in many applications the point-spread function is not known and has to be estimated. This introduces another source for error in the enhancement process.

In studying the inversion problem it becomes clear that a denoising step has to be coupled with the deblurring step. If no denoising is being performed scanner noise pixels may be enlarged and dominate the overall image appearance. Another category of noise is found in halftone areas. Common halftoning techniques are error diffusion and ordered dither. Whereas error diffusion introduces some randomness into the placement of halftone dots, ordered dither shows periodically, isolated dots. If those dots are enhanced by a deblurring filtering procedure, the enhancement can lead to severe moiré artifacts in the enhanced image. Those might become even worse after printing the enhanced page on a laser printer with a specific resolution (dpi) that interferes with the enhanced halftone frequency.

Therefore, a desired enhancement technology should be able to remove scanner and halftone noise while deblurring other areas such as text and photographic content. One approach would be to segment a document first and then apply different filters to the different segments. However, a segmentation technique will have difficulty performing a detailed segmentation in areas where text or image contours overlay noisy background or halftone pattern. In those areas modifications of pixels are required on a very fine scale. Another problem arises in the combination of different filtering results along the boundaries between segments.

3. WAVELET-BASED DENOISING, SHARPENING AND SMOOTHING IN BESOV SPACES

3.1 Wavelets and Besov spaces

One of the most important properties of orthogonal wavelet systems is that they form an unconditional basis in a large class of smoothness spaces, the *Besov* spaces. Besov spaces collect functions that have a specific degree of smoothness in their derivatives. The *Besov space* $B^{\alpha}_{q}(L^p)$ consists, roughly speaking, of functions that have α bounded ‘‘derivatives’’ in L^p. The parameter q is an additional refinement parameter. The exact definition of Besov spaces in the spatial domain is very complex.[8] In general it can be said that the smoothness of a function in $B^{\alpha}_{q}(L^p)$ increases with increasing α. Given the wavelet decomposition of a function *f* in scaling and wavelet coefficients computed using an orthogonal wavelet system (φ, ψ) of order D as in Eq. (1) there exists a wavelet-based characterization of the Besov norm of *f*, namely

$$\|f\|_{b^{\alpha}_{q}(L^p)} = \|c_J\|_{l^p} + \left(\sum_{j \ge J} (2^{j(\alpha p + p/2 + 1)} \|d_j\|_{l^p})^q\right)^{1/q} . \qquad (2)$$

This sequence norm $b^{\alpha}_{q}(L^p)$ is equivalent to the norm $B^{\alpha}_{q}(L^p)$ which is defined directly on the function *f.* A function belongs to the Besov space $B^{\alpha}_{q}(L^p)$ if the Besov norm is finite.

A closer look at the norm definition in Eq. (2) shows that the Besov norm can be interpreted as an L^p norm with special weighting of the detail coefficients. In contrast to the complicated and impractical formulation of a Besov norm in the spatial domain the wavelet representation of a function suddenly provides an easy way to compute the norm. The norm in Eq.(2) is only defined for a parameter α in the limited range $0 < \alpha < \gamma$, where γ reflects the regularity of the orthogonal wavelet system. If σ is the maximal Hoelder smoothness of ψ and D the order of the wavelet system, then the regularity γ of the wavelet system is defined as

$$\gamma = \min(\sigma, D). \tag{3}$$

In order to understand how Besov spaces compare to more familiar function spaces we give two examples of relations to other function spaces.

- Hoelder spaces C^α are equivalent to Besov spaces $B^\alpha_q(L^p)$ for $q = p = \infty$.
- A Besov space $B^\alpha_q(L^p)$ is embedded between the two Sobolev spaces $W^{\alpha+\varepsilon}_p$ and are $W^{\alpha-\varepsilon}_p$.

In general, the parameter α can be interpreted as characterizing the smoothness of a function. The larger α the smoother the function. A smooth function has typically a sharper decay of wavelet coefficients at larger scales then a function of little smoothness. Therefore, coefficients of a smooth function can be weighted at each scale by the weights $w_j = 2^{j(\alpha p + p/2 + 1)}$ and still maintain the finiteness of the sum in Eq. (1). If coefficients of a non-smooth function would be weighted by those weights, the sum would diverge.

Examples for the association of familiar functions to Besov spaces are the following.

- A step-edge function is in $B^\alpha_1(L^1)$ for all $\alpha < 1$ and in $B^\alpha_2(L^2)$ for all $\alpha < 0.5$, but not in $B^\alpha_\infty(L^\infty)$ for any $\alpha > 0$.
- The Gaussian kernel is in $B^\alpha_\infty(L^\infty)$ for all $\alpha > 0$.
- The hat function ($f(x) = |x|$) is in $B^\alpha_\infty(L^\infty)$ for all $\alpha < 1$.

The important point to notice is that a step edge still has some regularity in terms of $\alpha > 0$ measured in the spaces $B^\alpha_1(L^1)$ and $B^\alpha_2(L^2)$, compared to no measurable regularity in $B^\alpha_\infty(L^\infty)$. This facts leads to interesting results in the theory of nonlinear approximation of functions with applications to image compression.[7]

3.2 Denoising

One of the first applications taking advantage of Besov spaces was denoising studied by Donoho and Johnstone[6] and is known as *wavelet-shrinkage*. The two most commonly known shrinkage techniques are *hard-thresholding* and *soft-thresholding*. For hard-thresholding, all coefficients with magnitudes below a given threshold are set to zero. For soft-thresholding, in addition, all coefficients with magnitudes above the threshold are shrunk by the thresholding amount. If the noise is modeled as Gaussian white noise, the threshold is derived only from the number of samples and the noise variance and is independent of the function and the wavelet system.

One of the most powerful theoretical properties of wavelet-shrinkage is the smoothness of the denoised function. In summary: Start with noisy samples of an unknown function in a specific Besov space and perform soft thresholding on wavelet coefficients with the specifically derived threshold. Then there exists an interpolation of the reconstructed samples that is almost surely at least as smooth as the original unknown function its Besov space with decreased Besov norm. Since, in general, the reconstructed function is always in the same Besov space as the original unknown function Donoho et al. talk about "Adaption to unknown smoothness via wavelet shrinkage."[6] It is not possible to obtain a similar result using Fourier transform coefficients.

A big advantage of wavelet shrinkage in applications is its simplicity. The theoretically derived threshold for hard- or soft-thresholding is computed usually from one level of wavelet coefficients at fine scales and then applied to all coefficients.

3.3 Sharpening and smoothing

As mentioned above, the smoothness of functions in Besov spaces is mainly determined by the magnitude of the exponent α – increasing α means increasing smoothness. From the definition of the Besov norm in Eq. (2) it becomes clear that a change of the Besov space association of a function can be achieved by rescaling the wavelet coefficients in a special fashion.[2]

In detail, a function g is called a *smoothed version* of the function $f \in B^\alpha_q(L^p)$ if the wavelet coefficients $d_{j,k}[g]$ of g are rescaled versions of the wavelet coefficients $d_{j,k}[f]$ of f, namely

$$d_{j,k}[g] = 2^{-j\tau} d_{j,k}[f] \text{ for } \tau > 0. \quad (4)$$

This rescaling shifts the function *f* from the Besov space $B^{\alpha}_{q}(L^p)$ into the space $B^{\alpha+\tau}_{q}(L^p)$ as long as $\alpha+\tau < \gamma$, where γ is the regularity of the wavelet system. In contrast *g* is a *sharpened version* of *f* if

$$d_{j,k}[g] = 2^{-j\tau} d_{j,k}[f] \text{ for } \tau < 0, \text{ as long as } \alpha+\tau > 0. \quad (5)$$

As a consequence, using Besov space properties of wavelet systems it has become possible to unify the two image processing steps smoothing and sharpening as being a switch between smoothness spaces.[2]

This definition of sharpening or smoothing is very different from the classical approach using Fourier-domain representations. In the classical approach smoothing of function is being considered as a convolution with a smoothing kernel, e.g. as low-pass filtering, where smoothing means a suppression of high-frequencies. The opposite operation, sharpening, is typically considered to be high-pass filtering, i.e. a suppression of low-frequencies and enhancement of high-frequencies.

Sharpening and smoothing in Besov spaces influences the differentiability of a function by changing the scaling behavior of wavelet coefficients in some frequency bands, namely those that correspond to frequencies up to level J of the decomposition. The lowpass component at level J is neither enhanced nor suppressed. A deblurring of a function via sharpening in Besov spaces can be interpreted as inversion of the smoothness introduced by the blurring by rescaling the wavelet coefficients as in Eq. (5) using an exponent $\tau < 0$. If the original blurring is modeled by a smoothing in Besov spaces as described in Eq.(4) the smoothing is completely invertible. If the original blurring has been modeled by a convolution in the Fourier-domain, the sharpening in the wavelet domain may theoretically not fully invert the blurring in the case that the blurring kernel is invertible. However, in most experiments, the wavelet-domain deblurring of a Fourier-domain blurred signal is enough to obtain an visual acceptable deblurring result. Besov space deblurring of a convolution-blur has the advantage, that in case the convolution kernel is not invertible or in the presence of noise, the Besov space deblurring still inverts the smoothness of the function and is not at all effected by the ill-posed problem in the Fourier-domain. Another advantage of Besov space deblurring is for the case that the exact shape of the blurring kernel is not known, but its regularity can be estimated. In this case the regularity estimate can be used in the Besov space deblurring and no information on the kernel shape is necessary.[2]

4. MODELING OF DOCUMENT ENHANCEMENT IN BESOV SPACES

4.1 Modeling of documents in Besov spaces

The goal of this paper is to use thresholding of coefficients for denoising and rescaling of coefficients for deblurring of a scanned document. For achieving this goal an estimation of the smoothness of a document before and after the blurring procedure in terms of Besov space smoothness is required. A document typically contains white background, text, and photographic as well as halftoned image areas. The objects of minimal smoothness in such a document are step edges in text and dots in halftone areas. Using the results from the previous section, the original document can be considered to be in $B^{\alpha}_{1}(L^1)$ for all $\alpha < 1$ or in $B^{\alpha}_{2}(L^2)$ for all $\alpha < 0.5$ or in $B^{0}_{\infty}(L^{\infty})$. Depending on which space is chosen, the generalized energy of the document is given by the corresponding Besov norm.

Background or a photographic image can be modeled as a very smooth function at fine scales. The energy will mostly be contained in the scaling coefficients. Those will neither be thresholded nor rescaled.

During scanning blurring is introduced and scanner noise added. The smoothness degree of the blurring can be estimated as being at least of degree 1, i. e. the blurred document is at least one times differentiable. The noise can be modeled as independent additive Gaussian noise of a given standard deviation ε. The parameter ε can be estimated from wavelet coefficients at a high resolution level.

4.2 Modeling of enhancement processing in Besov spaces

In order to perform wavelet-based enhancement of a document the first step is to choose a wavelet system, transformation and maximal level of decomposition.

It has been shown that the translation invariant (redundant) transform using the Haar wavelet system (filter coefficients [1 1], [1 -1]) performs well in denoising of images with step edges.[4] In addition, the regularity of the overcomplete Haar system is $\gamma = 1$, i.e. it covers the range of regularity that we expect to be introduced by the blurring, namely differentiability of order 1. The approximation of an image using the overcomplete Haar systems uses piecewise constant and piecewise linear functions as basis functions.[3] Those basis functions match the characteristics of a document. It is important to notice that those properties are only valid for the redundant Haar system, not for the maximal decimated Haar system.

The maximal level of decomposition should be chosen with respect to providing enough levels to sharpen the text to the desired degree and to remove or suppress halftone noise. The choice depends on the dpi of the scanner and the dpi of the printer. In most experiments with 300 and 600 dpi resolution scanners and printers, a maximal decomposition level of two or three was sufficient. If the scanner resolution is much higher more levels would be required.

The actual processing is performed by soft or hard-thresholding of coefficients and by multiplying coefficients at each scale j with a factor $2^{-j\tau}$. In order to preserve the overall energy of the image a renormalization factor R has to be applied to all coefficients. This factor R is chosen such that the Besov norms before and after rescaling are maintained.[2] The following Table 1 gives a summary of the theoretical modeling of document enhancement in the wavelet domain and its practical implications. Even though this modeling seems to be very theoretical in the first place it turns out to lead to very practical parameter choices for an enhancement algorithm.

Table 1 — Summary of modeling of document enhancement in the wavelet domain.

Theoretical	Practical
Choose image Besov space, e.g $B^{\alpha}_1(L^1)$	Decide on norm for "energy" preservation, e.g $\|\ldots\|_{b_1^{0.5}(L^1)}$
Estimate degree of blurring β	For scanners, set $\beta = 1$
Compute forward wavelet transform	Compute overcomplete Haar transform (only adds and subtracts)
Determine Hoelder regularity γ of wavelet system	For overcomplete Haar set $\gamma = 1$
Perform denoising by wavelet shrinkage	Hard- or soft-thresholding of coefficients
Perform deblurring by rescaling of coefficients $d_{jk} \rightarrow R\, 2^{-j\tau} d_{jk}$, $-\tau = \min(\beta, \gamma-\alpha)$, R = Besov norm preserving factor	Multiply coefficients at level j with $R\, 2^{-j\tau}$, $-\tau = \min(1, 1-\alpha)$, R = "energy" preserving renormalization

5. CHARACTERIZATION OF ORDERED DITHER HALFTONE FROM WAVELET COEFFICIENTS

Wavelets have shown to be a very successful tool for denoising of images. In most cases, the noise is assumed to be some kind of random noise, most often Gaussian. Halftone patterns are often seen as being noise. Whereas error diffusion has some randomness in it, ordered dither clearly has to be treated as a deterministic noise. The theoretical results for Gaussian noise allow noise removal by simple thresholding of wavelet coefficients. In that case the threshold is chosen such that it is almost surely an upper bound on the magnitude of noisy wavelet coefficients. In applications it is always assumed that the noise level is not higher than the signal level, i.e. in general the noise pixels have a smaller amplitude than most of the real image edges. For dithered halftone patterns, this is not necessarily true, since halftone patterns usually consist of black and white dots of some diameter. The jump between black and white is considered to be a big edge. If a threshold had to be chosen such that wavelet coefficients associated with those edges disappear, a lot of text edges and real edges in images would be thresholded out. Therefore, the question has to be asked of how to distinguish real text and image edges from ordered dither dots in the wavelet domain such that those halftone wavelet coefficients can be thresholded, whereas coefficients associated to text or real image edges are maintained. A method for inverse halftoning in the wavelet domain has been presented that uses edge information across scales to derive a thresholding scheme.[12] The method has been derived for removing one specific halftone frequency from photographic images, not multicomponent documents. The approach presented in the following is specifically designed for distinguishing text from halftone patterns of various frequencies. Wavelet transform maxima in combination with probabilistic methods for estimation of character locations have been used to segment characters and use the output for other processing steps such as OCR.[9] Such a computational complex approach could also be used in a text segmentation step in the context of this paper, but it is designed only for segmentation of characters on background with scanner noise, not halftone noise. The approach chosen in this paper is suitable to distinguish between significant text and image edges on the one hand and scanner and halftone noise on the other hand while keeping the computational complexity low.

The difference between ordered dither halftone and real text or image edges in the spatial domain is that a real edge consists of a chain of points with rapidly changing intensity, where halftone is composed of isolated points of strong intensity change. In addition, ordered dither patterns are very periodic, including some offset in two neighboring rows of isolated dots. In order to derive criteria for characterization of halftone in contrast to text in wavelet coefficients we first consider the examples of horizontal, vertical and diagonal lines.

A vertical line in the image produces large wavelet coefficients of the same sign along a vertical line in the HL band, whereas the coefficients away from the line are small. A horizontal line produces large wavelet coefficients of the same sign along a horizontal line in the LH band. A diagonal line produces medium sized wavelet coefficients of the same sign along a diagonal line in the HL, LH and HH band.

On the first level of decomposition a halftoned area produces large coefficients of positive and negative sign in the HL and LH band. These coefficients are not connected, but rather isolated. In the HH bands a halftone area produces isolated large size coefficients. Fig. 1 shows an example of wavelet coefficient of diagonal lines and halftone in the different bands computed with the Haar wavelet system.

The difference in formation of wavelet coefficients of lines and halftone areas can be captured by computing the expected value and the variance in a window of a special length directed in horizontal, vertical and both diagonal directions (see Figure 2). In detail given a window of length 2N +1 the following parameters are computed.

$$\text{In LH: } E_{LH}(j,k) = 1/(2N+1)\, \Sigma_{m=-N,\ldots N}\, d_{j,k+m}, \text{ and } V_{LH}(j,k) = 1/(2N+1)\, \Sigma_{m=-N,\ldots N}\, (d_{j,k+m} - E_{LH}(j,k))^2, \quad (6)$$

$$\text{in HL: } E_{HL}(j,k) = 1/(2N+1)\Sigma_{m=-N,\ldots N}\, d_{j+m,k}, \text{ and } V_{HL}(j,k) = 1/(2N+1)\, \Sigma_{m=-N,\ldots N}\, (d_{j+m,k} - E_{HL}(j,k))^2, \quad (7)$$

$$\text{in HH: } E_{HH1}(j,k) = 1/(2N+1)\Sigma_{m=-N,\ldots N}\, d_{j+m,k+m}, \text{ and } V_{HH1}(j,k) = 1/(2N+1)\, \Sigma_{m=-N,\ldots N}\, (d_{j+m,k+m} - E_{HH1}(j,k))^2, \text{ and} \quad (8)$$

$$E_{HH2}(j,k) = 1/(2N+1)\Sigma_{m=-N,\ldots N}\, d_{j-m,k+m}, \text{ and } V_{HH2}(j,k) = 1/(2N+1)\, \Sigma_{m=-N,\ldots N}\, (d_{j-m,k+m} - E_{HH2}(j,k))^2. \quad (9)$$

In terms of those parameters the characterization of lines and halftone areas in Table 2 is derived. It becomes obvious that for the simplified line-vs.-halftone model the mean computed in a horizontal or vertical window is a sufficient criterion for separating horizontal and vertical lines from halftone or diagonal lines. For distinguishing diagonal lines from halftone the standard deviation has to be computed as an additional criterion.

From the information from Table 2 a thresholding scheme can be derived to classify coefficients as line/contour or non-line/contour coefficients, i.e. as *significant* or *non-significant* coefficients. If, e.g., $E_{HL}(j,k) > T_1$ and $V_{HL}(j,k) < T_2$ the coefficient d_{jk} is considered to be a significant coefficient. For such a coefficient a threshold $T_{significant}$ is used in the denoising step. Contrarily, if $E_{HL}(j,k) < T_1$ or $V_{HL}(j,k) > T_2$ the coefficient d_{jk} is considered to be a non-significant coefficient. In this case a threshold $T_{non\text{-}significant} > T_{significant}$ is used for denoising. A weaker criterion is to require only $E_{HL}(j,k) > T_1$ without any restriction on the variance $V_{HL}(j,k)$.

$$\begin{bmatrix} 1&1&0&0&1&1\\ 1&1&0&0&1&1\\ 0&0&1&1&0&0\\ 0&0&1&1&0&0\\ 1&1&0&0&1&1\\ 1&1&0&0&1&1 \end{bmatrix} \qquad \begin{bmatrix} 0&0&0&1&1&1\\ 0&0&1&1&1&0\\ 0&1&1&1&0&0\\ 1&1&1&0&0&1\\ 1&1&0&0&1&1\\ 1&0&0&1&1&1 \end{bmatrix}$$

halftone pattern in image — diagonal line pattern in image

$$\begin{bmatrix} 0&2&0&-2&0&2\\ 0&0&0&0&0&0\\ 0&-2&0&2&0&-2\\ 0&0&0&0&0&0\\ 0&2&0&-2&0&2\\ 0&0&0&0&0&0 \end{bmatrix} \qquad \begin{bmatrix} 0&0&0&0&0&0\\ 2&0&-2&0&2&0\\ 0&0&0&0&0&0\\ -2&0&2&0&-2&0\\ 0&0&0&0&0&0\\ 2&0&-2&0&2&0 \end{bmatrix} \qquad \begin{bmatrix} 0&0&0&0&0&0\\ -2&0&2&0&-2&0\\ 0&0&0&0&0&0\\ 2&0&-2&0&2&0\\ 0&0&0&0&0&0\\ -2&0&2&0&-2&0 \end{bmatrix}$$

halftone pattern in HL, LH, HH band (left to right)

$$\begin{bmatrix} 1&1&-1&-1&0&1\\ 1&-1&-1&0&1&1\\ -1&-1&0&1&1&0\\ -1&0&1&1&0&-1\\ 0&1&1&0&-1&-1\\ 1&1&0&-1&-1&0 \end{bmatrix} \qquad \begin{bmatrix} 1&1&-1&-1&0&1\\ 1&-1&-1&0&1&1\\ -1&-1&0&1&1&0\\ -1&0&1&1&0&-1\\ 0&1&1&0&-1&-1\\ 1&1&0&-1&-1&0 \end{bmatrix} \qquad \begin{bmatrix} -1&1&1&-1&0&-1\\ 1&1&-1&0&-1&1\\ 1&-1&0&-1&1&1\\ -1&0&-1&1&1&-1\\ 0&-1&1&1&-1&0\\ -1&1&1&-1&0&-1 \end{bmatrix}$$

diagonal line pattern in HL, LH, HH band (left to right)

Figure 1 — Formation of wavelet coefficients of halftone and diagonal line pattern using 2D Haar wavelet transform.

LH band

d_{11}	d_{12}	d_{13}
$\mathbf{d_{21}}$	$\mathbf{d_{22}}$	$\mathbf{d_{23}}$
d_{31}	d_{32}	d_{33}

HL band

d_{11}	$\mathbf{d_{12}}$	d_{13}
d_{21}	$\mathbf{d_{22}}$	d_{23}
d_{31}	$\mathbf{d_{32}}$	d_{33}

HH band

d_{11}	d_{12}	$\mathbf{d_{13}}$
d_{21}	$\mathbf{d_{22}}$	d_{23}
$\mathbf{d_{31}}$	d_{32}	d_{33}

HH band

$\mathbf{d_{11}}$	d_{12}	d_{13}
d_{21}	$\mathbf{d_{22}}$	d_{23}
d_{31}	d_{32}	$\mathbf{d_{33}}$

Figure 2 — Wavelet coefficients in 3x3 windows in the LH,HL,HH bands of one level of a wavelet decomposition are shown. Bold coefficients are those used for computing the expected values and variances from Eq. (6)-(9) for N=1.

Table 2 — Characterization of lines and halftone area in directional window in terms of wavelet coefficients parameters.

magnitude of	E_{LH}	V_{LH}	E_{HL}	V_{HL}	E_{HH}	V_{HH}
horizontal lines	**large**	*small*	*small*	*small*	*small*	*small*
vertical lines	*small*	*small*	**large**	*small*	*small*	*small*
diagonal lines	medium	*small*	medium	*small*	medium	*small*
halftone	medium	**large**	medium	**large**	medium	**large**

.

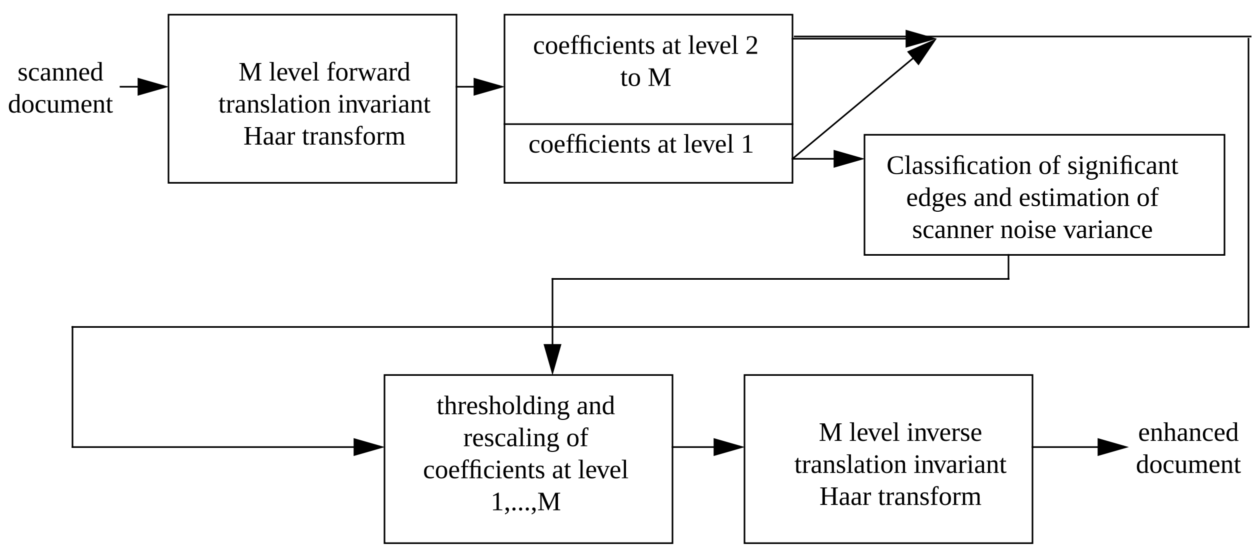

Figure 3 — Overview of wavelet-based enhancement system for deblurring and denoising in a digital copier.

6. EXPERIMENTAL RESULTS

Experiments have been performed using a 300 and 600dpi scanner and a variety of documents containing text of different font size as well as ordered dither halftone patterns of various frequencies. The enhancement has been performed following the practical instructions listed in Table 1 in combination with the line-vs-non-line classification from the previous section. An overview of the system used in the experiments is given in Figure 3.

The results of the enhancement of a scanned magazine page are displayed in Fig. 4-6. The page contained text, white background, and a picture represented by ordered dither halftone. The enhancement algorithm follows the practical instructions from Table 1 choosing $\tau = -0.5$, a renormalization constant of $R = 0.29$ and two levels of decomposition. The thresholds for distinguishing text versus non-text edges T_1 and T_2 from Section 5 were the result of some tuning ($T_1 = 46$, $T_2 = 18$). The parameter $T_{significant}$ was determined following the Donoho-scheme[6] including the median of wavelet coefficients at the finest scale ($T_{significant} = 11$). The parameter $T_{non\text{-}significant}$ for denoising of non-text areas was set by tuning ($T_{non\text{-}significant} = 100$ for scale 1, $T_{non\text{-}significant} = 30$ for scale 2).

Fig. 4 shows the enhancement results in an area with text, white background and some halftone patterns. The wavelet-based enhancement sharpens the text while simultaneously removing scanner noise from the white background and halftone patterns. For comparison the result of using the unsharp masking filter in Photoshop is given. The unsharp masking parameters were set such that the sharpness in the text area is similar to the one in the wavelet-enhanced image. Scanner and halftone noise are magnified after applying the unsharp masking filter.

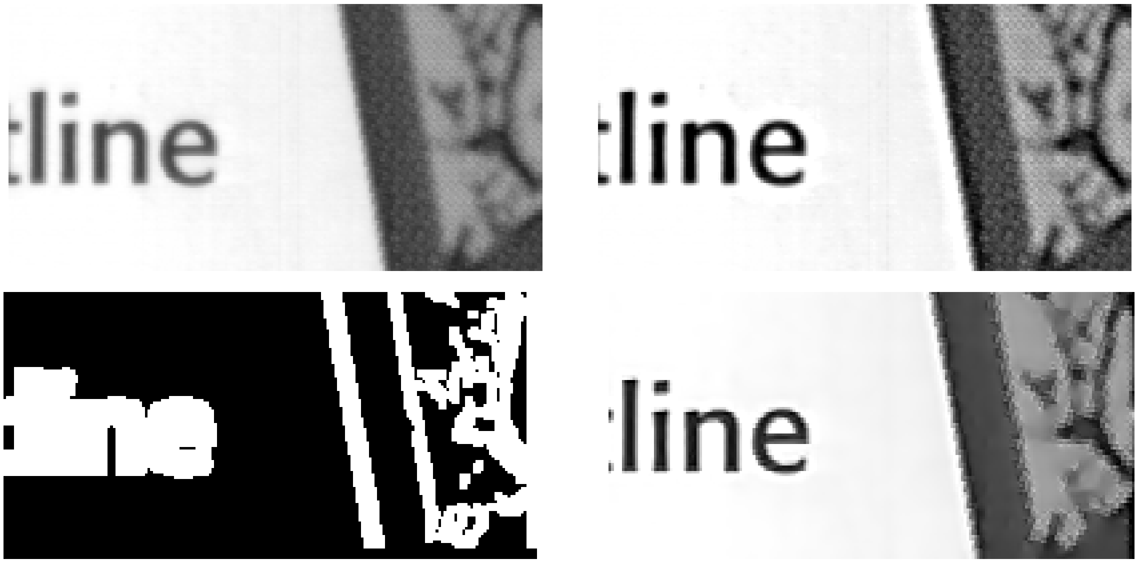

Figure 4 — Top left: Example of a scanned magazine page containing text, scanner noise, halftone pattern. Top right: Result after applying the Photoshop unsharp masking filter (pixel radius = 4, amount = 100%). Bottom left: Line classification as described in Section 5. Bottom right: Result after wavelet-based deblurring and denoising using classification data from bottom left image.

Fig. 5 shows the enhancement results in an area of halftone patterns and image edges. The wavelet-based enhancement suppresses halftone patterns while keeping the image edges sharp, whereas the unsharp masking filtering enhances the halftone everywhere in the image.

Fig. 6 shows the difference between a classifier using only the threshold $T_{significant}$ derived from scanner noise coefficients for classifying significant versus non-significant edges and a classifier using T_1 and T_2 for the classification. In the first case halftone edges are still classified as significant edges and would be enhanced. In the latter case, only image edges are classified as significant edges.

7. CONCLUSIONS

Wavelet bases and their connections to Besov spaces make it possible to approximate functions under consideration of their smoothness properties, such as differentiability. Using this framework a definition of increasing and decreasing of smoothness by shifting between different Besov spaces is derived. This shifting of smoothness spaces is used to model blurring or deblurring of images. The implementation of the shifting is simply a rescaling of wavelet coefficients. An application to a real-world problem of deblurring with wavelets in Besov spaces is shown at the example of the enhancement of a scanned document. In addition to the deblurring step characteristics of small groups wavelet coefficients are computed and used to derive classification data that allow to perform various degrees of denoising by thresholding in different areas of a document, e.g. in halftone and text areas. There are several advantages of the proposed wavelet-based technique compared to traditional Fourier-base techniques.

- The proposed technique takes advantages of the superior denoising performance of wavelets compared to Fourier-based techniques in keeping edges sharp while suppressing noise away from edges, e.g. in background areas.
- Most of the parameters used in the algorithm can be precomputed using an appropriate modeling of images and image processing steps in Besov spaces.
- For document enhancement the simple translation invariant Haar wavelet system has sufficient smoothness properties. This leads to very simple filtering routines for computing the wavelet coefficients.

.

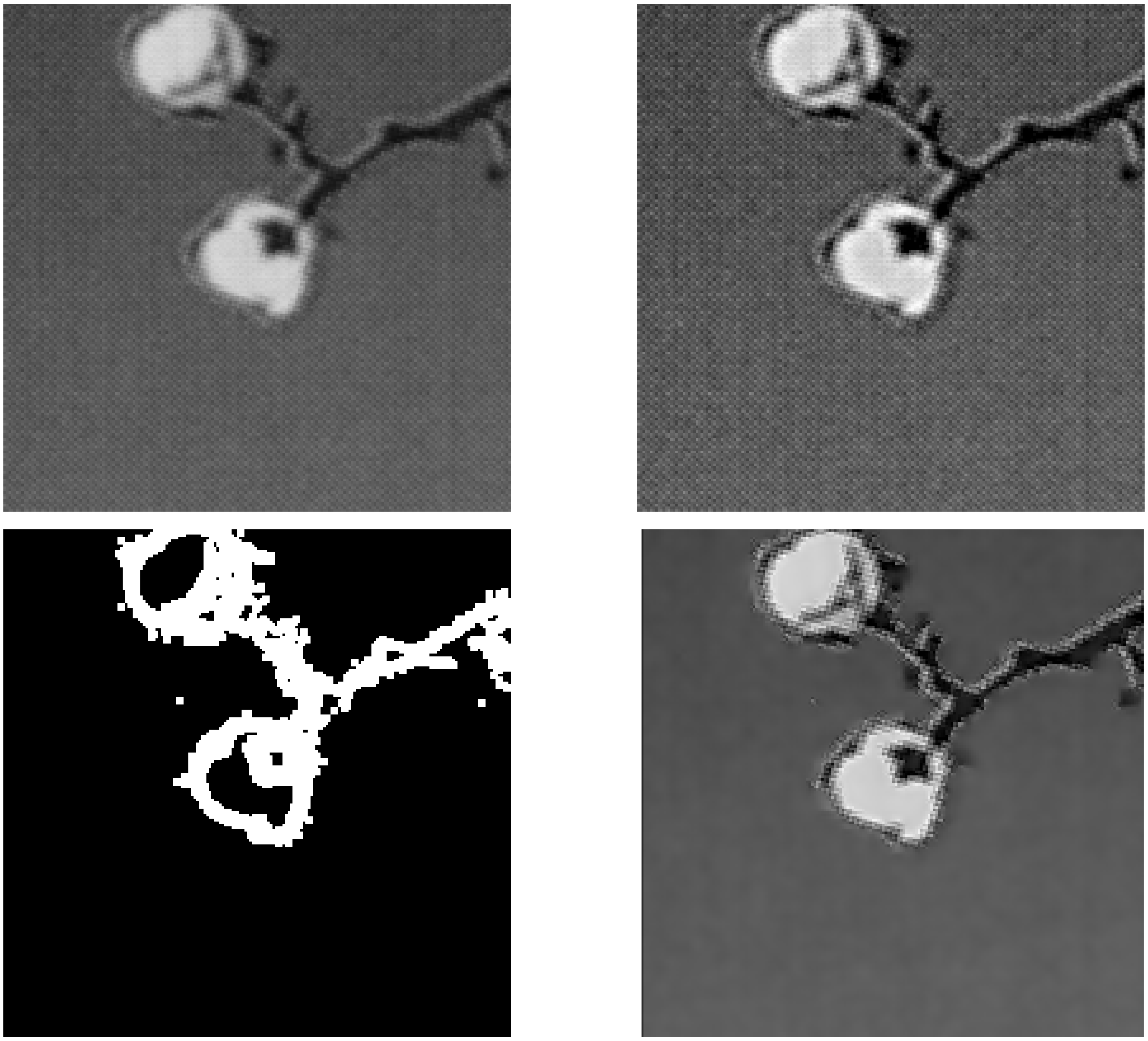

Figure 5 — Top left: Example of a scanned magazine page containing image contours and halftone pattern. Top right: Result after applying the Photoshop unsharp masking filter (pixel radius = 4, amount = 100%). Bottom left: Line classification as described in Section 5. Bottom right: Result after wavelet-based deblurring and denoising using classification data from bottom left image.

- The image processing steps deblurring and denoising are simple processing steps on wavelet coefficients, namely thresholding and rescaling.

Results of enhanced scanned images using the propose wavelet-based technique are compared to a traditional Fourier-based technique. The high quality performance and simple algorithmic implementation of the technique makes it an attractive tool for applications in digital imaging products.

.

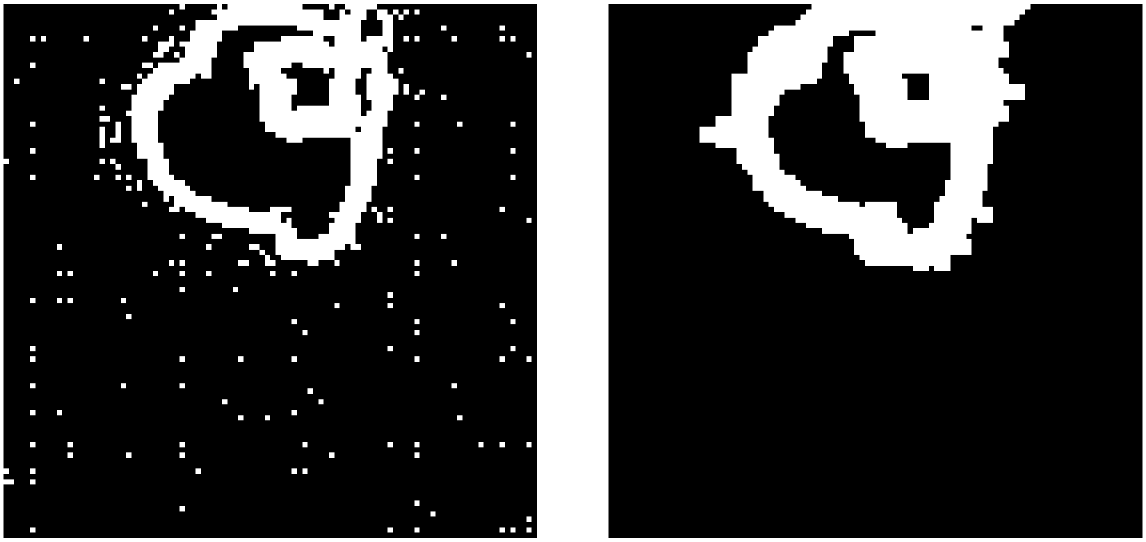

Figure 6 — Left: Edge classification using only absolute values of individual coefficients. Halftone patterns are miss classified as edges. Right: Line classification as described in Section 5.

REFERENCES

1. Andrews, H.C., Hunt, B.R., *Digital Image Restoration*. Prentice-Hall, Inc., Englewood Cliffs, New Yersey, 1977.
2. Berkner, K., Gormish, M., Schwartz, E., "Multiscale Sharpening and Smoothing in Besov Spaces with Applications to Image Enhancement," *Applied and Computational Harmonic Analysis*, vol. 11, pp.2-31, 2001.
3. Berkner, K., Wells Jr, R.O., "Denoising via nonorthogonal wavelet transforms," in *Wavelet transforms and time-frequency signal analysis*, (L. Debnath, ed.), pp. 67-80, Springer Verlag, 2001.
4. Coifman, R.R., Donoho, D.L., "Translation invariant denoising," in *Wavelets and Statistics, Springer Lecture Notes* (A. Antoniades, ed.), Springer-Verlag, 1995.
5. Daubechies, I., *Ten Lectures on Wavelets*. SIAM Publ., 1992.
6. Donoho, D.L., "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.
7. Donoho, D.L., Vetterli, M., DeVore, R.A., Daubechies, I., "Data compression and harmonic analysis," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2435–2476, 1998.
8. Frazier, M., Jawerth, B., Weiss, G., *Littlewood-Paley Theory and the Study of Function Spaces*, vol. 79 of *NSF-CBMS Regional Conf. Series in Mathematics*. American. Math. Society, 1991.
9. Hwang, W.L., Chang, F., "Character Extraction from Documents using Wavelet Maxima," *Image and Vision Computing*, vol. 16, pp. 307-315, 1998.
10. Jain, A.K., *Fundamentals of digital image processing*. Prentice Hall, Englewood Cliffs, NJ, 1989.
11. Mallat, S., *A Wavelet Tour of Signal Processing*. Academic Press, 1998.
12. Xiong, Z., Orchard, M., Ramchandran, K., "Inverse Halftoning Using Wavelets," IEEE Trans. Signal Processing, vol. 8, pp. 1479-1482, 1999.