IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727
PP 25-29
www.iosrjournals.org

Document Image Segmentation Using Wavelet Transform and Gabor Filter Technique

Manish T. Wanjari[1], Dr. Mahendra P. Dhore[2]
[1](*mwanjari9@gmail.com, Department of Computer Science, SSESA's, Science College, Congress Nagar, Nagpur, India)*
[2](*mpdhore@rediffmail.com, Department of Electronics & Computer Science, RTM Nagpur University Campus, Nagpur, India)*

Abstract : *Document Image Segmentation subdivides an document image into its constituent regions or objects. The level to which the subdivision is carried out depends on the problem being solved. The wavelet transform provides a compact description of document images and it is very helpful in description of edges and lines that are highly localized. 2-D wavelet decomposition is used for document images. The process of texture segmentation using Gabor filters involves proper design of a filter bank tuned to different spatial-frequencies and orientations to cover the spatial-frequency space; decomposing the document image into a number of filtered document images; extraction of features from the filtered document images; and the clustering of pixels in the feature space to produce the segmented document image. This paper deals with segmentation of document image based on wavelet transform and Gabor filter technique. The experimental results are obtained for wavelet transform using Discrete Cosine Transform coefficients & DWT, Inverse DWT and creating Gabor filter for document images.*

Keywords - *Image Segmentation, Document Image Segmentation, Wavelet Transform, Gabor Filter.*

I. INTRODUCTION

Wavelets Transform is the new approach to signal processing and analysis called multi-resolution theory. Multiresolution theory incorporates and unifies techniques from a variety of disciplines, including sub-band coding from signal processing, quadrature mirror filtering from digital speech recognition and pyramidal image processing. Multiresolution theory is concerned with the representation and analysis of document images at more than one resolution. In this paper, we examine wavelet based transformations from a multiresolution point of view. Although they can be presented in other ways, this approach simplifies both the mathematics and physical interpretations [1]. Document image segmentation is an essential step in many advance techniques of multi-dimensional signal processing and its applications. This paper describe the technique of wavelet transform used for features extraction associated with individual document image pixels and comparison of this method with application of the gabor filter technique. Wavelets are a mathematical tool for hierarchically decomposing functions in the frequency domain by preserving the spatial domain. Wavelets have found more and more applications in computer graphics, such as document image compression, digital image processing and feature detection [2, 3].

In image processing a gabor filter is a linear filter used for edge detection. Frequency and orientation representations of gabor filters are similar to those of the human visual system and they have been found to be particularly appropriate for texture representation and discrimination. The document image analysis by the gabor functions is similar to perception in the human visual system [4].

The Gabor Transform is also referred to as the Short Time Fourier Transform (STFT). Filtering the time-frequency content of a signal is indeed one of the main applications of Gabor multipliers. Gabor analysis is an essential part of time-frequency analysis, initiated by the seminal paper of Denis Gabor in 1946.

Some properties of Gabor filters:

i) A tunable bandpass filter
ii) Similar to a STFT or windowed Fourier transform
iii) Satisfies the lower-most bound of the time-spectrum resolution (uncertainty principle)
iv) It's a multi-scale, multi-resolution filter
v) Has selectivity for orientation, spectral bandwidth and spatial extent.
vi) Has response similar to that of the Human visual cortex (first few layers of brain cells)
vii) Used in many applications – texture segmentation; iris, face and fingerprint recognition.

viii) Computational cost often high, due to the necessity of using a large bank of filters in most applications [5].

II. WAVELET TRANSFORM

The wavelet transform has similar properties to Fourier transform as a mathematical technique for document image analysis, the basic difference between both is that wavelets are localized in both time and frequency, whereas the standard Fourier transform is only localized in frequency [6]. When digital document images are viewed or processed at multiple resolutions, the discrete wavelet transform (DWT) is the mathematical tool of choice. In addition to being an efficient, highly intuitive framework for the representation and storage of multiresolution document images, the DWT provides powerful insight into a document image spatial and frequency characteristics.

The wavelet transform is important to provide a compact description of document images that are limited in time and it is very helpful in description of edge and line that are highly localized. 2-D wavelet decomposition is use for document images. This 2-D wavelet transform requires two wavelets such as $\psi^1(x, y)$ and $\psi^2(x, y)$. At a particular scale s we have,

$$\psi_s^i(x,y) = \frac{1}{s2}\ \psi i\left(\frac{x}{s},\frac{y}{s}\right) \quad i = 1, 2 \ldots \quad (1)$$

By applying each one f(x, y), at a scale s=2j we will have a component

$$W_{2^j}^i f(x,y) = \left(f * W_{2^j}^i\right)(x,y) \; i= 1, 2\ldots \quad (2)$$

Then the original signal f(x, y) can be represented by the 2-D wavelet transform in terms of the two dual wavelets $\xi^1(x, y)$ and $\xi^2(x, y)$ [9]

$$f(x,y) = \sum\left(\left(W_{2^j}^1 f * \xi_{2^j}^1\right)(x,y) + \left(W_{2^j}^2 * \xi_{2^j}^2\right)(x,y)\right) \ldots \quad (3)$$

And it is required a scaling function $\varphi(x, y)$ for build a multistage representation, corresponding component at a scale 2j is

$$S_{2^j} f(x,y) = \left(f * \varphi_{2^j}\right)(x,y) \ldots . \quad (4)$$

These wavelet measure function variations along different directions [11]. We may interpret the component $S_{2^j} f(x,y)$ as a smoothed version of f(x, y) and the components for j = 1…J, as the document image details lost by smoothing going from $S_{2^0} f(x,y)$ to $S_{2^j} f(x,y)$.

By using wavelet on a document image for one level, document images will be obtained with correspond to the approximation and detail document images. [7]

2.1. Discrete Wavelet Transform

Like the Fourier expansion, the wavelet series expansion maps a function of a continuous variable into a sequence of coefficients. If the function being expanded is a sequence of numbers, like samples of a continuous function f(x), the resulting coefficients are called the discrete wavelet transform (DWT) of f(x).

Discrete wavelet transform is applied to document images because wavelets provide frequency information as well as time space localization. In addition their multiresolution text enables to visualize document image at various scales and orientation. The multiresolution property provides information about various high frequency components at different levels of decomposition. [8]

2.2. The Continuous Wavelet Transform

The natural extension of the discrete wavelet transform in the continuous wavelet transform (CWT), which transforms a continuous function into a highly redundant function of two continuous variables translation and scale. The resulting transform is easy to interpret and valuable for time-frequency analysis. Although our interest is in discrete images, it is covered here for completeness.

III. DOCUMENT IMAGE SEGMENTATION USING WAVELET TRANSFORM

We can now formally define several closely related wavelet transformations, the generalized wavelet series expansion, the discrete wavelet transform and continues wavelet transform. A computationally efficient implementation of the discrete wavelet transform called fast wavelet transform.

3.1. The Fast Wavelet Transform

An important consequence of the properties is that both φ(x) and ψ(x) can be expressed as linear combinations of double-resolution copies of themselves. That is the series expansions as follows

$$\varphi(x) = \sum_n h_\varphi(n)\sqrt{2}\varphi(2x-n)$$
$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\psi(2x-n)$$

Where, h_φ and h_ψ the expansion coefficients are called scaling and wavelet vectors, respectively. They are the filter coefficient of the fast wavelet transform (FWT), an iterative computational approach to the DWT.

3.2. The Inverse Fast Wavelet Transform

The inverse fast wavelet transform can be computed iteratively using digital filters.

3.3. Wavelets in Image Processing

The basic approach to wavelet based image processing is to

i) Compute the two-dimensional wavelet transform of a document image.
ii) Alter the transform coefficients.
iii) Compute the inverse transform.

As scale in the wavelet domain is analogous to frequency in the Fourier domain, most of the Fourier-based filtering techniques have an equivalent “wavelet domain” counterpart. [9].

IV. GABOR FILTER

Gabor filter is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. It is optimally localized as per the uncertainty principle in both the spatial and frequency domain. This implies Gabor filters can be highly selective in both position and frequency, thus resulting in sharper texture boundary detection .Gabor filter related segmentation paradigm is based on filter bank model in which several filters are applied simultaneously to an input document image. The filters focus on particular range of frequencies. If an input document image contains two different texture areas, the local frequency differences between the areas will detect the textures in one or more filter output sub-images. The process of texture segmentation using Gabor filters involves proper design of a filter bank tuned to different spatial-frequencies and orientations to cover the spatial-frequency space; decomposing the document image into a number of filtered document images; extraction of features from the filtered document images; and the clustering of pixels in the feature space to produce the segmented document image [10-11].

The Gabor filter is based on a multi-channel filtering theory for processing visual information in the early stages of the human visual systems. Gabor filters are directly related to Gabor wavelets, since they can be designed for a number of dilations and rotations. However, in general, expansion is not applied for Gabor wavelets, since this requires computation of bi-orthogonal wavelets, which may be very time-consuming. Therefore, usually, a filter bank consisting of Gabor filters with various scales and rotations is created. The filters are convolved with the signal, resulting in a so-called Gabor space.

The Gabor space is very useful in document image processing applications such as optical character recognition and text recognition. Relations between activations for a specific spatial location are very distinctive between objects in an document image. Important activations can be extracted from the Gabor space in order to create a sparse object representation [12, 13].

V. EXPERIMENTAL RESULT

The performance of wavelet transform by using Discrete Cosine Transform coefficient and Gabor filter using the different orientation of Theta values i.e. Theta = 0 or pi/4 or pi/2 is as below.

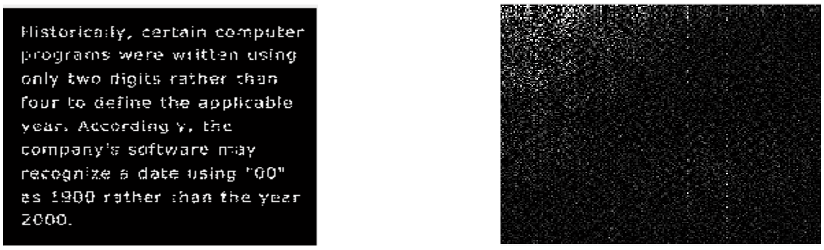


Fig. 1. a) Original Document Image & b) DCT coefficient

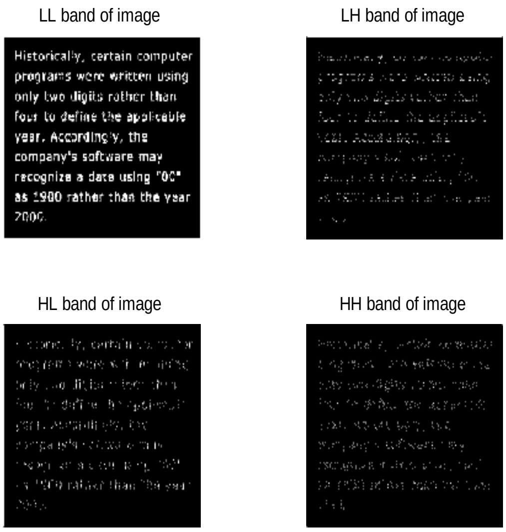
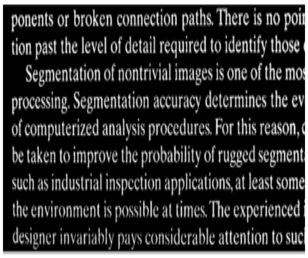
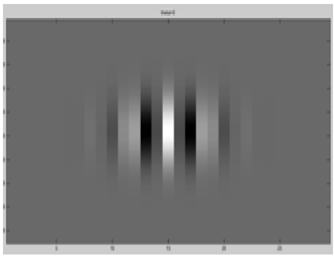


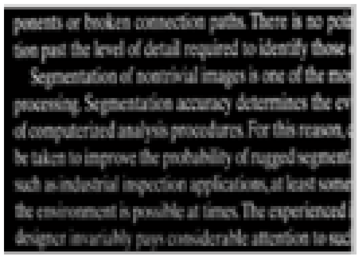
Fig. 2. DWT & Inverse DWT



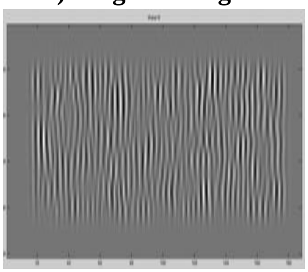
a) Original image



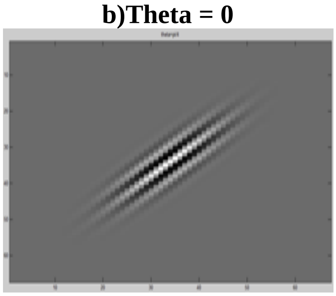
b)Theta = 0



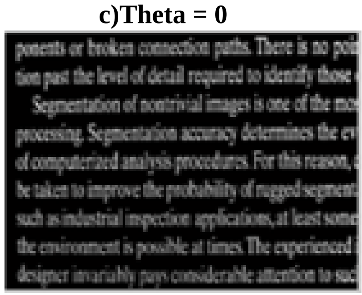
c)Theta = 0



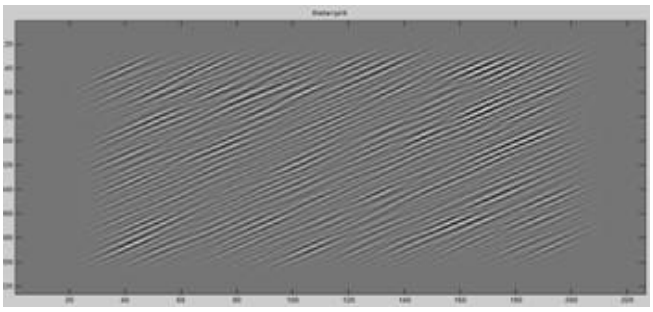
d) Theta = 0



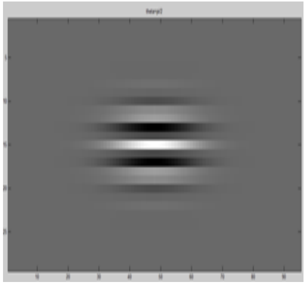
e)Theta = pi/4



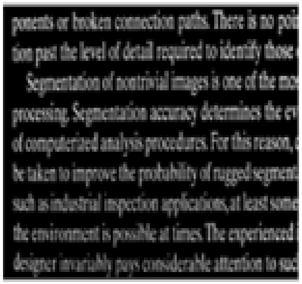
f) Theta = pi/4



g)Theta = pi/4

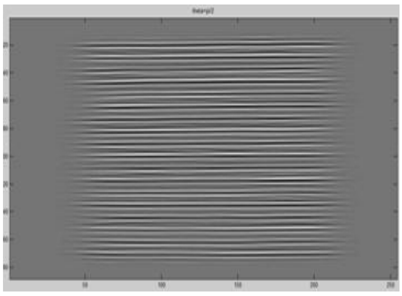


h)Theta = pi/2



ponents or broken connection paths. There is no poi
tion past the level of detail required to identify those
Segmentation of nontrivial images is one of the mo
processing. Segmentation accuracy determines the ev
of computerized analysis procedures. For this reason,
be taken to improve the probability of rugged segment
such as industrial inspection applications, at least some
the environment is possible at times. The experienced
designer invariably pays considerable attention to suc

i)Theta = pi/2



j) Theta = pi/2

Fig. 3. Gabor filter

Figure 1. shows, a) Original document image b) Discrete cosine transform coefficient. Figure 2. shows, Discrete wavelet transform and inverse discrete wavelet transform. Figure 3. shows Created gabor to the input document image and apply gabor filter with a specific scale and three orientations 0, pi/4 and pi/2 to a input gray scale document image.

VI. CONCLUSION

In this work, we have described document image segmentation performed by using wavelet transform and gabor filter technique. Wavelet transforms are used for increasing accuracy. Resolution reduction by wavelet is dependent on amount of noise in the document image and also the desired target size. This has been efficiently exploited by using a Gabor Filter.

REFERENCES

[1] R. C. Gonzalez, R. E. Woods, “Digital Image Processing, *Unoversity of Tennessee, MedData interactive, 2005.*

[2] S. Mallat, “A compact multiresolution representation: the wavelet model”, *Proc. IEEE Computer Society Workshop on Computer Vision, IEEE Computer Society Press, Washington, D.C (1987), p.2-7.*

[3] E. J. Stollnitz, T. D. Derose, & D. H. Salesin, “Wavelets for Computer Graphics”, Morgan Kaufmann Publishers, Inc., San Francisco, CA (1996).

[4] Hans G. Feichtinger, Thomas Strohmer: "Gabor Analysis and Algorithms", *Birkhäuser, 1998; ISBN 0-8176-3959-4.*

[5] A. R. Rao, G. L. Lohse, “Identifying high level features of texture perception,” *CVGIP: Graphical Models and Image Proc. 55(3), pp. 5218-5233, 1993.*

[6] R. D. da Silva1, W. R. Schwartz and H. Pedrini, “Image segmentation based on wavelet feature descriptor and dimensionality reduction applied to remote sensing”, *Chilean Journal of Statistics Vol. 2, No. 2, September 2011, 51-60.*

[7] Jung, C., R., “Combining wavelets and watersheds for robust multiscale image segmentation”, Image and Vision Computing 25(2007), pp. 24–33.

[8] N.Valliammal, Dr.S.N.Geethalakshmi, “Leaf Image Segmentation Based On the Combination of Wavelet Transform and K Means Clustering”, *(IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 1, No. 3, 2012.*

[9] M. T. Wanjari, K. D. Kalaskar & Dr. M. P. Dhore, “Wavelet and Watershed Transform based Document Image Segmentation”, International Journal of Advanced Computational Engineering and Networking (IJACEN), Special Issue 1, June 2015, pp. 519-524, ISSN No. 2320-2106 .

[10] A. R. Rao, G. L. Lohse, “Identifying high level features of texture perception,” *CVGIP: Graphical Models and Image Proc. 55(3), pp. 5218-5233, 1993.*

[11] D. Clausi, M. Ed Jernigan, “Designing Gabor filters for optimal texture separability”, *Pattern Recognition, vol. 33, pp. 1835-1849, 2000.*

[12] B. Allier, H. Emptoz, “Font Type Extraction and Character Prototyping Using Gabor Filters” *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR’03) pp.799-803, 2003.*

[13] V. K. Yeotikar, M. T. Wanjari & Dr. M. P. Dhore, “a Novel Approach for Text Extraction from Document Images”, *International Journal of Research in Computer Engineering and Electronics (IJRCEE), Vol. 1, issue 1, pp. 1-4, ISSN No. 2319-376X.*