Abstract

Digital image processing is a rapidly growing field with numerous applications in almost all areas of science and engineering. There is an abundance of these applications that can serve mankind with available and anticipated technology in the near future. Wavelets have emerged as a fruitful mathematical theory and a tool for development of various image processing applications due to the advantages of space-frequency localization and multi-resolution representation. Wavelets have proved their importance in various image processing applications such as denoising, compression, image enhancement, edge detection, image fusion and pattern recognition etc.

In this thesis, an effort has been made to investigate effectiveness and advance state of the art usage of wavelets as a tool in four very important applications of image processing. These four applications are Watermarking of grayscale and color images, denoising of documents images for OCR applications, Optical Character Recognition (OCR) of multi-font Roman characters and Content Based Image Retrieval (CBIR). A chapter wise description of the work done is as follows.

Digital Image Watermarking

Watermarking has evolved as a great tool for copyright protection of multimedia content such as images and videos. Watermarking is a process in which information in form of either text or image is permanently embedded in the host data. The process of watermarking should be able to provide enough robustness against the attacks, which are purposely done by attackers to destroy the watermark with the intention of unauthorized copy of the protected data. The prominent watermarking approaches are classified as Spread Spectrum techniques, Quantization based techniques etc. In this work the Spread Spectrum technique of watermarking which is known for better robustness is taken as the base technique to study the effectiveness of wavelets in this process of watermarking and development of new algorithms.

The performance of wavelet based watermarking is affected by various parameters such as appropriate selection of wavelet basis functions, level of decomposition and selection of various wavelet coefficients for watermark
embedding. Various types of wavelets are available. Each wavelet has its own specific properties and, therefore, its choice affects the watermarking performance differently. Selection of appropriate wavelet coefficients for watermarking is also a very important task as different wavelet coefficients have different properties.

In most of the wavelet based methods of watermarking proposed in the literature, the wavelet transform is just utilized as a transformation tool rather than emphasizing the selection of appropriate wavelet function, impact of level of decomposition and appropriate selection of embedding plane. In this work a detailed study is carried out regarding these issues for better watermarking performance. It is found that the performance of Bior6.8 wavelet is better than other wavelets for level 2 watermarking and the diagonal wavelet coefficients provide better choice as embedding plane. The other wavelets such as Coif5 and Db10 are slightly worse than Bior6.8.

Image contents also play an important role in determining watermarking performance. Effect of image contents on robustness of watermarking is also investigated and found that textured images produce larger wavelet coefficients as compared to smooth images at a given level of decomposition. Therefore, the robustness for textured images becomes poorer. The robustness can be increased by decomposing textured images at level 1 rather than decomposing them at level 2. A method based on Edge Histogram is proposed and implemented, which can distinguish between textured and non-textured images and takes the appropriate decision of decomposing the input image at either level 1 or level 2 for watermarking.

The effect of utilizing all the high frequency diagonal wavelet coefficients of a wavelet tree is also investigated. For this purpose Discrete Wavelet Packet Transform (DWPT) is employed for watermarking. The diagonal coefficients at level 2 decomposition with maximum entropy are selected for embedding. It is found that this scheme performs better in terms of imperceptibility and robustness and satisfies the characteristics of Human Visual Systems (HVS).

The problems of watermarking document and graphics images are also investigated. Watermarking of such images is not easy due to the large white background area. Any attempt to watermark this plane may cause visual artifacts. Therefore, a fuzzy based approach is presented which automatically selects the embedding strength based on image contents and modifies the wavelet
coefficients only related to the contents of the image and not the coefficients related to the white background area.

Watermarking of color images is also studied. A detailed analysis shows that Bior6.8 wavelet and diagonal wavelet coefficients at level 2 are better choices in this case too for chosen YCbCr color space. The issue of smooth and textured images remains same for color images also.

The choice of color space and selection of embedding color channel is critical. It is found that RGB model and watermarking in Blue plane is not suitable due to poor robustness against compression attacks because during JPEG compression Blue channel is heavily quantized. However, it is able to provide better imperceptibility. The YCbCr model is more appropriate if the watermark is embedded in Y channel because Y channel embedding can withstand compression type attacks. Y channel is more sensitive to human eyes and, therefore, embedding capacity for the Y channel is very limited. In this case, additional payload can be embedded in Cr channel without unduly compromising the visual quality of watermarked images.

**Document Image Denoising and Binarization**

Image denoising is an important process which finds a role in various image processing applications as main process or as sub process. A good image denoising algorithm must be able to remove the noise present while preserving image details such as edges. The majority of denoising algorithms are developed to deal with Gaussian white noise, Impulse and Speckle noise in natural images. The problem of denoising document images is different. In Optical Character Recognition (OCR), Image acquisition (by Scanner or Camera) and binarization are two main sources of noise introduced in document images. The acquisition process mainly introduces clutter, skew, undesirable impressions, Gaussian white noise etc. while binarization introduces various unwanted black spots, blobs etc. in binarized document images which make the images highly unsuitable for the OCR process. These binarization errors occur due to complexity of document images that is generally found in form of uneven illumination, color shading or complex images in the background.

In this work, a simple and effective document image binarization method is presented which does not require manual tuning of any parameter such as
adjustment of window size or value of some constant etc. This method can deal with poor binarization of document images caused by complex backgrounds or uneven illumination. This method is based on background estimation with the help of wavelet transform. The proposed method is tested by denoising a variety of document images considering various background complexities. The denoised images are then processed by OCR software. The results clearly show the superiority and effectiveness of the proposed method over other well-known binarization methods.

The document images can also be corrupted by Gaussian white noise. Generally video frames or pictures containing text may be corrupted by white Gaussian noise due to analog transmission. Capturing such images using a camera for OCR purpose is an example. Another case is text based indexing of pre-recorded video corrupted by white Gaussian noise. In this work, three different techniques of denoising are proposed and compared to remove white Gaussian noise. The first two methods use wavelets whereas the third method is based on curvelets with hard thresholding and modified universal threshold. The use of curvelets for denoising preserves edge features of characters of noisy document images which makes it suitable for OCR applications. All the three methods are compared for various types of document images of different size, font and font size. The results show that the curvelet based method is very efficient and able to preserve edge features of text up to standard deviation ($\sigma$) = 100. The proposed method is effective in removing impulse noise and a combination of Gaussian and Impulse noise too. The performance of the implemented method is compared with denoising performed by commercial OCR package ABBYY Fine Reader. The results of the comparison show the superiority of the proposed method.

The curvelet based method described above is implemented for document images with simple backgrounds but corrupted with white Gaussian noise. The problem becomes more difficult when document images with complex backgrounds are corrupted by white Gaussian noise. In this work, a method is implemented to remove white Gaussian noise as well as complex image backgrounds to achieve better character recognition with help of curvelet transform based approach. The results show that the proposed scheme is more effective when compared to other existing standard binarization techniques in terms of various evaluation parameters. Different types of test images are taken
for testing. Also in the case of high noise ($\sigma = 50$), the performance of proposed scheme is shown to be better than ABBYY Fine Reader and acceptable for effective and efficient character recognition.

**Multi Font Character Recognition**

In the OCR process, the document images which are either scanned by the scanner or captured by a camera are processed, interpreted and then converted into editable text. For example, the images of documents (*.bmp, *.jpg, *.tiff etc.) are not editable, when given to an OCR system, the written text is understood by the machine and converted into editable text format such as *.doc (MS-Word) or *.txt (Notepad) etc. Much work has been done in OCR particularly for Roman script but still the problem is not yet fully solved. For example, robustness of the OCR system against font shape variation is still a big challenge. Even best of commercial OCR packages such as ABBYY Fine Reader and Adobe Reader Pro are also not able to perform well on noisy and degraded images with varied fonts.

In this work, a more robust and efficient OCR scheme is developed and implemented for multi-font Roman script. The proposed technique extracts some geometrical features and wavelet features of the character image and using a proper combination, recognizes the characters in various fonts and sizes. Here wavelet features based on row and column wise standard deviation of wavelet coefficients provide good character recognition strength. The proposed scheme is able to recognize broken and degraded characters also. The scheme is tested for various unknown complex fonts and the performance is compared with sate of the art ABBYY Fine Reader and Adobe Reader Pro. The achieved recognition rate is up to 99.0% across 15 unknown fonts. An effort has also been made to select best wavelet basis for better recognition. It is found that performance of Db10 wavelet is better than other wavelets giving highest recognition rate. The results validate the superiority of the proposed scheme over existing methods. The proposed multi font scheme and previously described background removal scheme are also combined together and tested for various document images with complex backgrounds and varied fonts. The results are competitive when compared to those obtained from commercial OCR packages.
Content Based Image Retrieval

The explosive growth of the internet and the wide use of digital content necessitate the development of effective ways of managing visual information by its content and have increased the need for efficient image retrieval procedures. Due to the rapid proliferation of acquisition hardware, digital acquisition of information (images and videos) has become very popular in recent years. Every day, Giga bytes of images are generated world-wide. Proper management of images and their retrieval from a huge database based on image contents has proved to be a highly challenging problem.

In this work, a wavelet based image retrieval method is proposed based on color and texture features. A new global color feature based on wavelet transform and HSV color space is presented. The texture features are extracted with help of wavelet transform and edge histogram. The dominant edges are captured by calculating edge histogram on all the wavelet coefficients up to 3rd level of decomposition. This scheme utilizes the advantages of multi resolution capability of wavelet transform. The proposed scheme is tested on two different databases, i.e. Wang’s database of 1000 Images and Microsoft Research’s database of 5000 images of various categories. The implemented scheme is also tested under various alterations given to input images such as cropping, rotation, intensity variation, noise addition etc. The results of retrieval are expressed in terms of Precision and Recall. The results show the superior performance of the proposed scheme over various existing schemes. The image retrieval scheme is first tested for only color and only texture feature separately. It is also tested for both the features to test the strength of the combined feature. For various query images 100% relevant images are retrieved from the database up to recall of 50 images.