A Comparative Study in Wavelets, Curvelets and Contourlets as Feature Sets for Pattern Recognition

Angshul Majumdar[1] and Arusharka Bhattacharya[2]
[1]Department of Electrical and Computer Engineering, University of British Columbia, India
[2]Technology Advisory Service, PricewaterhouseCoopers India, India

Abstract: *There have been a number of recent works in computer vision that had used new age multiresolution multidirectional transforms like curvelets and contourlets for face and character recognition. Although these works produced high recognition accuracies they did not provide any comparative study against more well known techniques and hence could not justify the use of these new transforms as against more traditional methods. In this work we will compare the recognition accuracies of the aforesaid two transforms against a very well known multiresolution transform viz. the wavelet transform. this study aims at showing the research community how good or how bad the aforesaid transforms are when compared against wavelets as a feature set for pattern recognition.*

Keywords: *Curvelet, contourlet, face recognition, character recognition.*

Received July 1, 2007; accepted October 23, 2007

1. Introduction

Studies in human visual system and image statistics reveals that image representations should satisfy the following conditions:

- Multiresolution: the representation should allow images to be successively approximated, from course to fine resolutions.
- Localization: the basis elements in the representation should be localized in both the spatial and the frequency domains.
- Critical sampling: for some applications (e.g., compression), the representation should form a basis, or a frame with small redundancy.
- Directionality: the representation should contain basis elements oriented at a variety of directions, much more than the few directions that are offered by separable wavelets.
- Anisotropy: to capture smooth contours in images, the representation should contain basis elements using a variety of elongated shapes with different aspect ratios.

From the above list, the first three are successfully provided by separable wavelets, while the last two require new constructions.

Wavelets, due to this crude directional representation (primarily vertical, primarily horizontal and primarily diagonal), although are good at representing point discontinuities are not good at representing discontinuities along edge. Besides, wavelets and related classical multiresolution ideas exploit a limited dictionary made up of roughly isotropic elements occurring at all scales and locations. These dictionaries do not exhibit highly anisotropic elements. These two limitations of the wavelet transform, i.e., (1) limited directional representation, and (2) isotropic dictionary of bases inspired the vision researchers to propose new transforms that improved directional representation and anisotropicity; such as the steerable pyramids, cortex transforms curvelets [2] and contourlets [5]. steerable pyramids and cortex transforms do not allow for a different number of directions at each scale while achieving nearly critical sampling. this shortcoming is overcome in the curvelet and the contourlet transform. in this paper we will report a comparative study of curvelets and contourlets against the more standard wavelets for computer vision problems.

The curvelet transform [1] was developed initially in the continuous domain via multiscale filtering followed by a block ridgelet transform [3] on each bandpass image. Later, the authors proposed the second generation curvelet transform [2] that was defined directly via frequency partitioning without using the ridgelet transform. Both curvelet constructions require a rotation operation and correspond to a 2-d frequency partition based on the polar coordinate. This makes the curvelet construction simple in the continuous domain but causes the implementation for discrete images sampled on a rectangular grid to be very challenging. In particular, approaching critical sampling seems difficult in such discretized constructions. This fact motivates the development of a directional multiresolution transform like curvelets, but directly in the discrete domain, which resulted in the contourlet construction [5].

Wavelets have enjoyed a widespread exposure in applications of image processing and computer vision. So mush so that wavelets have become a household name today with the introduction of jpeg 2000 the new still picture compression standard that uses wavelet transform. the newer transforms viz. the curvelets and contourlets have enjoyed some applications in image denoising [16] and texture analysis [15], but apart from these the new transforms have enjoyed very limited applications in other areas like compression, super resolution, face recognition and optical character recognition to name a few. Of late, one of the authors has pioneered the application of curvelets and contourlets for face [12, 13] and character [4] recognition problems.

These studies [12, 13, 4] showed good results in terms of recognition accuracy, however they left an important question unanswered, i.e., "why one should one use the new transforms instead of more standard techniques like wavelets since the theoretical explorations into the new transforms do not hint at being better feature sets compared to wavelets?". It is not possible to answer analytically why one feature extraction technique should be better than another as far as computer vision is concrened. So instead, we have decided to tackle this question empirically. In this study we will compare the recognition accuracy of three transforms viz. wavelet, curvelet and contourlet at multiple resolutions as feature sets for face and character images. The purpose of our study is to compare two new transforms viz. the curvelet and contourlet as feature sets for pattern recognition vis-à-vis wavelets which have long been used as feature sets for pattern recognition.

The rest of the paper will be organized into a number of sections. The following section will have a brief description of the entire experimental procedure. In section 3 we will discuss the databases briefly. In section 4 we will tabulate the results. Finally in section 5, conclusion and future scope of work will be discussed.

2. Implementation

The limited scope of the paper does not allow us to delve in to the mathematical construction of curvelets, contourlets or wavelets; neither can we discuss the interesting properties of these transforms. The interested readers are requested to go through the original works on these transforms. what can only be told in short is that the wavelet transform has advantages over the fourier transform since it intrinsically act locally in space (or even in time), by utilizing the besov space that allows to include signals that are generally smooth except for some possible points of discontinuity (i.e., edges). However, a drawback of the wavelet transform is that it cannot recognize smoothness along contour discontinuities and yields ineffective representation. The curvelet transform deals better with contours but unfortunately the required interpolation in the fourier space can be its main drawback for general purpose applications. Also the algorithmic complexity is reasonably high. do and vetterli proposed the pyramidal directional filter-bank, also known as contourlet transform, which offers directionality and anisotropy to image representation that are not supported by wavelet transform. It appears to be more efficient than the curvelet transform. The pyramidal directional filter-bank is constructed by combining the laplacian pyramid and a directional filter bank. The laplacian pyramid is a set of bandpass filters and permits sub-space decomposition, whereas the directional filter-bank allows for a different number of directions at each decomposition level and is designed to capture high frequency directionality of the coefficients.

We carried out the experiments on a 64 bit AMD athlon, running windows XP. The environment we used for programming was matlab 6.5. The wavelets were constructed from the in-built wavelet toolbox in matlab. The curvelet transform was carried out with the curvelab 2.0 [8] toolbox while the contourlet transform was performed making use of the contourlet toolbox [9].

All the above mentioned transforms are multi-scale, multi-directional transforms. However, for object recognition purpose the main emphasis had been on the use of multiple resolutions but not on multiple directions. Multi-directional decompositions are exploited in problems related to texture analysis. In this study we are interested in the recognition ability of these transforms on face and character images. since both these problems fall under the more generic category of problem of object recognition we will be only exploiting the multi-resolution decomposition of the wavelet, curvelet and the contourlet transforms.

In [14] it was shown that the recognition accuracy for facial images does not decrease when the size is reduced to say one-fourth of the original. Guided by this study we reduced the facial images accordingly. Also we converted the facial images to greyscale. The character images were kept intact in size. In the first set of experiments the coarsest approximate coefficients of the wavelet, curvelet and the contourlet transforms formed the feature set. The training images were converted into the transformed domain and the coefficients were used as the prototypes for K-Nearest Neighbour classifiers (KNN). For the testing set, the images were similarly converted to the transformed domain and the approximate coefficients served as the feature set. The test images were fed into the KNN classifier for being classified. In the following set of experiments the same procedure was carried out but at a finer resolution. In subsequent experiments the resolution of the transforms was made finer still.

3. Databases

Our experiments were performed on two face databases from the set of Essex face databases, viz. the face94 [6] and the face95 [7] database. The datasets have both male and female subjects, and have representatives from 4 different races. The subjects are mostly university students but there are some subjects of a higher age range. The images for different individuals were captured by an S-VHS camcorder and stored in 24-bit JPEG format.

While capturing images for the faces94 database, the subjects sit at fixed distance from the camera and were asked to speak, whilst a sequence of images is taken. The speech was used to introduce variation in facial expression.

Figure 1. Samples from faces94 database.

While capturing images for the faces95 database using a fixed camera, a sequence of 20 images per individual was taken. During the sequence the subject takes one step forward towards the camera. This movement is used to introduce significant head (scale) variations between images of same individual. There is about 0.5 seconds gap between successive frames in the sequence.

Figure 2. Samples from faces95 database.

Both the databases had 20 facial images of each person. The first 10 images formed the training set and the remaining 10 the testing set. For the texture recognition we used the USPS database [11, 10] of handwritten numerals. The US Postal Service database (USPS) consists of 9298 handwritten numerals of size 16 x 16 pixel with intensity values varying between zero and two. 7291 samples constitute the training set and the rest 2007 images consist of the testing set.

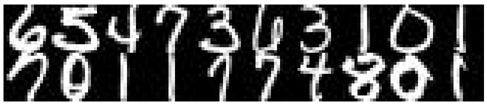


Figure 3. Samples from USPS database.

4. Experimental Results

This is arguably the most important section of our empirical study. The experiments were conducted as had been described earlier. Each of the databases was divided into testing and training sets. The images were transformed and decomposed into several resolutions. At each resolution, the results from the different transforms are compared. For each database the first table will correspond to the finest resolution and the in the following tables the resolution will be made coarser progressively. In all the sets, experiments were done with several values of *k*.

The results on the face94 database are tabulated as shown in Tables 1, 2, 3, and 4 and a similar set of experimental results for the face95 database is tabulated as shown in Tables 5, 6, 7, and 8.

Table 1. Results at finest resolution.

Transform	K=1	K=3	K=5
Contourlet	0.0217	0.0283	0.0375
Curvelet	0.0197	0.0289	0.0362
Wavelet	0.0184	0.0276	0.0362

Table 2. Results at fine resolution.

Transform	K=1	K=3	K=5
Contourlet	0.0211	0.0296	0.0382
Curvelet	0.0145	0.0237	0.0322
Wavelet	0.0086	0.0164	0.0243

Table 3. Results at coarse resolution.

Transform	K=1	K=3	K=5
Contourlet	0.0138	0.023	0.0322
Curvelet	0.0099	0.0171	0.0257
Wavelet	0.0178	0.0283	0.0309

Table 4. Results at coarsest resolution.

Transform	K=1	K=3	K=5
Contourlet	0.0138	0.023	0.0316
Curvelet	0.0092	0.0164	0.0243
Wavelet	0.0138	0.0289	0.0303

Table 5. Results finest resolution.

Transform	K=1	K=3	K=5
Contourlet	0.3861	0.4708	0.5194
Curvelet	0.3736	0.4556	0.5083
Wavelet	0.3708	0.4556	0.5

Table 6. Results at fine resolution.

Transform	K=1	K=3	K=5
Contourlet	0.3778	0.4542	0.4972
Curvelet	0.3403	0.4181	0.4667
Wavelet	0.3222	0.3931	0.4278

Table 7. Results at coarse resolution.

Transform	K=1	K=3	K=5
Contourlet	0.3431	0.4167	0.4736
Curvelet	0.3125	0.3958	0.4389
Wavelet	0.5153	0.5403	0.8917

Table 8. Results at coarsest resolution.

Transform	K=1	K=3	K=5
Contourlet	0.3444	0.4208	0.4736
Curvelet	0.3153	0.3903	0.4389
Wavelet	0.4653	0.5153	0.5889

We did a similar set of experiments for the USPS database. The only difference is that, we considered 3 resolution levels instead of the 4 levels we considered for the facial images. But here we experimented with a larger number of k-values since this database was larger in terms of number of samples.

Table 9. Results at finest resolution.

Transform	K=1	K=3	K=5	K=7	K=9
Contourlet	0.0528	0.0528	0.0528	0.0513	0.0528
Curvelet	0.0493	0.0488	0.0513	0.0523	0.0533
Wavelet	0.0478	0.0493	0.0488	0.0498	0.0548

Table 10. Results at fine resolution.

Transform	K=1	K=3	K=5	K=7	K=9
Contourlet	0.0528	0.0528	0.0528	0.0513	0.0563
Curvelet	0.0493	0.0478	0.0498	0.0498	0.0533
Wavelet	0.0618	0.0608	0.0573	0.0623	0.0628

Table 11. Results at coarsest resolution.

Transform	K=1	K=3	K=5	K=7	K=9
Contourlet	0.0528	0.0528	0.0528	0.0523	0.0563
Curvelet	0.0488	0.0478	0.0498	0.0508	0.0523
Wavelet	0.1126	0.1111	0.1111	0.111	0.1196

				6	

For the facial image databases, the best results from all the three transforms (contourlet, curvelet and wavelet) are always obtained for $k = 1$. but for the usps database the three transforms are individually producing their best results for different values of k. what we mean here is that for contourlets the best results are obtained for $k = 7$; where as for curvelets the corresponding value of $k = 3$ and for wavelets $k = 1$ shows the best results.

Apart from the aforementioned observations on the value of k the experiments reveal certain other interesting aspects. The contourlet coefficients never produce the best recognition results. When the resolution is finer the recognition accuracy from wavelet coefficients are the best. But when the resolution is made coarser the recognition accuracy from the approximate curvelet coefficients increase while those from the approximate wavelet coefficients fall behind curvelets. This behaviour is discernible for both facial and character images.

5. Conclusion

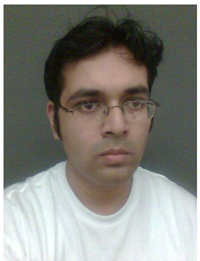
What we can conclude from this research is that at higher resolutions the wavelets serve as a good feature set for both the facial and the character images. But as the resolution is decreased the wavelets become progressively worse as feature descriptors and curvelets score better. But contourlets almost always are the worst feature descriptors, except at very coarse resolutions when the recognition by wavelets worsens to such an extent that contourlets are second best to curvelets. In general, since the recognition accuracy increases as the resolution is made coarse. This phenomenon points to the fact that for problems pertaining to computer vision, only the more prominent edges or discontinuities in the images are the discerning features.

This work is our first attempt to compare the recognition abilities of two new multiresolution multidirectional transforms viz. the curvelet and the contourlet against a well known and widely used transform such as the wavelet. In this work we have tested the recognition capacities of these three transforms only through KNN classification. It is not obvious if the same pattern of results will hold true for other well known classifiers like the artificial neural network or the support vector machine. In the future we wish to extend this work by carrying out the recognition scheme using SVMS and neural nets in place of KNN.

References

[1] Candès E. and Donoho D., *Curvelets: A Surprisingly Effective Nonadaptive*

Representation for Objects with Edges, Stanford University, USA, 1999.
[2] Candes E., Demanet L., Donoho D., and Ying L., "Fast Discrete Curvelet Transforms", *Macromolecular Theory and Simulations*, vol. 5, no .3, pp. 861-899, 2006.
[3] Candes E., *Ridgelets Theory and Applications*, Stanford, USA, 1998.
[4] Chaudhuri B. and Majumdar A., "Curvelet–Based Multi SVM Recognizer for Offline Handwritten Bangla: A Major Indian Script," *in the 7th International Conference on Document Analysis Research*, pp. 647-656, 2007.
[5] Do M. and Vetterli M., "The Contourlet Transform: An Efficient Directional Multiresolution Image Representation", *IEEE Transactions Image on Processing*, vol. 14, no. 12, pp. 2091-2106, 2005.
[6] Computer Vision Science Research Projects, http://cswww.essex.ac.uk/mv/allfaces/faces94.html, 2007.
[7] Computer Vision Science Research Projects, http://cswww.essex.ac.uk/mv/allfaces/faces95.html, 2007.
[8] Curvelet.org, http://www.curvelet.org/download/download.html, 2007.
[9] Image Formation and Processing, http://www.ifp.uiuc.edu/~minhdo/software/contourlet_toolbox.tar, 2003.
[10] Fedora Core Test Page, http://www-stat-class.stanford.edu/~tibs/ElemStatLearn/datasets/zip.test.gz, 2008.
[11] Fedora Core Test Page, http://www-stat-class.stanford.edu/~tibs/ElemStatLearn/datasets/zip.train.gz, 2008
[12] Majumdar A., Ray S., and Bhattacharya A., "Face Recognition by Multiresolution Contourlet Transform on Bit Quantized Facial Images," *Indian International Conference on Artificial Intelligence*, pp. 640-641, 2005.
[13] Mandal T., Majumdar A., and Wu J., "Face Recognition by Curvelet Based Feature Extraction," *in Proceedings of the International Conference on Image Analysis and Recognition*, pp. 806-817, 2007.
[14] Ruiz J., Acosta J., Salazar A., and Jaime R., "Shift Invariant Support Vector Machines Face Recognition System", *Transactions on Engineering, Computing and Technology*, vol. 16, no. 2, pp. 167-171, 2006.
[15] Shutalo L., Shawe J., Fred A., Caelli T., Duin R., Campilho A., and De D., "Texture Classification By Combining Wavelet and Contourlet Features," *Joint IAPR International Workshops SSPR*, pp. 1126-1134, 2004.
[16] Starck J., Candes J., and Donoho L., "The Curvelet Transform for Image Denoising," *IEEE Transactions on Image Processing* , vol. 11, no. 6, pp 670-684, 2002.



Angshul Majumdar received his BSc of engineering in electronics and telecommunication from Bengal Engineering College, Shibpur, India. Currently, he is pursuing his graduate studies at the University of British Columbia, Canada. His research interests are in the application of multiresolution tools to image processing and computer vision.

Arusharka Bhattacharya received his BSc of engineering in information technology from Jadavpur University, India. Currently, he is working at PricewaterhouseCoopers, India. His research interests are in computer vision and machine learning.