We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



7

Multiresolution Methods in Face Recognition

Angshul Majumdar[1] and Rabab K. Ward[2]
[1]*MASc student, Dept of Electrical and Computer Engg, University of British Columbia*
[2]*Professor, Dept of Electrical and Computer Engg, University of British Columbia*
Country

1. Introduction

Wavelets have been a prominent image analysis tool over the past decade. Face recognition researchers use it for varied reasons – pre-processing, compression and feature extraction. We refer the reader to [1] for a good review on the theory and applications of wavelets in face recognition. In this chapter we will concentrate on some new transforms that have emerged from the limitations in wavelets. First, we will outline the limitations of wavelets and show how the new image analysis tools overcome them. Next, we review some of the existing work in face recognition that has benefited from using these tools. Finally, we show how these new tools fit into the larger, newly developing arena of signal processing known as Compressive Sampling or Compressed Sensing (CS). We outline how CS can be used for face recognition which certainly will be a new direction in the field of face recognition.

2. The new image analysis tools

The literal meaning of the word 'transform' is 'change'. When we speak of an 'image transform', we refer to an alternate (changed) way of representing an image. Generally an image is represented in the spatial domain by pixels, but there are alternate representations, the most popular being the frequency domain representation obtained by the Fourier transform. However, the Fourier transform of an image is not very informative from the perspective of object recognition. Other transforms like wavelets, curvelets etc. provide alternative image representations (other than pixels or frequency). These transforms represent images in such a way that recognition is facilitated.

In [2], the five 'must-needed' properties of a multiresolution image transform are pointed out:

1. Multiresolution. The transform should allow images to be successively approximated, from coarse to fine resolutions.
2. Localization. The basis elements of the transforms should be localized in both the spatial and the frequency domains.
3. Critical sampling. For some applications (e.g., compression), the transforms should form a basis, or a frame with small redundancy. We will discuss more about it in section 2.2.
4. Directionality. The transform should contain basis elements oriented at a variety of directions, compared with the few directions offered by separable wavelets.

Source: Recent Advances in Face Recognition, Book edited by: Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett,
ISBN 978-953-7619-34-3, pp. 236, December 2008, I-Tech, Vienna, Austria

Open Access Database www.intechweb.org

5. Anisotropy. When a physical property changes with direction, that property is said to be anisotropic, e.g. in certain crystals the conductivity of heat is more in horizontal direction than in vertical direction; for such a case, the conductivity is said to be anisotropic. For image transforms, anisotropicity means that the basis elements of the transforms should not be circular (similar in all directions) but may be elliptical (more along the major axis and less along the minor axis).

From the above wish-list, the first three properties are provided by separable wavelets. To cover all the five properties, new transforms are required. This deficit inspired researchers to search for new methods for image representations, including curvelets, contourlets and surfacelets which we discuss below.

2.1 Motivation

In a seminal publication [3], it was shown that the widely accepted belief that an adaptive representation that in some sense 'tracks' the shape of the discontinuity of an image can be used to efficiently represent an otherwise smooth object with discontinuities such as edges, is incorrect. When the geometry of the object is known a priori, an ideal adaptive representation satisfies

$$\| f - \hat{f}_m \| \propto m^{-2}$$

where f and $\hat{f}_m$ are the image and its approximation using the m largest valued coefficients respectively. The error between the actual object (*f*) and the m-largest terms approximation ($\hat{f}_m$) is reduced quadratically in the number of terms. In practical scenarios, the geometry of the object is never known beforehand, so adaptive schemes like wavelets do not reach the ideal rate. The m-largest term approximation for wavelets satisfies

$$\| f - \hat{f}_m^W \| \propto m^{-1}.$$

This convergence rate is an order of magnitude smaller than the ideal rate. In [3] the authors proposed a non-adaptive (fixed) curvelet scheme which reached the ideal convergence rate asymptotically

$$\| f - \hat{f}_m^C \| \propto m^{-2} (\log m)^3$$

The crux of the above discussion is that, for an image, the discontinuities (edges) will be better approximated by curvelets than wavelets, i.e. to capture the edge information fewer curvelet coefficients are needed than wavelet coefficients. Looked at it differently, the same number of curvelet coefficients contain more edge information compared to wavelet coefficients. Since object recognition is driven by edge information, the more efficiently we can use this information, the better the recognition.

The curvelet transform was introduced in 1999 but its application was scarce till 2006. Following similar ideas to curvelets, the contourlet transform [2] was introduced in 2005. The contourlet transform was defined directly in the digital domain (while curvelets were defined in the continuous domain) and can be implemented efficiently through filter-banks. However it lacks the rich operator theory of curvelets and hence is not easily analyzed. The simpler and faster version of curvelets seen today was introduced in 2006 [4]. The problem

with curvelets is that they are over-complete and suffer from redundancy (redundancy factor is ~7.2 for 2D, and ~24 for 3D). Moreover, although theoretically it is possible to extend the curvelets to higher dimensions (more than 3), practically it has never been achieved. In 2007, surfacelets, a new transform that can be readily extended to higher dimensions was introduced [5]. Surfacelets are based on the filter-bank representation of contourlets.

Face recognition is mainly carried out using still images or video sequences, i.e. through 2D information. Thus we will discuss the 2D curvelet transform and the contourlet transform (surfacelets can be looked upon as an efficient higher-dimensional extension of contourlets).

2.2 Curvelet transform

Since curvelet transform is the pioneer of all the non-adaptive transforms available today, we will discuss it in a more detail. Once the reader understands the basics of curvelets, it will be easier to understand contourlets.

The core of the curvelet transform [4] is displayed in Fig. 1.

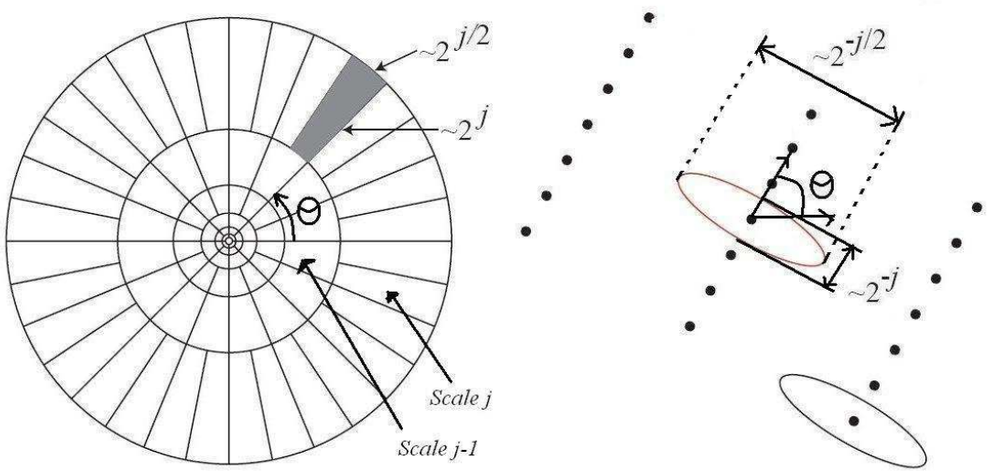


Fig. 1. Curvelet Transform: Fourier Frequency Domain partitioning (left) and Spatial Domain representation of a wedge (right) [4]

In Fig. 1 on the left, the Fourier plane is divided into wedges. The shaded area is one of the wedges. The wedge is formed by partitioning the frequency plane into radial and angular divisions. The radial divisions (concentric circles) are for band-passing the image at different resolution/scales. The angular divisions divide each bandpassed image into different angles. Therefore when we consider each wedge (say the shaded one), we are actually analyzing the bandpassed image at scale j and angle θ.

Property 2 of the aforesaid wish-list requires the transform to be localized both in the frequency and the spatial domain. However, a signal that is perfectly localized in one domain is spread out in the other. So, one can only expect the transform to be approximately localized in both domains. The image on the left of Fig. 1 shows how the Fourier plane is divided into wedges. If a wedge has an abrupt boundary in the frequency domain, it will be spread in the spatial domain. To avoid that, the boundary of the wedge is tapered as shown in Fig. 2.

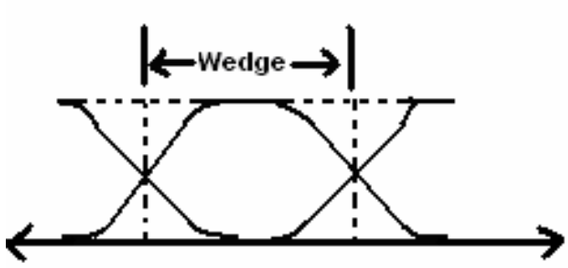


Fig. 2. Boundary of Wedge

The dotted line shows an abrupt wedge boundary, the continuous lines show the actual wedge boundary which is tapered off. This smooth tapering allows for localization in both the frequency and spatial domains.

Let us return to Fig. 1. Once the scale and the angle are defined the wedge (shaded region) is identified. The wedge is inverted to the spatial domain (right side of Fig. 1) by Inverse Fourier transform. The inverse Fourier transform of the wedge are the curvelets corresponding to the wedge (i.e. a particular scale and angle). The curvelets are periodic and repeated infinitely. In Fig. 1 they are shown as ellipses, the centres of the ellipses are shown as dots.

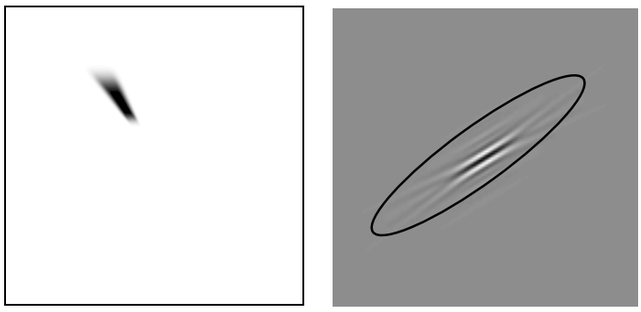


Fig. 3. Fourier Transform (left) and Curvelet in spatial domain (right) [4]

Fig. 3 shows an actual wedge (in the frequency domain) the left, and a curvelet (in spatial domain) corresponding to the wedge on the right. As we can see, the curvelet is not exactly an ellipse as depicted in Fig.1. It is elongated in one direction and wave-like in the other but its effective support is elliptical as shown in the right of Fig. 1 and 3. The relationship between the length of major and minor axes of the ellipse follow a parabolic scaling law, i.e. major axis ≈ minor axis^2. For wavelets, the corresponding shapes would be circular, because they are isotropic.

The curvelet shapes for all images are the same (elliptical). However, the values of the curvelet coefficients are determined by how much the curvelets and the actual image are aligned.

Fig. 4 shows how the curvelet coefficients are determined. The top left shows the image of an arbitrary object. The top right image shows the ideal band-passed image with prominent edges (at a particular concentric ring in Fig. 1). The bottom image shows some curvelets in the spatial domain corresponding to wedges at different orientations i.e. different wedges in the same concentric ring of Fig. 1. In this case, the values of the curvelets that are aligned with the edges will have high values e.g. curvelet c in Fig. 4. Curvelets that are not aligned to the bandpassed image i.e. a and b have very small coefficients. The greater the 'edginess'

of the image at a particular location, scale and orientation is, the higher the values of the curvelet coefficients. We will not discuss the exact algorithms for finding the curvelet coefficients. The interested reader can refer to the original paper [4].

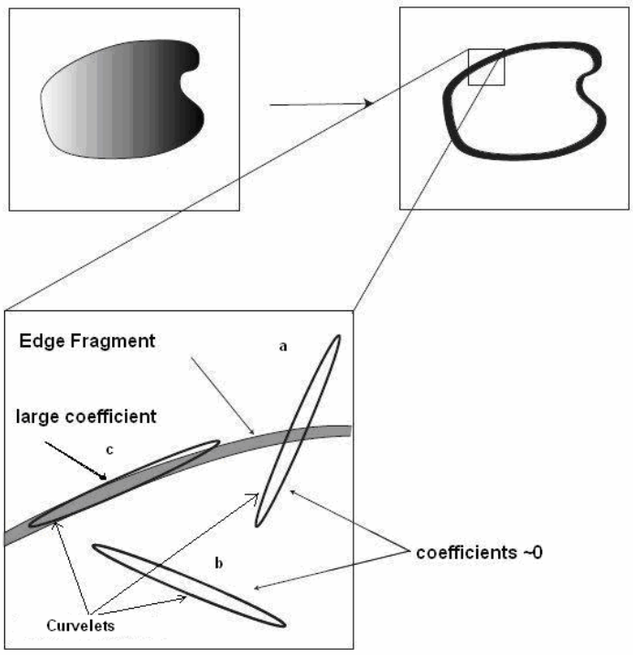


Fig. 4. Curvelet Alignment [6]

In the previous paragraph we qualitatively describe what the curvelet coefficients actually mean. This description applies to other transforms as well. For example, the wavelet representation in the frequency domain is not divided into angular wedges. The Fourier frequency plane is only divided into concentric rings corresponding to different scales. The spatial representation of such concentric rings is circles. When these circles are aligned with the bandpassed image the values of the coefficients is high, otherwise the values are low.

All the five properties of the wish-list are covered by the curvelet transform. Property 1, i.e. the multiresolution property of the curvelet transform, is easily discernible from Fig. 1. In the first step of the transform, the entire frequency plane is divided into concentric rings in order to facilitate bandpassing at multiple resolutions. Fig. 3. refers to property 2, it shows the actual localization related to the curvelets in both spatial and frequency domain. Curvelets do not have a compact support in the spatial domain, but the effective support is elliptical and can be compared with wavelets having infinite vanishing moments. As for property 3 of the wish-list (related to critical sampling) curvelets do not form a basis, but form tight frames[1]. Unlike wavelets, the curvelet transform is redundant having a

[1] For a basis, e.g. wavelets, we have $WW^T = W^TW$ = Identity; where W and W^T are the forward and inverse wavelet transform matrices respectively. For tight frames like Curvelets, C^TC = Identity but $CC^T \neq$ Identity; i.e. the forward transform followed by the inverse transform gets back to the original domain, but the reverse order (inverse transform followed by forward transform) does not.

redundancy factor around 7.2. The directionality of the curvelet transform (property 4) is easily discerned in Fig. 1. The entire frequency plane is divided into angular wedges. This allows for analyzing the image at various orientations. As for the last property of the wish-list, from the perspective of image processing, a transform is said to be anisotropic if its basis elements are not circular. Curvelets are elliptical in shape (the major and the minor axes related by a parabolic scaling law), thus they are anisotropic.

2.3 Contourlet transform

So far, our discussion on curvelets was completely for continuous signals. The curvelet literature defines its concepts in this domain and digitizes it during implementation. The algorithms for digital implementations are quite involved [4]. To overcome this (and also for combating the redundancy issues) researchers defined the contourlet transform directly in the digital domain [2]. The contourlet keeps all the desirable properties of curvelets including directionality and anisotropicity and at the same time reduces the redundancy of the curvelets. However, the contourlet transform does not follow the nice properties of operator theory and hence can not be easily analyzed.

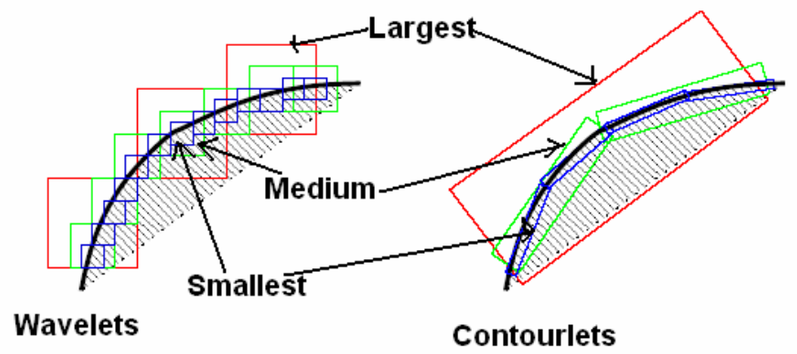


Fig. 5. Wavelets (left) vs Contourlets (right) [2]

As mentioned earlier, the coefficients of the curvelet transform are decided depending on how the curvelets align with the edge at a particular location, scale and orientation; the same argument is in fact true for all other transforms including wavelets and contourlets. In Fig. 5 the shaded area is an image of a portion of an arbitrary object where the blue boundary shows an image edge. Fig. 5 (left) shows how wavelets align with the edge, while the right shows alignment of contourlets with the edge. As mentioned above, in the spatial domain the support of wavelets is circular. In a digital domain the support of wavelets is approximated to be square. To analyze the image by wavelets, we examine how the wavelets align along the edge.

Fig. 5 (left) shows how wavelets arrange themsleves along the edge at different resolutions. The small blue squares represent the wavelets at the finest resolution, the green squares represent intermediate resolution and the red squares represent wavelets at the coarsest resolution. Fig. 5 on the right shows the alignment of contourlets. Notice that the squares are replaced by rectangles. We can see that at each resolution, the edge can be represented by a far less number of contourlets than wavelets. As Wavelets are isotropic they can not take advantage of the underlying geometry of the edge. They approximate the edge as a collection of dots (small squares) so many points are needed to represent an edge, but contourlets are representing the edge as a collection of small needles hence only a few needle shaped line segments can represent the edge).

Very loosely speaking, one contourlet in Fig. 5 may be assumed to be formed by grouping several wavelets at the same resolution. For example at the finest resolution, 4 or 5 consecutive wavelets may be grouped together to form a contourlet, at the next highest resolution, about 3 consecutive wavelets can be grouped to form a contourlet and so on. Based on this simple assumption the contourlet algorithm can be written as:

1. Apply a scale-space decomposition on the image, i.e. find the Laplacian of the image at several scales.
2. At each scale, represent the Laplacian of the image by wavelet elements (small squares as shown in Fig. 5).
3. At each scale, group the wavelet elements in a particular direction to form contourlets.

The filter-bank representation of the above three steps is shown in Fig. 6

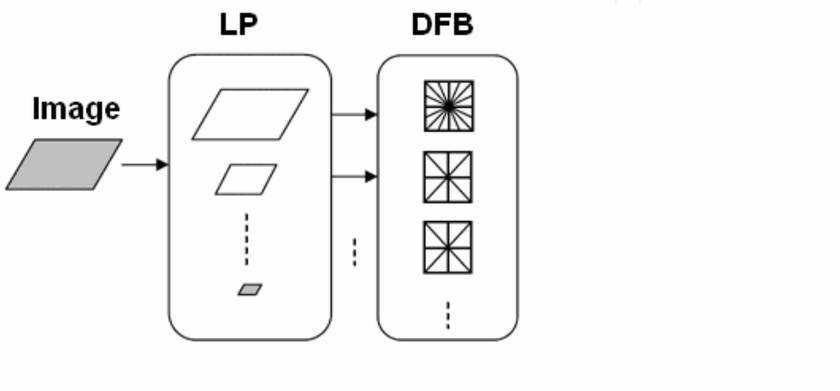


Fig. 6. Flow Filter-bank representation for Contourlets [7]

The image first undergoes a wavelet-like decomposition by a Laplacian Pyramid (LP). In the LP block of Fig. 6, the largest parallelograms represent the decomposition at the finest scale (corresponding to blue squares in Fig. 5). The smaller parallelograms of the LP denote coarser decompositions. Once the image is decomposed into different scales, they are represented by small square wavelet elements. The Directional Filter Bank (DFB) groups the wavelet elements at each scale and groups them into different orientation to form contourlets. The DFB block of Fig. 6 shows that, at the finest scale (corresponding to the largest parallelogram of LP block), DFB decomposes the image into the maximum number of orientations. As the resolution becomes coarser, DFB decomposes the image into a lesser number of orientations.

Although this contourlet implementation is simple to understand, it is not robust to noise; this version of the contourlet transform is implemented in [7]. The original implementation of the contourlet transform which is more robust is described in [2]. From this discussion, it is not clear how the redundancy is reduced in the contourlet implementation (compared to curvelets). We refer the interested reader to the original work in contourlets [2] to pursue the issue. Coarsely speaking, we can say that the contourlet transform uses a smart sub-Nyquist sampling scheme to check the redundancy.

3. Review of previous work

In a classic experiment reported in Nature [8], Field and Olshausen set up a computer experiment for empirically discovering the best representation for a database of 16 by 16 image patches. Although this experiment is limited in scale, they discovered that the best

way (in the sense of facilitating vision) to represent these image patches is a collection of needle shaped filters occurring at various scales, locations and orientations. These filters are shown in Fig. 7.

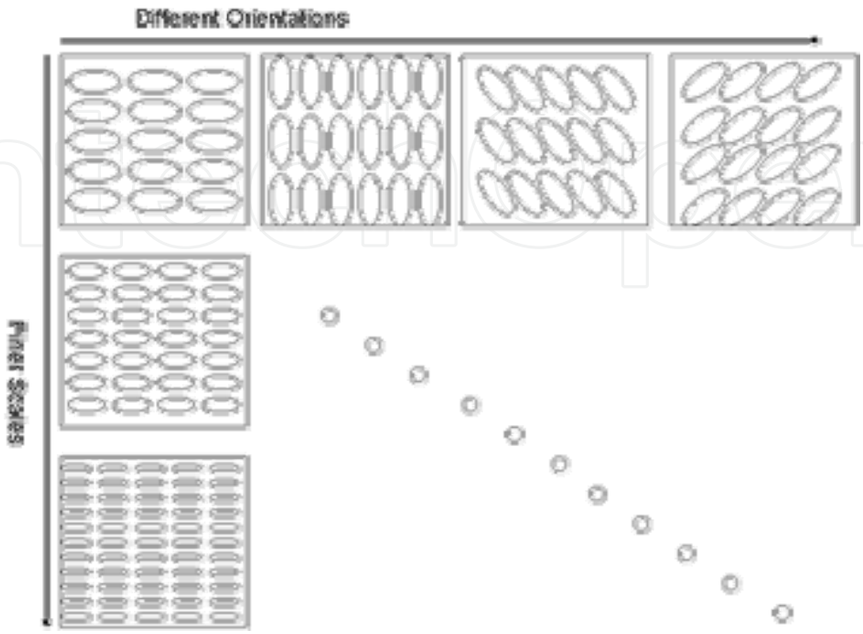


Fig. 7. Field and Olshausen's Filters

In Fig. 7, the first row shows the filters corresponding to the same scale but different orientations in the spatial domain and in first columns shows the filters at the same orientation but different scales. If an edge falls within a filter, the response of the filter is high. Fig. 8 shows the filter response of the filters at different scales for an arbitrary object.

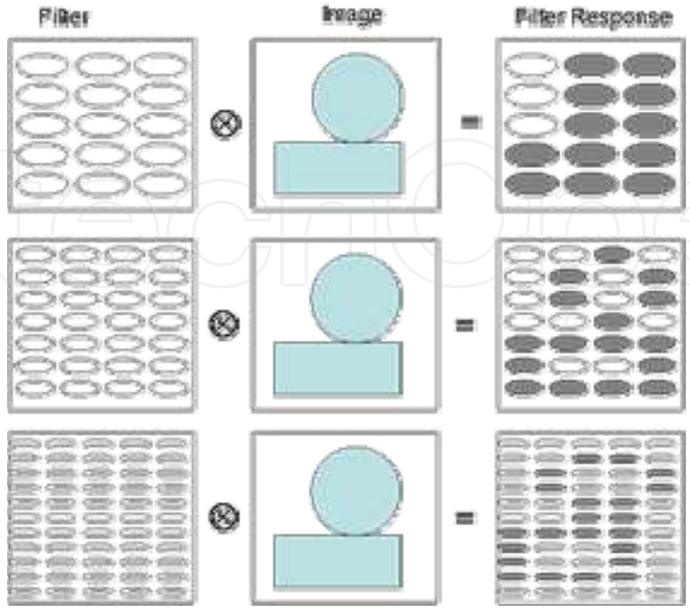


Fig. 8. Convolution of the filter-bank with an image

The researchers [8] found that, the response of such filter-banks facilitate object recognition. The stark similarity between curvelets (Fig. 1, 3 and 4), derived from mathematical analysis, and this empirical filter-bank arising from experiments (Fig. 7 and 8), is pointed out in [9].
The classic experiment [8] discovered that such filter responses (encoding edge information) were very suitable for computer vision tasks. Besides computer vision, human vision is also driven to a large extent by edge information. Neuro-scientists showed that edge processing neurons form the earliest and most fundamental stages of the pipeline upon which mammalian visual processing is built [10].
Face Recognition and other computer vision tasks in general are dependent on edge information. Successful multiresolution transforms mainly try to capture this information. Curvelets and contourlets capture the edge information more efficiently than wavelets. Once the implementation of these transforms were known, researchers in face recognition tried to use them to address different problems, but since the transforms are very new, their full potential is not yet completely achieved.

3.1 Comparative works

In [11] the author compared wavelets and curvelets as feature sets for face recognition. Curvelet and wavelet transform were applied to the images. The approximate coefficients i.e the non-directional transform coefficients (those corresponding to the innermost circle of Fig. 1) were used as features for recognition. It was found empirically that curvelets were better than wavelets for face recognition. A similar comparative empirical study, but between contourlets and wavelets [12] however showed that contourlets were not as good as wavelets for recognition.
A comprehensive empirical study comparing curvelets, contourlets and wavelets [13] decomposed the face images to several resolutions from fine to coarse using the three transforms. At each resolution the recognition accuracy of each of the three transforms were compared. An interesting observation was that the recognition accuracy from wavelets decreased as the scale was made coarser, where as for curvelets and contourlets the accuracy increased.
The studies [11 and 12] do not contradict [13] each other. In [11 and 12] the curvelet, wavelet and the contourlet transform was taken at the finest resolution. The studies showed that curvelet coefficients are the best for face recognition, followed by wavelet coefficients and lastly the contourlet coefficients. However [13] revealed that at coarser resolutions both curvelets and contourlets show better recognition accuracy than wavelets.

3.2 Feature extraction

Researchers have started using curvelet coefficients are used as features for face recognition only in the last few years [14-17]. The simplest application of using curvelets in face recognition can be found in [14]. The face images undergo a curvelet transform, and the approximate curvelet coefficients are used in a Nearest Neighbour classifier for recognition.
Face image is a 2D projection of the actual 3D human face. The 3D information is encoded as change in shade in a 2D image. Researchers in [15], wanted to capture some of this information from the face images. The number of bits in the original 8 bit images was quantized to 4 bit and 2 bit versions. 8 bit images with 256 levels of grayscale allow finer variations in shading. The variations in shading arising out of bit-quantization can be seen from Fig. 9. The 8 bit image shows fine variations in depth resulting in fine edges. The 4 bit image loses some of the finer depth information. The 2 bit image only shows the boldest

edges corresponding to large variations in depth. Curvelet transform of the different bit quantized versions encode varying edge information. The curvelet coefficients resulting out of these three versions of the same image were used as features for recognition. This approach was based on the assumption that different types of edge information (finer to coarser) will enhance the recognition accuracy. The recognition results obtained in [15] were better than those obtained in [14].



Fig. 9. 8 bit, 4 bit and 2 bit representations (from left to right) [15]

The work in [15] did not use the multiresolution properties of the curvelet transform. The transform was applied to bit-quantized images at a single resolution. Studies following the previous one [15] overcame this shortcoming [16, 17]. Each of the bit-quantized face images were transformed at three resolutions. For three levels of bit-quantization there were a total of nine sets of transform coefficients. These transform coefficients were used by separate classifiers for recognition. Finally the results from all the classifiers were fused using majority voting rule. In [16] the contourlet coefficients were used, and in [17] the same was done using curvelet coefficients. The results in [16, 17] showed that, curvelet and contourlet based schemes were better than wavelet based recognition schemes. The results also showed improvement over the previous approach [15].

The idea of using coefficients at multiple resolutions was probably borrowed from handwriting recognition literature [18-20], where the fine details and the coarse shape of the image are assumed to offer complementary information. The wavelet transform of handwritten characters at different resolutions were used to train neural networks. The outputs from the different neural nets were combined using majority voting. In this manner, very high recognition accuracy on handwritten characters was achieved. In [16, 17] the same idea was implemented on face images, but using curvelet and contourlet transforms instead of wavelets.

Subspace techniques like Eigenface and Fisherface are popular methods in face recognition. Experimental work on subspace techniques in the curvelet domain was carried out in [21]. The Eigenface and Fisherface techniques were applied on the curvelet coefficients. It was shown that such methods applied in the curvelet domain gave better recognition results compared to subspace methods in pixel domain and wavelet domain [22].

The idea of using Principal Component Analysis (PCA) followed by Linear Discriminant Analysis (LDA) for dimensionality reduction of face images is proposed in [23, 24]. The rationale behind this approach is somewhat eluding, but it was empirically shown that better recognition results are obtained by this combination rather than by PCA or LDA alone. The combined approach of PCA and LDA in the curvelet and in the wavelet domain is carried out in [21]. In this work, several standard subspace face recognition approaches are compared, and it is shown that the curvelet based ones always gave better results.

A challenging area in face recognition arises when there is only one image available for training. A recent work in this area used multiresolution transform for addressing this problem [25]. When pixel values or transform coefficients are used to represent a face image, the input vector is very high dimensional. This high dimensional space is called the face-space. In the aforesaid work, it is assumed that, there is a mapping from the high dimensional face-space to a lower dimensional discriminative subspace. If for a certain set of people, such a mapping can be found (by training) then this mapping is assumed be applicable to another set of people as well. The Fisher Linear Discriminant (FLD) was used for such a mapping in [25]. Instead of using pixel values to represent the high dimensional face-space, the image transform coefficients (wavelet/curvelet/contourlet) were used. The application of FLD mapping on the transform coefficients showed better results than previous discriminative subspace based approaches addressing the same problem.
For tackling the problem of having only one single training image for each person, another approach is proposed in [26]. In this paper new face images with varying edge information are generated from the single available image. Since vision is largely edge-driven, it is assumed that varying the edge information will synthesize new image samples that can be used for recognition by standard machine learning tools. In [26] after applying the curvelet transform to the image, a non-linear approximation algorithm is used threshold the transform coefficients. By varying the number of thresholded coefficients, slightly different versions of the original image are obtained. As thresholding the transform coefficients results in the loss of some edge information, the edge information in the synthesized image is thus different from the original one. The results in [26] show better recognition accuracy compared to other algorithms that generate new face images from a single available one in order to address the problem of identifying a person from a single available training image. It also shows better recognition results when compared to [25].

3.3 Pre-processing

The contourlet transform has been used to denoise face images [27]. The main problem in [27] was to transmit face images through wireless channels, and see how the channel affects recognition performance. The contourlet transform was used for denoising the received images. Several thresholding schemes such as Hard Thresholding, Soft Thresholding and Stein's Thresholding were tested. Stein's thresholding gave the best results. However, better denoising schemes other than thresholding are available today – Iterative Soft Thresholding, Basis Pursuit Denoising, Stagewise Orthogonal Matching Pursuit etc. to name a few.
A follow-up of [27] used the same contourlet based denoising scheme for face recognition, but the machine learning algorithm was different [28]. A sophisticated Active Learning framework called Swarm Intelligence was used to recognize faces.

4. Dimensionality reduction: random projection and compressive sampling

Face images are very high dimensional when represented as pixel values or transform coefficients. This makes the classification phase time consuming. In order to reduce the time during classification, a dimensionality reduction/feature extraction algorithm is usually applied to the face image. Standard dimensionality reduction algorithms are Eigenface (using PCA) and Fisherface (using LDA). However, these deterministic dimensionality reduction methods are computationally very complex and consequently time consuming. To

reduce the computational complexity, several studies [29-32] applied Random Projections (RP) for dimensionality reduction. In RP, the original high-dimensional data is projected onto a low dimensional subspace using a random matrix whose columns have unit norm. However, RP is instable. In this section we will show how RP can be used to make stable projections of face images.

There are two advantages of RP over conventional dimensionality reduction schemes like PCA and LDA:

- The conventional dimensionality reduction methods are data-dependent. The low dimensional components are chosen from the high dimensional input data by optimizing certain criterion e.g. maximizing variance in PCA. Random projection matrices are independent of the input data and are very easy to compute, they can be constructed just by normalizing the columns of a random matrix.
- RP is computationally simple and efficient to implement. Moreover theoretical results show that such projections approximately preserve pair-wise distances of points in Euclidean space. They also preserve volumes, affine distances and the structure of the data. Such properties make RP especially suitable for machine learning.

There is a problem however with RP and that the projections are not stable. A common practice to overcome this problem is to use multiple random projections so that the low dimensional projections of the input data form many random subspaces. Recognition is carried out in each of these low dimensional sub-spaces. Finally, the results from these multiple random subspaces are fused to arrive at the final decision. This approach is ad hoc. There is no theoretical framework for deciding how many random projections are required to stabilize the results. Moreover making multiple projections add to the computational complexity of the approach.

4.1 Choosing the subspace

In the Eigenface method, the number of principal components is chosen such that most of the energy of the original face image is retained. Formally, we chose the projection matrix P so that the energy is preserved

$$\frac{\| Px \|}{\| x \|} \approx 1 \tag{1}$$

where x is the original face image.

Unfortunately it is not possible to find a random projection matrix P, so that the above criterion is satisfied. Stability of projection is closely related to its energy conservation. Energy is approximately conserved in PCA projection (Eigenface method), but for RP the ratio (1) is much greater than unity, consequently the RP projections are instable.

In the next sub-section, we will show that an emerging paradigm in signal processing, called Compressive Sampling or Compressed Sensing (either way it is CS in short) attain energy conservation and and keep the advantages of RP. The solution is computationally fast and the recognition results are stable.

4.2 CS dimensionality reduction

The Uniform Uncertainty Principle (UUP) of the CS literature states that for a sparse signal x there exist some matrices $\mathbf{\Psi}$ such that the energy of x and that of Ψx are nearly the same; i.e.

$||\Psi x||\approx||x||$. This approximation is quantified by the "restricted orthonormality hypothesis" [33] which states that

$$(1-\delta_x)\,\|x\|^2 \leq \|\Psi x\| \leq (1+\delta_x)\,\|x\|^2 \tag{2}$$

δ_x (<< 1)is the restricted orthonormality constant.
Finding such a projection matrix Ψ deterministically is an NP hard problem. The CS literature however proves that, any random projection matrix will satisfy the restricted orthonormality hypothesis (2) with a very high probability.
When the signal is sparse, the restricted orthonormality hypothesis ensures that RP satisfies the energy conservation criterion approximately. This makes RP a stable dimensionality reduction technique for sparse signals.
The words 'restricted isometry', 'approximate isometry' comes up in compressive sampling frequently. Equation (2) says that the energy of a sparse vector x is conserved after it is projected by Ψ. This shows the near orthogonality of RP matrix with sparse vectors. CS literature provides even stronger property of Ψ and states that Ψ behaves as an identity matrix (isometry), and all the information content of the sparse signal x is preserved in the lower dimensional projection Ψx. The isometry is a stronger property compared to the orthonormality.
Equation (2) can be re-written as

$$\frac{\|\Psi x\|}{\|x\|} \approx 1 \pm \delta_x \tag{3}$$

Equations (1) and (3) are almost the same, i.e. when x is a sparse vector almost all the energy in the original signal is conserved by both RP (3) and the Eigenspace projection (1). This relation does not hold for dense signals. In the following table we show the simulation results we carried out to give the reader a feel of the numerical values involved in the discussion.

Length of signal	Number of non-zero coefficients		Length of Ψx	$\|\Psi x\|/\|x\|$ for sparse vector	$\|\Psi x\|/\|x\|$ for dense vector
	Sparse	Dense			
1000	50	1000	100	1.0056	3.5074
			200	0.9982	2.2744
			400	1.0022	4.0653
			800	0.9997	4.2277
	100		200	1.0034	2.7935
			400	1.0006	5.1309
			800	0.9999	3.9024
10000	500	10000	1000	0.9965	4.4473
			2000	1.0005	4.2946
	1000		2000	0.9994	3.9333

Table 1. Ratio between norms after and before projection

The results in column 5 in Table 1 were generated by creating sparse signals. The positions and amplitudes of the non-zero elements in the sparse signal were at random. An i.i.d

Gaussian matrix, with the columns normalized to unity, was used as the projection matrix. For the results in the last column of Table 1, the projection matrix remained the same, but the signal was dense. The dense signal was generated at random. The results shown in Table 1, are obtained by averaging the results of 100 test runs.

Table 1, is an experimental verification of compressive sampling. It demonstrates the theoretical result that for sparse vectors, the RP behaves as a restricted isometry.

We conclude that, if the input vector is sparse, we could rely on RP for dimensionality reduction. This dimensionality reduction has two advantages

1. RP is stable because the energy is conserved during projection. For dense vectors this condition was satisfied by PCA but not by RP.
2. RP is computationally efficient. The computational cost for PCA is $O(n^3) + O(nd)$, while for RP is it only $O(nd)$, where n denotes the length of the vector and d is the number of projections.

But, for our signal, the input face image is not sparse.

We now relate how the multiresolution analysis methods are related to CS (and RP) dimensionality reduction. Face images are not sparse in the pixel domain. However, their representation using the multiresolution transforms like wavelets, curvelets and contourlets are sparse. The following figure shows the sparsity of these transforms.

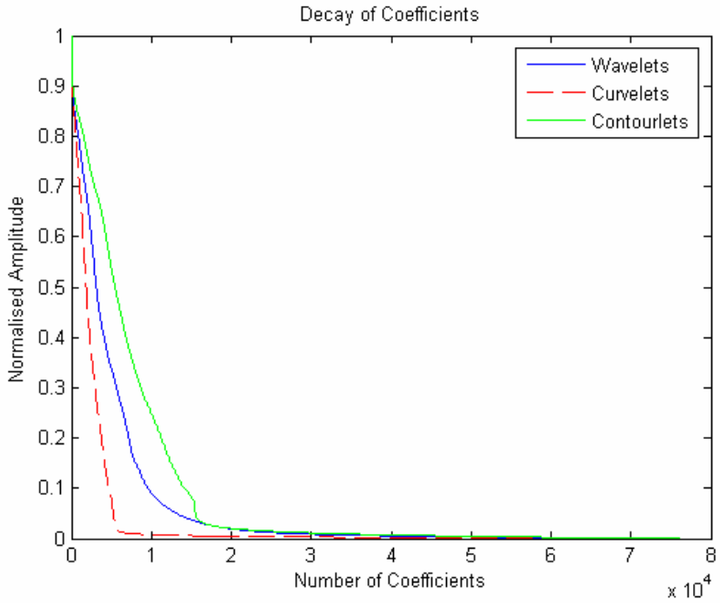


Fig. 10. Decay of Transform coefficients

In Fig. 10, in order to see how sparse the coefficients of this transforms are, the amplitude of the coefficients are normalized and plotted in descending order. The sparser the result of a transform is, the lesser is the number of high valued coefficients. Fig. 10 shows the decay of the transform coefficients. The results are from the ORL face database and are averaged over 400 images in the database. In the pixel domain each face images were of 256X256 resolution leading to 65,536 pixel values. From the figure, we can see that curvelets give the sparsest representation followed by wavelets and lastly contourlets.

A sparse representation of the image *I*, *x* can be obtained by a linear operation i.e. $x = \Phi I$, where Φ, can be any of the aforesaid three transforms.

We now quantify the limits of dimensionality reduction. Compressive Sampling (CS) states that if an n dimensional signal is s sparse then m projections are required where,

$$m = C \cdot s \cdot \mu_{\Phi,\Psi}^{2} \cdot \log(n) \tag{4}$$

C is a constant (between 2 and 5) and $\mu_{\Phi,\Psi}$ ($\in [1, \sqrt{n}]$) is the measure of coherence between the sparsifying transform Φ and the random sampling matrix Ψ.
From equation (4), it is deduced that smaller the μ i.e. the coherence between Φ and Ψ, lower will is the number of projections (dimensionality) m. In [33] it is shown that for Gaussian projection matrices are minimally coherent with wavelet, curvelet or contourlet basis. Hence whatever be the choice of Φ (wavelet/curvelet/contourlet), Ψ can be a Gaussian random matrix with columns normalized to unity. Other projection matrices like random permutations of the Fourier matrix and Bernoulli matrices (whose entries are formed by randomly placed 1 and -1) can be used instead of Gaussian projection matrix, but Gaussian projection matrices have been proven to be minimally coherent.
Equation (4) is derived for signal reconstruction and not for recognition. For the case of face recognition, the estimate m is pessimistic, i.e. we can afford to be greedy and can get good recognition accuracy with far less number of projections than indicated by (4). A study [34] showed that good recognition accuracy can be achieved by CS dimensionality reduction, but with far less number of samples that suggested by (4). However, there are no theoretical bounds for deciding the number of CS projections needed for recognition.

4.3 Some results

To apply CS dimensionality reduction in face recognition we need to follow three steps:
1. Sparsify the images using wavelets, curvelets or contourlets.
2. Create random projection matrices. These can be Gaussian, Bernoulli or Fourier.
3. Use the random projections of the sparse vectors as features in the classification stage.

In the following table, we show the results for CS dimensionality reduction and normal RP dimensionality reduction. A new classification framework proposed in [32] is used.

Method	Number of Projections	Dimensionality Reduction Ratio					
		1:1	2:1	4:1	8:1	16:1	32:1
RP applied to original image	1	0.08	0.09	0.095	0.085	0.145	0.21
	3	-	0.085	0.08	0.085	0.075	0.125
	5	-	0.085	0.085	0.085	0.075	0.105
RP applied to sparse wavelet coefficients	1	0.08	0.08	0.08	0.085	0.085	0.095

Table 2. Error Rates on ORL Face Database

The first column denotes the type of dimensionality reduction method used. The second column of the table shows, how many RPs were taken. As mentioned earlier RP on dense inputs results in instability. To counter the instability, multiple random projections need to be taken. RP applied to sparse vectors (wavelet coefficients) is stable and does not require multiple projections. The rest of the columns (from third to last) shows error rates for different dimensionality reduction ratios.

5. Conclusion

This chapter starts with discussion on curvelets and contourlets. These transforms are new and the research studies that make use of these transforms provide a formal mathematical description. This chapter gives a qualitative intuitive understanding of these transforms. Readers from areas of image processing other than pattern recognition may also find this section interesting. In section 2, a complete up-to-date summary of the different applications of these transforms in the face recognition area is provided. Current research in face recognition is yet to utilize the full potential of these transforms. Previous work only used the approximate transform coefficients of the transforms. The rich orientation information of the detail coefficients has yet to be used in feature extraction.

In the final section we relate state-of-the-art signal processing techniques to traditional feature extraction methods. Compressive Sampling is a new paradigm in signal processing with many promising prospects [35]. Current literature in CS mainly deals with signal reconstruction. CS's potential in signal recognition has not yet been studied. Investigating CS techniques in face recognition will be an exciting area of research.

6. References

D. Q. Dai and H. Yan, "Wavelets and Face Recognition" in Face Recognition, K. Delac and M. Grgic ed., I-TECH Education and Publishing, Vienna, 2007.

M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation", IEEE Transactions Image on Processing, Vol. 14 (12), pp. 2091-2106, 2005.

E. J. Candès and D. L. Donoho. Curvelets - a surprisingly effective nonadaptive representation for objects with edges. Curves and Surfaces, L. L. Schumaker et al. (eds), Vanderbilt University Press, Nashville, TN

E. J. Candès, L. Demanet, D. L. Donoho and L. Ying, "Fast discrete curvelet transforms", Multiscale Model. Simul., Vol. 5 pp. 861-899, 2006.

Y. M. Lu and M. N. Do, "Multidimensional directional filter banks and surfacelets", IEEE Transactions on Image Processing, Vol. 16 (4), pp. 918-931, Apr. 2007.

E.J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise-C2 singularities", Comm. Pure Appl. Math., Vol. 57, pp. 219-266, 2002.

R. Eslami and H. Radha, "The Contourlet Transform for Image De-noising Using Cycle Spinning", Signals, Thirty-Seventh Asilomar Conference on Systems and Computers, Vol.2, pp. 1982-1986, 2003.

B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," Nature, vol. 38, pp. 607–609, 1996.

E.J. Candes and F. Guo. "New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction," Signal Processing, Vol. 82 (2), pp. 1519-1543,2002.

H. Moravec, ROBOT: Mere Machine to Transcendent Mind, Oxford University Press, 1998

A. Majumdar, "Curvelets: A New Approach to Face Recognition", Proceedings of National Conference on Knowledge Based Frontier Technologies, National conference on Knowledge Based Computing Systems and Frontier Technologies, pp. 197-205, 2007.

A. Majumdar and J.N. Mazumdar, "Wavelets or Contourlets: Which is a better descriptor for Pattern Recognition?" Proceedings of INDIACom pp. 405-411, 2007.

A. Majumdar and A. Bhattacharya, "A Comparative Study in Wavelets, Curvelets and Contourlets as Feature Sets for Pattern Recognition", International Arab Journal of Information Technology (In Print).

Z. Jiulong, Z. Zhiyu, H. Wei, L. Yanjun and W. Yinghui, "Face Recognition Based on Curvefaces," Third International Conference on Natural Computation, Vol.2, pp.627-631, 2007.

T. Mandal, A. Majumdar and J. Hu, " Face Recognition Via Curvelet Based feature Extraction", International Conference on Image Analysis and Recognition, pp. 806-817, 2007.

A. Majumdar, S. Ray and A. Bhattacharya, "Face Recognition by Contourlet Based Feature Extraction" Indian International Conference on Artificial Intelligence, pp. 2010-2026, 2007.

A. Majumdar and A. Bhattacharya, "Face Recognition by Multiresolution Curvelet Transform on Bit Quantized Facial Images", IEEE International Conference on Computational Intelligence and Multimedia Applications, pp. 209-213, 2007.

A. Majumdar and B. B. Chaudhuri, "Curvelet-Based Multi SVM Recognizer for Offline Handwritten Bangla: A Major Indian Script", IEEE International Conference on Document Analysis and Research, Vol. 1, pp.491-495, 2007.

U. Bhattacharya and B. B. Chaudhuri, "Fusion of Combination Rules of an Ensemble of MLP Classifiers for Improved Recognition Accuracy of Handprinted Bangla Numerals", IEEE International Conference on Document Analysis and Research, Vol. 1, pp.322-326, 2005.

U. Bhattacharya and B. B. Chaudhuri, "A Majority Voting Scheme for Multiresolution Recognition of Handprinted Numerals IEEE International Conference on Document Analysis and Research, Vol. 1, pp.16-20, 2003.

T, Mandal, "A New Approach to Face Recognition Using Curvelet Transform", MASc Thesis, University of Windsor, ON, Canada.

H. K. Ekenel and B. Sankur, "Multiresolution face recognition." Image. Vision Comput., Vol. 23 (5), pp. 469–477, 2005.

G.L. Marcialis and F. Roli, "Fusion of LDA and PCA for Face Verification", Workshop on Biometric Authentication, pp. 30-37, 2002.

G.L. Marcialis and F. Roli, "Fusion of LDA and PCA for Face Recognition", Workshop on Machine Vision and Perception, Italian Association for Artificial Intelligence, pp. 215-227, 2002.

A. Majumdar and R. K. Ward, "Pseudo-Fisherface Method for Single Image per Person Face Recognition", International Conference on Acoustics, Speech, and Signal Processing, pp. 989-992, 2008.

A. Majumdar and R. K. Ward, "Single Image per Person Face Recognition with Images Synthesized by Non-Linear Approximation", accepted at International Conference on Image Processing (ICIP08).

Y. Yan and L. A. Osadciw, "Contourlet Based Image Recovery and De-noising Through Wireless Fading Channels",Annual Conference on Information Sciences and Systems, John Hopkins Unversity, 2005

R. Muraleedharan, Y. Yan and L. A. Osadciw, “Constructing an Efficient Wireless Face Recognitions System by Swarm intelligence”, Academic Graduate Excellence Symposium Program (AGEP 2007), Syracuse University, New York, June 2007.

N. Goel, G. Bebis and A. V. Nefian, “Face recognition experiments with random projections”, SPIE Conference on Biometric Technology for Human Identification, pp. 426-437, 2005.

T. A. B. Jin and T. Y Chong, "Cancelable Biometrics Realization With Multispace Random Projections," IEEE Transactions on Systems, Man, and Cybernetics, Part B, Vol.37 (5), pp.1096-1106, 2007.

Y. H. Pang and A. T. B. Jin, “Random Projection with Robust Linear Discriminant Analysis Model in Face Recognition”, Computer Graphics, Imaging and Visualization, pp. 11-15, 2007.

J. Wright, A. Y. Yang, A. Ganesh, S. Sastry, and Y. Ma. Robust face recognition via sparse representation. To appear in PAMI, 2008.

E. J. Candes, and T. Tao. “Decoding by linear programming”, IEEE Transactions on Information Theory, Vol. 51(12), pp. 4203-4215, 2006.

J. Haupt, R. Castro, R. Nowak, G. Fudge, and A. Yeh, "Compressive Sampling for Signal Classification," Asilomar Conference on Signals, Systems and Computers, pp.1430-1434, 2006.

"IEEE Signal Processing Magazine [Sensing, Sampling, and Compression]," Signal Processing Magazine, IEEE , Vol.25 (2), pp.c1-c1, March 2008

IntechOpen



Recent Advances in Face Recognition

Edited by Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett

ISBN 978-953-7619-34-3
Hard cover, 236 pages
Publisher InTech
Published online 01, June, 2008
Published in print edition June, 2008

The main idea and the driver of further research in the area of face recognition are security applications and human-computer interaction. Face recognition represents an intuitive and non-intrusive method of recognizing people and this is why it became one of three identification methods used in e-passports and a biometric of choice for many other security applications. This goal of this book is to provide the reader with the most up to date research performed in automatic face recognition. The chapters presented use innovative approaches to deal with a wide variety of unsolved issues.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Angshul Majumdar, and Rabab K. Ward (2008). Multiresolution Methods in Face Recognition, Recent Advances in Face Recognition, Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett (Ed.), ISBN: 978-953-7619-34-3, InTech, Available from: http://www.intechopen.com/books/recent_advances_in_face_recognition/multiresolution_methods_in_face_recognition

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2008 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen