Use of Wavelet-based Basis Functions to Extract Rotation Invariant Features for Automatic Image Recognition

SANTIAGO AKLE, MARIA-ELENA ALGORRI, ANTE SALCEDO
Department of Digital Systems
Instituto Tecnológico Autónomo de México
Río Hondo 1, Progreso Tizapán, 01080 México D.F.
tiagoakle@gmail.com, algorri@itam.mx

Abstract: - In this paper we explore the use of orthogonal functions as generators of representative, compact descriptors of image content. In Image Analysis and Pattern Recognition such descriptors are referred to as image features, and there are some useful properties they should possess such as rotation invariance and the capacity to identify different instances of one class of images. We exemplify our algorithmic methodology using the family of Daubechies wavelets, since they form an orthogonal function set. We benchmark the quality of the image features generated by doing a comparative OCR experiment with three different sets of image features. Our algorithm can use a wide variety of orthogonal functions to generate rotation invariant features, thus providing the flexibility to identify sets of image features that are best suited for the recognition of different classes of images.

Key-Words: - Rotation Invariant Features, Zernike Moments, Haar Wavelets, Daubechies Wavelets, Orthogonal Functions, OCR

1 Introduction

The problem of finding compact descriptors of the content of images for the use of automatic recognition of image information is nowadays more present than ever. Just as we have got used to browsing through textual data for words, terms or phrases, the demand for browsing through images in search of a given object, face, structure or any other 2D graphical representation is growing. Whether it is to render an image or illustration readable by the computer, or to be able to recognize objects for security or supervision purposes, image mining remains one of the big milestones to be settled by the information society. A popular technique to analyze images for the purpose of automatic recognition is supervised learning. In supervised learning a representative set of images is available from which relevant features that uniquely characterize the objects in the images can be extracted. The extracted features then allow the computer to recognize similar objects in images that are not contained in the representative set. Techniques for feature extraction in Image Analysis commonly use a basis of orthogonal functions to generate a subspace over which images are projected. The coefficients resulting from the projection are then, either directly or the result of a combination of them used as features to characterize the image. Zernike moments and orthogonal Fourier-Mellin moments, have been extensively used in Image Analysis since they allow the extraction of rotation invariant features for image recognition [1][2][3].Wavelets are also popular families of functions to build basis sets for feature extraction [4][5]. Because of their efficiency to identify temporal/spatial features in different types of data sets, wavelet functions have gained popularity in a large domain of data classification applications including data mining[19], classification of music[18], data stream mining[16], separation of linearly mixed signals[15], fingerprint verification[17] and medical image retrieval[14] among others.

In [6] and [7], Gabor Wavelets are used as basis functions to build a projection subspace. The image coefficients are obtained using the Gabor Transform and they characterize the orientations of local image structure. The method requires the choice of a number of parameters for the wavelets and the extracted coefficients are not rotation invariant. In [8] a log-polar wavelet transform is used to extract rotation and scale invariant features for texture analysis. [9] also proposes the extraction of shape descriptors using wavelet analysis. In this paper we propose a general framework to extract rotation invariant features from images. The features are compact descriptors of the image content and are well suited for image classification, analysis and supervised learning. The framework we propose is inspired in the popular Zernike moments which have been proved to provide robust features for image

description. We exemplify the use of our framework using wavelet functions and measure the quality of the extracted features as image descriptors in an OCR experiment.

2 Problem Formulation

Zernike moments are the coefficients that result from the projection of an image over a set of orthogonal functions in the unit disk called Zernike functions. The structure of these functions results in the extraction of features that are rotation invariant. When expressed in polar coordinates, a Zernike function $V_{pq}(r,\theta)$ can be factorized in two terms,

$$\begin{aligned} V_{pq}(r,\theta) &= R_{pq}(r)\exp\{-jq\theta\} \\ p-|q| &\, is\, pair \\ 0 \le |q| &\le p \end{aligned} \tag{1}$$

where $p,q \in Z$, and *j* is the imaginary unit. The projection of an image *I(x,y)* over a function of the basis V_{pq} is done using the inner product defined by

$$\mu_{pq} = \langle V_{pq}, I \rangle = \int_D \int V_{pq}(r,\theta)\overline{I(x,y)}\partial x \partial y \tag{2}$$

where the integral is over the unit disk with center on the origin. The image term in the integral is complex conjugated since the functions in the basis V_{pq} are complex valued. The resulting μ_{pq} is the coefficient that measures the similarity of one function of the basis V_{pq} to *I(x,y)*. *I(x,y)* must be projected over all the functions in the basis and the resulting set of coefficients μ_{pq} is then used as features that describe *I(x,y)* in terms of the basis V_{pq}. The projection process, together with the orthogonality of the basis functions under the inner product in (2), ensures that if all the functions in V_{pq} are weighted with their corresponding coefficient and then linearly combined, they will provide $\hat{I}(x,y)$, the best approximation to *I(x,y)* in terms of least square error, that is,

$$\hat{I}(x,y) = \sum_{(p,q)\in P} \mu_{pq} V_{pq}(x,y) \tag{3}$$

where the set *P* of integer pairs *(p,q)* is the set that holds the restrictions defined in (1).

A rotation of the image by an angle of θ_0 produces coefficients μ_{pq} that are multiplied by a unit modulus complex number with an angle dependent on θ_0. That is, the coefficients maintain their original modulus before rotation, therefore the coefficients' modulus can be used as rotation invariant features. A rotation of I(x,y) by an angle θ_0, can be represented in polar coordinates as $I(r,\theta+\theta_0)$. The projection of $I(r,\theta+\theta_0)$ over the Zernike basis V_{pq} is defined as,

$$\begin{aligned} \rho_{pq} &= \langle V_{pq}, I(r,\theta+\theta_0) \rangle \\ &= \iint_D V_{pq}(r,\theta)\overline{I(r,\theta+\theta_o)\partial x \partial y} \\ &= \int_0^1 \int_0^{2\pi} R_{pq}(r)\exp\{-jq\theta\}\overline{I(r,\theta+\theta_o)} r\partial\theta\partial r \end{aligned} \tag{4}$$

where ρ_{pq} is the coefficient of the projection of the rotated image over one Zernike function in V_{pq} and the new *r* term that appears is the integrand is the Jacobian of the transformation of *I(x,y)* to polar coordinates. By applying the change of variable $\theta' = \theta + \theta_0$, an $\exp\{jq\theta_0\}$ term can be factorized from the integral resulting in the following relationship,

$$\begin{aligned} &\rho_{pq} = \\ &\exp\{jq\theta\}\int_0^1 \int_0^{2\pi} R_{pq}(r)\exp\{-jq\theta'\}\overline{I(r,\theta')} r\partial\theta'\partial r = \\ &\exp\{jq\theta\}\mu_{pq} \end{aligned} \tag{5}$$

where the coefficients of the rotated image ρ_{pq}, provide rotation invariant features, since the modulus of the coefficients stays constant under rotations of the projected image.

2.1 Proposed Framework

To create a new basis of functions from which rotation invariant features can be extracted, the new basis must be orthogonal and must maintain an exponential term when expressed in polar coordinates. Therefore the elements of the new basis will take the form,

$$F_{pq}(r,\theta) = R_{pq}(r)\exp\{-jq\theta\} \tag{6}$$

To form an orthonormal basis using the above definition will require setting certain restrictions over the functions $R_{pq}(r)$. To be orthonormal, the inner product between two functions of the set must hold the following equality,

$$\begin{aligned} &\langle F_{no}, F_{pq} \rangle = \\ &\int_0^{2\pi} \int_0^1 R_{no}\exp\{-jo\theta\}R_{pq}(r)\exp\{jq\theta\} r\partial r\partial\theta = \\ &d_{np}d_{oq} \end{aligned} \tag{7}$$

where d_{ij} is the Kronecker symbol. Since complex exponential functions with integer period follow the equality,

$$\int_0^{2\pi} \exp\{-jq\theta\}\overline{\exp\{-js\theta\}}\partial\theta = 2\pi d_{qs} \quad (8)$$

the relationship in (7) can be rewritten as,

$$\left\langle F_{no}, F_{pq} \right\rangle = \int_1^0 R_{no}(r)R_{pq}(r)r\left(\int_0^{2\pi} \exp\{-jo\theta\}\exp\{jq\theta\}\partial\theta\right)\partial r = \int_1^0 R_{no}(r)R_{pq}(r)r\left(2\pi d_{oq}\right)\partial r = d_{np}d_{oq} \quad (9)$$

From (9) it can be inferred that if the $R_{pq}(r)$ terms follow the condition

$$\int_0^1 R_{no}(r)R_{pq}(r)r\partial r = \frac{1}{2\pi} d_{np}d_{oq} \quad (10)$$

the resulting set is orthonormal in the unit disk.

Our proposed strategy is to start with a set of orthonormal functions over the unit interval that achieves the condition

$$\int_0^1 R_i(r)R_j(r)\partial r = d_{ij} \quad (11)$$

we will then modify the functions in (11) to produce a set of orthonormal functions in the unit disk that complies with (10).

Given a set of orthonormal functions in the unit interval $\{R_0(r), R_1(r),..., R_M(r)\}$, we will modify them to take the form $\left\{\frac{1}{\sqrt{\pi}}R_0(r^2), \frac{1}{\sqrt{\pi}}R_1(r^2),..., \frac{1}{\sqrt{\pi}}R_M(r^2)\right\}$, and the members of the final family of orthonormal functions in the unit disk will take the form,

$$F_{pq}(r,\theta) = \frac{1}{\sqrt{\pi}} R_{pq}(r^2)\exp\{-jq\theta\} \quad (12)$$

To test for orthonormality, it suffices to evaluate the inner product between any two members of the set in (12),

$$\left\langle F_{no}(r,\theta), F_{pq}(r,\theta)\right\rangle = 2d_{oq}\int_0^1 R_n(r^2)R_p(r^2)r\partial r \quad (13)$$

which after a change of variable $r' = r^2$, equals

$$\left\langle F_{no}(r,\theta), F_{pq}(r,\theta)\right\rangle = 2d_{oq}\frac{1}{2}\int_0^1 R_n(r')R_p(r')\partial r' = d_{oq}d_{np} \quad (14)$$

This strategy allows the use of any, from a wide range, of known orthogonal functions in the unit interval, to form many different families of orthonormal functions in the unit disk, which can then be used to extract rotation invariant features.

We use families of wavelet functions to produce orthonormal basis in the unit disk, because they have the advantage that, using the DWT (Discrete Wavelet Transform) the projection of images onto the basis can be carried out in a much faster and easier way than the projection of images over the basis of Zernike functions. Our results also show that the quality of the features extracted from the wavelet based basis is as good as the quality of the features obtained from Zernike moments. We use Daubechies wavelets since they are generators of orthogonal basis of functions. We use Daubechies-2 and Daubechies-4 wavelets. Daubechies-2 wavelets, also called Haar wavelets, constitute the simplest of wavelet basis. The Haar Wavelet basis is defined in terms of a father and a mother function [11]. The father function is defined as

$$\Phi(t) = 1 \quad (15)$$

and the mother function is defined as

$$\Psi(t) = \begin{cases} 1, & 0 \le t \le \frac{1}{2} \\ -1, & \frac{1}{2} < t \le 1 \end{cases} \quad (16)$$

the rest of the wavelet basis is defined in terms of

$$\Psi_{j,s}(t) = 2^j \Psi(2^j t - s) \quad (17)$$

where *j,s,* are integers such that $0 \le s \le 2^j$. Fig. 1 shows the Haar wavelets for j = 0,1 and 2.

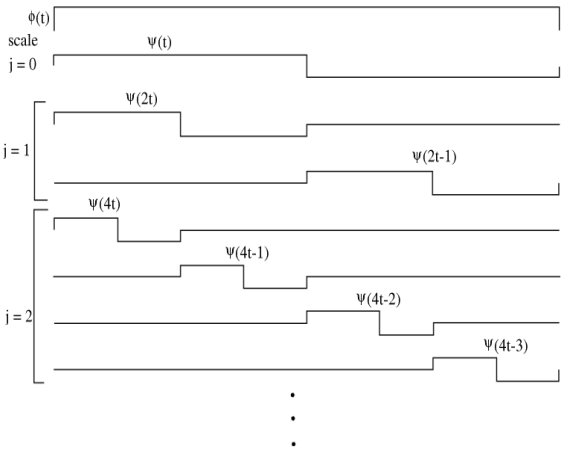

Fig. 1. The Haar wavelet basis

Fig. 2 shows the normalized value of the real part of some of the functions in the orthonormal basis that was created using the Haar wavelet basis. The first column shows functions based on the wavelet $\Phi(t)$, the second, third and fourth columns use functions based on the wavelets $\Psi(t)$, $\Psi_{1,0}(t)$, and $\Psi_{1,1}(t)$ respectively. Each row uses a different

value of *q* (the parameter of the exponential term), starting with *q=1* for row 1, and incrementing *q* by 1 thru *q = 4* for row 4.

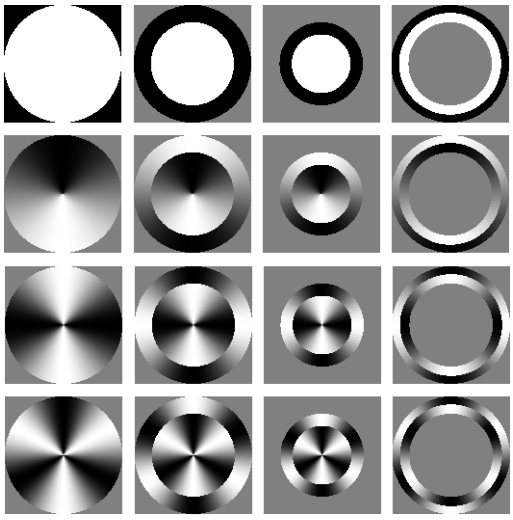

Fig. 2. Orthonormal basis functions in the unit disk.

The orthonormal basis in the unit disk takes the final form

$$F_{j,s,q}(r,\theta)=\frac{1}{\sqrt{\pi}}\Psi_{j,s}(r^2)e^{-jq\theta} \qquad (18)$$

We also created an orthonormal basis in the unit disk using Daubechies-4 wavelets. These functions are best described as a family of compactly supported functions, with four vanishing moments. For these, the scaling function (father function Φ(t)) can be represented as a linear combination of scaled and shifted versions of itself,

$$\phi(t)=\sqrt{2}\sum_{n=-\infty}^{\infty}h[s]\phi(2t-s) \qquad (19)$$

The series *h[s]* that defines the family of Daubechies-4 wavelets is:

$$h[0]=\frac{1+\sqrt{3}}{4}, h[1]=\frac{3+\sqrt{3}}{4} \qquad (20)$$

$$h[2]=\frac{3-\sqrt{3}}{4}, h[3]=\frac{1-\sqrt{3}}{4} \qquad (21)$$

The Daubechies-4 wavelet is illustrated in Fig.3. Although this wavelet is not discontinuous as the Haar wavelet is, its function has little regularity.

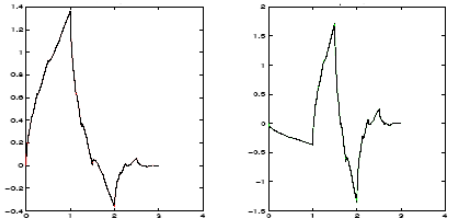

Fig.3. The Daubechies-4 wavelets

Just as we did with the Haar wavelets, we make the Daubechies-4 wavelets orthonormal over the unit disk by changing their domain from *t*, to (*r*, *θ)*, and then associating them an exponential term. The final basis of functions takes the form,

$$F_{s,q}(r,\theta)=\frac{1}{\sqrt{\pi}}\phi_s(r^2)e^{-jq\theta} \qquad (22)$$

with *q* the parameter of the exponential term, and θ the parameter that reflects image rotations. By expressing the basis functions in polar terms, their magnitude remains unaffected by rotations in the images.

2.2 Projection Process

To project an image *I(x,y)* over the basis in (18), we first binarize the image and then translate and scale the coordinates of the foreground pixels such that their centroid lies in (0,0) and all foreground pixels are at a distance smaller or equal to unity from (0,0). That is, all the foreground pixels lie within the unit disk. After binarization and normalization, the projection of *I(x,y)* over the proposed orthonormal basis, can be carried out in a 2 step process. First, the image *I(x,y)* is projected over the exponential part of the basis functions, then the image is projected over the wavelet functions. The projection process is carried out using the inner product expressed in polar coordinates,

$$\mu_{j,s,q}=\left\langle I(r,\theta),F_{j,s,q}(r,\theta)\right\rangle= \frac{1}{\sqrt{\pi}}\int_0^1\int_0^{2\pi}rI(r,\theta)\Psi_{j,s}(r^2)e^{jq\theta}d\theta dr \qquad (24)$$

where $\mu_{j,s,q}$ are the coefficients resulting from the projection. Integrating first over θ and then over *r*, (19) can be written as

$$\mu_{j,s,q}=\frac{1}{\sqrt{\pi}}\int_0^1 r\Psi_{j,s}(r^2)\left(\int_0^{2\pi}I(r,\theta)e^{jq\theta}d\theta\right)dr \qquad (25)$$

From (23) we can define the intermediate function

$$g_q(r)=\int_0^{2\pi}I(r,\theta)e^{jq\theta}d\theta \qquad (26)$$

and using the variable substitution $\rho=r^2$, the projection can be written as

$$\mu_{j,s,q}=\frac{1}{\sqrt{\pi}}\int_0^1 r\Psi_{j,s}(r^2)g_q(r)dr= \frac{1}{2\sqrt{\pi}}\int_0^1\Psi_{j,s}(\rho)g_q(\sqrt{\rho})d\rho \qquad (27)$$

which is the definition of the projection of a function $g_q(\sqrt{\rho})$ over a member of the wavelet basis $\Psi_{j,s}(\rho)$. To take advantage of the DWT (Discrete Wavelet Transform), the continuous $g_q(\sqrt{\rho})$ function must be approximated by a discrete series of finite length $g_q[x]$, formed by uniform sampling the continuous function $g_q(\sqrt{\rho})$. To evaluate $g_q[x]$, we define

$$T_q(\sqrt{\rho},\theta) = I(\sqrt{\rho},\theta)e^{jq\theta} \quad (28)$$

$$g_q(\sqrt{\rho}) = \int_0^{2\pi} T_q(\sqrt{\rho},\theta)d\theta \quad (29)$$

The evaluation of $I(\sqrt{\rho_o},\theta_o)$ at a foreground pixel coordinate $(r,\theta)_0$ equals unity, therefore, the evaluation of $T(\sqrt{\rho_o},\theta_o)$ at the same point equals $e^{jq\theta_0}$. The evaluation of the function $g_q(\sqrt{\rho_0})$ at a point ρ_0 is the integral of function T_q over the line defined by the circle of radius $\sqrt{\rho_0}$. To approximate the value of $g_q(\sqrt{\rho_0})$ we sum the values of T_q that are at a distance $\sqrt{\rho_0}$ from the origin. To enhance the performance of the DWT, we calculate the discrete series $g_q[x]$ to have a length N which is a power of two. To calculate $g_q[x]$ we subdivide the unit interval [0,1] in N subintervals of length 1/N. For a value $x = x_o$, the evaluation of $g_q[x_0]$ is the sum of values of T_q that are at a distance $\sqrt{\frac{x_0}{N}} \le \sqrt{\rho} \le \sqrt{\frac{x_0+1}{N}}$ from the origin, this is proportional to the number of foreground pixels in the image that are at a distance $\sqrt{\rho}$ from (0,0).

The projection algorithm proceeds iterating over as many values of q as coefficients are desired for the projection of the image over each wavelet function. At each iteration, the projection algorithm calculates the series $g_q[x]$ and uses the DWT to produce the desired coefficients.

3 Results

Using the framework proposed in the last section, we created two basis of functions that are orthonormal in the unit disk: one using the Haar wavelets and another one using the Daubechies wavelets. We then tested the quality and rotation invariance of the image features that were produced using the two basis of functions, by doing an OCR experiment. The OCR experiment was done with 6471 images of 58 different classes of symbols. To have some comparison metrics for the quality of the image features extracted, the OCR experiment was also run using the open source OCR program gOCR[12], and using Zernike moments as image features. The features extracted from Zernike moments also provided a comparison metric for the speed of calculation and rotation invariance of the features produced with our basis of functions.

3.1 OCR Results using the Haar Wavelet Basis

We extracted a set of 144 complex coefficients (288 features) for each of the 6471 images available for the OCR experiment using the basis functions built with the Haar wavelets. Each image contained only one symbol of 58 possible classes. We then used the support vector machine algorithm LibSVM[10] to implement a system of supervised learning. Support Vector Machines (SVMs) require a training set of known OCR classes. Using the *N* features of the known classes, an SVM creates partitions in the *N*-dimensional feature space which, ideally, cluster the features of the independent classes. If the feature space is well built, when a new image is presented to the SVM, its features will belong to the partition of the space that corresponds to the correct class.

We ran the OCR test doing Bootstraping, a well-known technique to estimate performance statistics of a system of supervised learning independently of the training set used. In Bootstraping, the set of 6471 known images is randomly divided into 2 sets: a training set and a test set. The training set is used to build the partitions of the features space in the SVM, and the test set is then classified using the feature space previously built. Each time the set of 6471 known images is partitioned into a training and a test set and the SVM algorithm is run is known as one replic in the Bootstraping process. In this part of the experiment we ran 200 replics of Bootstraping. From the results of the Bootstraping process we calculated the average error of the system and the average a posteriori error per class. The average error of the system is an estimate of the probability that the system makes an error when classifying any given image. The average a posteriori error per class is an estimate of the probability that the classification of an image into one specific class be wrong. Table 1 shows the average error and the

standard deviation of the OCR results and Table 2 shows the a posteriori errors per class.

Features from	Avg. Error	Std. Deviation
Haar wavelets	0.041501	.00441240

Table 1. Average error and standard deviation of the OCR system using Haar-wavelet features after 200 replics of Bootstraping.

Class	**!**	*****	**+**	**/**	**1**	**2**	**3**	**4**	**5**
APE*	.07	.05	0	.09	.02	.01	0	0	.01
Class	**6**	**7**	**8**	**9**	**=**	**>**	**?**	**A**	**B**
APE*	0	0	.01	0	.09	0.1	.52	.01	.01
Class	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**
APE*	.02	.01	0	0	.02	.01	.08	.18	.01
Class	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**
APE*	.01	0	0	.02	.01	0	0	.03	.04
Class	**U**	**V**	**W**	**X**	**Y**	**Z**	\	**_**	**a**
APE*	.08	0	0	.02	0	0	.07	.63	0
Class	**b**	**d**	**e**	**F**	**g**	**h**	**I**	**J**	**n**
APE*	.04	0	.02	.03	.04	.04	.1	.1	.01
Class	**q**	**r**	**t**	**~**					
APE*	0	.04	0	.34					

*APE = a Posteriori Error

Table 2. Probability of a posteriori errors per class in the OCR experiment using features extracted from functions built with Haar wavelets.

3.2 OCR Results using Daubechies-4 Wavelet Basis

We extracted 144 complex coefficients (288 features) for each of the 6471 images in the OCR experiment using the basis functions built with the Daubechies-4 wavelets. As in the experiment with the Haar-wavelet functions, every image contained only one of 58 possible classes of symbols. Using the Daubechies-4 features, we ran the OCR experiment using Support Vector Machines with 200 replics of Bootstraping. The 200 replics allowed us to calculate the average error of the OCR system, the standard deviation and the average a posteriori error per class. Tables 3 and 4 show these results.

Features from	Avg. Error	Std. Deviation
Daubechies-4 wavelets	0.103617264	.005475998

Table 3. Average error and standard deviation of our OCR system using Daubechies-4 features after 200 replics of Bootstraping.

Class	**!**	*****	**+**	**/**	**1**	**2**	**3**	**4**	**5**
APE*	.11	0.0	.06	.09	0.0	.01	.02	0.0	.03
Class	**6**	**7**	**8**	**9**	**=**	**>**	**?**	**A**	**B**
APE*	.03	.02	.01	0.0	.09	.25	.69	0.0	.01
Class	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**
APE*	.04	0.0	0.0	.04	.02	.03	.06	.02	.02
Class	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**
APE*	.01	.01	.02	.03	.03	.01	.04	.04	.12
Class	**U**	**V**	**W**	**X**	**Y**	**Z**	\	**_**	**A**
APE*	.03	.03	0.0	.03	.05	.04	.07	.68	.01
Class	**b**	**d**	**e**	**f**	**g**	**h**	**i**	**j**	**N**
APE*	0.0	.02	.05	.06	0.0	.02	.15	.08	.04
Class	**q**	**r**	**t**	**~**					
APE*	.02	.1	.01	.14					

Table 4. Probability of a posteriori errors per class in the OCR experiment using features extracted from functions built with the Daubechies-4 wavelets.

3.2 OCR Results using gOCR

gOCR is a robust open source OCR that is used in many non-commercial applications. From the 58 classes that we recognized with the Haar wavelet coefficients, gOCR can recognize 53, so we ran the OCR experiment with gOCR over 4595 images. Because gOCR cannot be bootstrapped, we only ran one pass over the images and were able to obtain an average error for the system and an a posteriori error per class but not a standard deviation. The results obtained with gOCR are shown in Tables 5 and 6.

gOCR	Avg. Error
	0.063731

Table 5. Average error in the OCR test using gOCR

Class	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
APE*	0	.01	0	0	.02	0	0	.02	0
Class	*****	**+**	**<**	**>**	**a**	**A**	**B**	**B**	**C**
APE*	.05	.25	0	.2	.64	0	0	0	.02
Class	**D**	**d**	**E**	**e**	**F**	**f**	**G**	**G**	**H**
APE*	0	0	0	0	0	.02	.67	Na	0
Class	**H**	**I**	**I**	**j**	**J**	**K**	**L**	**M**	**n**
APE*	0	0	0	Na	0	0	0	0	.02
Class	**N**	**O**	**P**	**q**	**Q**	**R**	**r**	**S**	**T**
APE*	0	.02	0	Na	.06	0	.74	.05	0
.04	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**		
APE*	.04	0	0	0	0	0	.2		

*APE = a Posteriori Error

Table 6. Probability of a posteriori errors per class using gOCR.

3.3 OCR Results using Zernike features

We ran the OCR test one more time using features extracted from Zernike functions. The objective was to evaluate the performance of the OCR in average error, standard deviation, rotation invariance and speed of computation when using Zernike features vs using the features obtained from the orthonormal functions built with the Haar and Daubechies-4 wavelets. Tables 7 and 8 show the results of the average error, standard deviation and a posteriori errors per class when the OCR experiment is done with Zernike features.

Features From	Avg. Error	Std. Deviation
Zernike Mmts.	0.08709277	0.00598263

Table 7. Average error and standard deviation of the OCR system using Zernike Moments as descriptors.

Class	**!**	*****	**+**	**/**	**1**	**2**	**3**	**4**	**5**
APE*	.12	0	.15	.11	.04	.02	.02	0	.01
Class	**6**	**7**	**8**	**9**	**=**	**>**	**?**	**A**	**B**
APE*	.03	.03	.02	.01	.22	0	.56	0	.01
Class	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**
APE*	.04	0	0	.02	.02	.03	.11	.03	.01
Class	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**
APE*	0	.04	.01	.02	.04	.03	0	.03	.06
Class	**U**	**V**	**W**	**X**	**Y**	**Z**	\	**_**	**A**
APE*	.06	.04	.01	0	.07	0	.09	.51	.01
Class	**B**	**d**	**e**	**f**	**g**	**h**	**I**	**J**	**N**
APE*	0	.01	.02	.06	.01	.01	.29	.07	.02
Class	**Q**	**r**	**t**	**~**					
APE*	.01	.13	.02	.31					

*APE = a Posteriori Error

Table 8. Probability of a posteriori errors per class for the OCR using Zernike Moments as descriptors.

3.4 Rotation Invariance of Image Features

To compare the rotation invariance qualities of the three feature sets used in the OCR experiment, we calculated the first 4 features of letters "A", "B" and "C" with 4 different rotations. Because the order of magnitude of the features extracted from Zernike moments, Haar wavelets and Daubechies-4 wavelets vary greatly, they cannot be directly compared for rotation invariance. Instead, for the set of rotations for each letter, we calculated the following dispersion measure,

$$Ri(fe_i(x)) = Var(|fe_{i,\theta_j}(x)|) / \mu(|f_{i,\theta_j}(x)|) \quad (30)$$

where *Ri* stands for rotation invariance and $fe_i(x)$ represents the feature ***i*** of class ***x***. So in (30), the rotation invariance of feature ***i*** of class ***x*** is given by the ratio of the variance of the magnitude of feature ***i*** subject to a set of rotations θ_j, $var(|fe_{i,\theta j}|)$, to the mean value of the magnitude of feature ***i*** subject to a set of rotations θ_j, $\mu(|fe_{i,\theta j}|)$. Fig. 4 shows the four rotations applied to letters "A", "B", and "C" that were used to calculate the rotation invariance of the first four features of the three sets of features.

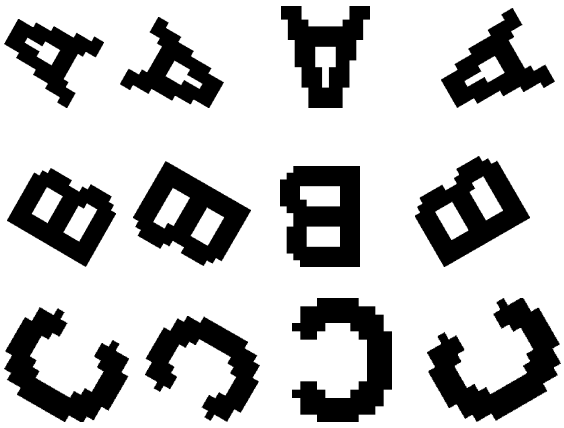

Fig. 4. Images of letters A, B and C with the 4 rotations used to compare the rotation invariance of the 3 different sets of features.

Tables 9, 10 and 11 show the rotation invariance measures obtained using (30). For each set (A, B, C) of rotated letters we calculated 4 features using the Haar, Daubechies-4 and Zernike basis. The measures indicate how much the modulus of the features vary when an image is rotated, the smaller the value, the less the variation of the feature modulus and the greater the rotation invariance associated to the feature.

Rotation Invariance of Features 1 to 4 for "**A**"			
Haar-wavelet features			
2.83E-07	1.4891E-06	1.7556E-06	3.2797E-06
Daubechies-4 features			
6.40E-06	1.7158E-06	6.4726E-06	3.2810E-06
Zernike features			
0.636619	0.63661977	0.00591442	0.13663572

Table 9. Measures of the rotation invariance of the first four image features of rotations of the letter **A**.

Rotation Invariance of Features 1 to 4 for "**B**"			
Haar-wavelet features			
2.056E-07	4.395E-06	1.814E-06	2.67E-06
Daubechies-4 features			
2.407E-06	6.450E-07	2.042E-05	2.03E-06
Zernike features			
0.6366197	0.6366197	0.0053833	0.124141

Table 10. Measures of the rotation invariance of the first four image features of rotations of the letter **B**.

Rotation Invariance of Features 1 to 4 for "**C**"			
Haar-wavelet features			
1.330E-06	5.725E-06	5.194E-05	9.13E-05
Daubechies-4 features			
2.362E-05	6.329E-06	1.212E-05	1.89E-05
Zernike features			
0.6366197	0.6366197	0.0769928	3.382378

Table 11. Measures of the rotation invariance of the first four image features of rotations of the letter **C**.

3.5 Speed of Computation of Image Features

The reported times in Table 12 measure the time taken to calculate the features used in the OCR experiment with the three different basis of functions (the image loading and normalization times are not taken into account). The total calculation time is divided over the total amount of features calculated for the entire corpus to obtain an average calculation time per feature.

OCR with **Haar-wavelet** Features	OCR with **Daubechies-4 wavelet** features	OCR with **Zernike** Features
Features per Image		
288	288	272
Total Images		
6471	6471	6471
Calculated Features for all Images in 1 Replic		
1863648	1863648	1760112
Avg. calculation Time for all Features in 1 Replic (ms)		
2253.25	4522.8	208648.4
Avg. calculation Time per Feature (ms)		
0.0012	0.002426853	0.137502

Table 12. Comparison of calculation times for Haar, Daubechies-4 and Zernike features used in the OCR experiment.

4 Discussion of Results

4.1 OCR Experiment

For the OCR experiment, we used as comparison metric the average error of the open source OCR program gOCR, which is a popular OCR application used in many non-commercial systems. gOCR also produced the best OCR results (smallest average error of classification) among three open source OCRs that we tested to be used as our comparison metric. The average error of the gOCR system was 6.37%, which is a very acceptable error for an OCR when identifying single characters that cannot be verified in the context of a word.

For the OCR experiment, we formed two basis functions, which were orthogonal over the unit disk using Haar and Daubechies-4 wavelets. Using these orthogonal basis, we extracted two sets of image features (288 features per image) from the set of 6471 images of 58 classes of symbols. We used the sets of image features in a supervised learning system, where a support vector machine algorithm classified the symbols in the images into the 58 classes. Since the gOCR system does not use supervised learning for the classification, we obtained a second comparison metric for our results by running the support vector machine algorithm using Zernike moments as image features. Since Zernike moments are well established, robust descriptors, they provided a good frame of comparison for our wavelet-based image features. In the OCR experiment, the results in term of average error are shown in Table 13, where it can be seen that the features obtained from the Haar wavelets performed best and, in general, the experiments that used supervised learning outperformed gOCR.

OCR results	Average Error
gOCR	6.37%
Zernike features	5.47%
Haar wavelet features	4.41%
Daubechies-4 features	5.98%

Table 13. Comparison of the average error in the OCR experiment obtained with the 4 systems tested.

4.2 Rotation Invariance results

One of the main motivations of the algorithmic methodology proposed in this paper was to produce image features that were rotation invariant since our research includes identifying randomly positioned objects in images. Again, as comparison metric for the rotation invariance quality of our produced sets of image features we used Zernike moments. Zernike moments are a standard for comparison for rotation invariance and it is this quality that makes them so popular in scientific literature. To test for rotation invariance we used three sets of images. Each set of images represented a symbol (letters "A", "B", and "C") with four rotations. We calculated 4 features for each set of images. If a feature were perfectly rotation invariant, its value would be the same for all the images in the set. We compared the variation of the first 4 features produced with the three basis of functions. For comparison we propose a dispersion measure that calculates the ratio of the variance of the magnitude of a feature vs the average magnitude of the feature,

(see (30)). Table 14 shows the results of rotation invariance, where the features obtained from Haar wavelets showed the best invariance to rotation followed by the Daubechies-4 features. Both sets of wavelet-based features outperformed the features obtained from Zernike moments.

Average Rotation Invariance of Image features	Letter "A"	Letter "B"	Letter "C"
Zernike moments	0.3539	0.350691	1.183152
Haar wavelet	1.701E-6	2.273E-6	3.758E-5
Daubechies-4	4.468E-6	6.377E-6	1.525E-5

Table 14. Results of the rotation invariance of the image features obtained from the wavelet basis vs the image features obtained from Zernike moments.

4.3 Speed of Computation Results

Another main objective of our algorithmic methodology to produce image features, was that their computation complexity be much lower than the computational complexity of Zernike features. Since our research includes automatic recognition of images over large databases of several thousand images, speed of computation was an issue. Table 15 shows the results of the average computation time in milliseconds needed to extract one feature using the three different sets of projection functions. Haar wavelet functions proved to have the smaller computational complexity and outperformed Daubechies-4 and Zernike in speed of computation.

Avg. computation time per feature	
Zernike moments	0.1375 ms
Haar wavelets	0.0012 ms
Daubechies-4 wavelets	0.0024 ms

Table 15. Results of speed of computation per feature for the three function basis.

5 Conclusion

We propose a framework for extracting rotation invariant image features from a wide variety of function sets. The function sets are first made orthogonal over the unit disk by a non linear scaling of their domain and then an exponential term is associated to them. We have tested the proposed framework using the Daubechies family of wavelets, in particular Haar and Daubechies-4 to obtain features for over 6000 images. We did an OCR experiment to compare the average error of a supervised learning system vs the error of an open source OCR. We also compared the quality of the wavelet features against the Zernike moments. The OCR experiment using supervised learning obtained a lower average classification error than the open source OCR, and in particular the features obtained from our wavelet based functions performed better than the features obtained from Zernike moments. As for rotation invariance and speed of computation, the wavelet based features also outperformed the Zernike features.

Of interest in the methodology we propose in this paper is that, for a given type of image classes, several sets of image features can be produced until the one that produces the best classification results is identified. The sets of image features will all have the property of rotation invariance. In our results, we observed that, for binarized images of letters, the Haar wavelet features produced the best results. Our current work includes running tests with gray-level and color images to find the most suitable set of image features for automatic recognition of the image content.

References:

[1]Belkasim S., Hassan E., Obeidi T., „Explicit invariance of Cartesian Zernike moments", *Patt. Recognition Letters*, Vol 28, 2007, pp 1969-1980

[2] Teoh Ben Jin A., Ngo Chek Ling D., Than Song O., "An efficient fingerprint verification system suing integrated wavelet and Fourier-Mellin invariant transform", *Image and Vision Computing*, Vol 22, 2004, pp. 503-513

[3] Sastry C.S., Pujar A.K., Deekshatulu B.L., "A Fourier-Radial Descriptor Algorithm for Invariant Feature Extraction", *International Journal of Wavelets, Multiresolution and Inf. Proc*, Vol. 4, No. 1, 2006, pp. 197-212

[4] Sastry C. S., Pujari A. K., Deekshatulu B.L., Bhagvati C., "A wavelet based multiresolution algorithm for rotation invariant feature extraction", *Pattern Recognition Letters*, Vol. 25, 2004, pp. 1845-1855

[5] Li C., Huang, J-Y., Chen C-M., „Soft computing approach to feature extraction", *Fuzzy sets and systems*, Vol. 147, 2004, pp. 119-140

[6] Shustorovich, A.,"Scale specific and Robust Edge/Line Encoding with Linear Combinations of Gabor Wavelets", *Pattern Recognition*, Vol. 27, No. 5., 1994, pp.713-725

[7] Shustorovich A., "A Subspace Projection Approach to Feature Extraction: The Two-Dimensional Gabor Transform for Character

Recognition", *Neural Networks*, Vol. 7, No. 8, 1994, pp 1295-1301

[8] Lee M-C., Pun C-M.,"Rotation and Scale Invariant Wavelet Feature for Content-Based Texture Image Retrieval", *Journal of the American Society. for Information Science and Technology,* Vol 54. No. 1, 2003, pp. 68-80

[9] Shen D., Ip H., "Discriminative wavelet shape descriptors for recognition of 2D patterns", *Pattern Recognition*, Vol 32, 1999 pp. 151-165

[10] Chang, C-C., Lin C-J., „LibSVM: A library for support vector machines", 2001, http://www.csie.ntu.tw/cjlin/libsvm

[11] Mallat, S., "A Wavelet Tour of Signal Processing", Elsevier, 1999

[12] Schulenburg J,gOCR,http://jocr.sourceforge.net

[13] Park L. A. F., Ramamohanarao K., Palaniswami M., "A Novel Document Retrieval Mehtod Using the Discrete Wavelet Transform", *ACM Transactions on Information Systems*, Vol. 23, No. 3, July 2005, pp. 267 – 298

[14]Karam O.H., Hamad A.M., Ghoniemy S., Rady S., "Enhancement of Wavelet-Based Medical Image Retrieval through Feature Evaluation Using an Information Gain Measure", *Proc. of the ACM Symposium on Applied Computing*, 2003, pp. 220 -- 226

[15] Kisilev P., Zibulevsky M., Zeevi Y.Y., "A Multiscale Framework for Blind Separation of Linearly Mixed Signals", *Journal of Machine Learning Research*, Vol. 4, 2003, pp. 1339—1364

[16] Papadimitriou S., Brockwell A., Faloutsos C., "Adaptive, unsupervised stream mining", *The VLDB Journal*, Vol. 13, 2004, pp. 222 – 239

[17] Andrew T.B.J., David N.C.L., "Integrated Wavelet and Fourier-Mellin Invariant Feature in Fingerprint Verification System", *Proc. of the ACM SIGMM workshop on Biometrics methods and applications*, 2003, pp. 82—88

[18] Grimaldi M., Cunningham P., Kokaram A., "A Wavelet Packet Representation of Audio Signals for Music Genre Classification Using Different Ensemble and Feature Selection Techniques", *Proc. of the ACM SIGMM intl. workshop on Multimedia info. retrieval*, 2003, pp. 102-- 108

[19] Li. T., Li, Q., Zhu S., Ogihara M., „A Survey on Wavelet Applications in Data Mining", SIGKDD Explorations Newsletter, Vol. 4, No. 2, 2002, pp. 49—68