WAVELETS AS FEATURES FOR OBJECTS RECOGNITION

Anca APATEAN (DISCANT) [1], Alexandrina ROGOZAN [2], Simina EMERICH [1], Abdelaziz BENSRHAIR [2]
[1] *Tehnical University of Cluj-Napoca, Cluj-Napoca, Romania*
[2] *National Institute of Applied Sciences, Rouen, France*
[1] *26-28, G.Baritiu Street, RO-400027, Cluj-Napoca, Romania, Tel: +40 264 401 568, Fax: +40 264 591 340*
{Anca.Discant@bel, Simina.Emerich@com}.utcluj.ro,{arogozan , Abdelaziz.BENSRHAIR }@insa-rouen.fr

Abstract: In this paper, we mainly concentrate on the recognition module of an object detection and recognition system. Two types of images, visible and infrared, are investigated in order to improve the objects detection and recognition process. Different types of mother wavelets (Haar, Daubechies, Coiflet, Symlet, Biorthogonal, Fractional causal or generalized, etc.) are used to extract the wavelet coefficients which will constitute the feature vector. The obtained feature vector then will be fed to a KNN classifier, in order to classify the object in one of the possible object's classes used in the training step.

Key words: *object recognition, wavelets, feature extraction, KNN classifier.*

I. INTRODUCTION

Object recognition in computer vision is a task of finding given object in an image. Humans recognize a multitude of objects in images with little effort; by the sensorial fusion process they combine visual with acoustic and tactile information to get more knowledge about the surrounding. Despite the fact that the image of the objects may vary somewhat in different view points, sizes/scale or even when they are translated or rotated, the respective object can be easy recognized by humans. This task is a true challenge for computer vision systems because a training and then a testing process is needed. For any object in an image, there are many "features" (interesting points on the object like textures, colours, symmetries, edges or other parameters) that can be extracted to provide a description of the respective object. This description extracted from a training image can then be used to identify the object in a test image.

Given an image, or a region within an image, we generate the wavelet features that will be fed to a KNN classifier, in order to classify the image in one of the 5 possible classes.

II. THEORETICAL BACKGROUND

Wavelet transform

A wavelet is a mathematical function used to divide a given function into different frequency components. A wavelet transform is the representation of a function by wavelets, which represent scaled and translated copies of a finite-length or fast-decaying oscillating waveform (known as the "mother wavelet").

Wavelet analysis represents a windowing technique with variable-sized regions. Wavelet analysis allows the use of long time intervals where more precise low-frequency information is needed, and shorter regions where high-frequency information is necessary.

All wavelet transforms may be considered forms of time-frequency representation for continuous-time (analog) signals. Given a signal with some event in it, one cannot assign simultaneously an exact time and frequency respective scale to that event. Almost all practically useful discrete wavelet transforms use discrete-time filter banks. These filter banks are called the wavelet and scaling coefficients and may contain either finite impulse response (FIR) or infinite impulse response (IIR) filters.

Wavelet transforms have advantages over traditional Fourier transforms because local features can be described better with wavelets that have local extent. Fourier analysis consist of breaking up a signal into sine waves of various frequencies, while wavelet analysis breaks a signal into shifted and scaled versions of the original (or mother) wavelet. Sinusoids do not have limited duration (they extend from minus to plus infinity), a wavelet is a waveform of effectively limited duration that has an average value of zero.

Mathematically, the process of Fourier analysis is represented by the Fourier transform, which is the sum over all time of the signal multiplied by a complex exponential. The results of the transform are the Fourier coefficients, which multiplied by a sinusoid of a specific frequency yield the constituent sinusoidal components of the original signal. Similarly, the continuous wavelet transform is defined as the sum over all time of the signal multiplied by scaled, shifted versions of the wavelet function. The results of the wavelet transform are many wavelet coefficients, which are a function of scale and position. Multiplying each coefficient by the appropriately scaled and shifted wavelet yields the constituent wavelets of the original signal.

Wavelet analysis consists of decomposing a signal or an image into a hierarchical set of approximations and details. For images analysis we used two-dimensional wavelets and corresponding scaling functions obtained from one-dimensional wavelets by tensorial product. The discreet wavelet transform (DWT) of a signal is calculated by passing it through a series of filters (high and low pass filters) and then down-sampled, as we can see from Figure 1.

At each level, the signal is decomposed into low and high frequencies, and this decomposition halves the resolution since only half the number of samples are retained to characterize the entire signal. The algorithm retains the even indexed columns respectively rows. Based on this scheme, two-dimensional DWT leads to a decomposition of approximation coefficients at level j in four components: the approximation (CA) at level j+1 and the details (CD) in three orientations (horizontal, vertical, and diagonal). Due to successive downsampling by 2, the signal length must be a power of 2, or a multiple of a power of 2, and the length of the signal determines the maximum levels in which the signal can be decomposed. An example of this wavelet decomposition on level two, applied on a vehicle image can be seen in Figure 2.

Some mother wavelet families implemented in Matlab (as Daubechies, Symlet, Coiflet, Biortogonal and Reverse biorthogonal wavelets) and the fractional B-spline functions are used to compute different feature vectors.

Orthogonal wavelets with FIR filters can be defined through a scaling filter. Predefined families of such wavelets include Haar, Daubechies, Symlets and Coiflets. The names of the Daubechies family wavelets are written *DbN*, where *N* is the order, and *Db* the "surname" of the wavelet. These wavelets have no explicit expression except for *Db1*, which is the Haar wavelet. Symlets (S*ymN*, where N is the order) are only near symmetric. Coiflet (*CoifN*) is a discrete wavelet designed by ingrid Daubechies to be more symmetrical than the DbN wavelet. In [1] a feature vector containing Haar wavelets together with features like statistical moments, discreet cosine transform, gray level coocurence matrix and other features was used in order to increase the recognition rates.

Biorthogonal wavelets with FIR filters include Biorthogonal (*BiorNr.Nd*) and Reverse Biorthogonal (*RbioNr.Nd*) wavelets, where Nr and Nd are the orders for the reconstruction and decomposition respectively. Both types of wavelets are compactly supported biorthogonal spline wavelets with FIR filters for which symmetry and exact reconstruction are possible with FIR filters.

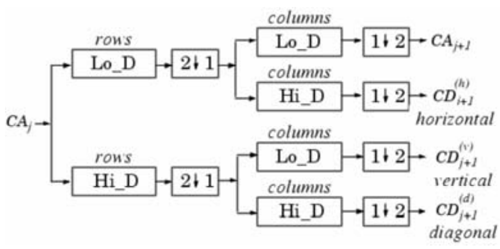

Figure 1. Multilevel 2-D wavelet decomposition

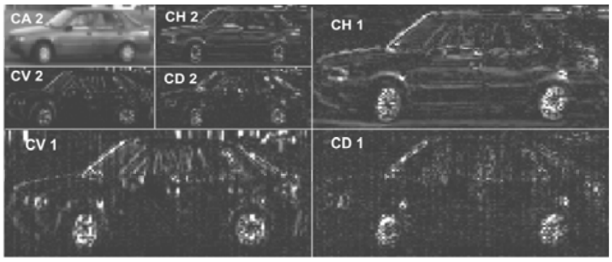

Figure 2. Wavelet decomposition

The fractional B-spline wavelets are tunable in a continuous fashion. By varying a parameter α, we have a direct control over a number of key wavelet properties: the parametric form of the basis functions, their smoothness, their space-frequency localization, and others. Different types of fractional B-splines wavelets have been investigated: causal (noted with symbol +), anti-causal (-), symmetric (*) and generalized (%). For more theory about wavelet analysis please see [2-6]. Different types of fractional mother wavelet were used in [2] together with statistical and Hu moments for objects recognition in visible color images.

Classifiers

We tested the wavelet algorithm using different functions on visible and infrared (IR) spectrum images containing cars or pedestrians in different arbitrary poses. The images from the recognition module contain a bounding box surrounded the object. In the classification task a feature vector of a fixed dimension is needed, thus in the preprocessing step a resize operation is performed. On these images the two-dimensional DWT was applied, based on wavelet families described before. The algorithm is iterative, the approximations being successively decomposed. Usually the decomposition level is chosen according on certain criteria.

The main purpose of our application is to obtain a small feature vector and a good classification rate. In this order a 3, 4 and 5 level wavelet decomposition was performed. For one image of 128x128 pixels, if a Haar wavelet transform is used, then in the first level decomposition results 64x64, in the second level 32x32, in the third level 16x16, in the fourth level 8x8 features, and so on. Comparing the results, we noticed level 4 give the best solution for our obstacle recognition problem considering the size of the feature vector and the achieved recognition rates. After the features are extracted from the image, they will be fed into a classifier, in order to classify the image in one of the 5 possible classes from Table 1.

TABLE I

Number of objects assigned to every class in visible and infrared domain

VISIBLE					**IR**			
Class	Learn	Test	Cross		Class	Learn	Test	Cross
1	105	67	172		1	300	221	521
2	34	22	56		2	63	48	111
3	50	33	83		3	69	46	115
4	1099	498	1597		4	1112	660	1772
5	118	71	189		5	115	75	190
Total	1406	691	2097		Total	1659	1050	2709

In order to detect and recognize some features in other new images, the system must be learned a priori. This step represents the learning procedure, and then a classification follows.

A very powerful system to perform the classification task is Weka [7] which provides implementations of different state-of-the-art learning and classification algorithms. We used Weka in order to apply a KNN (with k=1) learning method to our dataset and analyze its output. The k-nearest neighbor algorithm (k-NN) is a method for classifying objects based on closest training examples in the feature space. KNN is a type of instance-based learning where the function is only approximated locally. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. In the case k = 1, the object is simply assigned to the class of its nearest neighbor.

III. APPLICATION: OBJECTS RECOGNITION USING WAVELETS AS FEATURES

Two image databases, visible and IR were used for testing. These image datasets were provided at the Robin Competition [8] in order to test and compare the feature extraction and selection algorithms in both types of images, visible and IR. In Table I it can be seen there are 1406 objects (Learn), 691 objects (Test), so a total number of 2097 objects (Cross) in the visible domain, and 1659 objects (Learn), 1050 objects (Test) and a total number of 2709 objects (Cross) in the IR domain.

In order to test which wavelet family (used for the feature extraction task) give best results for our application, two evaluation methods were used:

- the 10-folds cross-validation technique (using images from Cross directory) and
- a classical scheme of learning with some instances (using Learn images) and testing with other instances (with Test images).

In a 10-fold cross-validation process, the original sample of data is partitioned into 10 sub-samples. From these 10 sub-samples, a single sub-sample is retained (the validation or learning set), and the remaining 9 sub-samples are used to train the classifier (the training set). The cross-validation process is then repeated 10 times and a combination of the 10 results (generally the average) is performed in order to obtain the accuracy parameter.

The summary of the results from the training data ends with a confusion matrix and the detailed accuracy parameter per classes. The confusion matrix shows how many instances of each class have been assigned to each class. If all instances have been correctly classified, only the diagonal elements of the matrix are non-zero. From the confusion matrix one can observe if the instances of a class have been assigned to another class. The accuracy is a statistical measure of how well the classifier correctly identifies the objects, and it is a parameter of test. The accuracy is the value appearing in Figure 3 and it represents the proportion of true results in the population. An accuracy of 100% means that the test assigns all the objects to the correctly class.

The number of CA4 (approximation coefficients) is around the value of 8x8, as in Table II. One of the analysis criteria was CA4 dimension and its connection with the classification ratio: the largest the feature vector dimension, the greater correct classification ratio. Figure 3 presents the classification rates. Different feature vectors with different wavelet families were used to classify the objects, in the following order:

- Daubechies: Db1 (Haar), Db2, Db3, Db4, Db5, Db6,Db7, Db8, Db9, Db10,
- Symlet: Sym2, Sym 3, Sym4, Sym 5, Sym6, Sym7, Sym8,
- Coiflet: Coif1, Coif2, Coif3, Coif4, Coif5,
- Biorthogonal: Bior1.1, Bior1.3, Bior1.5, Bior2.2, Bior2.4, Bior2.6, Bior2.8, Bior3.1, Bior3.3, Bior3.5,
- Reverse biorthogonal: Rbio1.1, Rbio1.3, Rbio1.5, Rbio2.2, Rbio2.4, Rbio2.6, Rbio2.8, Rbio3.1, Rbio3.3, Rbio3.5,
- Causal: + 0.2, + 0.4, + 0.9, + 1.4, + 1.8,
- Anti-causal: - 0.4, - 0.2, - 0.1, - 0.01, - 0.0001,
- Symmetric: * 0.2, * 0.4, * 0.9, * 1.4, * 1.8,
- Generalized: % 0.2, % 0.4, % 0.9, % 1.4, % 1.8.

TABLE III

Feature vector's dimension for different types of wavelets

Wavelet family									Feature vector
Db	Sym	Coif	Bior	Rbio	+ Causal	- Anticausal	* Symmetric	Generalized %	
1			1.1	1.1					8X8=64
2	2		3.1	3.1	0.2	0.4	0.2	0.2	10X10=100
					0.4	0.2	0.4	0.4	
					0.9	0.1	0.9	0.9	
					1.4	0.01	1.4	1.4	
					1.8	0.0001	1.8	1.8	
3	3	1	1.3	1.3					12X12=144
			2.2	2.2					
4	4		3.3	3.3					14X14=196
5	5		1.5	1.5					16X16=256
			2.4	2.4					
6	6	2	3.5	3.5					18X18=324
7	7		2.6	2.6					20X20=400
8	8								22X22=484
9		3	2.8	2.8					23X23=529
10									25X25=625
		4							29X29=841
		5							35X35=1225

The results from Figure 3 show that the highest recognition rates are obtained for visible images by crossvalidation technique. Here it seems that crossvalidation technique provide better results in both cases, visible and IR; that is because the images from the test directory were not well chosen. If compare different types of wavelets, it can be seen the accuracy above 91% are given by Daubechies, biorthogonal, reverse biorthogonal of first orders, causal and generalized fractional b-spline wavelets.

ACKNOWLEDGMENT

This paper was supported by the Education and Research Minister, under the GRANT “Researches in Processing and Fusioning Data Provided By Multiple Optical Sensors. Applications for Objects Detection and Classification”, Anca Apatean (Discant).

REFERENCES

[1] A. Apatean (Discant), A. Rogozan, A. Bensrhair, "Objects Recognition in Visible and Infrared Images from Road Scene", Proceedings of the IEEE International Conference on Automation, Quality and Testing, Robotics, Theta 16, Cluj-Napoca, Romania, 2008

[2] A. Apatean, S. Emerich, A. Rogozan, A. Bensrhair, E. Lupu, "Ruttier Obstacle Classification by use of Fractional B-spline Wavelets and Moments", Proceedings of the IEEE Region 8 Eurocon 2007 Conference "Computer as a Tool", ISBN 1-4244-0813-X, Varsovia, Poland, 2007

[3] S. Mallat, “A wavelet tour of signal processing”, New York, Academic Press, 1999.

[4] M. Unser and T. Blu, “Fractional splines and wavelets”, *SIAM Review*, vol. 42, no. 1, pp. 43–67, January 2000.

[5] M. Unser, “Splines: A perfect fit for signal and image processing”, *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, November 1999.

[6] T. Blu and M. Unser, “A complete family of scaling functions: The (α, τ)-fractional splines,” *Twenty-Eighth International Conference on Acoustics, Speech, and Signal Processing (ICASSP’03)*, Hong Kong SAR, People’s Republic of China, April 6-10, 2003, vol. VI, pp. 421–424.Lim, *Two Dimensional Signal and Image Processing*, Prentice Hall International, Inc., New York, 1992.

[7] G. Holmes, A. Donkin, and I.H. Witten, “Weka: A machine learning workbench”, In *Proc Second Australia and New Zealand Conference on Intelligent Information Systems*, Brisbane, Australia, 1994.

[8] http://robin.inrialpes.fr/robin.pdf

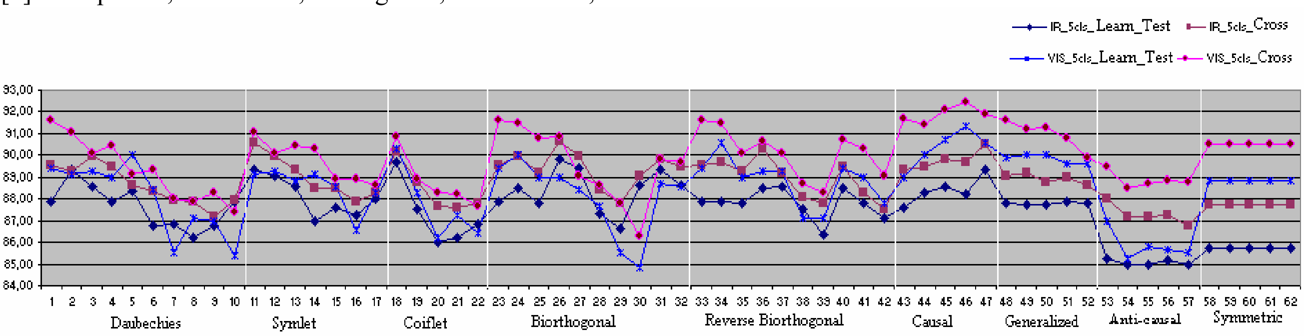

Figure 3. Recognition rates in visible and infrared images
using the two methods of testing: learning followed by testing (red) and crossvalidation (blue)

TABLE II

Recognition rates in visible and infrared images
using the two methods of testing: learning followed by testing (APP VAL) and crossvalidation (ALL)

			Learn Test	Cross	Learn Test	Cross
1	(Haar) Db1	64	87,90	89,55	89,43	91,60
2	Db2	100	89,33	89,24	89,14	91,03
3	Db3	144	88,57	89,95	89,29	90,12
4	Db4	196	87,90	89,47	89,00	90,46
5	Db5	256	88,38	88,66	90,01	89,12
6	Db6	324	86,76	88,37	88,42	89,36
7	Db7	400	86,85	87,96	85,52	88,03
8	Db8	484	86,19	87,89	87,12	87,88
9	Db9	529	86,76	87,15	86,97	88,31
10	Db10	625	87,90	87,96	85,38	87,41
11	Sym2	100	89,33	90,58	89,14	91,03
12	Sym3	144	89,04	89,95	89,29	90,12
13	Sym4	196	88,57	89,36	88,85	90,41
14	Sym5	256	86,95	88,51	89,14	90,27
15	Sym6	324	87,61	88,51	88,56	88,93
16	Sym7	400	87,23	87,85	86,54	88,93
17	Sym8	484	88,00	88,26	88,27	88,65
18	Coif1	144	89,71	90,14	90,30	90,84
19	Coif2	324	87,52	88,77	88,27	88,93
20	Coif3	529	86,00	87,70	86,25	88,31
21	Coif4	841	86,19	87,59	87,26	88,22
22	Coif5	1225	86,85	87,81	86,39	87,69

			Learn Test	Cross	Learn Test	Cross
23	Bior1.1	64	87,90	89,55	89,43	91,60
24	Bior1.3	144	88,47	89,92	90,01	91,46
25	Bior1.5	256	87,80	89,18	89,00	90,79
26	Bior2.2	144	89,80	90,66	89,00	90,84
27	Bior2.4	256	89,42	89,95	88,42	89,03
28	Bior2.6	400	87,33	88,44	87,69	88,65
29	Bior2.8	529	86,66	87,81	85,52	87,83
30	Bior3.1	100	88,66	89,03	84,80	86,26
31	Bior3.3	196	89,33	89,81	88,71	89,79
32	Bior3.5	324	88,66	89,44	88,56	89,69
33	Rbio1.1	64	87,90	89,55	89,43	91,60
34	Rbio1.3	144	87,90	89,66	90,59	91,51
35	Rbio1.5	256	87,80	89,25	89,00	90,08
36	Rbio2.2	144	88,47	90,29	89,29	90,65
37	Rbio2.4	256	88,57	89,22	89,29	90,08
38	Rbio2.6	400	87,52	88,11	87,12	88,74
39	Rbio2.8	529	86,38	87,78	87,12	88,26
40	Rbio3.1	100	88,47	89,44	89,43	90,70
41	Rbio3.3	196	87,80	88,29	89,00	90,27
42	Rbio3.5	324	87,14	87,56	87,84	89,03

			Learn Test	Cross	Learn Test	Cross
43	+ 0.2	100	87,61	89,33	89,00	91,65
44	+ 0.4	100	88,28	89,44	90,01	91,41
45	+ 0.9	100	88,57	89,84	90,73	92,13
46	+ 1.4	100	88,19	89,66	91,31	92,46
47	+ 1.8	100	89,33	90,51	90,59	91,89
48	% 0.2	100	87,80	89,07	89,86	91,60
49	% 0.4	100	87,71	89,18	90,01	91,22
50	% 0.9	100	87,71	88,77	90,01	91,27
51	% 1.4	100	87,90	88,99	89,58	90,79
52	% 1.8	100	87,80	88,63	89,58	89,86
53	- 0.4	100	85,22	88,00	86,97	89,46
54	- 0.2	100	84,95	87,19	85,23	88,50
55	- 0.1	100	84,95	87,15	85,81	88,69
56	- 0.01	100	85,16	87,22	85,67	88,84
57	- 0.0001	100	84,95	86,74	85,52	88,79
58	* 0.2	100	85,71	87,74	88,85	90,51
59	* 0.4	100	85,71	87,74	88,85	90,51
60	* 0.9	100	85,71	87,74	88,85	90,51
61	* 1.4	100	85,71	87,74	88,85	90,51
62	* 1.8	100	85,71	87,74	88,85	90,51