International Journal of the Physical Sciences Vol. 6(18), pp. 4428-4434, 9 September, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.370
ISSN 1992 -1950 ©2011 Academic Journals

Full Length Research Paper

Signature verification using rules 3-ext inductive learning system

Mehmet Sabih Aksoy[1]* and Hassan Mathkour[2]

[1]Department of Information Systems, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia.
[2]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia.

Accepted 29 April, 2011

This paper presents an alternative technique for "signature verification". The technique employs template matching for feature extraction and Rules 3-ext inductive learning algorithm to extract the necessary set of rules and to verify a signature. 15 of 3 × 3 masks were used to represent a signature. Each signature (or pattern) is presented by the frequencies of the masks used. The system was trained using 144 signatures (16 signatures belonging to 9 different persons each). The system has been tested using many unseen signatures and the ability to correctly classify them was found to be 97%.

Key words: Inductive learning, signature verification, machine learning, image processing.

INTRODUCTION

Signature is a very important evidence for persons. A formal letter, a cheque, a contract etc. is not valid without signature(s). The verification of a signature is very important as well. It can be done manually but an automatic verification system will save time and money. Especially for electronic transactions the need of such a system is obvious. That is why we need an automatic signature verification system. There has been many studies on signature verification and recognition. Nelson and Kishon used some of the dynamic features for signature data and their uses in augmenting the shape analysis of signatures in verification (Nelson and Kishon, 1991). Draouhard et al. (1996) developed a system which uses neural network approaches to on-line signature verification using directional "probability density function" (Draouhard et al., 1996). Lee et al. (1996) developed a reliable on-line signature verification system which uses several classifier types (Lee et al., 1996). Bajaj and Chaudhary have developed a system for signature verification using multiple neural classifiers (Bajaj and Chaudhary, 1997). Wu et al. (1998) developed an on-line signature verification system based on logarithmic spectrum and split-and-merge matching mechanism (Wu et al., 1998). Bovino et al. (2003) developed a multi-expert system for dynamic signature verification. The system uses a stroke-oriented description of signatures well-suited for multi-expert approach (Bovino et al., 2003). Coetzer et al. (2004) developed a system that automatically authenticates off-line handwritten signatures using the "discrete radon transform" and a hidden Markov model (Coetzer et al., 2004). Ozgunduz et al. (2005) developed an off-line signature verification and recognition system using the global, directional and grid features of signatures. They used support vector machine to verify and classify the signatures (Ozgunduz et al., 2005). Guru and Prakash have developed a system for online signature verification and recognition using an approach based on symbolic representation (Guru and Prakash, 2009). This paper presents an alternative technique for automatic signature verification.

The technique uses Rules 3-Ext inductive learning system for both learning and verification of signatures. The system employes template matching technique to represent a signature. 15 of 3 × 3 masks were used to represent a signature. Each signature (or pattern) is presented by the frequencies of the masks used. The system was trained using 144 signatures (16 signatures belonging to 9 different persons). The system has been tested using many unseen signatures and the ability to

*Corresponding author. E-mail: msaksoy@ksu.edu.sa. Tel: 00966 1 4677697. Fax: 0096614675423.

correctly classify them was found to be 97%. The organization of the paper is as follows: In this study, Rules 3-ext inductive learning algorithm is outlined. It further explains the new proposed technique and describes its operation. It also contains results and discussion of an application for the system. Lastly is the conclusion.

RULES 3-EXT INDUCTIVE LEARNING ALGORITHM

Rules 3-ext is a simple algorithm for extracting a set of classification rules from a collection of examples for objects belonging to one of a number of known classes (Mathkour, 2009). It is a descendant of Rules 3 algorithm (Pham and Aksoy, 1993). An object must be described in terms of a fixed set of attributes, each with its own range of possible values, which could be nominal or numerical. For example, attribute "length" might have nominal values (short, medium, long) or numerical values in the range (-10 and 10). An attribute-value pair constitutes a condition in a rule. If the number of attributes is N_a, a rule may contain between one and N_a conditions. Only conjunction of conditions is permitted in a rule and therefore the attributes must be all different if the rule comprises more than one condition. Rules 3-ext extracts rules by considering one example at a time. It forms an array consisting of all attribute-value pairs associated with the object in that example, the total number of elements in the array being equal to the number of attributes of the object. The rule forming procedure may require at most N_a iterations per example. In the first iteration, rules may be produced with one element from the array. Each element is examined in turn to see if, for the complete example collection, it appears only in objects belonging to one class. If so, a candidate rule is obtained with that element as the condition. In either case, the next element is taken and the examination repeated until all elements in the array have been considered. At this stage, if no rules have been formed, the second iteration begins with two elements of the array being examined at a time. Rules formed in the second iteration therefore have two conditions. The procedure continues until iteration when one or more candidate rules can be extracted or the maximum number of iterations for the example is reached. In the latter case, the example itself is adopted as the rule. If more than one candidate rule is formed for an example, the rule that classifies the highest number of examples is selected and used to classify objects in the collection of examples. Examples of which objects are classified by the selected rule are removed from the collection.

The next example remaining in the collection is then taken and rule extraction is carried out for that example. This procedure continues until there are no examples left in the collection and all objects have been classified. This algorithm is summarized in Figure 1.

THE PROPOSED TECHNIQUE

There are many techniques for visual inspection. Four of them are (1) image subtraction, (2) dimensional verification, (3) syntactic approach and, (4) feature matching (Davies, 1995; Jain et al., 1995; Roland, 1982). In “image subtraction”, the image subject to inspection is scanned and compared against the original image, which is stored before. The subtracted image is analyzed. This method needs large reference data storage, accurate alignment and sensitive illumination and scanner conditions. Also many images may not match identically even if they are acceptable. In “dimensional verification” method a determination for each measurement is made as to weather it falls within the previously established standards. The distance between edges of geometric shapes is the basic feature of this technique. Syntactic approach uses descriptions of a large set of complex objects using small sets of simple pattern primitive and structural rules. Primitives are small number of unique elements, such as lines or corners. A structural description of the primitives and the relationships between them can be determined to form a string grammar. In feature matching or template matching method, the required features are extracted by using a number of predefined templates from the image to be inspected. Then these features are compared with those defined for the perfect pattern. This method greatly compresses the image data for storage and reduces the sensitivity of the input intensity data. A number of predefined binary templates can be used to extract the necessary features for images to be inspected (Davies, 1995; Jain et al., 1995; Roland, 1982).

We use “feature (template) matching” method for this study because it is very suitable for Rules 3-ext algorithm, as it needs a set of examples that contain a number of attribute-value pairs and a class. Each template (mask) can be considered as an attribute and the frequencies of templates can be considered as values for those attributes. The class is the labeled image being processed.

Learning session

The first step is to set up a set of examples for Rules 3-ext to learn. Among 20 possible 3 x 3 binary templates (masks) (Figure 2), a set of 15 masks shown in Figure 3 was chosen because of their good performance after several tests on many “signatures” used for the study. The performance criterion is the frequency (number of matches) for masks. The masks that have the higher frequencies were chosen. Each template represents an attribute. 144 images for 9 different person’s signature (16 for each) were used for learning. The masks were applied to these patterns from left to right and from top to bottom in order to find their frequencies (Aksoy et al., 2000; Aksoy, 2004). Before applying the masks to scanned binary images, an edge detection operation is required. For this study the Laplacian edge detection operator was chosen because of its good performance than other operators such as Roberts, Sobel, Prewitt and Kirsch compass. The discrete form of Laplacian operator is as follows (Ballard, 1982):

$$L(x, y) = f(x, y) - 1/4[f(x, y + 1) + f(x, y - 1) + f(x + 1, y) + f(x - 1, y)]$$

Where (x, y) are the coordinates for the pixel being processed; L is the intensity to be calculated and f(x, y) is the intensity of the pixel having (x, y) coordinates. Figure 4 shows a signature before and after edge detection operation. The steps required to extract the necessary features (set of rules) are as follows:

Step 1. Scan a signature (binary image).
Step 2. Perform edge detection operation.
Step 3. Apply 15 binary masks to the image to find their frequencies.
Step 4. Set up the set of training examples.

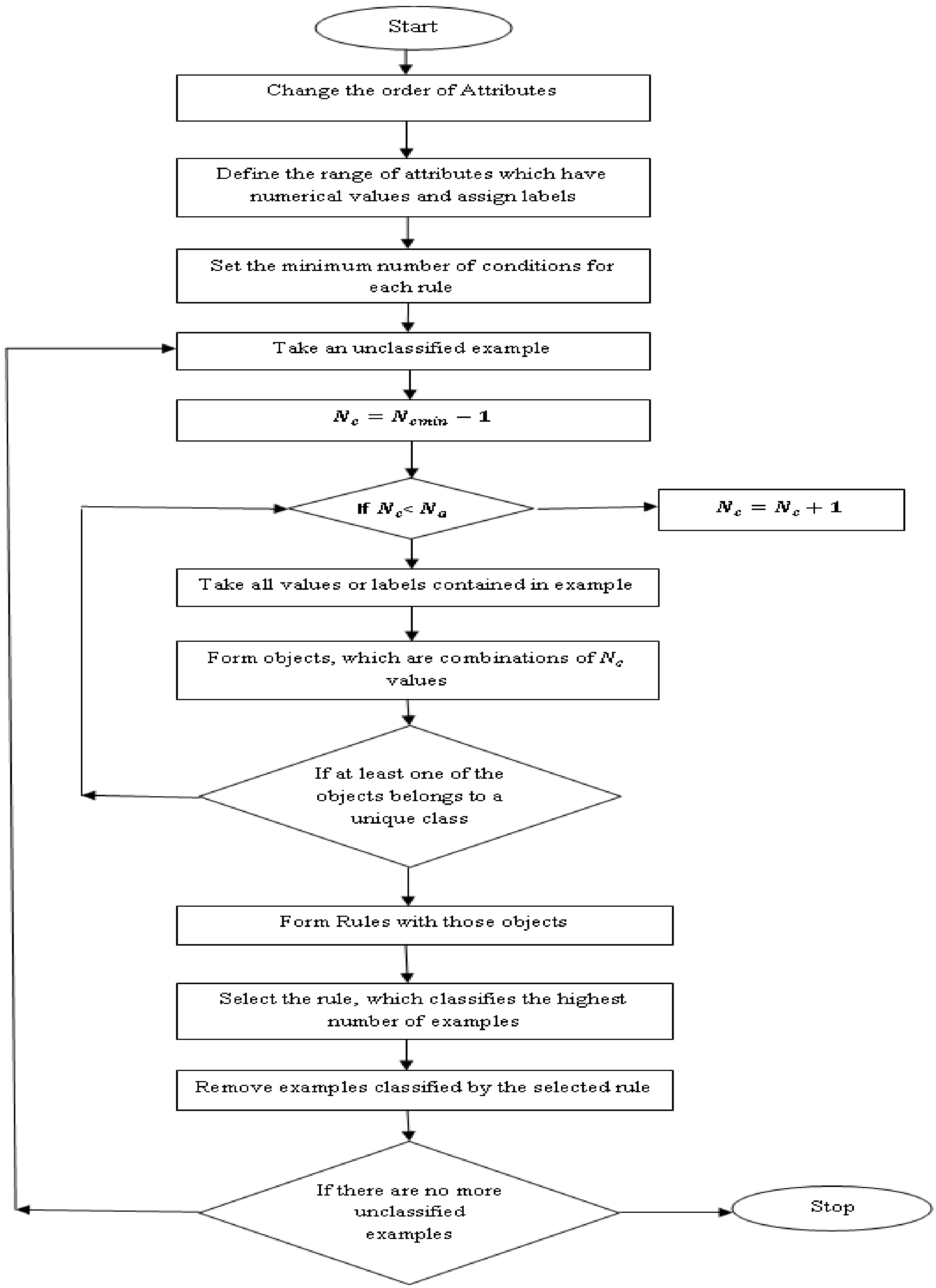


Figure 1. The flowchart for Rules 3-ext. (Where N_c is the number of condition(s) and N_a is the number of attributes).

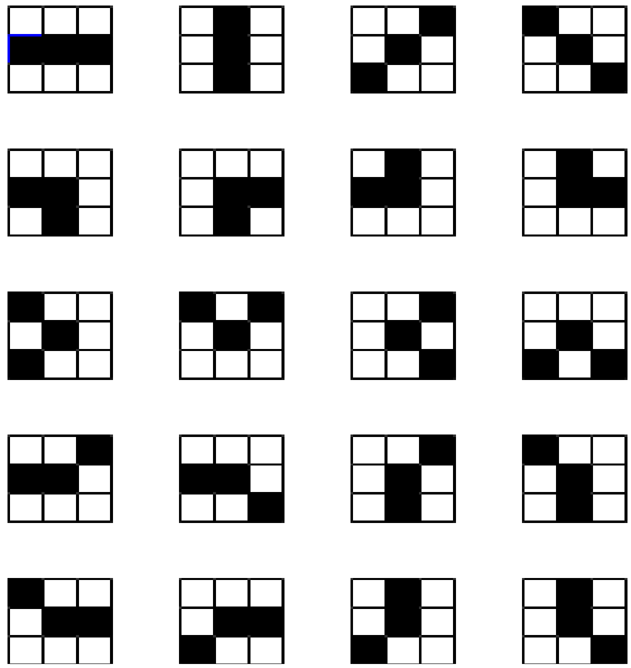


Figure 2. The complete set of 3 × 3 binary masks.

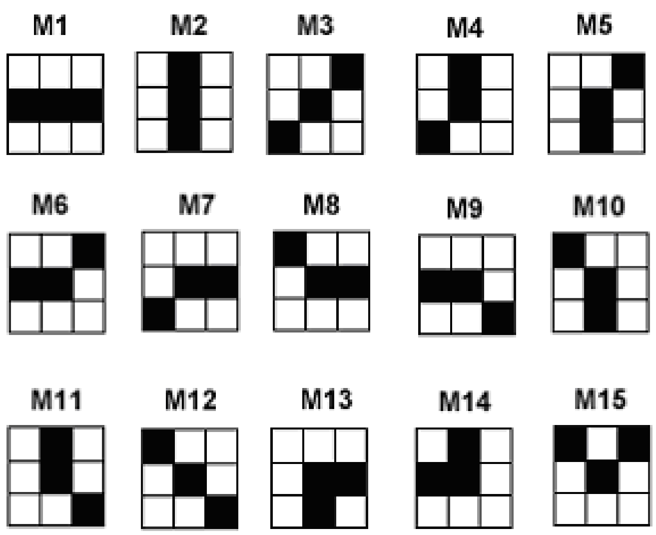


Figure 3. The selected set of 15, 3 × 3 binary masks used for the study.

Step 5. Invoke Rules 3-ext to extract the necessary rules from the set of examples.

For example the frequencies of 15 masks for Signature 1 to 1 were calculated as follows:

416, 53, 1, 2, 3, 39, 41, 23, 24, 13, 16, 10, 9, 2, 4, Signature 1 to 1.

This is an example which represents the Signature 1 to 1. Here the numbers represent the frequencies of Masks 1, 2, ..., 10 respectively. For example 416 means when the Mask 1 is applied to the image being processed it matches exactly 416 times.

Verifying/recognizing unseen signatures

The steps required to verify/recognize an unseen signature are as follows:

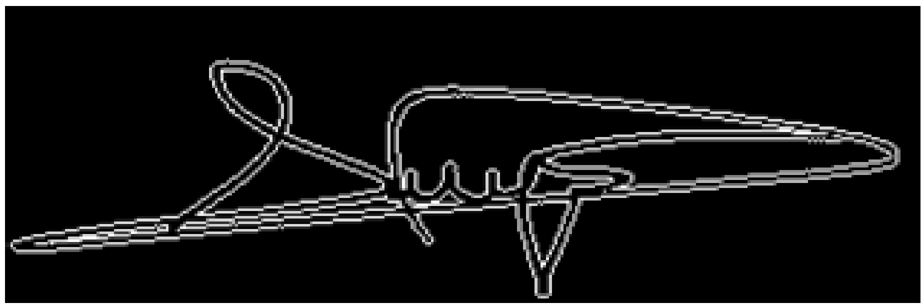
Step 1. Scan a signature (binary image).
Step 2. Perform edge detection operation.
Step 3. Apply 15 binary masks to the image to find their frequencies.
Step 4. Use the set of rules extracted by Rules 3-ext to recognize the signature.

Each rule is tested if it can be satisfied by using the frequencies of the masks used.

RESULTS AND DISCUSSION

The system was trained using 144 signatures (16 signatures belonging to 9 different persons each). A partially set of signatures selected randomly from the complete set of examples used for training are shown in Table 1. In order to find the best set of rules, a number of trials were made by using different number of examples, quantization levels and minimum number of conditions. Table 2 shows these trials. As it can be seen from Table 2, the best efficiency (that is the percentage of correctly classifying unseen or test examples) were reached to 97% when number of training examples were 60 and

(a)



(b)

Figure 4. A signature before and after edge detection operation. (a) A signature before edge detection; (b) A signature after edge detection.

Table 1. The partially set of signatures selected randomly from the complete set of examples used for training.

Frequency															
M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	
416	53	1	2	3	39	41	23	24	13	16	10	9	2	4	Signature1-1
588	110	55	54	52	133	130	39	40	58	59	26	4	4	5	Signature2-8
203	77	465	197	200	178	180	29	27	33	36	18	6	8	2	Signature3-4
428	268	129	78	79	146	147	57	54	68	68	75	4	3	3	Signature4-5
5	245	45	223	220	6	2	7	5	6	2	4	1	1	1	Signature5-13
19	47	328	124	124	47	47	8	8	5	6	9	2	3	1	Signature6-16
35	75	80	128	131	25	25	19	20	13	13	15	2	3	1	Signature7-2
30	597	96	171	169	40	38	66	64	17	17	34	3	2	1	Signature8-8
283	93	55	34	32	71	73	26	22	24	28	22	5	2	2	Signature9-3

Where M denotes mask.

100, quantization levels 9 and 5, minimum number of conditions 10 and 1, and number of test examples 100 and 160 respectively. Table 3 shows some of the rules extracted by Rules 3-ext. As is can be seen from Table 2, the user should make a number of trials (like many other techniques such as neural networks) by changing the parameters (number of learning examples, number of test examples, number of quantization levels and minimum number of conditions for rules) until the system reaches a satisfactory efficiency. As it is very difficult, so far we did not specify a specific way for this purpose.

CONCLUSION

In this paper we propose an alternative technique for signature verification using Rules 3-ext inductive learning system. The advantages of this technique can be summarized as follows:

i) The system uses the advantages of template matching technique. For example it does not need accurate alignment which is a problem when using other techniques in the area.

Table 2. The efficiency of the system using different number of examples, quantization levels and minimum number of conditions.

No. of training examples	Quantization level	Minimum no. of conditions	No. of test examples	Efficiency (%)
60	5	5	100	74
60	10	8	100	74
100	5	2	60	75
60	5	10	100	75
100	5	1	60	76
60	5	7	100	79
100	11	6	60	80
100	11	5	60	80
100	7	6	60	81
60	7	10	100	83
60	9	5	100	89
60	9	10	100	97
100	5	1	160	97

Table 3. Some selected rules extracted by Rules 3-ext.

IF 208 <= M1 < 413 and 7 <= M6 < 44 and 1<= M12< 22, then decision is Signature 1-2.
IF 413 <= M1 < 616 AND 59 <= M8 < 78, then decision is Signature 2-3
IF 280 <= M3 < 373 AND 170 <= M4 < 226, then decision is Signature 3-2.
IF 85 <= M12 < 106, then decision is Signature 4-1.
IF 143 <= M2 < 278 and 226 <= M4 < 282, then decision is Signature 5-5.
IF 3 <= M1 < 208 and 94 <= M3 < 187 and 4 <= M14 < 6, then decision is Signature 7-3.
IF 413 <= M2 < 548, then decision is Signature 8-1.
IF 44 <= M6 < 81 and 25 <= M10 < 46, then decision is Signature 9-3.

ii) This method greatly compresses the image data for storage and reduces the sensitivity of the input intensity data. Each signature is represented by a (number of) rule(s). So the original patterns (signatures) do not have to be stored in the memory as an image. This saves memory space.
iii) The decision can be made in a short time for an unseen example, because the number of conditions in each rule and the total number of rules are not big in number.
iv) It is easy and cheap to develop software and hardware for this technique since it is not complicated. Only a PC and a scanner could be enough (as hardware) to do the job.
v) The efficiency was found to be 97% for unseen examples which is high enough.

The disadvantage of the technique could be not having a specific way to decide the proper parameters to be used for a satisfactory efficiency. This is considered as future work.

ACKNOWLEDGEMENT

The authors would like to thank Research Center in the College of Computer and Information Sciences in King Saud University, Riyadh, Saudi Arabia for their support and fund provided to this work.

REFERENCES

Aksoy MS, Cagil G, Turker AK (2000). Numberplate recognition using inductive learning. Robotics and Autonomous Syst., 33: 149-153.
Aksoy MS (2004). Saudi Banknote Recognition Using Inductive Learning. Proceedings of the 2nd IIEC. Riyadh. Saudi Arabia.
Bajaj R, Chaudhary S (1997). Signature Verification Using Multiple Neural Classifiers. Pattern Recognition, 30: 1-87.
Ballard DH, Brown CM (1982). Computer Vision. Prentice-Hall. New Jersey.
Bovino L, Impdevo S, Pirlo G, Sarcinella L (2003). Multiexpert Verification of Hand-Written Signature. Proc. Int'l Conf. Document Analysis and Recognition, pp. 932-936.
Coetzer J, Herbst BM, Preez JA (2004). Off-Line Signature Verification Using Discrete Radon Transform and a Hidden Markov Model. EURASIP. J. Appl. Signal Proc., 4: 559-571.

Davies ER (1995). Machine Vision. Academic Press.

Draouhard JP, Sabourin R, Godbout M (1996). A Neural Network Approaches to On-Line Signature Verification Using Directional PDF. Pattern Recognition, 29: 415-424.

Guru DS, Prakash HN (2009). Online Signature Verification and Recognition: An Approach Based on Symbolic Representation. IEEE Trans. Patt. Anal. Mach. Intell., 31(6): 1059-1073.

Jain R, Kasturi R, Schunk B (1995). Machine Vision. McGraw Hill International Editions.

Lee LL, Berger T, Aviczer E (1996). Reliable On-Line Signature Verification Systems. IEEE Trans. Patt. Anal. Mach. Intell., 18(6): 643-649.

Mathkour H (2009). RULES3-EXT: Improvements of RULES3 Induction Algorithm. Math. Comput. Appl., 15(3).

Nelson W, Kishon E (1991). Use of Dynamic Features for Signature Verification. *Proc.* IEEE Int' Conf. Syst. Man Cybern., 1: 201-205.

Ozgunduz E, Senturk T, Karsligil ME (2005). Off-Line Signature Verification and Recognition by Support Vector Machine. 13th European Signal Processing Conference. September 4-8. Antalya. Turkey.

Pham DT, Aksoy MS (1995). A new algorithm for inductive learning. J. Syst. Eng., 5: 115-122.

Roland TC, Charles AH (1982). Automated visual inspection: A Survey. IEEE Transactions On Patt. Anal. Mach. Intell., PAMI, 4(6).

Wu QZ, Lee SY, Jou IG (1998). On-Line Signature Verification Based on Logarithmic Spectrum. Pattern Recognition, 31(12): 1865-1871.