Realization of a Signature Verification System Based on Morphological Transformation and Using the TMS320C6713 Digital Signal Processor

E. N. Zois*, E. Pikasis, Z. Kotsidimou, P. Tsitouras, and A. A. Nassiopoulos

Reseacrh.Development.Telecomunication.Laboratory, Electronics Department,
Athens Technological and Educational Institution.
12 Ag. Spiridonos Str. 12210, Aegaleo, Greece
Tel:+30 210 5385383, Fax:+30 210 5385 304,
**corresponding author email: elias@ee.teiath.gr*

Abstract: **In this paper a feature vector, which has been used for curve coding, is evaluated in case of a signature verification scheme using a real time digital signal processor. The feature extraction method is based on morphologically processing the vertical projections of prescaled signature images. Coding of the curve profiles is carried out using morphological opening which explores the pixel allocation along the directions of projection. Various lengths for the structuring element are employed to increase feature discrimination capabilities. The method has been implemented using the Texas Instruments TMS320C6713 digital signal processor in conjunction with the MATLAB and SIMULINK software. The efficiency of the method is examined using a new signature database which comprises 70 writers. The classification approach uses the weighted distance as a similarity measure. An error rate lower that 0.02% is obtained for the case of random and zero effort forgery.**

Index Terms: ***Image analysis, Verification, Signature, System implementation.***

I. INTRODUCTION

A handwritten signature as a behavioral biometric is the mean accepted method to declare someone's identity. Many documents necessitate a handwritten signature [1]. In general, there are two ways to process the signature sample. The first is on-line, where the image is captured directly as handwriting trajectory. The second is off-line, in which we use a digitizer in order to acquire a digital image. This amount of information is highly degraded when compared to on-line procedures but useful information may be recovered with the use of several techniques that have been reported in the domain, others heuristic and others based on ideas from signal processing [2-4].

In signature analysis we use the conclusive evidence from research in psychology and neurophysiology that boundary or contour information is essential for human perception of objects [5, 6]. This kind of information is used for recognition improvement in computer vision tasks such as character recognition, writer discrimination and object identification. Simple or complicated procedures, like dynamic programming, are employed for boundary extraction and contour following [7]. After that, contour feature extraction is performed.

Contour coding is based on local features such as discontinuities, tangent points, end points and curvature [8]. Parametric features like moments, Fourier descriptors, chain coding, polygonal approximation and syntactic representation have been used as well [9, 10]. Generally speaking, contours appear in image analysis in various ways. They can be as simple as a line, or complex as a set of lines, arcs and geometrically established primitives, which is constituted by closed or open curves. Levine [11] has introduced a set of non-geometrical primitives that are useful in applications like document processing and chromosome identification. Thus, letters within a phrase or the trace that a pen produces when a signature or sentence is generated can be modeled by means of the above formulation. Feature extraction from written patterns is very important but a rather difficult task, especially when dynamic information is not available.

Projection functions have been used widely for contour feature extraction [12, 13]. Pavlidis in his survey [13] has stated that projections can be used to deal with the problem of global shape description. There are two major types of projections, namely the Cartesian and the polar projections, which provide a kind of object encoding into a wide family of waveforms. The obtained waveforms are either processed statistically as being a global feature vector (shape profile) or further transformed, in order to extract salient features of the object. Projections have been used for signature image analysis as well as for preprocessing and recognition of handwritten characters (Latin, Asian) and numerals [14, 15].

In this paper, the characteristics and the efficiency of a feature vector for signature verification is realized using a dedicated hardware like a real time digital signal processor. The necessity that dictated us the creation of a hardware based signature verification system is that modern DSP systems provide integration of the entire design process, flexibility due to hardware implementation and portability. A hardware based DSP system outclasses convectional software based signature verification system since it provides dedicated and faster parts for image acquisition, image coding, feature extraction and classifier evaluation. Moreover, a DSP hardware system provides easy interface with other hardware modules like communication. For our application, we have chosen TMS320C6713 DSP floating-

This work is co-funded a) by 75% from E.U. and 25% from the Greek Government under the framework of the Education and Initial Vocational Training Program – **Archimedes** and b) The Research Committee of TEI of Athens under the framework of Project **Athena-04**.

point processor for implementation and performance evaluation. This processor is offered on a low-cost evaluation board (DSK) and the board can be targeted using TI's software tools and utilities such as CCS [16]. As DSPs are used more and more in everyday applications, their study has become imperative at university level. Matlab SIMULINK [21] is a perfect tool for DSP implementation, which allows to the engineer to make an easy transition from theory to practice. The method remains off-line, but we use the combined capabilities of a digital signal processing tool along with the Matlab SIMULINK environment in order to implement the entire image processing. This includes the preprocessing stage along with feature extraction, training and classification.

The feature extraction method is based on morphologically processing the projections of simple or complicated curves which are obtained from the binary signature samples. Morphological openings are applied to one dimensional projection functions in order to control and measure the information from shapes and waveforms by means of granulometries [17]. The length of the line structuring elements (SE), which controls the loss of information, has been selected and kept constant in order to extract the maximum amount of information which describes the pixel distribution. Another issue, which is studied, is the level of image partitioning in order to enhance the classification accuracy. The efficiency of the feature is examined using handwritten signatures on a large signature database.

II. THE DATABASE

The signature database consists of 70 sets of different signatures. Each set consists of 105 genuine and 35 forgery signatures. For the creation of this database each volunteer is asked to sign five pages of their own signatures in order to make the genuine subgroup of the database. The imitated subgroup has been created by selecting randomly persons from the genuine subgroup and asking them to imitate other people's signatures by signing three multiple pages. For creating the imitated subgroup the persons had been provided with one month time in order to train themselves. This is to ensure the highest level of trained forgery from non professional of the kind. The forgery samples represent various levels of imitation, ranging from simple freehand up to skilled. The first case represents the zero effort forgery and the random forgery while the second case represents skilled forgery. The images have been scanned at a resolution of 100 dpi, 8 bit grey scale. Figure 1 provides samples of the acquired images. From an early inspection we can conclude that the database contains signatures of various styles i.e clear and tided, cursive and oriental. A portable PC computer has been used (Pentium 4, 2.8 GHz) for the conduction of the experiments.

III. DESIGN AND IMPLEMENTATION OF IMAGE ANALYSIS AND VERIFICATION SYSTEM

This section describes the image processing algorithm which transforms the original image to a multidimensional feature vector. It also describes the building blocks that have been used in order to realize the proposed algorithm. This is accomplished by using the Matlab SIMULINK packet along with the TI code composer studio software. Finally, this section describes the recognition algorithm along with its hardware realization.

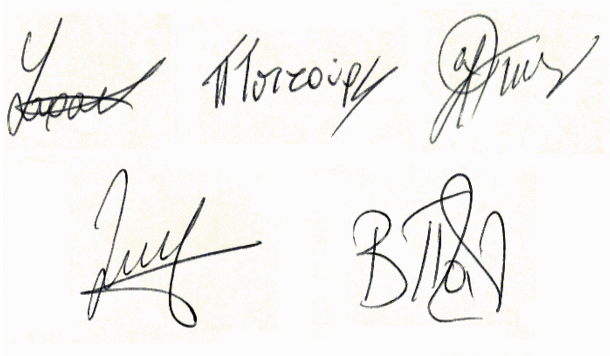


Figure 1. Samples from the signature database. It is observed that the database contains signatures of various styles.

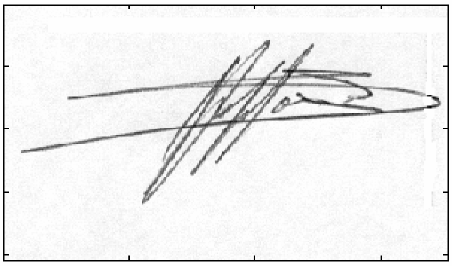
A. Preprocessing.

We are use a 'semi-gross' preprocessing procedure in order to avoid problems which could arise from various sources of the image (signature) creation. Among them we could refer to the line width (type of pen used) as well as to the degree of ink absorption (type of paper used). This semi-gross preprocessing includes thresholding in order to obtain a two-tone (black and white) image and edge detection in order to reject redundant pixels from the image. As a result we have for further elaboration a thinned trace describing the geometrical or structural allocation of the signature. Information which is lost through this preprocessing is not considered essential for describing the information of the geometrical and structural features. The double trace of the image after the edge detection algorithm carries the upper and lower profiles of the signature image, thus resulting to a clear and informative raw image data. Figure 2 provides the original along with the final binary image.

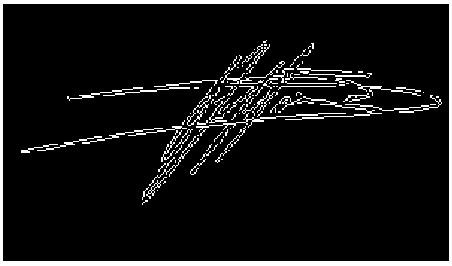
B. Feature extraction method.

Among the various shape descriptors that have been used for handwritten pattern representation and signature analysis are granulometries [18]. A granulometric feature vector is employed to signature representation [17]. It contains spatial information about the orientation of the line segments in a handwritten pattern. Accordingly, the binary image of each word is partitioned into five sub-blocks. Then, the projection function of each sub-block is defined. The vertical projection function f_i^V, i=1, 2, …, 5 (VPF) is defined as the sum of the black pixels with the same abscissa *k* inside the sub-block:

$$f^{V}(k) = \sum \text{ black pixels with abscissa } k \qquad (1)$$



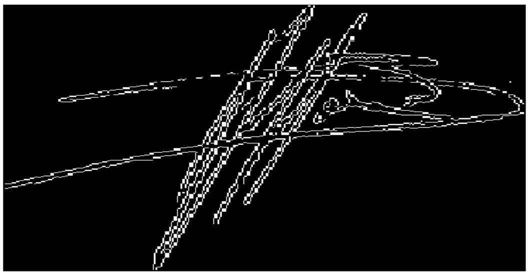
(a)



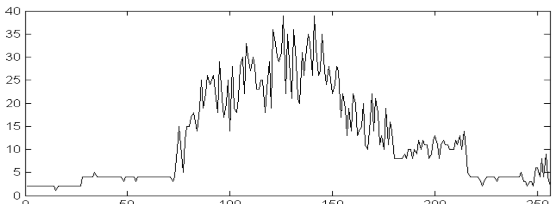
(b)

Figure 2. a) Original image. b) The final binary image.

Figure 3 demonstrates the projection waveform obtained when (1) is applied to a signature sample. The final feature vector is obtained when two successive morphological openings $f_i^{V} \circ g_k$ are performed on the projection functions f_i^{V} with a line Structuring Element (SE) g_k having two different lengths.



(a)



(b)

Figure 3. Binary image along with its corresponding projection. The vertical axis corresponds to the number of white pixels of each abscissa.

Figure 4 illustrates the effect of two morphological openings with SE length three and seven on the vertical projection function of a real image. As a result, the corresponding parameters *e* (fine details) and *c* (coarse details) are derived, which measure the gradual reduction in the area of each waveform according to:

$$e_i^{V} = \left[\frac{\text{mes}(f_i^{V}) - \text{mes}(f_i^{V} \circ g_1)}{\text{mes}(f^{V})} \right] \qquad (2.a)$$

$$c_i^{V} = \left[\frac{\text{mes}(f_i^{V} \circ g_1) - \text{mes}(f_i^{V} \circ g_2)}{\text{mes}(f^{V})} \right] \qquad (2.b)$$

where mes(.) is the area under the function in the argument and f^{V} is the original projection function. The normalization factor f^{V} which corresponds to the number of pixels in the initial image is used in order to achieve finite dynamic range for the obtained feature components as well as to make intraclass dispersion insensitive to natural variations of the genuine class. The set of all parameters:

$$\{ e_i^{V},\ c_i^{V} \},\ i=1,\ldots,5 \qquad (3)$$

constitutes the feature vector corresponding to each signature. It is obvious that the feature space dimensionality is determined by the partition level. In the general case of a $n \times m$ partition, the procedure results in a $4 \times n \times m$ dimensional feature vector.

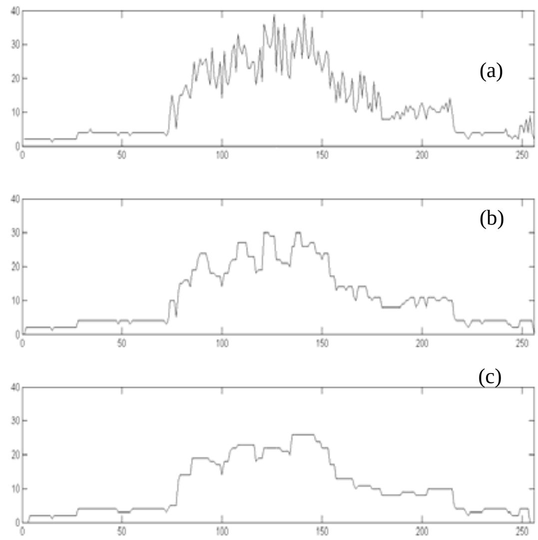


Figure 4. Morphological openings on the projection funtion. a) Initial projection function. b) Opening with line SE g_1 of length 3 provides the e^{V} parameter. c) Opening with line SE g_2 of length 7 provides the c^{V} parameter.

C. System realization.

The procedure followed at the previous section is realized using the combined capabilities of the Simulink environment and the power provided by the TI's C6000 floating point digital signal processor family.

The tool used is the TMS320C6713 starter kit which operates at 225MHz and provides a 16 MB SDRAM for data storage. Figure 5 provides the entire proposed process as it has been imprinted to the SIMULINK model builder.

As it is easily seen, the process of transforming the original grey-level image to a ten-dimensional feature vector is based on various sub-procedures. First, the

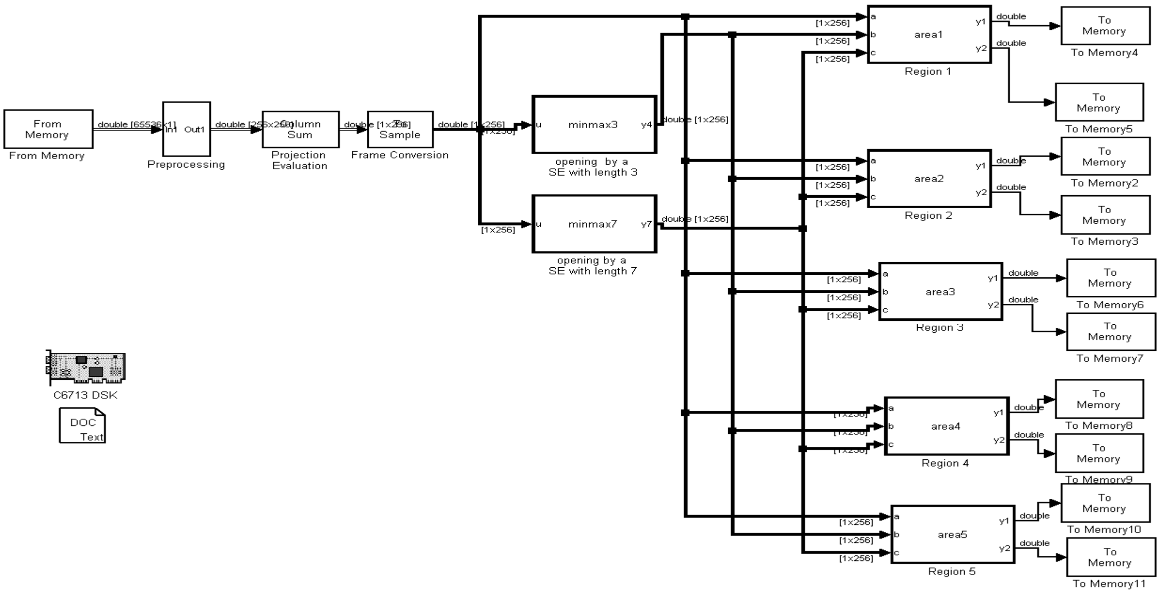


Figure 5. The feature extraction method implemented with the aid of SIMULINK Blocks

processor acquires from a predetermined memory location 65536 elements. This represents the primary image in a grey-level format and corresponds to a size of 256×256 pixels. However, the original signature images are not of this size due to variations to each person signature type and writing style. Therefore, a normalization algorithm must be applied to the family of images prior to DSP acquisition. The algorithm is based on the bounding rectangle [19] and therefore a scale vector is used for resizing this rectangle in both horizontal and vertical axes, resulting to a final image of 256×256 pixels. Another issue is CCS linking limits by means of CCS target memory read and write data. The total transfer should not exceed 32KB which corresponds to a loaded image size of 180×180 pixels. This obstacle has been overcome by using proper code with the MATLAB CCS link [20]. The total writing procedure, of a 256×256 image, in double precision format (8 bytes) takes about 35 seconds and it depends exclusively on software limitations. The 65536 elements that have been stored to the memory are reshaped in order to represent a matrix of 256×256 double precision pixels. Then, the image is transferred to the pre-processing sub-procedure block. As figure 7 shows, the original image within this block is thresholded while edge detection is applied afterwards. The final binary image is the signal which will be processed in order to evaluate the ten-dimensional feature vector. In order to provide the final edges we have convolved the binary image with the Sobel operator.

Then, the projection function f_i^V is evaluated and two morphological filters of length 3 and 7 are applied in order to provide the $f_i^V \circ g_1$ and $f_i^V \circ g_2$. These functions are fed to five processing algorithms which partitions the input signals to five regions. Then, according to (2) the set of parameters $\{e_i^V, c_i^V\}$, i=1,…,5 are evaluated thus resulting the final feature vector. These ten values are stored to different memory locations and a MATLAB based program has been employed in order to acquire them.

IV. Experimental Procedure.

A. *Experimental Protocol.*

The effectiveness of the proposed feature in discriminating signature specimens along with their implementation on the Digital Signal Processing Kit, is evaluated in this section by using a well known classification scheme.

Signature verification is a special case of the writer identification problem in which a person is identified among others based on his handwriting. In other words one must decide if the claimed identity can be accepted based on a specific specimen. Identification (declare the writer class) case-problem is addressed under the *random forgery detection* while the verification-case is studied under the *zero effort and skilled forgery* protocol. In this paper extensive work has been made for the first case while the second case has been examined partially. For each signature owner an individual classifier is designed, trained and evaluated for the case of random forgery. This design strategy offers reduction on the required training samples.

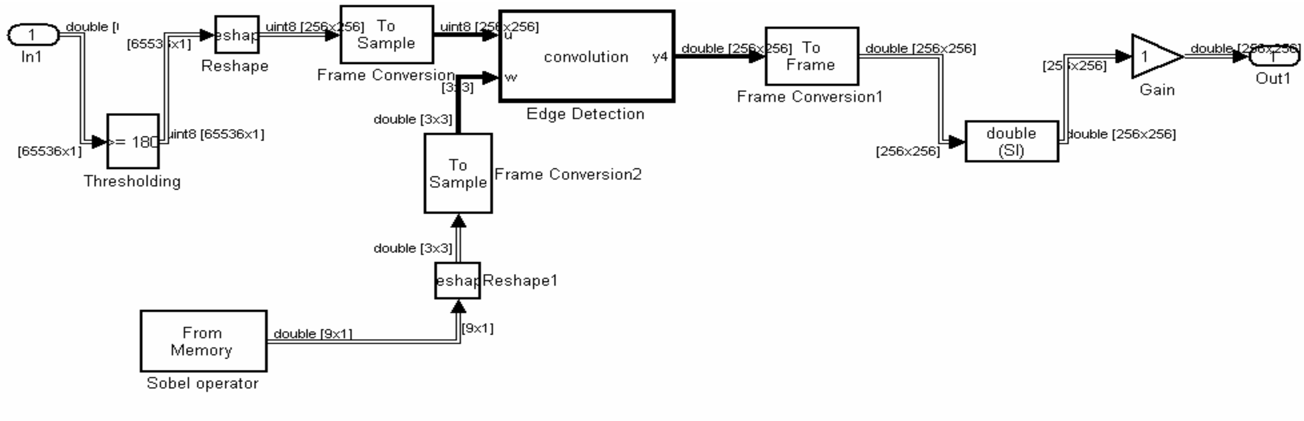


Figure 7. Thresholding and Edge Detection Sub-blocks

Moreover, it can be applied in case that another writer is added to the verification procedure. For the verification stage, the following simplified decision rule is adopted:

$$\phi(H_{1i}:f_d) \underset{H_{0i}}{\overset{H_{1i}}{\gtrless}} \phi(H_{0i}:f_d) \tag{4}$$

where $\phi(H_{1i}:f_d)$ and $\phi(H_{0i}:f_d)$ are expressions of the posterior probabilities of the genuine and forgery classes.

Essential to consider is the choise of the form of the posterior probabilities and the selection of the appropriate number of training and testing samples when a classification system is designed. For all cases, the classification criterion $\varphi(\bullet)$ used is the weighted distance from the center m_i of each cluster

$$d_i = (x-m_i)^t C_i^{-1}(x-m_i) \tag{5}$$

where C_i is the sample covariance matrix of the population. The classification rule assigns the test sample ***x*** into the class with the smallest distance. The covariance matrix is evaluated as follows:

$$C_i = \frac{1}{N_i-1}\sum_{j=1}^{N_i}(x_{i,j}-m_i)(x_{i,j}-m_i)^T \tag{6}$$

The number N_i of the sample population must be larger than the dimension of the feature space which equals 10 to avoid C_i being singular [20]. Equation (6) results to a poor estimate of C_i , unless the number of samples used is considerably larger than *n+1*. For limited training data, the use of the common covariance matrix $C_M = \frac{1}{70}\sum_{i=1}^{70} C_i$ can lead to higher accuracy than the sample estimate, even in the case that the individual C_i s are different. A solution to the problem, especially when the number of dimensions is large, is to use the leave-one-out method (LOOM) [20] in addition to the common covariance matrix. LOOM uses the majority of samples except one, for estimating the mean and covariance matrix of the population using (6), while the remaining sample is used for testing purposes. The LOOM process is repeated for the entire data and an average of the classification rate is evaluated. Thus, from (5) and for each writer we use 104 specimens for training and 1 specimen for testing.

B. Realization Procedure.

For each writer a SIMULINK model has been created that contains the mean vector m_i and their corresponding covariance matrix C_i , by means of its inverse, according to (6). The inverse operation of the covariance matrix has been made using MATLAB commands although there are many algorithms that allow implementation of that operation on a DSP system. The reason for this is mainly to keep as simple as possible the overall procedure. Figure 8 shows a system that has been build for the first writer.

The above procedure shows that for each writer the critical elements that will be stored on the DSK memory are the mean, the inverse covariance along with their corresponding values of image partition and length of the S.E. These elements are completely personal and they are used to describe each writer exclusively. The total SDRAM memory of the DSP kit is 16MB which indicates that all of the 70 writers can be implemented into the DSK. According to figure 7 a unknown specimen is acquired to the input of the DSK. Then, feature extraction is applied according to the material exposed to previous sections and for every writer. The DSK is also supplied with the inverse covariance and the mean values of the i-writer class. Final, the extracted feature is used in order to evaluate both the weighted as well as the Euclidean distances from the mean of i-writer class. After evaluation of all possible classes, a minimum selection algorithm decides the class that the unknown sample should be assigned. The overall error has been evaluated for the case of random forgery below 0.02% when the weighted distance is employed as a similarity measure and below 0.1% when the Euclidean distance is employed. For the case of skilled forgery preliminary results shows that the system overall error is towards 10%. This rate can be further reduced if we employ the horizontal projective profiles. Another way to enhance the overall verification efficiency is to employ

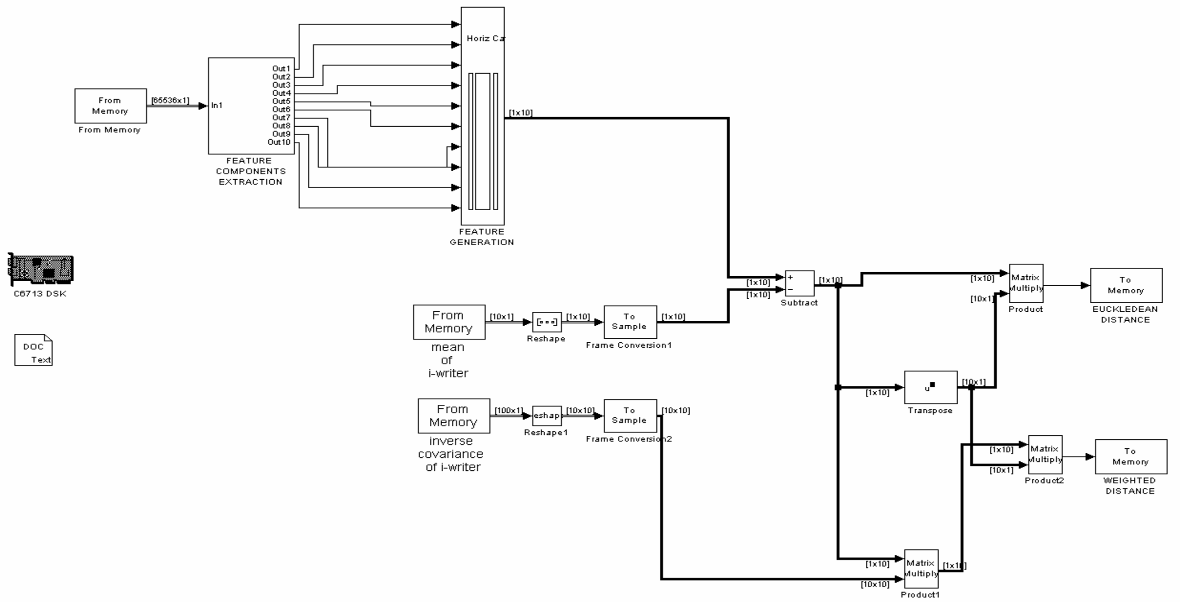


Figure 8. Discrimination Procedure for evaluation of the similarity measure.

fusion techniques using a secret sentence. Typical execution time of the signature verification process is about 1ms for an input size image of 256×256 pixels.

V. Conclusions

An automated handwriting verification system based on signatures has been realized using the combined capabilities of the Texas Instruments C6000 digital signal processing family and the SIMULINK model builder. The feature extraction process transforms the input handwritten signature to a ten-dimensional feature vector based on a non-linear morphological transformation. The similarity measure used is the weighted distance. For each writer a specific model is created and stored to the memory of the DSP. Results provided from a new signature database show that the proposed system is a fast, architecturally efficient and a low cost solution.

Acknowledgment

This work is co-funded a) by 75% from E.U. and 25% from the Greek Government under the framework of the Education and Initial Vocational Training Program – **Archimedes** and b) The Research Committee of TEI of Athens under the framework of Project **Athena-04**.

References

[1] G. Dimauro, S. Impedovo, G. Pirlo, A. Sazo, "Automatic bankcheck processing: a new engineered system", *Int, J. Pattern Recognition Artif. Intell.*, Vol. 11, 4, pp. 467-504, 1997.

[2] R. Plamondon and G. Lorette, ‘Automatic signature verification and writer identification-the state of the art’, Pattern Recognition, Vol. 22, No. 2, pp. 107-131, 1989

[3] R. Plamondon (Ed), "Special issue on automatic signature verification", *Int, J. Pattern Recognition Artif. Intell.*, Vol. 8, 3, pp. 1994.

[4] V. Clement, K. Steinke and R. Naske, "The application of image processing and writer recognition techniques to the forensic analysis of handwriting", *Proc. Int Conf. Secur,*. pp. 5-11, West Berlin, 1980

[5] T. Pavlidis, "Syntactic Pattern Recognition", New York, Springer –Verlag, 1977.

[6] N. A. Stillings, S. Weisler, C. H. Chase, M . H. Feinstein, J. L. Garfield, E. L. Rissland, "Cognitive Science: An introduction" , *MIT Press,* Cambridge, Massachusetts, 1995.

[7] A.A. Amini, T. E. Weymouth, R. C. Jain, "Using dynamic programming for solving variational problems in vision", *IEEE Trans. Pattern Analysis Machine Intell.*, Vol. **12,** 9, pp. 855-867, 1990.

[8] G. R. Wilson, "Properties of Contour Codes", *IEE Proc. Vision Image and Signal Processing*, Vol. **144**, 3, pp. 150-161, 1997.

[9] M. Ammar, Y. Yoshida, and T. Fukumura,, "Structural description and classification of signature images", *Pattern Recognition*, Vol. **23**, 7, pp. 697-710, 1990.

[10] S. W Lee and J. S. Park, "Nonlinear shape normalization methods for the recognition of large set handwritten characters", *Pattern Recognition*, Vol. **27,** 7, pp. 985-902, 1994.

[11] M. Levine, "Vision in Man and Machine", McGraw- Hill, 1985.

[12] C. S. Fuh and H. B. Liu, "Projections for Pattern Recognition", *Image and Vision Computing*, Vol. 16, pp.677-687, 1998.

[13] T. Pavlidis, "Algorithms for shape analysis of contours and waveforms", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-2, 4, pp. 301-312, 1980.

[14] R. Nagel, A. Rosenfeld, "Computer detection of freehand forgeries", *IEEE Trans. Comp.*, Vol. 26, 9, pp.895-904, 1977.

[15] T. Hildrbrand, W. Liu, "Optical Recognition of handwritten chinese characters: advances since 1980", *Pattern Recognition*, Vol. 26, 2, pp.205-225, 1993.

[16] Texas Instruments, "Code Composer Studio User’s Guide", SPRU328B, February 2000.

[17] E. N. Zois and V. Anastassopoulos, "Morphological waveform coding for writer verification". *Pattern Recognition,* Vol. 33, pp. 385-398, 2000.

[18] J. Serra, "Image analysis and mathematical morphology", *Academic Press, New York*, 1982.

[19] A. Jain, "Fundamentals of Digital Image Processing", *Prentice Hall*, London, 1989

[20] R. Duda and P. Hart, "Pattern Recognition and Scene Analysis", *John Wiley and Sons*, 1972.

[21] The Mathworks, "MATLAB Link for Code Compo-ser Studio Development Tools, User’s Guide", 2002