This article presents selected research on the development of complex fundamentals of building intelligent interactive systems for design of machine elements and assemblies on the basis of its features described in a natural language. We propose a new method for handwriting recognition that utilizes geometric features of letters. The article deals with recognition of isolated handwritten characters using neural networks. As a result of the geometrical analysis, graphical representations of recognized characters are obtained in the form of pattern descriptions of isolated characters. Selected parameters of the characters are inputs to the neural network for writing recognition which is font independent. In this article, we present a new method for off-line natural writing recognition and also describe our research and conclusions on the experiments.

**Key words:** human-machine interaction, intelligent automated design, intelligent system, handwriting recognition, artificial neural networks, computer aided design

1. INTRODUCTION

The presented research [6] involves the development of complex fundamentals of building intelligent interactive systems for design of machine elements and assemblies on the basis of its features described in a natural language.

In machine automation technology, the problem of writing recognition is quite complex [3], and even now there is no single approach that solves it both efficiently and completely in all contexts. In written language recognition processes, an image containing text must be appropriately supplied and preprocessed. Then the text must either undergo segmentation or feature extraction. Small processed pieces of the text will be the result, and these must undergo recognition by the system. Finally, contextual information should be applied to the recognized symbols to verify the result. Artificial neural networks, applied in handwriting recognition, allow for high generalization ability and do not require deep background
knowledge and formalization to be able to solve the written language recognition problem. Handwriting recognition can be divided by its input method into two categories: off-line handwriting recognition and on-line handwriting recognition. For off-line recognition, the writing is usually captured optically by a scanner. For on-line recognition, a digitizer samples the handwriting to time-sequenced pixels as it is being written. Hence, the on-line handwriting signal contains additional time information which is not present in the off-line signal.

In the proposed new method [7] of natural writing recognition in Figure 1, the handwritten text is subject to the following preprocessing: digitization, binarization, noise elimination, thinning, normalizing and segmentation. The next step is to find the center of mass of the character image. With the center of mass as a reference point, radiuses are drawn, creating a set of points describing the contour of the character so that its pattern description is made. In the proposed hybrid system, the pattern description of each isolated character, after the process of input value normalization and application of letter description rules using fuzzy logic, are the input signals for probabilistic neural networks for isolated character recognition. The recognized characters are grouped into more quantitative units with the letter string recognition module, which are coded as binary images of vectors and then become inputs of the module for recognizing words. The module uses a 3-layer Hamming neural network [11, 12]. The network of this module uses a training file containing patterns of words. The recognized vocabulary words represented by the output neurons are processed by the module for recognizing phrases which uses the Hamming Maxnet network equipped with a training file containing phrases built with contextual knowledge from linguistics.
2. THE STATE OF THE ART

The state of the art of automatic recognition of handwriting at the beginning of the new millennium is that as a field it is no longer an esoteric topic on the fringes of information technology, but a mature discipline that has found many commercial uses. On-line systems for handwriting recognition are available in hand-held computers such as personal digital assistants. Their performance is acceptable for processing handprinted symbols, and when combined with keyboard entry, a powerful method for data entry has been created.

Off-line systems are less accurate than on-line systems. However, they are now good enough that they have a significant economic impact on specialized domains such as interpreting handwritten postal addresses on envelopes and reading courtesy amounts on bank checks [1, 2, 3, 9, 10, 13, 19].

The success of on-line systems makes it attractive to consider developing off-line systems that first estimate the trajectory of the writing from off-line data and then use on-line recognition algorithms [15]. However, the difficulty of recreating the temporal data has led to few such feature extraction systems so far [1].

Research on automated written language recognition dates back several decades. Today, cleanly machine-printed text documents with simple layouts can be recognized reliably by OCR software. There is also some success with handwriting recognition, particularly for isolated handprinted characters and words. For example, in the on-line case, the recently introduced personal digital assistants have practical value. Similarly, some online signature verification systems have been marketed over the last few years and instructional tools to help children learn to write are beginning to emerge. Most of the off-line successes have come in constrained domains, such as postal addresses, bank checks, and census forms. The analysis of documents with complex layouts, recognition of degraded printed text, and the recognition of running handwriting continue to remain largely in the research arena. Some of the major research challenges in on-line or off-line processing of handwriting are in word and line separation, segmentation of words into characters, recognition of words when lexicons are large, and the use of language models in aiding preprocessing and recognition. In most applications, machine performance is far from being acceptable, although potential users often forget that human subjects generally make reading mistakes [2, 3, 5].

The design of human-computer interfaces [6, 7, 8, 12] based on handwriting is part of a tremendous research effort together with speech recognition, language processing and translation to facilitate communication of people with computers. From this perspective, any successes or failures in these fields will have an important impact on the evolution of languages [4, 14].
3. DESCRIPTION OF THE METHOD

The proposed system [6] attempts to combine two methods for natural writing recognition, neural networks and preprocessing for geometric features extraction. The system consists of the preprocessing subsystem, geometrical analysis subsystem, neural network subsystem for isolated characters as well as neural network subsystem for vocabulary and linguistics, as shown in Figure 2. The motivation behind that preprocessor is to reduce the dimensionality of the neural network input. However, another benefit given by the preprocessor is immunity against image translation, because all the information is relative to the image's center of mass.

The developed geometrical analysis is based on the processing of the images of letter shapes into their graphical representations in the form of pattern descriptions. The process of the geometrical analysis begins with determining of the center of mass of a letter with a gravity method in order to find the initial point of the analysis. The next step of the algorithm is based on drawing radiuses from the initial point, the lengths of which are equal to the length of the line segment created by the initial point and the point on the letter furthest from this point. The creation of a circle of that radius makes it visible that the analysis covers the whole letter. The precision of this geometrical analysis method is proportional to the number of radiuses.

Where the radiuses intersect with the letter, points are obtained, which makes it possible to obtain the measures of the line segment created by the initial point and the letter intersection point. The lengths of the created line segments obtained are represented in the form of pattern descriptions of isolated characters which are inputs of the probabilistic neural network. Geometrical analyses of characters for exemplary letters are shown in Figure 3, which are also used by probabilistic neural networks to recognize isolated characters.

The architecture of modified probabilistic neural networks for recognition of pattern descriptions of isolated characters is shown in Figure 4. It is composed of interconnected neurons organized in successive layers. The probabilistic neural network was first introduced by Specht [16, 17, 18]. The probabilistic network consists of input, pattern, summation and output layer.

In the proposed handwriting recognition system, because of the binary input signals, the Hamming neural network is chosen for both the word recognition and phrase recognition [11, 12] as shown in Figure 2. The network directly realizes the one-nearest-neighbor classification rule.
Figure 2. Methodology of the proposed system for handwriting recognition [6]
Figure 3. Geometrical analyses of characters: A) determination of the center of the mass for exemplary letter s; B) determination of intersection points of the letter and the radiiuses for exemplary letter o; C) measurement of the length of line segments \( l \) created by letter points in fragments \( f \) for letter e; D) summation of measurements in fragments \( f \) containing \( n \) radiuses for letter l; E) measurement of the length of line segments of each radius for exemplary letter e and letter l (F); G) measurements of differences of the radius lengths in each fragment \( f \) for exemplary letter e and letter l (H)
Figure 4. The architecture of the probabilistic neural network for recognition of pattern descriptions of isolated characters

4. EXPERIMENTAL RESULTS

The research on the developed method concerns the ability of the neural network to learn to recognize specific letters. The neural networks are trained with the model of isolated written language characters. Several geometrical analyses of isolated characters and their pattern descriptions were realized (Figure 5), which made it possible to draw significant conclusions and apply them in the proposed algorithms. The ability of the neural network to learn to recognize specific characters depends on the type of parameters of geometrical features. The specified type of parameters enables the network to minimize the error so that it can work more efficiently. Based on the research, the following parameters are the most significant for the recognition, as shown in Figure 6.
Figure 5. Geometrical analysis and pattern description of isolated characters

Figure 6. Geometrical analyses of characters for recognition of handwriting using neural networks
5. CONCLUSIONS AND PERSPECTIVES

The presented research involves the development of complex fundamentals of building intelligent interactive systems for design of machine elements and assemblies on the basis of its features described in a natural language. The advantages of this new method of natural writing recognition are flexibility with regards to writing style, geometrical analysis enabling font independent character recognition, possibility of application of other types of neural networks, extension of the range of geometrical analysis and other possibilities for further development.

ACKNOWLEDGEMENT

This project was financed from the funds of the National Science Centre (Poland) allocated on the basis of the decision number DEC-2012/05/B/ST8/02802.

REFERENCES

Streszczenie

W artykule przedstawiono wybrane prace badawcze dotyczące podstaw budowy inteligentnych systemów interakcji do projektowania elementów i zespołów maszyn na podstawie ich cech opisywanych w języku naturalnym. Zaproponowano nową metodę rozpoznawania pisma odręcznego, w której wykorzystano geometryczne cechy znaków. Artykuł dotyczy rozpoznawania izolowanych znaków pisma odręcznego za pomocą sieci neuronowych. W wyniku analizy geometrycznej otrzymuje się reprezentacje graficzne rozpoznawanych znaków w postaci opisów wzorców pojedynczych znaków. Wybrane parametry znaków stanowią wejścia sieci neuronowej do rozpoznawania pisma niezależnego od kroju. W artykule przedstawiono nową metodę rozpoznawania pisma naturalnego, a także opisano badania i podano wnioski wynikające z eksperymentów.

Słowa kluczowe: interakcja człowiek–maszyna, inteligentne zautomatyzowane projektowanie, inteligentny system, rozpoznawanie pisma odręcznego, sieci neuronowe, CAD