Robust Hough-Based Symbol Recognition Using Knowledge-Based Hierarchical Neural Networks

Alexander Wong[1] **and William Bishop**[2]
[1]Systems Design Engineering, University of Waterloo, Waterloo, Ontario, Canada
[2]Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada

Abstract - *A robust method for symbol recognition is presented that utilizes a compact signature based on a modified Hough Transform (HT) and knowledge-based hierarchical neural network structure. Relative position and orientation information is extracted from a symbol image using a modified Hough Transform (HT). This information is transformed and compressed into a compact, 1-D signature vector that is invariant to geometric transformations such as translation, rotation, scaling, and reflection. The proposed method uses a knowledge-based hierarchical neural network structure to reduce the complexity of the recognition process by effectively segmenting the search space into smaller and more manageable clusters based on a priori knowledge. The method achieved overall recognition rates of 96.7% on line graphic symbols from the GREC'05 symbol database under various models of image degradation and distortion.*

Keywords: *hierarchical neural networks, hough transform, symbol recognition*

1 Introduction

Symbol recognition is an important area in computer vision that involves the identification of symbols in an image or video. Symbol recognition is used in a large number of different applications such as:

- interpreting and converting scanned engineering drawings [4] and circuit diagrams [1] into other electronic document formats
- identifying and locating trademarked content [2, 3]
- querying images from databases based on shape [5], and
- recognizing characters and words within an electronic document [6, 7].

A number of issues make symbol recognition a difficult problem to solve. First, symbols in an image or video often undergo various geometric transformations such as translation, rotation, scaling, and reflection. The problem is further complicated by image degradation introduced by noise and lighting conditions. These issues make it difficult, if not impossible, to perform direct pixel-to-pixel similarity comparisons. Furthermore, symbols typically contain little or no color or texture information that can be used to distinguish between different symbols. This is particularly true for technical drawings such as circuit diagrams and engineering drawings. When circuit diagrams and engineering drawings are scanned, the resulting files are typically stored as binary raster images. Under such circumstances, a shape-based similarity method is needed to perform symbol recognition.

This paper introduces a new symbol recognition method that addresses the aforementioned issues. The binary symbol image is transformed into a modified Hough-space where relative position information and orientation information is obtained. The information in the modified Hough-space is transformed and compressed into a compact, 1-D feature vector that acts as the signature of the symbol image. The signature is fed into a knowledge-based hierarchical neural network structure along with optional a priori knowledge about the symbol to obtain to appropriate classification. The knowledge-based hierarchical neural network structure is used to reduce the complexity of symbol recognition by segmenting the search space based on a priori knowledge.

The main contribution of this paper is a new robust method for symbol recognition. Related work conducted in the area of symbol and shape similarity evaluation is presented in Section 2. The theory underlying the proposed system is described in detail in Section 3. The testing methods and experimental results is presented in Section 4. Finally, conclusions are drawn in Section 5.

2 Previous Work

A large number of approaches have been proposed for the purpose of symbol and shape similarity evaluation. A popular method of determining shape similarity is the use of Fourier descriptors [8, 9]. To obtain Fourier descriptors, the boundary contour of an object is represented in the form of a shape signature vector and transformed using the Fourier Transform. The resulting Fourier coefficients are used as a set of Fourier descriptors. There are a number of issues associated with using Fourier descriptors. First, a closed contour is required for Fourier descriptors to work. However, this is not always possible due to factors such as discontinuities caused by noise. Second, details contained within an object are effectively ignored by Fourier descriptors. Finally, since the object is represented by a single contour, Fourier descriptors cannot easily be used to represent objects that consist of multiple disjoint components. These two issues are particularly problematic in the context of symbol recognition since symbol objects often consist of multiple parts, some of which may be enclosed within another.

Another popular method for determining shape similarity is the use of moment invariants [10, 11]. In the simplest terms, moment invariants can be seen as moment-based descriptors that are invariant to deformations such as translation, scaling, and rotation. Moment invariants are region-based descriptors that are most applicable to situations where objects with filled regions exist [17]. This is often not the case in symbol recognition, particularly for symbols used in engineering drawings.

Other symbol and shape similarity techniques that have been introduced include: graph-based matching [12], heuristic-based methods [1], pixel-level constraints histograms [13], Force-based angle histograms [14], curvature scale space methods [15, 16] and methods based on the Hough transform and the related Radon transform [5, 17, 18].

The proposed method is most closely related to the Hough-based techniques proposed by Franti et al [17] and Vlachos et al [5]. In both methods, the orientation information from the resulting Hough representations are used to form a feature vector. The similarity between two symbol images is then determined using the Sum of Squared Differences (SSD) measure between the two images, and then finding the minimum distance measure. There are several problems with the two approaches. First, the methods are not reflection invariant. Furthermore, the similarity processes for both methods do not scale well as the number of symbol images in the system is increased. This is particularly true for the Franti method, which needs to evaluate the SSD measure for all possible rotations of a feature vector. Second, Finally, the methods do not account for the effects of noise and degradation. The proposed system addresses all of these issues in an efficient and robust manner.

3 Proposed Symbol Recognition Method

To narrow the scope of the paper, the proposed symbol recognition method focuses primarily on the signature extraction and classification of symbol images. Therefore, it is assumed that the input image is a binary representation of a symbol that has been segmented properly from the source image (i.e., technical drawing, circuit diagram). As a preprocessing step, the percentage of foreground pixels is calculated to determine whether a median filter can be used for noise reduction. Only if the percentage of foreground pixels exceed 10% of the input image is a median filter applied.

3.1 Signature Extraction

Signature extraction is one of the two key steps performed by the proposed method. The signature extraction process uses a modified Hough transform to transform a binary image in the x-y coordinate space into a relative $|\rho|$-θ parameter space. Then, further transformations are applied to make the resulting signature invariant to scaling and rotation.

3.1.1 Relative $|\rho|$-θ Hough Transform

The Hough transform is a common method used to describe the curve features of an image such as lines and circles. One major advantage of using the Hough transform for representing and identifying curve features is that it is highly robust against image noise and degradation. In the proposed signature extraction process, the classical Hough transform for line representation is modified to construct a global structural representation of a symbol that is invariant to translation and reflection. In the classical Hough transform for line representation, the parameters ρ and θ of a given line are defined relative to the normal vector between the line and the origin at (x, y)=$(0, 0)$. A translation of the symbol with respect to the origin results in a different normal vector and thus a different value of ρ. To make the resulting parameter space invariant to translation, the parameters ρ and θ of a given line can be defined relative to the centroid of the symbol object rather than the origin. Therefore, the modified line equation becomes

$$(x - x_c)\cos(\theta) + (y - y_c)\sin(\theta) = \rho \qquad (1)$$

where x_c and y_c are the x coordinates and y coordinates of the centroid of the symbol object. The centroid of the symbol object is approximated by finding the center of mass of the foreground pixels (each given uniform weight) in the input image.

Like the classical Hough transform, the resulting symbol representation in ρ-θ parameter space is sensitive to reflections. There are two factors that contribute to this sensitivity to reflection. First, the value of parameter ρ may hold either a positive and negative value depending on the orientation of the normal vector. Therefore, a positive value of ρ may become a negative value after reflection. Second, the orientation of the normal vector may change due to the reflection, leading to the change in the value of θ.

The approach taken in the proposed signature extraction process is to first remove the effect of reflection on the parameter ρ by utilizing the absolute value of ρ. The resulting parameter space is referred to as relative $|\rho|$-θ space, and is invariant to the effects of reflection and translation on $|\rho|$. The effect of reflection on the parameter θ is addressed in Section 3.1.2. An example of the resulting relative $|\rho|$-θ space for a given symbol image is shown in Fig. 1.

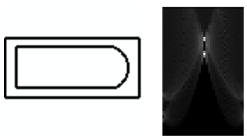

Figure 1. Example of the relative $|\rho|$-θ space representation of a symbol image.

To illustrate the advantage of using relative $|\rho|$-θ space for technical symbols, several circuit symbols and their reflected counterparts are shown in Fig. 2. Two different parameter space representations are created when the classical Hough transform is applied to the two symbol images. However, the same parameter space representation is created when the modified transform is used. This is important in technical drawings like circuit diagrams where the symbols may be flipped.

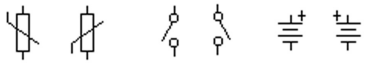

Figure 2. Examples of circuit symbols (and their reflected counterparts) that result in different parameter space representations when the classical Hough transform is applied.

3.1.2 Rotation Invariance

To obtain a symbol image signature that is invariant to the effects of rotation, the approach taken in the proposed method is to find a way to project the information from the relative $|\rho|$-θ space onto the $|\rho|$-axis. For the proposed algorithm, the resulting signature vector is constructed using the following equation

$$v(\rho) = \frac{1}{m} \sum_{\theta=0}^{\pi} A\left(|\rho|, \theta\right)^2 \tag{2}$$

where $A(|\rho|, \theta)$ is the value in the accumulator array at $(|\rho|, \theta)$ and m is the sample mean of $v(|\rho|)$ over all values of $|\rho|$. The values of $|\rho|$ range from 0 to the maximum distance from the centroid of the symbol object to the boundaries of the image. The projection approach is similar to that presented by Vlachos et al [5], with the major difference being that Vlachos et al performed a projection onto the θ-axis, thereby retaining explicit orientation information as opposed to explicit position information.

The explicit use of relative position information was chosen due to its benefits over orientation information with regards to strict symbol recognition. While orientation information provides a good generic shape representation that is effective for retrieving symbols with similar shapes, it does not provide a high level of discrimination between symbols with similar angular characteristics. A high level of discrimination is essential for the purpose of strict symbol recognition.

A number of important issues are addressed by performing the above projection. First, since orientation is not explicitly represented by the resulting signature vector, the symbol representation is made invariant to the effects of rotation and reflection. Yet, since a sum of squares approach is used to project the data onto the $|\rho|$-axis, the influence of angular distribution is still represented implicitly. By preserving the influence of angular distribution, the proposed method is able to handle situations where orientation plays a greater role in differentiating between symbol images. Therefore, invariance is gained using the proposed method without losing too much information that discriminates between individual symbol images. Finally, by compressing the information from the two-dimensional parameter space into a one-dimensional vector representation, the symbol signature is much more compact allowing for fast symbol similarity evaluation.

3.1.3 Scale Invariance

After a signature vector is obtained from the $|\rho|$-θ parameter space using the aforementioned projection, the signature vector is made invariant to the effects of scaling by performing the following modifications. First, the bounding distance of the symbol signature is determined as the maximum value of $|\rho|$ that contains 1% of the maximum of the signature vector. The signature vector is then truncated to the range of $|\rho|$ values between 0 and the bounding distance.

Finally, the truncated signature vector is normalized and compressed into a fixed number of bins to provide a scale-invariant representation of the symbol image. At this point, the signature vector is invariant to translation, scaling, rotation, and reflection. The bin size of the signature vector is selected such that it is small enough to provide a high level of discrimination between symbols, while large enough to be robust to the effects of noise and degradation. Based on the results of experiments with different types of symbols, it was determined that 30 bins for each signature vector provides good discrimination for neural network classification while maintaining robustness to the effects of noise and minor geometric distortions.

3.2 Knowledge-Based Hierarchical Neural Network

Once the symbol signature has been determined, the signature is fed into a knowledge-based hierarchical neural network structure to determine the best possible symbol match for the signature. The Hough-based methods proposed by Franti et al [17] and Vlachos et al [5] perform an exhaustive similarity analysis between the input signature and all signatures stored within a symbol signature database. For large symbol databases, an exhaustive search is intractable. To address this issue, the proposed method utilizes a neural network structure that can perform faster symbol recognition once the system has been trained. Furthermore, a neural network structure typically consumes less storage space when compared to instance-based methods used in the other Hough-based methods, which requires the storage of each instance trained. Finally, neural networks provide better generalization than these instance-based methods. Therefore, the proposed method scales better from a performance perspective than the other Hough-based methods when a large number of symbols are used.

One problem with the use of a global neural network for symbol recognition is that they get highly complex and hard to interpret for large number of symbol classes. To help alleviate this problem, the approach taken in the proposed method is to segment the symbol classes that needs to be learned into smaller clusters. The smaller clusters can then be used to train a collection of smaller neural networks and subsequently connected as a acyclic graph to form a hierarchical neural network structure [19, 20]. There are a number of advantages to using a hierarchical neural network structure. First, smaller local neural networks are easier to train than a single large global neural network. Furthermore, the addition and removal of symbol classes does not require the entire network structure to be retrained, making it more flexible than the use of a single global neural network.

An extension made by the proposed method to the hierarchical neural network structure is the integration of optional a priori knowledge about the symbol to adjust the flow of the structure. In a traditional hierarchical neural network structure, signal propagation always begins at a fixed starting point through the acyclical graph. One issue with this is that classification errors at the earlier neural networks in the graph can lead to signal propagation down the wrong path later in the graph. To help reduce such problems, the proposed method utilizes optional a priori knowledge to start signal propagation at a more appropriate point in the graph. For example, an engineer who scans in an analog circuit diagram for conversion into a digital form knows that the symbols within the drawing are analog circuit components such as resistors and capacitors and not digital circuit components. This type of knowledge can be used to reduce the search space by starting signal propagation at a later point in the graph.

3.2.1 Architecture

The proposed symbol classification process is consists of a collection of neural networks organized into a hierarchical tree structure. An example structure is shown in Fig. 3. Each group of symbols can be represented by a node in the tree, with subgroups being represented as child nodes. Each leaf node consists of a neural network that is trained to handle classification for a group of symbols. Each non-leaf node consists of a neural network that decides between which of its child nodes to propagate to. If no a priori knowledge is available for the input symbol image at execution time, signal propagation is initialized at the root node of the tree structure. However, if a priori knowledge is available, signal propagation is initialized at the appropriate node of the tree structure. This narrowing of search space helps guide the classification process towards the correct classification.

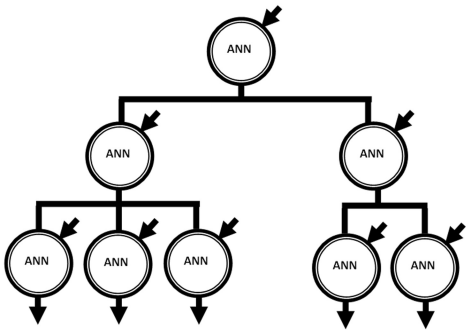

Figure 3. Sample knowledge-based hierarchical neural network structure

4 Experimental Results

To evaluate the effectiveness of the proposed symbol recognition method, tests where performed using symbols provided at the 2005 International Symbol Recognition Contest (ISRC 2005) held during GREC'05. This set of tests has been widely used to test symbol recognition performance and serves as a good test environment. The implementation of the proposed method utilizes multi-layer perceptron neural networks trained with the standard back-propagation algorithm. A total of 25 architectural and engineering symbols used during the contest were selected to use as a basis for the test sets. Some examples of the symbols are illustrated in Fig. 4.

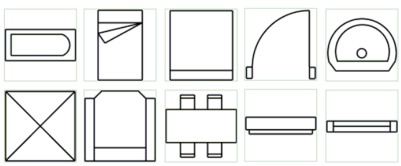

Figure 4. Examples of symbols from the ISRC 2005 symbol database.

Three test data sets were then generated based on the symbols chosen. These test data sets consisted of 50 symbol images with various types and levels of noise and degradation, as well as different rotation and scaling transformations. The noise and degradation model used was based on the document degradation model proposed by Kanungo et al [22]. This model is defined by a set of 6 parameters: $\{\alpha_0, \alpha, \beta_0, \beta, \eta, \kappa\}$. The parameters used for each of the three test data sets are summarized as follows:

- **TEST1**: $\{\alpha_0, \alpha, \beta_0, \beta, \eta, \kappa)\}$ = (0.8, 0.5, 0.02, 0.001, 0, 1.0)
- **TEST2**: $\{\alpha_0, \alpha, \beta_0, \beta, \eta, \kappa)\}$ = (0.5, 0.5, 0.5, 0.03, 0.1, 0)
- **TEST3**: $\{\alpha_0, \alpha, \beta_0, \beta, \eta, \kappa)\}$ = (0.5, 0.5, 0.8, 0.01, 0, 0)

Examples from each of the test data sets are shown in Fig. 5. The test data sets were chosen due to the high level of noise and distortion exhibited in the test images in these data sets, with some test images that are difficult even for a human expert to recognize. TEST1, TEST2, and TEST3 are good representations of replacement noise, salt and pepper noise, and diffusive noise respectively. The level of difficulty associated with these test data sets make them well suited for testing the robustness of the proposed system. The three test data sets provided a combined total of 150 test images. To quantitatively measure the effectiveness of the proposed system, each image in the test data sets was evaluated against each symbol in the symbol database and the best match was determined. Based on the experimental results, the recognition accuracy of the proposed system for each data set was calculated as the percentage of correct matches made out of the total number of test images in the data set.

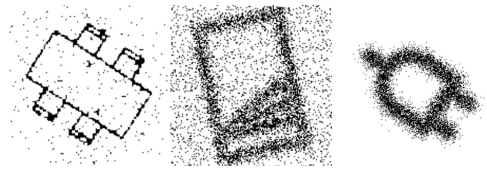

Figure 5. Test image from: a) TEST1, b) TEST2, c) TEST3

The recognition accuracy of the proposed method was compared to results obtained using the state-of-the-art symbol similarity algorithm described by Vlachos et al [5] which has been shown to have a significant performance advantage over the technique proposed by Franti et al [17]. The algorithm proposed by Vlachos et al has been shown to be competitive with widely used matching algorithms such as Chamfer distances [23] and Hausdorff distances [24]. The results of the comparison are presented in Section 4.

A second set of tests was conducted to perform parametric evaluation of the proposed method for symbol recognition under different levels of noise and degradation using hand-drawn symbol images. The two noise and degradation models used were:

1. salt-and-pepper noise, and
2. replacement noise.

The level of salt-and-pepper noise was controlled by a parameter α, which indicates the probability (in percent) that a pixel in the image is flipped. The level of replacement noise is controlled by a parameter β, which indicates the probability (in percent) that a pixel in the actual foreground symbol object is flipped. The test data set consists of hand drawn symbols based on the same set of model symbols from the ISRC 2005 symbol database. The symbol images were rotated randomly prior to the introduction of noise and distortions. The recognition accuracy of the proposed system was evaluated at increasing levels for each noise model, as well as at increasing levels of skew distortion. Examples of the symbol images degraded using the different noise models are shown in Fig. 6.

The recognition accuracy results for the first set of tests are shown in Table 1. It can be observed that the proposed method achieves a relatively high level of recognition accuracy. The proposed system significantly outperforms the algorithm described by Vlachos et al [5] in all of the test

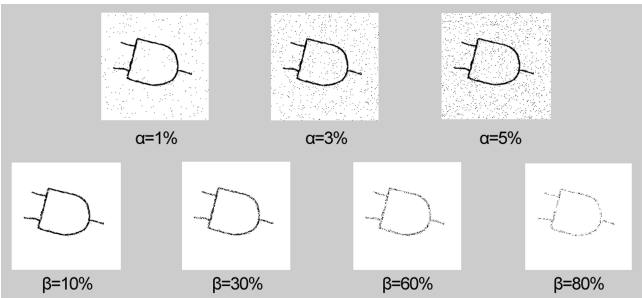

Figure 6. Top: Sample test hand drawn images with salt-and-pepper noise degradation Bottom: Sample test images with replacement noise degradation.

cases. The proposed method achieves an average recognition accuracy of 96.7%, while the Vlachos technique is able to achieve an average recognition accuracy of 58.7%.

Based on the high level of accuracy achieved by the proposed system, it is clear that the system is robust against geometric transformations such as rotation and scaling. Furthermore, it is evident that the proposed system is robust against noise and degradation, since it is capable of maintaining relatively high recognition accuracy under high levels of noise and degradation.

Table 1. Symbol Recognition Accuracy

Test Set	Recognition Accuracy (%)	
	Vlachos Technique	Proposed System
TEST1	64%	100%
TEST2	64%	98%
TEST3	48%	92%
Average	58.7%	96.7%

The recognition accuracies obtained with various levels of salt-and-pepper noise and replacement noise applied are shown in Fig. 7. It can be observed that the recognition accuracy of the Vlachos technique is more susceptible to both salt-and-pepper noise and replacement noise than the proposed method. The results clearly show the robustness of the proposed method under high levels of common document noise and degradation using hand-drawn symbol images.

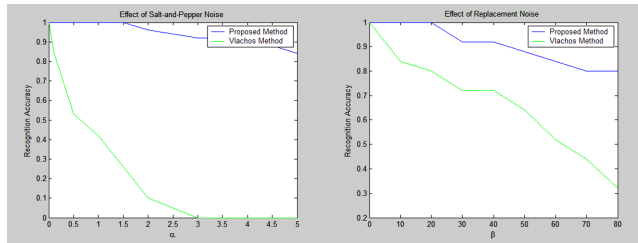

Figure 7. Effect of Salt-and-Pepper Noise and Replacement Noise on Recognition Accuracy

5 Conclusions and Future Work

This paper introduced a symbol recognition method using a modification of the classical Hough Transform and a knowledge-based hierarchical neural network structure. The proposed method is invariant to various geometric transformations such as translation, rotation, scaling, and reflection. It is also robust against common document noise and degradation. It has been shown that the proposed method is able to provide a high level of symbol recognition accuracy under various levels of noise and geometric distortions.

References

[1] Okazaki, A., Kondo, T., Mori, K., Tsunekawa, S. and Kawamoto, E., 1988. An Automatic Circuit Diagram Reader with Loop-Structure-Based Symbol Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 10(3), pp. 331-341.

[2] Lin, S., Hsu, B., and Yang, X., 2002. Content-Based Trademark Retrieval System Using a New Region Based Shape Description Method: The Distance-Angle Pair-Wise Histogram. 22nd International Conference on Distributed Computing Systems Workshops, p. 191.

[3] Leung, W., and Chen, T., 2002. Trademark retrieval using contour-skeleton stroke classification. IEEE International Conference on Multimedia and Expo, 2, pp. 517-520.

[4] Wendling, L., and Tabbone, S., 2004. A new way to detect arrows in line drawings. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(7), pp. 935- 941.

[5] Vlachos, M., Vagena, Z., Yu, P., and Athitsos, V., 2005. Rotation Invariant Indexing of Shapes and Line

Drawings. Proceedings of 14th ACM Conference on Information and Knowledge Management.

[6] Spitz, A., 1999. Shape-based word recognition. International Journal of Document Analysis and Recognition, pp. 178-190.

[7] Shi, D., and Damper, R., and Gunn, S., 2003. Offline handwritten Chinese character recognition by radical decomposition. ACM Transactions on Asian Language Information Processing, 2(1), pp. 27-48.

[8] Bartolini, I., Ciaccia, P., and Patella, M., 2005. WARP: Accurate Retrieval of Shapes Using Phase of Fourier Descriptors and Time Warping Distance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(1), pp. 142-147.

[9] Cai, J., and Liu, Z., 2001. Hidden Markov models with spectral features for 2D shape recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(12), pp. 1454-1458.

[10] Hu, M., 1962. Visual pattern recognition by moment invariants. IEEE Transactions on Information Theory, 8(2), pp. 179-187.

[11] Mokhtarian, F., and Abbasi, S., 2003. Robust Recognition based on Fusion of Multiple Shape Descriptors. Proceedings of IASTED International Conference on Signal and Image Processing, pp. 473-478.

[12] Llados, J., Marti, E., and Villanueva, J., 2001. Symbol Recognition by Error-Tolerant Subgraph Matching between Region Adjacency Graphs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(10), pp. 1137-1143.

[13] Yang, S., 2005. Symbol Recognition via Statistical Integration of Pixel-Level Constraint Histograms: A New Descriptor. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(2), pp. 278-281.

[14] Tabbone, S., Wendling, L., and Tombre, K., 2003. Matching of graphical symbols in line-drawing images using angular signature information. International Journal on Document Analysis and Recognition, 6(2), pp. 115-125.

[15] Mokhtarian, F., Abbasi, S., and Kittler, J., 1996. Efficient and Robust Retrieval by Shape Content through Curvature Scale Space. Proceedings of International Workshop on Image DataBases and MultiMedia Search, pp. 35-42.

[16] Mokhtarian, F., Abbasi, S., and Kittler, J., 1996. Robust and Efficient Shape Indexing through Curvature Scale Space. Proceedings of British Machine Vision Conference, pp. 53-62.

[17] Franti, P., Mednonogov, A., Kyrki, V., and Klviinen, H., 2000. Content-based matching of line-drawing images using the Hough transform. International Journal of Document Analysis and Recognition, 3(2).

[18] Tabbone, S., and Wendling, L., 2002. Technical Symbols Recognition Using the Two-Dimensional Radon Transform. 16th International Conference on Pattern Recognition, 3, pp. 200-203.

[19] Behnke, S., 2003. Hierarchical Neural Networks for Image Interpretation (Lecture Notes in Computer Science). Springer-Verlag New York, Inc.

[20] Simon, S., Schwenker, F., Kestler, H., Kraetzschmar, G., and Palm, G., 2002. Hierarchical object classification for autonomous mobile robots. International Conference on Artificial Neural Networks, pp. 831836.

[21] International Symbol Recognition Contest 2005, http://symbcontestgrec05.loria.fr/index.php, Accessed July 1, 2006.

[22] Kanungo, T., Haralick, R., Baird, H., Stuezle, W., and Madigan, D., 2000. Statistical, nonparametric methodology for document degradation model validation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, pp. 1209-1223.

[23] Barrow, H., Tenenbaum, J., Bolles, R., and Wolf, H., 1977. Parametric correspondence and chamfer matching: Two new techniques for image matching. International Joint Conference on Artificial Intelligence, pp. 659-663.

[24] Huttenlocher, D., Klanderman, D., and Rucklige, A., 1993. Comparing images using the Hausdorff distance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(9), pp. 850–863.