Recognition of Architectural and Electrical Symbols by COSFIRE Filters with Inhibition

Jiapan Guo[1], Chenyu Shi[1(✉)], George Azzopardi[1,2], and Nicolai Petkov[1]

[1] Johann Bernoulli Institute for Mathematics and Computer Science,
University of Groningen, Groningen, The Netherlands
{j.guo,c.shi,g.azzopardi,n.petkov}@rug.nl
[2] Intelligent Computer Systems, University of Malta, Msida, Malta

Abstract. The automatic recognition of symbols can be used to automatically convert scanned drawings into digital representations compatible with computer aided design software. We propose a novel approach to automatically recognize architectural and electrical symbols. The proposed method extends the existing trainable COSFIRE approach by adding an inhibition mechanism that is inspired by shape-selective TEO neurons in visual cortex. A COSFIRE filter with inhibition takes as input excitatory and inhibitory responses from line and edge detectors. The type (excitatory or inhibitory) and the spatial arrangement of low level features are determined in an automatic configuration step that analyzes two types of prototype pattern called positive and negative. Excitatory features are extracted from a positive pattern and inhibitory features are extracted from one or more negative patterns. In our experiments we use four subsets of images with different noise levels from the Graphics Recognition data set (GREC 2011) and demonstrate that the inhibition mechanism that we introduce improves the effectiveness of recognition substantially.

Keywords: COSFIRE · Trainable filters · Architectural and electrical symbols · Shape · Inhibition · Brain-inspired · Visual cortex

1 Introduction

The recognition of symbols in sketches or scanned documents facilitates the automatic conversion to digital representations that can be processed by computer aided design software. Examples of applications are the recognition of architectural and electrical symbols, optical music notes, document analysis, logo and mathematical expressions [13–15,18]. In such applications, it is common to find that a symbol is contained within another symbol that has a different meaning. Fig. 1 shows four patterns presented in the top images that are contained in the corresponding bottom images. The addition of extra strokes can radically change the meaning of a symbol.

Existing symbol recognition algorithms can be categorized into statistical and structural-based approaches. The former methods extract hand-crafted features from symbols and use them to form feature vectors and train classification
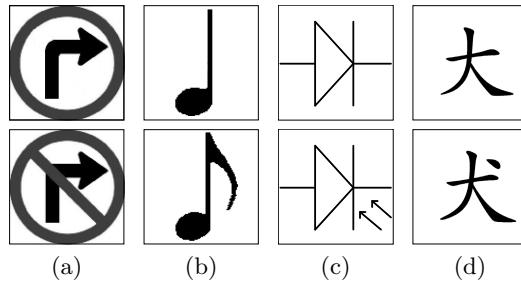
Fig. 1. Examples of pairs of symbols with different meanings. (a) Two traffic signs which give opposite instructions: permission and prohibition of turning right. (b) Two music notes: quarter and eighth. (c) Two electrical symbols: a normal and a light-emitting diode. (d) Two Chinese characters that are translated as "big" and "dog".

models [19]. While such methods may be effective, they require large numbers of training examples. Moreover, the selection of features is specific to the application at hand and typically requires domain knowledge. The structural-based approaches, which usually describe symbols by the geometrical relation between their constituent parts [9], are not suitable to distinguish symbols with similar shapes [17].

In this paper, we use inspiration from the function of shape-selective neurons in area TEO in visual cortex to develop an inhibitory mechanism that we add to the COSFIRE filters introduced in [3]. By means of single cell recordings on macaque monkeys, Connor et al. [4] discovered that such a neuron responds strongly for a certain arrangement of curvatures (Fig. 2a), but its response is suppressed by the presence of a specific curvature element (Fig. 2b). The effectiveness of COSFIRE filters have already been shown in various applications including detection of vascular bifurcations in retinal images [2], classification of handwritten digits [1] and localization and recognition of traffic signs [3]. A COSFIRE filter, as published in [3], can be configured to be selective for one of the symbols in the top row of Fig. 1. It will, however, also respond to the symbol underneath it, and thus it is suitable to distinguish between such patterns.

The response of a COSFIRE filter with inhibition that we propose is computed by subtracting a fraction of the combined responses of inhibitory part detectors from the combined responses of excitatory part detectors. The excitatory and inhibitory parts together with their spatial arrangement are determined in an automatic configuration procedure. A strong response by a COSFIRE filter indicates that the input pattern is similar to the positive prototype used to configure that filter.

The rest of the paper is organized as follows. In Section 2 we explain how a COSFIRE with inhibition is configured and applied. In Section 3 we describe our experiments on four subsets of images from the Graphics Recognition (GREC 2011) data set [16]. We provide a discussion in Section 4 and draw conclusions in Section 5.
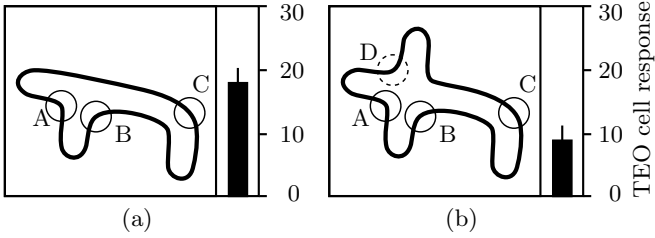
Fig. 2. Selectivity of a shape-selective neuron in posterior inferotemporal cortex [4]. (a) The contour segments marked with solid circles indicate curvatures that evoke excitation of the concerned cell, while (b) the segment marked with a dashed circle indicates a curvature that inhibits the activation of the cell. The bars specify the strength of the response.

2 Method

Let us consider the two symbols shown in Fig. 3a and Fig. 3b, which we refer to as a positive and a negative prototype, respectively. The configuration procedure, that we explain below, automatically determines the excitatory and inhibitory parts, and results in a filter which is able to respond selectively only to patterns similar to Fig. 3a and not to Fig. 3b.
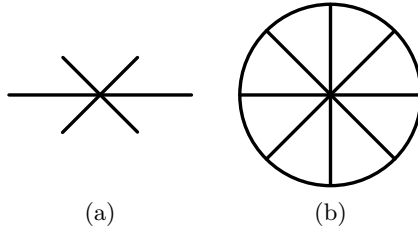
Fig. 3. Example of (a) a symbol that is contained within (b) another one. We use the former pattern as a positive example and the latter one as a negative example.

2.1 Configuration of a COSFIRE Filter with Inhibition

First, we use the publicly available implementation[1] to configure two COSFIRE filters of the type proposed in [3], one to be selective for the positive pattern in Fig. 3a and the other for the negative pattern in Fig. 3b. Such a filter uses a representation of the line segments and their mutual geometrical arrangement. A line segment i is described by the preferred scale λ_i and preferred orientation θ_i of a symmetric Gabor filter together with the polar coordinates (ρ_i, ϕ_i) of the position of the segment with respect to the center of the concerned COSFIRE filter. We denote by $P_f = \{(\lambda_i, \theta_i, \rho_i, \phi_i) \mid i \in 1 \ldots n_1\}$ and

---

[1] Matlab scripts: http://mathworks.com/matlabcentral/fileexchange/37395

$N_f = \{(\lambda_j, \theta_j, \rho_j, \phi_j) \mid j \in 1 \ldots n_2\}$ the configured COSFIRE filters for the positive and negative prototypes, respectively. The parameters n_1 and n_2 denote the number of tuples in the corresponding sets. Fig. 4a and Fig. 4b show the structures of the resulting two COSFIRE filters. For more technical details about the Gabor filters and the COSFIRE configuration method, we refer the reader to [3, 5–8, 10–12] and to an online implementation[2]. We only mention that the configuration of a COSFIRE filter takes as input three parameters, a set of wavelengths λ and a set of orientations θ that characterize a bank of Gabor filters, together with a set of radius values ρ that represent a number of concentric circles around the center of the filter along which the Gabor filter responses are considered. Fig. 4 illustrates the structure of the COSFIRE filter with inhibition that is configured by the positive and negative images in Fig. 3. The configuration uses $\lambda \in \{18, 36\}$, $\theta \in \{\frac{\pi i}{8} \mid i = 0\ldots7\}$ and $\rho \in \{0, 40, \ldots, 200\}$.

Second, we form a new set S_f by taking tuples from the two sets P_f and N_f and marking them with tags that represent the type of contribution (i.e. excitatory or inhibitory). We include all the tuples in P_f into the set S_f and mark them with a tag $\delta = +1$ to indicate that the involved Gabor responses provide excitatory input to the resulting filter. Then we compute the minimum distance $d(N_f^j, P_f)$ between the spatial coordinates of one tuple from N_f and the spatial coordinates of all the tuples from P_f:

$$d(N_f^j, P_f) = \min_{i \in \{1, \ldots, |P_f|\}} \left\{ \sqrt{(\rho_i \cos \phi_i - \rho_j \cos \phi_j)^2 + (\rho_i \sin \phi_i - \rho_j \sin \phi_j)^2} \right\}$$

If this distance $d(N_f^j, P_f)$ is larger than a threshold ζ, we conclude that the tuple N_f^j is sufficiently different from the tuples in P_f and we include the tuple N_f^j into the new set S_f and mark it with a tag $\delta = -1$ indicating that the corresponding Gabor response provides inhibitory input. We repeat this procedure for each tuple in set N_f. In this way we obtain a new set $S_f = \{(\lambda_k, \theta_k, \rho_k, \phi_k, \delta_k) \mid k \in 1 \ldots n_3\}$ of labeled excitatory and inhibitory tuples. The parameter n_3 denotes the number of tuples in set S_f. In Section 3, we provide the value of the parameter ζ that we use in our experiments. With this procedure we ensure that a line segment that is present in both the positive and negative prototypes in roughly the same positions is considered to give excitatory input. On the other hand, a line segment that is only present in the negative prototype is considered to give inhibitory input.

Fig. 4c shows the structure of the resulting COSFIRE filter with inhibition. The white ellipses indicate the line segments that provide excitatory input and the black ellipses indicate the ones that provide inhibitory input.

In order to extract as much detail as possible from a given prototype symbol, in our experiments we use a large set of radii values ($\rho = \{i \mid i = 0, 1, \ldots, 362\}$)[3]. Subsequently, we remove any redundant tuples from the resulting filter by computing the distances between the spatial coordinates of all pairs of tuples. For

---

[2] http://matlabserver.cs.rug.nl
[3] The maximum ρ value 362 is the largest diagonal distance with respect to the center of the image.

(a)   (b)   (c)
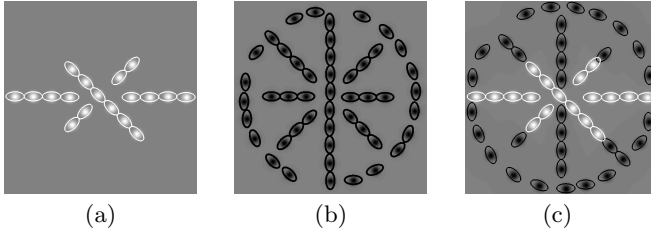
Fig. 4. (a-b) The structures of COSFIRE filters configured with the positive and negative prototype in Fig. 3a and Fig. 3b, respectively. (c) The structure of the resulting filter with inhibition. The ellipses illustrate the wavelengths and orientations of the selected Gabor filters and their positions indicate the locations at which their responses are used as input to the concerned COSFIRE filter. The blobs within the ellipses represent blurring functions that are used to provide some tolerance regarding the preferred positions. White and black ellipses and blobs indicate Gabor responses that provide respectively positive and negative inputs to the COSFIRE filter with inhibition.

any pair that has a distance lower than 12 pixels we delete one of the tuples randomly.

This configuration procedure is also applicable when multiple negative examples are used. In that case we have one COSFIRE filter P_f selective for the positive prototype and a number of COSFIRE filters $(N_{f_1}, N_{f_2}, \dots)$ selective for the negative prototypes. Then we apply the above procedure to determine the inhibitory tuples from each of the N_{f_i} sets. Each set of inhibitory tuples has a unique tag. For instance, the inhibitory tuples determined from the set N_{f_1} are assigned a tag $\delta = -1$, the inhibitory tuples determined from the set N_{f_2} are assigned the tag $\delta = -2$, and so forth.

2.2   Response of a COSFIRE Filter with Inhibition

The response of a COSFIRE filter with inhibition is computed as follows. First we compute the weighted geometric mean as defined in [3] for each group of tuples that share the same tag value. The intermediate representation defined by a tuple $(\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i)$ is computed by blurring the response map obtained by a Gabor filter (with parameter values λ_i and θ_i) with a Gaussian function[4]. The blurred response is then shifted by ρ_i pixels in the direction opposite to ϕ_i. In this way, all the Gabor responses described by different tuples meet at the same location. We denote by $r_{S_f^+}(x, y)$ the output of the group of excitatory tuples with tag $\delta = +1$. Similarly, we denote by $r_{S_f^{-1}}(x, y)$ the output of the group of inhibitory tuples with tag $\delta = -1$.

Finally, we denote by $r_{S_f}(x, y)$ the filter response, which we compute by subtracting a factor of the maximum response of all groups of inhibitory tuples

---

[4] For the blurring function we use a fixed standard deviation of 4, which we found empirically.

from the response of the group of excitatory tuples:

$$r_{S_f}(x,y) \stackrel{\text{def}}{=} |r_{S_f^+}(x,y) - \eta \max_{j=1}^{n}\{r_{S_f^{-j}}(x,y)\}|_{t_3} \tag{1}$$

where $S_f^+ = \{(\lambda_i, \theta_i, \rho_i, \phi_i) \mid \forall (\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i) \in S_f, \delta_i = +1\}$,
$S_f^{-j} = \{(\lambda_i, \theta_i, \rho_i, \phi_i) \mid \forall (\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i) \in S_f, \delta_i = -j\}$, $n = \max|\delta_i|$, η is a coefficient that we call inhibition factor and $|.|_{t_3}$ represents the thresholding operation of the response at a fraction t_3 of its maximum across all image coordinates (x, y).

2.3   Tolerance to Geometric Transformations

The proposed COSFIRE filters with inhibition can achieve tolerance to scale, rotation and reflection by similar manipulation of parameters as proposed for the original COSFIRE filters [3]. We do not elaborate on these aspects here as we do not use them in our experiments. We refer the reader to [3] for a thorough explanation.

3   Experiments

3.1   Data Sets

We use the GREC 2011 data set [16] that contains 150 model architectural and electrical symbols, and three data sets (called NoiseA, NoiseB and NoiseE) of images with different levels of degradation. In each of the three noisy data sets, there are 25 degraded images for every symbol class.

We configure 150 COSFIRE filters to be selective for the 150 models. Subsequently, we apply each resulting filter to the remaining 149 symbols that have different meaning than the symbol used for its configuration. It turns out that 26 COSFIRE filters give strong responses also to non-preferred symbols. In our experiments we only use test images that come from the 26 problematic symbol classes, Fig. 5. In practice, we form three subsets of test images, namely sub-NoiseA, sub-NoiseB and sub-NoiseE of $(25 \times 26 =)$ 650 images each. The model symbol images are of size 512×512 pixels, while the noisy images in the subsets are of size 256×256 pixels. The lines in the ideal models have a thickness of 9 or 18 pixels.

Fig. 6a shows a model symbol and Fig. 6(c-d) show three symbols of the same class from the subsets sub-NoiseA, sub-NoiseB and sub-NoiseE, respectively.

3.2   Pre-processing

We resize by a factor of 2 the degraded images in the sets sub-NoiseA, sub-NoiseB and sub-NoiseE in order to bring them to the same size of the model images. For the images in sub-NoiseA and sub-NoiseB we apply some morphological operations, which we explain below, so that the thickness of their lines becomes
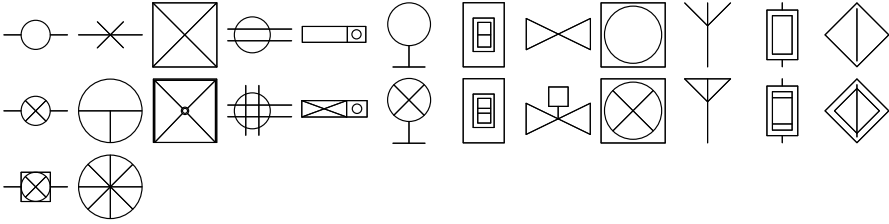
Fig. 5. A set of 26 symbols from the GREC 2011 data set. [16]. The symbol in the top left corner is contained within the symbol below it, which in turn is contained within the symbol in the bottom left corner. The top two symbols in the second column are both contained within the symbol in the third row of the same column. The symbols in the first row of the remaining columns are contained within the corresponding symbols of the second row.
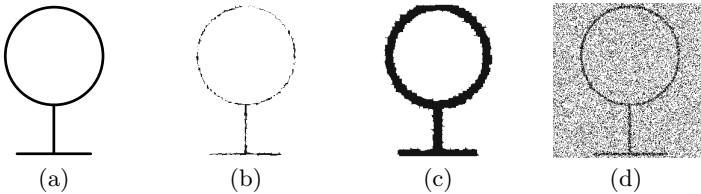
Fig. 6. Examples from the GREC 2011 data set [16]. (a) A model symbol image. (b-d) Degraded symbols of the same class from the data sets of noisy images; (a) sub-NoiseA, (b) sub-NoiseB and (c) sub-NoiseE.

similar to those of the model symbols. Since the images in sub-NoiseE have roughly the preferred line thicknesses, we do not apply any pre-processing.

For the images in sub-NoiseA, we first dilate them by six line-shaped structuring elements of 6 pixel length with different orientations ($\{0, \frac{\pi}{6}, \frac{\pi}{3}, \ldots, \frac{5\pi}{6}\}$). After that, we take the maximum value in every location across these six dilation maps. Then we perform a thinning operation followed by six dilations using line-shaped structuring elements of 4 pixel length with equidistant orientations. The final preprocessed image is obtained by taking the maximum value in every pixel location among the resulting dilation maps.

For the images in sub-NoiseB, we perform opening and thinning followed by a dilation operation using a series of line-shaped structuring elements of 4 pixels length in six orientations. Finally, we superimpose these six dilation maps by taking the maximum value in each location.

Fig. 7a and Fig. 7b show the preprocessed images corresponding to the noisy images in Fig. 6b and Fig. 6c.

3.3   Implementation

We configure 26 COSFIRE filters with inhibition to be exclusively selective for the 26 model symbols. For the configuration we apply the following approach. We configure a COSFIRE filter without inhibition for a given model symbol,
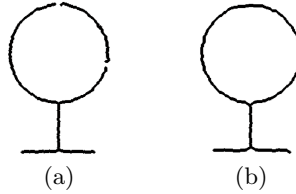
Fig. 7. Examples of the preprocessed images. (a) The improved image of Fig. 6b in sub-NoiseA and (b) the preprocessed image of Fig. 6c in sub-NoiseB.

apply the resulting filter to the remaining 25 model symbols and threshold the maximum values of each response map at a given threshold ε. The value of the threshold parameter ε is a fraction of the maximum response that the filter achieves when it is applied to its preferred model symbol. The symbol images that evoke responses greater than $(\varepsilon =)\ 0.3$ are considered as negative prototype patterns. Then we use these automatically selected negative patterns to determine the inhibitory line segments and configure a COSFIRE filter with inhibition ($\zeta = 12$ pixels) that is exclusively selective for patterns similar to the positive prototype symbol (of size 512×512 pixels). We perform this procedure for each of the 26 symbols.

Next, we apply the 26 filters with inhibition to each model symbol by using the method in Section 2.3. We investigate the inhibition factor by systematically varying the value of parameter η between 0 and 3 in intervals of 0.2. For $\eta = 1.4$ the filters give responses only to the preferred positive prototype patterns.

Then we apply these 26 inhibition-based COSFIRE filters with $\eta = 1.4$ to the preprocessed images in sets sub-NoiseA and sub-NoiseB and to the non preprocessed images of sub-NoiseE.

3.4 Evaluation and Results

In the first experiment we only use the 26 model images. Fig. 8 shows a comparison between the results obtained by the COSFIRE filters in their basic form and COSFIRE filters with the proposed inhibition mechanism. The results are shown in the form of a confusion matrix where the value at location (i, j) is the maximum response of the filter S_{f_i} (that was configured by model i) to a model image j. Fig. 8a shows the results of the COSFIRE filters without inhibition, the matrix of which is less sparse than that in Fig. 8b that is achieved by COSFIRE filters with inhibition. The off-diagonal non-zero elements in the left panel indicate that the corresponding COSFIRE filters without inhibition respond to more than one symbol. The absence of such elements in the right pannel means that each of the COSFIRE filters with inhibition responds only to one symbol, the positive pattern with which it was configured.

A given image is classified to the class of the positive prototype symbol by which the inhibition-based COSFIRE filter that achieves the maximum response
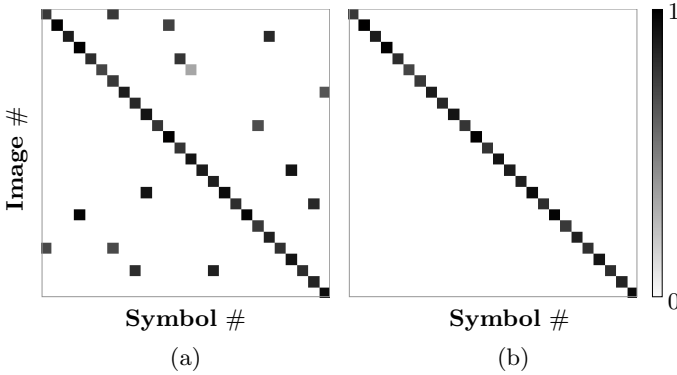
Fig. 8. Results of COSFIRE filters (a) without and (b) with inhibition. The matrices are of size 26×26; the columns represent model images and the rows represent configured COSFIRE filters. The elements of the matrices are the maximum responses of the filters to the 26 images.

was configured. The proposed inhibition-based approach achieves 100% accuracy while the one without inhibition achieves 97.78% accuracy.

In Table. 1 we report the results that we achieve for the three noisy subsets of test images.

Table 1. Accuracy (%) on three sub sets of $(26 \times 25 =)$ 650 noisy images each taken from the GREC data set [16]

	sub-NoiseA	sub-NoiseB	sub-NoiseE	Average
Inhibition-based COSFIRE	99.85	99.35	99.92	99.71
Original COSFIRE	97.83	97.46	91.51	95.60

4   Discussion

We propose an inhibition mechanism to the COSFIRE filters in order to increase their discimination ability. The inhibition mechanism involves the determination of line segments from training examples whose presence is used to suppress the response of the filter. This was inspired by the functionality of shape-selective TEO neurons in visual cortex. The firing rate of such a neuron can be suppressed by certain contour parts in specific positions within its receptive field [4].

The way we perform classification follows what is known as the grandmother cell hypothesis in neuroscience, in that the label of an input image is determined from one filter, the one that gives the strongest response. As already shown in [3], COSFIRE filters can be used in what is known as population coding in neuroscience, whereby a feature vector is formed by their maximum responses to

an input image. In this technique COSFIRE filters are configured with small parts of patterns of interest and it is suitable for applications where the involved patterns are more deformable, such as handwritten digits [1]. In future work we aim to investigate the effectiveness of feature vectors formed by the responses of the proposed COSFIRE filters with inhibition in classification tasks. Our speculation is that the resulting sparser vectors will improve the discriminating power. In principle, sparseness is a desirable feature as it increases storage capacity and allows the discrimination of more patterns.

Besides the recognition of architectural and electrical symbols the proposed methodology can be considered as a general framework for the classification of any sketched symbols. It first configures a number of COSFIRE filters, then it learns inhibitory input for each filter. At application stage it applies them simultaneously to a test image and classifies the image with the label of the filter that achieves the maximum response.

5   Conclusions

The proposed COSFIRE filters with inhibition mechanism are highly effective. In our experiments on architectural and electrical symbols we demonstrated that COSFIRE filters with inhibition improve the classification performance significantly on symbol classes that are contained within other classes. We achieve a recognition rate of 100% for the subset of 26 models and an average rate of 99.71% for the three subsets of noisy images.

References

1. Azzopardi, G., Petkov, N.: A shape descriptor based on trainable COSFIRE filters for the recognition of handwritten digits. In: Wilson, R., Hancock, E., Bors, A., Smith, W. (eds.) CAIP 2013, Part II. LNCS, vol. 8048, pp. 9–16. Springer, Heidelberg (2013)
2. Azzopardi, G., Petkov, N.: Automatic detection of vascular bifurcations in segmented retinal images using trainable COSFIRE filters. Pattern Recognition Letters **34**, 922–933 (2013)
3. Azzopardi, G., Petkov, N.: Trainable COSFIRE Filters for Keypoint Detection and Pattern Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **35**(2), 490–503 (2013)
4. Brincat, S., Connor, C.: Underlying principles of visual shape selectivity in posterior inferotemporal cortex. Nature Neuroscience **7**(8), 880–886 (2004)
5. Grigorescu, C., Petkov, N., Westenberg, M.: Contour detection based on non-classical receptive field inhibition. IEEE Transactions on Image Processing **12**(7), 729–739 (2003)
6. Grigorescu, C., Petkov, N., Westenberg, M.: The Role of Non-CRF Inhibition in Contour Detection. Journal of Computer Graphics, Visualization, and Computer Vision (WSCG) **11**(2), 197–204 (2003)
7. Grigorescu, S., Petkov, N., Kruizinga, P.: Comparison of texture features based on Gabor filters. IEEE Transactions on Image Processing **11**(10), 1160–1167 (2002)

8. Kruizinga, P., Petkov, N.: Non-linear operator for oriented texture. IEEE Transactions on Image Processing **8**(10), 1395–1407 (1999)
9. Lin, Y., Wenyin, L., Jiang, C.: A structural approach to recognizing incomplete graphic objects. In: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 1, pp. 371–375, August 2004
10. Petkov, N.: Biologically motivated computationally intensive approaches to image pattern-recognition. Future Generation Computer Systems **11**(4–5), 451–465 (1995). 1994 Europe Conference on High Performance Computing and Networking (HPCN Europe 94), Munich, Germany, 1994
11. Petkov, N., Kruizinga, P.: Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: Bar and grating cells. Biological Cybernetics **76**(2), 83–96 (1997)
12. Petkov, N., Westenberg, M.: Suppression of contour perception by band-limited noise and its relation to non-classical receptive field inhibition. Biological Cybernetics **88**(10), 236–246 (2003)
13. Rebelo, A., Fujinaga, I., Paszkiewicz, F., Maral, A.R.S., Guedes, C., Cardoso, J.S.: Optical music recognition: state-of-the-art and open issues. IJMIR **1**(3), 173–190 (2012)
14. Tang, P., Hui, S.C., Fu, C.W.: Online chemical symbol recognition for handwritten chemical expression recognition. In: ICIS, pp. 535–540. IEEE (2013)
15. Valveny, E., Dosch, P., Winstanley, A., Zhou, Y., Yang, S., Yan, L., Wenyin, L., Elliman, D., Delalandre, M., Trupin, E., Adam, S., Ogier, J.M.: A General Framework for the Evaluation of Symbol Recognition Methods. Int. J. Doc. Anal. Recognit. **9**(1), 59–74 (2007)
16. Valveny, E., Delalandre, M., Raveaux, R., Lamiroy, B.: Report on the symbol recognition and spotting contest. In: Kwon, Y.-B., Ogier, J.-M. (eds.) GREC 2011. LNCS, vol. 7423, pp. 198–207. Springer, Heidelberg (2013)
17. Yajie, Y., Zhang, W., Wenyin, L.: A new syntactic approach to graphic symbol recognition. In: Ninth International Conference on Document Analysis and Recognition, ICDAR 2007, vol. 1, pp. 516–520, September 2007
18. Zanibbi, R., Blostein, D., Cordy, J.R.: Recognizing Mathematical Expressions Using Tree Transformation. IEEE Transactions on Pattern Analysis and Machine Intelligence **24**(11), 1455–1467 (2002)
19. Zhang, W., Wenyin, L., Zhang, K.: Symbol Recognition with Kernel Density Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence **28**(12), 2020–2024 (2006)