computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

DETERMINING THE ELEMENTS OF THE TEXT ON IMAGE USING OF FUZZY COGNITIVE MODELS

Elena Pyatikop

Priazov State Technical University
7, Universitetska Street, Mariupol 87500 Ukraine, E-mail: Pjatikopelena@rambler.ru

Abstract: *This publication describes the modeling results of knowledge of cognitive psychology about the primary processing in the visual system to analyze the images in purpose to determine the elements of the text. Processing images at the level of primary visual cortex formalized using fuzzy sets.*

Keywords: *visual system, cells of the primary visual cortex, orientation columns, images, lines of letters.*

INTRODUCTION

In the field of artificial intelligence the challenge task is recognize text. Today, there is experience in developing methods of recognizing printed characters with different kinds of classifiers, neural networks [1-2]. But the volume of digital image information is increased the conditions of recognition are change and modern systems require a more in-depth intellectual analysis. Therefore for image processing the possibility of applying knowledge about the structure of the visual system and the processes occurring in it is explored [3-5]. The cognitive psychology describes this knowledge [6-7]. This article discusses the use of cognitive models, in particular the model ganglion cells and cell models of primary visual cortex to solve the identity of the elements of the text in the image.

1. THE PROBLEM

Suppose there is an image with text in 8 bit grayscale. The location of lines in the image may be at any angle. In addition, the letter size of the text is unknown. In this regard, an initial analysis of the image is intended to define the angle of slope and height of text lines. After that the line must be divided into possible words, and then make the selection of letters.

To investigate the images you need to use fuzzy model of the image cells of primary visual cortex.

If you glance at the image not fixing it on details, it will produce the perception of larger patterns. In this problem it will be a text lines. That is, if the image falls in the area of peripheral vision. After projecting the image on retina it is occurred the following processes:

1) for the first the image is processed by rods and cones. These rods and cones activate some receptive fields of ganglion cells through intermediate bipolar cells [6-7]. A lot of active ganglion cells form the first “projection” images.

2) next, the majority of ganglion cell axons form synaptic couplings with cells lateral geniculate nucleus (LGN). In the result it is an exact repetition of the retina in LGN, i.e. its retina “map”. The cells of this level by the structure of buildings are like the ganglion cells, but also differs by sensitivity and sizes.

3) In turn, LGN axons transmit signals to the primary (striate) visual cortex. The cells of visual cortex respond to specific orientation and the length of the stimulus. Cells that are responded to the some location of the stimulus are the orientation column. Thus it is formulated about 18-20 columns, with a gradual change in orientation from 0^0 to 180^0.

In that way, the most active column will reflect the lines of text. A lot of cells of this column are related to the whole line or the separate words of this one. The angle of this active column corresponds with the angle of slope of the text lines. And the size of the cell nucleus corresponds with the height of the line.

Formal description of the model ganglion cells and the possibility of its usage are discussed in detail in [9]. But in this publication it is considered the usage of models of primary visual cortex cells.

2. CELL MODEL OF PRIMARY VISUAL CORTEX

As for the activation of cells at this level must be closely related group of ganglion cells [8], for that we introduce the concept of the detector. The combination of adjacent cells is shown in Figure 1. Then the model of the detector, is the expression (1), where $K^{\alpha,r} =< K^{\alpha,r+2Rk}, K^{\alpha,r+4Rk} \ldots K^{\alpha,r+\ell^*\cdot 2Rk} >$; α – angle of orientation 0^0 to 175^0; ω – width of the detector, which is defined ω= *Rk*; ℓ^* – length detector, which is determined by the number of adjacent cells $K^{a,r}$, but not less 4; $\delta(D^\alpha)$ – the confidence function of the detector presence is calculated by the formula (2), where $\mu(K_i)$ is calculated by the formula [9].

Set of detectors will be $D^\alpha = \{D_j^\alpha\},\ j \in 1\ldots n_\alpha$ Based on the concept of the detector we will describe a model of different types of cells of the visual cortex.

Thus, a simple cell S^α, which responds only to a certain orientation, described by expression (3), where $\ell \geq \ell^*$, but $\delta(S^\alpha) = \min_{j\in 1\ldots n_\alpha}\{\delta(D_j^\alpha)\}$.

The most complicated visual cortex cells, as well as simple, are sensitive to the direction of the signal. But their basic characteristic is a reaction to a certain length ℓ. Then the most complicated visual cortex cell $G^{\alpha,\ell}$ can be described by the expression (4), where $\delta(G^{\alpha,\ell}) = \delta(D^a)$.

In one direction α the most complicated visual cortex cells by the same length can be quantity m_α. Then the set of all the most complicated cells will be described $G^{\alpha,\ell} = \{G_j^{\alpha,\ell}\},\ j \in 1\ldots m_a$. But these may be another length $\ell_1, \ell_2, \ldots \ell_z$. Then the set the most complicated cells reacting to specific lengths, expressed as

$$G^{\alpha,\ell_i} = \{G_j^{\alpha,\ell_i}\},\ i \in 1\ldots z,\ j \in 1\ldots m_a$$

As a result, you can build a model of the image as a set of orientation columns, where each column is described by expression (5).

Then the model of "projection" image, i.e. retina map is described by the expression (6), where α – angle of orientation 0^0 to 175^0; ω – width of the detector, which is defined ω= *Rk*; $\lambda(M^\omega)$ – the confidence function of the existence of the greatest number of the most complicated cells of the maximum length is calculated by the formula (7), where λ_1 – is the value of all active relay cells have enhanced to the most complicated cells is defined by (8); λ_2 – average value of the confidence of all the most complicated cell with width ω, is defined (9); α^* – angle of orientation 0 to 175, selected on the set of orientation columns C^α, determined by the expression $l_{\alpha^*} = \max_\omega\{\max_\alpha\{l_\alpha\}\}$.

You need to calculate the confidence function for each retina map retina map M^{ω_i} to select the retina map which has the maximum confidence function.

Then the cells of this map will correspond the whole lines or the separate words of the lines with angle α and height ω.

3. RESULTS

Consider an example application of the previous models at an image size 860 × 227, which is given in Fig. 2. Fig. 3 shows retina map obtained from the automatic determination of optimal angle of inclination. Detailed process of selecting the optimal parameters is given in [10]. Fig. 4 shows the localization of the basic elements of the text lines. Using the classical transformation of computer graphics and a calculated angle, you can rotate the image. And the result is shown in Fig. 5.

When the basic larger elements of images are located, you can make more detailed analysis. The following movement of the eye relates to each individual element and has a projection on the retina.

You can try to split each of the elements in the letters, but you can pre-examine whether this element has particular word characteristics. Under the features of the words it is implied the existence of a sequence of patterns of vertical and horizontal series [11]. To do this you need to re-submit images using models of cells of primary visual cortex only with a reduced receptive field. On Fig. 6 there is an example of the availability characteristics to the letters, as in Fig. 7 - their absence. Fig. 8 illustrates the next step - a breakdown in the letter, which is performed also on the basis of the images using the cells of primary visual cortex.

To each of the blocks need to perform the procedure of expanding the field for superscript or subscript elements of the letters. As well as you need make the procedure for cutting off the extra background. After that, the frame will closely match the width and height of character (s). The final result of localization of letters is shown on Fig.9.

4. CONCLUSION

The results of investigations have confirmed the possibility of applying knowledge about the structure of the visual system for processing image in aim to determine the elements of the text.

The usage of developed models of cells of primary visual cortex permits to show the image of located symbol as the set elements. Each element is characterized by its own length and orientation. They are the input data for the further identification symbol.