БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ . BULGARIAN ACADEMY OF SCIENCES

ПРОБЛЕМИ НА ТЕХНИЧЕСКАТА КИБЕРНЕТИКА И РОБОТИКАТА, 46
PROBLEMS OF ENGINEERING CYBERNETICS AND ROBOTICS, 46

София . 1997 . Sofia

Visual Image Preprocessing Using Neural Networks*

Valery Ilchev, Zlatolilia Ilcheva, Stefan Koynov

Institute of Information Technologies, 1113 Sofia

1. Introduction

The contemporary technological level admits that the visual images read in the computer by a TV camera or an optical scanner store in the memory even the slightest details of the optical image.

In fact the image which is encoded in the memory of the computer can contain even more information than the same optical image which is projected on the retina of the human eye. Nevertheless the optical system of the mammals solves extremely complex classification problems which are inaccessible by the contemporary computer technique. A lot of the visual information preprocessing which takes place in the retina is comparatively well studied.

This makes possible the application of some models for these processes in the real technical systems for image processing.

It is well known [1, 2, 3] that the image preprocessing in the neural environment of the eye's retina is performed by the so called receptive fields (RF). The generalization of the data from the neurophysiological experiments on such RF has lead to the conclusion that two of their basic functions are the filtration of the image which is projected on the retina and the detection of zones in which the brightness changes. Those two properties of the RF can be described with a great degree of conformity by $\nabla^2 G$-filters [3] and they can be included in a model constructed by an artificial NN which resembles to some extent the organization of the ganglious cells in the retina [1].

The present paper deals with the following questions:

1. Preprocessing of visual images by a ∇^2G-filter.

2. Construction of an artificial neural network which solves the problems postulated in point 1.

3. Recognition feature construction rules based on the NN-processed visual information (in point 2).

*This paper is partially supported by the Bulgarian National Science Foundation, Grant I-524/95.

The described method can be applied for non-standard font preprocessing (cyrillic and latin) which is different in type and size including a handwritten text with block and capital letters.

2. Theoretical background of the method

As it was mentioned earlier the neurophysiological and psychophysical investigations of the initial part of the visual channel and especially the preprocessing of the information from the RF of the retina may be described by the operator $\nabla^2 G$ [3]. This allows to consider in details its features with the following designation list:

∇^2 – Laplace operator which in the two-dimensional case is of the type $\partial^2/\partial x^2 + \partial^2/\partial y^2$ and the symbol G designates the Gauss's two-dimensional distribution:

$$G(x,y) = \exp\left(-\frac{x^2+y^2}{2\pi\sigma}\right). \tag{1}$$

The participation of the Gauss's function G in the operator for image preprocessing (OIP) ∇^2G makes the image fuzzy in its space domain. This leads to an effective deletion of all of the space structures in the image which belong to some level of its brightness less than the level determined by the space constant of the Gauss's function. It is well known that in the frequency space this is a filter for the low frequencies. Evidently it is not at all by accident that the information preprocessing by the retina RF uses filters that are very close to the G-filters. The preference of the nature for this type of filters may be explained with the smoothness and with the "good" localization of the function G not only in the two-dimensional space domain but also in the frequency domain. Those features of the function G which are valid in the both of its domains allow a space filtration with the minimum possible noise addition by the very filtering process.

The usage of the Laplacianne (∇^2) in the OIP permits this operator to discover the changes in the image brightness of those scaled levels which are an output from the filtration with the function G. One of the most important properties of ∇^2 is its isotropy which in the neurophysiological structure of the retina is a consequence of the very form of the RF.

As a whole the OIP in the frequency domain is a kind of a band filter which performs a selection of the image brightness. This is the way to eliminate the constant components of the image the background of which being most often ly one of them.

In the image space domain the filtration with the function G may be defined with the convolution $G*B(x,y)$, where $B(x,y)$ is the brightness function. Then the image processing with the OIP may be written in the following way:

$$F(x,y) = \nabla^2 [\, G(x,y) * B(x,y) \,] . \tag{2}$$

This form of the operator is unsuitable for practical realizations, so that it must be rationalized. We consider the expression with its detailed transcription:

$$F(x,y) = \nabla^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} B(x',y')\, G(x-x',y-y')\, dx', dy' = \tag{3}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} B(x',y')\, \nabla^2 G(x-x',y-y')\, dx', dy' .$$

Evidently (3) can be written in the following compact form:

(4) $$F(x,y)=B(x,y)*[\nabla^2 G(x,y)].$$

In this equation the expression in square brackets is the OIP thus allowing the operator to be applied directly for the image processing, not the sequential application of the G-filtering and the Laplace operator. The form of the OIP in (4) allows this operator to be treated as a filter and in this way the image processing problem can be treated simultaneously in the visual image space and in their spectral space.

One must bear in mind that the method describes just a share from the visual information processing functions of the retina RF.

3. Description of the NN for visual image preprocessing

The visual image which is loaded from a TV camera or from an optical scanner may be described by the two-dimensional image brightness function $I(x,y)$. This function is an input of a NN of the type shown in Fig. 1. The information which is included in $I(x,y)$ is parallelly sent to the elements that imitate the retina RF. Those RF form a mosaic structure which may be pledged in the input matrices A and B of the RF (Fig. 2). For example in Fig. 1 we have a mosaic of 5 RFs where RF3 can be treated as the central RF. Then the matrices A and B will be equal to: $A_{1i}=[a^i_{ps}]$, $B_{1i}=[b^i_{ps}]$ where $i=1,2,...,5$; a^i_{ps} and b^i_{ps} – the corresponding matrix elements. The information on the output of each RF could be compared to some extent with the information passing through a separate fibre of the optical nerve transmitting the retina preprocessed information towards the optical zones in the brain. In the NN this information scaled with the weight coefficients c_{1i} and c_{2i} ($i=1,...,5$) is an input for the neural elements NE11 and NE21 in Fig. 1 with the logistic functions of the same type as in Fig. 3, A with the ajustable thresholds t_{11} and t_{21}. As it is evident from Fig. 1 NE11 reacts to the bright fragments in the visual image and NE21 – to the dark fragments respectively. In this way NE11 and NE21 resemble one of the basic functions of the complex ON and OFF neurons in the optical zones in the brain [1, 2].

The last neuron NE02 unites the processed information from NE11 and NE21 which is multiplied by the coefficients c_{11} and c_{12}. In this way it completes the filtering which is accomplished with the operator $\nabla^2 G$:

(4) $$\bar{B}(x,y)=F(x,y),$$

where $F(x,y)$ is defined by the equation (4). The logistic function of NE02 is given in Fig. 3, B.

Any RF in the NN may be schematically represented with Fig. 2 from which it is clear that the output of the field consists of two mutually antagonistic channels ON and OFF. The one channel (ON) responds to bright fragments of the image and the other channel (OFF) – to the dark fragments. This complex reaction of the RF makes it comparatively universal in the information processing because such field has an "opinion" about the basic levels of the brightness function B(x,y) which is applied to its input. In Fig. 2 the ON-ouput is designated with the symbol o , and the OFF-output – with the symbol ●.

4. Forming of features for recognition

After the visual image processing with the operator $\nabla^2 G$ the symbols for recognition may be submitted to an additional (secondary) processing so as to form the recognition features [4].

Such a system of features which includes the structural properties of the symbols will be:

1) the number of the horizontal crossings,
2) the number of the vertical crossings,
3) the number of the left-side diagonal crossings,
4) the number of the right-side diagonal crossings,
5) the number of the upper-side concave arcs,
6) the number of the lower-side concave arcs,
7) the number of the left-side concave arcs,
8) the number of the right-side concave arcs.

The accepted recognition features are illustrated in Fig. 5 and in Fig. 6. Fig. 4 A, B depicts the letter "Ж": before the preprocessing with the operator $\nabla^2 G$ in Fig. 4 A and after it in Fig. 4 B. Fig. 5 explains the first four features of the letter "Ж" and Fig. 6 illustrates the features from the 5th up to the 8th for the same letter.

5. Conclusion

The described method in this paper allows the isotropic visual image processing by the $\nabla^2 G$-operator. The simultaneous space filtration and the border separation of the image have great advantages in comparison just with the gradient methods for the border separation.

The suggested NN reflects the ideas of this method. It can be realized not only by software but also by hardware.

References

1. Somjen, G. Sensory Coding in the Mammalian Nervous System. Moscow, Mir, 1975 (in Russian).
2. Szentagothai, J., M. Arbib. Conceptual Models of Neural Organization. Moscow, Mir, 1976 (in Russian).
3. Marr, D., Vision. A Computational Investigation into the Human Representation and Processing of Visual Information. Moscow, Radio i Svjaz, 1987 (in Russian).
4. Ilcheva, Zl., V. Alexandrov, V. Ilchev. A two stage method for recognition of alphanumerical symbols. – Comptes Rendus de l'Academie Bulgare des Sciences, **49**, 1966, No 10.
5. Pavlidis, T. Algorithms for Graphics and Image Processing. Moscow, Radio i Svjaz, 1986 (in Russian).

Предварительная переработка визуальных изображений, используя нейронные сети

Валери Илчев, Златолилия Илчева, Стефан Койнов

Институт информационных технологий, 1113 София

(Резюме)

В работе описан метод переработки визуальных изображений при помощи дву-димснсионного изотропического оператора Лапласа $\nabla^2 G$ фильтра Гауса G.

Предложенный метод фильтрации изображения используется при построении нейронных сетей, которые реализируют переработку.