UNSUPERVISED TEXTURE IMAGE SEGMENTATION BY IMPROVED NEURAL NETWORK ART2

Zhiling Wang, G. Sylos Labini, R. Mugnuolo and Marco De Sario

Center for Space Geodesy, Italian Space Agency, P.O. Box 11
75100 Matera, Italy
Fax:+39-835-339005 Tel:+39-835-3779 Email:zhiling@asimt0.ml.asi.it
Dept. of Electronic Engineering, University of Bari, Via Re David, 200
70125 Bari, Italy

Abstract

We here propose a segmentation algorithm of texture image for computer vision system on space robot. An improved Adaptive Resonance Theory (ART2) for sensing input patterns is adapted to classify the image based on a set of texture image features extracted by a fast Spatial Gray Level Dependence Method (SGLDM). The nonlinear thresholding functions in input layer of the neural network have been constructed by two parts: firstly to reduce the effect of image noises on the features, a set of sigmoid functions is chosen depending on the types of the feature; secondly, to enhance the contrast of the features, we adopt fuzzy mapping functions. The cluster number in output layer can be increased by an autogrowing mechanism constantly when a new pattern happens. Experimental results and original or segmented pictures are shown, including the comparison between this approach and K-means algorithm. The system written by C language is performed on a SUN-4/330 spare-station with an image board IT-150 and a CCD camera.

1. Introduction

Segmentation and classification of textured images have been considerable attention to contain significant discriminatory information for image segmentation in a variety of application, such as terrain classification, military surveillance and recognition, remote sensing images and biomedical image analysis. Although texture is a fundamental characteristic of images, the complexity involved in its quantification has presented its effective incorporation into the segmentation process.

In this paper, the neural network of an improved Adaptive Resonance Theory (ART2) is presented to segment an image consisting of several regions with different textures. Artificial neural networks offer several advantages over conventional classification techniques, due to their high computation rate, great degree of fault tolerance and unsupervised ability. The number of researches have engaged on the research by neural network.

In this paper, section 2 defines the texture feature types which are derived from co-occurrence matrices and selection of maximum and minimum measure window for feature extraction of the texture image. Section 3 describes an approach of improved ART2 neural network with alterable competitive layer (F2 layer). The nonlinear thresholding function in F1 layer is displaced by a fuzzy mapping function. Section 4 shows the results of experiments and illustration.

2. Feature extraction of texture image

Whether the segmentation of texture image is good or not depends on the extraction of texture features. There are number of the approaches to have been developed for feature extraction of the texture image: Fourier power spectrum method (FPSM), spatial texture-energy, Markov random field model, Gibbs random field model, zero-sum filter masks, gray level run length method (GLRLM), spatial gray level dependence method (SGLDM), gray level difference method (GLDM) and other methods. Some of these methods belong to statistical method, others to structural one. Among them, spatial gray level dependence method, which is introduced by Haralick et al. in their paper, is one of the most successful statistical representation for the texture. The feature measurement from co-occurrence matrices in the SGLDM is rather similar to the knowledge captured by the human eyes, and provides a convenient way to represent the properties of object textures. Weeks et al. experimentally compared feature on terrain images and found that SGLDM is more powerful than the GLDM, GLRM, and FPSM. Ohanian et al. also pointed that the features by SGLDM were better than Markov random field, multi-channel filtering features, and fractal based features. It is known, however, that the SGLDM requires much processing time and great number of memory. Only for mean probability distribution, 2^2^4^ times of multiplication in the SGLDM are done when a measured image is a size 64 x 64 with gray level 128, and the tendency will be raised at exponent rate with the enlargement of the image size, particularly, the increase of gray-level number.

In this paper, we use a set of simplified equations based on a fact that rows or columns around the current pixel are included or excluded almost simultaneously while the measured window is displaced in the horizontal or vertical direction of the image, so we could make the equation be simplified viewing from the pixels of rows(columns) both excluded and included from a window rather than a pixel method. Some calculations are done one time in a row or column instead of one in a pixel, so algorithm in the paper consumes much less time than Haralick's method.

We defined a co-occurrence matrix of relative frequencies with which two pixels separated by distance d at a specified angle occur on the image, one with gray level i and the other with gray level j.
A distance of one pixel, i.e., the measuring window slides over the image in one step length in both horizontal and vertical direction and angle quantized to 45° intervals, or 0°, 45°, 90°, and 135° will be used. We give a set of simplified equations:

(1) Mean

\[ m_k = m_{k-1} + \frac{1}{N} \sum (i_{kr}^k + j_{kr}^k - i_{kr}^* - j_{kr}^*) \]

(2) Variance

\[ \sigma_k^2 = \sigma_{k-1}^2 + \frac{1}{N} \sum (i_{kr}^k + j_{kr}^k - i_{kr}^* - j_{kr}^*)^2 \]

(3) Correlation

\[ C_k = \frac{(C_{k-1} - \sigma_{k-1}^2) + \frac{1}{2} \sum (i_{kr}^k + j_{kr}^k - i_{kr}^* - j_{kr}^*)}{\sigma_k^2} \]

(4) Energy

\[ L_\pm = \frac{N_{k-1}}{N_{k-1} \pm 2} \quad J_\pm = 1/(N_{k-1} \pm 2) \]

\[ E_k^\pm = (L_\pm)E_{k-1}^\pm + (J_\pm)\left( a - 4M_{k-1}^\pm \right) \quad (a=2, b=4, c=4) \]

where

\[ a = \begin{cases} 2 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad (6) \]

(5) Entropy

\[ E_k^p = L^+E_{k-1}^p - L^- \log(L^-) - A^- \quad (7) \]

\[ E_k^p = L^+E_{k-1}^p - L^- \log(L^-) - A^+ \quad (8) \]

\[ A^- = \begin{cases} 2J^-(M_k - (i_{kr}^*, j_{kr}^*)) - (L^-(M_k - (i_{kr}^*, j_{kr}^*)) - 1)^+ & i \neq j \\ J^-(M_k - (i_{kr}^*, j_{kr}^*)) - 2I^-(M_k - (i_{kr}^*, j_{kr}^*)) - 2J^- & i = j \end{cases} \]

\[ A^+ = \begin{cases} 2J^+(M_k - (i_{kr}^*, j_{kr}^*)) + (L^+(M_k - (i_{kr}^*, j_{kr}^*)) + 1)^+ & i \neq j \\ J^+(M_k - (i_{kr}^*, j_{kr}^*)) + 2I^+(M_k - (i_{kr}^*, j_{kr}^*)) + 2J^+ & i = j \end{cases} \]

(6) Contrast

\[ T_k = T_{k-1} + \frac{2}{N} \sum (i_{kr}^* - j_{kr}^*)^2 - (i_{kr}^* - j_{kr}^*)^2 \quad (9) \]

(7) Homogeneity

\[ H_k = H_{k-1} + \frac{2}{N} \sum (1 + i_{kr}^* - j_{kr}^*)^2 - (1 + i_{kr}^* - j_{kr}^*)^2 \quad (10) \]

where the \( \Sigma \) is the \( \sum_{k=1}^2 \). The \( L \) stands for the length of the row or column, i.e., the width of measuring square window, the \( M(i,j) \) is the element of a co-occurrence matrix, superscripts "+" and "-" express for a pixel \((x,y)\) included or excluded from the window. The equations for both energy and entropy features are used to the case considering a pixel included or excluded from the window because of the nonlinear decomposition for square and logarithm functions.

### 3. Improved ART2

Connectionist classification used here is called Adaptive Resonance Theory (ART) \[24-27\]. In general, ART is divided into two types depending on input patterns. ART1 is applied to solve binary input problem, ART2 is available to both binary and analog inputs. In the paper, the ART2 is used to classify the texture image because the 20 features (five for each angle) belong to gray-scale patterns. The classifier in the ART2 consists mainly of two subsystems: the attentional subsystem and the orienting subsystem. The former is composed of the Short-Term Memory (STM) and Long-Term Memory (LTM) elements.

#### 3.1 Short-Term Memory (STM)

The \( F_1 \), the input representation field, and \( F_2 \), the category representation field (competitive mechanism), are the two STM main components.

\( F_2 \) is composed of three layers with STM activation equations as (see Fig. 1.)

\[ p_i = u_i + \sum g(y_j)z_{ji} \quad (11) \]

\[ v_i = \frac{p_i}{\|p\|} \quad (12) \]

\[ u_i = f(v_i) + k_f(v_i) \quad (13) \]

\[ u_i = \frac{v_i}{\|p\|} + \epsilon \quad (14) \]

\[ u_i = I_i + \epsilon \quad (15) \]

\[ r_i = \frac{v_i}{\|p\|} \quad (16) \]

where \( a, b, c \) and \( \epsilon \) are constants, \( y_i \) is the STM activation of the \( j \)th \( F_2 \) neuron, \( \|p\| \) is the \( L_2 \) norm, \( f(x) \) is a nonlinear threshold function:

\[ f(x) = \begin{cases} 0 & \text{for } 0 \leq x < \theta \quad (17) \\ 2(\frac{x-\theta}{\alpha-\theta})^2 & \text{for } \theta \leq x \leq \alpha \\ \frac{1}{2} & \text{for } \alpha \leq x \leq \beta \\ 1 & \text{for } x \geq \beta \end{cases} \]

where the feature noises are suppressed by setting \( f(x) \) to zero when \( 0 \leq x < \theta \). The fuzzy mapping function is used to enhance the contrast among the features, and makes the input patterns classified.
more easily. The normalization mechanism keeps the pattern from saturation in spite of the constant presence of the pattern during the learning process. The $F_1$ layer provides internal feedback and a correlation between normalized bottom-up and top-down patterns to stabilize all activities in the STM before transmitting the output of the $F_1$ layer to the $F_2$ layer.

3.2 The search phase

In the $F_2$, a competitive mechanism is used to choose a winning neuron. Firstly, the input pattern of the $F_1$ is applied to the bottom-up adaptive filter by the bottom-up adaptive weight $Z_{ij}$.

$$T_j = \sum p_i Z_{ij} \quad \text{for } j = 1, 2, \ldots, K$$  \hspace{1cm} (18)

where $K$ is the total number of existing categories in the $F_2$, then the vector $T$ is put in the order from minimum value to maximum one. We here suppose the $j$th neuron in the $F_2$ is selected if this neuron becomes maximally active one among the neurons not to be reset in the trial, i.e.

$$T_j = \max(T_j) \quad j = 1, 2, \ldots, K_j$$  \hspace{1cm} (19)

where $K_j$ is the total number of the categories not to be used, then only winning neuron in the $F_2$ has nonzero outputs.

$$g(T_j) = \begin{cases} d & \text{if the } j \text{th } F_2 \text{ neuron is the winner based on } \\
\max(\sum p_i Z_{ij}) \text{ and it has not been reset in the trial} \\
0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (20)

The top-down pattern $g(T_j)$ is then feedback to the $F_1$ by top-down adaptive weight $Z_{ij}$ and compared to the original bottom-up pattern to see if a correct match has been made by an activated orienting subsystem.

3.3 Orienting subsystem

The orienting subsystem helps to directly search for the categories in the $F_2$. When the subsystem is activated, the bottom-up pattern vector $p$ and the top-down pattern vector $u$ are utilized to calculate the degree of match (vector $r$)

$$r_i = \frac{u_i + c p_i}{c + \| u \| + \| c p \|}$$  \hspace{1cm} (21)

if the choice in the $F_2$ is correct, i.e.

$$\| r \| > \rho$$  \hspace{1cm} (22)

where $\rho$ stands for the vigilance factor or match sensitivity parameter. At this time, adaptive resonance is considered to have occurred and entered to the categories in the Long-Term Memory (LTM). If the choice is incorrect, another neuron with maximum output value among the existing neurons not to be selected will be selected as a possible winner candidate. The new candidate may cause yet another mismatch, hence another reset happens and the selection of yet another neuron, eventually, either the bottom-up pattern will be placed in an existing category or learned as the first example of a new category in the $F_1$ layer. It is possible for an autogrowing mechanism to be activated to create a new category if no category in the $F_1$ could be used to save the new one.

3.4 The Long-Term Memory (LTM)

The LTM is made up of two components, the bottom-up adaptive weight $Z_{ij}$ and the top-down adaptive weight $Z_{ij}$. When the match operation in the orienting subsystem occurs successfully, the bottom-up and top-down weights should be adjusted. The weights can be obtained easily by

$$Z_{ij} = \frac{u_i}{1 - d}$$  \hspace{1cm} (23)

$$Z_{ij} = \frac{u_i}{1 - d}$$  \hspace{1cm} (24)

The procedure in the improved ART2 can be summarized as:

Step 1. Initialize bottom-up and top-down adaptive weights $Z_{ij}$ and $Z_{ij}$ in the LTM.
Step 2. Apply a new input pattern.
Step 3. Stabilize the output vectors of (or $p$) by the $F_1$ layer by repeated operating Eqs. 11 ~ 16, including noise reduction and contrast enhancement by a nonlinear thresholding function and fuzzy mapping function.
Step 4. Compute the output vector $p$ by Eq. 18.
Step 5. Select a winner neuron by Eqs. 18 and 19 if neurons not to be selected exist, else go to step 7.
Step 6. Apply Eq. 21 to determine whether the selected top-down winner matches the bottom-up input or within a certain acceptance level of vigilance. If Eq. 21 is not true, the selected winner neuron in the $F_2$ is disabled and return to step 5 in order to choose another winner neuron; else go to step 8.
Step 7. Autogrowing mechanism is activated to create a new category.
Step 8. Only adjust the bottom-up and top-down adaptive weights with respect to the matched winner neuron by Eqs. 23 and 24.
Step 9. Before taking the next new input pattern, neuron which has been disabled in step 6 will be enabled. The process return to step 2 if a new input pattern at least exists, else exit the system.

Viewing from the improved ART2 algorithm, if the network for an input pattern has learned previously to recognize the pattern, then
a resonant state will be achieved quickly when that pattern is presented, adaptive process will reinforce the memory of the stored pattern by formulas. If the pattern is not immediately recognized, the network will rapidly search through its stored patterns looking for a match. If no match is found, it will enter a resonant state whereupon the pattern will be stored as a new category for the first time if unused neurons in the competitive layer exist. Otherwise, a new neuron should be created automatically by the autogrowing mechanism the $F_2$ layer to store the new pattern. Thus, the network is able to respond fastly to previously learned data, yet learn novel data when those are presented.
4. Experiment results

The performance of the segment algorithm by improved ART2 is examined by a series of experiments on image containing different textures. The size of each image is $100 \times 100$ with 256 gray levels. The size of maximum and minimum measuring window is defined as $11 \times 11$ and $33 \times 33$, respectively.

For the texture features from the image by fast SGLDM algorithm, the K-means algorithm is used $^{17}$ (shown in Fig. 2 (b)). However, the K-means algorithm has some disadvantages:

- Supervised learning mode: the number of clusters must be set in advance. The different number may classify different results;
- Slow real-time ability: time of classification will raise at exponent rate with the cluster number increased;
- Unstability: the results of classification depends on the precision of feature extraction. When the extraction of the texture features has slightly change, the classifying result might be different.

Compared to the K-means algorithm, the ART2 has many advantages, such as unsupervised training, high computation rates, and great degree of fault tolerance (stability/plasticity).

In our test, the features, i.e. energy, entropy, correlation, homogeneity and inertia (or called as contrast), are used in texture analysis. The features have been proved to be a high degree of accuracy for the extraction of texture image features. The parameters $a$, $b$, $c$, $d$, $e$, $\theta$ and $\rho$ is selected in advance. $a=b=10$, $c=0.25$, $d=0.8$, $e=10^{-6}$, the selection of $\theta$ depends on different texture features and quantized angles of the features. For instance, the noise of each angle for the energy feature in the test is similar, so the value of $\theta$ is selected as 0.23 in every angle of the feature. On the other hand, the noise of each angle for the contrast feature is slightly different. The $\theta$ is set to 0.1, 0.12, 0.2, and 0.1 for the feature along to angle $0^\circ$, $45^\circ$, $90^\circ$, and $135^\circ$, respectively. The Fig 2. (a) is the original texture image. The Fig 2. (c) is the segmenting result of the improved ART2 only with noise reduction. It is seen from the Fig. 2. (c) to (d) greatly improve the segmentation of the texture image. The Fig 2. (d) shows that the segmentation operation is further good after not only the noise reduction but also the feature enhancement are done.

5. Conclusion

The SGLDM provides the most powerful statistical representation for segmentation and identification of texture images. Its problem, consuming time has been improved greatly by a fast algorithm.

An improved Adaptive Resonance Theory (ART2) for analog input patterns is adapted to classify the image based on a set of texture image features extracted by a fast SGLDM. The non-linear thresholding functions in the ART2 $F_t$ layer have been composed of two parts: to reduce the effect of image noises on the features, a set of sigmoid functions is chosen depending on the types of the feature; to enhance the contrast of the features, we adapt fuzzy multi-region mapping functions. The cluster number in output layer can be increased by an autogrowing mechanism constantly when a new pattern happens. The system written by C language is performed on a SUN-4/330 spare-station with an image board IT-150 and a CCD camera.

6. Acknowledgements

The authors acknowledge the discussion made by A. Acciani and G. Bianco to provide us with C library.

References