MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

8—18

Texture Image Segmentation using Wavelet Filters and Cellular Neural Network

Guoxiang Liu
Department of Information Science
and Intelligent Systems
Faculty of Engineering,
University of Tokushima,

Shunichiro Oe
Department of Information Science
and Intelligent Systems
Faculty of Engineering,
University of Tokushima,

Abstract

This paper presents a texture segmentation algorithm based on Discrete Wavelet Frames(DWF) and Cellular Neural Network(CNN).

DWF, zero-crossing, texture energy and selective local averaging are used to get a texture feature extraction and to form feature images.

Each feature image is segmented into parts by several gray range in its gray histogram. Resulting in the number of pixels that conform to each gray range composition of every feature image, segmentation can be got by those compositions that have a big pixel number. We call this method "Composition-Array method". A new CNN called "Multi-objective CNN" is developed using the fundamental theory of the traditional CNN, it is a multi-objective neural processing instead of 1 and -1 only. The noise of Composition-Array processing can be removed using this new CNN, and we can get a perfect final segmentation result.

1 Introduction

The unsupervised segmentation of textured images is a difficult and challenging low level vision problem with important applications in vision guided autonomous robotics, product quality inspection, medical diagnosis and in the analysis of remote sensing images, it has been investigated by many researchers using a diversity of approaches. Many approaches have been reported, but none of them provides a texture segmentation with high correctness, high processing speed and high adaptability like human being.

In general, each method consists two phases: feature extraction and segmentation, and features for texture representation are very important for accomplishing segmentation. Neural Networks and GA are usually used although they are computationally inefficient for there is no feature extraction method good enough to make a fine segmentation. Previous approaches for representing texture feature can be divided into two categories: 1.Statistical approaches: co-occurrence matrixes, autoregressive moving average, etc; 2.Spatial/spatial-frequency approaches: FFT, Gabor filters, Discrete Wavelet Transformation(DWT),etc. Recently, the Spatial/spatial-frequency approaches are remarkable, they can extract feature within a selected bandwidth along a selected orientation to provide more information. Although Gabor filters are easy to design and have desirable properties including orientation selectivity and filter bandwidth, they are computationally inefficient. Andrew Laine and Jian Fan presented a new method based on DWF and zero-crossing with several advantage such as fast processing speed[3].

On the other hand, there are so many kinds of texture and there is no texture definition that can be agreed by every researcher, it is impossible to segment or recognize all of those kinds of texture using only one texture feature or one kind of texture feature. So, the problem is how to integrate multiple features to produce a segmentation with a high processing speed. Cellular Neural Networks were introduced by Chua and Yang in 1988[1]. Since each cell is locally interconnected within its neighborhoods and it has feedback, this configuration gives a very high speed tool for parallel dynamic process. Despite of its many advantage, the traditional CNN hasn't been used in many fields, it has a fatal issue: it can only provide two kinds of result -1 and 1. In this paper, we propose a new method based on Gray histogram, Composition-Array and Multi-objective CNN to produce a segmentation using nine feature images got from the original image using DWF, zero-crossing and selective local averaging.

2 Texture Feature

2.1 Wavelet Frames

A Wavelet transform decomposes a 1-D signal $f(x)$ onto a basis of Wavelet functions:

$$(W_a f)(b) = \int f(x)\psi^*_{a,b}dx \qquad (1)$$

Such a basis, which is usually taken complete and orthogonal, is obtained translating and dilating a single mother Wavelet ψ:

$$\psi_{a,b} = \frac{1}{a^{\frac{1}{2}}}\psi(\frac{x-b}{a}) \quad (2)$$

The mother Wavelet ψ is localized in both spatial and frequency domain and it has to satisfy the constraint of having zero mean.

When a and b are restrained to a discrete lattice ($a = 2^n, b \in \mathcal{Z}$), the discrete Wavelet transform(DWT) is obtained. The DWT has an efficient implementation in the real space which uses quadrature mirror filters. Every Wavelet corresponds with a high and low pass filter. For the most common case with dilation by a factor of two, the scheme is called "dyadic" wavelet transform. The Wavelet decomposition of an 2-D image can be obtained by performing the filtering consecutively along horizontal and vertical directions, in general.

In this work, we use an over-complete Wavelet decomposition called a discrete wavelet frame(DWF). It is similar to discrete wavelet transformation(DWF) except that no down sampling occurs between levels. In other words, the size of transformed image is not changed. We want to get a segmentation in pixel, but the small difference of textures in high levels can not be represented by DWT for its down sampling, hence, DWT is not suitable. To get more feature information, in each level, we get three wavelet transformed images filtered along horizontal direction, vertical direction and horizontal direction + vertical direction with high-pass filter respectively. We call them *wavelet feature H*, *wavelet feature V* and *wavelet feature HV*. Nine filtered images in three levels are used in experiments. This is depicted schematically in Figure 1.

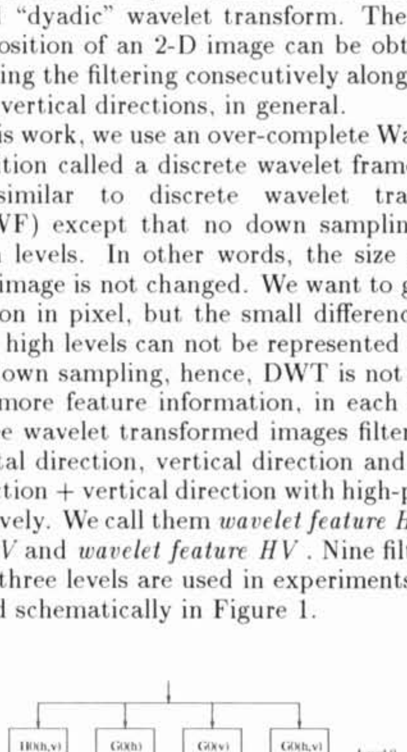

Figure 1: Tree structure for Wavelet frame used in this paper. G: high-pass filter. H: low-pass filter.

2.2 Zero Crossing

In order to get a fine feature extraction, we also perform zero-crossing transformation. To understand this extraction scheme, the concept of zero-crossing must be explained first. The result of wavelet transformation is zero-means, and if there is zero value somewhere in the result of wavelet transformation, we say there is a zero-crossing there. In this scheme, the maximum absolute deviation from the mean between two adjacent zero-crossing is found and assigned to all points within the interval. Actually, we perform zero-crossing transformation along vertical direction for *wavelet feature H*, along horizontal direction for *wavelet feature V* and *wavelet feature HV*.

2.3 Computing Feature Images

The following procedures are used to compute features from each zero-crossing transformed image. First, the grey-scale value of each transformed image is changed to the range of $[0, 1]$, then we use the following bounded nonlinearity:

$$\psi(t) = tanh(\alpha t) = \frac{1 - e^{-2\alpha t}}{1 + e^{-2\alpha t}} \quad (3)$$

where $\alpha = \frac{1}{av}$, av is the average of gray value of the transformed image. As a result, the application of the nonlinearity transforms the zero-crossing transformed images to square modulations.

We simply compute the average of the nonlinearity transformation results in small overlapping windows as the second stage of computing feature images. This is similar to the texture energy that was first proposed by Laws. The feature images $e_k(x, y)$ corresponding to transformed image $r_k(x, y)$ is given by:

$$e_k(x, y) = \frac{1}{M^2} \sum_{(a,b)\in W_{xy}} r_k(a, b) \quad (4)$$

where W_{xy} is an $M \times M$ window centered at the pixel with coordinates (x, y), and $M = 17$ is used in experiments.

After getting texture energy. We will segment the feature image by its gray histogram, but features near the edge between high gray value and low gray value perhaps have been changed to be in a wrong gray histogram segmentation result by texture energy processing. Hence, it is necessary to accomplish edge restoration processing. In this paper, selective local averaging is used to accomplish edge restoration processing.

3 Segmentation

3.1 Composition-Array Method

Having computed feature images(we call them single feature images), the main question is how to integrate those features to produce a segmentation. Traditionally, there are some methods such as M-Means, CLUSTER, etc, but none of them can provide a high speed computing. In this paper, we propose a new method to get segmentation to save time.

Using its gray histogram, one feature image can be segmented into several parts, and the number of pixels that conform to each gray range composition of each feature image can be computed after segmenting every feature image. This is given by:

$$
\begin{aligned}
C_{r_1 r_2 \cdots r_n} &= \sum_{(a,b)\in I} G^{(a,b)}_{r_1 r_2 \cdots r_n} \\
G^{(a,b)}_{r_1 \cdots r_i} &= \begin{cases} g_{r_1}(a,b) & i=1, \\ G^{(a,b)}_{r_1 \cdots r_{i-1}} \times g^{(a,b)}_{r_i} & otherwise \end{cases} \\
g^{(a,b)}_{r_i} &= \begin{cases} 1 & f_i(a,b) \in r_i \\ 0 & otherwise \end{cases} \\
& 1 \le i \le n; \\
& r_1 \in R_1, r_2 \in R_2, \cdots, r_n \in R_n
\end{aligned} \tag{5}
$$

where C is the number of pixels that conform to a composition, R_i is the gray ranges group segmented by the gray histogram of single feature image f_i, and r_i is one gray range in R_i. I is the original image. n is the number of single feature images used. Figure 2 shows the number of pixels that conform to every composition as a example used in section 4, you can see which is the one we want easily. We believe that compositions with too small pixel number are noise, and they are omitted in experiments. Only big ones are considered to be segmentation results. We call this method "Composition-Array" method.

Now, we know how many kinds of texture in original image, and one texture conform to which gray range of each single feature image. All of these information will be used in CNN processing as input.

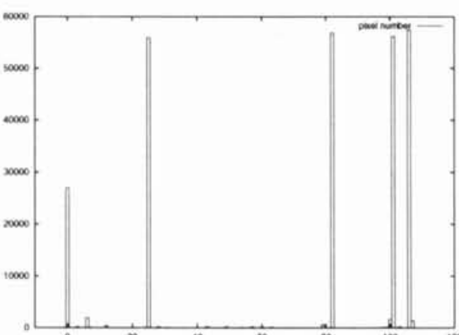

Figure 2: number of pixels that conform to compositions

3.2 A New CNN: Multi-objective CNN

A new 2-D CNN was developed using the fundamental theory of the traditional CNN. The input of the traditional CNN is the probability that one cell belongs to -1 or $+1$ in fact. In this new CNN, input, state and output of one cell have multiple vector instead of one only, the input of one vector of one cell is the probability that this vector(objective) represents the meaning of that cell. The main equations are shown in Equation 6, 7.

State equation of New CNN:

$$
\begin{aligned}
C\frac{dv_{xijp}(t)}{dt} &= -\frac{1}{R_x} v_{xijp}(t) \\
&+ \sum_{C(k,l)\in N_r(i,j)} A(i,j;k,l) v_{yklp}(t) \\
&+ \sum_{C(k,l)\in N_r(i,j)} B(i,j;k,l) v_{uklp}(t) + I, \\
& 1 \le i \le M; 1 \le j \le N. \\
& 1 \le p \le P
\end{aligned} \tag{6}
$$

Output equation of New CNN:

$$
\begin{aligned}
v_{yijp}(t) &= \frac{1}{2}\left(\left|\frac{Tv_{yijp}(t)\times 2}{\sum_{1\le q\le P} Tv_{yijq}(t)} + 1\right| - \left|\frac{Tv_{yijp}(t)\times 2}{\sum_{1\le q\le P} Tv_{yijq}(t)} - 1\right|\right) \\
Tv_{yijp}(t) &= \frac{1}{2}(|Tv_{xijp}(t)| + Tv_{xijp}(t)) \\
Tv_{xijp}(t) &= \frac{v_{xijp}(t)}{\sum_{1\le q\le P} v_{xijq}(t)} - \frac{1}{P} \\
& 1 \le i \le M; 1 \le j \le N. \\
& 1 \le p \le P; 1 \le q \le P.
\end{aligned} \tag{7}
$$

Where A and B are parameters; R_x is a linear resistor; I is an independent current source; P is the number of objective; $N_r(i,j)$ are neighborhoods of a 2-D CNN cell (i,j); and $M \times N$ is the size of CNN.

From Equation 6, you can see that a new state value is given by the output of last time, input and parameters. How much state value receive effect from input or output and the relation between center cell and its neighborhoods can be controlled by changing the parameter A and B. We call this new CNN "Multi-objective CNN".

4 Experimental results

We did lots of experiments using many mixed 512×512 texture images to test the correctness, the processing speed and the adaptability of this method. Figure 3-(a) shows a original image mixed with five kinds of texture, and its segmentation was produced using our new method as a example in this section. First, we got nine wavelet filtered images in three levels using Burt-Adelson wavelet basis(Figure 3-(b) is one wavelet filtered image in vertical direction level 0). Then performed zero-crossing transformation of every wavelet filtered images, got texture energy of them, accomplished edge restoration processing, and then computed the final features.

Gray histogram was used to segment each feature image into parts. Peak and valley were found first, and gray value of the deep valleys(whose histogram value were smaller than 20% of the shorter peak beside them) were the segmentation points. For example, feature image shown in Figure 3-(e) was segmented into three parts by its gray histogram shown in Figure 4. Then these segmentation results were used in Composition-Array processing as input. Figure 2 shows the pixel number that conform to every composition and Figure 3-(f) shows the result of Composition-Array processing, where the parts in white are noise. Finally we accomplished Multi-objective CNN processing to get final segmentation result shown in Figure 3-(g) with a error rate of 1.75% using $R_x = 1$, $I = 0$ and

$$A = \begin{bmatrix} 1.0 & 2.0 & 1.0 \\ 2.0 & 5.0 & 2.0 \\ 1.0 & 2.0 & 1.0 \end{bmatrix} \quad B = \begin{bmatrix} 0.1 & 0.2 & 0.1 \\ 0.2 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.1 \end{bmatrix}$$

In order to save processing time and omit the small different part, first, we used a 16×16 windows as a cell, then 4×4 and then pixel. The input of one vector of one cell is how much the number of pixels conform to one gray range composition be divided by the pixel number in that cell , and we only did CNN processing along edge. Some processing results are shown in Figure 3.

5 Conclusions

Composition-Array method and Multi-objective CNN were proposed. Using them, we can segment a 512×512 image in just a few minutes with high correctness. Many features can be used in this method, and, it is easy to apply more features(including other kinds of feature) to increase adaptability.

From experiments we know, Wavelet Frame methods have desirable properties about frequency, but they are inefficient to segment texture images with a same frequency but a different shape. So, in future work, we will apply other kinds of feature representation methods to get features, then use Composition-Array method and Multi-objective CNN to get the final segmentation result.

References

[1] L.O. Chua, L.Yang "Cellular Neural Networks", *IEEE Transactions on Circuit and Systems*, vol. 1.35, pp. 1257-1290, 1988.

[2] Anil K.Jan Farshid Farrokhnia "Unsupervised Texture Segmentation Using Gabor Filters", *Pattern Recognition*, vol. 24, No.12, pp.1167-1186,1991.

[3] Andrew Laine, Jian Fan "An Adaptive Approach for Texture Segmentation by Multi-channel Wavelet Frames".

[4] M. Vetterli and C.Herley. "wavelets and filter banks: theory and design", *IEEE Trans. Signal Processing*, vol. 40,pp. 2207-2232, 1992.

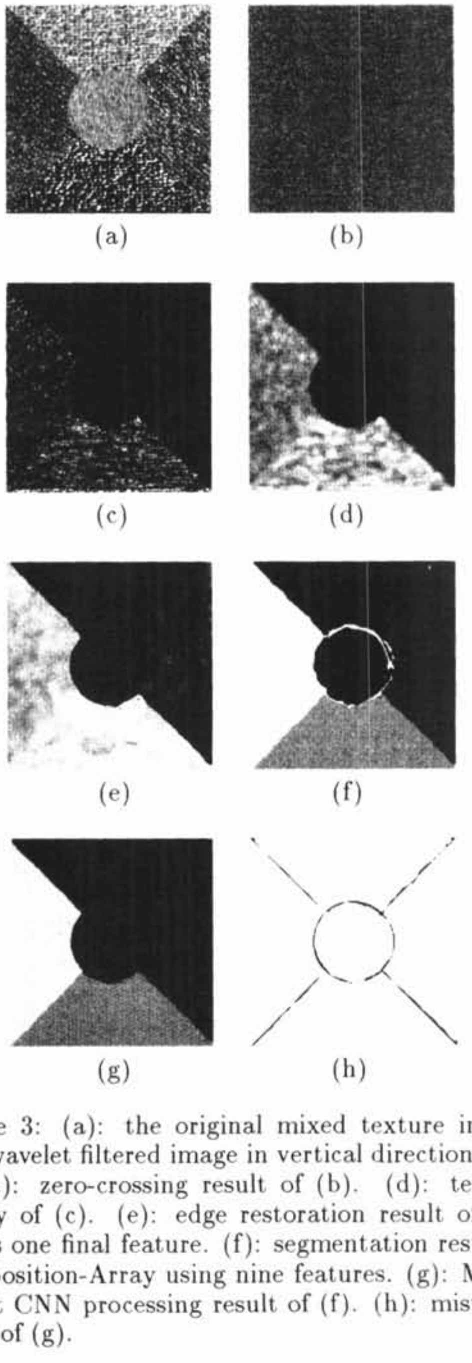

Figure 3: (a): the original mixed texture image. (b): wavelet filtered image in vertical direction level 0. (c): zero-crossing result of (b). (d): texture energy of (c). (e): edge restoration result of (d), this is one final feature. (f): segmentation result of Composition-Array using nine features. (g): Multi-object CNN processing result of (f). (h): mistaken parts of (g).

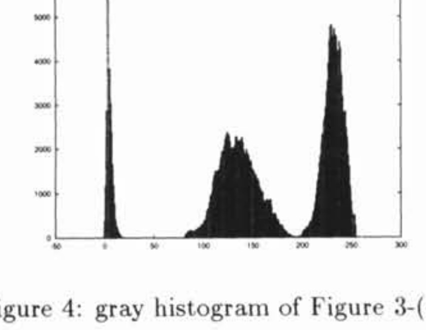

Figure 4: gray histogram of Figure 3-(e)