# Texture Discrimination Using an Adaptive Multiresolution Network Filter

Hiroshi SAKOU, Hitoshi MATSUSHIMA, and Masakazu EJIRI

Central Research Laboratory, Hitachi Ltd.,
1-280 Higashi-Koigakubo, Kokubunji, Tokyo 185, Japan
(Tel: 0423-23-1111, E-mail: sakou@hcrlgw.crl.hitachi.co.jp)

## ABSTRACT

This paper proposes a new discrimination method for textured images using an adaptive multiresolution network filter. First, the local transformation function is determined by a modified GMDH (Group Method of Data Handling) so as to have a high discrimination ability. Next, the method is extended to obtain a function in terms of multiresolution pixel densities for the discrimination of an image with an unknown texture element size. Finally, the grouping of areas having the same texture characteristics is achieved by applying the same procedure to the output of the function. The practicability of this filter is experimentally confirmed using several kinds of images.

## INTRODUCTION

The purpose of texture discrimination is to detect an area having a specific uniform texture from the entire image. Texture discrimination is important for image segmentation especially in remote sensing and shape-from-texture areas. Up to date, some useful methods for texture discrimination have been developed. These are based on cooccurence matrices [1], Fourier analysis [2], locally-defined filters [3],[4], a Markov random field model [5], and an autoregressive model [6]. In the discrimination method [4], a non-linear local transformation function which is expressed by the polynomial of some neighbouring pixels is obtained, and the texture class of each pixel is determined by estimating the output of the function at the pixel. To obtain the transformation function, GMDH (Group Method of Data Handling) [7] developed by Ivakhnenko is used. The advantages of this discrimination method are;

(i) The way to select neighbouring pixels and combine them so as to have the best discrimination ability is automatically determined in the process of making the transformation function. Other methods need two steps which are characteristics determination and discriminant function determination in feature space.

(ii) The method can be easily translated into hardware for high speed processing once the transformation function is obtained, because the function is composed of a number of multiply and add operations between neighbouring pixels and coefficients.

On the other hand, problems associated with this method are;

(i) The discrimination ability depends on the neighbourhood size to be considered.

(ii) The computational cost to obtain the transformation function becomes expensive because the number of neighbouring pixel combinations significantly increase when the large neighbourhood is required.

Though the discrimination ability is high only when the neighbourhood size equals to approximately that of the texture element size, as described in [4], it suggests that many trials are needed to decide the suitable neighbourhood size when the method is applied to an actual image which is composed of several textured patterns with unknown element sizes. It is also known that the neighbourhood size affects the discrimination ability in other methods [8].

This paper proposes a new discrimination method using GMDH, which is applied to actual textured images with unknown element sizes. The first part of this paper describes a new process to obtain the local transformation function, which is expressed by the polynomial of the selected pixels in a bounded neighbourhood so as to obtain the discrimination ability as high as possible. The next part describes the modification of the process to the discrimination at several resolutions. Finally, the proposed method is applied to several kinds of textured images. The practicability of the method is confirmed through these experiments.

## BASIC CONCEPT

When a human recognizes each textured pattern in an image, it seems that smaller areas having same texture characteristics are combined with each other to become a larger area. This is after he has observed each textured pattern at every resolution and preattentively selected several representative characteristics at a certain resolution to discriminate the object textured pattern among others. Three important points are learned from this human recognition process.

(1) Some characteristics which are useful to discriminate the object textured pattern from only neighbouring textured patterns and not from all other patterns in the image, are selected even if the entire image includes many kinds of textured patterns.

(2) The characteristics are selected from textured patterns at a certain resolution where the ability to discriminate the object textured pattern from others is the greatest.

(3) Smaller areas having the same characteristics are combined with each other to form the whole object textured pattern to be detected.

This paper proposes a new discrimination method using an adaptive multiresolution network filter based on GMDH to achieve this human recognition process.

## ADAPTIVE MULTIRESOLUTION NETWORK FILTER

Let an image be composed of two textured patterns Ta and Tb, and let I(x,y) denote the density of the pixel located at the coordinate (x, y) in the image. The original density I(x,y) can be changed into the density $I_k(x,y)$ at resolution k by the following transformation with a two-dimensional Gaussian filter;

$$I_k(x,y) = \frac{1}{\sqrt{2\pi}\sigma_k} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} I(\xi,\zeta) \exp\{-\frac{(x-\xi)^2+(y-\zeta)^2}{2\sigma_k^2}\} d\xi d\zeta , \quad (1)$$

where $\sigma_k$ is a standard deviation and k is used to express the difference of resolution. Next, as shown in Fig.1, $I_k(x,y)$ at several resolutions are calculated from $I(x,y)$ and several neighbouring pixels at $(x+\Delta x_i,\ y+\Delta y_i)$ are selected at each resolution. Approximate ranges of the filter operation are denoted by several squares and the selected neighbouring pixels are denoted by the + symbol in the figure. Consider a local transformation function F, hereinafter called LTF, which is expressed by a high-order polynomial in terms of pixels $I_{ki}$ $(=I_k(x+\Delta x_i,\ y+\Delta y_i))$ and satisfies the following condition;

$$F(I_{00}, I_{01}, \dots, I_{ki}, \dots) = \begin{cases} 1 & \text{for all } (x,y) \in Ta, \\ -1 & \text{for all } (x,y) \in Tb, \end{cases} \tag{2}$$

where i is used to express the difference of the selected neighbouring pixel. This LTF can realize three points in the human recognition process described in the previous section because the first and third points correspond to the LTF transforming every pixel density of Ta and Tb into 1 and -1, respectively. Moreover, the determination of LTF means not only the determination of the order and coefficients of the polynomial but also the selection of some useful neighbouring pixels at a certain resolution as variables. These determinations and selections correspond to the second point in the human recognition process.

Fig.2 shows the outline of an LTF, which is the multiresolution network filter this paper proposes. This filter is composed of a series of neigbourhood transformation functions, NTF's. Each NTF has the same structure and is a non-linear function in terms of neighbouring pixel densities at several resolutions of each input image. Partial areas having same characteristics on the original textured image are combined with each other gradually through each NTF, and finally the classified image which has pixel density areas approximately 1 and -1 corresponding to the oringinal Ta and Tb is obtained. Fig.3 briefly explains the structure of an NTF, and gives a case of detecting four neighbouring pixels each at four resolutions (k=0, 1, 2, 3). The NTF is composed of several 0th basic transformation functions, 0th BTF's, in terms of some useful neighbouring pixel densities on the input image at several resolutions. The NTF is also composed of several 1st through nth BTF's in terms of several outputs of their previous BTF's. The 0th BTF's contributes the multiresolution of the filter, and the 1st through nth BTF's make the discrimination ability of the filter as high as possible by combining some neighbouring pixel densities in the form of a high order polynomial.

The following sections describe how to determine each function corresponding to the portion 3.1 and 3.2 in Fig.3, and the portion 3.3 in Fig.2.

**Determination of the neighbourhood transformation function**: This section describes how to select and combine the pixels in a bounded neighbourhood of $I_0(x,y)$ at a resolution $\sigma_0$ so as to obtain the best discrimination ability. In addition, it describes how to make the neighbourhood function NTF by using those pixels. In short, the learning area ta and tb are first set within the textured pattern area Ta and Tb, respectively, and then the NTF, which can transform every pixel density on the former and the latter area into 1 and -1, is determined by the pixel density information from the learning areas.

[Step 1] Set the learning area ta and tb in the textured areas Ta and Tb, respectively.

[Step 2] Obtain some pixel densities $I_{00}, I_{01}, \dots, I_{0i}, \dots$ in a bounded neighbourhood of $I_0(x,y)$, and select two pixels $I_{0i}$ and $I_{0j}$ $(i \neq j)$ from them. Calculate the coefficients $a_{00}$, $a_{01}$, $a_{02}$ and $a_{03}$ of the 0th basic transformation function,

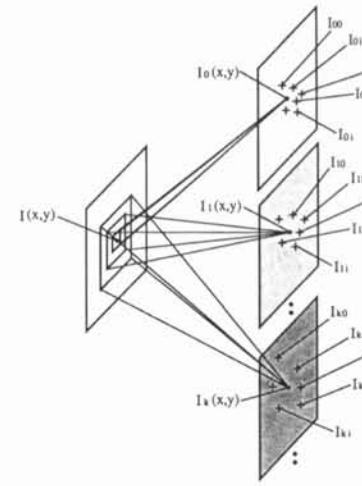



Fig.1 Selection of neighbouring pixels at every resolution.

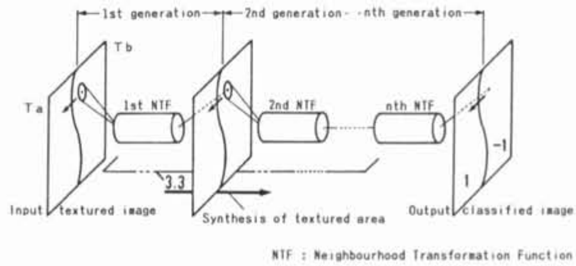



Fig.2 Outline of an adaptive multiresolution network filter.

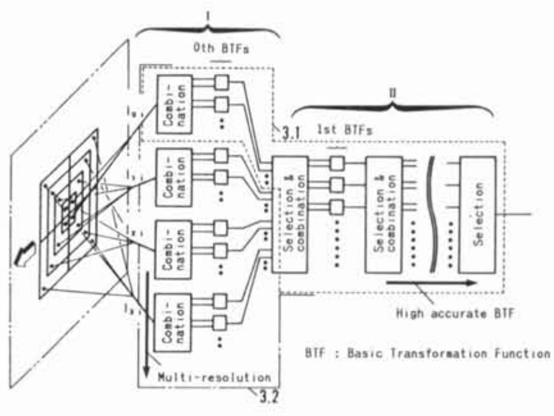



Fig.3 Outline of a neighbourhood transformation function.

0th BTF, $f'_{0ij}(x,y)$ expressed in the form of Equ.(3) by Lagrange's method of indeterminate coefficient so as to satisfy Equ.(4) and minimize the summation $s_{0ij}$ of dispersions $s^a_{0ij}$ and $s^b_{0ij}$ in Equ.(5) as shown in Fig.4.

$$f'_{0ij}(x,y) = a_{00} + a_{01}I_{0i} + a_{02}I_{0j} + a_{03}I_{0i}I_{0j}, \tag{3}$$

$$\frac{1}{A}\sum_{(x,y)\in ta} f'_{0ij}(x,y) = 1, \quad \frac{1}{B}\sum_{(x,y)\in tb} f'_{0ij}(x,y) = -1, \tag{4}$$

$$\begin{aligned} s_{0ij} &= s^a_{0ij} + s^b_{0ij}, \\ &= \frac{1}{A}\sum_{(x,y)\in ta} \{f'_{0ij}(x,y) - 1\}^2 + \frac{1}{B}\sum_{(x,y)\in tb} \{f'_{0ij}(x,y) + 1\}^2, \end{aligned} \tag{5}$$

where A and B are the total pixel numbers of the learning area ta and tb. The function $f'_{0ij}$ has x and y variables because $I_{0i}$ and $I_{0j}$ are pixel densities of which positional relationships against the processed pixel position (x, y) are preserved. This step corresponds to part I of the portion 3.1 in Fig.3, and is executed just for the number of combinations of two different pixels from the number of pixel densities $I_{00}, I_{01}, \ldots, I_{0i}, \ldots$ . Therefore, the number of 0th BTF obtained in this step is the same as the combination number. The purpose of this step is to select neighbouring pixels and determine the coefficients with high discrimination ability under the condition that the average output of the obtained 0th BTF becomes 1 in the area ta and -1 in the area tb.

[Step 3] Select h of the 0th BTF's which have a high discrimination ability, in other words, a small $s_{0ij}$. Regard the outputs of the selected 0th BTF's as the inputs to the following 1st BTF's. Let the indeces ij of the selected BTF's be $i_0j_0, i_1j_1, \ldots, i_{h-1}j_{h-1}$, and hereinafter $f'_{np}$ (n=0 in this step) is used instead of $f'_{0ij}$, where $1 \le p \le h$, p=1 corresponds to $i_0j_0$, p=2 corresponds to $i_1j_1$, ... , p=h corresponds to $i_{h-1}j_{h-1}$.

[Step 4] Calculate the coefficients $a_{m0}$, $a_{m1}$, $a_{m2}$ and $a_{m3}$ of the mth BTF $f'_{mpq}(x,y)$ (m=n+1, $1 \le p \le h$, $1 \le q \le h$, $p \ne q$) expressed in the form of Equ.(6) by Lagrange's method of indeterminate coefficients so as to satisfy Equ.(7) and monomize the summation $s_{mpq}$ of dispersions $s^a_{mpq}$ and $s^b_{mpq}$ in Equ.(8) as shown in Fig.5.

$$f'_{mpq}(x,y) = a_{m0} + a_{m1}f'_{np}(x,y) + a_{m2}f'_{nq}(x,y) + a_{m3}f'_{np}(x,y)f'_{nq}(x,y)\,, \tag{6}$$

$$\frac{1}{A}\sum_{(x,y)\in ta} f'_{mpq}(x,y) = 1\,, \quad \frac{1}{B}\sum_{(x,y)\in tb} f'_{mpq}(x,y) = -1\,, \tag{7}$$

$$\begin{aligned} s_{mpq} &= s^a_{mpq} + s^b_{mpq}\,, \\ &= \frac{1}{A}\sum_{(x,y)\in ta}\{f'_{mpq}(x,y) - 1\}^2 + \frac{1}{B}\sum_{(x,y)\in tb}\{f'_{mpq}(x,y) + 1\}^2\,, \end{aligned} \tag{8}$$

The purpose of this step is to obtain the next mth BTF by combining two nth BTF's. The mth BTF become a higher order polynomial than the nth BTF since Equ.(8) has a non-linear term, and is expressed only in terms of original neighbouring pixels $I_{00}, I_{01}, \ldots, I_{0i}, \ldots$ . The number of mth BTF obtained in this step is the same as the number of combinations of two differnt nth BTF's from h.

[Step 5] Check if the smallest $s_{mpq}$ is less than the predetermined small value or larger than the smallest value of the last $s_{npq}$. In either of these cases the final step is executed. Otherwise the next step is executed.

[Step 6] Select h' number of mth BTF's of which $s_{mpq}$ are small and less than the smallest value of the last $s_{npq}$, and regard their outputs as the inputs to Step 4. Steps 3 and 4, and the iterations of Step 4 from Step 6 correspond to part II of portion 3.1 in Fig.3. By these iterations of Step 4 the BTF gradually becomes a higher order polynomial with a smaller $s_{mpq}$.

[Final Step] Find the mth BTF with the smallest $s_{mpq}$, and construct a neighbourhood transformation function NTF in ascending order continuously from the BTF to its parent BTF's. The construction of the NTF means the fixing of the network route in Fig.3 or the acquisition of the formulas in the equation form.

One of characteristics of this discrimination method is based on the iterations of Step 4. By these iterations the BTF becomes a higher order polynomial with the pixels in a bounded neighbourhood. In the case of h=3, for example, three 0th BTF's, $f'_{01}$, $f'_{02}$ and $f'_{03}$, of which variables are pixel densities in the combinations of $(i_0, j_0)$, $(i_1, j_1)$ and $(i_2, j_2)$, respectively, are obtained in Step 3. In Step 4 three 1st BTF's, $f'_{112}$, $f'_{113}$, and $f'_{123}$, are obtained by the combinations of $(f'_{01}, f'_{02})$, $(f'_{01}, f'_{03})$ and $(f'_{02}, f'_{03})$, respectively. These BTF's are 4th order polynomials with pixel density variables in the combination of $(i_0, j_0, i_1, j_1)$, $(i_0, j_0, i_2, j_2)$ and $(i_1, j_1, i_2, j_2)$, respectively. By each iteration the order of the BTF polynomial is doubled and the polynomial is expected to have a higher discrimination ability than the last iteration. This is because the polynimial is non-linearly composed only of two less-ordered polynomials which have high discrimination abilities in the last iteration.

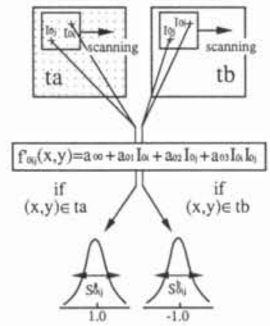



Fig.4 Determination of 0th transformation function.

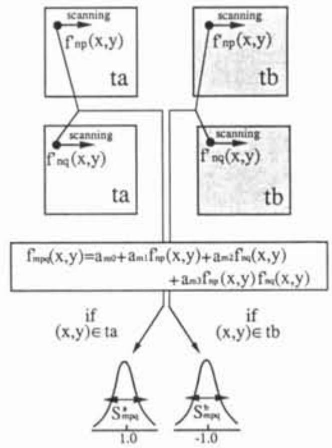



Fig.5 Determination of mth transformation function.

**Multiresolution network:** To deal with an image with an unknown texture element size, it is necessary to use a somewhat larger estimate of the neighbourhood size, and to use many pixel densities within the neighbourhood for discrimination. In this case the combination of pixel densities for obtaining characteristics of the texture is significantly increased and clearly the computational cost is expensive. To avoid this problem, the network described in the previous section is modified to the multiresolution network, which is able to adaptively use low or high resolutional pixel densities for the large or small texture element image. To achieve this, the function which can use several neighbouring pixel densities $I_{ki}$ for discrimination is added to the network as shown in the portion 3.2 of Fig.3. These pixel densities $I_{ki}$ can be obtained by applying a Gaussian filter with a standard deviation $\sigma_k$ to the input image at the coordinates $(x+\Delta x_i, y+\Delta y_i)$, where the distance $(\Delta x_i, \Delta y_i)$ from (x, y) is relative to its deviation value. By this modification, the density information of a larger neighbourhood is expressed by less number of pixel densities, and that prevents the explosion of combinations. The following is for the modification of the previous procedure.

[Step 2] Obtain some pixel densities $I_{k0}, I_{k1}, \dots, I_{ki}, \dots$ in a neighbourhood of $I_k(x,y)$, and select two pixels $I_{ki}$ and $I_{kj}$ ( $i \neq j$ ) from them at every resolution $\sigma_k$. Determine the 0th BTF $f'_{0kij}(x,y)$ at every resulution $\sigma_k$ so as to minimize the summation $s_{0kij}$ of dispersions. The conditions for this 0th BTF are the same as in Equ.(3), (4) and (5) except for thier notations.

[Step 3] Select h of the 0th BTF's which have a small $s_{0kij}$ and regard the outputs of the selected 0th BTF's as the inputs to the following 1st BTF's. Let the indeces kij of the selected BTF's be $k_0i_0j_0, k_1i_1j_1, \dots, k_{h-1}i_{h-1}j_{h-1}$, and hereinafter notate $f'_{np}$ (n=0 in this step) as $f'_{0kij}$, where $1 \leq p \leq h$, p=1 corresponds to $k_0i_0j_0$, p=2 corresponds to $k_1i_1j_1, \dots$, p=h corresponds to $k_{h-1}i_{h-1}j_{h-1}$.

This modification makes it possible to select the optimum combination of pixel densities for discrimination automatically even if the object image has a different texture element size.

**Synthesis of textured area:** It is necessary for the final discrimination to combine the partial areas having similar texture charactaristic to form one area. For this purpose the NTF is operated repeatedly on the output image of the last NTF as shown in the portion 3.3 of Fig.2. Hereinafter, one operation of the NTF is called a "generation". Since the operation with pixel densities within a certain neighbourhood is executed in Step 2 of every generation, the successive generation results in the combination of a series of pixel densities within a larger neighbourhood in the useful form for discrimination. The stop condition for generation is for the smallest $s_{nkij}$ in the present generation to become less than the predetermined value. The local transformation function F, is constructed by using all NTF's of the first to the present generation.

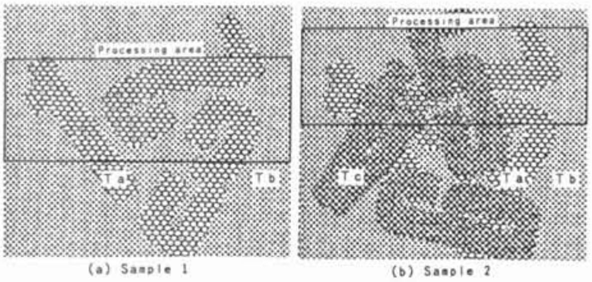



Fig.6 Texture pattern samples.

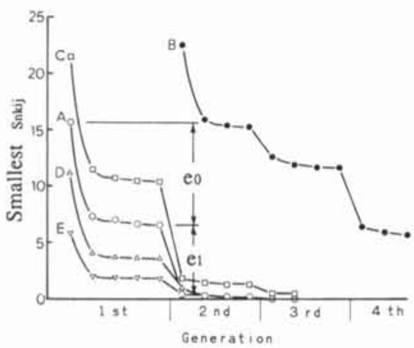



Fig.7 Convergence of the smallest Snkij.

## EXPERIMENTAL CONSIDERATION

In this section the discrimination ability of the multiresolution network filter is considered by using some experimental results. Fig.6 shows two textured samples used in the experiment which contain numbers in the processing area. Sample 1 has two textured areas Ta and Tb, and Sample 2 has three textured areas Ta, Tb and Tc. These textured images are digitized by a scanner with 16 gray levels and 8 lines/mm resolution, and the digitized data is fed into a computer. The proposed discrimination method is executed by this computer. The square area in the figure is the processing area (about 420×140 pixels). In the following, the neighbouring pixels are obtained by equ.(1) with resolutions $\sigma_0$=0.0, $\sigma_1$=2.0, and $\sigma_2$=4.0, respectively:

$$\begin{aligned}
(I_{00}, I_{01}, I_{02}, I_{03}) &= (I_0(x-1,y-1), I_0(x,y-1), I_0(x-1,y), I_0(x,y)),\\
(I_{10}, I_{11}, I_{12}, I_{13}) &= (I_1(x-1,y-1), I_1(x,y-1), I_1(x-1,y), I_1(x,y)),\\
(I_{20}, I_{21}, I_{22}, I_{23}) &= (I_2(x-2,y-2), I_2(x+1,y-2), I_2(x-2,y+1), I_2(x+1,y+1)).
\end{aligned} \tag{9}$$

They are used in Step 2 of each generation and both parameters h and h' are set to 4.

**Effects of step iterations and generation changes:** The learning areas ta and tb, each of which has 60×60 pixels, are set at their appropriate position in the textured areas Ta and Tb of sample 1, and the LTF is obtained by the process described in the previous section.

The line ( -O- ) indicated by A in Fig.7 describes the convergence of the smallest $s_{nkij}$ according to the iteration of Step 4 and each generation. A decrease in the value of the smallest $s_{nkij}$ means an increase of discrimination ability. One of the characteristics of this method is that the iteration of Step 4 makes the order of NTF higher and results in a high discrimination ability. $e_0$ and $e_1$ in the figure show the effect of the Step 4 iterations and the generation changes. The figure indicates that both of them play an important role to obtain an LTF with a high discrimination ability.

The transformation of the image by the NTF at the 3rd generation is shown in Fig.8(a). Since the output image at each generation has gray levels around 1 and -1, it is difficult to represent the image. Therefore, the binary images obtained by thresholding the output images are shown in the figure. The histogram of the output image pixel densities at each generation is shown in Fig.9. In the figure, each pixel density converges to 1 or -1, according to the generation.

**Effect of multiresolution:** In this chapter, changes of the discrimination ability in a single resolution are examined to compare with the effect of multiresolution. This effect is another characteristic of this method. The line indicated by B in Fig.9 shows the transition of the smallest $s_{nkij}$ in the case which uses the pixels ($I_{00}, I_{01}, I_{02}, I_{03}$) of a single resolution at every generation. The output image at the 3rd generation is shown in Fig.8(b). These figures show that the discrimination ability in a single resolution is lower than that in a multiresolution.


Fig.8 Transformed image at the 3rd generation in sample 1.

**Discrimination of a textured area from an image with three textured areas:** In the case of an image having three texured areas, the LTF which can discriminate one textured area from others is obtained by setting one learning area so as to have two textures. Fig.10(a), (b) and (c) show the discrimination results of sample 2. These results correspond to the textured area Ta, the textured area Ta$\cup$Tc and the textured area Tc, respectively. The changes of discrimination ability of these cases are indicated by lines C, D and E in Fig.7, respectively.

**Application to other textured images:** The filter for discrimination proposed here is applied to several kinds of images to show the practicability of this filter. Each original image in Fig.11 is made up of two cloth texture patterns, Herringbone weave and French canvas; and Herringbone weave and Cotton canvas, respectively. The images are digitized from the photographic album *"Textures"* [9]. Fig.11(c)(d) show the discrimination results. Fig.12 shows the output image at each generation in the case of another sample image. As shown in these figures, they are almost completely discriminated by the filter.

## CONCLUSIONS

The adaptive multiresolution network filter is proposed in this paper. It can be applied to the textured patterns of unknown element sizes. The filter utilizes a local transformation function (LTF) which is composed of a series of neighbourhood transformation functions (NTF's). These functions and the method to construct them are characterized as follows;

(1) The NTF is organized in the form of a non-linear high-ordered polynomial of useful pixel densities which are selected from pixel densities in a bounded neighbourhood. They are selected in a bounded neighbourhood so as to have the highest possible discrimination ability. The selection and combination of pixel densities are based on the modified GMDH method.

(2) To deal with the image of an unknown texture element, the pixel densities at several resolutions are detected and prepared for the selection and combination.

(3) The NTF is operated repeatedly to combine the partial areas with similar texture characteristics into one area.

The method is applied to the detection of a specific textured area from several kinds of images to confirm the practicability of the filter.


## Acknowledgements

The authors wish to thank Dr. Yoshito Tsunoda and Mr. Shigeo Nagashima and Dr. Hiromichi Fujisawa, Central Research Laboratory of Hitachi Ltd., for their encouragement and cooperation during the course of this study.


## References


[1] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural feature for image classification," IEEE Trans. Syst., Man, Cybern., vol.SMC-3, pp.610-621(1973).

[2] G. O. Lendaris and G. L. Stanley, "Diffraction-pattern sampling for automatic pattern recognition," Proc. IEEE, vol.58, pp.198-216(1970).

[3] K. I. Laws, "Texture energy measures," Proc. of DARPA Image Understanding Workshop, pp.47-51(1979).

[4] M. Sato and M. Ogata, "Texture analysis by the self-organization method," 7th ICPR, pp.1213-1215(1984).

[5] M. Hassner and J. Sklansky, "The use of random fields as models of texture," Computer Graphics and Image Processing, vol.12, pp.357-370(1980).

[6] C. W. Therrien, "Linear filtering models for texture classification and segmentation," 5th ICPR, pp.1132-1135(1980).

[7] A. G. Ivahknenko, "The Group method of data handling, A rival of the method of stochastic approximation," Soviet Automatic Control, vol.1, pp.43-55(1968).

[8] R. W. Conners, and C. A. Harlow, "Theoretical comparison of texture algorithms," IEEE Trans. Pattern Analysis & Mach. Intell., PAMI-2, 3, pp204-222(1980).

[9] P. Brodatz, *Textures: A Photographic Album for Artists and Designer*. Dover publication, Inc., New York (1966).

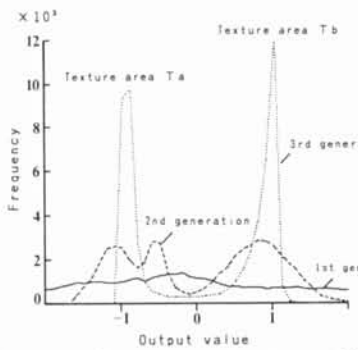
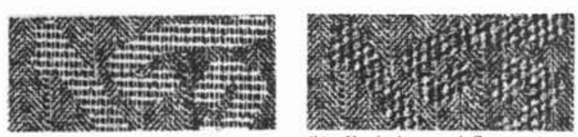




Fig.9 Aspect of convergence at every generation step.



(a) Herringbone and French canvas (b) Herringbone and Cotton canvas

(c) Classified image of (a) (d) Classified image of (b)

Fig.11 Application to cloth texture patterns


(a) Detection of Ta area (b) Detection of TaUTc area (c) Detection of Tc area

Fig.10 Transformed image at the 3rd generation in sample 2.

the 1st generation the 2nd generation the 3rd generation

Fig.12 Another application