SUMMARY

An extension of the conventional kernel-based method (Parzen density estimation method), which is one of fundamental pattern recognition methods [1], is proposed. In the conventional method, a single kernel function $K(\cdot)$ is used commonly in the superposition to make discrimination functions, or posterior probability functions $P(C_i)$. Figure 1(a) shows an example of the superposition in 1-dimensional space ($S_i$ and $d$ represent a sample pattern and a kernel region size, respectively).

In the extended method, each edge length of every kernel region is optimized individually. A difference in region between the conventional and the extended kernel-based methods in 2-dimensional space is shown in Fig. 1(b) ($d_{jk}$ in Fig. 1(b)-2 represents the $k$th edge length of the $j$th kernel region).

If there are $L$ candidates for each edge length, the number of candidates for the $n$-dimensional kernel region is $L^n$. In real applications $n$ is more than 50 and $L$ should be about 10 or more. Thus, we have to solve the combinatorial optimization problem for $10^{50}$ or more candidates and this cannot be done by using a brute-force search. We therefore apply a genetic algorithm to the combinatorial optimization problem.

In our GA operation, a chromosome represents the kernel region in an $n$-dimensional pattern space, and each locus corresponds to one of the candidates (genes) for an edge length of the kernel region. We have applied the extended method to a sonar spectrum recognition problem (http://www.boltz.cs.cmu.edu) and obtained a recognition accuracy of 83.9%, which is much higher than the 62.0% obtained using the conventional kernel-based method. We have also compared the recognition accuracy of the extended K-B method with those of the nearest neighbor method and a neural network (backpropagation algorithm). The results are listed in Table 1.

We have analyzed the individually optimized kernel regions and shown that in the GA process the features in

![Figure 1: (a) Estimation of posterior probability by using kernel functions in the K-B method. (b) Kernel regions in the conventional K-B method and the extended K-B method.](image)

Table 1: Recognition accuracies in the sonar spectrum recognition problem.

<table>
<thead>
<tr>
<th></th>
<th>Extended K-B</th>
<th>Nearest neighbor</th>
<th>Neural network</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample patterns</td>
<td>100.0%</td>
<td>100.0%</td>
<td>99.7%</td>
</tr>
<tr>
<td>test patterns</td>
<td>83.9%</td>
<td>82.7%</td>
<td>83.0%</td>
</tr>
</tbody>
</table>

References