sometimes, though the prime implicants necessary for finding the minimal solutions are always obtained. Further, this method, like that of Slagle et al., is directly applicable for finding the minimal forms of the function. An algebraic approach based primarily on successive expansion to generate all the prime implicants of a Boolean function utilizing the maxterm-type expression was first proposed by Nelson [2]. This basic idea of Nelson was subsequently utilized by Das and Choudhury [3] in developing a tabular method for a more efficient generation of all the prime implicants of a Boolean function starting from the maxterm-type expression represented in decimal mode. The semantic tree approach of Slagle et al. is almost similar to the method of Das and Choudhury, except that, in the method of Das and Choudhury, the expansion, unlike that by Slagle et al., is carried out successively about all the variables starting from the highest weighted one in different levels. The authors also extended their tabular method for generating prime implicants of functions having many unspecified fundamental products, utilizing a very novel idea suggested in a paper by McCluskey [4]. In developing their algorithm the authors of the aforesaid paper failed to mention these existing and closely related techniques. The idea of the present communication is thus to draw the readers' attention to the existence of these related papers.

REFERENCES

Comments on "An Algorithm for Finding Intrinsic Dimensionality of Data"

G. V. TRUNK

In the above paper,1 Fukunaga and Olsen present an alternative method of estimating the intrinsic dimensionality of data. Their proposed algorithm differs from others in that it relies heavily on operator interaction and provides a method of specifying variable local regions. The authors state: "This variability is critical as the practical problem of determining dimensionality depends on the size and number of samples in the local regions." This is illustrated in their summary Table II (B), in which, for local region sizes containing five and ten samples, the indicated dimensionalities are one and three, respectively, when using the 1 percent eigenvalue criterion; and one and two, respectively, when using the 10 percent criterion. While the authors may have a decision rule to select the correct answer from the summary table, I did not see it in their paper; and without such a rule, I do not believe the problem has been solved satisfactorily.

While the size of the local region is critical for Fukunaga and Olsen, it is not nearly as important for the statistical method [1]. In order to demonstrate this, consider the three Gaussian examples presented by Fukunaga and Olsen. One hundred cases of each example, an example consisting of 50 20-dimensional vectors, were analyzed by the statistical method; the results are shown in Table I—out of 300 cases, only two incorrect answers were obtained. For all these cases, the local region for each point was defined by its five nearest neighbors. The statistical method is also very fast: the running times for examples 1, 2, and 3 were 2.7, 2.9, and 3.1 s, respectively, on a CDC 3800 computer.2

Table I

<table>
<thead>
<tr>
<th>Gaussian Examples</th>
<th>Estimated Dimensionalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>in Fukunaga and Olsen</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

The authors state that previous investigators had not considered the noise problem and then attack the problem by using a large number of samples. They estimate very accurately the eigenvalues and note the small difference in the eigenvalues due to the parameters and those due to the noise. However, in 1968, a "filtering" method [2], which does not require a large number of samples, was proposed as a solution to the noise problem. This method defines a pseudo signal-to-noise ratio R:

\[ R = \frac{D}{\sigma(K-N)^{1/2}} \]

where K is the dimensionality of the vector space, N is the intrinsic dimensionality of the data, \( \sigma \) is the standard deviation of the noise on each basis vector, and D is the average distance from a point to its \((N+1)\)-nearest neighbor. It was shown that when \( R > 12 \), the noise does not affect the estimation of dimensionality. When \( R < 12 \), the data can be filtered in the following manner. The original M points are randomly divided into L subgroups, each subgroup containing \( M/L \) points. This filtering has increased the signal-to-noise ratio in each subgroup since the average distance D has been increased. The algorithm for finding the optimal number of subgroups L and the manner of combining the results of the various subgroups is presented in [2]. Whether or not the filtering method can be used in conjunction with Fukunaga and Olsen’s method is not known, since the latter requires a minimum number of points in the local region in order to estimate the covariance matrix.

REFERENCES

Authors’ Reply

K. FUKUNAGA AND D. R. OLSEN

In the statistical method proposed by Trunk [2], [3], he calculates the density function \( p(X \mid N) \) where \( X = [x_1, \ldots, x_n]^\top \) is the observed random vector and \( N = (1, 2, \ldots) \) is the intrinsic number of parameters for the random vector. The most likely \( N \) for the observed vectors is determined by applying the hypothesis test to these density functions. The number of random variables are reduced by using sufficient statistics such as the ratio of local distances and certain angles between local vectors, rather than the original \( x_1, \ldots, x_n \).

As far as the estimation of an intrinsic dimensionality is concerned, the statistical method is in its essence more a more accurate but more complex estimate than the local eigenvalue method of [1] where only second moments are considered. Although the statistical method requires the selection of some control parameters, these are readily set. Several assumptions were required in the derivation of the density

---

1 Manuscript received May 4, 1971.
2 Manuscript received May 25, 1971.
3 The authors are with the School of Electrical Engineering, Purdue University, Lafayette, Ind. 47907.
functions. However, the results presented in [2], [3] indicate that these assumptions are generally valid.

On the other hand, we have always wondered if a single number estimate of the intrinsic dimensionality has any practical value unless the method is somehow related to finding a functional form for the intrinsic parameters of a given data set. Also, it has been our experience that the intrinsic dimensionality of a given data set is sometimes very subjective, and that the dimensionality can vary from region to region.

Therefore, we attempted to present information which would give the operator a better understanding of the data structure. This is because our primary interest is to extend the dimensionality discussion to topics such as data reduction in communication and feature extraction in pattern classification. For example, let us consider the case where we send the data of Table II(A) of [1] with a specified accuracy ($D_1$ or 1 percent). From Table II of [1] we can obtain the design data for this communication system. If we decide to send the coefficients of the fixed basis vectors and reconstruct the original vector at the receiving end with the preset basis vectors, we need to send six coefficients. Using the local regions with ten samples, only three coefficients are needed along with an identification code of the local region. A different set of three basis vectors is provided for each local region and is chosen by the identification code.

The filtering technique proposed in [3] enlarges the size of local regions until the ratio of noise level to the region size is lowered to a certain value. Although it is not exactly the same, a similar procedure can be found in the local eigenvalue method as follows. Adjusting the size of local regions (for example, 150 samples and 300 samples in Table IV of [1]), we can observe the change of eigenvalues due to the signal while eigenvalues due to the noise remain relatively constant. Thus an operator may see the effect of the noise.

Finally, it should be pointed out that the number of samples needed to calculate the dominant eigenvalues in each local region does not depend on the dimensionality of the original space [4]. Even with added noise, the number of samples required is determined by the number of dominant eigenvalues for the signal with a certain constant multiplied.

The References:


Uniform Modular Realizations and Linear Machines

S. EROL GELENBE and N. ROSSI

Abstract—It is shown that a single-output Moore-type n-state linear machine may be realized with no more than 2^n copies of the AND-OR-delay (AOD) module of Newborn, Weiner, and Hopcroft. This bound is significantly lower than that for arbitrary single-output Moore-type machines, which is 2^n.

Index Terms—AND-OR-delay module, linear machines, uniform modular realization.

Recently there has been considerable interest in the realization of Moore-type [1] sequential machines by means of a finite number of copies of a single module, as shown by Newborn [2] and Weiner and Hopcroft [3]. In particular, it has been shown that any k binary input-signal binary output n-state Moore-type synchronous sequential machine can be realized using no more than 2^k copies of the AOD module [2] shown in Fig. 1. Since the main shortcoming of this method of realization is the exponential number of modules necessary, efforts have been made to design a machine with which this bound becomes linear in n [4], the price paid for this reduction being a more complicated module.

In this correspondence we point to a class of machines for which the bound is 2n when realized with the simple AOD module. In particular, we prove the following.

Theorem: Let M be a 2^n-state k binary input-single binary output Moore-type linear sequential machine over GF(2). Then M is realized (as in [2]) with no more than 2^{n+1} copies of the AOD module.

We shall only present a proof for the case in which M is a single binary input machine so that the module used is the AOD1, the generalization to the k-input case with the AODs module being straightforward.

First let us note the following basic properties of uniform modular realizations of an arbitrary single binary input-single binary output Moore sequential machine N using AODs [2], [3].

Property 1: To each module in the uniform modular realization of N there corresponds a unique subset of the set of states of N.

Property 2: Let $\mu_1, \mu_2, \mu_3$ be any (not necessarily distinct) modules in the uniform modular realization of N so that $g_1$ and $h_1$ are the intermodule input terminals to $\mu_1$, while $g_2$ and $z_2$ are the output terminals of $\mu_2$ and $\mu_3$, respectively. Let $S_1$, $S_2$, $S_3$ be the subsets of S (the set of states of N) corresponding to $\mu_1, \mu_2, \mu_3$, respectively. Since $\mu_1, \mu_2, \mu_3$ are not necessarily distinct modules, $S_1$, $S_2$, $S_3$ will not necessarily be different subsets of S. Then $z_2$ is connected to $g_1$ if and only if $S_1$ is the set of all states that "go" into $S_1$ under input $x=0$. Similarly, $S_3$ is the set of all states that go into $S_3$ with input $x=1$ if and only if $z_2$ is connected to $h_1$.

Property 3: The output of the uniform modular realization is taken from the output of one module $\mu_3$ which corresponds to the subset $S_3$ of S with the property that N is in a member of $S_3$, if and only if its output is 1 (and is in a member of $S-\overline{S_3}$ if and only if its output is 0).

It is clear that the procedure for the uniform modular realization follows directly from Properties 1, 2, and 3.

Proof of Theorem: The machine M (in the theorem statement) is fully described by the equations

$$s(t+1) = A_s(t) \oplus B x(t)$$

$$y(t) = C_s(t)$$

Fig. 1. The AODs ($k$-input AND-OR-delay) module.