A New Approach to the Objective Evaluation of Fabric Handle from Mechanical Properties

Part III: Fuzzy Cluster Analysis for Fabric Handle Sorting

N. PAN. K. C. YEN. S. J. ZHAO. AND S. R. YANG

Department oJ~Te_rtile Engineering, China Tertile University. Shanghai. Peoples Republic oJ~China

ABSTRACT

Fabric handle evaluation can be reduced to a typical clustering problem, and fuzzy cluster analysis is applied to a practical example in this paper. For comparison, the same example is also tackled using hierarchical cluster methods of multivariate analysis. Both kinds of results have been found to be consistent. To determine the optimum group number g of clustering, Marriot's *gz* I WJ criterion is suggested and the result is demonstrated as satisfactory; consequently it will be possible to use discriminant analysis to evaluate fabric handle.

There are many situations in scientific and business investigations in which a numerical classification method called cluster analysis is applicable. I'apers dealing with cluster analysis have occurred in a wide variety of journals treating diverse topics.

The cluster problem can be described as follows: Let the set I = (I, , Iz, - - - I„,) denote m individuals from a conceptual population T,. It is tacitly assumed that there exists a set of features or characteristics C = (Ci, Cz, - - - C„) that are observable or, more specifically,

0040-5175/88/58010-565$2.00

measurable and are possessed by each individual in the set I. The quantitative data of these features are at times called measurements or parameters of the individuals. We denote the value of the parameter on the ith feature of the individual I; by the symbol X;; and let X; _ [Xt;] denote the n X 1 vector of such parameters. Hence for a set of individuals I, there is available to the investigator a corresponding set of n X 1 parameter vectors X = {X,, X2, • • • X m} describing the set 1. Note that the set X can be thought of as m points in n-dimensional Euclidean space E„.

Let g be an integer less than n. Based on the data contained in the set X, the cluster problem is to determine g clusters (subsets) of individuals in 1, say G1, G2, • • • Gg, such that I; belongs to one and only one subset. Those individuals assigned to the same cluster are similar in the properties specified by n features, yet individuals from different clusters are different or not similar.

For a set of fabric samples, sorting or grading the fabrics by their differences and the ranges of fabric handle is an interesting problem with significant practical meanings. This is the main problem of assessing fabric handle, but most papers on this issue have approached it subjectively. We must point out that a sorting problem should not be processed arbitrarily. Only when the mean values of each group are demonstrated to be significantly different is the sorting result meaningful.

The procedure of fabric handle assessment or sorting can be reduced to a typical cluster problem. Suppose m fabric samples are to be assessed. Each of them is represented by n-dimensional variables (say the mechanical properties measured on Kawabata's KES-F instruments or other similar parameters). The samples are to be divided into different groups in accordance with fabric handle performance.

There are several categories of cluster analysis. One, called fuzzy cluster analysis, is a new development and in fact is a combination of fuzzy set theory [7J and conventional cluster analysis. The practical cluster problems are more or less accompanied by "fuzziness," so cluster analysis based on the fuzzy set theory seems more rational and realistic. In this paper, we have adopted fuzzy cluster analysis mainly to deal with the problem of fabric handle classification. We have also used another category, hierarchical cluster analysis, for comparison.

Theory of Fuzzy Cluster Analysis

Mathematically, a definite clustering is determined by an equivalence relation. So a fuzzy clustering is correspondingly determined by a fuzzy equivalence relation.

For the sake of saving space, we introduce the definitions and theorems concerned only briefly. For more details on fuzzy cluster analysis, the reader can refer to reference 3.

Definition 1, *the composition* ojfuzzy relations: Let R and S be two fuzzy relations. Define the composition Z of them as

$$Z = R \circ S, \tag{1}$$

and the membership function of Z as

$$taAi, k) = \text{sup}[\min(,u,(i,J),.u,(j, k)J$$

$$_ ; [R;; n \text{ SjJ} \tag{2}$$

Definition 2, *a-level sets* of a fuzzy set: Suppose R is a fuzzy relation on the conventional (nonfuzzy) set X, then one a-level set of X is also a conventional set Ra:

$$Ra = \{u: u \in X, A,(u) z a\} \tag{3}$$

the family of a-level sets (Ra: a E [0, IJ) is a monotonically decreasing series, *i.e.,*

a, s a2 w Rai D RaZ, a,, a2 E [0, IJ ; (4) and it is easy to understand that the 0-level set Ro is just the original set X.

Definition 3, fuzzy *equivalence relations:* Let R be a fuzzy relation on X; if R satisfies the conditions

(1) reflexive, i.e., A,(i, i) = I, Vi E X , (5)

(6)

(3) transitive, i.e.,;"(;,,(i, j) n *u,(j, k)J*

< , (7)

then *R* is called a fuzzy equivalence relation.

.*Theorem I:* If *R* is a fuzzy relation on the infinite set X, and R satisfies the conditions of *reflexive and* symmetric so that the membership function *u,* can be described in a matrix form, then there must exist k s n to make

$$\underbrace{R \circ R \circ \cdots \circ R}_{k} = RA \tag{8}$$

a fuzzy equivalence relation.

Theorem 1: The necessary and sufficient condition to make fuzzy relation R a fuzzy equivalence relation is that all a-level sets on X, Ra, *a* E [0, 1] are equivalence relations.

Tltourmnr 3: For an equivalence rclatiun Ra. define

$$Ra[i] _ \{ j: j \in X, iRaj \} \tag{9}$$

as the equivalence sorts of R_a determined by i. There must be

$$\bigcup_{i\in X} R_a[i] = X , \qquad (10.1)$$

and

$$R_a[i] \cap R_a[j] \neq \emptyset \Rightarrow R_a[i] = R_a[j] , \qquad (10.2)$$

where U = union operation, fl = intersection, and 0 = empty set. That is, if R_a is an equivalence relation on X, then X is the direct union of all equivalence sorts of R_a, and these equivalence sorts perform the only direct union deposition of X. In other words, X is divided by the equivalence sorts, which are disjoint, and each member of X certainly and only belongs to one of these equivalence sorts of R_a.

Definition 4: Define the partition of X based on the equivalence relation R_a as T_a. If $a_1 > a_2$, then T_{a_2} *is* the finer partition of T_{a_1}.

All these theorems provide the theoretical bases for the reasonableness of fuzzy clustering and the reliability of the results obtained.

Calculation and Example

SAMPLES AND THE ORIGINAL DATA MATRIX

The fabric samples used are all medium thickness, and the details are shown in Table I as the sample set I = {Ik, k = 1, 2, • • • 20}. Note that all samples are from five fabric types, each with four members, so the sorting result is known. The reason for such an arrangement is to make it easier to check the final clustering result. The mechanical property parameter can be measured by, say Kawabata's KES-1713 instruments. As an alternative, however, we suggest a new measurement proposal [SJ, which in light of the pattern recognition method, is used to obtain the mechanical parameters composing the data vector or matrix $X(m, n, m = 20, n = 10)$ of the samples.

BUILDING THE MEMBERSHIP FUNCTION

Let the relation R = "*with* similar fabric handle." Apparently R as a fuzzy relation for "similar fabric handle" is a somewhat indefinite description. The membership function of R consists of the similarity measures. A nonnegative real valued function $S(X_i, X_j) = S_{ij}$ is said to be a similarity measure if

$$\text{(a)}\ 0 \le S(X_i, X_j) \le 1 \text{ for } X_i, X_j , \qquad (11.1)$$

$$\text{(b)}\ S(X_i, X_i) = 1 , \qquad (11.2)$$

$$\text{(c)}\ S(X_i, X_j) = S(X_j, X_i) \qquad (11.3)$$

This means that the fuzzy relation R obeys *reflexive* and summetric conditions, that is, the membership function of R is a matrix (see theorem 1). This matrix is called the similarity matrix, and the quantity S_{ij} is called simply a similarity coefficient.

So, on the fabric samples set 1, the fuzzy relation R means possessing similar fabric handle, and the membership function of R describes the extent of possessing similar fabric handle of the individual fabric samples in the set I.

After normalizing the original parameters of the samples, we calculated the similarity coefficients, which are defined here as [8]

$$S_{ij} = 1 - D_{ij}, \quad S_{ij} \in [0, 1] , \qquad (12)$$

$$D_{ij} = \sum_{k=1}^{n} (X_{ik} - X_{jk}) , \qquad (13)$$

where i, j = 1, 2, • • • m and X =the mechanical parameters. The pairwise similarities may be arranged in the similarity matrix

$$S = (S_{ij}) \qquad (14)$$

EQUIVALIZATION OF THE FUZZY RELATION R

In general, fuzzy relations like R only obey the *reflexive* and *symmetric* conditions, excluding the *transitive* one, i.e., they are not fuzzy equivalence relations. According to theorem 1 of the last section, by means of the fuzzy composition operation in definition 1, we can make the fuzzy relation R an equivalence relation and get the corresponding similarity matrix S.

TABLE 1. Details of fabric samples.

Weave	Fiber content	Yarn count, Nm	Threads/ 10 cm W F	Weight, g/m2	Sample numbers	Sample code
Fancy suiting	wool/PET 92/8	70/2	513 362	281	4	5-8
Fancy suiting	wool/PET 93/7	60/2	503 354	325	4	
Gabardine	wool 100%	52/2	451 244	305	4	17-?0
						13-16
Fancy suiting	wool 100%	38/2	224 180	229	4	
Fancy suiting	wool/PET 45/55	48/2	235 213	198	4	1-4
						9-I?

CLUSTERING

The clustering is formed by means of the α-level sets of the fuzzy equivalence relation R, ranking the similarity coefficients S_{ij} of the final similarity matrix in terms of their values and then taking those unequal coefficients as the values of an α sequence, as shown below.

Group number 19 18 17 16 15 14 13
α .9997 .9995 .9992 .9988 .9987 .9981 .9970

Group number 12 11 10 9 8 7 6
α .9967 .9955 .9948 .9945 .9936 .9928 .9729

Group number 5 4 3 2 1
α .9631 .9575 .9077 .7414 .6124

At each stage of clustering, the value of α decreases gradually from 1 to 0, and we have merged the members with the similarity coefficients not less than the value of α to get the corresponding α-level sets:

$$R_\alpha = \{u : u \in X, A_{ij} = S_{ij} \geq \alpha\} ,$$

which provides the new group at each step of the clustering. In fact, all members are gradually merged in accordance with their similarities in fabric handle. The whole procedure can be drawn into a clustering tree, which may be viewed as a diagrammatic representation of the results of the clustering process (see Figure 1). Theorems 2 and 3 ensure the clustering results are reasonable and unique.

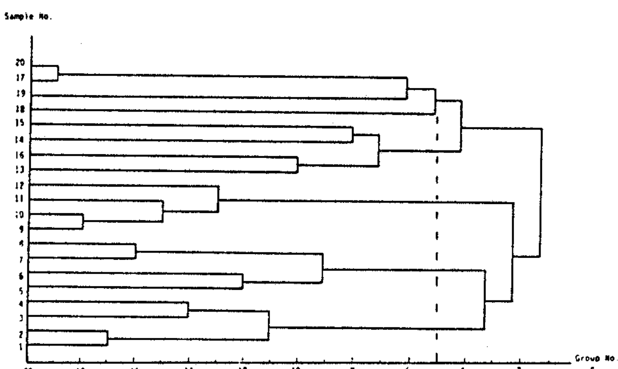


FIGURE 1. Fuzzy clustering tree.

The basic procedures of hierarchical clustering are as follows: (a) Begin with m groups, each group consisting of one member. (b) Define the distances between members and between groups respectively. (c) At each stage, reduce the number of groups by one through a merger of those two groups that are closest together, i.e., with the nearest distance between them, to form a new group. (d) Calculate the distances between the new group and other groups. (e) Repeat steps (c) and (d) until all members are merged into one group.

In defining the distance between members, the Euclidean distance is the most popular for cluster analysis, so we will use it here as the distance between members. Its definition was shown in Equation 13.

In defining the distance between groups, there are various definitions and therefore various algorithms. We have adopted four of them: Nearest Neighbor, Centroid, Group Average, and Ward's Method.

Wishart [6] proposed a general form of the hierarchical algorithm to calculate the distances between the new group G_k and other groups G_h after each merging of G_i into G_k with G_j:

$$D_{hk}^2 = a_i D_{hi}^2 + a_j D_{hj}^2 + b D_{ij}^2 + r \, | D_{hi}^2 - D_{hj}^2 | , \qquad (15)$$

$$G_k = G_i \cup G_j \quad k \neq i,j , \qquad (16)$$

where a_i, a_j, b, and r are determined by the particular definitions of group distances as shown in Table II. Using Equation 15, all four methods can be incorporated into the same computer program by varying parameters a, b, and r.

TABLE II. Values of the parameters in Equation 15.[a]

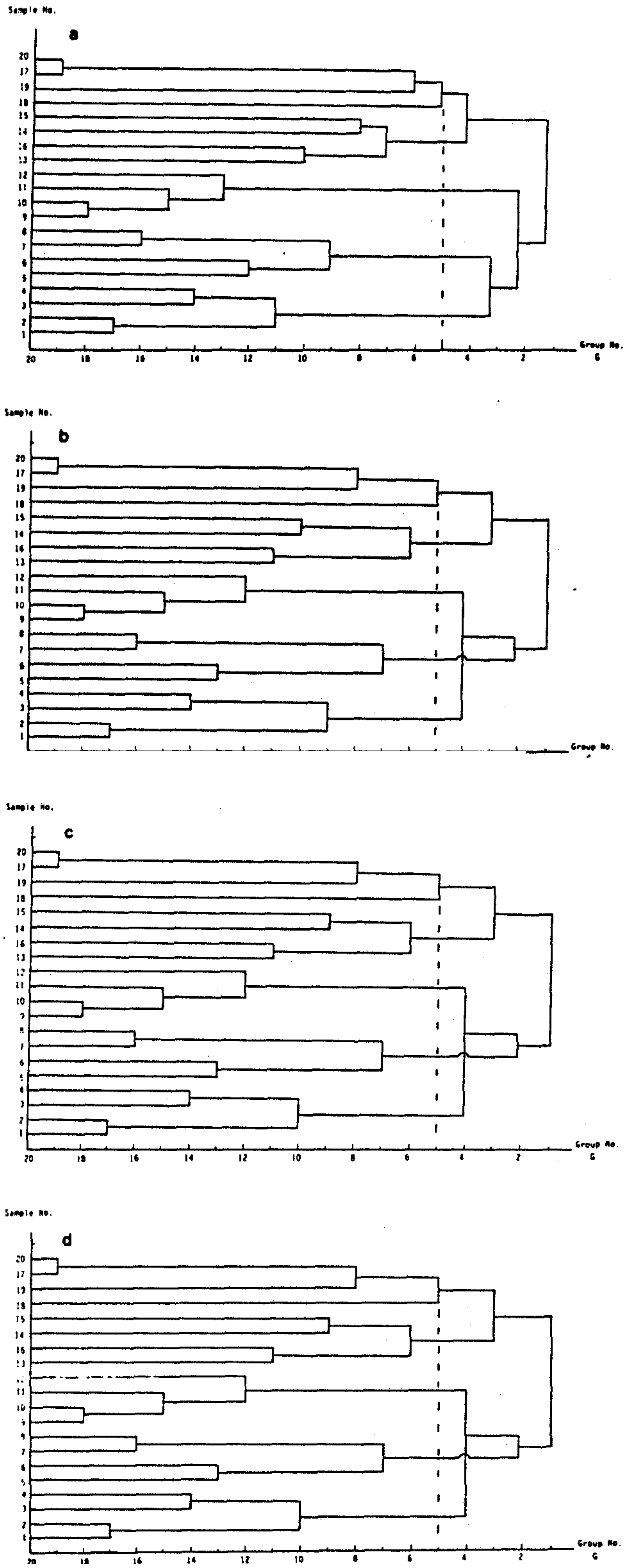
Method	Values
Nearest Neighbor	$a_i = a_j = 1/2$, $b = 0$, $r = -1/2$
Centroid	$a_i = e_i/e_k$, $a_j = e_j/e_k$, $b = -a_i a_j$, $r = 0$
Group Average	$a_i = e_i/e_k$, $a_j = e_j/e_k$, $b = r = 0$
Ward's Method	$a_i = (e_h + e_i)/(e_h + e_k)$, $a_j = (e_h + e_j)/(e_h + e_k)$, $b = -e_h/(e_h + e_k)$, $r = 0$

[a] Where e_i, e_j, e_k, and e_h are the member numbers in groups G_i, G_j, G_k, and G_h, respectively.

Hierarchical Cluster Analysis

Hierarchical cluster analysis is based on the theories of multivariate statistical analysis. For the sake of comparison, the example above is also treated by hierarchical cluster analysis; those who want more details of this cluster analysis can refer to reference 1.

Using the computer program of hierarchical cluster analysis and the parameters of the sample set, we have obtained the four results corresponding to these four methods, and we show them in Figure 2. The merging distances in each step of these methods are listed in Table III.



FiGUtte 2. Clustering trees of four methods: (a) Nearest Neighbor (NN) method, (b) Centroid (CE) method. (c) Group Average (GA) method, and (d) Ward's method (WM).

Determining the Optimal Number of Groups

The clustering trees of the various approaches only give a configuration for every number of clusters from one (the entire data set) up to the number of entities (each cluster has only one member) and have not provided the definite sorting result for the practical sorting problems.

In cluster analysis, however, determining the optimal group number is considered to be a formidable problem [2]. Everitt [2] recommended a less subjective but still essentially informal approach to the problem proposed by Marriot [4]. The latter suggested that a possible criterion for the assessing group number g is to take the value of g that minimizes g' J W1, where I W1 is the value of the determinant of matrix W containing the sum of squares and products within a group.

By this criterion, the complete sorting procedure is as follows: (a) cluster the members and draw the cluster tree, (b) for every possible choice of g, calculate the respective value of gZJ W1, (c) choose the value of g with minimum gzJ W1 value as the optimum number of groups, (d) from the clustering tree, g groups that are reasonably sorted can be obtained. To start with, suppose the possible values of g are

g=2> g=3, g=4, g=5,

g=6, g=7, g=8

Based on the clustering tree of fuzzy analysis, the corresponding results of gZJ W1 are listed in Table IV. Clearly g = 5 is the answer.

With the condition g = 5, we can derive the final clustering results from the clustering trees above. All results from either the fuzzy clustering or hierarchical trees are identical:

G, _ (1, 2, 3, 4), G, _ (5, 6, 7, 8>,

G3 = (9, 10, 11, 12), G4 = (13, 14, 15, 16),

G5 = (17, 18, 19, 20) ,

where G; indicates the ith group and the numbers in parentheses are the codes of the samples sorted in this group.

Because we cluster these fabric samples based on the mechanical parameters that determine fabric handle, this result is the same as that for fabric handle sorting. Five groups just correspond to five kinds of fabrics; that is, within each group, all members are from the same fabric type and should certainly have the same handle.

TABLE III. The merging distances of the four methods.

Method'	Group number						
	19	18	17	16	15	14	13
NN	.083	.138	.227	.331	.337	.544	.828
CE	.083	.138	.227	.33 I	.397	.544	.924
GA	.083	.138	.227	.331	.403	.544	.924
WM	.083	.138	.227	.331	.459	.544	.924
	12	11	!0	9	8	7	
NN	.924	1,257	1.471	1.542	1.807	1.816	
CE	1.091	1.471	1.637	1.807	2.029 2.103 '		
GA	1.108	1.471	1.664	1.807	2.029	2.160	
WM	1.336	1.471	1.807 2.316		2.343	2.975	
	6	5	4	3	2	I	
NN	2.025	7.599	10.342	11.913	25.901	61.320	
CE	3.441	8.921	13.812	19.464	35.231	77.438	
GA	3.633	8.972	13.863	19.952	35.921	78.660	
WM	4.866	10.933	27.632	38.917	81.349	187.208	

' NN = Nearest Neighbor, CE = Centroid, GA = Group Average, and WM = Ward's Method.

TABLE IV. Values ofg'l W1.

g	2	3	4	5	6	7	8
911 W1	5.84 X 10"	6.01 X 10'	4.34 X 10'	2111.57	5.59 X 108	3.14 X 10'2	4.32 X 10"

Discussion

In cluster analysis, little or no knowledge about the category structure of the sample set is required. All that is needed is a collection of measured parameters. The essence of cluster analysis might be viewed as sorting the samples into groups such that the degree of "property association" is high between members of the same group and low between members of different groups. Clearly it can be a useful approach to objective evaluation of fabric handle.

The final result of cluster analysis provides the category structure of the sample set, which is indispensable for discriminant analysis. The whole clustering procedure is in fact completed in an u-dimensional space (n is the number of parameters), this means that the difference of every parameter between samples is considered during clustering, so the results are much more realistic. In addition, the samples to be clustered compose a complete set; the results therefore are definite and the statistical estimate is not necessary. This is the unique advantage of cluster analysis compared with other multivariate statistical approaches.

As a rule, to determine the optimum group number in cluster analysis. Marriot's g'I W1 criterion, which we have used for several problems and found satisfactory, can be taken as a final rule for fabric handle sorting problems.

To compare these cluster methods, the theoretical bases of fuzzy cluster are mostly complete so that the result is more reliable. Among several kinds of hierarchical clusters, the Centroid is considered inferior to the others, for its merging distances change nonmonotonically [8J. Compared with Ward's Method, the Nearest Neighbor and Group Average methods are less sensitive but the results are more reliable [8J. These characteristics, however, did not have a significant influence in this steely because the category structure of the sample set used is relatively diverse.

Figures 1 and 2 show that all five clustering trees can be divided into three groups. The fuzzy cluster and the Nearest Neighbor compose the first group, there is only the Centroid method in the second group, and other two belong to the third group. Within the same group, the procedures of clustering and the clustering trees are the same. In fact, we can prove that the fuzzy cluster is equivalent to the Nearest Neighbor method.

In conclusion, cluster analysis combined with the 911 tVI criterion can be used in the objective sorting of fabric handle terms so as to eliminate the problems existing in the handle assessment process using the sensory perception approach.

Clearly, as a useful means of classification, this approach is also applicable to other similar practical problems of product performance evaluation. Meanwhile, we must admit there are limitations in the cluster analysis approach. The first problem is that the final result is only a diagrammatic one or something similar rather than a quantitative value. The second is a lack of comparability between the clustering results obtained in different situations. The last problem is the difficulty in dealing with a large number of samples, which will mean a great amount of calculation. Cluster analysis can thus be considered an additional approach for objective fabric handle assessment.

Literature Cited

1. Duran, B. S., and Odell, P. L., "Cluster Analysis - A Survey," Springer-Verlag, Berlin, 1974.
2. Everitt, B. S., Unresolved Problems in Cluster Analysis, *Biometrics 35*, 169 (1979).
3. Fu, K. S., and Whinston, A. B.. "Pattern Recognition Theory and Application," Sijthoff and Noordhoff International, The Netherlands, 1977, p. 453.
4. Marriot, F. H. C., Practical Problems in a Method of Cluster Analysis, *Biometrics 27, SO l (197* I).
5. Pan, N., and Yen, K. C., A New Measurement Proposal for Fabric Handle Evaluation, J. *China Textile Univ. Eng. Ed.*, no. 2, l (1986).
6. Wishart, W., An Algorithm for Hierarchical Classifica tion, *Biometrics 25*, 165 (1969).
7. Zadeh, L. A., Fuzzy Sets, *Inf. Control 8*, 338 (1965).
8. Zhang, Y. Z., and Fang, K. T.. "Introduction to Multivariate Statistical Analysis," Science Press, Beijing, 1982, pp. 393, 401.

Manuscript received October 15, 1986; accepted July 21, 1987.