Characterization of Korean Clays and Pottery by Neutron Activation Analysis(II). Characterization of Korean Potsherds

Chul Lee*, **Oh Cheun Kwun, and Seungwon Kim**

Department of Chemistry, Hanyang University, Seoul 133

Ihn Chong Lee

Department of Chemistry, Hallym College, Chunchun 200

Nak Bae Kim

Korea Institute of Energy and Resources, Seoul 150. *Received May 24, 1986*

Fisher's discriminant method has been applied to the problem of the classification of Korean potsherds, using their elemental composition as analyzed by neutron activation analysis. A combination of analytical data by means of statistical linear discriminant analysis has resulted in removal of redundant variables, optimal linear combination of meaningful variables and formulation of classification rules.

Introduction

The combination of multielement analysis and mathematical methods for the analysis of analytical data have recently been used to extract the useful information from a date set.[1-2] Pattern recognition(PR) approach involves the examination using quantitative and qualitative information available on the samples to find or predict a property of the objects, which is not directly measurable but is known to be related to the measurements via some unknown relationship. Archaeology is one of the major beneficiaries of such a PR approach.[1,5]

This paper reports on the work done in the development and application of Fisher's discriminant analysis (FDA) for grouping Korean ancient potsherds.

Fisher's Discriminant Analysis

The set of data on N samples with M variables, known to belong to two or more groups, can be represented as a set of N points in M dimensional space. The aim of the discriminant function analysis is to arrive at a set of rules, which will classify samples into one of these groups with a minimal error. Often the discriminant function derived is a linear combination of the original variables and of the form,

$$f(x_{i1}, x_{i2}, \cdots x_{ij} \cdots x_{iM}) = k_1 x_{i1} + k_2 x_{i2} + \cdots + k_j x_{ij} + \cdots + k_M x_{iM} \quad (1)$$

In vector notation, equation(1) can be written as $Y_i = K \cdot X_i$, where K is the set of coefficients, X_i is the sample vector of individual i and Y_i is called the discriminant score (DS_i)

Preprocessing usually involves the transformation of the original variables into a new set of variables. Autoscaling, *i.e.*, the transformation that produces a new set of variables with zero mean and unit standard deviation, is the often used technique for preprocessing of nonspectral data, so that all the variables are equally weighted. The original variables X_{ij} can thus be transformed into standardized new variables Z_{ij} as

$$Z_{ij} = \frac{X_{ij} - X_j}{\sigma_j} \quad (2)$$

when X_{ij} represents the sample vector and X_j and σ_j indicates overall mean and standard deviation of jth variable, respectively.

The discriminant function is a linear combination of new variables and is now given in the form,

$$f(z_{i1}, z_{i2} \cdots z_{ij} \cdots z_{iM}) = v_1 z_{i1} + v_2 z_{i2} \cdots + v_j z_{ij} \cdots + v_M z_{iM} \quad (3)$$

In vector notation equation(3) can be written as $Y_i = V \cdot Z_i$, where V is the set of new coefficients and Z_i is the new sample vector of individual i. Linear discriminants of the same form as equation(3) are vector in a M dimensional space. The discriminant scores Y_i give the projections of these M space data onto these vectors. The projections should be such that there is minimum overlap of the projected points of members of different groups. A suitable choice of the vectors is required to obtain an optimum set of coefficients V. Fisher suggested that a linear function that provides the maximum value for the ratio of the between-group dispersion to the within-group dispersion can be used to obtain the optimum coefficient values V.[6]

If there are K groups, each containing N_p(p=1,2 ... K) samples measured over M variables, the total number of samples is represented as $N = \sum_{p=1}^{K} N_p$. The total variance S for the new variables is given by,[7]

$$S = \sum_{p=1}^{K} \sum_{i=1}^{N_p} \sum_{j=1}^{M} (z_{pij} - z_{..j})^2$$

$$= \sum_{p=1}^{K} \sum_{i=1}^{N_p} \sum_{j=1}^{M} \left(\frac{x_{pij} - x_{..j}}{\sigma_{..j}}\right)^2$$

$$= \sum_{p=1}^{K} \sum_{i=1}^{N_p} \sum_{j=1}^{M} \left(\frac{x_{pij} - x_{p.j} + x_{p.j} - x_{..j}}{\sigma_{..j}}\right)^2$$

$$= \sum_{p=1}^{K} \sum_{i=1}^{N_p} \sum_{j=1}^{M} \left(\frac{x_{pij} - x_{p.j}}{\sigma_{..j}}\right)^2 + \sum_{p=1}^{K} \sum_{i=1}^{N_p} \sum_{j=1}^{M} \left(\frac{x_{p.j} - x_{..j}}{\sigma_{..j}}\right)^2$$

$$+2\sum_{p=1}^{K}\sum_{i=1}^{N_p}\sum_{j=1}^{M}\left(\frac{x_{pij}-x_{p\cdot j}}{\sigma_{\cdot\cdot j}}\right)\left(\frac{x_{p\cdot j}-x_{\cdot\cdot j}}{\sigma_{\cdot\cdot j}}\right) \tag{4}$$

where x_{pij} is the data corresponding to jth variable of ith individual in pth group, $x_{p\cdot j}$ is mean value of variables x_j in group p and $x_{\cdot\cdot j}$ and $z_{\cdot\cdot j}$ are overall mean values of variables x_j and z_j, respectively.

The last term in equation(4) is zero and hence S is given by,

$$S=\sum_{p=1}^{K}\sum_{i=1}^{N_p}\sum_{j=1}^{M}\left(\frac{x_{pij}-x_{p\cdot j}}{\sigma_{\cdot\cdot j}}\right)^2+\sum_{p=1}^{K}\sum_{i=1}^{N_p}\sum_{j=1}^{M}\left(\frac{x_{p\cdot j}-x_{\cdot\cdot j}}{\sigma_{\cdot\cdot j}}\right)^2$$

$$=\sum_{p=1}^{K}W_p+\sum_{p=1}^{K}B_p \tag{5}$$

W_p is in the form as,

$$W_p=\begin{pmatrix} w_{p11} & w_{p12} & \cdots & w_{p1M} \\ \vdots & & & \\ w_{pM1} & w_{pM2} & \cdots & w_{pMM}\end{pmatrix}$$

where each element of the matrix is

$$W_{pjk}=\sum_{i=1}^{N_p}\sum_{j,k=1}^{M}\left(\frac{x_{pij}-x_{p\cdot j}}{\sigma_{\cdot\cdot j}}\right)\left(\frac{x_{pik}-x_{p\cdot k}}{\sigma_{\cdot\cdot k}}\right) \tag{6}$$

The total within-group sum of squares and cross products(SSCP) is then given by

$$W=\sum_{p=1}^{K}W_p \tag{7}$$

Analogously, B_p is of the form

$$B_p=\begin{pmatrix} b_{p11} & \cdots\cdots & b_{p1M} \\ \vdots & & \\ b_{pM1} & \cdots\cdots & b_{pMM}\end{pmatrix}$$

where each element of the matrix is

$$b_{pjk}=\sum_{j,k=1}^{M}N_p\left(\frac{x_{p\cdot j}-x_{\cdot\cdot j}}{\sigma_{\cdot\cdot j}}\right)\left(\frac{x_{p\cdot k}-x_{\cdot\cdot k}}{\sigma_{\cdot\cdot k}}\right) \tag{8}$$

The pooled between-group SSCP matrix is then given by

$$B=\sum_{p=1}^{K}B_p \tag{9}$$

The above results can be represented in the matrix notation as,

$$T=B+W \tag{10}$$

where T is the total dispersion matrix.

The projections of new variables Z_{ij} onto the vector described by coefficients V are given by $Y=V\cdot Z$. The between-group and the within-group spreads for the projected points are given by $B_Y=V'BV$ and $W_Y=V'WV$, where B and W are obtained from original variables X_{ij}. If the ratio of the between group dispersion to within-group dispersion of the projected points is defined as

$$L=(V'BV)/(V'WV) \tag{11}$$

L can be maximized according to Fisher's criterion as

$$\delta L/\delta V=0.\quad \text{Thus,}$$

$$(BV/V'WV)-(V'BV/V'WV)(WV/V'WV)=0 \tag{12}$$

Since L is defined as equation(11), equation(13) can be derived

$$\frac{BV}{V'WV}-L\left(\frac{WV}{V'WV}\right)=0 \text{ or}$$

$$BV \quad =LWV \tag{13}$$

Multiplying both sides by W^{-1}, equation(13) is transformed as

$$W^{-1}BV=LV \tag{14}$$

From this equation, it is recognized that the coefficients V are given by the eigen vector coefficients of matrix $W^{-1}B$ and L is the corresponding eigen value. Since L is defined as the ratio of the between-group dispersion to the within-group dispersion to obtain the maximum discrimination, the eigenvectors associated with the largest eigenvalue of the matrix $W^{-1}B$ should be used as the discriminant coefficients.

The number of eigenvalues extracted from the matrix $W^{-1}B$ will be equal to M or $K-1$ (K is the number of groups), whichever is less. Hence M or $K-1$ discriminant functions, each with a different discriminating power, can be obtained. However in a two class problem, it is not necessary to solve for the eigenvalues of $W^{-1}B$, as the vector given by BV is in the same direction as the vector D of the difference between two means.[8] If the vector D of the difference is defined as

$$D=\mu_1-\mu_2 \tag{15}$$

a new set of discriminant coefficients V_2 is obtained by

$$V_2=W^{-1}D \tag{16}$$

A computational example of how to determine V_2 in practice is given by Kendall.[9]

Selection of Variables

A pattern consisting of many parameters often contains a lot of noise, *i.e.*, redundant parameters. These redundant parameters tend to obscure the difference between classes and therefore render the separation more difficult. The unnecessary parameters should be eliminated. To trace redundant variables several criteria are available.[7]

Criteria based on discriminant functions which are of the form as equation(3) give a larger importance to a variable when the absolute value of the corresponding weight coefficient is higher, provided that the variables have been standardized. A direct method is to determine the contribution percentage of each variable to the total distance D^2 in the discriminant space, which is the distance between the centroids of the groups considered. The contribution percentage of variable j is given by $100\times|v_j\delta_j|/D^2$, where v_j is the weight coefficient of the discriminant function for jth variable and

$$\delta_j = \left(\frac{x_{p\cdot j} - x_{q\cdot j}}{\sigma_{\cdot\cdot j}}\right).$$

$x_{p\cdot j}$ and $x_{q\cdot j}$ are the mean values of jth variable in group p and q, respectively. Thus $D^2 = \sum_{j=1}^{M} |v_j \delta_j|$.

Description and Analysis of Samples

Samples of potsherds from different sites in Korea were collected through museums. In Table 1, the sites where the specimens were found, are given together with the corresponding symbols. The whole samples were grouped into three classes according to geographical similarity as shown in Table 1. The samples found in Koryung were further grouped into three subclasses according to sites, using the symbols in parenthesis as shown in Table 1.

The elemental analysis of potsherds was carried out by thermal neutron activation analysis. The detailed analytical procedures had been described elsewhere.[4]

Results and Discussion

Twenty elements(Na, K, Sc, Cr, Fe, Co, Cu, Ga, Rb, Cs, Ba, La, Ce, Sm, Eu, Tb, Lu, Hf, Ta and Th) which were analyzed by neutron activation analysis, have been used in the present PR study for the classification of potsherds collected from various sites as shown in Table 1. The means and standard deviations of the variables, i.e., elemental contents, have been calculated for each group and overall samples and are given in Table 2 and 3, respectively.

The total within-group SSCP, *i.e.*, $W(M\times M)$ matrix, and pooled between-group SSCP matrix, *i.e.*, $B(M\times M)$ matrix, of equation(10) have been generated from original data set and the data given in Table 2 and 3. Characteristic roots of $W^{-1}B$ matrix, *i.e.*, 20 eigen values of equation(14), have been found. 20 sets of eigen vector coefficients corresponding to each eigen value have been calculated subsequently and 20 discriminant functions which are of the forms of equation(3) were finally obtained.

The discriminant function corresponding to the largest eigen value was selected. The contribution percentage of each element to the total distance D^2 in the discriminant space has been estimated, using eigen vector coefficient v_j, *i.e.*, the weighting factor corresponding to each element in equation(3) along with fractional distance δ_j between classes.

Table 1. Sampling Sites and Their Corresponding Symbols for Potsherds

Symbols	Number of samples	Sites	
△	4	Kwangju	Kyungki-do
△	22	Suwon	Kyungki-do
△	8	Yoju	Kyungki-do
△	2	Amsa-dong	Seoul
△	3	Yoksam-dong	Seoul
▲(▽)	23	Koryung A	Kyungsangbuk-do
▲(▼)	20	Koryung B	Kyungsangbuk-do
▲(×)	6	Koryung C	Kyungsangbuk-do
○	14	Kimhae	Kyungsangnam-do
○	12		Pusan

Data for the selection of variables are given in Table 4. The contribution percentage of each element to the distance between groups A and C is not given in Table 4 because the two groups are not seriously overlapped as shown in Figure 1. From the results on Table 4, the classification between groups have been found to be mainly attributed to 11 elements such as Cu, K, La, Na, Ce, Th, Cr, Cs, Sc, Rb and Co.

Using the data of the selected 11 elements $W^{-1}B$ matrix (11×11) has been generated again. Since a three-fold classification problem is involved, two discriminant functions have been computed similarly as described above. The first discriminant function f_1 corresponds to the largest eigen value L_1 of 2.46 and the second discriminant function f_2 to the eigen value L_2 of 0.61. The discriminant scores for individual i along the optimal discriminant function axes f_1 and f_2 have been generated and are given as follows.

$$L_1 = 2.460$$

$$\begin{aligned} DS_{1,i} = {} & -0.565z_{Cu,i} + 0.065z_{K,i} + 0.073z_{La,i} - 0.201z_{Na,i} \\ & + 0.138z_{Ce,i} - 0.010z_{Th,i} + 0.710z_{Cr,i} - 0.019z_{Cs,i} \\ & - 0.274z_{Sc,i} + 0.079z_{Rb,i} + 0.161z_{Co,i} \end{aligned}$$

$$L_2 = 0.609$$

$$\begin{aligned} DS_{2,i} = {} & 0.531z_{Cu,i} + 0.215z_{K,i} + 0.311z_{La,i} - 0.422z_{Na,i} \\ & + 0.272z_{Ce,i} - 0.340z_{Th,i} + 0.188z_{Cr,i} + 0.356z_{Cs,i} \\ & - 0.097z_{Sc,i} - 0.190z_{Rb,i} - 0.004z_{Co,i} \end{aligned}$$

Figure 1 shows a map of the individuals of the three groups and the corresponding group centroids in the 2-dimensional discriminant space together with a territorial diagram of each group. It is constituted of two diagrams. The territorial diagram contains linear boundaries drawn orthogonally on half the distance between each pair of group centroids. In this way three regions can be observed, one for each geographical site.

The present classification results are given in Table 5, which shows the number of correctly classified individuals. Table 5 shows that Kyungki and Koryung shards form distinct groups with efficiencies of 32/39 and 36/49, respectively. Kyungnam(Kimhae and Pusan) shards with an efficiency of 13/26 are mostly located near to the linear boundary between kyungnam and Kyungki. The overlap between two groups could be attributed to 5 samples located in the distance from the boundary. A further efficient separation between groups could be affected by eliminating the 5 samples.

The same procedure described above has been applied to 45 samples of Koryung for classifying into three subgroups which are based on geographical sites as follows.

Means and standard deviations have been calculated for each subgroup and overall samples and are given in Table 6 and 7, respectively. W matrix and B matrix have been generated using original data set together with the data given in Table 6 and 7. Discriminant functions have been obtained by using the set of eigen vector coefficients corresponding to each calculated eigen value. Data for the selection of variables have been derived and are given in Table 8. As the results, the classification between subgroups were found to be mainly attributed to 8 elements such as Sm, Ga, La, Cs, Tb, Sc, Rb and Fe.

Using the data of the selected 8 elements $W^{-1}B$ matrix (8×8) has been generated again and eigen vector coefficients were calculated by using corresponding eigen values. Since

Table 2. Mean Values (ppm) of Elemental Contents in Potsherds

Sites \ Elements	Sm	Cu	Ga	K	La	Na	Ce	Lu	Th	Cr
Kyungki	2.26	83.6	14.8	6.19×10^3	29.1	4.78×10^3	72.2	0.900	14.8	69.3
Koryung	2.47	67.9	14.8	6.64×10^3	34.5	3.44×10^3	88.5	1.06	13.3	99.5
Kyungnam	2.36	159	17.2	5.75×10^3	29.9	5.02×10^3	71.7	0.133	13.5	53.8
Total	2.38	94.2	15.4	6.19×10^3	31.6	4.26×10^3	79.1	0.794	13.9	78.9

Sites \ Elements	Hf	Ba	Cs	Tb	Sc	Rb	Ta	Fe	Co	Eu
Kyungki	5.43	449	5.38	0.549	10.2	110	1.09	1.72×10^4	10.5	0.825
Koryung	5.96	474	7.13	0.699	12.2	116	1.24	2.04×10^4	12.2	0.846
Kyungnam	5.53	663	5.67	0.646	11.1	95.0	1.84	1.89×10^4	51.8	0.847
Total	5.68	508	6.20	0.635	11.3	109	1.33	1.89×10^4	20.6	0.839

Table 3. Standard Deviation of Elemental Contents in Potsherds

Sites \ Elements	Sm	Cu	Ga	K	La	Na	Ce	Lu	Th	Cr
Kyungki	0.571	28.3	3.56	2.48×10^3	10.1	2.19×10^3	26.6	2.20	3.34	21.4
Koryung	0.577	30.9	4.72	2.79×10^3	12.0	1.59×10^3	35.5	2.44	2.25	41.1
Kyungnam	1.33	126	15.4	3.36×10^3	17.4	2.64×10^3	42.9	0.068	5.80	24.3
Total	0.808	74.3	8.20	2.82×10^3	13.0	2.18×10^3	35.4	2.07	3.71	36.9

Sites \ Elements	Hf	Ba	Cs	Tb	Sc	Rb	Ta	Fe	Co	Eu
Kyungki	1.26	140	1.86	0.240	2.91	33.5	1.08	6.21×10^3	16.4	0.221
Koryung	1.89	156	2.36	0.523	4.55	33.0	2.82	8.04×10^3	13.5	0.457
Kyungnam	2.90	658	2.85	0.396	4.05	50.5	3.60	8.91×10^3	92.1	0.377
Total	1.99	346	2.45	0.418	4.01	38.4	2.59	7.74×10^3	48.3	0.370

Table 4. Data for the Selection of Variables

Distance \ Elements	Sm	Cu	Ga	K	La	Na	Ce	Lu	Th	Cr
$\lvert v_1^a\ \delta_{AB}^b\rvert$	0.003	0.020	0.000	0.12	0.088	0.068	0.12	0.001	0.085	0.045
C^c (%)	0.40	2.7	0.00	17	12	9.1	15	0.31	11	6.0
$\lvert v_1^a\ \delta_{BC}^b\rvert$	0.002	0.12	0.006	0.26	0.077	0.080	0.12	0.007	0.011	0.068
C^c (%)	0.17	9.8	0.50	21	6.4	6.7	9.9	0.59	0.92	5.6

Distance \ Elements	Hf	Ba	Cs[d]	Tb	Sc	Rb	Ta	Fe	Co	Eu
$\lvert v_1\ \delta_{AB}\rvert$	0.027	0.002	0.018	0.007	0.057	0.048	0.008	0.013	0.008	0.005
C (%)	3.6	0.27	2.4	0.94	7.6	6.5	1.1	1.7	1.1	0.67
$\lvert v_1\ \delta_{BC}\rvert$	0.022	0.015	0.015	0.002	0.032	0.15	0.031	0.007	0.18	0.000
C (%)	1.8	1.3	1.3	0.17	2.7	13	2.6	0.59	15	0.00

[a]The coefficients of discriminant function corresponding to eigenvalue L_1. [b]$\delta_{AB} = \left(\frac{m_{A\cdot j} - m_{B\cdot j}}{\sigma\cdot\cdot j}\right)$ $m_{A\cdot j}$, $m_{B\cdot j}$: The mean values of the jth variable in group A and B. $\sigma\cdot\cdot j$: the overall standard deviation of the jth variable. A, B and C are denoted to Kyungki, Koryung and Kyungnam, respectively. [c]Contribution percentage $\left(= \frac{|v_j\cdot\delta_{AB}|}{D^2} \times 100,\ D^2 = \sum_{j=1}^{M} |v_j\cdot\delta_{AB}|\right)$. [d]Cs is especially contributed to the second eigenvalue L_2.

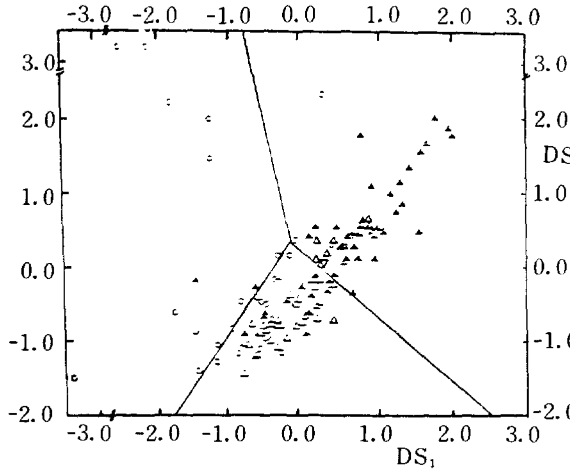


Figure 1. Plot and territorial map of discriminant score 1 versus discriminant score 2 for the Kyungki/Koryung/Kyungnam potsherds. For the symbols, see Table 1.

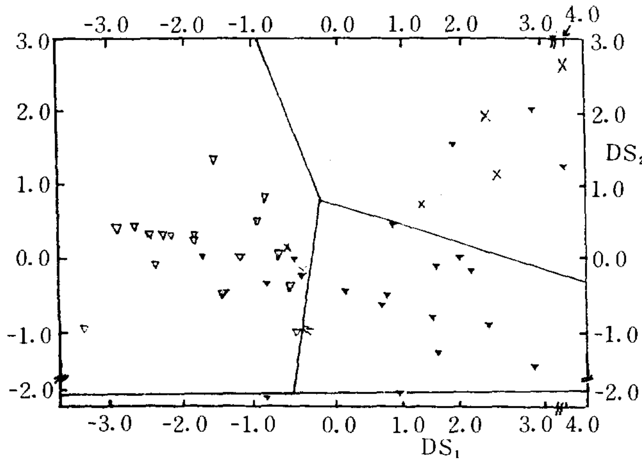


Figure 2. Plot and territorial map of discriminant score 1 versus discriminant score 2 for the subgrouping of Koryung potsherds. For the symbols, see those in parentheses in Table 1.

Table 5. Prediction Results for the Kyungki/Koryung/Kyungnam Potsherds

A priori group membership	Number of samples	A posteriori (predicted) group membership		
		Kyungki	Koryung	Kyungnam
Kyungki	39	32(28 %)	12(11 %)	10(8.7%)
Koryung	49	6(5.2%)	36(32 %)	3(2.6%)
Kyungnam	26	1(0.9%)	1(0.9%)	13(11%)

Table 6. Mean Values (ppm) of Elemental Contents in Potsherds Collected from Koryung

Elements / Sites	Sm	Cu	Ga	K	La	Na	Ce	Lu	Th	Cr
Koryung A	2.09	52.2	12.7	4.83×10^3	26.1	3.08×10^3	70.3	0.912	11.5	76.2
Koryung B	2.77	66.6	16.1	6.88×10^3	36.9	3.12×10^3	95.6	1.65	14.4	107
Koryung C	2.43	100	14.6	9.79×10^3	42.7	4.69×10^3	103	0.127	14.1	126
Total	2.44	65.0	14.5	6.40×10^3	33.2	3.31×10^3	85.9	1.13	13.2	96.9

Elements / Sites	Hf	Ba	Cs	Tb	Sc	Rb	Ta	Fe	Co	Eu
Koryung A	5.63	426	5.38	0.465	9.19	90.3	1.76	1.44×10^4	9.16	0.591
Koryung B	5.91	508	8.40	0.934	13.1	131	0.961	2.19×10^4	15.0	0.979
Koryung C	6.92	423	7.98	0.662	16.1	129	0.779	3.01×10^4	12.6	1.04
Total	5.93	463	7.07	0.700	11.9	114	1.27	1.98×10^4	12.2	0.793

a three-fold classification problems is also involved in this case, two discriminant functions have been computed similarly as described above. The first discriminant function f_1 corresponds to the largest eigen value L_1 of 2.61 and the second function f_2 to the eigen value L_2 of 0.86. The discriminant scores for individual i along the optimal discriminant function axes f_1 and f_2 have been generated and are given as follows.

$$L_1 = 2.610$$
$$DS_{1,i} = 0.257z_{Sm,i} + 0.194z_{Ga,i} + 0.297z_{La,i} + 0.347z_{Cs,i} + 0.321z_{Tb,i} + 0.106z_{Sc,i} + 0.393z_{Rb,i} + 0.648z_{Fe,i}$$

$$L_2 = 0.863$$
$$DS_{2,i} = -0.292z_{Sm,i} - 0.197z_{Ga,i} - 0.045z_{La,i} - 0.244z_{Cs,i} - 0.512z_{Tb,i} + 0.587z_{Sc,i} + 0.054z_{Rb,i} + 0.452z_{Fe,i}$$

Figure 2 shows a map of the individuals of the three subgroups and the corresponding group centroids in the 2-dimensional discriminant space with a territorial diagram of each group. The classification efficiencies are shown in Table 9, which shows that shards of Koryung A, B and C form each isolated group with efficiencies of 17/19, 12/20 and 4/6,

Table 7. Standard Deviation of Elemental Contents in Potsherds Collected from Koryung

Sites \ Elements	Sm	Cu	Ga	K	La	Na	Ce	Lu	Th	Cr
Koryung A	0.546	19.9	2.36	1.17×10^3	5.46	1.08×10^3	31.3	2.21	1.76	25.6
Koryung B	0.446	29.7	4.01	2.71×10^3	11.7	1.73×10^3	35.3	3.09	1.78	42.1
Koryung C	0.196	32.7	4.05	3.05×10^3	12.2	1.68×10^3	32.9	0.04	1.24	50.5
Total	0.561	30.1	3.69	2.74×10^3	11.3	1.55×10^3	35.4	2.53	2.18	40.9

Sites \ Elements	Hf	Ba	Cs	Tb	Sc	Rb	Ta	Fe	Co	Eu
Koryung A	2.19	177	1.40	0.282	2.04	22.4	4.50	5.04×10^3	10.1	0.311
Koryung B	1.71	109	2.20	0.684	4.29	32.1	0.601	5.59×10^3	18.5	0.427
Koryung C	1.58	83.7	1.38	0.135	5.95	16.7	0.184	1.04×10^4	3.87	0.368
Total	1.92	143	2.30	0.534	4.46	33.0	2.94	8.07×10^3	14.1	0.436

Table 8. Data for the Dimension Reduction of Koryung Potsherds

Distance \ Elements	Sm	Cu	Ga	K	La	Na	Ce	Lu	Th	Cr
$\|v_1^a\ \delta_{AB}^b\|$	0.36	0.043	0.22	0.028	0.23	0.002	0.15	0.028	0.26	0.073
C^c (%)	10	1.3	6.4	0.81	6.7	0.06	4.4	0.81	7.6	2.1
$\|v_1\ \delta_{BC}^b\|$	0.18	0.10	0.10	0.040	0.12	0.065	0.047	0.057	0.023	0.042
C (%)	11	6.5	6.5	2.6	8.0	4.2	3.0	3.7	1.5	2.7

Distance \ Elements	Hf	Ba	Cs	Tb	Sc	Rb	Ta	Fe	Co	Eu
$\|v_1\ \delta_{AB}\|$	0.015	0.054	0.33	0.26	0.14	0.70	0.053	0.27	0.002	0.23
C (%)	0.44	1.6	9.7	7.4	4.0	20	1.5	7.8	0.06	6.6
$\|v_1\ \delta_{BC}\|$	0.054	0.056	0.047	0.15	0.11	0.034	0.012	0.29	0.001	0.028
C (%)	3.5	3.6	3.0	9.6	6.8	2.2	0.77	19	0.06	1.8

[a] The coefficients of discriminant function corresponding to eigenvalue L_1. [b] $\delta_{AB} = \left(\frac{m_{A\cdot j} - m_{B\cdot j}}{\sigma\cdot\cdot j}\right)$ $m_{A\cdot j}$, $m_{B\cdot j}$: the mean values of the jth variable in group A and B. $\sigma\cdot\cdot j$: the overall standard deviation of the jth variable. A, B and C are denoted to Koryung A, Koryung B and Koryung C, respectively. [c] Contribution percentage$\left(= \frac{|v_j\cdot\delta_{AB}|}{D^2} \times 100,\ D^2 = \sum_{j=1}^{N} |v_j\cdot\delta_{AB}|\right)$.

Table 9. Prediction results for the Subgrouping of Koryung Potsherds

A priori group membership	Number of samples	A posteriori (predicted) group membership Koryung A	Koryung B	Koryung C
Koryung A	19	17(38 %)	5(11 %)	2(4.4%)
Koryung B	20	2(4.4%)	12(27 %)	0(0.0%)
Koryung C	6	0(0.0%)	3(6.7%)	4(8.9%)

respectively.

The discriminant functions derived and given above can be used as classification rule for unknown samples.[7] A further study for this purpose is in progress.

Acknowledgement. The present studies were supported(in part) by the Basic Science Research Institute Program, Ministry of Education, 1985.

References

1. M.J. Karson, "Multivariate Statistical Methods", Iowa State University Press, 1982.
2. K. Varmuza, "Lecture Notes in Chemistry", "Pattern Recognition in Chemistry", New York, 1980.
3. D.N. Lawley and A.E. Maxwell, "Factor Analysis as a Statistical Method", Butterworths, London, 1963.
4. C. Lee, O.C. Kwun, N.B. Kim and I.C. Lee, Bulletin of Korean Chem. Soc., **6**, 241 (1985).
5. C. Lee, O.C. Kwun and H.T. Kang, *ibid*, **7**, 73 (1986).
6. R.A. Fisher, *Ann. Eugenics*, **7**, 127 (1936).
7. D. Coomans and D.L. Massart, *Analytica Chimica Acta*, **112**, 97 (1979).
8. W.W. Cooley and P.R. Lohnes, "Multivariate Data Analysis", Wiley, New York, 1971.
9. M. Kendall, "Multivariate Analysis", Charles Griffin and Co., London, 1975.

Thermotropic Compounds with Two Terminal Mesogenic Units and a Central Spacer, 8. Mutual Miscibility between the Dimesogenic, Nematic Compounds

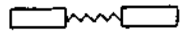
Jung-Il Jin,* E-Joon Choi, and Joo-Hoon Park

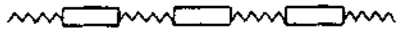
Department of Chemistry, Korea University, Seoul 132. *Received June 2, 1986*

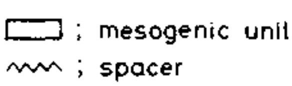
Mutual miscibility between thermotropic, nematic compounds with two terminal mesogenic units and a central spacer was studied by differential scanning calorimetry (DSC) and on a polarizing microscope. It was found that the isomorphous, nematic dimesogenic compounds with wide variety of structures are miscible in mesophases with each other over the whole range of composition and that Schröder-van Laar equation almost correctly predicts the melting temperature and composition of eutectic mixtures. There was a pair of compounds which were exceptional and did not form a eutectic mixture and, instead, revealed a monotonous change in melting (T_m) and isotropic transition temperatures (T_i) as the composition of the mixture was varied. The compounds were of almost same structure in shape and seemed to undergo formation of solid solution.

Introduction

We have been investigating the liquid crystalline properties of a wide variety of series of thermotropic dimesogenic compounds with two terminal mesogenic units and central polymethylene spacers.[1-3] It was found that melting (T_m) and isotropization temperatures (T_i) of this type of compounds (I) show odd-even dependence on the number of carbon atoms in the central spacer and that these compounds can be taken as models for main chain thermotropic polymers (II) with similar mesogenic units and the same spacers.[3]

 (I)

 (II)



Other thermodynamic parameters such as change in enthalpy (ΔH_i) and entropy (ΔS_i) for isotropization also showed a regular odd-even dependence on the number of methylene units in the spacers.

On the other hand, it is well known that selective miscibility rule originally developed by Sackmann and Demus[5] is applicable to low molecular weight thermotropic compounds.[6,7] This rule states, "all liquid crystalline modifications which exhibit an uninterrupted series of mixed crystals in binary systems without contradiction can be marked with the same symbol". This rule, however, has been tested up to now mainly for monomesogenic compounds, *i.e.* those with only one mesogenic unit.

It is also well known that Schröder-van Laar equation (III or IV) predicts correctly the eutectic compositions of binary mixtures of monomesogenic compounds.[8,9]

$$-\ln x_1 = \frac{\Delta H_m^1}{R}\left(\frac{1}{T} - \frac{1}{T_m^1}\right) \qquad (\text{III})$$

$$-\ln x_2 = -\ln(1-x_1) = \frac{\Delta H_m^2}{R}\left(\frac{1}{T} - \frac{1}{T_m^2}\right) \qquad (\text{IV})$$

where ΔH_m's are heats of melting and x's mole fractions. The numbers 1 and 2 stand for each component.

While we were trying to establish the structure-property relationship of dimesogenic compounds of type I, we became interested in their mutual miscibility in mesophase. Whether