Int J Adv Manuf Technol (1997) 13:637–648

The International Journal of
Advanced Manufacturing Technology

A Flexible Classification Method for Evaluating the Utility of Automated Workpiece Classification System

S. H. Hsu, T. C. Hsia and M. C. Wu

Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu, Taiwan, Republic of China

In group technology, workpieces are categorised into families according to their similarity in design or manufacturing attributes. This categorisation can eliminate design duplication and facilitate the production of workpieces. Much effort has been focused on the development of automated workpiece classification systems. However, it is difficult to evaluate the utility of such systems. The objective of this study was to develop a benchmark classification system based on global shape information for use in evaluating the utility of workpiece classification systems. A classification system has a high level of utility if its classification scheme is consistent with users' perceptual judgement of the similarity between workpiece shapes. Hence, in the proposed method, the consistency between a classification system and users' perceptual judgements is used as an index of the utility of the system. The proposed benchmark classification has two salient characteristics:

1. *It is user-oriented, because it is based on users' judgements concerning the similarity of the global shape of workpieces.*
2. *It is flexible, allowing users to adjust the criteria of similarity applied in the automated workpiece classification.*

The development of this classification consisted of three steps:

1. *Gathering row data on global shape similarity from a group of representative users and modelling the data by fuzzy numbers.*
2. *Developing benchmark classification for various similarity criteria by using fuzzy clustering analysis.*
3. *Developing indices for evaluating the appropriate number of workpiece categories and homogeneity within each group.*

The applicability of the benchmark classification system in evaluating the utility of automated workpiece classification systems was examined.

Correspondence and offprint requests to: Dr S. H. Hsu, Department of Industrial Engineering and Management, 1001 Ta-Hsueh Road, Hsinchu, Taiwan, Republic of China.

Keywords: Automated workpiece classification; Benchmark classification; Flexible classification method; Group technology

1. Introduction

Workpiece coding schemes are widely used in the implementation of group technology (GT) to classify workpieces according to the similarity of their design and manufacturing attributes. The results of workpiece classification can be used to establish design and manufacturing databases, which facilitate the retrieval of similar designs and the standardisation of manufacturing processes and thus enhance design and manufacturing productivity.

The design and manufacturing attributes used in coding workpieces generally involve shape (i.e. geometric form and size), function, material, and other manufacturing characteristics. Among these attributes, in recent years shape-related attributes have drawn much attention from researchers because of the increasing demand for fully integrated CAD/CAM systems.

Manual coding of workpieces on the basis of their shape is a time-consuming and error-prone process. The operator has to memorise all the template-shapes and then match a particular template-shape with each workpiece. Few operators can perform such matching accurately and reliably, especially when a large number of workpieces are involved. To overcome this problem, researchers have developed several automated classification systems [1–4]. Most of these approaches use individual local geometric features as the descriptors for workpiece classification, and approaches of this type have been shown to be useful in the planning of manufacturing processes.

However, there are two shortcomings to using individual features as classification criteria:

1. As Fig. 1 shows, similarity of isolated individual features does not necessarily entail similarity in global shape.
2. Isolated individual features cannot be used for identification during the early stages of the design phase, because the designer's conceptual model evolves from an overall, global picture to individual details.

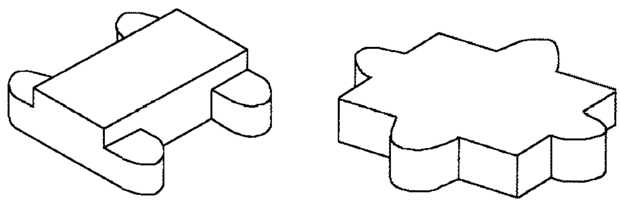


Fig. 1. Two workpieces with similar individual features but significantly different in global shape information [5].

Thus the use of local features as workpiece classification criteria is generally limited to information retrieval and practical applications.

In more recent GT research, workpieces have been described and classified on the basis of the overall contour of the workpiece instead of local attributes [6–8]. This approach enhances performance in the design phase and increases efficiency on the manufacturing and assembly lines. One of the key criteria in choosing a practical automated classification system for design, manufacturing, and assembly is whether the classification results are compatible with the user's own classification. To ensure that this criterion is met, benchmarks reflecting the user's classification are needed to evaluate the performance of automated classification systems.

The purpose of this research was to establish a system for generating such benchmark classifications with which to measure the utility of automated workpiece classification systems. The utility of an automated classification system depends on the ease with which users can store and retrieve information and on the extent to which the classification system is consistent with the user's intuitive judgement of the similarities and differences between different workpieces. In this work, a set of sample workpieces was selected and classified according to users' intuitive perception of the workpieces' global shapes.

A flexible system for generating benchmark classifications was then established, which allows users to adjust the similarity evaluation criteria to suit their particular requirements. If a user adopts a stringent criterion, so that only workpieces with a high degree of similarity are classified as belonging to the same group, then a benchmark classification will be obtained that includes more groups, each containing a small number of workpieces. On the other hand, if the user decides that for a particular application, even workpieces with a low degree of similarity can be grouped together, then a benchmark classification will be obtained that includes fewer groups, but each group will contain more workpieces. Since the criteria for forming the benchmark classification can be adjusted freely, we call the proposed benchmark classification system a "flexible classification method". Each particular classification produced by the system is called a "benchmark classification". A classification of the same set of workpieces generated by a particular automated classification system is called a "test classification". The utility of an automated classification system can be assessed by comparing the results of the test classification with one or more benchmark classifications.

After the benchmark classification is determined, two indices are used to measure the level of consistency between the test classification and the benchmark classification. The first is an index of the number of workpiece groups, which is used to check whether there are too many or too few groups. If there are too many groups, then the similarity criteria used in the test classification system are too stringent. On the other hand, if there are too few groups, then the similarity criteria are too loose. The second index is an index of the level of homogeneity within each group. This index can be used to compare the level of similarity between corresponding groups in the test classification and the benchmark classification. A one-to-one correspondence is found between groups in the benchmark classification and the test classification by working from the classification with fewer groups. For instance, if the benchmark classification has fewer groups, then a correspondence is assigned between the groups in the benchmark classification and the groups in the test classification that are most similar to them. The higher the level of similarity between each pair of corresponding groups, the more accurate the test classification is.

2. Establishing the Benchmark Workpiece Classification

The process of establishing the benchmark workpiece classification can be divided into two stages:

1. Collecting and aggregating information.
   (1) Selection of subjects.
   (2) Method of representing workpieces for comparison.
   (3) Definition of linguistic terms and membership functions.
   (4) Aggregation of membership functions and defuzzification.
2. Establishing the benchmark classifications.
   (1) Fuzzy clustering analysis.
   (2) Using aggregated comparison crisp numbers for workpieces clustering.

2.1 Collecting and Aggregating Information

The aim of this stage is to establish the benchmark workpiece classification. In this research, 30 subjects were asked to make pair comparisons of the global shape of the 36 sample workpieces shown in Fig. 2. When subjects are asked to compare the similarity of various objects, their judgements are often limited by their attention span, memory capacity, and previous experience. Their responses are generally fuzzy and cannot be expressed by crisp numbers from 0 to 9. To cope with the fuzziness inherent in human cognitive processes, we used linguistic variables for the similarity comparison and integrated the subjects' responses by means of a fuzzy number operation. The fuzzy numbers were then changed into crisp numbers through a defuzzification process. The detailed procedures are described below.

2.1.1 Selection of Subjects

Since the aim of this project was to employ users' intuitive classification of workpieces as criteria for evaluating the per-

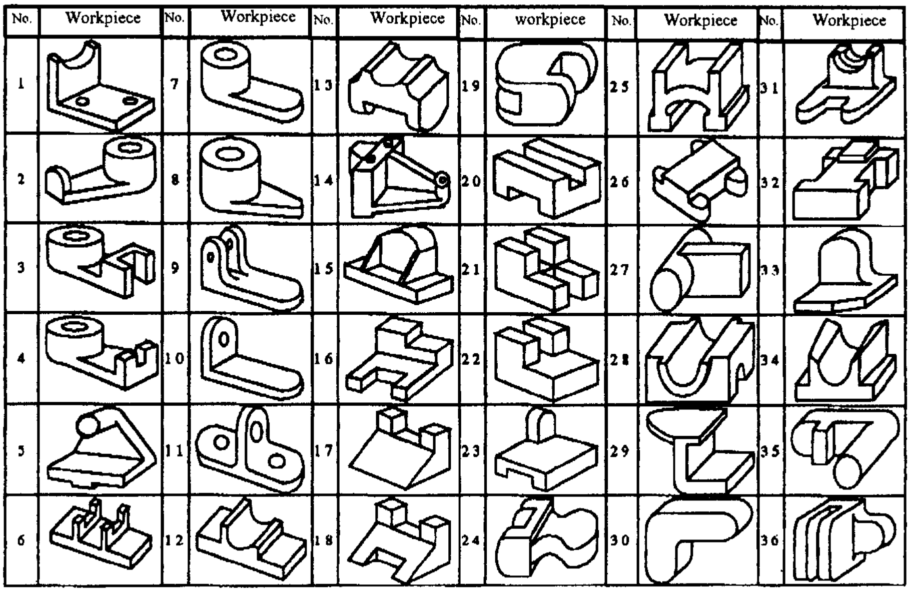


Fig. 2. The 36 sample workpieces used for pair comparison.

formance of automated classification systems, it was important that the subjects be truly representative of the intended users of such systems. Ideally, in a project of this type, the subjects should be randomly and proportionally selected from all departments of the factory that will use the automatic classification system. In this case, 30 subjects were selected from various departments of an aircraft plant. The sample was designed to reflect the actual proportions of the firm's workers employed in various tasks: 22 of the subjects were operators who worked directly on the production line (5 in the tooling design section and the other 17 in manufacturing and structural assembly), and 8 were employed in program management, production control, and other departments.

2.1.2 Method of Representing Workpieces for Comparison

Workpieces are 3D objects. There are two methods of representing a 3D object in a 2D drawing. First, the object can be represented by means of three 2D orthogonal drawings (front view, side view, and top view) as shown in Fig. 3(*a*). This method depicts the shape of the object exactly, but makes it difficult for the subject to picture the entire 3D object mentally. Secondly, the object can be presented by means of an isometric drawing, as shown in Fig. 3(*b*). This type of drawing conveniently depicts both the overall structure and individual details of the 3D object, and because the drawing conveys a great deal of information in a compact form, it facilitates recognition and comparison of objects [9]. However, because isometric drawings provide only one view of an object, they can be misleading in at least two ways. First, two objects with a similar shape may appear more alike than they really are, because their differentiating features may be hidden from a certain angle of view. Secondly, as mentioned by Arnhiem [10], an object presented in a 2D drawing will carry more visual weight in the upper or left part of the drawing. To prevent these two facts from biasing the subjects' judgement

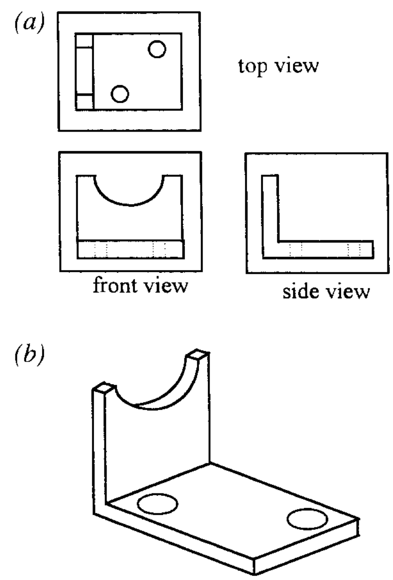
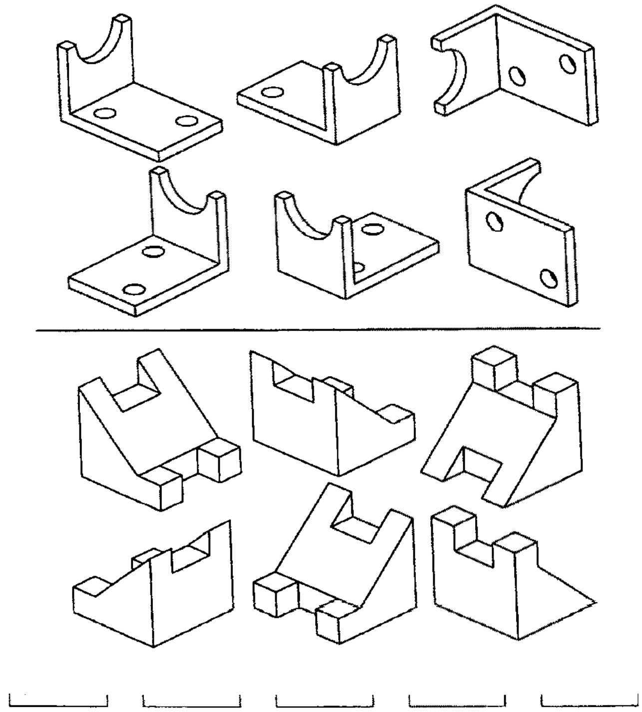


Fig. 3. Workpieces represented in (*a*) 2D orthogonal drawing and (*b*) isometric drawing.

of the similarity of the sample workpieces, in this research the subject were presented with front, back, top, bottom, left, and right isometric views of each of the workpieces, as shown in Fig. 4. In this way the subjects could easily determine the overall shape of the workpieces, and they then selected linguistic terms on an answer sheet to indicate the level of similarity between various workpieces. The evaluation scale used offered a choice of five terms: very low similarity, low similarity, medium similarity, high similarity, and very high similarity.

To compare all possible pairs of the 36 sample workpieces, the subjects had to make a total of 630, or C_2^{36}, individual comparisons. In order to ensure that the subjects could make consistent judgements, each subject worked independently,



U O I E A

Fig. 4. Workpieces represented in 6 isometric views for comparison. U, very low similarity; O, low similarity; I, medium similarity; E, high similarity; A, very high similarity.

without a time limit. The 630 answer sheets were randomly distributed to the subjects. The results were coded by using the letters {A,E,I,O,U} to denote the linguistic terms {very high similarity, high similarity, medium similarity, low similarity, very low similarity}, as shown in Table 1.

2.1.3 Definition of Linguistic Terms and Their Membership Functions

As proposed by Chen and Hwang [11], the membership function for each of the five linguistic terms was taken to be a trapezoidal fuzzy number, as shown in Fig. 5 (see the Appendix).

2.1.4 Aggregation of Membership Functions and Defuzzification

In this step of the analysis, fuzzy number operations were employed to aggregate the information collected in preceding stages. The membership function for each linguistic term can be seen as a fuzzy number. Linguistic terms chosen by more than two subjects were integrated into membership functions through operations with fuzzy numbers. There are many methods of aggregating a decision-maker's fuzzy assessments, such as mean, median, maximum, minimum, and mixed operators [12]. The most commonly used aggregation method, however, is the average operation. In the present research, the average operator was used to aggregate the subjects' judgements of the similarity of various pairs of workpieces by using the formula:

$$\tilde{S}_{ij} = \left(\frac{1}{n}\right) \otimes (\tilde{S}_{ij1} \oplus \tilde{S}_{ij2} \oplus \cdots \oplus \tilde{S}_{ijk} \oplus \cdots \oplus \tilde{S}_{ijn}) \quad (1)$$

where

- $\oplus$ represents addition of fuzzy numbers
- $\otimes$ represents multiplication of fuzzy numbers
- $\tilde{S}_{ijk}$ represents the membership function for the linguistic term, which is obtained by taking the number of subjects k comparing the similarity of workpiece i and workpiece j, where $k = 1, \ldots, n$, $1 \leq i < j \leq w$, $n = 30$ (number of subjects), and $w = 36$ (number of workpieces)
- $\tilde{S}_{ij}$ represents the membership function after aggregation of the n subjects' similarity comparison between workpiece i and workpiece j, where $i < j$.

In order to perform a crisp number calculation of the aggregated membership functions, it is necessary to determine a crisp value for the fuzzy sets concerned. This type of a transition process is called *defuzzification*, which generally uses the centre of gravity or α-cut element average method [13]. In this research the centre of gravity method was used to seek the solutions, because, in terms of geometry, the centre of gravity is the most representative point of the fuzzy set. The aggregated membership function is expressed in equation (2), and its value is denoted by X_G [14].

$$X_G(\tilde{S}) = \frac{\int_0^1 x\mu_{\tilde{S}}(x)\mathrm{d}x}{\int_0^1 \mu_{\tilde{S}}(x)\mathrm{d}x} \quad (2)$$

Because the linguistic term chosen by the subject to express the degree of similarity between workpieces is a trapezoidal fuzzy number, the aggregated membership function remains a trapezoidal fuzzy number. If the defuzzification with the centre of gravity formula (2) is simplified, so that it is denoted by a trapezoidal fuzzy number (a,b,c,d), it is easy to determine a crisp comparison number for the fuzzy relations among the aggregated 36 sample workpieces. The simplified formula can be expressed as follows:

$$X_G(\tilde{S}) = \frac{-a^2-b^2+c^2+d^2-ab+cd}{3(-a-b+c+d)} \quad (3)$$

2.2 Establishing the Benchmark Classifications

In this stage, fuzzy clustering analysis was employed to classify the sample workpieces into groups based on their similarity levels after being compared and aggregated. These groups represent the users' judgments about the degree of similarity between various workpieces. As described in the Introduction, different benchmark classifications can be formed, depending on the user's needs. Each particular classification can then be used as a benchmark to evaluate the utility of the automatic classification system under test. The clustering procedures are described below.

Table 1. The fuzzy relation of 36 workpieces by pair comparison.

Work-piece	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	Work-piece
1	1	I	I	I	I	O	I	I	E	E	E	E	I	I	I	I	I	I	I	I	I	I	I	E	I	E	I	E	I	O	I	I	E	E	I	I	1
2		1	A	A	E	I	A	A	E	E	I	I	I	I	I	O	O	O	I	O	O	O	I	E	O	I	E	O	I	O	I	O	I	I	E	I	2
3			1	A	I	I	E	E	E	I	E	E	I	E	I	E	E	I	I	E	E	E	E	I	I	I	E	I	I	I	I	E	I	I	E	I	3
4				1	I	I	A	A	I	I	E	I	I	E	E	E	E	I	I	E	E	E	E	I	O	I	E	O	I	O	I	E	I	I	E	I	4
5					1	I	E	E	I	I	I	E	E	E	E	E	E	I	I	E	E	E	I	I	I	E	E	I	E	I	E	E	I	I	E	I	5
6						1	O	O	I	I	I	E	I	I	E	A	E	E	I	E	A	A	E	I	E	I	I	E	I	I	I	A	I	E	I	E	6
7							1	A	E	E	E	I	E	E	I	O	O	O	I	O	O	O	I	E	I	I	E	I	I	I	E	O	E	I	E	I	7
8								1	E	E	E	I	I	E	E	O	O	O	I	O	O	O	I	I	I	I	E	I	I	I	I	O	E	I	E	I	8
9									1	A	A	I	E	I	E	I	I	I	E	I	I	I	E	A	I	I	E	I	I	E	I	I	E	I	I	E	9
10										1	A	I	I	I	I	I	I	I	E	I	I	I	E	E	I	E	E	I	E	E	E	I	E	I	I	I	10
11											1	I	E	I	I	O	O	I	E	I	I	I	I	E	I	E	I	I	I	E	E	I	E	I	E	E	11
12												1	E	I	I	E	E	E	E	E	E	E	I	I	E	I	I	E	E	I	E	E	I	A	I	E	12
13													1	I	E	E	I	E	A	I	I	I	E	A	I	I	I	E	I	I	E	O	E	E	I	E	13
14														1	I	O	O	O	I	I	I	I	I	O	O	O	I	I	I	I	E	I	I	I	E	I	14
15															1	I	I	I	E	I	I	I	E	E	I	E	I	I	I	E	E	I	E	I	E	E	15
16																1	A	A	E	A	A	A	E	I	E	I	I	E	E	I	E	A	I	E	I	E	16
17																	1	A	I	E	A	A	E	I	E	I	I	E	E	I	I	A	I	E	I	I	17
18																		1	I	E	E	E	E	I	E	I	I	I	I	I	E	E	I	E	I	E	18
19																			1	I	I	I	E	A	I	E	I	I	E	I	E	I	E	E	I	E	19
20																				1	A	A	E	I	E	I	I	E	E	I	E	A	I	E	I	E	20
21																					1	A	E	I	E	O	I	E	E	I	E	A	I	E	I	E	21
22																						1	E	I	E	I	I	I	E	I	I	A	I	E	I	E	22
23																							1	E	E	A	I	I	E	I	E	E	E	E	E	E	23
24																								1	E	E	I	I	E	E	E	I	E	I	E	E	24
25																									1	I	O	E	I	I	I	E	I	A	I	I	25
26																										1	I	I	E	E	E	I	E	I	E	I	26
27																											1	I	I	I	I	E	O	I	E	I	27
28																												1	E	I	I	I	I	E	I	I	28
29																													1	I	E	E	E	E	I	E	29
30																														1	E	I	E	I	E	I	30
31																															1	I	E	E	E	E	31
32																																1	I	E	I	E	32
33																																	1	I	O	E	33
34																																		1	I	E	34
35																																			1	I	35
36																																				1	36
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	Work-piece

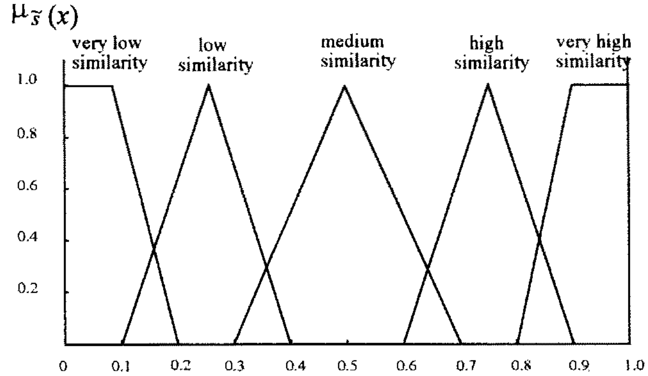


Fig. 5. Membership functions for linguistic terms indicating level of similarity.

2.2.1 *Fuzzy Clustering Analysis*

To form clusters on the basis of a fuzzy relation $\tilde{R}$, $\tilde{R}$ must satisfy the similarity relation [15], which is defined as follows:

Definition. Consider $X = \{x_1, x_2, \ldots, x_n\}$ to be a reference set. A binary fuzzy relation $\tilde{R}$ on X is a fuzzy subset of the Cartesian product $X \times X$.
Let $\mu_{\tilde{R}}$: $X \times X \rightarrow [0,1]$
denote the membership function of $\tilde{R}$ and let

$$\tilde{R} = [r_{ij}]_{n \times n}$$
$$r_{ij} = \mu_{\tilde{R}}(x_i, x_j) \quad (x_i, x_j \in X)$$

A fuzzy relation $\tilde{R}$ on X is said to be reflexive if $\mu_{\tilde{R}}(x_i, x_j) = 1$ for all $x_i \in X$, $\tilde{R}$ is symmetric if $\mu_{\tilde{R}}(x_i, x_j) = \mu_{\tilde{R}}(x_j, x_i)$ for all $x_i, x_j \in X$, and $\tilde{R}$ is (max–min) transitive if for any $x_i, x_k \in X$

$$\mu_{\tilde{R}}(x_i, x_k) \geq \max\{\min\{\mu_{\tilde{R}}(x_i,x_j), \mu_{\tilde{R}}(x_j,x_k)\}\} \quad (\forall x_j \in X)$$

If the fuzzy relation $\tilde{R}$ on X is reflexive, symmetric, and transitive, then $\tilde{R}$ is said to be a *similarity relation* on X, denoted by $\tilde{\tilde{R}}$ [15].

If $\tilde{\tilde{R}}$ is a similarity relation on X, and α-cut of the membership matrix ($\tilde{\tilde{R}}^{\alpha}$), $\alpha \in [0,1]$, is a similarity relation on X. These similarity relations can be used to classify elements in X [15,16]. Such a classification is called *fuzzy clustering analysis*. The following is an example of fuzzy clustering analysis.

Example 1. Assume $\tilde{R}$ is a fuzzy relation on the set $X = \{x_1, x_2, \ldots, x_5\}$ with its membership matrix as shown below. The elements in this matrix represent the similarity level between x_i and x_j ($i = 1, 2, \ldots, 5$; $j = 1, 2, \ldots, 5$; $i \neq j$, when $i = j$ the level of similarity is 1).

$$\tilde{R} = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{bmatrix} 1 & 0.1 & 0.8 & 0.5 & 0.3 \\ 0.1 & 1 & 0.1 & 0.2 & 0.4 \\ 0.8 & 0.1 & 1 & 0.3 & 0.1 \\ 0.5 & 0.2 & 0.3 & 1 & 0.6 \\ 0.3 & 0.4 & 0.1 & 0.6 & 1 \end{bmatrix} \end{array}$$

By the above definition, $\tilde{R}$ is reflexive and symmetric. We now need to calculate the max–min transitivity. To do so, we search $\tilde{R}^2 = \tilde{R} \circ \tilde{R}$, …, until we find $\tilde{R}^{2k} = \tilde{R}^k$. In this example, we calculate $\tilde{R}^2$, $\tilde{R}^4$, and then $\tilde{R}^8$. Finally, we obtain $\tilde{R}^8 = \tilde{R}^4$, and thus $\tilde{R}^4$ is a similarity relation on X, which we denote by $\tilde{\tilde{R}}$.

$$\hat{R} = \tilde{R}^4 = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{bmatrix} 1 & 0.4 & 0.8 & 0.5 & 0.5 \\ 0.4 & 1 & 0.4 & 0.4 & 0.4 \\ 0.8 & 0.4 & 1 & 0.5 & 0.5 \\ 0.5 & 0.4 & 0.5 & 1 & 0.6 \\ 0.5 & 0.4 & 0.5 & 0.6 & 1 \end{bmatrix} \end{array}$$

In general, if $\tilde{\tilde{R}}$ is a similarity relation on X, different clusters can be formed by taking different α-cuts on $\tilde{\tilde{R}}$. For instance, if $\alpha = 0.8$, there will be four groups: $\{x_1, x_3\}$, $\{x_4\}$, $\{x_5\}$, and $\{x_2\}$. If $\alpha = 0.6$, there will be three groups (as shown in Fig. 6). Thus the value of α can be viewed as an index of flexibility. When the value is high, only elements of X that have a high degree of similarity will be placed near to each other; when the value is low, elements that have a low degree of similarity will also be placed together. The value of α can be used to represent subjects' judgements of the degree of similarity between various workpieces. Different values of α will produce different numbers of clusters containing different elements.

2.2.2 *Using Aggregated Comparison Crisp Numbers for Workpiece Clustering*

After the subjects compared the 36 sample workpieces, the crisp numbers obtained from the fuzzy relations representing the comparisons between the workpieces were determined through membership function aggregation and defuzzification. These crisp numbers can be seen as the fuzzy relation on $\tilde{R}$ of the 36 workpieces in the Cartesian product $X \times Y$ of a given set $X = \{x_1, x_2, \ldots, x_{36}\}$. As in the example above, this fuzzy relation is reflexive and symmetric. If the calculation satisfies the max–min transitivity and we find $\tilde{R}^{16} = \tilde{\tilde{R}}$, then this fuzzy relation is also a similarity relation on X. Arranging the α-cuts on $\tilde{\tilde{R}}$ in order from largest to smallest (1, 0.864, 0.823, 0.818, …, 0.533), we obtain different similarity relations on X. The classification results corresponding to different similarity relations for the 36 sample workpieces are shown in Fig. 7.

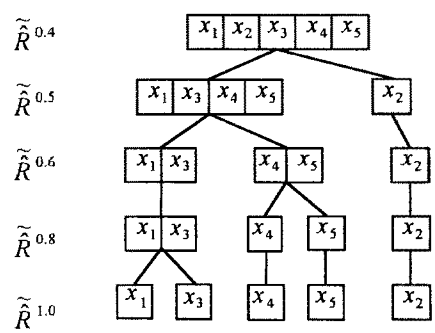


Fig. 6. Results of each cluster for the similarity relation $\tilde{\tilde{R}}$ (example 1).

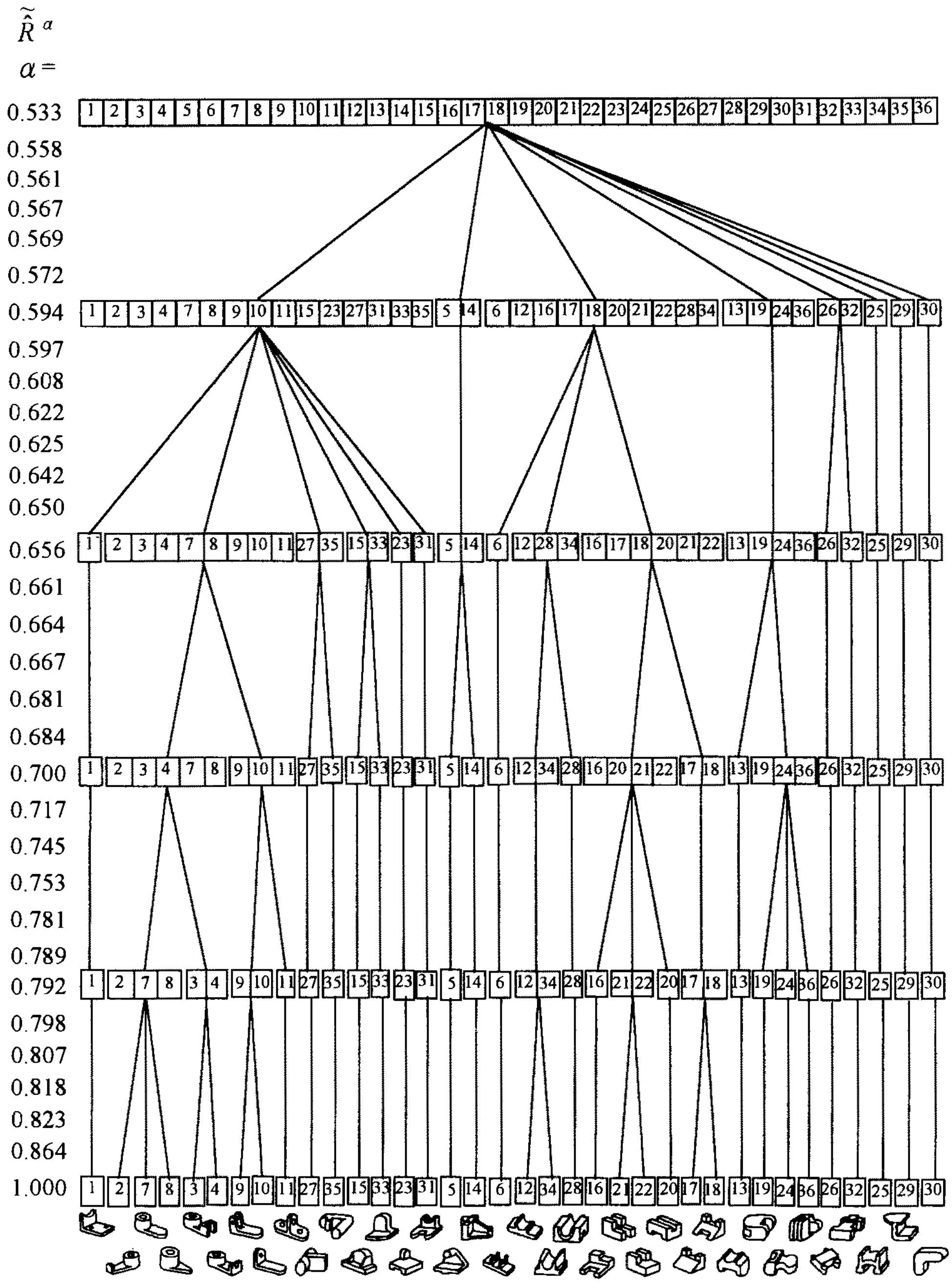


Fig. 7. The benchmark classification system (only classification for selected values of α are shown).

This chart represents the entire benchmark classification system, and the different classifications shown, which are generated by using different similarity relations, can each be used as a benchmark classification. The user can select a particular benchmark classification or number of workpiece groups (and the corresponding α-level) depending on his/her intuitive judgement of the degree of similarity between different workpieces. The indices then allow us to measure the level of consistency between the benchmark classification selected and the test classification.

3. Application of Benchmark Classification System

After establishing the benchmark classification system, we developed two indices – the appropriate number of workpiece groups and the level of homogeneity within each group – with which to evaluate the utility of test classifications. In this section, we will explain these indices and describe how we applied them to evaluate the automated workpiece classification technique developed by Wu and Jen [8].

3.1 Index of Appropriate Number of Workpiece Groups

This index is used to compare the number of groups classified by the test classification system with the number in the benchmark cluster classification, and thus to determine whether the test classification generates too many or too few groups. If too many groups are formed, the criteria of similarity applied in the system are too strict; if too few are formed, the criteria are too loose. The formula for calculating the index is shown below:

$$r = \begin{cases} \dfrac{q-1}{q^*-1} & (1 \le q \le q^*) \\ \dfrac{x-q}{x-q^*} & (q^* \le q \le x) \end{cases} \tag{4}$$

where

q = the number of groups in the test classification system

q^* = the user-specified number of groups in the benchmark classification

x = the number of workpieces being classified

r = the index of appropriate number of workpiece groups

The level of consistency between the number of groups in the test classification and in the benchmark classification as measured by this formula is indicated by a value from 0 to 1, with 0 representing the largest difference and 1 representing completely identical groups. The following example illustrates the application of this index.

Example 2. Consider a sample containing 36 workpieces ($x = 36$). If the user selects a benchmark classification with 29 groups and there are 24 groups in the test classification, then the index of the appropriate number of workpiece groups for the test classification will be:

$$r = \frac{24-1}{29-1} = 0.82$$

By the same principle, if the number of groups in the benchmark classification is 23, then the index for the test classification will be 0.92. If the number of groups in the benchmark classification is 16, then the index for the test system will be 0.60, and if the number of groups in the benchmark classification is 8, then the index will be 0.43.

3.2 Index of Homogeneity within Each Group of Workpieces

This index measures the average degree of similarity within each group in the test classification and the corresponding group in the benchmark classification. As described earlier, a one-to-one correspondence is found between groups in the two classifications by starting from the classification with fewer groups. Each pair of corresponding groups from the two classifications is called a corresponding pair. The algorithm used to calculate this index is explained below.

Step 1: Arrange the classifications being compared in a matrix.

1. x workpieces are classified by means of the two classification methods. The first method produces m groups. The classification results are denoted by $A_1, \ldots, A_i, \ldots, A_m$; A_i denotes the set of workpieces within group i. The second classification method produces n groups. The classification results are denoted by $B_i, \ldots, B_j, \ldots, B_n$; B_j denotes the set of workpieces within group j. Suppose $m \le n$.
2. The classification results from the first method are arranged in columns, and those from the second method in rows (see Table 2).

Step 2: Calculate the degree of similarity of any two groups in A_i and B_j.

1. Take A_i as the basis. a_{ij} denotes the degree of similarity between A_i and B_j. The value of a_{ij} is calculated as follows:

$$a_{ij} = \frac{N(A_i \cap B_j)}{N(A_i)} \tag{5}$$

where $N(A_i \cap B_j)$ is the same number of workpieces in A_i and B_j

$N(A_i)$ is the number of workpieces among A_i, $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

2. Take B_j as the basis, b_{ij} denotes the degree of similarity between B_j and A_i. The value of b_{ij} is calculated as follows:

$$b_{ij} = \frac{N(A_i \cap B_j)}{N(B_j)} \tag{6}$$

where $N(B_j)$ is the number of workpieces among B_j, $j = 1, 2, \ldots, n$.

3. Let c_{ij} denote the degree of similarity between A_i and B_j.

$$c_{ij} = \min\{a_{ij}, b_{ij}\} \tag{7}$$

Step 3: Determine the correspondences between groups.

1. Arrange all the values of c_{ij} in rank from large to small, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$
2. Take the largest value of c_{ij}. The corresponding A_i and B_j are the first corresponding pair of groups from the two classification methods. Delete the row i and column j that intersect at this c_{ij} in the matrix.
3. If there are two or more identical largest values of c_{ij}, then compare the second largest value in the ith row and the jth column corresponding to the largest c_{ij}, and choose the smallest of these values. Then delete the row and column corresponding to the value chosen. If the second largest values are also identical, then compare the third largest

Table 2. Results of two classification arranged in a matrix.

2nd classification / 1st classification	B_1	$\cdots$	B_j		B_n	
A_1	c_{11}	$\cdots$	c_{1j}	$\cdots$	c_{1n}	d_1
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
A_i	c_{i1}	$\cdots$	c_{ij}	$\cdots$	c_{in}	d_i
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
A_m	c_{m1}	$\cdots$	c_{mj}	$\cdots$	c_{mn}	d_m

values, and so on until a usable value is obtained. If the last values for comparison are still identical, then return to the original largest values and choose one arbitrarily.

4. Repeat the above procedure for all undeleted c_{ij} until all the rows and columns in the matrix have been deleted.
5. A one-to-one correspondence will then have been found between each group in the classification with fewer groups and a group in the classification with more groups.

Step 4: Calculate the degree of homogeneity of each pair of corresponding groups.

1. Let d_i denote the value of c_{ij} for each pair of corresponding groups.
2. Let h denote the index of homogeneity within each group. The value of h is calculated as follows:

$$h = \left(\sum_{i=1}^{m} d_i\right)/m \tag{8}$$

The index of homogeneity within each group falls between $1/x$ and 1, that is, $1/x \leq h \leq 1$. If there is a large number of workpieces in a group, that is, $x >> 0$, then the value will be between 0 and 1. The following example shows how the index works.

Example 3. Assume there are 10 workpieces ($x = 10$). Suppose there are 4 groups in the first classification and 5 in the second. The hypothetical classification results of the two classification methods are shown in Table 3. We now use the above algorithm to calculate the index of homogeneity within each group.

First, we calculate the level of similarity c_{ij} between any group of A_i and B_j. Take A_2 and B_3 as example. In this case, $a_{23} = 0.5$ and $b_{23} = 0.67$, so $c_{23} = 0.5$. Next, we set the largest value c_{ij} to be 0.5. There are three values of c_{ij} in the same group: c_{11}, c_{22}, and c_{23}. The second largest value in the second row and second column is 0.33, and that in the first row and first column and that in the second row and third column are both 0.25. Therefore, the first pair of corresponding groups should be either A_1 and B_1 or A_2 and B_3. Now we compare the next largest values in the corresponding rows and columns of c_{11} and c_{23}. Because the values are all equal to 0, then c_{11} and c_{23} can be chosen at random. If we choose c_{11} first, we will delete the first row and first column. We repeat the above procedure until all the values of c_{ij} are deleted. Then four pairs of corresponding groups can be found: A_1 and B_1, A_2 and B_3, A_3 and B_2, and A_4 and B_4 or A_4 and B_5. Finally, we calculate the index of homogeneity within each pair of two groups, $h = (0.5 + 0.5 + 0.33 + 0.25)/4 = 0.40$.

Table 3. Homogeneity between corresponding pairs of groups in example 3.

2nd classification / 1st classification	B_1 = {1,9}	B_2 = {2,3,4}	B_3 = {5,6,7}	B_4 = {8}	B_5 = {10}	
A_1 = {1}	0.5	0	0	0	0	0.5
A_2 = {2,4,5,6}	0	0.5	0.5	0	0	0.5
A_3 = {3}	0	0.33	0	0	0	0.33
A_4 = {7,8,9,10}	0.25	0	0.25	0.25	0.25	0.25

3.3 Practical Example

To demonstrate the use of the proposed benchmark classification system, we took Wu and Jen's automated classification technique [8] as the test classification system. It divides the 36 sample workpieces into 24 groups, as shown in Table 4. We then used the benchmark classification system to evaluate this system.

If the user applies a strict criterion of similarity in classifying the workpieces, then a classification obtained from a high value of α should be used as the benchmark to evaluate the test classification (see Fig. 7); if a looser criterion is needed, then a classification obtained with a lower value of α can be chosen. In this research, we chose α = 0.792, 0.700, 0.656, and 0.594 to represent four different levels of strictness in judging similarity and used the two indices described in previous sections to evaluate the utility of the test classification. The results of our calculations are shown in Table 4. When $\alpha = 0.594$ there are 8 groups in the benchmark classification, and when $\alpha = 0.656$, there are 16 groups. In both cases, there are far fewer groups than the 24 identified by the test classification. The index of the appropriate number of groups is only 0.43 and 0.60, respectively, for these two cases, which means that the test classification applies a stricter criterion of similarity within groups than the benchmark classification does. The index of homogeneity within each group is 0.42 and 0.56, respectively, which indicates that the similarity criteria applied by the benchmark classification are very different to those applied by the test classification. When $\alpha = 0.700$, on the other hand, there are 23 groups in the benchmark classification, which is almost the same as the number in the test classification. Thus, in this case the criteria of similarity applied by the benchmark classification and the test classification system in clustering the benchmarks are equally strict. However, the index of homogeneity within each group is only 0.64, indicating that the specific criteria applied by the two systems are very different. This may be because the test classification employs three 2D orthogonal drawings for benchmark clustering first and then aggregates the results. Using aggregated information in this way is more difficult than directly comparing global shape information and then clustering the workpieces. When $\alpha = 0.792$, the benchmark system produces 29 groups, and the index of the appropriate number of groups is 0.82, which means that the test classification is now applying a looser criterion of similarity than the benchmark classification. The index of homogeneity within each group is 0.81. This value is higher than for the other values of α because the benchmark system now divides the workpieces into 29 groups. Since many of these groups now include only one workpiece, the index of homogeneity within each workpiece group will tend to be higher.

Table 4. Comparison of results of benchmark classification and automatic classification.

Classification type		Workpiece clustering	Number of groups	Appropriate number of groups (r)	Homogeneity within each group (h)
Benchmark classification (denotes the user's judgement of workpiece similarity)	$\alpha = 0.792$	1 [2 7 8] [3 4] 5 6 [9 10] 11 [12 34] 13 14 15 16 [17 18] 19 20 [21 22] 23 24 25 26 27 28 29 30 31 32 33 35 36	29	0.82	0.81
	0.700	1 [2 3 4 5 8] 5 6 [9 10 11] [12 34] 13 14 15 [16 20 21 22] [17 18] [19 24 36] 23 25 26 27 28 29 30 31 32 33 35	23	0.92	0.64
	0.656	1 [2 3 4 7 8 9 10 11] [5 14] 6 [12 28 34] [13 19 24 36] [15 33] [16 17 18 20 21 22] 23 25 26 [27 35] 29 30 31 32	16	0.60	0.56
	0.594	[1 2 3 4 7 8 9 10 11 15 23 27 31 33 35] [5 14] [6 12 16 17 18 20 21 22 28 34] [13 19 24 36] 25 [26 32] 29 30	8	0.43	0.42
Automatic classification system	Wu and Jen [8]	[2 4] 3 [7 8 10 30] [12 13 34] [1 9 17 22] 6 33 [11 27] 5 14 15 16 [18 21] 19 20 [23 29] 24 25 26 28 31 32 35 36	24		
Remarks		1. [ ] denotes groups of workpieces 2. Workpieces not included within [ ] each constitute their own group			

4. Discussion and Conclusion

In this research, a systematic method was developed to evaluate the utility of automated workpiece classification systems. This method employs benchmark classifications based on information concerning the global shape of workpieces. An empirical example was presented to demonstrate the application of this method.

Two salient characteristics of the proposed evaluation method are that it is based on users' judgements of the degree of similarity between samples and that it allows flexible similarity criteria to be applied. The benchmark classification system is based on the judgements of actual users concerning the similarity of various workpieces, so it enables us to examine the degree to which a particular automated classification system is consistent with the judgements of human experts. Moreover, since a range of stricter or looser classification criteria can be applied, the method allows us to evaluate the utility of an automated classification system in terms of specific user requirements, instead of in comparison to an arbitrary standard. The proposed evaluation method provides a convenient means of evaluating and choosing an automated system for workpiece classification, and can help users to generate workpiece classification databases based on different degrees of similarity. The flexibility of such databases and their consistency with expert judgements makes them robust practical tools for information retrieval and other applications.

In this research, a considerable amount of time was spent collecting information on human subjects' judgements of the degree of similarity between various workpieces in order to establish the benchmark classification system. Subjects were asked to make a very large number of similarity comparisons, which may have led to a certain degree of bias or error in their judgements owing to fatigue. In order to resolve this difficulty, all procedures in the human comparison stage must be strictly controlled. In addition, we plan to develop a systematic method that will reduce the number of comparisons needed to produce a benchmark classification system while maintaining the accuracy of the classification. In this new method, subjects will need to compare sample workpieces with only a few typical workpieces, instead of exhaustively comparing every pair of workpieces in the sample. By applying mathematical deduction to this partial comparison information, we hope to obtain classification results approximately equivalent to those obtained from exhaustive comparisons of every sample in the set.

Acknowledgement

The authors thank the employees of the Aircraft Manufactory, Aerospace Industrial Development Corporation (AIDC), Taichung, Taiwan, ROC, for their participation in the workpiece pair comparison experiment.

References

1. A. Bhadra and G. W. Fischer, "A new GT classification approach: a data base with graphical dimensions", *Manufacturing Review*, **11**, pp. 44–49, 1988.
2. C. S. Chen, "A form feature oriented coding scheme", *Computers and Industrial Engineering*, **17**, pp. 227–233, 1989.
3. M. R. Henderson and S. Musti, "Automated group technology part coding from a three-dimensional CAD data-base", *Transactions of the ASME Journal of Engineering for Industry*, **110**, pp. 278–287, 1988.
4. S. Kaparthi and N. Suresh, "A neural network system for shape-based classification and coding of rotational parts", *International Journal of Production Research*, **29**, pp. 1771–1784, 1991.
5. S. R. Jen, "A neural network approach to the classification of 3D prismatic parts", Master Thesis, National Chiao Tung University, Hsinchu, Taiwan, 1995.
6. T. Lenau and L. Mu, "Features in integrated modeling of products and their production", *International Journal of Computer Integrated Manufacturing*, **6**(1&2), pp. 65–73, 1993.

7. M. C. Wu, J. R. Chen and S. R. Jen, "Global shape information modeling and classification of 2D workpieces", *International Journal of Computer Integrated Manufacturing*, **7**(5), pp. 216–275, 1994.
8. M. C. Wu and S. R. Jen, "A neural network approach to the classification of 3D prismatic parts", *International Journal of Advanced Manufacturing Technology*, **11**, pp. 325–335, 1996.
9. L. A. Cooper, "Mental representation of three-dimensional objects in visual problem solving and recognition", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, pp. 1097–1106, 1990.
10. R. Arnheim, *The Power of Center*, Berkeley and Los Angeles: University of California Press, 1988.
11. S. J. Chen and C. L. Hwang, *Fuzzy Multiple Attribute Decision Making – Method and Application, A State-of-the-Art Survey*, Springer, New York, 1992.
12. J. J. Buckley, "The multiple judge, multiple criteria ranking problem: a fuzzy set approach", *Fuzzy Sets and Systems*, **13**, pp. 25–37, 1984.
13. S. Murakami, H. Maeda and S. Imamura, "Fuzzy decision analysis on the development of centralized regional energy control system", Preprints IFAC Conference on *Fuzzy Information, Knowledge Representation and Decision Analysis*, pp. 353–358, 1983.
14. F. Bouslama and A. Ichikawa, "Fuzzy control rules and their natural control laws", *Fuzzy Sets and Systems*, **48**, pp. 65–86, 1992.
15. S. K. Tan, H. H. Teh and P. Z. Wang, "Sequential representation of fuzzy similarity relations", *Fuzzy Sets and Systems*, **67**, pp. 181–189, 1994.
16. G. J. Klir and T. A. Folger, *Fuzzy Set, Uncertainty, and Information*, Prentice-Hall, 1992.
17. L. A. Zadeh, "Fuzzy Sets", *Information and Control*, **8**, pp. 338–353, 1965.
18. H. J. Zimmermann, *Fuzzy Set Theory: and Its Application, 2nd edn*, Kluwer Academic Publishers, 1991.
19. R. Jan, "Decision-making in the presence of fuzzy variables", *IEEE Trans. Systems Man Cybernetics*, **6**, pp 698–703, 1976.
20. D. Dubois and H. Prade, "Operations on fuzzy numbers", *International Journal of System Sciences*, **9**, pp 613–626, 1978.
21. A. Kaufmann and M. M. Gupta, *Introduction to Fuzzy Arithmetic: Theory and Applications*, Von Nostrand Reinhold, New York, 1985.
22. L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning", Part 1, 2 and 3, *Information Science*, **8**, pp. 199–249, 301–357, 1975; **9**, pp. 43–58, 1976.

Appendix

Four major concepts from fuzzy set theory were used in this research: fuzzy sets, fuzzy numbers, linguistic variables, and the α-cut of the membership matrix.

A1. Fuzzy Sets

Fuzzy set theory was introduced by Zadeh [17]. It can be used to deal with problems in which a source of vagueness is present. It can be considered a modelling language that approximates situations in which fuzzy phenomena and criteria exist. Consider a reference set X with x as its element. A fuzzy subset $\tilde{A}$ of x is defined by a membership function $\mu_{\tilde{A}}(x)$ which maps each element x in X to a real number in the interval [0,1]. The value of $\alpha_{\tilde{A}}(x)$ denotes the grade of membership, that is, the degree to which element x is a member of set $\tilde{A}$. A fuzzy subset is often referred to briefly as a fuzzy set [18].

A2. Fuzzy Numbers

A fuzzy number is a special fuzzy subset in R (real line) which is usually represented by a special membership function over a closed interval of real numbers. In this research, a special class of fuzzy numbers known as trapezoidal fuzzy numbers (TrFN) developed by Jain [19] and Dubois and Prade [20] was used. As shown in Fig. 8, a TrFN has a trapezoidal shape and can be denoted by (a,b,c,d), where a can be semantically interpreted as the lower bound, b and c are the most probable value, and d is the upper bound, with the membership function defined as follows:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & (x \leq a) \\ \dfrac{x-a}{b-a} & (a < x < b) \\ 1 & (b \leq x \leq c) \\ \dfrac{x-d}{c-d} & (c < x < d) \\ 0 & (x \geq d) \end{cases} \tag{9}$$

By the extension principle proposed by Zadeh [17], the addition and subtraction operations on TrFNs definitely yield a TrFN. Multiplication, inverse, and division operations on TrFNs do not necessarily yield a TrFN. However, the results of these operations can be reasonably approximated by TrFNs [21], as illustrated below:

Addition $\oplus$

$$\tilde{A}_1 \oplus \tilde{A}_2 = (a_1,b_1,c_1,d_1) \oplus (a_2,b_2,c_2,d_2) = (a_1+a_2, b_1+b_2, c_1+c_2, d_1+d_2) \tag{10}$$

Subtraction $\ominus$

$$\tilde{A}_1 \ominus \tilde{A}_2 = (a_1,b_1,c_1,d_1) \ominus (a_2,b_2,c_2,d_2) = (a_1-d_2, b_1-c_2, c_1-b_2, d_1-a_2) \tag{11}$$

Multiplication $\otimes$

$$k \otimes \tilde{A}_1 = k \otimes (a_1,b_1,c_1,d_1) = (ka_1, kb_1, kc_1, kd_1) \quad \text{if } k \geqq 0 \tag{12}$$

$$\tilde{A}_1 \otimes \tilde{A}_2 = (a_1,b_1,c_1,d_1) \otimes (a_2,b_2,c_2,d_2) \cong (a_1a_2, b_1b_2, c_1c_2, d_1d_2) \quad \text{if } a_1 \geqq 0, \quad a_2 \geqq 0 \tag{13}$$

Division $\oslash$

$$\tilde{A}_1 \oslash \tilde{A}_2 = (a_1,b_1,c_1,d_1) \oslash (a_2,b_2,c_2,d_2) \cong (a_1/d_2, b_1/c_2, c_1/b_2, d_1/a_2) \quad \text{if } a_1 \geqq 0, \quad a_2 > 0 \tag{14}$$

In the special case when $b=c$, the trapezoidal fuzzy number (a, b, c, d) equals a triangular fuzzy number. The extended algebraic operations on triangular fuzzy numbers are the same as those on trapezoidal numbers.

A3. Linguistic Variables

Linguistic variables are variables whose values are represented in words or sentences in natural languages. Each linguistic value can be modelled by a fuzzy set [22]. For example, let $\tilde{S}$ be a linguistic variable with the name "similarity" (the pair comparison between any two workpieces), and let the set of its linguistic terms be {very low

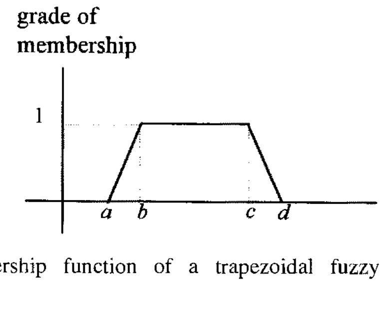


Fig. 8. Membership function of a trapezoidal fuzzy number $\tilde{A} = (a,b,c,d)$.

similarity, low similarity, medium similarity, high similarity, very high similarity}. Each of these linguistic terms can be represented by a TrFN with its membership functions, as shown in Fig. 9. Note that these linguistic terms are trapezoidal fuzzy numbers in the interval [0,1].

From the figure, the membership function of "low" is (0.1, 0.25, 0.25, 0.4). That is, an expression of "low similarity" between two workpieces is between 0.1 and 0.4, and the most probable value is 0.25. The membership function for the degree of similarity at 0.2 in low is 0.67, while for the degree of similarity at 0.35 in low it is 0.33. Linguistic variables are useful for allowing experts to express uncertain judgements, such as those concerning the workpiece pair comparisons.

A4. α-cut of membership matrix

Let X be the universe of discourse, $X = \{x_1, x_2, \cdots, x_n\}$. Suppose that $\tilde{A}$ is a fuzzy number with membership function $\mu_{\tilde{A}}$. Then for every $\alpha \in [0,1]$, the set $\tilde{R}^{\alpha} = \{x | \mu_{\tilde{A}}(x) \geqq \alpha\}$ is called an α-cut of $\tilde{A}$.

Similarly, suppose $\tilde{R}$ is a fuzzy relation, $\tilde{R} = [r_{ij}]_{m \times n}$ for every $\alpha \in [0,1]$, $\tilde{R}^{\alpha} = [r_{ij}{}^{\alpha}]$ is called an α-cut of the membership matrix

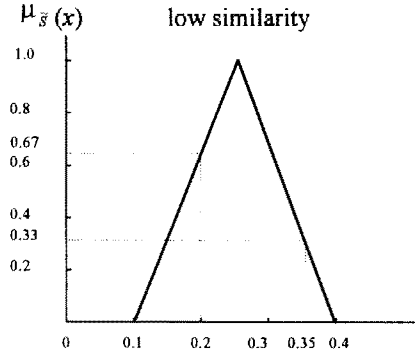


Fig. 9. The membership function for linguistic term "low similarity".

where

$$r_{ij}{}^{\alpha} = \begin{cases} 1 & \text{if } r_{ij} \geqq \alpha \\ 0 & \text{if } r_{ij} < \alpha \end{cases} \tag{15}$$