A case-based reasoning system for interior design using a new cosine similarity retrieval algorithm

Kuo-Sui Lin

Department of Information Management, Aletheia University, New Taipei City, Taiwan, R.O.C.

ABSTRACT

During concept design stage, a case-based reasoning (CBR) system is good for searching the most suitable interior design drawing case as a preliminary concept drawing design aid to bridge the design requirements gap between the consigner and the designer. In the case base of the system, features of the pre-stored design drawing cases could be characterized by interval, nominal, ordinal and/or ratio scale data. However, the cosine similarity measure used for case reasoning scarcely deals with mixed measurement scales of interval, nominal, ordinal and ratio scale types. Therefore, the objective of this study has been twofold. The first goal was to propose a new cosine similarity retrieval algorithm for case reasoning, in which mixed measurement scales are considered and applied. The second goal was to propose a CBR system and apply the new cosine similarity retrieval algorithm to the search engine of the CBR system for interior design drawing retrieval. Finally, numerical experiments were carried out to demonstrate the capability and effectiveness of the proposed similarity matching algorithm on the case-based reasoning system.

ARTICLE HISTORY

Received 15 March 2017
Accepted 30 November 2019

KEYWORDS

Case-based reasoning; case-based design; cosine similarity measure; interior design

1. Introduction

Case-Based Design (CBD) is the application of Case-Based Reasoning (CBR) to develop new design solutions by adapting previous solved solutions. A CBR is especially suitable for domains with a weak domain theory, i.e. when the domain is difficult to formalize and is empirical (Begum, Barua, Filla, & Ahmed, 2014). A CBR system can start working with a few reference cases in its case base and then learn day by day by gradually adding new solved cases into the case base. The domain of interior design exhibits both the ill-defined and complex design generation problems, thus, application of case based design systems in this domain is of increasing necessary. In our previous study (Lin & Ke, 2015), the authors proposed a virtue reality based recommender system to retrieve an interior design drawing case from a pre-stored drawings case base. The retrieved design drawing case can be treated as a prototype drawing for developing a new design solution. A binary bit measurement scale is applied in that study and the cosine similarity measure function is used for binary string pattern matching between two
sequences of 0’s and 1’s in the retrieval engine of the recommender system, 0 for absence of a specific design feature and 1 for presence of a specific design feature. In further practical similarity measurement, the author found that features of the pre-stored design drawing cases could be characterized by interval, nominal, ordinal and/or ratio scale data. Since cosine similarity measure is the most commonly used similarity measure when dealing with numerical features, this measure fails as the Euclidean distance is not defined for features of ordinal, nominal or mixed scale types. No studies is reported in the literature on the use of cosine similarity measure function in developing a search engine for dealing with mixed measurement scales of interval, nominal, ordinal and ratio scale types at the same time. Therefore, the objective of this study has been twofold. The first goal was to propose a new cosine similarity retrieval algorithm for case reasoning, in which mixed measurement scales are considered. The second goal was to propose a CBR system and apply the new cosine similarity retrieval algorithm to the search engine of the CBR system for interior design drawing case retrieval.

2. Related works

2.1. Case-Based Reasoning

Case-Based Reasoning (CBR) is a problem solving system that uses past experiences as knowledge used to solve similar new problems (Aamodt & Plaza, 1994; Weenawadee & Sarun, 2013). Case-Based reasoning deals with a new problem by first retrieving a past similar case, and then reusing and adapting that case for solving the new problem. For a CBR System, the four main phases of the CBR 4 R cycle are iteratively performed: retrieval, reuse, revision and retention. The knowledge of previous experience is represented as a case that consists of problem descriptions and solutions. Several cases are collected into the case base. Domain experts store previous experiences as cases in case base. Typically, each case contains a description of the problem, plus a solution and/or the outcome. A new case or unsolved case is the description of a new problem to be solved. When a new problem comes, it becomes an input to the system. In the retrieval phase, cases that are to be retrieved from the case database will be matched with the new case based on several features. The most similar case to the target problem from the case database is retrieved. In the reuse phase, a solution for the new case from the available solutions of the cases that were retrieved from the case database is suggested and reused to solve the new problem. In the revise phase, if there is a very similar past case whose solution needs little adjustment to the query, revise or adapt the proposed solution for the new problem if necessary. If the new case is judged to be possibly revised/adapted for future problem solving in the next retrieval cycle, then it can be retained in the retain phase. In the retain phase, the revised/adapted new problem and its solution/experience is kept as a new case embodied in the case base to use in future problem solving.

2.2. Cosine similarity measure for fuzzy sets

A commonly used similarity measure for the retrieval engine is cosine similarity (Larsen & Aone, 1999; Nahm, Bilenko, & Mooney, 2002). To deal with vague and uncertain human knowledge, the fuzzy set based information retrieval methods are increasingly recognized
as more realistic than the classical set based information retrieval methods (Baeza-Yates & Ribeiro-Neto, 2011; Cross, 1994; Donald & Erin Colvin, 2017; Zadrożny & Nowacka, 2009).

A cosine similarity measure for fuzzy sets (Salton & McGill, 1983) is defined as the inner product of two vectors divided by the product of their lengths. Let \( A = (\mu_A(x_1), \mu_A(x_2), \ldots, \mu_A(x_i), \ldots, \mu_A(x_n)) \) and \( B = (\mu_B(x_1), \mu_B(x_2), \ldots, \mu_B(x_i), \ldots, \mu_B(x_n)) \) be two fuzzy sets in the universe of discourse \( X = \{x_1, x_2, \ldots, x_n\} \). The cosine similarity measure (angular coefficient) between \( A \) and \( B \) can be defined by the following definition:

**Definition 2.1:** Cosine Similarity Measure for Fuzzy Sets (Salton & McGill, 1983; Ye, 2011):

\[
\cosineSimilarity(A, B) = \cosSim(A, B) = \frac{\text{Dot}(A, B)}{||A||||B|||} = \frac{\sum_{k=1}^{i} (\mu_A(x_k) \ast \mu_B(x_k)) / \left( \sqrt{\sum_{k=1}^{i} \mu_A(x_k)^2} \right)}{\sqrt{\sum_{k=1}^{i} \mu_B(x_k)^2}} = \frac{\mu_A(x_1) \ast \mu_B(x_1) + \mu_A(x_2) \ast \mu_B(x_2) + \mu_A(x_3) \ast \mu_B(x_3) + \ldots}{\sqrt{(\mu_A(x_1)^2 + \mu_A(x_2)^2 + \mu_A(x_3)^2 + \ldots)}}
\]

where \( 0 \leq \cosSim(A, B) \leq 1 \).

For improving the performance efficiency of the retrieval engine, based on the cosine similarity measure, the author proposed a new cosine similarity matching algorithm, which will be described in the later section.

### 2.3. Interior design

Design is a problem solving process. Interior design plays a major role in the quality of building design. The layout and space division, type and colour of coated surfaces (floors, ceilings, and walls), lighting, furniture and other important elements are influential factors that determine the beauty and work efficiency of interior spaces (Azad, 2015). An interior design prototype drawing is better to be created earlier during concept design stage; thereafter, any adjustment, correction, or addition to the design prototype drawing can further be developed and implemented, bridging the design requirements gap between the consigner and the designer. ‘Seeing design as redesign’ has long been a popular cliché in interior design industry. There is no interior design that starts from the very original. Thus, interior design can be seen as an improvement process and can follow a continuous redesign of the base design. Normally, interior designers draw upon their past experiences and expertise to solve new design problems. They always work on design solutions that are already at hand and which had existed as a case. Interior design cases are reusable solutions that are used as an effective tool for solving the recurring interior design problems. Interior design cases provide efficient design by not having to spend much time to find the solutions to problems that have already been solved.

### 3. New fuzzy similarity retrieval algorithm

The most important part of a CBR system is the retrieval engine. Thus, improving the performance efficiency of the retrieval engine is critical for the CBR system (Leake, 1996). A commonly used similarity measure for the retrieval engine is cosine similarity (Larsen &
Aone, 1999; Nahm et al., 2002). For improving the performance efficiency of the retrieval engine, the author proposes a new cosine similarity retrieval algorithm, which is described as follows.

Step 1: Translating Various Measurement Scales into Fuzzy Measurement Scales
From a practical perspective, four different formats of measurement scales are considered for interior design drawing case retrieval in a CBR system. Thus, we need to translate the four traditional measurement scales into fuzzy measurement scales, including nominal, ordinal, interval and ratio measurement scales. The pre-stored cases and query case which are rated by various numerical or categorical scales will then be translated into their corresponding fuzzy linguistic terms. Then, the translated fuzzy linguistic terms will be transformed into their corresponding membership functions. By referring to several types of triangular fuzzy numbers in linguistic terms (Chen & Ku, 2008), the author parameterizes these linguistic terms with Triangular Fuzzy Numbers.

Step 2: Computing Local Similarity Measure
The local similarity is the similarity measure on each feature of the new query vector \( Q \) and the stored case vector \( S_i \). Let \( \text{sim}(r_{qj}, r_{ij}) \) denote the local similarities (feature similarities) between the stored case vector \( S_i \) and the new query vector \( Q \) with regard to feature \( F_j \). \( r_{qj} \) and \( r_{ij} \) is the rating of \( F_j \) in the query case \( Q \) and the historical case \( S_i \) respectively. Depending on different measurement scales, the local similarity can be calculated as follows.

1) Cosine Similarity Measure for Fuzzy Nominal Scale
   A nominal scale is a scale with no order in rank and an ordinal scale is a scale with an order in rank. Let \( \text{sim}(r_{qj}, r_{ij}) \) denote the feature similarity measure between the stored case \( S_i \) and the query case \( Q \) with regard to the feature \( F_j \). If the feature \( F_j \) is a nominal variable, the local similarity between two \( j \)-th feature ratings \( r_{qj} \) and \( r_{ij} \) is given by:
   \[
   \text{sim}(r_{qj}, r_{ij}) = 1, \text{ if } r_{qj} = r_{ij}; \text{sim}(r_{qj}, r_{ij}) = 0, \text{ if } r_{qj} \neq r_{ij}.
   \]
   It assigns value 1 if the two features match and value 0 otherwise.

2) Cosine Similarity Measure for Fuzzy Ordinal, Interval and Ratio Scale
   The ordinal, interval and ratio scale can be translated into corresponding triangular fuzzy number. Thus, a cosine similarity measure for triangular fuzzy numbers is proposed in an analogous manner to the cosine similarity measure (angular coefficient) between fuzzy sets (Salton & McGill, 1983; Ye, 2011). Let \( A = (\mu_A(x_1), \mu_A(x_2), \mu_A(x_3)) \) be a triangular fuzzy number in the set of real numbers \( R \), the three parameters in \( A \) can be considered as a vector representation with the three elements. Assume that there are two triangular fuzzy numbers \( A = (\mu_A(x_1), \mu_A(x_2), \mu_A(x_3)) \) and \( B = (\mu_B(x_1), \mu_B(x_2), \mu_B(x_3)) \) in the set of real numbers \( R \). Based on the extension of the cosine similarity measure for fuzzy sets, a cosine similarity measure between \( A \) and \( B \) is proposed as follows:
   \[
   \text{sim}(A, B) = \frac{\text{Dot}(A, B)}{||A|| \cdot ||B||} = \frac{A \cdot B}{||A|| \cdot ||B||} = \sum_{k=1}^{3} (\mu_A(x_k) \ast \mu_B(x_k)) / \sqrt{\left( \sum_{k=1}^{3} \mu_A(x_k)^2 \right) \times \left( \sum_{k=1}^{3} \mu_B(x_k)^2 \right)}
   = (\mu_A(x_1) \ast \mu_B(x_1) + \mu_A(x_2) \ast \mu_B(x_2) + \mu_A(x_3) \ast \mu_B(x_3)) / \sqrt{\left( \mu_A(x_1)^2 + \mu_A(x_2)^2 + \mu_A(x_3)^2 \right) \times \left( \mu_B(x_1)^2 + \mu_B(x_2)^2 + \mu_B(x_3)^2 \right)}
   \]
Step 3: Computing Global Similarity Measure

To retrieve proper pre-stored case, it is necessary to measure the global similarities. Thus, the local similarities (feature similarities) can be aggregated into a global similarity. A global similarity measure is a normalized similarity measure resulting from the summation of the individual similarity measure multiplied by the weight factors. The global similarity measure $Sim(Q, S_i)$ aggregates the local similarity measures $sim(r_{qj}, r_{ij})$ into one similarity value. Let $Sim(Q, S_i)$ denote the global similarity between the pre-stored case $S_i$ and the query case $Q$, then a global similarity measure can be derived by the weighted summation of the local similarity matching measures: $Sim(Q, S_i) = (\sum_{j=1}^n w_j \times sim(r_{qj}, r_{ij}))/\sum_{j=1}^n w_j$, where $w_j$ are the local weights defined by domain experts to reflect importance of the corresponding feature, $n$ is the number of features in each case, and $sim(r_{qj}, r_{ij})$ is the local similarity measure for individual feature $F_j$ in query case $Q$ and stored case $S_i$.

4. Architecture of the proposed CBR system

After proposed the cosine similarity retrieval algorithm, the author further developed a virtual reality based CBR system for interior design drawing case retrieval. As shown in Figure 1, the proposed CBR system consists of a case base, a virtual reality based display module, a query module and a retrieval engine. The case base stores a group of historical interior design drawings drawn by major interior design software tools. The retrieval engine calculates the similarity of the cases in the case base to the query using the new similarity matching model and retrieves the case which is most suitable to the user’ request and match his design preferences. The case retrieval phase is typically the main step of the CBR cycle and the majority of CBR systems can be described as sophisticated retrieval engines. The query module generates the preference query from the designer and transfers the designer query into a query profile. The display module

![Figure 1. The architecture of the proposed CBR system (modified after Aamodt & Plaza, 1994).](image-url)
provides a virtual reality based platform for displaying retrieved design drawings by using FancyDesigner® interior design drawing tool (http://www.fancydesigner.com.tw/). A designer can requests a solution pre-stored as a case, a similarity measure is computed based on the current problem statement and pre-stored cases with the associated solution to support retrieval.

5. Numerical experiments of the proposed CBR system

When a designer needs to find an interior design drawing case that matches the consigner’s design requirements, the user inputs some preference ratings about the design features of the design drawing case through a questionnaire query module. The designer’s query is then taken to retrieve a most suitable case and displayed to the consigner through a display module. If the consigner is not satisfied with the retrieved case, the designer can modify and adapt the case to form a new case to satisfy the current query and store the new case to the case base.

5.1. Experimental setup

To show the effectiveness of the proposed cosine similarity retrieval algorithm, the author applies the proposed algorithm to the proposed CBR system for searching interior design drawing case. The numerical experiments were carried out following the following steps:

Step 1: Collecting Old Cases to the Case Base

In CBR systems, cases have been regarded as an important source of knowledge. A design case is a design solution that can be applied to a design problem (query). Designers’ expertise is embodied in a collection of past design cases (seed cases) pre-stored in the case base. The information used to search the case base for most matched cases is a design problem (query). Using the information of a given design query, a user can browse throughout the case base and search the most suited case for reuse.

Step 2: Extracting Features of Cases

Case indexing is essential to assign features to cases for use in the retrieval process. The extracted features are applied to represent the characteristics of indexes that are used to identify case problem descriptions. Let \( n \) be the number of extracted feature set \( F = \{F_1, F_2, \ldots, F_n\} \) and \( F_j \) be the \( j \)-th feature \( (j = 1,2,\ldots,n) \). These features will be used for case based reasoning in the CBR system. In most situations, cases are represented as feature-value pairs. It is a group of related records of data set. Each record is a group of related design features of drawings. A feature is a distinctive property (or characteristic, attribute) used to distinguish two objects. Feature values are ratings of features that are provided by the users. An individual case \( S_i \) is assumed to be comprised of a finite list of feature-value pairs: \( S_i = \{(F_1, r_{i1}), (F_2, r_{i2}), \ldots, (F_j, r_{ij}), \ldots, (F_n, r_{in})\} \). Let the case base, \( CB \), be a finite set of individual cases, \( CB = \{S_1, S_2, \ldots, S_m\} \) and \( S_j \) be the \( i \)-th pre-stored case \( (i = 1,2,\ldots,m) \), where \( m \) is the number of cases. From careful analysis, a set of ten extracted design features are used for a case base and are developed for classifying the interior design drawings. The ten features are: \( F = \{F_1, F_2, \ldots, F_{10}\} = \{\text{Budget, Working Duration, Design Style, Design Theme, Design Hue, Brightness, Saturation, Colour temperature, Day-lighting source, Lighting type}\}. \)
Step 3: Establishing Case Profiles and Query Profile

In this study, a case profile is a collection of records representing important features and descriptions of the case. There are two classes of entities in the CBR system: a pre-stored case profile consisting of pre-specified feature-value pairs and a query profile consisting of user’s preferences feature-value pairs. These feature-value pair entries are represented as a utility matrix. With profile vectors for both query profile and past cases profiles, we can estimate the degree to which a user would prefer a case by computing the similarity between the user’s and a case’s vectors. Both feature ratings of the stored cases and the user’s query case are transformed into a set of old case-feature-rating vectors (case profiles) and user’s query-feature-rating vector (query profile). A set of old case-feature-rating vectors is represented as a utility matrix.

(1) Collecting case-feature rating matrix and query-feature rating vector

The profiles of pre-stored cases can be represented by a case–feature-rating matrix. Each row in the matrix is the vector representation of a design case. As shown in Figure 2, case-feature-rating entries usually are described as an \(m \times n\) ratings matrix \(R_{mn}\), where each entry \(r_{ij} (1 \leq i \leq m, 1 \leq j \leq n)\) means the rating assigned to the case \(i\) on the feature \(j\); the row represents \(m\) cases and the column represents \(n\) features.

The designer’s preference query can be described as the following feature-value pairs: \(Q = \{(F_1, r_{q1}), (F_2, r_{q2}), \ldots, (F_j, r_{qj}), \ldots, (F_n, r_{qn})\}\). Table 1 shows a set of extracted features, a set of collected content profiles in the case base and a query profile of a problem description. As shown in Table 1, example of values of the case–feature-rating matrix and query-feature-rating vector can be nominal, ordinal, interval or ratio scale.

(2) Translating traditional measurement scales into fuzzy measurement scales

A linguistic variable representing natural language expression whose value is not numerical value. Instead, a linguistic variable can be decomposed into a set of linguistic terms and quantified by different membership functions. Based on the available a priori information or the phenomenon intending to describe, a membership function may take many forms. Typical shapes for membership functions include triangular, trapezoid and Gaussian membership functions. The shape of each membership function used in this study is triangular due to its popularity in specifying fuzzy sets (Nahm et al., 2002). Because we have more knowledge about Budget and Working Duration than other rational features in this study, the author defined five linguistic terms. The semantics of the linguistic terms are given using fuzzy numbers. By referring to Chen and Ku (2008), the author further

\[
\begin{pmatrix}
F_1 & F_2 & \ldots & F_j & \ldots & F_n \\
W & \omega_1 & \omega_2 & \omega_j & \omega_n \\
S_1 & r_{11} & r_{12} & \ldots & r_{ij} & \ldots & r_{1n} \\
S_2 & r_{21} & r_{22} & \ldots & r_{2j} & \ldots & r_{2n} \\
S_m & r_{m1} & r_{m2} & \ldots & r_{mj} & \ldots & r_{mn}
\end{pmatrix}
\]

Figure 2. A case–feature rating matrix.


| Case | $F_1$ Budget (S) Ratio | $F_2$ Working Duration (Days) Ratio | $F_3$ Design Style Nominal | $F_4$ Design Theme Nominal | $F_5$ Colour Hue Nominal | $F_6$ Brightness Ordinal | $F_7$ Saturation Ordinal | $F_8$ Colour Temperature Ordinal | $F_9$ Daylighting Source Nominal | $F_{10}$ Lighting Type Nominal |
|------|------------------------|-----------------------------------|-----------------------------|-----------------------------|---------------------------|------------------------|------------------------|-----------------------------|-----------------------------|-----------------------------|
| 1    | 10,000                 | 12                                | 100000                      | 00100000                    | 000010                    | B                      | M                      | W                          | H                          | 010                         |
| 2    | 20,000                 | 14                                | 001000                      | 00100000                    | 000010                    | M                      | M                      | M                          | M                          | 100                         |
| 3    | 20,000                 | 14                                | 001000                      | 00001000                    | 000010                    | D                      | D                      | C                          | L                          | 001                         |
| 4    | 60,000                 | 21                                | 000010                      | 00001000                    | 000010                    | B                      | I                      | W                          | M                          | 001                         |
| 5    | 50,000                 | 21                                | 000010                      | 00100000                    | 000010                    | B                      | M                      | W                          | H                          | 100                         |
| 6    | 80,000                 | 42                                | 010000                      | 00100000                    | 000010                    | D                      | D                      | C                          | L                          | 100                         |
| Query| 30,000                 | 20                                | 010000                      | 00100000                    | 000010                    | B                      | I                      | M                          | H                          | 001                         |
translated the linguistic variables and terms into five triangular fuzzy numbers, ranging between zero and one (Tables 2 and 3 and Figure 3).

For translating feature of Design Style, Design Theme and Colour Hue, the author uses binary bit string of nominal scale for measurement (Tables 4–6).

For translating feature of Brightness, Saturation, Colour Temperature and Daylighting Source, we categorized three linguistic terms of ordinal scale and the range is defined between zero and one (Table 7 and Figure 4).

For translating feature of Lighting Type, the author uses binary string of nominal measurement scale (Table 8). Lighting Type is classified as task, accent, or general lighting, depending largely on the distribution of the light produced by the fixture.

Example of calculation sheet for translating measurement scales into linguistic terms is shown in Table 9.

**Step 4: Computing Similarity Measures**

(1) **Computing local similarity measures on features between a case and a query**

**Table 2.** Linguistic variables and terms of budget.

| Description (Days) | Linguistic Term  | Triangular Fuzzy Number |
|--------------------|------------------|-------------------------|
| Budget ≤10         | Very Low (VL)    | TFN(0, 0, 0.25)         |
| 11 ≤ Budget ≤ 30   | Low (L)          | TFN(0, 0.25, 0.5)       |
| 31 ≤ Budget ≤ 50   | Medium (M)       | TFN(0.25, 0.5, 0.75)    |
| 51 ≤ Budget ≤ 70   | High (H)         | TFN(0.5, 0.75, 1)       |
| 71 ≤ Budget        | Very High (VH)   | TFN(0.75, 1, 1)         |

**Table 3.** Linguistic variables and terms of working duration.

| Description (Calendar Days) | Linguistic Term   | Triangular Fuzzy Number |
|-----------------------------|-------------------|-------------------------|
| Duration ≤ 10               | Very Short (VS)   | TFN(0, 0, 0.25)         |
| 11 ≤ Duration ≤ 20          | Short (S)         | TFN(0, 0.25, 0.5)       |
| 21 ≤ Duration ≤ 30          | Medium (M)        | TFN(0.25, 0.5, 0.75)    |
| 31 ≤ Duration ≤ 40          | Long (L)          | TFN(0.5, 0.75, 1)       |
| 41 ≤ Duration               | Very Long (VL)    | TFN(0.75, 1, 1)         |

**Figure 3.** Membership functions for features with nominal scales.

**Table 4.** Binary string of Design Style.

| Linguistic Term     | Binary string |
|---------------------|---------------|
| Royal               | 100000        |
| Zen                 | 010000        |
| Nautical/Coastal    | 001000        |
| High Tech           | 000100        |
| Modern              | 000010        |
| Country             | 000001        |
The local similarity calculates the similarity for each feature $j$ of the case $S_i$ and the query $Q$. By applying the local similarity measure $\text{sim}(r_{ij}, r_{ij})$, the similarity on feature $j$ between stored case $S_i$ in the case base and the new query $Q$ can be calculated. Taking query $Q$ and case $S_6$ for examples, calculation sheet for translated fuzzy number and calculated local similarity measures and global similarity measure is shown in Table 10.

### Table 5. Binary string of Design Theme.

| Linguistic Term | Binary String |
|-----------------|---------------|
| Greece          | 10000000      |
| Middle East     | 01000000      |
| European        | 00100000      |
| American        | 00010000      |
| Taiwanese       | 00001000      |
| Chinese         | 00000100      |
| Japanese        | 00000010      |
| Korean          | 00000001      |

### Table 6. Binary string of Colour Hue.

| Linguistic Term | Binary String |
|-----------------|---------------|
| Red             | 100000        |
| Yellow          | 010000        |
| Green           | 001000        |
| Cyan            | 000100        |
| Blue            | 000010        |
| Magenta         | 000001        |

### Table 7. Brightness, saturation, colour temperature and daylighting source.

| Description   | Linguistic Term | Linguistic Term | Linguistic Term | Linguistic Term | Triangular Fuzzy Number |
|---------------|-----------------|-----------------|-----------------|-----------------|--------------------------|
| ordinal #1    | Dark            | Cool            | Dull            | Low             | TFN(0, 0, 0.5)           |
| ordinal #2    | Medium          | Medium          | Medium          | Medium          | TFN(0, 0.5, 1)           |
| ordinal #3    | Bright          | Warm            | Intense         | High            | TFN(0.5, 1, 1)           |

### Figure 4. Membership functions for features with ordinal scale.

### Table 8. Binary string of Lighting Type.

| Linguistic Term    | Binary String |
|--------------------|---------------|
| General Lighting   | 100           |
| Accent Lighting    | 010           |
| Task Lighting      | 001           |
Table 9. Example of calculation sheet for translating measurement scales into linguistic terms.

|   | $F_1$ Budget | $F_2$ Working Duration | $F_3$ Design Style | $F_4$ Design Theme | $F_5$ Colour Hue | $F_6$ Brightness | $F_7$ Saturation | $F_8$ Colour Temperature | $F_9$ Daylighting Source | $F_{10}$ Lighting Type |
|---|-------------|------------------------|-------------------|-------------------|-----------------|-----------------|------------------|-------------------------|------------------------|----------------------|
| Case 1 | VL | S | 100000 | 001000000 | 000010 | 3 | 2 | 3 | 3 | 010 |
| Case 2 | L | S | 001000 | 001000000 | 000010 | 2 | 2 | 2 | 2 | 100 |
| Case 3 | L | S | 000100 | 000100000 | 000010 | 1 | 1 | 1 | 1 | 001 |
| Case 4 | H | M | 000001 | 000100000 | 000010 | 3 | 3 | 3 | 3 | 001 |
| Case 5 | M | M | 000010 | 001000000 | 000010 | 3 | 2 | 3 | 3 | 100 |
| Case 6 | VH | VL | 010000 | 001000000 | 000010 | 1 | 1 | 1 | 1 | 100 |
| Query | L | S | 100000 | 001000000 | 000010 | 3 | 3 | 2 | 3 | 011 |

Table 10. Example of calculation sheet for fuzzy numbers and local similarity measures.

|   | $F_1$ Budget | $F_2$ Working Duration | $F_3$ Design Style | $F_4$ Design Theme | $F_5$ Colour Hue | $F_6$ Brightness | $F_7$ Saturation | $F_8$ Colour Temperature | $F_9$ Daylighting Source | $F_{10}$ Lighting Type |
|---|-------------|------------------------|-------------------|-------------------|-----------------|-----------------|------------------|-------------------------|------------------------|----------------------|
| $S_a$ | VH | VL | 100000 | 001000000 | 000010 | #3 | #2 | #3 | #3 | 010 |
| $Q$ | M | M | 100000 | 001000000 | 000010 | #3 | #3 | #2 | #3 | 001 |
| $S_b$ | TFN(0.75, 1, 1) | TFN(0.75, 1, 1) | 000110 | 001000000 | 000010 | TFN(0.5,1,1) | TFN(0.5,1,1) | TFN(0.0,0.5) | TFN(0.5,1,1) | 010 |
| $Q$ | TFN(0.0,25,0.5) | TFN(0.0,25,0.5) | 100000 | 010000000 | 000010 | TFN(0.5,1,1) | TFN(0.5,1,1) | TFN(0.0,0.5) | TFN(0.5,1,1) | 001 |
| $sim$ | 0.84 | 0.84 | 1 | 0 | 1 | 1 | 0.89 | 0.89 | 1 | 0 |
Computing global similarity measure between a stored case and a query case

The global similarity function measures similarity between each of stored cases $S_i$ ($i = 1 \ldots m$) to the new query case $Q$. Suppose the weightings are equally weighted, each being equal to 1. Then the global similarity between the stored case $S_6$ and the query case $Q$ is calculated as:

$$\text{Sim}(Q, S_6) = \frac{\sum_{j=1}^{n} w_j \times \text{sim}(r_{qj}, r_{6j})}{\sum_{j=1}^{n} w_j} = 1 \times (0.84 + 0.84 + 1 + 0 + 1 + 0.89 + 0.89 + 1 + 0)/10 = 0.75.$$  

**Step 5: Retrieving the Case**

In the previous step, the query case and the stored cases are fed into the retrieval engine to retrieve the most similar case to meet the user’s preferences. The transformed query vector and each case vector of the case-feature-rating matrix are compared through the new fuzzy similarity retrieval algorithm. In this step, according to the calculation results of the similarity retrieval algorithm, the global similarity values in descending order are derived and the case with the highest degree of similarity measure is retrieved and displayed to the user. The target design drawing which is associated with the retrieved case is then triggered and displayed to the user as a prototype drawing through a Fancy-Designer® based virtual reality platform (http://www.fancydesigner.com.tw/). Thereafter, necessary review/revision to the retrieved drawing can be made.

**Step 6: Revising the Case**

Typically, the initial formulated case base might not be complete or accurate; thus, continuous learning is crucial to the CBR. Learning in the CBR system means revising/extending/refining previous experiences and incorporating them into the case base to make them more usable and accessible. Normally, the case retrieved from the case base will not be an exact match with a current design problem; it is necessary to adapt the retrieved case to better meet the current new problem (design requirements).

**Step 7: Retaining the Case**

A CBR system should seek to increase its knowledge by retaining new cases. The CBR system retains a revised problem description and its solution as a new case. The revised new case is stored into the case base for future reference case. A CBR system has the ability of learning by retaining useful cases during this step. Retaining step is a necessary prerequisite for enabling a system to learn from experiences. As long as the CBR system retains new cases, it learns and grows. The revising and learning characteristics of a case-based reasoning (CBR) are useful when the new query is hard to articulate their requirement. Thus, a CBR system can begin with a few number of reference cases and allows the case database to be developed incrementally, regarded as an initial version that could be revised, retained and restored to increase the number of reference cases.

5.2. Experimental results and discussion

*Precision* is a vital measure used to evaluate the efficiency of the case retrieval process of a CBR system. In this study, the performance of the similarity matching algorithm is evaluated by *precision*, where *precision* is defined as the measure of how many correct cases ($N_{\text{correct}}$) hits in the total retrieved cases ($N_{\text{retrieved}}$):

$$\text{Precision} = \frac{N_{\text{correct}}}{N_{\text{correct}} + N_{\text{false}}} = \frac{N_{\text{correct}}}{N_{\text{retrieved}}}$$. It represents the proportion of the relevant retrieved cases to all the
retrieved cases and is a direct measurement of the success hit counts of the retrieval process. Number of retrieved cases which achieves user satisfaction is used to measure the success hit counts of search engines. The more satisfied cases are retrieved the better the precision will be. For the comparison and evaluation of the performance of our previous matching algorithm (Lin & Ke, 2015) with the proposed matching algorithm, experiments were carried out and experimental results were calculated. Table 11 showed the experimental results comparing two similarity matching algorithms in terms of precision. The experimental results showed that the similarity matching efficiency of the proposed matching algorithm improved significantly and demonstrated its ability with 73% Precision during retrieval. The results proved that the proposed similarity matching algorithm on the proposed case-based reasoning system is effective and capable of handling case based retrieval of interior design drawings.

6. Conclusions and future recommendations

A new fuzzy cosine similarity retrieval algorithm has been proposed for searching interior design drawing case in a case-based reasoning system, in which mixed measurement scales are considered and applied to the case-based reasoning. Numerical experiments were carried out and the retrieval performance of the proposed similarity retrieval algorithm was measured by precision measurement. Compared with previous study, the increase in precision percentage proved the efficiency of the proposed algorithm. The experimental results demonstrated that the proposed similarity matching algorithm on the case-based reasoning system is effective and is suitable for handling case based retrieval of interior design drawings.

In our ongoing future work, as the number of query cases increases, more cases will be revised and restored in the case base, so that the availability of choices on cases can be increased. Besides, more design features will be extracted from more data sources to characterize complex interior design drawing case. Thus, with increased number of stored cases and extracted features, the next step of this study will be further to enhance more accurate retrieval performance and better user satisfaction.

Notes on contributor

Kuo-Sui Lin received his MSc. degree in industrial and management from National Tsing Hua University, Taiwan, in 1987, and his PhD degree in the field of management from University of Paisley, United Kingdom, in 1999, respectively. After nearly 25 years of carrier in the practice fields of manufacturing, system integration and venture capital, he changed his career to the academic circle. Since December 2010, he has been with the college of management of Aletheia University. Now he is an assistant professor in the department of information management and his current research interests include Fuzzy Set Theory and Applications, Engineering and Managerial Economics, Intelligent Information Systems, Project Management and Multiple Criteria Decision Making (MCDM).

### Table 11. Experimental results comparing two similarity matching algorithms in terms of precision.

|                          | Precision |
|--------------------------|-----------|
| In our previous study (Lin & Ke, 2015) | 57%       |
| In this study            | 73%       |
Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Kuo-Sui Lin http://orcid.org/0000-0003-3502-6705

References

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1), 39–59.

Ahmed, M. U., Begum, S., & Funk, P. (2012). Case studies on the clinical applications using case-based reasoning. In *Proceedings of 2012 federated conference on computer science and information systems (FedCSIS)* (pp. 3–10). Wroclaw.

Azad, B. (2015). Process of interior designing: A case study. *MAGNT Research Report*, 3(4), 125–131.

Baeza-Yates, R., & Ribeiro-Neto, B. (2011). *Modern information retrieval: The concepts and technology behind search*. Harlow: Addison-Wesley.

Begum, S., Barua, S., Filla, R., & Ahmed, M. U. (2014). Classification of physiological signals for wheel loader operators using multi-scale entropy analysis and case-based reasoning. *Expert Systems With Applications*, 41(2), 295–305.

Chen, T. Y., & Ku, T. C. (2008). Importance-assessing method with fuzzy number-valued fuzzy measures and discussions on TFNs and TrFNs. *International Journal of Fuzzy Systems*, 10(2), 92–103.

Cross, V. (1994). Fuzzy information retrieval. *Journal of Intelligent Information Systems*, 3(1), 29–56.

Kraft, D. H., & Colvin, E. (2017). Fuzzy information retrieval. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 9(1), i–63.

Larsen, B., & Aone, C. (1999). Fast and effective text mining using linear-time document clustering. In *Proceedings of the fifth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 16–22). San Diego, CA.

Leake, D. B. (1996). CBR in context. The present and future. In D. B. Leake (Ed.), *Case-based reasoning: Experiences, lessons, and future directions* (pp. 3–30). CA: American Association for Artificial Intelligence.

Lin, K. S., & Ke, M. J. (2015). A virtual reality based recommender system for interior design prototype drawing retrieval. In D. Barbucha, N. T. Nguyen, & J. Batubara (Eds.), *New Trends in Intelligent information and database systems, studies in computational intelligence* (Vol. 598, pp. 141–150). Heidelberg: Springer.

Nahm, U. Y., Bilenko, M., & Mooney, R. J. (2002). Two approaches to handling noisy variation in text mining. In *Proceedings of the ICML-2002 workshop on text learning* (pp. 18–27). Sydney.

Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*. Auckland: McGraw-Hill.

Weenawadee, M., & Sarun, I. (2013). Case-based reasoning for design patterns searching system. *International Journal of Computer Applications*, 70(26), 16–24.

Ye, J. (2011). Multicriteria decision-making method based on a cosine similarity measure between trapezoidal fuzzy numbers. *International Journal of Engineering, Science and Technology*, 3(1), 272–278.

Zadrożny, S., & Nowacka, K. (2009). Fuzzy information retrieval model revisited. *Fuzzy Sets and Systems*, 160(15), 2173–2191.