An Approach to Generating Scientific Techniques Based on Compositional Adaptation

Kambiz Badie, Maryam Tayefeh Mahmoudi, Mahsa Kiani
Info-Society Dept.,
Iran Telecom Research Center
Emails : { k_badie, mahmodi}@itrc.ac.ir
mahsa_kiani @yahoo.com

Abstract

This paper presents a new approach for generating scientific techniques based on a kind of compositional adaptation. In this respect, problem situation and solution are represented in terms of appropriate frames and tree structures respectively, and the final solution is then obtained through composing the solutions of the similar cases using a global distance function. The workability of this approach has been examined for the domain of Image Processing. The proposed approach can be applied to any scientific domain within which, goals can be defined in terms of appropriate actions.

Keywords : Technique generation, scientific techniques, frame, compositional adaptation

1. Introduction

The basic motive for using Case Based Reasoning is to make use of similar situations already experienced in the past, whose solutions can somehow be usable with respect to current problem, through some sort of adaptation. Regarding this, the approaches to all aspects of case representation, case retrieval and case adaptation can highly influence the final problem solving performance of CBR.

With respect to design or planning type problems, generating scientific techniques is one of these crucial issues, which can have lot of importance for those involved in research & development activities. The significance of this issue is particularly due to the fact that the very techniques suggested by human researchers/ engineers in a certain domain, though demonstrated to be satisfactory in certain default contents, may fail when the content is changed. It is therefore significant to generate techniques in such a manner that can best fit the current problem content. This will eventually lead to a remarkable saving in the amount of time and energy to be spent on research and development activities. Although CBR has widely demonstrated its capabilities in a wide range of design or planning issues, its ability in solving the problem of technique generation is still in the offing. In this paper, we present a CBR-based approach to generating scientific techniques, based on adapting previously-generated techniques through composition, and test it for generating techniques in the domain of image processing.

2. Some Existing Applications of Case-Based Reasoning to Solving Problems in the Domains of Design and Planning

CADRE is a prototype design system for building design [1], which is based on two forms of adaptation; dimensional adaptation based on the notion of, dimensionality expansion to make conflicts resolvable and dimensionality reduction to limit the complex of modifications and topological adaptation with case combination. Another application called CLAVIER which is a system for curing composite airplane parts by laying out in autoclave, solves a configuration problem. Here the adaptation is performed on the basis of a secondary case-based substitution for repair of cases [2]. JULIA is also a system, basically for meal planning as a sort of design problem, which uses a range of techniques for adaptation; such as simple substitution or special purpose structure modifying rules as a kind of transformational adaptation, It expects significant constraint changes during a problem solving session [2]. In the meantime, STEPPC has been developed as a generic tool to design process tracing and reuse, based on a substitutional adaptation for the parameters [3]. CBModeler is also a system, which applies CBR techniques to information systems design, and consists of two sorts of adaptation: internal transformational adaptation using integrity rules, that define the criteria for a consistent design in general, as well as reference model rules, that preserve consistency in the information system domain, and external adaptation which is to be performed by system designer [4]. These is also a CBR-based design system that assists human engineers in performing mechanical bearing design, which combines parametric adaptations and constraint satisfaction adaptation by both global and local rules [5]. Meanwhiles, COMPOSER, which is used in engineering design, applies constraint satisfaction algorithm for adaptation [6]. PLAKON is also a domain-independent configuration system, which is based on obtaining a partial solution by combining multiple parts of cases, provided that they do not overlap. Here, a sort of generative or derivative adaptation is used to complete the partial solutions [7]. A CBR-based system has also been proposed, which is able to make design patterns under constraints for engineering systems in general, and electronic systems in particular, based on a sort of transformational adaptation that takes into account the status of incompatibility of an old solution toward the new problem context and tries to overcome it at each stage of transformation [8]. It is seen that, all types of adaptation; i.e. substitutional, transformational, derivational (generative), as well as compositional, are suitably used to materialize the CBR-oriented design of some system.

3. The Proposed Technique Generation Process Based on Compositional Adaptation

3.1 Basics

The process of developing a scientific technique is not necessarily based on a comprehensive knowledge on the entire possibilities of the current problem; let say, no systematic effort is made to formalize the context for which the proposed technique can act efficiently. Having admitted to this fact, a question is that whether there is any systematic method for developing techniques based on observing the performance of previous experiences, taking into account the fact that formalizing the goals context is a controversial issue. To make a suitable response to this question, it is first necessary to develop a plan for each category of problems within which some predefined techniques are scheduled to respond to possible modes that can occur within the problem-solving process. Within this regard, both the technique and the mode, can be respectively described in terms of some propositions. Respecting this viewpoint, a problem may be represented in terms of a hierarchical structure within which the individual techniques are to be layered according to their related modes. Once a new problem is faced, its plan should be constructed based on a process of composing the plans of similar modes in such a manner that the similarity between the resulted plan and the plan of each similar problem can be reasonably proportional to the similarity between the current and the related similar problems. Regarding this, problems with complex nature should first be mapped onto some prototypical problems already experienced, and the related solutions should be adapted to them. In this view, Case-Based Reasoning is the ultimate alternative. With respect to using CBR for generating novel techniques, the situation in a case is the problem context which can preferably be represented in terms of a frame with appropriate attributes. The case solution is a tree-structure, within which each node represents the pair of an action and the very technique which is to handle it in some way, taking into account the fact that each technique located at a higher level of hierarchy as a parent node, is benefited by a number of child nodes comprising of appropriate pairs of action and technique, where actions should preferably be generic indicating the prime aspects that the parent technique is involved in. The hierarchical tree structure for representing the solutions is continued until the stage where the technique parts in the nodes are in practice turned into either a parametric structure within which parameters values constitute the identity of the related techniques, or some structures which cannot be further propounded in terms of appropriate actions, and it is the human expert, which can give the final idea on the way these structures can be composed.

3.2 Process of Technique Generation Based on Compositional Adaptation

The entire process of technique generation for a new problem context is as follows: First, past experiences of researchers in generating techniques for their problems are represented in terms of appropriate cases including case situation, representing the problem context, and case solution. Now, facing a new problem, it is first checked which case or cases in the library can be sufficiently close to the current problem context, making use of the normalized distance of the current problem with the situations in the stored cases. The cases retrieved in such a way are then stored on the basis of their normalized distance, and out of them, those which have a variance less than a certain threshold are selected for adaptation. The process of adaptation, in our approach, goes back to two different adaptations; one for adapting the values of related techniques for similar actions (or generic actions) and the other for adapting the sequences of actions justifying the related techniques. The adaptation in our approach is based on two types of compositional adaptation, respectively considered for the adaptations discussed above.

An approach, in this respect, is to minimize the summation of the distances between the final solution and a case solution, taking into account the normalized similarity between the current problem and the corresponding case situation as the weight. This sort of adaptation is performed for composing the values of the related attributes. Another approach is to apply a global distance function, where the major concern is to compose the solutions in similar cases in such a manner that the outcome can lead to a minimum value with respect to the sum of distances between the current situation and each case situation on one hand, and the distances between the final solution and each case solution on the other hand [9]. The mathematical expression in this regard is as follows:

$$\text{Global\_Dis}_k = \sum_{\substack{i=\text{For any case used}\\ \text{in the composition}}} [(\text{Diff}_i^{\,k}).(\text{Normalized\_Case\_Distance}_i)] \qquad (1)$$

then the resulted Global_Dis_k (GDF) is normalized with following formulas:

$$\text{Temp} = \sum_{k=1}^{z} \text{Global\_Dis}_k \qquad (2)$$

$$\text{Normalized\_Global\_Dis}_k = \frac{\text{Global\_Dis}_k}{\text{Temp}} \qquad (3)$$

$\text{Diff}_i^{\,k}$ is the number of differences between sequences of actions in proposed solution tree_k and solution of Case_i that is used for composition, $\text{Normalized\_Case\_Distance}_i$ is normalized distance between the current situation and situation of Case_i, and (z) is the number of proposed solutions with the system. Since the solutions in our approach have tree-like structure, it is important to consider them in a way that they can get ready for adaptation. Regarding this, those nodes in the solution tree which follow a similar action, are composed in their technique values. Obviously, since a technique itself is the root of a sub-tree, the same process of composition is repeated with respect to these sub-trees as well. The entire process is repeated until the stage where a technique is represented in terms of some parameters values which can become simply subject to composition, or some predefined structures whose composition can be carried out by a human expert who is in interaction with the system.

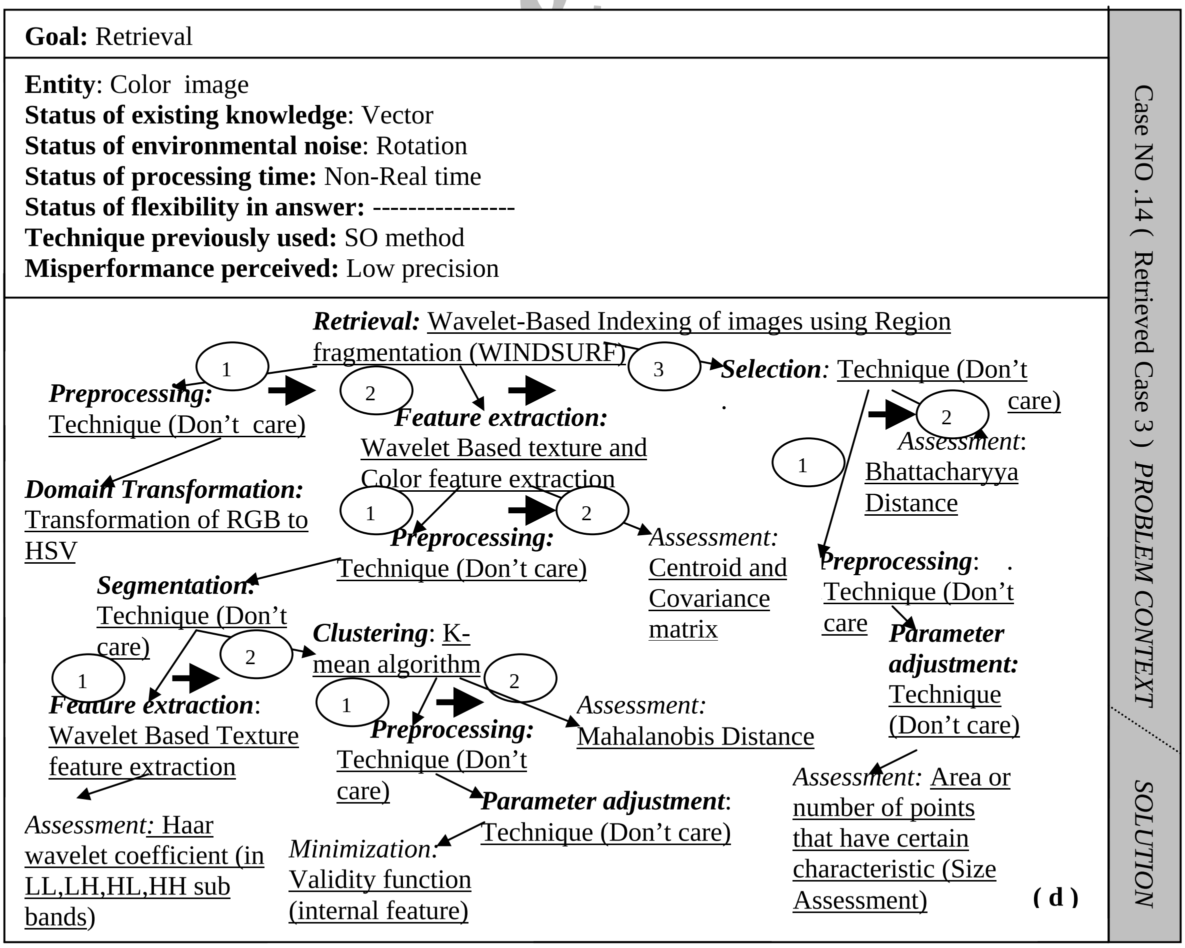
4. An Example in the Domain of Image Processing

4.1 Representation of Problem Context & Solution

Regarding frame as a means for representing problem context, entities such as "goal", "entity", "status of existing knowledge", "status of environmental noise", ”status of processing time", "status of flexibility in answer", ”technique previously used", and "misperformance perceived" can be selected as frame attributes, each with its predefined appropriate range of values. Now, considering the case of Image Processing for generating new techniques, the possible actions, which are in some sense the possible goals for technique generation, include entities such as "classification", "segmentation", "retrieval", "enhancement", "restoration", etc. Each technique in our approach, is equivalent to the set of some other techniques, each proposed to respond to an action or a generic action. Generic actions in our approach have been defined in such a manner that can respond to a wide range of problems. Regarding this, we have realized that entities such as "mode analysis", "assessment", "mapping", ”modification", ”transformation", ”fusion", ”reasoning", etc. can be good alternatives. An example is illustrated in Figure 1, for case situation and case solution in the domain of Image Processing.

4.2 An Example

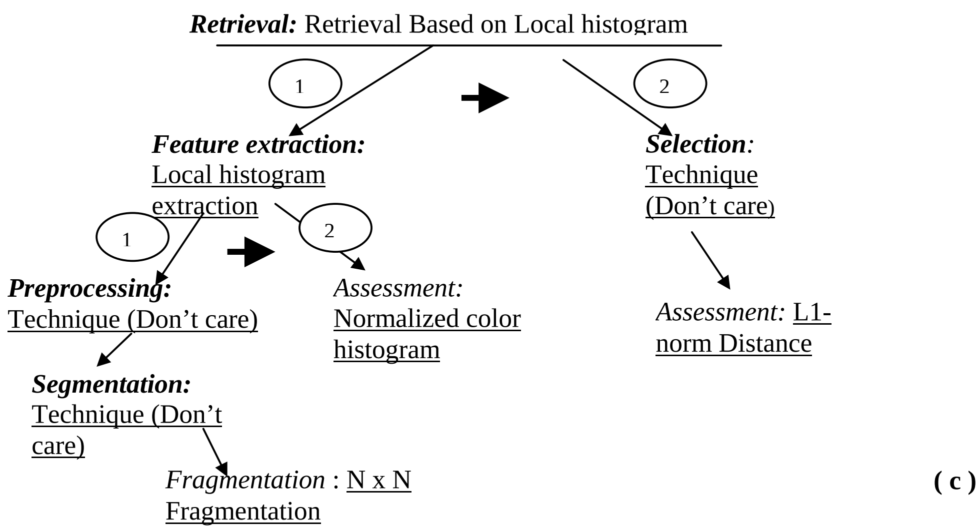
Suppose that we want to generate a technique to respond successfully to the problem context of Figure 1(a). Taking into account the discussion of 3-2, whose goal is retrieval, and considering a suitable threshold (0.3) for assessing distance between the current problem context and the situation in the stored cases, the similar cases to be retrieved will be case No.6, case No.13 , and case No.14 as illustrated in Figure1(b) , (c) and (d). It is to be noted that, each technique not only depends on a set of actions (and/or generic actions), but may also depend on the status of priority between these actions. In other words, sequence of actions play a significant role in describing a technique. For instance, in the situation of case No.6 (Figure 1(b)), to actualize the technique based on "gray-scale histogram" for "Retrieval" as the goal, the actions "feature extraction", "clustering" and "selection" are to be performed in a sequential manner. It should also be noticed that this sequence is not necessarily the same in all the retrieved cases; let say, the sequence "feature extraction, clustering, selection" (from Case No.6) is different from the sequence "feature extraction, selection" (from Case No.13), and the sequence "preprocessing, feature extraction, selection" (from Case No.14), and therefore a sort of adaptation is needed to optimize these sequences through compositional adaptation. The global distance function, already discussed in 3-2, is used for this purpose. It is seem that all these three cases share in two actions of Feature extraction and Selection, and since Assessment is the generic action propounded for the techniques of Feature extraction, Selection as well, compositional adaptation will be finally performed for “gray-scale histogram”, “normalized color histogram”, and “centroid and covariance matrix”, which are the corresponding techniques for Assessment under Feature extraction, and in the meantime the same compositional adaptation will be performed for “L2-Norm distance”, “L1-Norm distance”, and “Bhattacharyya distance”, which are the corresponding techniques for Assessment under Selection. Applying the global distance function to the sequences such as “feature extraction, clustering, selection” (from Case No.6), “feature extraction, selection” (from Case No.13), and “preprocessing, feature extraction, selection”(from Case No.14), etc. as the method for composing sequences, and taking into account the normalized distances “0.277”, “0.288” and “0.3” respectively belonging to Case No.6, Case No.13 and Case No.14, the final solution tree is illustrated in Figure 1(a). The ratios indicated in the figure, stand for the normalized similarities of the corresponding retrieved cases with respect to the current problem context. However, since the related functions hold different attributes, composition can not be performed at value level, and instead it is the human expert, whom at this stage will be asked to give his final idea on the appropriate functions, based on the status of their values.



Case NO .13 (Retrieved Case 2) *PROBLEM CONTEXT* / *SOLUTION*

Goal: Retrieval

Entity: Color image
Status of existing knowledge: Vector
Status of environmental noise: ------------
Status of processing time*:* ------------
Status of flexibility in answer: ------------
Technique previously used: Global color histogram
Misperformance perceived: Low precision

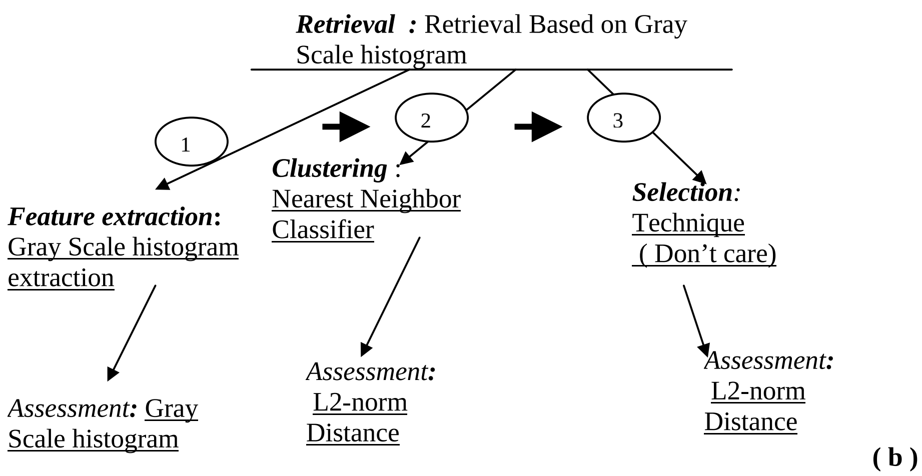


(c)

Case NO .6 (Retrieved Case 1) *PROBLEM CONTEXT* / *SOLUTION*

Goal: Retrieval

Entity: Gray Scale image
Status of existing knowledge: Vector
Status of environmental noise: Noise free
Status of processing time: --------------
Status of flexibility in answer: Medium
Technique previously used: --------------
Misperformance perceived:- --------------



(b)

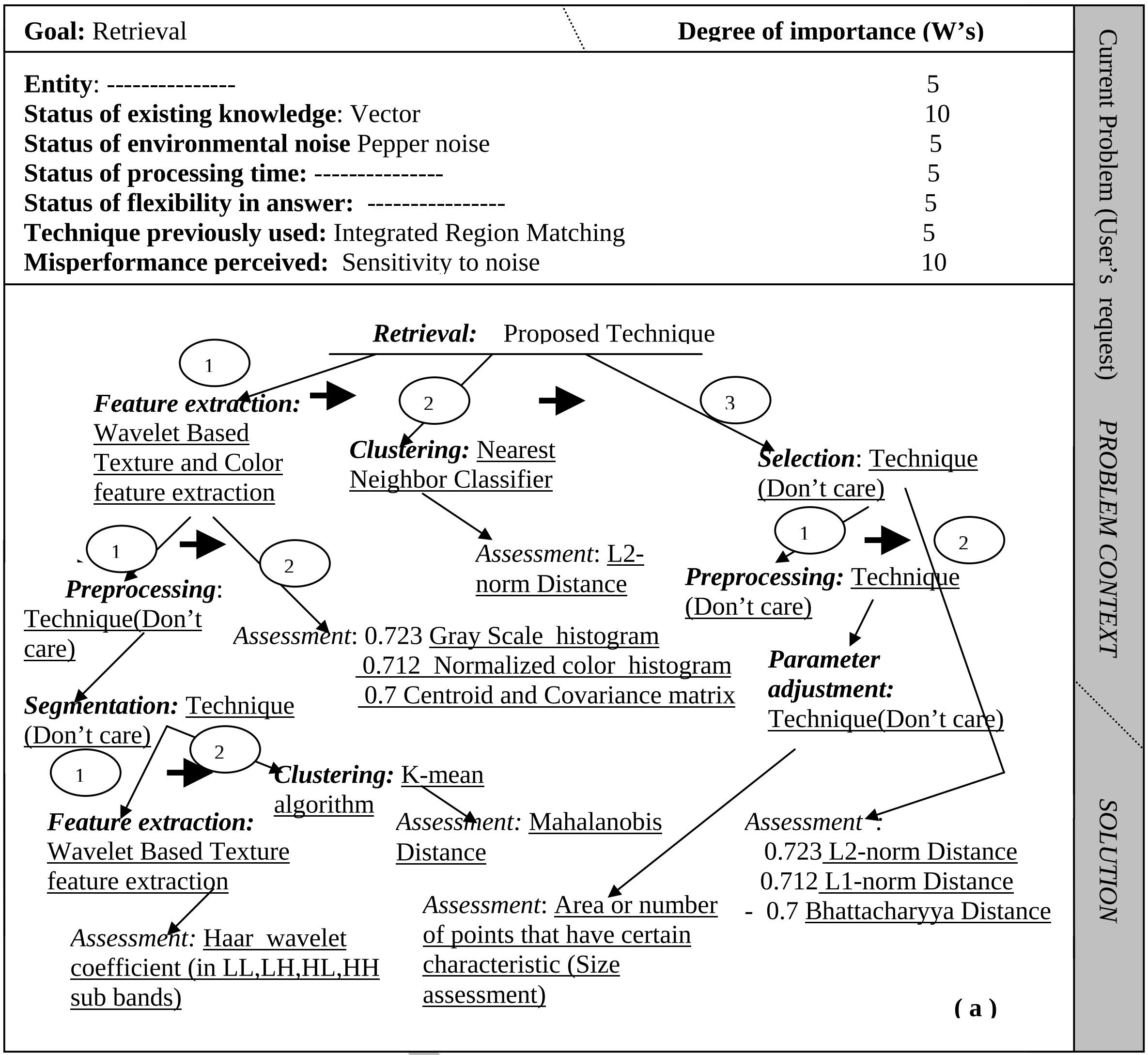


Figure1.An example

The textual expression for the final solution of Figure1 (a) is as follows:

… The appropriate technique for the current problem of retrieval, whose context is already described in Figure1 (a), is to first perform "feature extraction" with an appropriate technique of "wavelet-based texture and color feature extraction", next to perform "clustering" with "nearest neighbor classifier", and finally to perform "selection" with the following details: first a preprocessing (with its own details) is to be performed, and next to that, an assessment based on a choice out of "L2-Norm distance", "L1-Norm distance", and "Bhattacharyya distance", should be done. With respect to "wavelet-based texture and color feature extraction", first we need a pre- processing based on its own technique and "assessment", which is to be performed based on a choice from "gray-scale histogram", "normalized color histogram" and "centroid & covariance matrix". For "preprocessing" technique, we need a "segmentation", whose phases are respectively "feature extraction" and "clustering". The technique for "feature extraction" is "wavelet-based texture feature extraction", and the technique for "clustering" is "K-mean algorithm". "wavelet based texture feature extraction" is based on an action of assessment whose technique is "Harr wavelet…", and "K-mean algorithm" is again based on an action of "assessment", which is based on "Mahalanobis distance". "nearest neighbor classifier" itself is based on an action of "assessment", which itself is supported by "L2-Norm distance". And so on……..

It is seen that, through a compositional adaptation of the pre-stored cases, which are sufficiently similar to the current problem context, we have been able to generate a novel technique, whose phases can be described in a sequential manner. It is interesting to notice that, since techniques on our approach are treated as some sort of architectures, within which generic actions together with the related local techniques are used at different phases, the chance will exist to merge them systematically in a way that the final solution can reasonably include the benefits of each technique. Taking this point into account, we may end up with the conclusion that, the symbolic representation of techniques as the solutions for research and development situations, may only be feasible if they are planned to be obtained as architectures with different phases including different generic actions and corresponding local techniques.

5. Concluding Remarks

It was discussed in the paper that, adapting the previously- experienced techniques, in the form of tree structures, using compositional adaptation, can finally lead to formation of an appropriate technique for the current problem context. Here, adaptation is performed in two different aspects: (i) optimizing the sequences of actions in the solution tree, and (ii) adapting the most appropriate values for the terminal techniques, which can be numerically described. The proposed approach can be used as a potential scheme for generating techniques automatically in a wide range of engineering applications. It can therefore be used as a helpful decision & idea generation support tool for research and development engineers working in R&D units.

6. References

[1] I.Smith and B.Faltings, "Spatial design of complex artifacts using cases", *IEEE Proc. of 10th Intl.Conf. on Artificial Intelligence for Applications*, pp.70-76, 1994.

[2] M.Sasikumar, "case based reasoning", NCST, http://staff.ncst.ernet.in/1/1/sasi/clor.html.

[3] N.Prat, "Applying case-based reasoning in a design environment: issues and lessons learned", *6th Intl. Conf. on Artificial Intelligence in Design (MLinD'oo)*, Massachusetts, June 2000.

[4] D.Krample and M.Lusti, "case-based reasoning for information system design", *ICCBR*, USA, pp.63-73, 1997.

[5] X.Qin and W.C.Regli, "Applying case-based reasoning to mechanical bearing design", *Proc. of Design Engineering Technical Conf. (DETC'2000)*, pp.1-13, Maryland, 2000.

[6] L.Purvis and P.Pu, "COMPOSER: A case-based reasoning system for engineering design", *Robotica*, Vol.16, No.3, pp. 285-295, 1998.

[7] R.Lopez and E.Plaza, "case-based learning of plans and goal states in medical diagnosis", *Artificial Intelligence in Medicine*, Vol.9, No.1, pp.26-60, 1997.

[8] M.Hejazi and K.Badie, "A new approach to solution adaptation and its application for design purposes", *6th European Conf. (ECCBR 2002)* Vol. 2416, pp.549, 2002.

[9] N.Reyhani, K.Badie and M.Kharrat, "A new approach to compositional adaptation based on optimizing the global distance function and application in an intelligent tutoring system", *IEEE Intl. Conf. On Information Integration & Reuse*, 2003.