An automatic Case Based Reasoning System Using Similarity Measures between 3D Shapes to Assist in the Design of Metal Castings.

Miltos Petridis, Soran Saeed, Brian Knight
University of Greenwich, School of Computing and Mathematics, Park Row, Greenwich, London SE10 9LS UK
{m.petridis, s.saeed, b.knight} @greenwich.ac.uk

Abstract

In this paper, we present current research using the ShapeCBR system that automates the process of creation and selection of cases to populate a CBR system for retrieval of 3D shapes to assist with the design of metal castings. The special feature of this system is that similarity is derived primarily from graph matching algorithms. The particular problem of such a system is that it does not operate on simple search indices that may be derived from single cases and then used for visualisation and principal component analyses. Rather, the system is built on a similarity metric defined directly over pairs of cases and is primarily structural. An overview of previous research in this area is presented. This demonstrates the feasibility of a CBR approach to the design of metal castings. This paper describes further research into the use of the traditional componentisation as used in method engineering to provide a shape representation suitable for efficient retrieval of design knowledge. The paper presents current work aiming mainly at enhancing the efficiency and accuracy of the similarity metrics used in the ShapeCBR system. The architecture of the ShapeCBR system is presented. Finally, performance measures for the CBR system and the metrics used are given, and the results of trials of the system are presented and compared to results obtained from previous research.

Keywords

Case-Based Reasoning, Spatial reasoning, Shape recognition, Casting design, Knowledge Based Systems, 3D Shapes, Casting, Foundry.

1 Introduction

This paper reports on current research at Greenwich that aims at automating the process of retrieving and reusing design knowledge involved in the design of sand castings. In the heart of this problem lies the problem of retrieval of arbitrary 3D shapes from a case base designed as a design assistant to be used in the metal casting industry. In an early paper [7] a support tool was described, based upon a traditional componentisation of 3D shapes, with components of known cooling modulus. An initial evaluation of the tool through its application to a family of rotationally symmetrical casting shapes has shown the feasibility of a CBR approach to assist the design of sand castings [25],[26].

Out of a range of solids processing methods for the mass production of components; for example, casting, forging and machining, casting is the generally the cheapest. However, the problem with casting is one of quality, which depends on the existence of casting design knowledge. The advantages of a CBR system, capable of containing detailed information on the design process for products, devolve from its ability to realise casting know-how as a valuable asset. The knowledge of how to cast a product soundly within tight cost constraints is the result of a huge investment on the part of industries, universities and government over many years. Although the value of design knowledge is widely recognised throughout the industry, the management of design knowledge is often ad hoc in some respects. Design histories are often lost, or banished to paper files that are difficult to search. Also, design engineers retire [23], or move away leaving inadequate design records.

There are many problems faced by a casting design engineer, revolving around the physical freezing processes. Foremost among these is shrinkage in the mould, which can give rise to porosity and areas of structural weakness [1]. Other practical problems arise during pattern making and subsequent machining of the cast part. Many software tools have been developed to assist the designer. Jolly [2] found in his survey that the foundry industry is looking for software that can not only predict problems that occur during metal solidification (such as shrinkage porosity) but also, having predicted these problems, to propose intelligent solutions to problems found.
Current commercial casting software can be classified into two broad areas: intelligent knowledge-based systems (IKBS)[22,24], and numerical simulations based on physical process models[3-5].

IKBS systems attempt to support an earlier stage in the design process. Numerous software tools such as those discussed in [7] have clearly demonstrated the usefulness of knowledge-based and other advanced heuristic-based programs for designing castings. Some of the commercial software packages available can calculate the position of feeders (NOVACAST [8]) and also analyse geometric properties to give suggestions to improve the design further (AutoCast [9]).

Although many prototype tools have demonstrated the efficacy of CBR in the domain of engineering and design [10-15], there is a scarcity of research for its use in the foundry industry. CBR can play an important role in intelligent casting software. One commercial CBR system [16] called Wayland, is used for the setting of parameters in pressure die-casting. This research has demonstrated that CBR has an exciting future in casting software.

The main problem for a CBR system is how to retrieve cases, where the retrieval must be based on shape. Although there are other possible search indices, for example the type of casting alloy, weight and general description of part (wheel, sea-gland, valve, engine bearing cap, etc.), these descriptions are too general for accurate retrieval. General classifications of shape components have been proposed; for example, Biederman's geons [17]. However, during this research, it became apparent during knowledge elicitation that a decomposition of shapes specific to the casting industry already existed in practice, [7, 18]. The research described here uses a graphical representation of shapes based on this decomposition as a foundation for shape retrieval.

A first evaluation of this approach was conducted by Mileman using the CASTAID system[26]. Mileman showed that a case based reasoning system can retrieve competently some useful relevant knowledge from a case base of past 2D casting shapes and provide casting advice that is comparable to that elicited from a human expert using the CASTAID CBR system. Although Mileman demonstrated the feasibility of this CBR approach and evaluated its effectiveness, there were three key limitations in the approach adopted in the research:
1. The similarity metrics used did not take into account the geometrical proportions of shape components.
2. The approach was applied and evaluated only on axisymmetric shapes and worked on componentised representations of 2D cuts through the axis of symmetry. The CASTAID system could not cope with non-axisymmetric 3D shapes that are common in casting problems.
3. The CASTAID system relied in the manual componentisation of a 2D cut through a casting shape. In order to make this approach usable in the context of a casting design exercise, it is important to automate the componentisation process that is necessary to enable the calculation of similarity metrics.

In [27] we proposed an extension of the similarity metrics to take into account the geometrical proportions of individual components and to cope with 3D shapes. Additionally, we proposed a way to automate the process of decomposing the geometry of a real 3D geometrical casting shape to the graphical representation needed to allow for the efficient retrieval of similar shapes and thus reuse relevant casting design knowledge. This paper discusses in some depth the new similarity metrics and presents the key results of a systematic evaluation of the new approach and system and a comparison of the efficiency of the new metrics to that achieved by Mileman.

In section 2 of this paper, the graphical representation and the extended similarity measures used for retrieval are explained. Section 3 discusses the automatic process of encoding real 3D casting shape into the case base to allow efficient retrieval of similar shapes and reuse of casting knowledge. Section 4 gives an evaluation of the current system based on experimental results from a trial domain of rotationally symmetric objects.

2 Graphical representation and similarity metrics

In [7] a decomposition of a shape into a set of joined components was described. The decomposition is a natural one, used over many years by casting design engineers. It is based on a set of component types of significance in casting design. There are 8 main component types including Bar, L, T, X, Taper, Flange, Bespoke-Taper, and Bespoke-T. The componentisation process distinguishes two sets of component type: those that define the structure (L’s, T’s and X’s) and those that join the first set together (bars and tapers). Using this classification, we may abstract a graphical representation of the structure of any shape S where the nodes are elements of either set, and the arcs represent interfaces between components.

As an illustration, consider the rotationally symmetric shape shown in cross section in figure 1. A graph representation of this figure is given in Figure 2.

Retrieval of shapes for casting design is an example of structure based case retrieval, as defined by Gebhardt [15]. For these systems, attributes representing complex structures are difficult to define, and similarity must be derived from structure directly. For the sub-class of graphical structures, Gebhart reviews several retrieval systems. These include clique detection as in the Fabel component Topo [19], largest common subgraph [20] and hamming distance [21].
In the research described here, we have used similarity measures based on features extracted from the structural graphs. Perfect similarity between shapes S1 and S2 is obtained when they have identical structural graphs. However for graphs that do not match completely, there are a number of features that can be extracted and compared. Each feature gives rise to a different similarity measure, representing a different case retrieval.

Correspondingly, there are a number of different problems associated with casting a shape, each connected with a different structural feature. Porosity tends to depend on specific local features, whereas machining problems tend to depend on global structure. The approach of this research has been to construct a retrieval tool to investigate the efficacy of the various metrics with respect to different casting design problems. The tool employs a generalised similarity measure $\sigma(S1,S2)$ between shapes S1 and S2, representing a weighted sum of the similarity measures based on different features extracted from the graphs of S1 and S2:

$$\sigma(S1,S2) = w_{\text{comp}}\sigma_{\text{comp}} + w_{\text{mcs}}\sigma_{\text{mcs}} + w_{\text{cycle}}\sigma_{\text{cycle}} + w_{\text{leaf}}\sigma_{\text{leaf}} \quad (1)$$

Variation of the weights in this formula allows a general test of retrieval against any given casting problem. The individual similarity metrics in (1) are defined as follows:

- $\sigma_{\text{comp}}(S1,S2)$ is a measure based on the number of component types that are common to the two graphs. If two graphs are nearly identical, $\sigma_{\text{comp}}$ will be close to 1. The length function is defined as $\text{length}(S) = \text{number of components in } S$, and the value of this metric is given by:

$$\sigma_{\text{comp}}(S1,S2) = \frac{1}{\text{CompTypesNo}} \sum_{\text{Comp}} \frac{\text{length}(S'_{\text{comp}})^2}{\text{length}(S1)\text{length}(S2)} \quad (2)$$

where $S'_{\text{comp}}$ is the maximal number of common components of a particular type to graphs S1 and S2. Nevertheless, this metric does not take into account of the graph’s topology. This is taken care of by the metric in (3) below:

- $\sigma_{\text{mcs}}(S1,S2)$ is a measure based on the length of the maximum matching common subgraph (MCS). If two graphs are nearly identical, $\sigma_{\text{mcs}}$ will also be close to 1. This similarity metric is given by:

$$\sigma_{\text{mcs}}(S1,S2) = \frac{\text{length}(S')^2}{\text{length}(S1)\text{length}(S2)} \quad (3)$$

where $S'$ is the maximal common subgraph of S1 and S2, i.e. the largest graph which is a subgraph of both S1 and S2. The problem of finding $S'$ is related to that of the well-known graph isomorphism problem.

For small graphs of up to 10 arcs, a search based on direct recursive comparison of all subgraphs of S1 with those of S2 using a “greedy” algorithm is possible. For larger graphs a strategy based on a preliminary comparison of node types and degree can help to reduce the search time. (3a) shows the calculation of the MCS similarity metric for the graphs shown in Fig. 3.
\[ \sigma_{\text{mcs}}(S1, S2) = \sum_{\text{comp}} \frac{(2)^2}{(7)X(6)} \times 100 = 9.52\% \] (3a)

\[ \sigma_{\text{cycle}}(S1, S2) = 1 - \frac{|\text{ncycles}(S1) - \text{ncycles}(S2)|}{\max(\text{ncycles}(S1), \text{ncycles}(S2))} \] (4)

\[ \sigma_{\text{leaves}}(S1, S2) = 1 - \frac{|\text{nleaves}(S1) - \text{nleaves}(S2)|}{\max(\text{nleaves}(S1), \text{nleaves}(S2))} \] (5)

**The extended metrics used**

The metrics defined in equations (2) and (3) above assume 100% match between two components of the same type irrespective of their geometry. However, it is arguable that components of the same type, but of considerably different geometry may require different casting advice. For example, a thin and long bar may need to be cast differently to a short and fat one. In order to take this into account, the definition of the length of the common part of the graph was adapted to take this into account. It was thus multiplied by the average geometrical similarity of the overlapping components between the corresponding components of the same type in the two graphs. So, for example, now the component metric from equation (2) above reads:

\[ \sigma_{\text{comp}}(S1, S2) = \frac{1}{\text{CompTypesNo}} \sum_{\text{Comp}} \frac{(\text{AvSim}_{\text{comp}} \times \text{length}(S'_{\text{comp}}))^2}{\text{length}(S1) \text{length}(S2)} \] (6)

Where \( \text{AvSim}_{\text{comp}} \) is the average similarity between components of type \( \text{comp} \) in the two graphs \( S1 \) and \( S2 \).

Additionally, the metrics were extended to deal with 3D shapes. Typically, casting shapes are stored as files produced by CAD packages such as AutoCAD[6]. These files contain all geometrical information and most CAD packages provide facilities for providing 2D sections through the 3D shape. The case base in the first system contained only one 2D section through each shape. However, in many cases two or more substantially dissimilar 2D sections could provide a more accurate description of a 3D shape. These would need to be taken into account for a more efficient retrieval of 3D shapes. The selection of dissimilar 2D sections can be achieved with the use of a similarity threshold to define substantially dissimilar sections.

Arbitrary 3-D shapes, taking as an example the mug in Figure 4, can be treated as the two or more cross-sections as depicted in Figure 5. These two ‘views’ of the mug can provide valuable identifiers to enable accurate retrieval.

In this case, the overall measure of similarity between two 3D shapes \( S1 \) and \( S2 \) needs to be redefined as:

\[ \sigma_{\text{3D}}(S1, S2) = \sum w_i \sigma(S1_{2D\text{section } m}, S2_{2D\text{section } m}) \] (7)

**Fig. 1.** Example MCS comparison
- \( \sigma_{\text{cycle}}(S1, S2) \) is based on a count of elementary graph cycles:
- \( \sigma_{\text{leaves}}(S1, S2) \) is based on a count of leaf nodes, and gives the number of branches in the graph that are not connected on both ends:

© BCS Specialist Group on Artificial Intelligence
where each 2D section \((n,m)\) of each 3D session is used only once so that the above measure is maximised.

The CBR process and metrics can in fact be used to determine a number of dissimilar cuts through a 3-D shape (Figures 4 and 5 above).

Fig. 3. Arbitrary 3-D Shape (mug) Fig. 2. View 1 of Mug (of Figure 4.10) made of Bars, L-junctions and T-Junctions

3 Automating the decomposition process

The *ShapeCBR* software system has been developed at Greenwich to automate the process of matching a given target shape to a case in the Case Base. The case base is populated with cases containing information relevant to real metal casting experience. The information contained in each case relates to both a geometrical description of a real shape and domain specific information about the way that the shape was actually cast. Additionally, some cases may contain general expert advice relevant to casting the shape in a textual form. The system allows the user to retrieve a shape from the case base to match a target case, according to a match on the four contributing features as described in the previous section. Weighing factors can be applied by the user to attach varying importance to each of the similarity measures.

Figure 6 shows an example exercise of matching a target case to a retrieved case from the case base. Advice on positions of feeders and chills is annotated on the picture of the retrieved case. Early feedback from the use of the system has been promising [25]. However, CASTAID the first generation of the system as described by Mileman [26] relied on the manual decomposition of shapes to the generic components identified. This was an inefficient and cumbersome process that could hamper the practical use of a commercial system. Additionally, only the type and not the actual geometrical dimensions of a component were stored. This prevented us from increasing the sensitivity of the similarity criteria to take into account a similarity measure between components of the same type. For example, it makes sense that the aspect ratio of a Bar component would affect its similarity to another Bar component for purposes of casting. The positioning of feeders and chills can be affected, so that the knowledge associated with a shape may be contingent not only on the types, but on actual geometrical features of the constituent components.

An algorithm was devised to provide for the automatic decomposition of shapes into the generic components used in this research. This algorithm is made up of a decomposition and a classification algorithm. The decomposition algorithm decomposes a shape into a number of components. The classification algorithm is then used to classify each component under one of the generic types (bar, L, T etc). The algorithms start by projecting each vertex to any sides that are directly opposite it. This provides a decomposition of the area of the shape into a set of rectangles and triangles (fig. 6 second shape). These are then reconciled and merged if sides defined by internal points only connect them. For example two consecutive Bar components can be merged into one longer one. A set of rules then identifies each element as one of the generic components needed for the componentisation of the shape. For example, a rectangular component that has two opposite internal sides is resolved to be a Bar. Finally, the components are created by adding “stems” where appropriate (typically to joins, such as L, T and X).

Figure 7 shows an example of such a decomposition. Notice the top left L component. In the middle figure, the algorithm has identified there a rectangle. The rule that identifies this as an L component fires on the fact that this rectangle has two adjacent sides (right and bottom) that are internal lines. This identifies the component as an L.

A more detailed description of the algorithms can be found in [28]. An additional advantage of automating the decomposition into components is that the output of this process is not only the graph of connected components representing the structure of the shape. Each component is now associated with geometrical information describ-
ing the exact dimensions of the component. This allowed us to extend the definition of similarity between shapes taking into consideration the actual geometry in addition to just the layout of the components in each shape.

Fig. 4. Matching a case to a target in the ShapeCBR system

Fig. 5. The decomposition process

Figure 8 shows the architecture of the system and an overview of the process of importing cases into the case base.

4 Evaluation of the system

The integrated system with the new algorithm that automates the componentisation of the shapes into cases in the case base was tested against the 100 manually generated cases used to evaluate the previous system. It was then found that the componentisation produced by the system was identical to the one produced manually. The evaluation of the decomposition and classification algorithms will need to continue, by trying more complex shapes. However, the current case base already covers a wide range of real casting shapes. This shows that these algorithms are fit for purpose at least for the types of casting shapes that the system is dealing with.

The automation process has thus not affected the overall performance of the system as all cases were componentised exactly as they were manually in the previous research. The performance measures for the whole system were based on the trial case base consisting of 100 cases and tested against 20 new problems. For this trial, we took the domain of shapes with rotational symmetry: wheels, armatures, cylinders, etc. This domain is coherent from a practical point of view, so that we can attempt to cover it with a limited case base. It is however sufficiently varied to encompass a wide range of casting problems.

Performance of the case base was assessed on several different measures. For each case we obtained expert advice from a domain expert. This is compared to the advice given by the system for each target case. For a given target the retrieved set should provide the solution to (I) correct orientation, (II) the number and positions of feeders (III) the position of possible chills, (IV) the need for chills. (V) special problems encountered with this shape. For each of these problems, we can score how well the retrieved case presents the answer. In situations where no obvious visual match may be made with the nearest case, we can widen the search to retrieve more cases, and leave the user to select the one with the best visual match. In such a mode of operation, the user is allowed to browse the nearest matches to look for the best advice. A full discussion of this trial can be found in the paper by Knight et al [25] and in the thesis by Mileman [26].

© BCS Specialist Group on Artificial Intelligence
Figure 9 summarises the results of this trial using various combinations of weights for all similarity features and an optimisation of the weights showing the weights combination that can optimise the advice for the metrics used over all 100 cases in the case base.

**Fig. 7. ShapeCBR advice on 3D shapes**

**Fig. 8. Comparison between the CASTAID and the ShapeCBR casting advice efficiency**

**Fig. 9. Shape CBR advice using various values for k**

Figure 10 shows the results of the experimentation using casting design advice applied to a number of 3D shapes. ShapeCBR gave casting advice for the 3D shape by matching a number of independent 2D “cuts”
through the 3D shape to their closest neighbours in the case base. The advice was then aggregated/averaged for the various cuts to provide an overall advice for the 3D shape. The advice was compared to expert advice provided by a domain expert. It can be seen that the efficiency of the advice is not as good as the one obtained for 2D shapes, but it is still competent. Further analysis discussed by Saeed [28] determined that even in cases where the advice obtained by ShapeCBR was different to the one obtained from the expert, the expert consequently assessed this advice in the majority of the cases as “competent”.

Figure 11 shows the results from further experiments. These showed that the efficiency of the system can be further improved by varying the number of nearest neighbours used to obtain casting advice. The advice was obtained using a weighted voting system between the K nearest neighbours to the target case.

6 Conclusion

In this paper we have described recent work on a case based system for the design of metal casting procedures. The key problem addressed by the work is the retrieval of typical families of 3D casting shapes. The method proposed is based on a shape componentisation, which is particular to the domain of casting problems. The shape componentisation gives rise to a graphical representation of shapes, from which similarity metrics may be abstracted. We presented an overview of previous research that showed the feasibility of a CBR approach to assist the design of metal castings in the foundry industry. Current research work on automating the process of eliciting the cases from real CAD drawings of 3D shapes has been discussed here. Our evaluation shows that it is possible to automate this process and produce a componentisation similar to the one produced by domain experts. This componentisation can then be used by the CBR system to determine the similarity between shapes and thus retrieve competent solutions for a given target case shape.

Additionally, the inclusion of this missing link into the proposed process allows more detailed geometrical information relating to the shape and its components to be linked directly to the cases in the case base. In particular, the inclusion of actual geometrical data describing each component in the graph representation of a shape allowed an extension of the similarity metrics that increase the competence of the system. Additionally, the evaluation showed that it is now possible to automate the process of eliciting distinct 2D sections from 3D partially rotationally symmetric shapes. This allows the use of the extended similarity metrics between 3D shapes that draw on the similarity between numbers of 2D sections from each shape and reuses expert casting advice associated with them. Finally, a further line of research is under way to use more effectively casting knowledge associated with past cases, It is now possible to link more relevant information to cases in the database in ShapeCBR. For example documents such as test documents, photographs and blueprints can be associated with a retrieved case, providing richer contextual knowledge and thus improving the usefulness and relevance of the advice given by the system. Future work is also planned to extend the trials to wider domains, including general 3D systems. Work is also being planned for the integration of the system with physical modelling systems, such as SOLSTAR, or CFD software packages to assist, seed and optimise the resulting casting design.

References

2. Jolly, M. “Overview and appraisal of numerical analysis software for the simulation of casting processes”, BICTA bulletin 3, 1996
4. Simular. Aluminium Pechiney, Aluval, BP27 38340 Voreppe, France
5. SOLSTAR. Corbett, C.F. “Right first time...with Solstar”, Foundry Practice, Number 218, December, 1989
6. AutoCAD. Autodesk Inc. USA.
7. Knight, B; Cowell, D.; Preddy, K. “An object-oriented support tool for the design of casting procedures”, In Engineering Applications of Artificial Intelligence, Volume 8, Number 5, pp: 561-567, 1995

(C) BCS Specialist Group on Artificial Intelligence
20. Tammer, E.C; Steinhofel, K; Schoner, S; Matuschek, D. “Anwendung des Konzepts der strukturellen Ähnlichkeit zum Fallvergleich mitterns Term- und Graph-representationen”, Fabel report, 38, GMD, Sankt Augustin, 1995