*Artificial Intelligence for Engineering Design, Analysis and Manufacturing* (1999), *13*, 145–156. Printed in the USA.


# A voxel-based representation for evolutionary shape optimization

PETER BARON,[1] ROBERT FISHER,[1] ANDREW TUSON,[2] FRANK MILL,[3] AND ANDREW SHERLOCK[3]

[1]Department of Artificial Intelligence, University of Edinburgh, 5 Forrest Hill, Edinburgh EH1 2QL, Scotland, UK.
[2]Department of Computing, City University, Northampton Square, London, EC1V 0HB, UK.
[3]Department of Mechanical Engineering, University of Edinburgh, Sanderson Building, The King's Buildings, Mayfield Road, Edinburgh EH9 3JL, Scotland, UK.

(RECEIVED June 6, 1998; REVISED November 27, 1998; ACCEPTED December 7, 1998)

**Abstract**

A voxel-based shape representation when integrated with an evolutionary algorithm offers a number of potential advantages for shape optimization. Topology need not be predefined, geometric constraints are easily imposed and, with adequate resolution, any shape can be approximated to arbitrary accuracy. However, lack of boundary smoothness, length of chromosome, and inclusion of small holes in the final shape have been stated as problems with this representation. This paper describes two experiments performed in an attempt to address some of these problems. First, a design problem with only a small computational cost of evaluating candidate shapes was used as a testbed for designing genetic operators for this shape representation. Second, these operators were refined for a design problem using a more costly finite element evaluation. It was concluded that the voxel representation can, with careful design of genetic operators, be useful in shape optimization.

**Keywords:** Shape Optimization; Evolutionary Algorithms; Voxel Representation

## 1. INTRODUCTION

Shape optimization attempts to find an optimal shape for a component subject to design constraints. Typical problems that are of interest to the research community in this area have been concerned with structural load-bearing components and aerodynamic profiles. Some work has also been reported in areas such as thermal conduction for heat sinks and manufacturing cost minimization. In structural shape optimization, often these studies aim to minimize the amount of material (and hence perhaps cost and weight) needed to support a given load. In aerodynamic optimization, often the aim is to minimize drag subject to constraints on lift and geometry. Almost all of the work to date has described shape representations for single-criterion optimization, although many researchers are interested in multicriteria problems.

Structural shape optimization can be usefully characterized as the integration of geometric modelling, structural analysis, and optimization algorithms (Hsu, 1994). The finite element (FE) method is popularly used to analyze candidate shapes. In early research in shape optimization, the FE mesh itself was used as the geometric model to be manipulated by the optimizer. Optimization techniques then available were based on mathematical methods of function optimization, typically gradient based. The nodal coordinates of the FE mesh were used as design variables. However, it soon became apparent that use of the mesh as the geometric model was impractical due to difficulties in ensuring that the mesh could adequately calculate stresses and in keeping the shape's boundary smooth. Researchers moved to separating the geometric modeller and the FE mesh. Commonly the boundary of the component is modelled using splines, with control point coordinates used as design variables. Splines have the useful property of smoothness and local shape control. Mesh generation techniques then generate an adequate mesh given a description of the candidate shape's boundary.

Gradient-based optimizers can find optima with few design evaluations. This is often extremely important in engineering problems, where the time taken to perform one

Reprint requests to: Andrew Tuson, Department of Computing, City University, Northampton Square, London, EC1V OHB, UK. E-mail: andrewt@soi.city.ac.uk

design evaluation is often many orders of magnitude greater than the time taken to produce candidate designs. However, such optimizers can often have difficulties in dealing with local optima, discrete design variables, and with noise generated when small changes in the design variables cause changes in mesh topology. Recently, to address these problems, the use of stochastic optimization techniques, such as genetic algorithms (GAs) due to Holland (1975), and simulated annealing (Kirkpatrick et al., 1983), in shape optimization (Chapman et al., 1994; Smith, 1995*b)* has been a popular area of research. Generally, this research has still retained a parameterized description of the shape's boundary as the geometric model.

The work described in this paper investigated the possibility of replacing this boundary representation of the shape with a cellular representation. The cellular representation chosen in this work used voxels that partition the design space into rectangular regions or boxes that are then assigned a binary full or empty value. This approach was motivated by a number of potential advantages (Smith, 1995*b)*:

- any shape can be represented to an arbitrary accuracy by increasing resolution;
- it is straightforward to convert existing engineering solutions into voxels;
- they map naturally to the representations frequently used by GAs;
- domain knowledge can be readily incorporated;
- geometric constraints can easily be applied; and,
- the topology of candidate shapes is not predefined.

However, in contrast to the successful application of this technique in (Farrell, 1998) for the inversion of geographical and potential-field data, earlier work by Watabe and Okino (1993) states the following objections to the scalability of voxel representations:

- the occurrence of small holes in the final shape;
- the long length of the chromosomes;
- the expectation that crossover operators would be ineffective; and
- the lack of smoothness in the shapes' outlines.

Given the potential advantages of a voxel representation, the Authors considered it worthwhile to address these difficulties. Specifically, the aims of this work were:

- to determine the suitability of voxels as a geometric model for use in shape optimization and any difficulties, such as those outlined above, that may arise;
- to design suitable operators for a GA optimizer to use with such a representation to overcome such difficulties; and
- to investigate and identify issues that will have to be confronted by the practitioner in scaling up this representation to real-world problems.

Therefore this work does not aim to produce a system that returns a usable, improved solution to a real-world problem. Instead, it concerns itself with the more strategic and scientific question of investigating and, where possible, resolving issues that pertain to *how* a practitioner is to construct such a practical system.

### 1.1. Experiments

Two experiments were devised to investigate the voxel representation. First, a simplified beam design problem was formulated for which the cost of evaluation would be small. Using this problem as a testbed, a number of operators were designed. Second, an annulus design problem was tackled using a finite element analysis. Thus, the computation cost of evaluation in this case was much greater. The usefulness of the operators designed in the first experiment could then be evaluated with a more difficult design problem and related scalability issues investigated. Baron (1997) gives comprehensive details of all experiments undertaken.

Finally, this investigation will restrict itself to examples where 2D voxels (pixels) are used. This is for reasons for convenience and speed of solution evaluation as FE analyses in three dimensions are more computationally demanding. However, no assumptions are made in this study regarding the dimensionality of the problem and so the results presented here should be generalizable to higher dimensional problems.

## 2. SIMPLIFIED BEAM DESIGN

A prototypical mechanical engineering problem is that of optimizing a beam to support various loads with a minimal amount of material. Evaluation of the candidate cross sections was made using bending theory for symmetrical beams, considering only normal stresses (Gere & Timoshenko, 1984). This is an oversimplified model, but is sufficient to test whether the potential problems with a voxel representation outlined above do pose a problem in practice. The maximum stress constraint imposed by the physics model used in these experiments is summarized below.

$$\left|\frac{My_i}{I}\right| < \sigma_{\max} \textit{ for all voxels,}$$

where $\sigma_{\max}$ is the maximum stress allowed within any given area (voxel); $M$ is the bending moment; $y_i$ is the distance of the voxel $i$ from the neutral axis of the shape; and $I$ is the second moment of area of the candidate cross section. The neutral axis of a shape is defined as a horizontal line that passes through the center of the mass of the shape. As a voxel representation uses areas that are all of uniform size and density, the center of mass can be found by taking the average of the positions of all occupied voxels. The second

https://doi.org/10.1017/S0890060499133031 Published online by Cambridge University Press

moment of area is approximated in the discrete representation by summing the moments of each voxel, that is:

$$I = \sum_{i=0}^{n} a y_i^2,$$

where $a$ is the area of a voxel.

In the real world, the solution to this problem would correspond to an *I*-beam, but that also requires a web to connect the two flanges of the beam together. In a design based on a full calculation with shear stress, the web would be necessary so to counteract this additional stress. However, as shear stress is not represented in this problem, a connectivity requirement in the form of a repair step was added, whereby all pixels must be connected to a seed pixel in the center top edge of the beam. In addition, all vertically central voxels were enabled to provide a straight web before the connectivity repair step. This was found, in formative experiments, to prevent the formation of a crooked web (as the physics model used does not prevent this), and improve slightly the results obtained.

To try to ensure that the alterations and improvements made to the GA here will also prove beneficial to the real-world problem, it was decided not to concentrate on fine-tuning any of the various parameters available, but rather to focus on the design and operation of various new operators. Therefore, parametric variations were restricted to an absolute minimum and were used only to determine the approximate values required to gain reasonable advantages from the new operators. Therefore, in the following experiments, the following parameter settings remain constant unless mentioned otherwise:

| | | |
|---|---|---|
| Beam dimensions | = | $0.05 \times 0.10$ m |
| Bending moment | = | 13,000 Nm |
| Voxel grid | = | $32 \times 64$ voxels |
| Max. stress allowed | = | $2 \times 10^8$ Nm$^{-2}$ |

## 2.1. Experiments using the naïve GA

The first set of experiments with a 2D representation treated the chromosome as a long 1D binary string that wrapped around at the vertical edges onto new lines to form the 2D cross section. Standard two-point crossover ($p_c = 0.35$) and bitwise mutation ($p_m = 0.001$) were used in conjunction with a generational GA with a population of size 20. GENITOR-style rank-based selection (Whitley, 1989) was used throughout. From the above, the fitness function, $F$, to be minimized was of the following form:

$$F = V + S/(1000 \times \sigma_{\max}) + k \times \max\{(S - \sigma_{\max}), 0\},$$

where $V$ was the count of active voxels (proportional to weight), $S$ the maximum stress of any voxel, $\sigma_{\max}$ the value of the maximum stress constraint, and $k$ the constraint penalty multiplier (set to $5 \times 10^{-5}$ according to the results of formative experiments).

With this particular optimization problem, the difficulty lay not in getting a valid solution, but in getting a near optimal-mass solution. The first experiments were relatively unsuccessful in this regard: the results after 2000 generations were full of small holes and had extremely uneven inner edges. This can be seen in the typical end-of-run results shown in Figure 1 (the numbers represent the fitness values of each individual).

The stresses were concentrated at the vertical extremes of the beam, so the material in the middle contributes less toward the beam's ability to withstand the load, and therefore as we are trying to minimize the mass of the beam, the material is more usefully used at the extremes of the beam. The GA, even in this simple standard form, rapidly removed material from the middle of the cross section, and in the later stages of the experiments was observed to be moving material from low-stress areas into high-stress areas where holes were left near the extremities.

However, this first naïve GA approach took an extremely large number of evaluations to make significant progress,

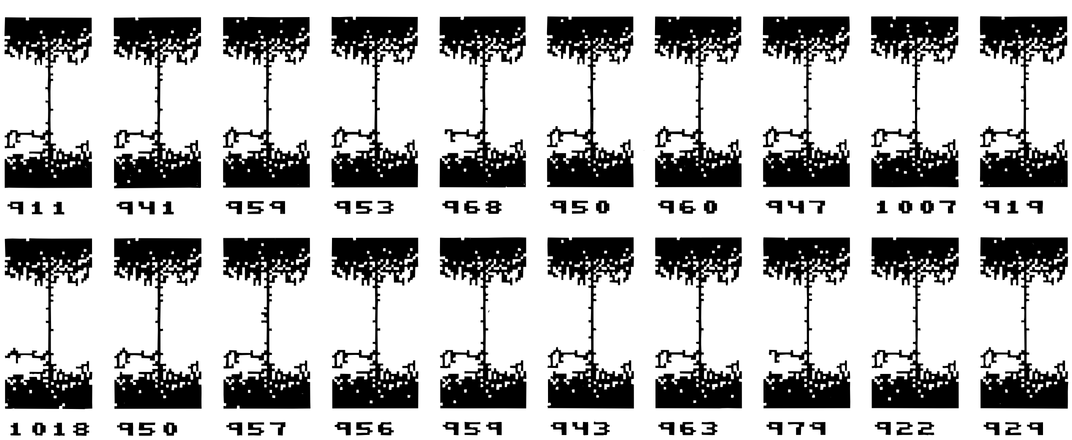


**Fig. 1.** Typical end of run results from the naïve GA.

https://doi.org/10.1017/S0890060499133031 Published online by Cambridge University Press

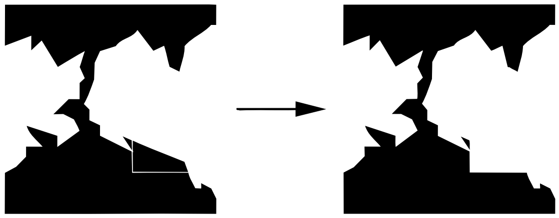

**Fig. 2.** The smoothing operator.

and this is not acceptable as later experiments would have a greatly increased evaluation time due to the integration of the FE package. The rate of improvement was also seen to decrease as the run continued, levelling off to almost none at all by the end of the run. This means that the GA was not finding any further improvements to the chromosome and, as the results are visibly poor, it indicates a general weakness in the operators being applied.

Attention was therefore concentrated towards improving the GA operators to achieve greater benefits during the early search period, and to produce better quality final results.

### 2.2. The smoothing mutation operator

The smoothing operator experiments were an attempt to address directly some of the weaknesses of the voxel representation by devising a new specialized operator, which should aid the search by reducing the number of small holes and ragged edges produced by the GA. The new operator was intended to be capable of easy expansion from two dimensions to *n*-dimensions that it would continue to be useful in the case of higher dimensional problems using the voxel representation.

This operator selects a rectangle with random position and size ranging from 2 pixels to one quarter of the dimensions of the grid. The most common value for the pixels in the area selected was then found and written to all of the pixels in that area (Fig. 2).

The GA parameters used were the same as before and the new operator was applied in addition to the previous mutation and crossover operators—application of this operator to 60% of the chromosomes in the population was found, in formative experiments, to give the best results. The GA configuration was otherwise unchanged, though the number of generations was limited to 1500 in this case.

Comparing Figure 3, which displays some typical end-of-run population members with earlier results (shown in Fig. 1), shows just how effective this domain specific approach to operator design has been, especially at eliminating isolated holes and reducing ragged edges.

### 2.3. UNBLOX: An *N*-dimensional crossover operator

The two-point crossover operator, which had been used up to this point, treated the chromosome as a 1D string of bits and therefore suffered from a problem with linkage—voxels that are adjacent in a 2D grid are not necessarily adjacent in the 1D string. This separation increases the possibility that useful building blocks (areas of the grid which contribute to a higher overall fitness evaluation) will be disrupted during the crossover procedure.

Cartwright and Harris (1993) describe the use of the UNBLOX crossover operator, which was specifically designed to overcome these limitations with conventional two-point crossover. This operator swaps a rectangular area of the grid instead of the substring swapped by two-point crossover. If the area overlaps an edge of the grid then it is made to "wrap-around" to the opposite side—this convention was adopted from the original paper, though its effect on edge smoothing is somewhat unclear. The size and location of the area


**Fig. 3.** Typical end-of-run results with the smoothing operator.

https://doi.org/10.1017/S0890060499133031 Published online by Cambridge University Press

to be swapped are selected at random, and in this implementation the area was restricted to a minimum size of two voxels per dimension so that the operator would always have some effect when applied.

The crossover operators were used with the standard probability of 0.3 per chromosome and no changes were made to the standard algorithm or to any of the other parameter settings described earlier. The graph in Figure 4 shows the results of 10 trials using three alternative crossover operators, including the UNBLOX operator. The other two crossover operators were the standard two-point crossover and uniform crossovers (Goldberg, 1989).

The results confirm that the UNBLOX operator does indeed perform better than either the two-point crossover or the uniform crossover techniques on this problem. The rate of descent of the UNBLOX line is quicker, indicating that the population converged to good solutions faster with this approach than with the other operators, and the eventual end result after 1500 generations had a slightly better fitness value than those produced by the other techniques.

## 2.4. Two-dimensional mutation operators

A new mutation operator was designed which scrambles the contents of a randomly selected rectangular area of the voxel grid, it is referred to here as the "two-dimensional" operator. This operator can be easily modified to work in *N*-dimensions, and affects a relatively small area of the chromosome rather intensively in the selected rectangular selected area in the same way as for the smoothing mutation operator. A second, somewhat altered version of this mutation operator was also designed and tested in these experiments called the "two-by-two" area mutation operator. This operator uses a fixed mutation square of two-by-two voxels and was designed to be applied only if at least one voxel in the mutation area is already active. The theory behind this operator is that most of the modifications need to be made to the surface or interior of the evolving shape and that little benefit will result from flipping isolated voxels in the middle of the void areas. The choice of a fixed two-by-two area was motivated by the observation that most of the irregularities on the surfaces would fit into such an area and that with only 16 permutations possible (4 binary bits), the probability of mutating a poor-quality area into a more fit variation would be reasonably high.

The new operators were again applied in addition to the original bitwise mutation operator, with a probability of 0.25 per chromosome of being applied. After each application there was a decreased probability of the same operator being applied again, with the probability of a further application being decreased to one half of its previous value each time. The experiments were performed 10 times for each of the 3 alternative mutation combinations, over a period of 1500 generations.

The graph in Figure 5 shows the effect of the two new mutation operators alongside the results obtained when neither of them was applied. The generation number is plotted along the horizontal axis and the average fitness of the best individual from the population at each generation is plotted vertically.

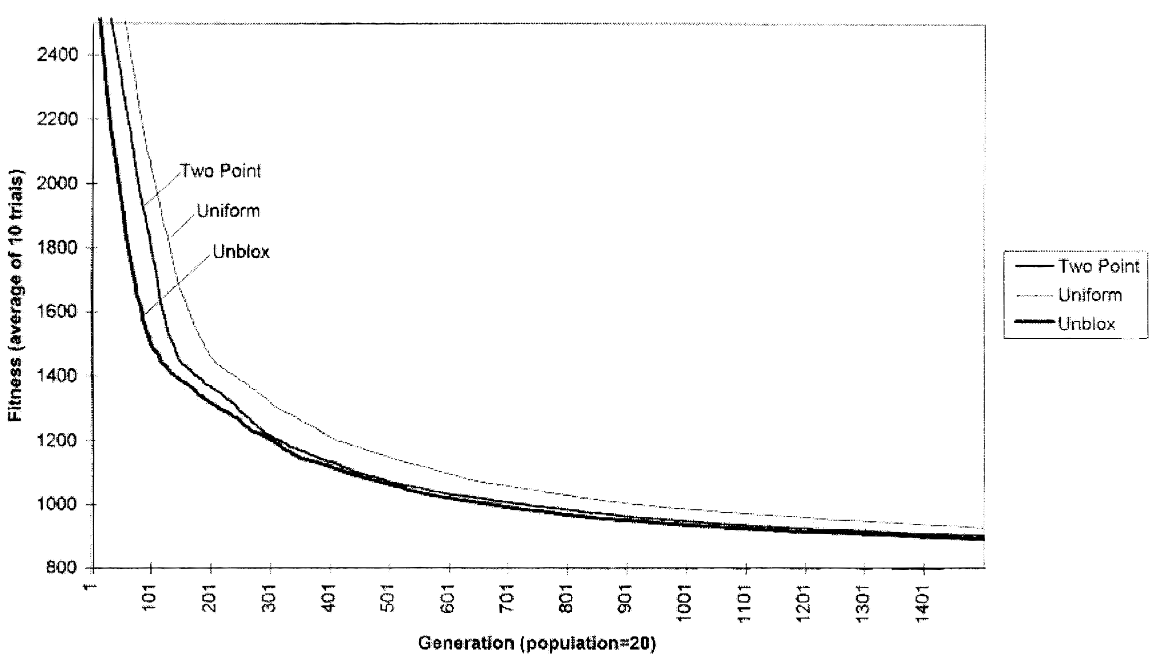


**Fig. 4.** The effectiveness of various crossover operators.

https://doi.org/10.1017/S0890060499133031 Published online by Cambridge University Press

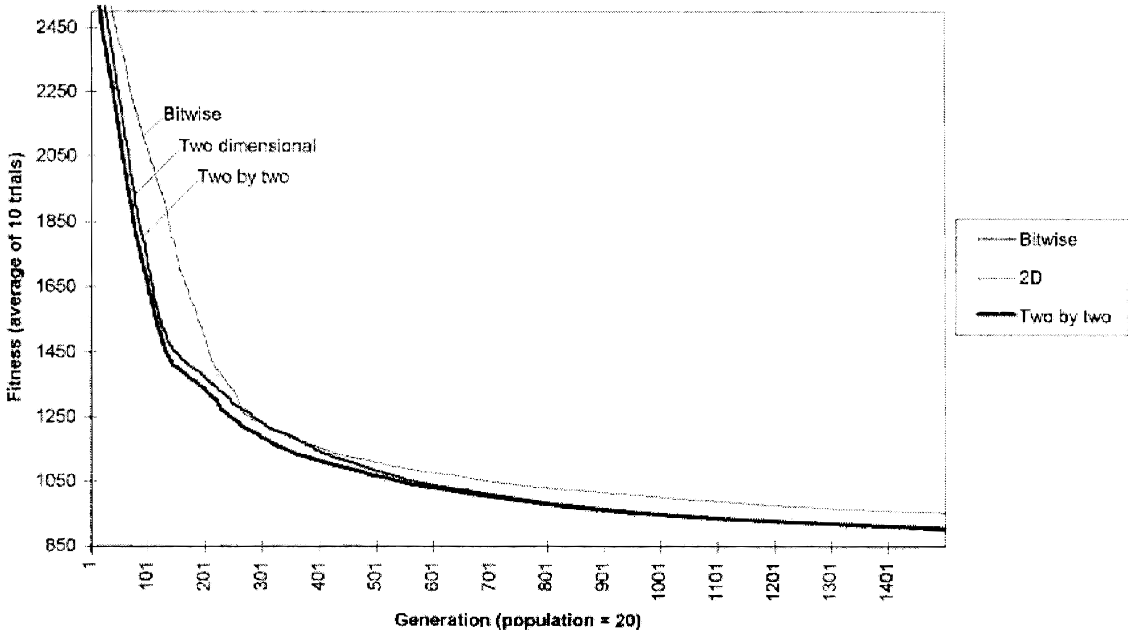


**Fig. 5.** The effectiveness of various mutation operators.

The addition of the 2D operator generally results in better performance than the bitwise operator alone, though the two lines do meet between generations 300 to 400. The steeper descent of the 2D operator line indicates that early performance was especially improved, and the final result after 1500 generations is significantly better than previously. The two-by-two operator offers a similar rate of improvement during the early stages of the trial, a slightly better performance between generations 100 to 600 and finally converges with the 2D operator's line at about generation 1000. This seems to indicate that although offering early benefits to the optimization, it is not better than the 2D operator in the long run.

In conclusion, two new mutation operators were designed with the particular intention of directly addressing the perceived problems with the prior optimizations. Both of the new operators were found to be more effective than the previous uninformed bitwise mutation, producing benefits to the rate of early improvement and the final quality of solution generated.

In the absence of any other clearly distinguishing features, the two-by-two operator will be used during the further experiments, as it offers a speed advantage over the 2D mutation operator outlined above.

### 2.5. Conclusions about the beam design problem

The results have shown that although a naïve GA does indeed suffer from the problems suggested by Watabe and Okino (1993), a small selection of operators informed only by domain knowledge about the representation will effectively solve each of these difficulties.

To see whether the above improvements can be usefully combined to produce the desired behavior, and improve further upon Figures 1 and (especially) 3, Figure 6 depicts a number of typical end-of-run results for the complete system with all operators active. Comparison with the earlier results shows that the complete system produces superior results with no holes or large proturberances. In addition, the dramatically improved performance of the final system in terms of the solution quality-time tradeoff surface it exhibits is shown clearly by Figure 7.

In summary, the final system uses a normal bitwise mutation operator in addition to the two new mutation operators, smoothing, and two-by-two. The smoothing operator rapidly cuts away unwanted areas of material during the early stages of the optimization and can help to smooth ragged edges and fill small holes later on. The two-by-two mutation operator is highly effective at smoothing off ragged edges and at filling in small holes in the material if they occur in undesirable places. Finally, the two-point crossover operator has been replaced by the *n*-dimensional UNBLOX operator, to fully exploit the 2D structure of the problem.

## 3. ANNULUS DESIGN PROBLEM USING FE ANALYSIS

The experiments undertaken with the simplified beam design problem outlined in Section 2 led to the design of ef-

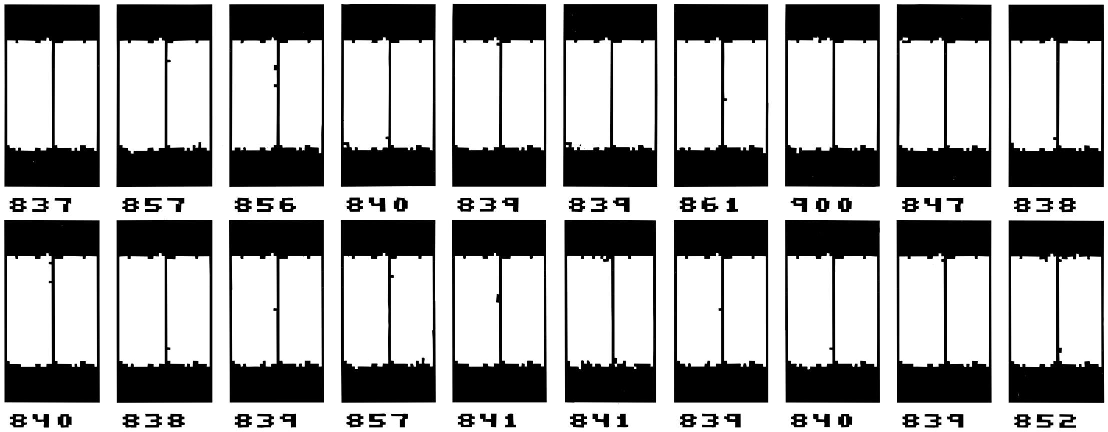


**Fig. 6.** Typical end-of-run results for the complete system.

fective GA operators for manipulation of 2D shapes. This section details further experiments undertaken to apply these operators to a more difficult design problem. The problem chosen was to design a jet-engine annulus. The finite element method was chosen as the design evaluation/analysis technique. Initially, for ease of implementation, the voxel shape description was directly used as the finite element mesh.

### 3.1. The annulus design problem

The full original specification of this problem was taken from Smith (1995*a*). The problem is to design a jet-engine annulus that is subjected to loading due to rotation and due to the attachment of the turbine blades to its outer circumference. The part is axisymmetric around the axis of rotation, and consequently it reduces to the 2D shape optimization problem shown as Figure 8.

The optimization involved reducing the mass of the annulus while observing a series of four separate stress constraints at discrete locations in the annulus. The constraints relate to the hoop stresses at the inner and outer circumferences and the radial stresses along the center line of the annulus. The stress constraints to be observed were, in descending order of importance:

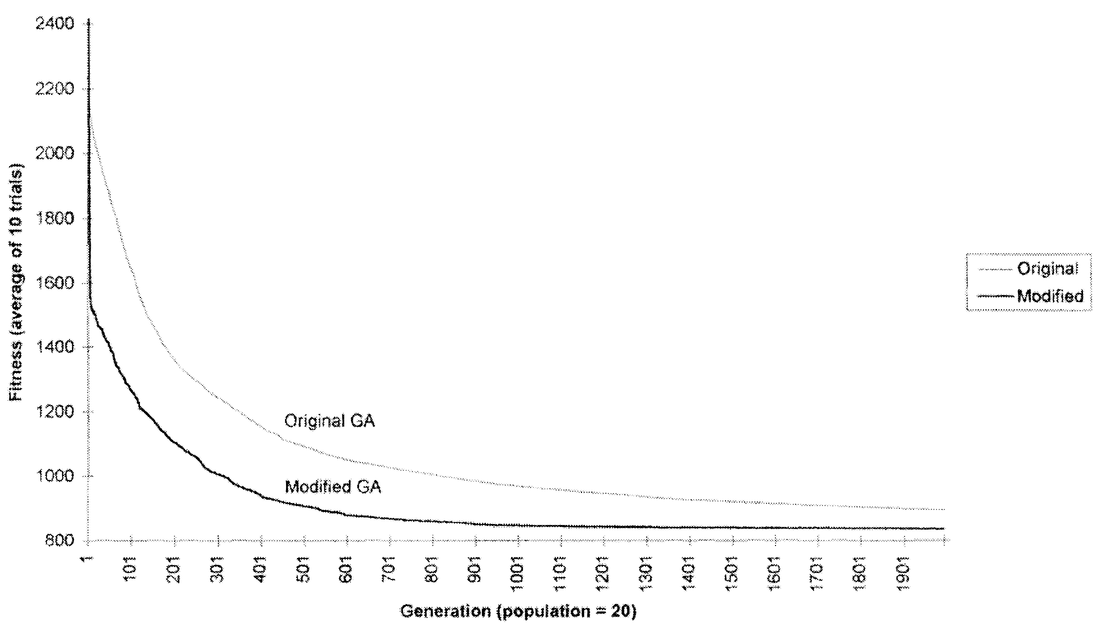


**Fig. 7.** Performance comparison between the naive and final GAs.

https://doi.org/10.1017/S0890060499133031 Published online by Cambridge University Press

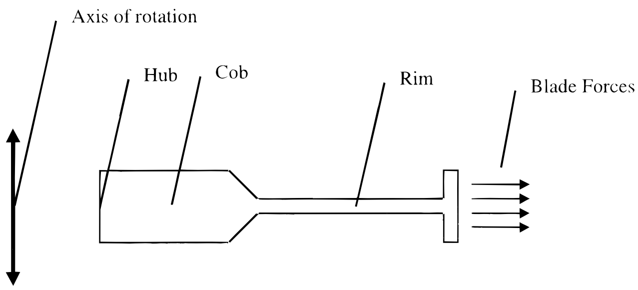


**Fig. 8.** Annulus axisymmetric cross section.

| | |
|---|---|
| Hub hoop stress | $< 1330$ MPa |
| Rim hoop stress | $< 396$ MPa |
| Inner radial stress | $< 741$ MPa |
| Outer radial stress | $< 334$ MPa |

### 3.2. The fitness function

The GA fitness function was defined as an objective (the weight of the annulus in kg, and a factor to minimize the *total* stress, in MPa) plus a sum of penalty terms if one of the 4 stress constraints was broken. The function maximized

$$F = \Sigma_i \sigma_{\max(i)} / (\Sigma_i 1000 \times S_i) - \mathit{annulus\_weight}$$
$$- \Sigma_i k \times i \times \max\{S_i - \sigma_{\max(i)},, 0\}.$$

Constraint penalties were applied if any of the four constraints limits $\sigma_{\max(i)}$ were exceeded by the stress, $S_i$, measured (in MPa). The constraints were ordered in importance by using $4 \times k$ for the most important, $3 \times k$ for the second most important, $2 \times k$ for the next and $1 \times k$ for the least important constraint, the (decreasing) order of importance was as for the constraints limits listed above.

### 3.3. Results from the basic system

Again, a generational GA with a population of size 20 and GENITOR-style rank-based selection was used. The UNBLOX, smoothing mutation, and 2-by-2 mutation operators were applied sequentially with probabilities 0.3, 0.8, and 0.8 respectively (on the basis of formative experiments). A 62-by-27 voxel grid was used to represent the annulus and the constraint penalty, $k$, was set to 0.00005. The settings used for the annulus were:

| | | |
|---|---|---|
| Dimensions of design space | = | $0.25 \times 0.05$ m |
| Radius of hole | = | 0.10 m |
| Blade force | = | $10 \times 10^5$ N rad$^{-1}$ |
| Young's modulus | = | $2.238 \times 10^{11}$ N m$^{-2}$ |
| Material density | = | $8.221 \times 10^3$ kg m$^{-3}$ |
| Revolution speed | = | 1571.0 rad s$^{-1}$ |

The basic system was first applied without further modifications to the annulus optimization. However, the problem, as specified, was very tightly constrained, which meant that the attempts to solve this problem using random population initialization violated all of the stress constraints by large amounts. Also, the rate of improvement in the population, when extrapolated beyond the time period allocated to the experiments, indicated that a valid solution would not be found for some considerable number of generations.

To circumvent this problem, the population was instead initialized with a selection of variations on the annulus design supplied with the original specification, which were modified further by an aggressive random mutation operator that added and removed small areas of material over the surface of the annulus design. This kind of intelligent initialization was thought reasonable as a user will often want to start the GA with existing designs to see what improvements can be made. Even when a totally new shape is being designed, the user would normally have some expectation about the final form, which could easily be used to initialize the population. The intelligent initialization approach meant that the initial population was not unreasonably far outside of the stress constraints, yet supplied the optimization with sufficient variation that the population did not rapidly converge onto a single solution. Some of the results from this basic system can be seen in Figure 9, which shows six members of the population after 75 generations.

The results shown in Figure 9 were poor. The lack of symmetry around the horizontal axis and the uneven edges were just the most visible failings in this set of results. A second problem was the occurrence of large stresses at the corners of elements on the edge of the shape. These failing need to be addressed if any claims as to this representation's scalability can be made.

### 3.4. Improvements made to the system

Attention was now turned to resolving the issues and shortcomings highlighted by the above investigation in turn.

#### *3.4.1. Use of symmetry*

It was known that a solution to the annulus design problem should be symmetric about a radial axis. It was there-

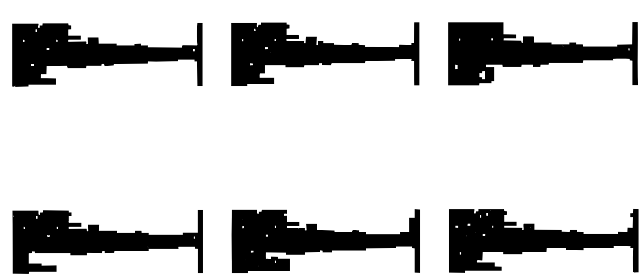

**Fig. 9.** Results of the basic annulus optimization after 75 generations.

https://doi.org/10.1017/S0890060499133031 Published online by Cambridge University Press

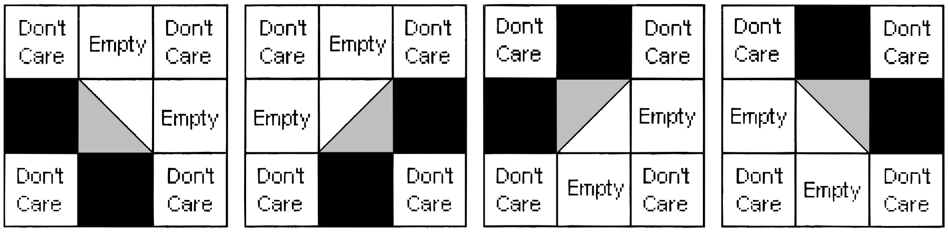


**Fig. 10.** Convolution masks for triangle insertion process.

fore decided to utilize this domain knowledge and thus reduce the search space of the problem. The GA was modified to reconstruct the final shape in its entirety only when producing the element definition files to be accessed by the FE package. This simple modification reduces the search space from a typical size of $2^{2542}$ for a 62-by-41 voxel grid, to $2^{1302}$, which represents a 62-by-21 voxel half-grid. The central line of voxels along the axis of symmetry is not mirrored as it is now enforced by the GA to be always turned on—this also provides a guaranteed central line of elements for the stress measurements to be taken from.

#### *3.4.2. Mesh improvement*

It was found in the initial experiments for the annulus design problem that directly using the voxel description of the geometry as the FE mesh caused problems with high stresses caused by corners in the mesh. It was therefore decided to separate the geometry model and mesh. There were several possible approaches that could have been taken. An approach that was considered was to use interpolation splines to form a smoothed edge. The voxels would then act as a "skeleton" and the spline as a "skin." A mesh generator could then produce a mesh whose density could then be independent of the voxel model. However, for this prototype system it was decided simply to add triangular elements at the corners. While this was a far less elegant solution, it was much simpler to implement.

These new triangular elements were created by specifying connections between groups of three nodes in the element connection file. These triangular elements were added to the shape at all suitable "steps," which were identified by convolving the voxels in the shape against a series of four matching template masks. If each square in the mask matched the value of the voxels surrounding an empty voxel then the appropriate triangular element was created in the "step." The convolution masks and the triangles that they caused to be inserted are shown in Figure 10.

#### *3.4.3. Design of operator to remove holes*

The two-by-two mutation operator (which can either fix holes or cause them to appear) was modified to only mutate areas where, as well as at least one voxel being turned on, at least one of the four voxels is also turned off. The result of this modification is that the two-by-two mutation operator can now only mutate at the boundaries of the shapes being formed, and consequently it should also help reduce the number of small protuberances.

### 3.5. Results of improved system

The improved GA for annulus optimization used the same settings as the basic system for all parameters except that the chromosomal grid was set to 21 voxels high, which is mirrored due to the symmetry used to produce a voxel grid height of 41 voxels. The analysis was permitted to continue for 114 generations and this took approximately 24 hr in total. Some of the final population created by the improved GA are shown in Figure 11. This displays 3 of the 20 individuals and shows a clear improvement in quality over the results generated previously. The small protuberances have been totally eliminated and only a few members of the population contain small holes. The rate at which a valid solution was found is considerably faster than the basic implementation, and once found, the GA continued to improve upon this solution even to the very last pass of this trial.

The annulus shapes produced can be seen to be unusual. It is proposed that the "overhangs" present at the cob and the thinness of the neck are due to the inadequate specification used for the annulus and the method used to penalize

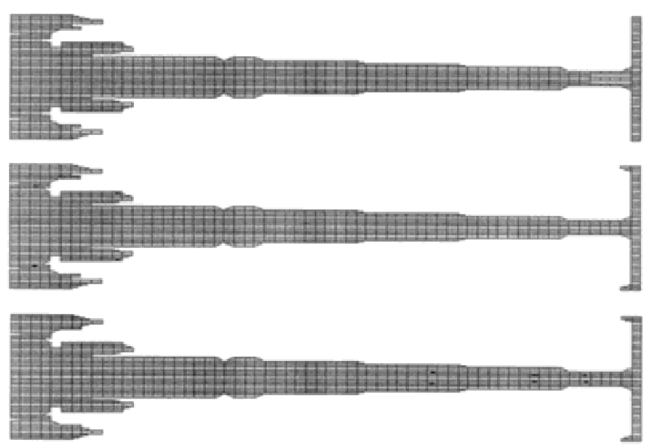

**Fig. 11.** Final annulus cross sections from improved GA.

https://doi.org/10.1017/S0890060499133031 Published online by Cambridge University Press

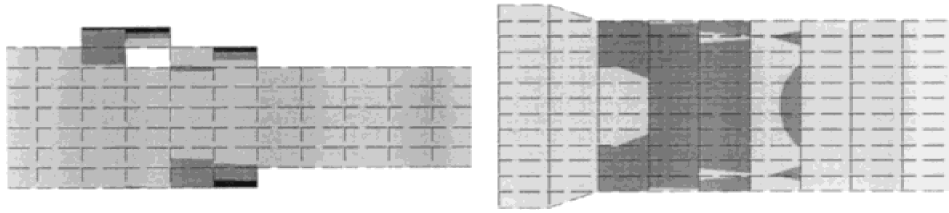

**Fig. 12.** Results without and with smoothing triangles.

constraint violation. Stress constraints were defined for four discrete points in the specification that was intended to be used with a parameterized shape description. This specification would be adequate for such a representation. However, with the voxel representation the optimizer was able to remove material with greater flexibility. At an optimal solution one of the stress constraints is just inactive. Removing more material would then increase the stress to above the maximum value. However, the GA could improve the fitness value if, by adding material elsewhere, the position of high stress was moved from the point at which the constraint was assessed, as long as the amount of material added was less than that removed. Given that this explanation is correct, the problems do not lie with the voxel representation and could be solved by improving the specification and method of penalizing constraint violation.

After using the FE package to examine the solutions produced by this optimization, it was possible to confirm that the use of the triangular elements to smooth the boundary worked as expected in reducing the amount of stress in the regions immediately surrounding a step. Figure 12 shows the stress values calculated by the FE package for the voxels surrounding steps in two typical runs and clearly shows how the triangles permit the excess stress to be distributed in a more even pattern. Darker shades indicate higher stress levels in both of these pictures.

### 3.6. Conclusions for the annulus design problem

It was found that the use of unmodified operators from the beam design problem was unsuccessful. However, when the operators were modified, taking into account knowledge held about the annulus design problem, the results were more successful (Fig. 13).

Difficulties were encountered in the direct use of the voxel shape representation as the FE mesh. These were, to some extent, alleviated by the use of smoothing triangular elements. However, the full decoupling of the primary voxel-based shape description and FE mesh would be desirable in future studies.

Unfortunately, due to the flexibility of the voxel representation in removing and adding material coupled with the GA's ability to exploit the whole search space, it was found that the specification of the problem needed to be more tightly defined, as unwanted overhangs were present in the final solution. In response it should be noted that, in the authors' experience, there are often a number of possible problem formulations for a parametric approach, each with differing suitability to the problem at hand and ability to represent only feasible solutions. Therefore, the above should not be taken to be a severe criticism of the voxel representation—for any approach, a significant amount of experimentation will be required to identify a suitably constrained problem formulation.

The unwanted overhangs aside, a comparison of the mass of the annulus produced by the voxel representation (41 kg), compares well against the original annulus design (68.6 kg), and that produced by the parametric GA described in (Smith, 1995*a)* which achieved an annulus of mass 40.9 kg. All of these annulus designs satisfied the stress constraints, though given that these designs were evaluated using different FE packages, a fine-grained comparison needs to treated with some caution.

Finally, and rather unfortunately, the voxel GA did not perform as well in regard to time to solution. The parametric GA found its solution in 400 evaluations compared to the 1000 evaluations required by the voxel-based GA—this was felt to be a result of the GA having to search a much larger and less constrained search space when using a voxel representation.

## 4. CONCLUSION

Voxels were found to be a viable representation for shape optimization with an evolutionary algorithm in 2D problems. They have a number of potential advantages over other

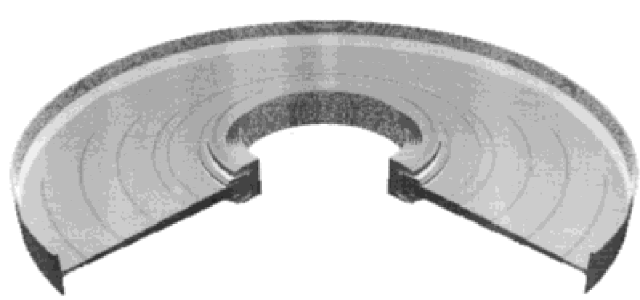

**Fig. 13.** The best annulus design from the final set of experiments.

https://doi.org/10.1017/S0890060499133031 Published online by Cambridge University Press

representations such as parameterized boundary descriptions. Topology is not predefined, domain knowledge is easy to incorporate, geometric constraints can be easily applied, and it is straightforward to convert existing solutions into such a description to "seed" an initial population of shapes.

Experiments were undertaken on two design problems to investigate the effectiveness and scalability of this representation: a simplified beam design and a jet-engine annulus design using finite element analysis. During these experiments, a number of difficulties inherent with this representation were addressed, primarily by use of specifically designed genetic algorithm operators that utilized domain knowledge held about the problems tackled. An $N$-dimensional crossover operator was used that provided linkage between adjacent rows of voxels and thus avoided the slow convergence found with a conventional crossover operator. An operator was designed to remove unwanted holes produced in candidate shapes and to smooth boundary edges.

On the annulus design problem, the direct use of the voxels as the finite element mesh was found to be inadequate, and a convolution mask-based solution to this issue was devised. That said, further work in this regard will involve the further decoupling of the voxel representation and mesh.

Furthermore, the flexibility of the voxel representation, along with the GA's exploitation of a much expanded search space uncovered deficiencies in the specification used for the annulus design problem, leading to unwanted "overhangs" in the solutions obtained. Although the results obtained were roughly equivalent in terms of the mass of annulus produced, they compared poorly with regard to the number of evaluations required to find such a solution.

Finally, it should be noted that GA optimizers can easily be modified to be used as interactive optimization systems (Tuson et al., 1997). In this case, the computer would rely on an engineer's practical experience and knowledge of the problem domain to direct key choices in the optimization process. Given the diversity of possible shape optimization problems, such flexibility will be required to deal with the constraint handling issue noted above. The lack of initial assumptions in the voxel representation could be seen to be an advantage here as the engineer has, in effect, a *tabula rasa* to start work from, and constraints on the solutions obtained can be expressed directly. Given the amount of experimentation required to find a good problem formulation for both parametric and voxel approaches, such an interactive approach would be highly desirable in any case. Further research into principled methods for allowing the user to interact with such a system is therefore recommended.

## ACKNOWLEDGMENTS

The authors thank the Engineering and Physical Sciences Research Council (EPSRC) for partly funding this work *via* two research studentships (95303677 and 95306458).

## REFERENCES


Baron, P.J. (1997). *A voxel representation for evolutionary shape optimisation*. MSc Thesis. Department of Artificial Intelligence, University of Edinburgh, Edinburgh, UK.

Bentley, P.J. (1996). *Generic evolutionary design of solid objects using a genetic algorithm*. PhD Thesis. University of Huddersfield, Huddersfield, UK.

Cartwright, H.M., & Harris, S.P. (1993). Analysis of the distribution of airbourne pollution using genetic algorithms. *Atmospheric Environment 27*, 1783–1791.

Chapman, C.D., Saitou, K., & Jakiela, M.K. (1994). Genetic algorithms as an approach to configuration and topology design. *Trans ASME, J. Mechanical Design 116*, 1005–1012.

Farrell, S.M. (1998). Gamin: A new approach to the inversion of geographical and potential-field data. PhD Thesis. Monash University, Monash, Australia.

Gere, J.M., & Timoshenko, S.P. (1984). *Mechanics of materials*, 2nd ed., Brooks/Cole Engineering Division, Monterey, California.

Goldberg, D.E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, New York.

Holland, J.H. (1975). *Adaptation in natural and artificial systems*. The University of Michigan Press, Ann Arbor.

Hsu, Y.-L. (1994). A review of structural shape optimization. *Comput Industry 25(1)*, 3–13.

Husbands, P., Jeremy, G., Ilhagga, M., & Ives, R. (1996). Two applications of genetic algorithms to component design. *Selected Papers: AISB Workshop on Evolutionary Computing*, pp. 50–61. Springer Verlag, Berlin.

Kirkpatrick, S., Gelatt, C.D. Jr., & Vecchi, M.P. (1983). Optimization by simulated annealing. *Science 220*, 671–680.

Smith, R. (1995*a*). *A parameterised Rolls Royce annulus*. Internal Report No. 21. Manufacturing and Planning Group, Department of Mechanical Engineering, University of Edinburgh, Edinburgh, UK.

Smith, R. (1995*b*). *A first investigation into a voxel based shape representation*. Technical Report. Manufacturing and Planning Group, Department of Mechanical Engineering, University of Edinburgh, Edinburgh, UK.

Tuson, A., Ross, P., & Duncan, T. (1997). On interactive neighbourhood search schedulers. *The 16th Workshop of the UK Planning and Scheduling SIG*, 207–212.

Watabe, H., & Okino, N. (1993). A study on genetic shape design. *Proc. Fifth Int. Conf. Genetic Algorithms*, pp. 445–450. Morgan Kaufmann, San Mateo, California.

Whitley, D. (1989). The GENITOR algorithm: Why rank based allocation of reproductive trials is best. *Proc. Third Int. Conf. Genetic Algorithms*, pp. 116–121. Morgan Kaufmann, San Mateo, California.


---

**Peter Baron** is a professional video games programmer who received his M.Sc. in Information Technology: Knowledge-Based Systems at Edinburgh University through part-time study, completing his thesis in 1997. His thesis "Evolutionary Shape Optimization using a Voxel Based Representation" forms the basis for this paper. He has spent a year working on a "Virtual Reality" advertising system for Intelligent Sensor Systems in East Kilbride and has recently completed a computer-based version of the board game "Stratego," released by Hasbro-Interactive in December 1998.

**Andrew Tuson** first studied for his M.A. in Chemistry at Oxford University (1990–1994). In 1995, he obtained a M.Sc. degree (with distinction) in Information Technology: Knowledge-Based Systems at Edinburgh University and continued his research there by receiving his Ph.D. in Artificial Intelligence (1995–1998). He is currently a lecturer in Computing at City University, London, U.K., where his research interests include methodologies for the principled design of neighborhood search and evolutionary optimizers, and their application to domains such as operational research and engineering.

**Robert B. Fisher** received a B.S. with honors (Mathematics) from California Institute of Technology (1974) and a M.S. (Computer Science) from Stanford University (1978). He received his Ph.D. from University of Edinburgh (awarded 1987), investigating computer vision in the Department of Artificial Intelligence. Dr. Fisher is a Reader in the Division of Informatics at the University of Edinburgh. His research covers topics in high-level computer vision, and he directs a research project investigating three-dimensional model-based vision and automatic model acquisition. He teaches general and industrial vision courses for undergraduate, M.Sc.- and Ph.D.-level students.

**Frank Mill** is a senior lecturer in manufacturing engineering at the University of Edinburgh. He has published over 30 papers on optimization in design, process planning, and manufacturing systems, for example scheduling, in the last 12 years. His primary interests are in the use of evolutionary algorithms to such problems.

**Andrew Sherlock** is presently a Ph.D. student in the Manufacturing Planning Group in the School of Mechanical Engineering at Edinburgh University. His interests lie in the application of biologically inspired processes to Mechanical design. Specifically, he is using morphogenetic approaches to shape generation, coupled with a genetic algorithm, for structural shape optimization.