Generative computer-aided design: multi-modality large-scale direct physical production

Lujie Chen\textsuperscript{a} and Lawrence Sass\textsuperscript{b}

\textsuperscript{a}Singapore University of Technology and Design, Singapore; \textsuperscript{b}Massachusetts Institute of Technology, USA

ABSTRACT
A rapid prototyping system for creating large-scale physical objects directly from computer models is introduced. Several production modalities incorporated in the system can be used to produce objects in different scales and types. Parts of the objects are generated by computational algorithms with appealing features that enable CNC manufacturing and facilitate manual assembly. Experiments have shown that man is an important factor of the production process even though the system has automated majority of part generation and fabrication. We show several large models created from the system, discuss problems associated with large-scale prototyping, and present potential applications of the proposed methods.

KEYWORDS
Generative CAD; large-scale prototyping; direct physical production

1. Introduction

Computer-Aided Design has come a long way since the advent of computer technology. Design, a human-centered cognitive activity, has been facilitated, advanced, and even partly replaced by computers through the development of sophisticated numerical algorithms. Computer technology that enables design is a consequence of extraction of replicable work embedded a design process. In early days, such work pertained to drafting, printing and error checking. Now, deeper investigation of the design process has motivated the creation of computer-based design tools that are able to automate model design and manufacturing with little human intervention. Offloading complicated but replicable design work from designers to computers leads to Generative CAD, an approach that will have considerable impact in design and manufacturing fields.

Through decades of development in the manufacturing industry, many production methods have seen a transition from a manual or semi-automatic approach to a full digital approach. Material cutting is one of most affected area; for any reasonably-sized cutting work, a CNC machine such as a waterjet cutter, a laser cutter or a plasma cutter is used instead of a semi-automatic machine such as a table saw. Automation in production also brings a new method of 3D printing, which is often considered as a prototyping tool. However, 3D printing is relatively high cost, slow in production, and hardly scalable with a limited build volume. The research community has begun addressing the need for methods to generate large-scale prototypes as alternatives to 3D printing. Novel methods are graphics based algorithms used to generate models as a collection of objects from a starting model. Current approaches are used to manufacture furniture, toys and models as interlocking, interlacing and sewn physical objects [1, 5, 10, 11].

The objective of this paper is to introduce a design-fabrication framework that is able to generate modal parts for direct physical production. The input to our generative CAD system is a computer model of a triangle mesh format (e.g. a STL file). The system can be configured to produce a physical model in three modalities: 1D contouring, 2D contouring and plate forming. Each of the modality can generate physical models of an arbitrary size, exceeding the working dimension of a particular machine. All modalities are based on Scalable Planar Structure (SPS) [14] with parts ready for fabrication by a CNC cutter, and with features to facilitate manual assembly. We show a variety of experiments to demonstrate the interactions between man, machine, and objects involved in physical production. Connection of the system with industrial applications is discussed.

2. Related work

2.1. Early work in architecture

Producing physical shape from 3D computer models was initially explored in architecture [6], where interlocking
planar structures were created to quickly form an assembly of models. Systematic methods of this approach were demonstrated as a grammar to form physical shape; the outcome was a plywood cabin built of interlocking parts [12, 13]. These early methods were different from those appeared later as rules within the system were meant to guide human during modeling; there was little automation in the process of partitioning a large model into parts, and fabricating the parts from CNC machines; nevertheless, the grammar-based methods were seminal to the idea of direct physical production; they were initial attempt to extract replicable work from such a design process.

2.2. Automated approach
An automated system to generate planar structures was first reported by Oh et al. [11]; the system worked at a small scale in generating toys and furniture. It freed a designer from tedious modeling tasks, such as drawing slots and joints to connect parts. Some limitations of the system motivated further studies and resulted in more robust computational algorithms that could work with complex shapes from a wide variety of 3D models [7, 16]. The rules incorporated dealt with different materials and different connection mechanism, including wheels, screws and hinges.

2.3. Interlocking structures
Many computational methods have been reported in the literature to generate interlocking planar structures, which are a class of physical shapes built of intersecting layers of material across two or three directions. The methods expect some sort of input to capture the shape of a computer model and then generate vastly different types of parts for manual assembly, where the difference is determined by goals of the resultant structures. For example, in an interactive system [9] a user could selectively choose to fabricate a subset of planes of a model based on relative importance of representation. In a chair-design system [15], the goals were chair strength and comfort, which determined the orientation of the slots for interconnection of parts. There was also a method to generate planes in locations based on visual goals while maintaining structural goals [4]. Visual representation of a physical model might even be enhanced by generating unequally spaced slots [8]. A more flexible algorithm allowed for the connection of non-perpendicular planar parts [17]. It was also showed that by using bendable material, physical models could be assembled from ribbon-like strips. The commonality of these automated modeling systems is that they produce an open structure from the original computer model.

2.4. Closed structures
In contrast to open structures, closed structures are physical models as a complete volume. Such structures may have real applications not achievable through an open structure. A method was described recently [1] to partition a 3D model into planar surface patches, which were physically joined by interior connectors. There are limitations in model fabrication and assembly when using the interior connectors. First, the connectors must be fabricated by a separate process from the planar parts. Second, as the connectors are inside the surface volume, assembly is difficult for the access area to the connectors is limited. The topic of automatic generation of closed structures is sparsely explored while they have strong potential for industrial applications such as mold making.

2.5. Large-scale structures
Despite lots of research activities in direct physical production, little has been done in generating large-scale physical models whose size can be many times greater than the working dimension of a machine. We started trial projects in 2012; through a series of investigations, we have gradually created a generative CAD system as a design tool for large-scale rapid prototyping. Two of our methods were published last year [2, 14], describing details of computational algorithms and human factor in the process of direct physical production. In this paper, we demonstrate the complete system at the current state; instead of focusing on algorithm realization, we show features of the system that enable large-scale prototyping and facilitate human cognition in physical production.

3. Large-scale direct physical production

3.1. 1D contouring
Assuming a 3D Cartesian coordinate where z axis points upwards, the 1D contouring method slices a triangle mesh at multiple z positions with a plane perpendicular to the z axis. The intersections of the plane with the mesh are contour lines, which can be treated as planar parts to be fabricated from a CNC cutter, e.g. a laser cutter. The parts are glued layer by layer to produce a physical model. Figure 1 shows an overview of this production modality.

3.1.1. Recursive partitioning
Several heuristic policies are applied in the method for large-scale model generation. If a contour exceeds the working dimension, L, of a cutter, the system partitions it
into smaller parts. The partition is recursively performed; for example, in Figure 2(a) a contour of the body exceeds $L$; it is first partitioned along $a-a$ section as its $x$ dimension is longer than its $y$ dimension. Both of the resulting sub-contours are still longer than $L$; therefore, they are partitioned along $b-b-b$ section as their $y$ dimension is longer than that of $x$. Subsequently, the same partitioning strategy is recursively applied on every resulting contour, and sections $c-c$ and $d$ are generated. The process stops when no contour is longer than $L$. The contours of the arm and wrist are shorter than $L$; hence, no partitioning is performed.

### 3.1.2. Staggered partitioning
Second, the locations of partition at each layer are designated rather than simply halfway through a contour. The locations are staggered at adjacent layers so that when assembled they produce a strong overlapping block [Figure 2(b)]. This heuristic policy follows the principle of masonry: when building a wall of bricks, adjacent layers are always staggered to ensure the robustness of the wall. Violating the policy would result in structures that may easily break along partitions closed positioned in several layers.

### 3.1.3. Reference holes
Third, reference holes are automatically generated in contours to facilitate alignment (Figure 3). In large-scale prototyping, alignment is the most critical problem during assembly. When a person assembles a layered structure, it is intuitive to align based on the outer rim of the contours as adjacent contours have very similar shapes; however, the intuition would cause cumulated alignment errors because despite the similarity, each layer is progressively different from the others. Over time, the cumulated error could be considerable and distorts the structure from the original computer model. Reference holes enable a person to put reference bars into the structure. During assembly, as long as the contour parts are put through the bars, relatively accurate alignment is achieved with little effort. In addition, each contour piece is generated with multiple reference holes, and because of the staggering policy, a reference hole often alternates between a left and a right piece at adjacent layers. The alternation is useful in pulling different parts together and naturally maintains a robust structure. The system is configured to generate roughly three reference holes per part; so the total number of reference holes are about three times the number of parts.

---

**Figure 1.** A statue of David produced by the 1D contouring method. (a) Mesh model. (b) 1D contours extracted from the mesh model. (c) Assembled bust made of 1.5 mm cardboard. (d) Assembled physical model consisting of 5425 parts.

**Figure 2.** (a) Recursively partition a contour longer than the working dimension of a machine. (b) Staggered partitioning locations at adjacent layers to ensure robustness.
3.2. 2D contouring

The 2D contouring method slices a triangle mesh in two directions. The extracted contours intersect in 3D space. An inner rim of each contour is generated; along with a corresponding outer rim, they form a contour part that can be fabricated. The intersections of the contours in the two directions are extended to become slots used for interlocking the parts. Figure 4 shows an overview of the 2D contouring production modality.

3.2.1. Partitioning of contours

As horizontal and vertical contours are interlocked, it is compulsory to partition one of the contours so that assembly is feasible; otherwise, contours are closed rings by default and cannot be interlocked. In our system, we choose to partition vertical contours regardless of whether their sizes are smaller or larger than the dimension of a machine. A vertical contour is partitioned into two or more parts at places where there are local extrema of the inner rim, as illustrated in Figure 5. The breaking line starts from a local extreme on the inner rim and extends to a nearby vertex on the outer rim. An in-plane slot is then created to constrain the in-plane motion of the parts.

In practice, an inner rim may contain many local extrema; in this case, a contour may be partitioned into unnecessarily too many parts; hence, some criterion has to be imposed to reduce the number of partitions. In our system, if two partitioning locations are closer than a threshold, one of the location is ignored [Figure 5(b)]. The threshold is arbitrarily set to 10% of the length of a part.

For a contour part, be it horizontal or vertical, that exceeds the machine’s working dimension, recursive and staggered partitioning strategies are applied in the same way as in 1D contouring (Sections 3.1.1 and 3.1.2).

3.2.2. Tolerance of slots

The size of a cross-plane slot (Figure 6) is determined by the thickness of a cut sheet, while a number of factors may affect the fabricated slot size. For instance, if a laser cutter is used to fabricate the parts, the laser beam will change the slot size by burning out a thin layer of material. Usually, the higher the laser power, the wider the slots. At a

Figure 4. (a) Mesh model of a horse. (b) Contours and slots generated by the 2D contouring method. (c) A physical model made of 5 mm plywood.
fixed power level, the slot size is affected by the cutting speed; the slower the speed, the wider the slots. The slot size is also affected by the material of the sheets as different materials have different response to the laser beam. Finally, the slot size is affected by the precision of the CNC machine.

Figure 5. (a) Partitions are generated at the global minimum and maximum. (b) Partitions are generated at some local extrema, not all.

Figure 6. Tolerance affects the ease of assembly.

All the above factors can be catered in one control parameter: tolerance. Assuming a 5 mm thick sheet, the slot size generated in our system is 5 mm minus a user-specified tolerance value, e.g. 0.1 mm; hence, the generated slot size is 4.9 mm. If the tolerance accurately reflects the combined effect of the aforementioned factors, the fabricated slot size will be 5 mm.

Determination of a suitable tolerance can be done by a few trials of varying tolerances. For making a large-scale model, the rule of the thumb is that when interlocking one horizontal and vertical part, a slot should be relatively loose, i.e. a person should not have to press hard in order to put the parts in place. During assembly when more parts are interlocked by multiple slots, the structure becomes increasingly tighter because each slot produces some friction; the overall friction produced by many slots can be significant, making it difficult to assemble subsequent parts. Experience shows that to successfully make a large model tolerance is the key parameter that a user should determine rather than leaving a computer algorithm to figure out.

3.3. Plate forming

The plate forming method creates a closed structure of interlocking planar parts from a solid mesh model [2]. The system extracts planar surfaces from a triangle mesh model and generates finger joints at the intersection of the surfaces to enable interlocking connection. The output is digital data of the planar surface patches with joints ready for fabrication. Basic user controls define material thickness, the number of interlocking fingers along an edge and the tolerance of cutting. The physical structure can then be assembled manually to produce a representation of the original model. Figure 7 shows an overview of the plate forming production modality.

3.3.1. Part labeling

Unlike 1D and 2D contouring, the plate forming method does not slice a mesh model in fixed directions, while the planar parts maintain the original positions in the computer model and may not have consistent orientation. Consequently, it is less intuitive to see how parts should be labeled. (In 1D contouring, parts can be labeled based on their z positions; in 2D contouring, the vertical and horizontal parts can be labeled based on their x and z positions respectively.) A feasible strategy of labeling the parts is to follow the default input sequence of the planar surface patches. As the default sequence may not indicate any connectivity between the parts, organizing the parts requires all parts to be laid out spatially; then during assembly, whichever part is needed can be retrieved from the entire layout. An example is shown in Figure 8.

The strategy becomes problematic when the number of parts increases with the size of the physical model. Figure 9 shows the parts of a larger model. Organizing the parts has to be done before assembly begins. More space is needed to store the parts and it takes longer time to retrieve a part for assembly.
An alternative labeling strategy is to label the parts based on their centroid’s $z$ coordinate, regardless of their orientation. The advantage is twofold; firstly it means that a part number is associated with the part’s spatial location; secondly it means that assembly of a structure goes naturally from bottom to top. This would result in lower demand in space because not all parts have to be fabricated and organized before assembly. Time is also saved when dealing with a subset of the entire part stock.

### 3.4. Part nesting

Part nesting refers to the process of laying out parts on planar sheets for fabrication. It is an indispensable process for large-scale model manufacturing and is achieved by computer algorithms fully automatically. One of the main goals of part nesting is to save material [3]; nevertheless, in our system, there are three nesting methods for a user to specify depending on their preference. An example of applying the three methods is illustrated in Figure 10.

The first method nests parts in an area-descending order, i.e. large parts are nested first and smaller parts are nested at a later stage. In so doing, the unoccupied regions in the sheets can fit subsequent parts with maximum likelihood. The advantage of the method is in saving material; however, the parts are completely disordered after nesting; a part may be nested on any sheet. If a large-scale model ends up in hundreds of sheets, it is extremely time consuming to find a specific part.

The second method uses the part number as the nesting sequence. Parts are nested onto existing sheets as much as possible. If a part cannot fit into any of the existing sheets, a new sheet is generated. The sequence of the parts after nesting is still disordered but parts of consecutive numbers tend to be nested on the same sheet; therefore, less effort is needed to find a part and to organize the parts.
The third method also uses the part number as the nesting sequence; in addition, it does not allow jump of parts to earlier sheets: when a new sheet is generated, any unoccupied regions in the earlier sheets are discarded from nesting subsequent parts. After nesting, the order of parts is intact with respect to the sheet number, which means that any part on the second sheet is guaranteed to have a bigger part number than any of those on the first sheet. Organizing the nested parts requires trivial effort.

With the three options, a user can determine whether to save material at the cost of some effort in organizing parts, or to save time in organizing parts at the cost of some waste in material, or even a balanced approach. We believe the best solution is application dependent and should be determined by a man, not a machine.

### 3.5. System view

There are five key factors in a large-scale direct physical production process: man, machine, object, space and time, where the last two refer to the space and time constraints. The key factors are interrelated and interact in profound ways. A successful design tool for large-scale prototyping should have features that address the key factors in order to facilitate production. We described six features of several multi-modality generative CAD methods. The relationship between the features and the key factors is depicted in Figure 11. All features are inevitably related to man, the most important factor, as they are designed to serve people during the production process.

Recursive partitioning is carried out based on the machine dimension. It affects the space and time needed to organize and to assemble the parts. It also affects the appearance and strength of the resulting object. Staggered partitioning is primarily used to strengthen the structure of the object. Reference holes are designed to facilitate alignment, thereby improving the quality of the object. Tolerance of slots is related to the machines’ configurations. A suitable tolerance simplifies assembly, thereby saving time of production. Effective part labeling and nesting saves space and time to organize the parts.

After all, the system view suggests that addressing the key factors is an essential design approach to computer algorithm and user interface development.
Figure 11. Five key factors, shown in red circles, of large-scale prototyping. Six features, shown in cyan boxes, incorporated in our system that address the key factors.

4. Results and discussions

Models of various sizes and types have been produced based on our generative CAD system. In this section, we show results of large-scale models, discussions on issues in the production process, and potential applications of the methods.

4.1. 1D contouring

A $3.4 \times 1.7 \times 1.3$ m$^3$ (length, height, width) triceratops model made of 1.5 mm cardboard was produced by five students in 45 days (Figure 12). The model consisted of more than twenty thousand parts, which were nested in the sequence of the part number and were allowed to jump in sheets. More than 1800 sheets were used in fabrication. Organizing the parts turned out to be very time consuming and required a 9 m$^2$ room [(b) and (c)].

Assembly of the parts was done manually with $5 \times 5$ mm$^2$ reference holes to guide the alignment (d). Parts were glued layer by layer. Due to the size of each layer, it is efficient to have several students working on assembly at the same time, each on a region of the layer, so that they do not have to move around. Four students can finish about 50 layers in a day while if one student did it alone, he could only finish 5 layers before getting exhausted because moving around is time consuming and surprisingly exhausting.

Issues arose when making the head of the triceratops (f). It was overhanging and started to bent down due to weight. At first, the students remade about 80 layers, trying to correct the deviation. It did not work well as the remade layers could not be joined with the lower layers seamlessly. The most effective fix turned out to be support the head by a stool (g). In this case, the deviation was corrected gradually by each added new layer. When the parts were pushed into the reference bars, they gradually pulled the head back to the desired position, given that the support made the entire layer relatively free from bending.

4.2. 2D contouring

A $1 \times 0.8 \times 0.4$ m$^3$ (length, height, width) horse model made of 5 mm plywood was produced by one student in 5 days (Figure 13). The tolerance of slots was calibrated based on machine settings. When the student assembled the model, he felt that it was straightforward to fit the parts into the slots at the beginning but was increasingly difficult to do so. This actually suggested that the tolerance was suitable. Had the tolerance been too small or too big, he would feel that the structure was too loose or too tight to assemble.

A larger, two-meter long model made of the same material was shown in Figure 14. Nearly 500 parts were generated. Many in-plane slots were generated to partition complete contours into smaller parts so that they could be nested on 600 by 450 mm$^2$ sheets. After one week of assembly, we found that this model could not stand on its own. The instability of the structure is due to three factors. First, each leg of the horse was partitioned into several sections, which were connected by in-plane slots. As such slots do not constrain out-of-plane motion, a twist of the model can lead to dislocation of the parts. Second, the head of the horse was too heavy to be sustained by its body; external support had to be introduced, as shown in (a). Third, it is difficult to fix the midsection of the horse; when working on one side of the midsection, parts on the other side may detach, as shown in (b).

Instability of the large-scale models produced by 2D contouring suggests an inherent drawback of the method.

4.3. Plate forming

Plate forming may be applied to concrete casting, widely used in building construction, to create formwork, i.e. mold. There exist two approaches to concrete casting: first, casting based on system formwork for regular-shaped structures (e.g. straight wall or staircase, rectangular pillar, etc.); second, casting based on customized formwork for irregular-shaped structures (e.g. curved wall or staircase, dome, etc.). At present, customized formwork is made by skilled workers using semi-automatic tools, such as table saws. The process is slow, labor demanding and prone to large precision errors.

We have applied the plate forming method to create various molds. Figure 15 shows a dome made from our system. Parts of the dome was assembled based on
Figure 12. The process of making a $3.4 \times 1.7 \times 1.3$ m$^3$ triceratops model.

Figure 13. A medium-sized, $1 \times 0.8 \times 0.4$ m$^3$, horse model.

Figure 14. A two-meter long horse model.
Figure 15. The process of making a dome.

Figure 16. The process of making a modular house.
automatically generated assembly map (a). Each part was interlocked with its neighbors through finger joints (b). Edges of the parts were sealed using glue to prevent leaking during casting (c). Plaster was pulled into the mold (d) and left there for about an hour to settle down. The parts of the mold were peeled off from the plaster dome (e). Finished dome and its mold (f). It is said that a dome is one of the most difficult structures in concrete casting because of its irregular shape and difficulty in support. Though we did not address the issue of support, the formwork parts produced by our system based on the digital methods are more accurate than those produced by hand, and the fabrication process is more efficient.

A three-story modular house was also produced using the system (Figure 16). Several students modeled the components of the house in CAD software, exported the components in STL files, and then used the plate forming method to generate a mold for each component. The molds were fabricated and assembled (a). After concrete casting, the physical components were de-mold (b) and (c). They were further assembled to produce the basic structures of the house (d) and (e). (f) to (h) show sections of the first floor. The entire house model was 1.5×1×1 m³. Through this experiment, we explored the feasibility of using plate forming to rapidly produce modular structures. For now, a complex structure has to be partitioned by human, while our generative system produces the formwork parts. Algorithms of higher intelligence may be developed in future to automate the process of modular partitioning.

5. Conclusion

A large-scale direct physical production system based on the generative CAD approach is described. Compared to conventional CAD systems that are designed to facilitate keyboard-and-mouse-based drafting, the new system does not require manual drafting as it automatically generates parts from a computer model and interfaces with a CNC machine for fabrication. Three production modalities are introduced. 1D contouring resembles an additive manufacturing process while features such as partitioning and reference holes are created to guide manual assembly of large models. 2D contouring produces interlocked structures; stability of a large model presents to be an issue that deserves further research. Plate forming produces closed structures and is promising to be used as a mold-making method with applications in construction.

Acknowledgement

This work was supported by the International Design Center of Singapore University of Technology and Design (SUTD). We thank students of MIT and SUTD for their contribution in creating the physical models. They are Ines Ariza, Elisabeth Boles, June Kim, and Calvin Zhong from MIT, Yanchao Wang, Nuo Lei, Cong Liu, Huijuan Zhou, Wei Pan, Lu Xing, Shaun Phua, and Darren Lee from SUTD.

ORCID

Lujie Chen http://orcid.org/0000-0002-9153-9348
Lawrence Sass http://orcid.org/0000-0003-0584-6382

References


