A survey of manufacturing oriented topology optimization methods

Jikai Liu, Yongsheng Ma*

Department of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada

ARTICLE INFO

Article history:
Received 12 February 2016
Revised 18 July 2016
Accepted 24 July 2016
Available online 29 July 2016

Keywords:
Topology optimization
Manufacturability
Machining
Injection molding/casting
Additive manufacturing

ABSTRACT

Topology optimization is developing rapidly in all kinds of directions; and increasingly more extensions are oriented towards manufacturability of the optimized designs. Therefore, this survey of manufacturing oriented topology optimization methods is intended to provide useful insight classification and expert comments for the community.

First, the traditional manufacturing methods of machining and injection molding/casting are reviewed, because the majority of engineering parts are manufactured through these methods and complex design requirements are associated. Next, the challenges and opportunities related to the emerging additive manufacturing (AM) are highlighted. SIMP (Solid Isotropic Material with Penalization) and level set are the concerned topology optimization methods because the majority of manufacturing oriented extensions have been made based on these two methods.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In the past two decades, topology optimization has become an active research field and the related algorithms developed create a powerful approach to perform innovative and efficient conceptual design activities [148]. To be specific, topology optimization is powerful because the related algorithms have been applied to a broad range of design problems governed by different physical disciplines, i.e. solid mechanics [1,8,121,134], fluid dynamics [9,151], and thermal dynamics [43,135,154] etc. Many algorithms developed are innovative because they can help engineers to think out of the box to generate innovative design ideas, even for those designs of already highly engineered products [148]. Furthermore, topology optimization is efficient because automated optimization processes are employed to generate the conceptual designs instead of the conventional trial-and-error approach.

Currently, SIMP (Solid Isotropic Material with Penalization) [8], ESO (Evolutionary Structural Optimization) [134], and level set [1,8,121] topology optimization methods represent the main streams. These methods have their unique characteristics and at the same time, are tightly associated. There have been a few comprehensive reviews in the literature [26,28,89,98,99,112,117]. ESO is categorized as a hard-kill method which iteratively removes or adds a finite amount of material. Heuristic criteria are employed which may or may not be based on the stringently calculated sensitivity information. Therefore, ESO is relatively simple in implementation which demonstrates advantages for topology optimization problems involving complex physical processes. For instance, Naceur et al. [90] designed the initial blank through ESO which involved finite element analysis of the sheet metal forming process. Azamirad and Arezoo [6] optimized the stamping die through ESO which performed numerical simulation of the sheet metal forming process through Abaqus. Shao et al. [104,105] optimized forging preforms though ESO and the forging process was simulated through the DEFORM 2D. However, as summarized in [89], there is almost no implementation of the ESO method to address other manufacturing oriented topology optimization problems. Therefore, this survey paper pays more attention to SIMP and level set methods, because the majority of manufacturing oriented extensions have been so far developed based on these two methods.

The typical compliance minimization problem based on SIMP method is demonstrated (see [8] for more details) in Eq. (1).

\[
\begin{align*}
\text{min. } C &= U^T K U = \sum_{e=1}^{n} \rho^e \mathbf{u}^e \mathbf{u}^e = \sum_{e=1}^{n} \left( \rho^e \right)^p \mathbf{u}^e \mathbf{u}^e \\
\text{s.t. } V &= \sum_{e=1}^{n} \rho^e v_0 \leq V_{\text{max}} \\
K U &= F \\
0 < \rho_{\text{min}} \leq \rho^e \leq 1
\end{align*}
\]

where \( U \) and \( F \) are the global displacement vector and loading vector, respectively. \( K \) is the global stiffness tensor. \( \mathbf{u}^e \) is the element displacement vector and \( k^e \) is the element stiffness tensor after density interpolation. \( k_0 \) and \( v_0 \) are the stiffness tensor and
material volume of a solid element, respectively. \( \rho^* \) is the element density and \( \rho_{\text{min}} \) is the lower bound. \( V_{\text{max}} \) is the upper bound of the in total material volume.

It is worth noting that, compared to direct density optimization, interpolation of the nodal or point-wise densities with material properties is also an effective approach [36,56,81] which has contributed to achieving the manufacturing-oriented topology design, such as minimum length scale control for machining and bidirectional material change for injection molding/casting. More details will be discussed in later sections.

Comparatively, the typical problem formulation based on level set method is demonstrated in Eq. (2).

\[
\min C = \int_D A e(u) H(\Phi) d\Omega \\
\text{s.t. } a(u, v, \Phi) = \langle 1(v, \Phi), \forall v \in U \rangle \\
V = \int_D H(\Phi) d\Omega \leq V_{\text{max}}
\]

\[
a(u, v, \Phi) = \int_D A e(u) e(v) H(\Phi) d\Omega
\]

\[
l(v, \Phi) = \int_D \rho \mathbf{F}(\Phi) d\Omega + \int_D \tau \mathbf{F}(\Phi) d\Omega
\]

in which \( u \) is the displacement vector, \( v \) is the test vector, and \( e(u) \) is the strain. \( U = \{ v \in H^1(\Omega) : \forall v = 0 \text{ on } \Gamma_D \} \) is the space of kinematically admissible displacement field. \( A \) is the Hooke's law for the defined isotropic material. \( p \) is the body force and \( \tau \) is the boundary traction force. \( \Phi \) is the level set function, which is defined by Eq. (3).

\[
\begin{align*}
\Phi(X) &> 0 \quad \forall X \in \Omega \quad \text{(material)} \\
\Phi(X) &< 0 \quad \forall X \in \Gamma \quad \text{(interface)} \\
\Phi(X) &< 0 \quad \forall X \in \partial \Omega \quad \text{(void)}
\end{align*}
\]

The adopted Heaviside function \( H \) and the Dirac Delta function \( \delta \) are defined in Eq. (4) and Eq. (5), respectively.

\[
\begin{align*}
H(\Phi) &= \begin{cases} 
1 & \Phi > 0 \\
0 & \Phi \leq 0
\end{cases} \\
\delta(\Phi) &= \begin{cases} 
0 & \Phi \neq 0 \\
\int_{-\infty}^{+\infty} \delta(\Phi) \, d\Phi = 1
\end{cases}
\end{align*}
\]

By comparing Eqs. (1) and (2), the distinctions between SIMP and level set can be observed. SIMP method employs the element or nodal densities as the optimization variables, which is referred to an element based method. It can freely generate topology changes and have fast and stable convergence; however, the derived structural boundary tends to be blurred and staggered. Level set method defines the material domain by the positive level set field and the structural boundary by the zero-value level set contour, which is categorized as a boundary based method. However, because of the boundary based structural evolution, the employed shape derivative generally only leads to shape deformations if no interior void exists inside the design domain, and topology changes are usually forced by predefining interior holes or applying topological derivative.

On the other hand, as revealed by other authors [112], the distinctions between the two methods are in fact not that fundamental. It is typically claimed that level set method employs clear-cut and smooth boundary representation, but in most implementations, the boundary elements are modeled through the approximated Heaviside projection and what passed into the finite element model is actually blurred. In addition, density field projection is quite commonly applied these days [36,110] which makes it similar to the Heaviside projection of the level set function. As mentioned in [112], hybrid application of these two methods may be a trend of future research which could benefit from the advantages of both methods.

From the perspective of software implementation, topology optimization has been embedded as a module of most commercial CAD/CAM systems, e.g. the OptiStruct [4] from Altair HyperWorks, and the SIMULIA Tosca Structure [25] applied in Abaqus, ANSYS, and MSC Nastran. Additionally, some advanced toolkits have been released by academic research groups, e.g. the TopOpt (http://www.topopt.dtu.dk) from TopOpt research group and the PareTOWorks (http://www.ersl.wisc.edu) from Engineering Representations and Simulation Laboratory, etc. Other than that, several Matlab programs can be found in academic publications [5,17,73,109,133,138,139,157].

In summary of the published research works and the released software tools, the existing effort concentrates on the following aspects [112]: (1) low CPU time; (2) generality of applicability; (3) reliability; (4) simplicity of implementation; and (5) simplicity of topologies obtained. From the authors’ interest, the aspects (2) and (5) are highlighted, because they could make topology optimization friendly to manufacturing. Topology optimization pursues the result optimality and generally produces complex topologies, which can only be manufactured through additive manufacturing (AM). However, in practice, AM is only an emerging technique while the conventional manufacturing methods such as machining and injection molding/casting still dominate the manufacturing sector. Therefore, coming back to the aspects (2) and (5), topology optimization problems should be formulated and solved with careful considerations of the manufacturing requirements, in order to generate simple topologies which are manufacturable through the conventional manufacturing methods.

The main body of this paper is organized as: Section 2 reviews the efforts to parameterize the topology design; Section 3 looks into the machining-oriented topology optimization methods which highlights two aspects, i.e. length scale control and geometric feature-based design; Section 4 reviews the injection molding/casting-oriented topology optimization methods which highlights two aspects of part ejection and rib thickness control; In Section 5, future research directions are proposed and the underlying challenges and opportunities are discussed. At the end, the conclusion is given.

2. Parameterization

Topology optimization originates as a discretized computational design method. For this reason, it generates topology designs in tessellated and even blurred form, which may not be manufacturable or very costly to do so [12,24]. To make the topology design manufacturing-friendly, post-processing is usually required to identify, smooth and parameterize the structural boundary. Surface smoothness enables smooth tool path and thus the fast machining process. In contrast, a tessellated surface requires many tool path turnings and yet the ravines are non-machinable by economical cutters. Parameterization is also important because a smooth surface does not always guarantee good manufacturability. For instance, undercuts are non-machinable or require special tools. Therefore, further shape editing is generally required which needs to be facilitated by the parameterized surface definition.

A simple approach to is manually reconstruct the topology design through parameterized solid geometry modeling. However, because of the uncertainties of the manual operations, a following sizing/shape optimization is required to ensure the result optimality [18]. A more desirable approach is to automatically smooth and parameterize the topology design through an integrated optimization algorithm. Beneficially, design efficiency could be improved and the subsequent sizing/shape optimization may be eliminated. Therefore, many research efforts have been spent on this advanced
approach and this section is dedicated to the parameterization aspect of topology optimization.

2.1. SIMP method

SIMP method employs the discrete element densities as the optimization variables and therefore tends to generate a staggered grey image of the topology design. To make it manufacturable, three post-treatment steps are required: identify the topology design as a black-and-white image, smooth the structural boundary, and then realize the parameterization.

In early 1990s, the topology design was purely applied as a conceptual idea [792] and it relied on the designers to make the interpretation. This approximated approach is somehow arbitrary and inaccurate. To overcome this limitation, image processing techniques were applied in some following works. Bremicker et al. [11] processed the grey image into a binary map, which could be post-treated into a truss structure through skeleton extraction or a continuum structure through boundary smoothing. Lin and Chao [67] applied a similar process to obtain the binary map and did the repairing to remove the noisy elements and voids, after which the external boundary was interpreted by B-splines and the interior holes by predefined shape templates (see Fig. 1). Yildiz et al. [136] developed a neural network based image processing method to recognize the hole features from the topology design. In summary of these works, image processing techniques have been employed to translate the blurred topology design into a black-and-white image according to the threshold density value. All trusses, shape templates and B-splines could be used to smooth and parameterize the binary map which is application-dependent.

Other than image processing, interpretation based on the iso-density contour is direct and was also extensively investigated. Maute and Ramm [84] interpreted the topology design by extracting the iso-density contours through cubic or Bezier splines. Because of the checker-board pattern, the modeling process is cumbersome and could produce impractical results. Young and Park [137] developed a density redistribution algorithm to suppress the checker-board pattern, through which the iso-density contours could be more easily extracted. Hsu et al. [51] fully automated the iso-density contour extraction for both 2D and 3D cases [52]. To be specific, the 2D iso-density contours were interpreted by B-spline curves; for 3D cases, the sectional 2D contours were swept to form the 3D surface. A novel continuity analysis technique was developed to ensure the reasonability of the extracted 2D iso-density contours [53]. This method is feasible but has limitations in dealing with complex topologies, e.g. small topology details are ignored and it is non-trivial to determine the sweeping direction. Koguchi and Kikuchi [61] developed a surface reconstruction algorithm, which applied the marching cubes method to identify the iso-density surfaces. Crease and corner features were detected and therefore, the boundary segments could be individually modeled by spline surface patches. This method employs better accuracy in recovering the 3D solid geometry compared to the contour sweeping method. For instance, C0-continuity surface could be modeled. More importantly, it matches the B-rep (boundary representation) based solid geometry modeling which is the basis of many CAD tools.

Other than the iso-density contour extraction, Tang and Chang [114] and Chang and Tang [18] developed a model reconstruction method. Based on a bitmap, the staggered structural boundary was reconstructed by averaging the boundary nodes and the least square fitting was performed to fit the averaged boundary nodes (see Fig. 2). Jang et al. [55] implemented a similar process to design compliant MEMS mechanisms.

2.2. Level set method

As mentioned earlier, level set method belongs to a boundary based structural optimization method and the smooth structural boundary can always be identified throughout the optimization process, which is an advantage compared to the SIMP method. For instance, the maximum local curvature can be constrained according to the smallest cutting tool radius to ensure the derived surface machinable. On the other hand, in implementation, the level set field is mapped to the finite element model through the approximate Heaviside projection, where intermediate densities exist among boundary areas and the analysis accuracy is reduced, such as stress evaluation in high-curvature boundary areas. To fix this issue, a popular approach is to implement the X-FEM (Extended Finite Element Method) under the level set framework, which realizes the solid part integration of the boundary crossed elements through local enrichment and is effective in modeling the moving discontinuity [118,119,127]. In this way, the clear-cut and smooth
boundary representation can be better claimed by the level set method. It is noted that there is a large number of topology optimization implementations by combining X-FEM and level set. They will not be specified here and interested reader can refer to the review papers [112,117].

Other than the clearly identified structural boundary, level set function is possible to be parametrically defined, with which coefficients of the basis functions will be optimized instead of the discrete level set values. One group of the parameterized level set functions is the radial basis functions (RBFs) in different forms and orders [13,48,49,74,77–79,122,123]. Smoothness of the RBFs ensures the inherited smoothness of the structure boundary. As presented in Eq. (3), the Wendland’s CS-RBF with C2-smoothness is demonstrated [77,78].

\[
\Phi(x, a) = \sum_{i=1}^{n} \psi_i(x) a_i
\]

\[
\psi_i(x) = [\max(0, (1 - r))]^2 (4r + 1)
\]

\[
r = \frac{d_i}{d_p} = \frac{\sqrt{(x - x_i)^2 + (y - y_i)^2}}{d_p}
\]

where \(\psi_i(x)\) and \(a_i\) are the ith RBF and the ith expansion coefficient, respectively; \(r\) represents the radius of support and \(d_p\) is the predefined support radius. The RBF-based level set method has also been applied to solve fluid flow problems [63,94], metamaterial problems [124], multiphase actuators [80], and multi-material problems [126], etc. Alternatively, B-splines can also be applied for the parameterization purpose [19,20].

The other benefits of the parameterized level set method include: creation of new holes (2D), free of velocity extension and no regular re-initializations [77,122].

2.3. Spline based methods

In SIMP method, splines are employed to post-process the topology design, while in level set method, the level set function can be parametrically defined based on the B-spline basis functions. Other than that, there are also purely spline based topology optimization methods. These methods represent the structural boundaries by spline curves/surfaces and certain hole generation rules are employed for topology changes, e.g. the bubble method [27], the ESO method, and the topological derivative method, etc. Spline based methods are very promising because the smooth and parameterized boundary representation is naturally there and the spline based curve/surface representation is the basis of most commercial CAD tools. However, spline based methods still need more research, especially compared to the popular SIMP and level set methods.

Cervera and Trevelyan [14,15] employed the ESO method to generate voids. The structural boundary and interior voids were modeled by NURBS (non-uniform rational B-splines), of which the control points were optimized for shape control. This method was successfully applied to both 2D and 3D cases. Lee et al. [65] employed the B-spline curves for boundary representation and a topological derivative based selection criterion to generate voids. This method was also used to solve electrostatic problems [60]. Weiss [129] developed a CAD based structural optimization method. The CAD specification tree was defined as the design object and its configuration was optimized by the evolutionary algorithm, which addressed both shape and topology changes.

These days, the iso-geometric topology optimization is emerging as a popular approach. Iso-geometric analysis unifies the CAD model and the numerical analysis model through the same spline information [54] and therefore, promotes the CAD/CAE integration. Seo et al. [102,103], for the first time, extended the iso-geometric shape optimization to topology design. It employed a similar way as the bubble method to insert holes and the trimming technique was used to attach new spline patches (see Fig. 3). Currently, iso-geometric topology optimization is under active exploration and it can be foreseen a burst of publications in the near future.

2.4. Limitations

The primary motivation of parameterization is to identify and smooth the structural boundary and facilitate the following detail editing under the CAX system for a polishing process. This process includes the CAD model construction, shape optimization if necessary, CAE validation, and CAM simulation, etc. The CAD/CAE/CAM tools are seamlessly integrated and can work in a collaborative manner [82], which ensures the design efficiency. However, topology optimization, as part of the design process, acts as a standalone tool without effective interface with the CAX system. Intermodel information transfer and translation rely heavily on human intervention. Therefore, it is still under exploration about how to transform the parameterized topology design into a “ready” CAD model. The “ready” word here means the CAD model could be fully supported by a widely-adopted commercial CAD tool, e.g. Solidworks, NX, Pro/Engineer, and Catia, etc. Few publication can be
found in this aspect. For instance, Chacon et al. [16] developed a program to translate the topology design into IGES format, a neutral CAD format compatible to most commercial CAD tools. In addition, the SIMULIA Tosca Structure [25] has the post-treatment module named “Tosca.Smooth”, which smoothes the iso-surface and could generate STL/IGES/STEP files. As well, OptiStruct [4] software package has a similar module named “OSSmooth”. However, commercial CAD tools read in the neutral files in “frozen state” and cannot perform the effective editing. For the reason, freeform surfaces may be formulated by Bezier, B-spline or NURBS with different limits for degree [33] and the difference causes compatibility issues between the neutral formats and the proprietary CAD formats. Another problem is that, low-order boundary elements of C0-/C1-continuity rarely appear in the topology design, even though they are widely adopted by mechanical part design and friendly to model manipulation and manufacturing.

3. Machining oriented topology optimization

Manufacturing rule violations are very common in topology optimization based conceptual design solutions, which negatively impact the manufacturability and even make them non-manufacturable. For instance, interior holes are non-manufacturable by either machining or injection molding/casting, but they frequently appear in the topology design for superior structural performance. These violations complicate the CAD model construction and the following shape optimization because it is non-trivial and somehow arbitrary to remove these violations. Therefore, it is necessary to carefully consider the manufacturing rules when configuring the topology optimization problem and its solution strategy.

In this section, machining oriented topology optimization methods are carefully reviewed and commented.

3.1. Length scale control

Length control is significant in guaranteeing the machinability: the void size should be controlled bigger than the minimum cutter size and too small components should be avoided because they may cause machining difficulties.

About length scale control in topology optimization, the pioneering works can be tracked back to the filtering method [107] and the local gradient constraint method [93]. These methods were developed mainly to eliminate the checker-board
In order to strictly constrain the minimum length scale, Guest et al. [36] developed the Heaviside projection method. Nodal densities were defined as the optimization variables and the element densities were calculated through Heaviside projection of the nodal values, where the minimum length scale is embedded in the operator. However, this method was only effective in controlling the minimum length scale of components but not the voids. Later, this method was modified into double projections, which effectively restricts the minimum length scales of both the component and void phases [37]. Sigmund [110] developed a morphology-based density filtering strategy, which enables black-and-white design solutions and well-constrained minimum length scales. Four morphology operators: erode, dilate, open, and close, were developed to control the single-phase minimum length scale. Two enhanced operators: open-close and close-open, were developed for double-phase minimum length scale control (Table 1), but the related sensitivity analysis cost is outweighed because “the sensitivity result of one element is a function of its neighbors, which are functions of their neighbors, and so on four times” [110]. This task is computationally even heavier than the involved finite element analysis. Fig. 4 demonstrates a few examples of applying these operators. Later, a robust topology optimization method [100,111,120] was developed according to these morphology operations. Multiple design realizations were evaluated while the worst case was optimized. Better local length scale control can be realized for both phases [100,120] if the multiple realizations keep a consistent topology, even though it is not always true [150]. In addition, this method had the drawback of performing multiple finite element analyses in each optimization loop.

Another obstacle in the density filters, it is also widely applied of constraints. Poulsen [95] developed the MOLE (MOnotonicity based Minimum LEnghth scale) method for minimum length scale control. This method relied on the local integral constraints checking the monotonic density variations. By satisfying the local constraints, the minimum length scale is explicitly satisfied for both the components and voids. Zuo et al. [156] applied a minimum hole size constraint to remove the small hole features from the topology design. Guest et al. [38] constrained the maximum component length scale by restricting any circular area in diameter of the length scale upper bound not fully filled. More recently, Zhang et al. [143] realized the simultaneous maximum and minimum component length scale control through addressing the structural skeleton based constraints. An image processing technique was employed to extract the structural skeleton and local constraints were constructed based on the maximum sizes of the skeleton-centered and structure-inclusive circles. In Zhou et al. [150], the structural indicator function based geometric constraints were developed based on the filtered and physical density fields. Strictly-satisficed minimum length scales were realized for both phases and more importantly, it does not require multiple finite element analyses in each optimization loop as compared to the robust topology optimization [100,111,120].

Level set method is also capable of length scale control, and in some aspects, it has demonstrated unique characteristics. Chen et al. [21] and Luo et al. [76] applied the quadratic energy functional as part of the objective function for shape feature control, which successfully realized the strip-like design in expected component length scale. Liu et al. [68] developed a simple thickness control functional in deriving constant rib thickness. Guo et al. [41] realized the concurrent maximum and minimum component length scale control through the structural skeleton based method, which is similar to [143]. The signed distance information facilitated the structural skeleton extraction of a narrow band, and level set values of the skeleton elements were constrained for length scale control. Xia and Shi [131] further explored the structural skeleton based method, in which the trimmed structural skeleton and the concept of maximal inscribable ball were employed to measure the length scale. The structural skeleton was still extracted based on the signed distance information, but discretized points were identified instead of a narrow band, through which the length scale constraints could be more directly applied to the structural boundary points. Allaire et al. [3] explored the thickness control mechanism in depth, with diversified schemes of maximum thickness only, minimum thickness only and also the hybrid manners; additionally, a comparative discussion between thickness control constraints and functional was given.

In comparison of the methods for length scale control, it seems more natural to apply the level set method, because it defines the entire design domain by signed distance information which eases the component/void size measurement. For instance, some level set based implementations have realized the concurrent maximum

<table>
<thead>
<tr>
<th>Density filter</th>
<th>Equation</th>
</tr>
</thead>
</table>
| Heaviside projection [36] | \[
\rho^v = 1 - e^{-\beta \rho_j / \rho_j^v} + \mu^v \rho_j e^{-\beta} \\
\mu^v(\rho_j) = \sum_{i=1}^{n} \frac{\rho_i \rho_j}{\rho_i + \rho_j} e^{-\beta} \\
\alpha(x_j - x_k) = \begin{cases} 
\frac{\sum_{i=1}^{n} \rho_i x_i}{\rho_j} & \text{if } x_j \in S_r \\
0 & \text{otherwise}
\end{cases}
\]
| Double Heaviside projection [37] | \[
\rho^v = 1 - e^{-\beta \rho_j / \rho_j^v} + \mu^v \rho_j e^{-\beta} \\
\mu^v(\rho_j) = \sum_{i=1}^{n} \frac{\rho_i \rho_j}{\rho_i + \rho_j} e^{-\beta} \\
\mu^v(\rho_j) = \sum_{i=1}^{n} \frac{\rho_i \rho_j}{\rho_i + \rho_j} e^{-\beta} \\
\alpha(x_j - x_k) = \begin{cases} 
\frac{\sum_{i=1}^{n} \rho_i x_i}{\rho_j} & \text{if } x_j \in S_r \\
0 & \text{otherwise}
\end{cases}
\]
| Dilate [110] | \[
\rho^v = \log \left( \frac{\rho_j}{\rho_j^v} \right) / \beta 
\]
| Erode [110] | \[
\rho^v = 1 - \log \left( \frac{\rho_j}{\rho_j^v} \right) / \beta 
\]

*Please refer to the original works for more specific meanings of the symbols.
and minimum length scale control. However, SIMP method in fact has the potential to realize equivalent control effects, especially given the active research in this field. It would not be surprising to find more works for the simultaneous maximum and minimum length scale control in the near future. In addition, it is interesting to find out that the level set method has only been applied to constrain the component length scale but not the voids, even though it has the capability. Another point worth noticing is that the signed distance field employed by level set method facilitates not only the evaluation of length scales but also the generation of cutting paths [155], which makes it possible to embed some quantitative manufacturability measure into the topology optimization problem, e.g. cutting efficiency. At the end, we need to mention that not all level set methods employ the signed distance information, e.g. the parametric and phased-field methods generally do not.

3.2. Geometric feature based topology optimization

Design-for-manufacture (DFM) is a feature based conceptual design method, which improves the product competitiveness by concurrently considering its functioning and manufacturing requirements during the early design stage [58]. In this way, the required cutting methods, tools, suggested tolerances, surface finish specifications, and estimated manufacturing cost can be determined before passed into the CAM module [50]. In a recent work [69], the authors proposed an updated approach named optimization-for-manufacture (OFM). It attempts to numerically incorporate the manufacturing evaluations such as time and cost into the optimization problem formulation, and therefore, to achieve the concurrent functionality and manufacturability design through topology optimization. A conceptual framework of OFM is demonstrated in Fig. 5. To realize the OFM, feature technology should be involved because both the geometric and semantic information included in the machining feature definition [83] is mandatory to quantitatively evaluate the manufacturing feasibility. Therefore, this subsection investigates the geometric feature based topology optimization methods: their current status and future research directions.

Based on the SIMP method, the earlier works mainly focused on multi-component layout design [152,153] and a review can be found in [141]. The main approach was to optimize the components’ positions through parametric sensitivity analysis and the support structure through SIMP method. A finite circle method was developed to prevent the components from being overlapped. However, this approach has the limitation that re-meshing was repeatedly performed around the component areas which reduced the overall computational efficiency. To fix this issue, some improvements were made in [57,132,142] by adopting the level set based geometric feature representation and the X-FEM to trace the material/material interfaces, which altogether eliminated the repeated re-meshing. Wang et al. [125] studied the piezoelectric structure design by combining level set representation for moving piezoelectric actuators and SIMP for the host structure. Zhang et al. [144] performed the multi-component design in a similar way by combining the SIMP and level set representations, where the structural skeleton based non-overlap constraints were developed.

In some recent works of the SIMP method, Clausen et al. [23] realized the fixed-area void feature in the topology design by
employing several penalization measures as part of the objective function. Ha and Guest [44] and Guest [40] developed a Heaviside projection based component layout design method, which enabled the creation of small components and realized the non-overlap control. Recently, Norato et al. [91] inherited the idea from [42] by filling the design domain completely with components. A geometry projection method was developed to optimize sizes and spatial positions of these components, which finally derived the geometric feature based design.

Geometric feature based topology optimization under the level set framework can trace back to Chen et al. [19,20]. They fully parameterized the level set functions to implicitly model the geometric features, which later, were combined through R-functions to form complex CSG models. Parametric sensitivity analysis and design update are enabled by the differential properties of the implicit function [106]. Cheng et al. [22] and Mei et al. [87] employed a similar way to perform geometric feature modeling and parametric sensitivity analysis. More importantly, they developed an initial procedure to topologically generate geometric features inside the design domain, which made it possible to derive geometric feature based design from an arbitrary input. Gopalakrishn and Suresh [35] contributed the feature-specific topological derivative to insert both internal and boundary features. This work provided a good theoretical basis but the implementation under a topology optimization framework has not been verified yet. Zhou and Wang [149] did the geometric feature manipulation in a different way that, boundary velocity fields of the geometric features were regulated via least squares fitting to reserve the shape characteristics; by doing so, they accomplished the concurrent geometric feature control and freeform support structure design. This least squares fitting idea was inherited by Liu and Ma [69] to enable the geometric feature insertion during the early optimization loops, which has been used to realize the 2.5D machining feature based design (see Fig. 6). Recently, Guo et al. [42] and Zhang et al. [145] contributed a novel level set method. The design domain was initially distributed with geometric feature components, which had the freedom of scaling, movement, and rotation. By optimizing these degrees of freedom, a purely geometric feature based design could be derived.

In summary, the existing methods, developed under the SIMP or level set framework, have partially or fully realized geometric feature based design, where the geometric features can be either solid or void and have the freedom of movement, scaling and rotation. However, these methods also have some shared limitations. The majorities of these methods are only capable of manipulating the pre-specified geometric features but cannot generate new ones during the optimization process. In other words, quantity and types of the geometric features have to be pre-determined while it...
Fig. 6. 2.5D machining feature based level set topology optimization [69].
is somehow arbitrary and cumbersome to do so. For the only a few exceptions, the method reported by Cheng et al. [22] and Mei et al. [87] is still not well-developed for several reasons: (1) they developed a compound feature type to be inserted, which could evolve into any specific feature type but requires a slow converging process [87]; (2) the scale control is far from ideal for the infinitesimal nature of topological derivative, which causes the topology design composed of too many small boundary segments; and (3) this method has not been proven effective in 3D scheme. Gopalakrishn and Suresh [35] provided a potential theoretical basis enabling feature insertion during topology optimization process. However, the possibility and effectiveness of its application has not been verified yet. The method developed by Liu and Ma [69] was proven effective under the 2.5D machining background, but at this stage, multi-direction 2.5D machining cannot be addressed. The method presented in [40] so far can only generate circular components.

Therefore, an ideal geometric feature based topology optimization method is still in need and expected to have the following characteristics: (1) in-process generation of geometric feature primitives; (2) good scale-control of the generated feature primitives; (3) eliminated/reduced need of post-treatment; (4) effective application to 3D problems.

4. Injection molding/casting oriented topology optimization

Injection molding and casting are two similar manufacturing processes. They share some common procedures that: liquefy the material, inject/pour the liquid material into the cavity of the desired shape, perform cooling and solidify the material, open the mold and eject the part. These procedures are distinctive from the material removal based machining process, which makes the injection molding/casting oriented topology optimization methods strategically different.

4.1. Rib thickness control

For injection molding/casting, part cooling and solidification are carried out with external devices such as cooling channels and air blowers. These devices speed up the cooling process but intensify the cooling imbalance, which causes residual thermal stress and distortion that negatively impacts the part quality, especially for the areas of varying rib thickness. Therefore, small rib thickness gradient or even constant rib thickness is required to relax the cooling imbalance.

It is non-trivial to realize the small rib thickness gradient through topology optimization. Extra control efforts are mandatory to affect the shape and topology evolution. As reviewed in Section 3, there are several efforts of the SIMP method to realize the minimum/maximum component length scale control. Theoretically, the small rib thickness gradient could be realized by simultaneously constraining the maximum and minimum component length scales. However, there is few implementation of the concurrent length scale control under the SIMP framework and the only exception is found in [143]. Comparatively, it is more mature of applying the level set method because numerous implementations of the concurrent length scale control can be found in literature [3,21,41,68,76,128,131]. For instance, an example from [131] is presented in Fig. 7. Form the authors’ opinion, these levels set based implementations benefit from the signed distance information which greatly eases the length scale measure and control.

In summary, there is a common limitation that the concurrent length scale control easily traps the topology design at a local optimum which means the optimality strongly depends on the initial guess. A strategy to fix this issue could be alleviating the length scale control in early optimization loops to derive the optimal topology and then gradually intensifying the control effect to realize the targeted length scales [3].

4.2. Interior voids and undercuts

Another requirement is to avoid interior voids and undercuts (see Fig. 8), because these details can only be molded by using extra devices such as movable mold inserts and will complicate the ejection process. In literature, there have been effective modifications of the optimization algorithm in satisfying this requirement.

Based on SIMP method, Zhou et al. [147] and Schramm and Zhou [101] developed the casting constraints that elements were only allowed of monotonous density changes along the casting direction. Stromberg [113] applied the molding/casting constraints for unilateral contact problems. Lu et al. [75] applied the molding/casting constraints in a multi-direction manner. Gersborg and Andreasen [34] modified the SIMP method by using a single material density variable to decide the solid-void interface for each row of elements along the casting direction, which avoided the large number of constraints especially for a refined mesh. Guest and Zhu [39] extended the projection-based algorithm to satisfy the no undercut and void restrictions. Additionally, short reviews about the casting part design through topology optimization can be found in [45–47].

As for level set method, Xia et al. [130] satisfied the no undercut and void restrictions for casting parts by adjusting the design velocity only parallel to the pre-defined casting direction. In this way, once the initial design satisfies the casting restrictions, the topology design is guaranteed castable. In Xia’s work (2010), the velocity was aligned to the casting direction and the algorithm cannot recover the material portions once removed. An enhanced version was reported by Allaire et al. [2]. They added a minimum thickness constraint in the casting direction, and therefore avoided the overly removed material portions. Liu et al. [70] addressed the non-interior void constraint for multi-material injection molding. It is worth noticing that, parts manufactured through extrusion employ even stricter restrictions that the cross-sections in the extrusion direction should be identical, and topology optimization for extruded parts was also addressed under level set framework through the boundary velocity projection [66].

In summary, both the SIMP and level set methods can effectively satisfy the no interior void and undercut restrictions. In addition, it is worth noting that interior voids and undercuts sometimes can cause difficulties in machining as well, so the modifications discussed in this sub-section can be equally applied to machining-oriented topology optimization.

5. A promising future research direction: topology optimization for additive manufacturing

In recent years, there has been an explosive growth of additive manufacturing (AM) research [29]. Different from the subtractive machining, AM employs the layer-by-layer material deposition process. Due to this layered approach, the engineering parts can be designed with great complexity while the manufacturing cost would not accordingly increase and sometimes even decrease [10]. Therefore, the complexity constraint for the conventional manufacturing methods is removed and the design creativity is greatly liberated.

Hence, in this section, the topology optimization methods for AM are summarized. We define this topic as one of the future research directions because there are only a limited number of publications and more challenges than already well-developed solutions.

Because of the removed complexity constraint, the full energy of topology optimization has been released. The direct benefit is
that intermediate densities can now be effectively mapped to 3D printing units and several research groups [59,97,146] have proved its feasibility through testing the 3D printing samples. More importantly, advanced topology designs can be realized as real products. For instance, topology optimization of meso-scale porous structures (see Fig. 9) has achieved the extremely light-weight designs which are very promising to replace the conventional solid mechanical parts [96,115]; the achievements of the multi-material topology optimization are manufacturable and validations have been performed by some recent works [31,85]; even the two-scale topology optimization has been emerging into an active research field owing to the rapid development of AM. There are plenty of research works presenting advanced topology optimization methods, the results of which can only be produced through AM. However, these works mainly focus on the computational design aspect while address little or nothing related to the newly-introduced AM-specific rules and constraints; therefore, they will not be further specified in this survey.

As mentioned by many authors [10,88], topology optimization and AM should be closely bonded to address the newly-arising AM-specific rules and constraints. A few authors have conducted the literature survey [10,88] and in this paper, we would make the further summarization and give our comments and suggestions.

A typical one is the part’s buildability constraint [62]. It is violated by surfaces whose inclination angles from the platen are smaller than the threshold value, and consequently, these areas are not buildable through AM without support structures. Due to this constraint, a topology optimization problem is decomposed into two sub-problems [30]: the part design and the support structure design. However, so far, there is no clearly constructed algorithm for the support structure optimization and this problem is intrinsically very complicated. In addition, as shown in Fig. 10, the topology optimization process generates even more unbuildable areas where more support structures are required. So many support structures cause some disadvantages, including: waste of material, increased building time, and extra effort to remove them. Therefore, from the authors’ opinion, a support-free topology optimization method is preferable for the community. A recent work can be found in [64], which contributed a post-treatment method to linearize the structure boundaries and remove the support-free violations by adding materials. However, this post-treatment is somehow arbitrary and would severely sacrifice the result optimality. Brakett et al. [10] proposed one solution to iteratively linearize the structural boundary, measure their lengths and orientations, and iteratively penalize the unbuildable areas. However, this method is still only a conceptual idea. Gaynor [32] realized the maximum overhang control through an additional layer of design variable projection. However, the derived topology would be drastically changed compared to the regular compliance minimization result and, the sensitivity analysis is too costly which is even heavier than the involved finite element analysis. Therefore, a better support-free topology optimization method is still needed.

Another AM-specific constraint is the restricted minimum component size, because too small component size is non-manufacturable [10,30] and thin ribs have the risk of breakage when removing the support structure. Especially for compliant mechanism design, the point joint appears frequently [31] through the conventional topology optimization. Methods of realizing the minimum component size control have been well developed under both the SIMP and level set frameworks. They have been discussed in early sections, and therefore, will not be repeated here.

Interior voids are non-manufacturable through AM and should be restricted during the optimization process. For instance, Liu et al. [72] developed a virtual temperature method incorporated into the topology optimization algorithm, which successfully achieved the no interior void design. Other than that, the no undercut and interior void solutions developed for injection molding/casting parts also apply, but they would generate too conservative results as undercuts are allowed by AM. Therefore, there is still room for a better solution to satisfy the no interior void constraint.

Another opportunity created by AM is the topology optimization for hybrid manufacturing. AM produces sub-standard surface
finish which is often further polished by post-machining. In this situation, both the AM and machining constraints should be addressed when solving the topology optimization problem. On the other hand, in practice, not all boundary segments require the high-standard surface finish. Therefore, a topology optimization method for hybrid manufacturing would be useful. Briefly speaking, the boundary segments can be distinguished into two categories: requiring or not requiring post-machining. The former can be optimized through 2.5D machining feature based method [69], while the latter will employ the freeform evolution. Feasibility of this type of combination has been proven by [149].

In summary, it can be concluded that the AM-specific rules and constraints have created great opportunities for topology optimization development. Other than the ones discussed above, there are always further space for digging, e.g. the directional material properties [71,116], the topology design interpretation for AM [140], and the part consolidation [71] etc. On the other hand, these opportunities are also very challenging research topics and as pointed out by [10], a lack of solutions is the current situation of most AM-related problems.

6. Conclusion

This paper presents a survey of manufacturing oriented topology optimization methods. Several aspects have been covered including topology design parameterization, length scale control and feature based design of machining parts, rib thickness control and ejection of injection molding/casting parts. For each aspect, the state-of-art is summarized, limitations are pointed out, and potential solutions are proposed.

For future research, topology optimization for AM is emphasized. Owing to the rapid development of AM, quite a few new rules and constraints arise which bring new challenges and opportunities to topology optimization. Some important issues have been discussed and “a lack of solutions” can be summarized as the current status. It can be predicted a burst of publications in the near future.

At the end, a few promising but also challenging research topics are summarized for future development of manufacturing-oriented topology optimization. Especially, their tight connection to industrial practice is discussed.

(1) The concurrent minimum and maximum component length scale control should be paid more attention and be extended to 3D structures. This is important for injection molding parts which require the nearly constant rib/wall thickness in tight tolerance. So far, there are few implementations for the concurrent component length scale control [3,41,131,143] and big intervals are employed between the lower and upper bounds. It is worth a deep exploration about how they perform in case of 3D examples and reduced intervals. In addition, the optimized shape and topology with length scale control is drastically changed compared to the regular compliance minimization result and therefore, the practical applicability deserves further exploration.

(2) Quantitative evaluations of the manufacturability should be addressed by the topology optimization implementations. So far, manufacturing-oriented topology optimization implementations have produced manufacturable topology designs. However, quantitative evaluation of the manufacturability is rarely focused, e.g. the manufacturing time and cost for machining parts, and the mold cost for injection molding parts. Practically, these quantitative evaluations are even more concerned by the engineers.

(3) For AM, topology optimization for hybrid additive-subtractive manufacturing is a promising topic. Some details about this topic have been discussed in the last section. In fact, several commercialized CNC machines have been developed which embedded AM as part of the functions. However, a dedicated part design methodology is still vacant.

(4) Again for AM, concurrent optimization of the structural topology and the related deposition path would be an interesting topic. These two aspects are tightly bonded because the deposition path determines the anisotropic material properties adopted by topology optimization and inversely, the structural shape and topology affect the deposition path planning. Therefore, these two aspects should be concurrently optimized while to the best of the authors’ knowledge, there is no related implementation.

Acknowledgements

The authors would like to acknowledge the financial support: NSERC Discovery grants and China Scholarship Council (CSC) student scholarship. All the research works were carried out at University of Alberta.

References


Bracke D, Ashcroft H, Hague R. Topology optimization for additive manufacturing. 22nd annual international solid freeform fabrication symposium Austin, TX; 2011.


Dassault Systemes. SIMULIA Tosca Structure documentation 8.0; 2013.


Optimization for Shao Schramm Petersson


