Proceedings of the ASME 2013 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2013
August 4-August 7, 2013, Portland, USA

DETC2013-13048

SHAPESIFT: SUGGESTING SUSTAINABLE OPTIONS IN DESIGN REUSE FROM PART REPOSITORIES

Devarajan Ramanujan[†][*], **William Benjamin**[†]
[†]School of Mechanical Engineering
Purdue University
West Lafayette, Indiana 47907

William Z. Bernstein[†], **Niklas Elmqvist** [✠], **Karthik Ramani**[†] [✠]
[†]School of Mechanical Engineering
[✠]School of Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana 47907

ABSTRACT

ShapeSift is a framework for supporting sustainability-based decision making during selection of similar previous designs from part repositories. Our framework is designed for 3D part repositories that contain metadata pertaining to materials and manufacturing processes as well as the functionality of a given part. To demonstrate the usefulness of the proposed framework we develop an example multi-dimensional visualization that encodes part similarities as well as a calculated environmental sustainability indicator. This visualization is incorporated in a prototype interface that focuses on enhancing the intuitiveness of the exploration process by the use of sketch-based retrieval.

INTRODUCTION

Life Cycle Assessment (LCA) has become one of the most objective methods for quantifying environmental sustainability of a product or a process. Even so, conducting a detailed LCA for every part archived in a repository is impractical due to the entailed time and resource intensiveness. An alternative to the above is to simplify the assessment process by reducing the scope of the analysis and approximating materials and manufacturing by the closest available alternative from a standardized inventory [1]. Although such an approach entails several approximations and assumptions regarding the nature of the part, it can be argued that an approximate measure for environmental impact at the design stage can be used as a preliminary check against developing designs that have a significant negative effect on the environment. Another important consideration is the ability to automate the estimation of an environmental indicator such that it scales well with regards to larger repositories. One such method for automating the estimation of life cycle impacts for parts present in a design repository is detailed by Haapala et.al. [2]. The authors identify challenges such as 1) the availability of accurate mass and volume data, 2) levels of specificity in the description and material and manufacturing process and 3) estimates of part features relevant to the process, which prevent a more accurate computation of life cycle impact. We try to address some of these challenges in the current paper through a taxonomy-based description for part attributes. Using a taxonomy-based approach facilitates variable levels of data abstraction as well as quantifying pairwise similarities. Our framework also incorporates a geometric inference module to extract volumetric as well as shape data from 3D models of mechanical parts. Finally, we develop a visualization pipeline that presents these results in a manner that aids part retrieval, selection and repository exploration.

The primary contribution of this paper is the development of a novel support framework for selection of similar previous designs in a 3D part repository that is guided by sustainability principles. Our framework allows automated computation and visualization of 1) similarities in part attributes and 2) corresponding environmental indicators. In the literature that was reviewed, we

[*]Address all correspondence to this author. Email: dev@purdue.edu

could not find a system that merged these concepts in the context of mechanical designs. We make this integration possible by building on our insight that environmental impact is an inherent part attribute that can be derived from other part attributes such as geometry, material and manufacturing. The complex nature of the relationship between environmental impact and other part attributes is difficult to explicitly quantify. However, allowing the designer to develop an intuition about 'impact-attribute' relationships can significantly aid the sustainable design process. We address this challenge by encoding these relationships as 'visual variables'. We envisage our method to enable designers to hypothesize and evaluate their mental models pertaining to 'impact-attribute' relationships and lead to better insights regarding the factors correlating environmental sustainability and design decisions. Additionally, we quantify similarities on multiple product dimensions such as shape, function, material and manufacturing. By allowing the designer to explore existing part data through multiple lenses, our framework provides a richer context in the design exploration process. Inferring volumetric properties from 3D models in vendor neutral file formats and using standard taxonomies for material and manufacturing process data allows our framework to be easily scaled to larger part repositories. We also present a prototype interface that is designed for sustainability-aware exploration of 3D repositories by visualizing similarities in part attributes. This prototype interface uses sketch-based querying for enhancing the intuitiveness of the interaction.

BACKGROUND

An estimated 75% of design activity involves reuse of previous design knowledge in order to solve new design problems [3]. One significant challenge for estimating life cycle impact of previous designs archived in repositories is clearly defining the nature and the quality of archived data. Although efforts such as the National Institute of Standards and Technology (NIST) repository [4] and the University of Missouri-Rolla (UMR) design repository [5] have been successful at defining data standards for design information, they are not designed to contain an exhaustive lifecyle inventory (LCI) that can be used for environmental impact assessment. In order to make our framework relatively independent of the type of part repository, we use representations that can be derived from part data contained in existing repositories. Although this independence comes at the price of information loss in lifecycle data, developing a more general approach can help reach a larger user base. Additionally, users can customize our framework towards their specific requirements by altering the chosen data scheme.

Feature level information is often absent in existing repositories. Therefore we work with 3D part repositories that do not contain a mapping of manufacturing processes to specific part features. Previous research in sustainable design has looked at bridging gaps in lifecycle related information by techniques that leverage implicit knowledge embedded in existing parts. Approaches that are preferred among researchers in this area include using surrogate measures of environmental impact [6, 7], developing indices that relate environmental impact to part attributes [8, 9] and extrapolating impact on the basis similar existing products [2, 10]. Although these methods develop solutions from a computational perspective, most of them stop short of developing visual representations that effectively communicate this data. The process of design inherently deals with decision making that involves people. Therefore, using the right representation for data can significantly alter the quality of the analysis. The importance of placing data in the right context and allowing decision makers to make quantitative comparisons among them is emphasized by Tufte [11].

We posit that integrating meaningful visualization schemes with sustainability assessment can help designers observe covariation among product attributes and enable better decision making in the context of product reuse. Research has shown that visualization can act as a powerful enabler of environmental sustainability. Creative real-time visualizations that quantify energy consumption and carbon loads have been used to promote resource conservation [12]. Developing meaningful visualization of sustainability indicators presents a challenge due to its high dimensionality. An interface for visualizing the QUEST environmental sustainability model is presented in [13]. The authors provide insights into the successes and challenges in designing visualization schemes required for engaging communities in environmental policy making. Quay et al. [14] propose a hierarchical data organization where users can select regions of spatial, temporal or topical interests. An additional requirement for a visualization scheme applicable to 3D repositories is the ability to query and convey shape information. A 3D visualization of shape search results achieved by multi-dimensional scaling (MDS) based dimensionality reduction of shape distances is presented in [15]. Kim et al. [16] present a method to explore shapes in a 3D repository using a user specified region of interest. While these methods characterize 3D models purely based on shape, our goal is to explore 3D repositories based on combination of shape, environmental sustainability and relevant part attributes.

METHODOLOGY

The focus of this paper is to enable sustainability-based selection during exploration of similar parts archived in 3D part repositories. Therefore, the preliminary necessity for applying our framework is access to 3D product data with metadata regarding the process plan and functionality. All or most of the data necessary to support our framework is stored in different representations within various existing PLM systems and repositories. Since we plan to integrate our framework with such repositories

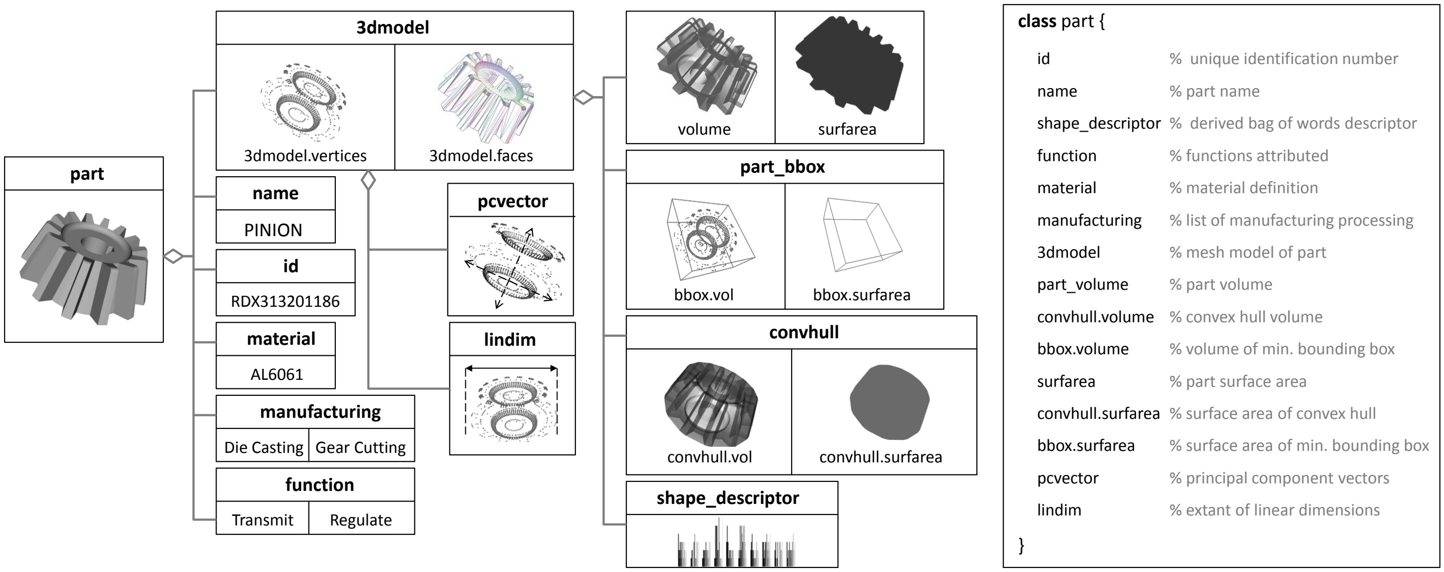

FIGURE 1. DIAGRAM OF DEPENDENCIES IN PART DEFINITION. ARROWS REPRESENT AN AGGREGATION RELATIONSHIP BETWEEN ATTRIBUTES. ATTRIBUTES CONTAINED IN THE CLASS ARE EITHER SPECIFIED AS INPUT DATA DURING INSTANTIATION OR DERIVED FROM INPUT DATA.

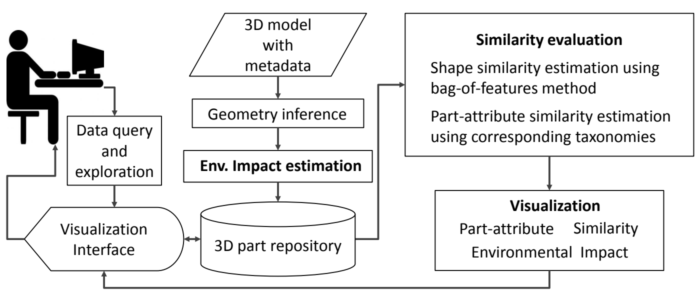

FIGURE 2. OVERVIEW OF THE PROPOSED FRAMEWORK FOR SUPPORTING SUSTAINABILITY BASED DECISION MAKING IN 3D PART REPOSITORIES WITH MATERIAL, MANUFACTURING AND FUNCTION DATA. CORE COMPONENTS OF THE PIPELINE ARE IN BOLD FONT.

and not necessarily develop an exhaustive stand-alone system, a pilot database consisting of only 50 engineering parts was constructed to demonstrate our method. 3D parts used for constructing the pilot database were obtained from the Engineering Shape Benchmark (ESB) [17]. The ESB contains a total of 479 models in Stereolithography (.STL) file format that are classified into 45 shape classes. Synthetic data regarding material, manufacturing and functionality was added to the part data. We tried to ensure that the selected parts have a reasonable degree of variability in their shape, material and manufacturing definitions. In the pilot database, each mechanical part is represented as an object with three major attributes. Figure 1 illustrates the definition scheme for the part class. The primary inputs to the database are 1) 3D model of part, represented by a mesh 2) material definition 3) ordered list of manufacturing processes and 4) functional description of the part. Metadata such as part name and identification number can be provided for better indexing and retrieval. Other part attributes including approximate environmental impact and the proposed measure of similarities are derived from these inputs.

A high level overview of our framework is illustrated in Fig. 2. The primary interaction mode for users of our framework is query-based exploration of part similarities. A range of visualizations can be designed to guide these processes. An interface with one such visualization scheme is discussed in this paper. We also discuss interaction modes that are currently implemented in our prototype interface. We start this discussion by detailing the methodology behind the three core modules in our framework, namely the 1) environmental impact assessment module 2) similarity evaluation module and the 3) visualization module and prototype interface.

Environmental Impact Assessment Module

Within this study, we limit the focus of our assessment from cradle to gate (resource extraction to end of manufacturing). Reducing the scope of the life cycle assessment can result in higher uncertainties in the calculated value of environmental impact.

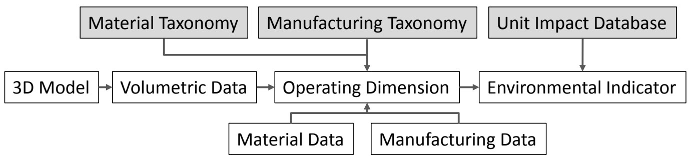

FIGURE 3. PIPELINE FOR ESTIMATING ENVIRONMENTAL IMPACT INDICATOR FROM INPUT DATA. GRAY SQUARES REPRESENT REFERENCE TAXONOMIES AND DATABASES USED FOR STANDARDIZING DATA DESCRIPTION.

However for traditional mechanical parts it can be argued that use phase impacts are comparatively low and therefore a cradle to gate analysis covers a significant portion of the overall impact. Furthermore, the primary objective of our framework is to safeguard designers against selection of environmentally unsustainable designs. Thus, our design goal is to establish an indicator that enables mechanical parts that have a high value of cradle to gate impact to be filtered out of the selection process.

Given a 3D model of a mechanical part with corresponding metadata, we start by extracting volumetric as well as shape data using the geometry inference module. This information along with process data is used for estimating the approximate environmental impact of the product. It should be noted that since a mesh-based representation of the solid model is used, feature level information is considered unavailable.

Manufacturing processes are specified in our framework as per the Allen and Todd taxonomy described in [18]. This taxonomy categorizes processes into fourteen major families. This classification taxonomy takes into account workpiece geometry, resulting tolerances, workable materials and cost. This taxonomy was preferred as the classifications described correspond closely with volumetric data corresponding to the part. One of the reasons for incorporating a taxonomy-based specification for manufacturing is the flexibility that it allows in the level of specificity of a process. For example, a repository might contain a part that is described as being 'cast' without further information on the exact nature of the casting process (e.g. die casting, investment casting, sand casting...). However, estimating the environmental impact data requires a more specific unit process. In such cases, an approximate measure for environmental impact can be established by averaging the unit impacts of the set of manufacturing process in the induced sub-tree. Similarly, it is possible that unit process information regarding a specific process is unavailable in the used LCI database. Here, we can approximately estimate impact by substituting it for the most similar manufacturing process that has data available in the LCI. This is achieved using a measure for similarity among manufacturing processes based on the structure of the taxonomy.

A corresponding taxonomy for material specification described by Ashby [19] is also incorporated in our framework. Within this scheme materials are grouped into five classes: ceramics and glasses, fibers and particulates, hybrids, metals and alloys and polymers. Each material class is classified into several material groups. A complete classification scheme is available within the CES Edupack software [20]. The sheer number of cataloged engineering materials presents a challenge with regards to implementation. Currently, we have a total of 39 material subgroups implemented and future work will focus on expanding the same.

The pipeline for estimating the environmental impact indicator is described in Fig. 3. The first step involves extraction of volumetric properties from a 3D model of the mechanical part stored in the database. Properties such as volume, surface area, convex hull volume, minimum bounding-box volume are calculated using standard algorithms. The next step in our pipeline involves the estimation of an operating dimension (O_{dim}) for each manufacturing process associated with the part. O_{dim} is defined as the physical variable pertaining to part geometry (i.e volume, surface area) that is processed by a manufacturing operation. For example, the O_{dim} for a material removal process such as machining is the volume of removed material. The operating dimension is used as a multiplicative scaling factor for unit process impacts in order to compute the net impact of the process. In an ideal setting, the operating volume for each process is specified as input data or encoded as changes in shape of the 3D model. Although a well defined PLM system might have such data archived, most repositories today do not provide and means for obtaining this data. Therefore, within this study we try to estimate the associated O_{dim} for a specific manufacturing process based on the following approximations:

◦ If the volume of the starting stock/blank is not specified, it is taken to be equal to the smaller value of 1) the convex hull volume of 3D part and 2) the volume of the minimum bounding box of the 3D part.

◦ If there are more than one material removal operation in the list of manufacturing processes, the total removed volume is divided equally among these processes.

◦ The Allen and Todd taxonomy is used to categorize manufacturing processes into one of the following four types:

- Mass conserving volumetric. e.g., forging, annealing
- Mass reducing volumetric. e.g., turning, drilling
- Surficial. e.g., anodizing, electrocoating, dust coating
- Joining. e.g., welding, adhesive bonding

Thus, any process that appears before the first 'mass reducing' process has occurred always operates on the convex hull volume or surface area. Similarly, any processes that occur after a 'mass reducing' process operate on the reduced volume. Although units such as volume and surface area are easily computable from a 3D model, extracting feature level information for calculating the operating dimension for 'joining' processes present significant challenges. Therefore, information about the operating dimension (i.e length of weld, surface area of bonded surfaces) is required to be specified by the user as input to the framework.

Once the O_{dim} for each manufacturing process is estimated, the environmental impact is computed as a linear sum of the impact of material extraction and manufacturing processing, as shown in Eq.(1).

$$EI = e * b_v + \sum_{i=1}^{n} p_i * (O_{dim})_i \quad (1)$$

Where,
$EI =$ Net environmental impact
$e =$ Environmental impact associated with the unit process for material extraction
$b_v =$ Blank/Initial volume of material used for manufacturing the part
$p_i =$ Environmental impact associated with the i^{th} unit manufacturing process. Note that this quantity is also dependent on the type of material that is manufactured
$O_{dim} =$ Operating dimension of the i^{th} manufacturing processes
$n =$ Total number of unit manufacturing processes associated with the part

In this study, we have chosen Cumulative Energy Demand (CED) as the representative environmental impact indicator for each unit process. Cumulative energy demand for a product is defined as the total quantity of primary energy needed to produce, use, transport and dispose of that particular product. A lookup table is hard coded into our environment that contains CED values of unit processes for material extraction as well as a given material-manufacturing process combination. The data for these entries have been referenced from the methods library available through SimaPro 7.1 [21]. In future implementations, we plan to include indicators for global warming such as kg eq. CO_2 and Eco Indicator points.

An additional point of concern while estimating environmental indicators is the change in variables related to the process plan. Methodologies such as environmentally conscious process planning (ECPP) have the potential to optimize process selection as well as control with respect to environmental sustainability. In order to account for these changes our framework is designed to link to efforts such as the unit process life cycle inventory (UPLCI) database [22]. By adopting the same manufacturing taxonomy as per the UPLCI database, any changes made in the impact estimation for unit manufacturing processes can be readily updated within our framework. Future efforts in this direction will look at achieving a greater level of interoperability between the systems.

The function taxonomy that we use is adapted from the categorization of functions by Hirtz et al. [23]. The authors develop a reconciled functional bases where functions grouped into eight primary classes. The eight primary classes are divided into many more sub-classes. The authors also provide a list of correspondents that allow users to correlate their functional basis with related efforts.

Similarity Evaluation

A natural way of quantifying similarity between elements of a set is by establishing a measure of similarity or a distance between them. The similarity between two objects is a function of the commonality and the differences they share [24]. We capture these properties using a distance function $d : \varepsilon X \varepsilon \rightarrow R$ that operates on elements of a taxonomy ε and returns a real valued distance measure. Although we do not enforce that condition that the distance function d meets the required conditions to be defined as a metric, we develop a function that possesses the following properties:

1. Non negativity : $d(e_1,e_2) \geq 0; \{e_1,e_2\} \in \varepsilon$
2. Symmetry : $d(e_1,e_2) = d(e_2,e_1)$
3. Identity : $d(e_1,e_2) = 0 \Leftrightarrow e_1 = e_2$

Since the developed function does not conform to either the ultrametric or the triangle inequality, by definition our distance function can be termed as a semimetric. We begin the discussion on similarity computation by defining the terms involved.

All mechanical parts are considered to be elements of a set ρ, each having a material $m \in M$, a function $f \in F$ and a shape $s \in S$. Here, M, F, R are the respective taxonomies adopted to represent these attributes. A manufacturing process $r \in R$ is treated as an operator $r : \rho X \rho \rightarrow P$ such that it operates on a certain part and returns another part with either same or different material and shape properties. Thus the entire sequence of manufacturing processing can be viewed as a composition of operators that transform an initial blank $P_0\{m_0,s_0,f_0\}$ to the final part $P_n = \{m_n,s_n,f_n\}$.

The material, manufacturing, function and shape definition represent significant decisions towards framing design intent. Therefore we interpret the similarity among parts as a composition of similarities in these four attributes. Thus, we define a set of distance functions, namely, d_m, d_f, d_r, and d_s associated with these attributes respectively. Since material, manufacturing and function definitions are represented using corresponding taxonomies, we develop a generalized similarity measure that can be adapted to taxonomies. The distance function for shape is defined using the bag-of-features (BoF) method.

Material, manufacturing and function similarities:

Classification trees and taxonomies increase in specificity as we proceed lower down the hierarchy. Therefore elements in the same subtree at a lower level are more similar than elements higher then them. For example in a manufacturing classification scheme, two types of single point cutting processes are more similar to each other than any two mass reducing processes.

Exploiting this property for similarity computation requires making use of the structure of the hierarchy tree. The distance measure developed in the current study builds on concepts that are discussed in [26] and applies them towards the used material and manufacturing taxonomies. Given any two elements in a taxonomy, we calculate a distance measure using the following process:

◦ *Tree Depth Equalization:* When comparing the similarity between any two elements within the same taxonomy tree, we first ensure that only elements at the same depth from the root are evaluated. The algorithm for the same is illustrated below:

if $depth(a_1) > depth(a_2)$
  then $a_1^* = ancestor(a_1)$ at $depth(a_2)$ && $a_2^* = a_2$
else if $depth(a_2) > depth(a_1)$
  then $a_2^* = ancestor(a_2)$ at $depth(a_1)$ && $a_1^* = a_1$
else $a_1^* = a_1$ && $a_2^* = a_2$

◦ *Distance Estimation:* The next step is to calculate the numerical value of similarity between the entities substituted in the first step. The corresponding distance function is defined in Eq.(2):

$$D(a_1, a_2) = \frac{pathlength(a_1^*, a_2^*)}{depthlca(a_1^*, a_2^*) + pathlength(a_1^*, a_2^*)} \tag{2}$$

Here,
$depthlca(a_1^*, a_2^*)$ = Depth of the lowest common ancestor of node a_1^* and node a_2^* from the root. Note that the depth of the root node is equal to zero.
$pathlength(a_1^*, a_2^*)$ = Length of traversed path (number of hops) to reach node a_2^* from node a_1^*. Although a taxonomy tree is directed from top to bottom (root to leaves), for the computation of *pathlength* we allow edge traversals in both directions.

As both *pathlength* and *depthlca* and lie in the interval $[0, \infty)$, the distance measure D is restricted to the interval $[0, 1]$. However, when $pathlength = depthlca = 0$ the similarity measure is indefinite. These cases occur only when comparisons are made among elements of taxonomy and its root. As these comparisons do not hold any meaning, we exclude them from the set of allowable comparisons. It can be easily verified that our distance function satisfies the non-negativity, symmetry and identity conditions mentioned earlier. The distance between two elements in a taxonomy $D(a_1, a_2) = 1 \Leftrightarrow depthlca = 0$. In other words two elements in the taxonomy are considered to be entirely dissimilar if their least common ancestor is the root node of the taxonomy.

We proceed to define our method for composing a scalar measure of distance for specified material, manufacturing and function definitions. An important point to note is that a definition of an attribute can consist of either a single element, or in other cases a set of elements from the taxonomy. Additionally, the ordering of the associated elements may hold significance in cases such as the definition of a manufacturing sequence. Therefore, we develop a measure of each of the attributes that encodes dissimilarity as the maximum deviation of one set of attributes from the other.

In our repository each part is associated with a single material type. Therefore for any two materials $m_1, m_2 \in M$ the distance function d_m is directly given by the function operating on the material taxonomy as shown in Eq.(3). A manufacturing description $r = \langle e_{r1}, e_{r2} ..., e_{rn} \rangle, r \in R$ is considered as an ordered n-tuple of manufacturing processes. Given two manufacturing descriptions r_1 and r_2, we define a set $r_1 \circ r_2$ whose elements are 2-tuples formed by the element-wise product of r_1 and r_2. The intuition behind performing an element-wise operation is that, given two manufacturing descriptions it only makes sense to compare primary production processes with other primary processes, secondary processes with other secondary processes and so on. For example a comparison between a surface treatment process such as 'nitriding' with a primary process such as 'die casting' hold little significance. In cases where the cardinalities of r_1 and r_2 are different, we restrict the similarity computation to the first n elements, where n is the lower of the two cardinalities. The distance function d_r is defined as the maximum possible value of dissimilarity among the sets of descriptions as given in Eq.(4). A function description $f = \{e_{f1}, e_{f2} ... e_{fn}\}$ is considered as a set of functions wherein the ordering of the elements are immaterial. Like d_m the dissimilarity between two sets of function descriptions are governed by the maximum possible value of dissimilarity among the descriptions. Given two sets of function descriptions f_1 and f_2 the distance function d_f as detailed in Eq.(5). Here, $f_1 X f_2$ represents the Cartesian product of the sets f_1 and f_2. Note that although a Cartesian product is not symmetric, the distance function D that operates on that product is symmetric. Therefore the distance function d_f is also symmetric

$$d_m(m_1, m_2) = D(m_1, m_2) \tag{3}$$
$$d_r(r_1, r_2) = max(D(r_1 \circ r_2)) \tag{4}$$
$$d_f(f_1, f_2) = max(D(f_1 X f_2)) \tag{5}$$

Estimation of shape similarity: User sketches cannot be directly correlated to 3D models of parts available in repositories. The reason is that 2D sketches and 3D models should be compared using a compatible representation. Therefore, we convert 3D models into 2D projections of sketch-like renderings using 'suggestive contours' [27]. Even so, directly comparing user sketches to these 2D projections will result in erroneous matching due to noise introduced in user sketches by affine deformations and malformed strokes. Addressing these problems requires a shape similarity metric that is robust to both noise and distortion. Therefore, we use the bag-of-features method

(BoF) discussed in [25] to develop a metric for shape similarity. Previous literature [28, 29] has shown that the BoF method has commendable performance with regards to 2D shape classification and retrieval. The core idea of the BoF method is to represent images as a histogram of occurences of 'visual words'. Our method for computing shape similarity using the BoF method is described below:

◦ *Feature Detection:* In this step, we compute locations of interesting features, given by computing the 'feature points' on the image using the Harris Detector [30]. Finding such discriminative locations helps in identifying differences between shapes. For better discrimination, we look for repeatability of these feature points. Here, repeatability is defined as correspondence in the location of the features points after applying a transformation which maps the one shape to the other.

◦ *Feature Description:* In this step, we compute patch descriptors for each detected feature using the Scale Invariant Feature Transform (SIFT) [31]. SIFT embeds these features in a high dimensional space by assign a 128 dimensional descriptor to the features.

◦ *Quantizing Features using Visual Vocabulary:* Although the feature descriptors computed using SIFT are informative, the high dimensionality of the resulting data and the large number of detected features adds to the complexity of computation. To reduce some of the involved complexity we compute a 'visual vocabulary' by clustering features in the database. The clustering matches each feature to the nearest cluster center or 'visual word'.

◦ *Image Descriptor Generation:* In this step, we transform the image data into a histogram of count of occurrences of cluster center matches for computing the similarity metric. Given any two histograms x and y which represent two images S_x and S_y respectively, a p-norm distance can be computed by Eq.(6).

$$d_s(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right) \tag{6}$$

In this implementation we use a simple Euclidean norm by setting $p = 1$. Additionally, in the interest of supporting fast retrieval we use the fast approximate nearest neighbor method described in [32] to index queries.

Thus, the overall distance between two parts is given by $\{d_m, d_r, d_f, d_s\}$ which is a set comprising of pairwise distances among corresponding part attributes. Although it is possible to compose a scalar pair-wise distance measure from this set, there is a possibility that drastically reducing the dimensionality of the data might result in excessive loss of similarity information. Interpreting whether two parts are more similar due to similarities in material, function or any such attribute is largely decided by the context of the application and therefore by the user. So, we focus on creating meaningful multi-dimensional information visualization schemes that aids users in exploring a repository. This scheme overlays computed environmental indicators on the similarity visualization for enabling sustainability-aware design exploration.

Visualization and Interface Prototype

Although there are numerous schemes for visualizing sustainability related data, a handful of them try to merge these visualizations with the design process. For creating a seamless interface between the two, we develop a list of the following design goals that are sensitive to needs of the designer.

◦ *Ability to explore product repositories from a design similarity and sustainability perspective:* The process of exploration should allow the user to build mental models of the relationship between shape, material/manufacturing data and environmental sustainability. This approach assumes that a user has previous knowledge regarding the contents of the repository and supports him/her to arrive at a solution by exploring similar alternatives.

◦ *Intuitive Interaction:* One of our focus is to simplify the design exploration process by providing intuitive means of navigating through and searching for alternate design solutions from a given database.

◦ *Decision Support for Design Process:* We posit that human spatial and visual reasoning skills can be leveraged for effective decision making in the design process. Therefore, we adopt cognitively prominent visual elements such as variations in shape,size, color and spatial location in our representation.

For developing a solution that meets the above requirements, we develop a visualization scheme that contains the following elements:

Sketch based input: Adopting sketching as the primary method for query within our framework gives us the advantage of utilizing one of the dominant modes of artifact creation among designers. Sketching is shown to provide a visible graphic memory that facilitates creativity by providing an easily accessible repository of generated ideas and by stimulating building on earlier ideas [34]. Using sketch-based input gives users the ability to modify their input with relative ease.

Squarified layout visualization: Squarified layouts are useful for visually providing a summary of the search results. They can also provide visual cues that allow users to aggregate and discriminate search results. Fig.4-2, shows an example result that is automatically generated from a sketch-based retrieval process. In this visualization, environmental impact calculated through Eq.(1) is divided by part volume to develop an indicator for representing the sustainability of a unit shape. Squares larger in area contain parts with a lower value of the environmental indicator. By representing the indicator using a prominent visual

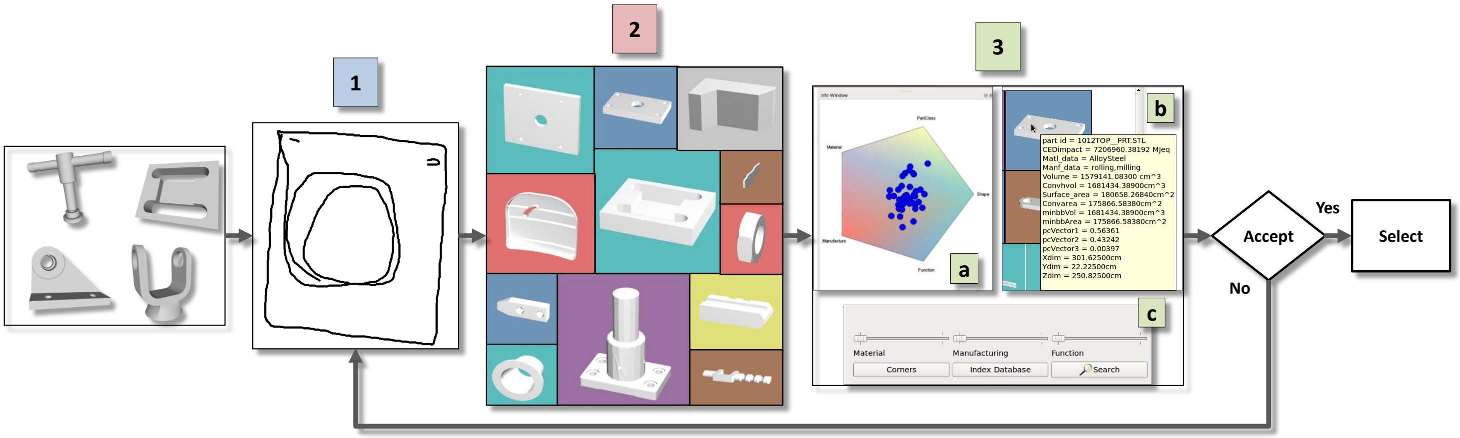

FIGURE 4. VISUALIZATION PIPELINE FOR EXPLORING 3D REPOSITORIES. USERS BEGIN BY SKETCH BASED QUERYING AS SHOWN IN 4-1. OUR SYSTEM ORGANIZES SEARCH RESULTS USING A SQUARIFIED LAYOUT 4-2 THAT IS CONSTRUCTED USING BRUL'S SQUARIFIED TREEMAP ALGORITHM [33]. HERE, EACH COLOR CORRESPONDS TO A PARTICULAR MATERIAL CLASS. THE AREA OF A CELL IS SCALED IN NEGATIVE PROPORTION TO THE CALCULATED ENVIRONMENTAL IMPACT INDICATOR. PARTS ARE ORDERED BY SHAPE SIMILARITY RELATIVE TO THE QUERY. USERS CAN EXPLORE THESE RESULTS FURTHER USING A SIMILARITY POLYGON 4-3a, SLIDERS THAT SET VALUES FOR DISTANCE THRESHOLD 4-3b AND INTERACTIVE TOOLTIPS 4-3c. THE COMBINATION OF THESE STEPS FORMS A UNIT ITERATION THAT CAN BE REPEATED AS DESIRED BY THE USER.

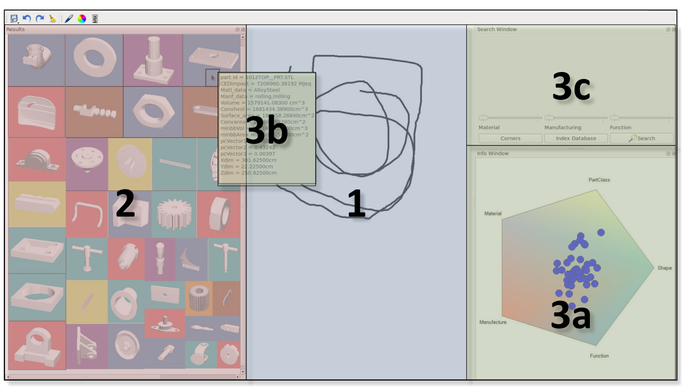

FIGURE 5. SCREENSHOT OF THE PROTOTYPE INTERFACE THAT INTEGRATES ELEMENTS DISCUSSED IN THE VISUALIZATION PIPELINE. THE NUMBERS ON THE INTERFACE REPRESENT IMPLEMENTATIONS OF CORRESPONDING ELEMENTS IN FIG.4. A DEMONSTRATION VIDEO DETAILING THE CURRENT CAPABILITIES OF THE INTERFACE IS AVAILABLE AT http://youtu.be/jICSAX3Lgc8

variable, we hope to nudge designers away from selecting unsustainable options. The color of each cell corresponds to the taxonomic class that the part's material belongs to. For example, in the current visualization a red background indicates that the part material is a kind of 'Alloy Steel'. A legend of colors is available to the user, and we ensure that we always maintain color continuity for the same material class in the visualization. The coloring information can be changed to represent classification along other attributes such as manufacturing or function class. Users can also filter results either by setting individual or multiple thresholds for the set of computed similarities (d_m, d_r, d_f). Parts that are dissimilar to a query part in terms of these attributes are represented with a grayed out background as seen in the top right corner of Fig.4-2. The current visualization is organized according to the shape similarity (d_s) relative to the best sketch query. Similarity decreases as we move from top-left to bottom-right in a horizontal raster. More specific information about part attributes such as volume, surface area, material are displayed using interactive elements such as tooltips.

Prototype interface: A screen capture of the developed prototype interface is shown in Fig.5. The current version of the interface contains the following panels:

○ *Results viewer:* This panel displays results of the query using the squarified layout as shown in Fig.5-2. Apart from implementing navigation aids such as pan and zoom, we have added interactive elements such as tooltips Fig.5-3b, and clickable tiles that allow users to explore the generated results in more detail.

○ *Sketcher Panel:* The sketcher panel shown in Fig.5-1 implements sketch-based query. Our interface also offers additional functionalities such as input customization by changing color and thickness of pens and ability save and load sketches.

○ *Similarity Viewer:* This panel contains sliders shown in Fig.5-3c that can be used for setting a distance threshold on a $0-1$ scale for the computed distances $\{d_m, d_r, d_f\}$. Results that do not fall within this threshold are interactively grayed out in the squarified layout. This panel also contains a 'similarity polygon' shown in Fig.5-3a that allows quantitative comparison of attribute similarities. Here, each vertex of the polygon corresponds to a particular part attribute. Scalar distance data are used to compute barycentric coordinates for a point inside the polygon. Each plotted point corresponds to a specific search result in the squarified layout. Using our 'similarity polygon', designers can assess the presence of a dominant distance metric in the set of attribute dimensions.

CONCLUSIONS AND FUTURE WORK

This paper has presented shapeSift, a novel framework for sustainablity-aware selection from 3D repositories. Part similarities are quantified on multiple dimensions such as material, manufacturing and function-based on the structure of their corresponding taxonomies. The framework describes methods for automating the computation of environmental impact indicators and similarities in part attributes. This data is visualized using a squarified layout which provides an overview of similar parts and their attributes. Finally, this paper discusses a prototype interface that integrates the visualization with sketch-based querying for supporting intuitive exploration of 3d part repositories.

Future work will be focused on extending the capabilities of the current interface by developing alternate visualization and interaction schemes. We plan on conducting full fledged user studies to evaluate our interface from a human computer interaction perspective. An important direction for future work is allowing users the additional flexibility of assessing the environmental impact of novel design concepts. Supporting environmental and similarity assessment for novel 3D parts, requires estimation of their geometric properties. Recently developed natural user interfaces for rapid virtual prototyping such as [35, 36] are particularly applicable in the this context. These interfaces allow designers to iterate over several designs and can be possibly used to guide sustainability-based decision making. Another important consideration that we wish to address is more accurate estimation of sustainability indicators. We will research methods for quantifying and visually representing uncertainties present in sustainability assessment.

ACKNOWLEDGEMENTS

The authors would like to convey their gratitude to Fu Zhao and Vinayak for their constructive feedback. This research work is partially supported by NSF under grants EEC 0935074 and CMMI 110619. The contents of this manuscript do not necessarily reflect the views or opinions of the funding agency.

REFERENCES

[1] Hunt, R., Boguski, T., Weitz, K., and Sharma, A., 1998. "Case studies examining lca streamlining techniques". *The International Journal of Life Cycle Assessment,* ***3***, pp. 36–42.

[2] Haapala, K., Poppa, K., Stone, R., and Tumer, I., 2011. "Automating environmental impact assessment during the conceptual phase of product design". In AAAI 2011 Spring Symposium: Artificial Intelligence and Sustainable Design, pp. 21–23.

[3] Ullman, D., 2002. *The Mechanical Design Process*, 3 ed. McGraw-Hill Science/Engineering/Math, July.

[4] Szykman, S., 2002. "Architecture and implementation of a design repository system". In Proceedings of DETC2002, DETC2002/CIE-34463, Montreal, Canada.

[5] Bohm, M., Stone, R., Simpson, T., and Steva, E., 2008. "Introduction of a data schema to support a design repository". *Computer-Aided Design,* ***40***(7), pp. 801–811.

[6] Sousa, I., Wallace, D., and Eisenhard, J. L., 2000. "Approximate life-cycle assessment of product concepts using learning systems". *Journal of Industrial Ecology,* ***4***(4), pp. 61–81.

[7] Park, J.-H., and Seo, K.-K., 2006. "A knowledge-based approximate life cycle assessment system for evaluating environmental impacts of product design alternatives in a collaborative design environment". *Advanced Engineering Informatics,* ***20***(2), pp. 147 – 154.

[8] Dewulf, W., and Duflou, J., 2006. "A web based application for the eco-pas tool". In CIRP - 13th International Conference on Life Cycle Engineering location:Leuven, Belgium, Vol. 1, pp. 143–147.

[9] Huang, H., Liu, Z., Zhang, L., and Sutherland, J., 2009. "Materials selection for environmentally conscious design via a proposed life cycle environmental performance index". *The International Journal of Advanced Manufacturing Technology,* ***44***, pp. 1073–1082.

[10] Devanathan, S., Ramanujan, D., Bernstein, W., Zhao, F., and Ramani, K., 2010. "Integration of sustainability into early design through the function impact matrix". *Journal of Mechanical Design,* ***132***(8), p. 81004.

[11] Tufte, E., and Weise Moeller, E., 1997. *Visual explanations: images and quantities, evidence and narrative*. Graphics Press Cheshire, CT.

[12] Holmes, T. G., 2007. "Eco-visualization: combining art and technology to reduce energy consumption". In Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition, C&C '07, ACM, pp. 153–162.

[13] Munzner, T., Barsky, A., and Williams, M., 2009. "Reflections on questvis: A visualization system for an environmental sustainability model". *Scientific Visualization: Interactions, Features, Metaphors,* ***2***, pp. 240–259.

[14] Quay, R., and Hutanuwatr, K., 2009. "Visualization of sus-

tainability indicators: A conceptual framework". In *Visualizing Sustainable Planning*, H. Hagen, S. Guhathakurta, and G. Steinebach, eds., X.media.publishing. Springer Berlin Heidelberg, pp. 203–213.

[15] Pu, J., Kalyanaraman, Y., Jayanti, S., Ramani, K., and Pizlo, Z., 2007. "Navigation and discovery in 3d cad repositories". *Computer Graphics and Applications, IEEE,* ***27***(4), july-aug., pp. 38 –47.

[16] Kim, V. G., Li, W., Mitra, N. J., DiVerdi, S., and Funkhouser, T., 2012. "Exploring collections of 3d models using fuzzy correspondences". *ACM Trans. Graph.,* ***31***(4), July, pp. 54:1–54:11.

[17] Jayanti, S., Kalyanaraman, Y., Iyer, N., and Ramani, K., 2006. "Developing an engineering shape benchmark for cad models". *Computer-Aided Design,* ***38***(9), pp. 939 – 953.

[18] Todd, R., Allen, D., and Alting, L., 1994. *Manufacturing processes reference guide*. Industrial Press Inc.

[19] Ashby, M., and Cebon, D., 1993. "Materials selection in mechanical design". *Le Journal de Physique IV,* ***3***(C7).

[20] EduPack, C., 2012. *version 11.9.9*. Granta Design, Cambridge, United Kingdom.

[21] Goedkoop, M., Oele, M., Schryver, A., and Vieira, M., 2008. *SimaPro Database Manual-Methods library*. Pré Consultants, Netherlands.

[22] Zhao, F., Murray, V., Ramani, K., and Sutherland, J., 2012. "Toward the development of process plans with reduced environmental impacts". *Frontiers of Mechanical Engineering*, pp. 1–16.

[23] Hirtz, J., Stone, R., McAdams, D., Szykman, S., and Wood, K., 2001. "Evolving a functional basis for engineering design". In Proceedings of the ASME IDETC2001, Pittsburgh, PA.

[24] Lin, D., 1998. "An information-theoretic definition of similarity". In Proceedings of the 15th international conference on Machine Learning, Vol. 1, San Francisco, pp. 296–304.

[25] Squire, D., Müller, W., Müller, H., and Pun, T., 2000. "Content-based query of image databases: inspirations from text retrieval". *Pattern Recognition Letters,* ***21***(13), pp. 1193–1198.

[26] Ganesan, P., Garcia-Molina, H., and Widom, J., 2003. "Exploiting hierarchical domain structure to compute similarity". *ACM Transactions on Information Systems (TOIS),* ***21***(1), pp. 64–93.

[27] DeCarlo, D., Finkelstein, A., Rusinkiewicz, S., and Santella, A., 2003. "Suggestive contours for conveying shape". In ACM Transactions on Graphics (TOG), Vol. 22, ACM, pp. 848–855.

[28] Csurka, G., Dance, C., Fan, L., Willamowski, J., and Bray, C., 2004. "Visual categorization with bags of keypoints". In Workshop on statistical learning in computer vision, ECCV, Vol. 1, p. 22.

[29] Sivic, J., and Zisserman, A., 2006. "Video google: Efficient visual search of videos". *Toward Category-Level Object Recognition*, pp. 127–144.

[30] Harris, C., and Stephens, M., 1988. "A combined corner and edge detector". In Alvey vision conference, Vol. 15, Manchester, UK, p. 50.

[31] Lowe, D., 2004. "Distinctive image features from scale-invariant keypoints". *International journal of computer vision,* ***60***(2), pp. 91–110.

[32] Muja, M., and Lowe, D., 2009. "Fast approximate nearest neighbors with automatic algorithm configuration". In International Conference on Computer Vision Theory and Applications (VISSAPP?09), pp. 331–340.

[33] Bruls, M., Huizing, K., and Van Wijk, J., 2000. "Squarified treemaps". In Data Visualization, Citeseer, pp. 33–42.

[34] McKim, R. H., 1972. *Experiences in visual thinking*. Brooks/Cole Pub. Co, Monterey, CA.

[35] De Araùjo, B., Casiez, G., and Jorge, J., 2012. "Mockup builder: direct 3d modeling on and above the surface in a continuous interaction space". In Proceedings of the 2012 Graphics Interace Conference, Canadian Information Processing Society, pp. 173–180.

[36] Vinayak, Murugappan, S., Liu, H., and Ramani, K., 2013. "Shape-it-up: Hand gesture based creative expression of 3d shapes using intelligent generalized cylinders". *Computer-Aided Design,* ***45***(2), pp. 277 – 287.

