Exploring Entities in Text with Descriptive Non-photorealistic Rendering

Meng-Wei Chang and Christopher Collins

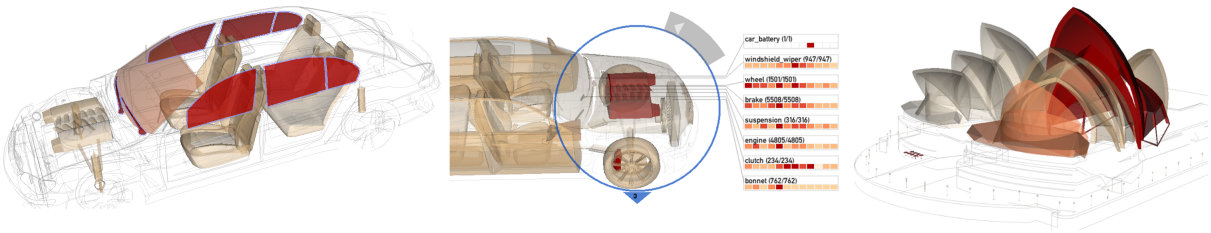

Fig. 1. Left: Multiple entity selection showing parts that failed along with wiper and window components. Middle: Exploring issues in the front of the vehicle. Right: A potential usage scenario for building maintenance.

Abstract— We present a novel approach to text visualization called *descriptive non-photorealistic rendering* which exploits the inherent spatial and abstract dimensions in text documents to integrate 3D non-photorealistic rendering with information visualization. The visualization encodes text data onto 3D models, emphasizing the relative significance of words in the text and the physical, real-world relationships between those words. Analytic exploration is supported through a collection of interactive widgets and direct multitouch interaction with the 3D models. We applied our method to analyze a collection of vehicle complaint reports from National Highway Traffic Safety Administration (NHTSA), and through a qualitative evaluation study, we demonstrate how our system can support tasks such as comparing the reliability of different makes and models, finding interesting facts, and revealing possible causal relations between car parts.

Index Terms—Integrating Spatial and Non-Spatial Data, Text Visualization, Non-photorealistic Rendering.

1 Introduction

There is an ever-increasing number of document collections concerning physical entities, objects we can relate to in our daily lives. Business intelligence applications rely on automatic summarization of such collections for tasks such as tracking customer comments about a product or service. The most common approach to visual summarization, the word cloud (e.g., [25]), renders words in an abstract array, out of context and without considering the meanings of the words. However, many nouns have easily discovered semantic relationships, such as meronymy or the *part-of* relation. Using a database of meronyms, it is possible to extract and relate keywords from documents based on their part-whole relationship. In this work, we focus on meronyms which have an inherent spatial relationship; that is, they are parts of a physical whole. For example, "handle" is a part of "toaster," located on the front, and "exhaust" is part of "car," located on the underside.

The major contribution of this paper is a new approach to text analytics which we call *descriptive non-photorealistic rendering* (d-NPR). Inspired by calls for closer integration of Scientific Visualization and Information Visualization [13, 14], we investigate a hybrid approach for text analytics by presenting both abstract and spatial semantics in a single view, thus creating an engaging visualization that has roots in the real world. Our approach consists of the analysis of descriptive texts, using a lexical database, to extract mentions of physical entities and compute their occurrence and co-occurrence scores. We encode the results onto segmented 3D models that correspond to the mentioned objects. Our rendering process creates information-rich visuals such that significant parts of the model appear to "pop-out" to the viewers, while retaining their easily recognizable shapes and forms. Interactive widgets are linked to the visualization and help analysts navigate the 3D visualization space, as well as discover trends and other interesting patterns in the data.

To demonstrate our method within a realistic analytic scenario, we applied our techniques to analyze a text corpus of 600,000 vehicle complaints from the US National Highway Traffic Safety Administration (NHTSA). Our intention is to help consumers make better, more informed choices when they purchase a vehicle. In order to find out how people respond to the visualization approach, we conducted an evaluation study with 12 participants. In particular we assess their accuracy and how they perform open-ended analytical tasks.

- *Meng-Wei Chang is with University of Ontario Institute of Technology, Canada. E-mail: meng-wei.chang@uoit.ca*
- *Christopher Collins is with University of Ontario Institute of Technology, Canada. E-mail: christopher.collins@uoit.ca*

2 Related Work

This work is informed by past research which falls primarily into two categories. We first describe works that deal with scene construction based on text data, we then walk through focus + context techniques and discuss how they are related to our visualization design.

2.1 Descriptive Illustration

Descriptive illustration is a sub-field of rendering focused on generating static illustrations or animated sequences based on text input. Unlike other text visualizations with abstract outputs, descriptive visualization is concerned with a literal representation of the text.

WordsEye [9] demonstrates one of the more fully-realized text-to-scene generation systems. It uses a combination of grammatical rules and heuristics, along with a large repository of 3D models to render an image which represents the literal meaning of the text. WordsEye is restricted to short descriptive sentences composed of adjectives, nouns and spatial prepositions which relate them, e.g."The tawny cat is on the red couch." CarSim [19] uses a similar logical model and rendering process, but deals with animated scenes for reconstructing car crashes from incident reports. These approaches contrast with text visualizations which use aesthetic arrangements of word forms to reveal content (e.g.[15, 25, 28]). These systems render an array of words using random packing or by fitting them into pre-determined geomet-

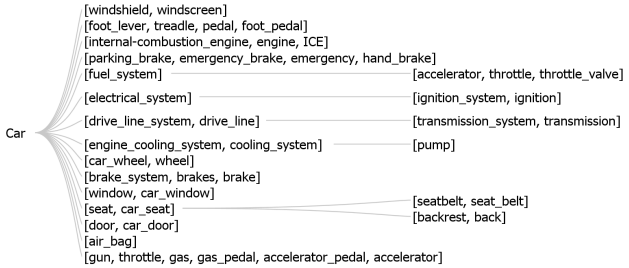

Fig. 2. A subset of the component keywords hierarchy. Words in square brackets are synonyms. The components in this hierarchy were isolated in the 3D model of a car, and extracted from the text data and counted.

ric shapes designed to be semantically meaningful, such as rendering words about flowers into the shape of a blossom.

Our application differs in that we aim for an interpretive view which summarizes common nouns in a collection of text, rather than deeply processing short sentences to construct representable logical forms. The spatial positions in our system are not determined by the prepositions in the text, but rather determined by their inherent positions as part of the 3D model which serves as the substrate of our visualization.

Another related area of descriptive rendering is systems which already have a defined scene geometry and use input data to modify how a scene is rendered. For example, IBIS is one of the first systems to use dynamic rendering based on user-specified input [22]. IBIS parses a set of text-based rules to determine the optimal perspective and lighting parameters that satisfy them. More recently, Sedlmair *et al*. augmented a system that visualizes communication among vehicle components with a 3D model of the vehicle itself, using the model to show the communication paths [21]. Our work is related in that we augment a non-photorealistic rendering of a car to emphasize important regions, however our visualization is driven by unstructured text data and we provide for direct interaction on the 3D visualization itself.

2.2 Focus + Context

Focus + Context is a well-known technique used to draw attention to areas of interest while maintaining visibility of contextual information. There are various ways to discriminate these areas, for example Hauser [13] reviews a general approach in which the use of colour, space, opacity and other resources are demonstrated.

A lens widget can be considered a focus + context technique that maps to a real-world metaphor of a magnifying glass. The area or objects "under" the lens can take on additional properties or receive additional interactions, as seen in examples such as Magic Lens [5], 3D Magic Lens [24] and NPRLens [18]. Our lens system drew inspiration from these examples. In addition, our lens widget is location-aware as our scene is spatially segmented, which is somewhat similar to the BrainGazer visual query paths [6].

Lenses can also be used for dynamic labelling, in which the lens is used to reveal additional data that would have otherwise created too much visual clutter on the display [11, 12]. In addition, summaries and other abstract information can be displayed as part of the labelling scheme [4]. Our system uses a similar approach, but additional small visualizations are shown instead of textual labels.

3 Problem

Within documents which describe physical objects, the words may have two types of context. First, the context of the word in the sentences of the document (i.e., co-occurrence relationships). Second, the context of the word in terms of its location in the physical world and its spatial relationship with other words mentioned in the document. For example, the sentence "The brake lights would not turn on, but the horn continued to sound." indicates a co-occurrence relationship between "brake lights" and "horn," but the physical contexts of these objects in a real vehicle are quite different.

Traditional text visualization approaches such as word clouds could be used to summarize the reports about a particular subject matter over a given time period, but they present the common terms out of context—both co-occurrence context and physical space context. Some approaches to text visualization spatially organize words based on semantic relationships such as the "is-a" relation [8], but we are unaware of any visualization which approaches text analysis by visualizing the real-world spatial contexts of the words in the text.

Revealing the spatial dimensions may have several benefits. Foremost, the familiarity of the form makes the subject matter immediately recognizable to experts and novices alike, combined with the message-carrying capability of NPR illustrations, we argue that our approach creates a rich, engaging experience. Second, it is possible to conduct a different type of data exploration: the spatial dimension allows us to explore proximal relations and filtration by spatial volumes, possibly allowing new insights to be formed.

Consider product quality reports for a musical instrument. Visualizing these report allows one to see exact location of the problems on the 3D model, for example: which valves are failing. Seeing the instrument in physical form may promote conjectures that are less apparent with text or abstract visualization, for example: perhaps the valves failed because they are encased in a faulty housing. There are many applicable datasets which carry this sort of physically mappable vocabulary: hotel and consumer product reviews, technical manuals, and technical support logs are examples.

3.1 Use Case: Vehicle Complaint Reports

We choose to demonstrate our approach on a dataset of vehicle complaint reports. Each year thousands of reports are submitted to the NHTSA database, consisting of consumer complaints, defect reports and manufacturer recalls. Each report has fixed fields describing the details of the incident (date, make, model, etc.), and a free-form text field, typically containing several sentences which describe the incident in detail, including what physical parts were damaged or broken. Thus this data can be mined for frequency counts as well as co-occurrence counts of car parts. All together the meta data and free text offer a wealth of information on safety and reliability issues of vehicles. Consumers can access this data online to support car-buying decisions. The current interface uses a conventional search form, returning long lists of textual results; there are no mechanisms to support concise overviews or dynamic details-on-demand.

In order to derive concrete requirements for our design, we needed to determine the considerations which are most important to a consumer in a purchasing decision. To determine this, we surveyed car-experts' columns, user forums, and review services such as Consumer Reports and Edmunds. Our findings revealed that aside from price, safety, trends and capability to relate problems to each other are of high concern. Based on these findings, we have four design requirements for our visualization:

1. Provide an intuitive representation and make important items clearly visible;
2. Facilitate finding of trends, interesting facts and causal relations in the reports;
3. Allow multiple types of comparisons such as time, components and make/model of vehicles;
4. Provide for reading of the original complaint report in the context of the visualization.

4 Data Processing

We applied several processing steps to each text record. First, we parsed the metadata fields. Then we parsed all nouns, and filtered this list to create a hierarchical vocabulary of components mentioned in vehicle defect reports. Each component reference was then counted in all documents to calculate a word score and co-occurrence scores for that (component, document) pair. Finally, we segmented 3D models to match our keyword ontology. These steps are explained in greater detail below.

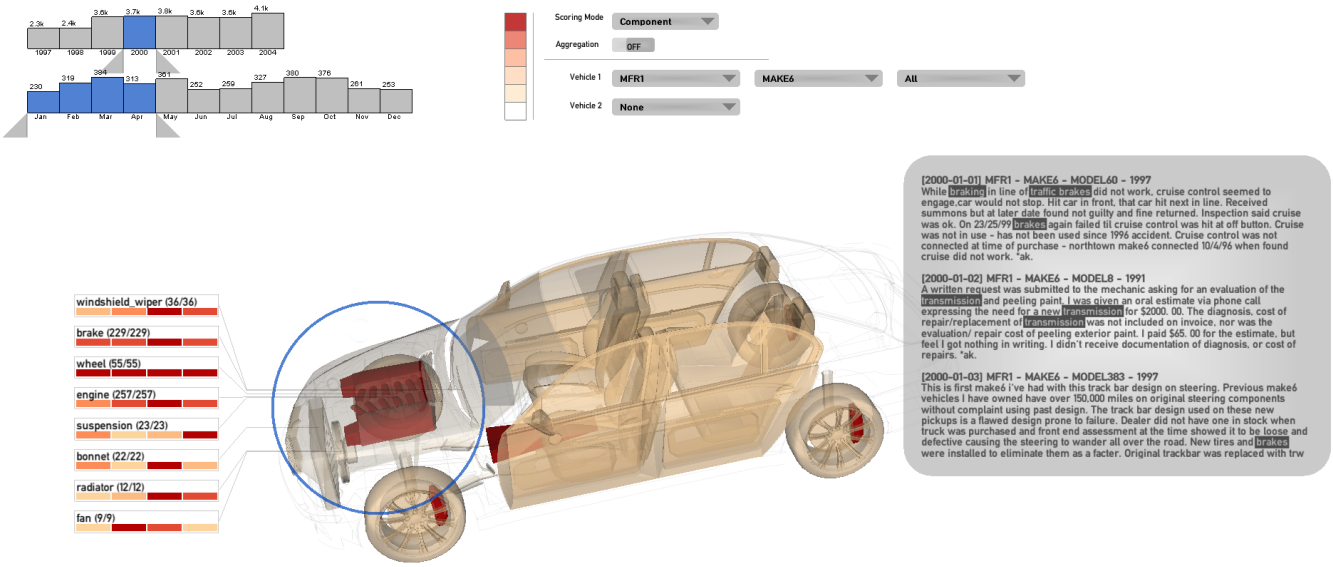

Fig. 3. The Descriptive NPR visualization in use. The lens widget is used to solicit more details in the front portion of the vehicle. The time frame is set to include the first four months of the year 2000. The document panel at right shows some texts matching the filter query.

4.1 Hierarchy Generation

For keyword extraction, we leverage the concepts of meronym and synset from WordNet [16]. A meronym in WordNet describes a *part-of* relationship between two objects. For example: wheel is a part of (a meronym of) a vehicle. Using "vehicle" and "automobile" as root words we created a hierarchy of part-of relations from the most dominant object down to the most specific sub-components. A synset represents a set of semantically equivalent words, for example "bonnet" and "cowlings" are semantically equivalent to the component "hood." For each node in the hierarchy, we also extracted its synsets in order to increase our vocabulary coverage. Thus, our initial hierarchy was based only on WordNet.

As an additional step to expand our dictionary of keywords, we performed part-of-speech (POS) tagging on our document corpus. POS taggers look at the grammatical structure of text and break down sentences into lexical categories such as nouns, verbs and adjectives. We then collected the most frequently occurring nouns and manually selected the relevant terms, adding them into our parts hierarchy when they did not already exist.

Lastly, we manually pruned the hierarchy to remove any keywords that may cause false-positives in our matching scheme, for example: "first", "second" and "third" refer to first, second and third gears respectively, but including them will likely result in counting more occurrences of "gears" than we should. A sample of the final component hierarchy is in Figure 2.

4.2 Text Scoring

Each text document may contain multiple component keywords (e.g., "brake"), and each component keyword can also appear multiple times within a single document. Let G be a (possibly empty) list of currently selected components. The score for component c, $S(c,G)$ is defined to be the total number of documents that have at least one mention of c and G. Thus when G is the empty set we get the absolute occurrences for c alone, when G is non-empty the score reflects the co-occurrence strength among a set of components. Each text document is only counted once per component to discourage biases coming from longer documents where parts are repetitively mentioned.

Early in our design, we considered applying sentiment analysis to our data, so that we would know if mentions of a particular component were positive or negative. Given the poor performance of currently-available sentiment taggers, and rather uniformly negative nature of the sentiment in the car defect report database, we make the assumption that when a car part is mentioned, that mention indicates a failure of that part. We know this assumption not to hold in all cases, but given the large number of documents in our dataset, we believe the errors will be subsumed by the volume of data. This problem is made less serious by the ability to drill down to read the underlying documents to confirm the problems with (or praise about) the mentioned components for oneself.

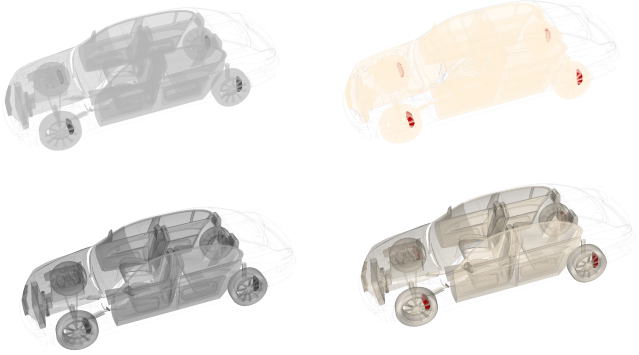

Fig. 4. Design options using hues and lighting effects. Top-left: Single-hue, solid colour. Top-right: Multi-hue, solid colour. Bottom-left: Single-hue, simulated lighting. Bottom-right: Multi-hue, simulated lighting (used in our prototype).

4.3 3D Model Preparation

We manually segmented the 3D model to match the major components in the hierarchy. We organized segments into mesh groups and assigned a unique identifier to each group. Where there are missing components we inserted placeholder geometries, using vehicle schematics as a guide.

In this work we have chosen to use a sedan model to represent the vehicle population, as the majority of the reports are related to sedans. While there are variations in how parts are placed spatially among

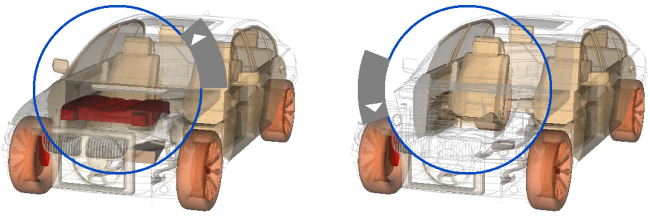

Fig. 5. Left: Lens widget with no depth plane specified. Right: Lens widget with active depth plane cutting into the model. The wheel and part of the radiator are in front of the plane, thus rendered as outlines.

different brands of vehicles, we believe in general there are sufficient similarities to be represented by a single model. We could provide specific 3D models as they become available.

5 Design of the Descriptive NPR Interface

The interface is composed of four components, as seen in Figure 3:

1. A stylized rendering of a 3D model forms the central point of our visualization. It emphasizes the most highly scored components to create a *pop-out* effect.
2. A lens widget to explore the 3D visualization; the lens extracts detailed information of entities under the lens.
3. Heatmap visualizations for showing trend and outliers for individual entities.
4. The document panel displays the source text documents that match the current query.

In addition, we created a set of filter widgets that provide domain-specific navigation for our vehicle dataset. These filters are fixed in place at the top of the display space.

We designed the visualization to run on large, multi-touch displays. We see potential deployment of our system in a walk-up-and-use or office meeting scenarios. Though the design of the touch interactions was an extensive part of the research, our focus for this paper will be on the visualization technique.

5.1 3D Visualization

The central contribution of this work is the mapping of abstract semantics onto realistic looking, 3D models. But this is also a source of complication: we have to deal with the additional difficulties of navigating in three-dimensional space, as well as work around perceptual limitations. So why use 3D models in the first place? We make the argument that the entities are inherently in 3D space, thus it is more natural to observe them in their real-life representation. While a series of linked 2D images may provide some form of realism, a viewer would have to observe multiple images and mentally connect picture to picture to get a sense of overview.

5.1.1 Rendering

Non-photorealistic Rendering (NPR) can be considered to be any computer generated graphics that do not involve the accurate simulation of light. We incorporated several NPR effects into our rendering process, in particular, we use hue/transparency to denote importance and switch between outline and non-outline styles for semantics. Examples of such scene segmentation strategies are often employed from the Scientific Visualization field (e.g., [23, 26]).

Effects on the 3D model are assigned based on scores of entities in the text. Entity scores are dual encoded as both hue and opacity on the 3D geometry. We map the scores to a yellow-to-red hue scale which is further divided equally into six bins. While this setup has a limited granularity, it is easier to perceive values from a small number of hues than a continuous scale. We vary opacity values to partially mitigate occlusion issues and to emphasize highly scored entities, thus the amount of opacity is proportional to the entity's score. To ensure at least partial visibility at all times, the opacity is limited to a mid-range between 40 percent and 80 percent. Entities with zero scores are rendered qualitatively differently, with an outline style in a just noticeable colour so they are visible, but not overly distracting [2]. This indicates an 'inactive' state, but maintains spatial context. Outlines are computed by comparing dihedral angles between neighbouring polygons [20]. To show entity selections, we draw a silhouette-like contour around the selected geometries to simulate a glowing effect. We rendered the silhouette by creating blurred versions of the 3D geometries as textures, then superimpose them back into the final scene.

5.1.2 Selection

By default the application has no entities selected, thus the visualization reflects the absolute number of occurrences of each entity. As selections are made, each entity's score is recomputed to show co-occurrence relations with the selection. Note since selected objects fully co-occur with themselves, they are promoted to the highest bin. In this manner, high correlations are red and highly opaque, while low correlations are yellow and highly transparent. For example, if we select the windshield wiper, the visualization is rerendered to show entities that co-occur with the wiper component and the strength of this relation. If we also select the windows, the visualization will show co-occurring relations to both windshield wiper *and* windows. This example is reflected in Figure 1. The rerendering process is facilitated by an animated transition that interpolates the graphical effects.

5.1.3 Design Trade-offs

There are several design trade offs with our technique. Blending in 3D space may produce artifacts as there is no guarantee of hue preservation. A preservation scheme does exist [7], however we did not implement it due to additional performance complexity of per pixel adjustments. Subjectively, we did not find any visual distractors thus decided this was not necessary. A single hue approach was tried with varying saturation and opacity, but it lacked the *pop-out* effect visible in multi-hue schemes. The effect of lighting is another design trade off. Lighting effects enhance the surface shapes so objects are easily distinguishable from one another. However, the colour properties are modified such that they no longer match those in the scale legend. A complete lack of lighting produces the opposite problem: the colours match exactly, but it is difficult to distinguish objects in the model, especially if they are near or partially occlude each other. Adding object outlines in non-lighting situations helps, but is not aesthetically pleasing and can produce visual clutter. Ultimately we decided that object recognition and familiarity outweigh the colour offsets. We contend that with only six buckets on the scale the lighting effects do not disturb the colour perception enough to obscure which hue-bin the component belongs to. Our experimental results support this contention. The various trade-offs can be seen in Figure 4.

A specific challenge of using transparency in 3D rendering is that modern GPUs do not provide native support for out-of-order blending: objects that are positioned behind others can appear to be in front if they are not rendered in depth order, whether it be back-to-front or front-to-back. Sorting geometries by their distance away from the camera can help but is expensive, view-dependent and does not solve certain pathological cases where geometries intersect. Recently there are quite a few developments in order-independent-transparency that yield better results, e.g. [3, 10, 17, 29]. We implement a version of dual-depth peeling [3] for our prototype. It yields accurate and pleasing results at sufficient speed, with less reliance on specialized hardware features.

5.2 Lens

Using a metaphor of looking through a magnifying glass to reveal more details about a specific object, we created an interactive lens to extract and show detailed information about entities in the text.

The lens operates in a hybrid 2D / 3D space: the lens itself exists on a flat 2D plane and casts a cylindrical query volume into the scene. To be able to query different entities, each entity object is tested to see if its centroid is in the querying volume. Entity objects in the lens

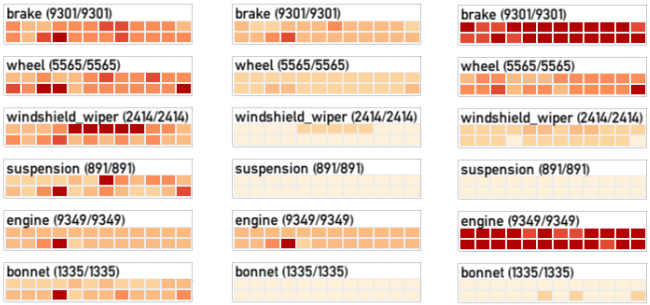

Fig. 6. Heatmap widgets for several entities. Left column: Component-max perspective. Centre column: Global-max perspective. Right column: Month-max perspective.

will activate their heatmap charts, which are displayed alongside the lens' circumferences in a flush right / flush-left manner. To associate 3D geometry and the heatmap chart, we connect them together with line segments similar to the technique presented in [11]. More advanced labelling algorithms exist and may produce more eye-pleasing layouts [1, 12], though they are not implemented in our prototype and are considered future work.

The lens widget utilizes its own rendering pipeline, object geometries are sent into the pipeline as normal, the rasterized result is then stored in an intermediate buffer and later combined in fragment shaders with the default rasterization scene. This is an independent process, and thus allows us to render the lens' scene in different rendering styles and semantics. To visualize the lens widget itself, we draw a semi-transparent border around its circumference so viewers are aware of its existence. When interacting with the lens widget, the widget is active and we render the border in blue, otherwise we use the default grey colour. The semantics of the lens is not impacted by whether the lens is active or inactive.

The lens enables three different actions. The position of the lens can be moved by dragging within the lens, impacting the currently selected filters and the heatmap charts. The lens can be resized by dragging on the border of the lens, increasing or decreasing the query area. Lastly, the depth plane can be adjusted by rotating the depth selector tab around the circumference of the lens (see Figure 5). The depth plane function provides a method for people to reduce occlusion, as all entities that are the cut by the plane are drawn in an outline style, allowing viewers to see through them and into the object. Objects that are cut off are excluded from any scoring calculations, they also have their heatmaps hidden to reduce visual clutter. These three interactions can be combined together to create a rich, flexible query mechanism.

5.3 Heatmaps

Heatmap visualizations on the entity labels communicate how each entity's score changes over time. We project the time-series data onto two dimensional grid much like a calendar. We considered other designs, such as small bar charts, but decided on heatmaps because of their space-efficiency and ability to support yearly (row-to-row) and monthly (column-to-column) comparisons . For example, it is easy to compare spring seasons across multiple years, as the relevant cells are adjacent to each other. Each grid cell is colour coded with respect to the 6-bin colour encoding to denote the scoring strength. The heatmap is labelled with the name of the entity, and the overall occurrence and co-occurrence scores. The entity labels could be easily adjusted for other analysis scenarios to show other forms of data or a different type of visualization.

The heatmap is selectable, which is equivalent to making a selection on the 3D visualization. Moving one's finger over a heatmap cell will toggle a brushing effect, drawing a border around the cell and all cells in the same time period across all visible heatmaps, allowing for quick comparison. A tooltip is also displayed, showing the cell's data in numerical format.

5.4 Document Panel

The document panel is the final stage of our drill-down process, providing the complete texts which drive the higher level views. Documents matching the current filter query (hierarchy, time, component selections) are shown in chronological order in the panel. Each document is divided into two sections: the header section shows each documents fixed attributes and the content section shows the raw text descriptions. We denote the selected entity words and co-occurring entity words using blue and grey highlights respectively. Scrolling is enabled by a single finger drag along the right border of the document panel. The document widget is hidden by default, and can be activated by a touch and hold gesture on the background. Once it is activated it can be moved around by the viewer to a desired screen location.

5.5 Filter Widgets

Two types of filters are available: time and organizational hierarchy. The time filter consists of two sliders representing year and month, the selection of month and year are independent of each other. The hierarchy filters are shown as a set sequence of drop-down lists that allow viewers to refine their queries successively, from most general to most specific: manufacturer, make, model and model year. Histograms are embedded into both widgets to enhance visual cues [27], with each bar representing the volume of documents for each time period, or each organization.

6 Enabling Analysis

Aside from query-based manipulation, we provide higher level interactions to ease finding trends and making data-driven decisions.

6.1 Heatmap Views

The heatmaps data is, in essence, time series data. There are multiple ways we can view this time series to derive interesting patterns and trends. For this prototype we want to concentrate on comparability across different components and across time. For example, we want to enable questions such as: *Which vehicle component had more complaints? Which month had the most complaints during the year?* To do this, we need to have multiple perspectives where data can be viewed from different functional needs (see Figure 6). We have provided the following perspectives, each using a different denominator to normalize component scores:

- Month-Max: A monthly perspective where the score of each month is divided by the maximal score for that month, over all components.
- Component-Max: A component perspective where the score of each month is divided by the maximum score for that component during the selected time.
- Global-Max: A global perspective where the score of each month is divided by the maximum score of all components over the selected time.

Each of the perspectives above answers different questions and has its own advantages and disadvantages. The month-max perspective allows us to see which component had the most occurrences in a given month, but comparison of adjacent cells is meaningless because each month uses a different base value. The component-max perspective is the opposite, it allows us to see trends within a single entity, but it does not allow comparison across components. Lastly, the global-max perspective is good at showing the outliers and supports both month-to-month and component-to-component comparisons, but it is difficult to see overall trends because the outliers, if any, will dominate and push all non-outliers into the same scoring bin.

For consistency, the perspective is the same for all heatmaps and can be changed using an on screen drop-down selector.

6.2 Comparison

Comparison mode allows people to compare entity occurrences across two different subsets of the data. To select data to compare, we provide two sets of filter widgets which can be used to specify manufacturer, make, model and model year. Each set of filters specifies a query, which we will call Q1 and Q2, and each query is assigned a colour,

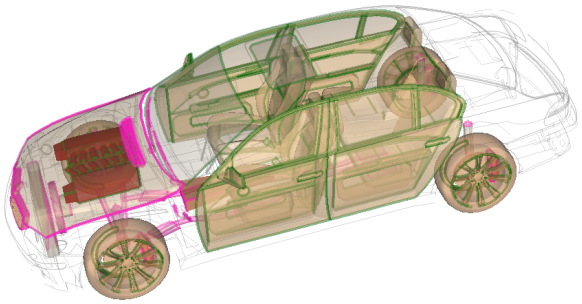

Fig. 7. In the comparison view, the outline on each component indicates whether vehicle one (pink) or vehicle two (green) had a higher rate of failure. The fill colour indicates how frequently the component failed across both vehicles (the sum). The engine is mentioned frequently in this comparison, and more often in relation to vehicle two.

which is used in the visualization. For example, we can compare Honda Civic (Q1) to Toyota Corolla (Q2), or we can compare Ford Focus (Q1) against all other Ford vehicles (Q2), by not fully specifying Q2. Comparison mode is implicitly activated by specifying Q2. When Q2 is empty, normal mode is enabled.

Two separate measures are used to render the comparison view. The *contribution sum* is the aggregated component score from the two query sets: it reflects the overall importance of the component by emphasizing the most frequently occurring components matching Q1 and Q2. The *percentage difference* describes the relative frequencies of a component, whether it occurs more frequently under Q1 or Q2 relative to the total contributions from Q1 and Q2 respectively. The percentage score is calculated as the component score divided by the total contribution. Then the percentage difference follows as percentage score Q1 minus percentage score Q2, with the sign and magnitude indicating which query set has the stronger presence of that component. We made the decision to use percentage based comparisons because it enables the comparison of query results of different sizes.

These scores are used to render the 3D view (see Figure 7). Using the percentage difference, the colour of the outline of a component indicates which query set has the higher rate of complaints, and the opacity of the outline indicates the strength of the difference. Using the contribution sum, the standard hue and opacity encoding is used to indicate the sum of the two query sets, giving an impression of the overall importance of that component. Thus, a highly problematic component from both queries will have a strong presence overall but with a faint outline, while a lopsided but infrequently mentioned component will have strong outline but barely visible interior colour.

6.3 Aggregation

By default, the system treats each object individually rather than object groups. For example "seatbelt," "backrest," and "seat" are all scored separately, even though they are logically under the group "seat." This setting allows people to isolate and identify unique problems accurately. There are times, however, when this level of information is unnecessarily detailed and a higher level of abstraction is desirable.

Aggregation mode mimics the type of high level rating system found on consumers review websites. When aggregation mode is enabled, individual objects, and their scores are aggregated up to the first level entities. In our specific case, the first level are the major subsystems in a vehicle. Aggregated components respond to interaction events as a single group (e.g. selecting "seatbelt" selects the entire "seat" sub-system). Figure 8 shows a before and after illustration of using aggregation mode. We can observe that the engine and seat subsystems have much higher severity when viewed as a whole rather than their individual parts.

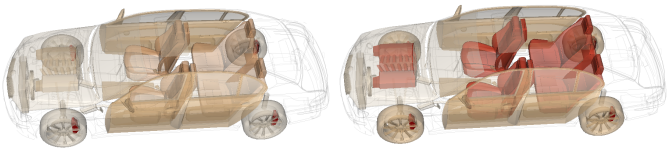

Fig. 8. Aggregation mode. Left: disabled. Right: enabled.

Fig. 9. A participant using the lens to examine the vehicle.

7 Evaluation

We conducted a preliminary evaluation study of our prototype. The study was largely qualitative in nature: our goal was to assess if, and how a person can use the visualization to facilitate his/her analytical tasks, which are framed around scenarios of assessing safety and reliability issues to support purchasing decisions. We further present several use case scenarios to demonstrate the potential of our visualization.

7.1 Methodology

We recruited 12 participants from the student population at a technical university. All had experience with touch interfaces, and six had previously used a 3D interface. Only two participants owned their own vehicle, but six had previously purchased a car and seven had previously investigated vehicle safety issues in some way. Participants were asked to perform analytical tasks using the visualization, which ran on a 60 inch multitouch enabled display. Our study setup can be seen in Figure 9. Note that in order to reduce personal bias, we removed identifiers and replaced them with placeholders, for example we replaced "Toyota" with "MFR1."

After a brief tutorial on how to interpret the visualization and how to use the interface, participants were asked to perform three sets of tasks. The first set consisted of warm-up exercises aimed to help participants become familiar with the interactions (these are excluded from our analysis). Next came a set of focused questions regarding interpretation of the visualization. For example, "Select the most complained about component in the year 1999," or "What components in the vehicle are associated with complaints about windshield and wheel?". Subjects then had to adjust the visualization and identify the correct entities. Finally, the last set of questions were more open-ended in nature. For these, specific scenarios were presented to the subjects, they were asked to analyze the scenarios and come up with a decision based on the visualization and its widgets. For example, we asked: "Between 1997 and 2000, which of the vehicles X and Y would you purchase and why? Assume they are similarly priced." After completion of the computer-based tasks, we conducted a semi-structured interview to solicit opinions from participants about their experiences using our system.

Each study session took approximately one hour to complete, though participants were allowed to take as much time as they wanted on any task. Each session was recorded on video, and touch interactions were logged by the system.

7.2 Discussion

12 participants took part in the study, however one session was excluded from our analysis; the participant exhibited a lack of English language skills which lead to difficulty understanding the instructions and contradictory statements during the interview portion.

In general, feedback was favourable and most tasks were completed reasonably well, in the sense that correct answers were derived based on findings from using our system. There were a few exceptions: some participants did not correctly respond to the focus questions that asked them to identify outliers (3/11 and 1/11). This may partly be attributed to initial unfamiliarity with what the visualization is trying to show, as one participant (P5) revealed later that at first the answers were not based on the visualization, but rather on personal opinion about automotive vehicles. The other exception was a task that required a person to switch among multiple visualization views and compare them in sequence. The added cognitive load of memorizing states likely contributed to the varied answers.

Participants enjoyed using the 3D visualization along with the lens widget. Several participants made explicit comments with regards to the usage of familiar form factor: "Nicer to look at a picture than a bunch of numbers."(P1), "everything is in detail, very interactive. [...] The visual, is self-explanatory" (P8), "I can see clearly each part in the car, so I can know what to choose"(P6) and "it is relatable, I've been in cars and I've had the opportunity to see some of the components." (P2). Using the lens widget for dynamic focusing of interesting data was voiced by several participants: "kind of cool, being able to dissect with it.' (P1) and "You can zoom in to the parts that you cannot really understand, for example the transmission." (P7).

The heatmap widget received the least favourable responses, with four participants mentioning that it provided too much low-level detail, in particular P11 suggested using a pie-chart or bar-chart for each component in comparison view. The design implication here may be to provide different levels of granularity that can be dynamically adjusted. We observed that three participants had difficulties understanding co-occurrence, however, we are not sure if this is due to insufficient explanations at the start of the study, or the fact that the same colouring scheme that is used for both occurrence and co-occurrence caused the confusions. Several interaction issues such as executing touch gestures were noted during the study, however participants seemed to enjoy using the visualization: "even though there are some interaction problems, it just looks really good!' (P10).

In Figure 10, we present a summary of the interactions logs for the task of comparing two different types of vehicles. In this task participants were free to use any widgets they want to investigate which of the two vehicles is more reliable. From the figure, we see that participants first remove irrelevant data using the filters, then they either interact through the lens or directly with the 3D scene to explore the visualization. The navigation actions were mostly executed in short bursts, followed by an idle period. This behaviour appears to correspond to finding interesting data and then spending time to assess his/her findings. We thought for most participants the lens widget would be used only after the 3D model is moved into a desired orientation, however this is not supported by the logs. The participant strategies seem to group into three types: use of both the 3D scene and the lens widget for exploration (P3,6,8,9,10), only the lens (P2,7), and only the 3D scene (P1,4,5,11). Surprisingly, participants did not make use of the document widget during this task. Further investigation of the role full-text details can play in decision-making using d-NPR is warranted.

7.3 Analysis Scenario: Toyota Recall

In addition to experimental testing, we can examine how our visualization system works by reviewing example scenarios. The Toyota vehicle recall happened between September 2009 to February 2010 and had to do with faulty accelerators and brakes. We wondered if

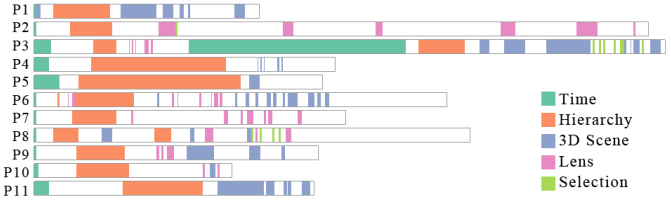

Fig. 10. Widget usage for the task of comparing two vehicles.

there were any leading or lagging indicators in our data, which may be indicative of the problem. We set the visualization time frame to be between 2008 and 2010 and selected Toyota as the vehicle manufacturer. The first thing we noticed is that the engine, usually one of the highest occurring components in the complaint reports, no longer dominated the visualization. Instead, two components pop out: brakes and accelerator. A closer examination with the lens widget raised more questions. The heatmaps show that there are two outlier months where a huge amount of complaints were registered: February 2010 and March 2010, after the recall was announced. Perhaps the widely publicized event triggered a loss in consumer confidence, which in turn led to an over-reporting of problems. This is supported by the sharp drop-off after March 2010.

7.4 Analysis Scenario: Using Spatial Dimension

This scenario describes how a regular consumer, Daniel, may use the visualization to research a problem. Daniel has about three years of driving experience but does not know a lot about cars. Recently, while driving, he noticed a rattling sound coming from the front passenger side of his vehicle. He decides to conduct some research on his own before taking the car back to the dealership.

Using the visualization, Daniel filters the dataset to focus on his vehicle model. Since he is not sure exactly where the noise came from, he uses the lens widget to focus on components near the front-passenger region. Using the lens, he can see that the suspension component has a higher number of complaints registered against it than the other components in the focused area. Wondering if the suspension is normally a problematic component, Daniel activates the comparison mode to view his vehicle against all other models by this manufacturer. He finds that the suspension for his model is more frequently reported than for other cars. He then selects the suspension component using the heatmap, and toggles the document widget so he can read through the actual complaint reports. After a few minutes of reading, Daniel notes that there are at least eight or nine reports that seemed to document similar noise issues and point to defective suspension setup. He decides that he should contact his dealership to have them check it out.

8 Implementation

Our prototype is written in Java and Java OpenGL. The dataset is freely available on NHTSA's website and consists of data from 1995 to present day. Our 3D model is taken from Google 3D Warehouse[1] and has approximately 160K vertices. Our application runs at 1680x1050 resolution and was able to achieve between 15-30 FPS, depending on model complexity and size of the dataset. We use a 3.2GHz I5 Quad Core CPU.

9 Discussion and Future Work

The system described here is a working prototype that is capable of allowing people to explore text entities through the use of 3D NPR-renderings and spatially-aware UI widgets. However, there are still some outstanding opportunities for improvement which can form the basis for future work.

Our parsing uses a simplistic dictionary approach. Creating a parser capable of deciphering more complex relations, such as dependency relations, would enrich the data. Another path to explore is sentiment

[1] http://sketchup.google.com

analysis. We did not pursue this path as our dataset has an overall negative connotation, however in the more general sense, sentiment analysis may add value, particularly if we were to apply our approach to consumer product reviews where mentions are often both positive and negative.

The initial setup cost of procuring the hierarchy and 3D models can be time consuming and labour intensive. However, our semi-automatic hierarchy construction method, using widely accepted lexical resources, is certainly faster and more accurate than fully manual creation. The hierarchy and model generation is a one-time cost and the hierarchy can be adjusted or expanded for similar objects without starting anew.

In terms of rendering, varying opacity and hues only partially solves the occlusion issue; where there are densely packed geometries it can still be difficult to distinguish entities from one another. Alternative approaches, such as using exploded views to expose densely packed regions may help, though exploded views can also distort spatial relationships. It is also important to note that due to the physical size differences between objects in the model, it may be that large objects (e.g.windows) receive additional attention while they are arguably less important than small objects (e.g.brakes). We considered enlarging/shrinking parts based on scores, but this design would require readers to know the baseline (undistorted) size of objects. The use of distortion as a descriptive rendering technique is a promising avenue for future investigation.

Several participants experienced some difficulty selecting objects which were occluded or in dense regions. The lens and heatmaps offer one way to improve selection in these situations. Another possibility to ameliorate difficulties with selection of components in dense regions is to use a priority-based selection based on component scores.

Lastly, our prototype does not take into account the social aspect of analysis—we allow a variety of exploration capabilities, but did not offer any functions for people to save, restore and share their findings. Allowing people to save, share and search other people's findings will certainly make the application more useful. Analyzing long-term usage patterns may also yield more insights.

10 Summary

In this paper, we introduced a text visualization approach which we call *descriptive non-photorealistic rendering*. Our contributions are: (1) A novel approach for text visualization that combines spatial and non-spatial attributes to create a visualization that is based on real world semantics. (2) An interactive visualization prototype using our approach for the analysis of vehicle defect reports. An initial study of our system showed favourable reception by study participants, and that they were able to use the visualization to analyze the text corpus. However, a thorough comparative study is needed to gauge the effectiveness of our visualization against standard approaches.

While the examples given in this paper are framed around the automotive industry, d-NPR is generalizable to other domains. It would be interesting to explore new datasets to see what type of insights can be gleaned from our technique. For example, we mocked up an example of exploring a visualization of building maintenance records, as seen in the far right of Figure 1. In the future, we would like to explore alternative descriptive rendering techniques, refine the interaction design, and apply our approach to analyze different datasets.

References

[1] K. Ali, K. Hartmann, and T. Strothotte. Label layout for interactive 3D illustrations. *Journal of the WSCG*, 13(1):1–8, jan/feb 2005.

[2] L. Bartram and M. Stone. Whisper, don't scream: Grids and transparency. *IEEE Trans. on Visualization and Computer Graphics*, 17(10):1444–1458, Oct. 2011.

[3] L. Bavoil and K. Myers. Order independent transparency with dual depth peeling. Technical report, NVIDIA, 2008.

[4] E. Bertini, M. Rigamonti, and D. Lalanne. Extended excentric labeling. *Computer Graphics Forum*, 28(3):927–934, 2009.

[5] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and Magic Lenses: The see-through interface. In *Proc. of SIGGRAPH*, pages 73–80. ACM, 1993.

[6] S. Bruckner, V. Šolteszova, E. Groller, J. Hladůvka, K. Buhler, J. Y. Yu, and B. J. Dickson. Braingazer - visual queries for neurobiology research. *IEEE Trans. on Visualization and Computer Graphics*, 15(6):1497–1504, Nov. 2009.

[7] J. Chuang, D. Weiskopf, and T. Möller. Hue-preserving color blending. *IEEE Trans. on Visualization and Computer Graphics*, 15:1275–1282, 2009.

[8] C. Collins, S. Carpendale, and G. Penn. Docuburst: Visualizing document content using language structure. *Computer Graphics Forum (Proc. of the Eurographics/IEEE-VGTC Symp. on Visualization (EuroVis))*, 28(3):1039–1046, 2009.

[9] B. Coyne and R. Sproat. Wordseye: An automatic text-to-scene conversion system. In *Proc. of the Annual Conf. on Computer Graphics and Interactive Techniques, SIGGRAPH*, pages 487–496, 2001.

[10] C. Everitt. Interactive order-independent transparency. Technical report, NVIDIA, 2001.

[11] J.-D. Fekete and C. Plaisant. Excentric labeling: Dynamic neighborhood labeling for data visualization. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI)*, pages 512–519. ACM, 1999.

[12] M. Fink, J.-H. Haunert, A. Schulz, J. Spoerhase, and A. Wolff. Algorithms for labeling focus regions. *IEEE Trans. on Visualization and Computer Graphics*, 18(12):2583 –2592, dec. 2012.

[13] H. Hauser. Toward new grounds in visualization. *SIGGRAPH Computer Graphics*, 39:5–8, May 2005.

[14] H. Hauser, D. Weiskopf, K.-L. Ma, J. J. van Wijk, and R. Kosara. SciVis, InfoVis - bridging the community divide?! In *IEEE Visualization Conf. Compendium*, pages 52–55, 2006.

[15] R. Maharik, M. Bessmeltsev, A. Sheffer, A. Shamir, and N. Carr. Digital micrography. In *ACM SIGGRAPH 2011 papers*, SIGGRAPH '11, pages 100:1–100:12, New York, NY, USA, 2011. ACM.

[16] G. A. Miller, C. Fellbaum, R. Tengi, S. Wolff, P. Wakefield, H. Langone, and B. Haskell. WordNet: A lexical database for the English language, Mar. 2007.

[17] K. Myers and L. Bavoil. Stencil routed a-buffer. In *Annual Conf. on Computer Graphics*, 2007.

[18] P. Neumann, T. Isenberg, and M. S. T. Carpendale. *NPR Lenses: Interactive Tools for Non-Photorealistic Line Drawings*. 2007.

[19] O. Åkerberg, H. Svensson, B. Schulz, and P. Nugues. CarSim: An automatic 3D text-to-scene conversion system applied to road accident reports. In *Proc. of the Conf. of the European Chapter of the Association for Computational Linguistics*, pages 191–194, 2003.

[20] R. Raskar. Hardware support for non-photorealistic rendering. In *Proc. of the ACM SIGGRAPH/Eurographics workshop on graphics hardware*, pages 41–47, 2001.

[21] M. Sedlmair, K. Ruhland, F. Hennecke, A. Butz, S. Bioletti, and C. O'Sullivan. *Towards the Big Picture: Enriching 3D Models with Information Visualisation and Vice Versa*. 2009.

[22] D. D. Seligmann and S. Feiner. Automated generation of intent-based 3d illustrations. In *Proc. of SIGGRAPH*, pages 123–132. ACM, 1991.

[23] C. Tietjen, T. Isenberg, and B. Preim. Combining silhouettes, surface, and volume rendering for surgery education and planning. In *Proc. of the Eurographics/IEEE-VGTC Symp. on Visualization (EuroVis)*, pages 303–310, 2005.

[24] J. Viega, M. J. Conway, G. Williams, and R. Pausch. 3D Magic Lenses. In *Proc. of the ACM Symp. on User Interfaces and Technology (UIST)*, pages 51–58. ACM, 1996.

[25] F. B. Viégas, M. Wattenberg, and J. Feinberg. Participatory visualization with Wordle. *IEEE Trans. on Visualization and Computer Graphics (Proc. of the IEEE Conf. on Information Visualization)*, 15(6):1137–1144, Nov./Dec. 2009.

[26] I. Viola, A. Kanitsar, and M. E. Groller. Importance-driven volume rendering. In *Proc. of the IEEE Conf. on Visualization*, pages 139–146. IEEE Computer Society, 2004.

[27] W. Willett, J. Heer, and M. Agrawala. Scented Widgets: Improving navigation cues with embedded visualizations. *IEEE Trans. on Visualization and Computer Graphics*, 13:1129–1136, 2007.

[28] J. Xu and C. S. Kaplan. Calligraphic packing. In *Proc. of Graphics Interface*, GI '07, pages 43–50. ACM, 2007.

[29] J. C. Yang, J. Hensley, H. Grün, and N. Thibieroz. Real-time concurrent linked list construction on the gpu. *Computer Graphics Forum*, 29:1297–1304, 2010.