Spatial Interaction: Exploring Interaction within the Spatial Metaphor

Alex Endert Chris North

Virginia Tech

ABSTRACT

Visual analytics emphasizes sensemaking through interactively exploring visualizations. In this paper, we discuss how instead of adjusting model parameters to change the view, *spatial interaction* allows users to explore information in the visual metaphor, while the system adjusts the parameters based on the user's analytical reasoning.

KEYWORDS: interaction, spatialization, sensemaking, analytics.

INDEX TERMS: H.5.2 [User Interfaces]: Interaction styles

1 INTRODUCTION

Visual analytics bases its success on combining the abilities of statistical models, visualization, and human intuition for users to gain insight into large, complex datasets [1]. This success often hinges on the ability for users to interact with the information, manipulating the visualization based on their domain expertise, interactively exploring possible connections and investigating hypotheses. It is through this interactive exploration that users are able to make sense of complex datasets, a process referred to as sensemaking [2]. However, directly interacting with and changing parameters of statistical models can be complicated, and is often outside of the expertise of the user (i.e., the analyst). How can we design visual analytic tools that embrace all three abilities (statistical models, visualization, and users) without separating the interactive visual exploration of the information from the adjustment of the underlying mathematical model? In other words, how can we allow users to stay *within the metaphor*?

Visualizations such as IN-SPIRE's "Galaxy View" (shown in Figure 1) present users with a spatial layout of textual information where similar documents are proximally close to one another [3]. An algorithm creates the layout by mapping the high-dimensional collection of text documents down to a two-dimensional view. In these spatializations, the spatial metaphor is one in which users can infer meaning of the documents based on their location. For instance, a cluster of documents represents a group of similar documents, and documents placed between two clusters implies those documents are connected to both clusters. The notion of distance between documents represents how similar the two documents are (i.e., more similar documents are placed closer together). These views are beneficial as they allow users to visually gain an overview of the information, such as what key themes or groups exist within the dataset. The complex statistical models that compute similarity between documents are based on the structure within the data, such as term or entity frequency. However, given the domain expertise of the user, inconsistencies in the spatial layout are often found with respect to the user's expert mental model (e.g., this document should not belong to this cluster, that document fits better over here, etc.). Therefore, users need to adjust the layout in order to test their hypotheses and gain insight into the data.

The interactions required to perform the change in the layout are typically accessible to users through various menus, sliders, and other visual controls designed to directly change parameters of the statistical algorithm responsible for generating the layout. These parameters include adding or removing keywords, changing the relative weight of specific keywords, and filtering the raw data upon import. However, analysts typically want to interact with such spatial information by interacting spatially through rearranging information, highlighting important terms or phrases, searching and receiving search results in context of the spatial layout, and annotating the documents as previously evidenced [4, 5]. In both studies, users were successfully able to solve intelligence analysis scenarios using these familiar, but manually performed, interactions. Therefore, the challenge of this work is how to enable users to augment the statistical models for generating a spatial layout through interacting directly in the spatial metaphor.

In this paper, we discuss the opportunities and challenges involved in providing users the ability to interact at an intermediate, semantic level – positioned between the formality of interacting directly with the model parameters, and the flexibility of directly manipulating the spatial locations of information in a layout. We refer to this design space for interaction as *spatial interaction*. Spatial interaction is grounded in the principles of how users are familiar with analyzing and exploring information spatially [4, 5]. The opportunity spatial interaction offers over manual spatial repositioning is a tight coupling between the visualization and the underlying mathematical model. Thus, users are able to interact within the spatial metaphor, abstracted from the complexities of directly augmenting parameters of the model. We have developed an application, ForceSPIRE, which serves as a testbed for spatial interaction (details of ForceSPIRE can be found in [6]).

This is achieved through *capturing* the spatial interaction, *interpreting* the analytical reasoning associated with the interaction, and *updating* the corresponding statistical parameters. Hence, users are able to leverage spatial interaction within the metaphor to explore and analyze the data interactively, while the algorithm is responsible for properly updating the complex combination of corresponding statistical parameters. Through spatial interaction, users can focus on their task of spatially organizing their information, while the system responds to provide and updated visualization. Thus, interaction takes on a deeper, more integrated role in the exploratory spatial analytic process. Essentially, users are able to input their domain knowledge by modifying the spatial layout, which in turn informs the layout models to respond and produce a better overall layout.

{aendert, north}@vt.edu

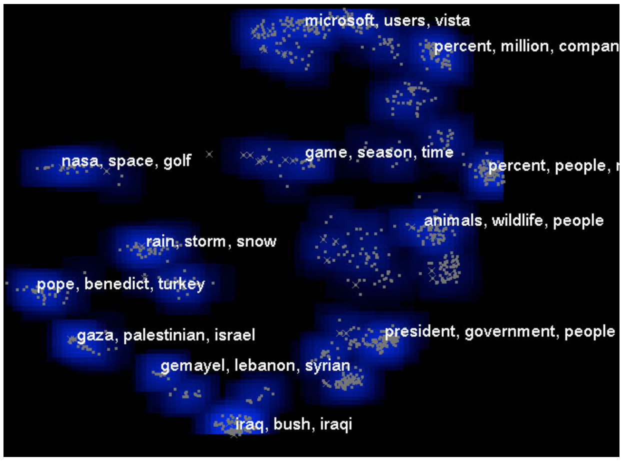

Figure 1. The IN-SPIRE Galaxy View showing a spatializtiation of documents represented as dots. Each cluster of dots represents a group of similar documents.

2 Related Work

The following previous work is relevant to the discussion of spatial interaction and the design of ForceSPIRE. Visual analytic tools exist that represent textual information spatially, such as the "Galaxy View" within IN-SPIRE [3], InfoSky [7], and others. In the Galaxy View, text documents are mapped to spatial locations based on the terms (entities) they contain. Thus, similar documents are clustered or grouped, and users can observe spatial relationships between documents (e.g., close documents are similar, distant documents are dissimilar, etc.). In order to interactively change the view, users are required to directly adjust keyword weights, add or remove documents/keywords, or provide more information on how to parse the documents upon import.

Similarly, an interactive visualization tool called iPCA uses Principal Component Analysis (PCA) to reduce high-dimensional data down to a two-dimensional plot, providing users with sliders and other visual controls for directly adjusting numerous parameters of the algorithm, such as individual eigenvalues, eigenvectors, and other components of PCA [8]. Through adjusting the parameters, the user can observe how the visualization changes. This allows users to gain insight into a dataset, given they have a thorough understanding of PCA, necessary to understand the implications behind the changes they are making to the model parameters.

Alsakran et al. presented a visualization system, STREAMIT, capable of spatially arranging text streams based on keyword similarity [9]. Again, users can interactively explore and adjust the spatial layout through directly changing the weight of keywords that they find important. In addition, STREAMIT allows for users to conduct a temporal investigation of how clusters change over time. While an interaction such as highlighting is available to users, the purpose is "to track the semantic evolution around the documents or keywords highlighted", not adjust the parameters of the statistical model in any way.

We believe this reveals a trend in how interaction is currently handled in many visual analytic systems where complex statistical models are used – users are required to *go outside of the metaphor*. That is, while the visual representation given to users is spatial, the methods of interaction require users to step outside of that metaphor and interact directly with the parameters of the statistical model using visual controls, toolbars, etc.

There has been some work in providing more easy to use interactions for updating statistical models. For example, relevance feedback has been used for content-based image retrieval, where users are able to move images towards or away from a single image in order to portray pair-wise similarity or dissimilarity [10]. From there, an image retrieval algorithm determines the features and dimensions shared between the images that the user has determined as being similar. We view this as one example where the interaction stays in the spatial metaphor of the visualization.

Additionally, spatializations of document sets exist that allow users to place "points of interest" into the spatial layout. In VIBE, users are allowed to define multiple points of interest in the spatial layout that correspond to a series of keywords describing a subject matter of interest to the user [11]. Similarly, Dust & Magnet allows users to place a series of "magnets" representing keywords into the space and observe how documents are attracted or repelled from the locations of these magnets. Through both of these systems, users can interact in the spatial metaphor through these placements of "nodes" representing keywords. However, spatial interaction focuses on interacting with data (i.e., documents), an important distinction discussed below.

From the sensemaking loop presented by Pirolli and Card [2], we learn that in intelligence analysis, that analytic process consists not only of the information that is explicitly within the dataset being analyzed, but also the domain knowledge of the analyst performing the analysis. It is through this domain knowledge that analysts interact and explore the dataset to "make sense" of the information. Thus, we believe this interaction (and the domain knowledge associated with it) is equally important as the raw data, and must be incorporated into the visualization by tightly coupling the model with the interaction.

Therefore, from this body of work, we most notably come away with an understanding that 1) analysts fundamentally understand the spatial metaphor used in many spatial visualizations, 2) many of these systems are constructed using complex mathematical algorithms to transform high-dimensional data to two dimensions, and 3) in most cases these algorithms can be controlled by analysts largely through visual controls (e.g., sliders, knobs, etc.) to directly adjust parameters of the algorithms, updating the spatial layout.

3 Spatial Interaction: Interactive Exploration Within the Metaphor

Given the previous work, spatial interaction occupies a new design space for interaction. It merges the ability to change the algorithmic parameters with the flexibility and familiar methods for interacting within the metaphor of spatial visualizations. Users can benefit from spatial interactions in that they can interact within a metaphor which they are familiar with, performing interactions which are part of the analytic process in such metaphors [4], without having to focus on formal updates to model parameters.

3.1 What *is* Spatial Interaction?

In the purest sense, a spatial interaction refers to an interaction occurring *within* a spatial visualization (such as the IN-SPIRE "Galaxy View"), with the added requirement that it is tightly coupled to the algorithm calculating the spatial layout. As such, the interaction functions *in the spatial metaphor*, and is performed directly on the data. Through tightly coupling the interaction with the model (or system) performing the algorithmic data transformation generating the visualization, it is the responsibility of the system to determine which parameters to update.

Spatial interaction differs from direct manipulation of parameters in that it is based on familiar interactions to data yet enable complex parameter updates. The concept is grounded in the ability for users to correctly interpret the meaning of information in spatial layout, and in making the interaction in which the layout can be changed more usable. The following intelligence analysis scenario is representative of the strategies and interactions of analysts when performing an intelligence analysis task of textual documents in a spatial visualization, as previously found by Andrews et al. [4], and further motivates and explain the concept of spatial interactions:

> *During her analysis, an intelligence analyst finds a suspicious and interesting phrase within a document. While reading through the document, she highlights the phrase "suspicious individuals were spotted at the airport", in order to more easily recall this information later. After she finishes reading the document, she moves the document into the bottom right corner of her workspace, in the proximity of other documents related to an event at an airport. Now, with the goal of further examining the events at the "airport", she searches for the term, continuing her investigation.*

This scenario showcases how traditional interactions that analysts are familiar with using, such as highlighting an interesting phrase, repositioning a document, and performing a search frequently occur within the analytic process when working within a spatial metaphor. That is, relative proximal distance between documents represents similarity between those documents. Thus, the movement of a document towards a region (or cluster) of other documents implies the analyst's reasoning behind the interaction is to extend the similarity between the documents in the cluster to the document being added to the cluster (i.e., this document is similar to those documents, therefore it is being moved closer). The other interactions (highlighting and searching) are also interactions present in the analytic process within such a visualization, as they are basic interactions present in sensemaking.

That these interactions are spatial make them more meaningful to the user based on the following. For instance, in highlighting the phrase, she is able to "mark up" the document and is able to re-find it more easily at a later time. The re-finding takes place through recalling the visual appearance of the document (with visual marks of highlights throughout the text) in relation to the spatial layout. In performing a search, the user is able to ask or query not only the dataset, but the space. The result of a search is therefore important not only because of the list of documents that contain the search phrase, but also where those resulting documents relative to the spatial layout.

Herein lies the opportunity for a new design space for interaction in spatializations of textual information. With spatial interaction, we leverage spatial abilities of human cognition, and provide a set of interactions that are tightly integrated into the statistical model. This marks a key difference from how users interact with information in previous tools (e.g., iPCA [8], STREAMIT [9], IN-SPIRE[3], etc). In each of these examples, users are given a spatial visualization of information, where an underlying mathematical model calculated similarity in a high-dimensional space, and ultimately outputs a two-dimensional view of the information, in which closeness is mapped to similarity. However, in order to interact with the information, direct adjustments are required to the mathematical model through various visual controls (e.g., slider, knobs, etc.). Hence, interactively exploring the data involves translating their insight from the spatial visualization, into changing formal statistical parameters. We see this step (or often series of steps) as a hindrance to the analytic process. One could postulate that since users are aware of patterns and relationships between information spatially, having to transform these insights from the spatial metaphor into parametric adjustments is ultimately pulling user out of their "cognitive zone" [12]. In contrast, users staying within this "cognitive zone" can be described as being cognitively engaged in their task, rather than focused on the tool. Therefore, we claim that through leveraging the spatial interactions in order to handle the parameter updates, users are able to stay within their spatial analytic process (i.e. their cognitive zone).

Table 1. Forms of spatial interaction. Each interaction corresponds to reasoning of users within the analytic process.

Form of Spatial Interaction	*Associated Analytic Reasoning*
Document Movement	• Similarity/Dissimilarity • Create spatial construct (.e.g timeline, list, story, etc) • Test hypothesis, see how document "fits" in region
Text Highlighting	• Mark importance of phrase (collection of entities) • Augment visual appearance of document for reference
Pinning Document to Location	• Give semantic meaning to space/layout
Annotation, "Sticky Note"	• Put semantic information in workspace, within context
Document Coloring	• Create visual group/cluster • Mark group membership
Level of Visual Detail	• Change ease of visually referencing information (e.g. full detail = more important = easy to reference)
Query Terms	• Expressive search for entity

In addition to the three instances of spatial interaction in the scenario, Table 1 provides a list of various forms of spatial interaction, including how each can be used within the analytic process of investigating textual information spatially. We do not claim that this list is complete, but instead point out that each of these interactions can relate to a user's reasoning within the analytic process. In general, the analytic reasoning can be categorized as either *entity extraction* or *connection making*. These two categories relate to the two general stages in the sensemaking loop [2] (foraging and sense-making) in that users are either collecting and identifying key entities or making connections between a collection of entities, respectively.

3.2 A Need for a New Model

In order for analysts to interact with information in a spatial metaphor, it must first be created. Following the model of the visualization pipeline [13], this creation calls for a series of mathematical transformations, turning raw data into a spatial layout – much the way many of the visualizations mentioned previously are constructed. The challenge then, is how to incorporate spatial interactions in such a model, where interacting with the algorithm must come through the visualization, not directly on algorithm parameters. A new model is needed.

Spatial interaction requires a fundamentally different model for how tools design and plan for user interaction – one that can *capture the interaction*, *interpret the associated analytical reasoning*, and *update the appropriate mathematical parameters*. We will discuss each of these stages in more detail below.

3.2.1 Capturing the Spatial Interaction

A non-trivial first step in the model is capturing the spatial interaction. Much research has been done in capturing user interaction, mainly for the purpose of maintaining a history of their process (e.g., [14], [15], [16], etc.). When considering how to capture interaction, one decision to be made is at what "level" to capture it. For example, GlassBox [17] captures interaction at a rudimentary level (i.e. mouse clicks and key strokes), while Graphical History [18] keeps track of a series of previous visualizations as a user changes the visualization during the exploration of the data.

Spatial interaction is captured at a *data level*. As the interactions occur on the data, and within the spatial metaphor, this is an appropriate level at which to capture the interaction. Using the earlier analytic scenario, the interaction being captured would be:

- The highlighted **phrase**
- When the highlighting occurs (**timestamp**)
- The **color** chosen for the highlight
- The **document** in which the highlight occurs
- The new document **location**

By capturing the interaction and treating it as data, we can interpret the interaction in order to associate the proper analytical reasoning of the user. Thus, we not only capture the interaction, but also *use it*.

3.2.2 Interpreting the Associated Analytical Reasoning

In interpreting the interaction, the goal is for the algorithm to determine the analytical reasoning associated with the interactions and update the proper parameters. From previous findings [4], we can associate analytical reasoning with spatial interactions (see Table 1). It is essentially the model's task to determine *why*, in terms of the data, the interaction occurred. To answer this question, we do not propose that this model can accurately gauge user intent. Instead, the goal is to calculate, based on the data, what information is consistent with the captured interaction. For instance, we associate text highlighting with adding importance to the text being highlighted. We do not claim that we can associate the interaction of highlighting to the intuition that spurred the analyst to highlight the text, which is far more challenging, and arguably impossible.

We refer to the captured and interpreted interactions as *soft data*, in comparison to the *hard data* that is extracted from the raw textual information (e.g., term or entity frequency, titles, document length, etc.). We define soft data as the stored result of user interaction as interpreted by the system. In treating interaction as soft data, the algorithm can calculate and reconfigure the spatial layout accordingly. For example, when a user highlights a phrase, the interpreted analytical reasoning is to increase the importance of the entities within that phrase, changing the spatial layout to reflect this added importance. The updating of the algorithm, and ultimately the spatial layout is described in section 3.2.3.

There has been previous work in capturing and interpreting reasoning from user interaction. For instance, Dou et al. [19] performed a study where financial analysts were asked analyze a dataset using WireVis, an interactive financial transaction visualization. The tool developers then analyzed the captured interaction, and assumptions were made about the reasoning of the analysts at specific points in the investigation. These results were compared to the analysts' self-recorded reasoning, and found to be accurate up to 82%. While our work has similar goals (i.e., interpreting the analytical reasoning associated with the analysts through an evaluation of the interaction) our model does so through tightly integrating the interaction with the underlying mathematical model. In doing so, the interpretation can be done algorithmically.

3.2.3 Updating Appropriate Mathematical Parameters

Finally, the soft data must be mathematically applied to the appropriate spatial layout parameters. In essence, the algorithm that generates the spatial layout must be adjusted according to the analytic reasoning associated with the spatial interaction. We perform this update using two methods.

First, using the soft data generated in the previous steps, the algorithm is designed to incorporate specific instances of this data into its calculation. For instance, in annotating a document through the addition of a sticky note, the text of the note is parsed into entities, appended to the entities of that document, thus allowing the algorithm to include those entities into the calculations associated with that document.

Second, algorithms often incorporate weighting of dimensions or attributes as part of how spatial layouts are calculated. We view these parameters as an opportunity to adjust the weights according to the soft data. For example, in highlighting a set of terms within a document, the model has the ability to increase the "weight" of those terms, ultimately increasing the degree to which those terms impact that spatial layout. Thus, highlighting "airport" would indicate to the model that "airport" is more important, and allow the dynamic spatial layout to adjust according to the user interaction, creating a layout where "airport" has a higher impact. The resulting layout has clusters form more tightly around documents with the term "airport", which may not have been the case prior to the interaction.

4 Discussion

4.1 Unifying the Sensemaking Loop

With the addition of soft data, and algorithms capable of interpreting the analytical reasoning, we leverage interactions that

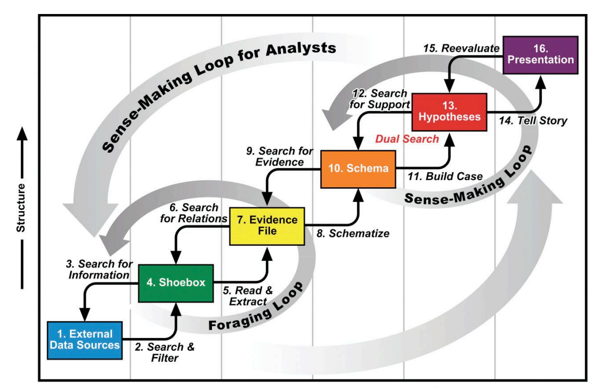

Figure 2. The sensemaking loop, illustrating the complex sequence of steps used by intelligence analysts in order to gain insight into data.

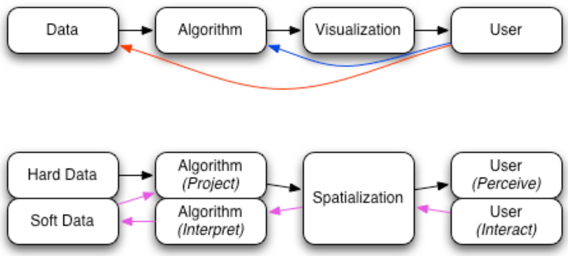

Figure 3. (top) The basic version of the "visualization pipeline". Interaction can be performed on directly the Algorithm (blue arrow) or the data (red arrow). (bottom) Our modified version of the pipeline, where the user interacts spatially, and within the metaphor (purple arrow).

are already occurring in the spatial analytic process to further aid users in their sensemaking process. The sensemaking loop (shown in Figure 2) illustrates how this process is primarily made up of "foraging" (gathering information from the data) and "sense-making" (building hypotheses about the data). Users proceed to forage for information to either confirm or deny a hypothesis, spawning another hypothesis, and so on. Between these two stages, users must formalize their hypothesis in order to forage for relevant information. Many existing tools focus either on foraging (e.g., IN-SPIRE [3], etc.) or sense-making (e.g., Analytst Notebook [20], etc.). ForceSPIRE can combine the flexibility of a sensemaking tool (i.e., where users organize their thoughts through creating a spatial representation) with the expressiveness of a foraging tool (i.e., where users can explicitly change parameters of a filter, etc. to quickly filter through massive amounts of data).

With spatial interaction, the amount of formalization between these two stages on the part of the user is reduced. For instance, in moving a document, users can formulate a hypothesis based on that document, expecting similar documents to follow. ForceSPIRE attempts to update the layout based on the interaction, and gives the user feedback. Thus, the foraging stage occurs as a result of the hypothesis being formed through spatial interaction. By not forcing users to over-formalize their analytic reasoning in order to forage for the relevant information, spatial interaction creates a more seamless transition between generating hypotheses and foraging, unifying the sensemaking loop.

4.2 Coupling Sensemaking and Interaction

In general, interaction in visual analytic tools functions as the medium through which communication between the user and the system occurs. Through interaction, users can communicate desired changes in the visualization based on a hypothesis or domain expertise in the data. While establishing a clear taxonomy of interaction for information visualization is challenging [21], we can see from a number of interactive visualization systems that it is through sliders and other visual controls of parameters that much of the interaction dealing with statistical models occurs. It is thus through interaction, that users can control and change algorithms, parse the data, and change the visualization according to those inputs.

The visualization pipeline models this role of interaction for interactive information visualization, showing how users can directly interact with various data filters, algorithm parameters, and visual mappings. At the basic level, the pipeline can be reduced to a user interacting directly with the model or the data. Figure 3 shows this basic pipeline (top) in comparison to the model used for spatial interaction (bottom). The model used with spatial interaction enables users to interact within the spatialization, makes use of modified algorithms capable of interpreting the analytical reasoning associated with the interaction, and creates soft data. The modified algorithm is also able to create a spatial projection based on both hard and soft data.

Through expanding the pipeline to accommodate for the concept of spatial interaction, it is a more appropriate match to the user's sensemaking process. The sensemaking process is based largely on connecting domain expertise with the data, through iteratively interacting with the data [2]. Thus, the spatialization acts as the medium through which the user can communicate (interact), and where the algorithm can respond (show new layout). With spatial interaction, the opportunity exists to more closely couple interaction with the statistical model and the data. As users are more familiar with spatial interaction techniques, as opposed to directly adjusting statistical parameters, we in turn narrow the gap between these fundamental aspects of visual analytics (visualization, algorithm, data, and domain expertise).

4.3 Complexities of Document Movement

Of the forms of spatial interaction we listed, we view document movement as one of the more complex interactions in regard to what it *means* to move a document. Is the user moving away from a group of documents, or towards another group of documents? Which group of documents are being targeted in the move? In ForceSPIRE, we used the most simplistic form of movement, in that it is "anchoring" the document in a spatial location, and the layout adjusts according the location.

There are more complex uses of movement, in that the analytic reasoning associated with the move is to portray similarity or dissimilarity. For instance, a pair-wise movement of two documents towards one another may imply that there are characteristics within the structure of these two documents that the user finds important. Therefore, it is up to the system to calculate what this important structure is, and adjust the layout accordingly. This particular case has been explored be Leman et al., who show that a modified version of the Multidimensional Scaling (MDS) algorithm is successful in determining the dimensions of similarity or dissimilarity when a select group of documents are moved closer or apart, respectively [22]. The algorithm responds by changing the weight, or importance value, on those dimensions to account for the analyst's reasoning.

4.4 Future Work

Spatial interaction, as a concept, opens up many possibilities for further research. In tightly coupling the interaction with the underlying model generating the spatialization, design decisions have to be made, such as: what information to capture from an interaction, which parameters of the model to update, and by how much to update those parameter. We plan to explore other variations of parameter updates based on forms of spatial interaction.

In order to make more concrete claims regarding the usability and effectiveness of spatial interacion, a formal user study is needed. Our plan is to introduce ForceSPIRE [6] to professional intelligence analysts and have them solve scenarios that model their daily task, such as one of the VAST datasets. The observations and feedback from these users will allow us to tweak the mapping between the analytical reasoning of the user and the parameter update (as well as what the relative weights are for these mappings).

5 CONCLUSION

In this paper we have discussed how the concept of spatial interaction leads to a new design space for interaction in spatializations of textual information. We describe spatial interaction, discussing the three components required in the design of a system that would incorporate forms of spatial interaction – capturing the interaction, interpreting the analytical reasoning, and updating the appropriate mathematical parameters. We discuss how spatial interaction has the opportunity to unify the sensemaking loop, creating a more seamless analytic process.

Spatial interaction has the ability to more closely match the sensemaking process of users, in that the interaction is performed onto the data and within the spatial metaphor (i.e., spatialization). Therefore, users are able to leverage the power of both visualization and statistical models, using interactions that have been shown to be familiar and common during a spatial analysis of text documents (e.g., highlighting, searching, moving of documents, adding annotations) [4]. In allowing users to interact within the spatial metaphor, they can remain more focused on their analysis of the data, without having to become experts in the underlying mathematical algorithms of the system.

ACKNOWLEDGEMENTS

This work is funded by the NSF, grant #CCF-0937071.

REFERENCES

[1] Thomas, J. J., Cook, K. A., National, V. and Analytics, C. *Illuminating the path*. IEEE Computer Society, City, 2005.

[2] Pirolli, P. and Card, S. Sensemaking Processes of Intelligence Analysts and Possible Leverage Points as Identified Though Cognitive Task Analysis *Proceedings of the 2005 International Conference on Intelligence Analysis, McLean, Virginia*2005), 6.

[3] Wise, J. A., Thomas, J. J., Pennock, K., Lantrip, D., Pottier, M., Schur, A. and Crow, V. *Visualizing the non-visual: spatial analysis and interaction with information for text documents*. Morgan Kaufmann Publishers Inc., City, 1999.

[4] Andrews, C., Endert, A. and North, C. Space to Think: Large, High-Resolution Displays for Sensemaking. In *Proceedings of the CHI* (2010), [insert City of Publication],[insert 2010 of Publication].

[5] Robinson, A. C. *Design for Synthesis in Geovisualization*. PhD thesis, Pennsylvannia State University, University Park, PA, 2008.

[6] Endert, A., Fiaux, P. and North, C. Interacting in the Metaphor: Spatial Interaction in Spatializations of Textual Datasets. In *Proceedings of the IEEE VAST* (Under Review), [insert City of Publication],[insert Under Review of Publication].

[7] Andrews, K., Kienreich, W., Sabol, V., Becker, J., Droschl, G., Kappe, F., Granitzer, M., Auer, P. and Tochtermann, K. The InfoSky visual explorer: exploiting hierarchical structure and document similarities. *Information Visualization*, 1, 3/4 2002), 166-181.

[8] Jeong, D. H., Ziemkiewicz, C., Fisher, B., Ribarsky, W. and Chang, R. iPCA: An Interactive System for PCA-based Visual Analytics. *Computer Graphics Forum*, 282009), 767-774.

[9] Alsakran, J., Chen, Y., Zhao, Y., Yang, J. and Luo, D. STREAMIT: Dynamic visualization and interactive exploration of text streams. In *Proceedings of the IEEE Pacific Visualization Symposium* (2011), [insert City of Publication],[insert 2011 of Publication].

[10] Torres, R. S., Silva, C. G., Medeiros, C. B. and Rocha, H. V. Visual structures for image browsing. In *Proceedings of the Proceedings of the twelfth international conference on Information and knowledge management* (New Orleans, LA, USA, 2003). ACM, [insert City of Publication],[insert 2003 of Publication].

[11] Olsen, K. A., Korfhage, R. R., Sochats, K. M., Spring, M. B. and Williams, J. G. Visualization of a document collection: the vibe system. *Inf. Process. Manage.*, 29, 1 1993), 69-81.

[12] Green, T. M., Ribarsky, W. and Fisher, B. Building and applying a human cognition model for visual analytics. *Information Visualization*, 8, 1 2009), 1-13.

[13] Heer, J. *prefuse manual*. City, 2006.

[14] Callahan, S. P., Freire, J., Santos, E., Scheidegger, C. E., Cl\, \#225, Silva, u. T. and Vo, H. T. VisTrails: visualization meets data management. In *Proceedings of the Proceedings of the 2006 ACM SIGMOD international conference on Management of data* (Chicago, IL, USA, 2006). ACM, [insert City of Publication],[insert 2006 of Publication].

[15] Shrinivasan, Y. B. and Wijk, J. J. v. Supporting the analytical reasoning process in information visualization. In *Proceedings of the Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (Florence, Italy, 2008). ACM, [insert City of Publication],[insert 2008 of Publication].

[16] Gotz, D. *Interactive Visual Synthesis of Analytic Knowledge*. City, 2006.

[17] Cowley, P., Haack, J., Littlefield, R. and Hampson, E. Glass box: capturing, archiving, and retrieving workstation activities. In *Proceedings of the Proceedings of the 3rd ACM workshop on Continuous archival and retrival of personal experences* (Santa Barbara, California, USA, 2006). ACM, [insert City of Publication],[insert 2006 of Publication].

[18] Heer, J., Mackinlay, J., Stolte, C. and Agrawala, M. Graphical Histories for Visualization: Supporting Analysis, Communication, and Evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 14, 6 2008), 1189-1196.

[19] Dou, W. *Recovering Reasoning Processes from User Interactions*. City, 2009.

[20] *i2 Analyst's Notebook*. City.

[21] Yi, J. S., Kang, Y. a., Stasko, J. and Jacko, J. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13, 6 2007), 1224-1231.

[22] Leman, S. C., House, L., Maiti, D., Endert, A. and North, C. *A Bi-directional Visualization Pipeline that Enables Visual to Parametric Interation (V2PI)*. NFS FODAVA Technical Report (FODAVA-10-41), 2011.