Extended thesis abstract, accepted by CHI '98 Doctoral Consortium Program and to appear in CHI '98 Summary

Automated Visual Discourse Synthesis: Coherence, Versatility, and Interactivity

Michelle X. Zhou
Department of Computer Science
Columbia University
500 West 120th St., 450 CS Building
New York, NY 10027
+1 212 939 7000
zhou@cs.columbia.edu

ABSTRACT

We summarize the motivation, the approach, and the significance of a thesis that is concerned with automating the design of visual discourse. To synthesize effective visual discourse, three design criteria are used: coherence, versatility, and interactivity. The thesis focuses on finding comprehensive and systematic methodologies that meet these criteria, and on developing practical ways to implement visual discourse synthesis systems.

Keywords: Automated visual presentation design, knowledge engineering, planning, interaction handling.

INTRODUCTION

Visual design automation is an emerging research field that has demonstrated some of its early promise in coping with large volumes of information in a timely, customized manner. Previous research in this area has focused on designing individual presentations (e.g., [3, 4]), or planning sequences of predominantly static presentations (e.g., [1, 5]). However, full attention has not been paid to designing *visual discourse.* We use the term *visual discourse* to refer to a series of connected displays, of which a single display or a series of static displays may be considered special cases.

This thesis focuses on investigating and establishing comprehensive and systematic methodologies for creating coherent, versatile, and interactive visual discourse. *Coherent* visual discourse has smooth transitions between displays, consistent designs within and among displays, and effective integration of various components. *Versatile* discourse is capable of presenting heterogeneous information (e.g., quantitative and qualitative) through different visual techniques (e.g., 2D charts and graphs or 3D animation sequences), using coordinated visual media (e.g., generated graphics, image and video). *Interactive* discourse presents visual affordances [2] that indicate that it can react to user interaction.

To demonstrate the feasibility of our approach, this thesis also describes how to implement a working visual discourse system. As part of the result, a general reference model for visual discourse generation systems will be provided, allowing different implementations to be compared or evaluated against the reference model. Our research involves four areas: *knowledge engineering*, *inference modeling*, *visual realization*, and *interaction handling*. Structural components resulted from the four areas are needed to assemble a generation system: a *knowledge base*, an *inference engine*, an *visual realizer*, and (optionally) an *interaction handler.* As an instantiation of the reference model, we have developed a working prototype called IMPROVISE (Illustrative Metaphor Production in Reactive Object-oriented VISual Environments) that automates visual discourse generation in two different application domains.

In the rest of this summary, we briefly describe the research has been conducted in each of the four areas in relation to visual discourse generation. We also highlight some of the results and indicate the remaining work.

KNOWLEDGE ENGINEERING

An automated visual discourse synthesis system generates visual representations based on information it knows about. How effective and efficient the generation would be depends upon what types of knowledge the system has and how efficiently the knowledge can be utilized. Thus, there are two tasks involved in knowledge engineering: identifying knowledge sources and determining representation mechanisms.

The task of identifying knowledge sources not only requires discovering needed information, but must also establish criteria and develop measurements to analyze and abstract the information comprehensively and systematically. In this area, research tasks include but are not limited to:

Data Characterization. To represent various information effectively, we are developing a comprehensive data characterization taxonomy to analyze presentation-related properties for heterogeneous information [8].

Visual Hierarchy. We are designing a visual hierarchy to systematically define various visual forms and their relations at different levels of abstraction, based on their syntax, semantics, and pragmatics.

Visual Lexicon. To capture visual representations of atomic objects and use them as building blocks in visual discourse synthesis, we are developing a compact but powerful visual lexicon representation to parametrize various visual forms for primitive domain objects [9].

Visual Design Principles. We are summarizing a set of *expressiveness* and *effectiveness* design rules for creating coherent, versatile and interactive visual discourse.

Situation Modeling. To tailor the discourse to different users and situations, we are modeling the relationship between the presentation situation (e.g., audience identity) and desired visual discourse properties and timing constraints of the design process.

While the task of identifying knowledge sources focuses on the content of the knowledge, the task of determining knowledge representation paradigms is concerned with the efficiency of knowledge utilization. Due to the diverse nature of the knowledge that we are dealing with, this thesis suggests a mixed representation formalism to facilitate knowledge acquiring, retrieving, and reasoning.

INFERENCE MODELING

Visual discourse design is an inferencing process. Our approach achieves design efficiency by modeling visual discourse as a set of temporally constrained visual actions. A *visual action* is an encoded graphic design technique (e.g., Move ?x). A hierarchical planning approach [7] would be particularly suitable for the design task. Our research adapts a generic hierarchical planning algorithm to this domain by adding visual design constraints based on coherence, versatility, and interactivity requirements [10]. As a result, the inference engine achieves computational efficiency by reducing the amount of search needed compared to search-based approaches. It also eases the task of knowledge encoding by reusing the actions that are common to many design tasks. In addition, we adopt a top-down hierarchical decomposition strategy to facilitate the interleaving of planning and execution and to ensure both local and global coherence within a visual discourse.

VISUAL REALIZATION

A complete graphics generation cycle includes planning a design and realizing it. In order to be realized, the design must be represented in a form that can either be understood by a rendering package directly, or be easily translated into the rendering language. It is nonetheless desirable to separate the process of design from the process of realization. In other words, both the design and realization processes should not be affected by the other's representation or functionality. In addition, the separation will contribute towards the desired system portability or parallelism. Bearing these criteria in mind, an *intermediate language*, which accommodates both high-level design specifications and low-level graphic primitives, is developed to interface the design and the realization processes.

INTERACTION HANDLING

At any given time, the user wishes to have certain control over what s/he is viewing. In our case, the user should be able to explore the information, or to stop or resume the ongoing visual discourse, at any point. Thus, an interactive but interruptible visual discourse should be produced.

While designing interactive visual presentations involves analyzing various interaction metaphors and styles to produce visual affordances, providing an interruptible visual discourse requires reactive planning [6]. In order to resume, the system must maintain a presentation execution history to record when the interruption occurred, what actions have been performed, and what actions are yet to be carried out. Furthermore, the system replans based on the result of the current user interaction.

CONCLUSIONS

This thesis aims to establish comprehensive and systematic methodologies for automating visual discourse design and approaches to assembling synthesis systems. It also tries to demonstrate the feasibility and generality of these approaches through the implementation of a working system called IMPROVISE. The completion of this thesis requires further research: using a more accurate and consistent model (e.g., a probabilistic model) for representing various design principles, developing a reactive planning algorithm that suits our needs in dealing with user interactions in the course of visual presentation, and conducting user studies to evaluate the effectiveness of the working system.

ACKNOWLEDGMENT

I would like to thank my advisor, Professor Steven Feiner, for all of his advice. This research is supported in part by DARPA Contract DAAL01-94-K-0119, the Columbia University Center for Advanced Technology in High Performance Computing and Communications in Healthcare (funded by the New York State Science and Technology Foundation), the Columbia Center for Telecommunications Research under NSF Grant ECD-88-11111, and ONR Contract N00014-94-1-0564.

REFERENCES

[1] S. Feiner. APEX: An experiment in the automated creation of pictorial explanations. *IEEE Computer Graphics and Applications*, 5(11):29–37, 1985.

[2] J. Gibson. Notes on direct perception and indirect apprehension. In E. Reed and R. Jones, editors, *Reasons for Realism: Selected Essays of James J. Gibson*, pages 289–293. Lawrence Erlbaum Associates, Hillsdale, NJ, 1977.

[3] J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Trans. on Graphics*, 5(2):110–141, 1986.

[4] S. F. Roth and J. Mattis. Automating the presentation of information. In *Proc. IEEE Conf. on AI Applications*, pages 90–97, 1991.

[5] D. Seligmann. *A Visual Language for Automated 3D Graphics Generation*. PhD thesis, Columbia University, New York, NY, 1993.

[6] D. Wilkins, K. Myers, J. Lowrance, and L. Wesley. Planning and reacting in uncertain and dynamic environments. *Journal of Experimental and Theoretical AI*, 6:197–227, 1994.

[7] R. Young, M. Pollack, and J. Moore. Decomposition and causality in partial-order planning. In *2nd Int. Conf. on AI Planning Systems: AIPS-94*, pages 188–193. Chicago, IL, June 1994.

[8] M. Zhou and S. Feiner. Data characterization for automatically visualizing heterogeneous information. In *Proc. IEEE InfoVis'96*, pages 13–20, San Francisco, CA, October 1996.

[9] M. Zhou and S. Feiner. The representation and use of a visual lexicon for automated graphics generation. In *Proc. IJCAI'97*, pages 1056–1062, Nagoya, Japan, August 1997.

[10] M. Zhou and S. Feiner. Top-down hierarchical planning of coherent visual discourse. In *Proc. IUI'97*, pages 129–136, Orlando, FL, January 1997.