Information Dense Summaries for Review of Patient Performance in Biofeedback Rehabilitation

Weiwei Xu   Hari Sundaram

Arts, Media and Engineering Program, Arizona State University, Tempe, Arizona 85251.

E-mail: { wwxu, hari.sundaram }@asu.edu

ABSTRACT

In this paper, we present a novel visual design for information dense summaries of patient data with applications in biofeedback rehabilitation. The problem is important in review of large medical datasets where the clinicians require that both summary and all the performance details be shown at the same time. There are two main ideas (a) Summarizing data along the conceptual facets (accuracy / flow / openness) and the temporal facets (session / set / trial) in the biofeedback therapy. The conceptual facets represent key information needed by the experts to review patient performance. (b) Effectively present the data trends and the details in context of the entire performance. The summary incorporates ideas from graphic design and reveals the performance data at two time scales. Our preliminary user study is promising with some significant results. They indicate that our visual summaries are useful, functional and promote inquiry.

Categories and Subject Descriptors

H.5.3 [**Group and Organization Interfaces**] *Computer-supported cooperative work*

General Terms

Algorithms, Experimentation, Human Factors

Keywords

Summarization, multimedia, stroke rehabilitation

1. INTRODUCTION

This paper focuses on designing visual summaries of patient performance in biofeedback rehabilitation. The biofeedback therapeutic process generates large spatial-temporal datasets. The fundamental challenge in summarizing these clinical datasets lies in the following observation – clinicians want *both* the summary (i.e. insight) and the details at the same time. Importantly, since we have been using our tool with actual stroke patients, the clinicians do *not* want any details to be automatically removed from the presentation. This problem bridges the reductionist summarization problem in multimedia [3] with information dense displays found in Tufte's work [4] in graphic design Hence we term this problem as *information dense summarization*.

There has been prior work in faceted dataset browsing [5] and clinical data visualization [2,4]. In [2], the authors visualize temporal relationships in clinical data using novel silhouette representations. Tufte ([4], pp.111) summarizes patient data using non-linear time scales. There are several limitations of prior work: (a) handling multiple time-scales (e.g. 10 sec. – 2 hours) in continuous data; (b) handling different clinical data types (e.g. spatial / temporal data) simultaneously.

MM'07, September 23–28, 2007, Augsburg, Bavaria, Germany.

There are two main ideas in our design of summarizing the patient performance at the set and the trial level: (a) Summarizing data along the conceptual facets (accuracy / flow / openness) and the temporal facets (session / set / trial) in the biofeedback therapy. We organize the conceptual facets vertically and the temporal facets hierarchically and horizontally. This organization reveals data trends within a conceptual facet and enables efficient data comparison across temporal facets. (b) Effectively present the data trends and the details in context of the entire performance dataset. At the set overview level we use aggregates of the spatio-temporal data with increasing transparency in time order. We reveal conceptual cross-facet correlation using a river metaphor. On the expanded set view, we use small multiples to allow for inter-trial comparisons. At the trial level, different trial segments (i.e. reaching, grasping, and returning) are colored with decreasing transparency, to reveal trial level segmentation. Details are revealed on demand through user interactions. Our preliminary user study is promising with some significant results. They indicate that our visual summaries are useful, functional and promote inquiry.

The rest of the paper is organized as follows. In the next section, we briefly review the biofeedback system. In section 3, we present the summarization challenges. In section 4, we discuss our design of the visual summarizations. We analyze user study results in section 5, and then present conclusions.

2. BIOFEEDBACK SYSTEM OVERVIEW

We briefly review the biofeedback system for stroke rehabilitation [1]. The system aims to help stroke patients reach for a target. It situates the patients in a multi-sensory engaging environment, where physical action (therapeutic task: reach for a cup) is closely coupled with media feedback. The system integrates five computational subsystems: (a) motion capture; (b) motion analysis; (c) audio feedback; (d) visual feedback; and (e) database archiving. Motion capture tracks the three-dimensional marker positions on the patient's arm. Motion analysis derives an expanded set of task specific quantitative features from the raw sensing data. The audio and visual subsystems adapt their feedback dynamically to selected motion features. The database archiving continuously stores the large-scale multimodal dataset. There are two important aspects to the therapy:

Conceptual facets: There are three conceptual facets (semantic messages) that the multimedia feedback environment must communicate to patients to encourage accomplishment of rehabilitation goals: *reaching accuracy*, *flow,* and *opening*.

Therapeutic Process: An approximately two-hour *session* is organized into *sets*, where a set has consistent physical setup and media feedback. A set comprises *trials*, where the patient performs the reaching task. Subsystem adaptations are made between sets based on domain expert understanding of patient performance.

3. SUMMARIZATION CHALLENGES

The problem in information dense summarization is to reveal key variable changes while preserving the details of the patient performance. This is challenging because of the sheer magnitude of captured (e.g. hand trajectory) and derived (i.e. performance analysis such as reaching error) time-series data. A typical patient session can last two hours, with the patient performing 80~90 therapeutic reaches. For each reach (lasting a few seconds), we gather three-dimensional patient performance variables, one-dimensional error variables and one-dimensional media synthesis variables. There are two main problems:

- *Faceted organization*: the summarization should reveal the conceptual organization of the biofeedback therapy (*accuracy, flow, openness*). It should also reveal the temporal organization (*session / set / trial*).
- *Efficient data ink*: The presentation needs to present summaries as well as detailed patient information. This needs to be accomplished without overwhelming the domain experts with information..

4. VISUAL SUMMARIZATION DESIGN

In this section we will discuss our design of faceted organization and efficient data-ink, in summarizing the patient performance.

4.1 Faceted Organization

We organize the biofeedback dataset into conceptual and temporal facets. *Facet*s are clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject. In our approach, conceptual facets are organized vertically. Temporal facets are organized *hierarchically* and horizontally in the correct time order of the therapy (Figure 1).

The conceptual facets in biofeedback rehabilitation are *accuracy, flow* and *openness*. These conceptual facets represent the key information that the experts need to review patient performance. For each conceptual facet, we present the raw captured data and the derived errors (Figure 1a). We now examine one conceptual facet (accuracy) in some detail. The explanations for the other two conceptual facets are similar.

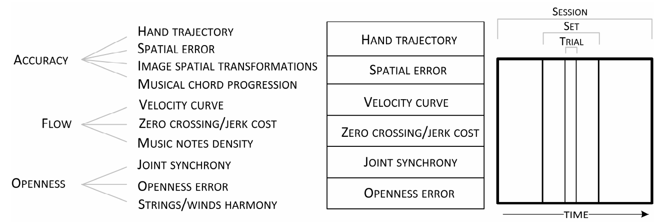

Figure 1: (a) Conceptual facets; (b) conceptual facets organization; (c) temporal facets organization

The biofeedback system aims at improving stroke patient's reaching *accuracy* (ref Section 2). The motion capture subsystem captures patient's hand trajectory that is the raw, detailed description of the patient reach. The motion analysis subsystem computes spatial error by examining the relationship of the trajectory to the target location. The visual subsystem maps the reaching *accuracy* into spatial transformation of images. The audio feedback subsystem maps it into music chord progression. Therefore, generalizing the reaching *accuracy* as a common conceptual facet embraces different subsystem perspectives, and can help the domain experts effectively understand the presented data and relate it to their role in the therapy.

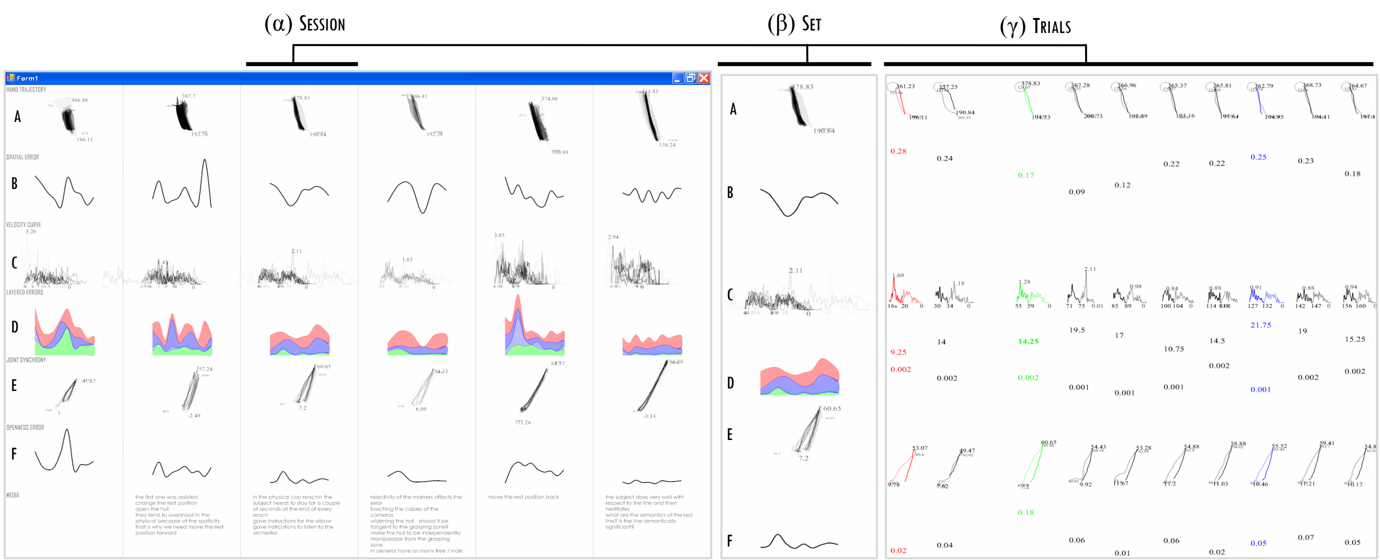

Figure 2 Screenshots of the visual summarizations that reveal key ideas in our design: *faceted organization and efficient data-ink.* In **faceted organization**, the conceptual sub-facets of patient performance – (A) hand trajectory; (B) spatial error; (C) velocity curve; (D) layered errors; (E) joint synchrony; (F) openness error – are organized vertically. The temporal facets – (α) session; (β) set; (γ) trials – are organized hierarchically and horizontally. (α) gives an overview of the patient session, comprising the collapsed overviews (β) of sets, in time order. Double-clicking any collapsed overview (β) will expand it into a detailed view (γ) of trials within the set. In our **visual design**, we reveal details differently in set and trial level summarizations. On the set-level collapsed overview (β), trials are colored with increasing transparency in their time order within the set. Aggregation of the trials reveals spatial and temporal stylization of the patient performance, e.g. hand trajectory (A) converges into a straight line. The subtle, dynamic correlations of the derived errors (D) are revealed in an unsaturated river of errors. in the set-level expanded detailed view (γ), horizontal lengths are proportional to the actual rehab time, e.g. trial duration and break can be observed to converge in the last four trials on (γ)-(C). At the trial level, different trial segments (reaching, grasping and returning) are colored with decreasing transparency. Trials with critical derived errors are highlighted using primary RGB colors across all conceptual facets. We have also used numbers directly to mark data points for error (γ)-(D).

The temporal facets are *session, set and trial* (Figure 1-(c)). They reveal the procedural organization of the biofeedback therapy, at different levels of temporal granularity. Each level requires a different data summarization / presentation design, which will be further discussed in the next subsection.

4.2 Efficient Data-Ink

The main idea in our visual summaries is to reveal data trends in the patient performance *in context of the entire performance dataset.* It is an important consideration for summarizing and presenting the dataset in our clinical, experimental and exploratory setting. It is very difficult to develop a utility based summarization measure (e.g. [1]), as it is critically dependent on the domain experts (doctors, physical therapists, musicians, etc.). It is not sufficient to create a "summary curve", as this does not capture the highly detailed variations in the patient data. It is these details that are of interest to the experts while exploring the dataset. They are keen to see the entire dataset, at high detail and then draw conclusions on the patient performance. Indeed since we work with actual stroke patients it is critical not to make therapeutic mistakes. Hence in our current design approach, we have adopted the "human in a loop" approach. We include the domain experts as integral to the process of determining insights from the data and focus on developing effective data presentation techniques that help reveal details and data trends.

4.2.1 Displaying details is hard

Our visual design is strongly motivated by Tufte's principles of designing statistical graphics on print media [4]. Tufte's work deals with effective information dense displays for print media. He focuses on maximizing information in print via formalized graphic design principles to allow for effective communication. There are three differences with our work – display capacity, interactivity, and time for review. The first difference is the display capacity of print media (1200dpi) and LCD displays (72dpi). Print media can also reveal much higher contrast, allowing for subtlety, than do LCD displays. To overcome the low information density on electronic displays, we must take advantage of their interactivity, to reveal details on demand. Implicit in Tufte's work is the significant time taken to understand the information dense display – in our case, the clinician time for data review is limited.

4.2.2 Efficient data-ink to reveal details

We organize the conceptual facets vertically (Figure 2.(a)-(f)) and the temporal facets hierarchically and horizontally (Figure 2.(α),β),(γ)). This organization reveals data trends within a conceptual facet and enables efficient data comparison across temporal facets. Each temporal facet requires a different data summarization / presentation design. In this paper, we focus on designing *set* and *trial* level summarizations.

At the set level, we have designed both a collapsed overview and an expanded detailed view of the set. On the collapsed overview, trials are aggregated with increasing transparencies in time order (Figure 2.(β)-(a),(c),(e)). The aggregated contours of hand trajectory / velocity curve / joint synchrony immediately summarize the temporal progression (i.e. spatial and temporal stylization) of the patient performance within a set. We reveal the subtle, dynamic correlations of the derived errors in the river of errors, allowing for comparison of errors over the set. (Figure 2.(β)-(d)). On the expanded detailed view, we use small multiples to allow for inter-trial comparisons (Figure 2.(γ)-(a),(c),(e)). Horizontal lengths on the expanded view are always proportional to the actual time (i.e. trial duration and break), revealing the temporal details of the therapeutic process, and highlighting any anomalies within a set. Within a set show errors as numeric graphs – numbers themselves are used to mark data points, to help reviewers focus on the data ((Figure 2.(γ)-(b),(d),(f)).

At the trial level, different trial segments (i.e. reaching, grasping, and returning) are colored with decreasing transparency, to reveal trial level segmentation (Figure 2.(γ)-(a),(c),(e)). We consistently use primary RGB colors to highlight trials with critical errors, and show the error across all conceptual facets (Figure 2.(γ)). Details are revealed on demand through user interactions. Double-clicking the collapsed overview of a set expands it into an expanded detailed view (*semantic zooming*). The application allows zooming via mouse movement + keyboard – reaching hulls are revealed above a certain zoom level. Our application is realized using Piccolo.NET. The application loads the biofeedback dataset from a SQL Server database.

5. EVALUATION

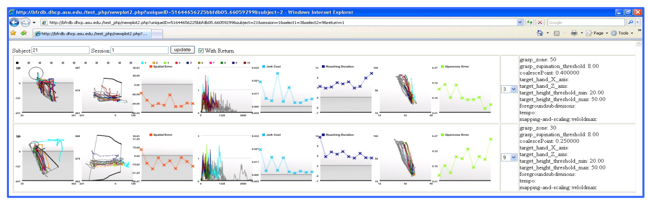

Figure 3 The web-based baseline interface

In this section we present a preliminary user study of our proposed application. We are at present working on a longitudinal evaluation of the application, which analyzes the detailed interactions of the domain experts with the application.

We compared two interfaces, with two groups of users. The baseline interface was web-based (ref. Figure 3) that merely plotted the data. The second interface was the application that is the study of this paper. The web interface was extensively used by the biofeedback domain experts during 18 real-time patient sessions and 13 post data reviews. However, its scalability to the large-scale biofeedback dataset is poor – it can only load two patient sets simultaneously, and the experts have expressed their frustrations when data loading exceeded one minute. It does *not* support advanced data presentation techniques, e.g. flexible data-ink placement, hence precludes effective visual design. It is also non-interactive. Our new application focuses on improving the usability, functionality and visual design. It can load the entire patient session data at once. Faceted organization and efficient data-ink are realized.

The first user group comprised three members of the biofeedback group (ref Section 2), who were highly familiar with the therapeutic process as they were participants in those sessions. The second group comprised graduate students unfamiliar with the project. The presence of the biofeedback team members is important to this study as the interface is domain specific, and most valuable to researchers to that domain.

The one-hour user study was conducted as follows. Users of the external group were first briefly introduced to the project through sample videos and verbal scripts. All users were asked to review a patient session with both the web interface and the new application. They were asked to compare the two interfaces, in aspects of usability, functionality and visual design on a 7-scale questionnaire. Questions with their mean score ± standard deviation are shown in Table 1. Higher scores indicate higher evaluation of the new application, compared to the web interface.

Table 1 Quantitative results of the user study

The tool compared to the baseline interface	Score1 biofeedback	Score2 external	Score3 all users
Ease of learning to use	5.67	6.00	5.83 ± 0.98 / 7
Ease of interface navigation	6.33	6.67	6.50 ± 0.55 / 7
Response to user interaction	6.67	6.33	6.50 ± 0.55 / 7
data exploration at different time scales	5.67	5.67	5.67 ± 1.21 / 7
Clear procedural organization	5.50	5.83	5.67 ± 1.27 / 7
Comparison between sets / trials	6.33	6.33	6.33 ± 0.52 / 7
Correlation across variables (facets)	6.00	5.67	5.83 ± 0.98 / 7
Help discover interesting variation over sets	6.00	5.67	5.83 ± 1.17/ 7
Help discover important variation over sets	6.00	6.00	6.00 ± 1.00 / 7
Help explore the patient session data	6.67	6.33	6.50 ± 0.55 / 7
Help compare errors through the river view	5.34	6.50	5.91 ± 1.32 / 7
Removing coordinates help focus on the data	3.67	5.67	4.67 ± 1.63 / 7
Use numbers to mark data points help focus on the data	6.00	5.00	5.50 ± 1.52 / 7
Text and graphics integration	4.33	6.00	5.17 ± 1.47 / 7
Consistent and informative use of color	5.50	6.17	5.84 ± 0.73 / 7
Expanded details in context of collapsed overview is needed	6.00	5.67	5.83 ± 1.47 / 7
Overall information organization and display	6.67	6.67	6.67 ± 0.52 / 7

The quantitative analysis reveals a statistically significant difference (t-test, $p=0.03$) in the means between the biofeedback users ($\mu=5.60$) and the external users ($\mu=6.03$). This test is over the entire set of questions for the two groups as a whole. This difference suggests that experiences with the system for the biofeedback group lead them to form different opinions than the group unfamiliar with the rehab process. The biofeedback users had stronger opinions on how the data should be summarized and presented, than the external users. We note that we do not have enough data to measure statistical significance across the two groups for a *specific* question.

For all users, the new application better supports exploration of the entire session data (6.50 ± 0.55 / 7) and comparison of trials / sets (6.33 ± 0.52 / 7). Note that it can present the entire session at once, which facilitates data comparison and trend discovery. The application has better information organization / display (6.67 ± 0.52 / 7), and is easier to use and navigate (6.50 ± 0.55 / 7).

The qualitative feedback from the users provides rich details not captured by the questions. In aspect of **usability (**top section, Table 1). all users agreed that the tool "has faster access (to data)" and "is much easier to use". The biofeedback users were more careful in using the new application and pointed out that "zooming is great but is a little hard to control", "if a column A is expanded and you fold it up by double-clicking on its very right, the cursor will fall on another column B after the shrink of A". In aspect of **functionality** (middle section, Table 1) all users preferred the new application, for "I can have a whole view of the dataset and catch more meanings of the data", "I can look at all sets immediately. Then if I'm interested in a single set, I just click on it and get the details", "gives more time information. It gives three obvious colors – red, green and blue, telling me about three kinds of maximum error", and "it is very convenient to tell which error affects most (through the river view)". However only the biofeedback user made a definite judgment that "patient get improvement in the sense of spatial error and smoothness; patient has consistent reaching at the end of experiment", and also pointed out that "it is not easy to compare detail of the first set and the last set". In aspect of **visual design**, the feedback varied the most (bottom section, Table 1). From biofeedback users, one expressed that "colors are used a bit more effective in the baseline interface, I like the idea of using transparency but perhaps color can help as well"; another questioned that "we cannot compare high level info (on the overview) with low level info (on the detailed view) directly". From external users, one expressed that "the error river looks pretty and clear, also makes the whole process interesting"; another was that "it's better to add a curve to connect the numbers to be better understood". Overall, the new interfaces were well liked by both groups, but familiarization with the tool takes time. It points to the necessity of a longitudinal study with the experts to understand tool use.

6. CONCLUSIONS

In this paper we presented an information dense summary of patient performance in biofeedback rehabilitation. This bridges reductionist approaches in multimedia with information dense approaches in graphic design. There are two main ideas in our design – (a) the conceptual / temporal facets organization; (b) efficient data-ink to reveal details and data trends. Our preliminary user study indicates that the visual summaries are useful, functional and promote inquiry. We are at present working on a longitudinal evaluation, to improve our design.

7. REFERENCES

[1] Y. CHEN, H. HUANG, W. XU*, et al.* (2006). *The design of a real-time, multimodal biofeedback system for stroke patient rehabilitation*, Proc. of ACM multimedia 2006, 763-772, Santa Barbara, CA, USA.

[2] C. PLAISANT, B. MILASH, A. ROSE*, et al.* (1996). *LifeLines: visualizing personal histories*, Proc. of the SIGCHI conference on Human factors in computing systems: common ground, 221-ff., Vancouver, British Columbia, Canada.

[3] H. SUNDARAM, L. XIE and S.-F. CHANG (2002). *A utility framework for the automatic generation of audio-visual skims*, ACM Multimedia 2002, ACM Press, 189-198, Dec. 2002, Juan-les-Pins, France.

[4] E. R. TUFTE (2003). Visual explanations : images and quantities, evidence and narrative. Graphics Press Cheshire, Connecticut.

[5] K.-P. YEE, K. SWEARINGEN, K. LI*, et al.* (2003). *Faceted metadata for image search and browsing*, Proc. SIGCHI conference on Human factors in computing systems, 401-408, Ft. Lauderdale, Florida, USA.