MULTI-DIMENSIONAL VISUALIZATION OF PROJECT CONTROL DATA

Anthony D. Songer[1], Benjamin Hays[2], and Christopher North[3]

ABSTRACT: The construction industry produces voluminous quantitative data. Much of this data is created during the controls phase of projects and relates to cost, schedule, and administrative information. Recent storage and processing advances in computers as well as display capabilities afforded by computer graphics increase the opportunity to monitor projects fundamentally different than existing project control systems. However, changes in project control methods have been slow to evolve. The lack of a fundamental model of project control data representation contributes to the inadequate application and implementation of visual tools in project control methods.

Difficulties associated with the graphical representation of data can be traced to the diversity of skills required in creating visual information displays. Due to the reality that not all engineers/constructors possess these attributes in great strength, streamlining the process of how to best visualize data is important.

Visual representations of data hold great potential for reducing communication difficulties fostered by industry fragmentation. However, without information structure, organization, and visual explanations, the massive amount of data available to project managers results in information overload. Therefore, improved information displays are needed to overcome the possibility of information overload with the capability of human perception.

This paper discusses research to create a framework for visual representation of construction project data. Underlying visualization theory, the visual framework, and a detailed implementation are provided.

KEYWORDS: Construction, Construction Management, Project Controls, Project Management Visualization.

INTRODUCTION

Computing tools are available to greatly enhance collaborative decision-making and project performance among construction project participants. However, current use of computing tools among practioners does not capitalize on the available capacity (Kunz et.al., 2001). Causes of this mis-alignment include lack of training and understanding (Griffis and Struts 2000, Russel 1997). Additionally, simple automation of existing processes fails to provide breakthrough advances in construction performance.

These issues are very prevalent in the area of construction project controls. Principles of data graphics offer visual capabilities well beyond those currently employed by the construction industry to display appropriate data in a manner that enhances comprehension of control processes. Prior research in the area of visualization demonstrates that within the construction project control domain, visual representation of data is often oversimplified or data is too detailed to provide adequate visual representation (Subramanian et .al, 2000). Within the domain of construction project control, graphical representation for construction control processes such as scheduling, budgeting, and change management

[1] Associate Professor, Department of Civil Engineering, Virginia Polytechnic Institute and State University, 200 Patton Hall, Blacksburg, VA 24060. email: adsonger@vt.edu
[2] Project Manager, Los Angeles Water District, Los Angeles CA.
[3] Assistant Professor, Department of Computer Science, Virginia Polytechnic Institute and State University, 660 McBryde Hall, Blacksburg, VA 24060.

follows no predefined method. Many graphics neglect relevant information necessary to discover trends or relationships between processes.

The underlying problems with the prevailing practice are the disparity between visualization capacity available and used, and the lack of an established framework to investigate visual representation of data.

Resolving the disparity between data visualization needs and data visualization reality within the construction industry provides a myriad of benefits to both researchers and practitioners. These include:

- Reducing ambiguity with which graphics are sometimes created.
- Elucidating which visual principles should be employed to describe particular data types.
- Making the process of creating visual representations simpler.
- Providing managers information rich overviews regarding the status of various pertinent aspects of projects.
- Decreasing the time spent explaining and understanding various aspects of the project's status, particularly in review meetings.
- Increasing the speed at which day to day decisions can be made.
- Helping create proactive management environments.
- Reducing stress associated with managing large numbers of paper documents.
- Improving the accuracy and completeness with which users achieve data specific objectives.
- Amplifying cognition of quantitative data.

Need-based practices within construction project control promote reactive development of visual tools. For example, responding to a particular question about project performance rather than developing fundamental visualization theory creates prevailing project control data graphic representations. E.G. ‘how is the budget?’, ‘what is the RFI turn around?’ among many others. These query-based representations only provide answers to the specific queries. While user-scenarios are important to visual strategy development, strategies must also consider fundamental principles in visualization theory, data type and density. Data type-based representations create improved understanding of the data [Bertin 1983, Card 1999, Shneiderman 1996, Spence 2001]. Additionally, current data graphics are typically limited to representing 1 dimensional or temporal plus one dimensional data types. These graphical layouts lack information density required to provide relationships and perceptual content required for effective decision making. A formal, theory-based framework for investigating and creating project control data graphics is needed.

Therefore, the authors set out to investigate underlying visualization theory, develop a visual framework to apply when considering possible project control data representation, implement the framework for a specific project control process and visual representation, and observe reactions to the new data portrayals.

This paper describes a formal framework for developing visual strategies and establishes a taxonomy for visualizing construction project control data. The framework is illustrated for one data rich, information poor construction problem. Charrette testing validates the strategy and subsequently the framework.

METHOD: A FORMAL FRAMEWORK USING VISUALIZATION THEORY

Using underlying visualization theory, the authors developed a formal framework for developing visual strategies for construction project controls. Visualization theory stresses several principles when creating data graphics. These include structuring and filtering, editing for honesty and density, and communicating efficiently. The visual framework shown in figure 1 illustrates the iterative process of implementing each of these principles when creating data graphics for project control data.

The first step in the framework is structuring data. A primary facet of visual theory, structure must be guided by data type (Bertin 1983, Shneiderman 1996, Card 1999, Spence 2001). Data types include uni-, bi-, tri-, and multidimensional, network, temporal, and hierarchical (Shneiderman 1996). Two or more data types often categorize construction project control processes. Current data graphics in construction often only represent one data type. Complex data sets require multi-data type consideration for adequate representation. For example, scheduling data is characterized as having at least three types of data, temporal, hierarchical and multi-dimensional.

Analyzing the necessary level of detail required for a data set is accomplished through filtering. Investigating user scenario requirements establishes applicable levels of detail. This step filters and specifies the structure, resulting in an initial data graphic. User scenarios often suggest an overview, zoom, filter, detail on demand capacity for data graphics (Shneiderman 1996).

Over-filtering often results in one-dimensional visual structures. Instead of beginning with simple, traditional questions, information displays must seek to present as much data as possible prior to filtering. When related though possibly non-traditional information is displayed, unexpected findings and the development of new questions may result.

Editing is an ongoing, iterative process in the framework. Primary considerations for editing include insuring data graphics accurately represent the data set (honesty) and monitoring graphics for adequate density of information. More information in the same space (data density) is typically desired (Tufte 1983).

The framework includes considering the perceptual component of visualization. This insures efficient communication of data. For example, when data is represented with graphical efficiency comparisons become simple. Visual efficiency is inversely related to the number of images needed for the perception of a data set (Bertin 1983).

Throughout the process, iterative evaluations provide feedback. Depending on the stage of development, evaluations range from an informal user survey to rigorous statistical analysis.

As noted, the formal graphic framework is centered on data-types. Data types and associated graphic strategies are represented in Figure 2. Data types include 1-dimensional, 2-dimensional, 3-dimensional, multi-dimensional, temporal, network, and hierarchical. Figure 3 provides a representative depiction of current industry uses of visualization. It illustrates taxonomy of current graphic representations of typical construction project control processes, the associated data-type represented by the graphic and the data-type requirements for each process.

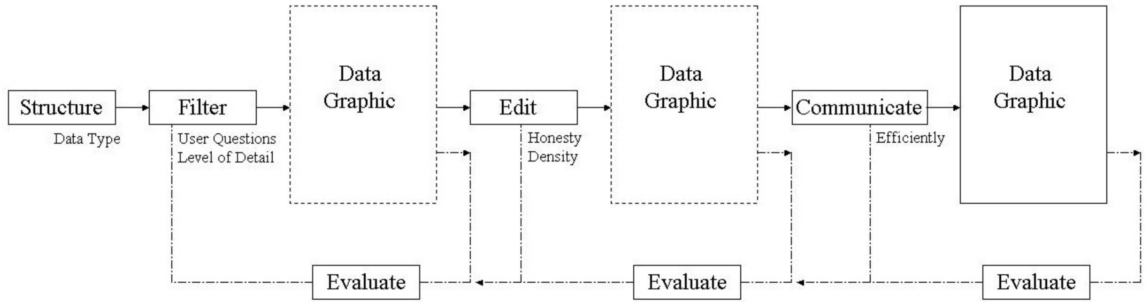

Figure 1. Visual Development Framework

This taxonomy was developed through a literature review and validated by interviewing 10 general contracting organizations on their use of graphics for project control. Construction control processes are listed horizontally. As noted, data type drives the visual layout of a graphic. Therefore, critique of current data graphics and analysis of controls processes must begin with discussion of this feature. Consequently, data types are listed by increasing complexity on the vertical axis. The vertical location of each outlined box corresponds to the one or more data types related to each specific process. For instance, the scheduling box overlaps multidimensional-temporal, network and hierarchical data types.

A number of existing graphics focus on one dimensional or one-dimensional-temporal overviews (budgeting, RFI, submittal, change order and resource processes). Existing graphics of many processes lack data types they intend to represent; notably the budgeting, submittal, change order and equipment processes. Where existing graphics overlap the data types present in the process (i.e. overlap the outlined box), the current strategy employs an adequate layout but may be improved in terms of its level of detail, efficiency or density. The gaps between existing graphical representations and process data types illustrate existing development potential. For example, figure 3 illustrates that budgeting representations currently do not adequately address the hierarchical or multi-dimensional plus temporal data types.

Using the budget/cost project data for a mid-sized general contractor this paper illustrates the each step of the visual framework, data structure, filter, edit, and communicate, to develop new visual strategies for budgeting/cost.

DATA STRUCTURE: A MULTI-DIMENSIONAL VIEW OF COST DATA

Within the construction industry, the project budgeting/cost control issue represents one of many *data rich, information poor* problems. Numerically rich construction problems exist, but value adding, visually communicated information is lacking. Due to the complexity of budget data and industry wide differences in accounting standards, no general graphical strategy exists. Using an existing data set for an 18.8 million dollar highway project (Figure 4), the authors implemented the visual framework to investigate four visualization strategies for cost/budget. The cost report

Data Type	Visual Strategies
1-Dimensional	Histograms, Piecharts, Tukey Box Plots
2-Dimensional	Scatterplots, Linked Histograms
3-Dimensional	*Physical Data*: Three Orthogonal Axes
	Relational Data: See Multi-D below
Multi-Dimensional	*Glyphs and Cleveland's Rules*: Position, Length, Direction, Area, Angle, Volume, Curvature, Shading, Color Saturation
	Statistical Methods: Perturbation / Scatterplot Matrix, Mosaic Glyphs, Small Multiples and Multidimensional Icons
	Visual Separation: Micro/Macro Strategies, Color, Layering
Temporal	*1D + Time*: Time Series
	Multi-D + Time: Small Multiple, Parallel Axis Time Series
Hierarchical	Tree (Orthogonal or Circular), Treemap
Network	Simple, Circular, Three Dimensional, Aggregated

Figure 2. Visual strategies by data type

in Figure 4 was the only 'graphic' used to monitor the project budget/cost by the participant organization. Figure 4 illustrates the data rich, information poor problem typical among construction industry organizations. The spreadsheet format represents multi-dimensional, hierarchical, and temporal data types. The column headings (cost code, description, unit of measure, type, budget quantity, budget dollars, job to date quantity, and job to date dollars) represent the multi-dimension data types. The hierarchical data types are identified as 60 pay items and 1040 work items in the cost report. The temporal data type is represented by the month. The multi-dimensional, hierarchical, and temporal data types illustrated in the case example are discussed below.

Multi-dimensional Data Type

The horizontal layout of the dataset reveals the cost report's multidimensional features. Corresponding to each cost code entry are the dimensions of cost code number, description, work area (equipment, labor, subcontractor), the unit of measure as well as estimated and job to date (JTD) measures.

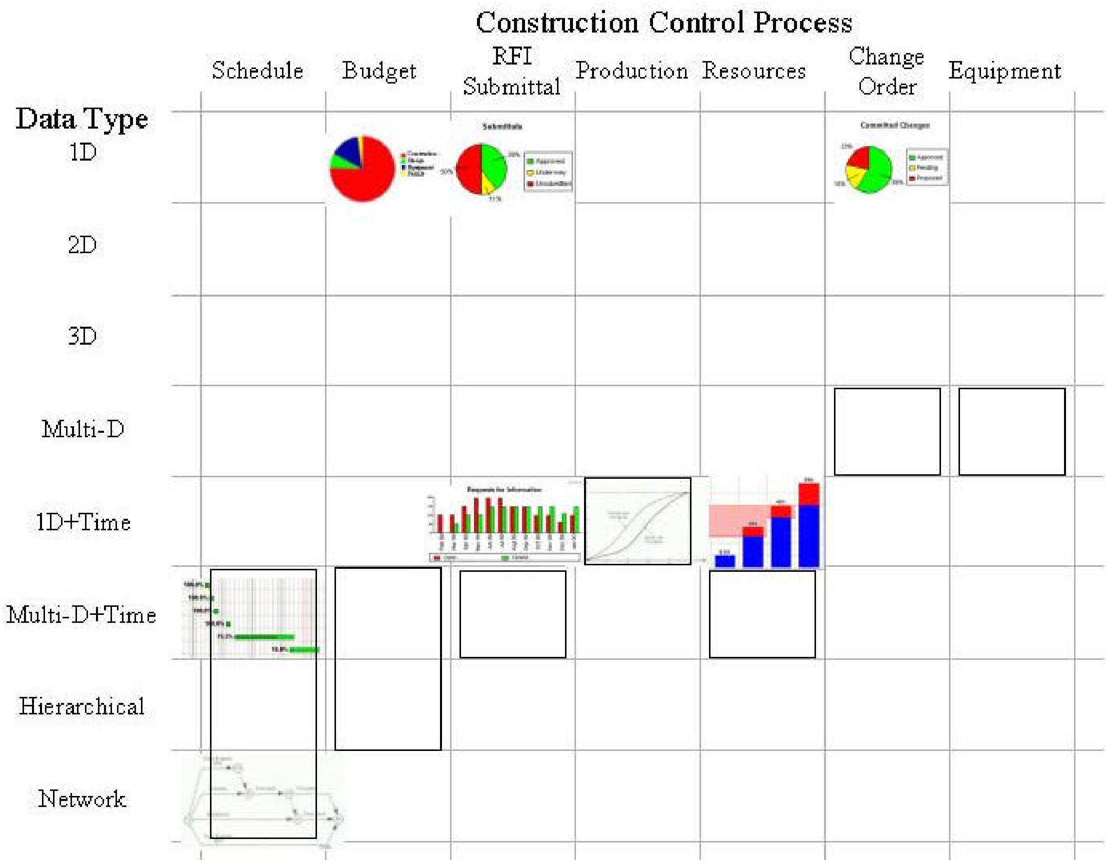

Figure 3. Graphical state of construction project controls: a taxonomy

CO#	Cost Code	Cost Code Description	UOM	Type	Bdgt Qty	Bdgt $$	JTD Qty	JTD $$
20	1020200123000	EARTH ADT 0-3000 FT	M3	E	155000	205031	131177.7	285809
20	1020200123000	EARTH ADT 0-3000 FT	M3	L	155000	81920	131177.7	146781.3
20	1020200124000	EARTH ADT 3-6000 FT	M3	E	312000	575531	32247.28	73538.49
20	1020200124000	EARTH ADT 3-6000 FT	M3	L	312000	208204	32247.28	33495.03
20	1020200125000	EARTH ADT +6000 FT	M3	E	180000	360002	62520	135466.2
20	1020200125000	EARTH ADT +6000 FT	M3	L	180000	128649	62520	60144.95
20	1020200160000	EARTH - TANDEMS	M3	E	0	0	59290	85017.44
20	1020200160000	EARTH - TANDEMS	M3	L	0	0	59290	60431.54
20	1020200169000	INACTIVE	M3	E	0	0	0	0
20	1020200169000	INACTIVE	M3	L	0	0	0	0
20	1020200523000	ROCK ADT 0-3000 FT	M3	E	104000	216668	156080	288950.3
20	1020200523000	ROCK ADT 0-3000 FT	M3	L	104000	81327	156080	129299.3
20	1020200524000	ROCK ADT 3-6000 FT	M3	E	208000	580422	250164.8	441461.4
20	1020200524000	ROCK ADT 3-6000 FT	M3	L	208000	202089	250164.8	202648
20	1020200525000	ROCK ADT +6000 FT	M3	E	120000	359715	167539	351673.8
20	1020200525000	ROCK ADT +6000FT	M3	L	120000	114355	167539	152520.3

Figure 4. Monthly Cost Report Data

Thus, in addition to the hierarchy, the cost report consists of eight separate dimensions. A number of dimensions directly correspond to the designation of each cost code. For example, the code description ***Earth ADT + 6000*** falls under the parent pay item ***Mass Earthwork*** , carries the code number *125000*, work area *E* (for equipment) and unit *M3* (cubic meters). Each cost code description carries a unique set of values for these four dimensions. The other four dimensions, two each for the estimated and JTD measures, are described below.

Estimated dollar value and quantity are not directly related to each cost code as the other dimensions above. When determining quantities and costs for a job, the construction company's estimating department accesses internal accounting cost codes. As the scope of work is understood and communicated through contract drawings and specifications estimators have liberty to open the cost codes necessary to the job. Occasionally, additional cost codes are opened though no money is initially allocated to them. This occurs when a job may need a specific item, but the certainty of this need is variable. For example, ambiguity regarding the necessity of an extra crane during latter stages of construction may cause an estimator to open an extra crane cost code but budget $0 therein. Similarly, during the course of construction, a project manager may open a cost code when new construction methods call for unbudgeted work items. The estimated values for this cost code will show up as $0. These features affect the calculation and graphical representation of dimensions needed to answer user tasks.

After identifying necessary cost codes, estimates for cost and quantity of each are generated entered, employing historical and published data as well as good intuition. Estimated unit costs are determined by dividing the estimated cost by estimated quantity. Unit costs serve as the baseline for comparing project cost status throughout the life of the project.

Data entry related to Job to Date (JTD) is often less precise than that for the Estimated measures. JTD quantity, cost and unit costs are tracked similar to estimated values though the methods for acquiring this data are quite different.

For many cost code items, quantity is a simple calculation; the sum total of quantities like concrete poured or hours worked. Frequently however, quantity is more subjective. The amount of dirt hauled away from a job site provides an example. Not only can differences in bank and haul cubic yards, produced from varying soil densities, create less than exact estimates of JDT quantities, but incorrect assumptions as to how full or how many trucks leave the jobsite can compound accuracy problems. Particularly with large-scale earthmoving operations, the quantity of work performed under a particular cost code can vary greatly from the quantity estimated. Quantity overruns appearing on a report could therefore be attributed to problems ranging from poor contract documents to poor initial or contemporaneous estimates.

Job to date costs are also subject to wrongful data entry. To "balance" costs associated with work items that are overrun with ones under-run, expenses incurred for one cost code are sometimes placed under another. This practice leads to a false sense of security about certain work items. Similarly, if work needs to be performed where no related cost code is open, the cost may be placed under a different or unrelated code rather than opening a new one.

The above discussion underscores the need for accurate and honest data entry, from the estimating department to field operations manager. Without quality data entry, comparison between budgeted unit costs and JTD unit costs is meaningless. As noted, incorrect assumptions about quantities or dishonest budget balancing can lead to wrong interpretations of the JTD unit costs, resulting in equally dishonest graphical representation of the data.

Hierarchical Data Type

A hierarchy, necessary to link the internal accounting systems of the owner and general contractor, governs the vertical structure of the cost report. The broad hierarchy contains two types of nodes: pay items and cost codes. The project contract outlines a number of specific work items designated for payment upon progress or completion. In general, these items may reflect the owner's accounting system or may be based on explicitly outlined work packages. Particular to this dataset, the ***Pay Item*** is the quantity of work monitored by the contractor in terms of receiving payment from the owner, in this case the Virginia Department of Transportation. Not all pay items have the same method of reimbursement. Full treatment of reimbursement schemes for each pay item lies outside the scope of this work. It is sufficient to note that some pay items are fully compensated on completion while others are incrementally paid as work is performed.

In addition to the pay item designation set by the owner, the general contractor breaks up costs per an internal accounting system. As money is spent on work throughout the course of the project, the dollar amounts are appropriated to the corresponding cost code.

To bridge the gap between the owner designated pay items and general contractor specified cost codes a hierarchy is established. The contractor categorizes internal cost code items under specific owner designated pay items. Each cost code therefore has one "parent" pay item while pay items contain multiple cost code "children".

Temporal Data Type

The temporal nature of the cost report becomes apparent when comparing As Budgeted and Job to Date quantity, costs and unit costs. Though current cost reports do not include such data, changes in the unit cost over time is a useful indicator for each work item's cost progression. Cost history of codes over time is particularly important in light of certain user requirements and tasks.

FILTERING: USER SCENARIOS/LEVEL OF DETAIL

Filtering determines appropriate levels of graphical detail. Investigating user scenarios for results in identifying two primary types of individuals that interact regularly with the monthly budget/cost report: accountants and project managers. A series of interviews conducted with both accountants and project managers from the participating company helped determine the questions regularly asked of the cost report. In understanding how both project managers and accountants use the cost report, the appropriate level of detail can be presented within the graphical representation. Each of these levels is discussed below.

Project Management Level

In current practice, project managers receive a paper cost report related to their specific project each month. Should cost data need to be reviewed in the interim, an online version of the cost data is available. From the project manager's point of view, the most pertinent information in the report is the budgeted and job to date unit cost for current activities. Changes in unit costs are also important, though current methods necessitate having to return to previous cost reports to access this data. The project manager also monitors the presence and degree of quantity overruns to ensure that payment is received for all work performed. Visual representations most useful from the project manager's perspective include filtering the report down to currently active cost codes.

Accountant Level

The accounting department tracks costs for multiple projects in a less detailed fashion. The same parameters, the budgeted and job to date unit costs, are used to perform different tasks between the project manager and accountant. The accountant desires an overview to check if financial statements have been quoted appropriately. The overview helps to recognize profits evenly and take losses immediately by quickly identifying large problems.

Graphical representation of the cost report should effectively answer questions posed by both project managers and accountants. The primary dimensions for both groups, *as budgeted* compared with *job to date* costs require a slightly different level of detail. Identifying the size (accountants) and immediacy (project managers) of problems are also questions that the graphical representation should be able to easily answer. Project managers' interest additionally lies with the status of ongoing work items, necessitating some display of the percent complete. Before looking at specific visual strategies however, it is necessary to discuss the practical aspects of manipulating and filtering current data into a form necessary for integrated project control reporting.

Filtering and editing continues with organizing the original cost reporting data in a manner that meets both project manager and accountant needs. The current format does not readily enable such analysis. Organization proceeds by establishing the hierarchy and unit costs, formulating measures to answer given questions, and removing inaccurate or erroneous values.

Hierarchy and Unit Costs

As initially arranged, the spreadsheet format of the cost report contains 1040 cost code items, along with their corresponding description, unit of measure, type, budgeted quantity and dollar value and JTD quantity and dollar value; 1040 rows with eight dimensions. The cost code number reflects both the pay item and cost code designations. Division of this column using a *Text to Columns* function yields a hierarchy format. The first seven numbers relate to the pay item designation—the parent level of the hierarchy, while the last five are specific to cost codes. Unit costs present in the paper cost report are calculated by dividing *As Budgeted or Job to Date* dollar values by their respective quantities.

Answers to Questions

To quickly answer *comparison* questions posed by project managers and accountants, the cost index is computed. As defined, this measure compares the *budgeted* with *as built* costs in terms of the work performed:

Cost Index = Actual Cost of Work Performed / Budgeted Cost of Work Performed

To determine the cost index, the work performed is defined as the job to date (JTD) quantity. The budgeted cost is simply the unit cost corresponding to each cost code. Thus the budgeted cost of work performed is the product of the unit cost and number of units (quantity).

Budgeted Cost of Work Performed = BC (unit cost) * WP (JTD quantity)

The Actual Cost of Work Performed is defined as the Job to Date costs. As an indicator of how far work has progressed, the percent complete is defined as:

Percent Complete = JTD Quantity / Budgeted Quantity

The percent complete can therefore be greater than 100% for items whose JTD quantity exceeds budgeted quantities. Work items with a percent complete greater than 100 are labeled as quantity overrun.

Removal of Erroneous or Extraneous Data

Modifications described to this point yield a number of problems. For cost code items with no budgeted quantity the budgeted unit cost does not exist. Similarly, for items with no work performed, the actual unit cost does not exist. To make the information manageable, rows with no work scheduled and no work performed are removed, a pragmatic decision in light of the project manager's desire to track current work items.

Data Selection

A remaining subset of 890 remaining cost codes allows for simpler study and feedback of the graphical representations. To explore both the multidimensional and hierarchical features of the report, the top-level hierarchy (approximately 60 pay items) is selected. Aggregating contractor cost codes into their corresponding pay items provides summary level percent complete and cost index formulas.

Summary Cost Index = ACWP / BCWP; where
ACWP = Actual Cost of Work Performed = S (JTD Cost)
BCWP = Budgeted Cost of Work Performed = S (Budgeted Unit Costs * JTD Quantity)
Summary % Complete = S (BCWP) / S (Budgeted Cost)

The cost index is used for its intuitive meaning: a high cost index reflects a cost overrun situation. Summary percent complete values reflect a weighted average of all lower level cost codes. The cost indices and percent complete values for 60 pay items are used in the graphical strategies discussed in the next chapter. These values affect the most pertinent user tasks, those of comparing the job to date with budgeted costs. Additionally, the budget has been filtered to emphasize an overview of currently active cost codes.

VISUAL STRATEGIES

For developing possible visual strategies and editing data graphics (honesty and efficiency), the authors investigated four visual layout strategies for representing the structured budget/cost control data. The strategies were centered on the two primary types of information found within the cost report, the multivariate and hierarchical data types. Additionally, the temporal importance of the data was added to each of the strategies. Note that this step in creating appropriate visual representation merges several visual strategies noted in Figure 1. The strategies include scatterplot, linked histogram, hierarchical tree, and treemap layouts (Figure 5). The treemap strategy is explored in detail in this paper.

Treemap

A Treemap's visual structure allows for a number of dimensions to be displayed in an efficient manner. The most prominent feature is area; large branches in the hierarchy are given large areas. Because the cost report is hierarchically aggregated around the budgeted dollar (the cost code's budget sums to the corresponding pay item's budget) displaying the budget size in proportion to area is selected. A color scale is selected to show the cost index information.

The percent complete of each pay item is displayed by the degree of shading for each rectangle. Thus, pay items with an aggregated completeness of 50 percent are shaded 50 percent. (A threshold level of 20 percent shading is selected to show enough cost data for each pay item). For quantity overrun items, those past the budgeted quantity, shading with the color black begins from the center of the rectangle out. Dark shading

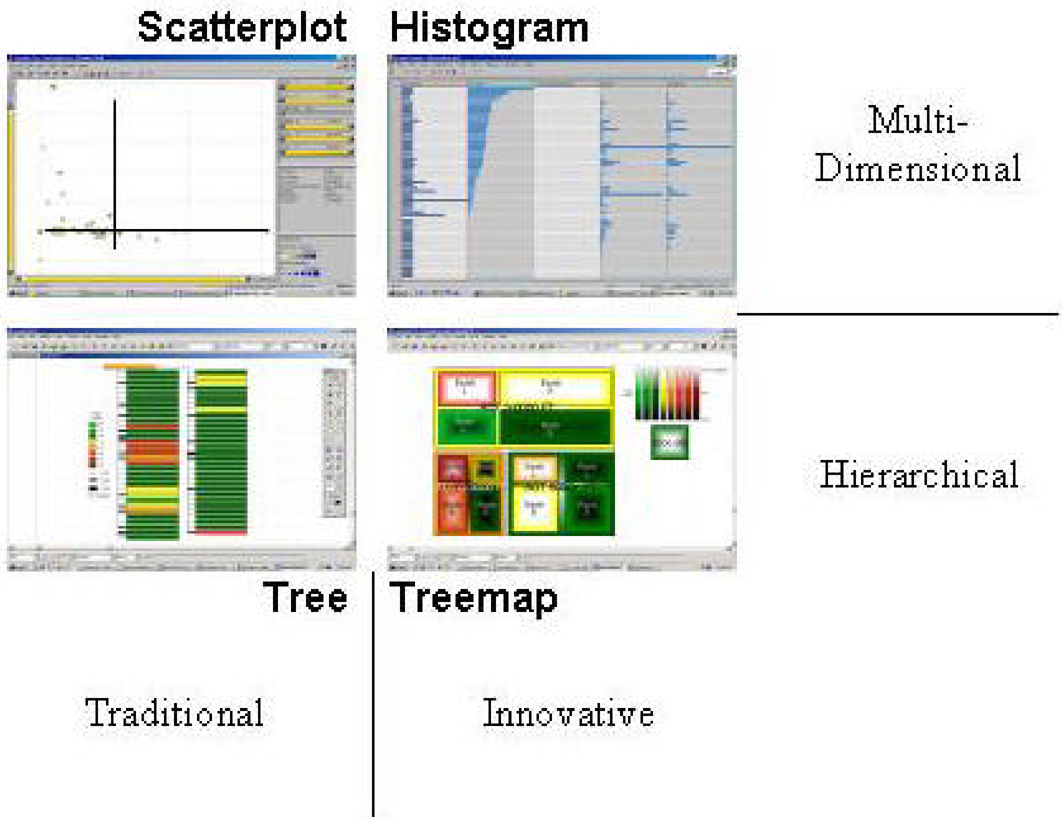

Figure 5. Visual Strategies

provides a powerful indicator for all quantity overrun items (see Fig. 6, second from the top right).

Interaction between color and shading emphasizes reasons for the color scale selected. Colors need to be distinct enough to minimize interaction with shading so as to not cause confusion between the cost index and percent complete. Though a continuous scale of two colors is usually better, distinction between colors necessitated using a color scheme similar to that used to show precipitation intensity by radar.

Visual inclusion rather than branching allows for the display of multiple hierarchical levels; levels deeper in the hierarchy are included within the area assigned to higher levels (see Figure 6, upper left corner). Hierarchical levels may be turned on and off to allow for efficient observation within the same layout strategy. Data pertinent to specific work areas are similarly filtered. The treemap thus provides details on demand of problem areas within the context of the higher levels (see Figure 7, two levels of detail).

For determining causes of cost or quantity problems, small multiple treemaps can be used. Dimensions other than cost and quantity can be associated with color and shading to explore relationships between cost or quantity overruns. Shading with respect to unanswered RFIs or procurement delays in one diagram could reveal reasons why the schedule or cost index is a particular shade in another diagram. The temporal history of the cost or

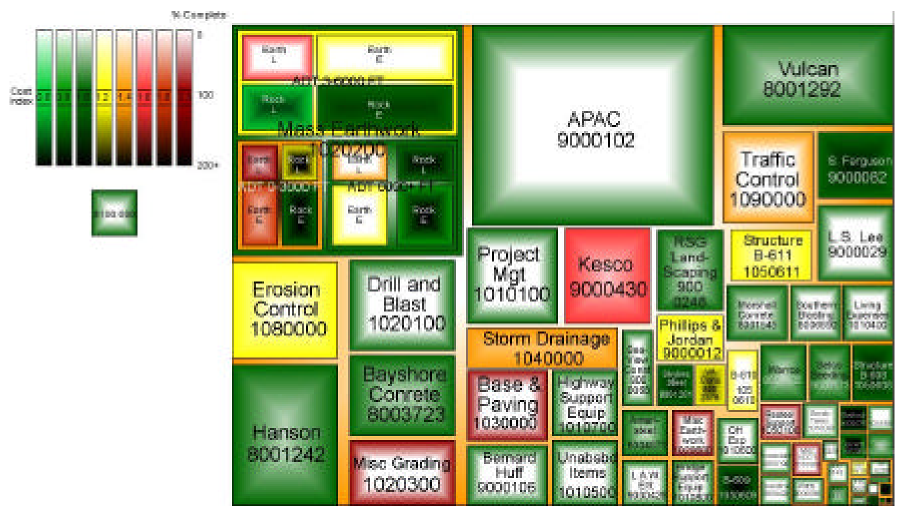

Figure 6. Treemap representation of monthly cost report, pay item level

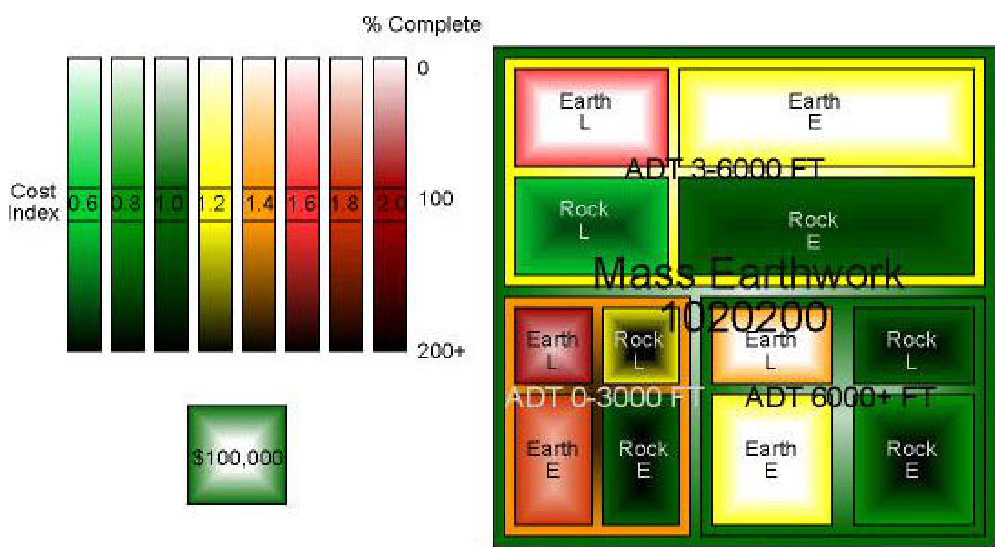

Figure 7. Treemap representation of the Cost Code level of detail

schedule index using color more purposely introduces the temporal nature of the data. Thus, the treemap layout lends itself to reuse when exploring the impact of processes on cost or scheduling problems.

Treemap Performance

The layout described directly answers the most pertinent user questions. A colored cost index and shading provide comparative answers to cost issues for currently active work items. (In an integrated system, a query of in progress activities would produce a similar layout). The eye is quickly drawn to burgundy outlined rectangles—those work items currently overrun and largely incomplete; pay items for which intervention is possible. Likewise, the relative size of problems and presence of overruns are efficiently displayed on the initial overview graphic.

The treemap, when fully populated with all levels of detail in the hierarchy, encourages micro/macro reading and has a high data/ink ratio. Visual structure is similar for all but the database level of the cost report, increasing the overall efficiency of information displayed in this manner.

CHARETTE TESTING

User survey results provide support to validate the proposed framework. The method used to evaluate the effectiveness of graphics follows a process similar to charrette software user testing formalized by Clayton, Fischer and Kunz (Clayton et. al. 1998). The charrette test method empirically compares newly created visual strategies with conventional graphics.

Graphic strategies in this research were tested for accuracy and user perception. Testing for accuracy provides an objective result of each graphic's effectiveness. User perception rating is more subjective, however, it assess individual responses to each graphical layout. Users evaluated visual strategies by responding to six questions. Recall, the authors' investigated scatterplots, histograms, trees, and treemaps (figure 5).

The variability of a comparison between the prototype visual strategies and conventional budget/cost graphical reporting need not involve complicated statistical methods. Unanticipated patterns must be considered. Accurate answers for each graphic are calculated as deviations from a correct response. The analysis provides evidence to support conclusions regarding the effectiveness of the framework (Clayton et. al. 1998).

Test Description

A series of six questions for field and office personnel were used to test each graphic for accuracy. The questions arise from user scenarios identified during interviewing construction organizations. Participants receive the traditional tabular cost report and the four prototype visual representations. Each question solicits an accurate response based on assessment of a graphic or the table. After responding to each question, participants are asked to estimate the ease at which the graphic helps answer the question on a scale from 1 to 10; one being most difficult, ten being least.

Test Questions

Three topical groupings of two questions each test the graphic strategies with respect to cost, quantity and integrated cost-quantity. The six questions are given to participants in the following order (general answers to some questions are provided here for clarity).

1. *Integrated Cost-Quantity:* At this point in time, which five Pay Items are the most cost overrun with respect to where spending should be?
   Answer: Pay items with the highest Cost Index
2. *Cost:* As a whole is this project: far over, over, at, under or far under budget? (*far over* being more than 1.5*budget, *over* being between 1.5* and 1.1*budget, *at* being between 1.1* and 0.9*budget, etc).
   Answer: Over budget
3. *Cost:* Which five Pay Items have the largest budget?
4. *Quantity:* How many Pay Items have quantity overruns?
   Answer: 10
5. *Quantity:* Which Pay Items are nearest to completion (assuming no quantity overruns)?
6. *Integrated Cost-Quantity:* To which *cost critical* Pay Items should attention be given due to a *large percentage of work remaining*?

   Answer: Pay items whose current Cost Index, if continues though the life of the item, will cause the largest cost overrun.
   .

After completing the questions, participants are invited to qualitatively discuss any graphic. The scorecard and graphics with recorded answers are collected at the end of the trial and, along with any qualitative feedback, provide preliminary assessments of the graphic strategies.

Test Participants

Calibration trials to determine whether procedures are clear and complete were conducted. The number and occupation of test participants also adhered to Clayton et al's method (Clayton et. al. 1998). Although a large number of participants is always desirable, as few as four participants provide useful information to the researcher. Ideally the participants should be from the targeted population with experience in budget/cost reporting. A within-subjects approach is recommended to allow a smaller number of participants (Clayton et. al. 1998).

Multiple iterations and the continuous improvement necessary for visual prototypes mean that large numbers of participants are unnecessary. Thus, seven industry professionals participated in this research. One-to-one testing between the researcher and participant took approximately 90 minutes.

RESULTS

Figure 8 illustrates average values from the seven participants. Results are tabulated for the textual cost report and each of the four graphics (scatterplot, histogram, tree and treemap). Although the Treemap strategy is the focus of this paper, for discussion and comparison, figure 8 illustrates charrette results for all 4 graphic strategies that are part of the larger study.   The graphics are rated for accuracy on scale from zero to five. The user perception assessment is scored from one to ten.

		Accuracy Test (averages)						Intiutive Test (averages)			
Question	Report	Scatterplot	Histograms	Tree	Treemap	Question	Report	Scatterplot	Histograms	Tree	Treemap
1	0.2	5.0	5.0	3.7	3.9	1	2.0	7.0	7.0	2.9	4.4
2	5.0	4.8	4.5	4.5	4.8	2	6.3	6.3	4.6	3.9	4.3
3	4.7	4.6	4.9	3.4	4.6	3	8.3	4.4	8.0	3.3	7.1
4	4.8	4.6	4.2	-	4.6	4	8.9	6.7	8.4	-	5.6
5	4.6	5.0	4.3	-	1.4	5	8.1	6.6	8.9	-	3.7
6	2.3	4.1	3.0	-	3.7	6	6.1	8.0	7.4	-	6.1

Figure 8. Quantitative results from Charrette testing

Test Results: Accuracy and Opinion Scores

In general, all graphics provided accurate responses to most questions. From observation, the graphics offer answers more readily than did the tabular cost report. Significant deviations in accuracy occurred only a few times. The traditional textual report had low accuracy scores for questions one and six; the questions that focused on integrated cost/quantity responses. Lower scores for these questions arose from the lack of providing the cost index value on the spreadsheet (current reports lack this comparative measure). The mental work of calculating and ranking all pay items in order proved to be a difficult task.

The histogram also scores lower than normal for question six (To which cost critical Pay Items should attention be given due to a large percentage of work remaining?). Without the interactive features common to this dynamic tool, participants found it difficult to relate a high cost index histogram bar with a low percent complete bar.

Finally, the treemap scored with a low accuracy for question five (Which Pay Items are nearest to completion—assuming no quantity overruns?). Interaction between color and shading, particularly with smaller cost items, makes it difficult to assess which pay items are nearing completion.

Opinion ratings proved more variable than accuracy scores. The traditional report, scatterplot and histograms had a low rating for one question each. The tree had the lowest average opinion while the treemap has the highest variation in opinion scores possibly due to its unconventional format.

Qualitative Test Results

Participants were invited to discuss the testing or graphical representations upon completing the questions detailed above. During such conversation two interesting physical phenomena, highlighted by the treemap but unknown to the researcher, presented themselves.

Level of Detail: Mass Earthwork

The pay item *Mass Earthwork* is comprised of twelve cost codes. In the treemap, cost codes are displayed proportionally by budget size within the area assigned to their pay item. Details at the *cost code level* (one level deeper than the pay items) include the *work area* (here E or L for Equipment or Labor) and *cost code description* (ADT 3-6000 FT corresponds to

Articulated Dump Truck for excavation stations 3000 ft to 6000 ft. (Figure 7.

Figure 7 reveals that all *Earth* cost codes have a Cost Index greater than 1. Additionally, all *Rock* cost codes enjoy Cost Indices at or below 1 with the exception of "Rock, ADT 0-3000 FT, Labor". In terms of percent complete, all *Earth* codes are less than 100% complete while all *Rock* codes have some degree of overrun.

Visual indicators thus show the presence of a *cost overrun* with the earth-moving operations and *quantity overrun* with rock moving. It appears that *rock* quantity overruns indicate future *earth* quantity overruns. Instead, test participants with knowledge of the job hypothesize that incorrect estimates were made about the original quantity of rock and earth. The amount of material to move was approximately correct in total, but actual site conditions contained proportionally more rock and less earth than estimated.

Therefore, all earth quantities are under-run while rock quantities are overrun. The physical situation revealed by the treemap underscores the opportunity to use the visual strategy as a small multiple. Were the scheduling data available for these cost codes, another treemap could be constructed and visually queried to determine the schedule completeness of these work packages. If all cost codes are complete the above hypothesis would be verified.

Drill and Blast vs Kesco

The project's accountant highlighted another physical artifact displayed by the treemap. Though most of the earthwork is complete (see Mass Earthwork's rectangle), only a small portion of the Drill and Blast pay item has been performed (Figure 6). Nearly complete earthwork combined with little drilling and blasting signals some type of an accounting problem. Initial methods for the project dictated that blasting would be self-performed. As the project progressed, blasting costs escalated and the task was eventually subcontracted out to Kesco (possibly on account of the large amount of rock discovered—see discussion above). Observation of the treemap shows that the size of the initial budgets for 'Drill and Blast' and 'Kesco' are similar. Thus, the treemap provides a good visual explanation to other participants for certain budgeting situations.

CONCLUSION

This paper provides a framework for developing visualization strategies for multi-dimensional construction control data. The framework encourages an interactive process of Structure-Filter-Communicate while considering level of detail, density, and efficiency of data representation.

The framework is illustrated for developing cost control/budget visualization layouts using a treemap strategy. Through charrette testing, both the framework and treemap strategy were found to be appropriate tools for enhancing the process of visually displaying construction control data.

APPENDIX I: REFERENCES

Bertin, J. (1983). Semiology of Graphics. Madison, WI., University of Wisconsin Press.

Card, S. K., Mackinlay, J., and Shneiderman, B. (1999). Readings on Information Visualization: Using Vision to Think. San Francisco, Morgan Kaufmann Publishers, Inc.

Clayton, M.J., Fischer, M.A., and Kunz, J.C. (1998). CAD Prototype Testing: Worked Examples, Demonstrations, Trials, and Charrettes, Proceedings from ASCE International Computing Congress,.Boston, MA, pp. 106-116.

Griffis, BF.H., and Struts, C., (2000). FIAPP and the 3Dimensional Computer Model, Proceedings from ASCE Construction Congress V., Orlando FL, pp. 996-1006.

Kunz, J., Fischer, M., Haymaker, J., and Levitt, R. (2001). Integrated and Automated Project Processes in Civil Engineering; Experiences of the Center for Integrated Facility Engineering at Stanford University, Proceedings from ASCE Specialty Conference. Blacksburg, VA., pp. 96-105.

Russel, A. (1997). Challenges and a Vision for Computer-Integrated Management System for Medium Sized Contractors, Canadian Journal of Civil Engineering, 24 (1), pp. 180-190.

Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization, Proceedings of IEEE Symposium on Visual Languages. Los Alamitos, CA, IEEE.

Spence, R. (2001). Information Visualization. Essex, England, ACM Press / Addison-Wesley.

Subramanian, P., Songer, A., and Diekmann, J. (2000). Visualizing Process Performance, Proceedings from ASCE International Conference on Computing in Civil and Building Construction. Stanford, CA., pp. 1-6.

Tufte, E. R. (1983). The Visual Display of Quantitative Information. Cheshire, CT., Graphics Press.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. CMS-9410683.

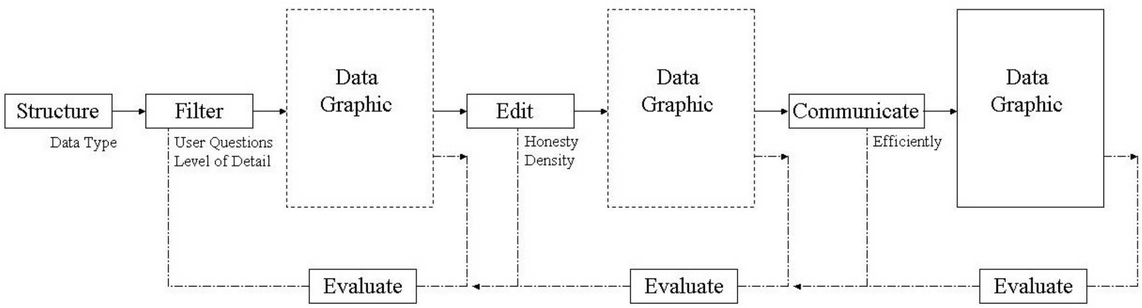

Figure 1. Visual Development Framework

Data Type	Visual Strategies
1-Dimensional	Histograms, Piecharts, Tukey Box Plots
2-Dimensional	Scatterplots, Linked Histograms
3-Dimensional	*Physical Data*: Three Orthogonal Axes
	Relational Data: See Multi-D below
Multi-Dimensional	*Glyphs and Cleveland's Rules*: Position, Length, Direction, Area, Angle, Volume, Curvature, Shading, Color Saturation
	Statistical Methods: Perturbation / Scatterplot Matrix, Mosaic Glyphs, Small Multiples and Multidimensional Icons
	Visual Separation: Micro/Macro Strategies, Color, Layering
Temporal	*1D + Time*: Time Series
	Multi-D + Time: Small Multiple, Parallel Axis Time Series
Hierarchical	Tree (Orthogonal or Circular), Treemap
Network	Simple, Circular, Three Dimensional, Aggregated

Figure 2. Visual strategies by data type

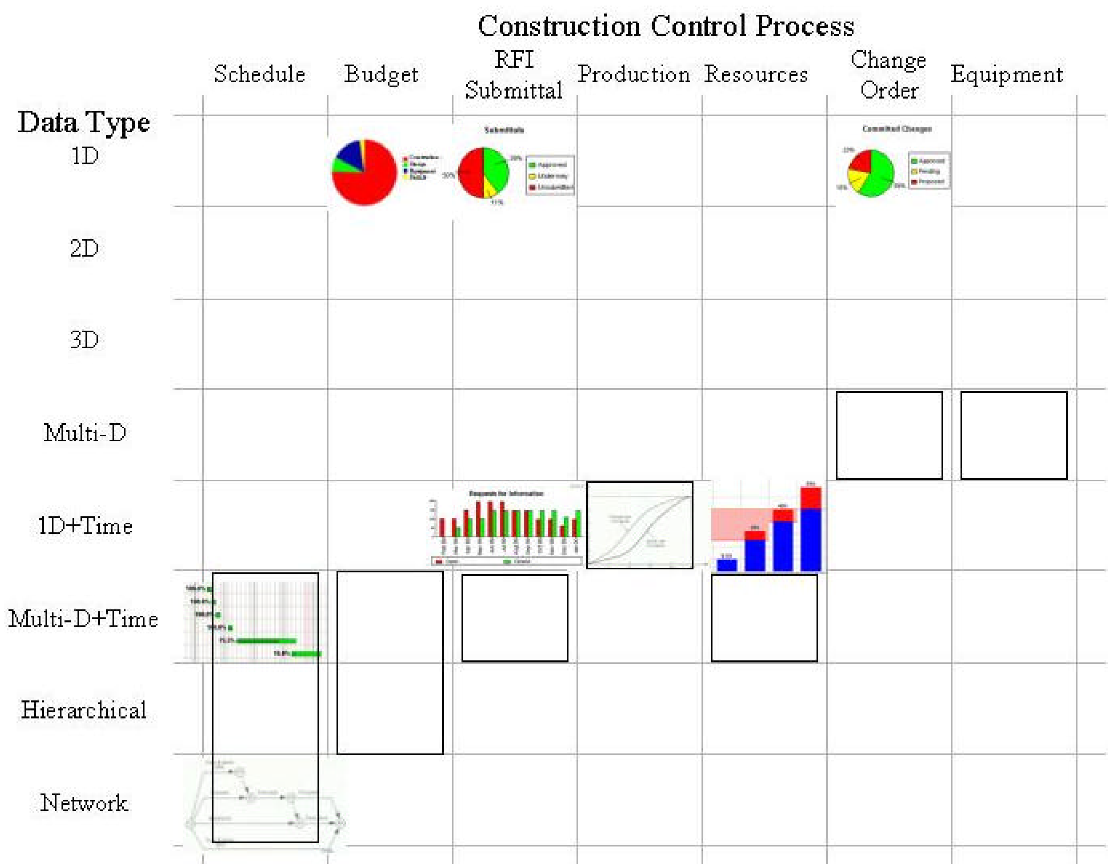

Figure 3. Graphical state of construction project controls: a taxonomy

CO#	Cost Code	Cost Code Description	UOM	Type	Bdgt Qty	Bdgt $$	JTD Qty	JTD $$
20	1020200123000	EARTH ADT 0-3000 FT	M3	E	155000	205031	131177.7	285809
20	1020200123000	EARTH ADT 0-3000 FT	M3	L	155000	81920	131177.7	146781.3
20	1020200124000	EARTH ADT 3-6000 FT	M3	E	312000	575531	32247.28	73538.49
20	1020200124000	EARTH ADT 3-6000 FT	M3	L	312000	208204	32247.28	33495.03
20	1020200125000	EARTH ADT +6000 FT	M3	E	180000	360002	62520	135466.2
20	1020200125000	EARTH ADT +6000 FT	M3	L	180000	128649	62520	60144.95
20	1020200160000	EARTH - TANDEMS	M3	E	0	0	59290	85017.44
20	1020200160000	EARTH - TANDEMS	M3	L	0	0	59290	60431.54
20	1020200169000	INACTIVE	M3	E	0	0	0	0
20	1020200169000	INACTIVE	M3	L	0	0	0	0
20	1020200523000	ROCK ADT 0-3000 FT	M3	E	104000	216668	156080	288950.3
20	1020200523000	ROCK ADT 0-3000 FT	M3	L	104000	81327	156080	129299.3
20	1020200524000	ROCK ADT 3-6000 FT	M3	E	208000	580422	250164.8	441461.4
20	1020200524000	ROCK ADT 3-6000 FT	M3	L	208000	202089	250164.8	202648
20	1020200525000	ROCK ADT +6000 FT	M3	E	120000	359715	167539	351673.8
20	1020200525000	ROCK ADT +6000FT	M3	L	120000	114355	167539	152520.3

Figure 4. Monthly Cost Report Data

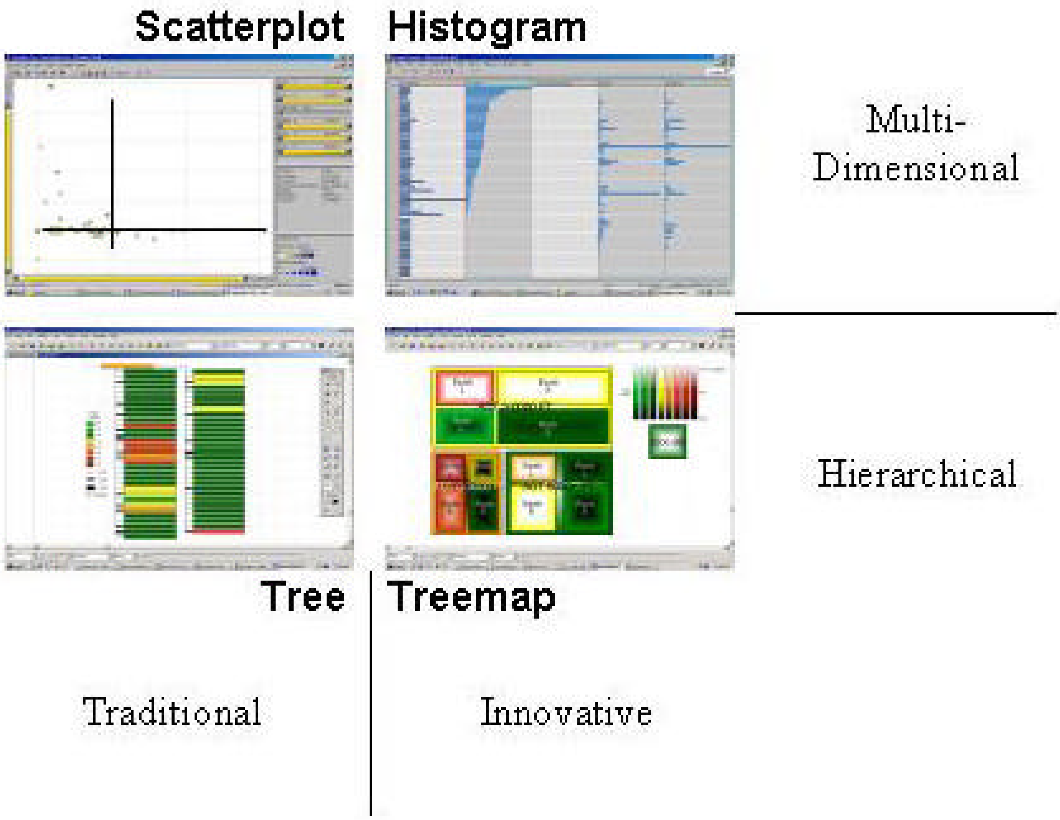

Figure 5. Visual Strategies

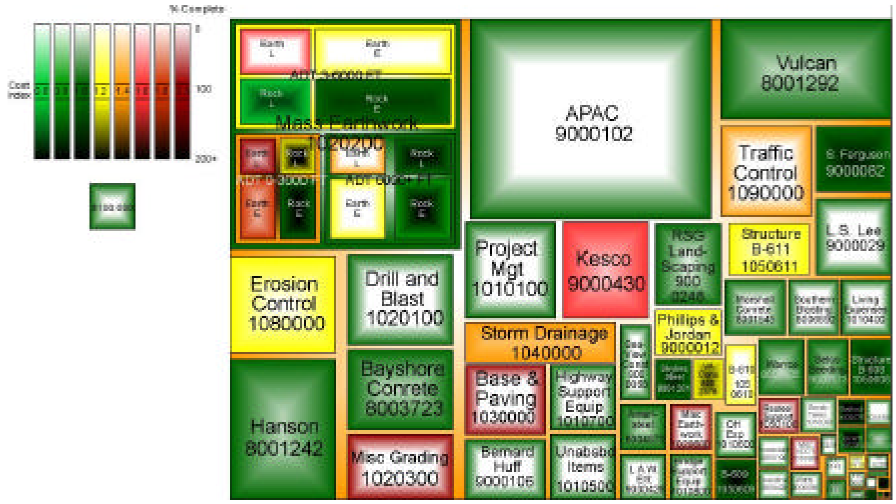

Figure 6. Treemap representation of monthly cost report, pay item level

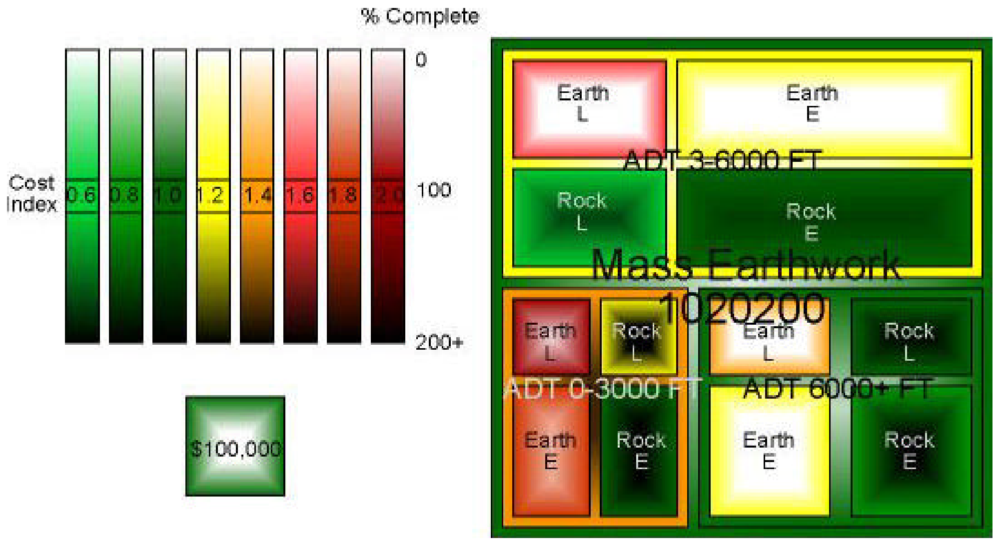

Figure 7. Treemap representation of the Cost Code level of detail

		Accuracy Test (averages)						Intiutive Test (averages)			
Question	Report	Scatterplot	Histograms	Tree	Treemap	Question	Report	Scatterplot	Histograms	Tree	Treemap
1	0.2	5.0	5.0	3.7	3.9	1	2.0	7.0	7.0	2.9	4.4
2	5.0	4.8	4.5	4.5	4.8	2	6.3	6.3	4.6	3.9	4.3
3	4.7	4.6	4.9	3.4	4.6	3	8.3	4.4	8.0	3.3	7.1
4	4.8	4.6	4.2	-	4.6	4	8.9	6.7	8.4	-	5.6
5	4.6	5.0	4.3	-	1.4	5	8.1	6.6	8.9	-	3.7
6	2.3	4.1	3.0	-	3.7	6	6.1	8.0	7.4	-	6.1

Figure 8. Quantitative results from Charrette testing