Visual Analytics for Multivariate Sorting of Sport Event Data

Conference Paper · January 2013

7 authors, including:

Phil Legg
University of the West of England, Bristol
37 PUBLICATIONS **428** CITATIONS

Iwan Watts Griffiths
Swansea University
42 PUBLICATIONS **759** CITATIONS

Robert S. Laramee
Swansea University
113 PUBLICATIONS **2,546** CITATIONS

Visual Analytics for Multivariate Sorting of Sport Event Data

D. H. S. Chung, P. A. Legg, M. L. Parry, I. W. Griffiths, R. Bown, R. S. Laramee, and M. Chen

Abstract— A critical job coaches and sport analysts are tasked with is the planning of key match videos for analytical coaching sessions. Each session may focus on a diverse range of topics, such as the strengths and weaknesses of a game. This needs to be tailored further based on a player's tactical position or skill. Hence, the criteria for sorting video is dynamic. This motivates a sorting criteria beyond individual attributes of a multi-dimensional data set. We propose a knowledge-assisted, event ranking framework to interactively model implicit sorting as formal parameters that can be used to perform multivariate sorting. We incorporate knowledge in the form of a user's event ranking which we formalize using regression analysis. Depending on the ranking criteria, the resulting function can be customized to many forms such as importance, or other performance metrics. We use visual analytic methods to depict the set of sortable attributes and weights determined by the model. Visual feedback helps the user comprehend the function, and aids in choosing the most appropriate model. We find that this approach significantly increases the usability of multivariate sorting and allows domain experts to incorporate their knowledge and expertise into the analysis. This work is undertaken in conjunction with a national rugby team. To demonstrate the effectiveness of our sorting system, we present a use case scenario in rugby event analysis, where coaches and analysts need to re-organize match videos in order to study and evaluate team and player performance.

1 Introduction

Event sorting is a fundamental task in visual analytics. Such a task becomes challenging when sorting involves several data dimensions, and the way in which each dimension influences the sorting is not well defined. Such a sorting task is commonplace in practical visual analytics, where one often encounters ad hoc request for organizing data in to some kind of order without precise specification of the relevant *sort keys* and a *sorting function*. Although some analytical methods such as multidimensional scaling (MDS) [6] or principle component analysis (PCA) [16] may help in some applications (e.g., [14]), they focus on the discovery of the most influential attributes in the data, rather than the discovery of a *sorting function* for an ad hoc requirement of sorting task. This work addresses this challenge in the context of sports event analysis by using a knowledge-assisted visual analytics process.

We notice that when given an ad hoc requirement of event organization, a user normally knows well how the sorting outcome should look like, without knowing explicitly about the sort keys and a sorting function with a visualization. In a knowledge framework [4], we can summarize the situations as follows:

- Users have *tacit* knowledge about sorting a set of events, but do not have the *formal* knowledge as to a sorting function. They may have partial knowledge about sort keys as they typically speculate a set of attributes that may influence the sorting.

- Although users can organize a given set of events in an 'accurate' manner using their tacit knowledge (because they define the expected sorting outcome), this does not scale up to a large number of events. It is generally easy for users to place a few most representative events (e.g., success, neutral, failure) into order. The task becomes inefficient when the number of events increases significantly, and ineffective (i.e., less 'accurate') for events with a similar principle criterion (e.g., how successful), but a diverse set of conditions (left or right, earlier or later, different players involved, etc.).

- On the other hand, the system does not have any a priori knowledge about the expected sorting outcome, since the sorting requirement is not predefined. Of course, it does not have the formal knowledge about a sorting function either. If the system has a sorting function, it can perform event sorting in a more scalable and consistent manner.

We thereby developed a visual analytics system that enables users to provide the system with some of their *tacit knowledge* by selecting a small set of events (typically 3-7), and placing them in an order as an example for the system. The users may also provide their *partial knowledge* about possible attributes (i.e., data dimensions) that should be considered. This partial knowledge is not essential, but can reduce the amount of computation significantly. The system uses an analytical method to convert the *tacit* and *partial knowledge* to some *formal knowledge* in the form of a potential sorting function and a measure of influence of different sort keys. The system then provide users with a visualization of the sorted results in relation to the potential sorting function and the weights of different sort keys. The former is shown in a glyph-based sorting canvas, and the latter in a parallel coordinates plot. Users can interactively refine the sorting results and the weights of different sort keys, or re-activate the knowledge discovery process by refining their initial specification of the example set or the speculated data dimensions. Satisfactory results can normally be obtained within a few iterations, and users can produce a sorted set of events (i.e., video clips) for supporting further analytical tasks such as compiling various statistical indicators in relation to the sorted events, and analyzing video clips in a coaching session. Our contributions are:

- We introduce a novel visual analytic approach to sorting multi-dimensional events by converting users' tacit and partial knowledge to formal knowledge.

- We develop a system that supports such a process iteratively through a close integration of interaction, analysis and visualization.

- We demonstrate the efficiency and effectiveness of visual analytics for multivariate sorting through a real-world application, and we evaluate our work objectively with a user consultation.

The remainder of the paper is organized as follows: In Section 2, we provide a brief overview of related work. Section 3 gives some background on rugby and outlines our motivation. The pipeline of the visual analytic system is described in Section 4. In Section 5 we describe

- *D. H. S. Chung, P. A. Legg and M. L. Parry are with the Department of Computer Science and the College of Engineering, Swansea University, UK.*
- *I. W. Griffiths, is with the College of Engineering, Swansea University, UK.*
- *R. Bown, is the head performance analyst at the Welsh Rugby Union.*
- *R. S. Laramee, is with the Department of Computer Science, Swansea University.*
- *M. Chen, is with Oxford e-Research Centre, Oxford University, UK.*

Manuscript received 31 March 2013; accepted 1 August 2013; posted online 13 October 2013; mailed on 27 September 2013.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

the method for converting tacit and partial knowledge to an explicit sorting function in our knowledge-assisted, event ranking framework. Section 6 details the visual mappings and interaction of multivariate sorting. We present our results in Section 7 and evaluate our work in Section 8. In Section 9 we discuss the limitations of the system and draw our concluding remarks in Section 10.

2 Related Work

Sorting is the computational process of rearranging a sequence of items in order, or categorizing entities with similar properties [18]. We utilize a concept similar to card sorting [26], a user-centered design that allows a user to decide how to categorize a set of items into groups or structures they are familiar with. This approach has been effectively used for classifying symbols in cartography [25], organizing online course sites [7], and clustering multivariate glyphs [2, 17], where the glyphs are treated as cards. The fundamental difference here is that we express different levels of knowledge as a sorting function to order events with an implicit sorting criteria.

The incorporation of knowledge in visual analytics is still a early research topic. Wang *et al.* [29] propose a framework to support communication between both, domain and individual knowledge structure. Lipford *et al.* [20] explore the use of visualization to help recall users reasoning. Mistelbauer *et al.* [21] introduce Smart Super Views for analysis of different data sets in medical visualization. More work has been studied on using knowledge to tune parameters in statistical models. Heimerl *et al.* [10] present a user-study comparing three approaches to interactive classifier training for text search and filtering. Hoferlin *et al.* [11] describe the use of inter-active-learning in visual analytics to allow users to adjust and understand complex classifier models through visualization. Other approaches include using a distance function [3], semantic interaction [8] that captures the analytical reasoning of users and multiple views for cross-filtering analysis [31]. To the best of our knowledge, this is the first work of its kind introducing visual analytics to sorting events by converting users' tacit and partial knowledge to formal parameters.

Sports Visualization is an area that is gaining interest within the community. Parry *et al.* [23] propose a framework for hierarchical event visualization that encodes event importance, and demonstrate this in application to sports. Jin and Banks make use of treemaps for visualizing scoring results and match statistics of tennis matches [15]. Moore *et al.* [22] look at the potential of using visualization techniques for spatial temporal analysis of rugby data. Legg *et al.* [19] conducted a design study to show the effective use of glyph-based visualization within sports performance analysis. Pileggi *et al.* [24] introduce SnapShot: a system that enhances hockey analytics through visualization.

3 Understanding the Problem

Analyzing performances and devising strategies is a critical process in every team sport. Teams will study details with fine granularity in order to gain a competitive edge over their opponents. We have worked in close collaboration with the Welsh Rugby Union (WRU) team over a 2 year period. During our initial consultation, the performance analysis team outlines their current work flow and existing problems:

"In our current practice, getting the data and generating charts (through spreadsheets) is very time consuming. Once a chart is plotted, we often get "What if we take this variable into account?", which then requires us to go back to the raw data and process it all again."

This suggests that an intelligent sorting mechanism beyond sorting by individual attributes of a multi-dimensional dataset is needed. To fully consider the challenges involved with rugby event analysis and how a visual analytic system for sorting can be of significant advantage, we provide a background to the game in this section.

3.1 Rugby Union

Rugby Union is a popular team sport which consists of two teams (of 15 players) who advance an oval ball across a rectangular field (up to 144m long by 70m wide) with two H-shaped goal posts at either end. The game is played primarily by carrying the oval ball from one end of the pitch to the other. Points can be scored in several ways: A *try*, which involves grounding the ball in the opposition goal area, or through kicking the ball between the H-Shaped post from a *conversion*, *penalty kick* or *drop goal*. Each match is played in two 40-minute halves, where the objective is to score more points than the opponent.

3.2 Rugby Event Analysis

Analysts and sporting coaches heavily rely on using notational data [12] for player and team analysis. Notational analysis consists of "*tagging*" video footage with key events and semantic notations from which key performance indicators can be derived. One resulting output is a set of video clips that capture moments in a game of when a team receives and loses/changes possession. In rugby, such events are known as a *phase ball event*, and involves smaller *phases* that describe the period of play. Thus, a single match is decomposed into a sequence of phase ball events that may lead to points scored. With the advancement of new digital media, these events are enriched further with a multitude of data sources (e.g., nutrition data, player data, and ball tracking) that increases the complexity of the analysis. The event descriptors (or attributes) we use in this work are:

- **start event** — the type of event in which play is started.
- **gain** — the distance gained towards the goal area.
- **territory start position** — the starting position of ball when a team receives possession in relation to the the goal area.
- **time** — the starting time of the event.
- **tortuosity** — the tortuosity of the ball path.
- **number of phases** — a count of the phases.

Coaches, players and sports analysts have to watch the video clips frequently, usually in a selective but comparative manner. For instance, an analyst can often be asked to present the 10 most significant events of a series of matches for a coaching or tactical session. One desirable tool is to enable analysts to collect all the events (metaphorically as a deck of cards) and to lay the events (cf. cards) out in different ways. This would empower human analysts to make use of their knowledge and experience within the analysis. The conventional video editing software does allow selecting clips easily, but it is time consuming to select clips according to some criteria (cf. finding the cards) and combining them together into a form in which events can be compared (cf. lay the cards out). Since the criteria for sorting video can vary for each coaching staff (e.g., the attacking coach or defence coach), this becomes a challenging task when sorting can only be performed on individual attributes of a multivariate space. This is particularly difficult when the criteria is not well-defined (i.e., selecting the most significant events). It is the card metaphor that led us to think about incorporating a user's knowledge in visual analysis for sorting. We introduce a framework to model a user's tacit ordering requirement as a multivariate sorting function based on the event attributes.

3.3 Tacit knowledge vs Formal knowledge

The incorporation of human knowledge in visual analytics describes the process of transferring *knowledge* into some explicit form (e.g., a function) [29]. For the purpose of this paper, we define the following:

Tacit knowledge. User's input of event ranking. The know-how cannot easily or explicitly be transferred to another user.

Partial knowledge. User's input of possible attributes that may affect the ranking. This narrows the search space.

Knowledge discovery. Using regression analysis for determining sortable attributes, how the attributes are combined into a ranking function, and evaluating the accuracy of the function.

Formal knowledge. The function (including the weights on various attributes) is formal knowledge as it can be explicitly written down and transferred to others.

Knowledge externalization. Visualizing the sortable attributes and their weights and their impact on sorting various events.

Knowledge application. Sorting video clips to be watched in an analytical coaching session.

In the next section, we use the above definitions to describe the pipeline to our visual analytic framework.

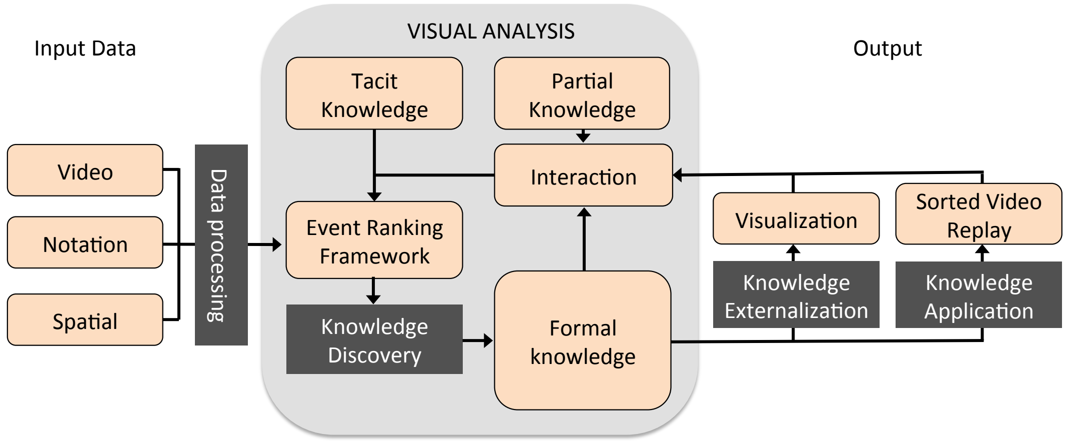


Fig. 1. A graphical pipeline illustrating our visual analytic system for multivariate sorting of rugby event data. It consists of four steps: processing the input rugby event data, a knowledge discovery process to derive a set of sortable attributes combined into a function (formal knowledge), knowledge externalization to determine the function's impact on sorting various events in a visual form, and finally, using formal knowledge to interactively sort and replay match videos (knowledge application).

4 Overview

Fig. 1 shows the process to our visual analytic sorting system. It consists of four key steps: *processing* of rugby event data, *knowledge discovery*, *knowledge externalization* and *knowledge application*.

The first step involves processing and integrating each data source into an event data structure in order to extract the underlying sorting attributes. We combine three types of data, namely: video, notation and spatial ball tracking data. Rugby event data is a mixed data set containing both quantitative attributes (e.g., gain, time, ball tortuosity) and categorical attributes (e.g., start event). In addition, these events may be associated with other types of information such as a video clip.

In the next step (detailed in Section 5), we use a knowledge-assisted, event ranking framework to determine a set of sort keys and a ranking function that expresses the user's ordering requirement. Tacit knowledge is stored in the form of the user's implicit ranking of events. We use regression analysis to develop a model that represents the tacit ordering as formal parameters. As the ordering specification may be based on prior-knowledge, the ranking function can be customized to many forms such as importance, or other types of performance metrics. We refer to this as *knowledge discovery*. The goal is to formally describe the user's ranking criteria as a function which can be explicitly written and transferred to other users. Optionally, the user may refine the model using partial knowledge by adjusting weights on various sort keys that affects their ranking.

In the third step, we visualize the set of sortable attributes and their weights in order to provide visual feedback to the user. Since there are many possible types of analytical algorithms, we use visual analysis as a method for choosing and optimizing the model (see Section 5.3). The visualization which is to be detailed in Section 6.1, enables analysts to understand how the data is sorted and its impact on sorting various events. We refer to this process as *knowledge externalization*. This allows the domain expert to incorporate their knowledge and experience into the analysis.

In the final step, we demonstrate an example of *knowledge application* through sorting of video events. We focus on two aspects: 1) using glyph-based visualization as an interface to interactively select events that need to be sorted, and 2) the ability to playback match videos based on the sorting criteria (see Section 6.3).

5 Knowledge-Assisted Event Ranking

The knowledge-assisted, event ranking framework involves defining a relationship between the user's sorting criteria and the set of sort keys (i.e., data dimensions) of the input data. Fig. 2 provides an overview of the process. In order to model tacit knowledge for sorting, we store the user's initial specification based on their ranking of events. Let $e_1, e_2, \ldots, e_n$ be events and $e_{i,j}$ be its j-th attribute value. The user may wish to sort the n events with implicit criteria by choosing a subset of events $n_s \leq n$, and placing them in order $e_{s_1} < e_{s_2} < \ldots < e_{s_n}$. Here, both the selection and ranking of events is used to express the users' tacit knowledge. Now assume we have m attribute axes. The objective is to determine a ranking function $f(x_1, x_2, \ldots, x_m)$ with the set of normalized attributes $x_j \in [0,1]$, such that the above order is maintained. The function can take many analytical forms. The simplest case is to choose a function as a weighted additive model of each parameter:

$$f(x_1, x_2, \ldots, x_m) = \frac{1}{\sum w_i} \sum_{i=1}^{m} w_i x_i \tag{1}$$

where the weights $w_i \in [0,1]$ are configured by the user. For simplicity, we shall denote $f(x_1, x_2, \ldots, x_m)$ with the term $f(e)$. One can then sample the weights at d intervals, for example 11 intervals as $[0, 0.1, 0.2, \ldots, 1]$ in a naive search for optimizing the solution. There will be $m \times d$ combinations. Given the large combination of parameter weightings, requiring a user to optimize such a function is not feasible. In addition, an implicit sorting criteria can be dynamic, and therefore it is often difficult for a user to quantify precisely the influence each sort key has on their ordering.

One effective approach for modeling the relationship of such a function is through multiple regression analysis. In the context of multivariate sorting, the aim of regression analysis is to model the user's ranking value y of each event, in terms of the independent attribute vector $\mathbf{x}$ under some functional kernel. Examples of kernels include linear and non-linear models [1]. Since the system has no predefined knowledge, we cannot assume which model is most appropriate. To cover a set of possible models, we deploy three main regression techniques: multiple linear regression, polynomial regression and logistic regression.

5.1 Regression Techniques

Linear regression is an established technique for modelling statistical data [1] using linear predictor functions, specified in the form:

$$f(e_i) = \beta_0 + \beta_1 e_{i,1} + \beta_2 e_{i,2} + \ldots + \beta_m e_{i,m} + \varepsilon_i \tag{2}$$

where the coefficients β_j for $j = 1, \ldots, m$ are the unknown contribution of each attribute (or regressor) $e_{i,j}$, β_0 is the curve intercept and the error term ε_i measures the deviations from the model. Suppose we let $y_i = f(e_i)$ be the expected event ranking outcome as determined by the user. Since $e_{i,j}$ are known parameters, a solution to the above system can be approximated using a least squares fitting [1]. This is often denoted in matrix form by $\mathbf{Y} = \mathbf{E}\beta + \varepsilon$, such that:

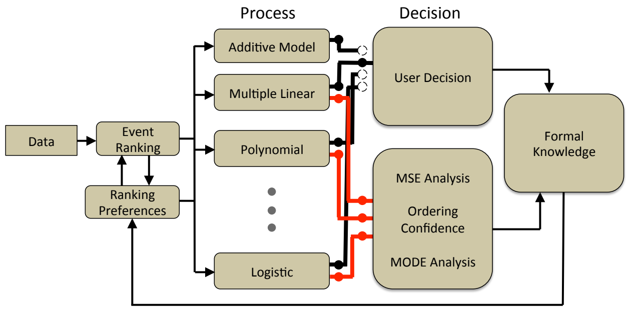


Fig. 2. A graphical pipeline for modeling tacit and partial knowledge as formal parameters using regression analysis. The black path indicates an under-determined event ranking specification. Here, the user chooses the most appropriate model based on visual feedback of the sort keys and weights (see Fig. 4 for an example). For an over-determined system (red path), the best model is selected automatically using regression evaluation metrics.

$$\mathbf{Y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},\quad \mathbf{E}=\begin{pmatrix}1 & e_{1,1} & e_{1,2} & \dots & e_{1,m}\\1 & e_{2,1} & e_{2,2} & \dots & e_{2,m}\\\vdots & \vdots & & \vdots & \\1 & e_{n,1} & e_{n,2} & \dots & e_{n,m}\end{pmatrix}$$

$$\beta=\begin{pmatrix}\beta_0\\\beta_1\\\vdots\\\beta_m\end{pmatrix},\quad\text{and}\quad \varepsilon=\begin{pmatrix}\varepsilon_1\\\varepsilon_2\\\vdots\\\varepsilon_n\end{pmatrix} \tag{3}$$

Hence, the least squares estimate β can be generalized to:

$$\hat{\beta}=(\mathbf{E^TE})^{-1}\mathbf{E^TY} \tag{4}$$

It is clear in Eq. 4 that a solution to β exists as long as the matrix $\mathbf{E^TE}$ is invertible. Algebraically, the matrix is non-invertible given any fixed regressor i.e., $e_{i,j}=c_j$ for all $i\le n$ for some constant c_j. To resolve this, we simply remove the regressor from the model, since any constant attribute will not impact the event ranking $e_{s_1}<\ldots<e_{s_n}$.

Regression analysis typically relies on an *over-determined* system of equations $\mathbf{E}$ (i.e., for $n>m$) to estimate the unknown parameters. The solution obtained from linear least squares are then optimal. Conversely, $\mathbf{E}$ is *under-determined* if there are more unknowns than equations. Generally, such a system may have infinitely many or no solutions. We can pick one of these solutions by finding the smallest one such that $\hat{\beta}$ is minimized subject to the constraint $\mathbf{Y}=\mathbf{E}\beta$. This can then be solved using the method of Lagrange multipliers:

$$\hat{\beta}=\mathbf{E^T}(\mathbf{EE^T})^{-1}\mathbf{Y} \tag{5}$$

If the matrix is square (i.e., $n=m+1$), then there exists only one unique solution. This becomes an important aspect when it comes to evaluating regression models which we discuss in Section 5.2.

Polynomial Regression is a form of linear regression for fitting a non-linear model to a set of data [27]. Such a model is ideal in many settings where a linear relationship does not hold. For multivariate data, the general polynomial regression model is defined by:

$$\begin{aligned}f(e_i)= \;& \beta_0+\beta_{1,1}e_{i,1}+\ldots+\beta_{1,m}e_{i,m}+\ldots+\\ & \beta_{k,1}e^k_{i,1}+\ldots+\beta_{k,m}e^k_{i,m}+\varepsilon_i\end{aligned} \tag{6}$$

where k is the degree of the polynomial. In our work, we restrict the degree of the polynomial to $k\ge 2$ which can be configured by the user. We then estimate the model parameters using Eq. 4 and Eq. 5.

Logistic regression is a type of regression analysis that is very well suited when data consists of categorical data and/or continuous variables [27]. The model estimates the data using a sigmoid function. There are various types of functions that produce sigmoid curves, but the most common has the form:

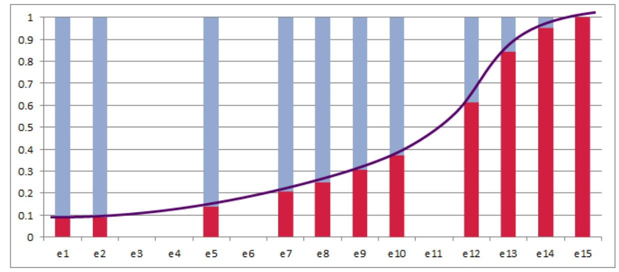


Fig. 3. Applying a Gaussian moderation as a weighting parameter for determining the event ordering confidence. The blue values depict events that are ordered successfully. The adjusted values are shown in red. This emphasizes the importance of ordering higher ranked events, and de-emphasizes the ordering of lower ranked events.

$$f(e_i)=\frac{1}{1+\exp\left(-\beta_0-\sum_{j=1}^{m}\beta_j e_{i,j}\right)}+\varepsilon_i \tag{7}$$

Geometrically, the sigmoid kernel is bounded between $(0,1)$. If we apply a logistic transformation to each side of the equation, we get:

$$logit(f(e_i))=\log\left(\frac{f(e_i)}{1-f(e_i)}\right)=\beta_0+\sum_{j=1}^{m}\beta_k e_{i,j} \tag{8}$$

which describes $logit(f(e_i))$ as a linear relationship of the event attributes which can then be solved using the methodology above.

5.2 Regression Evaluation

As part of the knowledge discovery step, we evaluate the quality of the regression model for defining formal knowledge. In statistical modeling, a wealth of different criteria (e.g., RMSE, RSE and MAE) has been used to evaluate and compare models. The most common comparison test is by Mean Squared Error (MSE), which quantifies the difference between the estimated rank and the true value of the quantity being estimated [1, 27]. This is described by the following:

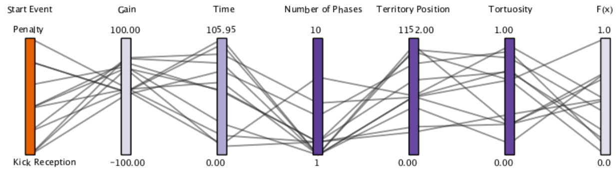
$$MSE=\frac{1}{n-dof-1}\sum_{i=1}^{n}(\hat{y}_i-y_i)^2 \tag{9}$$

where n is the number of events and dof is the degrees of freedom. We note that the term $(\hat{y}-y)$ is equivalent to the error ε in the regression models we described in Section 5. The MSE incorporates both the variance of the estimator and its bias, which is effective for evaluating the accuracy of the fitted model. In order to compare different models by MSE, the unit of error in each model must be the same. This is important when comparing bounded (e.g., logistic regression) versus unbounded models (e.g., linear regression). We address this challenge by scaling the user-defined event ranking values y_i, to the co-domain of the bounded function. If we were to solely rely on MSE to determine the best model, this formal knowledge may not be optimal, since the criterion does not include any information on the prospective ordering of events ranked by the sorting function. Hence, to accurately obtain the best model we introduce two novel comparison metrics: an ordering confidence σ, and a Mean Ordering Distance Error (MODE).

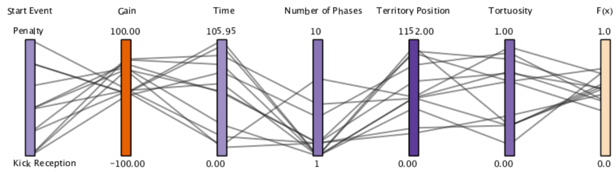
Since the goal is to preserve the ordering $e_{s_1}<e_{s_2}<\ldots<e_{s_n}$, we introduce the confidence metric σ that measures the accuracy in which the order is maintained. Let $\phi:\mathbb{R}^2\mapsto\{0,1\}$ be a binary function that returns 1 if an ordering is maintained such that $f(e_{s_i})<f(e_{s_{i+1}})$ for all $i=1,\ldots,n-1$. We derive the ordering confidence as:

$$\sigma=\frac{1}{n-1}\sum_{i=1}^{n-1}\phi(f(e_i),f(e_{i+1})) \tag{10}$$

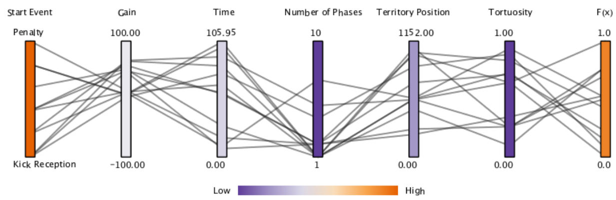
For extensively large event subsets n_s, we find that a user's tacit knowledge for organizing events (i.e., the accuracy of ranking) do not scale up well. In general, important or higher ranked events can be more easily ordered in comparison to ones that are less significant



(a) Linear Regression



(b) Polynomial Regression



(c) Logistic Regression

Fig. 4. Visual comparison of (a) Linear, (b) Polynomial and (c) Logistic regression model for estimating tacit sorting using parallel coordinates. The user's partial knowledge w_j, and weight of each sort key (i.e., the model coefficients $||w_j \beta_j||$) of the ranking function $F(X)$ is mapped to axis width and color respectively. We color map the $F(X)$ axis based on the resulting ordering confidence σ.

(i.e., the lower ranked events). The quality of the formal knowledge should therefore take into account this bias. We incorporate this into our confidence metric by applying a gaussian moderation as a weighting parameter to the initial confidence. Fig. 3 illustrates this process. We can see that the unsuccessful ranking of events e_3, e_4, e_6 and e_{11} have significantly less impact to the overall quality of the model after moderation. As a result, the model would now be considered as a viable option to the user.

The third comparison metric we use is the Mean Ordering Distance Error (MODE) to describe the volatility of the estimated outcome in each regression model. Let us consider the ordered event subset $A = \{e_{s_1}, e_{s_2}, \ldots, e_{s_n}\} = \{e_{s_i}\}$ with total order. Each event is ranked according to its i^{th} index position (i.e., $e_{s_1} = 1, e_{s_2} = 2$, etc.), and have corresponding model estimates $f(A) = \{y_i\}$ for all $i = 1, \ldots, n$. By sorting the values $f(A)$ in ascending order, we obtain a list $sort(f(A)) = \{y_{t_i}\}$, where t_i denotes the sorted position of e_{s_i} in context with formal knowledge $f(e)$. We calculate the MODE between the events expected tacit position i ranked by the user and its new sorted position t_i in the model by:

$$MODE = \frac{1}{n} \sum_{i=1}^{n} ||i - t_i|| \tag{11}$$

such that if $i - t_i = 0$, then the ordering is maintained. The MODE is useful for describing how much movement (or re-ordering) occurs when events are organized by the defined formal knowledge. An optimal model is given when these two ordering metrics are minimal.

5.3 Model Selection

The selection of regression model for defining formal knowledge largely depends on the user's input of event ranking. Fig. 2 illustrates two possible case scenarios. If the system is over-determined, we perform an analysis of the comparison metrics to automatically select the optimal model. Let $m_1, m_2 \in F = \{F_{linear}, F_{polynomial}, F_{logistic}\}$ be two models in a class of regression models F, with least MSE. We determine the optimal regression model using the algorithm outlined in Fig. 5 as a series of *if*-conditions that compares the ratio $\delta_{MSE}, \delta_{\sigma}$ and

Algorithm deriveOptimalRegressionModel(*RegressionModels reg*[])

Model m_1, m_2
$getModelByMinMSE(reg, m_1, m_2)$
$\delta_{MSE} = \frac{\min(MSE(m_1), MSE(m_2))}{\max(MSE(m_1), MSE(m_2))}$
$\delta_{\sigma} = \frac{\min(\sigma(m_1), \sigma(m_2))}{\max(\sigma(m_1), \sigma(m_2))}$
$\delta_{MODE} = \frac{\min((MODE(m_1), MODE(m_2))}{\max(MODE(m_1), MODE(m_2))}$
if ($\delta_{MSE} < T_{MSE}$) **then**
  return $\min_{MSE}(m_1, m_2)$
else
  if ($\delta_{\sigma} < T_{\sigma}$) **then**
    return $\max_{\sigma}(m_1, m_2)$
  end if
else
  if ($\delta_{MODE} < T_{MODE}$) **then**
    return $\min_{MODE}(m_1, m_2)$
  end if
end if

Fig. 5. Algorithm for automatically choosing the optimal model.

δ_{MODE} of each error with threshold values $T_{MSE}, T_{\sigma}, T_{MODE} \in [0, 1]$. The ratio test measures how significantly different each model is under the three comparison methods. To produce the results in this paper, we set a threshold of $T = 0.7$ for each metric. However, this can be customized according to the user's preference.

In the second case, the system is under-determined. When the system is under-determined, the MSE is not derivable since the number of chosen events minus one (observations) is less than the degrees of freedom (see Eq. 9). Statistically, we are unable to measure how trustworthy the derived formal knowledge is. Instead, we provide visual feedback of the ordering confidence (see Section 6.1 and Fig. 4) and allow the user to interactively inspect and determine the most appropriate model. This approach has been shown to be effective in many other visual systems [11, 20]. For the example shown in Fig. 4, a logistic regression model would be the most preferred option.

5.4 User Interaction

Some times, the solution to the system can be trivial, e.g., if events happen to have been ranked in a temporal sequence, then the user would suspect time to be one primary sort key. However, users can often make intuitive or educated guesses on specific sort keys that may or may not affect their ranking criteria. We liken this to *partial knowledge*. To facilitate this in our visual analytic system, we allow the user to interactively tune the model parameters by applying weights $w_j \in [0, 1]$ to the unknown contributions β such that:

$$y_i = F(\mathbf{w}, \beta, e_i) \tag{12}$$

By de-emphasizing a specific attribute axis, the user can analyze alternate sorting attributes and how they impact the ordering of events. Optionally, users can choose to remove a sort key completely ($w_j = 0$). We incorporate this function into our system as a series of interactive sliders which the user can adjust (see Fig. 8 for example).

6 Visual Mapping and Interaction

We have described the process of converting tacit and partial knowledge into formal parameters of a ranking function. To illustrate the associated model (i.e., knowledge externalization), we focus on three important aspects: 1) visualization of sort keys and their various weights, 2) informing the quality of formal knowledge through visualization, and 3) the interaction and visualization of sorted rugby events.

6.1 Visual Mapping

In order to visually convey the model parameters of the ranking function, we adopt the use of parallel coordinates which is proven to be effective in multivariate analysis [13]. Each attribute dimension is

Fig. 6. Screenshot of the video playback of sorted events. We integrate four different broadcasting feeds that correspond to the event.

plotted as vertical axes and the events are drawn as polylines. To illustrate the contribution of each attribute, we color map the axes according to the magnitude of the weighted model parameters $||w_j \beta_j||$ for $j > 0$ (see Fig. 4). We use a diverging color scheme chosen from Color Brewer [9] to emphasize the attributes that are least and most influential. The ranking function $F(X)$ is plotted as an additional axis and color mapped to depict the ordering confidence σ from the model. This enables the user to assess the quality of the formal knowledge in a visual manner. We find that parallel coordinates offers a holistic externalization of the formal knowledge which allows the user to comprehend model parameters and its impact on sorting events. This supports the user in making an informed decision on choosing the most appropriate model through visual feedback.

6.2 Interaction and Glyph-based Visualization

Glyph-based visualization is an effective tool for representing multivariate data [30]. Glyphs are graphical entities that convey one or more data values using visual features such as size, shape and color. We take advantage of the recent work by Legg *et al.* [19], who demonstrate the usability of glyphs in rugby. We position the glyphs along two primary axes (see Fig. 7(c)). Although interactive multivariate sorting is the focus of this work, we are careful not to confuse the end-user with an unfamiliar visual design. To facilitate this, we adopt their glyph [19] to encode our event properties (see Section 3.2) as shown in Fig. 7(d). The glyphs highlighted in purple within the glyph-based canvas indicate events that resulted to a point scored. Other visual design choices (e.g., Chernoff Faces [5] and Star Glyphs [28]) may be used depending on its application context. Due to the inherent occlusion of using large glyphs [30], we provide interactive sliders that enables the user to adjust the length of the sorting axes. This can significantly reduce the amount of visual clutter. The user can then select the events (i.e., the glyphs) and import them into a ranking table. The table view provides an interface where the user can specify the event ordering by drag-and-drop. We found glyphs to be an intuitive mechanism for selecting and ranking events. This is due to similarity to our card metaphor.

6.3 Sorted Event Replay

Sporting analysts often rely on making semantic observations that can only be gained through studying video in order to determine the importance of an event and its event ordering. To support the transfer of tacit knowledge, our visual analysis system facilitates the inspection of key events by brushing the sorted results within the parallel coordinate or glyph-based view. Since the data are associated with single or multiple video clips, we incorporate a video playback user-option for viewing the sorted events (see Fig. 6). The playback of ordered video clips enable users to choose, view, and rank the events in a much more effective manner than the results of a typical search query.

7 Results

Fig. 7 presents our visual analysis system analyzing a rugby match. The system contains four main views: (a) the parallel coordinate view for depicting the ranking function based on the the example ordering specification shown in table view (b). The table interface allows the user to configure, or modify their event ranking. Subsequently, we update the model according to the users' ranking and preference. The resulting model is then conveyed by the parallel coordinate view. In (c), we depict the sorted events using glyph-based visualization. We control the primary sorting axes within this view by clicking on the corresponding glyph component in the graphical interface (d). In order to sort the glyphs using formal knowledge, a drop down option menu is available for users to export the current ranking function to one of the X or Y axes. For this example, the X axis corresponds to the ranking function $F(X)$. This allows the analyst to explore how all events in the match are reorganized according to the their derived sorting criteria.

There are many strategies towards using the system. The most generalized approach starts by choosing the glyphs (or events) of interest in the glyph-based visualization which the user wishes to order. Different layout methods can be used by changing the primary axes to enhance the searchability of such glyphs. These are imported into the ranking table view where the user can specify the event rank in a top-to-bottom (best-to-worst) order. This defines the event ranking which is parsed as a parameter to our knowledge-assisted, event ranking framework. The user can then visually assess the quality of the ranking function in the parallel coordinate view. In some cases, this may require several refinement phases by tuning the attribute weights (see Fig. 8) in order to achieve an optimal solution.

Each of the views are linked such that the user can interactively explore and brush the data in an intuitive manner. Selecting events in the table view will highlight the polyline red within the parallel coordinates, and render the corresponding glyph in focus (see Fig. 7). This enables the user to closely inspect and understand the relationship between the events, the event ordering, and the derived sorting function. Similarly, we provide brushing in both the parallel coordinate view and glyph-based visualization. The focus+context interaction offers a convenient method for analysts to choose and play back the sorted video clips associated with each event for further analysis.

8 Evaluation

To evaluate the work of this paper, we organized a consultation session consisting of 5 participants (3 computer scientists and 2 sport scientists). Each participant had reasonable knowledge of both rugby and visualization. The question we propose is, *"What are the most important positive outcomes?"*. For this example, importance is the tacit knowledge we are trying to formalize, and positive outcome are the measurable parameters. We consider one positive outcome in rugby as an event that leads to points scored, and if a penalty or free kick is awarded. The same rugby match was used throughout the study. During each session, the users were presented with several tasks outlined in Fig. 9. With each task, we provide the additional optional meta-answers that describe the following: (a) I am reasonably confident about my answer, (b) I am unsure about my answer and (c) I do not know how to do this. The user is required to give an answer for only (a) and (b). This would later help us analyze whether participants were merely guessing. For tasks 1 and 2, users were able to use a basic version of the system that consists of the three views (b), (c), and (d) along with a video playback functionality as shown in Fig. 7. We consider this as our benchmark for analyzing and sorting rugby events.

In task 1, we asked the users to identify and rank five events based on importance. The participants generally felt confident with their choices. We anticipated this bias to have some affect in the following task for ranking ten events. Instead, we noticed that users became less confident of their answers. Looking at the event choices closely, there were only two instances of which more than one participant ranked equivalently. These are rank 10 (best) event, and rank 5 respectively. The evidence here supports our hypothesis and application of a gaussian-moderated ordering confidence σ. On average, task 1 took longer by 5.1 minutes, since users had to study all 12 possible video events for the first time. The most time-consuming process came from the random approach towards selecting events. Upon completion of task 1, users could formulate some implicit strategy for identifying more important events. For example, one observed method was

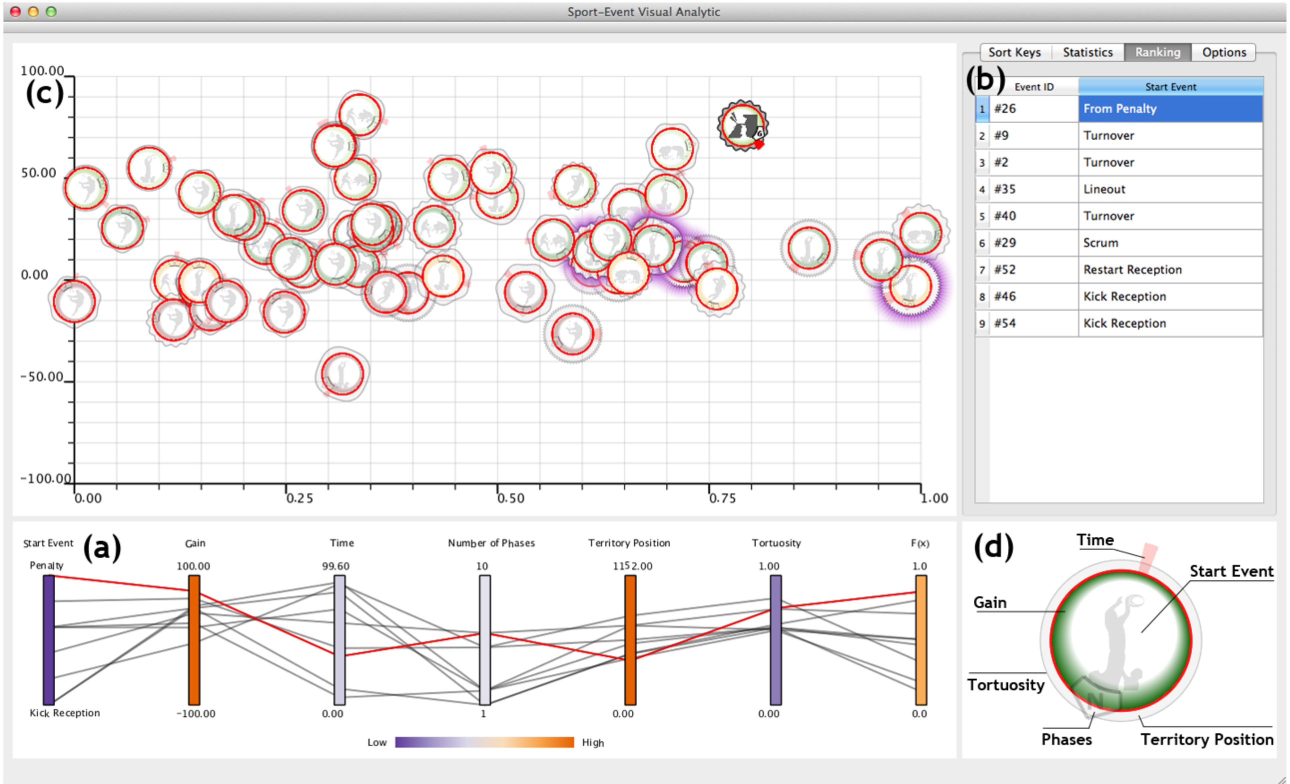


Fig. 7. A visual analysis sorting system used for multivariate sorting of rugby event data. It contains four main views: (a) is the parallel coordinate view of the ranking function. This allows the user to see the composition of the ranking function, and the accuracy in which the ordering of the event subset shown in table (b) is maintained. The user can adjust or refine the event ranking within the table view. (c) displays the sorted results using glyph-based visualization and (d) is an glyph-based interface for selecting the primary axes in (c).

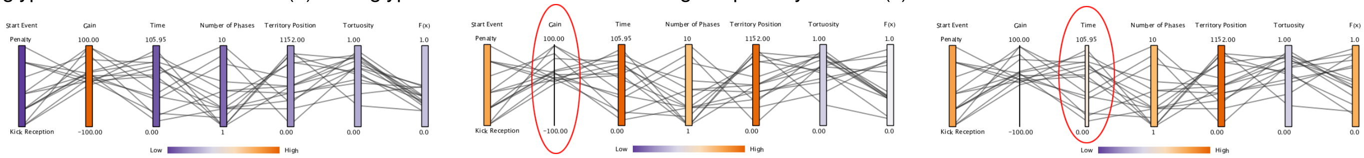


Fig. 8. Refining the knowledge discovery process (from left to right) using partial knowledge. The initial model (left) starts with a low ordering confidence indicated by the blue $F(X)$ axis. After several iterations of refinement (centre and right), one solution is obtained that accurately preserves the user's ordering since $F(X)$ is colored orange. The adjusted sort key weights (highlighted with a red circle) correspond to the values (centre) $w_{gain} = 0$ and (right) $w_{time} = 0.50$ respectively.

to choose events based on highest gain. This made the second task much faster and simpler to do. Even so, we notice a diverse variation in event ranking as illustrated in Fig. 9. We conclude from both tasks that each participant had different interpretations of importance.

Feedback from the third task proved to be very interesting. A number of participants understood their sorting strategy used in the previous tasks did not easily generalize to importance (e.g., the most important event is not determined solely by the largest gain). All users recognized the multitude and combination of factors that potentially affected their importance criteria. However, the participants could speculate several attributes that are most significant. Tortuosity and number of phases are two attributes that were often coupled together. The amount of gain is also considered influential. It shows that there are some tacit similarities when quantifying importance. Furthermore, we note that none of the participants chose *time* as a sortable attribute, which supports the use of partial knowledge for narrowing the search space. It is also clear that defining importance is multi-dimensional in nature, and that ranking rugby events by importance would require a tool for multivariate sorting. The final task involved formalizing their ranking function. Whilst most participants attempted to write their strategy, they agreed that this is too complex to do. The feedback from the study demonstrate the need for formalizing a user's tacit knowledge in order to perform more effective sorting of rugby data. It also reveals the benefits of using a visual analytic system to allow the user to understand, refine, and incorporate their knowledge within the analysis.

9 Discussion

Among several modeling techniques, we chose three types of regression methods to formalize a user's tacit and partial knowledge. A fundamental step in many machine learning methods involves training the model. Based on the study in Section 8, we demonstrate training off a relatively small sample (5-10 events) which makes it difficult to judge how accurate and robust the derived formal knowledge is. Clearly, a more extensive training and validation process of multiple matches would yield a better sorting function. With enough training, the formal knowledge can then be stored externally as a template metric.

We have demonstrated our system for sorting rugby events. The framework can easily be applied to other sports (e.g., football, basketball and tennis) since both regression analysis and parallel coordinates are generalizable. For our glyph-based visualization, a multivariate glyph such as Star Glyphs [28] can be used instead. However, one limitation of such designs is that a greater learning process may be required in comparison to domain specific glyphs that are semantically richer [19]. Extensions to higher dimensionality is also scalable, giving more (and possibly more accurate) solutions to the user's sorting criteria. A potential issue here is the increase in computational cost can affect the interactivity of the system for very large dimensions.

Task	(optional) meta-answer	Result											
1. Identify and rank 5 events from best-to-worst	(a)	3		1					5	4	2		
	(a)		4	5		2		3	1				
	(b)		2					4	3	1		5	
	(a)		4					1	5	2			3
	(b)			1					5	3		4	2
2. Identify and rank 10 events from best-to-worst	(a)	4	8	2		1		7	10	9	3	6	5
	(b)	10	7	8		5	4	6	1	2	9	3	
	(b)	2	7		10	1	4	9	3	5		6	8
	(b)	7	8	2	3			4	10	5	1	9	6
	(b)	2		3	6	4	8		10	5	1	9	7
3. Identify a set of attributes that may affect the ranking	(a)	Gain (high), Tortuosity (low), Number of Phases (low)											
	(b)	(Tortuosity + Number of Phases), (Gain + Territory Position)											
	(b)	Tortuosity, Number of Phases, Start Event											
	(a)	Gain, Start Event, Number of Phases											
	(a)	Gain, Number of Phases											
4. Formulate a ranking based on the set of attributes	(c)	N/A											
	(c)	N/A											
	(c)	N/A											
	(a)	Combination of high gain, low tortuosity and a weighted start event (e.g., turnover is more important than scrum)											
	(b)	Sequences containing high gain or high number of phases from various start events											

Fig. 9. Table showing the consultation session results for sorting rugby events. Each sub-row within the four primary tasks correspond to five participants along with their optional meta-answer (see Section 8 for details). For task 1 and 2, each cell in the result column indicate 12 possible events. These are labelled and color-mapped from worst-to-best with 1-5 and 1-10 respectively according to the user's ranking criteria.

10 Conclusion

We proposed a knowledge-assisted, event ranking framework for interactive multivariate sorting of sport event data. Users can choose and rank an event subset in order to define a ranking function using regression analysis as a sortable criteria. Tacit knowledge by domain experts can be incorporated into the ordering rule. This is reflected within the ranking function, meaning it can take on many complex forms such as importance. Such metrics are often multivariate in nature and are difficult to formally encode. Alternatively, the framework can be used to query data patterns by investigating how one can sort a set of data in a respective order. We have found that our visual analytic approach significantly enhances the usability of multivariate sorting.

References

[1] F. S. Acton. *Analysis of Straight-line data*. Wiley, 1966.

[2] I. Borg and T. Staufenbiel. Performance of snow flakes, suns, and factorial suns in the graphical representation of multivariate data. *Multivariate Behavioral Research*, 27(1):43–55, 1992.

[3] E. Brown, L. Jingjing, C. Brodley, and R. Chang. Dis-function: Learning distance functions interactively. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 83–92, 2012.

[4] R. Chang, C. Ziemkiewicz, R. Pyzh, J. Kielman, and W. Ribarsky. Learning-based evaluation of visual analytic systems. In *Proceedings of the 3rd BELIV'10 Workshop: BEyond time and errors: novel evaluation methods for Information Visualization*, BELIV '10, pages 29–34, 2010.

[5] H. Chernoff. Using faces to represent points in *k*-dimensional space graphically. *Journal of the American Statistical Association*, 68:361–368, 1973.

[6] T. Cox and M. Cox. *Multidimensional Scaling*. Chapman and Hall, 2001.

[7] A. Doubleday. Use of card sorting for online course site organization within an integrated science curriculum. *Journal of Usability Studies*, 8(2):41–54, 2013.

[8] A. Endert, P. Fiaux, and C. North. Semantic interaction for sensemaking: Inferring analytical reasoning for model steering. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2879–2888, 2012.

[9] M. Harrower and C. A. Brewer. *ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps, in The Map Reader: Theories of Mapping Practice and Cartographic Representation*. Wiley-Blackwell, 2011.

[10] F. Heimerl, S. Koch, H. Bosch, and T. Ertl. Visual classifier training for text document retrieval. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2839–2848, 2012.

[11] B. Hoferlin, R. Netzel, M. Hoferlin, D. Weiskopf, and G. Heidemann. Inter-active learning of ad-hoc classifiers for video visual analytics. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 23–32, 2012.

[12] M. D. Hughes and I. M. Franks. *Notational analysis of sport*. London: E. & F.N. Spon., 1997.

[13] A. Inselberg. The plane with parallel coordinates. *The Visual Computer*, 1:69–91, 1985.

[14] D. H. Jeong, C. Ziemkiewicz, B. Fisher, W. Ribarsky, and R. Chang. ipca: an interactive system for pca-based visual analytics. In *IEEE VGTC conference on Visualization*, pages 767–774, 2009.

[15] L. Jin and D. C. Banks. Visualizing a tennis match. In *Proceedings of the 1996 IEEE Symposium on Information Visualization (INFOVIS '96)*, INFOVIS '96, pages 108–, Washington, DC, USA, 1996. IEEE Computer Society.

[16] I. T. Jolliffe. *Principal Component Analysis*. Springer, second edition, 2002.

[17] A. Klippel, F. Hardisty, R. Li, and C. Weaver. Colour-enhanced star plot glyphs: Can salient shape characteristics be overcome? *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44(3):217–231, 2009.

[18] D. E. Knuth. *The Art of Computer Programming, Vol. 3: Sorting and Searching, Second Edition*. Addison-Wesley, Reading, Mass., 1998.

[19] P. A. Legg, D. H. S. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths, and M. Chen. Matchpad: Interactive glyph-based visualization for real-time sports performance analysis. *Computer Graphics Forum*, 31(3pt4):1255–1264, 2012.

[20] H. Lipford, F. Stukes, D. Wenwen, M. Hawkins, and R. Chang. Helping users recall their reasoning process. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pages 187–194, 2010.

[21] G. Mistelbauer, A. Kochl, H. Bouzari, S. Bruckner, R. Schernthaner, M. Sramek, I. Baclija, and M. E. Groller. Smart super views: A knowledge-assisted interface for medical visualization. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 163–172, 2012.

[22] A. Moore, P. Whigham, C. Aldridge, A. Holt, and K. Hodge. Rugby: (a) union of space and time. In P. A. Whigham, editor, *Proceedings: Thirteenth Annual Colloquium of the Spatial Information Research Centre*, pages 183–194, 2001.

[23] M. L. Parry, P. A. Legg, D. H. S. Chung, I. W. Griffiths, and M. Chen. Hierarchical event selection for video storyboards with a case study on snooker video visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1747–1756, 2011.

[24] H. Pileggi, C. D. Stolper, J. M. Boyle, and J. T. Stasko. Snapshot: Visualization to propel ice hockey analytics. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2819–2828, 2012.

[25] R. E. Roth, B. G. Finch, J. I. Blanford, A. Klippel, A. C. Robinson, and A. M. MacEachren. Card sorting for cartographic research and practice. *Cartography and Geographic Information ScienceCartography and Geographic Information Science*, 38(2):89–99, 2011.

[26] G. Rugg and P. McGeorge. The sorting techniques: a tutorial paper on card sorts, picture sorts and item sorts. *Expert Systems*, 14(2):80–93, 1997.

[27] A. Sen and M. Srivastava. *Regression Analysis: Theory, Methods, and Applications*. Springer, 1990.

[28] J. Siegel, E. Farrell, R. Goldwyn, and H. Friedman. The surgical implication of physiologic patterns in myocardial infarction shock. *Surgery*, 72:27–35, 1972.

[29] X. Wang, D. H. Jeong, W. Dou, S.-W. Lee, W. Ribarsky, and R. Chang. Knowledge assisted visualization: Defining and applying knowledge conversion processes to a visual analytics system. *Computers and Graphics*, 33(5):616–623, 2009.

[30] M. O. Ward. A taxonomy of glyph placement strategies for multidimensional data visualization. *Information Visualization*, 1(3-4), 2002.

[31] C. Weaver. Multidimensional visual analysis using cross-filtered views. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pages 163–170, 2008.