

Visualizing Running Races Through the Multivariate Time-Series of Multiple Runners

Guilherme Oliveira, João Comba
UFRGS - Brazil

Rafael Torchelsen
UFFS - Brazil

Maristela Padilha
IPA - Brazil

Claudio Silva
NYU-Poly - USA

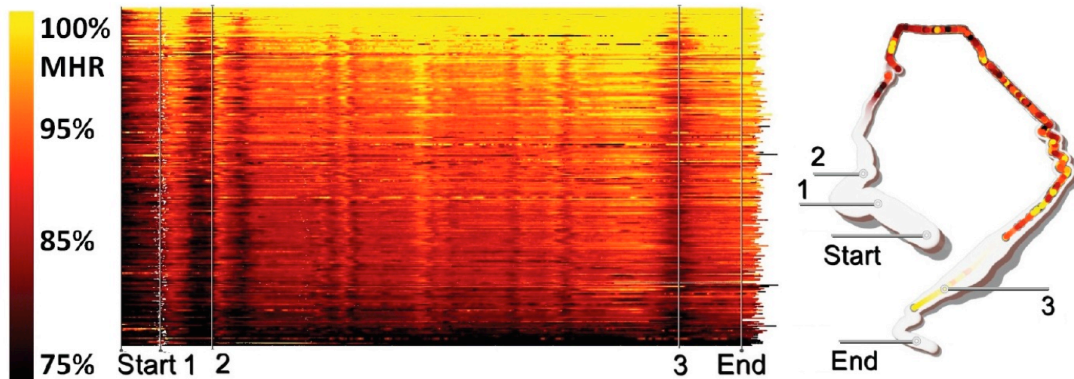


Figura 1. Visualization design used to analyze a 15 km running race composed of activities of multiple runners. We show in the left the linear heatmap design, where the time-series for the heartbeat of each runner is normalized against his maximum heart rate (MHR), and drawn as color-coded particles from left to right in a single line, with the x-axis corresponding to distance. Heatmaps were sorted from top to bottom based on the average effort level to group runners with similar effort levels to make it easier to identify patterns, such as the number of runners exercising close to their MHR. In addition, we placed distance markers to highlight vertical patterns of common variation. On the right, we show the augmented track view of the same race, with the same distance markers to allow us to correlate patterns.

Abstract—The recent widespread of heart rate (HR) monitors is allowing people to measure body response during and after exercise, which produces a collection of time-series on multivariate aspects, such as heartbeat, speed, geolocation, etc. Such monitoring can be extremely important for people with low fitness levels, since they are susceptible to cardiovascular diseases or other physical injuries when exercising at high heartbeat frequencies. Even though most monitors provide tools to export and display this information for each individual, the ability to visualize the collection of multiple runners in a given running race is mostly unexplored. In this work, we present a design study that aims to support analysis and answer several questions raised by an expert on exercise physiology about a given running race. We describe each visualization design and how they individually, or in collaboration, can be used to reveal interesting aspects of the data. We illustrate our results with use cases that provide evaluation and feedback about the visualization designs proposed.

Keywords—Time Series; Races;

I. INTRODUCTION

Physical activity is an essential component for a healthy lifestyle. There are several studies that correlate low fitness levels to high risk of cardiovascular problems [1], [2], [3] and regular physical activity is essential to reduce such risks. Although there is a strict recommendation for a preliminary doctor check-up before starting any period of physical activity, there is no way to enforce this recommendation and it is not uncommon for people to start exercising without visiting a

doctor. This can be dangerous, specially when people with low fitness level engage in intense activities such as running.

HR monitors were introduced in the 70's [4] to help athletes record heart-rate activity during competition which could be used in a subsequent analysis to improve performance. Such devices comprise a HR monitor (incorporated into a wrist receiver) and a chest strap transmitter. Data recorded can be as complex as the time-series containing heartbeat during the entire exercise, as well as other information such as speed, geolocation, etc. The affordability of such devices has made them popular recently. In addition, activity recorded from monitors can be uploaded to computers, where they can be inspected or shared with others (e.g. a person's physician).

Current visualization tools for HR data focus on the visualization of a single activity, and often lack the ability to compare multiple activities. The ability to inspect multiple activities at the same time can be very useful to compare the effort of different runners in a given activity, or to compare the effort of a single person against others in a shared activity. The analysis of this data is challenging since it contains the multivariate time-series data generated by HR monitors for multiple runners. We constrain ourselves to data collected from people in a given running race (e.g. 10 km, marathon, etc). Although we use only data from running races, similar analysis can be performed for other sports, for example, cycling.

The main contribution of this work is a design study comprised of three visualization designs that support the analysis of a given running race, using data from the HR monitors of several competitors. Such data includes heart rate, speed and gps readings only; no further information about the athlete (e.g.: age, weight, height, etc) is available. In collaboration with an expert in exercise physiology, we formulated questions to be answered about the running race. Such questions are detailed in Section III-A and guided the conception of the following visualization designs we will present. To validate our visualization designs, we created use cases with data from different running races, and assessed how helpful they were in answering the questions posed by the expert. Figure 1 gives a preview of the results we obtained for a 15 km running race.

II. RELATED WORK

Below we review related work on time-series visualization and physiology regarding physical fitness and races.

Time-series Visualization

There is a vast literature about techniques for time-series visualization in [5] and [6]. Time-searcher [7] describes how to visualize queries over time-series data to identify interesting patterns. Similar filtering ideas were also described by LiveRAC [8], which uses data re-ordering to improve the visualization. CloudLines [9] describes a technique to visualize multiple time-series, where values are represented by the width of a line. Applying this approach to datasets composed of a few dozen series becomes confusing due to the visualization area required to represent the line thickness. Horizon Graphs [10] reduces the spatial requirement by stacking sections of the area graphs, but it still cannot represent sets of hundreds of series. Braided graphs [11] succeeds at visualizing multiple series in a single frame through sectioning, ordering, and color labels. However, the scheme quickly leads to confusion when the number of series increases. One of our visualization designs is similar to the line graph panel presented in [12], where time-series are straight lines and values are represented by colors. Sequence Surveyor [13] uses a similar linear layout for genome visualization. Hao et al.[14] also describes techniques for visualizing large time-series data using color scales, similar to the concept of Heatmaps[15]. The usual approach for representing multiple variables of a series is to plot a line graph per variable. The TimeWheel and MultiComb presented in [16] follow this approach. However, they were designed to handle a single multivariate time series. Visualization techniques, such as the line graph, are usually adapted to the specific characteristics of the data at hand. In that sense, we adapted the filter lens concept, which is similar to those found in [17] and [18]. Pathline [19] describes a comparative visualization of functional genomics data, which shares with our work the idea of looking at collections of time-series to identify trend patterns. The work of Viau and McGuffin [20] explores the use of different data charts to visualize multidimensional data using visual connections between charts to correlate information; strategy we used for

HR Zone	Effort Index	Effort Level	Pace	Fuel Source
1	60-75%	Easy	Slow	Primarily Fats
2	75-85%	Moderate	Moderate	Carbs and Fats
3	85-95-%	Difficult	Fast	Primarily Carbs
4	95-100%	Very Hard	Sprint	All Carbs

Figure 2. Heart Rate Training Zones. Effort index given as percentage of the maximum heart rate. (from [26])

our distance marks (section III-E2). The track view (section III-D) shares with the work of Tominski et al. [21], the concept of visualizing multidimensional data along a track. However, our design is confined in to a 2D visualization, as opposed to the 3D visualization of the mentioned work. Finally, the work by Legg et al. [22] also produces visualizations for sports data, specifically for performance analysis. However, their focus is on real-time analysis and glyph-based visualization, while this work focuses on post-activity analysis and heart health.

Physical Fitness and Races

There are several studies that correlat fitness levels with well-being [1], [2], [3]. The Physical Activity Guidelines Report [23] presents proofs of the importance of physical fitness. In [3] consistent evidence is given about the direct association of myocardial infarction to physical inactivity and that people with low fitness levels have a higher risk of developing cardiovascular diseases. Running helps improve physical fitness, and there are studies [24], [25] that correlate heart health to the time one takes to run a given distance. Running data, therefore, can provide important indicators of overall fitness. For this analysis, the heartbeat of each individual can be normalized as a percentage of the individual's maximum heart rate (MHR), and different effort levels can be identified using this information. In Figure 2, from [26], four main HR zones are identified. Training programs often rely on defining how long or how far a runner should stay in each HR zone. Exercising at HR zone 4 for a long period can be dangerous, and there is a great concern about mortality and cardiac diseases [27], [28], [29]. Related to this is the study on human limits [30], [31] and strategies in different aspects like hydration [32] and energy consumption [33].

III. VISUALIZATION DESIGNS FOR RUNNING DATA

In this section we describe visualization designs for running data. We first pose the driving questions we discussed with an expert in exercise physiology. Once our goals were defined, we created distinct visualization designs, which are described first individually, and later by their common functionalities.

A. Driving Questions

The conception of the visualization designs was driven towards helping answering questions about a street race using public data from several participants. The questions are usually asked by someone organizing a race or by a physiological specialist. Those have been selected based on the experience of one of the authors, who is a physiological researcher, that

usually searches for the answers without a visualization tool. Below, we enumerate such questions:

- 1) Is there a predominant effort level in the race? Where does the effort level changes?
- 2) Which parts of the race require more effort?
- 3) Are there any common patterns among runners during the race? Can we identify the source of such patterns?
- 4) Can we identify the running strategies for a given race?
- 5) Are there people running at dangerous effort levels?
- 6) How can we compare races?

B. Visualization Design 1: Line Graph Heatmap

The building blocks of the visualization designs are the activities' trackpoints. Trackpoints are samples of the athlete's state collected by the monitor device at regular intervals. Common to the visualization designs is the rendering of consecutive trackpoints using a solid circle (particle). Each trackpoint is rendered only if the current visualization time is within the range of the trackpoint. After rendered, each trackpoint fades according to user parameters.

The first visualization displays multiple time series (runners) similar to a conventional line graph with a set of improvements. The vertical axis represents the effort level, while the horizontal axis represents the distance from the start line, see Figure 3. The color of the trackpoint represents the HR value of the trackpoint in relation to the MHR and is mapped to a color gradient (black, red, orange, and yellow). The color mapping scheme is designed to relate HR to the respective effort zones. HR training corresponds to the use of HR data to customize a given workout to improve a runner's performance [34], [35], [26]. The cardiovascular system reflects the body stress at any given moment, and by keeping track of the HR one can estimate the effort at any given time. Differences in HR measurements, when compared to rest state, can provide immediate feedback on how tired the body is. Those differences indicate how hard the body is working, and how adapted it is to a given workload.

The intensity of the exercise is closely related to the actual HR reading (as a percentage of the MHR) [26]. Based on the HR it is possible to identify the effort level, energy source, and performance-related fitness that will benefit from the activity. Notice from Figure 2 that the range of the different effort zones is not the same. We used layers with the same thickness to lay emphasis on the variation in the zones of higher effort, which is the focus of most of the questions presented before. In other words, the changes in the low effort zone (black layer) become smoother while the change of the visualization in the high effort zone (yellow layer) is more noticeable.

The color of each trackpoint is linearly interpolated to a given color-scale based on the layer it belongs. Each layer keeps track of how many runners are inside its effort zone at any given time. This is visually represented by filling a part of the layer area, bottom to top, proportional to the percentage of runners inside it, with a faded version of the layer color (see Figure 3). This allows one to keep track of a runner and see if his condition is similar to others. Furthermore, we can

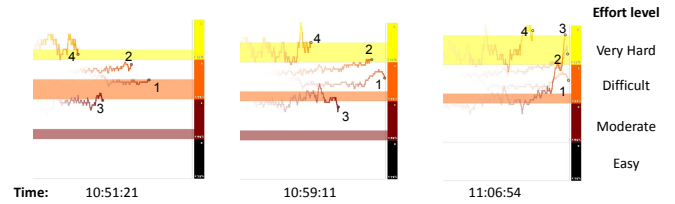


Figure 3. **Line Graph Heatmap** Consecutive frames of the line graph heatmap showing how the number of runners in each effort zone is represented. In the four vertical layers, from yellow to black, we display the different effort levels. From left to right we display the HR for 4 runners (illustrated as points 1 to 4) as function of distance. We used horizontal strips in each effort layer with varying height to encode the number of runners in a given time at that effort level. Such design allows to verify how many runners are in a given effort zone, and how the runners relate to the others.

see the distribution of effort at different times during the race (Q1). One problem of the line graph heatmap is information overlap. To highlight the actual state of each runner, we add a border to the current trackpoint. To reduce the clutter that comes from the overlap with past trackpoints, an adjustable decay factor provides a tradeoff between history length and readability that can be modified during visualization.

C. Visualization Design 2: Linear Heatmap

The second visualization design is called Linear Heatmap since it represents each runner's activity as a horizontal line with colors to indicate the effort level and horizontal space to represent distance covered. The prior design, the line graphs heatmap, is useful to see the overall effort level at different times, but not at different places along the course. If the decay was reduced until the trails become the whole line graphs, such comparison would be possible, however, the overlap of lines would still persist and complicate such task. The purpose of this new design is to provide a way to compare the effort and speed of the runners, while avoiding data overlap. The trackpoints are positioned from left to right according to the distance covered and rendered when the animation time surpass the trackpoint's key time. The result is a set of horizontal lines starting as dots in the left and increasing in length to the right as the runners get closer to the goal.

We also used a different colormap to represent the speed in the trackpoints. The datasets usually contain outliers (measurement errors in the monitor device) in the speed values way above the average speed. In this case, normalizing the values based on the minimum and maximum would bias the result, bringing almost every trackpoint to the same small portion of the color range. To solve this issue, we use a blue-white-red colormap, where white is the average speed in the whole dataset, red marks the trackpoints that are N times the standard deviation above the average speed, and blue marks the ones that are N times the standard deviation below it, with N being a parameter that can be modified in real time. Without the overlapping of runners, now, we can: identify the predominant effort level in the race, as well as in a given part of the course (now in distance, not in time as in the line graph heatmap) (Q1); find the section of the course that demands more effort

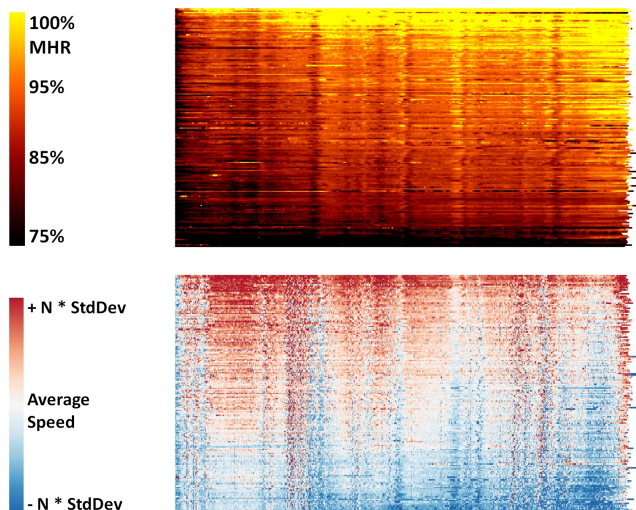


Figure 4. **Linear Heatmap** for the heart rate, measured as percentage of maximum heart rate, (top) and speed (bottom). (Average speed = 10 km/h; Std Dev = 1.4 km/h; N = 3)

(Q2); identify common patterns of effort variation among the runners and look for an explanation to them in the slope and speed variation (Q3); and use the linear heatmap with the speed color mapping to see the running strategy of the race, i.e. how runners change their speed during the race (Q4) (see Figure 4). One drawback of this design is that neighboring time series' representation will begin to overlap if their number is greater than the vertical resolution of the available display are in pixels. The same happens to the samples of a single runner if there are more of them than the horizontal resolution.

D. Visualization Design 3: Augmented Track View

The third visualization design focuses on the geo-spatial data of the dataset. The main purpose is to view and annotate the race course to understand the patterns and events found in the other views. This design is basically a sketch of the race course with the trackpoints representing the runner's state along the competition (see Figure 5). The trackpoint position (in the visualization) is based on the trackpoint latitude and longitude to form the shape of the race course. This course sketch is created using the geo-spatial data from a single chosen runner, which is usually very similar among runners. Inclination and even altitude itself have great influence in the runners' performance and are usually the main cause for the variation in effort patterns shown in the other visualizations. To represent altitude we encode it using shadows. We draw the full course three times, with different offsets in some direction to simulate shadowing. The offset is proportional to the trackpoint altitude at that geo-spatial location. The three layered layouts with the altitude-based offsetting makes the course look like a surface extruded from the plane and under a directional light source. A problem that may arise from the directional offset is that only the top layer will be visible when the course direction is too close to the light direction.

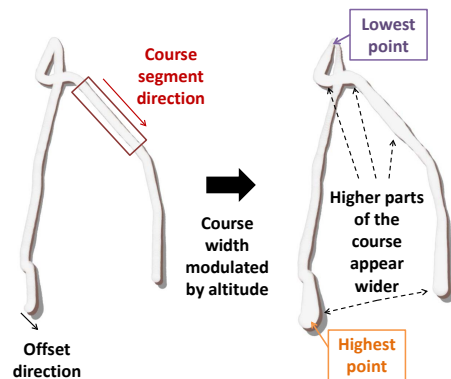


Figure 5. **Extruded course with proportional altitude:** In cases where regions of the course match the offset direction, shadows are occluded by the top layer itself, hiding the slope information. We create a second representation of the altitude by making the course width relative to the altitude at each point, and higher altitude sections become wider.

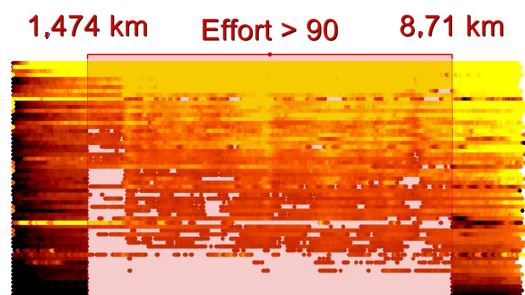


Figure 6. **Filter** to remove trackpoints with effort rate below 90% of MHR.

To overcome this, the altitude is also mapped to the course width as an additional hint, making the higher sections of the course wider, as it would be if a bird eye view of a 3D representation of the course was used (Figure 5). Finally, the runners trackpoints are rendered, on top of the extruded course, with the same motion trail animation of the line graph heatmap. High number of runners may lead to overlap in the track, which makes it difficult to see the changes in effort level over time and at different parts of the course as well. The linear heatmap is more suitable to such purpose.

E. Correlated Features among Visualization Designs

We used common features among the visualization designs. The goal was to correlate information between designs in such a way that a pattern that is only visible in one design can be exposed in another, as shown below.

1) *Filters:* Filters can be applied to the linear and line graphs heatmaps to reduce the visual overload to reveal interesting patterns. By choosing a distance interval the user can customize the visualization to show, in that range, only the trackpoints whose data are situated in a defined span of values. The range and parameters of each filter can be defined at any moment during the visualization and, since the filtering is computed at runtime, the result is immediately visible. Filters can be dragged along the horizontal axis and overlapped to create more complex filtering schemes. The filters are

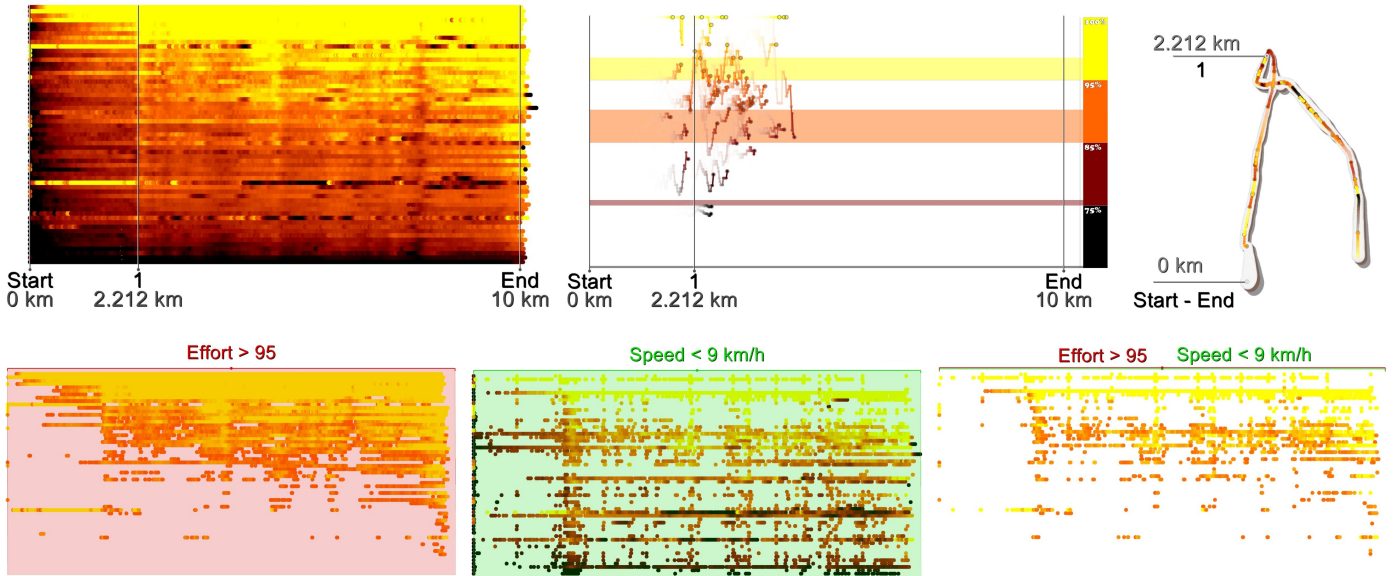


Figure 7. **Adidas Summer Run in São Paulo** Top row: one alignment of increase in effort marked in the linear heatmap; the same position is marked in the line graph heatmap, the thickness of the stripes of the effort zones shows that the majority of runners is above 85% MHR; the same position marked in the track view turns out to be the beginning of a ramp. Bottom row: combination of filters to find trackpoints that indicate a possible risk state.

important to identify behaviors, for example, athletes pushing their HR close to their MHR (see Figure 6). From this we can identify regions of the course where the behavior begins and ends. Based on that, we can redefine the course to avoid stressing too much the majority of the athletes or an athlete can prepare an activity strategy based on past experiences.

2) *Distance markers*: Distance markers are used to correlate patterns found in the different views. They are tools for annotation along the race course, for example, to insert a description label for a section of the race, or to mark a point where there is an increase in effort level. Figure 1 shows an example where distance markers are used to correlate events in the linear heatmap effort view to the location in the track view, thus allowing one to investigate causes of some of the patterns found in the linear heatmap (Q3).

3) *Runner trackers*: The focus of the visualization is the entire dataset (all runners). However, there are situations in which it is interesting to follow a particular runner or make a detailed comparison between performances of two runners. The runner tracker is a tool to provide a summary of a runner's state during the race. The information that can be visualized are the MHR and the variation of effort rate, speed, and altitude. The properties chart keep a small history of the last values of the variable and extreme values. The area chart baseline is defined by the runner's mean value during the race, and it appears upside down when the value is below the average. In the altitude chart the base line is always 0, representing the sea level. We show the altitude itself, since the inclination becomes evident in the altitude area graph. Runner trackers can be used in all the three views (Figure 8 shows a runner tracker in the course view) and always point to the last rendered trackpoint of the runner in the animation.

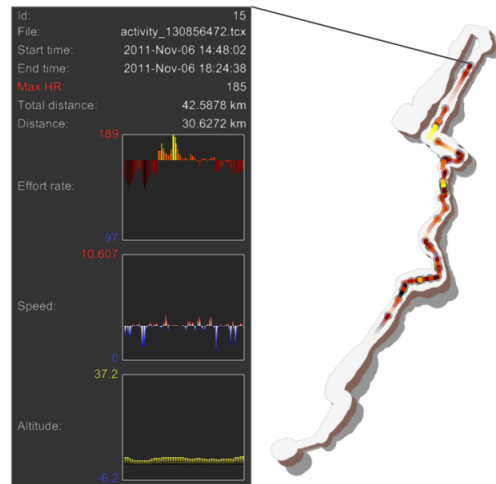


Figure 8. **Runner trackers** Tracker following a runner in the course view, displaying runner id, name of the activity tcx file, MHR, among other info.

IV. DATASETS AND PROTOTYPE CONSIDERATIONS

A prototype was created to validate our design proposals. We describe below considerations on the data we used and details on how we created renderings for each design.

A. Running Data: Garmin Connect

The Garmin Connect website [36] stores a massive amount of public training data, uploaded by users, that can be filtered according to date, total distance, location, etc. Each activity has a time-ordered sequence of trackpoints exported in the Garmin's Training Center XML format (TCX). Trackpoints are samples of a time series, each with a time instant as

key property plus several other variables. The number of variables available depends on the model of the HR monitor. We constrained this range to HR, speed, altitude, latitude, and longitude. Different monitors and other accessories can store other data (e.g. cadence and calories), but their low availability could significantly reduce the size of our datasets.

Each dataset is comprised of a set of TCX files, and each file stores the data of a single runner in the same race. We used datasets of sizes ranging from 60 to 483 runners. The number of trackpoints of each runner in a dataset depends on the monitor sampling rate, and the distance of the race. In the smallest dataset, the number of trackpoints per runner was between 200 and 1000 with an average of 89 trackpoints per minute (1.48 Hz). For the largest dataset, in both distance and number of runners, it ranged from 3000 to 7000 with an average of 218 trackpoints per minute (3.6 Hz).

B. Implementation Details

The prototype import TCX files to be displayed using the visualization designs proposed. The interface allows the user to set the parameters for each design and control how the time-series will be displayed and animated. Common to all designs is the rendering of color-coded particles for each trackpoint. To avoid rendering every trackpoint at each frame, we render to an accumulative texture (called history texture). Trackpoints are rendered only once per animation frame.

For each animation frame we render the background of the actual visual design (horizontal layers for the line graph and course sketch for the course view) followed by the history texture containing the trackpoints, and finally the distance markers, filters, and runner trackers. As a result, we can render a dataset composed of almost 500 runners and more than a million trackpoints (particles) in real-time. To create the fading trail effect for both the line graph and the course view, we render a white rectangle with transparency in the history texture, over the trackpoint, at regular time intervals. The decay factor controls how often the rectangle is rendered and its opacity value. Decreasing the interval among renderings, or increasing opacity will result in a faster trackpoint decay. Due to the order in which the trackpoints are rendered, older trackpoints will fade first, creating the motion trail effect.

V. USE CASES

We downloaded training data from GarminConnect of three different races: 10 km (also called 10K), 15K, and 42K. To allow the comparison of a race in different years, we obtained data for the 15K and 42K from two consecutive years.

The 10K race evaluated was the Adidas Run in São Paulo, Brazil. In Figure 7, we observe a high number of runners running close to their HR limit (third and fourth HR zones). This behavior changes when the race distance increases, since it is harder to sustain higher effort levels for longer time periods. Another interesting result is the aligned increase (or decrease) of effort levels at specific points in the distance axis of the linear heatmap. In Figure 7 we marked one such alignment with a vertical line, and looked for the respective

place along the race course using the augmented track view. We found it to be the start of a steep slope in the lowest part of the course. We also evaluated the results to find patterns of low physical fitness. For this purpose, we used a combination of two filters, one that only accepts trackpoints with speed below 9 km/h, and a second that only accepts trackpoints with effort level above 95% MHR. Such combination of high effort and low speed might be an indication of health hazards. The bottom row of Figure 7 shows, from left to right, the trackpoints filtered only by speed, only by effort, and by both variables. Observe that there are a substantial number of runners that satisfy both filters, a clear sign of alert.

The second race we studied was the São Silvestre race, a 15K race that takes place every year in São Paulo, Brazil, on December 31st. We downloaded two datasets of this race (2010 and 2011). Since the race course changed in 2011, it is possible to verify different patterns when comparing the visualization for the years of 2010 and 2011 (Figure 9). The distance markers in this figure are used to correlate changes in the effort and speed in the linear heatmap to the respective spots in the course view. In comparison to the 10K race studied before, we observe that high effort readings decrease, since runners have to sustain effort for a longer period, and therefore run at a lower intensity level. This use case illustrates the ability to compare different races, as raised in (Q6).

The last use case was the 2010 and 2011 New York City Marathons (42K). These were the largest datasets with 473 runners in 2010 and 483 in 2011, with more than a million trackpoints in each race. In Figure 10 we compare the linear heatmaps for both years. We observe that runners sustain an even smaller effort level throughout the race in comparison to the 10K and 15K races, revealing the running strategies for different race distances (Q4). We also observe that there is a similar pattern in the linear heatmap that is kept nearly unchanged for both 2010 and 2011 years, even though they come from a distinct set of runners. Since the course track was the same, we conclude that it has a great impact in the changes of the heart rate measurements for the runners.

VI. EVALUATION AND FEEDBACK

The visualization designs were driven by questions posed in Section III-A by an expert on exercise physiology. We evaluate below how our designs helped answering these questions:

Q1: Effort levels can change during the race, and becomes important to identify where and why such changes occur. Using the line graph heatmap, we keep track of the number of runners in each effort zone at any time, independent of their location in the track. This information is conveyed in a stripe with its thickness proportional to this number. Alternatively, the linear heatmap allows one to observe changes in effort level throughout race length by looking at the color changes along the horizontal axis. The use cases we used allowed us to observe such patterns in races of different distances. For example, comparing the 10K race (Figure 7) against the marathon (Figure 10) we observe that runners sustain high effort levels in the shorter race.

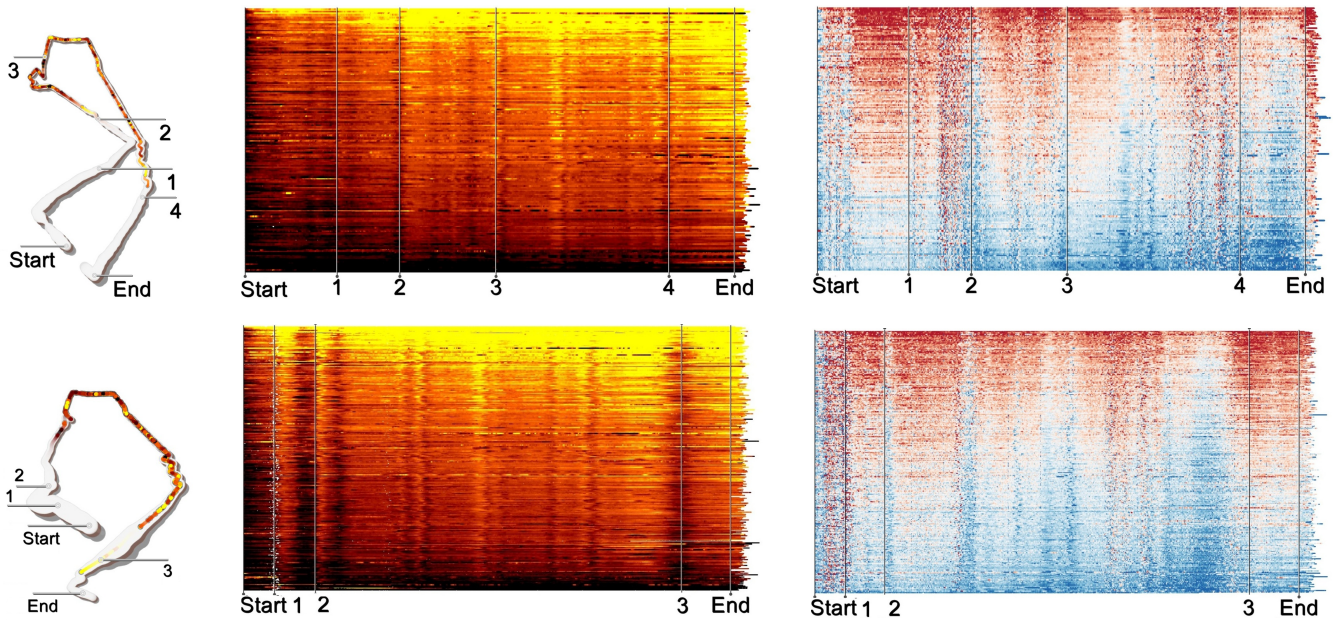


Figure 9. **São Silvestre Race in São Paulo** Augmented track view and linear heatmaps used to compare the datasets from 2010 (top) and 2011 (bottom). Distance markers are used to point alignment patterns in the linear heatmaps, showing that different courses create very different profiles of effort and speed.

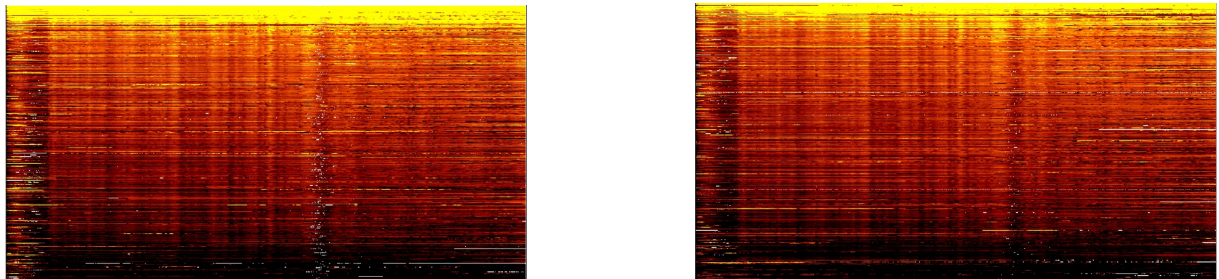


Figure 10. **Effort profile from New York City Marathons** Linear heatmaps showing the effort profile of the datasets from 2010 (top) and 2011 (bottom). Both show a small yellow area when compared to the effort profiles in the shorter races, indicating a more moderate effort level. Even though the data are from different years, since the course remained the same, the effort profiles are almost identical.

Q2: The most effort demanding part of a race can be identified in the linear effort heatmap as the vertical section where the particles have the brightest overall color. The use cases revealed that they are usually sections of long or strong increase in altitude, or the race final dash, where competition increases toward a better classification.

Q3: In the linear heatmap we observe an overall behavior with noisy vertical stripes created by changes in color shades. The distance at where such changes occur can be marked in the race track, and appears in both the effort and speed linear heatmaps. We verify that an increase in effort is caused by an increase in speed or at an uphill section of the track.

Q4: The speed linear heatmap reveal the running strategy of the race. The alignments in color changes represent an increment or decrement in the overall speed, and helps identifying parts of the track with high and low overall speed.

Q5: Identifying people running at risky levels is a major concern. It can help defining safer race tracks and running

strategies, as well as creating different pace groups. With filters in the linear heatmap, we could see people running slow but at high effort level and identify the section where such behavior begins. Further analysis of the same section in the speed heatmap and track view tells whether such speed is below the predominant value for the section and whether the cause is an increase in elevation. Speed below the local average, especially in sections of constant altitude is a indicator of fatigue.

Q6: Each of the three designs is useful to compare different races. The line graph heatmap shows the overall effort variation in time, the linear heatmap shows the effort and speed outcomes, and the track view display the race topography. This allows one to compare different races, but also to find differences in different in distinct instances of the same race. In the use cases, we compared races from consecutive years. As the course of the São Silvestre race changed from 2010 to 2011, the effort and speed profiles are very different, while, in the the New York marathons, they were basically the same.

VII. CONCLUSIONS AND FUTURE WORK

We presented three different designs, to visualize data collected from multiple runners in a given running race, that can be used together to bypass their individual limitations. They were motivated by a set of questions formulated by an expert on exercise physiology, and use cases to evaluate the visual analysis of different running events. The designs we created proved to be a useful tool for the comparison of different races, allowing, for instance, to assess the physical condition of the studied population. We expect such designs to be useful for elaborating new running strategies, or to help organizers to better plan running events by understanding how the race course affects the runners. The analysis of related datasets, like similar groups of runners on the same, and also on different races, can allow the comparison between those groups, establishing training goals for groups of lesser performance, for example. Other similar datasets that could benefit from our designs include the history of activities of a single person and groups of runners in different races. Also, the use of more variables and the analysis of other exercise modes, like cycling, are examples of topics for further research.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for valuable feedback. This work was supported in part by CNPq (Brazil) (140983/2011-2, 309483/2011-5, and 476685/2012-5), the NSF (USA) (CNS-1229185, CNS-1153503, IIS-0905385, IIS-1142013, and AGS 0835821), and the Department of Energy.

REFERENCES

- [1] S. N. Blair, H. W. K. III, J. Ralph S. Paffenbarger, D. G. Clark, K. H. Cooper, and L. W. Gibbons, "Physical Fitness and All-Cause Mortality. A Prospective Study of Healthy Men and Women," *Journal of The American Medical Association (JAMA)*, vol. 262, 1989.
- [2] S. N. Blair, H. W. K. III, C. E. Barlow, J. Ralph S. Paffenbarger, L. W. Gibbons, and C. A. Macera, "Changes in Physical Fitness and All-Cause Mortality. A Prospective Study of Healthy and Unhealthy Men," *Journal of The American Medical Association (JAMA)*, vol. 273, 1995.
- [3] M. R. Carnethon, M. Gulati, and P. Greenland, "Prevalence and Cardiovascular Disease Correlates of Low Cardiorespiratory Fitness in Adolescents and Adults," *Journal of The American Medical Association (JAMA)*, vol. 294.23, 2005.
- [4] R. M. Laukkanen and P. K. Virtane, "Heart rate monitors: State of the art," *Journal of Sports Science*, vol. 16, 1998.
- [5] W. Aigner, S. Miksch, W. Muller, H. Schumann, and C. Tominski, "Visualizing time-oriented data: A systematic view," *Computers & Graphics*, vol. 31, no. 3, 2007.
- [6] W. Aigner, S. Miksch, H. Schumann, and C. Tominski, *Visualization of Time-Oriented Data*, 1st ed., ser. Human-Computer Interaction. Springer Verlag, 2011.
- [7] H. Hochheiser and B. Shneiderman, "Interactive Exploration of Time Series Data," in *Proc. of the 4th Inter. Conf. on Discovery Science*, ser. DS '01. Springer-Verlag, 2001.
- [8] P. McLachlan, T. Munzner, E. Koutsofios, and S. North, "LiveRAC - interactive visual exploration of system management time-series data," in *Proc. ACM Conf. Human Factors in Computing Systems (CHI)*, 2008.
- [9] M. Krstajic, E. Bertini, and D. Keim, "CloudLines: Compact Display of Event Episodes in Multiple Time-Series," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 12, 2011.
- [10] H. Reijner, "The Development of the Horizon Graph," *Electronic Proceedings of the VisWeek Workshop From Theory to Practice: Design, Vision and Visualization*, 2008.
- [11] W. Javed, B. McDonnell, and N. Elmqvist, "Graphical perception of multiple time series," *IEEE Trans. Vis. Comput. Graph.*, vol. 16, 2010.
- [12] R. Kincaid and H. Lam, "Line graph explorer: scalable display of line graphs using Focus+Context," in *Proc. of the working conf. on advanced visual interfaces*. ACM, 2006.
- [13] D. Albers, C. Dewey, and M. Gleicher, "Sequence Surveyor:Leveraging Overview for Scalable Genomic Alignment Visualization," *IEEE Trans. Vis. Comput. Graph.*, 2011.
- [14] M. Hao, D. A. Keim, U. Dayal, and T. Schreck, "Multi-resolution techniques for visual exploration of large time-series data," in *Symposium on Visualization (EuroVis)*, 2007.
- [15] M. Friendly, "The history of the cluster heat map," *The American Statistician*, 2009.
- [16] C. Tominski and J. Abello, "Axes-Based Visualizations with Radial Layouts," *Proceedings ACM Symposium on Applied Computing*, 2004.
- [17] J. Zhao, F. Chevalier, and R. Balakrishnan, "KronoMiner: using multi-foci navigation for the visual exploration of time-series data," in *Proc. of the 2011 annual conference on Human factors in computing systems*, ser. CHI '11, 2011.
- [18] J. Zhao, F. Chevalier, E. Pietriga, and R. Balakrishnan, "Exploratory analysis of time-series with ChronoLenses," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 12, 2011.
- [19] M. D. Meyer, B. Wong, M. P. Styczynski, T. Munzner, and H. Pfister, "Pathline: A Tool For Comparative Functional Genomics," *Comput. Graph. Forum*, vol. 29, no. 3, 2010.
- [20] C. Viau and M. J. McGuffin, "ConnectedCharts: Explicit Visualization of Relationships between Data Graphics," *Comput. Graph. Forum*, vol. 31, 2012.
- [21] C. Tominski, H. Schumann, G. L. Andrienko, and N. V. Andrienko, "Stacking-Based Visualization of Trajectory Attribute Data," *IEEE Trans. Vis. Comput. Graph.*, vol. 18, 2012.
- [22] 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," *Comput. Graph. Forum*, 2012.
- [23] W. Haskell, M. E. Nelson, R. Dishman, E. Howley, W. Kort, W. Kraus, I. Lee, A. McTiernan, R. Pate, K. Powell, J. Regensteiner, J. Rimmer, and A. Yancey, "Physical Activity Guidelines Advisory Committee Report, 2008," June 18, 2008 2008.
- [24] S. Gupta, A. Rohatgi, C. R. Ayers, B. L. Willis, W. L. Haskell, A. Khera, M. H. Drazner, J. A. De Lemos, and J. D. Berry, "Cardiorespiratory fitness and classification of risk of cardiovascular disease mortality," *Circulation*, vol. 112, 2005.
- [25] J. D. Berry, B. Willis, S. Gupta, C. E. Barlow, S. G. Lakoski, A. Khera, A. Rohatgi, J. A. De Lemos, W. Haskell, and D. M. Lloyd-Jones, "Lifetime risks for cardiovascular disease mortality by cardiorespiratory fitness levels measured at ages 45, 55, and 65 years in men. The Cooper Center Longitudinal Study," *Journal of the American College of Cardiology*, vol. 57, 2011.
- [26] R. Benson and D. Connolly, *Heart Rate Training*, 1st ed. Human Kinetics, 2011.
- [27] J. H. Kim, R. Malhotra, G. Chiampas, P. d'Hemecourt, C. Troyanos, J. Cianca, R. N. Smith, T. J. Wang, W. O. Roberts, P. D. Thompson, and A. L. Baggish, "Cardiac Arrest during Long-Distance Running Races," *New England Journal of Medicine*, vol. 366, no. 2, 2012.
- [28] S. C. Mathews, D. L. Narotsky, D. L. Bernholt, M. Vogt, Y.-H. Hsieh, P. J. Pronovost, and J. C. Pham, "Mortality Among Marathon Runners in the United States, 2000-2009," *Am J Sports Med*, 2012.
- [29] M. Wilhelm, L. Roten, H. Tanner, J. Schmid, I. Wilhelm, and H. Saner, "Long-term cardiac remodeling and arrhythmias in nonelite marathon runners," *Am J Cardiol*, vol. 110, no. 1, 2012.
- [30] A. Murray and R. J. S. Costa, "Born to run. Studying the limits of human performance," *BMC Med*, vol. 10, no. 1, 2012.
- [31] M. J. Joyner, J. R. Ruiz, and A. Lucia, "The Two-Hour Marathon: Who and When?" *Journal of Applied Physiology*.
- [32] L. Beis, M. Wright-Whyte, B. Fudge, T. Noakes, and Y. Pitsiladis, "Drinking behaviors of elite male runners during marathon competition," *Clinical Journal of Sport Medicine*, vol. 22, no. 3, 2012.
- [33] E. Tam, H. Rossi, C. Moia, C. Berardelli, G. Rosa, C. Capelli, and G. Ferretti, "Energetics of running in top-level marathon runners from Kenya," *Eur J Appl Physiol*, 2012.
- [34] E. R. Burke, *Precision Heart Rate Training*. Human Kinetics, 1998.
- [35] J. Friel, *Total Heart Rate Training*. Ulysses, 2006.
- [36] Garmin, "Garminconnect," <http://connect.garmin.com/>.