Collecting Information Needs for Egocentric Visualizations while Running

Ahmed Elshabasi*
University of Calgary

Lijie Yao†
Xi'an Jiaotong-Liverpool University
University of Calgary
Université Paris-Saclay, CNRS,
Inria, LISN, France

Petra Isenberg‡
Université Paris-Saclay CNRS, Inria,
LISN, France

Charles Perin§
University of Victoria

Wesley Willett¶
University of Calgary

Figure 1: Four information needs, plus accompanying video stills captured during real-world runs and races highlight opportunities for new egocentric visualizations.

ABSTRACT

We investigate research challenges and opportunities for visualization in motion during outdoor physical activities via an initial corpus of real-world recordings that pair egocentric video, biometrics, and think-aloud observations. With the increasing use of tracking and recording devices, such as smartwatches and head-mounted displays, more and more data are available in real-time about a person's activity and the context of the activity. However, not all data will be relevant all the time. Instead, athletes have information needs that change throughout their activity depending on the context and their performance. To address this challenge, we describe the collection of a diverse corpus of information needs paired with contextualizing audio, video, and sensor data. Next, we propose a first set of research challenges and design considerations that explore the difficulties of visualizing those real data needs in-context and demonstrate a prototype tool for browsing, aggregating, and analyzing this information. Our ultimate goal is to understand and support embedding visualizations into outdoor contexts with changing environments and varying data needs.

Index Terms: Situated visualization, visualization in motion, sports.

*e-mail: ahmed.elshabasi@ucalgary.ca

†e-mail: yaolijie0219@gmail.com

‡e-mail: petra.isenberg@inria.fr

§e-mail: cperin@uvic.ca

¶e-mail: wesley.willett@ucalgary.ca

1 INTRODUCTION

The integration of situated visualizations into sports and physical activities presents an exciting frontier for enhancing athletic performance and experience. With the proliferation of wearable technology and real-time data collection, athletes have unprecedented access to detailed information about their activities. Such information can offer athletes immediate, actionable insights that can inform their training and performance. However, the challenge lies in effectively identifying the athletes' data needs and presenting this information in a manner that is contextually relevant and dynamically responsive to their changing needs and environments.

Our work presents an initial corpus of contexts and information needs captured live by amateur athletes during outdoor physical activities (primarily running). The corpus includes biometric data, egocentric videos captured by the athletes, think-aloud audio documenting their in-context information needs, and written transcripts of those needs. Additionally, we highlight early progress on a tool for analyzing these needs and their surrounding context and supporting the design of new visualizations to address them. Our initial analysis of information needs from the corpus reveal considerable variability in needs across individuals, even for similar activities, along with a number of other potential challenges and opportunities for designing immersive visualization-in-motion applications.

2 RELATED WORK

Our work is closely related to the topic of situated and embedded visualizations [23], which are visual representations located close to a data referent [3]. In addition, we are inspired by the research direction of visualization in motion in which visualizations are used during relative movement between a viewer and a visualization [25]. We also build on visual analytics research for outdoor physical activities, focusing on the connection between situated data needs and their contexts under dynamic situations and how to visualize such needs best.

Figure 2: Our lightweight workout equipment for capturing runs, information needs, and biometric data: a Ray-Ban Meta Smart Glasses and a Garmin smartwatch.

Figure 3: Our initial prototype interface showing video snippets and biometrics for a single activity.

Physical activities in dynamic outdoor scenarios have gained a lot of traction in the visualization community. The research ranges from analyzing the data outside of its context [1, 13–17, 22, 24] to embedding visual representations into scenarios, such as swimming [26], table tennis [5, 6], soccer [19, 24], basketball [2, 4, 11, 12, 20], cycling [9], badminton [7, 10, 27], and tennis [18]. These works mainly focus on providing technologies and tools that support injecting visual representations into various scenarios.

Only few past works elicited viewers' actual data needs [12, 26]. Yao et al. [26] investigated the audiences' real data interests when watching swimming races and provided a technology probe, *Swim-Flow*, that allows viewers to embed visualizations that can move with swimmers in real-time into swimming videos according to their real data needs. Lin et al. [12] conducted a survey with basketball fans to understand important moments in basketball games and provided concrete visualization solutions for these spot moments. Our work is also based on real data needs but it is unique in that, instead of considering the audience, we focus on the athletes themselves and how their needs change over time, based on their performance, and the context of their activity.

Other works have employed various approaches in analyzing running data for prospective design applications. A study on designing for active environments [21] used data from running apps to explore urban design. Through the collective lens, the analysis revealed general running patterns and environmental preferences, while the individual lens focused on unique user behaviors. Contextual data such as weather, neighborhood info, and lighting conditions were also considered to understand environmental influences. A similar study examined how environmental characteristics and runners' motives affect their perception of running environments [8], collecting data from half marathon runners through the Eindhoven Running Survey 2015. The researchers used linear regression analyses to assess the impact of intrapersonal characteristics and perceived environmental traits.

3 PREPARATION: COLLECTING INFORMATION NEEDS

To study different in-situ data needs in various contexts of outdoor physical activities, we are currently working on a database of first person videos of these activities. The database is meant to collect short video snippets together with audio recordings of information needs. We selected running as our first sports activity for the following reasons: (a) Prevalence: Running is one of the most prevalent sports activities, does not require costly equipment, and due to the number of runners it would be easy for us to find participants.

(b) Context variety: Unlike other activities such as skating, skiing, or swimming, which require a specific court/pool, running can be performed almost anywhere. As such, we may be able to collect an extensive range of contexts outside of indoor and outdoor courts, such as urban areas, forests, beaches, in the mountains, or on flat land—but also in varied weather and daylight conditions.

We divided our data collection process into three phases (recording, cleaning up, and analysis).

For **recording**, we used either a GoPro HERO 10 or Ray-Ban Meta Smart Glasses, paired with a Garmin or Coros smartwatch. Participants (all of whom are authors on this publication) captured first-person videos of their own typical activities over a 3-month period, including a variety of runs and races. We used a self-initiated think-aloud protocol where participants would turn on their recording device and record a brief spoken note along with contextualizing video whenever an information need occurred to them.

For **cleaning up**, each individual reviewed their own audio and video and removed potentially-identifying or otherwise sensitive content, with a particular attention to other bystanders and other non-participant athletes. We then extracted transcripts from all videos for indexing and analysis.

For **analysis**, we manually coded the extracted transcripts, identifying emergent themes for each individual and looking for patterns both within and across activities.

Each activity in the resulting corpus includes a `.fit` file containing biometrics for the run, a text transcript documenting information needs (stored as a `.txt` file), and a series of video snippets (as `.mp4` files). Observations from our initial corpus (141 videos over 18 runs by 3 runners) show that information needs vary significantly between individuals, even when conducting similar activities. Over a three-month period, we examined our own information needs and found minimal overlap. Specifically, one of us identified 56 distinct information needs, another identified 37 distinct needs, and the third identified 42 distinct needs. Out of these, only 11 information needs were common across all three runners. These common needs included access to a route map, heart rate, speed, elevation, distance, elapsed time, water intake tracking, performance comparisons to other workouts, progress percentage, weather information, and pace planning. Figure 1 shows examples of data need transcripts along with a screenshot of the video at the time that provides the context for the data needs.

4 WORK IN PROGRESS: ANALYSIS PROTOTYPE

To complement our process, we also initiated the development of a tool to help us analyze and aggregate first-person videos alongside the collected biometric data (Figure 3). This tool allows us to view the video and corresponding biometric data in a synchronous manner, providing contextually-relevant information for each data need. Currently in its early phases, the tool allows to plot biometric

Figure 4: An early-stage sketch on top of a video snippet.

data along with a timeline indicating when the videos were taken during the workout. We plan to continue refining and expanding its features to further support our research and development efforts.

Additionally, with this prototype, we created early-stage sketches on top of stills captured from these recorded videos. This approach helped us explore how these sketches can be overlaid on first-person points of view and what implications this might have for future tools and applications in similar contexts.

5 FUTURE RESEARCH: OPPORTUNITIES & CHALLENGES

Our initial data collection and analysis surfaced several recurring themes and opportunities related to information needs. These opportunities suggest potential approaches as we move forward with our prototype design. Given that this research project is in its early stages and our data collection corpus is limited, our findings are preliminary and require further validation. Nonetheless, these are the main opportunities we have identified so far:

- **Variability Across Individuals:** Information needs can vary considerably across individuals and activity types, influenced by factors such as weather, time-of-day, mindset, and location. It remains to be seen how we can identify regular patterns even for individuals. Additionally, there is a need to explore whether some needs are highly specific to particular circumstances and mindsets; for instance, whether users require detailed guidance and feedback or prefer minimal, high-level data. How needs differ based on whether the activity is solitary or involves social or competitive elements also remains uncertain.

- **Variability Across Time:** Information needs can change considerably over time, even for the same individual. For example, a person training for a marathon may have different needs during the early stages of training compared to the final weeks before the event. Developing strategies and data references that adapt over time and context is an area that requires further investigation. Additionally, similar information needs from different participants might look different on the surface or vice versa, which complicates the process of identifying and addressing these needs effectively.

- **Relevance of Comparisons**: Many information needs involve comparisons to past data, other activities, or other individuals, raising the question of how to determine which comparisons are relevant given the vast range of possible comparisons. Information needs often extend beyond biometrics to include safety considerations (like the presence of bears), environmental conditions (weather conditions, trail closures), and personal preferences (such as music or historical context).

- **Importance of Certain Needs and Avoiding Overload:** Some analyses may not be practical or necessary to perform in real-time during the activity. Identifying which analyses are meaningful and when they should be performed remains an open question. Additionally, there are times when individuals might prefer not to have access to data, focusing instead on connecting with nature or the activity itself. Balancing the provision of useful data with the option for minimal or no data is another area that needs careful consideration.

- **Influence of Data Collection on Data Needs:** The process of collecting data can influence the data needs, prompting new ideas and needs that might not have been considered otherwise. Managing this self-nurturing and iterative effect without complicating the design process is a valuable opportunity for innovation.

- **Broadening the Corpus:** The extensive data collection requirements, while crucial for detailed insights, limited participant diversity and may affect the generalizability of the findings. Moving forward, one potential avenue worth exploring could be leveraging accessible platforms like Strava for data collection. For example, a Strava user might store a recorded running workout privately for later processing. Although Strava only captures biometric data and would need to be aided by a pipeline for think-aloud information needs, there remains significant potential in finding ways to simplifying the data collection process and integrating less intrusive methods which could help in ensuring a wider range of participants.

These opportunities highlight the potential for designing tools that address varied and evolving information needs during physical activities. As we continue our research, these insights will inform our approach to developing effective and adaptable solutions.

REFERENCES

[1] G. Andrienko, N. Andrienko, G. Anzer, P. Bauer, G. Budziak, G. Fuchs, D. Hecker, H. Weber, and S. Wrobel. Constructing spaces and times for tactical analysis in football. *IEEE Transactions on Visualization and Computer Graphics*, 27(4):2280–2297, 04 2021. doi: 10.1109/TVCG.2019.2952129 2

[2] L. Bai, C. Efstratiou, and C. S. Ang. wesport: Utilising wrist-band sensing to detect player activities in basketball games. In *Proc. of 2016 IEEE International Conference on Pervasive Computing and Communication Workshops*, pp. 1–6, 2016. doi: 10.1109/PERCOMW.2016.7457167 2

[3] N. Bressa, K. Wannamaker, H. Korsgaard, W. Willett, and J. Vermeulen. Sketching and ideation activities for situated visualization design. In *Proc. Conf. Design. Interact. Syst.*, pp. 173–185. ACM, 06 2019. doi: 10.1145/3322276.3322326 1

[4] Z. Chen, Q. Yang, J. Shan, T. Lin, J. Beyer, H. Xia, and H. Pfister. IBall: Augmenting Basketball Videos with Gaze-Moderated Embedded Visualizations. In *Proceedings of the Conference on Human Factors in Computing Systems*, number 841, pp. 1–18, 2023. doi: 10.1145/3544548.3581266 2

[5] Z. Chen, Q. Yang, X. Xie, J. Beyer, H. Xia, Y. Wu, and H. Pfister. Sporthesia: Augmenting sports videos using natural language. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):918–928, 10 2023. doi: 10.1109/TVCG.2022.3209497 2

[6] Z. Chen, S. Ye, X. Chu, H. Xia, H. Zhang, H. Qu, and Y. Wu. Augmenting sports videos with VisCommentator. *IEEE Transactions on*

Visualization and Computer Graphics, 28(1):824–834, 2022. doi: 10.1109/TVCG.2021.3114806 2

[7] X. Chu, X. Xie, S. Ye, H. Lu, H. Xiao, Z. Yuan, C. Zhu-Tian, H. Zhang, and Y. Wu. Tivee: Visual exploration and explanation of badminton tactics in immersive visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):118–128, 2022. doi: 10.1109/TVCG.2021.3114861 2

[8] I. Deelen, M. Janssen, S. Vos, C. B. Kamphuis, and D. Ettema. Attractive running environments for all? a cross-sectional study on physical environmental characteristics and runners' motives and attitudes, in relation to the experience of the running environment. *BMC public health*, 19:1–15, 2019. 2

[9] O. Kaplan, G. Yamamoto, Y. Yoshitake, T. Taketomi, C. Sandor, and H. Kato. In-situ visualization of pedaling forces on cycling training videos. In *Proceeding of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 000994–000999, 2016. doi: 10.1109/SMC.2016.7844371 2

[10] T. Lin, A. Aouididi, C. Zhu-Tian, J. Beyer, H. Pfister, and J.-H. Wang. Vird: Immersive match video analysis for high-performance badminton coaching. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):458–468, 2024. doi: 10.1109/TVCG.2023.3327161 2

[11] T. Lin, Z. Chen, J. Beyer, Y. Wu, H. Pfister, and Y. Yang. The ball is in our court: Conducting visualization research with sports experts. *IEEE Computer Graphics and Applications*, 43(1):84–90, 2023. doi: 10.1109/MCG.2022.3222042 2

[12] T. Lin, Z. Chen, Y. Yang, D. Chiappalupi, J. Beyer, and H. Pfister. The quest for Omnioculars: Embedded visualization for augmenting basketball game viewing experiences. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):962–971, 2023. doi: 10.1109/TVCG.2022.3209353 2

[13] C. Perin, J. Boy, and F. Vernier. Using Gap Charts to Visualize the Temporal Evolution of Ranks and Scores. *IEEE Computer Graphics & Applications*, 2016. doi: 10.1109/MCG.2016.100 2

[14] C. Perin, R. Vuillemot, and J.-D. Fekete. Soccerstories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, 2013. doi: 10.1109/TVCG.2013.192 2

[15] C. Perin, R. Vuillemot, and J.-D. Fekete. À Table!: Improving Temporal Navigation in Soccer Ranking Tables. In *Proceedings of the Conference on Human Factors in Computing Systems*, pp. 887–896. ACM, 2014. doi: 10.1145/2556288.2557379 2

[16] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and S. Carpendale. State of the art of sports data visualization. *Computer Graphics Forum*, 37(3):663–686, 2018. doi: 10.1111/cgf.13447 2

[17] G. Pingali, A. Opalach, Y. Jean, and I. Carlbom. Visualization of sports using motion trajectories: providing insights into performance, style, and strategy. In *Proceeding of the IEEE Visualization Conference*, pp. 75–544, 2001. doi: 10.1109/VISUAL.2001.964496 2

[18] T. Polk, D. Jäckle, J. Häußler, and J. Yang. CourtTime: Generating actionable insights into tennis matches using visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):397–406, 01 2020. doi: 10.1109/TVCG.2019.2934243 2

[19] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim. Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):13–22, 01 2018. doi: 10.1109/TVCG.2017.2745181 2

[20] L. Tica, S. Rishi, Y. Yalong, N. Carolina, B. Johanna, A. S. Maurice, and P. Hanspeter. Towards an understanding of situated ar visualization for basketball free-throw training. In *Proceedings of the Conference on Human Factors in Computing Systems*, number 461, pp. 1–13, May 2021. doi: 10.1145/3411764.3445649 2

[21] L. van Renswouw, S. Bogers, C. Lallemand, P. van Wesemael, and S. Vos. Changing perspective on data in designing for active environments. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, pp. 2403–2419, 2024. 2

[22] J. Wang, J. Wu, A. Cao, Z. Zhou, H. Zhang, and Y. Wu. Tac-miner: Visual tactic mining for multiple table tennis matches. *IEEE Transactions on Visualization and Computer Graphics*, 27(6):2770–2782, 2021. doi: 10.1109/TVCG.2021.3074576 2

[23] W. Willett, Y. Jansen, and P. Dragicevic. Embedded data representations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):461–470, Jan. 2016. doi: 10.1109/TVCG.2016.2598608 1

[24] Y. Wu, J. Lan, X. Shu, C. Ji, K. Zhao, J. Wang, and H. Zhang. iTTVis: Interactive visualization of table tennis data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):709–718, 01 2018. doi: 10.1109/TVCG.2017.2744218 2

[25] L. Yao, A. Bezerianos, R. Vuillemot, and P. Isenberg. Visualization in motion: A research agenda and two evaluations. *IEEE Transactions on Visualization and Computer Graphics*, 28(10):3546–3562, Oct. 2022. doi: 10.1109/TVCG.2022.3184993 1

[26] L. Yao, R. Vuillemot, A. Bezerianos, and P. Isenberg. Designing for visualization in motion: Embedding visualizations in swimming videos. *IEEE Transactions on Visualization and Computer Graphics*, 30(3):1821–1836, March 2024. doi: 10.1109/TVCG.2023.3341990 2

[27] S. Ye, Z. Chen, X. Chu, Y. Wang, S. Fu, L. Shen, K. Zhou, and Y. Wu. ShuttleSpace: Exploring and analyzing movement trajectory in immersive visualization. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):860–869, 02 2021. doi: 10.1109/TVCG.2020.3030392 2