Diving Deep into Time: Temporal Arrangements for Embedded Visualization in Swimming Videos

Junxiu Tang, Lijie Yao, Lu Ying, Romain Vuillemot, Petra Isenberg

***Abstract*—We introduce a temporal arrangement framework for embedding visualizations in sports videos with a focus on swimming. Our work is inspired by strategies used in current TV broadcasts, where visualizations are selectively displayed to provide meaningful and engaging information to audiences. We began with a systematic review of TV broadcast practices, through which we identified recurring temporal combinations of visualizations and competition statuses, which we define as patterns of temporal arrangement for embedded visualizations. To move beyond the constraints of existing broadcast practices, we then conducted a formative study with a general population. Based on this broader perspective, we designed a configuration framework that allows us to formally specify when and for how long, related to swimming context metadata, visualizations appear in a video. We instantiate the framework in a technology probe, *SwimChrono*, for applications with real-world swimming context videos. Through audience-customized configurations, *SwimChrono* supports novel arrangements beyond those used in existing professional settings, is adaptable to various swimming contexts, including different lengths and swimming styles, and key events. Furthermore, we conduct user studies and contribute use cases to illustrate how our framework can be well applied for diverse needs.**

***Index Terms*—Embedded visualization, visualization in motion, sports visualization, augmented sports video, and design pattern.**

I. Introduction

Embedded visualizations in professional swimming race videos [6], [56] are seen by millions of viewers. Fig. 1, for example, shows broadcast visualizations[1] from the Olympics, (a) ranking the top 3 swimmers after a turn and (b) speed and nationality labels that moved with swimmers when the race was close to the end. Visualizations such as these provide additional information to the audience and can help them follow a race and the performance of specific swimmers. These embedded visualizations, however, need to be carefully designed for dynamic content and timed to appear at chosen moments for a precise duration. Carefully designing these arrangements is important not only because of limited screen space but also because not all available data are relevant or interesting to an audience at once. For example, the yellow world record (WR) line matters most when the leading swimmer has a chance to approach or break the world record (Fig. 1 (b)), while the elapsed time info box in the bottom right corner is informative during the entire race (Fig. 1 (c)).

Therefore, sports broadcasting is a storytelling process [35], [36] in which emphasis can be placed on certain race states, events, or actors by adding data visualizations to the scene at the right moments. In addition, audiences may want to craft their own stories by tailoring visualizations to their own needs and preferences [5], [29]. In summary, these characteristics point to concrete user pain points with current systems: broadcast directors must balance the tension between limited screen space and the availability of a large amount of race metadata, deal with changing data relevance over the course of a rance, and diverse audience needs, while viewers have little agency to adapt visualizations to their own evolving inerests during a race. These pain points, together with the absence of a flexible data-driven temporal framework for sports videos, motivate our work.

In this paper, we treat temporal arrangement as a distinct design dimension in embedded sports visualizations and formalize it as a configurable framework that captures and generalizes recurring patterns from broadcasting practice. Our work is based on a deep, data-driven analysis of existing broadcasting practices and, therefore, supports industry conventions. Yet, in our analysis, we also found that current practices are limited in their creative expression and in the data and temporal events the visualizations relate to. We address this limitation and allow to easily go beyond existing practices by providing a set of definitions that can be combined to generate novel, previously unseen temporal arrangements and event-based dependencies among embedded visualizations in sports videos. Our rules define "*when* and for *how long* to show" a visualization and, therefore, characterize the *temporal arrangement* of any embedded visualization in a sports video. Specifically, we focus on how to configure these temporal arrangements to be data-driven and therefore transferable across different events within the same sport. Our temporal arrangements work alongside other arrangements that decide *where* a visualization appears, *what* a visualization shows, and *how* it looks.

We started our work in the domain of swimming, a speed-dominant sport that contains a rich set of data to track and visualize. Swimming broadcasts already contain simple embedded visualizations that are either static or in motion [56], providing us with practical professional scenarios to start from, improve, and extend. To **ground our contributions in**

Junxiu Tang is with Northwestern University, E-mail: junxiu.tang@kellogg.northwestern.edu. Lijie Yao is with Xi'an Jiaotong-Liverpool University, E-mail: yaolijie0219@gmail.com. Lu Ying is with the National University of Singapore, E-mail: yiyinyingl@outlook.com. Romain Vuillemot is with LIRIS, École Centrale de Lyon, E-mail: romain.vuillemot@ec-lyon.fr. Petra Isenberg is with Université Paris-Saclay, CNRS, Inria, LISN, E-mail: petra.isenberg@inria.fr. Part of this work was done when Junxiu Tang, Lijie Yao and Lu Ying were with Université Paris-Saclay, CNRS, Inria, LISN.
Corresponding author: Lijie Yao.

[1]We use the term visualization here to refer also to visual representations that use text and tables to "amplify cognition" [12]

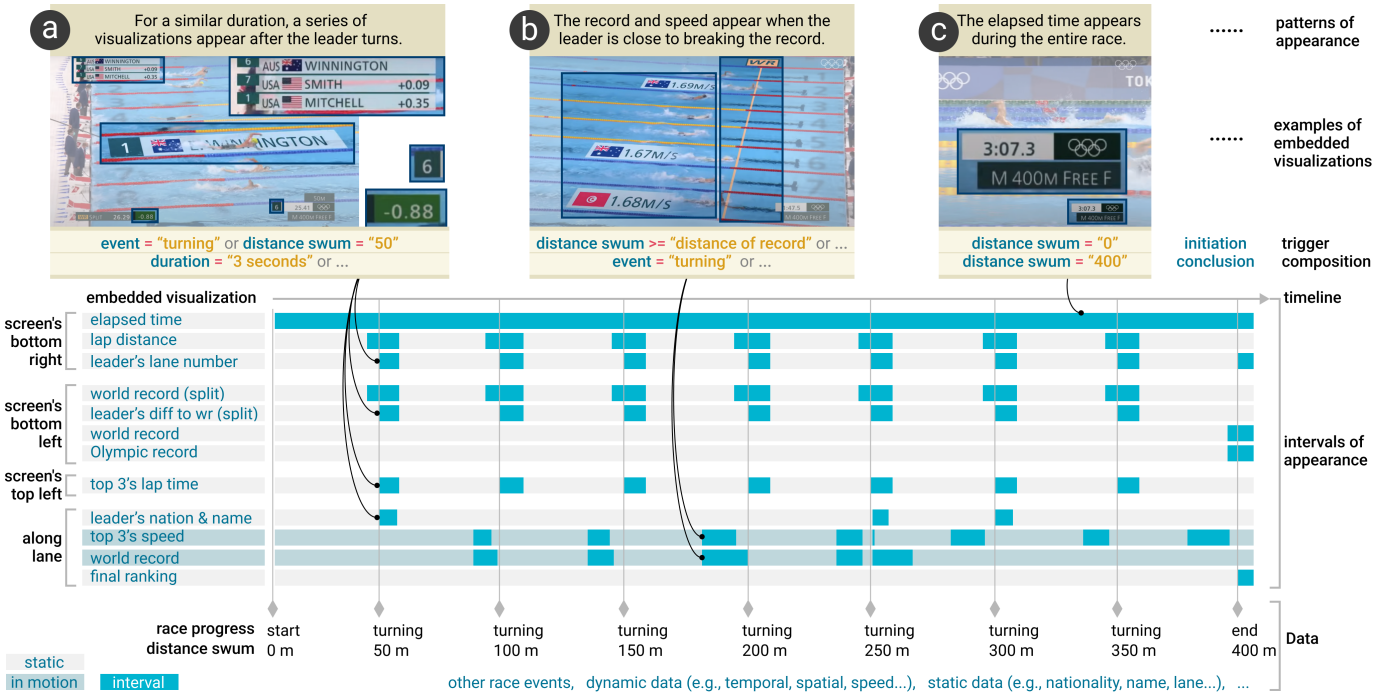


Fig. 1: Temporal arrangement of embedded visualizations in an Olympic swimming race video [6]. Visualizations appear depending on race progress and metadata, and remain visible for a predetermined duration. The visualizations are designed to provide relevant and timely insights and context to an audience. For example, (a) a series of visualizations appear after turning; (b) the record line appears when the leading swimmer is close to breaking the record; (c) the elapsed time at the corner appears during the entire race.

existing practice and ensure relevance to end-users, we engaged in multiple research activities: a) We systematically reviewed and labeled video data before describing patterns of temporal arrangements for embedded visualizations in recent international and regional swimming race broadcasts. b) We surveyed the literature and commercial solutions to learn about current workflows. c) We conducted a formative study to determine when and which information would be of interest to audiences, and the triggers that should make this information visible during a swimming race.

In order to **ensure practical relevance**, we conceptualized temporal arrangements based on race data using a framework that both captures the existing patterns we found and enables potentially new temporal arrangements of embedded visualizations desired by audiences. We implemented our configuration framework into *SwimChrono*, a technology probe that allows users to define flexible temporal arrangements of visualization in a data-driven way. By using our framework and *SwimChrono*, the predefined temporal arrangements can be smoothly applied to other swimming videos of varying lengths or race types. And we even show how it can be applied to a climbing video.

Our contributions are threefold: First, **a dataset** of labeled embedded visualizations and identified design patterns from **a systematic review** of temporal arrangements of embedded visualizations in professional broadcasts. Second, **a formative study** to explore audience interests, which inform our **configuration framework** of flexible condition triggers for the temporal arrangement of embedded visualizations. Third, **a technology probe** called *SwimChrono*, which allows users to design their own temporal arrangements of visualizations across different swimming videos based on data. *SwimChrono* is open-source and available `online`. We last conducted **case studies** across real-world swimming races, an **expert interview** and a **comparative study** to evaluate the expressiveness and effectiveness of our configuration framework.

II. Related Work

Embedded visualizations are visual data representations attached to data referents [51]. Existing studies relevant to our work on embedded visualizations in sports videos have targeted various types of sports [13]–[15], [21], [30], [47], [54], [56]. These past studies have in common with ours that researchers investigated multiple video corpora and proposed design spaces or approaches for embedding visualizations for these sports. Generally, they focused on *what data* to visualize and *how* and *where* to display it, while we focus on the *when* and *for how long* questions — what we call the temporal arrangements of visualizations. When temporal arrangements were considered in the past, they targeted key events in a competition (like a series of strokes in table tennis) as temporal anchors for postgame commentary [15] or used the narration of professional commentary to schedule visualization presentations [14]. In contrast, we treat temporal arrangement as an independent dimension abstracted away from specific visualization designs or geometric encodings. More broadly, our research is highly related to embedded visualization with

a specific focus on visual narrative, cinematic theory, as well as sports videos, and temporal authoring of animated visualization.

A. Narrative and Cinematic Theory

Visual narrative and cinematic theory play important roles in shaping storytelling in data videos [3], [17], [40], [53], including those videos that include embedded visualizations. Amini et al. [3] investigated data videos through four narrative categories, *Establisher*, *Initial*, *Peak*, and *Release*, based on Cohn's visual narrative theory [16], which are linear on the video timeline. The authors highlight the importance of flexible configurations of a temporal narrative structure in data video authoring. Prior studies have also explored pre-scheduled camera templates for movie narratives and virtual environments [20], [22]. Sports races contain recurring phases, such as the start, competition progress, and final moments, as well as unexpected events like lead changes. The ability to schedule visualizations to appear or disappear during such phases can support narrative framing through highlighting and annotation. Temporal arrangements of visualizations in sports videos can, therefore, be considered one narrative technique among many others, such as verbal narration, camera movements, etc [35], [36]. We did, however, not design our techniques through the lens of narration support, but focused on the technical aspects of providing temporal triggers for embedded visualizations. We enabled the timing to present visualizations so people are not overwhelmed, and the full story can unfold naturally.

B. Embedded Visualizations in Sports Videos

Embedded visualizations in sports videos have been the subject of research for professional analysis [15], [47] and to enhance personal viewing experiences for the general public [5], [13], [29], [30], [56]. Previous research on professional sports visual analytics has yielded diverse insights. Stein et al. [46] proposed a workflow framework that helps to decide on appropriate visualizations for specific situations in soccer analysis. Chen et al. [15] identified a design space for augmenting sports videos and developed *VisCommentator* as a tool that enables semi-automatic commentary generation for table tennis videos. Follow-up work from Chen et al. on *Sporthesia* [14] extended *VisCommentator* by using keywords in commentary as temporal anchors and mapping them with relevant visual content in videos. Similarly, we use key events as temporal references for embedding visualizations into swimming videos. We additionally explore features of sports metadata (distance and time) to develop a data-driven framework of triggers for the temporal arrangement of embedding visualizations in swimming videos.

Regarding viewing experience enhancement, Lin et al. [30] designed approaches to embed personalized visualizations for basketball broadcasting. The authors highlighted the benefits of embedding personalized visualizations in sports: increased engagement and deeper insights. Our work complements previous research by providing personalizable and transferable embedded visualization configurations adaptable to various swimming races. Our approach maintains consistent spatial and temporal visualization settings, aiming to support similar viewing experiences even across different swimming videos.

Researchers also explored scenarios where motion factors have been heavily involved. Yao et al. introduced visualization in motion [55] as a research direction that studies situations where relative movement exists between an entire visualization and its reviewer. Next, Yao et al. [56] investigated the data interests of the general public on swimming races and provided *SwimFlow*, a technology probe meant to show the importance of providing a mechanism to prototype embedded visualizations directly on videos in real-time, as opposed to traditional design methods. However, the authors focused on what to show and how to show it, but neither included temporal features for embedding visualizations nor the ability to transfer designs between videos of different races. While we reused the successful basic visualization authoring functions from *SwimFlow* (which are not contributions of this paper), *SwimChrono* is an entirely new implementation that enables embedding visualizations in swimming videos with flexible, data-driven, transferable, and customizable temporal arrangements.

Apart from research, industry has developed various sports video tools and general video editing tools. Several commercial companies propose solutions tailored to specific sports [10], [21], [44], [45], [50] and have deployed their products in major sports events. For instance, Vizrt [50] supports a pipeline including camera recording, real-time analysis, and broadcasting. General video editing tools, like Premiere [2] and CapCut [11], allow users to manually add custom designs and compose clip timelines according to their specific needs. However, whether focused on sports or not, these commercial tools have not specifically addressed *data-driven* temporal arrangements for visualizations. Our complete survey of current commercial sports video editing tools is available in the supplementary materials. In addition to providing data-driven arrangements and a technology probe, our work focuses on speed-dominant sports and explores how visualizations in motion can be temporally embedded within this context. Our work might inspire features of future professional products.

C. Temporal Authoring of Animated Visualization

Many visualizations in motion in swimming videos can be considered animated because they refresh synchronously with real-time updating sports data. For our work, we specifically take inspiration from how the animation literature has more broadly considered authoring for motion. Existing studies on animation can be divided into two primary categories: those with data-defined duration and those with user-defined duration. Data-defined animations are automatically generated from updates of the visualized data. Animated Vega-Lite [59] supports visualization animation specification through time-varying data queries. As data query results change over time, smooth transitions in the visualization present new data, forming animated visualizations. Another example is Data GIF [43], where data is illustrated by frames and the frames are concatenated into a short, automatically repeating animation.

Alternatively, animated visualization can be defined by users based on a timeline concept. Gemni [28], Canis [24]

and CAST [23] allow users to set the order of visualization animation, while other tools such as DataClips [4], DataAnimator [49], AutoClips [42], and DataPlayer [41], all feature a timeline panel. The order configuration or timeline function in these tools supports users in modifying the order of a narrative sequence, adjusting timestamps for visual element appearance, and setting animation duration, thereby facilitating various animated visualization sequences.

Our technology probe features both timeline and trigger mechanisms to support user-required and data-driven temporal arrangements of visualizations. This unique combination of data-defined and user-defined visualization updates distinguishes our approach from previous work. Besides, we focus on the overall temporal arrangement of visualizations rather than the animation of individual elements within them.

III. Systematic Review

To ground our research on temporal patterns in current practice, we reviewed 31 professional swimming broadcasts and categorized how visualizations are currently temporally arranged in them. Based on this data, we identified 8 temporal arrangement patterns.

A. Preparation: Data Collection and Labeling

To create our video corpus for review, we narrowed our focus to swimming races for the following reasons: (a) Swimming has rich data with race-related events; (b) The visualizations embedded in swimming include both static representations and visualizations in motion [56], which provide rich usage scenarios; (c) Swimming races have various durations (50 m, 100 m, 1000 m, ...), allowing us to investigate whether race length currently plays a role in data presentation; (d) The visualizations embedded in swimming are rendered in real-time with the race's progress. This differs from commentary and analysis considered in prior work about team or racket sports [15], [30], [47] and can offer novel insights. While we focus on swimming, we will later discuss how our work may apply to other performance-related sports and beyond.

We chose the official semifinal/final races from recent Olympics, World Championships, and European Championships because the latest broadcasts of major events represent the most advanced embedded visualizations, giving us the best impression of the state-of-the-art [56]. We focused on official broadcasts because they were designed to be engaging and interesting to a general audience. We collected the videos from official race websites [32] or race recordings of official broadcasting accounts on YouTube [58].

Given that not all kinds of swimming race videos are complete and available online, we chose a subset for detailed labeling. PRISMA-like details regarding video collection are provided in our supplementary material. To ensure a reasonably diverse corpus, we counterbalanced multiple race characteristics, including genders, stroke styles, race lengths, and race types. Finally, we selected 31 swimming race broadcasting videos. All races had standard swimming pools of 50 m or 25 m, with ten lanes, among which eight were used.

TABLE I: Statistics of video durations (in seconds) in our corpus, categorized by swimming event length.

Event (m)	Count (N)	Mean (s)	Min (s)	Max (s)
50	6	42.68	25.93	72.00
100	5	72.31	58.28	86.04
200	7	144.54	118.70	193.89
4×50	2	112.49	105.60	119.38
400	5	247.90	227.75	279.27
4×100	2	265.64	255.43	275.85
4×200	1	486.13	486.13	486.13
800	1	540.35	540.35	540.35
1500	2	938.27	925.02	951.52

We reviewed videos from the moment that all swimmers were positioned at the starting point and awaited the start signal, to the moment all swimmers had completed the race. Table I details the statistical distribution of video durations across different swimming events in our corpus. The dataset covers a broad spectrum of race distances. Short-length events, such as the 50 m, exhibit a mean duration of 42.68 s, whereas endurance events like the 1500 m extend to an average of 938.27 s. This wide range of durations spans from a minimum of 25.93 s to a maximum of 951.52 s.

To develop a set of codes for our review, two authors independently open-coded three videos. They then discussed their codes, coding challenges, and observations with co-authors. All authors are researchers in visualization and HCI, and have more than two years of experience in research on sports visualization. Through these discussions, we removed conflicts between the codes and coded visualizations. After two rounds of discussions, we finalized the criteria for data labeling, based on which the two coders completed the coding. We observed that the visualizations in our corpus could be coded using a design space outlined in previous work [56]. To make the coding manageable, we decided that the data behind the smallest visualization object we would code should pertain to race-related data, such as swimmers' metadata, time, ranking, and distance; in contrast to visual objects unrelated to the race itself and swimmers' performance, like sponsor logos, which we subsequently excluded from our review. In addition, the visualizations we coded as a whole appeared and disappeared visually as a unit, with or without entering and leaving animations. For visualizations that included staged animations, in contrast, our coding considered all visual elements that appeared within such animations. We also coded temporal patterns by focusing on: *when a visualization appeared* and *when it disappeared* relative to the race length. We labeled this temporal information at the video frame level. We coded the appearance of a visualization when its first component became visible (e. g., colored background or text); conversely, we coded disappearance on the frame when the last component disappeared from the video.

In addition to timing, we coded attributes of the visualizations' presentation, including the data visualized, visual components, bounding boxes, movements, animations, morphing, and placement on the screen. We selected these attributes after careful discussions among all the authors as the most relevant defining components for embedded visualizations in video that

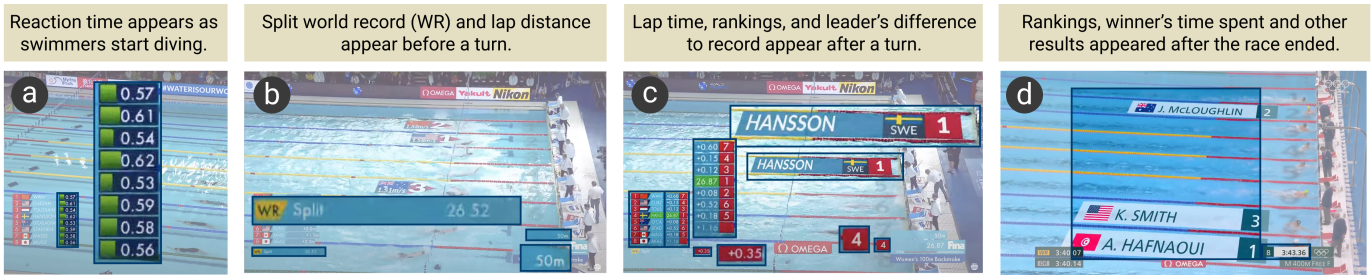


Fig. 2: Examples of temporal arrangements for embedded visualizations.Embedded visualizations appeared and disappeared according to specific race progress, including events like diving, turning, or the moment when swimmers were near a record.

are generalizable beyond swimming broadcasts to other sports contexts. We also coded the timing of key events (e. g., dolphin kick, turn) and camera shots used in the broadcasts. To improve the efficiency and accuracy of our coding, we developed a specialized coding tool with an interface for the coders to concretely mark the bounding box, starting and ending video frames for embedded visualizations. Further information about this coding tool and the specifics of the data coding criteria can be found in our supplementary materials. In the case of slight disagreement in duration values between the two coders ($< 1s$), we recorded the final duration as the average of the two values. If the difference exceeded 1 second or if two coders disagreed on data attribute options, they discussed the issue to reach consensus. If the disagreement could not be resolved through discussion, it was further discussed with all co-authors during our regular meetings.

B. Results

Overall, we coded 1544 visualizations from 31 videos that showed swimming races that took place between 2021 and 2024. The races of 13 videos took place at the Olympics, 12 at the World Championships, and 6 at the European Championships. Table I shows statistics about these videos. On average, 57 visualizations of 15 types appeared in each video, showing how prevalent these visualizations are in professional broadcasts. We associated each coded visualization with a distinct time interval in the video. Compared to the total video duration we reviewed, visualizations were visible for nearly 100% of the time. 361/1544 ◧ were visualizations in motion, defined as "visual data representations that are used in contexts that exhibit relative motion between a viewer and an entire visualization [55], [56]." The remaining 1183/1544 ■ visualizations were static. We provide more details on our reviewed videos, collected visualizations, and coded results in the supplementary material. Next, we report our findings, with a specific focus on temporal features. We categorized our collected visualizations according to their data updating frequency, the relevant events, and the embedding locations.

1) Data: We categorized the visualized data according to its updating frequency: *dynamic data* (used in 1056/1544 ◧ visualizations) would update during the race process, and *static data* (used in 912/1544 ◧ visualizations) did not update during the race. A visualization could combine both dynamic and static data. We followed the data classification of Yao et al. [56] but only report on categories of data we saw visualized.

Dynamic data included (a) time, (b) distance, and (c) speed. ►(a) *Time-related data* (used in 688/1056 ◧ dynamic visualizations) included reaction time, elapsed time, time of current lap, and lap time differences to a record. Apart from the elapsed time, which was displayed throughout the race, the current lap (and lap counts) and the lap time differences to a record always appeared together and were only shown after a turn. Similarly, the current ranking of swimmers was typically shown at the same time after a turn. The average duration of lap data per video was 11.30 s. The total display duration of lap data out of the total duration (timing ratio) of our analyzed videos was 19.72%.

►(b) *Distance-related data* (used in 282/1056 ◧ dynamic visualizations) included the distance swum, the remaining distance, and the distance difference to the leader for the two closest swimmers. The distance swum and remaining distance data served as an indicator of the race's progress. For the distance differences to a leader, this information appeared along a lane when no other specific event occurred (no turning or diving), but often when there was a tight competition among the first 3 swimmers. The average appearance duration for intermittently appearing swimming distances was 27.29 s in short and medium-length races and 168.45 s in longer races. Their duration was 42.16% of the race length for races <= 400 m while 66.82% for longer ones.

►(c) *Speed-related data* (used in 202/1056 ◧ dynamic visualizations) refers to the current speed of a swimmer. The speed label appeared typically when the current leader was close to the end of the race or when swimmers were close to a turn. The average duration of a visualization showing current speed was 7.57 s, and these visualizations were shown for 21.73% of the race lengths.

Static data included (a) swimmers' metadata and (b) record. ►(a) *Swimmers' metadata* (used in 506/913 ◧ static visualizations) included their nationalities, names, and lane numbers, and was shown in all reviewed videos. This data appeared at the start of the race (87×), when turning (326×), during the swimming phase (13×), and when close to the end (43×). Its average duration was 28.97 s, with a total duration of 33.27% of the race length on average.

►(b) Record-related data (used in 354/913 ◧ static visualizations) varied depending on the race type. In the Olympics, both Olympic and world records, along with lap splits, were shown near turns or the race's end. For other races, Olympic records were sometimes removed or replaced by the competition

record, while the world record always appeared. The average duration of record displays across races was 9.81 s, with an appearance timing ratio of 16.85% of the race length.

2) Race Events: We noticed frequent display of visualizations driven by swimming-specific events such as (a) the start, (b) underwater dolphin kicks, (c) turns, (d) the finish, and (e) regular swimming phases, depending on the distance to the side of the pool or the race length [9], [25]. These race events served as anchors for the appearance of specific visualizations. For instance, reaction times that indicate how quickly swimmers reacted to the start signal were typically reported after swimmers jumped into the water. Reaction-time visualizations appeared briefly after the race started and disappeared to offer the screen space for other dynamically updating data. The majority (1181/1544 ◧) of our collected visualizations were displayed before or after a *turn event*. The duration for visualizations around *turns* was also high with 21.02 s and a timing ratio of 10.70% of the race lengths.

3) Screen Placement: It determined when and for how long a visualization was shown, as visualizations required available space on the video. Static visualizations were usually positioned at some corners of the screen in fixed positions (see Fig. 1), sometimes during the entire race. We found 1069/1544 ◧ visualizations had constantly fixed positions on the *screen* (e.g., bottom left corner). The remaining 475/1544 ◧ visualizations were embedded in the lane, attached to swimmer in the pool, and did not have a dedicated *screen* position. Some of the visualizations embedded in the lane were not in motion, for example, the swimmer's name and nation shown at the start or end of the race. The average appearance timing of visualizations with fixed positions (62.77 s) was approximately 2× higher than the visualizations without a fixed placement (29.83 s).

C. Findings: Temporal Patterns

In our corpus, visualizations appeared and disappeared throughout the whole race. The data representations were mainly changed at turning events, and most had fixed positions on the screen. On the one hand, the timing and duration of visualizations depended on the meaning of the data and its associated events. On the other hand, continuously visible visualizations were placed on the side of the screen, which avoided cluttering the screen space that shows the swimmers. Overall, we found 8 general temporal patterns.

P1 The swimmers' metadata, including nationality (22/31), name (20/31), and lane number (22/31), appeared before the race started for an average duration of 6.09 s. This pattern exhibits substantial variability across videos ($SD = 3.56$ s, $IQR = 4.03$ s), with durations ranging from 1.60 s to 15.22 s. Olympics videos showed longer and more consistent durations ($M = 8.04$ s, $SD = 1.65$ s), whereas European Championships ($M = 6.34$ s, $SD = 5.15$ s) and World Championships ($M = 4.91$ s, $SD = 3.27$ s) displayed markedly greater dispersion.

P2 The reaction time appeared for an average duration of 5.68 s after swimmers jumped into the water (17/31) (Fig. 2 (a)). Across all videos, this pattern exhibited moderate variability ($SD = 1.02$ s, $IQR = 2.01$ s), with durations ranging from 3.86 s to 6.58 s. While European Championships showed relatively consistent reaction-time displays ($M = 6.25$ s, $SD = 0.36$ s), World Championships exhibited greater dispersion ($M = 5.36$ s, $SD = 1.14$ s). Notably, reaction-time visualizations were not shown in the Olympics.

P3 The elapsed time (31/31) was shown throughout the entire race (Fig. 1 (c)). On average, it appeared 1.38 s after the race started and disappeared 13.05 s after the race ended. The onset timing after race start was relatively consistent across videos ($SD = 0.28$ s, $IQR = 0.22$ s; range $\in$ [0.66 s, 2.02 s]). In contrast, the post-race persistence exhibited substantial variability ($SD = 13.87$ s, $IQR = 17.37$ s), ranging from 0.51 s to 47.42 s after race completion. Olympics showed short and consistent post-race durations ($M = 3.73$ s, $SD = 0.85$ s), whereas European ($M = 27.92$ s, $SD = 13.82$ s) and World ($M = 8.30$ s, $SD = 9.09$ s) Championships retained elapsed-time displays for substantially longer and more variable periods.

P4 Lap counts (27/31), split records (27/31), and the leader's distance swum (30/31) appeared before the turn with an average lead time of 4.06 s. This pattern showed moderate variability across videos ($SD = 1.48$ s), with European Championships being the most consistent ($SD = 0.58$ s) and Olympics ($SD = 2.11$ s) exhibiting occasional longer lead times.

P5 Rankings (15/31), leader's lap time (15/31) and the leader's lane number (20/31), the time differences between leader's lap time and the record lap split (26/31), and the lap time differences between others and the leader (26/31), appeared all together after a turn, with substantial variability across broadcasts ($M = 9.14$ s, $SD = 7.32$ s) (Fig. 2 (c)). European Championships showed the longest and most variable durations ($M = 16.69$ s, $SD = 8.68$ s), while Olympics videos exhibited earlier and more consistent post-turn appearances ($M = 4.63$ s, $SD = 1.15$ s). In our coded Olympics videos, there was no direct appearance of ranking numbers; instead, the swimmer's lap time was shown in a table, so the vertical order of visualization indicated a ranking.

P6 Record lines (21/31, $M = 9.21$ s, $SD = 7.62$ s) and speed labels (19/31, $M = 7.57$ s, $SD = 6.10$ s) typically appeared when the leader was close to the record (Fig. 1 (b)). World ($M_{record_lines} = 10.65$ s; $M_{speed_labels} = 10.00$ s) and European Championships ($M_{record_lines} = 8.78$ s; $M_{speed_labels} = 8.67$ s) tended to retain these visualizations slightly longer than Olympics ($M_{record_lines} = 7.90$ s; $M_{speed_labels} = 5.92$ s).

P7 Final visualizations that showed race results included rankings (31/31), winner's race result (30/31) and their lane number (31/31), and for other swimmers their time differences to the winner's result (18/31) appeared after the race ended for an average duration of 11.03 s, with substantial variability across broadcasts ($SD = 9.60$; range $\in$ [2.74 s, 47.09 s]) (Fig. 2 (d)). European Championships ($M = 24.56$ s, $SD = 12.34$ s) retained these visualizations markedly longer than World Championships ($M = 10.44$ s, $SD = 6.33$ s) and Olympics ($M = 5.32$ s, $SD = 1.68$ s).

P8 In races (27/31) whose lengths were equal to or less than 400 m, the swimming distance appeared at the current leader's turn. For even longer races [7], swimming distance was shown at several milestones (such as 500 m and 1000 m).

As the above data illustrates, while some patterns exhibited

relatively stable timing (P2 P3 P4), others demonstrated substantial variability (P1 P5 P6 P7). This variance highlights the presence of flexibility in broadcast production decisions. Furthermore, the context of broadcasts also matters. Specifically, the duration of embedded visualizations in the Olympics was shorter and more consistent compared to those in European and World Championships.

IV. Formative Study

To better investigate the specific requirements for developing a framework of temporal arrangements that is useful for real audiences, we conducted a formative study based on the re-analysis of data collected in prior work [56] and newly conducted structured interviews. We finished the two parts of the analysis independently. To give a complete picture, we report the merged findings below. All details are available in the supplemental materials. During the formative study period, we also met with and observed existing workflows at a local broadcasting company that helped us to get an informal picture of practices and live broadcasting challenges. We later conducted a formal practitioner interview with this broadcaster to evaluate our framework (see Sec. VII-A).

A. Procedure

Part 1 - Analysis of Existing Data: We first re-analyzed participants' answers to open-ended questions regarding their data preferences, sourced from a previous study [56] that investigated what data participants found interesting when watching swimming races. Eighty participants were included in this study (27 ♀, 52 ♂, 1 unspecified), with an age range from 18 to 75. More detailed demographics can be found in [56]. Our analysis focused on the temporal features mentioned in participants' answers, which had not previously been analyzed. Example responses we coded included *"Number of underwater kicks (after starts and turns)"* or *"Speed after turns (to evaluate turn efficiency)"*. In total, we found 35 descriptions from 18 participants that mentioned temporal arrangements from Yao et al.'s survey data.

Part 2 - Audience Interview: To complement the existing data analysis and gain deeper insights into audience preferences for temporal arrangements of data visualization in swimming races, we conducted interviews with actual swimming race viewers. Our interview received ethics approval from our institution's ethics review board (XJTLU, № ER-LRR-11000180720240828231725) and pre-registered the experiments on OSF (`osf.io/w2ykr`).

Our interview began with a pre-questionnaire collecting participants' gender, age range, and self-reported experience with swimming (on a 7-point Likert scale, 1: Novice and 7: Expert, noted as EwS below), familiarity with watching swimming race broadcasts, and experience with data visualizations. Participants then watched two professionally produced swimming race broadcast videos. We selected one race from the recent Olympic Games [34] and another from the Swimming World Cup [8], as these events represented the most advanced and standardized practices for data visualization in competitive swimming. After watching, participants were asked a set of questions regarding their perceptions of existing broadcast data and visualizations. Specifically, we asked whether the currently presented data aligned with their interests and preferences, whether the data were understandable, and which data types they found interesting but missing or insufficiently represented. To support reflection, we provided a reference list of swimming race data categories adapted from prior work [56]. We asked for participants' feedback on (1) the data they were interested in, (2) the moments during a swimming race at which they would like to view such data, and (3) the preferred duration or end moment of the data presentation. Additionally, participants rated their overall satisfaction with existing swimming broadcast visualizations using 7-point Likert scales and provided open-ended feedback on how the timing and temporal arrangement of visualizations could be improved. The whole study lasted approximately 45 minutes.

Participants were recruited through social media and university bulletin boards and received payment aligned with the local legal hourly salary. We had 11 participants with various experience with swimming (EwS) as a sport (ranging from novices to experts, 7/11 ♀, 4/11 ♂; age range from 18 to 40).

B. Findings

Two authors independently coded Parts 1 and 2. We merged the results through a consensus-building process for each part. We coded the data according to: (1) the category of data from Yao et al.'s study [56], and (2) the key events based on our metrics identified in the systematic review. Based on the results, we found and report several temporal arrangements of data that audiences were interested in but current broadcasts did not include (noted as FN below). The notation (x/18, y/11) indicates x occurrences from 18 participants (P1_1 to P1_18) in **Part 1** and y occurrences from 11 participants (P2_1 to P2_11) in **Part 2**:

F1 **Proper alignment of key events and related data.** Most participants (12/18 and 8/11) wanted to see relevant data displayed at the moment when key events occur. For example, P2_6 (swimming fan, EwS: 6/7) mentioned, *"I'd like to see the time taken at certain points, like when they touch the wall"*, highlighting the need for timely data presentation. P2_11 (a swimming coach, EwS: 7/7) emphasized, *"Put the data on screen as soon as it is collected"*, reinforcing the expectation for real-time updates and indicating a demand for more comprehensive data. This interest provides evidence for the usefulness of existing patterns P2, P4, and P5, where a series of results are presented for each key event. We further collected evidence of what these key events might be.

F2 **Qualitative and quantitative descriptions for events.** 29 participants (18/18, 11/11) described key moments using descriptive terms, such as *"overtaking"* and *"last lane,"* or precise data constraints, like *"first lane,"* "25 m," or *"per lap."* From an audience's perspective, an event may be primarily defined by the action or occurrence itself rather than the exact distance, even though swimming data is fundamentally based on distance and time, e. g., *"overtaking"*. Therefore, a transformation is needed between the audience's mental model

of race events and the collected data (which consists mainly of distance and time).

F3 **Adjustment on demand.** 6/11 participants in Part 2 highlighted the wish for flexibility in adjusting the duration of visualizations. As indicated in P1 to P7, existing embedded visualizations adopt fixed durations for all audiences, but such a one-size-fits-all design fails to accommodate diverse audience needs. For instance, *P*2_4 (a swimming amateur, EwS: 3/7) suggested, *"The display time might need to be a bit longer"*, while *P*2_9 (EwS: 2/7) mentioned that *"(Real-time ranking updates) should stay on the screen the whole time"*. Additionally, the audience expressed varying preferences for visualizations depending on the length of race, which reflects P8. For example, *P*2_2 (EwS: 4/7) noted, *"Because the original swimming time is long, it can appear in longer distance races, while short distance competitions may be more intense, and it can be replayed after the race"*. As a consequence, the viewing experience should be customizable to accommodate the diverse needs of different viewers. As *P*2_4 (a swimming amateur, EwS: 3/7) pointed out, *"Athletes could check boxes for data like stroke frequency... Amateurs might not need to check these. Allowing them to choose could also prevent the screen from becoming too cluttered."*

F4 **Global settings.** 6/11 participants from Part 2 mentioned issues regarding settings to manage aspects like appearance frequency and information load. *P*2_9 (EwS: 2/7) noted, *"After a turn, there is too much data to process, which feels overwhelming and frustrating"*, highlighting the need for more control over visual information load. *P*2_7 (EwS: 5/7) mentioned, *"If it's just a single leg, like a 50-meter segment, I don't think there's a need for too much data in such a short distance"*, emphasizing the importance of adjusting the data displayed based on the race context.

In summary, some temporal patterns align with the data–moment–duration combinations that participants explicitly expect, such as using key events as triggers for visualization changes. In contrast, other patterns reflect data facts observed in current broadcasting practice that do not conform to these expectations. By comparing the aligned and misaligned cases, we identify not only where existing conventions accommodate audience needs, but more importantly, where arrangement mismatches emerge. Such mismatches provide a foundation for deriving new design implications from a temporal and structural perspective.

V. Configuration Framework

Building on our findings from the systematic review, literature survey [9], [15], [25], [46], and formative study, we propose a three-level description structure for key moments in swimming races. This structure informs the development of a basic configuration framework that enables the flexible and precise definition of temporal arrangements in swimming race videos, with potential for adaptation to other sports contexts (see Sec. VIII-A).

A. Layered Description for Key Moments

Driven by F1 and F2, this three-layer structure is designed to encompass the full range of possible descriptions and requirements for key moments in swimming races, while being defined in a way that can be practically implemented and correctly transferred. Specific implementation examples of each layer can be found in our supplementary material.

Layer 1: Meta moment—basic data. At this foundational layer, key moments are described based on quantifiable data attributes, such as swimming style, distance, or time points, etc. These can be directly derived from raw race data without considering temporal sequences or trends. For example, a turn event can be defined as occurring when a swimmer's distance swum equals $n \times l$, where n is the lap count and l denotes the lane length. This layer supports binary or threshold-based detection, e.g., a particular swimming style, duration, or when distance swum enters/leaves a predefined range.

Layer 2: Time series—temporal dynamics. This intermediate layer characterizes short-term evolutions of basic data metrics, transforming static values into dynamic conditions. This intermediate layer extends beyond static data points to consider how key attributes evolve over time within a defined interval. By analyzing dynamic changes in basic data, it can describe events such as "approaching," "overtaking," and "acceleration." For instance, an approaching event occurs when the distance between two swimmers gradually decreases to a certain value over a period. Similarly, an overtaking event can be defined as the moment when a swimmer's position changes from N to $(N-1)$ within Δt. Moreover, acceleration detection requires analyzing velocity changes across consecutive time steps rather than isolated speed values.

Layer 3: Event sequences—semantic narratives. This is the highest abstraction layer, combining lower-level moments, time-series events, or interactions between swimmers into semantic narratives. For example, a "chasing" event can be defined as a swimmer first approaching another and then overtaking them. It not only connects basic events into meaningful sequences, but also allows for qualitative enhancements, such as describing an event as "tight chasing" or "fierce competition", which can be transferred as the event when ranking changes happen frequently. These descriptions also require thresholds for values or frequency to quantify the intensity of qualitative attributes.

B. Framework for Atomic Descriptions

The design considerations for our framework revolve around practical questions about embedding visualizations in a video. To embed a visualization, it is necessary to make several design decisions:

- *Content*: which specific data is visualized?
- *Placement*: where on the video is the visualization placed?
- *Design*: which encoding, layout, etc., are used?
- *Appearance*: when does the visualization appear?
- *Disappearance*: when does the visualization disappear, or for which duration is the visualization visible?

Driven by these factors and the interest expressed by our participants (F1–F4), our configuration framework is structured into three key parts: visualization design, temporal triggers, and global constraints (Fig. 3 (a)). The visualization design configuration includes the first three design decisions. The

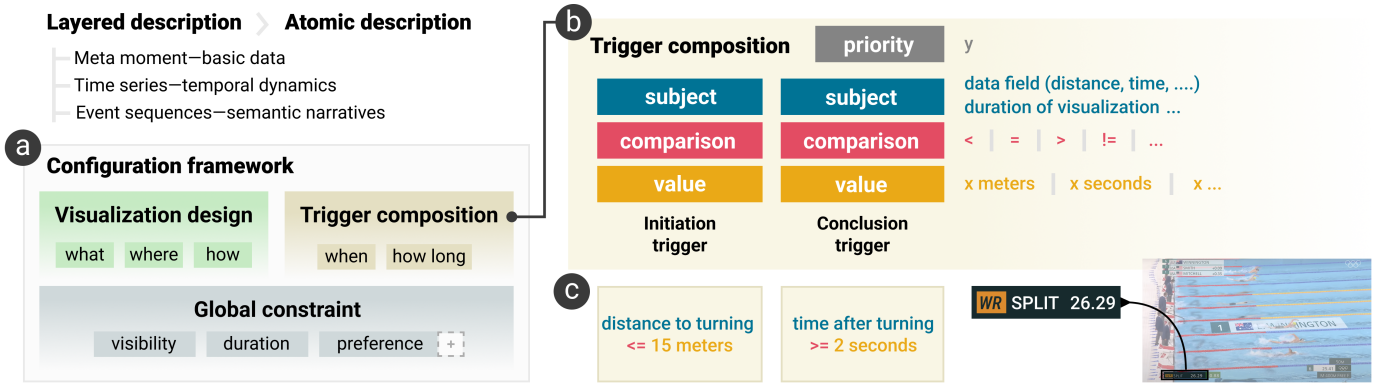


Fig. 3: (a) The configuration framework consists of three parts. The visualization design configures *what* data to show, *where* to embed, and *how* to visualize the data. The temporal trigger composition defines *when* and for *how long* to show the visualization. Global constraints ensure the visibility and duration of visualizations maintain reasonable and readable values. These constraints can also be expanded to include audience preferences or other necessary constraints tailored to different usage scenarios. (b) Each temporal trigger is defined by subject, comparison, value, and priority. (c) The showing intervals of each visualization are decided by a composition of triggers, including an initiation trigger and a conclusion trigger.

temporal trigger configuration concerns the last two related to timing. The global constraint configuration defines global rules to maintain an audience-friendly viewing experience.

In addition, we aim to make these configurations transferable between videos of similar sports events that may differ in specific conditions, such as the length of the race or the participating athletes. This goal is motivated by the understanding that individual viewers have diverse interests in data presented in sports videos [30], [56] and adjustment on demand (F3). Transferable visualization arrangements enable the audience to create customized sports videos once tailored to their unique data and visualization preferences, thereby enhancing their dedicated viewing experience [30] while reducing effort in creating tailored experiences for similar events.

1) Visualization Design: Visualization design includes non-timing-related factors that need to be configured for embedded visualizations. The three factors (what, where, and how) do not need to be considered in a fixed order (Fig. 3 (a)). First, we need to specify the data to be shown in the visualization. For example, we may want to filter the swimmers from one nationality and focus on their speed. Second, we need to determine a location on the video to place the visualization. For static visualizations, a specific area on the video needs to be chosen (such as the bottom left corner). For visualizations in motion, a data referent needs to be chosen [55]. This referent needs to have position metadata. Third, choosing the visualization design involves making choices regarding readability, aesthetics, and possibility to engage an audience.

2) Temporal Trigger Composition: Temporal triggers organize the timing of embedded visualizations in videos as a sequence of intervals. Specifying the temporal information of such an interval requires the definitions of both start and end [48]. Therefore, a temporal trigger should determine the exact moment when a visualization will appear or disappear.

A temporal trigger is essentially a data-driven condition that, once it meets a specific data statement, initializes or concludes the showing of a visualization. We include both initiation and conclusion triggers (Fig. 3 (b, c)). An initiation trigger determines the start moment when a visualization appears. The conclusion trigger is a condition that determines when the visualization should disappear. Note that the specific condition for a conclusion trigger can be either based on dynamic data related to sports or a pre-set duration (P1 - P7 and F1 - F3). Therefore, there are two kinds of trigger compositions: *"start-duration"*, and *"start-end"*. Given the dynamic data available in a swimming race and to make the triggers transferable across videos, each trigger in our framework is organized as a data statement. These data statements follow the layered description structure introduced in Sec. V-A, specifically at the meta moment level. The data statement comprises three key components, ***Subject***, ***Comparison***, and ***Value*** (Fig. 3 (b)):

Subject can be either the visualization duration (such as 10 s) or a data field (speed, distance swum, the difference between two numerical data, etc). If the ***Subject*** is a duration, this trigger is a conclusion trigger, of type *"start-duration"*. If the ***Subject*** is a dynamic data field, this trigger is either an initiation trigger or a conclusion trigger of type *"start-end"*. Based on the factors we previously identified in Sec. III-C, we categorize the variables for activating the triggers across three dimensions: spatial data, temporal data, and event data. Several spatial and temporal data are directly accessible in advanced detection techniques applied in sports scenarios, such as distance, position, and time. Based on the source data provided by detection methods, more data, such as speed and rankings, can be calculated for visualization and analysis. Event data served as a commonly used temporal anchor in previous sports visual analysis studies [15], [19], marking the key moments or periods in sports. In swimming, given the fixed length of both the pool and the race, the event data is highly related to race progress, which can be inferred from the distance swum [9], [25].

Comparison is the comparative relation between the ***Subject*** and ***Value***. When the ***Subject*** is the duration of a visualization, the ***Comparison*** should be "equal". ***Value*** is a threshold to activate the trigger, while ***Subject*** decides its unit.

In addition, each trigger composition should have a defined activation ***Priority***. The ***Priority*** is a relative value that decides, in case multiple triggers are defined that conflict with each other for the same visualization. For example, in a 400 m race, an audience wants to see swimmers' speeds per 50 m lasting for five seconds, and also wants to see the speed when the swimmers from their nation rank in the top three. These two trigger definitions might generate timing conflicts, e.g., the distance swum in the race is now 50 m, but the swimmer whom the audience focuses on is not in the top three. Under such circumstances, only the trigger with higher priority will fire, deciding the time intervals for the speed visualization.

3) Global Constraints: They are used to ensure the readability and preferred viewing experience of embedded visualizations (F4). We abstract our findings into three factors: ***visibility***, ***duration***, and ***preference constraints***. The scope of each constraint could be further expanded and tailored to meet the practical requirements for specific scenarios:

Visibility constraint refers to the possible visibility of the embedded visualization in sports videos and its viewport. For example, visualizations may only appear during specific camera shots, or they must fit within available screen space when shown. Moreover, on devices of different sizes, visualizations may require layout adjustment to preserve more space for the main video content [52]. Such considerations need to be clarified globally, as they create hard constraints for embedding visualizations. Visibility constraints influences the placement, readability, aesthetics, and visibility of visualization, i. e., the *where* and *how* of a visualization design (Fig. 3 (a)).

Duration constraint refers to the appropriate duration for which the visualizations should appear. While trigger compositions strictly determine the time intervals for embedded visualizations on video timelines, they may not always result in reasonable or effective temporal arrangements. For example, if the ***Subject*** in a trigger is a frequently updating data field, such as speed or ranking, especially in highly competitive scenarios, the predefined ***Comparison*** and ***Value*** can cause visualization to flicker on and off at very fast intervals. To prevent this, appropriate durations ensure viewers have sufficient time to notice and interpret the visualizations for insights. Thus, a global minimum duration value is essential to ensure visualization readability. Duration constraints, thus, influence the duration of visualization, i. e., the *how long* of a trigger composition (Fig. 3 (b)). When conflicting with trigger compositions, this constraint will have higher priority to decide the duration of visualizations.

Preference constraint refers to general audiences' interests in seeing certain embedded data visualizations [30], [56] (F3). These preferences can span a wide range, including a focus on specific athletes and relevant data, etc. Therefore, we consider this constraint as a global filter to support possible preference configurations for diverse audiences. Preference constraints influence the data to be shown in the visualization, i. e., the *what* of visualization design (Fig. 3 (a)).

VI. SwimChrono: A Technology Probe

We implemented a technology probe called *SwimChrono* to integrate our configuration framework. Different from *SwimFlow* [56], *SwimChrono* defines data-driven temporal arrangements for embedded visualizations.

A. Panel Introduction

SwimChrono has three primary panels: video panel, visualization panel, and timeline panel (Fig. 4):

1) Video Panel: Users can preview, play, and pause both the video and embedded visualizations (Fig. 4 (a1), 6).

2) Visualization Panel: Users can configure visualization design parameters (Fig. 4 (a2), 3). In *SwimChrono*, we extend the visualizations and configuration settings of *SwimFlow* [56] with those commonly observed in current broadcasts.

3) Timeline Panel: It presents an editable timeline with layers of embedded visualizations (Fig. 4 (a3)). This panel consists of three primary functions: ***timeline***, ***trigger***, and ***template***:

Timeline represents the temporal arrangements of visualizations as multiple interval event sequences. Our interface draws inspiration from Adobe Premiere Pro [2]. The horizontal axis is the video timeline, from start (left side) to end (right side). Visualizations are arranged in vertical layers, with higher layers having higher visibility priority (Fig. 4 (a): 2 5). Each layer corresponds to a single visualization configuration, including the visualization design and temporal triggers. On the timeline, the visibility of each visualization layer is shown by a sequence of intervals. Each interval represents its duration, and the moment the visualization appears and disappears. The left and right borders of an interval indicate the triggers that define its timing. All intervals have a green left border, which means an initiation trigger starts them. The right border is either purple or grey, depending on whether the end is determined by a *"start-end"* or *"start-duration"* trigger. By default, an embedded visualization lasts for the entire video.

Trigger activates after users click the trigger button (Fig. 4 (a): 4). Users can assign multiple triggers with different priorities to one visualization layer through data selection or AI interpretation. Through these, users can define the triggers with both specific and abstract descriptions (F1 and F2). For instance, users can customize the trigger's name, type, and priority. They can also specify the ***Subject***, ***Comparison***, and ***Value*** for both the initiation and conclusion triggers independently. We integrate large language models (LLMs) to organize temporal arrangements based on natural language input, for their excellence in such tasks [41], [57]. Each AI interpretation returns triggers in the format of our configuration framework. This AI functionality assists users in translating their ideas into a structured trigger configuration within our framework. If the generated configuration does not fully meet their needs, users can further refine it manually. We also suggest triggers based on the patterns P1 - P7. Once users confirm the trigger compositions, *SwimChrono* calculates the temporal arrangements based on the data and then updates the visualization layer's display intervals accordingly.

Template allows users to export their customized arrangement templates and import a predefined template for videos (Fig. 4 (a): 7). A template captures all the detailed specifications of embedded visualizations in the race video, based on the configuration framework. Users can export a JSON file in a

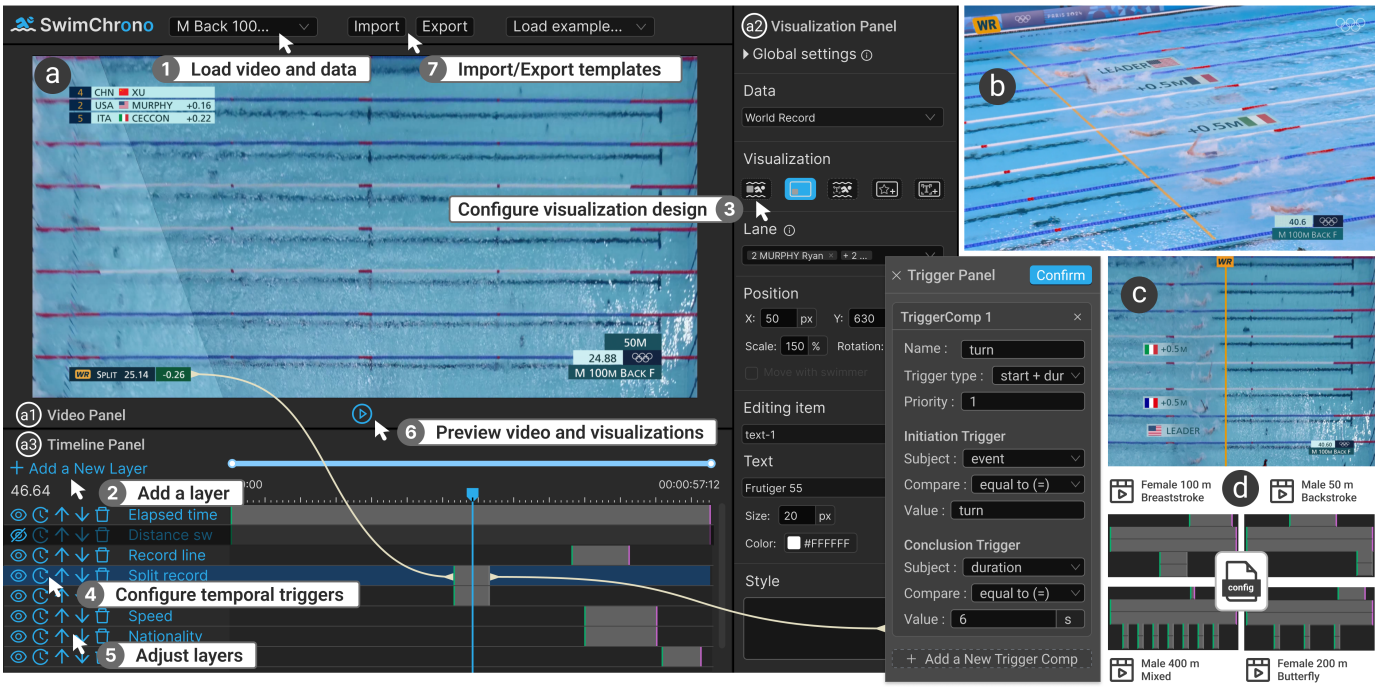


Fig. 4: *SwimChrono* consists of three primary panels: (a1) **Video**, (a2) **Visualization**, and (a3) **Timeline** panels. Users can configure the temporal arrangements of embedded visualizations. For example, (a-c): *SwimChrono* supports multiple visualization designs and the temporal triggers to arrange them; and (d): *SwimChrono* allows users to transfer their own visualization templates across more diverse races.

Domain-Specific Language format commonly used in toolkits (e. g., [39]). Later, they can apply similar arrangements and personal preferences across different race videos.

B. Implementation

SwimChrono is an entirely new web-based technology probe, implemented from scratch by using TypeScript and React, and is to support the trigger-based display of the visualizations and a timeline. In *SwimChrono*, the AI function utilizes DeepSeek's "deepseek-chat" model API [18] (default parameters, temperature=1.0) with tailored prompts that convert natural language inputs into framework-compliant trigger compositions. Comprehensive prompt details and event detection rules are available in the supplementary materials. The video dataset follows the same pre-processing Computer Vision techniques as *SwimFlow* to synchronize and generate orthographic projections of two videos into one. We expanded the dataset with three additional self-recorded videos to prevent any copyright issues. As such, these videos do not have the quality and viewpoints of professional broadcasts but can be freely used for research and publication. We manually annotated the videos to track swimmers' positions and cycle events using the same annotation tool [31] as in *SwimFlow*. Moreover, we incorporated metadata (e. g., world record, swimmer names) from various public swimming websites.

To ensure the readability of embedded visualizations, we set several global constraints based on findings from our systematic review. For example, we set the minimum visualization duration to 2 s, as 95% (1439/1544) of our coded visualizations lasted longer. We also assigned default placements for commonly used visualizations based on their most common positions in our statistics. These constraints remain adjustable in the visualization panel.

VII. Evaluation

To evaluate our proposed framework, we interviewed a professional sports broadcaster, presented several illustrative use cases, and conducted a comparative user study. Both our comparative user study and expert interview received ethics approval from our institution's ethics review board (XJTLU, № ER-LRR-11000180720240828231725).

A. Expert Interview

To evaluate the proposed framework and its implementation as a technology probe, we conducted a semi-structured interview with a professional TV broadcaster to gather qualitative expert feedback (pre-registered on OSF: `osf.io/px3nk`). The interview was conducted via video call, with two researchers (authors) present.

1) Participants: Swimming broadcast experts are difficult to reach. We contacted four broadcast experts in our collaboration network, one of whom ($E1$) agreed to participate. $E1$ is the leader of a broadcasting company with 10 years of experience in broadcasting swimming races. Participation in this interview was entirely voluntary without cash compensation.

2) Interview Procedure: Before starting, our participant signed a consent form, answered demographic questions, and described their general broadcasting experience. We then introduced our project background, configuration framework,

and technical implementation, and demonstrated how *SwimChrono* supported the customization of temporal arrangements for embedded visualizations based on race data. Then, we asked the participant a set of open-ended questions about the framework's alignment with live broadcast logic, its potential impact on existing broadcasting workflows, and the feasibility of integrating *SwimChrono* into current production pipelines. We also asked additional questions to understand which workflow steps could benefit from integration, which might be unnecessary, and what additional features could further improve broadcasting practice. The whole interview lasted approximately 50 minutes.

3) Data Analysis & Results: We transcribed and analyzed the participants' responses with the following results:

Comprehensive coverage: $E1$ confirmed that the framework captures *"every parameter and trigger that is needed"* and addresses gaps in existing tools. He noted that current software (e.g., Sport In The Box [27]) offers few adjustable temporal parameters, relies on basic triggers (e.g., Start, Split Time, Stop), and often requires manual intervention, whereas the proposed framework provides greater flexibility, including support for continuous and mid-lap graphics.

Transferability: According to $E1$, directors actually have "no time" to change visualization parameters during a race, making full pre-configuration and automation essential. The participant further emphasized the value of integrating functionality into a single system to avoid managing multiple computers. This aligns with our support of pre-configured templates for different races (see Sec. VII-B).

Long-distance races' complexity: $E1$ noted that long-distance races (e.g., 1500 m) differ from shorter races (e.g., 100 m), as the leader may finish while others are still racing. This creates conflicts: a winner graphic must be shown for the leader while simultaneously displaying lap splits for trailing swimmers, requiring expert prioritization logic to manage overlapping states within the framework.

Overall, $E1$ was excited about configuring visualizations using our framework and *SwimChrono* in practice, if they can be smoothly integrated into the software he was familiar with.

B. Illustrative Use Cases

To demonstrate the capabilities and wide applications of *our configuration framework implemented in* SwimChrono (hereafter just referred to with the probe name *SwimChrono*), we present three illustrative use cases: recreating current practice, transferring configurations, and exploring novel usage scenarios beyond the state of the art. The demonstration video of these cases is included in the supplementary material.

1) Case 1: Recreating Olympic TV Broadcasts. This case illustrates that *SwimChrono* is expressive enough to replicate standard temporal arrangements used by current TV broadcasters. We selected the Men's 100 m Backstroke final at the Paris 2024 Olympic Games [33], as it was expected to be a widely watched event with high production quality and precise visual presentation. Screenshots from the TV broadcast (Fig. 4 (b)) showcased visualizations such as elapsed time, distance difference, and world record. To recreate the temporal arrangement, we loaded a self-recorded video of the same race into *SwimChrono* and then added several visualizations (Fig. 4: ❶❷❸) to match the original timing. For each visualization, we refined their style and placement to make them resemble the broadcasting design and adjusted temporal triggers for data, including the top 3 rankings, the leader's lap time difference to the record, and the top 2 and top 3's lap time difference to the leader, which were all triggered by the turn event (Fig. 4: ❹). To achieve this, we configured a *"start-duration"* trigger composition, setting "key event" as the ***Subject*** and "turn" as the ***Value*** for the initiation trigger. We set the duration of this visualization to six seconds after each turn. Then, we configured other visualizations and their temporal triggers based on the recommendation patterns P1, P3, and P6. By playing the video, we noticed an overlap between the record line and the bottom-right visualizations. To minimize visual clutter, we fine-tuned the layer order (Fig. 4: ❺❻). After all adjustments, we successfully configured embedded visualizations in the Olympic broadcast style (Fig. 4 (a, b, c)). This demonstrated that our configuration framework and its implementation enabled us to capture, and automate the visualization timing practice commonly seen on TV.

2) Case 2: Transferring Configuration. This case showcases that our configuration framework is flexible enough to transfer a design from one race (e. g., 100m) to another (e. g., 50m), or to a different race type (e. g., swimming style). The architecture of our framework enables this transferability, which separates race properties and events (data) from visualizations and triggers. Thus, when the data changes, the trigger compositions remain valid and adapt automatically. Still, some minor adjustments might be necessary, especially for races with significantly different pacing, where visualizations might overlap or need timing modifications. These adjustments could be easily specified in *SwimChrono* and handled by global constraints that add new trigger conditions. To achieve this in *SwimChrono*: we first select a race video with an existing temporal arrangement (e. g., 100m) and adapt it for another race (e. g., 50m). As shown in Fig. 4 (d), the timeline updates accordingly. One difference is that, since 50m races have no turn events, triggers associated with lap times are not activated. Additionally, because the race results are displayed on the left side (where swimmers finish in a 50m race), the finish rankings need to be repositioned (on the opposite side). Other minor adjustments might include shortening visualization durations and reducing the number of visual elements due to the race's fast pace (F4). All these modifications are documented as an updated configuration that would apply easily to other 50m races. The same process would be undertaken to adapt the designs for longer races (e.g., 200m, 400m, 800m, 1500m) or tailor visualizations to specific audiences. Overall, the templating ability of *SwimChrono* allows to save considerable time because visualizations and timings have to only be slightly adjusted than entirely recreated.

3) Case 3: Going Beyond TV Broadcast Standards. SwimChrono has demonstrated exceptional flexibility and adaptability in aligning with current sports TV broadcast standards. Yet, *SwimChrono* also supports a broader level of temporal

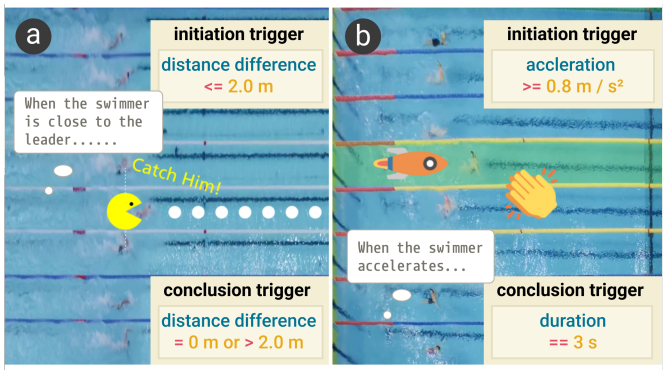


Fig. 5: Screenshots of the embedded visualizations in Case 3: Pacman style. We configured visualizations to highlight (a) the distance difference and (b) the focused swimmer.

arrangements. To illustrate this benefit, we introduce a non-standard humorous visualization using a Pac-Man-inspired design, driven by a *"start-end"* trigger composition (Fig. 5 (a)). In this visualization, the white circles represent the second fastest swimmer's remaining distance, while a yellow Pac-Man in the same lane indicates his or her position. The Pac-Man appears behind the leader when the second-fastest swimmer is within a specific distance (e. g., 2 m), alerting the audience to a possible overtake. Beyond this simple example, *SwimChrono*'s architecture supports even more diverse visual augmentations to enhance the race-watching experience. For example, in Fig. 5 (b), triggers highlight swimmers during acceleration, activating visualizations when acceleration exceeds a threshold, usually for a short duration. These triggers and accompanying visualizations are not currently available in professional broadcasts. The definition of these triggers is simplified through a natural language interface (see the demo in the supplementary material), and once confirmed by the audience, *SwimChrono* updates the timeline accordingly. These visual designs could reveal subtle competitive and tactical aspects of the sport. Additionally, novel themes for embedded visualizations—whether gamified, fun, calm, or serious—could be explored to enhance audience engagement. Importantly, *SwimChrono*'s ability to go beyond the state of the art enables researchers and practitioners to generate test videos for empirical studies, including evaluations of how new temporal arrangements and designs affect audience perception, readability, and understanding.

C. Comparative User Study

In order to understand the value of showing visualizations continuously versus scheduling them, we conducted a short first experiment with representative temporal patterns informed by our study of the Olympic broadcast. We primarily wanted to measure participants' preferences for each technique to understand design trade-offs and potentially confirm the value of offering a temporal framework. Our user study (pre-registered on OSF: `osf.io/mcsw6`) was a within-subject experiment to compare two temporal arrangements of embedded visualizations in swimming race videos: *Scheduled* and *Always-on* modes. In *Scheduled* mode, visualizations are temporally arranged using the proposed framework, while for *Always-on* mode, visualizations remain continuously visible without event-based temporal configuration. Both modes are supported by our framework.

1) Stimuli Selection: The stimuli that we used were two different 100 m swimming race videos with embedded visualizations edited by *SwimChrono*. We selected 100 m races because they include a turning event that serves as a key event for presenting race data, while remaining short enough to avoid excessive viewing time per trial. The data presented, and the visualization designs (e.g., embedding locations, size, shape, and color) remain consistent across modes and follow a style similar to Olympic broadcasts in terms of data type, visualization type, and appearance. The only differences between modes were the timing of appearance and disappearance and the duration of display. The data we selected included the race name, current lap, and elapsed time. We also showed information for the top three swimmers, such as lane number, name, nationality, speed, ranking, and distance to the leader. In addition, our data included the Olympic record, the world record, the split world record, and the difference between the split world record and the overall world record.

2) Study Procedure: Our study consisted of a consent form, a pre-questionnaire, the formal trials, and a post-questionnaire. In the pre-questionnaire, we collected participants' gender, age, and experience with swimming, watching swimming races, and interpreting visualizations. Participants completed two controlled and counter-balanced trials, each involving video watching followed by comprehension questions and subjective ratings of interference, visual clutter, immersion, and overall satisfaction. Additionally, we included two attention checks (e. g., selecting a specific number on a Likert scale) to ensure participants concentrated on the study. In the post-questionnaire, participants indicated an overall preference and provided open-ended feedback. The whole study lasted approximately 10 minutes.

3) Participants: We recruited participants through Qualtrics [38] as well as the authors' networks. We excluded participants who did not complete our study or failed attention checks and did not use their data in the analysis. In total, we had 34 valid participants. Each participant received remuneration aligned with their local legal hourly salary.

4) Data Analysis: To analyze subjective measures, we used the paired Wilcoxon signed-rank test, as the paired differences violated normality for most outcomes (Shapiro–Wilk tests). We applied the Holm–Bonferroni correction to the subjective measures to control the family-wise error rate. We tested comprehension accuracy through a non-inferiority analysis with a margin of $\Delta = 0.05$, defined as $D = Accuracy_{Always-on} - Accuracy_{Scheduled}$. We conducted a thematic analysis of the open-ended comments that we received. Two authors independently initiated coding, then discussed disagreements and merged their coding books.

5) Results: Participants' **preference** between *Scheduled* and *Always-on* modes was relatively balanced; *Scheduled* mode was slightly more often preferred (18/34) than *Always-on*

mode (16/34). When asked whether they wanted to adjust temporal arrangements, most participants (23/34) answered yes: 14/23 participants preferred *Scheduled* while the remaining 9/23 preferred *Always-on*. Our thematic analysis further explains this preference and the trade-offs between the two modes: *Scheduled* was valued for supporting race understanding and focus on swimmers' performance (7/18), with less visual interference (4/18) and more manageable information load (3/18). For example, P1 expressed, "I want to watch the players' performance for around 70% of all time and the rest 30% for some specific timing with visualizations (like very beginning, turning, and near the end point)." Participants favoring *Always-on* emphasized continuous awareness (4/16) and persistent visual reference to avoid losing context (4/16), with some also noting greater immersion (2/16), less distraction (1/16), and finding this mode easy to follow (1/18, 2/16). For example, P15 wrote, "I like staying informed at all times." Comments such as these suggest a trade-off between focused, context-aware presentation and continuous information availability. *Scheduled* helps reduce visual clutter and supports attention to key race moments, while *Always-on* provides persistent references that maintain situational awareness but may increase information load. Other participants reported that their choice was primarily driven by personal preference (2/18, 1/16) or that they had no clear preference (1/18, 2/16).

The **comprehension accuracy** for *Scheduled* mode ($M = 0.66$) was marginally higher than that for *Always-on* mode ($M = 0.63$), though the available sample size limits definitive conclusions about the reliability of this difference. The point estimate favored *Scheduled* ($M_{diff} = -0.03$), but the one-sided upper confidence bound (0.15) exceeded the non-inferiority margin. We measured **subjective experience** using a 7-point Likert scale (1=strongly disagree, 7=strongly agree) with 7 items: interference (s_1), visual clutter (s_2), immersion (s_3), naturalness (s_4), information sufficiency (s_5), overall satisfaction (s_6), and whether they want this in broadcasts (s_7). Four Likert items—interference (s_1), visual clutter (s_2), immersion (s_3), naturalness (s_4)—showed directional trends that *Scheduled* provides a better experience, while the overall satisfaction nearly has no difference ($M_{diff} = 0$). However, when applying one-sided tests with Holm correction for multiple comparisons, none of the items reached statistical significance.

6) Summary: Our quantitative results suggest that the difference between *Scheduled* and *Always-on* was inconclusive, while *Scheduled* has slight observable advantages over *Always-on* across preference, comprehension accuracy, and several subjective experience dimensions. However, these trends do not survive conservative multiple-comparison correction and therefore should be treated only as suggestive patterns warranting further investigation. Such further investigations could be inspired by the qualitative findings, which provide a clearer picture of possible tradeoffs between the two conditions. Most participants desired user-adjustable temporal arrangements of embedded visualizations, particularly those who preferred *Scheduled* mode. Overall, this user study showed that both arrangements involved trade-offs and elicited different audience preferences and wishes. This, in turn, demonstrates the necessity of configurable temporal arrangements and motivates further user studies on this topic (see Sec. VIII-B and Sec. VIII-C).

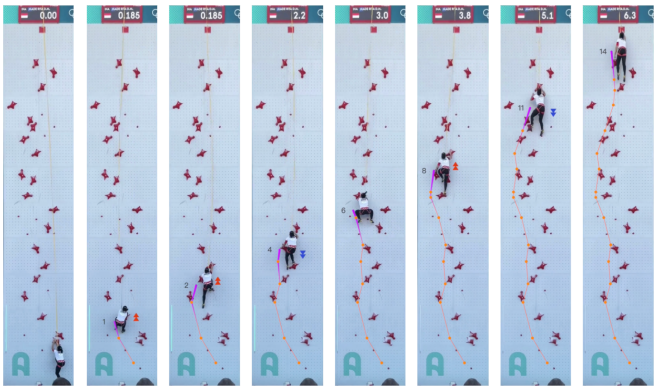


Fig. 6: Screenshots of the embedded visualizations in a climbing video. The embedded visualizations present: (1) an action count updated with changes in the athlete's movement, (2) acceleration/deceleration visualized through red/blue icons, and (3) the athlete's movement trajectory over time.

VIII. Discussion

We discuss how our work informs situated and sports visualizations, its benefits and limitations, and future work.

A. Extendability and Generalization

The patterns we identified reflect strategies for temporally arranging embedded visualizations in current practices. These patterns serve as baseline evidence that enables subsequent studies and practitioners to build upon. For example, the distribution of fixed and dynamic positioning can establish a starting point for future work on audience comprehension, perception, and design trade-offs. The formative study went beyond the state of the art and provided us with additional requirements, which we incorporated into our configuration framework. The framework is powerful and easily extensible to integrate new visualization designs and temporal triggers. As *SwimChrono* implements such patterns, it enables users to experiment with visual designs efficiently while supporting extensions for more diverse embedded visualizations in swimming race videos.

Our framework is designed for broad applicability across various swimming races. Yet, we hypothesize that the framework could potentially be extended to other speed-dominant sports with sufficient space and time for embedded visualizations, such as rowing [1], climbing, or certain running events. In these contexts, athletes remain in constrained trajectories and share similar key metrics (e.g., time, distance, speed, rankings). However, such extensions would require domain-specific adaptations and have not yet been validated beyond swimming. As a preliminary illustration, we implemented temporal arrangements for a speed-climbing video (see Fig. 6 and supplementary materials).

The ***Subject*** can also extend to multiple modalities, such as camera shots and sound, enabling a more context-aware,

immersive, and accessible broadcasting experience [26]. We used top-down videos we recorded during a national event and the Paris 2024 Olympics, which naturally limited available camera perspectives. However, this limitation is not important to our main contribution, which focuses on race metadata triggers. Adding triggers to switch between camera feeds and synchronize them with our framework is straightforward. Moreover, professional broadcasts already tie camera switches to sports events, which can be readily integrated into our framework in future extensions [37].

Our configuration framework holds potential for applications in situated sports visualizations, particularly in AR/VR scenarios. In these scenarios, visualizations can appear, update, and disappear in response to changes in detected environmental data. Such a dynamic nature of embedded visualizations, triggered by surrounding data, is similar to animated embedded visualizations in videos and falls within the scope of our framework. However, AR/VR introduces additional challenges, such as the impact of viewing angles, viewports, and movement speed, which require further consideration.

B. Design Implication

Our work treats temporal arrangements as an independent design dimension for embedded visualizations in sports broadcasts. We abstracted away from specific visualization designs or geometric encodings. This separation allows us to systematically characterize recurring temporal arrangements that generalize across different visualization types. Our work cannot, without risk of unjustified overgeneralization, derive prescriptive or universally optimal design guidelines. More empirical research is needed to test different temporal arrangements across different audiences, as indicated by our formative study and user study, where audience preferences varied across users and contexts. Such studies are clear future work, pointed out below. Our descriptive framework provides a flexible and extensible foundation that supports future temporal arrangement design innovations and controlled evaluation of perceptual effectiveness.

C. Limitations and Future Work

We limited the number of surveyed videos due to the time-intensive labeling process and focused on swimming, as visualization timing varies across sports. No user study is without limitations. We were only able to recruit one broadcasting expert, which necessarily limits our qualitative domain assessment to the potential and integration of our workflow to this one very positive example. The exploratory user study in Sec. VII-C was useful in uncovering the first tradeoffs between showing information continuously versus time-limited. Further studies and dedicated papers are needed to assess other points of the design space or further research questions regarding how much information can be tracked at a time, how visually salient information should be, or how best to fade information in and out. Our work lays the groundwork for new research avenues in embedded sports visualizations. For example, upcoming studies could use our configuration framework and *SwimChrono* to generate trials for empirical research, thereby expanding understanding of how embedded visualizations facilitate audiences' data-driven insights and viewing engagement.

IX. Conclusion

Currently, audiences watch swimming races with uniformly preset temporal arrangements of embedded visualizations provided by TV broadcasters. However, the structure of data-driven temporal trigger conditions governing the appearance and disappearance of these visualizations over time has not been thoroughly studied. In response to this gap, we explored patterns of temporal arrangements for embedded visualizations in current practice and surveyed audiences on their preferences regarding which information to view at which specific key moments. Based on our findings, we proposed a framework for configuring visualizations in a data-driven way. This framework enabled the transferable and data-driven configuration of embedded visualizations in swimming videos. To make this framework practical and accessible, we implemented it in a free technology probe, *SwimChrono*, to support multiple usage scenarios related to embedded visualizations. A comparative user study using videos edited with our framework revealed that most participants preferred having user-adjustable temporal arrangements of embedded visualizations. Furthermore, an expert interview and several illustrative use cases demonstrated the powerful capabilities and expressiveness of our configuration framework, showcasing its ability to create, transfer, and customize temporal arrangements of embedded visualizations for swimming race videos. Our work contributes to the understanding of temporal visibility and arrangements for situated visualizations, providing a labeled dataset on embedded visualization labeling in current swimming broadcasting practices, a series of design patterns, a formative study and its findings, a framework, and a technology probe for future studies related to visualization in motion.

Acknowledgments

The authors wish to thank reviewers for all their feedback and suggestions. This research is partially supported by the Open Project Program of the State Key Laboratory of CAD&CG (Grant № A2514), Zhejiang University; XJTLU RDF, grant № RDF-24-01-062; ANR, the French National Research Agency, with the SPORTSVIZ project, grant № ANR-24-CE33-4945.

References

[1] L. V. 37th America's Cup Barcelona. Preliminary Regatta Jeddah - Day 3 LIVE. Youtube, 2023. Retrieved March 9th, 2024 from https://www.youtube.com/watch?v=A7CikJZvgKI&t=3214s.

[2] Adobe. Adobe Premiere Pro. Website, 2024. Retrieved March 9th, 2024 from https://www.adobe.com/products/premiere.html.

[3] F. Amini, N. Henry Riche, B. Lee, C. Hurter, and P. Irani. Understanding data videos: Looking at narrative visualization through the cinematography lens. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, p. 1459–1468, 2015. doi: 10.1145/2702123.2702431

[4] F. Amini, N. H. Riche, B. Lee, A. Monroy-Hernandez, and P. Irani. Authoring data-driven videos with DataClips. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):501–510, 2017. doi: 10.1109/TVCG.2016.2598647

[5] P. Andrews, O. E. Nordberg, S. Zubicueta Portales, N. Borch, F. Guribye, K. Fujita, and M. Fjeld. Aicommentator: A multimodal conversational agent for embedded visualization in football viewing. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*, p. 14–34, 2024. doi: 10.1145/3640543.3645197

[6] O. Aquatics. 18 year-old Ahmed Hafnaoui's road to gold in men's 400m freestyle! Youtube, 2022. Retrieved March 9th, 2024 from https://www.youtube.com/watch?v=ZWbxgzDNyzg.

[7] O. Aquatics. 1500m freestyle final at Tokyo 2020 — FULL LENGTH. Youtube, 2023. Retrieved March 9th, 2024 from https://www.youtube.com/watch?v=9sHPbtj5Tj8.

[8] W. Aquatics. Morning heats budapest — day 3 — world aquatics swimming world cup 2023. Website, 2023. Retrieved Jan 4th, 2026 from https://youtu.be/YXGC4c2A0CE?list=PLg25OJAWYpDLUl95p8D_so1wUIckZzfTO&t=5040.

[9] T. M. Barbosa, A. C. Barbosa, D. Simbaña Escobar, G. J. Mullen, J. M. Cossor, R. Hodierne, R. Arellano, and B. R. Mason. The role of the biomechanics analyst in swimming training and competition analysis. *Sports Biomechanics*, 22(12):1734–1751, 2023. doi: 10.1080/14763141.2021.1960417

[10] V. S. Camera. Veo Sports Camera. Website, 2024. Retrieved March 9th, 2024 from https://launch.veo.co/.

[11] CapCut. CapCut. Website, 2025. Retrieved March 20th, 2025 from https://www.capcut.com/.

[12] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in Information Visualization: Using Vision to Think*. The Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann Publishers, San Francisco, CA, USA, 1999.

[13] 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 ACM Conference on Human Factors in Computing Systems (CHI)*, article no. 841, 18 pages, 2023. doi: 10.1145/3544548.3581266

[14] 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, 2023. doi: 10.1109/TVCG.2022.3209497

[15] 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

[16] N. Cohn. Visual narrative structure. *Cognitive Science*, 37(3):413–452, 2013. doi: 10.1111/cogs.12016

[17] N. Cohn. *From Visual Narrative Grammar to Filmic Narrative Grammar: The narrative structure of static and moving images*, pp. 94–117. Routledge, 2016. doi: 10.4324/9781315692746

[18] DeepSeek. DeepSeek. Website, 2025. Retrieved March 28th, 2025 from https://api-docs.deepseek.com/.

[19] D. Deng, J. Wu, J. Wang, Y. Wu, X. Xie, Z. Zhou, H. Zhang, X. L. Zhang, and Y. Wu. EventAnchor: Reducing human interactions in event annotation of racket sports videos. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, article no. 73, 13 pages, 2021. doi: 10.1145/3411764.3445431

[20] S. M. Drucker and D. Zeltzer. Camdroid: a system for implementing intelligent camera control. In *Proceedings of the 1995 Symposium on Interactive 3D Graphics*, p. 139–144, 1995. doi: 10.1145/199404.199428

[21] M. T. Fischer, D. A. Keim, and M. Stein. Video-based analysis of soccer matches. In *Proceedings of the International Workshop on Multimedia Content Analysis in Sports*, p. 1–9, 2019. doi: 10.1145/3347318.3355515

[22] Q. Galvane, R. Ronfard, C. Lino, and M. Christie. Continuity editing for 3d animation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 29(1), 2015. doi: 10.1609/aaai.v29i1.9288

[23] T. Ge, B. Lee, and Y. Wang. CAST: Authoring data-driven chart animations. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, article no. 24, 15 pages, 2021. doi: 10.1145/3411764.3445452

[24] T. Ge, Y. Zhao, B. Lee, D. Ren, B. Chen, and Y. Wang. Canis: A high-level language for data-driven chart animations. *Computer Graphics Forum*, 39(3):607–617, 2020. doi: 10.1111/cgf.14005

[25] T. Gonjo and B. H. Olstad. Race analysis in competitive swimming: A narrative review. *International Journal of Environmental Research and Public Health*, 18(1), 2021. doi: 10.3390/ijerph18010069

[26] G. Gourmelen, S. Toriya, E. Miya, N. Shioura, and H. Iwata. Miruoto: Sports event atmosphere visual rendering through real-time image and sound processing system. In *ACM SIGGRAPH 2024 Emerging Technologies*, article no. 14, 2 pages, 2024. doi: 10.1145/3641517.3664393

[27] S. in the Box. Sport in the box support. Website, 2024. Retrieved Jan 20th, 2026 from https://support.sportinthebox.com/en-US.

[28] Y. Kim and J. Heer. Gemini: A grammar and recommender system for animated transitions in statistical graphics. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):485–494, 2021. doi: 10.1109/TVCG.2020.3030360

[29] C. Lee, T. Lin, H. Pfister, and C. Zhu-Tian. Sportify: Question answering with embedded visualizations and personified narratives for sports video. *IEEE Transactions on Visualization and Computer Graphics*, 31(1):12–22, 2025. doi: 10.1109/TVCG.2024.3456332

[30] 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

[31] Neptune. Neptune. Website, 2025. Retrieved March 28th, 2025 from https://centralelyon.github.io/annotation/?competition=2024_JO_Paris&course=2024_JO_Paris_dos_hommes_100_finaleA.

[32] Olympic. Olympic. Website, 2025. Retrieved March 20th, 2025 from https://www.olympics.com/.

[33] Olympics. Men's 100m freestyle - final — swimming — olympic games paris 2024. Website, 2024. Retrieved March 29th, 2025 from https://www.olympics.com/en/video/men-s-100m-freestyle-final-swimming-olympic-games-paris-2024?uxreference=playlist.

[34] Olympics. Men's 4x100m medley relay - final — swimming — olympic games paris 2024. Website, 2024. Retrieved Jan 4th, 2026 from https://www.olympics.com/en/video/men-s-4x100m-medley-relay-final-swimming-olympic-games-paris-2024.

[35] J. Owens. *Television Sports Production*. Routledge, 6 ed., 2021.

[36] Y. Pan, Y. Chen, Q. Bao, N. Zhang, T. Yao, J. Liu, and T. Mei. Smart director: An event-driven directing system for live broadcasting. *ACM Trans. Multimedia Comput. Commun. Appl.*, 17(4), article no. 119, 18 pages, 2021. doi: 10.1145/3448981

[37] Y. Pan, Y. Chen, Q. Bao, N. Zhang, T. Yao, J. Liu, and T. Mei. Smart director: An event-driven directing system for live broadcasting. *ACM Trans. Multimedia Comput. Commun. Appl.*, 17(4), article no. 119, 18 pages, 2021. doi: 10.1145/3448981

[38] Qualtrics. Qualtrics. Website, 2026. Retrieved Jan 20th, 2026 from https://qualtrics.com.

[39] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer. Vega-Lite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350, 2017. doi: 10.1109/TVCG.2016.2599030

[40] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010. doi: 10.1109/TVCG.2010.179

[41] L. Shen, Y. Zhang, H. Zhang, and Y. Wang. Data Player: Automatic generation of data videos with narration-animation interplay. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):109–119, 2024. doi: 10.1109/TVCG.2023.3327197

[42] D. Shi, F. Sun, X. Xu, X. Lan, D. Gotz, and N. Cao. AutoClips: An automatic approach to video generation from data facts. *Computer Graphics Forum*, 40(3):495–505, 2021. doi: 10.1111/cgf.14324

[43] X. Shu, A. Wu, J. Tang, B. Bach, Y. Wu, and H. Qu. What makes a data-GIF understandable? *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1492–1502, 2021. doi: 10.1109/TVCG.2020.3030396

[44] M. Sports. Metrica Sports. Website, 2024. Retrieved March 9th, 2024 from https://metrica-sports.com/?_go=nicholas21.

[45] W. Sports. WSC Sports. Website, 2024. Retrieved March 9th, 2024 from https://wsc-sports.com/.

[46] M. Stein, T. Breitkreutz, J. Haussler, D. Seebacher, C. Niederberger, T. Schreck, M. Grossniklaus, D. Keim, and H. Janetzko. Revealing the invisible: Visual analytics and explanatory storytelling for advanced team sport analysis. In *International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, pp. 1–9, 2018. doi: 10.1109/BDVA.2018.8534022

[47] 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, 2018. doi: 10.1109/TVCG.2017.2745181

[48] J. Tang, F. Yang, J. Wu, Y. Wang, J. Zhou, X. Cai, L. Yu, and Y. Wu. A comparative study on fixed-order event sequence visualizations: Gantt,

extended Gantt, and stringline charts. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2024. doi: 10.1109/TVCG.2024.3358919

[49] J. R. Thompson, Z. Liu, and J. Stasko. Data Animator: Authoring expressive animated data graphics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, article no. 15, 18 pages, 2021. doi: 10.1145/3411764.3445747

[50] V. L. Vizrt. Viz Libero - Vizrt. Website, 2024. Retrieved March 9th, 2024 from https://www.vizrt.com/products/viz-libero/.

[51] W. Willett, Y. Jansen, and P. Dragicevic. Embedded data representations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):461–470, 2017. doi: 10.1109/TVCG.2016.2598608

[52] A. Wu, W. Tong, T. Dwyer, B. Lee, P. Isenberg, and H. Qu. Mobilevisfixer: Tailoring web visualizations for mobile phones leveraging an explainable reinforcement learning framework. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):464–474, 2021. doi: 10.1109/TVCG.2020.3030423

[53] X. Xu, L. Yang, D. Yip, M. Fan, Z. Wei, and H. Qu. From 'wow' to 'why': Guidelines for creating the opening of a data video with cinematic styles. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, article no. 599, 20 pages, 2022. doi: 10.1145/3491102.3501896

[54] J.-H. Yao, J. Feng, X. Tian, K. Xu, G. Amirkhanova, and S. Chen. Beyond the broadcast: Enhancing vr tennis broadcasting through embedded visualizations and camera techniques. *IEEE Transactions on Visualization and Computer Graphics*, 32(1):802–812, 2026. doi: 10.1109/TVCG.2025.3634638

[55] 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, 2022. doi: 10.1109/TVCG.2022.3184993

[56] 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, 2024. doi: 10.1109/TVCG.2023.3341990

[57] L. Ying, Y. Wang, H. Li, S. Dou, H. Zhang, X. Jiang, H. Qu, and Y. Wu. Reviving static charts into live charts. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–16, 2024. doi: 10.1109/TVCG.2024.3397004

[58] YouTube. YouTube. Website, 2025. Retrieved March 20th, 2025 from https://www.youtube.com/.

[59] J. Zong, J. Pollock, D. Wootton, and A. Satyanarayan. Animated VegaLite: Unifying animation with a grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):149–159, 2023. doi: 10.1109/TVCG.2022.3209369

BIOGRAPHY SECTION

Junxiu Tang is currently a postdoctoral fellow at Northwestern University. He received his Ph.D. in Computer Science from Zhejiang University. His research focuses on the design and interaction of data-driven communication, storytelling, spatial-temporal visual analysis, and event sequence visualization. For more information, please visit https://ahugh19.github.io/.

Lijie Yao is an Assistant Professor of Computer Science at Xi'an Jiaotong-Liverpool University. She is interested in investigating how motion characteristics and spatial factors affect the perception, design, and interaction of visualizations. For more information, please visit `lijieyao.com`. Contact her at yaolijie0219@gmail.com.

Lu Ying received her Ph.D. in Computer Science from Zhejiang University. She is currently a postdoctoral fellow at Nation University of Singapore. Her research lies at the intersection of human–AI interaction, visualization, and machine learning, where she design human-centered AI methods that empower people in sensemaking and decision-making. For more information, please visit https://yiyinyinguu.github.io/.

Romain Vuillemot is an Assistant Professor of Computer Science at Centrale Lyon, Université de Lyon, France, in the LIRIS laboratory. His research interests include information visualization and visual analytics, for application domains such as sport, trajectory-based data, and deep learning explicability.

Petra Isenberg is a research director(DR) at the Inria Saclay Centre at Université Paris-Saclay, France in the Aviz team and part of the Computer Science Laboratory (LISN) of the University Paris-Saclay. She is particularly interested in exploring how to design and present visualizations to the general public in novel device contexts, such as non-planar displays, tabletops, or smartwatches and in new application contexts like sports or video games.