A Case Study on the Design and Use of an Annotation and Analytical Tool Tailored To Lead Climbing

Bruno Fruchard
bruno.fruchard@inria.fr
Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL
Lille, France

Cécile Avezou
c.avezou@ffme.fr
Fédération Française de la Montagne et de l'Escalade
Paris, France

Sylvain Malacria
sylvain.malacria@inria.fr
Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL
Lille, France

Géry Casiez
gery.casiez@univ-lille.fr
Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL
Lille, France

Stéphane Huot
stephane.huot@inria.fr
Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL
Lille, France

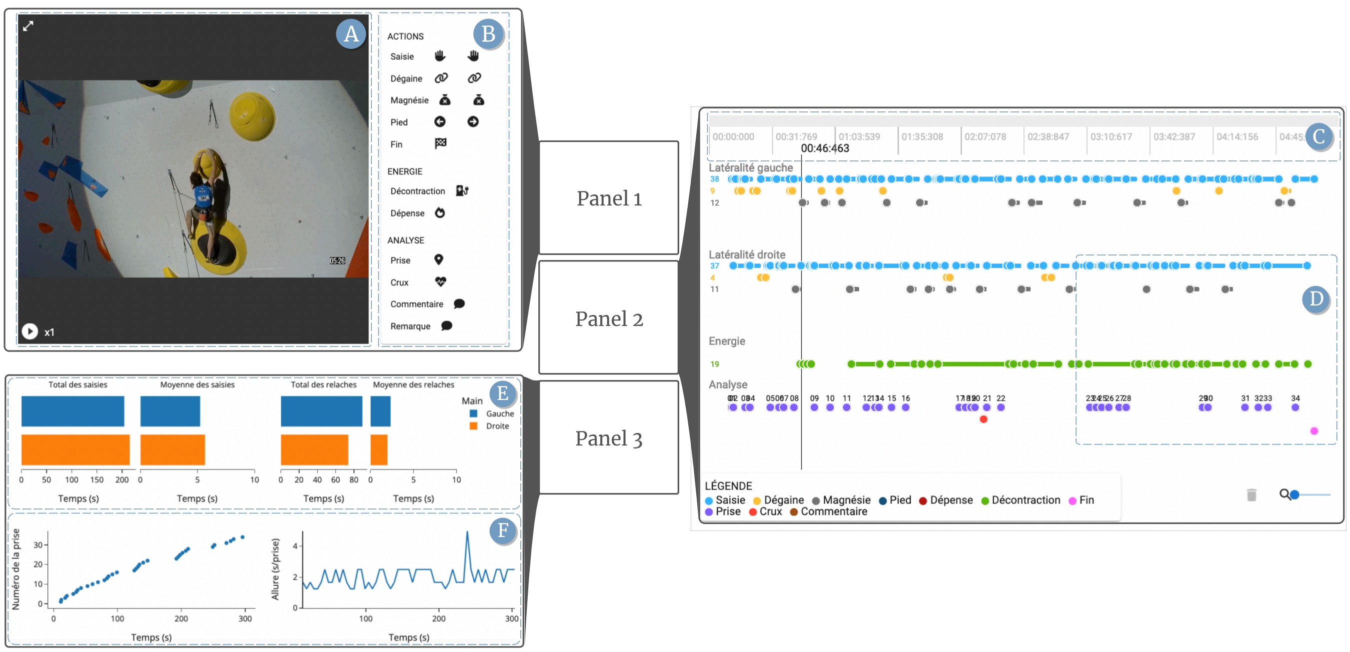

Figure 1: The tool consists of three panels: one to play the video stream and create annotations, a second to visualize them, and a third to summarize the quantitative information they represent. The first panel presents the performance video stream (a) and a sequencing panel to create annotations (b). The second panel displays the timeline (c) with annotations depicted as circles and lines that represent transient or lasting actions (d). The last panel presents plots that depict the cumulated or average holding and resting times for both hands (e), or the score evolution and climbing speed (f).

ABSTRACT

Annotating sport performances enables to quantitatively and qualitatively analyze them, and profile athletes to identify their strengths and weaknesses. We present the case study of the design and use of an annotation and analytical tool tailored to lead climbing analysis, developed with and for the French climbing federation. We used an iterative design cycle mostly fueled by virtual meetings with the federation trainer and analyst to identify requirements and implement essential features over time. We complemented these meetings with two workshops involving them, as well as French athletes competing at the international level, to identify the tool advantages and limitations. We contribute a list of insights based on the design process and feedback from stakeholders that inform the design of annotation and analytical tools for lead climbing and potentially other sports.

CHI EA '23, April 23–28, 2023, Hamburg, Germany

The definitive Version of Record was published in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*, https://doi.org/10.1145/3544549.3573876.

CCS CONCEPTS

• **Human-centered computing** → *Human computer interaction (HCI).*

KEYWORDS

case study, video annotation, sport analysis, user-centered design

ACM Reference Format:
Bruno Fruchard, Cécile Avezou, Sylvain Malacria, Géry Casiez, and Stéphane Huot. 2023. A Case Study on the Design and Use of an Annotation and Analytical Tool Tailored To Lead Climbing. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3544549.3573876

1 INTRODUCTION

Producing quantitative data related to sport performances enables to statistically analyze them and provide objective measures on the athlete efforts [18, 25]. Commercially available tools such as DartFish [10] or Hudl Sportscode [1] enable to create and associate such data to video sequences. They facilitate reviewing sequences of an athlete performance by using time-coded annotations, and ultimately facilitate the production of large data sets to profile them. The primary limitations of commercial tools lie in the fact that they are not tailored to specific sports, are not easily customizable or open source, are expensive, and do not provide direct means for analyzing a performance.

Ideally, analysts annotate all possible events happening during a performance, but this is highly time-consuming when videos require many annotations. A common compromise consists in leveraging crowdsourcing approaches to split the workload between workers [22, 42], use the crowd in public events to annotate them [31], or extract data from social networks [41]. These approaches are focused on public events and do not accommodate to sensitive data such as training sessions of international athletes. Another compromise is to identify and focus on a small set of relevant metrics that one can annotate in an acceptable time frame. Identifying the right metrics to annotate is challenging for all sports as many factors can influence an athlete performance [29, 35, 44].

This work focuses on lead climbing, a discipline with significant history that was included only recently to the Tokyo Olympics 2020 [8]. The goal of this discipline is to climb a wall composed of artificial holds under 6 minutes[1]. Athletes climb with a rope they clip on quickdraws attached to the wall to ensure security in the event of a fall. The score is evaluated based on the number and type of holds grasped, and the ascent duration[2]. As part of the PerfAnalytics project [30] founded by the ANR (Agence National de la Recherche [12]), we worked closely with the French climbing federation (Fédération Française de la Montagne et de l'Escalade - FFME [11]) to design an annotation tool tailored to lead climbing performances. Before starting this project, the FFME had only a precedent with the Meta-Video [28] annotation tool that seemed challenging to customize to their needs. Because of the cost and limitations of existing tools, we tailored one to specifically answer their needs for supporting data production and analyzing performances.

[1]this time can vary; we take the Tokyo Olympics as a baseline
[2]for further details, see [16]

We present our design process and contribute a list of insights based on the lessons we learned along the way. We applied user-centered design [38] through an iterative design process involving the official lead climbing trainer and analyst to identify the right metrics to annotate, and design functionalities supporting the analysis of climbing performances. This process also involved athletes competing at the international level to understand how this tool can support them in reviewing their performances. We collaborated through weekly virtual meetings over a span of 5 months, and organized two physical workshops with staff members and athletes to evaluate the efficacy of the tool for both annotation and analysis. The outcome of this collaboration is a list of insights that inform the design of annotation tools for lead climbing and possibly other sports to some extent. These insights span from the level of details provided by different annotation types, to the iterative process for annotating performances, and the need for functionalities to compare performances.

2 RELATED WORK

In this section, we discuss the literature on notational and physiological analyses of climbing sports. We present existing annotation tools for sport, image, or behavioral analysis and tools supporting the analysis of sport performances. We then present methods for automatically creating data about athlete performances based on computer vision and wearable approaches.

2.1 Manual Annotation

Various measures can be used to evaluate the climbing performance of an athlete. Physiological tests, carried out before or after a performance, provide detailed data on how athletes use their body. They enable, for instance, comparing oxygen intake between rock climbers and artificial climbers [5], comparing the climbing-specific strength between boulder climbers and lead climbers [15], or identify determinants for success in climbing [35, 44]. However, they do not always adapt to competitions as they may require specific equipment, can induce fatigue, and may conflict with the competition schedule.

Other approaches consist in annotating performances by hand with or without tool support [18]. Several commercial annotation tools focus on sport analysis. They support annotating time events during a performance [1, 10], or add visual annotations on a video stream to highlight specific sections [9] or compute trajectories and movement speeds [26]. Academic tools propose similar functionalities to analyze video content for sport performances [19, 40], support behavioral analyses [2, 17, 20, 24], or build training data sets for machine learning [3, 4, 14]. The main drawback of manual annotation is that it is time-consuming. Crowdsourcing offers a solution by splitting the workload between a set of workers [22], but raises privacy issues that are essential when dealing with videos of professional athletes.

Based on discussions with the FFME staff, and to the best of our knowledge, there currently exist no conventional tools for analyzing lead climbing performances. Most tools offer customizing their sequencing panels (manually or by asking the provider) to adapt to various sports, but characterizing the panel features and updating it based on the evolving needs is tedious and can be expensive. We tailored a tool to the FFME needs to propose short design cycles fueled by frequent discussions with the analysts that enabled quickly adapting the tool features based on their feedback.

2.2 Limitations of Automatic Approaches

Data such as body or hand postures can also be extracted from video feeds using computer vision (CV) approaches [7, 39]. Seifert et al. [37] used this type of data to study skill transfer in different climbing environments, and Reveret et al. [33] used it to estimate the speed and position of an athlete for speed climbing. Unfortunately, these visual approaches remain error-prone: occlusions are frequent when the climber moves their hand in front of their body, and when the video captures the entire wall, a frequent case in competitions to capture several athletes climbing at the same time, the size of the climber is likely too small to provide a good resolution.

Data can also be collected from wearables fixed on the athlete's body while climbing. Analyzing data from Inertial Measurement Units attached to the pelvis and limbs enables to recognize one's movements [6, 13] or the route climbed [21], and identify new indicators of efficiency [36]. The ClimbAX system [23] proposes to leverage this type of data for self-training purposes. Augmenting athletes with sensors is, however, not possible in most competitions and can hinder training sessions (e.g., preparation time, props may hinder movements).

We designed the annotation tool to facilitate characterizing a climbing performance happening in the context of a training session or a competition. This tool solely relies on a video recorded before starting the annotation session, and does not use any kind of visual or wearable automation at the moment. Analysts, trainers, or athletes can upload a video and start annotating immediately.

3 DEFINING THE REQUIREMENTS AND DESIGNING THE TOOL

We have been working continuously with the FFME since May 2022 by organizing virtual meetings online every two weeks. We organized two physical workshops during this period to test various versions of the tool and get valuable feedback from professionals. Figure 2 summarizes our schedule for the last months by marking significant events.

The tool was integrated in a suite of applications designed for the FFME as a part of the PerfAnalytics project. These tools build on the Dash full-stack framework [32] that relies on React.js [27] for the front-end. They facilitate uploading and indexing videos to easily search for performances to annotate. It is important to note that the schedule in Figure 2 presents only important events with regard to interaction design, and does not list technical problems faced along the way that slowed down the development of the tool (e.g., integration of custom React code to Dash, adapting the app to desktop and mobile environments, etc.).

Initial virtual interviews with the trainer and analyst of the federation in May outlined that the tool should have two main purposes: support analysts in quickly annotating a performance, and help them analyze the outcomes of a performance after producing annotations. In the following, we present design considerations that emerged in initial meetings, then present the type of annotations the tool supports, and explain how functionalities evolved through time.

3.1 Design considerations

We dedicated the first virtual meeting to identifying design considerations through an interview with the FFME analyst and the trainer. The meeting consisted in the experts explaining their annotation needs and sharing their experience with the Meta-Video [28] tool to point out its limitations in this regard. The main drawbacks were the lack of features to quickly review sequences in a video and features to facilitate the analysis of a performance. The following design considerations are also inspired by the various tools in the literature (e.g., [2, 4, 14, 20, 24]).

DC 1 *The tool should facilitate both annotating and analyzing climbing performances.*
Most annotation tools focus solely on producing and gathering data, but do not provide visualization means to analyze a video and ideally reveal interesting visual patterns. The tool should automatically produce reports supporting the analysis.

DC 2 *Annotations should represent transient and lasting events to enable reviewing specific sequences.*
The tool should support two types of events and help visualize them. Lasting events are represented by two annotations (start and end) and correspond, for instance, to grasping holds. Transient events are represented by a single annotation and correspond, for instance, to a change in the score. The former enables identifying duration and sequences of events (e.g., both hands holding together), and review them quickly by replaying sequences of the video.

DC 3 *Annotations should be frame-precise and the tool should provide precise time control to support that.*
Precisely evaluating the duration of events is essential to analyze a performance. Annotations must be set precisely at the start and end of actions, for instance, when grasping and releasing a hold. Interactions must support quickly skipping sections of the video as well as providing finer control when needed.

DC 4 *Annotations should account for both quantitative and qualitative data.*
Objective (physical actions) and subjective (interpretation of actions) measures are both important for analysis. Annotations should represent factual events as well as athlete comments entered in post hoc reviewing sessions.

DC 5 *The tool should be accessible to any stakeholders (trainers, analysts, athletes) to facilitate rapid adoption*
Sport professionals have particularly tight schedules that must accommodate competitions and training sessions, and offer little time to experiment with interactive tools that might, without certainty, help them improve their skills. A supporting tool should therefore provide few and simple interactions, and straightforward means for analysis to facilitate its adoption.

3.2 Annotation Types

Climbing essentially consists in efficiently using the 4 limbs to make use of the holds and the wall in its entirety to ascend. The most essential information is how are the limbs used at a given point

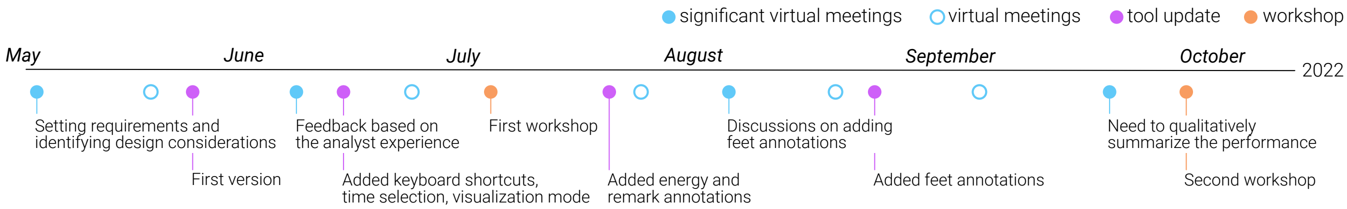

Figure 2: Interaction process with the FFME over 5 months. This process builds on one-hour virtual meetings every other week, and two physical workshops used to evaluate the tool with the analysts and athletes.

in time, which is of quantitative nature. Another is to understand the athlete's decisions leading to more or less physical movements, which is of qualitative nature.

In the following, we present the initial annotation types supported by the tool, and the ones that emerged through experience and discussions.

3.2.1 Quantitative annotations. The tool supports three types of lasting hand actions (DC 2): 1) grasping a hold[3], 2) clipping the rope, 3) dipping the hand into the chalk bag. These annotations provide the lowest possible level of information for each hand, but have the advantage of being factual and do not require the analyst's subjective view. More descriptive annotations could, for instance, describe the type of grasping and the clipping quality but would require more time to analyze.

The other low-level annotation related to the athlete's performance describes how they finished the route. They can either top out the route, fall while climbing, or are required to stop climbing for technical reasons.

Higher-level annotations depend on both the athlete's performance and the route characteristics. The score associated to a hold, for instance, varies depending on the competition. Analysts must be aware of the details of the route setting (information conventionally provided during competitions) to correctly annotate an athlete performance. Similarly, the cruxes of a route, i.e., the sections identified as being the most challenging, depend on the route setting and can be defined for a specific route. These annotations relate to metadata about a route that is external to the athlete performance.

A new category of quantitative annotations emerged after annotating an initial set of videos and completing the first workshop (see Figure 2). This category relates to the energy consumption and includes annotations that mix objective and subjective measures; the analyst must interpret the athlete movements as being relaxing or effortful. The first annotation type denotes an effort due to a mistake, and the other denotes muscle relaxations. This type of annotation was added to complement the low-level ones, and provide more indicators of the performance quality. This especially helps providing detailed feedback on an athlete climbing strategy, or their mistakes. For instance, no relaxation periods might indicate a decision in the climbing strategy or a mistake.

After the trainer and analyst gained experience with the tool, discussions led to creating annotations for the feet (see Figure 2). The analyst annotated this information on several videos and remarked it required significantly more time to completely annotate a performance. The main reason was that foot actions are error-prone; how the foot is used on and off the wall (e.g., balancing, resistance, heel hook) is difficult to assess. While this annotation is still present in the tool, it is currently not used. This revealed the importance of the trade-off between the annotation production time and their outcome, and showed that some annotation types are not worth focusing on.

[3]colors correspond to the color mapping used by the tool, see Figure 1

Figure 3: Keyboard shortcuts for creating annotations and controlling time in the video based on the analyst requirements.

3.2.2 Qualitative annotations. We added athlete comments as annotations in the initial version of the tool, following (DC 4). These annotations were a request from the FFME trainer who wanted support for reviewing sessions with athletes. They consist of free-form textual information associated with categories. Initial categories consisted, for instance, in the athlete's *emotions*, how they *perceived* external events, their *physical shape*, and their overall *feelings*. More categories were added later such as *technique*, *strategy*, or *state-of-mind*, and were informed by several feedback sessions where these categories did not match the athlete's comments.

We added a second type of qualitative annotation after the first workshop (Figure 2) to let analysts add remarks (text without categories) based on the athletes' comments or to describe something remarkable in the performance.

3.3 Implementation: Layout and Functionalities

At the moment of writing, the tool cannot be made open-source for confidentiality reasons, but we hope to open it soon after clearance. The tool consists of three panels that serve distinct purposes (Figure 1).

3.3.1 Panel 1: Video and Sequencer. The first panel presents a video of the performance and a sequencing panel that contains all annotation types.

The video can be played at half or double the normal speed, and the spacebar can be used to pause and play it (DC 3).

The sequencing panel consists of one button for each annotation type. Clicking on a button adds an annotation at the exact frame currently displayed (DC 3). Based on the analyst feedback

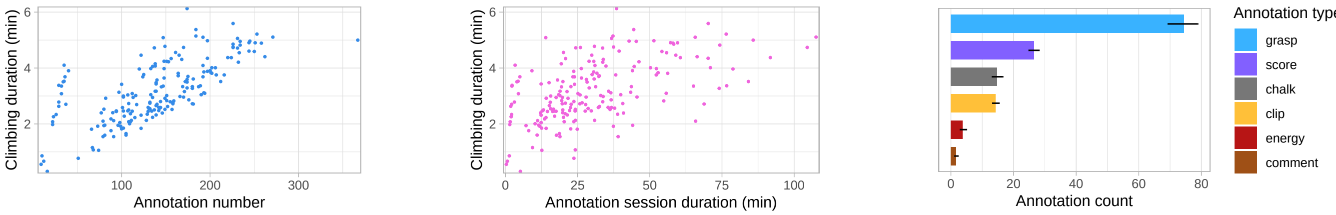

Figure 4: Annotation data for 197 videos. Plots depict the annotation count (left) and duration of annotation sessions (middle) in relation to the climbing duration for each video, and the annotation count for several annotation types (right). Error bars depict 95% confidence intervals.

through the design process, we added keyboard shortcuts for each annotation type to support full keyboard control when annotating a performance (Figure 2). We leveraged a spatial mapping often used for video game controls to mimic the position of icons in the sequencing panel, respect the action side (actions performed by the left side are on the left in the keyboard), and keep keys close to each other to avoid stretched movements (Figure 3).

3.3.2 Panel 2: Visualizing Annotations. This second panel displays all annotations positioned on a timeline that spans the entire duration of the video.

Each annotation is represented in the visualization panel as a colored circle. Lasting events are represented by two annotations linked with a colored rectangle. We added the number of these events to the left of each line based on the trainer request to better grasp the number of actions performed at a glance.

The tool enables time control using the mouse or keyboard. The user can click on an annotation to jump to the associated frame, or click (and drag) on the timeline to directly control the time cursor. To control their precision, they can modify the timeline granularity by modifying the zoom level using a slider (lens icon at the bottom right). They can also use keyboard shortcuts to skip 1 or 5 frames, see Figure 3. Feedback from the analyst showed more interest in using the keyboard for time control which led to the implementation of keyboard shortcuts also for creating annotations discussed in section 3.3.1.

3.3.3 Panel 3: General Statistics. This last panel displays a list of plots that summarize the data produced by the annotations. These plots are automatically updated when entering annotations DC 1.

Plots depict the cumulated and average holding and resting time for each hand, the evolution of the score over time, and the climbing speed as the number of seconds spent grasping holds over steps of 5 seconds. These plots facilitate identifying asymmetrical behaviors that could denote the route style or progression patterns (e.g., linear vs. staircase).

Early feedback on the tool identified the lack of data filtering that would enable focusing on specific sections of the route. We added a time selection feature (Figure 2) that enables to drag the mouse cursor over the gray area above the timeline to filter out data represented by the plots. Based on subsequent discussions, this feature is now currently used to collect data on specific route sections, but we currently do not have data quantifying this usage.

3.4 Quantifying the tool Usage

At the time of writing, the tool is used on a daily basis by the FFME analyst to produce quantitative annotations, and often by the trainer to produce qualitative annotations. This section presents global data on its usage. We compute all numbers based on the annotation files; each annotation is as saved a single file on the server. To compute the duration of an annotation session, we sort the files based on their creation time, then calculate the sum of intervals between consecutive files and exclude intervals longer than 10 minutes. Figure 4 depicts the number of annotations and the duration of annotation sessions in relation to the climbing duration, and the average annotation types created per performance.

This data indicates the number of annotations seems to linearly correlate with the climbing duration. This aligns with the fact that the longer the athlete ascends, the more annotations are produced. The distinct group on the left of the left-hand plot on Figure 4 consists mostly of performances with only score annotations. Counting the types of annotation produced provides evidence that grasp and score annotations are the most frequent. The duration of the annotation sessions seems, in contrast, to not directly correlate to the athlete climbing duration. We do not have further information on the context of these sessions and cannot identify what factors influenced this variable.

These results indicate that annotation sessions are tedious; analysts must spend considerable time creating hundreds of annotations for a single performance, and that they produce some annotation types much more frequently than others.

4 WORKSHOP 1: EVALUATING THE PROTOTYPE EFFICACY FOR ANNOTATING QUANTITATIVE AND QUALITATIVE DATA

We conducted a first workshop over a day following a few iterations of the first version of the tool (Figure 2). The goal of the workshop was to identify missing features for quantitative annotations, and use the tool for the first time in collaboration with athletes to review their performance. This workshop involved the French lead climbing trainer, the FFME video analyst, and three athletes competing at the international level.

4.1 Structure

The analyst had annotated a set of 33 videos before the workshop. The first half of the day was dedicated to discussions on the efficacy of the tool based on their experience. The second half was dedicated to evaluate how the tool supports athletes in reviewing their performance and comment on it. We recorded videos of their performances during the session, then the trainer reviewed them with the athletes, while asking questions on their feelings and adding annotations accordingly.

4.2 Lessons learned

The analyst and the trainer used the tool in distinctive ways: on the one hand to annotate objective measures alone and on the other hand to produce subjective data in collaboration with athletes. Discussions further emphasized these two steps are independent; the trainer prefers to have the athletes feedback unaltered by quantitative data. This revealed that the tool should provide an interface with a complete sequencing panel, and another focused on comments only to clearly differentiate the annotator purpose.

Insight 1

Annotations can address various levels of information and be created by different type of users, requiring to adapt the type of interface to match the annotator's intention

The video of the athlete performances were shot in a portrait format to capture the entire wall. This revealed a limitation of the tool as athletes could barely see their movements on the video. We fixed this issue by adding a functionality to enlarge the video. The feedback session also pointed out the lack of annotations the trainer could use to react to the athlete comments. We added remark annotations to serve this purpose. More importantly, the trainer wanted to summarize the performance with the athletes using closing remarks and comments, and ask the athletes to rank the performance in terms of quality, likely in a Likert-scale format.

Insight 2

The tool should support different comment types to enable multiple stakeholders to react, and provide means for summarizing the performance based on the feelings of both the trainer and the athlete

5 SECOND WORKSHOP: OBSERVING AN ANNOTATION AND AN ANALYSIS SESSION

The second workshop took place as part of a FFME bootcamp to prepare international athletes for incoming competitions. We planned two sessions focused on the tool in coordination with the trainer and analyst. In the first, the experimenter observed an annotation session, and in the second, athletes using the tool to analyze their performance.

5.1 First session: Observing an Annotation Session

Before this session, the analyst had annotated 167 videos (134 since the last workshop). This session lasted around an hour and a half. The analyst chose a video of a performance and was asked to completely annotate it. We asked them to think-aloud while annotating to explain their process and train of thoughts, and the experimenter asked questions in reaction to the analyst interactions. The session was video recorded.

5.1.1 Results and Insights. The annotation process consisted of four iterations. In the first iteration, the analyst played the video at double the speed to quickly skim through the performance and identify landmarks such as rests and when the performance ends. They explained that identifying these landmarks helps knowing whether sequences of the performance can be analyzed faster (less annotations to produce). In the second iteration, they focused on low-level annotations for both hands. They solely used the keyboard to do so; the mouse was used to delete annotations in case of mistakes. They played the video at normal speed until identifying an action to annotate, then used frame control (±1 and ±5 frames) to select the exact frame and create the annotation. In the third iteration, they used the previous annotations to come back to specific hold events to create the *score* annotations. The last iteration was used to add *energy* annotations, for which the analyst played the video at normal speed once again.

The analyst adopted a bottom-up approach building on iterations to focus on specific annotation types and levels of detail. The sequence of iterations was important to respect as each iteration seemed to build on the others. The analyst clarified that this was the ideal strategy they were using to completely annotate videos. When they need to produce data faster, they sometimes focus only on the score and energy, thus skipping iterations.

Insight 3

The annotation process consists of multiple iterations building on each other that increasingly raise the level of details covered by the annotations

During the session, the analyst raised questions on the time granularity. They wondered about the advantages of setting annotation to the exact frame, and whether capping frame control to a minimum of ±5 would provide enough accuracy and speed up the process. In this regard, the literature proposes efficient time control techniques such as subpixel interactions [34] that could potentially improve annotation time. We implemented a version of this technique, but faced strong limitations due to video downloading times when using the tools in competitions, rendering it unusable. More research is required to evaluate the efficiency of various time controls.

Insight 4

Time granularity depends on a trade-off between the annotation session duration and the accepted error margin; the optimum granularity should minimize annotation times without inducing more mistakes

5.2 Second Session: Observing Athletes Review their Performance Data

In this second session, we invited French athletes to review their annotated performance and compare their feelings to objective data. We split the session in groups of men (7) and women (5) as each group climbed the same route.

5.2.1 Structure. We first asked the athletes to estimate several characteristics of their personal performance such as the average holding time per grip, how many times they applied chalk on their hands, how many times they clipped the rope, etc. Then, they were given a computer each or created pairs to review their performance. Once done, we showed them the data of finalists from the last international competition that took place in Munich (European Championships 2022). This session lasted around 30 minutes per group.

5.2.2 Results and Insights. We observed discrepancies between the athlete's feelings and the objective data. Some athletes overestimated their climbing time while others underestimated it, and the average holding duration sometimes showed asymmetries between the left and right hand that athletes did not predict. While we miss data to run further analyses, these preliminary results suggest that the tool supports athletes in analyzing their performance precisely and provides them indicators on their climbing styles they might not be aware of.

We observed some athletes comparing their performance to others by looking at each other screens and either visually compare the sequences of annotations, or compare absolute values of average holding times. Further discussions revealed a common will to compare performances involving the same route (e.g., two training performances) to identify any type of progress.

Several athletes were interested in identifying the climbing styles of the best international athletes to identify winning strategies. They clarified that identifying this strategy would help them adopt a specific climbing style. Conversely, others said they would rather keep their own style and that data on their performance might likely not help them. Quickly analyzing the finalists data at the end of the session did not help identifying a specific behavior the athletes could compare to. Overall, this highlighted a limitation of the tool to perform one-to-one or one-to-many performance comparisons.

Insight 5

Visualizing performance data helped athletes identifying discrepancies with their feelings, but did not facilitate comparing performances to each other (one-to-one or one-to-many)

6 CONCLUSION

This case study presented the design and use of an annotation and analytical tool tailored to lead climbing, designed with and for the French climbing federation. Two workshops involving the official lead climbing trainer and analyst, as well as athletes competing at the international level, helped identifying the tool limitations and its efficacy to support the annotation and analysis process. A significant limitation was that analyzing a performance requires comparing it to others, but the tool only considered single performances. We contributed a list of insights based on our observations that can inform the design of annotation tools for lead climbing and likely other sports. Several insights are based on a single analyst's method, thus might not be representative of all annotation approaches. The tool was used to annotate 197 videos to date and is used on a daily basis: data shows a median of 11 videos per week for the last 7 weeks at the time of writing (σ=7.52). We reported usage data on the duration of annotation sessions and the number of annotations produced in relation to the climbing time analyzed. This data demonstrates the tediousness of video annotation and highlights some annotation types are produced more frequently than others. We plan on leveraging assistive tagging [43] to automate such annotations, e.g., by detecting the end of an athlete's action using computer vision approaches once the annotator defines the start, and ideally reduce the duration of annotation sessions.

ACKNOWLEDGMENTS

This project was funded by the ANR PPR STHP 2020 (PerfAnalytics project, ANR 20-STHP-0003). We would like to thank Gaël Duprat as well as all international athletes for providing feedback on the tool, and thank all the FFME staff to welcome us in various events that enabled this work.

REFERENCES

[1] Inc. Agile Sports Technologies. Last visited in October 2022. Hudl Sportscode. https://www.hudl.com/products/sportscode.
[2] Olivier Aubert and Yannick Prié. 2007. Advene: an open-source framework for integrating and visualising audiovisual metadata. In *Proceedings of the 15th international conference on Multimedia - MULTIMEDIA '07*. ACM Press, Augsburg, Germany, 1005. https://doi.org/10.1145/1291233.1291451
[3] Simone Bianco, Gianluigi Ciocca, Paolo Napoletano, and Raimondo Schettini. 2015. An interactive tool for manual, semi-automatic and automatic video annotation. *Computer Vision and Image Understanding* 131 (Feb. 2015), 88–99. https://doi.org/10.1016/j.cviu.2014.06.015
[4] Tewodros A. Biresaw, Tahir Nawaz, James Ferryman, and Anthony I. Dell. 2016. ViTBAT: Video tracking and behavior annotation tool. In *2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. 295–301. https://doi.org/10.1109/AVSS.2016.7738055
[5] J. Booth, F. Marino, C. Hill, and T. Gwinn. 1999. Energy cost of sport rock climbing in elite performers. 33, 1 (1999), 14–18. https://doi.org/10.1136/bjsm.33.1.14
[6] Jeremie Boulanger, Ludovic Seifert, Romain Herault, and Jean-Francois Coeurjolly. 2016. Automatic Sensor-Based Detection and Classification of Climbing Activities. 16, 3 (2016), 742–749. https://doi.org/10.1109/JSEN.2015.2481511
[7] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. 2019. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019).
[8] International Olympic Comittee. Last visited in October 2022. The IOC approves five new sports for the Tokyo 2020 Olympic games. https://olympics.com/ioc/news/ioc-approves-five-new-sports-for-olympic-games-tokyo-2020?
[9] ChyronHego Corporation. Last visited in October 2022. Tracab. https://tracab.com/.
[10] Dartfish. Last visited in October 2022. Dartfish. https://www.dartfish.com/.
[11] Fédération Française de la Montagne et de l'Escalade. Last visited in October 2022. Official website. https://www.ffme.fr/.
[12] Agence Nationale de la Recherche. Last visited in October 2022. Agence Nationale de la Recherche. https://anr.fr/en/.
[13] Vladislavs Dovgalecs, Jérémie Boulanger, Dominic Orth, Romain Hérault, Jean François Coeurjolly, Keith Davids, and Ludovic Seifert. 2014. Movement phase detection in climbing. 7, 3 (2014), 174–182. https://doi.org/10.1080/19346182.2015.1064128
[14] Abhishek Dutta and Andrew Zisserman. 2019. The VIA Annotation Software for Images, Audio and Video. In *Proceedings of the 27th ACM International Conference on Multimedia*. ACM, Nice France, 2276–2279. https://doi.org/10.1145/3343031.3350535
[15] Maurizio Fanchini, Frédéric Violette, Franco M. Impellizzeri, and Nicola A. Maffiuletti. 2013. Differences in Climbing-Specific Strength Between Boulder and Lead Rock Climbers. 27, 2 (2013), 310–314. https://doi.org/10.1519/JSC.0b013e3182577026
[16] International Federation for Sport Climbing. Last visited in October 2022. Official rules of all climbing disciplines. https://www.ifsc-climbing.org/index.php/world-competition/rules.
[17] Brais Gabin, Oleguer Camerino, M. Teresa Anguera, and Marta Castañer. 2012. Lince: Multiplatform Sport Analysis Software. *Procedia - Social and Behavioral Sciences* 46 (2012), 4692–4694. https://doi.org/10.1016/j.sbspro.2012.06.320
[18] Mike Hughes and Ian M Franks. 2004. *Notational analysis of sport: systems for better coaching and performance in sport.* Routledge, London; New York. OCLC: 208524179.
[19] Kinovea. Last visited in October 2022. Kinovea. https://www.kinovea.org/.
[20] Michael Kipp. 2012. Multimedia Annotation, Querying, and Analysis in Anvil. In *Multimedia Information Extraction*, Mark T. Maybury (Ed.). John Wiley & Sons, Inc., Hoboken, NJ, USA, 351–367. https://doi.org/10.1002/9781118219546.ch21
[21] Felix Kosmalla, Florian Daiber, and Antonio Krüger. 2015. ClimbSense: Automatic Climbing Route Recognition using Wrist-worn Inertia Measurement Units. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul Republic of Korea, 2015-04-18). ACM, 2033–2042. https://doi.org/10.1145/2702123.2702311
[22] Adriana Kovashka, Olga Russakovsky, Li Fei-Fei, and Kristen Grauman. 2016. Crowdsourcing in Computer Vision. *Foundations and Trends® in Computer Graphics and Vision* 10, 3 (2016), 177–243. https://doi.org/10.1561/0600000071
[23] Cassim Ladha, Nils Y. Hammerla, Patrick Olivier, and Thomas Plötz. 2013. ClimbAX: skill assessment for climbing enthusiasts. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing* (Zurich Switzerland, 2013-09-08). ACM, 235–244. https://doi.org/10.1145/2493432.2493492

[24] Bokyung Lee, Michael Lee, Pan Zhang, Alexander Tessier, and Azam Khan. 2019. Semantic Human Activity Annotation Tool Using Skeletonized Surveillance Videos. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers* (London, United Kingdom) *(UbiComp/ISWC '19 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 312–315. https://doi.org/10.1145/3341162.3343807

[25] Adrian Lees. 2003. Science and the major racket sports: a review. *Journal of Sports Sciences* 21, 9 (Sept. 2003), 707–732. https://doi.org/10.1080/0264041031000140275

[26] Quintic Consultancy Ltd. Last visited in October 2022. Quintic Software. https://www.quinticsports.com/portfolios/quintic-software/.

[27] Inc Meta Platforms. Last visited in August 2022. React. https://reactjs.org/.

[28] Meta-Video. Last visited in October 2022. Meta-Video. https://www.meta-video.com/.

[29] Dominic Orth, Graham Kerr, Keith Davids, and Ludovic Seifert. 2017. Analysis of Relations between Spatiotemporal Movement Regulation and Performance of Discrete Actions Reveals Functionality in Skilled Climbing. 8 (2017), 1744. https://doi.org/10.3389/fpsyg.2017.01744

[30] PerfAnalytics. Last visited in October 2022. Programme Prioritaire de Recherche PerfAnalytics. http://perfanalytics.fr/.

[31] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. 2013. Real-Time Crowdsourcing of Detailed Soccer Data. (2013), 9.

[32] Plotly. Last visited in October 2022. Dash: full-stack framework. https://dash.plotly.com.

[33] Lionel Reveret, Sylvain Chapelle, Franck Quaine, and Pierre Legreneur. 2020. 3D visualization of body motion in speed climbing. *Frontiers in Psychology* 11 (2020), 2188.

[34] Nicolas Roussel, Géry Casiez, Jonathan Aceituno, and Daniel Vogel. 2012. Giving a hand to the eyes: leveraging input accuracy for subpixel interaction. In *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*. ACM Press, Cambridge, Massachusetts, USA, 351. https://doi.org/10.1145/2380116.2380162

[35] Dominik Saul, Gino Steinmetz, Wolfgang Lehmann, and Arndt F. Schilling. 2019. Determinants for success in climbing: A systematic review. 17, 3 (2019), 91–100. https://doi.org/10.1016/j.jesf.2019.04.002

[36] Ludovic Seifert, Dominic Orth, Jérémie Boulanger, Vladislavs Dovgalecs, Romain Hérault, and Keith Davids. 2014. Climbing Skill and Complexity of Climbing Wall Design: Assessment of Jerk as a Novel Indicator of Performance Fluency. 30, 5 (2014), 619–625. https://doi.org/10.1123/jab.2014-0052

[37] L. Seifert, L. Wattebled, M. L'Hermette, G. Bideault, R. Herault, and K. Davids. 2013. Skill transfer, affordances and dexterity in different climbing environments. 32, 6 (2013), 1339–1352. https://doi.org/10.1016/j.humov.2013.06.006

[38] H. Sharp, J. Preece, and Y. Rogers. 2019. *Interaction Design: Beyond Human-Computer Interaction*. Wiley. https://books.google.fr/books?id=UDeQDwAAQBAJ

[39] Tomas Simon, Hanbyul Joo, Iain Matthews, and Yaser Sheikh. 2017. Hand Keypoint Detection in Single Images using Multiview Bootstrapping. In *CVPR*.

[40] Alberto Soto, Oleguer Camerino, Xavier Iglesias, M. Teresa Anguera, and Marta Castañer. 2019. LINCE PLUS: Research Software for Behavior Video Analysis. *Apunts Educació Física i Esports* 137 (July 2019), 149–153. https://doi.org/10.5672/apunts.2014-0983.es.(2019/3).137.11

[41] Fabio Sulser, Ivan Giangreco, and Heiko Schuldt. 2014. Crowd-based Semantic Event Detection and Video Annotation for Sports Videos. In *Proceedings of the 2014 International ACM Workshop on Crowdsourcing for Multimedia - CrowdMM '14*. ACM Press, Orlando, Florida, USA, 63–68. https://doi.org/10.1145/2660114.2660119

[42] Anthony Tang and Sebastian Boring. 2012. #EpicPlay: crowd-sourcing sports video highlights. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 1569–1572. https://doi.org/10.1145/2207676.2208622

[43] Meng Wang, Bingbing Ni, Xian-Sheng Hua, and Tat-Seng Chua. 2012. Assistive tagging: A survey of multimedia tagging with human-computer joint exploration. *Comput. Surveys* 44, 4 (Aug. 2012), 1–24. https://doi.org/10.1145/2333112.2333120

[44] Phillip B. Watts. 2004. Physiology of difficult rock climbing. 91, 4 (2004), 361–372. https://doi.org/10.1007/s00421-003-1036-7