Torazame: Annotation and Analysis of Shooting Competition Videos

Zehua Zeng*
University of Maryland, College Park

ABSTRACT

Although sports data analysis has become increasingly common, no academic work on video analysis for shooting competitions has been done before. Existing data analysis tools for shooting competitions are limited in functionality. They do not allow users to visualize their performance, which makes it difficult for them to draw meaningful conclusions about how to improve their results. We introduce *Torazame*, an interactive visualization tool that enables users to upload, annotate, and analyze shooting competition videos. The tool automatically marks gunshots in the soundtrack, after which the user can annotate and label the actions between gunshots. Using the event-annotated data, the tool generates multiple visualizations of the shooters performance. Users can also upload another video to compare their performance with other shooters in the same stage. Through a user study, we demonstrate how *Torazame* improves the process of data analysis on shooting competitions and enhances users' understanding of their shooting performance.

Index Terms: Human-centered computing—Visualization—Visualization techniques—sports analysis; Visualization design and evaluation methods

1 INTRODUCTION

In the past two decades, there have been significant advances in the field of sports data analysis and visualization [5]. Sports data analysis allows players and coaches to develop strategic and tactical insights into player performance, and sports visualization can enhance the exploration and understanding of sports data [3, 8, 10].

A primary challenge in sports data analysis is data acquisition, given that analysis is typically performed on video recordings. Data acquisition often involves some degree of video processing, which is a time-consuming process that might require the analyst to memorize and annotate events [7, 9]. Existing video analysis tools used for other sports are not applicable to shooting videos due to the uniqueness of each stage in the shooting competition, as well as limited recording resources.

Notably, sports data analysis efforts tend to focus on major sports such as soccer, basketball, baseball, and tennis. Despite the overall growth of sports data analysis, no academic work on video analysis for shooting competitions has been done before [6, 7]. Although there are several commercially available tools for the video analysis of shooting matches, these tools are limited in functionality. For example, the mobile app *Clip Shot* identifies the sound of all shots fired and displays split times on user uploaded videos [2]. Given that the data is displayed as a table of numbers, it might be difficult for players to draw meaningful conclusions about their performance. Another popular tool, *Shot Coach*, allows users to manually annotate the events between gunshots, but still only presents the data as a table of numbers [4].

Torazame explores the novel domain of video analysis and visualization of shooting competitions. It builds upon existing tools by offering an automatic gunshot detection feature and allowing users to annotate the events between gunshots. More importantly, it generates visualizations from the collected data in the form of grouped bar charts, timelines, and stacked bar charts.

We first provide some background about shooting competitions and related work in video analysis and data extraction. Next, we describe the design of *Torazame* and explain its functionality. Then, we discuss the design and results of our user study to test the impact of *Torazame* on users ability to analyze their shooting performance. Finally, we discuss the significance of our results and suggest some topics for future work.

2 BACKGROUND & RELATED WORK

2.1 IPSC/USPSA Shooting Competition

IPSC (International Practical Shooting Confederation) is the oldest and largest association within practical shooting, while USPSA (United States Practical Shooting Association) was incorporated as the US Region of IPSC. Unlike other types of shooting, practical shooting requires shooters to use a firearm to score as many points as possible within the shortest amount of time. IPSC/USPSA developed classifiers to evaluate the shooting skills, and shooters are separated into the following 6 classes: GM (Grand Master, performance >95% of all shooters), M (Master, performance 85 - 94.9% of all shooters), A (performance 75 - 84.9% of all shooters), B (performance 60 - 74.9% of all shooters), C (performance 40 - 59.9% of all shooters), and D (performance 2 - 39.9% of all shooters).

In IPSC/USPSA shooting, no course of fire (also called stage) is ever the same from one competition to the next, except for classifier stages, and competitors do not know what exactly stages look like in advance. The score for the stage is determined by dividing the points scored by the time it took to complete the stage, with any accrued penalties being added. That results in a Hit Factor (HF) and the highest HF wins the stage. The stage winner gets the whole stage points while others get a fraction of the stage points, which is calculated by dividing his HF by the highest HF. Therefore, to improve the match rankings, shooters would like to compare their stage performance with other competitors' by analyzing shooting videos.

2.2 Video Annotation

Video annotation plays an important role in developing insights about player performance. In shooting competitions, video annotation can be difficult due to the manner in which most shooting videos are recorded. Due to safety constraints, most videos are taken from behind shooters, which can make it difficult to analyze their actions in the video. Each shooting stage is unique in terms of location and design (i.e. physical configuration), which makes a standardized approach to video annotation challenging.

While video content analysis techniques that are used for other sports may not be applicable to shooting competitions, there have been significant improvements in the field of video annotation and analysis, which could be applicable to shooting competitions. Track Xplorer, for example, is an interactive visualization system that visualizes motor activities collected from sensor data [1].

2.3 Sports Analysis

In sports analysis, there are three main types of data that are used in sports visualizations. These include box score data, which contains

*e-mail: fiona.zehua.zeng@gmail.com

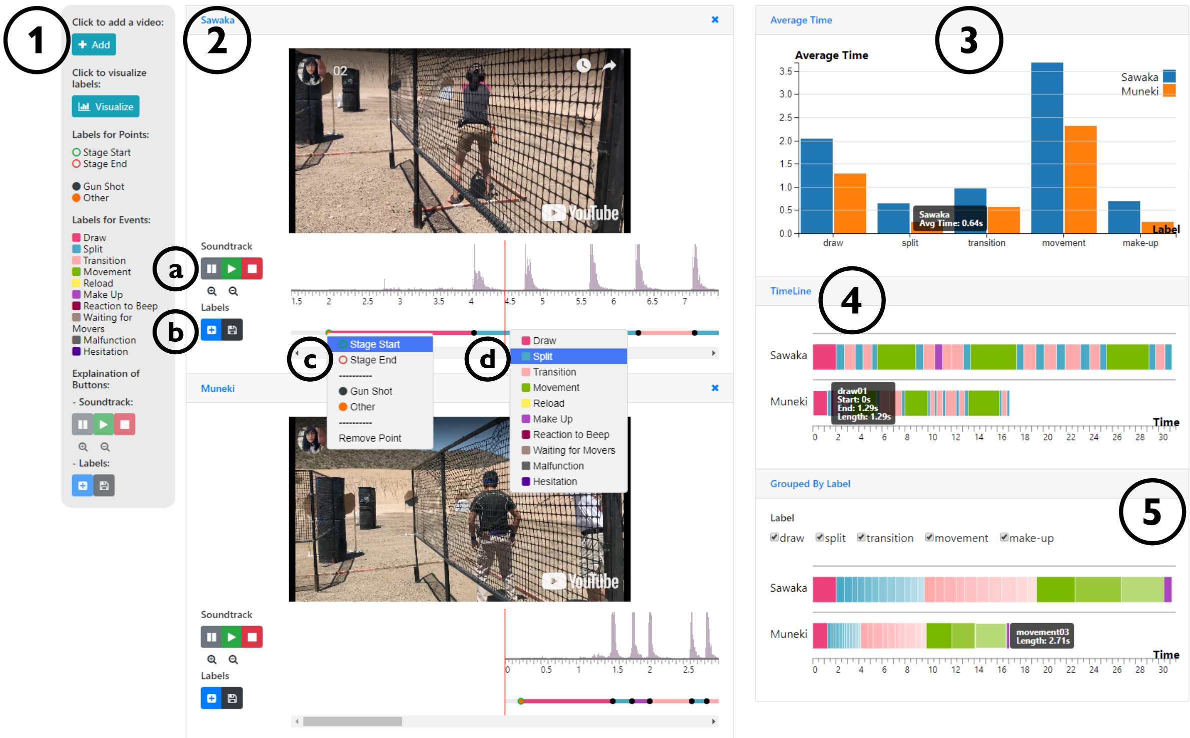

Figure 1: The user interface of Torazame: (1) control panel contains the legend of labels for points and for events, as well as an explanation of the functionality of buttons; (2) a video tab consists of an embedded video with a soundtrack and a label-track below it; (3) a grouped bar chart displays average times for each event; (4) a timeline chart depicts events in chronological order, and color encodings for labels are preserved; (5) a stacked bar chart displays individual event time grouped by labels.

statistical summaries of a sport event; tracking data, which relates to in-game actions and trajectories; and meta-data, which is data about the sport and its participants, but not a specific game or player [7]. The analysis of shooting competition videos primarily involves the use of tracking data.

Tracking data is typically collected by using machine vision to gather spatio-temporal information about players and equipment in real-time [7]. In soccer for example, there are several tools that extract information about players' movements and the ball's trajectory from video recordings [9]. Video cameras installed on stadium roofs allow for an analysis of the game from multiple angles and assist in tracking players on the field [8]. In comparison, videos of shooting competitions are mostly taken from one angle using a single camera, which can make it more difficult to accurately track player position. Videos may also not be professionally recorded like they are in other sports, and many videos are taken using smartphones. In the future, a possible solution might be to attach sensors directly to players; however, current analysis tools have to work with the present way in which shooting videos are recorded.

Even the use of sensors does not present a complete solution to the challenge of analyzing shooting videos. While sports with a primary game object (e.g. a soccer ball) use sensors for tracking purposes, it may be difficult to outfit every stage with sensors, given that stages are constantly changing [7].

3 The Design of Torazame

The purpose of the work we present here has been to create a new visualization tool that allows users to better understand their performance compared with other shooters through analyzing shooting videos. In this section we describe our motivating design considerations and the design of the *Torazame* user interface (see Fig. 1).

3.1 Design Considerations

As mentioned previously, one of the most time-consuming aspects of video analysis is data collection. In order to allow users to focus on discovering insights from the data, we wanted to minimize the amount of time and effort that users have to spend on uploading and annotating their videos as part of the data collection process.

We also wanted to improve upon existing commercial tools, which only offer the data as a table of numbers. Users may not necessarily be comfortable with visualization tools, so rather than enabling user exploration, we decided to offer a pre-defined set of visualizations.

C1. Facilitate easier creation of annotations. In the shooting matches, there exists a certain portion of long-course stages (more than 24 rounds), thus it would be tedious to mark down all gunshots manually. *Torazame* minimizes the manual work on video annotation by providing a semi-auto gunshot detection system, which labels all detected gunshots. By marking gunshots and all other important time points, *Torazame* reduces the amount of time that users need to spend on event annotation.

C2. Facilitate easier verification of annotations. Given the rapid pace of shooting stages, we considered that users may need to watch a video several times while verifying their annotations. *Torazame* retains the fundamental features of a video player and provides an interface for users to execute basic video player functions. In case the gunshot detection system makes mistakes, *Torazame* allows users to add, remove or adjust falsely marked gunshots.

C3. Enable analysis of annotated data by users who are non-experts in data science. The main drawback of existing applications for shooting video analysis is the lack of an accompanying visualization, which makes it difficult for users, especially those who are not experts in data science, to draw meaningful conclusions. To improve that, *Torazame* produces a set of interactive visualizations displaying the stage performance of each shooter based on the video annotation data labeled by the user.

3.2 The User Interface of Torazame

3.2.1 Control Panel

With the control panel (1), users can upload a video in a matter of seconds by providing a YouTube link after clicking the "Add" button, as well as visualize annotated data by clicking the "Visualize" button. The control panel also contains the legend of labels for points and events, as well as an explanation of the functionality of all buttons within the interface.

3.2.2 Video Player Tab

Once a video has been uploaded, a video tab (2) identified with the input shooters name appears. It consists of an embedded video with a soundtrack and a label-track below it. To make the process of annotation verification easier (C2), *Torazame* enables users to to play, pause, and stop the video as well as zoom in and zoom out the soundtrack with the buttons below the soundtrack (2a).

The primary purpose of the soundtrack is for users to quickly and clearly annotate different events (C1). The tool automatically detects gunshots and displays them on the label-track using black points. Users can also manually add gunshots missed by the detection system or important time points to the label-track (2b). Since time points mark the start and end of different events, all users have to do at this point in time is label each event. By right-clicking on a line segment between two points, the user can assign a color-coded label to mark that event, as seen in (2d). Users can also label the stage start and stage end or remove unnecessary points (2c).

3.2.3 Set of Interactive Visualizations

In order to facilitate users' ability to analyze their shooting data (C3), we produced a set of interactive visualizations.

The first visualization (3) displays a grouped bar chart so the user can compare their average times for each event. By displaying the averages, the tool allows users to first develop general insights. For example, a user might discover that his performance is comparable to the other shooter in most categories, but significantly slower during movements.

The timeline (4) depicts events in chronological order, and color encodings for labels are preserved. While the average time visualization provides a general overview of user performance, it is important for users to be able to visualize each individual event. For example, by checking the order of the colors, a user might notice that he had a different stage plan compared to the other competitor. On the other hand, by checking the frequency and length of undesirable labels, like make-up, hesitation and malfunction, a user would develop insight into how much time he could have reduced during the stage.

In the last visualization (5), users can view their times grouped by event. They can select which labels to display (draw, split, transition, movement, make-up, etc.) by selecting the corresponding checkboxes. They can choose to display one, some, or all the labels. This visualization allows users to develop insight into their performance on an individual event or some event combinations over time. For example, a user might discover that he lost more time on movement rather than other shooting events, like split and transition.

4 User Study

To evaluate the effectiveness and utility of *Torazame*, we conducted an open-ended exploratory user study.

4.1 Study Design

Participants. We recruited 4 participants (1 female, 3 male) between 25 and 45 years old with more than 3 years of IPSC/USPSA shooting competition experience. All subjects vary from C class to GM class in terms of shooting proficiency, and their main divisions cover Open, Limited, and Carry Optics, which are popular divisions[1] among USPSA matches. All subjects self-reported as proficient computer users, however only half of them have used other software or tools to analyze their shooting videos.

Dataset. To simulate a real-world experience in the study session, we asked participants to provide shooting videos they wanted to analyze. In each study session, participants analyzed two shooting videos in which both shooters were shooting the same stage. All participants brought videos of their own and another shooter's (who they wanted to compare with).

Study Protocol. We wished to ensure that participants could explore *Torazame* with their real goals in analyzing shooting videos. Therefore, in preparation for the session, we asked each participant to formulate three questions based on the two videos that they selected. We gave them some example questions as an idea, and they could select the provided example questions or come up with their own questions. The example list includes questions such as:

- Was there any difference between my draw (split, transition, movement) and the other shooter's?
- What was the time difference between my shooting speed and the other shooter's?
- Was there any difference between my stage plan and the other shooter's? If so, which stage plan was better?

Procedure. In the pre-session, participants first completed two surveys, one of which asked for their general information and the other one which asked them to think of three questions that they wanted to solve through analysis of shooting videos. After that, they went through a 10-minute tutorial, with a shooting video distinct from those used for actual analysis. During the session, participants proceeded with their own videos and tasks. We asked them to think-out-aloud and tell us about any good features, drawbacks or interesting discoveries they found. Participants were free to explore the whole visualization tool with no time limits. They ended the session whenever they were satisfied with their exploration. Each study session lasted approximately 60 to 80 minutes. After the analysis session, participants were asked to complete an exit survey and short interview in which we asked for their comments on *Torazame*.

Collected Data. All sessions were done remotely. An experimenter video chatted with each participant through Google Hangouts. During the session, the experimenter observed the analysis process through the screen shared by participants and took notes. We also collected data from all three questionnaires and the exit interview, including Likert scale ratings and participant quotes.

4.2 Real-World Usage Scenarios

As mentioned above, participants were asked to think of the sorts of questions that they wanted to answer through the analysis of shooting videos. The goals they set up were similar but different.

Andrew: USPSA Open GM class shooter. Andrew selected a stage from a big match that he completed one week ago, in which he did not perform well. He wanted to compare his performance with another Open GM class shooter, Ted, to see which shooting part he lost time in. Andrew knew that he had a different stage plan than Ted, so one of his goals in analyzing shooting videos was to find out whether his stage plan was better than Teds or not. On the other hand, he also wanted to know the difference of shooting speed between him and Ted.

Ryan: USPSA Limited A class shooter. Ryan shot a match with his friend this year, and he picked up one stage from the match to see why his friend beat him on time. They had almost the same

[1] IPSC/USPSA recognizes different firearms and equipment into different divisions.

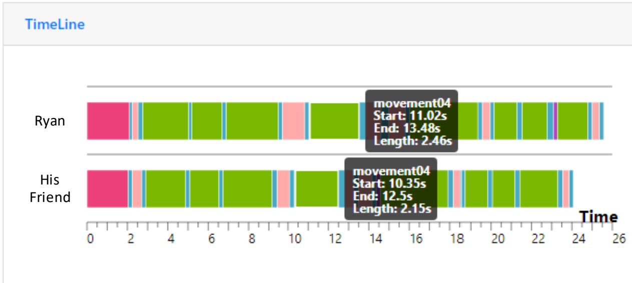

Figure 2: Ryan noticed that he lost 0.3 second compared to his friend because he did reloading by mistake during *movement04*.

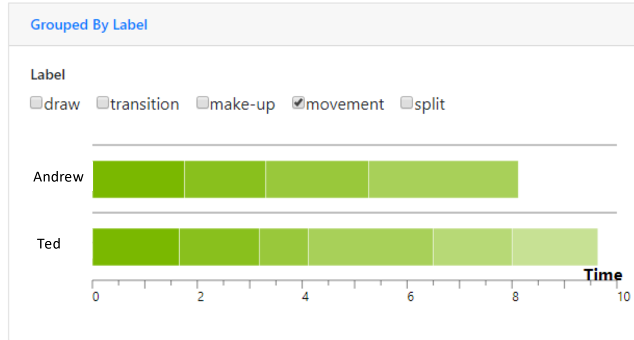

Figure 3: Andrew concluded that his stage plan was better than Ted's, since the number of movements was lesser in his plan.

stage plan, but Ryan made a small mistake during the stage. Therefore, Ryan wanted to know how much time the mistake cost and whether there was any time difference between their shooting speed.

Alice: USPSA Limited B class shooter. Alice did very well on one stage of a local match, and she wanted to know which shooting part she could work on to improve her overall performance. Therefore, she compared her video with her friend who has won IPSC Handgun World Shoot in Limited division to know the difference between their shooting time.

Sam: USPSA Carry Optics C class shooter. Sam is still a new competitive shooter, so he wanted to explore all kinds of time difference between him and experienced shooters. He chose a video that a GM class shooter posted on YouTube, and compared a video in which he shot the same stage as the GM class shooter.

4.3 Analysis & Results

In this section, we present the major insights that we collected from the four user studies.

4.3.1 Multifunctional Video Annotation System

During the user study, participants explored all functions provided in the video annotation system. They first **play**ed the video to check whether all gunshots were detected by the system correctly. They **pause**d the video when they noticed mistakes of detected gunshots. They would **remove** gunshots that the system detected by mistake. While adding a missing gunshot, they wanted to **zoom in** to see the detail of the soundtrack, and then **add** a point to the location where the gunshot happened. They would also drag to **move** the point until they were satisfied with the position. After the confirmation that all gunshots were correctly marked, participants preferred to **zoom out** the soundtrack and watch the video again to annotate shooting events, since the soundtrack moves slowly when it is minimized. During the annotation process, participants would also click the save button to **save** their current labelling to the server to prevent the data loss due to misoperations.

4.3.2 Enlightening Visualizations

Average Time Chart The average time chart enabled users to discover the average time difference in each event category between two shooters. Ryan was surprised to see that he had a slightly faster average split time, but a relatively slower average transition time than his friend. He could not predict this result. He said *"I thought there was no difference between our average split and transition time."* On the other hand, Sam realized that there was a large time difference for draw, transition and movement events between himself and the other experienced shooter.

TimeLine Chart The timeline chart helped users to distinguish the difference between stage plans and ordered shooting events. From the timeline chart, Ryan confirmed that he had the almost same stage plan with his friend. However, he did reloading by mistake during *movement04*, which caused his slower movement. He said that *"I should have done the reloading during movement05 instead of movement04. It was a longer distance and I would have more time to reload."* He realized that he lost around 0.3 second because of this small mistake by checking the time difference on *movement04* (Fig. 2).

Stacked Bar Chart The stacked bar chart helped users to check the detailed time of each category. Users can include or exclude events by checking or unchecking the corresponding check-boxes. In all study sessions, participants spent the most time exploring the stacked bar chart. They checked only one event category at a time to see the total time difference of that event. Some of them also checked multiple events as a combination, like (split + transition) combination for figuring out the total time difference of shooting speed.

By checking the movement time in the stacked bar chart, Andrew concluded that his stage plan was better than Teds, since he only needed to move four times while Ted moved six times to complete the stage (Fig. 3). He stated that with the better stage plan, he saved around one and a half seconds on movement.

4.3.3 Participant Feedback

Participants were impressed by the labelling system. Alice said that *"I really liked the labelling system. It was easy-to-use and straightforward. I had a hard time to learn how to do the labelling with Shot Coach. It doesn't allow me to add any time points besides gunshots, and it is not user-friendly to assign or change event labels."* Sam also commented that *"The gunshot detection was accurate, I did not even need to add or remove any points. I could start labelling events right away."*

In terms of the effectiveness, all participants reported in the exit survey that they got all their questions answered by exploring visualizations in *Torazame*. Some participants also stated that they found some interesting discoveries that they would not have realized without checking the charts *Torazame* provided. Andrew noted that *"I knew we had different stage plans, however, until I used the tool, I could not know that his movement time was more than mine, which confirmed myself that my stage plan was better. "* Sam also said that *"I was glad to see that there was not as much difference on split time between me and the experienced shooter as I thought, but I need to work hard on to improve my draw and transition time."*

We also asked participants who have used other tools to analyze their shooting videos before to give us comments on how *Torazame* compared to other tools. Alice said that *"The visualizations are neat.*

When I was using Shot Coach, even I marked down all event labels, it only produces numbers. I had to do some math to figure what was going on." Sam also appreciated that *"I want to keep using your tool after the study session. I have used Clip Shot when it came out since some of my friends recommended me to use it. However, I did not continue to use it after I put in a video to test it out. It did not allow me to compare with any other shooters, and only showed the time lapse between each gunshot, but I have no idea what I can learn from those information."*

4.4 Implications

Our work shows that users are better able to understand their shooting performance by using *Torazame*, particularly compared to using other tools. Visualizations were shown to be effective in helping users discover insights about their stages, which affirms the idea that sports data analysis and visualization are critical in understanding player performance. Based on results from the user study, it is also clear that the ease of video annotation plays a significant role in the efficiency and effectiveness of a tool. Future tools should make annotation as convenient as possible in order to allow users to focus on data analysis.

4.5 Limitations

Although *Torazame* helped users develop deeper insights on their stage performance by comparing their shooting videos with others', we have discovered some limitations during the study session.

Gunshot detection is not accurate in some cases. Andrew complained that our gunshot detection system could not detect all of the double taps[2]. Since both shooting videos he provided were from Open GM class shooters, there were a few double taps in the stage, where he had to pause the video to add the missing gunshots to the label-track. On the other hand, Ryan and his friend were shooting at a range which was not sound proof; in this case, the gunshot detection system detected several gunshots from the shooter who was shooting next door at the same time. During the process of gunshot checking, Ryan had to pause the video and remove some extra gunshots from the soundtrack.

Event labels are not exhaustive. We did not anticipate that users would complain about the non-exhaustive event labels, since we attempted to cover all possible labels. However, different shooters have different definitions on some shooting event concepts. While Andrew was labelling events, he commented that he was not quite sure whether to label an event as a split or a make-up. Usually shooters need to have two hits on each paper target, but some shooters took a third shot just in case there was a miss. He said that *"Make-up is more like you knew there was a miss then you shot it one more time, however, in my case, I shot the paper target three times just in case, but I was not sure if there was a miss. I would call it an extra-shot instead of a split or a make-up."*

Playback speed is unchangeable. Some participants wanted to mark down important time points which cannot be recognized by sound, like the time they started to move their hands to grab the gun, from which they could check their reaction time to the starting signal. They said that everything happened too fast in the stage, and they could not get the precise time point by checking the video with normal speed. They suggested that we should add a function to support playing the video with half of or a quarter of the speed.

Insufficient information is displayed in visualizations Some participants complained that they could not check the precise time difference. One participant stated that he could see the time difference was approximately two and a half seconds, but could not narrow the value down within 0.01s. One the other hand, Sam noted that he wanted to see the proportion of each event time in comparison to the total stage time. He said that *"It would be helpful for me to check how much percentage of the stage time I spent on movement. Was it more or less than other experienced shooters? Then I could get more general ideas of how experienced shooters shooting looks like and know more about how I should perform the match next time."*

[2]A double tap is a shooting technique where two shots are fired in rapid succession at the same target with the same sight picture. [11]

5 Conclusion & Future Work

Sports data analysis and visualization are critical in helping coaches and players improve their performance. Our work allows shooters to better understand their performance and compare to other shooters by uploading and annotating videos, and then using the collected data to generate grouped bar charts, timelines, and stacked bar charts. A user study demonstrated that *Torazame* was effective in helping users develop new insights, particularly compared to existing tools for shooting analysis.

While there have been significant advances in the field of sports data analysis and visualization, there is still a lot of work to be done. By developing tools for novel domains such as shooting competitions, it is possible to discover new video content analysis techniques. In the future, we would like to extend *Torazame* with a more accurate gunshot detection system and an automatic event-labelling system. In addition, we would also like to add more functionality into the video player system as well as explore more types of visualizations and apply to the user interface of *Torazame*.

References

[1] M. Cavallo and . Demiralp. Xplorer: A system for visual analysis of sensor-based motor activity predictions. In *2017 IEEE Workshop on Data Systems for Interactive Analysis (DSIA)*, pp. 1–4, Oct 2017. doi: 10.1109/DSIA.2017.8339085

[2] Double-Alpha Academy BV. Clipshot. https://www.doublealpha.biz/us/clipshot.

[3] A. G. Losada, R. Theron, and A. Benito. BKViz: A basketball visual analysis tool. *IEEE Computer Graphics and Applications*, 36(6):58–68, Nov. 2016. doi: 10.1109/mcg.2016.124

[4] NIftyBytes. Shot coach. https://itunes.apple.com/us/app/shot-coach/id605565597.

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

[6] T. Polk, J. Yang, Y. Hu, and Y. Zhao. TenniVis: Visualization for tennis match analysis. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2339–2348, Dec. 2014. doi: 10.1109/tvcg.2014.2346445

[7] H.-C. Shih. A survey of content-aware video analysis for sports. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(5):1212–1231, May 2018. doi: 10.1109/tcsvt.2017.2655624

[8] M. Stein, H. Janetzko, T. Breitkreutz, D. Seebacher, T. Schreck, M. Grossniklaus, I. D. Couzin, and D. A. Keim. Director's cut: Analysis and annotation of soccer matches. *IEEE Computer Graphics and Applications*, 36(5):50–60, Sept. 2016. doi: 10.1109/mcg.2016.102

[9] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlucke, 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, Jan. 2018. doi: 10.1109/tvcg.2017.2745181

[10] M. Stein, H. Janetzko, D. Seebacher, A. Jger, M. Nagel, J. Hlsch, S. Kosub, T. Schreck, D. A. Keim, and M. Grossniklaus. How to make sense of team sport data: From acquisition to data modeling and research aspects. *Data*, 2(1), 2017. doi: 10.3390/data2010002

[11] Wikipedia. Double tap. https://en.wikipedia.org/wiki/Double_tap.